Study of Climate Variability Patterns at Different Scales – A Complex Network Approach

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

d o c t o r   r e r u m   n a t u r a l i u m
(Dr. rer. nat.)
im Fach Physik
Spezialisierung: Theoretische Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Shraddha Gupta, M.Sc.

To Naanu and Sanjay uncle

Saudade

Abstract

The Earth's climate system consists of numerous interacting subsystems varying over a multitude of time scales giving rise to highly complicated spatio-temporal climate variability. Understanding processes occurring at different scales, both spatial and temporal, has been a very crucial problem in numerical weather prediction. The variability of climate, a self-constituting system, appears to be organised in patterns on large scales (such as the Indian Ocean Dipole, El Niño-Southern Oscillation, North Atlantic Oscillation). Over the last two decades, the holistic complex network-based approach to climate has been developed as a tool to complement the constructionist approach of climate modelling. Termed as *climate networks*, this approach has been very successful in providing innovative and efficient ways to predict large scale climate phenomena by detecting spatial propagation of these patterns of variability in the climate system. However, its applicability to the study of climate variability at smaller spatio-temporal scales, such as weather or intraseasonal time scales, has been very limited. Although incorporating spectral techniques such as wavelet decomposition or singular spectrum analysis into climate networks is one possible way of exploring the patterns of climate variability at different time scales, it requires a very long time series. Furthermore, such an approach reveals information about the average pattern of variability and is not able to distinguish individual events from each other.

In this thesis, it is demonstrated using climate network approach that climate variability at shorter scales are also organized in patterns. By suitably tailoring the functional network framework, its capability to detect individual extreme weather events, in particular tropical cyclones, is shown. Network-based indicators are found to exhibit significant signatures of tropical cyclones and have striking similarities with their tracks. Furthermore, an innovative complex network-based methodology is proposed to understand the evolving vortical interactions between the two cyclones when in close proximity. The method is successful in providing deeper insight into the rare event of interaction between binary cyclones, a complete understanding of which has been a challenge for weather forecasters. Network measures derived from this method can effectively quantify the changes in binary cyclone interaction leading to a complete merger.

Next, using a combination of a nonlinear synchronization measure with functional network techniques, the spatio-temporal organization of extreme rainfall events in the Asian Summer Monsoon (ASM) system, and its variability at intraseasonal time scales is investigated. This sheds light on the intricate relationship between the Indian and the East Asian summer monsoon systems, revealing the various climate processes which modulate the interaction at subseasonal time scales. Thereafter, the role of the El Niño-Southern Oscillation on the spatial co-variability of the convective processes that drive the ASM system is analyzed at interannual timescales.

Finally, the applicability of the climate network framework to analyse the spatial variability of error correlations in forecast data is demonstrated. Understanding error properties is crucial for improvement of forecasts. Correlated errors arise due to the presence of a predictable relationship between errors of different regions because of some underlying systematic or random process. The

analysis of the error network topology of a climate variable provides a preliminary understanding of the dominant source of error which may not be clearly evident from the interaction pattern of the variable itself. This shows the potential of climate networks as a very promising diagnostic tool to study error co-variability.

Zusammenfassung

Das Klimasystem der Erde besteht aus zahlreichen interagierenden Teilsystemen, die sich über eine Vielzahl von Zeitskalen hinweg verändern, was zu einer äußerst komplizierten räumlichen und zeitlichen Klimavariabilität führt. Das Verständnis von Prozessen, die auf verschiedenen räumlichen und zeitlichen Skalen wirken, ist ein zentrales Problem der numerischen Wettervorhersage. Die Variabilität des Klimas, ein sich selbst konstituierendes System, zeigt sich in großflächigen Strukturen (wie dem Indischen-Ozean-Dipol, der El-Niño-Southern Oscillation und der Nordatlantischen Oszillation). In den letzten zwei Jahrzehnten wurde der ganzheitliche, auf komplexen Netzwerken basierende Ansatz für das Klima als Instrument zur Ergänzung des konstruktivistischen Ansatzes der Klimamodellierung entwickelt. Dieser als *Klimanetzwerke* bezeichnete Ansatz hat sich als sehr erfolgreich erwiesen, da er innovative und effiziente Möglichkeiten zur Vorhersage großräumiger Klimaphänomene bietet, indem er die räumliche Ausbreitung dieser Variabilitätsmuster im Klimasystem aufdeckt. Seine Anwendbarkeit auf die Untersuchung der Klimavariabilität auf kleineren räumlich-zeitlichen Skalen, wie z. B. Wetter oder subsaisonale Zeitskalen, ist jedoch sehr begrenzt. Die Einbeziehung von Spektraltechniken wie der Wavelet-Zerlegung oder der Singularspektralanalyse in Klimanetzwerke ist zwar eine Möglichkeit, die Muster der Klimavariabilität auf verschiedenen Zeitskalen zu untersuchen, erfordert aber sehr lange Zeitreihen. Außerdem liefert ein solcher Ansatz Informationen über das durchschnittliche Muster der Variabilität und ist nicht in der Lage, einzelne Ereignisse voneinander zu unterscheiden.

In dieser Arbeit wird anhand des Klimanetzwerkansatzes gezeigt, dass die Klimavariabilität auf kürzeren Skalen ebenfalls in Mustern organisiert ist. Durch entsprechende Anpassung der funktionalen Netzwerktheorie wird gezeigt, dass diese in der Lage ist, einzelne extreme Wetterereignisse, insbesondere tropische Wirbelstürme, zu erkennen. Es hat sich gezeigt, dass netzwerkbasierte Indikatoren signifikante Signaturen tropischer Wirbelstürme aufweisen und verblüffende Ähnlichkeiten mit deren Zugbahnen haben. Darüber hinaus wird eine innovative, auf komplexen Netzwerken basierende Methode vorgeschlagen, um die sich entwickelnden Interaktionen zwischen zwei Wirbelstürmen zu analysieren, wenn sie sich in unmittelbarer Nähe befinden. Die Methode ermöglicht einen tieferen Einblick in das seltene Ereignis der Interaktion zwischen zwei Wirbelstürmen, dessen vollständiges Verständnis eine Herausforderung für die Meteorologen darstellt. Die aus dieser Methode abgeleiteten Netzwerkmaße können die Veränderungen in der Interaktion binärer Wirbelstürme, die zu einer vollständigen Fusion führen, wirksam quantifizieren.

Als Nächstes wird durch die Kombination eines nichtlinearen Synchronisationsmaßes mit funktionalen Netzwerktechniken die räumlich-zeitliche Organisation extremer Niederschlagsereignisse des Asiatischen Sommermonsuns (ASM) und seine Variabilität auf intraseasonalen Zeitskalen untersucht. Die Resultate verdeutlichen die komplizierte Beziehung zwischen dem indischen und dem ostasiatischen Sommermonsun-System und zeigen die verschiedenen Klimaprozesse auf, die die Interaktion auf subsaisonalen Zeitskalen modulieren. Danach wird die Rolle der El Niño-Southern Oscillation (ENSO) auf die räumliche sowie jährliche Variabilität der konvektiven Prozesse, die das ASM-System antreiben, analysiert.

Schließlich wird die Anwendbarkeit des Klimanetzwerkansatzes zur Analyse der räumlichen Variabilität von Fehlerkorrelationen in Vorhersagedaten demonstriert. Das Verständnis der Fehlereigenschaften ist entscheidend für die Verbesserung von Vorhersagen. Korrelierte Fehler entstehen durch die Existenz einer vorhersagbaren Beziehung zwischen Fehlern verschiedener Regionen aufgrund eines zugrunde liegenden systematischen oder zufälligen Prozesses. Die Analyse der Netzwerktopologie der Abweichungen einer Klimavariablen ermöglicht ein erstes Verständnis der vorherrschenden Fehlerquelle, die aus dem Interaktionsmuster der Variablen selbst möglicherweise nicht klar ersichtlich ist. Dies zeigt das Potenzial von Klimanetzwerken als vielversprechendes Diagnoseinstrument zur Untersuchung der Kovariabilität von Fehlern.

Acknowledgements

"I am a part of all that I have met."
– Alfred Lord Tennyson, *Ulysses*

Looking back into the years, it has been quite an eventful journey – from the bustling city life of Kolkata to co-existing with monkeys in Chennai; adapting from the deafeningly silent life at Potsdam during the pandemic to my exciting literary (and scientific!) stay at Reading, just halfway between London and Oxford. The COVID lockdown taught all of us a lot about life. I found great solace in trying to lead a life by the philosophy of *wabi-sabi*. Looking at the seasons change from my studio apartment as the year passed on, I learnt to absorb the beauty of the seemingly mundane things around me, living life a bit more slowly to feel the joy of what it means to be alive. It was a good time to introspect and rediscover myself. And this would not have been possible without understanding the contributions of all the people I have met who have shaped me the way I am today. I take this chance to express my heartfelt gratitude to them.

Working with Prof. Jürgen Kurths was always a dream come true, since I took a great interest in the field of synchronization. I consider myself extremely fortunate to have received a positive response from Prof. Kurths when I inquired for a Ph.D. position. After some interviews over a noisy internet connection, I finally started my Ph.D. journey at the Potsdam Institute for Climate Impact Research (PIK), a beautiful place offering a relaxing research environment. I am deeply grateful to Prof. Kurths for this opportunity and his unwavering belief in me since the very beginning. I thank him for his constant encouragement and guidance throughout this journey, especially during the pandemic.

For me, good things came in big packages. I thank the EU H2020 MSCA ITN CAFE project for funding my Ph.D., and allowing me to meet some of the most amazing people. Foremost among them is my co-supervisor, Prof. Florian Pappenberger, a person I grew to deeply admire over time for his intellect and great wit. Discussions with Florian were always very insightful and motivating, and left me in high spirits. I am grateful to Florian for including other colleagues from ECMWF, Linus Magnusson, David Richardson and Rebecca Emerton, in our discussions since I begun my first online secondment, which gave me a different perspective on the work I was doing. I am grateful to all of them for helping me to bring the climate perspective into my work and validate my findings. I thank Linus, who is also a part of CAFE, for his really interesting questions which triggered engaging discussions and often helped me explore a bit more on the topic. I thank David and Rebecca for their meteorological inputs. I am also thankful to Rebecca for taking the time to answer my sometimes very naive questions about meteorology. I thank other colleagues from ECMWF from whom I learnt a lot in the weather discussions and especially those with whom I interacted at the coffee discussions during my stay at ECMWF.

I was lucky to have arrived at PIK few months before COVID and witnessed the free, trustful and vibrant research atmosphere of Prof. Kurths' research group. I enjoyed being part of their lunch and coffee discussions, some of which turned out to be great collaborations.

List of Publications

This dissertation is partly based on the following publications. The author contributions corresponding to each publication are provided based on the CRediT taxonomy[1]. The identifiers [P1–P7] are cited in the text to highlight passages that are connected to these papers. The dissertation partly uses passages from the papers written by the author of this thesis, few of which are published, some are under review while some are in preparation. All passages provided by the co-authors have been thoroughly rewritten.

[P1] **S. Gupta**, N. Mastrantonas, C. Masoller, and J. Kurths. "Perspectives on the importance of complex systems in understanding our climate and climate change—The Nobel Prize in Physics 2021". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 32.5 (2022), p. 052102.

　　Author Contributions: **SG:** Data curation; software; investigation; formal analysis; methodology; visualization; validation; resources; writing – Original Draft Preparation (lead); writing – review and editing. **NM:** Validation; resources; writing – Original Draft Preparation; writing – review and editing. **CM:** Validation; resources; supervision; writing – review and editing. **JK:** Conceptualization; supervision; resources; writing – review and editing.

[P2] **S. Gupta**, N. Boers, F. Pappenberger, and J. Kurths. "Complex Network Approach for Detecting Tropical Cyclones". In: *Climate Dynamics* 57 (2021), pp. 3355–3364.

　　Author Contributions: **SG:** Conceptualization; data curation; software; investigation; formal analysis; methodology; visualization; validation; resources; writing – Original Draft Preparation (lead); writing – review and editing. **NB:** Conceptualization; formal analysis; methodology; validation; resources; writing – Original Draft Preparation; writing – review and editing; supervision. **FP:** Formal analysis; validation; supervision; writing – review and editing. **JK:** Supervision; writing – review and editing.

---

[1] *CRediT – Contributor Roles Taxonomy.* https://credit.niso.org/

[P3] S. De, **S. Gupta**, V. R. Unni, R. Ravindran, P. Kasthuri, N. Marwan, J. Kurths, and R. I. Sujith. "Study of Interaction and Complete Merging of Binary Cyclones Using Complex Networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 33.1 (2023), p. 013129.

Author Contributions: **SD:** Conceptualization (equal); Data Curation (lead); Formal Analysis (lead); Investigation (equal); Methodology (equal); Software (equal); Resources (equal); Writing/Original Draft Preparation (equal); Writing/Review & Editing (equal). **SG:** Conceptualization (equal); Methodology (equal); Resources (equal); Writing/Original Draft Preparation (equal); Writing/Review & Editing (equal). **VRU:** Conceptualization (equal); Writing/Review & Editing (equal). **RR:** Software (equal). **PK:** Resources (equal); Writing/Review & Editing (equal). **NM:** Writing/Review & Editing (equal). **JK:** Conceptualization (equal); Supervision (equal); Writing/Review & Editing (equal). **RIS:** Conceptualization (equal); Methodology (equal); Resources (equal); Funding Acquisition (lead); Project Administration (lead); Supervision (equal); Writing/Review & Editing (equal).

[P4] **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". In: *International Journal of Climatology* 43.2 (2023), pp. 1034–1049.

Author Contributions: **SG:** Conceptualization; data curation; investigation; formal analysis; methodology; software; visualization; validation; resources; writing – Original Draft Preparation; writing – review and editing. **ZS:** Data curation; formal analysis; software; investigation; visualization; validation; writing – review and editing. **NB:** Conceptualization; formal analysis; methodology; software; validation; writing – review and editing; supervision. **JK:** Supervision; writing – review and editing. **NM:** Conceptualization; validation; supervision; writing – review and editing. **FP:** Formal analysis; validation; supervision; writing – review and editing.

[P5] **S. Gupta**, A. Banerjee, N. Marwan, L. Magnusson, F. Pappenberger, C. López, E. Hernández-García, and J. Kurths. "Impact of ENSO on spatial connectivity patterns of convection during Asian Summer Monsoon". In: *In Preparation* (2022).

Author Contributions: **SG:** Conceptualization; data curation; investigation; formal analysis; methodology; software; visualization; validation; resources; writing – Original Draft Preparation (lead). **AB:** Formal analysis; methodology; visualization. **NM:** Formal analysis; methodology. **LM:** Investigation;

validation. **FP:** Investigation; validation; supervision. **CL:** Conceptualization; formal analysis; methodology; supervision. **EHG:** Conceptualization; formal analysis; methodology. **JK:** Conceptualization; resources; supervision.

[P6]   **S. Gupta**, S. De, M. S. Janaki, and A. N. S. Iyengar. "Using wavelet analysis to investigate synchronization". In: *Phys. Rev. E* 100.2 (2019), p. 022218.

Author Contributions: **SG:** Software; investigation; formal analysis; methodology; visualization; validation; resources; writing – Original Draft Preparation (lead); writing – review and editing., **SD:** Data curation (lead); software; investigation; formal analysis; methodology; visualization; validation; resources; writing – Original Draft Preparation; writing – review and editing. **MSJ:** Conceptualization; resources; validation; writing – review and editing, **ANSI:** Conceptualization; supervision; resources; formal analysis; methodology; validation; writing – review and editing.

[P7]   **S. Gupta**, A. Banerjee, N. Marwan, D. Richardson, L. Magnusson, J. Kurths, and F. Pappenberger. "Analysis of spatial variability of error correlations – A complex network approach". In: (2023). (Under Review in Quarterly Journal of the Royal Meteorological Society).

Author Contributions: **SG:** Conceptualization; data curation; investigation; formal analysis; methodology; software; visualization; validation; resources; writing – Original Draft Preparation (lead); writing – review and editing. **AB:** Formal analysis; methodology; investigation; visualization; resources; writing – review and editing. **NM:** Methodology; validation; resources; writing – review and editing. **DR:** Conceptualization; formal analysis; investigation; validation. **LM:** Conceptualization; formal analysis; investigation; validation; resources; supervision; writing – Original Draft Preparation; writing – review and editing. **JK:** Supervision; methodology; validation; writing – review and editing. **FP:** Conceptualization; formal analysis; investigation; validation; supervision; writing – Original Draft Preparation; writing – review and editing.

Contents

Contents

List of Figures

List of Tables

1. Introduction

The climate of the Earth is governed by the dynamics of both the larger solar system to which it belongs, and the several subsystems it is constituted of such as the atmosphere, the hydrosphere, the cryosphere, the biosphere, to name a few. The myriad of multi-scale interactions among these internal spheres along with the external astronomical forcing give rise to the complexity of the climate system, which is manifested in its heterogeneity. Large-scale motions, such as the ocean currents, produce smaller-scale ones like the eddies through a hierarchical cascade of instabilities. The complicated interplay among the small-scale processes, on the other hand, impacts the dynamics of the large-scale processes. The quest to understand the magnificence of nature has led mankind on to a journey from the worshipping of nature out of fear of the unknown to that of scientific rationalism.

In the following, the motivation of the scientific approach taken in this dissertation to understand the Earth's climate system are briefly outlined along with the central themes and the organization of this dissertation.

1.1. Motivation

There have been mainly two approaches to study the dynamics of the Earth system, both of which follow the 'digital-mimicry' principle [2], although from different perspectives. The first lays more emphasis on perfectly imitating the nature of the components and, to some extent, the interactions in order to build up the whole system part by part. The second, which is more recent, however utilizes the bird's eye view of the Earth, i.e., the layer of information surrounding the Earth, to understand the pattern of interactions. While the latter approach has been opted in this thesis, both these approaches are discussed in brief below to elucidate the motivation of the work.

Constructing the whole system

Due to its complexity, explaining the entire climate variability was a challenging problem. So the most feasible approach was reducing the climate system into its various components and processes, then formulating physical laws to model each of them, and finally putting all of them back together to construct the whole system.

This chapter is partly based on the associated publication Gupta et al. [P1], licensed under a Creative Commons Attribution (CC BY) license. Parts of Section 1.1 closely follow this work.

1. Introduction

This approach, known as climate modelling, has been the central tool to understand and predict climate variability. While it has been the cumulative efforts of various scientists in different disciplines, there were some key milestones that substantially improved our understanding of the climate [P1].

The notion that the laws of physics could be used for weather prediction was first conceived by Vilhelm Bjerknes [3] which later inspired Lewis Fry Richardson to lay the foundations of present-day numerical weather prediction models using the more tractable 'finite-difference equations' in the 1920s [4]. The first numerical weather forecast using the first programmable computer, the ENIAC, was produced in 1950 by John von Neumann and Jule Gregory Charney together [5]. However, a new challenge in climate modelling came up when Edward Lorenz, in 1963, discovered the chaotic nature of weather while studying his strongly simplified model of convective processes in the atmosphere [6]. His findings imposed a limit on the predictability of the weather, as the weather has a sensitive dependence on initial conditions which implies that extremely small errors in the initial state amplify rapidly and lead to large uncertainties in the forecasts.

The importance of the atmosphere, in particular, the greenhouse gases, in making our planet habitable by maintaining a livable temperature was first understood by Svante Arrhenius (Nobel Laureate in Chemistry in 1903) [7]. One of the most fundamentally important steps towards the development of a comprehensive general circulation model of the atmosphere [8] was the development of the one-dimensional, single-column radiative-convective model of the atmosphere with positive feedback effect of water vapour [9] which Syukuro Manabe (Nobel Laureate in Physics in 2021) realised was an excellent conceptual tool to study the greenhouse effect. From the very first zero-dimensional model, introduced by Arrhenius, to the currently used state-of-the-art general circulation models (GCMs) [10–13], there is a long history of contributions that can be seen from the perspective of model hierarchies [14].

The problem of scales There is often a direct relationship between spatial and temporal scales of variability in geophysical systems. A formidable problem in climate modelling arising from the heterogeneity of climate is the broad range of time scales associated with the different components of the climate system, ranging from seconds to tens of thousands of years [15]. Due to limitations on computational resources, it is not feasible for GCMs to deterministically model all dynamically important components of the climate system on all relevant time scales. Klaus Hasselmann (Nobel Laureate in Physics in 2021) introduced his stochastic model of climate variability which coupled the statistically varying, short-time-scale atmospheric system to the more slowly changing components of the climatic system [16]. He argued that the climate system is analogous to the Brownian motion problem, exhibiting the same random-walk response characteristics as large particles interacting with an ensemble of much smaller particles. His model interpreted the red-noise spectrum observed in climate signals as the natural response of the oceanic surface layers to short-time-scale atmospheric forcing which acts as a white-noise generator, thus providing insights

2

into the origin of natural climate variability [17]. Hasselmann's ingenious idea of separating physical processes by time scales was later applied to explain the glacial-interglacial climate transitions with an apparent periodicity of 10^5 years, and the corresponding temperature variations, when Benzi et al. [18, 19] incorporated the effects of internal noise arising due to atmospheric and oceanic circulations along with the periodic changes in the solar radiation due to Milankovitch forcing [20], terming the phenomena as *stochastic resonance*.

Humans being an integral part of the biosphere, have strongly influenced the climate through their various activities [21]. A broad time-scale range is typically of relevance in practical applications of climate models to quantitatively assess the climatic influence of human activities [15]. Hasselman's stochastic model played an important role in the considerations of the signal-to-noise problem in the application of climatic models to climate impact studies [22–24].

Currently, the world is heading towards a climate crisis. There is overwhelming evidence that climate change is causing a change in the frequency, intensity, spatial extent, duration, and timing of weather and climate extremes resulting in unprecedented impacts [25]. With the advent of the satellite era, the amount of observations available to scientists have also increased manifold. Despite improved observations and advancements in climate modelling, many important physical processes have not yet been resolved primarily due to the multiscale nature of the climate system. Importantly, more research is needed to improve our understanding of the interactions between different tipping elements. With the available sea of information, innovative data analysis methods offer promising approaches in addition to dynamical systems theory and non-equilibrium statistical physics [26–29]. Such approaches provide alternative perspectives of the functioning of climate system and can arguably complement numerical modeling methods. In the following, we discuss one such approach, *climate networks* [30–32], employed in this thesis to obtain insights into the Earth's climate system.

Collective dynamics of an aggregated system

Modelling processes and components of the climate system to construct a 'digital twin' of the Earth has unequivocally been a successful approach [33, 34]. Yet, the richness in dynamics imparted by the synergistic character of the Earth's climate, being a nonlinear complex dynamical system, has baffled climate scientists. Understanding how spontaneous order at different scales in the form of patterns of climate variability, such as atmospheric teleconnections and oceanic circulations, emerges from the spatio-temporal coupling among its nonlinear interacting units is important to gain insights into many physical processes. The key is to comprehend the self-organization behaviour of the climate by examining the whole system as a coherent entity.

The concept is similar to that of the emergence of consciousness in living beings from billions of unconscious neurons [35], and perhaps points towards the philosophical concept of the Earth as a conscious being [2]. Nevertheless, the logical approach to study the collective dynamics of such large interlinked systems is via the theory

of *complex networks* [36–38]. Every complex system can be thought of as an intricate network that encodes the interactions between the system's components. The architecture of the relationships is given more importance than the precise nature of the components. Despite the diversity of complex systems in various disciplines, the structure and the evolution of the networks behind each system is governed by the same organizing principles [39]. When applying networks to climate, it is assumed that its dynamics can be approximated reasonably well by a grid of low dimensional nonlinear dynamical systems, an assumption implicitly made in climate modelling [30]. A direct application of networks to climate requires specific knowledge of two aspects: first, the components which will represent the nodes and second, the coupling topology which represents the interaction between the nodes. However, our knowledge of both these aspects of the climate system is incomplete. Therefore, this calls for more indirect approaches to infer the interaction architecture of the climate system.

One way is to utilize the availability of big data in climate. The nodes of the network are taken as the regular latitude-longitude discretization of the Earth's surface, as is commonly found in the spatially gridded climate data [40]. Each node thus represents a dynamical unit whose dynamics can be measured in time via climate observations or model simulations. The interactions in this case are inferred through the ability of the subsystems to influence each other. This is estimated either by certain physical laws, if known, which may govern the interactions, or by measuring the degree of synchronization between the interacting units. The patterns of climate variability are then expected to emerge due to the spontaneous synchronization among the different subsystems. These can then be identified through measures used to quantify the network topology [41].

The problem of scales   The climate network approach is particularly attractive due to its inherent nonlinear and data-integrative nature [P1]. Climate networks have been very successful to study patterns of climate variability at larger scales such as the Indian Ocean Dipole [42], El Niño-Southern Oscillation [43–45], the North Atlantic Oscillation [46, 47], the Indian summer monsoon [48, 49], and the South American monsoon [50, 51], to name few. By quantitatively studying the spatial organization of these patterns, they have provided innovative and efficient ways to predict these large scale climate phenomena [43, 52, 53]. However, very limited attention has been given to the study of the patterns of variability at smaller scales [54]. This is mainly due to the problem that most statistical interdependence estimators to measure the degree of synchronization require sufficiently long time series for their successful application [55]. Therefore, this is disadvantageous to the application of climate networks to the study of the dynamically changing pattern of interactions at finer resolutions for phenomena occurring at weather time scales which are non-stationary and exhibit higher variability. This also hinders applications related to the study of specific patterns of variability for individual extreme weather events as opposed to the average general patterns, in which case the data available is limited

(very short time series or only a few observations) [31]. Therefore, new and innovative methods of climate network construction must be devised to overcome these problems.

Complementary to climate modelling In view of the above discussion, it may be argued that the limitation of the predictability of existing numerical climate and weather models is partly due to the fact that the models do not perfectly mimic the nature of interactions in the real climate system. Recent applications of climate networks to evaluate global teleconnections in CMIP6 Climate Projections [56, 57] and to validate the representation of boreal tropical-extratropical teleconnections in seasonal forecast data with respect to the observational data [58] have shown the potential of climate network approaches to detect model deficiences. Moreover, aside from the systematic inter-dataset comparison studies [59], climate network approaches have been used to study changes in teleconnection patterns in future climate projections [60, 61]. Therefore, besides providing innovative ways of constructing prediction schemes for large scale climate phenomena, climate networks can complement numerical modelling approaches. However, there is still a lot of scope to extract the full potential of climate networks and skillfully integrate them in different areas of numerical weather prediction.

1.2. Central themes of the thesis

In view of the aforementioned discussion, there are two central themes that surround the work presented in this thesis, as discussed below.

Application to phenomena at different time scales

One of the main themes is to implement different ways of constructing climate networks to address the problem of short time scales. In this context, time-evolving climate networks are specifically designed to study the dynamically changing patterns of weather variability of short-lived extreme weather events like tropical cyclones [P2, P3]. This is done by choosing the resolution and time window of the data, the geographical domain of analysis, and appropriate estimators to compute the interactions in accordance with the scale and variability of the extreme event. Furthermore, the existing climate network methods are applied to study phenomena at larger scales such as the spatial synchronization patterns of extreme rainfall in the Asian Summer Monsoon region at seasonal and monthly time scales [P4]. The variability of these spatial synchronization patterns at intraseasonal and interannual time scales are of interest. In this regard, the role of tropical intraseasonal oscillations on the extreme precipitation and the impact of El Niño-Southern Oscillation on the convective processes driving the precipitation are investigated [P5], respectively.

On the theoretical side, detecting different kinds of synchronization between chaotic time series and identifying the associated relevant scales are discussed on the basis

of wavelet analysis. Such scale-based synchronization measures were previously proposed only for phase synchronization [62, 63] and applications thereof [64–66]. New wavelet-based synchronization measures to extend the applicability of wavelet analysis to other kinds of synchronization, such as complete synchronization in dissipative systems and measure synchronization in Hamiltonian systems, are devised [P6]. Synchronization is known to be one of the important mechanisms involved in climate variability. However, as the climate subsystems are not exactly identical, forms of strong synchronization like complete synchronization is not possible [67]. Instead, weak forms of synchronization such as generalized or phase synchronization is more likely. Correspondingly, only wavelet-based phase synchronization techniques are applied to study phase synchronization between the precipitation of different regions at intraseasonal time scales, as other kinds of synchronization were not clearly identified in the climate examples studied here. Furthermore, the wavelet-based measures, could not be integrated with the applications concerning short-lived weather events discussed above due to the brevity of the time series associated with them. In lieu of the aforementioned arguments, measures of degree of synchronization are instead employed in climate network applications to infer the extent to which climate subsystems synchronize.

Integration with numerical weather prediction

Another primary theme is to devise methods to skilfully combine climate networks with numerical weather prediction to facilitate the improvement of the forecasts. As already discussed earlier, climate network applications have recently been extended to validate models and to study future projections of teleconnection patterns. Aside from model evaluation purposes (comparison between forecast and reanalysis/observation), the application of climate networks as promising diagnostic tools to obtain an understanding of the error properties is proposed [P7]. While the former enables us to identify the overestimation or underestimation of certain climate interactions, the latter helps to identify the dominant source of error in the climate variable. It is also of interest, in this context, to investigate whether the interaction structure of a climate variable influences that of the errors. Such applications of climate network can be integrated in data assimilation techniques to analyse the spatial variation of error correlations, as well as error assessment in other fields.

1.3. Organization of the thesis

In accordance with the themes of the dissertation described above, the present thesis is organized as follows. Part I (Chapters 2 and 3) lays down the theoretical foundations of this thesis and Part II (Chapters 4-6) focuses on their applications to climate data.

In Chapter 2, we first discuss the measures of synchronization which can be used to detect various kinds of synchronization between two chaotic time series. A special emphasis is placed on wavelet-based synchronization measures which allow us to

analyse synchronization in our preferred frequency range. Thereafter, we introduce the common measures used to quantify the degree of synchronization between two time signals.

Chapter 3 introduces the relevant elements of complex network theory and the framework of its application to analysis of climate data. It discusses the suitable choice of measure to quantify the the degree of co-variability among the wide range of available synchronization measures (Chapter 2) for construction of climate networks based on the characteristics of the data.

A first application of climate network approach to study spatial organization patterns of climate variability at weather time scales is presented in Chapter 4. It is shown that the regional weather system undergoes a spatial reorganization during the presence of a tropical cyclone, which could be then used to detect tracks of individual cyclones. Subsequently, an innovative network-based approach which can quantify the evolution of the interaction structure of the instantaneous vorticity field is applied to identify the transitions in the interaction between two cyclones when they are in close proximity. The method is shown to be able to successfully classify the stages of binary cyclone interaction leading to a complete merging of the two cyclones.

In Chapter 5, the spatial synchronization pattern of extreme rainfall events in the Asian summer monsoon (ASM) region is investigated, especially focusing on the complex relationship between the Indian and the East Asian summer monsoon systems. The relationship between the two components of the ASM system is shown to vary at intraseasonal time scales. The mechanisms underlying the respective modes of interaction between the Indian and the East Asian monsoon systems earlier and later in the season are revealed. Furthermore, by investigating the impact of the El Niño-Southern Oscillation on the ASM, which is the primary source of its interannual variability, a distinction is made between the dominant convective processes that power each of the two systems.

In Chapter 6, the applicability of climate networks to analyse spatial co-variability patterns of forecast errors is demonstrated. It is shown that the error network topologies of different climate variables reveal information about the dominant underlying process which gives rise to correlated errors in those variables.

Finally, Chapter 7 concludes this dissertation by recapitulating the achievements attained, and outlining the challenges and promising future directions of research.

Part I.

Theoretical Foundations

This first part of the thesis offers an introduction to measures of synchronization for chaotic time series. Scale-based measures are proposed to identify different kinds of synchronization. Existing measures applied in this thesis to quantify the degree of synchronization for continuous and discrete data are explained. Furthermore, the basic concepts of complex network theory and network-based analysis of spatio-temporal data are provided.

2. Measures of Synchrony

"Out of discord comes the fairest harmony."
– Heraclitus

2.1. Introduction

Synchronization is ubiquitous in nature, occurring in a wide range of physical systems across multiple disciplines, ranging from electronic circuits, communication devices and power grids, to ecological systems and human physiology and, even climate and weather forecasting [68–70]. In terms of coupled dynamical systems, synchronization may be defined as a form of coordinated behaviour among the various units constituting a dynamical system, due to some coupling or forcing. First discovered by Huygens in the 16th century in two coupled pendulum clocks hanging from a common beam [71], synchronization was known to occur only in non-chaotic systems. The reason being that the emergence of synchronization in chaotic systems is, in general, counter-intuitive due to their sensitive dependence on the initial conditions which implies that two trajectories evolving from nearby initial conditions should diverge exponentially with time, i.e., evolve in an unsynchronized manner. The discovery of chaotic synchronization [72] led to a paradigm shift in the study of dynamical systems. Since its discovery, the study of synchronization in chaotic systems has expanded tremendously and has helped in the understanding of the dynamics of various systems of immense physical relevance, climate being one of them [40, 67, 73].

Synchronization of chaotic systems is broadly classified into two categories — strong synchronization and weak synchronization [68]. Strong synchronization can occur in coupled identical systems, wherein different chaotic trajectories having different initial conditions come closer with time and, later, evolve as a single trajectory on a chaotic synchronization manifold. This is called complete (or identical) synchronization [37, 72, 74]. On the other hand, in case of non-identical coupled chaotic systems, it is possible to observe various kinds of weak synchronization, such as, phase synchronization [75] (phase difference between the coupled systems remain bounded while their amplitudes do not match), lag synchronization [76] (the coupled systems differ with

---

This chapter is partly based on the associated publication Gupta et al. [P6]. Some figures and text of Section 2.2.1 are adapted with permission from [P6]. Copyright 2019 by American Physical Society.

respect to each other with a certain phase lag), generalized synchronization [77, 78] (the trajectory of one of the coupled system can be written as an analytic function of that of the other), to name a few. In all these cases, the systems have dissipative dynamics and dissipation plays an important role in bringing the individual systems into sync. However, a weak form of synchronization, termed as measure synchronization [79, 80], has also been observed in coupled chaotic non-dissipative (or Hamiltonian) systems, wherein, the phase space domains of the individual systems overlap, i.e., their trajectories define identical invariant measures on the portion of the phase space they share, giving rise to interesting collective dynamics.

Depending on the kind of synchronization, it is possible to define various parameters to identify the emergence of synchronization, such as, synchronization error [69] (complete synchronization), phase difference between trajectories [75] (phase synchronization), similarity function [76] (lag synchronization), similarity of probability of recurrence in phase space [81] (generalized synchronization), frequencies of the trajectories visiting a particular phase space region [80] (measure synchronization), and so forth. In case of chaotic systems, it is often useful to work with quantities derived from mode-decomposition techniques, such as Fourier analysis, wavelet decomposition and empirical mode decomposition, to name a few, rather than the raw data itself. Such techniques can provide new insights into the properties of synchronized chaos. In order to illustrate this, here, we use wavelet analysis to identify some of the different types of synchronization mentioned above (Section 2.2.1).

In most real-world systems, such as the brain or the Earth's climate system, the exact governing equations are unknown. However, one can measure the state of the system at different time instances to obtain a time series representing the dynamical evolution of the system. One can then compute the statistical dependencies indicating functional interactions between the different components of the system in order to gain insight into its underlying structural connectivity. To put it in other words, the functional connectivity between two time signals imply the likelihood of synchronized dynamic activity. There are various linear and non-linear measures that can be used to quantify the degree of synchronization. In this chapter (Section 2.2.2), we discuss some of these estimators of synchrony or similarity measures which have been used in Part II of this thesis.

2.2. Synchronization measures

Over the years, many measures have been used to detect different kinds of synchronization in chaotic systems. Synchronization error $e(t)$, defined as norm of the separation between the individual trajectories, serves as a useful tool for identification of complete synchronization in identical systems [69]. The difference between the phases of the chaotic trajectories derived using techniques such as Hilbert transform, phase-space projections, Poincaré section [75], etc., or other quantities derived therefrom, are used to identify phase synchronization. Measures based on recurrences in phase

space have also been used to indicate the onset of phase synchronization, generalized synchronization and measure synchronization [80, 81].

However, in most real-world scenarios, the only accessible information is the recorded time series of the subsystems, with limited information about the governing equations. Under such circumstances, it is difficult to directly characterize the collective dynamical behaviour of the system. Wavelet analysis is one of the techniques which can be used to identify the underlying synchronization between the various modes of the components of the dynamical system [63]. In the following sub-section, we exemplify a few measures derived from Continuous Wavelet Transform (CWT) analysis which can be used to identify various types of synchronization.

2.2.1. Wavelet-based measures to identify synchronization

A wavelet may be physically understood as a waveform of finite length in space or time. Mathematically, it can be defined as a square-integrable function $\psi(t)$, that satisfies both the conditions of a nonzero but finite norm, $0 < \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$, and admissibility, i.e., $\int_{-\infty}^{\infty} \psi(t) dt = 0$ [82]. In wavelet analysis, instead of expanding a function $F(t)$ in terms of sines and cosines as in Fourier analysis, the function is expressed as a superposition of orthonormal wavelet basis functions, $\psi_{s,t_0}(t)$,

$$F(t) = \sum_{s,\, t_0 \,\in \mathbb{Z}} C(s, t_0) \psi_{s,t_0}(t) \tag{2.1}$$

The wavelet basis functions, $\psi_{s,t_0}(t)$, are dilatations and translations of the *generating* wavelet $\Psi(t)$,

$$\psi_{s,t_0}(t) = \frac{1}{\sqrt{s}} \Psi\left(\frac{t - t_0}{s}\right) \tag{2.2}$$

Here, s is the scaling parameter which indicates the degree of compression, and is inversely related to the wavelet frequency f. t_0 is the time translation parameter which determines the temporal position of the wavelet. For the given wavelet system, the CWT coefficients $\{C(s, t_0)\}$, are defined as

$$C(s, t_0) = \langle \psi_{s,t_0}(t) | F(t) \rangle = \int \psi_{s,t_0}(t) F(t) dt \tag{2.3}$$

The map $F \mapsto \{C(s, t_0)\}$ is the wavelet transform defined by the wavelet system, $\psi_{s,t_0}(t)$.

The temporal localization of the wavelet basis functions makes this technique much more advantageous than Fourier analysis in the study of synchronization. From the spectrum of CWT coefficients, it is possible to visualize the time-evolution of the various frequency modes which comprise the signal. One can then identify the frequency modes responsible for synchronization, thereby leading to a better insight into the collective dynamics. Some examples of identification of synchronization using wavelet-based measures are illustrated below.

2. Measures of Synchrony

Complete synchronization

The following wavelet-based order parameter can be used to quantify complete synchronization [P6]:

$$\mathcal{M}_D(f,t) = \left[\frac{1}{N(N-1)} \sum_{\kappa=1}^{d} \sum_{i,j=1}^{N} [C_{\kappa i}(f,t) - C_{\kappa j}(f,t)]^2 \right]^{1/2} \quad (2.4)$$

Here, N is the number of coupled sub-systems, d is the phase space dimensionality, κ represent the phase space components, and the indices i and j denote the sub-systems. $C_{\kappa i}$ (and $C_{\kappa j}$) are the magnitudes of the CWT coefficients of the κth component of the ith (and j^{th}) subsystem. In the synchronized state, $|C_{\kappa i} - C_{\kappa j}| \approx 0$, for all components κ of all pairs of subsystems i, j, at all frequencies f. Consequently, $\mathcal{M}_D(f,t) \approx 0$ during complete synchronization. Synchronization error [69], defined in terms of the phase space coordinates $x_{\kappa i}$,

$$e(t) = \left[\sum_{\kappa=1}^{d} \sum_{i,j=1}^{N} [x_{\kappa i}(t) - x_{\kappa j}(t)]^2 \right]^{1/2} \quad (2.5)$$

is similar to Eq. (2.4), although less advantageous as it does not provide information about the synchronization of the different frequency components.

To demonstrate the applicability of $\mathcal{M}_D$, we consider a pair of identical Rössler systems coupled in a drive-response (master-slave) manner [72] as follows,

$$
\begin{aligned}
\dot{x_1} &= -(y_1 + z_1) & \dot{x_2} &= -(y_1 + z_2) \\
\dot{y_1} &= x_1 + ay_1 & & \\
\dot{z_1} &= b + z_1(x_1 - c) & \dot{z_2} &= b + z_2(x_2 - c)
\end{aligned}
\quad (2.6)
$$

We choose the parameters $a = 0.2$, $b = 0.2$, and $c = 9$ for which the system exhibits chaotic complete synchronization. From Figure 2.1, it can be seen that both $e(t)$ and $\mathcal{M}_D(f,t)$ (Figure 2.1b) approach zero as the trajectories of the respective subsystems merge (Figure 2.1a). $\mathcal{M}_D(f,t)$ is also a robust order parameter for chaotic complete synchronization obtained via other coupling schemes such as active-passive decomposition [74], as shown in Figure 6 of Gupta et al. [P6].

Since complete synchronization mainly occurs between coupled identical systems, as in communication systems [83], the above measure may find application in such systems. However, climate subsystems are seldom identical, and therefore complete synchronization is not possible. Instead other weaker forms of synchronization such as generalized or phase synchronization are more likely to be found [67].
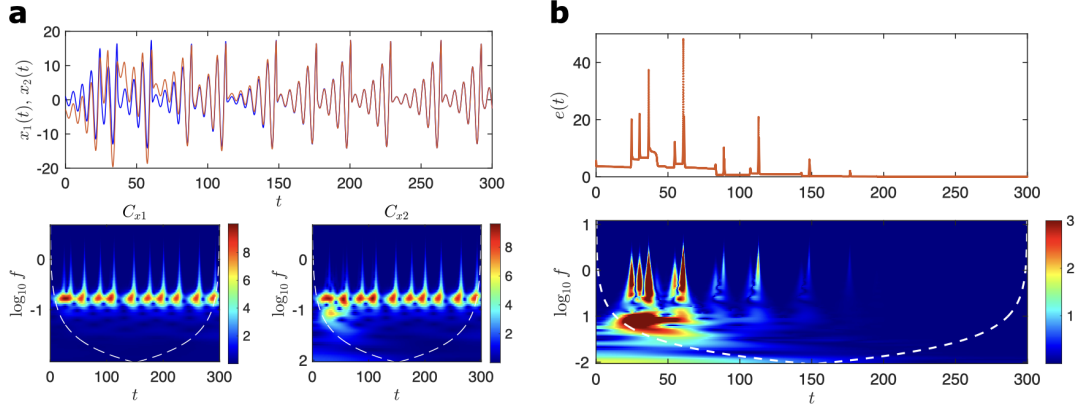
FIGURE 2.1.: (a) *Upper panel:* Time series of the x-component of the driver (blue) and response (red) systems of the coupled y-driven Rössler systems in Eq. (2.6); *Lower Panel:* Magnitude of the CWT coefficients, C_{x_1} and C_{x_2}, as functions of the wavelet frequency f and time t, with the dashed lines representing the cones of influence. (b) Comparison between the synchronization error $e(t)$ (*upper panel*) and wavelet-based order parameter $\mathcal{M}_D(f,t)$ (*lower panel*). $\mathcal{M}_D(f,t)$ not only agrees well with $e(t)$ but also shows how the various frequencies synchronize as the individual trajectories approach the synchronization manifold. (Adapted with permission from Gupta et al. [P6]. Copyright 2019 by the American Physical Society.)

Phase Synchronization

The instantaneous phases of the various frequency components of a chaotic time series can be derived from the CWT coefficients obtained by convolving the time series $F(t)$, with a complex wavelet function $\psi_{f,t_0}(t)$, as given below,

$$C(f,t) = \int \psi_{f,t}(t')F(t')dt' = A(f,t)e^{i\phi^W(f,t)} \tag{2.7}$$

where $A(f,t)$ and $\phi^W(f,t)$ are the amplitude and phase of the frequency component of frequency f at time instant t, respectively. Therefore, an instantaneous scale (frequency) dependent phase synchronization measure can be defined by considering the wavelet phase difference, $\Delta\phi^W(f,t) = |\phi_x^W(f,t) - \phi_y^W(f,t)|$, between a pair of chaotic oscillators, $x(t)$ and $y(t)$, computed from Eq. (2.7) [63]. If $\Delta\phi^W(f,t)$ remains constant with time for certain frequencies, f, then those frequency components of the oscillators are said to be phase synchronized. However, their amplitudes $A_{x,y}(f,t)$ may not be synchronous.

As an example, consider the following coupled non-identical Rössler systems [75],

$$\begin{aligned}
\dot{x}_{1,2} &= -\omega_{1,2} - z_{1,2} + \gamma(x_{2,1} - x_{1,2}), \\
\dot{y}_{1,2} &= \omega_{1,2} + 0.15y_{1,2}, \\
\dot{z}_{1,2} &= 0.2 + z_{1,2}(x_{1,2} - 10).
\end{aligned} \tag{2.8}$$

where the parameters $\omega_{1,2} = 1 \pm \Delta\omega$ govern the frequency mismatch and γ is the
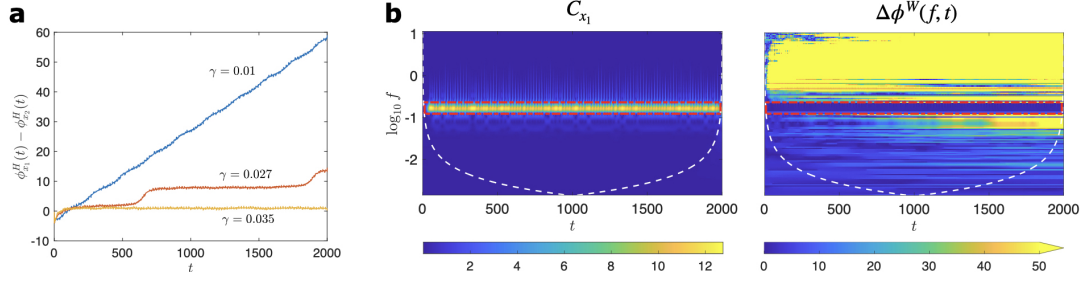
FIGURE 2.2.: (a) Phase difference $|\phi_{x_1}^H(t) - \phi_{x_2}^H(t)|$ between time series of $x_1(t)$ and $x_2(t)$, for the coupled non-identical Rössler systems in Eq. (2.8), calculated using the Hilbert Transform for different values of the coupling γ. (b) *Left*: Magnitude of CWT coefficients of $x_1(t)$, C_{x_1}; *Right:* Scale-dependent phase synchronization measure $\Delta\phi^W(f,t)$, as functions of frequency f and time t for $\gamma = 0.035$. The dashed white lines mark the cones of influence while the red boxes denote the scales which contain the largest wavelet energies. These scales are found to be phase-locked, i.e., $\Delta\phi^W(f,t)$ is close to zero and roughly constant for all of them, implying phase synchronous behaviour.

strength of coupling. For a fixed mismatch, $\Delta\omega$ (here $\Delta\omega = 0.015$), transition to phase synchronization occurs on increasing γ, as seen from Figure 2.2a, wherein the phase difference between x_1 and x_2 is computed using the Hilbert transforms of the respective time series. We then compute the above scale-dependent phase synchronization measure, $\Delta\phi^W(f,t) = |\phi_{x_1}^W(f,t) - \phi_{x_2}^W(f,t)|$, for $\gamma = 0.035$ when the oscillators are phase synchronized. It can observed from Figure 2.2b that the dominant frequency components are phase locked, i.e., $\Delta\phi^W$ is close to zero and fairly constant at those scales which account for the greatest fraction of the wavelet-spectrum energy.

It is important to note that the wavelet-based measure allows us to analyse phase synchronization in our preferred frequency range, as opposed to other approaches of computing the instantaneous phase difference such as the Hilbert transform. This is of particular advantage when phase synchronization occurs in certain frequency ranges, i.e., fewer frequency components are entrained, which cannot be revealed using Hilbert transform [62]. Moreover, the method allows the freedom to choose a suitable wavelet function according to the nature of the time series from a wide range of available wavelets, making it particularly useful for systems whose phases are not well-defined [62, 63, 68].

Many real-world systems such as fireflies [84], neurons [85] and the human cardio-respiratory system [86, 87] exhibit phase synchronization. Several examples of phase synchronized behaviour have also been identified among different limate phenomena such as that between the El Niño-Southern Oscillation (ENSO) and the Asian Summer Monsoon [64, 65, 88], between the ENSO and the North Atlantic Oscillation [66], between the South American Monsoon and Rossby waves [89], etc. In Chapter 5, we use the above wavelet-based technique to identify phase synchronization between precipitation over India and north China.
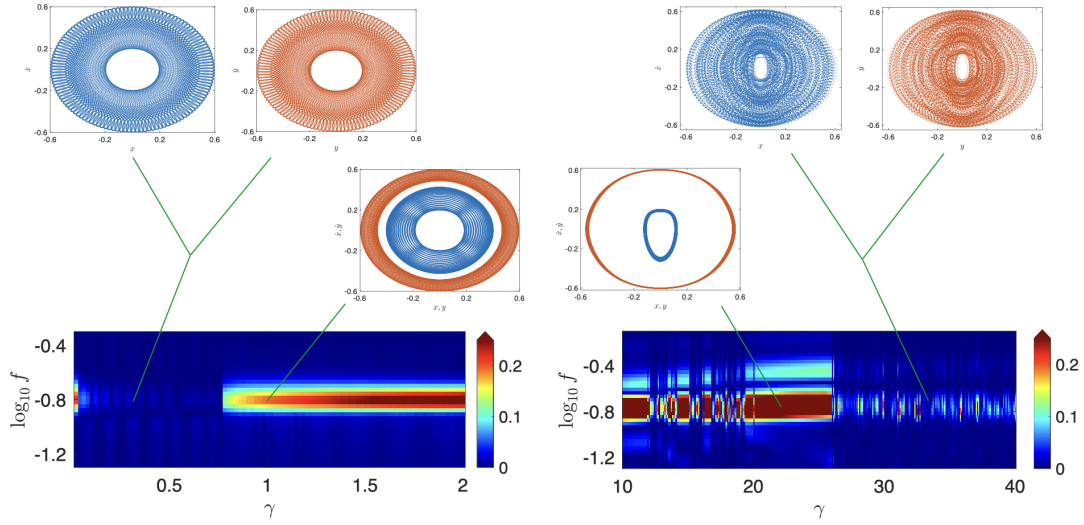
FIGURE 2.3.: The color plots *(below)* exhibit the variation of wavelet-based order parameter for identifying measure synchronization $\mathcal{M}_H(f; \gamma)$, as a function of frequency f and coupling γ for different γ. $\mathcal{M}_H(f; \gamma)$ clearly distinguishes between the measure synchronized (quasiperiodic and chaotic) and unsynchronized states and the dominant scales involved. Also shown are phase space plots of the oscillators x (blue) and y (red) for quasiperiodic measure synchronized ($\gamma = 0.3$), unsynchronized ($\gamma = 1, 22$) and chaotic measure synchronized ($\gamma = 33$) states. (Adapted with permission from Gupta et al. [P6]. Copyright 2019 by the American Physical Society.)

Measure Synchronization

Coupled Hamiltonian systems exhibiting measure synchronization have similar frequency components and root-mean-square (r.m.s.) values of amplitudes because the phase space volumes of their individual subsystems overlap [80]. Thus, a generic wavelet-based order parameter for identifying measure synchronization is,

$$\mathcal{M}_H(f; \{\gamma_i\}) = \frac{1}{N(N-1)} \sum_{\alpha, \beta = 1}^{N} |C_{\alpha,\text{rms}} - C_{\beta,\text{rms}}| \tag{2.9}$$

where $C_{\alpha,\text{rms}}$ is the r.m.s value of the CWT coefficients of the α_{th} subsystem at frequency f, N is the number of coupled subsystems, and $\{\gamma_i\}$ is the set of control parameters. Unlike dissipative systems, the subsystems in this case do not synchronize with time. Since conservative systems are ergodic, the concept of time-averaging (i.e., taking r.m.s. values) over a long duration is important to measure the degree of overlap of the phase space domains of the individual subsystems. The greater the degree of measure synchronization, the smaller are the values of $|C_{\alpha,\text{rms}} - C_{\beta,\text{rms}}|$ at the dominant frequencies, and hence $\mathcal{M}_H$ is smaller at those frequencies.

As an example, we consider a two-oscillator system [80] with the Hamiltonian $\mathcal{H} = \frac{1}{2}\left(p_x^2 + p_y^2\right) + \frac{\beta}{2}\left(x^2 + y^2\right) + \frac{\gamma}{2}x^2 y^2$, where (x, p_x) and (y, p_y) are the phase space

coordinates of the individual oscillators, respectively. $\beta = 1$ is the bare frequency of each oscillator. γ is the coupling strength by varying which we observe different dynamical regimes, such as quasiperiodic measure synchronization, chaotic measure synchronization as well as unsynchronized states, as shown in Figure 2.3. We compute $\mathcal{M}_H(f; \gamma)$, using Eq. (2.9), which distinctly shows the different synchronization regimes in Figure 2.3.

Since measure synchronization is exhibited by conservative systems, the above measure may find applications in quantum mechanics, such as in bosonic Josephson junction [90] and molecular system models, in stellar astronomy, and in studying coherent dynamical behaviour of crossed magnetic fields in plasma physics [80]. However, this form of synchronization is unlikely to be observed in climate which is a forced, dissipative, nonlinear, complex system [26].

As described above, synchronization in chaotic time series can manifest itself in many different ways. It can involve various spatial and temporal scales. Changes in synchronization can occur on rather large temporal (spatial) scales or in a more transient (gradual) manner. Following this, on the macroscopic level, the degree of synchronization (or desynchronization) is typically measured to investigate whether spatio-temporal changes in the level of synchronization can provide information about predictable relationship between two variables, for example, a common underlying climate phenomena leading to interaction between two variables. In the following subsection, we describe different linear and nonlinear measures that can be applied to quantify the degree of synchronization between two continuous or discrete time series.

2.2.2. Measures of degree of synchronization

The measures to estimate the degree of synchronization can be divided broadly into two classes: measures that are applied to continuous time series (e.g., temperature and pressure data in climate, electroencephalogram (EEG) data in neuroscience, etc.) and measures that are applied to discrete (or event-like) data (e.g. precipitation and flood data in climate and neuronal spike trains in neuroscience). The approaches to quantify the degree of synchronization between two continuous signals comprise both linear measures, like cross-correlation and spectral coherence function, as well as nonlinear measures like mutual information [91, 92], transfer entropy [93, 94], Granger causality [95, 96], and nonlinear interdependence [62, 97]. In climate science, some of these measures have been successfully applied to reveal both linear and nonlinear relationships between climate observables. For instance, one-point correlation maps or teleconnectivity maps are commonly used to study teleconnection patterns in large-scale fields [98, 99], and non-symmetric bidirectional interdependence and phase coherence between El Niño and Indian Monsoon were uncovered using nonlinear measures [88, 96].

Below we describe in detail few of these bivariate measures of synchronization for continuous time series which has been used later in this thesis.

Cross-correlation

The simplest and most widely used measure of synchronization is the linear measure of *cross correlation*. The term is generally used to refer to the *Pearson product-moment coefficient of linear correlation* [100]. The parametric empirical Pearson's correlation coefficient, $R_{ij} = \langle \hat{a}_i(t) \hat{a}_j(t - \tau) \rangle_t$, estimates the strength of the linear relationship between two normalized time series, $\hat{a}_i(t)$ and the time-lagged $\hat{a}_j(t - \tau)$, each of length N and distributed normally with zero mean and unit variance. It is defined in the time domain as a function of the time lag, $\tau = -(N-1), ..., -1, 0, 1, ..., (N-1)$, with $R_{ij}(\tau) = R_{ji}(-\tau)$. It can be viewed as the ratio of the sample covariance of the two variables, x and y, to the product of their respective standard deviations,

$$R_{xy} = \frac{Cov(x,y)}{s_x s_y} = \frac{\frac{1}{N-1} \sum_{i=1}^{N} [(x_i - \bar{x})(y_i - \bar{y})]}{\left[\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2\right]^{1/2} \left[\frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{y})^2\right]^{1/2}} \tag{2.10}$$

or as the (nearly) average product of the variables after conversion to standardized anomalies by moving the standard deviations in the denominator inside the summation of the numerator,

$$R_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} \left[\frac{(x_i - \bar{x})}{s_x} \frac{(y_i - \bar{y})}{s_y} \right] \tag{2.11}$$

It is bounded between -1 and 1, where the maximum value $R_{xy} = 1$ implies perfect positive linear association which can happen for complete or lag synchronization, while the minimum value $R_{xy} = -1$ means that there is a perfect, negative linear association. If $R_{xy} = 0$, then the variables are linearly independent. However, it must be noted that the correlation coefficient does not provide additional information about the relationship between the two variables, x and y, at least not in any physical or causative sense.

The Pearson's correlation coefficient is not robust because it may not recognize strong but nonlinear relationships between the two variables, x and y. Also, it can be extremely sensitive to one or a few outlying point pairs and produces spurious results for variables which are not normally distributed. The non-parametric Spearman's rank correlation coefficient, which does not depend on the assumption of normality, is a good alternative to Pearson's correlation coefficient [100] (Section 5.3 and Chapter 6). It is computed by calculating the Pearson's correlation coefficient using the ranks of the data. While Pearson's correlation coefficient reflects the strength of linear relationships, Spearman's rank correlation reflects the strength of monotone relationships [101]. Although, either Eq. 2.10 or Eq. 2.11 can be applied to the ranks of the data rather than to the data values themselves, the computation can be

simplified because it is known in advance that the ranks consist of integers ranging from 1 to N. The Spearman's rank correlation can then be calculated as

$$R_{xy}^{\mathrm{rank}} = 1 - \frac{6 \sum_{i=1}^{N} D_i^2}{N(N^2 - 1)} \tag{2.12}$$

where D_i is the difference in ranks between the ith pair of data values. If there are ties in ranks, it can be resolved by assigning the average rank to all the equal values before computing D_i.

The Kendall's τ coefficient [102, 103] is another rank-based correlation coefficient alternative to the standard Pearson coefficient. It is calculated by considering the relationships among all possible matchings of the $N(N-1)/2$ data pairs (x_i, y_i), and then finding the number of concordant (both members of one pair are larger than the corresponding elements in the other pair) N_C, and discordant (not concordant) N_D pairs. Identical pairs contribute $1/2$ to both N_C and N_D. Kendall's τ is then given by,

$$\tau = \frac{N_C - N_D}{N(N-1)/2} \tag{2.13}$$

Like Spearman's correlation, it can measure the strength of monotone relationships. It is more robust to outliers (large-amplitude noise) and it imposes weaker requirements on the sample size than both Pearson's and Spearman's correlation, i.e., performs well for shorter time series [104, 105]. The latter advantage makes it a suitable choice of similarity measure for investigating extreme weather events like tropical cyclones (Section 4.2), which occur over short time scales, and therefore require a shorter time window.

Mutual Information

Contrary to cross-correlation, mutual information quantifies not only linear but also nonlinear dependencies between the observables, $\hat{a}_i(t)$ and $\hat{a}_j(t)$. Originating from information theory, mutual information is based on the Shannon entropy H, which quantifies the uncertainty of a probability distribution. The mutual information M_{ij} between two observables can be interpreted as the excess amount of information generated by falsely assuming the two time series, $\hat{a}_i$ and $\hat{a}_j$, to be independent. Hence, by definition, $M_{ij} = 0$ for two independent random variables. M_{ij} is large if the two time series are highly linearly (anti)correlated. On the other hand, M_{ij} is also large if $\hat{a}_i$ and $\hat{a}_j$ have a strong nonlinear relationship in which case the Pearson's correlation coefficient yields a smaller value. M_{ij} can be estimated using the Shannon

entropies of the marginal distributions, i.e.

$$M_{ij} = H_i + H_j - H_{ij}$$

$$= -\sum_{\mu} p_i(\mu)\log p_i(\mu) - \sum_{\nu} p_j(\nu)\log p_j(\nu) - \left(-\sum_{\mu\nu} p_{ij}(\mu,\nu)\log p_{ij}(\mu,\nu) \right)$$

$$= \sum_{\mu\nu} p_{ij}(\mu,\nu)\log \frac{p_{ij}(\mu,\nu)}{p_i(\mu)p_j(\nu)}$$

$$(2.14)$$

where $p_i(\mu)$ is the probability density function (PDF) of the time series $\hat{a}_i$, and $p_{ij}(\mu,\nu)$ is the joint PDF of the pair $(\hat{a}_i, \hat{a}_j)$. Mutual information is non-negative, i.e. $M_{ij} \geq 0$, and symmetric, so that $M_{ij} = M_{ji}$. The standard unit of measurement of mutual information is the bit, if logarithms to base 2 are used.

The conventional approach for estimating mutual information from two time series, $\hat{a}_i$ and $\hat{a}_j$, consists of partitioning their supports into bins of finite size and counting the numbers of points falling into the various bins. We however use a more sophisticated, data-efficient estimator having minimal bias based on entropy estimates from the $k-$nearest neighbor distances (Section 5.3). It adapts the resolution by using bins whose sizes are adjusted according to the local data density in the joint space and then kept equal in the marginal subspaces [92]. As per the $k-$nearest neighbor formalism, H_i is written as,

$$H_i = -\psi(k) + \psi(N) + \ln(c_{d_i}) + \frac{d}{N}\sum_{\mu}\ln[\epsilon(\mu)] \tag{2.15}$$

where $\epsilon(\mu)$ is twice the distance of μ from its kth neighbor, N is the number of samples in the time series $\hat{a}_i$, d_i is the dimensionality of $\hat{a}_i$, c_{d_i} is the volume of the d_i-dimensional unit ball and ψ is the digamma function. H_j can be expressed in a similar manner. Additionally, H_{ij} can be written as,

$$H_{ij} = -\psi(k) + \psi(N) + \ln(c_{d_i}.c_{d_j}) + \frac{(d_i + d_j)}{N}\sum_{\mu}\ln[\epsilon(\mu)] \tag{2.16}$$

Using Eqs. 2.14, 2.15 and 2.16, we calculate M_{ij}.

Although there exist conceptual differences between measures of synchronization for continuous data and those for discrete data, they are closely related since very often the time stamps of the discrete events first have to be extracted from continuous recordings of an observable, e.g., extraction of flood events from river discharge data. This can lead to a significant reduction of information because, in this case, synchronization is only evaluated between defined events and not between the time series as a whole. Different choices of events can generally yield different values. Common methods of transforming a continuous time series to a discrete event series

are using peak-over-thresholds such as zero crossings, or block maxima/minima. The measures of synchronization for discrete data can either be time scale independent, such as the original definition of event synchronization [106], or time-scale dependent, such as edit-distance and other Victor-Purpura distance measures used to define the (dis)similarity between two spike trains [107–110]. We have used in this thesis a updated version of event synchronization for estimating the degree of co-variability between event-like data which is discussed as follows.

Event Synchronization

Event synchronization (ES) is an event-based similarity measure which gives an objective and comparable estimate of the synchronization between two event series at varying time scales. It was originally devised to be parameter-free by making them computationally less expensive as no parameter optimization is required [106]. It is time-scale adaptive and therefore, preferable in applications to real data for which there is no validated knowledge about the relevant time scales. In climate observables, such as precipitation, the delay between synchronous rainfall events at different locations may vary in time, even for fixed spatial distances, due to varying scales of the driving atmospheric processes, for example, changing group or phase velocities of atmospheric waves. ES has been successfully applied to analyse global and regional extreme rainfall patterns [48–51, 111] due to its capability of dealing with the technical challenges of event-like time series and the additional complication of varying temporal delays. ES can also be combined with wavelet transform to analyse synchronization between two event series at different temporal scales [112]. In Section 5.2, we use the following definitions of ES to study spatial synchronization patterns of extreme rainfall in the Asian Summer Monsoon region and to identify the large-scale atmospheric circulation anomalies associated with such synchronization patterns.

The improved version of ES, proposed in Boers et al. [111], is employed in this thesis and is defined as follows. Consider two event series $\{e_i^\mu\}_{\mu=1,..,l_i}$ and $\{e_j^\nu\}_{\nu=1,..,l_j}$ with l_i and l_j events, respectively, corresponding to grid points i and j, where e_i^μ denotes the time stamp of the μth event observed at grid point i (Figure 2.4). Two events e_i^μ and e_j^ν can be uniquely associated with each other if the absolute value of the temporal delay between them ($t_{i,j}^{\mu,\nu} := |e_i^\mu - e_j^\nu|$) is less than a *dynamical delay* defined by $\tau_{i,j}^{\mu,\nu} := \frac{\min\{t_{i,i}^{\mu,\mu-1}, t_{i,i}^{\mu,\mu+1}, t_{j,j}^{\nu,\nu-1}, t_{j,j}^{\nu,\nu+1}\}}{2}$. The allowed maximum temporal delay, $\tau_{\max}$, a parameter later introduced, is fixed to confine the synchronization time scale and avoid large waiting times. ES_{ij} gives the number of such uniquely associable, i.e. synchronous, pairs between the two event series. The time-adaptive choice of temporal delay in the range $[0, \tau_{\max}]$ instead of a static choice of delay for the entire time series makes ES advantageous over ordinary lead-lag correlation analyses. Such a dynamic choice of delay enables the method to take into account a potentially changing density of events.
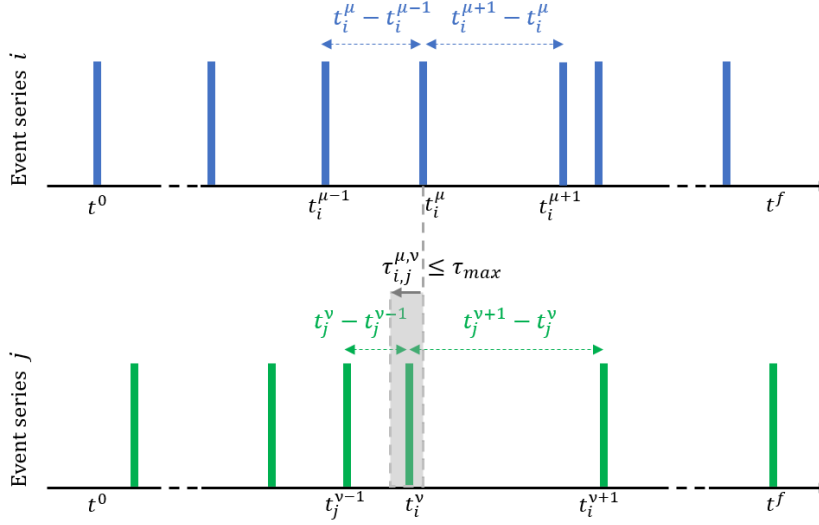
FIGURE 2.4.: Schematic diagram to illustrate the measure of the degree of synchronization between two event series using Event Synchronization.

Identification of specific time of high event synchronicity The aforementioned definition of ES can be further modified to enable us to determine the specific times when high event synchronization between two event series occurs, while keeping track of the temporal ordering [111]. Let us consider two sets of event series, A and B. We can compute $\mathrm{ES}^{\mu}_{A \to B}$ ($\mathrm{ES}^{\nu}_{B \to A}$) by counting the number of events in A (B) that have a subsequent uniquely associable event in B (A). Then, a low-pass filter can be applied to the resultant time series of $\mathrm{ES}^{\mu}_{A \to B}$ ($\mathrm{ES}^{\nu}_{B \to A}$). Finally, the specific times of high synchronization between the two event series are determined by identifying the local maxima of the time series that are above a certain threshold (here, 90th percentile) of the entire time series.

Many oscillators coupled to each other can be modelled as a network of interacting elements, where the network structure indicates which of the elements interact with one another. Due to the coupling, the oscillators appear to self-organize, leading to the emergence of synchronization which then describes a spontaneous transition to order because of the interaction [70, 113]. Such a synchronization is possible even in a network of chaotic oscillators [36]. In the real world, we seldom have a detailed understanding of the coupled system. If the dynamics of the individual oscillators are known, then one can measure the synchronization between every pair of oscillators in the network. This synchronization then translates into links between the individual oscillators that define the structure of the network of the coupled dynamical systems [114]. One can then get an understanding of the collective behaviour of the interactive system by studying the network structure, using the elements of graph theory. In the

following chapter, we discuss the framework of complex network theory and how we can study the Earth's climate using this approach.

3. Complex Networks

"Learn how to see. Realize that everything connects to everything else."

– Leonardo da Vinci

3.1. Introduction

One of the most fundamental and pervasive questions in all of science is: How does collective behaviour *emerge* from individual behaviour? For example, how does connection between billions of neurons in the brain give rise to properties such as consciousness and memory, or how interactions between different components of the Earth's climate system give rise to teleconnections. The Nobel Laureate, Phillip Anderson, explained in his 1971 paper, "More Is Different" [115], that the ability to reduce everything in terms of simple fundamental laws does not necessarily imply the capability to reconstruct the universe starting from those laws, as the constructionist hypothesis fails when faced with the twin challenges of scale and complexity. The fact that the components of complex systems don't add up simply to make up the whole is what makes them complex. Instead the interaction between large and complex aggregates of individuals can generate entirely new properties at each level of complexity. As Friedrich Nietszche said, "In individuals, insanity is rare; but in groups, parties, nations and epochs, it is the rule." The way of interaction between components of a complex system can have profound consequences, such as spread of epidemics, population genetics, global synchronization, political revolution, cascading failure in power grids and so on. Using complex networks to model complex systems is one such perspective for studying the collective behavior of interactive systems.

A *network* is a collection of objects connected to each other in some manner. Although structurally simple, networks are powerful representations of many complex systems which are constituted of large numbers of intricately interacting elements, pervading all of science ranging from physics to biology, technology, ecology, sociology and so on [39]. Some examples are food webs [116], spread of epidemics [117], electrical power grids [118], citation networks [119], the Internet backbone [41, 120], online social networks such as Facebook, Twitter [121], and the like. The fundamental essence of complex network theory is to discover what aspects of the observed complex behavior of real-world networked systems can be explained on the basis of the interaction structure alone, neglecting most other details like the dynamics of individual components or various types of connections. Mathematically, networks have been studied as *graphs*, as long ago as in 1736, by one of the greatest mathematicians,

Leonard Euler, who solved the topological problem of whether there exists a route to cross the seven bridges of the Prussian city of Königsberg passing through each bridge only once, using methods now attributed to *graph theory* [122].

The study of network topology as graphs is important in order to characterize its anatomy because it provides hints about the collective dynamics of the network, as structure always affects function [36]. As an example, flight routes in the network of airlines played an important role in the spread of the COVID disease [123]. However, unlike graphs which are objects of pure structure fixed in time, a real-world network is a dynamic object because it is constituted of numerous dynamical systems. Therefore, a network is a continuously evolving and self-constituting system leading to the emergence of some global behaviour, such as synchronization, from the interaction of its numerous units [124].

The Earth's climate is one such highly complex dynamical system wherein the intricate interactions between the various constituents of the Earth (atmosphere, hydrosphere, lithosphere, cryosphere, biosphere, etc.) give rise to dynamical patterns, teleconnections, critical elements and extreme weather events, among many others. Application of complex networks to understand these phenomena as emergent behaviour of the climate system is a macroscopic approach to its study, and is complementary to traditional approaches such as climate modelling [31]. Particularly, in this thesis, we try to understand some of the aforementioned climate and weather phenomena using network-based data analysis.

In the following sections, we first introduce the necessary concepts of graph theory (Section 3.2) and describe the types of graphs we will mainly consider in this thesis. Thereafter, in Section 3.3, we describe the network measures mainly used in this thesis to characterize the network topology. Next, in Section 3.4, we explain the classification of networks into a few broad categories based on some of their properties. Finally, Section 3.5 focuses on how complex network theory can be applied to study spatio-temporal climate data with special emphasis on the data and methodology used in this thesis.

3.2. Mathematical Representation

Definition: *A graph G consists of a nonempty set of elements, called vertices, and a list of unordered pairs of these elements, called edges. The set of elements of G is called the vertex set of G, denoted by $V(G)$, and the list of edges is called the edge list of G, denoted by $E(G)$. If $i, j \in V(G)$, then an edge of the form $\{i, j\}$ is said to join or connect i and j* [125].

The number of vertices $V(G)$ denotes the order of the graph (N), while the number of edges $E(G)$ is termed as the size of the graph (M). Graphs are have been used to represent different kinds of networks, wherein the vertices of the graph denote the network elements or *nodes*, which could be anything, for example, humans in a network of friendships, animals in an ecological network, cities in an air route network, routers in a computer network, banks or organizations in a finance network and so
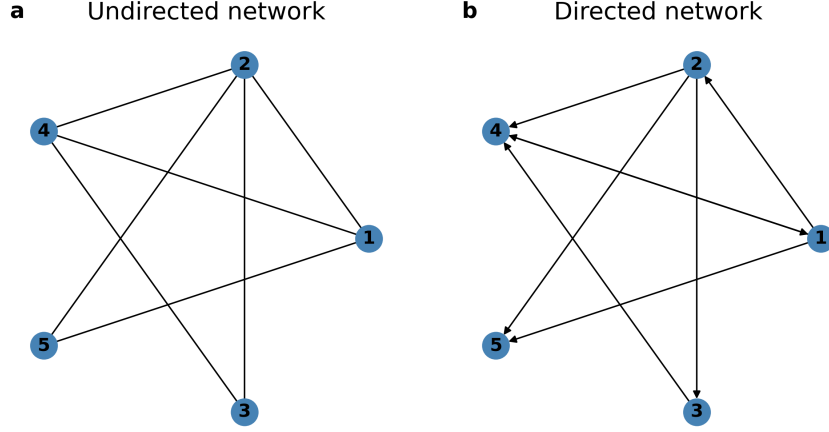
FIGURE 3.1.: Schematic diagram of a network with (a) undirected and (b) directed links respectively.

forth. The edges of the graph depict some predefined relationships or *connections* between the nodes of the network, such as friendship, prey-predator relationship, existence of a flight route, Ethernet connections, business transactions, etc., and are commonly termed as *links* of the network. In a network of oscillators, each oscillator is a node of the network, while the couplings between them could be considered as links.

The graphs or networks mainly considered in this thesis have the following common properties:

(i) *Undirected or Directed:* Based on the whether the edges have some inherent directionality, the network can be directed or undirected. If the relationship between nodes i and j is symmetric (asymmetric) then the edges are undirected (directed) (Figure 3.1).

(ii) *Unweighted:* The edges of the network are not assigned any strength. Any importance a particular edge may assume in a later stage depends solely on its relationship with other edges.

(iii) *Simple:* At most one edge $\{i, j\} \in E$ can exist between a pair of vertices and self-loops of the type $\{i, i\}$ are not allowed.

(iv) *Sparse:* A simple graph is called dense if the number of edges M is close to the maximum number of edges $M_{\max} = \binom{N}{2}$ which corresponds to a fully connected or complete graph. The networks considered here are sparse, implying, $M \ll M_{\max}$. In this respect, the *edge density* ρ of the network can be defined as $\rho = M/M_{\max}$.

(v) *Finite:* The graph has a finite number of vertices N.

(vi) *Connected:* Any vertex can be reached from any other vertex by traversing a path consisting of only a finite number of edges.

(vii) *Spatially embedded:* The vertices of a spatially embedded graph are placed in some general metric space, which may be an abstract space, like the phase space of dynamical systems, or a physical space like the Earth's surface, therefore inducing the notion of the length of edges. Hence, a spatially embedded or spatial network $G = (V, E)$ carries a mapping $V \to S : i \mapsto x_i$, assigning each vertex $i \in V$ to an element x_i of a set S and a metric $l : S \times S \to \mathbb{R} : (x_i, x_j) \mapsto l_{ij}$. x_i is called the coordinate vector of vertex i, and l_{ij} can be used as a measure of edge length whenever $i, j \in E$.

It may be argued that these assumptions greatly simplify the resulting analysis since some relationships in real world networks may be more important than the others, there may exist isolated nodes or groups of nodes, or often multiple types of relationships may exist between the same pair of nodes. However, these assumptions provide a good starting point to model complex systems, and as we shall see later in the thesis (Part II), they reveal a great deal about the behaviour of highly complex systems.

Graphs are commonly represented pictorially. However, the order and size of most real world networks is too large for one to be able to extract any meaningful information from it. An *adjacency matrix* or an *adjacency list* is used to represent the interactions encoded in a network which can then be used to compute the graph properties. The adjacency matrix $\mathrm{A}(G)$ is an $N \times N$ matrix in which the value of A_{ij} denotes the strength of the interaction between nodes i and j. In case of an unweighted network, the elements A_{ij} is either 1 or 0 indicating the presence or absence of a link between the corresponding pair of nodes, i.e.,

$$A_{ij} = \begin{cases} 1, & \text{if } i, j \in E. \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

For an undirected network, the adjacency matrix A is symmetric. The adjacency matrix of the undirected network corresponding to Figure 3.1a is given as:

$$A_{ij} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} \tag{3.2}$$

For a directed network like in Figure 3.1b, the adjacency matrix is asymmetric as below:

$$A_{ij} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{3.3}$$

It is important to mention here that the adjacency matrix representation is primarily used for only formal purposes because the memory required to store a dense adjacency matrix grows as N^2, and therefore, it is computationally feasible for only relatively small graphs. Most calculations therefore rely on the sparse adjacency list representation which is a collection of unordered lists where the ith list is a list of vertices adjacent to the ith vertex.

In order to characterize the topological structure of the network, network measures can be defined which allow us to study different aspects of the topology at local, mesoscopic and global scales [30, 41]. In the following section, we list some of these measures which have been used extensively in Part II of this thesis.

3.3. Network Measures

Degree: Centrality measures usually assign a number to each node which tells us how important that node is [41]. The *degree centrality* uses local information of the direct neighborhood of a node. The number of neighbours adjacent (connected) to a given vertex i is called the *degree* of i, denoted as k_i. For an undirected network, k_i is computed as,

$$k_i = \sum_{j=1}^{N} A_{ij} \tag{3.4}$$

where N is the total number of nodes in the network and A_{ij} is the adjacency matrix. For example, the degree of node 1 in Figure 3.1a is 3, that of node 2 in 2 and so on. The probability that a node in the network is connected to k other nodes is given by its *degree distribution* p_k.

The *degree assortativity coefficient* r is a measure of preferential connectivity in networks based on the node degree, i.e., whether a node of high degree preferentially is linked to other nodes of high degree, and vice-versa. It is calculated as the Pearson correlation coefficient of degree between pairs of linked nodes, with its value between $-1 \leq r \leq 1$ [126]. Positive values of r indicate a correlation between nodes of similar degree, with $r = 1$ meaning the network has perfect assortative mixing pattern. Negative values indicate relationships between nodes of different degree – $r = -1$ implies that the network is completely disassortative. When $r = 0$ the network is non-assortative. The degree assortativity could be used as a way to estimate the level of homogeneity/heterogeneity of real networks, where a more assortative network is more homogeneous and vice-versa.

In case of directed networks, counting of the outgoing links from a node have to be differentiated from the incoming links to the node. The *in-degree* k_i^{in} which counts the number of incoming links to a node is defined as,

$$k_i^{\text{in}} = \sum_{j=1}^{N} A_{ji} \tag{3.5}$$

The *out-degree* k_i^{out} which counts the number of outgoing links from a node can be computed as,

$$k_i^{\text{out}} = \sum_{j=1}^{N} A_{ij} \tag{3.6}$$

Clustering: Measures such as *local clustering coefficient* and *global clustering coefficient* depend upon the information of the neighbours and next neighbours of the nodes. The clustering coefficient is a measure of the degree to which nodes in a network tend to cluster together.

The *local clustering coefficient* [124] of a node i in a network quantifies how close its neighbours are to being a clique (i.e., a complete graph), that is, the average probability that a pair of node i's neighbours, j and h, are connected. Mathematically, we calculate the ratio of the number of links connecting the direct neighbours of i to the number of all possible connections between them,

$$C_i = \frac{\sum_{j,h=1}^{n} A_{ij} A_{ih} A_{jh}}{k_i \, (k_i - 1)} \tag{3.7}$$

The local clustering coefficient, C_i, measures the control over flows between just the immediate neighbours of a node [41]. It indicates spatial continuity in network.

The *global clustering coefficient C*, also known as *transitivity* [41], measures the average probability that two neighbours of a vertex are themselves neighbours, for the whole network. It measures the density of triangles in the networks and is defined as the fraction of paths of length two in the network (triplet of nodes) that are closed. This is equivalent to the ratio of the number of closed triplets to the total number of triplets. As a triangle graph includes three closed triplets, one centred on each node, the number of closed triplets is equal to thrice the number of triangles.

$$C = \frac{3 \times \text{the number of triangles}}{\text{number of all triplets}} \tag{3.8}$$

For an undirected network with an adjacency matrix A, the global clustering coefficient is expressed as:

$$C = \frac{\sum_{i,j,k} A_{ij} A_{jk} A_{ki}}{\sum_i k_i (k_i - 1)} \tag{3.9}$$

and $C = 0$ when the denominator is zero. If $C = 1$, perfect transitivity occurs in the network, i.e., the components of the network are all cliques. The global clustering coefficient is of interest because a higher C than expected by chance indicates the formation of localized structures of high connectivity in a network, e.g., the presence of tightly knit groups characterized by a high density of ties in a social network.

Spatial length scale associated with a node: Sometimes an estimation of the spatial length scale involved with each node of a spatially embedded network can provide deeper insight into the collective dynamics, e.g. if the metric embedding space is the Earth's surface [P2, 48–50]. In such cases, we can define the mean spatial distance associated with each node based on the spatial distance between the node and its connected neighbours. For a network embedded on the Earth's surface, the *mean geographical distance* $\mathcal{L}_i$ of a node i can be computed by calculating the mean of the spatial distances of i from all its connected neighbours j, along the corresponding great-circles, i.e.,

$$\mathcal{L}_i = \frac{1}{k_i} \sum_{j=1}^{n} \mathcal{L}_{ij} A_{ij} \tag{3.10}$$

where $\mathcal{L}_{ij}$ is the great-circle distance between nodes i and j calculated using the Haversine formula for a spherical Earth projected on to a plane [127].

Common Component Function (CCF): This network measure enables us to measure the similarity in network topology between two different networks. For two unweighted, undirected networks, $G_a(V, E_a)$ and $G_b(V, E_b)$, defined on the same set of nodes V but with links E_a and E_b respectively, the common component function CCF(G_a, G_b) counts the number of common links $n(E_a \cap E_b)$ between the pair of networks [128].

$$\text{CCF}(G_a, G_b) = \frac{\sum_{j,i=1(i<j)}^{N} A_{ij}^a A_{ij}^b}{\sum_{j,i=1(i<j)}^{N} A_{ij}^a A_{ij}^a} , \tag{3.11}$$

where N is the number of nodes in the set V, A^a and A^b are the adjacency matrices of networks G_a and G_b respectively, and $\sum_{j,i=1(i<j)}^{N} A_{ij}^a A_{ij}^a = \sum_{j,i=1(i<j)}^{N} A_{ij}^a$ is simply the number of links in the network G_a. Therefore, the normalised CCF(G_a, G_b) takes values between [0,1]. This implies that if the networks G_a and G_b are completely identical then CCF$(G_a, G_b) = 1$, and CCF(G_a, G_b) is 0 if they have no common links.
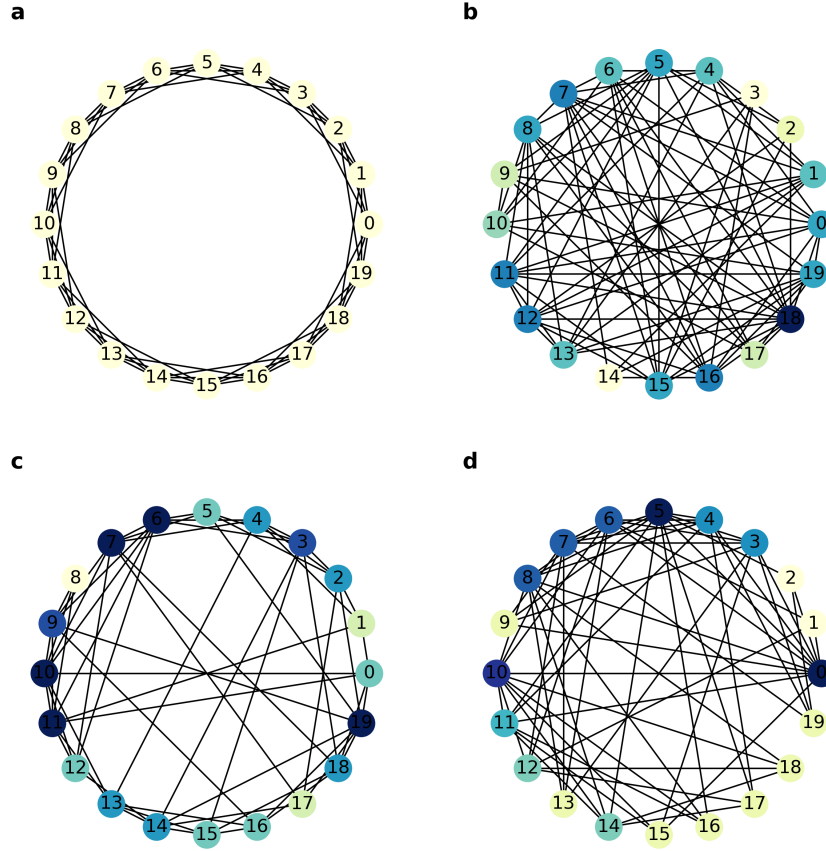
FIGURE 3.2.: Schematic representation of (a) Regular network, (b) Erdős–Rényi random network, (c) Watts-Strogatz network have small world property, and (d) Barabási-Albert network having a scale free degree distribution.

3.4. Types of Networks

Networks can be broadly classified into a few categories based on various aspects such as topological structure, type of connectivity, whether defined on a spatial or temporal domain, time evolution, etc. Below we describe briefly some of these classifications.

3.4.1. On the basis of topology

The pattern of connections between the nodes of a network can range from being completely ordered to completely random. Real world networks have a complex structure which is neither completely regular nor completely random. On the basis of topological structure, four basic types of networks are typically used to model real world networks, as described below.

Regular (ordered) networks: These networks consist of a fixed number of nodes, each of which is connected in a specific way to a number of its neighbors via the same number of links (Figure 3.2a). The network becomes fully connected if each node is linked to every other node. Regular networks are locally clustered but do not exhibit efficient transfer of information.

Random networks: In the classical or Erdős–Rényi networks [129], the nodes are connected at random (Figure 3.2b). The degree distribution of the nodes of the network is a Poisson distribution. However, a more realistic random network in space could also be defined by considering the relationship between the internode spatial distance, which could be the *great-circle* distance if the metric embedding space is the Earth's surface, and the connection probability [130]. Random networks do not exhibit local clustering but can transfer information efficiently. On the other hand, they are not stable as communication can easily be hampered by removal of nodes or links.

Small-world networks: A desirable property in real-world networks is that they should be efficient in processing information but, at the same time, be stable. A small-world network not only exhibits a high degree of local clustering, but also has a small number of long-range connections added at random for efficient information transfer. The Watts-Strogatz networks is one such family of small-world networks [124] (Figure 3.2c).

Networks with a given degree distribution: The network structure may be characterised by a variety of degree distributions such as truncated power-law distributions [129], Gaussian distributions [131], power-law distributions [120], and distributions consisting of two power laws separated by a cutoff value of degree [132, 133]. An interesting property of most real world networks is that there exists a small number of nodes with unusually high degree called as 'hubs', which in some cases are known to have a dominant effect on the behavior of the network as a whole [41]. By having a few nodes with a large number of links and many nodes with fewer links, these networks exhibit the small world behaviour. This property of the network architecture wherein the degree distribution has a long tail indicating the presence of a few very high degree nodes (supernodes) is often associated with power law distributions, $p_k \propto k^{-\gamma}$, in which case they are termed as scale-free networks [120, 134] (Figure 3.2d). In Chapter 6, we use the property of existence of hubs in the real-world networks to test the significance of the degree distribution of the forecast error networks against that of random networks.

3.4.2. Based on the type of connectivity

Based on whether the interactions in the network represent actual physical connectivity or statistical dependency, networks can be fundamentally classified into *structural* and *functional* networks [30, 32].

Structural networks: These networks reflect the topological architecture, on a certain level of abstraction, of existing connections between objects (e.g., computers, neurons, columns of neurons), which may either be physical connections (e.g., internet, power grids, neuronal networks) or abstract relations (e.g., world wide web, social networks, citation networks).

Functional networks: In most cases, the complete knowledge of the interactions in a complex system is lacking, and thus the structural network topology is not known. In such cases, knowledge of functional dependence is inferred solely from the measured or simulated dynamics of subsystems by detecting and assessing similarities in their dynamical behavior. The resulting topological interconnections between the components of the network are referred to as functional networks, which may not necessarily represent but are expected to have close resemblance to the real physical connectivity of the system. Therefore, unlike structural networks, they do not directly allow us to draw conclusions on a causal interrelationship between the dynamics of the different components. Brain networks [135] and climate networks [40] are two archetypal examples of functional networks. This implies that special emphasis has to be given to physical arguments when interpreting topological features of these networks. Almost all reconstructed complex networks from climate data in Chapters 4-6 are functional networks.

3.4.3. Spatial versus temporal similarity networks

Spatial similarity networks: These networks are based on fields of time series, such as a spatio-temporal climate dataset [30], wherein similarity is measured between two nodes distributed in space. Climate networks regarded in this thesis (Chapters 4-6) are examples of spatial networks.

Temporal similarity networks: Such networks are sophisticated techniques used in nonlinear time series analysis. Examples include recurrence networks [136] which make use of graph representations of similarity relationships between state vectors or groups of state vectors to represent a single time series, visibility graphs in time domain [137], etc. Such networks are outside the current scope of this thesis.

3.4.4. Static versus Evolving Networks

Real-world networks are dynamical objects. Therefore, the dynamics of nodes are continuously evolving and so are the relationships between the different pairs of nodes in the network. In this thesis, we consider the number of nodes to be fixed, while the connections between nodes may vary with time. Depending on whether or not the assumption of stationarity of the dynamics of the nodes is taken into account, two types of similarity networks are considered in this thesis:

Static networks: For a stationary system, the functional dependence between its components does not explicitly depend on time. In such cases, similarity between the nodes of the network are computed for the entire time period, and the functional network can essentially be considered to represent the time-averaged interactions between the components of the networks for that time period. Such networks have been used in Sections 5.2 and 6.3.

Evolving networks: In order to detect dynamical transitions or changes in the interaction pattern in the network, sliding window-based analysis could be done using the functional network approach [138–140]. In this case, the entire time period is divided into overlapping or non-overlapping time windows, and then spatial similarity networks are computed for each window. The time-evolving networks can be treated as successive snapshots of the interaction structure of the complex systems and the dynamics are assumed to be stationary for the duration of a single time window. Evolving complex networks capture the mechanisms that contribute to the system's evolution. They are of particular importance to study real-world networks [138, 139, 141, 142] which undergo structural changes (addition or deletion of nodes and connections) as the system dynamically evolves (such as growth or aging processes). They exhibit rich dynamics, such as, structure formation and evolving collective behaviour among the elements. Sections 4.2, 5.3 and 6.5 employ such time-evolving networks to study changes in the network topology due to some climate or extreme weather phenomena. In Section 4.3, we define time-evolving networks defined in a way such that each network truly represents instantaneous interactions between the nodes.

3.5. Complex network approach for spatiotemporal analysis of Climate data

3.5.1. Overview

Inferring relationships between the climate observables of two regions using correlation, among other statistical methods, has been a common approach for meteorologists. Correlations between the climate fields of different places are popularly represented using one-point correlation maps or teleconnectivity maps and have been extensively used to detect teleconnections [98, 99, 143]. However, due to the atmospheric data being excessively huge, the very large number of pairwise comparisons does not allow us to directly obtain a complete visual appreciation of the patterns of relationship in them. In recent decades, complex networks have been used extensively for inferring interrelationships from spatio-temporal data of spatially extended systems like the human brain [135] or the Earth's climate system [40]. Although there have been major advances in the modelling of the Earth's climate [P1], our knowledge on the various components and the intricate interactions among them has been lacking. Functional network approaches to analyze the climate data from observations or
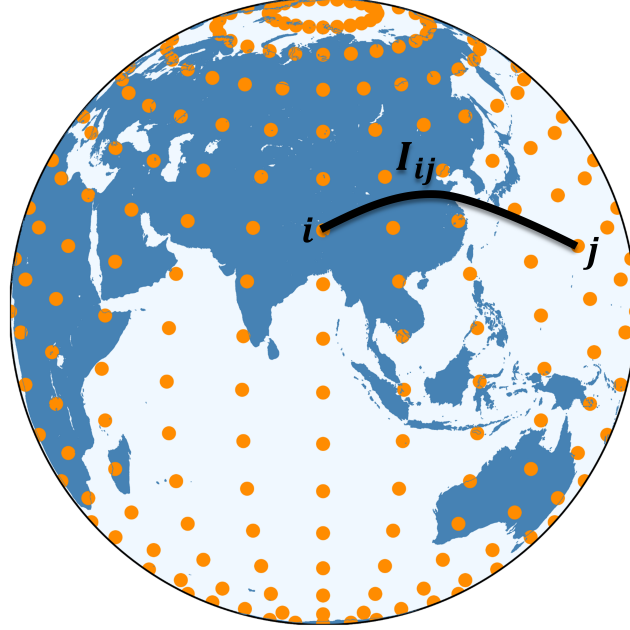
FIGURE 3.3.: Schematic representation of climate network: spatial grid points of an underlying climate dataset are considered nodes (orange dots). The interaction I_{ij} between the dynamics of any two nodes i and j is estimated using some similarity measure or physical law to obtain the links.

models, termed as *climate networks*, have proven useful to understand the functioning of the highly complicated climate system, and thus gain insight into the real physical connectivity of the system [31].

The nodes of a climate network are identified with the spatial grid points of an underlying global climate data set, and therefore are, in general, embedded in the four-dimensional space-time. If networks are constructed only for a particular surface above the Earth's surface, then the three dimensional embedding space-time consists of the space coordinate in terms of the latitude and longitude of a grid point and the time coordinate, although the two spatial coordinates now can be further be condensed into a single dimension. The central idea behind climate networks is akin to that of inferring functional dependencies from a network of oscillators when the coupling topology is not known [30, 114, 144–146]. However, the climate network approach is based on the assumption that the dynamics of the climate system can be approximated by a grid of low dimensional nonlinear dynamical systems interacting only with their spatial neighbors according to the locality principle of classical physics. Each node is considered to represent a dynamical system whose dynamics can be measured, and in this case, provided by the corresponding time series of the climate observable at that location. The degree of synchronization between any pair of nodes can be measured (Section 2.2.2) and if two nodes are significantly synchronized with each other, then they are connected by links. The synchronization pattern so

obtained consists of nodes that are locally correlated as well as those with long range correlations which can be useful to identify teleconnections. Thereafter, the collective dynamics of these interacting grid points can be studied using network measures, as defined in Section 3.3.

In must be noted that, throughout this thesis, we will construct climate networks using only a single climate variable. However, links based on statistical interdependencies between multivariate time series describing the dynamics of multiple observables recorded at the same locations/nodes can be defined using more general multilayer network approaches [147], which are not covered in this thesis.

In the following, we illustrate the general recipe for constructing climate networks, discuss the similarity measures used according to the type of the climate data, and discuss the effects of spatial embedding.

3.5.2. Construction and Analysis of Climate networks

Each geographical grid-point of the climate dataset has an associated time series or value at a particular time instant. The climate network is a representation of the interactions or interrelations I_{ij} between the grid points i and j, computed by using some synchronization (similarity) measure (Section 2.2.2) or physical laws (Figure 3.3). The climate network can then be constructed by thresholding the interaction matrix I_{ij}, such that only edges $\{i, j\}$ that satisfy $I_{ij} > \tau$ are considered as linked, where τ is the threshold.

The choice of the threshold τ, is made such that only connections that are statistically significant with respect to a reasonable test are maintained while the rest which do not meet the criterion are rejected. The significance of the similarity or interaction value for a given pair of time series is tested against randomly shuffled surrogates of the time series. Other methods for constructing surrogates such as Fourier surrogates and twin surrogates can also be considered. Furthermore, the choice of τ reflects a trade-off between the statistical significance of links and the richness of network structures revealed. In particular, it is desirable to include the long distance links with high similarity/interaction value which are responsible for many interesting and non-trivial features of climate networks, such as the small world property or hubs, and may potentially be associated with teleconnections [30]. When comparing the properties of different climate networks, e.g., time-evolving networks (Chapter 4, and Sections 5.3 and 6.5) or networks generated with different similarity measures (Section 5.3), we choose to fix the edge density ρ, whose choice is still guided on the basis of the above discussion, which results in different thresholds τ for the networks.

The adjacency matrix A_{ij} of the climate network can therefore be written using the Heaviside function $\Theta(x)$ as $A_{ij} = \Theta(|I_{ij}| - \tau) - \delta_{ij}$, where δ_{ij} is the Kronecker delta (subtracted in order to remove self-loops).

Following the construction of climate networks, the network architecture can be analysed using network measures, some of those that are used in this thesis are defined in Section 3.3.

3.5.3. Choice of similarity measure

As already discussed in Section 2.2.2, depending on the type of data, different synchronization measures may be used to measure the degree of synchronization between two time series, and hence to construct the corresponding functional network. Furthermore, depending on whether the climate observable is a scalar or vector, the quantity computed to estimate the interaction between the nodes may differ, and thus the climate network construction. For example, as discussed in Section 2.2.2, for node pairs with long continuous time series, Pearson's or Spearman's correlation may be used depending on whether the data is normally distributed or not. For short time series, Kendall's τ is a better choice. For event-like data, event synchronization may be used. However, to estimate instantaneous interactions, some physical law may be necessary to compute the interaction based on the instantaneous spatio-temporal values of the node pairs. In Table 3.1, the climate variables used in this thesis along with the characteristics of the data and the corresponding measure of interaction chosen to construct the climate network representation are listed.

TABLE 3.1.: Climate variable and the choice of measure of interaction for Climate network construction.

Climate variable	Data Characteristics	Measure of Interaction
Mean Sea Level Pressure	Continuous data, Normally distributed	Kendall's τ coefficient is used because analysis is performed for short time series (Section 4.2).
Relative vorticity	Continuous data	Biot-Savart Law to estimate directed instantaneous interactions (Section 4.3)
Extreme Precipitation	Event-like data	Event Synchronization (Section 5.2)
Outgoing long wave radiation	Continuous, non-Gaussian distribution	Spearman's rank correlation Coefficient (Sections 5.3 and 6.3)
Zonal and meridional components of wind	Continuous, non-Gaussian distribution	Spearman's rank correlation Coefficient (Sections 5.3 and 6.3)
Geopotential height	Continuous data	Spearman's rank correlation Coefficient (Section 6.3)
Forecast error of climate variables	Continuous data, non-Gaussian distribution	Spearman's rank correlation Coefficient (Chapter 6)

3.5.4. Correction due to Spatial embedding

The climate network analysis performed in this thesis is mostly limited for a certain region of the Earth's surface. The introduction of such spatial boundaries cuts the links that would connect the considered region with outside regions. This artificially reduces the degree of the nodes and the number of long links, and also influences the spatial patterns of any other network measure. Boundary effects depend on the distribution of link lengths and on the network measures themselves. As more links are cut for nodes closer to the boundaries than those deep inside the region, the degree of the nodes near to the boundaries has a stronger reduction compared to those in the interior. In the case of the clustering coefficient, which depends on topological paths of length three, it is seen that nodes along the boundaries tend to have a higher tendency to cluster, while for mean geographical distance the effects of boundaries become more complex (see Fig. 1 in Supporting Information of Boers et al. [50]).

In order to avoid any spurious conclusion arising solely from the effects of the spatial embedding [130], a correction procedure needs to be adopted [148] for the considered network measures, especially if the climate phenomena under investigation is a highly localized event and affects a much smaller area compared to that of the whole network, as in the case of detection of tropical cyclones (Section 4.2). The correction procedure is as follows:

- First, 1000 spatially embedded random networks (SERN) that preserve both the node positions in space and the link probability are constructed, depending on the spatial link lengths of the original network.

- Then, each of the considered network measures are computed for all the SERN surrogates.

- The boundary effects on the network measure is estimated by taking the average of that measure over the ensemble of surrogates.

- Finally, the corrected network measure is obtained by dividing the network measure of the original network by the corresponding average measure of the SERN surrogates.

As the corrected network measure gives the value of the network measure relative to the value expected from the spatial embedding, it is dimensionless. The SERN surrogates can also be used to test the significance of the network properties of the original climate network against that of the spatially embedded random networks.

In the subsequent part of the thesis, we apply the theoretical methods discussed in this part to study various climate phenomena from observation and model data.

Part II.

Applications

This second part of the thesis applies the concepts and techniques explained in Part I to real-world problems and data in the context of Earth system analysis. It focuses on application of complex network approaches to characterizing patterns of weather variability during the occurrence of cyclones and quantifying the stages of interaction between binary cyclones. Next, the intraseasonal variability of spatial synchronization patterns of Asian Summer Monsoon is investigated which yield insights into the underlying climate processes. The impact of the El Niño-Southern Oscillation on the spatial co-variability patterns of convection anomalies associated with the Asian Monsoon system is investigated. Finally, the applicability of complex networks to gain understanding of the origin of forecast errors, in the context of the Asian Summer Monsoon, is demonstrated.

4. Tropical Cyclones: Detection and Binary Cyclone Interaction

> *"'Tis the Whirlwind, that has driven out Zeus and is King now."*
>
> — Aristophanes, *The Clouds*

4.1. Introduction

Cyclones, also known as typhoons, hurricanes, etc., in different parts of the world, are amongst the most destructive natural hazards on Earth. Cyclones over tropical or subtropical waters are highly organized warm-core non-frontal vortical systems that lie within the synoptic scales and are formed as a result of significant convective activity. They are characterized by a low pressure center [149] that produces strong inward-spiraling winds. A tropical cyclone can cause severe widespread damage to land infrastructure, especially when it makes landfall along densely populated coastlines, disrupting human lives and even resulting in numerous casualties as it is accompanied by torrential rains and violent winds [150]. The increase in occurrence of severe destructive cyclones in the recent years has raised genuine concerns worldwide. Numerous studies have shown that this may be attributed to human-induced climate change which has led to an increase in sea surface temperature as well as the moisture carrying capacity of the atmosphere, therefore possibly causing rapid intensification of tropical cyclones, higher precipitation rates and increased probability of occurrence at higher latitudes [151]. Understanding the behavior of cyclones is thus of paramount interest to weather forecasters and policy makers.

As the climate network approach enables us to study the interaction structure of the climate system, it has been used to investigate the topological and dynamical evolution of several climate phenomena which occur over annual or seasonal time scales such as the El Niño-Southern Oscillation (ENSO) [43–45], Indian Summer Monsoon (ISM) [48, 49], South American Monsoon [152], etc. Such an understanding has aided the construction of prediction schemes, such as forecasting the magnitude and occurrence of ENSO [43–45], the onset and withdrawal of ISM [53] and of extreme floods in eastern Central Andes [52]. In this context, understanding the temporal evolution of

---

This chapter is based on the associated publications Gupta et al. [P2] and De et al. [P3], both licensed under a Creative Commons Attribution (CC BY) license. The sections in this chapter closely follow these publications.

network topology of the regional weather system during individual tropical cyclones is an important step towards the study of its predictability. However, one of the main challenges faced when applying the functional network framework to study individual cyclone events is that cyclones are highly localized extreme weather events with a short lifespan (7-10 days on an average), which makes it challenging to devise a suitable method to quantify their spatial patterns of variability. In terms of climate networks, it is difficult to describe the strength of the pairwise interactions between the nodes, instantaneously or averaged over a short time period. While some earlier works based on climate network approaches have attempted to investigate the spatial characteristics of extreme rainfall synchronicity [54, 153, 154], limited attention has been given towards understanding the temporal evolution of network topology of the regional weather system during individual tropical cyclones.

In order to understand whether the climate variability at weather time scales are organized in patterns, the topological and dynamical evolution of the regional weather conditions has to be studied. One possible way to approach the problem involves the construction of climate networks over rather small spatial regions (cyclone basins), which evolve in time according to the time scale of the extreme weather event, i.e., the underlying interaction structure of the meteorological fields cannot be considered as static (see Evolving networks in Section 3.4). Time-evolving complex networks have been used to investigate failure propagation in power-grids [155, 156], hierarchical structures in the brain [36, 157, 158], structural differences in the interconnectivity of the climate system between El Niño and La Niña conditions [140], early-warning of El Niño events [43–45], transition of regional connectivity during the South American Monsoon onset [152], and the multiscale nature of Australian Summer Monsoon development [159].

Characterizing the spatial organization pattern of the weather variability during the occurrence of an extreme weather event, such as a cyclone, is different from studying how this organization takes place. The former is the generalization of the problem to identify the track of a cyclone in the cyclone basin in a given time span, while the latter involves tracking the reorganization of the interaction structure at almost instantaneous time scales. Therefore, different methods of constructing evolving network are required in the two cases. In this chapter, we attempt to solve both these problems in the context of cyclones. In the first case, we use functional networks constructed over overlapping short-length sliding time windows to compare the spatial patterns of the various topological properties, such as degree and clustering coefficient (Section 3.3), and the spatial scales involved, for short time frames around the occurrence of individual cyclones (Section 4.2, [P2]). For the second case, we deal with the specific problem of interaction between two cyclones when in close proximity, known as the *Fujiwhara interaction*, which eventually leads to the cyclones completely merging into one cyclone. We study the evolution of the instantaneous vorticity field as a directed spatio-temporal network by estimating the induced velocity between the flow elements by means of the Biot-Savart law (Section 4.3, [P3]). The adoption of induced velocity network based on the Biot-Savart law has been successfully used to

study the turbulent flow dynamics [160], which we extend to investigate flow dynamics in cyclonic systems here. It must be mentioned here that this approach can also be used to study the dynamical changes in the vorticity field during the genesis of cyclones. In contrast to the previous network approach which depicts only statistical relationships, the induced velocity networks represent real physical links indicating the induction of velocity by one flow element on the others. We use directed network measures such as in-degree and out-degree (Section 3.3) to quantify the mutual interaction between two cyclones when they are in close vicinity. Finally, we summarize our findings in Section 4.4.

4.2. Detection of Tropical Cyclones

In this section, we study the topological and dynamical evolution of the regional weather conditions over a particular cyclone season. We characterize the spatial patterns of weather variability by quantifying the topological properties of the time-evolving functional networks constructed over overlapping short-length sliding time windows. Our analyses show that the regional system undergoes a characteristic spatial reorganization in the connectivity structure during a cyclone in such a way that the network measures are in close correspondence with the cyclone tracks. We also confirm that our inferences hold true for different cyclone basins irrespective of the differences in the complexity of their dynamics.

In Section 4.2.1, we list the employed datasets, explain our choices of the spatial and temporal resolutions, and then outline our methodology. We then discuss our results in Section 4.2.2.

4.2.1. Data and Methodology

Data

In this study, we use the state-of-the-art ERA5 reanalysis data for 3-hourly mean sea level pressure (MSLP) [161, 162] over the sea. As the cyclones can undergo both a rapid intensification and weakening within a span of a few hours, the use of the 3-hourly temporal resolution ensures a high enough temporal auto-correlation. Moreover, as the cyclones are short-lived, with a typical lifespan of $\sim$ 3-10 days, the sub-daily resolution adds more time points to the period in consideration. MSLP exhibits stronger variability at higher frequencies than sea surface temperatures (SSTs) or surface air temperatures (SATs), which enhances its sensitivity towards cyclone signals and thereby increases the possibility of capturing cyclones in MSLP networks for the duration of their lifetime. The entire availability period of the dataset is from 1950 to present, available at hourly resolution. The daily climatology is computed as the mean of the daily MSLP values over a period of 40 years (1979-2018). We remove the seasonal cycle from all time series of the dataset, by calculating the anomaly time series, i.e., subtracting the daily climatology of each day from all the hours of that day.

FIGURE 4.1.: Evolving networks for the Sep–Oct–Nov–Dec season. Networks are constructed over a time window of 10 days. Successive windows have 9 days of overlap. (Taken from Gupta et al. [P2])

The spatial resolution of the MSLP dataset plays a significant role in the identification and tracking of cyclones [163]; the probability of detecting cyclones increases with an increase in the resolution of the dataset. We use a high spatial resolution of $0.75° \times 0.75°$, which proves to be sufficient for our analysis. As the MSLP large-scale patterns are not so well determined by the land-sea boundary, and most cyclones originate over the sea and dissipate shortly after landfall, we analyse the MSLP spatiotemporal dataset over the sea only. Furthermore, it should be noted that the MSLP over land is estimated by extrapolation of surface pressure in the models used in the ERA5 reanalysis and therefore may introduce artificial inconsistencies if compared to sea values. Our results in Section 4.2.2 show that the omission of land points does not affect the analysis of land-crossing cyclones (Figures 4.2 and 4.4), rather the network measures over land can get affected by the orography (Figure A.1).

We generate cyclone tracks from the Best Tracks data available over the north Indian Ocean basin (entire availability period of 1982-2020, Indian Meteorological Department) and the north Atlantic Ocean basin (entire availability period of 1851-2019, HURDAT2, NOAA [164]) to compare them with the results obtained from our analyses.

Time-Evolving Functional networks

In accordance with the idea of evolving networks (Section 3.4.4), we divide the reanalysis data into overlapping short time windows and construct a climate network for each of these windows. The length of the time window is taken to be 10 days, which is of similar time scale as that of the typical lifespan of cyclones, to capture the effect of cyclones on the dynamical and structural evolution of the network better. The successive time windows have 9 days of overlap, i.e., the climate network evolves in daily steps (see Figure 4.1). It should be noted, that the obtained results in Section 4.2.2 do not have a strong dependence on the chosen parameters – networks constructed for time windows spanning up to 15 days yielded similar results. Following the method of reconstruction of evolving climate networks, every node or spatial grid point of the cyclone basin is associated with a 3-hourly 10-day anomaly time series, i.e., 80 time points at each grid point.

We employ the functional network representation (Sections 3.4.2 and 3.5.2) of the spatio-temporal climate dataset to encode the strong statistical linkages between every pair of involved time series. We choose Kendall's τ coefficient computed at zero lag as the similarity measure for the network construction as it is known to perform better than other measures such as Pearson's correlation coefficient for short time series (Sections 2.2.2 and 3.5.3, and Table 3.1). It should be noted that a positive time lag may be used for cyclones with a slower translation speed, such as those in the Atlantic Ocean, in which case the information transfer cannot be assumed to be instantaneous. We only take into account correlation values that are statistically significant at a confidence level of 0.05 and set all other values to zero. The cross correlation matrix so obtained is symmetric.

The time-varying climate network adjacency matrices $A(t)$ for every time window are constructed by considering the strongest 5% of the significant correlations to define the links (Sections 3.2 and 3.5.2). Among the thresholds ranging from 80th to 99.5^{th} percentile of the correlation matrix, 95th percentile was found to be the optimal choice for all our networks. The evolving networks so obtained are undirected and unweighted.

Analysis of network topology

We analyze the time variation of the topology of the interaction patterns in the regional climate system of the cyclone basin by using global and local network measures to characterize the climate networks [30, 50, 145, 146, 165]. Several commonly used network measures [39, 41] are adopted, namely, the degree centrality (Eq. (3.4)), the mean geographical link distance (Eq. (3.10)), and the local and global clustering coefficients (Eq. (3.7)-(3.9)) which are explained in Section 3.3. The degree centrality k_i enables us to identify the important regions based on their number of connections. Regions with higher connectivity have larger values of k, while regions of low k values are indicative of a small-scale atmospheric process and are often related to large topographic barriers [48, 49, 51]. The mean geographical distance $\mathcal{L}_i$ provides the information of the spatial length scale involved with the cyclones. The local and global clustering coefficients indicates whether localized structures of high connectivity are present which indicates regions of spatial continuity in the network.

Necessity of correcting boundary effects

As cyclones are highly localized extreme weather events, the networks are constructed over areas of cyclone formation, i.e., cyclone basins, instead of taking the full globe into consideration to enable a detailed understanding of the regional weather system. As mentioned earlier, we only consider grid points at sea. Therefore, in addition to the boundaries of the cyclone basin, the coastlines also spatially confines the regional networks. However, the introduction of such spatial boundaries affect the network measures [148] (Section 3.5.4) as the effect of spatial embedding now becomes anisotropic on the network. Furthermore, the area covered by the nodes affected by
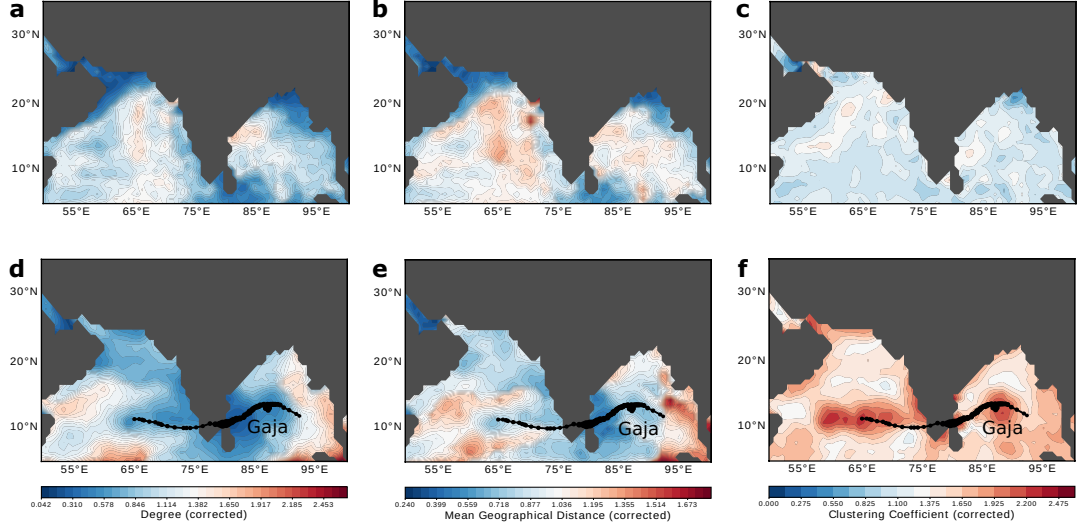
FIGURE 4.2.: Comparison of degree (a) and (d), mean geographical distance (b) and (e) and local clustering coefficient (c) and (f) fields before and during Very Severe Tropical Cyclone Gaja (Nov 10-19, 2018). Figures in a given column have the same colour scale. (a)–(c) shows the network measures before the cyclone for the period Oct 29-Nov 7, 2018. (d)–(f) shows the network measures for the period Nov 10-19, 2018 during the cyclone. The cyclone tracks are represented by solid black circles whose sizes are scaled according to the cyclone intensity. (Taken from Gupta et al. [P2])

the cyclone is much smaller than the area of the cyclone basin, and therefore the detection of cyclones are more prone to boundary errors. In view of the above argument, correction of effects due to spatial embedding is important. We therefore apply the correction procedure described in Section 3.5.4 based on Rheinwalt et al. [148] on the network measures before interpreting the results in Section 4.2.2.

4.2.2. Results and Discussion

We use the network-based methodology, described above, to study some recent cyclones in the North Indian Ocean (NIO) basin, extending from 49.5°E to 100°E and from 34.5°N to 4.5°N. Due to the annual cycle of the background vertical shear of the horizontal winds [166, 167], the NIO basin has a bimodal cyclone season [168–170] — pre-monsoon (March-April-May; MAM) and post-monsoon (September to December; SOND) — both of which are during the monsoon transition periods. The cyclone frequency in the post-monsoon season is comparatively higher than that in the pre-monsoon season mainly because of the difference in the mean relative humidity between the two seasons [169]. We concentrate on the cyclones which have occurred in the SOND season over the last decade (2009–2018).

We show a comparison between the network measures before and during the Very Severe Cyclonic Storm (VSCS) Gaja, which formed in the Bay of Bengal and later
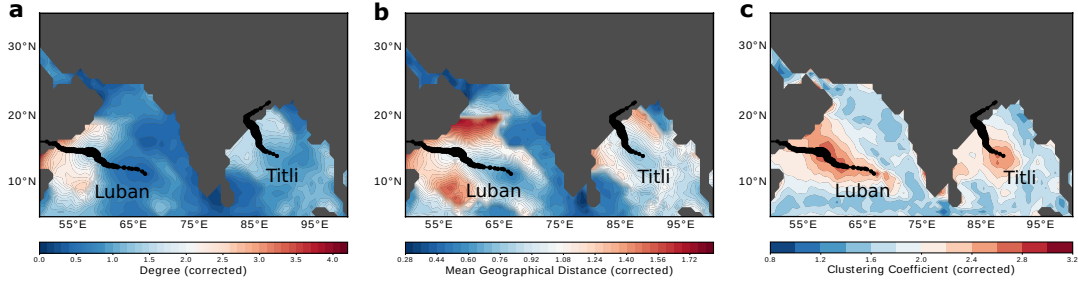
FIGURE 4.3.: (a) Degree, (b) mean geographical distance and (c) local clustering coefficient fields for network constructed over the period Oct 2-11, 2018, during the Very Severe Tropical Cyclones Luban (Oct 6-15, 2018) in the Arabian Sea and Titli (Oct 8-12, 2018) in the Bay of Bengal. The cyclone tracks are represented by solid black circles whose sizes are scaled according to the cyclone intensity.(Taken from Gupta et al. [P2])

crossed the Indian peninsula into the Arabian Sea, in Figure 4.2. There are no definite patterns in the spatial distributions of the degree, mean geographical distance and local clustering coefficient fields in the absence of a cyclone, as seen, in Figures 4.2a–c, from the plots of these fields for the network corresponding to the period Oct 29–Nov 7, 2018. However, the networks measures undergo a spatial organization and exhibit definite patterns, as shown in Figures 4.2d–f, during the period Nov 10-19, 2018, when the cyclone had occurred. Along the cyclone track, the nodes have lower degree k_i than those in the surrounding regions (Figure 4.2d), and the spatial distribution of the mean geographical distance, $\mathscr{L}_i$, is very similar to that of k_i (Figure 4.2e). This is because the area affected by cyclones is much less than that of the surrounding regions since cyclones are highly localized mesoscale convective events. Hence, the nodes along the cyclone track are only connected to each other and not to the other nodes, as a result of which, the both k_i and $\mathscr{L}_i$ in the affected regions are much less than those in the unaffected regions. While there are shorter links along the cyclone tracks (lower values of $\mathscr{L}_i$), longer links connect surrounding regions separated by the track, as seen from their higher $\mathscr{L}_i$ values. Therefore, the cyclone track separates a region of high connectivity into two. Note that the mean geographical distances in Figures 4.2-4.5 and Figure 4.7 are dimensionless quantities because of the correction of bias due to spatial embedding, as mentioned in Section 4.2.1.

On the contrary, the local clustering coefficient field, C_i, (Figure 4.2f) is relatively high in the localised region around the cyclone track, implying spatial continuity in the network along the track. This leads us to infer that the nodes along the track form a small, tightly-knit group wherein most nodes have connections with only those which belong to the group. Naturally, such a group is detached from the rest of the network and tends to behave as an isolated sub-network. This explains why the nodes along the cyclone track have high values of C_i and lower values of k_i and $\mathscr{L}_i$, thereby distinguishing the dynamics of a cyclone from those of the surroundings.

We show the spatial patterns of the network measures for a few other cyclones of the NIO basin — VSCS Luban and Titli in Oct 2018 (Figure 4.3), VSCS Vardah
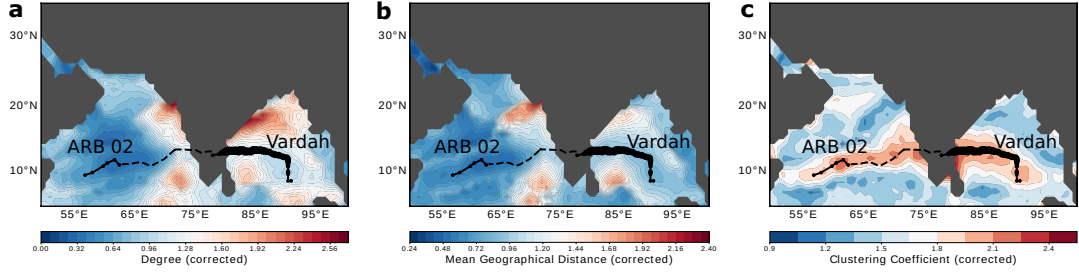
FIGURE 4.4.: (a) Degree, (b) mean geographical distance and (c) local clustering coefficient fields for network constructed over the period Dec 9-18, 2016, during the Very Severe Cyclonic Storm Vardah (Dec 6-13, 2016) in the Bay of Bengal which crossed the Indian peninsula and formed depression ARB 02 (Dec 17-18, 2016). The cyclone tracks are represented by solid black circles whose sizes are scaled according to the cyclone intensity. (Taken from Gupta et al. [P2])
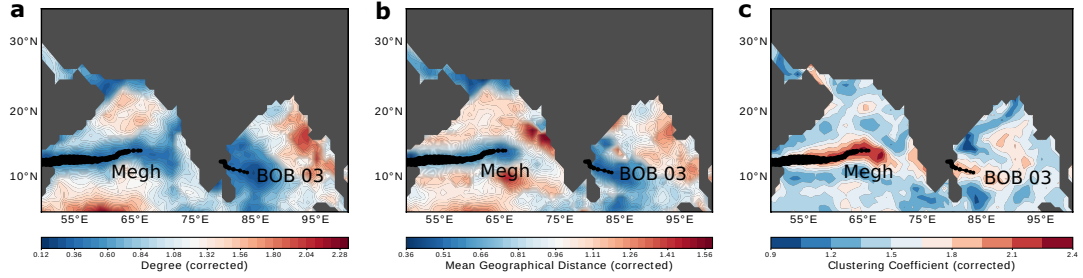
FIGURE 4.5.: (a) Degree, (b) mean geographical distance and (c) local clustering coefficient fields for network constructed over the period Nov 3-12, 2015, during the Extremely Severe Tropical Cyclones Megh (Nov 5-10, 2015) in the Arabian Sea and Deep Depression BOB 03 (Nov 8-10, 2015). The cyclone tracks are represented by solid black circles whose sizes are scaled according to the cyclone intensity. (Taken from Gupta et al. [P2])

in Dec 2016 which crossed the Indian peninsula and formed the Depression ARB 02 (Figure 4.4), and the Extremely Severe Cyclonic Storm Megh and Deep Depression BOB 03 in Nov 2015 (Figure 4.5). In all these cases, we observe the similar characteristic behaviour described above, i.e., lower k_i and $\mathcal{L}_i$ with higher C_i along the respective cyclone tracks, thereby providing further support to our inference in the previous paragraphs.

A time series of global clustering coefficients, C, of the successive networks can be constructed for each of the cyclone seasons. In Figure 4.6, we plot the values of C for the 2018 SOND season of the NIO basin versus the dates corresponding to the middle of the respective network periods. We observe that at least one cyclone event is associated with those networks having relatively large values of C. This implies that networks have greater transitivity during cyclones because of the presence of localised structures of high connectivity. Note that the choice of plotting C versus the middle date of a network period brings in a temporal tolerance of ± 5 days since

each network spans over 10 days.
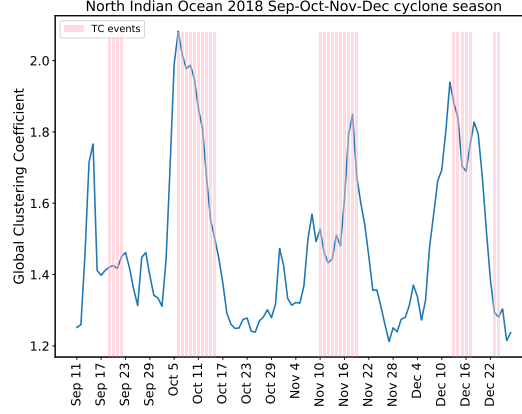
FIGURE 4.6.: Global Clustering Coefficient (blue) for evolving networks of 2018 NIO cyclone post-monsoon (Sep–Oct–Nov–Dec) season plotted against the date corresponding to the middle day of the network period. Networks containing cyclone events show high global clustering coefficients. (Taken from Gupta et al. [P2])

Furthermore, we confirm our findings from analyses of cyclones in the North Atlantic Ocean (NAO) cyclone basin during the period August-September, when the hurricane season is at its peak in the basin, using our network-based methodology. The region under consideration extends from 10°N to 42°N and 42°W to 100°W. With the same choice of network window length and time lag as for the cyclones in the NIO basin, we observe similar spatial patterns of degree, mean geographical distance, and local clustering coefficient for Hurricane Irma which took place during the period Aug 30–Sep 13, 2017 (see Figure 4.7). Hence, a key observation is that although the basin properties (e.g., sea surface temperature) of the NIO and NAO basins are different, similar topological evolution of the structure of the underlying network occurs during a cyclone. Since Atlantic hurricanes typically have a longer lifespan and are slow-moving, one can opt to choose a longer network period (> 10 days) and a positive time lag while constructing the networks.

Our topological study of evolving climate networks in a cyclone basin shows that the network connectivity structure goes through a specific rearrangement during a cyclone. Although there will be a rearrangement in the connectivity structure of the network during any low pressure system, the above observed signatures are strongest for cyclones. Since the degree and mean geographical distance are proportional to the size of the low pressure system, for larger low pressure systems such as the monsoon trough (see Figure 9 in Stolbova et al. [49]), these measures would have comparable values with the other regions. Hence, the above observations cannot be uniquely associated with any general low pressure system formation.
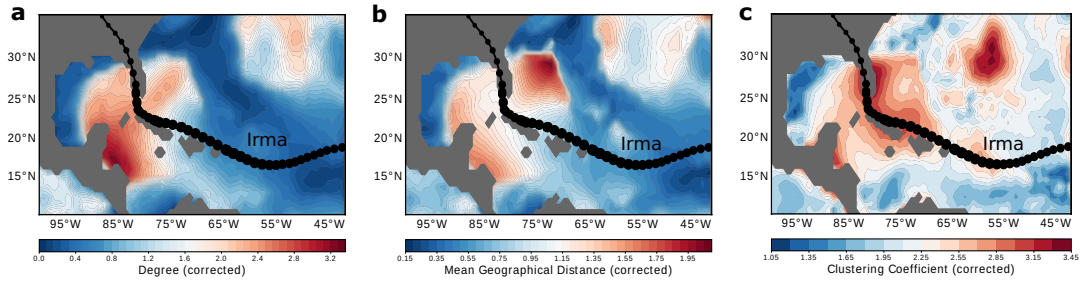
FIGURE 4.7.: (a) Degree, (b) mean geographical distance and (c) local clustering coefficient fields for network constructed over the period Sep 1-10, 2017, during Hurricane Irma (Aug 30–Sep 13, 2017) in the North Atlantic Ocean. The cyclone tracks are represented by solid black circles whose sizes are scaled according to the cyclone intensity. (Taken from [P2])

4.3. Binary Interaction of Cyclones leading to Complete Merger

In some very active cyclone basins, such as the northwestern Pacific and Atlantic, multiple cyclone systems can be formed simultaneously. The increase in the occurrence of severe cyclones in the recent years due to global warming has also increased the probability of occurrence of such multiple cyclone events in the Indian Ocean [171]. Although rare, two cyclones can come within close proximity and interact, beginning an intense dance about their common center. A number of such weather events have been previously recorded [172–174]. Known as the *Fujiwhara interaction*, this often alters the tracks of the cyclones, making them difficult to forecast [175–178]. Inaccuracies in predicting cyclone tracks increase the threat to life and property due to unpreparedness caused by misinformation and the lack of early warning. For instance, unforeseen heavy rainfall occurred in Taiwan, and the same region of the Luzan Island of the Philippines experienced landfall of typhoon Parma thrice due to its interaction with another typhoon Melor in October 2009, causing significant fatalities and economical losses [179]. Although very rare, the binary interaction may lead to a re-strengthening of the cyclone, as in the case of Category 3 severe tropical cyclone Seroja in April 2021 due to its complete merger with Odette [171]. Interaction of a cyclone with other cyclonic vortices may also prolong its life span, e.g., the Super Typhoon Noru in July 2017 lasted for 19 days due to its successive dual vortex direct and indirect interactions with typhoons Kulap, Haiting and Nesat [180], and the prolonged southward trajectory of Seroja was highly unusual compared to cyclones of similar intensity in the past 5000 years [171, 181].

Several studies [174, 175, 182, 183] based on observational data found that there were some notable exceptions to binary cyclone interaction events that do not follow the classical Fujiwhara model of cyclone merger. The presence of large-scale clockwise circulation patterns masks the Fujiwhara effect, sometimes even at separation distances where the Fujiwhara forces are quite strong. Further, large-scale circulation due to the presence of subtropical high or monsoon depression [183–186] and the

presence of multiple weak cyclonically-rotating meso-vortices [178] pose significant challenges towards cyclone track forecasts. Till date, a complete understanding and incorporation of the Fujiwhara effect in numerical weather prediction models to improve cyclone forecasts have not been achieved. Hence, understanding cases of binary cyclone interaction remains highly relevant.

A complete merger (CM) of two cyclones involves the transfer of energy and vorticity across the different scales in turbulent flow [187]. Previous studies [187–189] based on theoretical calculations have shown that the diffusion of vorticity from the inner core region to the inner and outer recirculation regions in a system of co-rotating vortices is the reason for their merging. However, such inner- and inter-layer fluid exchanges are not confirmed in real-world binary cyclone interactions. Several numerical and analytical studies on the interactions of binary cyclones have been attempted in an effort to understand both two-dimensional [172, 173, 179, 187] and three-dimensional dynamics [190, 191] of the CM phenomena. DeMaria and Chan [191] demonstrated that the mutual attraction can be explained using vorticity advection alone, and is strongly dependent on the initial wind profile of the vortices. A number of studies [175, 192–196] underlined the significant role of the separation distance, finding that merging occurs when the sizes of the vortex cores of co-rotating vortices increase beyond a critical fraction of the separation distance due to viscous diffusion. Further, several dissipative and convective stages [189, 197] are identified based on the separation distance in the vortex merging process. However, there has not been much observation or reanalysis data-based investigation on the dynamics of CM to compare with these model-based findings. This is partly due to the paucity of the occurrence of such merging events in nature.

In this section, we suggest an innovative framework by combining the data-driven complex network approach with elements of fluid dynamics to explore the evolution of the interactions in the vorticity field between binary cyclones when in close proximity. We construct time-evolving directed network representation of the instantaneous relative vorticity field in which we estimate the instantaneous interaction between flow elements by computing the induced velocity using the Biot-Savart law [160]. The results show that as the two cyclones approach each other, the ensuing changes in the network topology can be used to classify the complete merging process into several interaction stages.

We select two recent examples of binary cyclone systems for our study which eventually lead to a CM event – Noru-Kulap (during Jul 23-26, 2017) [180, 198] occurring in the Northern Hemisphere, and Seroja-Odette (Apr 5-10, 2021) in the Southern Hemisphere. Due to its Fujiwhara interaction with Kulap and indirect interactions with other cyclone systems, the Category 4 Super Typhoon Noru which was the second most intense tropical cyclone of the Northwestern Pacific Ocean basin in 2017, became the third longest-lasting cyclone on record in the Northwest Pacific Ocean in 2017 [180]. Noru brought torrential rainfall to southern and western parts of Japan that triggered widespread flooding and caused huge economic losses [199]. Similarly, following the CM event of the severe tropical cyclone Seroja with the tropical storm
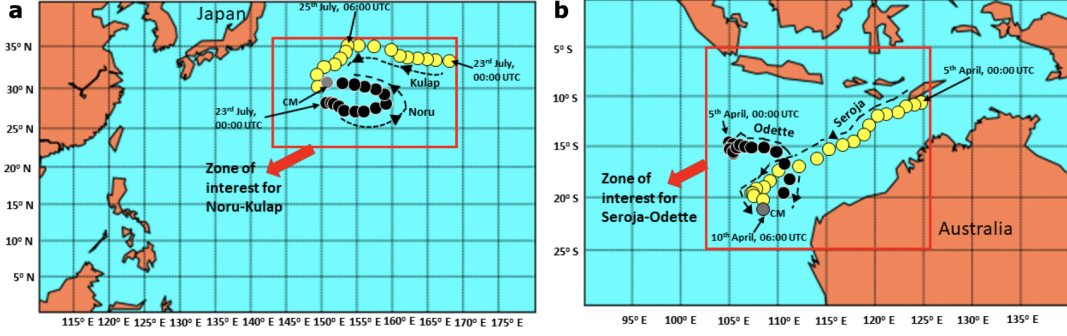
FIGURE 4.8.: The regions of interest considered for the network for the analysis of (a) Noru-Kulap and (b) Seroja-Odette interactions. In the case of Noru-Kulap interaction, the area extends from 143°E to 169.5°E and from 23.5°N to 35.5°N. For the Seroja-Odette interaction, the region of interest extends from 102°E to 125.5°E and from 5°S to 25°S. The trajectories of the cyclones are also shown to justify the selection of the spatial domain. (Taken from De et al. [P3])

Odette, the merged cyclone Seroja was further strengthened and steered southward towards Australia. Thereafter, it made landfall as a Category 3 severe tropical cyclone on the west coast of Western Australia causing significant damage.

This section is organized as follows. In Section 4.3.1, a detailed description is provided about the source of data and the method of the construction of the network, which is used in the present study. In Section 4.3.2, first a topological analysis of the evolving networks is performed to understand the temporal evolution of vortical interactions between the two converging cyclones (Section 4.3.2), and then we find that the transitions exhibited by the network measures enable us to classify the stages of the mutual interaction between the two cyclones which eventually lead to their merging (Section 4.3.2).

4.3.1. Data and Methodology

Data

We use the relative vorticity (ω) data obtained from the state-of-the-art ERA5 re-analysis dataset [162] to understand the interaction dynamics between two co-rotating cyclones. Relative vorticity at a particular height above the sea level is defined as the rotation of air about a vertical axis, relative to a fixed point on the Earth's surface and calculated as $\omega = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$, where, u and v corresponds to the zonal and meridional components of the wind velocity at that height, respectively.

Relative vorticity is reported to be more suitable than the MSLP field for capturing the local features in the evolution of cyclones. This because the features of the small and weak circulations, e.g., that during the onset of a cyclone, are not adequately represented in MSLP field as compared to that in relative vorticity field at 850 hPa [200, 201]. Moreover, use of relative vorticity at heights higher than the sea

level instead of MSLP will reduce errors when including land points as explained in Section 4.2.1. Furthermore, large-scale relative vorticity at lower atmospheric levels (500-850 hPa), are known to significantly affect cyclones [202–204] and influence their relative motion in the presence of another cyclone [182]. Many previous studies [172, 179, 190, 205] on binary cyclone interaction found it difficult to correctly incorporate these large-scale circulations in cyclone models, leading to erroneous predictions of cyclone tracks. The use of relative vorticity from reanalysis data ensures the inclusion of these large scale wind circulations.

As already mentioned in Section 4.2.1 that the probability of detecting cyclones improves with increase in spatial resolution [163], it is important to choose a high spatial resolution for the data. Furthermore, the relative vorticity, being a wind-based field, is sensitive to the spatial resolution of the data set [201]. Since here we are not interested in detecting cyclone tracks but instead want to study the binary interaction between cyclones, we choose a higher spatial resolution of $0.5^o \times 0.5^o$ for this part of the study compared to the previous analysis (Section 4.2.1). The spatial domain is chosen in a manner that ensures the elimination of any other neighbouring weaker cyclonic or anticyclonic vortices apart from the considered cyclone pair. So, inherently, we have made the assumption that the cyclone pair is not affected by the climate behavior outside the selected spatial region. For the analysis of the Noru-Kulap interaction, the spatial region of interest extends from 143° E to 169.5° E and from 23.5° N to 35.5° N (Figure 4.8a). Similarly, in the case of the Seroja-Odette interaction, the spatial region of interest extends from 102° E to 125.5° E and from 5° S to 25° S (Figure 4.8b). Furthermore, in order to study the rapid intensification and weakening of the cyclones, and the changes in their mutual interactions, we use a 3-hourly temporal resolution for the relative vorticity data set, similar to that chosen for detecting cyclone tracks in Section 4.2.1 [P2, 206].

We perform our analyses to obtain the interaction structure of the two-dimensional relative vorticity field at the lower tropospheric level of 850 hPa, as commonly used for cyclone forecasts [207, 208]. Vorticity at 850 hPa has a stronger magnitude compared to vorticity at near surface heights (1000 hPa), especially for weaker circulations and therefore, is more robust when representing the strong upward motion of air. Hence, the 850 hPa relative vorticity field exhibits better continuity in the course of cyclone evolution [200] which is essential to deal with a CM event of two cyclones. Moreover, weaker cyclones have a shallow-lower tropospheric vertical depth (850-500 hPa) while only the most intense cyclonic systems move with a deeper layer flow (850-200 hPa) [209, 210] which should be taken into account for producing optimal forecasts of cyclone tracks with the lowest mean forecast errors [211]. Therefore, we also investigate the evolution of the network connectivity structure for other higher tropospheric levels (650 hPa and 700 hPa) such that it includes most cyclones, which not only allows us to verify the consistency of our results, but also to identify the transitions in the interaction structure of the binary cyclone system in the three-dimensional column of the atmosphere.

Instantaneous Time-evolving Directed networks
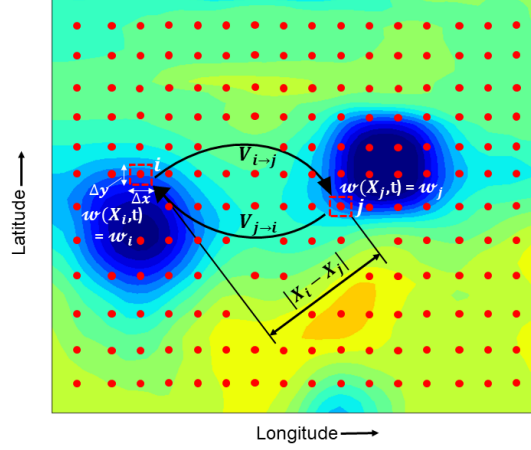
FIGURE 4.9.: Schematic illustration of the method of construction of vorticity network for a binary cyclone system at a given time t. The solid red circles in the spatial domain represent the grid points or nodes of the underlying reanalysis data set. The velocity, induced by the flow element at node i on node j is shown in terms of $V_{i \to j}$. ω_i and ω_j represent the relative vorticities of the ith and j^{th} flow elements, respectively. The dashed square boxes denote the sizes of the fluid elements at the ith and j^{th} grid points. (Taken from De et al. [P3])

We use the climate network framework described in Section 3.5. The network-based approach is used to study the two-dimensional vortical interactions in binary cyclone systems at a particular geopotential height. The spatial grid point of the data set represent the nodes (Figure 4.9). However, here the links between two nodes represent the instantaneous interaction between the fluid elements at the corresponding grid points [160, 212, 213] computed using the Biot-Savart law (Table 3.1). The Biot-Savart law is widely used to calculate the magnetic field induced by a current-carrying wire in electromagnetic theory [214, 215] and aerodynamic forces exerted by the flow on complex geometries such as wings using vortex panel methods [216]. Strictly, the Biot-Savart law is only applicable for incompressible flows, that is, when the velocity field is divergence-free [160, 213]. The velocity field associated with cyclonic flows are not divergence-free, especially close to the center of the cyclone. Therefore, the value of induced velocity found using this method is not necessarily accurate at all points but is adequate to compare the strength of connections between different spatial locations. Furthermore, to a good approximation, atmospheric flows can be considered as incompressible. The assumption of incompressibility have been used in earlier studies in the modelling of cyclones [217, 218].

Following the Biot-Savart law, we estimate the magnitude of the velocity induced by the vorticity of a flow element at the ith grid point on another flow element at the

j^{th} grid point, $V_{i \to j}$ (Figure 4.9) [214] as,

$$V_{i \to j} = \frac{\mid \gamma_i \mid}{2\pi \mid X_i - X_j \mid} \tag{4.1}$$

where, X_i and X_j are the spatial location of the ith and j^{th} grid point respectively. We take the absolute value of the circulation ($\gamma_i = \omega(X_i)\Delta x \Delta y$) of the flow element at the grid point (node) i as mentioned in Taira et al. [160]. Note that in this case, the interaction is not computed between the relative vorticity time series of nodes i and j, but between the relative vorticity values at the two nodes at a given time. Treating the spatial domain as planar (2D), we compute the Euclidean distance between the ith and j^{th} nodes represented by $\mid X_i - X_j \mid$. If the number of grid points (nodes) in the flow domain is N, then the size of the induced velocity matrix is $N \times N$. The velocity induced by the flow element by the ith node on the element at the jth node ($V_{i \to j}$) is different from that induced by the element at the jth node on the element at the ith node ($V_{j \to i}$), and therefore the matrix is asymmetric.

We choose the threshold corresponding to an edge density $\rho = 0.05$, which translates into considering only the highest 5% of the induced velocities to define the links in our network. This choice of threshold τ is found to be the optimum choice to retain connections corresponding to both cyclones, ensuring that the network is not too dense. Then, the adjacency matrix $\mathrm{A}(t)$ at a given time t is constructed as explained in Section 3.5.2. The velocity induced by a flow element on itself is considered to be zero, i.e., self connections are neglected (Eq. 4.2). Thus, the unweighted directed network at a given time whose adjacency matrix A_{ij} is represented as,

$$A_{ij} = \begin{cases} 1, & \text{if } i \neq j \text{ and } V_{i \to j} > \tau \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

In this manner, we construct a time-varying spatial network from the vorticity field at every time instant to understand the evolution of the binary cyclone interaction.

It is important to emphasize the difference between this method of constructing the networks and the correlation-based sliding time window network-approach used in Section 4.2 [P2]. The networks in Section 4.2 are time-averaged networks constructed over a period of 10 days which are unable to capture the evolution of a mutual interacting binary cyclone system which varies over hourly to daily time scales. Therefore, instantaneous time-varying vorticity networks are a better alternative to study the formation of cyclone and study their interaction with other cyclones. However, the network-approach using Biot-Savart law is only limited to vorticity data, and unlike correlation networks has not yet been generalized to other climate variables.

Analysis of network topology

We use the network measure *degree* to measure the centrality of the nodes in the interacting flow domain. As our instantaneous vorticity network is a directed net-

work, we distinguish the number of incoming and outgoing links, to and from a node respectively, in terms of its *in-degree* (k_i^{in}) and *out-degree* (k_i^{out}), both of which have been defined in Eq. (3.5) and (3.6) of Section 3.3 respectively. The in-degree k_i^{in} is used to describe the impact of the induced velocities of the neighboring nodes at the ith node in the interaction domain. On the other hand, the out-degree k_i^{out} can identify the strong vortices which induce velocities over other nodes in the interaction domain.

Separation distance between cyclones

The separation distance is a commonly used metric to classify the interaction stages of binary cyclones [183, 219] and the vortex merging process [189]. In the present study, the position of each cyclone is obtained by tracking the geographical latitude and longitude of the center as provided by Weather Underground's Online database [220]. We use the Haversine formula [127] for spherical Earth projected on to a plane to calculate the separation distance (d) between two cyclones.

4.3.2. Results and Discussion

We first outline the evolution of the network connectivity structures of the two binary cyclone systems under consideration and relate it to the variations observed in their respective relative vorticity fields. Following that, the network-based parameters are used to characterize the different stages of the merging process.

Degree analysis

1. ***Noru-Kulap system***
   We study the interaction between the binary cyclones, Noru and Kulap, and the effect of the neighbouring air flow patterns in the northwest Pacific [180] Ocean basin during July 2017. In the colour plots of relative vorticity, ω, at 850 hPa in Figures 4.10a1-d1 and Figures 4.11a1-d1, large positive values of ω imply the strong counter-clockwise rotation of winds, typical for Northern Hemisphere cyclones [221]. Hence, the two distinct blobs of positive ω in each of these figures denote the two cyclones — Noru marked with the letter 'N' and Odette marked with 'O'. During the period Jul 23-24, 2017, Kulap moves slightly towards west while Noru has a slight towards eastward movement [198]. Next, during Jul 25-26, their direction of motion changes significantly leading to a decrease in their separation distance, d, and finally leading to a complete merger.

   We first discuss the binary cyclone interaction during the period Jul 23-24, as shown in Figure 4.10. On Jul 23, Kulap and Noru are far away from each other ($d \sim 1510$ km). The network connectivity structure, at this time, shows a higher in-degree, k_i^{in}, of the nodes near the centre of Kulap compared to that of Noru, implying the greater the dominance of vortical influence from the
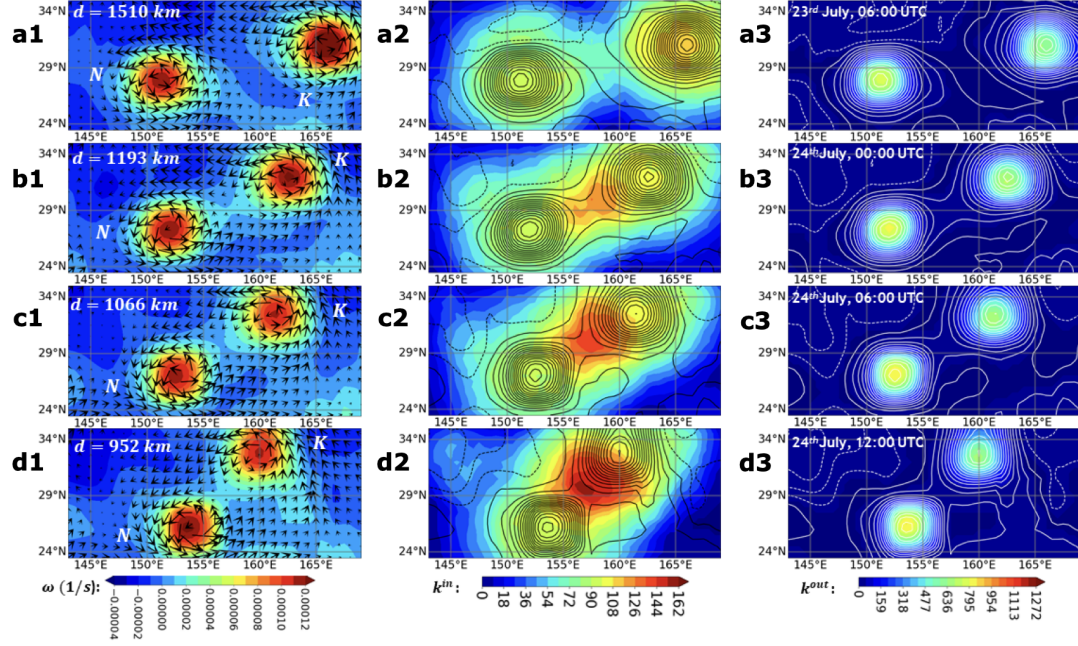
FIGURE 4.10.: The distributions of ω [(a1)-(d1)], k_i^{in} [(a2)-(d2)], and k_i^{out} [(a3)-(d3)] during the interaction of Noru (N) and Kulap (K) at the geopotential height of 850 hPa from Jul 23-24, 2017. The snapshots shown here are for the time steps, Jul 23, 06:00 UTC [(a1), (a2), (a3)], Jul 24, 00:00 UTC [(b1), (b2), (b3)], Jul 24, 06:00 UTC [(c1), (c2), (c3)] and Jul 24, 12:00 UTC [(d1), (d2), (d3)]; the arrows in (a1)-(d1) are the wind velocity vectors. The vorticity contours corresponding to (a1)-(d1) are overlaid on the distribution of k_i^{in} and k_i^{out} for a better understanding of the changes during the Noru-Kulap interaction; positive vorticity contours are represented by the solid lines while dotted lines indicate the negative vorticity contours. During the interaction, k_i^{in} increases significantly between two cyclones [(a2)-(d2)] as the cyclones come closer, while k_i^{out} is high over both the cyclones. The higher value of k_i^{out} at the center of Noru [(a3)-d(3)] implies its stronger impact on the neighboring nodes. (Taken from De et al. [P3])

surrounding regions on the nodes near Kulap (Figure 4.10a2). The nodes in the region between the cyclones have low values of k_i^{in} (Figure 4.10a2). As the cyclones rotate about each other, the values of k_i^{in} in the region between the two cyclones gradually increase (Figures 4.10b2-d2). This may be attributed to the inter-layer vorticity advection between the cyclones [187, 188]. Also, during this interaction phase, the nodes in the outer layers of Kulap facing Noru have higher k_i^{in} than those in the outer layers of Noru which face Kulap, thereby indicating the higher influence of Noru on Kulap. The spatial distribution of the out-degree, k_i^{out}, of the nodes in each of these stages seems to resemble that of relative vorticity ω (compare Figures 4.10a1-d1 with Figures 4.10a3-d3). At all times, the central region of Noru has higher k_i^{out} than that of Kulap. For both cyclones, the values of k_i^{out} decrease significantly with distance from
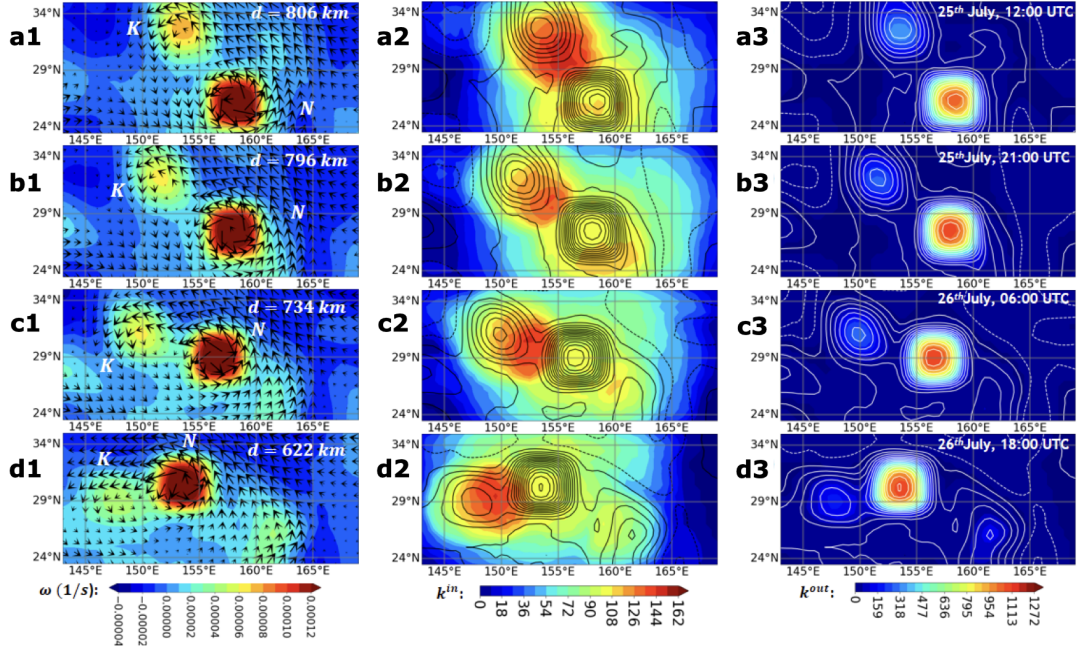
FIGURE 4.11.: The distributions of ω [(a1)-(d1)], k_i^{in} [(a2)-(d2)], and k_i^{out} [(a3)-(d3)] during the interaction of Noru (N) and Kulap (K) at a geopotential height of 850 hPa from Jul 25-26, 2017, just before they merge completely. The snapshots shown here are for the time steps, Jul 25, 12:00 UTC [(a1), (a2), (a3)], Jul 25, 21:00 UTC [(b1), (b2), (b3)], Jul 26, 06:00 UTC [(c1), (c2), (c3)], and Jul 26, 18:00 UTC [(d1), (d2), (d3)]. The vorticity contours corresponding to (a1)-(d1) are overlaid on the distribution of k_i^{in} and k_i^{out} and the wind velocity vectors are represented in (a1)-(d1) by arrows. The cyclones are at a closer proximity to each other [(a1)-(d1)]. k_i^{in} in the region between the cyclone reduces as the merging process begins [(a2)-(d2)]; k_i^{out} is higher over both the cyclones but its magnitude at the centre of Noru increases while that at Kulap decreases most likely due to vorticity advection [(a3)-(d3)]. (Taken from De et al. [P3])

their respective centers with the non-cyclone nodes having negligible k_i^{out} in comparison. As a result, the effect of the non-cyclone nodes on the dynamics of the network is inconsequential.

During the period Jul 26-27, Noru propagates northwestwards while Kulap moves slightly towards the southwest (Figures 4.11a1-d1). During this time interval, the vorticity core of Kulap diminishes as the inter-layer vorticity interaction between the cyclones results in the formation of an unstable connected structure [222] (Figures 4.11b1-c1). Comparing Figures 4.11b2-d2 with Figure 4.11a2, we observe a considerable reduction in (i) the average in-degree of and (ii) the area covered by the high-k_i^{in} nodes in the region between Noru and Kulap. This occurs due to the closer proximity between the two cyclones, which reduces the area between them that is occupied by their respective outer layers. At the same time, there is a significant decrease in the values of k_i^{out} of nodes of

Kulap, accompanied by a simultaneous increase of those of the nodes of Noru (Figures 4.11a3-c3). When complete merging occurs, k_i^{out} values at the location of Kulap reduce to almost zero (Figure 4.11d3). Note that, in Figure 4.11b2 the nodes on the side of Noru, opposite to that of Kulap, have high k_i^{in}, although of a lower magnitude. This may be due to the interaction of Noru with a nearby vortex which we do not consider in the current study.

In summary, k_i^{in} provides a quantitative measure of the binary interaction through vorticity advection between Noru and Kulap, while the relative influence of one cyclone on another is captured by spatial distribution of k_i^{out}.

2. **Seroja-Odette system**
 In Figures 4.12a1-d1 we plot ω at 850 hPa in the region 102°E to 125.5°E and 5°S to 25°S, during the period Apr 6, 2021, 06:00 UTC to Apr 8, 2021, 09:00 UTC, which demarcates the interaction period of the pair of cyclones, Seroja and Odette, which have been denoted by the letters 'S' and 'O', respectively. The regions of large negative ω have a strong clockwise wind circulation, as observed for a cyclone in the Southern Hemisphere [223]. At 06:00 UTC on Apr 6, these two cyclonic systems are $\sim$ 1690 km apart. Around this time, diffusion of vorticity from the center of the cyclone to the outer layers occurs [187, 188], which dynamically changes the shapes of the cyclones. Throughout the interaction period, Apr 6-7, the geographical location of Odette remains fairly constant while Seroja continuously approaches it (see Figures 4.12a1-c1). As a result, the separation distance (d) between the cyclones decreases during the interaction period.

At the initial stage, the nodes near Odette have high values of k_i^{in} (Figure 4.12a2). As the cyclones approach each other, the values of k_i^{in} of the nodes which are between the two cyclones increase due to the inter-layer vorticity advection [187, 188] (Figures 4.12b2-d2) with Odette lying closer to these high-k_i^{in} nodes. During this time, k_i^{out} of the non-cyclone nodes are negligible compared to those within the cyclones, the nodes of Seroja having higher k_i^{out} than those of Odette (Figures 4.12b3-d3). This implies that (i) the non-cyclone nodes have marginal effects of the network dynamics and, (ii) the vortical influence of Seroja on its surrounding regions is stronger than that of Odette, as also inferred from similar network characteristics for Noru and Kulap previously.

Next, we study the final stages of the Seroja-Odette interaction up to their complete merger. The spatial distributions of ω, k_i^{in} and k_i^{out} corresponding to these stages are shown in Figure 4.13. We observe the formation of a dumbbell-shaped structure[222] (Figures 4.13b1-c1). At this time, the area of the total region of high-k_i^{in} nodes between the cyclones and their average in-degree decrease (Figures 4.13a2-c2), as also seen earlier for the Noru-Kulap system. Also, the spatial distribution of k_i^{out} resembles that of ω, with nodes corresponding to Seroja having higher k_i^{out} values because of the advection of vorticity from Odette (Figures 4.13a3-c3). Finally, when Seroja and Odette merge completely,
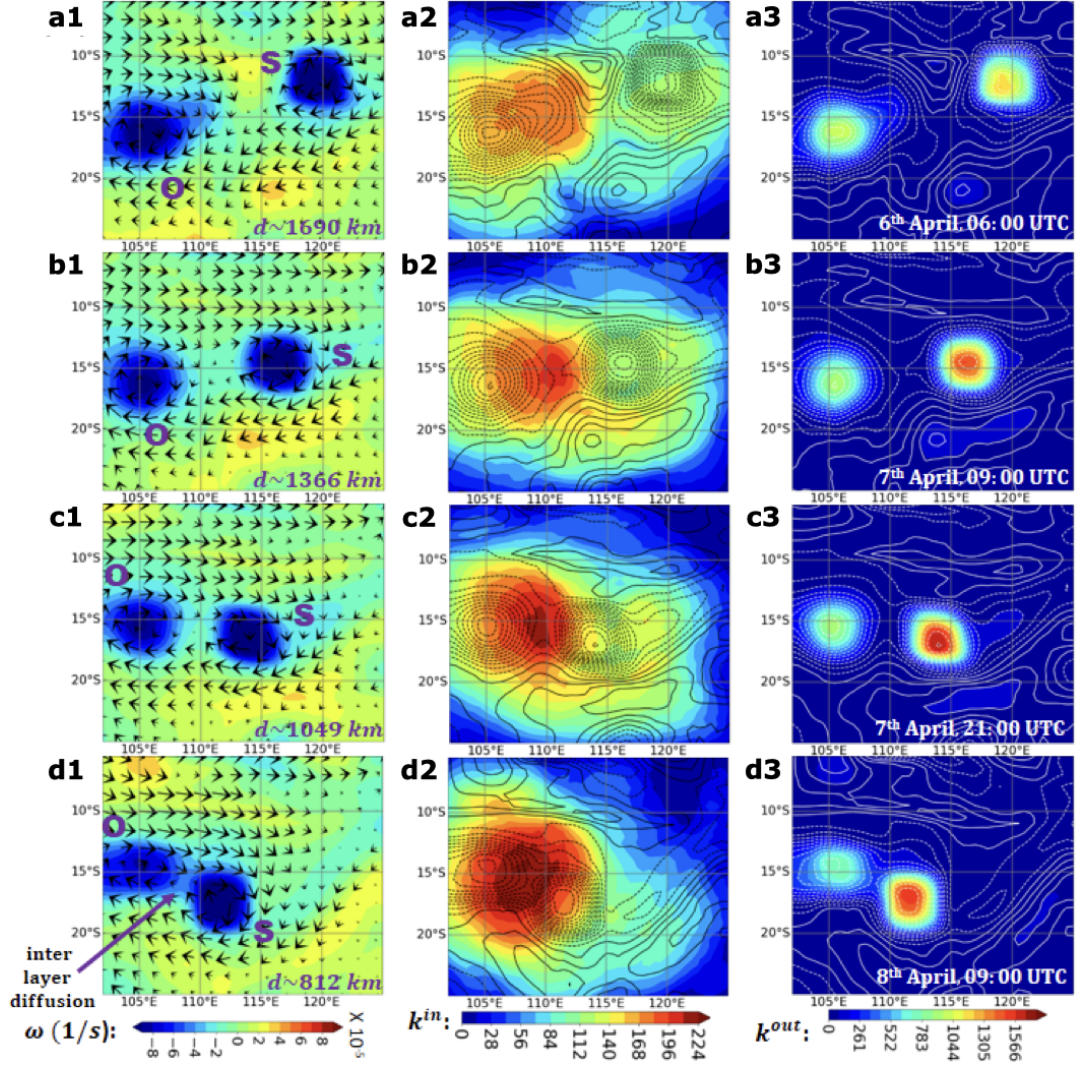
FIGURE 4.12.: The distributions of ω [(a1)-(d1)], k_i^{in} [(a2)-(d2)], and k_i^{out} [(a3)-(d3)] during the interaction of Seroja (S) and Odette (O) at the geopotential height of 850 hPa during Apr 6-8, 2021. The snapshots shown here are for the time steps, Apr 6, 06:00 UTC [(a1), (a2), (a3)], Apr 7, 09:00 UTC [(b1), (b2), (b3)], Apr 7, 21:00 UTC [(c1), (c2), (c3)], and Apr 8, 09:00 UTC [(d1), (d2), (d3)]. The wind velocity vectors are represented by arrows in (a1)-(d1). The vorticity contours corresponding to the Figures (a1)-(d1) are overlaid on the distributions of k_i^{in} and k_i^{out} for a better understanding of the changes of the interaction between the two cyclones; negative vorticity contours are represented by the dotted lines while solid lines indicate the positive vorticity contours. Here too, k_i^{in} in between the cyclones increases as they approach each other [(a2)-(d2)], while k_i^{out} remains high over the cyclone nodes. The magnitude of k_i^{out} at the centre of Seroja is higher than that at the centre of Odette [(a3)-(d3)]. (Taken from De et al. [P3])
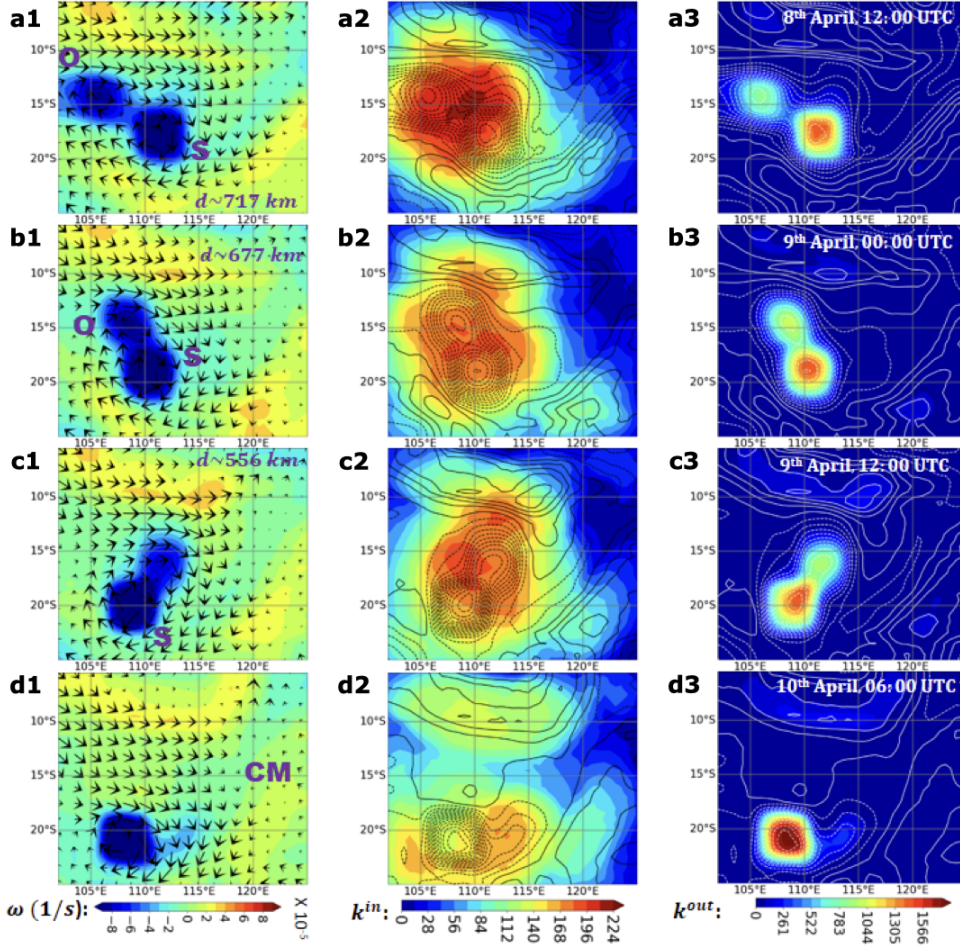
FIGURE 4.13.: The distributions of ω [(a1)-(d1)], k_i^{in} [(a2)-(d2)], and k_i^{out} [(a3)-(d3)] during the interaction of Seroja (S) and Odette (O), prior to the CM, at the geopotential height of 850 hPa. The snapshots shown here are for the time steps, Apr 8, 12:00 UTC [(a1), (a2), (a3)], Apr 9 , 00:00 UTC [(b1), (b2), (b3)], Apr 9, 12:00 UTC [(c1), (c2), (c3)], and Apr 10, 06:00 UTC [(d1), (d2), (d3)]; wind velocity vector of the wind is shown in (a1)-(d1) by the arrows; the vorticity contours corresponding to the Figures (a1)-(d1) are overlaid on the distributions of k_i^{in} and k_i^{out}. There is a gradual reduction in the area of the high-k_i^{in} nodes [(a2)-(d2)] while k_i^{out} is much higher at the centre of the merged cyclone [(a3)-(d3)]. (Taken from De et al. [P3])

there is only a single strong vortex in the system (Figure 4.13d1), and the blob of high k_i^{out} values which was originally centred at Odette vanishes (Figure 4.13d3).

Thus, from our findings, we make the following generic inferences which are true for binary cyclone pairs in either hemisphere:

– k_i^{in} provides a quantitative measure of the interaction via vorticity advection between a pair of binary cyclones.

– k_i^{out} captures the relative vortical influence of a cyclone on its surrounding regions.

– The average value of k_i^{in} of the nodes in between the binary cyclones increases, reaches a maximum value and then decreases during the merging process.

– Most of the links to the high k_i^{in} nodes in between binary cyclones during the intermediate stages of their merger emanate from the stronger of the two cyclones.

– The simultaneous decrease of k_i^{out} of the nodes in the central region of one cyclone and the increase of k_i^{out} at the core of the other during the merger implies the advection of vorticity from one cyclone to the other.

– The sharp decline of k_i^{out} beyond a certain radius of a cyclone implies that the nodes within a cyclone form a highly-connected isolated group within the network [P2]. This also makes the cyclones clearly identifiable from vorticity networks.

– The values of k_i^{out} are an order of magnitude higher than those of k_i^{in}, implying that the stronger interactions in the network come from the nodes concentrated at the centres of the cyclones while, the nodes in between the cyclones are primarily influenced by the cyclone having higher k_i^{out}.

Identification of interaction stages leading to cyclone merger

The classification of the various stages of a binary cyclone merger is a complicated task because the wind shear across the different vorticity layers and the changing wind speeds at different geopotential levels add additional complexity to the analysis. In the previous section we noticed, in particular, that k_i^{in} increases at the nodes between the binary cyclones during the initial stages of the merger and decreases thereafter. Therefore, in order to characterize the different interaction stages we compute the mean of the 95th percentile of k_i^{in} of the network, which we denote as $\left\langle k_{95\text{th}}^{in} \right\rangle$, and study its variation during the entire merging processes of the binary cyclones analyzed above.

1. ***Noru-Kulap system***
   In Figure 4.14a, we plot the values of $\left\langle k_{95\text{th}}^{in} \right\rangle$ for the vorticity fields at 850 hPa, 700 hPa and 650 hPa, during the interaction period of the cyclones Noru and Kulap. All of them show a similar overall trend – increases, reaches a maximum value and then decreases. We observe four distinct stages of interaction prior to CM, marked as Stages I-IV, which we explain in the following paragraphs.

   At the beginning of the interaction, since Kulap is the weaker of the two cyclones and is spread over a larger area (see Figure 4.10a1), it provides the dominant contribution to $\left\langle k_{95\text{th}}^{in} \right\rangle$. During the period Jul 23, 03:00 UTC to July 23, 21:00
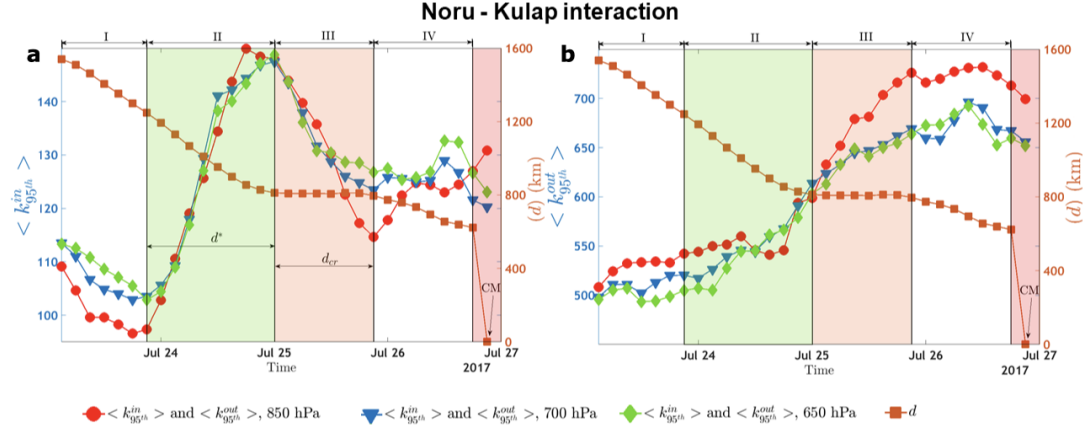
FIGURE 4.14.: (a) The variation in $\left\langle k_{95\text{th}}^{in} \right\rangle$ (mean of 95$^{\text{th}}$ percentile of k_i^{in} in the vorticity network) during the Noru-Kulap interaction; stages -I, II, III and IV are demarcated on the basis of this variation. (b) The variation in $\left\langle k_{95\text{th}}^{out} \right\rangle$ (mean of 95$^{\text{th}}$ percentile of k_i^{out} in the vorticity network) during the Noru-Kulap interaction; stages I-IV are marked as in (a). The changes in $\left\langle k_{95\text{th}}^{out} \right\rangle$ quantify those of the strength of the dominant cyclone. Also shown in (a) and (b) is the temporal variation of separation distance d between the two cyclones. (Taken from De et al. [P3])

UTC, although the vorticity of Kulap decreases due to vorticity advection, Noru is still very far away to significantly affect the region surrounding Kulap. As a result, $\left\langle k_{95\text{th}}^{in} \right\rangle$ decreases slightly during this time interval which we denote as Stage-I.

Thereafter, there is a sharp rise in $\left\langle k_{95\text{th}}^{in} \right\rangle$ during the period Jul 23, 21:00 UTC to Jul 25, 00:00 UTC, when $d \sim 1100$ km. We demarcate this rising phase as Stage-II. During this stage, k_i^{in} in the region between the cyclones rises and reaches a maximum value and so does the area covered by the high-k_i^{in} nodes in this region, as seen from the spatial in-degree distributions earlier. The maximum of $\left\langle k_{95\text{th}}^{in} \right\rangle$ corresponds to the time at which the connected structure of high vorticity (see Figure 4.10d1-d2) emerges. At the end of this stage, $d \sim 812$ km. The range of values of d covered during this stage of the interaction is denoted by d^* in Figure 4.14a.

Next, $\left\langle k_{95\text{th}}^{in} \right\rangle$ decreases rapidly from its maximum value during the interval Jul 25, 00:00 UTC to Jul 25, 21:00 UTC. In this phase, which we refer to as Stage-III, the cyclones form a connected region of high vorticity and come close to each other and, consequently, the number of high k_i^{in} nodes in the region between the cyclones reduces. During this stage, the variation of d is very slow as it decreases to ~ 797 km at the end of this period. This range of separation distance in which $\left\langle k_{95\text{th}}^{in} \right\rangle$ decreases from its maximum value is termed as the range of critical separation distance (d_{cr}) and is shown in Figure 4.14a.

65

Prior to the complete merger, $\left\langle k_{95\text{th}}^{in} \right\rangle$ varies slowly and remains small which corresponds to the Stage-IV shown in Figure 4.14a. At this time, Kulap gradually disappears and the only contribution to $\left\langle k_{95\text{th}}^{in} \right\rangle$ comes from Noru.

Therefore, we observe four distinct stages in the Noru-Kulap interaction before the merging phenomena occurs. There have been a number of numerical studies of the pairing and merging of vortices on the basis of their separation distance d [197, 224, 225]. It has recently been shown using the separation distance, d, that there are three phases before a CM event [197] – a first diffusion (d slowly reduces), convection (d decreases rapidly), and a second diffusion phase. In Figure 4.14a, we also plot d as a function of time. We observe that d decreases during stages I and II, remains almost constant in stage III and then again decreases slightly slowly during stage IV. Importantly, d is unable to distinguish between stages I and II, contrary to $\left\langle k_{95\text{th}}^{in} \right\rangle$.

We also calculate the mean of the 95$^{\text{th}}$ percentile of the out-degree of the network, $\left\langle k_{95\text{th}}^{out} \right\rangle$, at different geopotential heights and plot its temporal variation in Figure 4.14b. We mark the various stages of interaction as per our observations of $\left\langle k_{95\text{th}}^{in} \right\rangle$ in Figure 4.14a. In stage I, $\left\langle k_{95\text{th}}^{out} \right\rangle$ at 850 hPa increases slightly faster than those at 700 hPa and 650 hPa. This is attributed to the greater increase in the vorticity of Noru at 850 hPa than at higher geopotential heights. In stages II and III, we observe a more rapid increase of $\left\langle k_{95\text{th}}^{out} \right\rangle$ at all geopotential heights, which occurs when $d \sim 855$ km. This occurs due to the faster strengthening of Noru due to increased vorticity advection from Kulap. In stage IV, the vorticity of Kulap being extremely small, a further advection of vorticity does not affect Noru appreciably. Consequently, $\left\langle k_{95\text{th}}^{out} \right\rangle$ is roughly constant in this stage. After CM occurs, the entire contribution to both $\left\langle k_{95\text{th}}^{in} \right\rangle$ and $\left\langle k_{95\text{th}}^{out} \right\rangle$ come from the single vortical structure that remains.

2. ***Seroja-Odette system***

We now try to visualize the various stages of the Seroja-Odette interaction using the same methodology. In Figure 4.15a1-a2, we plot the variation $\left\langle k_{95\text{th}}^{in} \right\rangle$ at geopotential heights of 850 hPa, 700 hPa and 650 hPa, along with that of d. The stages of interaction are distinguished as per the variation of $\left\langle k_{95\text{th}}^{in} \right\rangle$. Here, Odette is the weaker cyclone and is spread over a bigger area. Hence, initially, the major contribution to $\left\langle k_{95\text{th}}^{in} \right\rangle$ comes from the nodes in and around Odette, similar to what we saw earlier for the Noru-Kulap system. However, the stage I interaction between Seroja and Odette seems considerably different from that of Noru and Kulap – there are small oscillations in $\left\langle k_{95\text{th}}^{in} \right\rangle$. This is due to fluctuations in ω in the region around Odette, which are stronger at 650 hPa and 700 hPa (Figure 4.15a2) than at 850 hPa (Figure 4.15a1). During stage I, the separation distance d decreases from 2034 km to 1750 km.
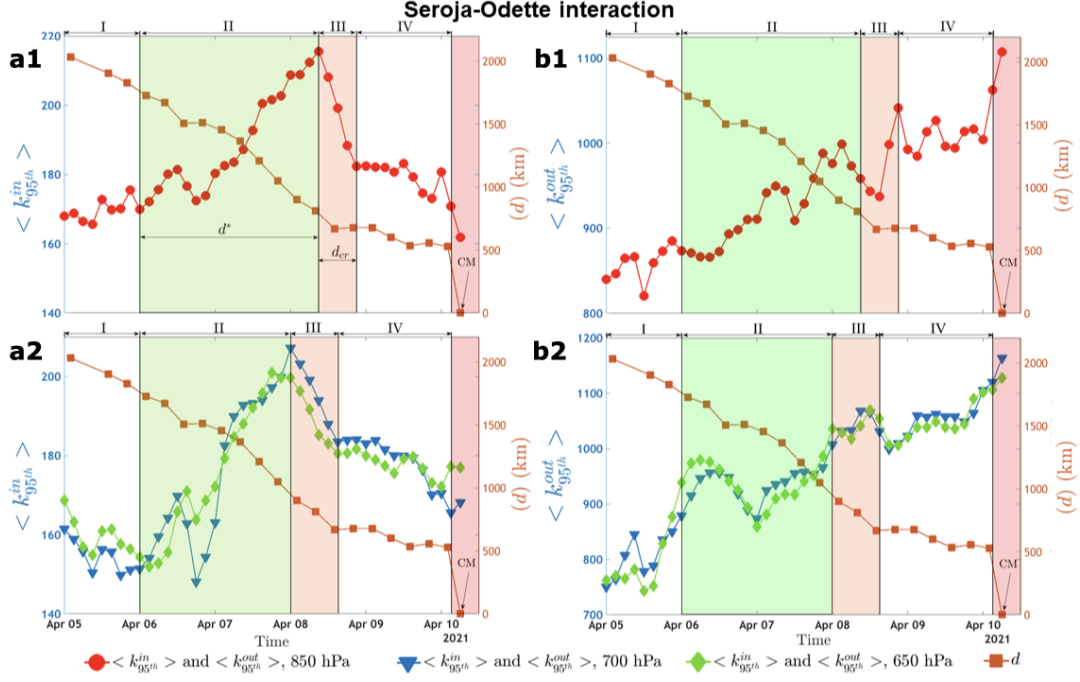
FIGURE 4.15.: (a1)-(b1) The variations in $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ and $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ during the Seroja-Odette interaction at 850 hPa. (a2)-(b2) The variations of $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ and $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ for the same binary interaction at higher geopotential levels of 700 hPa and 650 hPa; stages I-IV, in each case, classified on the basis of the change of $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ that geopotential level. (Taken from De et al. [P3])

In stage II, $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ at all geopotential heights show an overall increase and reach their respective maximum values at the end of this stage when the two cyclones form a large connected structure. Note that, the time of formation of this structure (i.e. time corresponding to the maximum of $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$) is different for different geopotential levels, with those for 650 hPa and 700 hPa happening earlier than that for 850 hPa. Correspondingly, $d \sim 812$ km at 850 hPa while $d \sim 950$ km at 650 and 700 hPa at the end of this stage. Also, there is a significant drop in $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ at 650 hPa before Apr 7, 00:00 UTC.

In stage III, $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ decreases due to reduction in the area of the high-k_i^{in} nodes as the cyclones begin to merge. Finally, in stage IV, the variation of $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ is relatively slower as the influence of Odette on the network almost vanishes before it merges with Seroja entirely. The variation of d across the different stages is qualitatively similar to that seen in the Noru-Kulap interaction i.e. decreases almost uniformly and then saturates for a certain duration before the CM.

We estimate $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ for the Seroja-Odette system at geopotential heights of

850 hPa, 700 hPa and 650 hPa during the various stages of interaction and plot their variations in Figure 4.15b1-b2. As Seroja intensifies and becomes more compact, its vortical interaction with rest of the network increases because of which $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ increases steadily in stage I. The increasing trend continues in stage II due to vorticity advection. At the end of this stage at 850 hPa, when the connected structure of high vorticity just forms, there is a slight drop in the values of $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$. This is due to the slight decrease in the vorticity of Seroja accompanied by a simultaneous increase in that of Odette (see Figure 4.12d1 and 4.13a1) at this time. This is not observed in the Noru-Kulap system. In stage III, a rapid increase in the vorticity of Seroja leads to a significant rise in $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$. Finally in stage IV, $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ is shows small fluctuations at 850 hPa, while at 700 hPa and 650 hPa, it rises very slowly, as Odette gradually merges into Seroja. The trend in $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ differs from what was observed in the Noru-Kulap system.

Thus, the network-based measures uncover the significant differences in the interaction between the binary cyclones during various stages of their merging process. In particular, $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ can be potentially used as an early indicator in the forecast of CM events. Previously, a value of d within the range 1050 km to 2250 km was found to be the critical separation distance for a pair of cyclones to interact. In our study, an estimate of this critical distance from the variation of $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ for different cases of cyclone mergers can improve our understanding of Fujiwara interaction than what is known using conventional methods. Moreover, the positive correlation between $\left\langle k_{95^{\text{th}}}^{in} \right\rangle$ and $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ during stage II, as they both increase, can also be used as an indicator to detect the growing interaction between two nearby cyclones. $\left\langle k_{95^{\text{th}}}^{out} \right\rangle$ by itself provides useful insight into the behaviour of the more dominating cyclone.

4.4. Summary

We have shown that the time-evolving climate network approach has been very promising to study the evolution of weather extremes occurring over very short time scales of only a few days. Through a complex network analysis of the spatio-temporal pressure or vorticity field, we can extract insightful information about the underlying dynamical organization of the regional weather system during cyclones.

In the first part of the study (Section 4.2), we showed the applicability of the method to distinguish between the interaction structure of the mean sea level pressure (MSLP) field when normal conditions prevail and when a cyclone is present in the basin. The temporal evolution of MSLP co-variability patterns is quantitatively studied by utilizing various topological properties of functional climate networks, namely, degree, mean geographical distance and clustering coefficient which characterize the changes in connectivity structure from three different perspectives – centrality of nodes, associated spatial length scale and tendency to form clusters. We employed sliding

windows of 10 days, successively shifted by 1 day over the period of a cyclone season – post-monsoon Sep–Oct–Nov–Dec cyclone season of north Indian Ocean basin, and the Aug-Sep season of the north Atlantic Ocean basin. We found that cyclone-affected regions exhibit an increase in the local clustering values C_i along with a decreasing degree k_i as compared to their surroundings, implying the formation of an almost isolated system within the network based on mean sea level pressure. The highly localized nature of these tropical storms leads to such a behaviour, as was evident from the lower values of mean geographical distance $\mathcal{L}_i$ along the nodes affected by the cyclone. The results revealed that there is a close resemblance of cyclone tracks with high C_i nodes if the cyclone event occurs during the span of the network time window. Such networks also tend to have relatively high global clustering coefficient values. This indicated that the regions along the cyclone track are localized structures in the network with high connectivity, along which there is a continuous flow.

Thus, provided the length of the time window and the temporal resolution of the data is chosen in accordance to the weather phenomenon in consideration, evolving climate networks can be used to study weather variability that occurs over much shorter, daily time scales. Application of this approach over time scales relevant for cyclones allowed us to gain deeper insights into the individual local signatures of changes in the flow structure of the regional weather system, in contrast to generic long-term topological changes reported in earlier works [153, 154]. This methodology using network-based indicators has a strong potential to detect cyclones and their tracks from MSLP outputs from models as well as other inputs.

In the second part of this chapter (Section 4.3), we explored the topological dynamics during the interaction and complete merging of binary cyclone systems by adopting an innovative data-driven complex network analysis of the instantaneous relative vorticity field at a particular geopotential height. The interaction between two flow elements located at two grid points of the network was computed on the basis of velocity induced by one flow element on another using the Biot-Savart law. Such a framework allowed us to comprehend the changes in the connectivity structure during the interaction between two cyclones when they are close to each other. We characterized the topology of the constructed time-evolving directed induced velocity networks by performing a degree analysis. By analysing the spatial distributions of *in-degree*, we quantified the extent of binary interaction between the cyclones, while the distributions of *out-degree* enabled us to identify the dominant cyclone during each time step of the interaction until the merging of the two cyclones finally occurred. Further, the distribution of high out-degree nodes could be used to clearly identify the cyclone, indicating the occurrence of strong interactions within it. The most interesting result was that we were able to classify the transitions of the binary cyclone interaction into four stages before cyclone merging takes place based on the quantification of mean of the 95th percentile of in-degree in the vorticity network. The trends of the mean of the 95th percentile of in-degree and out-degree could also be used to differentiate between the interaction stages of different binary cyclone systems. Thus, the complex network approach enabled us to directly study the evolution

of the interaction structure of the vorticity field, making it a suitable method to gain incisive insights into the interaction process of binary cyclones.

The findings presented in this chapter points towards the importance of the detailed characterization of the interaction structure of a tropical cyclone or a binary cyclone merging event as an essential step towards improving cyclone track forecasts. While the correlation-based networks in the first part were important to demonstrate the capability of network measures to identify cyclone signatures, the Biot-Savart law based induced velocity networks goes a step further to enable us to track instantaneous interactions. It is important to mention here that although the Biot-Savart law is applicable strictly for velocity fields that are divergence-free, we use it to primarily estimate the strength of connection between two points in the flow-field, and our results clearly show the effectiveness of such networks.

Further investigations involving the proposed network tools to study different types of cyclone interactions, such as partial merger, partial straining out and elastic inter-action in different cyclone basins, and differentiating between the interaction structure between co-rotating and counter-rotating (such as cross-equatorial twin cyclones) cyclone pairs can be outlined as relevant topics for future research. Studying the inter-action of the cyclone with large-scale low-level cyclonic vortices such as the monsoon gyre [226] could also be a possible application of the work. Furthermore, this complex network approach, in combination with the physics-inspired machine learning algorithms, can also be used to improve the prediction of cyclone tracks.

5. Asian Summer Monsoon: Spatial synchronization patterns of Extreme Precipitation and Convection

"...the key to it all – the colours, the moods, the scents, the subtle, mysterious light, the poetry, the heightened expectations, the kind of beauty that made your heart miss a beat – well, that remained the monsoon."

– Alexander Frater, *Chasing the Monsoon*

5.1. Introduction

The Asian summer monsoon (ASM) is a planetary-scale phenomenon involving the annual migration of the Intertropical Convergence Zone and is an important component of the global climate system. The ASM is characterized by a distinct seasonal reversal of low-level wind fields accompanied by heavy rainfall. The Indian summer monsoon (ISM) and the East Asian summer monsoon (EASM) are the two main subsystems of the ASM. A large amount of literature has contributed to the study of both monsoon systems and their connections [227–230]. While both subsystems have profound differences, they strongly influence each other [231–235]. The ASM has a huge socioeconomic impact in South and East Asia; therefore, the understanding of its underlying complex interconnectivity structure is crucial.

The ISM is known to have significant impact on rainfall over northern China [236–240]. Several studies also investigated the relationship between ISM onset and the onset of the Meiyu season over the Yangtze River Valley in China [241, 242]. Some authors have referred to the above relationships between rainfall over India and that over East Asia as the two modes of ISM-EASM teleconnection [238, 241–243]. Furthermore, the various modes of the tropical intraseasonal oscillations (ISO) [227, 244] – the eastward propagating Madden-Julian oscillations (MJO) [245, 246] and the northward propagating boreal summer intraseasonal oscillation (BSISO) [247–251] – are known to have considerable influence on the variability of the global monsoon system at intraseasonal timescales. Numerous studies have investigated the impact

---

This chapter is based on the associated publication Gupta et al. [P4], licensed under a Creative Commons Attribution (CC BY) license, and the manuscript in preparation Gupta et al. [P5]. The sections in this chapter closely follow these works.

of the ISO on the different aspects of the ISM and the EASM, individually. However, not much attention has been given to the possible influence of the ISO on the intraseasonal variability of the ISM-EASM connection. Understanding the variability of the spatial patterns of rainfall associated with the ASM at intraseasonal time scales is essential to improve the subseasonal forecasting of extreme precipitation events.

The interannual variability of ASM in terms of the impacts of the El Niño-Southern Oscillation (ENSO) is another important aspect [235]. The mechanism by which the ENSO remotely influences the dynamics of the ASM is complex and varied for different components of the ASM. The effect of the phase of the ENSO is mainly via the modulation of the convective processes that drive the ASM. Therefore, the study of the impact of the ENSO on the convection anomalies is another point of interest for our long-term climatological analysis, to corroborate previous studies that investigate the impact of the ENSO on the ASM [252, 253]. These studies mostly consider the model or satellite observations of sea surface temperature, winds, and outgoing long wave radiation (OLR) which can be taken as a proxy for detecting deep convection and hence heavy precipitation in the tropical and subtropical regions, as cloud top temperatures (colder is higher) are an indicator of cloud height [254].

Most of the studies on the spatial variability of the ASM at different time scales primarily rely on methods of correlation, empirical orthogonal functions (EOFs), and composite analyses. However, these methods are often not sufficient for analysing the spatial characteristics of extreme rainfall co-variability. Furthermore, analyses based on correlations alone typically do not allow to identify the specific times associated with a particular climate interaction pattern. In fact the analyses based on correlations, such as one-point correlation or teleconnectivity maps, give a limited understanding of the climate mechanisms based on the selective impact of a reference region. Complex networks approaches have proven to be an effective tool to study the intricate spatiotemporal coupling structures due to the large-scale seasonal changes in precipitation, wind and other observables in monsoon systems. They have been applied in combination with a nonlinear synchronization measure to analyse both the regional and global patterns of extreme rainfall [48–51, 111]. In particular, separate studies based on complex networks have been conducted to reveal spatiotemporal patterns of synchronous heavy rainfall associated with the ISM [48, 49, 255] and the EASM [256–258].

In the first part of this chapter (Section 5.2, [P4]), we use a similar method of constructing climate networks using extreme precipitation spatiotemporal satellite data [259] for the monsoon season, to go a step further by investigating the interaction between these two subsystems of the ASM and its intraseasonal variability. We show that the connection between ISM and EASM has two different modes which are separated both spatially and temporally. Moreover, we provide a comprehensive picture of the large-scale atmospheric circulation patterns associated with each mode of connection. We further investigate the potential role of the different modes of the ISO by studying the distribution of the MJO/BSISO phases of those days when there is high synchronization of extreme rainfall events between the ISM and the EASM.

However, the aforementioned approach based on complex network analysis of extreme precipitation events may not be suitable to study the interannual variability of the ASM due to the ENSO as it may lead to loss of information regarding the mechanism of the ENSO-ASM relationship. This is because the phase of the ENSO modulates the ASM at monthly time scales by a series of processes that can outlast a season [260, 261]. Therefore, it is more suitable to instead employ the continuous temperature, OLR and wind data which carries the signature of the impact ENSO on tropical convection. Although climate networks based on sea surface temperature or surface air temperatures have been used to identify and distinguish different types of ENSO events [43–45, 262], networks based on wind have been found to be effective to investigate the impact of ENSO on monsoon systems in other continents [263]. In this context, we investigate the influence of the ENSO on the interannual variability of the convection anomalies in the Asia-Pacific region using correlation-based complex networks of the reanalyses-based OLR and the lower tropospheric wind components [162] in the second part of this chapter (Section 5.3, [P5]). Such a complex network analysis of OLR data has not been conducted in earlier studies to the best of our knowledge. This multivariate OLR and wind based climate network analysis reveals important features of the ASM, such as the prominent distinction between the convective sources that drive the different components of the ASM. Our results also reveal that the phases of the ENSO impact the different convective sources differently and hence shed light on the spatial variability of its relation with the various regions of the ASM.

Finally, our findings on both the intraseasonal and interannual variability of the ASM are summarized in Section 5.4.

5.2. Interconnection between the Indian and the East Asian Summer Monsoon

Here, we use a method based on the combination of a nonlinear synchronization measure and complex network theory to unravel the profound and complicated relationship between the ISM and the EASM, as well as its variability as the season progresses. In Section 5.2.1, we describe the data employed in the analysis as well as outline the necessary pre-processing steps. Thereafter, the methodology based on event synchronization and complex networks is explained. The results so obtained along with their climatological interpretation is provided in Section 5.2.2.

5.2.1. Data and Methodology

Data

We analyse the satellite-derived Tropical Rainfall Measurement Mission (TRMM 3B42 V7) [259] total precipitation data, with daily temporal resolution, provided on a spatial grid with resolution of $0.25° \times 0.25°$, ranging from 50°N-50°S, for the

time period 1998-2019 (`https://disc.gsfc.nasa.gov/`, last accessed: 8 Sep., 2020). The spatial domain of this study is confined to the region 0° to 50°N and 60°E to 160°E for the functional network analysis, which gives 80,000 grid points. It is worth mentioning here that, although satellite rainfall products are known to underestimate extreme precipitation [264], their high spatiotemporal resolution and almost global coverage make them convenient for studying spatial patterns of precipitation.

Data for atmospheric variables, such as wind components and geopotential height (GPH) at pressure levels of 250 hPa, 500 hPa and 850 hPa, vertically integrated water vapour flux and outgoing long-wave radiation (OLR), are obtained from the ERA5 reanalysis [162] at daily temporal and $1° \times 1°$ spatial resolutions for the same time period (`https://cds.climate.copernicus.eu/`, last accessed: 9 Dec., 2020). Analysis of their composite anomalies is done over an extended spatial region (15°S to 50°N and 0° to 160°E).

The Real-time Multivariate MJO indices (RMM1 and RMM2) of Wheeler and Hendon [246], based on a pair of empirical orthogonal functions (EOFs) of the combined fields of near-equatorially averaged 850-hPa zonal wind, 200-hPa zonal wind and OLR data, were used for defining the various phases of the MJO (`http://www.bom.gov.au/climate/mjo/`, last accessed: 22 Nov., 2021). The phases of the BSISO were defined using the real-time indices BSISO1 and BSISO2 proposed by Lee et al. [251], which are based on a multivariate EOF analysis of daily anomalies of OLR and zonal wind at 850 hPa (U850) in the region 10°S-40°N, 40°-160°E (`https://apcc21.org/ser/moni.do/`, last accessed: 22 Nov., 2021).

Extreme Rainfall Events   Extreme rainfall events (EREs) are identified as days with total rainfall sums greater than the 90th percentile of wet days (rainfall > 1 mm) at each grid point, for the particular month (June, July or August) or the season under consideration (here, the monsoon season from June to August, JJA). The events are de-clustered by counting consecutive days with rainfall above the threshold as a single event placed on the first day of occurrence.

Functional Network Analysis

Network Construction   We construct the extreme precipitation network using method outlined in Section 3.5.2. The nodes of our functional network are the spatial grid points of the TRMM precipitation data with their corresponding ERE series whose construction is described above. We employ here event synchronization (ES) (Section 2.2.2) as an event-based similarity measure to quantify synchronous EREs at different locations occurring at varying temporal delays [48–51, 106, 111]. Although ES uses an adaptive delay to find synchronous pairs of events, we confine the synchronization time scale by allowing a maximum temporal delay of $\tau_{\max} = 7$ days, which is equivalent to the time period of the Rossby waves. ES_{ij} is computed for all pairwise combinations of grid points i and j, i, $j = 1, ..., N$, where $N = 80,000$. The statistical significance of each empirical value ES_{ij} is determined on the basis of a null model

distribution which is numerically obtained by computing ES for 2000 pairs of surrogate event series with l_i and l_j uniformly and randomly distributed events. Finally, a network link is placed between grids i and j if ES_{ij} is significant at a significance level of 0.05 (i.e, $ES_{ij} > $ 95th percentile of the corresponding null model distribution). The network adjacency matrix A_{ij} is then constructed by setting $A_{ij} = 1$ if there is a link between nodes i and j, and $A_{ij} = 0$ otherwise (Sections 3.2 and 3.5.2). The spatial patterns of network degree depend only weakly on the choice of the significance threshold, ranging from the 95th to the 99.5th percentile of the null model distribution. Our results are also robust to the choice of different maximum temporal delay between synchronous events, $\tau_{max} \in [3, 15]$ days. All our results remain similar for the case when EREs are defined as events above the 95th instead of the 90th percentile (not shown).

We construct three separate networks using the above method for the successive months of June, July and August, using EREs calculated from the daily rainfall of the respective month to observe monthly evolution of the connectivity structure of the region during the monsoon season. A separate network is also constructed for the whole JJA season using EREs computed from the JJA daily rainfall.

Network Measures After constructing the extreme precipitation network, we compute the node-base network measure *degree k_i* for every node i of the network as defined in Eq. (3.4) of Section 3.3. In this case, the degree at a given grid point gives the number of those grid points where extreme rainfall occurs synchronously. Regions of higher degree are of particular importance for identifying large-scale synchronization of extreme rainfall and their causes [48, 49, 51].

We also calculate the *partial degree $\{k_i\}_R$* of the nodes in the network linked to a particular region R, which yields the number of links connecting a node i outside R with the nodes within R. This gives a selective view of the degree plot to identify regions in the network connected to a specific area. This approach leads to a visualization that reveals the various modes of connection among the different regions.

Identification of days of high rainfall synchronicity As explained in Section 2.2.2, we use a modification of ES to determine the specific days when high event synchronization of extreme rainfall occurs between two regions of interest A and B, while keeping track of the temporal order [111]. Then, a Butterworth low-pass filter with a cutoff frequency of 7 days is applied to the time series of $ES^{\mu}_{A \to B}$ ($ES^{\nu}_{B \to A}$) so obtained. Finally, we obtain the days of high synchronization between the two regions by identifying the local maxima of the time series that are above the 90th percentile of the entire time series. The specific time points so obtained are then used to compute the composite anomalies of rainfall (from TRMM) and other atmospheric variables (from ERA5).

Lead-lag correlation analysis

We get two time series by spatially averaging the daily number of EREs in the two given regions of interest, and the high-frequency noise is then eliminated by passing them through a Butterworth low-pass filter with a cutoff period of 7 days. A lead-lag correlation analysis is then performed between both time series by computing the Spearman rank correlation (Eq. (2.12)) between the first time series and the lagged values of the second time series, with lags varying in the range $[-30, 30]$ days. The results remain very similar when the cutoff period of the Butterworth filter is varied within a range of 7-11 days (not shown). This analysis helps us to obtain the duration by which one series leads or trails behind the other which gives us a preliminary idea about how long it may take for the effect to propagate from one region to the other and, hence, establish the connection. It serves as an additional step to see if the two connected regions obtained for the network are indeed significantly correlated. The information of the time lags at which high correlation occurs, is useful to compute the composites of vertically integrated water vapour flux in order to trace the path of the moisture transport between the interacting regions.

5.2.2. Results and Discussion

In this section, we first present our results based on the network analysis (see Section 5.2.1). From the spatial patterns of degree of the extreme precipitation networks, we identify the regions associated with the ISM and the EASM that show the strongest connections. Thereafter, we investigate the mechanism underlying these synchronization pathways based on the associated atmospheric circulation features. Finally, we study the potential influence of ISO on variability of the ISM-EASM interconnection at intraseasonal timescales.

Network analysis – Spatial degree patterns

The spatial pattern of degree yielded from the climate network analysis (see Section 5.2.1) of the chosen bounded region for the successive months of June, July and August, and the entire JJA season are shown in Figures 5.1a-d, respectively. From the monthly evolution of the degree configuration of the extreme precipitation network (Figures 5.1a-c), it is evident that – while there occurs an overall change in degree over the ISM, EASM and Pacific Ocean regions – two specific regions of the ISM show a pronounced change over these months. First, the Arabian Sea region (ARB, red box in Figure 5.1) and adjacent coastal region of India exhibit a discernible decrease in degree when going from June to July. In addition, the northern and central part of India, which is often referred to as the core monsoon zone (CMZ, blue box in Figure 5.1) [229, 265] of the ISM, shows a significant increase in degree in July and August in comparison with June. Interestingly, although the spatial degree pattern for the entire JJA season (Figure 5.1d) is qualitatively rather similar to that of the monthly networks (Figures 5.1a-c), we do not see a high degree patch over ARB in the
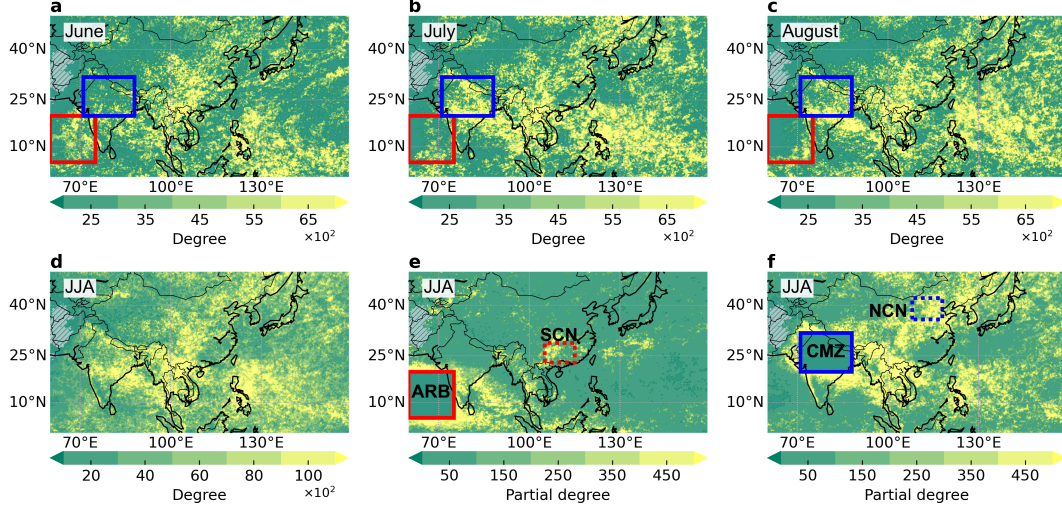
FIGURE 5.1.: Monthly evolution of network degree for networks constructed for (a) June, (b) July and (c) August, respectively. The solid red and blue boxes are positioned at the Arabian Sea (ARB) and core ISM zone (CMZ), respectively, to indicate noticeable changes in degree during each month at the different ISM regions. (d) Spatial pattern of degree for the network constructed for the entire JJA season. (e) and (f): Partial degree for the regions, ARB (solid red box: 5°-20°N, 60°-75°E) and CMZ (solid blue box: 20°-32°N, 71°-88°E) respectively, based on the network in (d) to indicate specific regions connected to each of them. From (e), it is seen that ARB has connections to Southern China (SCN; dashed red box: 23°-29°N, 105°-115°E) while from (f), CMZ is seen to be connected to parts in Northern China (NCN; dashed blue box: 36°-42°N, 108°-118°E). (Taken from Gupta et al. [P4])

JJA network (Figure 5.1d), implying that links connecting ARB with other regions occur earlier but get overshadowed later in the season when the monsoon activity is at its peak and has moved inward into the subcontinent.

Therefore, we compute the partial degree (see Section 5.2.1) corresponding to the JJA network in Figure 5.1d for the regions ARB and CMZ separately. This will help us to identify any changes in the spatial connectivity of these regions, which may occur over the course of the season. From the partial degree distribution of the region ARB (red solid box) shown in Figure 5.1e, we observe that apart from local connections with adjacent regions of the Indian peninsula and the Bay of Bengal, it has long-range connections to a region in south-eastern China (SCN, red dashed box), roughly located in the middle and lower reaches of the Yangtze River basin, which receives persistent Meiyu rainfall [230, 231]. On the other hand, the CMZ region (blue solid box) is well connected to the northern parts of China in the Yellow River basin (NCN, blue dashed box) (Figure 5.1d). We hence identify two spatially and temporally separated ERE synchronization pathways between ISM and EASM regions: (i) ARB with SCN, and (ii) the CMZ with NCN. These two pairs form the Southern (ARB-SCN) and the Northern (CMZ-NCN) modes of the ISM-EASM connection [228, 238, 241, 242].

The distribution of the spatial distances across which significant synchronizations occur in case of the ASM regional network constructed for the JJA season exhibits a scale break at approximately 2000 km (see Figure A.2a), similar to that shown by [111] for a global extreme precipitation network, indicating a physical regime shift. For distances (d) less than 2000 km, the distribution shows a power-law decay [$p(d) \propto d^{-\alpha}$], with an exponent $\alpha = 0.83$, signifying links associated with regional weather systems. On the other hand, the distribution for longer distances ($d > 2000$ km) follows closely the distribution of all possible great-circle distances on Earth's surface, implying that the links are potentially associated with global-scale teleconnections. It must be noted that the spatial patterns of the partial degree of ARB and CMZ, computed by considering only links longer than 2000 km (see Figures A.2b-c), still show the significant connections to SCN and NCN, respectively, as seen in Figures 5.1e-f, thereby implying that these two modes of connection between the ISM and the EASM are also part of the global-scale teleconnections.

Moreover, from the partial degree plots of Figures 5.1e-f, it should be noted that there are almost no links connecting ARB and SCN to NCN (Figure 5.1e and A.3a), as well as CMZ and NCN to SCN (Figure 5.1f and A.3b). The asynchronicity of EREs between SCN and NCN (also seen from Figures A.3a-b) along with significant change in the degree patterns of ARB and CMZ when evolving from June to July clearly indicate that the two connection modes occur at different times within the JJA season. The higher connectivity of nodes in ARB in June and its subsequent decrease in July, during times when connectivity is enhanced over CMZ, implies that the Southern mode precedes the Northern mode. The simultaneous occurrence of EREs in the regions ARB and SCN in June and that in the regions CMZ and NCN later in July and August can also be observed by defining an index for the northward movement of ISM and EASM based on the number of EREs, as described in Appendix A.2.1 and Figure A.4. In the following subsection, we show using event synchronization that the pairs of regions (ARB-SCN and CMZ-NCN) indeed have high synchronization of EREs at different times in the JJA season. We will focus on understanding the mechanism of the two modes of the ISM-EASM connection. The red and blue pairs of boxes (solid and dashed) shown in Figures 5.1e-f are the representative regions chosen for ARB, SCN, CMZ and NCN for all further analysis.

CMZ also exhibits a very high number of links (Figure 5.1d) to regions of the western Pacific Ocean, the Philippine Sea and South China Sea, which also occur as patches of overall high degree in Figure 5.1e. The synchronous EREs in these regions are mostly due to typhoons, which are prevalent during the JJA season [266]. We do not consider these regions in our further analysis, as we are more interested in monsoon-type rainfall than that from the thunderstorms. High total degree is also observed over northern Bay of Bengal, parts of north-east India, Tibetan Plateau and parts of Mainland Southeast Asia such as South Vietnam and Thailand (Figure 5.1e). These regions also experience rainfall from the southwesterly monsoon winds which explain their linkage with CMZ (Figure 5.1f).

Although extreme rainfall over southern Japan due to the Baiu front is also a part
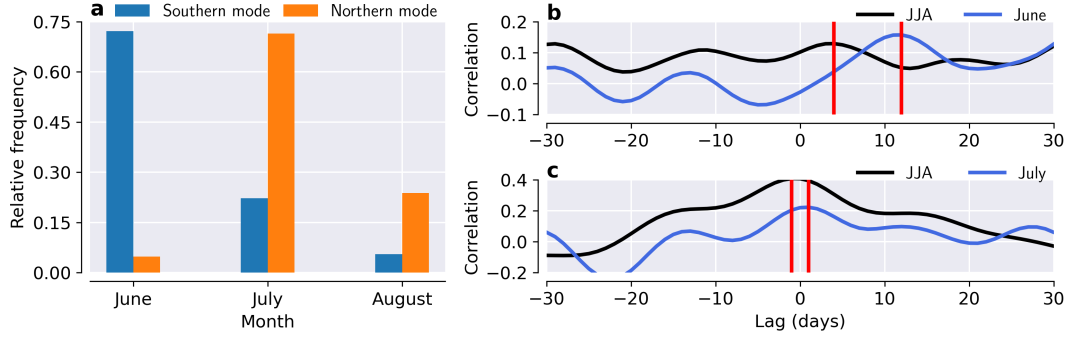
FIGURE 5.2.: (a) The month-wise distribution of days with high ERE synchronization for the Southern mode (ARB → SCN) and the Northern mode (CMZ → NCN). Lead-lag correlations of low-pass filtered time series of spatially averaged daily numbers of extreme rainfall events (> 90th percentile) between (b) ARB and SCN, i.e., Southern mode; and (c) CMZ and NCN, i.e.,Northern mode. (Taken from Gupta et al. [P4])

of the EASM system, it is less synchronous with ISM, as seen from Figure A.3c [237, 239, 242]. However, there are links connecting southern Japan to SCN, as seen from Figures A.3a and A.3c, indicating the occurrence of synchronous frontal rainfall in both regions [230, 231].

Specific times of high extreme rainfall synchronicity

Next, we identify specific days during which extreme rainfall events occur synchronously in each of the two pairs of regions, viz., ARB-SCN and CMZ-NCN (see Sections 5.2.1 and 2.2.2). The month-wise distribution of the days of highest rainfall synchronicity is shown in Figure 5.2a. We see that the frequency of synchronous extreme rainfall occurrences for the Southern mode of ISM-EASM connection is the highest in June and decreases sharply in July and August, while that for the Northern mode peaks in July. This is consistent with our earlier observations in Figure 5.1. Moreover, the frequency of the high ERE synchronization times is seen to be higher at the beginning of June, which is the period of ISM onset. It also corroborates with previous studies of the interrelationship, which claimed that the Southern mode is associated with the onset of the ISM in June and the Meiyu in Yangtze basin [241]. Ding and Wang [228] speculated that the circumglobal teleconnection (CGT) plays an important role to connect ISM with rainfall over northern China and made an indirect inference about the time of the establishment of this relation being in late summer (July-August). Our results in Figure 5.2a directly confirm this conjecture.

Lead-lag analysis of rainfall Figures 5.2b-c show a lead-lag analysis of the time series (Section 5.2.1) obtained by spatially averaging the number of EREs in each pair of regions, for the JJA season (*black curve*) and for the month when the highest synchronization is observed for each mode (*blue curve*). We find from Figure 5.2b

that the correlation between ARB and SCN peaks at a lag of 4 days for the entire JJA season but is maximum at 14 days when only June is considered. However, interestingly, in case of the Northern mode of ISM-EASM connection, we see from Figure 5.2c that the maximum correlation occurs at lag 0 during the JJA season and at a lag of 1-2 days for only July. The positive correlation between CMZ and NCN is in agreement with that observed by Kripalani and Kulkarni [237] between summer monsoon rainfall over India and north China, indicating an in-phase relationship. Only those lags which have significantly high correlations ($P < 0.05$) were used to determine the local maximum peak. The rainfall composite anomalies (see Figure A.5) also reflect the above findings.

Atmospheric Circulation Composite Anomalies

Next, we compute the composites of the anomalies of geopotential height, meridional wind speeds, and the horizontal wind field at different pressure levels (see Section 5.2.1), based on the days of strong synchronization (Day 0) for both modes of the ISM-EASM connection (refer Sections 5.2.1 and 2.2.2). The anomalies are calculated with respect to the JJA seasonal climatology.

Southern mode

Geopotential height and wind Figure 5.3 provides the composite anomalies of the geopotential height (GPH) and wind field at 250 hPa, 500 hPa and 850 hPa, respectively, on Day 0 for the Southern mode of the connection. We also show the corresponding composite anomalies of the wind direction field during those times on the GPH-plots. In the upper-level of the troposphere (250 hPa) (Figure 5.3a), we find an anomalous high-pressure system over the tropics, spanning over the Indian Ocean up to South Asia – covering the Indian subcontinent, the Mainland Southeast Asia up to the Yangtze basin and the Maritime continent. A deep low-pressure zone is present over the rest of the continent. There is also a strengthened upper-westerly jet stream flowing further north of the high GPH region. Strong westerly winds mark the northern boundary of the high GPH region, north of the Yangtze basin, as also seen from the upper-level meridional (v) wind component in Figure 5.4a. The position of the anomalous high GPH in the upper troposphere controls both the ISM and the rainfall over SCN, which explains the relationship between ARB and SCN.The wave pattern seen in the upper level meridional wind composite anomalies (Figure 5.4a) originates from North Africa to east Asia along the mid-latitude westerly jet [267]. The wavenumbers $k = 7$, 11 are associated with this pattern, as determined from the spatial power spectral density of the latitude belt from 25°N to 35°N (see Figure A.6a). The westerly zonal flow is anomalously strong, which may explain the high correlation between rainfall in ARB and SCN at a lag of 4 days during JJA (black curve in Figure 5.2b).
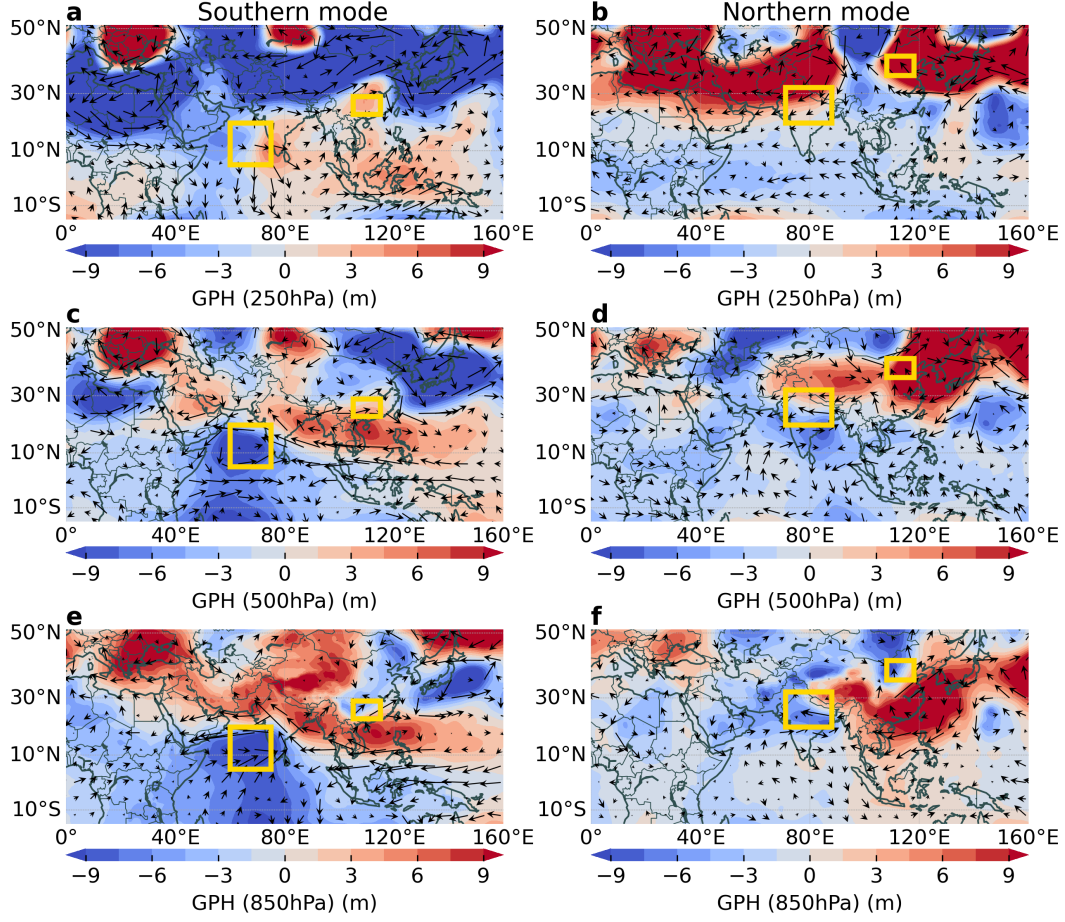
FIGURE 5.3.: Composite anomalies of geopotential height at 250 hPa, 500 hPa and 850 hPa, with respect to JJA climatology, based on the days of high ERE synchronization on Day 0, for the Southern ((a), (c), and (e)) and Northern ((b), (d), and (f)) modes of the ISM-EASM connection. (Taken from Gupta et al. [P4])

In the mid-(500 hPa; see Figure 5.3c) and lower-level atmosphere (850 hPa; Figure 5.3e), there is a cyclone formation over the ARB, and also an enhanced western North Pacific subtropical high (WNPSH) south of 20°N and extending to the Bay of Bengal, whose position determines the region of convergence of winds. At low levels we see a cross-equatorial south-westerly wind coming from Somalia hitting the south-western coast of India. The wind direction in the southern part of the WNPSH over the Bay of Bengal is easterly. This easterly wind meets the above mentioned westerly wind flow at the Malabar coast leading to the ISM onset [241, 242, 268]. As the location of both the cyclone and the western margin of WNPSH moves north-eastward over the course of the next few days, the region of convergence shifts from ARB and South India via the Bay of Bengal to the South China Sea and SCN [268,
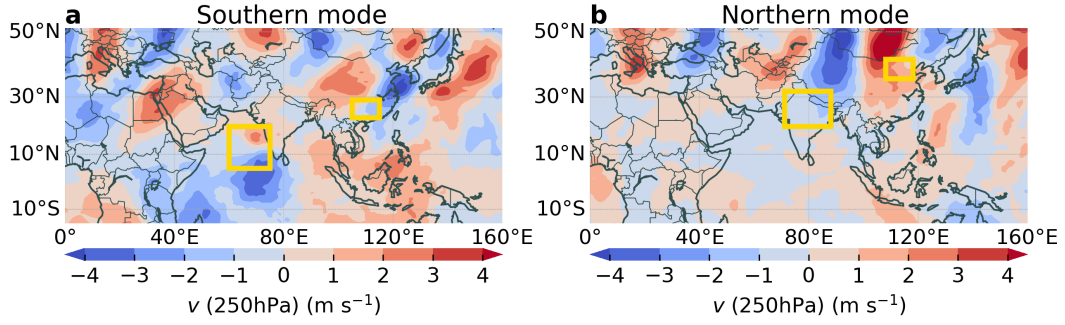
FIGURE 5.4.: Composite anomalies of upper-level meridional wind component v at 250 hPa, with respect to JJA climatology, based on the days of high ERE synchronization on Day 0, for the Southern (a) and Northern (b) mode of the ISM-EASM connection. (Taken from Gupta et al. [P4])
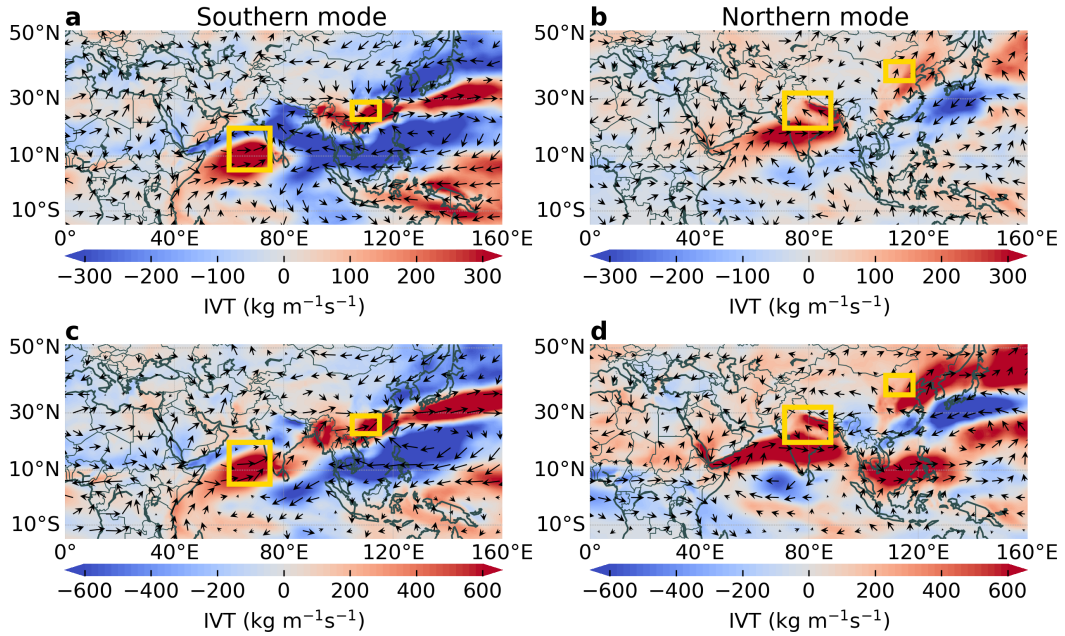
FIGURE 5.5.: Aggregates of composite anomalies of vertically integrated water vapour flux, with respect to JJA climatology, for the Southern mode for days (a) 0-4 and (c) 0-14. The same is also shown for the Northern mode for days (b) 0-2 and (d) 0-14. Day 0 corresponds to days of high ERE synchronization for each mode of the ISM-EASM connection. (Taken from Gupta et al. [P4])

269]. As we will discuss later, the WNPSH shifts over to Japan in about a month as seen from the GPH 500 hPa (Figure 5.3d) and 850 hPa (Figure 5.3f) plots of the Northern mode which peaks in July.

Vertically Integrated Water Vapour Flux   Although there is a high correlation between ARB and SCN at a lag of 4 days for the entire JJA season (Figure 5.2b), the aggregates of IVT composite anomalies (Figure 5.5a) for Days 0-4 still show separate moisture sources for South India and the Yangtze basin. The water vapour transport pattern highlights the Somalian low-level jet and the Zanzibar current along it as the main sources of moisture to ARB and hence the onset of ISM (Figure 5.5a) [270]. On the other hand, the predominant moisture sources of SCN during this time are the South China Sea and the adjacent gulf regions [271].

However, when we compute the aggregates of IVT composite anomalies for Days 0-14 (since the highest correlation is seen at a lag of 14 days between ARB and SCN for June; blue curve in Figure 5.2b), a continuous path of anomalously high moisture transport is established between the two regions via the Bay of Bengal (Figure 5.5c). Therefore, a convergence of the intense water vapour transport from the western Indian Ocean and the Pacific Ocean sources occurs in about 2 weeks via the Bay of Bengal, establishing a moisture corridor which coincides with the dominant route of the East Asian atmospheric river [272–274]. The structure and period of formation of the moisture corridor is similar to that of the third stage of the climatological annual cycle of the East Asian atmospheric river catalogued by Pan and Lu [275], which lasts approximately from 11 Jun to 1 Jul. The moisture pathway transports water vapour from the western Indian Ocean to the ARB, India, the Bay of Bengal, and then extends to the north of the Indo-China peninsula, South China Sea, eastern China, and further northeastward to Japan (Figure 5.5c and A.7a). Hence, the establishment of the moisture corridor over the course of two weeks between ARB and SCN explains the Southern mode of the connection between ISM and EASM. Furthermore, we see from Figures 5.5a and 5.5c that, although the heavier rainbelt along the SCN follows from a convergence of south-westerly tropical and mid-latitude northeast moisture sources, the tropical water vapour transport actually originates from the Philippine Sea, as was also noted by [272].

Northern mode

Geopotential height and wind   In the upper troposphere (250 hPa), we see a strong anomalous high in Figure 5.3b over Eurasia, roughly north of 20°N. However, the GPH 250 hPa is not smooth over the entire continent. In the 250 hPa wind field (Figure 5.3b), we identify an anticyclone-cyclone-anticyclone (A-C-A) circulation pattern, along a southwestern-northeastern direction [276, 277]. There is an anomalous high over west central Asia and parts of the Tibetan Plateau west of 90°E [278], and a westerly wind is found north of the Tibetan Plateau. The other subtropical anticyclone is formed over parts of north-east China and predominantly over Japan [231, 279].

At 500 hPa (Figure 5.3d), a high in the east and low in the west is seen. The WNPSH is further up in the northeast with a protruding tilt in towards the southwestward direction. There is a sustained low over the Indian subcontinent [280, 281]

with a formation of a deep trough over Kazakhstan. In the lower-level (Figure 5.3f), we observe a deepened low pressure trough over India, a higher ridge over the eastern Tibetan Plateau and a low over NCN. The enhanced WNPSH is now further north-east as in GPH 500 hPa with a protruding part over Japan as is also seen for the mid-level. The above circulation features are known to be related to strong ISM [282] and are also favourable conditions for causing abundant rainfall over NCN [231, 283–285].

The composite anomalies of the upper-level meridional wind component v (Figure 5.4b) show a large-scale wave train stretching across Eurasia. The dominant wavenumbers associated with this wave pattern (Figure 5.4b), $k = 5, 8$, are lower than those for the Southern mode, as determined from the spatial power spectral density of the latitude belt from 40° N to 50° N (see Figure A.6b). This denotes the propagation of quasi-stationary Rossby wave energy along the east Asian jet, also named as the "Silk Road teleconnection pattern" [279, 286]. The Eurasian wave train, originating in the jet exit region of the North Atlantic, may be affecting the west central Asian high and, thus, the intensity of the ISM [228]. The upper-level anti-cyclone over west central Asia results in an enhanced convection over India, inducing rainfall [287, 288]. The anticyclone is perturbed by the stationary waves along the Asian jet in late July [289]. The propagation of the quasi-stationary Rossby waves downstream along the jet and their accumulation in the jet-exit region near Japan, forms the anticyclone over Japan [279]. In the meantime, the WNPSH moves northward in the lower and mid troposphere as Meiyu/Baiu fades away. This forms a deep-barotropic ridge throughout the troposphere over Japan with a slight westward tilt called the Bonin high [279, 289], as seen from the GPH plots (Figures 5.3b, 5.3d, and 5.3f). However, the deepened low over NCN in the lower-troposphere in contrast to the high in Japan, indicates enhanced convection over NCN, inducing increased rainfall, while Japan experiences dry spells during this period. Ambrizzi et al. [98], through an upper-tropospheric streamfunction teleconnectivity map and time-lag cross-correlation analysis, showed that the wave pattern emanates downstream eastward to the Pacific from the base point (35°N, 75°E), in about 3 days (see Fig. 7 in [98]). Therefore, the propagation of the quasi-stationary Rossby wave establishes the connection between the strong ISM rainfall over CMZ and the concurrent extreme rainfall over NCN [237, 288]. The amplification of the same Rossby wave train upstream due to strengthening of the upper-level anticyclone over west central Asia was also identified as the plausible reason behind the connection of extreme rainfall between Europe and ISM by Boers et al. [111], who noted that the specific times of the Europe-ISM extreme rainfall synchronization are more in July and August than in June and September.

The in-phase relationship between the precipitation at CMZ and NCN at intraseasonal timescales is also verified using wavelet-based analysis for identifying phase synchronization between two time series (Section 2.2.1). The significant scales for the time series of JJA daily total precipitation (smoothened using a 3-day mov-
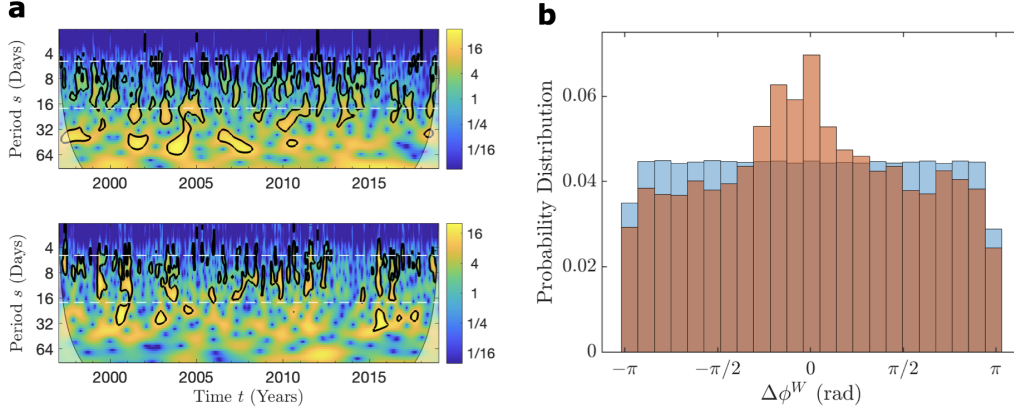
FIGURE 5.6.: (a) Spectra of continuous wavelet transform (CWT) coefficients of the CMZ (upper panel) and the NCN (lower panel) time series of their respective JJA precipitation for the years 1998-2019. The values exceeding the 95% significance level are marked with black contour lines. The high energy scales common to both time series lie approximately between the range of 6-16 days. (b) Probability distribution of the wrapped wavelet phase difference, $\Delta\phi^W$ computed for the common scales (orange), in a background of the distribution obtained from 200 surrogates of the data (blue); The observed histogram differs from the surrogate distributions at a significance level of $\alpha < 0.0001$ compared using a Kolmogorov-Smirnov test. The observed phase difference distribution has a distinct peak close to zero, implying that the CMZ and NCN precipitation time series are phase synchronized at these scales.

ing average), common to both regions, are found approximately in the range of 6-16 days (Figure 5.6a). The distribution of the wavelet phase differences wrapped in the interval $[-\pi, \pi]$ shows a visible peak close to 0 at these scales (Figure 5.6b), thereby indicating that the rainfall at the two regions are phase synchronized (also see Figure A.8).

Vertically Integrated Water Vapour Flux   We find that although there might be a slight lag of about 1-2 days in the occurrence of rainfall at NCN with respect to the CMZ (Figure 5.2c), there is mostly a local recycling of moisture from the Pacific Ocean and Indian Ocean sources respectively, as seen from the aggregate of IVT composite anomalies, not only for Days 0-2 (Figure 5.5b) but also for Days 0-14 (Figure 5.5d). Unlike most previous studies [238, 242], which identify the south-westerly moisture flow in the monthly IVT composites as the cause of the connection between rainfall over India and Northern China, we uncover from the composites of IVT anomalies for days of high synchronization that sufficient moisture is not transported directly via a tropospheric path. As expected, the above inference is not clearly evident from the IVT composites of Day 0 (see Figure A.7b), where the general direction of flow is still south-westerly. Dethof et al. [290] observed moistening of the lower stratosphere in the Asian monsoon region and suggested a northward moisture transport towards the Pacific from the monsoon anticyclone in July along the upper tropospheric jet stream.
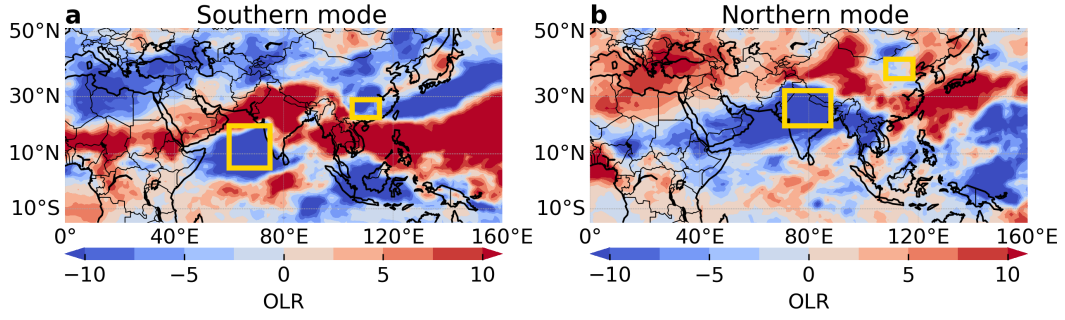
FIGURE 5.7.: Composite anomalies of outgoing long-wave radiation, with respect to JJA climatology, based on the days of high synchronization on Day 0, for (a) Southern and (b) Northern modes of the ISM-EASM connection. (Taken from Gupta et al. [P4])
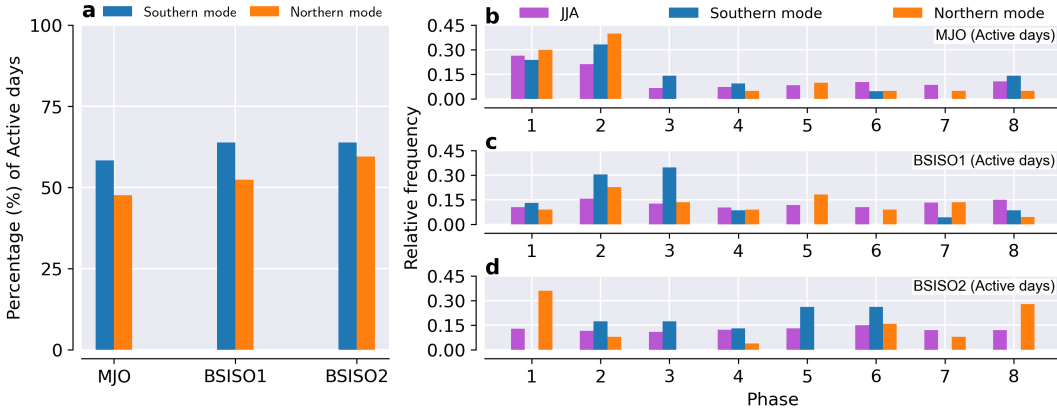
FIGURE 5.8.: (a) Percentage of active (amplitude > 1) MJO, BSISO1 and BSISO2 days among the specific days of high ERE synchronization for the Southern and Northern modes. MJO/BSISO phase distribution for days during the JJA season (for reference) and days of ERE synchronization for Southern and Northern modes which had *active* MJO (b), BSISO1 (c) and BSISO2 (d). (Taken from Gupta et al. [P4])

From the above observations of the composite anomalies of atmospheric variables, it is highly likely that the quasi-stationary Rossby waves are the common driver of extreme precipitation in CMZ and NCN, affecting the intensity of rainfall in both regions.

Role of intraseasonal oscillations

Tropical intraseasonal oscillations (ISOs) are an important source of variability of the Asian Summer Monsoon at short time scales [244]. They are mainly given by two modes: the Madden-Julian Oscillations (MJO) and the boreal summer ISO (BSISO). The MJO, propagating eastward along the equatorial region, in spite of having a weaker variability during boreal summers as compared to that in the boreal win-

ter [245, 246, 291], has a considerable influence on the variability of the monsoon at intraseasonal timescales [292–294]. However, BSISO [251, 295, 296], which has a prominent northward propagation and extends further northward from the equator, is the prevalent ISO mode responsible for the short-term climate variability of the monsoon [247–250, 297–299]. While the MJO has a periodicity of 30-60 days, the BSISO exhibits two different periodicities: BSISO1 with a periodic cycle of 30-60 days and BSISO2 recurring every 10-20 days. Several works have previously investigated the influence of the ISO on the onset of the ISM and EASM [251, 295, 300], the active/break phases of the monsoon [255, 301–303] and seasonal mean rainfall [227, 304, 305] of the ISM and EASM separately. However, less attention has been paid to its possible role in establishing a connection between the ISM and EASM. Many previous works have suggested that the ISO can not only be a possible triggering mechanism for the sudden onset of monsoon, but can also play a crucial role in the northward progression of the EASM [231, 306]. This poses an important question about the possible role of ISO in the shifting of the synchronization between the ISM and the EASM from the Southern to the Northern mode. In the following, we investigate the impact of the MJO and BSISO on each of the two discovered modes of ISM-EASM connection by classifying the days of high rainfall synchronization between both pairs of regions (ARB-SCN and CMZ-NCN) into active (amplitude > 1) and inactive MJO/BSISO days, and by classifying the active days according to the different BSISO/MJO phases (see Section 5.2.1).

Figure 5.7 shows the outgoing long-wave radiation (OLR) composite anomalies of both the Southern and the Northern mode of the ISM-EASM connection. The presence of a negative OLR anomaly, i.e., enhanced convective activity, simultaneously in both ARB and SCN (CMZ and NCN) indicates the synchronization of rainfall in case of the Southern (Northern) connection mode. From Figure 5.8a, we see that $\sim 40\%$ of the days of high rainfall synchronization are inactive in BSISO/MJO. This suggests that ISO is not the sole factor causing the two modes of ISM-EASM connection. This can also be inferred from the OLR composite anomalies of inactive BSISO/MJO days of high rainfall synchronization, which show simultaneous negative OLR anomalies in the pair of regions for each mode (see Figure A.9) aside from similar observations made from the OLR composite anomalies of active days (see Figure A.10).

Next, we classify the MJO/BSISO active days of high ERE synchronization (Figures 5.8b-d) on the basis of their different phases. We see that particular phases are associated with enhanced rainfall synchronization between ISM and EASM. MJO Phases 1 and 2 favour synchronization for both Northern and Southern modes (Figure 5.8b). Similar observations were made by [111] in case of link bundles connecting ISM with regions of East Asia and the northwest Pacific Ocean. In case of BSISO1, which has a similar periodicity as the MJO and is significantly correlated with it, we observe that there is a higher occurrence of extreme rainfall synchronization in Phases 2 and 3 for the Southern mode, and in Phase 2 for the Northern mode (see Figure 5.8c). However, it is interesting to note that in case of BSISO2, which has a higher variability, an opposite polarity in the phase distribution is seen (Figure 5.8d).

There is a relatively higher frequency of synchronized days for Phases 2 to 6 in case of the Southern mode. On the other hand, synchronization is enhanced during Phases 6 to 1 for the Northern mode. This suggests that while particular phases of the MJO/BSISO1 lead to an overall increase in the extreme rainfall synchronicity between ISM and EASM, BSISO2 possibly supports the shifting between the modes.

We have successfully investigated the spatial patterns of extreme rainfall synchronization in the Asian monsoon region using a combination of nonlinear event-based synchronization measures and a complex network-based approach, with a special emphasis on the interrelationship between the Indian and East Asian summer monsoons. We have also studied intraseasonal variability of the relationship between the ISM and the EASM, in particular the role of the intraseasonal oscillations which are associated with variations in deep tropical convection at these time scales. The main features of the ISM-EASM connection are summarized in Table 5.1. In addition to the above, it is of interest to also understand further the low frequency variability in tropical convection and its impact on ASM (annual, decadal, etc.). The El Niño-Southern Oscillation (ENSO) and the Indian Ocean Dipole are the dominant sources of interannual and interdecadal variability in the monsoon. In the following section, we investigate the role of ENSO on the spatio-temporal variability of the convective processes that drive the Asian monsoon system using a climate network approach.

5.3. Impact of El Niño-Southern Oscillation on the spatial connectivity pattern of Convection during Asian Summer Monsoon

ENSO has considerable impact on the spatial variability of ASM at interannual timescales. One possible way to investigate this could be by employing multiscale event synchronization (Section 2.2.2) on the precipitation dataset and analysing the resulting network at the timescales relevant to ENSO [112, 307]. However, the ENSO influences the ASM by directly or indirectly modulating the convective heat sources that power the monsoon system [278]. Therefore, in order to investigate the effect of the ENSO on ASM variability, it is more appropriate to understand the variability of the interaction between the convection anomalies during the monsoon season for different phases of the ENSO instead of that of precipitation which is the result of this modulation. A fairly straightforward way to do so is by analysing the interannual variability of the interaction structure of convection during the summer season by constructing separate networks of outgoing longwave radiation (OLR) for periods when ENSO is positive, negative and neutral. An important advantage of OLR networks over precipitation networks is that, unlike precipitation the OLR data does not have many zeros in its time series, and hence one can employ measures of synchrony for continuous time series instead of event-based synchrony measures (Section 2.2.2) which prevents any loss of information contained in the data due to the conversion

TABLE 5.1.: Summary of ISM-EASM connection.

Feature		Southern mode	Northern mode
Synchronization pathway		Arabian Sea and south-west coast of India (ARB) with parts of southeastern China (SCN; near the middle and lower reaches of Yangtze River)	Core monsoon zone of India (CMZ; central and northern parts) with northern China (NCN; near the Yellow River valley)
Time of high synchronization		June	July
Anomalous Atmospheric Circulation	250 hPa	Anomalous high over upper Indian Ocean, peninsular India and Maritime continent; strong westerly jet stream current north of high GPH region in the mid-latitudes; wave train originating from North Africa along the westerly jet	Undulating anomalously high GPH over Eurasia with anticyclone (Tibetan Plateau)-cyclone-anticyclone (central China) pattern; Eurasian wave-train, originating in the jet-exit region of the North Atlantic – dominant Silk Road teleconnection
	500 hPa	Western North Pacific subtropical high (WNPSH) stretching up to Bay of Bengal. Convergence of south-westerly winds from Somalian Coast with easterly winds in the southern edge of the WNPSH over the southwest coast of India	WNPSH further northwest over Pacific and Japan with a southwest protruding tilt, sustained low over Indian subcontinent
	850 hPa	Same as 500 hPa	Indian low, high over eastern Tibetan Plateau and low over NCN
Anomalous vertically integrated water vapour transport		Establishment of moisture corridor in about 2 weeks, transporting moisture from ARB to SCN via the Bay of Bengal and South China Sea	Mostly local circulation of moisture – CMZ from Indian Ocean sources and NCN from Pacific Ocean sources
Intraseasonal Oscillation	MJO	Phases 1 and 2	Phases 1 and 2
	BSISO1	Phases 2 and 3	Phases 2 and 5
	BSISO2	Phases 2 to 6	Phases 6 to 1

to an event series. In the following, a detailed explanation of the data and methods (Section 5.3.1), and the results so obtained (Section 5.3.2) are provided in this section.

5.3.1. Data and Methodology

We use the ERA5 reanalysis data for the JJA season during the period 1980-2020 at a spatial resolution of $1° \times 1°$ for OLR, and zonal and meridional components of wind at 850 hPa (U850 and V850 respectively). We compute the seasonal anomalies of the variable based on the seasonal climatology of the entire whole period. We then divide the whole JJA time series of the above considered climate variables into three separate time series containing the JJA seasons when ENSO was in the warm (El Niño), cold (La Niña) and neutral periods, respectively, based on the Oceanic Nino Index (ONI) [308] values for that season (Table 5.2). The index is computed from the 3-month running mean of sea surface temperature (SST) anomalies in the Niño 3.4 region (5°N-5°S, 170°W-120°W), and is based on centered 30-year base periods updated every 5 years.

TABLE 5.2.: Classification of JJA seasons for the period 1980-2020 into ENSO phases based on ONI.

El Niño	1982, 1987, 1991, 1997, 2002, 2004, 2009, 2015
La Niña	1985, 1988, 1998, 1999, 2000, 2007, 2010, 2011
Neutral	1980, 1981, 1983, 1984, 1986, 1989, 1990, 1992, 1993, 1994, 1995, 1996, 2001, 2003, 2005, 2006, 2008, 2012, 2013, 2014, 2016, 2017, 2018, 2019, 2020

We construct three separate networks of OLR by grouping together the El Niño, the La Niña and the neutral monsoon seasons separately (Section 3.5.2). To check if the statistical relationships are predominantly nonlinear, we make a comparison between networks constructed from the Spearman's rank correlation coefficient (Section 2.2.2) and those constructed from mutual information estimated from the k-nearest neighbour distances (Section 2.2.2) for the same region as considered for the precipitation networks. Both these measures are suitable for OLR data whose distribution is non-Gaussian. We consider both positive and negative correlations, while mutual information only gives positive values. We only retain those values in the similarity matrix which are statistically significant ($p < 0.05$) by testing against randomly shuffled time series surrogates. Finally, if the value is higher than the 95th percentile of the significant values in the similarity matrix, the corresponding pair of nodes are considered to be linked. The results presented here do not consider lagged synchronization, i.e., lag is zero. The networks are therefore unweighted and undirected. However, the patterns remain mostly unchanged when lags are allowed to vary up to 14 days which might be because ENSO can have a delayed effect on ASM by more than a season. Then, we compute and plot the *degree* of the nodes (Eq. (3.4) in Section 3.3) for each network (Figure 5.9). It is of interest to know which regions are connected with a region of high degree, in which case we compute the *partial degree* for that region (Section 5.2.1).
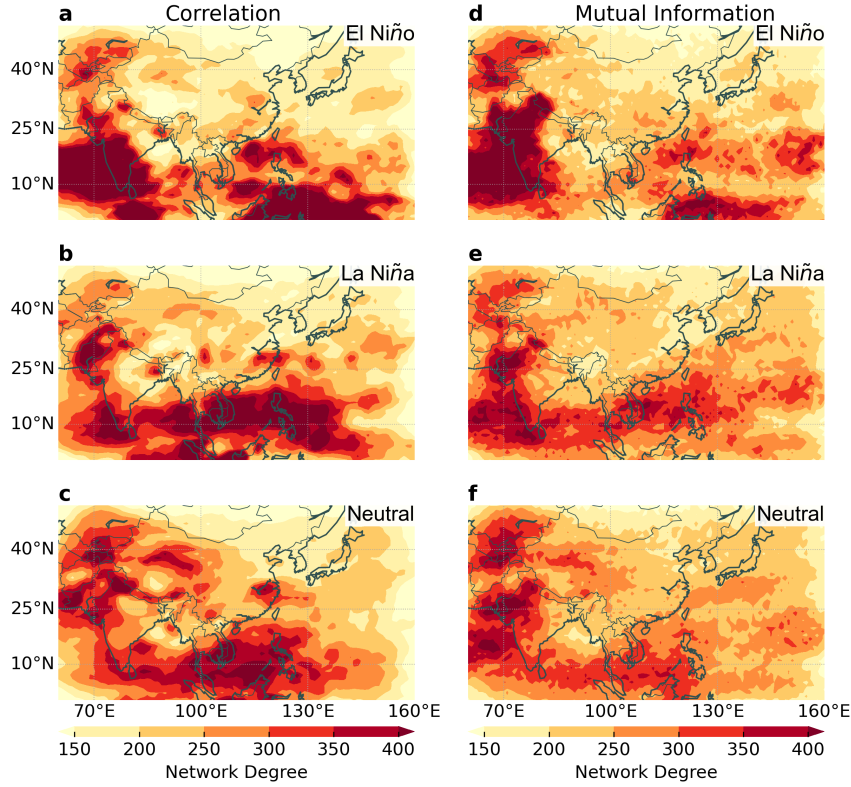
FIGURE 5.9.: Spatial patterns of degree for the ASM region for functional networks of OLR constructed using (a-c) correlation and (d-f) mutual information for different ENSO phases during the JJA season. (Taken from Gupta et al. [P5])

As seen from Figure 5.9, the spatial patterns of degree for the correlation-based networks fairly resemble those of the mutual information based networks. Therefore, most of the connections are not strongly nonlinear [30]. It must be noted that most high degree features in Figure 5.9, are connected to the boundaries of the network. Therefore, it can be assumed that the pattern is prone to boundary effects. The impact of the boundaries on the network measures can be checked by using the correction procedure outlined in Section 3.5.4 to correct the effect of spatial embedding on the network [130, 148]. However, the patterns remain fairly unchanged even after employing the correction (Figure A.11). A probable reason for this is that monsoon is a large-scale phenomenon unlike tropical cyclones which are highly localized events (Chapter 4), so the error induced due to the boundaries have significantly much less impact on the large-scale features of the monsoon than the spatially smaller features of tropical cyclones. Furthermore, this also implies that the patterns of degree of the OLR networks are part of a larger pattern which extends outside the considered region. So it is more suitable to extend the analysis to a larger region which includes the adjacent Indian and Pacific ocean regions along with the Nino 3.4 region (35°N-

10°S, 35°E-120°W). However, this will significantly increase the computation cost if the networks are calculated using the computationally expensive mutual information and if the network measures need to be boundary corrected. Based on the aforementioned analysis and arguments, we can safely conduct all further investigations using only correlation-based networks without correcting for boundary effects.

Since convection and circulation are closely related, we also refer to observations from correlation networks of the lower tropospheric wind components (Sections 3.5.2 and 3.5.3), constructed for the same period and region, in the interpretation of the results.

5.3.2. Results and Discussion

The spatial patterns of degree of the OLR networks during JJA season for the different ENSO phases are shown in Figure 5.10. The common features for all three phases are: (i) a continuous high degree region along the tropical east coast of Africa, the Horn of Africa, the Arabian Sea, west coast of India extending up to north-western India and a part of the Bay of Bengal, (ii) regions of high degree in the tropical eastern Indian Ocean and Maritime continent, and (iii) patches of high degree in the eastern Pacific Ocean – one in the subtropics in the Northern Hemisphere and the other close to the equator in the Southern Hemisphere.

However, there are several topological differences as well – the degree in the western Indian Ocean is higher during El Niño (Figure 5.10a) than during other phases. The high degree region over the Maritime continent is predominantly over the Indonesian archipelago around the equator during El Niño (Figure 5.10a) while it is split into two during La Niña, one over the eastern Indian Ocean and parts of Indonesia and the other over the South China Sea and the Philippines Sea (Figure 5.10b). The ENSO neutral OLR network has comparatively lesser degree over the Maritime continent (Figure 5.10c). The high degree regions in the eastern Pacific also has higher values during La Niña (Figure 5.10b). Next, we investigate the regions which are connected to the high degree regions in the OLR networks. The partial degree corresponding to the major high degree regions in the OLR network for the ENSO positive phase (Figure 5.10a) is shown in Figure 5.11. We see from Figures 5.11a,b that the convection anomalies in the stretch of high degree nodes of the western Indian Ocean from the Somalian coast to lower Bay of Bengal are connected to the high degree nodes of the Indonesian archipelago. However, this connection is highly reduced to almost zero in the La Niña network (Figure 5.12a,b). Instead, the Indonesian archipelago shows higher connections to the equatorial central Pacific (close to 170°W; Figure 5.12b). The dominant high degree region over Philippines and adjacent seas in the La Niña network however is only locally connected and seems to be associated roughly with the position of the WNPSH (Figure 5.12c). Interestingly, the region also seems to have connections, although very less, with the high degree patch in the subtropical north-eastern Pacific (Figure 5.12c,d).

The highly connected high degree region in the western India Ocean stretching from the lower Bay of Bengal to Arabian Sea, up to the Somalian coast, can be understood
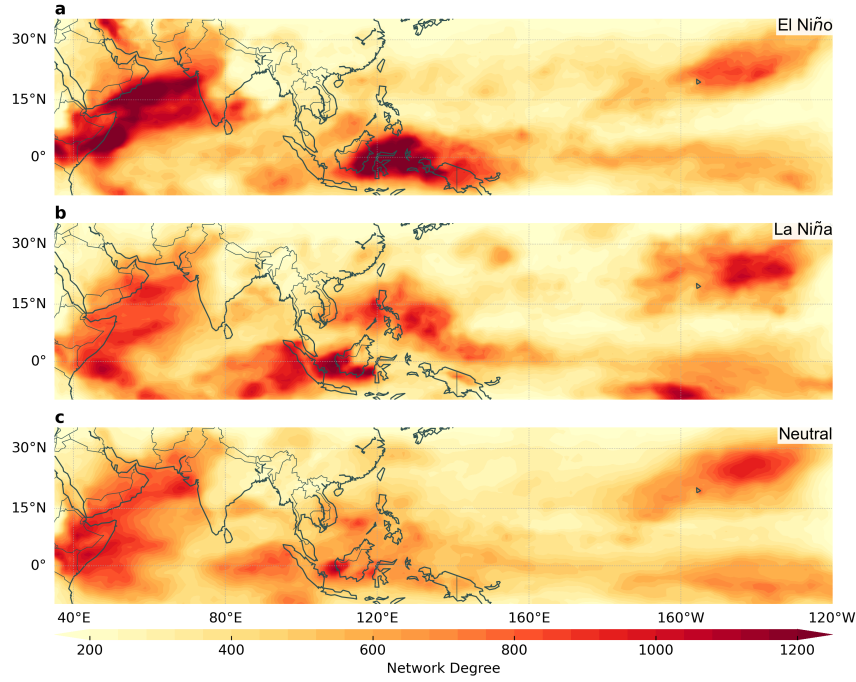
FIGURE 5.10.: Spatial patterns of degree for functional networks of OLR constructed for the extended region of ASM, which includes the adjacent oceans, for JJA season when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral. (Taken from Gupta et al. [P5])
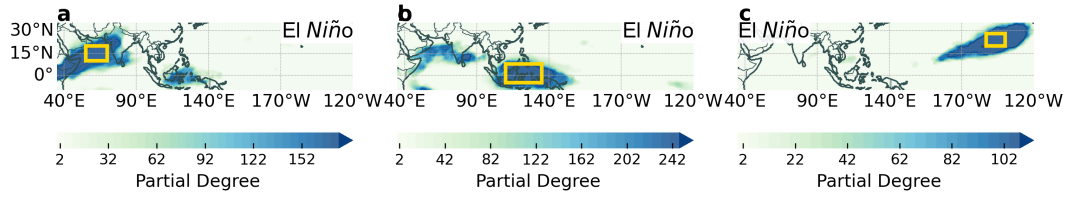
FIGURE 5.11.: Partial degree plots corresponding to the high degree regions of Figure 5.10a: (a) Arabian Sea, (b) equatorial Maritime Continent and (c) subtropical north-eastern Pacific Ocean. (Taken from Gupta et al. [P5])
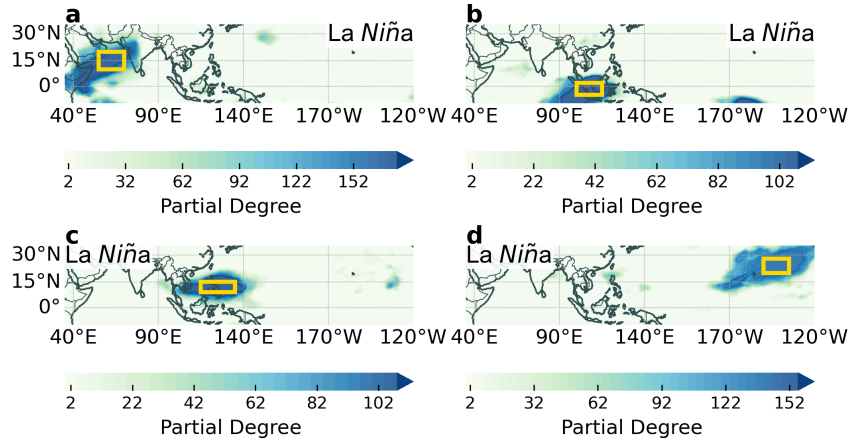
FIGURE 5.12.: Partial degree plots corresponding to the high degree regions of Figure 5.10b: (a) Arabian Sea, (b) equatorial Maritime Continent, (c) Philippines Sea and (d) subtropical north-eastern Pacific Ocean. (Taken from Gupta et al. [P5])

in the following way: Enhanced convection over the ISM region is related to the magnification of the 850hPa Somalian jet and the southerlies over the Bay of Bengal, as also observed from the high degree regions in these locations in the V850 network (Figure A.15). As a response to this enhanced convection over the ISM region, the 850hPa westerlies strengthen to the west of 80°E over a narrow latitude band between 5°-15°N (see high degree in this region in the U850 network in Figure A.12 and also Figure A.13a).

The connection between the Somalian jet-Arabian Sea and the Indonesian archipelago regions in the OLR network of the ENSO warming phase (5.10a) is likely to be related to the longitudinal shift in the upward branch of the Walker circulation along with the suppressed convection over these regions [256, 309], which may cause a latitudinal shift in the position of the intertropical convergence zone (ITCZ) in the eastern Indian and western Pacific Ocean, which in turn would influence summer monsoon circulation in south Asia and western North Pacific [252].

Although the high degree region over Philippines in the OLR La Niña network does not have connections to other regions, the region is well connected to Malaysia with a tail extending westward to southern India and the Arabian Sea in the 850 hPa zonal wind network (Figures A.12, A.13b and A.14a). This connection in the U850 network, however, is reduced during the El Niño phase (compare Figure A.13b and A.14a). Philippines and the adjacent seas are known to play a crucial role in the mechanism by which the ENSO impacts the ASM [261, 278]. A large-scale low-level anticyclone forms over the Philippine Sea in the post-El Niño summers resulting from a Rossby-wave response to suppressed convective heating [310], which has a prolonged impact on the ensuing early summer Meiyu/Baiu, for example the devastating 1998 floods in Yangtze River valley. However, it is not clear why the region shows up as an isolated high degree area in the La Niña network and probably requires a more

in-depth analysis taking into account whether the ENSO is emerging or continuing, including the role of the upper tropospheric waves. Nonetheless, it is fairly instructive that the OLR network analysis distinctly identifies the key regions associated with the ENSO-ASM relationship.

Moreover, the role of the high degree region over subtropical eastern Pacific Ocean in the OLR networks is not clear and might be associated with climate anomalies over North America which has not been included in the analysis. It may be speculated that its low connectivity with the South China Sea-Philippine sea region may be associated with the Pacific-North American teleconnection [278].

It is thus clear that the high degree regions of OLR networks in the Indian and western Pacific regions highlight the two main convective heat sources that power the Asian summer monsoon system. These are known to distinguish between the ISM and western North Pacific-EASM variability [253]. The ISM variability is governed by the Bay of Bengal, India and Arabian Sea region's heat sources. On the other hand, the South China Sea and the Philippine Sea's convection center govern the western North Pacific-EASM variability. The topological differences in the OLR networks of the different ENSO phases suggest that ENSO impacts ASM variability by altering these convective heat sources either directly or indirectly. This is consistent with the findings of Wang et al. [278].

Additionally, the network topology of the U850 and V850 networks (Figures A.12, A.15) bear good resemblance to the one-point correlation maps in Figure 5 of Wang and Fan [253] and can be explained similarly.

5.4. Summary

We have analyzed the spatial synchronization patterns of extreme rainfall in the Asian monsoon region and the convective processes during the monsoon season. Firstly, we have constructed complex networks of extreme rainfall events using event synchronization for the June-July-August season. Using the network measure, degree, i.e., the number of network links attached to each location, we identified two distinct synchronization modes between the Indian and the East Asian monsoon systems along with the specific times when each synchronization pathway becomes dominant. This allows us to distinguish between the specific large-scale atmospheric circulation patterns related to each mode (Table 5.1):

1. *Southern mode:* The synchronization pathway between the Arabian Sea and southeastern China (middle and lower reaches of the Yangtze River valley) is dominant in June. The associated atmospheric circulation patterns are those which lead to the onset of summer monsoon over India. A moisture corridor between the west Indian Ocean sources (Somalian jet and the Zanzibar current) and the western Pacific sources (South China Sea) is established via the Bay of Bengal in about two weeks from the Indian monsoon onset, leading to the onset of Meiyu.

2. *Northern mode:* As the Indian summer monsoon progresses inland in July, the path of extreme rainfall synchronization shifts northward, between the core monsoon zone of India and northern China (near the Yellow River valley). While there is no substantial transport of anomalous moisture directly via a moisture pathway, the strengthening of the upper-level wave train due to stationary Rossby waves (the *Silk Road teleconnection* [279, 286]) leads to synchronous extreme rainfall conditions over both regions. The rainfall in the two regions are also found to be phase synchronized at intraseasonal timescales.

We also investigated the role of the tropical intraseasonal oscillation in the modulation of extreme monsoon precipitation over India and East Asia and in the mutual interaction between the two monsoon systems by using the phase distribution of the days of high rainfall synchronization. Through our analysis, we have shown that extreme rainfall events over the Asian monsoon region are favoured by certain phases of the lower frequency mode (MJO, BSISO1), while the higher frequency mode (BSISO2) may support the switch between the two connection modes. However, more detailed investigation is needed to gain a deeper understanding of the extent of the role played by the intraseasonal oscillation in the synchronization between the Indian and the East Asian summer monsoon.

This study has provided valuable insights into the intricate relationship between the Indian and East Asian components of the Asian summer monsoon system. Further extension of this approach, e.g. to shed new light on the relation between the Indian monsoon and rainfall associated with the Baiu season over Japan, understanding the role of Rossby waves as a common driver of extreme rainfall over India and East Asia, as well as the construction of prediction schemes using the above findings can be outlined as promising future lines of research.

Additionally, we have studied the spatial connectivity patterns of the convection anomalies during the monsoon season which has enabled us to identify the major convective heat sources which differentiate the variability of Indian summer monsoon from that of the western North Pacific-East Asian summer monsoon. The Arabian Sea, southern Bay of Bengal and the Indian subcontinent are the main convective heat sources of the Indian monsoon which is also related to the Somalian jet. On the contrary, the convective system over the Philippines and the adjacent seas regulate the western North Pacific and East Asian monsoon systems.

We further investigated the impact of the El Niño-Southern Oscillation (ENSO) on the interactions among these convective sources by distinguishing between the functional network structures of outgoing longwave radiation during different ENSO phases. Our analysis revealed that changes in the connection between convection anomalies of the Arabian sea and those over the Indonesian archipelago occur as the phase of the ENSO changes which implies that the ENSO-ISM relationship is related to shift in limbs of the Walker Circulation, and thus the position of the inter-tropical convergence zone. On the other hand, the relation between the ENSO and western North Pacific-East Asian monsoon systems occurs through changes in the lower tropospheric circulation pattern over the Philippines and South China Sea. Here, we

have explored the relation between the connections among the convective processes and those between the lower troposphere wind anomalies. However, from the results of extreme rainfall connectivity patterns, we found that the Indian monsoon affects its East Asian counterpart via the upper tropospheric Rossby waves. It may be of interest to also investigate the impact of these changes in the convection anomalies due to the ENSO on the Silk Road pattern, or in general, geopotential height and wind patterns in the upper atmosphere, as a future scope of this work.

It must be mentioned here that the convection connectivity structure observed here is most likely the accumulated result of the series of processes triggered by the sea surface temperature anomalies in the equatorial Pacific in the previous months because the whole mechanism by which the ENSO impacts the Asian summer monsoon outlasts a season. In this context, a possible future direction of the investigation is understanding the changes in the interaction between the convection anomalies in the spring or winter months due to ENSO, prior to the summer monsoon. Moreover, the difference in the impact of the decaying or emerging ENSO phases, and further distinction between the central Pacific and eastern Pacific episodes of the ENSO could also be explored.

To summarize, the complex network analysis of extreme precipitation events and convection anomalies provides a comprehensive understanding of the interaction within the Asian Summer monsoon region at intraseasonal and interannual time scales. It helps us identify the crucial regions which play a major role in the variability of the different components of the ASM. Such a detailed understanding of the underlying mechanisms of the spatio-temporal variability of the Asian monsoon is important to improve the subseasonal and seasonal forecasting of extreme precipitation.

6. Spatial variability of error correlations

"Until I know this sure uncertainty, I'll entertain the offered fallacy."

– William Shakespeare, *The Comedy of Errors*

6.1. Introduction

Understanding the properties of errors is an integral part of numerical weather prediction in order to produce better forecasts. The state of the system at one location can have an impact on the state at another location because the different components of the Earth are intricately linked to each other. Similarly, the error in forecasting the state of the system at one place can also affect that at another place. Therefore, error correlations may occur, often due to common error between measured values due to a common measurement [311], which are indicative of systematic or structured random errors [312]. Estimating error correlations is a crucial step for producing quality forecasts and is a key issue for data assimilation [313]. Particularly, proper estimation of the spatially varying component of the error is important to account for inhomogeneities and anisotropies in order to further improve the forecasts [314–316]. However, it has been a significant challenge to diagnose the full geographical variations of error correlations because the total number of elements of the correlation matrix is overwhelmingly large (square of the number of grid points), making it challenging to effectively represent them on geographical maps [317]. An economical approach commonly used to obtain a synthetic view of the spatial variability of error correlations is to estimate the local correlation length scale and then study its latitudinal variations [317–319].

In this chapter, we propose a different methodology for examining and visually understanding the spatial variability of error correlations. We utilize the climate network framework (Section 3.5.2) which has been used to study patterns of statistical [30] and causal interactions [256] in different spatio-temporal climate datasets such as temperature, pressure, geopotential height, wind and precipitation [111, 144, 146, 262, 320] (Chapter 4-5). As the network representation of a spatio-temporal climate dataset essentially represents the pairwise interaction among the grid points estimated

---

This chapter is mostly based on the associated work Gupta et al. [P7], which is currently under review in the Quarterly Journal of the Royal Meteorological Society. All sections in this chapter except Section 6.5 closely follow this work.

using causality measures like correlation, mutual information, event synchronization, etc., the concept of error correlation is similar to the network representation of forecast errors. Therefore, the wide range of complex network measures derived from graph theory that are used to analyse the topological features of the complex network [30] can be readily employed to analyse the spatial structure of error correlations.

Using several examples, we show that the analysis of error correlations using complex network-based approach can provide an understanding of the primary source of error affecting the variable. The method enables us to identify the important areas which exhibit high influence on the error properties of the whole region under consideration for a given climate variable. Although, it may be argued that the most erroneous regions of forecast could be identified instead by comparing the depiction of climate interactions in the reanalysis data with that in forecasts, such a method helps to identify the statistical links which are underestimated or overestimated, and hence the area affected due to them allowing us to detect systematic errors, especially model deficiencies, which can otherwise get obscured in error correlations due to smaller orders of magnitude [312]. Therefore, the latter method is effective for evaluating model performances [56–58, 321–323]. This is different from the proposed methodology of error networks as the spatial interaction pattern of error network may be different from that of the climate variable itself. This is shown by comparing the topology of the error correlation network with that of the respective reanalysis network, where they are found to exhibit varying levels of differences between each other for different variables, thereby implying that the spatial variability pattern of errors may not be directly inferred from that of the corresponding climate variable.

Additionally, it is also of interest to see how the different phases of the El Niño-Southern Oscillation (ENSO) affect these error correlation structures, as ENSO is known to impact predictability of different regions. Our analysis is primarily focused on analysing error properties of the Asia-Pacific region in the Northern Hemisphere in the summer months of June-July-August during which the southern and eastern parts of Asia experience the monsoon. However, it is possible to extend the analysis for the whole globe and also investigate the seasonal dependence of the errors. Our findings demonstrate the potential of the climate network approach as a diagnostic tool for gaining a preliminary understanding of the origins of forecast errors. We are unaware of any previous research that has used the climate network framework to examine the spatial characteristics of forecast error.

The remainder of the chapter has been organized as: In Section 6.2, we list the employed datasets and then outline the methodology. In Section 6.3, we first show that error correlation networks of different climate variables exhibit significant spatial patterns, and then we discuss our findings on the properties of errors mostly affecting the Asian Summer monsoon from the analysis of these networks. Thereafter in Section 6.4, we make a comparison with the error information inferred from the difference between the reanalysis and forecast networks of the same variables, thus demonstrating that the analysis of error correlations using network measures reveal crucial information about the underlying sources of errors. Next, we divide the period
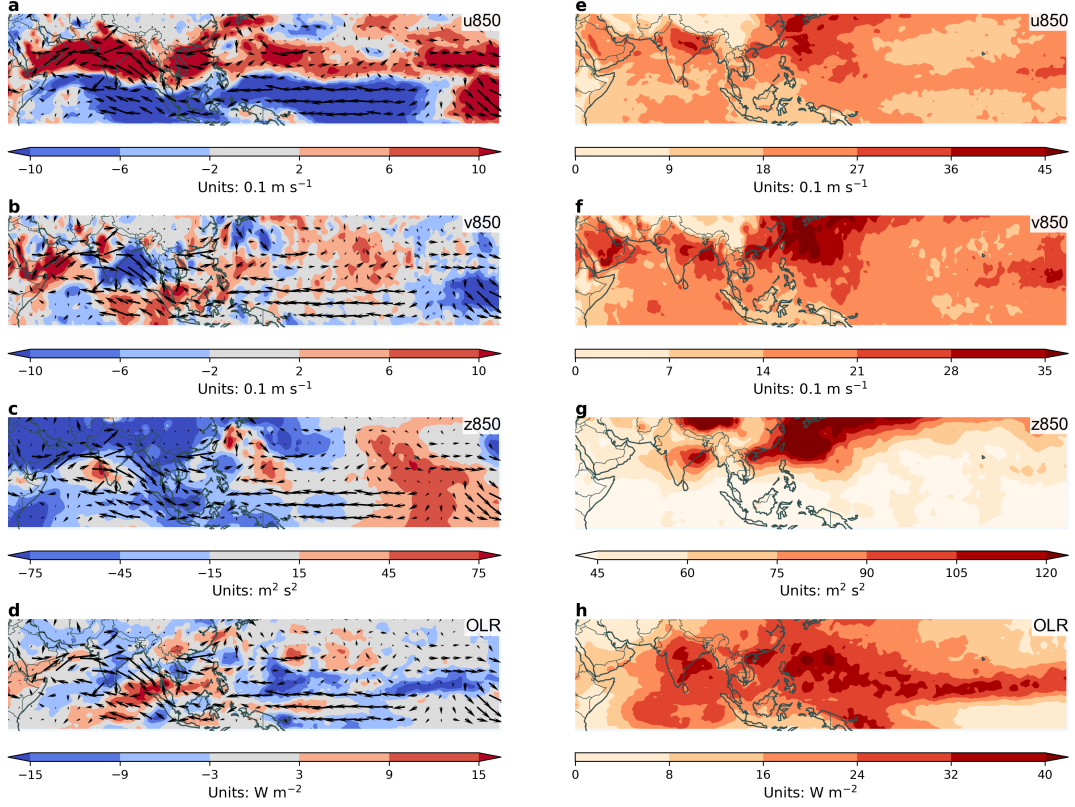
FIGURE 6.1.: (a-d) Day 5-6 mean forecast error and (e-h) mean absolute error for day 5-6 bias corrected forecasts for JJA 2018 for (a) U850, (b) V850, (c) Z850 and (d) OLR. (Taken from Gupta et al. [P7])

of analysis according to the phases of the ENSO and compare the corresponding error network topologies in Section 6.5. Finally, in Section 6.6, we provide some concluding remarks regarding the relevance and the future scope of the work.

6.2. Data and Methodology

6.2.1. Data and pre-processing

We use the ERA5 reanalysis data [162] at a spatial resolution of $1° × 1°$ and daily mean of the hourly values for the period 1980-2020, for outgoing long wave radiation (OLR) as well as a few lower tropospheric (at 850 hPa) climate variables, namely, the geopotential (Z850), and the meridional (V850) and zonal (U850) components of wind. Forecast data for the same variables is obtained from the 10-day forecasts produced from the same system by averaging over the time interval 120-144 hours (day 5-6). The forecast lead-time of day 5-6 is chosen to focus on large-scale errors in the medium-range and also to get error propagation that is beyond linear advection of

error structures [324]. Daily anomalies of the variables were computed with respect to the daily climatology of the whole period of analysis. The forecast errors are computed by subtracting the reanalysis from the forecast data (Figure 6.1a-d). The effect of the mean bias in the forecast model is removed by the subtracting the mean error from the forecast error time series. The absolute of the resultant values in the error time series is used for our subsequent analysis (see Figure 6.1e-h).

The error properties during the June-July-August season are analysed using our climate network approach. The analysis is restricted to the Asian Summer Monsoon region and the adjacent Indian and Pacific Oceans, which play an important role during the monsoon. The region of interest extending from 35°N to 10°S and from 30°E to 120°W, includes the Nino 3.4 region (5°N-5°S, 170°W-120°W) as El Niño-Southern Oscillation is known to considerably affect the inter-annual variability of the Asian Monsoon [252] (Section 5.3). The Oceanic Nino Index is used to identify ENSO phases, which is based on variations in 3-month running means of sea surface temperatures in the Niño 3.4 region [308].

6.2.2. Methods

The spatial coherence pattern of the reanalysis, forecast and forecast error data are obtained using climate network based approach outlined in Section 3.5. The specifications of the network construction and the measures used to analyze and compare the network topologies are described below.

Network Construction

We construct the functional network representation (Section 3.5.2) of the spatio-temporal climate data sets described above, wherein the spatial grid points of the ERA5 data are the nodes of our network. We calculate the links of our network by computing the Spearman's rank correlation coefficient (refer to Eq. (2.12) in Section 2.2.2) between different pairs of nodes at zero lag. We perform a significance test on our data and construct the correlation matrix by retaining only those correlations whose $p-$values are less than 0.05. We choose the threshold τ such that strongest 5% correlations (ignoring the sign of the correlation) are retained and the corresponding pairs of nodes are considered to be connected. Then the adjacency matrix A representation of the network is obtained as described in (Section 3.5.2). We note that $A_{ij} = A_{ji}$, i.e., the adjacency matrix is symmetric. This results in an undirected, unweighted network.

Network Measures

Following the construction of the climate networks, we compute the network measure *degree k_i* (Eq. (3.4)). As in most cases, the high degree nodes play an important role in the functioning of the system, the degree centrality can be a useful guide for focusing our attention on the system's most crucial regions, which in this case have

either strong influence on the forecast errors of other regions or are strongly influenced by them. In order to find the areas interacting with a region R of high degree in the climate network, we calculate the *partial degree* $\{k_i\}_R$ of the nodes in the network, which yields the number of links connecting a node i outside R with the nodes within R.

If the errors are uncorrelated, as in case of independent random errors, the forecast error network would be completely random. However, if the errors of different grid points have a partially predictable relationship between them due to an underlying process which is either deterministic or a random process with structures, then the errors exhibit correlated structures which will get depicted as definite spatial patterns in the forecast error network. Since only the highest correlations to construct our network are preserved, the most dominant structured errors causing the strong spatio-temporal correlation in the forecast error of the climate variable are expected to show up in the spatial distribution of high degree nodes in the network. The significance of the degree distribution of the forecast error network can be tested against that of random networks (Section 3.4), as they do not contain nodes with unusually high degree ('hubs') unlike real-world networks [41]. Therefore, the degree distribution of the original network can be compared with the mean degree distribution of Erdős–Rényi networks [129] with the same number of nodes and average degree obtained by rewiring the links of the original network entirely randomly. However, it is more appropriate to test the significance of the degree distribution of the original climate network against that of spatially embedded random networks [130, 148], as climate networks are spatial networks, and hence the influence of spatial embedding on the network structure due to distance-based costs of links should be taken into account (Sections 3.5.4 and 3.2).

We also measure of preferential connectivity in the networks based on the node degree by computing the *degree assortativity coefficient r* in order to estimate the level of homogeneity/heterogeneity of real networks (Section 3.3), as a more assortative network is more homogeneous and vice-versa.

Finally, we calculate the *Common Component Function* (CCF), which measures the number of common links between a pair of networks [128] in order to quantitatively compare how similar the network topologies of the two networks are (refer Section 3.3). This will be useful here to quantify the degree of similarity between the reanalysis, forecast and forecast error networks.

6.3. Spatial patterns of Error correlation networks

The forecast error networks of different climate variables are shown in Figure 6.2a-d. They exhibit definite significant spatial patterns of degree unlike random networks (Figure 6.3). This indicates the presence of an underlying systematic or random process that leads to spatio-temporally interrelated errors. We see that the degree distribution of the forecast error networks (Figure 6.3) have long right tails, indicating that fewer nodes have very high degrees. We focus on the high degree nodes in the
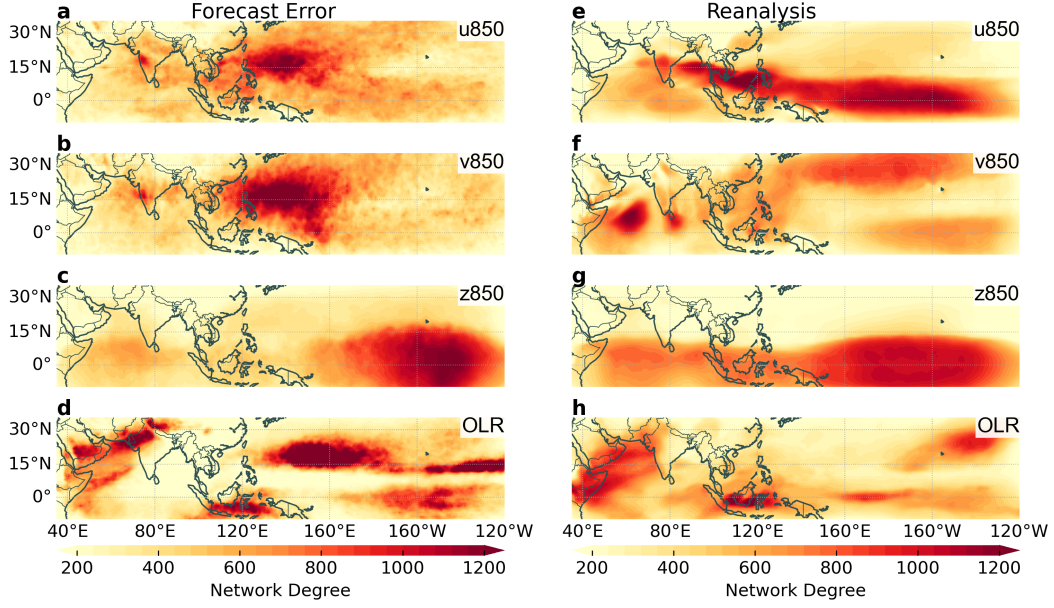
FIGURE 6.2.: Spatial patterns of network measure degree of forecast error networks for (a) U850, (b) V850, (c) Z850 and (d) OLR. Spatial pattern of network degree from reanalysis data for (e) U850, (f) V850, (g) Z850 and (h) OLR. (Taken from Gupta et al. [P7])

forecast error networks of Figure 6.2 which are important regions affecting the behaviour of the forecast error patterns of the whole region. Since the network measure degree counts the number of links of each node, representing pairwise interactions, the dominant errors causing a pattern of influence on the forecast of multiple regions show up as a particular pattern of locally or distantly connected high degree nodes in the network.

In case of both the wind components at 850 hPa (U850 and V850), the part of western North Pacific Ocean adjacent to Southern China and the Maritime Continent exhibits the highest degree (Figures 6.2a-b). Furthermore, we find that the connections of these high degree nodes are limited to the nodes in this region only (Figures 6.4a-b). The mean absolute forecast error of U850, V850 and Z850 in Figures 6.1e-g shows that the region a bit more north of the highest degree regions of the wind error networks is associated with large error in the Western Pacific Subtropical High (WPSH). While the area of the WPSH in Figures 6.2a-b shows higher degree than most of the other regions, it appears that the southern boundary of the WPSH exhibits the highest correlated error structures as seen from the region of the highest degree. WPSH is an important circulation pattern which affects the Asian summer monsoon and the tropical cyclone activities in the highly active western North Pacific Ocean basin. Recent studies [184, 325, 326] have exhibited that subsequent track forecast errors of tropical cyclones are strongly sensitive to the small initial errors in the predictions of WPSH, which can fluctuate on synoptic time scales, therefore, is prone
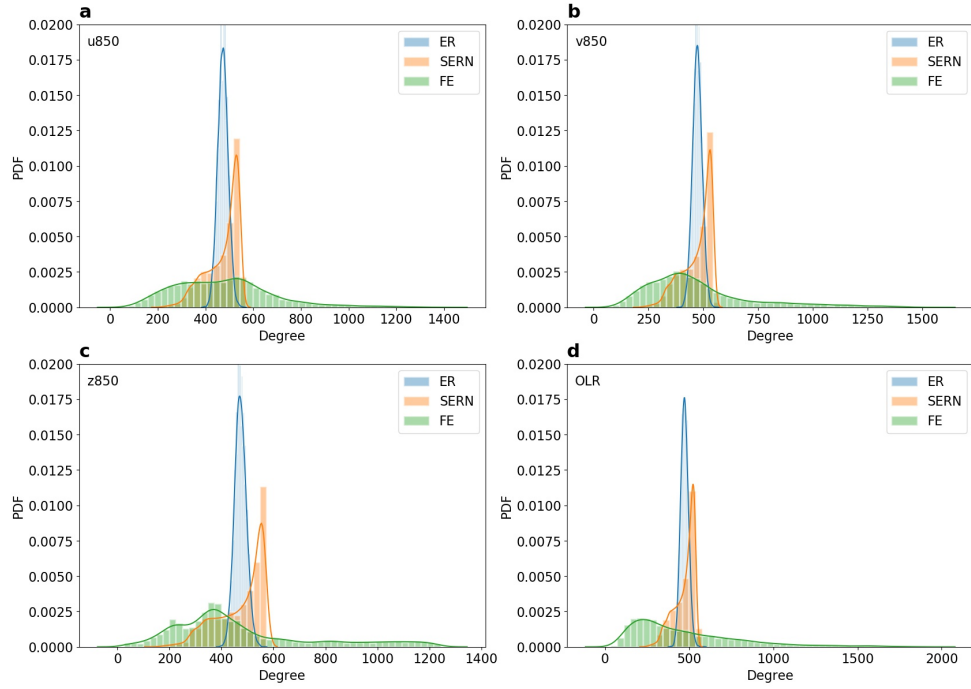
FIGURE 6.3.: Degree distribution comparison of forecast error network (FE) with the mean distribution of 100 Erdős–Rényi (ER) networks with the same number of nodes as FE and random rewiring of the original links, and that of 100 spatially embedded random networks (SERN) with the same number of nodes and link distance distribution as FE for (a) U850, (b) V850, (c) Z850 and (d) OLR. (Taken from Gupta et al. [P7])
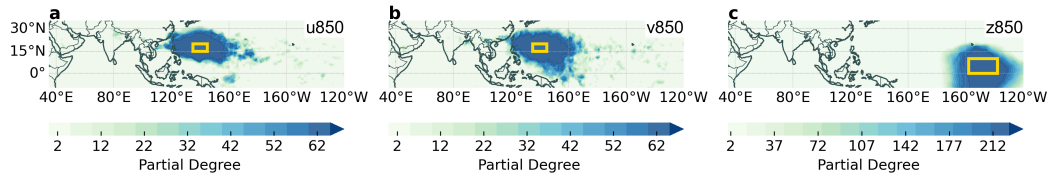
FIGURE 6.4.: Partial degree associated with the regions in yellow boxes showing the areas connected to those regions in the (a) U850, (b) V850 and (c) Z850 forecast error networks. (Taken from Gupta et al. [P7])
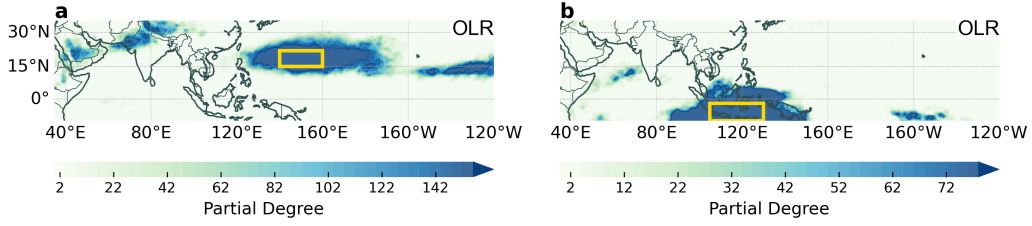
FIGURE 6.5.: Partial degree associated with the regions in yellow boxes showing the areas connected to those regions in the OLR forecast error networks. (Taken from Gupta et al. [P7])

to non-systematic errors [184]. Gao et al. [327] also showed that there are systematic biases in the WPSH forecasts, such as a smaller area, eastward and southward shift of location. The predictability of WPSH is a pre-requisite for the improved prediction of not only western North Pacific tropical cyclones but also that of the Asian summer monsoon rainfall [328]. The small high degree region over south-west India is locally systematic or correlated and may be related to the wind errors associated with onset of monsoon over India.

However, the region of WPSH is not seen to exhibit high degree in the forecast error network of Z850 (Figure 6.2c). Instead, we see a high degree in the region of equatorial central Pacific Ocean, along with a lighter patch of comparatively lower degree in the equatorial Indian Ocean. These regions are also associated with large mean forecast errors in Z850 (Figure 6.1c), and show diverging wind error fields. This suggests that the error might be related to the inter-tropical convergence zone (ITCZ) as also seen in the mean absolute forecast error of OLR (Figure 6.1h). The connections of the high degree nodes in the equatorial central Pacific region is limited to the nodes in the same region (Figure 6.4c), although, the connectivity structure appears to be much smoother than those of the wind error networks (Figures 6.4a-b). Furthermore, even though all the forecast error networks (Figure 6.2) are assortative, i.e., high degree nodes have a tendency to be connected to high degree nodes, as seen from Table A.1, the Z850 forecast error network (Figure 6.2c) particularly has a very high degree assortativity of $r = 0.8$ indicating more homogeneity than the other networks. Such a homogenous large scale connectivity structure in the geopotential height error network is possibly due to errors in the forecasts of the tropical circulation associated with a lack of direct observations of wind profiles and with complex tropical dynamics. As the coupling between the geopotential height and winds (quasi-geostrophic balance) is weaker in the tropics, the flow has to be decomposed into a balanced component, represented by the equatorial Rossby waves, and an unbalanced component consisting of inertio-gravity (IG) waves [329]. The representation of the unbalanced tropical circulation, which is used to describe tropical variability in both the atmosphere and the oceans on all scales, is a source of significant uncertainties in both weather forecasts and climate models [329, 330]. In particular, the spatial pattern of degree of the Z850 error network (Figure 6.2c) is similar to the geopotential height perturbation

at 850 hPa of the zonal wave number 1 equatorial Kelvin wave [329, 331], which is believed to be the main driver of the Madden-Julian oscillation system [332, 333].

The forecast error network of OLR (Figure 6.2d) shows multiple regions of very high degree. There are three separate patches of high degree in the Pacific Ocean – two in the Northern Hemisphere and one just below the equator. The grid points in the Maritime Continent, northwest India and Pakistan, along the moisture-bearing cross-equatorial south-westerly low-level jet coming from the Somalian coast also exhibit high degree. We find the regions connected to these high degree regions by computing the partial degree of some nodes within the region (Figures 6.5 and A.16) and observe that the three high degree regions in the Northern Hemisphere are interconnected to each other. Similar behaviour is observed for the three high degree regions in the Southern Hemisphere, although the number of links are comparatively less. These regions coincide with the position of the ascending/descending limbs of the Walker Circulation (Figure 6.6) of the ENSO neutral years, classified based on the Oceanic Niño Index [308].

The association of the spatial pattern of degree of the OLR forecast error network to the Walker circulation points towards cloud biases in the forecasts which leads to systematic errors in the simulation of the ITCZ, as seen from Figure 6.1h. The mis-representations of ocean-atmospheric interactions and the thermodynamic processes in the equatorial Pacific such as too strong equatorial zonal surface winds give rise to the cold tongue bias in the forecast models which extends westward and is stronger in the boreal summer [334]. [328] in their investigations using empirical orthogonal function (EOF) analyses of boreal summer geopotential height at 850 hPa showed that an anomalous central Pacific cooling shifts the rising limb of the Walker circulation cell westward. This suppresses convection high degree region around 160°E (Figure 6.2d), and enhances convection over the Maritime continent [335] (high degree region in Figure 6.2d), both of which can strengthen the WPSH, the former via generating descending Rossby waves and the latter by inducing equatorial easterlies over western Pacific [336]. The above suggests a possible mechanism by which conditional biases in one of the processes might induce error in different variables, which then show up as different topological patterns in their forecast error networks.

6.4. Effect of statistical relationships in Reanalysis/Forecast data on Error correlations

Although, from the aforementioned discussion, it is evident that the spatial connectivity pattern of forecast error mostly reveals the conditional biases in the forecast of the variable, it is important to understand whether the spatio-temporal connectivity pattern of the error is inherited from that of the observed or predicted climate observable. In other words, for a given climate variable, whether an existence of a high statistical dependency between two regions may cause a correlation between their predictability skills due to common errors arising from the same process which connects them. In terms of climate network, this question transforms to whether the
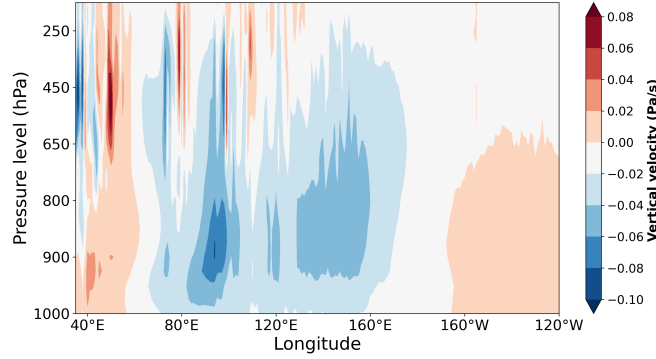
FIGURE 6.6.: Longitude-height plots of ERA5 reanalysis climatology of vertical velocity (in Pa s^{-1}) for June-July-August for ENSO Neutral years between 1980–2020, 35°E–120°W, averaged over 35°N–10°S. Negative values indicate upward motion (ascent) while positive values indicate downward motion (subsidence).

interaction structure of reanalysis (and therefore, forecast) data resembles that of the forecast error data.

We compute CCF to count the fraction of common links between the reanalysis, forecast (Figure A.17) and forecast error networks (Figure 6.2) of the different climate variables (Table 6.1). We see that the connectivity structure of forecast error networks have a varying degree of similarity with reanalysis or forecast networks for different climate variables. For instance, the similarity between the forecast error network and reanalysis network is the highest for Z850 ($CCF(R, FE) \approx 0.8$), less similar for OLR, and least similar for the wind components, U850 and V850. This indicates that for a given climate variable if two regions have a high statistical interdependency between them, their predictability skills may not be correlated. It can be inferred that for Z850, the underlying climate phenomena responsible for the interaction pattern in the reanalysis data is highly likely the cause of the correlations between the forecast errors. This is however, only partially the case for OLR and even less so for the wind components. The underlying climate interactions responsible for the spatial coherence patterns of the reanalysis (and forecast) data of OLR, U850 and V850 are described in Section 5.3 and the general pattern of geopotential height networks in Tsonis and Roebber [144].

It must be mentioned that, although not exact but there is high resemblance between the reanalysis and forecast networks ($CCF(R, F) \approx 0.9$), i.e., the state-of-the-art ERA5 system well-simulates the real climate system. However, it must be clarified that the 5-day forecast error networks do not simply reproduce the excess/missing links between the reanalysis and forecast network (Figure A.17). This can be verified by comparing the pattern of the difference in degree of reanalysis and forecast networks (Figure A.18) with Figure 6.2. Except for Z850 (Figure A.18c), the error networks of other variables bear little resemblance with the degree difference pat-

TABLE 6.1.: Common Component Function (CCF) values between Forecast error (FE), Reanalysis (R) and Forecast (F) networks of U850, V850, Z850 and OLR.

Variable	$CCF(R, FE)$	$CCF(F, FE)$	$CCF(R, F)$
U850	0.580	0.579	0.909
V850	0.535	0.527	0.881
Z850	0.808	0.784	0.937
OLR	0.620	0.620	0.858

tern between reanalysis and forecast. The degree difference for U850 (Figure A.18a) shows missing interactions in the western North Pacific Ocean. Several differences between reanalysis and forecast networks of wind (Figures A.18a-b) also occurs in the Northern Indian Ocean, and monsoon affected regions of south India, southern China and Maritime Continent. Errors related to overestimation/underestimation of links in the ITCZ can also be seen in the degree difference between reanalysis and forecast networks of OLR (Figure A.18d). However, as our purpose here is not the evaluation of climate interactions predicted by models [56, 57, 321, 323], we do not seek a detailed understanding of the differences between the reanalysis and forecast network connectivity structure. But from our aforementioned discussion, it is clear that the topological structure of forecast error correlation network of the climate variable indeed highlights the primary source of structured error for that variable, which might not be revealed from the connectivity structure of the variable itself.

6.5. Impact of El Niño-Southern Oscillation on Error Correlations

We divided the mean absolute model bias corrected forecast error time series (JJA season 1980-2020) of the above considered climate variables into three separate time series containing JJA seasons when ENSO was in the warm (El Niño), cold (La Niña) and neutral periods respectively based on the ONI values [308] for that season (Table 5.2). Figures 6.7-6.9 and A.19 show the respective forecast error networks of the variables for the different ENSO phases.

In case of the error networks of the wind components for different ENSO phases (Figures 6.7 and A.19), the error networks show a high degree at the region corresponding to the WPSH similar to the networks for all years (Figures 6.2a-b). However, the values of degree are different along with the shape of the high degree region. Both the wind components show a higher error correlation in the neutral years (Figures 6.7c and A.19c) and the pattern resembles the most to that of all years among the three phases. The values of degree are comparatively lower for the JJA seasons when El Niño is active, and the pattern extends more southeastwards (Figures 6.7a and A.19a). The degree is lowest during the La Niña periods when the high degree region
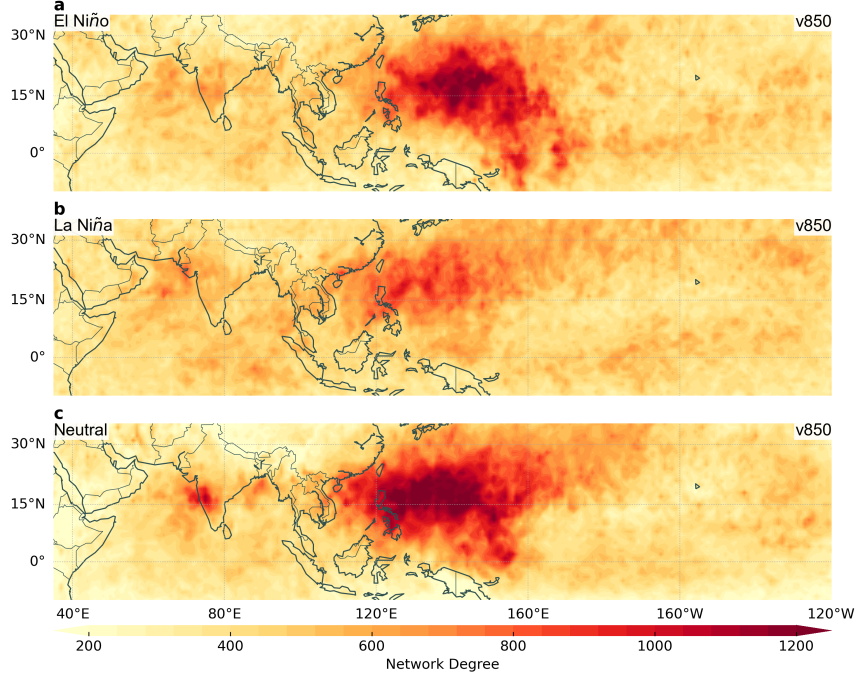
FIGURE 6.7.: Spatial patterns of degree for functional networks of V850 forecast error when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral.

extends northeastwards Figures 6.7b and A.19b). There is also higher degree over southwestern India in the neutral periods Figures 6.7c and A.19c).

The ENSO neutral Z850 error network (Figure 6.8c) resembles that for all years (Figure 6.2c). However, there are some difference for the El Niño and La Niña periods. Compared to the neutral periods, the high degree values in the equatorial central Pacific Ocean is reduced while that in the equatorial Indian Ocean is increased during the El Niño periods (Figure 6.8a), while the reverse happens during the La Nina periods (Figure 6.8b).

The OLR error networks for different ENSO phases (Figure 6.9) show similar spatial patterns of degree as that for all years (Figure 6.2d), except that the values of degree are lower for the El Niño and La Niña periods (Figure 6.9a-b), while the connectivity pattern for the neutral period (Figure 6.9c) bears most resemblance to that of all years. There are also some topological variations in the OLR error connectivity structure during El Niño and La Niña. During El Niño, the high degree region in the Pacific Ocean is longitudinally more spread than during other phases (Figure 6.9a). There is also lesser degree over the Maritime Continent. On the other hand, there is higher degree over northwestern India and the Maritime Continent during the La Niña period (Figure 6.9b) compared to that during El Niño (Figure 6.9a).

ENSO is the primary source of predictability at seasonal time scales. The error correlation features are observed to be stronger during the neutral periods proba-
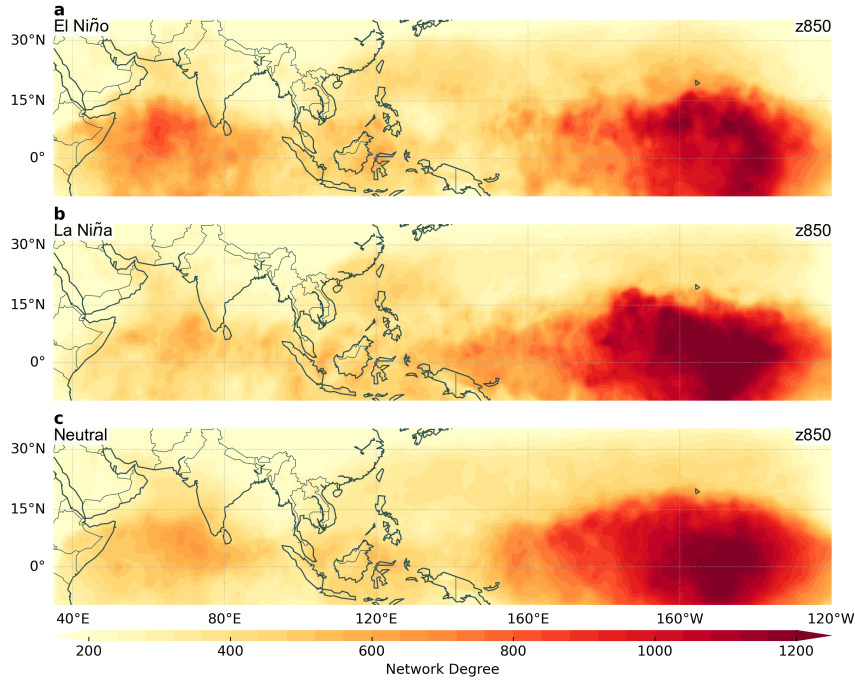
FIGURE 6.8.: Spatial patterns of degree for functional networks of Z850 forecast error when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral.

bly due to the decrease in predictability in ENSO neutral years. The variation in the error network topology under different ENSO forcing sheds light on how the interrelationship between predictability skill of different locations changes due to the ENSO.

6.6. Summary

The results presented in this chapter highlight the applicability of the climate network framework as an effective tool to study the spatial variability of error properties of different climate variables. They can find applications in numerical weather prediction, particularly in data assimilation, and assessment of errors in other fields. Such an approach provides a way to effectively visualize the $N \times N$ error correlation matrix on a N-grid point geographical space along with performing an in-depth analysis of the local and global effects of errors using the multitude of graph theory measures.

We demonstrate the effectiveness of the network-based approach to study error correlations using one of the basic network centrality measures, the degree. Using a few examples, namely, the zonal (U) and meridional (V) components of wind and geopotential (Z) at 850 hPa, and outgoing long wave radiation (OLR), we illustrate that the corresponding forecast error correlation networks of the climate variables show definite spatial patterns of degree which are significantly different from those of
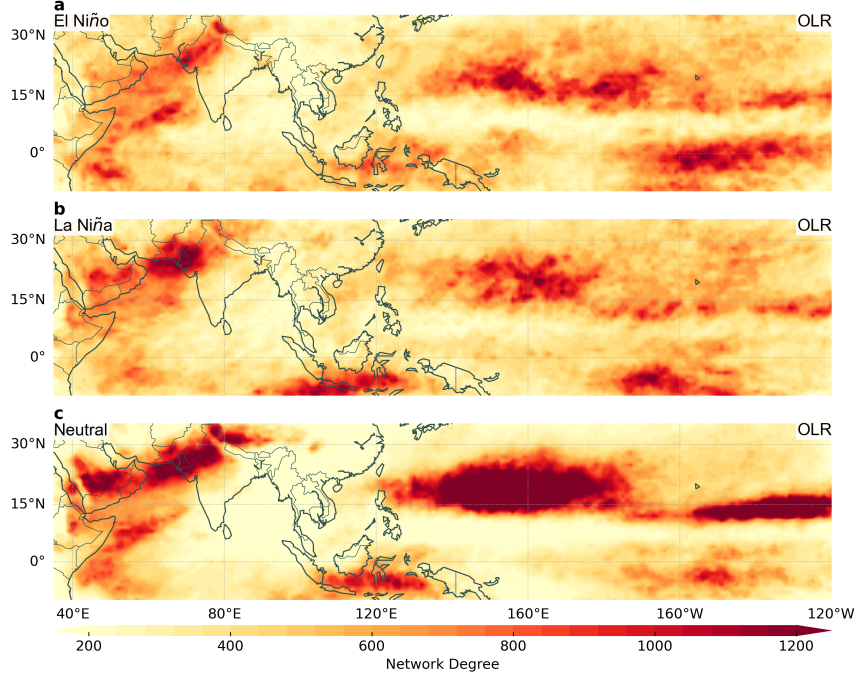
FIGURE 6.9.: Spatial patterns of degree for functional networks of OLR forecast error when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral. (Taken from Gupta et al. [P7])

random networks. Such patterns indicate the existence of underlying systematic or structured random processes that give rise to a relationship between the errors. Our analysis performed on the Asia-Pacific region for the June-July-August sheds light on the dominant conditional biases in the 5-day forecasts of a particular climate variable. In particular, the wind forecast error networks reveal that the southern boundary of the western North Pacific Subtropical High exhibits highly correlated errors, while the Z850 error network depicts errors in the representation of the tropical circulation. Although the wind and geopotential error networks exhibit a large cluster of connected errors, the OLR error network reveals several distant areas of high degree connected to each other. It is found that the high degree regions considerably overlap with the ascending/descending limbs of the Walker Circulation during the neutral El Niño-Southern Oscillation (ENSO) years and may have relations to cloud biases in the inter-tropical convergence zone.

The complex network analysis of the errors thus enables us to identify the most crucial regions which exhibit dominant influence on the predictability skill of whole geographical area. This helps us to conduct a preliminary diagnosis of the structured errors responsible for causing the deviation of forecast values from the reanalyses. Furthermore, we show that such a clear pattern of correlated error structures is not always directly evident from the statistical interaction of the climate variable itself

as the connectivity structure of error networks may be different from that of the reanalysis network to varying extent for different climate variables. This implies that the underlying driving mechanism giving rise to statistical dependency in the climate variable of two regions may not be the common source of error causing the error correlation. Our further investigation on the variation in the connectivity structure of the error network for different phases of the ENSO provides insight into how the ENSO affects the relationships between the predictability skills of various locations.

Though our results are based on only one network measure, several other network measures (not shown here) can be used to derive other quantities of interest from the error correlation matrix. Some of them maybe outlined as follows. Mean geographical link distance (Figures A.20 and A.21), which calculates the mean of the spatial great-circle distances of a node to all its connected neighbours [P2, 48, 50], can be used to obtain an estimate of correlation length scale associated with each grid point [317, 319]. The triangular or hexagonal structure of error correlations can be studied using measures of cliques and clustering in networks [P2, 30, 41]. Diagnosis of the main direction and intensity of the local correlation anisotropies is an interesting way to evaluate the properties of heterogeneous covariance formulations. This can be studied using network measures characterizing the spatial directedness of connections such as edge anisotropy [337] and edge directionality [338]. It must be mentioned here that these network measures may be affected to varying levels by the artificial boundaries introduced to conduct a regional analysis requiring a correction procedure to remove boundary effects [148]. However, it is more appropriate to compute measures such as the mean geographical link distance and edge anisotropy for a global network to get a correct estimation, in which case boundary correction is not necessary. Furthermore, vertical correlations may be studied in a similar fashion as the horizontal correlations shown here. The method can also be extended to analyse multivariate correlations using the concept of network of networks [147] instead of studying univariate correlations as done here.

7. Conclusion

Against the backdrop of Earth system analysis, the larger goal of this dissertation was to contribute to our understanding of climate variability at different time scales. The path taken to achieve this understanding was through a holistic approach by applying complex network theory to the analysis of climate data. In the following, we review the main contributions of this thesis, examine the main challenges faced and the related open questions, and outline the potential directions of future research.

7.1. Contributions of this thesis

The first part of the thesis (Part I) provides an introduction to the necessary concepts on synchronization measures and complex network theory, which are applied in Part II. In particular, Chapter 2 defines measures to detect various kinds of synchronization based on Continuous Wavelet Transform analysis which can be used to identify hidden synchronization from measured time series of chaotic systems within any preferred frequency range [P6]. However, since changes in synchronization can occur on rather large temporal or spatial scales in the macroscopic level, it is essential to know whether there exists a predictable relationship between the variables before identifying the kind of synchronization. So, some common measures to quantify the degree of synchronization or co-variability such as correlation, event synchronization, etc., which are used in this thesis have been explained. Chapter 3 builds on the idea of synchronization as a form of collective behaviour in a complex system which consists of various interacting components and can be represented using a complex network. The concept is further extended to explain the framework under which spatio-temporal data such as different climatological variables can be analysed, and meaningful information about interaction structure and collective dynamics exhibited by the underlying complex system can be extracted.

Complex networks have provided insights into the underlying climate mechanisms along with innovative ways of prediction of climate variability at large scales (El Niño-Southern Oscillation [43–45], Indian Ocean Dipole [42], Indian Summer Monsoon [53], South American Monsoon [52], etc.) by detecting the spatial propagation of coherent patterns. The first aim of this thesis was to investigate whether this approach could be used to capture spatio-temporal coherent structures of variability at smaller time scales, such as at weather or intraseasonal time scales. Such an application could be helpful to construct prediction schemes for extreme weather events, which will have significant implications for the meteorological community in particular, and society in general. The second goal was to integrate the complex network approach

with climate modelling to provide alternative perspectives which can contribute to the improvement of numerical weather forecasts. In accordance with the above, the contributions of this thesis mainly address the following questions:

Can complex network approach be used to study the dynamical changes of spatial patterns of variability of individual extreme weather events?

The framework of time-evolving complex networks can be applied to characterize patterns of weather variability which are particularly important during the occurrence of extreme weather events. In Chapter 4, we applied this framework to understand the rearrangement of the underlying interaction structure of the pressure and vorticity fields during some particular tropical cyclone cases in different cyclone basins. The main challenge here was that tropical cyclones are highly localized extreme events having a short lifespan of approximately 7-10 days and can undergo rapid intensification or dissipation at hourly time scales. Using correlation-based evolving networks constructed over sliding 10 day-time windows, we showed that network based indicators can capture the signatures of the cyclone. In particular, the local clustering coefficient showed a high resemblance to the cyclone tracks. The application of Kendall's τ coefficient, which performs well for short time series, instead of the conventionally used Pearson's or Spearman's correlation coefficient to estimate the link strength, enabled us to construct these networks which could capture the formation of statistically significant localized structures of high connectivity during a cyclone.

In the next step, we made an attempt to study the evolution of the interaction between cyclones when two cyclones are in close proximity. In this case, the time-averaged correlation-based networks are not of much help to study the dynamical changes at hourly time scales. So we proposed an innovative approach to estimate the directed interaction between any two nodes of the vorticity network by using the Biot-Savart law, which gives the velocity induced by one flow element on another based on the instantaneous value of the relative vorticity of the former. Although this law applies strictly to incompressible flows, we found that networks constructed in such a manner can quantitatively describe the transitions during the interaction between binary cyclones. Network measures derived from these directed evolving vorticity networks were able to successfully classify the various stages of binary cyclone interaction prior to complete merger for two such recent examples – the Noru-Kulap interaction in the North Pacific Ocean in July 2017 and the Seroja-Odette interaction in the Southern Indian Ocean in April 2021. This thereby establishes the utility of complex networks in the characterization of weather variability during individual extreme weather events.

Associated publications: Gupta et al. [P2] and De et al. [P3]

Can complex network theory be applied to studying climate variability at intraseasonal timescales? How has our understanding of the Asian Summer Monsoon system improved in this regard?

The spatial variability of the Asian Summer Monsoon (ASM) is broadly dictated by many factors such as the differential heating due to the topography of the region, the surrounding water bodies, the upper level winds, and the tropical oscillations which are intricately linked with each other. The two components of the ASM, the Indian and the East Asian Summer monsoon (ISM and EASM respectively), therefore exhibit a complicated relationship with each other despite affecting different regions, which varies as the season progresses. In Chapter 5, our event synchronization based complex network analysis revealed the spatial coherence of extreme precipitation in the Asian Summer Monsoon region. We identified two dominant synchronization pathways between the ISM and the EASM: (a) a southern mode connecting onset of the ISM over the Arabian Sea and southern India in June to the onset of Meiyu over south-eastern China, i.e., lower and middle reaches of the Yangtze river valley, and (b) a northern mode relating the occurrence and intensity of rainfall over the northern and central parts of India to those in northern China during July. A slight modification of event synchronization allowed us to determine the specific times of high synchronization of extreme precipitation, using which we were able distinctly identify the particular large-scale atmospheric circulation and moisture transport patterns associated with each mode. The southern mode of interaction was found to occur mainly due to the convergence of the south-westerly moisture bearing low level winds coming from the Somalian coast, with the easterly winds associated with the position of the Western Pacific Subtropical high. On the other hand, for the northern mode, while the moisture sources were identified to be primarily local for the Indian and East Asian systems, the in-phase relationship between the India and north China is due to the Silk Road Teleconnection. Furthermore, we showed that the northward progression of the synchronization mode between the ISM and the EASM as the season progresses may not be solely due to the tropical intraseasonal oscillations as thought previously, although there is a considerable influence. Particular phases of the lower frequency modes (40-50 days variability) of the intraseasonal oscillations were found to favour the overall occurrence of extreme rainfall events in the ASM region, while the higher frequency mode (10-20 days time period) likely supports the shifting between the ISM-EASM interaction mode. Therefore, identification of days of high rainfall synchronization using event synchronization was particularly useful in unravelling the climate mechanisms which cause the intraseasonal variability of the spatial synchronization patterns of extreme rainfall in the ASM region.

Associated publications: Gupta et al. [P4]

How do the spatial co-variability patterns of the convection processes that power the Asian Summer Monsoon change at interannual time scales depending on the phases of the El Niño-Southern Oscillation (ENSO)?

The spatial co-variability patterns of the convection anomalies during the ASM season revealed the dominant convective sources which power the monsoon system (Chapter 5). Although outgoing long wave radiation (OLR) is a good proxy for heavy precipitation in the tropics and subtropics, the interaction structure was quite different from that of extreme precipitation, possibly due to the inclusion of more information in the continuous time series of OLR than the event-like extreme precipitation data. The OLR spatial co-variability patterns also showed a distinct interannual variability depending on the phases of the ENSO. The main convective sources associated with the ISM were identified to be the western Indian Ocean and the Indonesian region, while the Philippine and the South China Sea was found to be the primary convective source for the western north Pacific and East Asian monsoon systems, hence confirming earlier studies. ENSO was found to modulate both components of the ASM via their respective convective sources. The primary mechanism behind the ENSO-ISM relationship was found to be associated to the shift in the limbs of the Walker circulation. In contrast, the ENSO impacts the northwestern Pacific-EASM system through changes in the lower tropospheric circulation patterns. Such a comprehensive study of the interannual variation in the spatial co-variability pattern of convection processes has not been conducted earlier using complex network approaches. However, the results showed the effectiveness of the analysis which could also be applied to other monsoon systems.

Associated publications: Gupta et al. [P5]

Can complex network-based approaches provide an understanding of error properties in forecast data ?

Apart from evaluation of teleconnections in climate models, we showed in Chapter 6 that the spatial correlation matrix of forecast errors for a particular climate variable can effectively be represented using the climate network of the spatio-temporal forecast error data, This could then be useful in analysing its spatial co-variability. As presence of correlated errors imply some associated underlying systematic or random process causing the relationship between the errors, the analysis of the topological structures in the error correlation network unveiled information about the dominant source of error. As an example, the dominant error source in the lower tropospheric winds associated with the ASM was the western Pacific Subtropical High, whereas the error network of OLR data revealed systematic biases associated with the Walker Circulation. The phases of the ENSO were also found to affect the relationship between the errors of different locations. The relevance of this approach lies in the fact that the properties of error are not always immediately evident from the interaction pattern of the climate variable itself because the mechanism causing the relationship

between errors might be different from that leading to a statistical dependency of the climate variables. This application of complex networks is first of its kind and shows its potential to be a very promising diagnostic tool to obtain an initial understanding of the origin of forecast errors.

Associated publications: Gupta et al. [P7]

7.2. Challenges and Outlook

As illustrated by the various applications presented in this thesis, complex network approaches to analyse climate data have a great potential to distill meaningful information from it. They can be used complementary to other approaches to gain better insight of the functioning of the climate system at various scales. However, despite the appeal of its holistic ethos, representation of data with functional networks has several potential limitations. The most important point which must be kept in mind while applying the method is that the functional networks do not represent the actual representation of the underlying complex system, but the functional connectivity so obtained is manifested indirectly in the observations. Consequently, the links estimated using similarity measures may be erroneous. Prediction of missing links from an observed network is an active area of research in this direction [339, 340]. Another source of error is the thresholding of the interaction matrix to obtain the unweighted network. Here, although the choice of the threshold is guided based on several arguments and significance tests (Section 3.5.2), it may be affected by biases in estimating the similarity measures (e.g., whether or not binning is used to compute mutual information), and the fact that some link strengths may be overestimated or underestimated as correlation does not imply causation. Therefore, our interpretation of the links are also guided by their physical interpretations which may be readily available or might involve further analysis (Chapter 5). Another possible approach to this problem may be via application of causal inference methods [256] or by performing a link bundle test to indicate significantly important connections [111]. However, the latter approach may be unreliable when the underlying data is locally correlated [341] because this may potentially lead to spurious link bundles even if one spurious long edge is present. Furthermore, the nodes of the network may also contain errors related to uncertainty in measurements and therefore require error assessment. This issue may be partly addressed by construction of error networks (Chapter 6) for different types of errors such as those related to observations, models or background which would be helpful in understanding the systematic biases.

One of the most challenging steps of representing data with networks involves the choice of relevant variables, i.e., how to define the nodes and the links, because the complex systems may have multiple representations [342]. In case of climate networks, this involves two central choices, one of the climate variable relevant to the phenomena under study and the other of how to define the interaction between the grid points. This should be a thoughtful consideration and may involve several trials, as is evident

from Chapters 4 and 5. In Chapter 4, among several possible choices of variables such as sea surface temperature, surface air temperature, pressure, relative vorticity, it was found that pressure and relative vorticity were the most suitable variables to capture signatures of the cyclone. Similarly, in Chapter 5, while precipitation data itself was useful to study the patterns of extreme rainfall synchronization associated with the ASM, it was not a good choice to understand the impact of the ENSO on the ASM. Furthermore, we had to consider different choices of similarity measure to reduce errors due to the property of the data: correlation was not a suitable choice for event-like data, Kendall' τ coefficient performed better than Pearson's correlation coefficient for short time series, and vorticity links calculated on the basis of the Biot-Savart law instead of correlation enabled us to track instantaneous interactions. Moreover, in case of spatial networks, the choice of the relevant spatial and temporal scales is important and is dependent on the scale of the phenomena which is of interest.

Another important problem encountered while representing climate data with networks is the choice of network measures. The complex network theory, and its extension to spatially embedded networks, offers a wide range of network measures to characterize various properties of the network, such as identifying important nodes using centrality measures, or characterizing coherent structures using clustering measures. Nevertheless, it is essential to refrain from applying them in a black-box manner, but instead to consider the context, and carefully weigh their limitations and interpretability [343]. This assumes more importance in interdisciplinary applications of network theory such as in climate data where network analysts often struggle to provide an interpretation of the network measures to the meteorologists, in which case the approach may lose its appeal. Keeping this in mind, in this dissertation, we have used only those network measures which can be effectively interpreted in climatological terms for a particular application. As an example, to indicate the presence of spatial continuity in networks which is observed during cyclones (Chapter 4), we have preferred the use of clustering coefficient to that of betweenness coefficient [41, 146] owing to several reasons. First, the interpretation of betweenness coefficient in functional networks is debatable (some work to resolve this has been done by Molkenthin et al. [344]); second, betweenness centrality is extremely dependent on the sparseness of the network [341]; and third, it significantly increases the computation cost for time-evolving networks.

Finally, construction of a randomized null model to testing a certain network property is also an issue because it requires assumption of a testing hypotheses of what the network is not, which is often guided according to the given situation [343]. An example is when the Erdős–Rényi networks are not a good choice to test network property for spatially embedded networks. In such cases, we typically preserve the link distribution of the original network when constructing random networks (Section 3.5.4, Chapter 4) [130, 148]. In Chapter 6, the choice between suitable testing hypothesis is guided by the presence of hubs or the approximate scale free property of the degree distribution. However, the assumptions of what property the network does not have is often unclear. As an example, the testing hypotheses of the existing

correction procedure for boundary errors is not suitable for network measures such as edge anisotropy [337] which would probably require some assumption about the angular distribution of the edges.

Apart from the aforementioned unresolved issues regarding the technical aspects of network construction from climate (in general, spatio-temporal) data, there are several possible future scopes of the work presented in this thesis. Some of the open questions are outlined as follows:

– The application of vorticity networks based on Biot-Savart law in Chapter 4 may be applied to study other cases of interaction between vortices in climate, such as comparison between different types of Fujiwhara interactions and interaction of a cyclone with large-scale low-level cyclonic vortices. Also, network-based indicators can be combined with physics-inspired machine learning algorithms to improve the prediction of cyclone tracks. Furthermore, as Biot-Savart law is used strictly for incompressible turbulence, its application to merger of cyclones may not be completely accurate. Therefore, one open question that remains is how compressibility may be taken into account while constructing the networks for studying binary cyclone interactions.

– In Chapter 5, the extreme rainfall network approach may be extended to understand the relationship of the Indian Summer Monsoon with the Baiu season over Japan, and the western north Pacific monsoon system. Furthermore, the analysis of the spatial co-variability of the convection anomalies to investigate the impact of the ENSO on the ASM may be extended to other monsoon systems. It is also interesting to look at the co-variability patterns in the previous spring and winter months to distinguish between the effects of the continuing and emerging phases of ENSO. A truly multivariate coupled network analysis of the OLR and wind components can be performed [345]. It is also interesting to investigate the combined impact of the phases of the intraseasonal oscillations (Madden-Julian Oscillation or the boreal summer intraseasonal oscillation) and the ENSO on the spatial coherence patterns of the convective processes related to the ASM. Finally, construction of novel prediction schemes to forecast the monsoon using the functional network topology can be outlined as a highly relevant future research direction.

– In Chapter 6, it may be of interest to analyse error network properties of the whole globe. Also, as already discussed in Section 6.6, network measures other than degree should be explored to provide meaningful interpretation of the error properties. Furthermore, multivariate correlations such those of geopotential height and wind, and variation of vertical correlations (i.e., spatial correlation at different heights) can also be studied under the framework of network of networks [147].

– Lastly, as climate time series exhibit multiscale dynamics, it is quite challenging to design prediction schemes that can truly predict this character. The limited

availability of high resolution data also poses a serious challenge. Advanced machine learning-based methods that can learn the dynamics from limited data must be developed to predict the multiscale dynamics.

The realms of synchronization, complex networks and climate dynamics are vast in themselves, each having a plethora of unsolved questions, which make them highly active areas of modern day research. The achievements of this dissertation highlight the fact that the domain of intersection of these subjects is very promising, and more intensive multidisciplinary research in this direction may enable us to unravel some of the deeper mysteries of the Earth System Dynamics.

Appendices

A. Supporting Information

A.1. Tropical Cyclones: Detection and Binary Cyclone Interaction
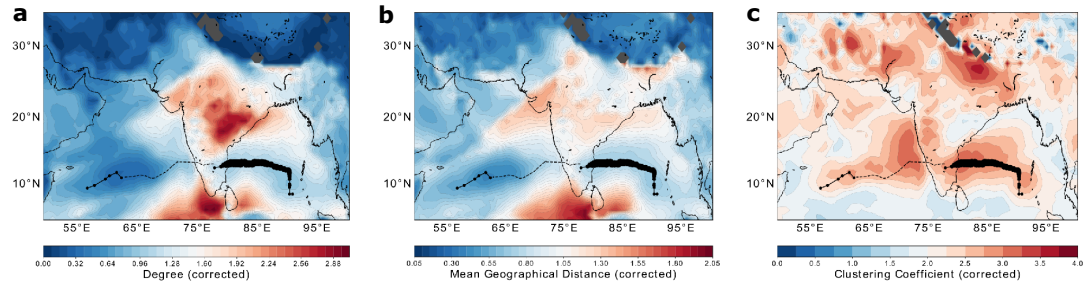
FIGURE A.1.: Mean Sea Level Pressure network analysis including land points for the land-crossing Very Severe Cyclonic Storm Vardah in Dec 2016. The result maintains the inferences drawn from that performed after masking the land as in the main text (see Figure 4.4 in Section 4.2). Although the (a) degree and (b) mean geographic distance performs well, the quality of results for the (c) local clustering coefficient seems to be affected by orography of the region (high clustering near the Himalayas), justifying the reason for removing land points from the analysis, apart from reasons related to the reduction of computational cost.

A.2. Asian Summer Monsoon: Spatial synchronization patterns of Extreme Precipitation and Convection

A.2.1. Interconnection between Indian and East Asian Summer Monsoon
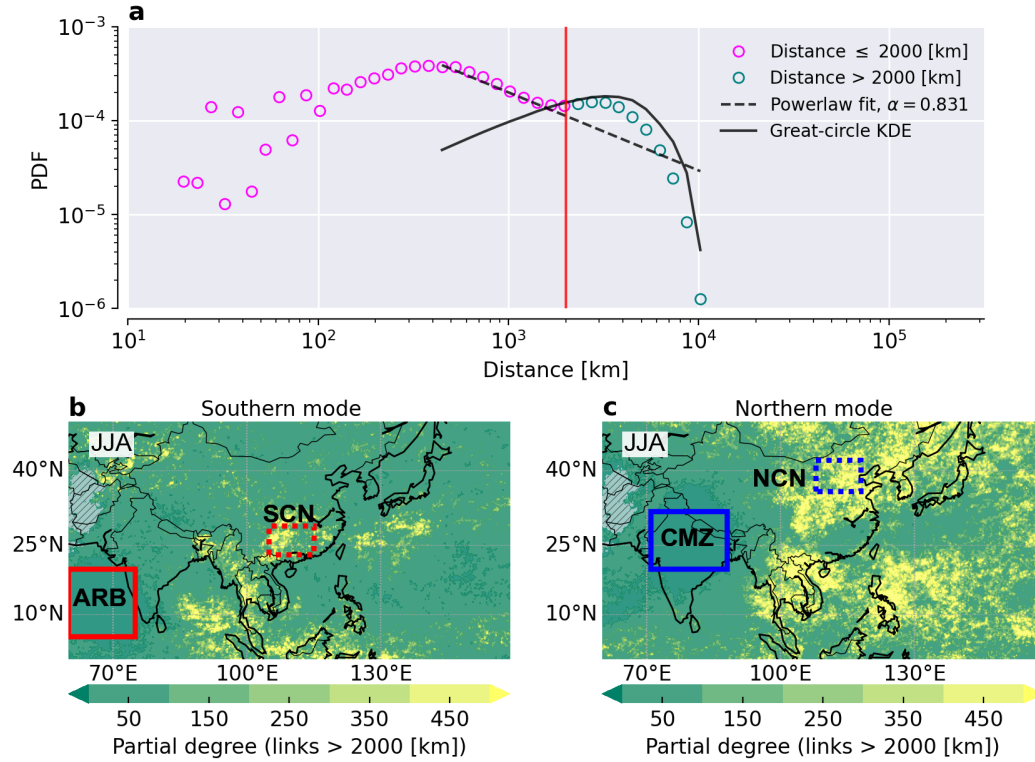
FIGURE A.2.: (a) The link distance distribution for the network constructed for the entire JJA season (Figure 5.1d in the main text). The distance of red vertical line at 2000 km marks the regime shift from regional weather systems to large-scale teleconnections. Spatial patterns of partial degree computed on the basis of links longer than 2000 km, for (b) ARB and (c) CMZ, are similar to Figures 5.1e and 5.1f, respectively, in the main text. It indicates that the two pairs of ISM-EASM connection modes, i.e., the Southern mode (ARB-SCN) and the Northern mode (CMZ-NCN), are part of the global-scale teleconnections. (Taken from Supporting Information of Gupta et al. [P4])
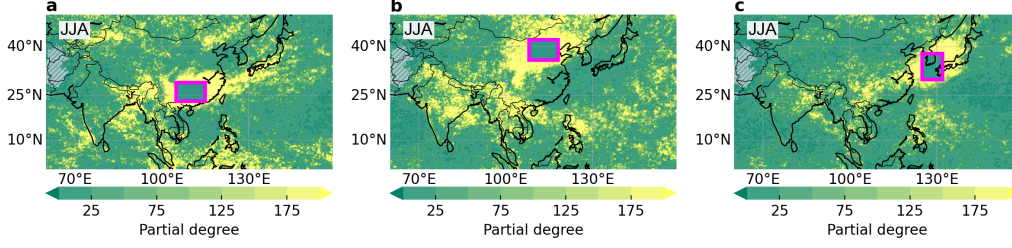
FIGURE A.3.: Partial degree for (a) SCN, (b) NCN, and (c) southern Japan obtained from network constructed for the JJA season (Figure 5.1d in the main text). (a) and (b) show the partial degree for the counterpart regions over China, of the Southern and Northern modes. SCN (NCN) has relatively more connections with ARB (CMZ). Rainfall at southern Japan, is seen to be synchronized with SCN from (c). (Taken from Supporting Information of Gupta et al. [P4])

Explanation based on ISM and EASM indices

We can define indices describing the northward progression of ISM (from ARB to CMZ) and EASM (from SCN to NCN) based on the extreme rainfall event (ERE) series of ARB, CMZ, SCN and NCN. We show using these indices this northward propagation of EREs occur simultaneously for ISM and EASM. However, one should note that due to the dependence of EREs on the percentile chosen, these indices are not as general as the ones defined based on the zonal component of wind at 850 hPa (U850) component [278]. Also, the traditional indices defined using U850 can have daily values and in real-time, which is not the case with the potential one defined using EREs as the event series can have zeros in their daily values. We describe below the method for calculating the index using EREs and discuss their intra-seasonal variations.

The calculation process for the indices is as follows:

1. *Averaging EREs over ARB, CMZ, SCN and NCN regions.* For each box, we compute the daily average number of EREs during JJA. The rest of seasons are all set to zero, leading to four event series each of length 8034 for these four regions.

2. *Compute the daily EREs climatology over ARB, CMZ, SCN and NCN regions.* We then compute the mean of EREs for the same day over all the years. This process converts the four event series from a length of 8034 to 366.

3. *The indices for ISM and EASM are then defined for the JJA season as:*

$$\text{Index(ISM)} = \frac{\text{CMZ}}{\text{ARB}} - 1$$
$$\text{Index(EASM)} = \frac{\text{NCN}}{\text{SCN}} - 1$$

(A.1)

A. Supporting Information

We compute indices for both ISM and EASM, based on different percentiles of EREs, shown below in Figure A.4a-b to verify the robustness. It is clear from the negative values of the indices that in June that both ARB and SCN have relatively more EREs, which potentially indicate a possibility of synchronization between them. Starting from July, both indices become positive around the same time, indicating that CMZ and NCN have more EREs, and are potentially synchronized. The switch for the positive to the negative values of the index denotes the possible northward shift of the synchronization between ISM and EASM, i.e. from ARB-SCN to CMZ-NCN synchronization mode.
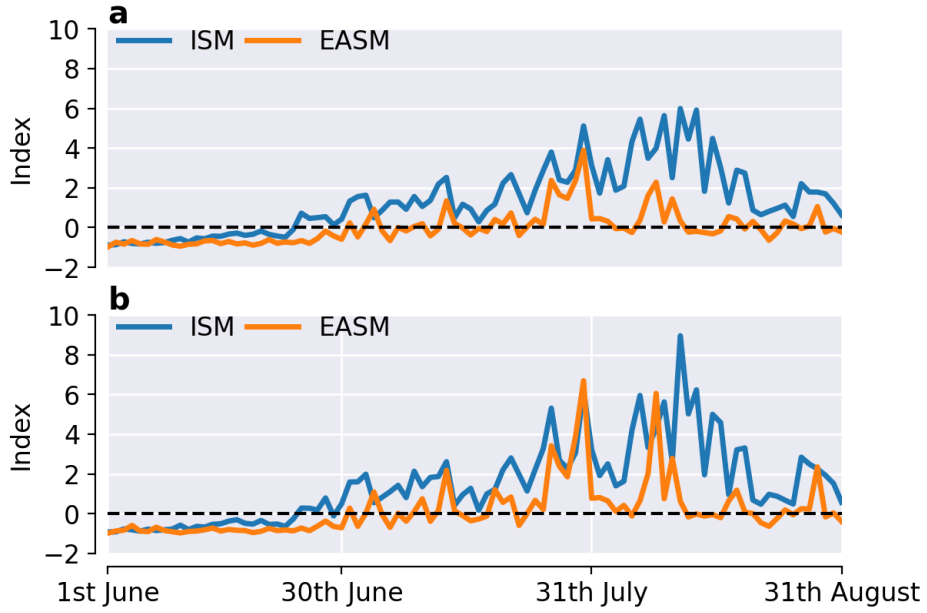
FIGURE A.4.: Indices for ISM and EASM based on EREs defined using (a) 90th percentile and (b) 95th percentile of wet days. Both indices switch from negative to positive values almost simultaneously indicating a possibility of synchronization. (Taken from Supporting Information of Gupta et al. [P4])
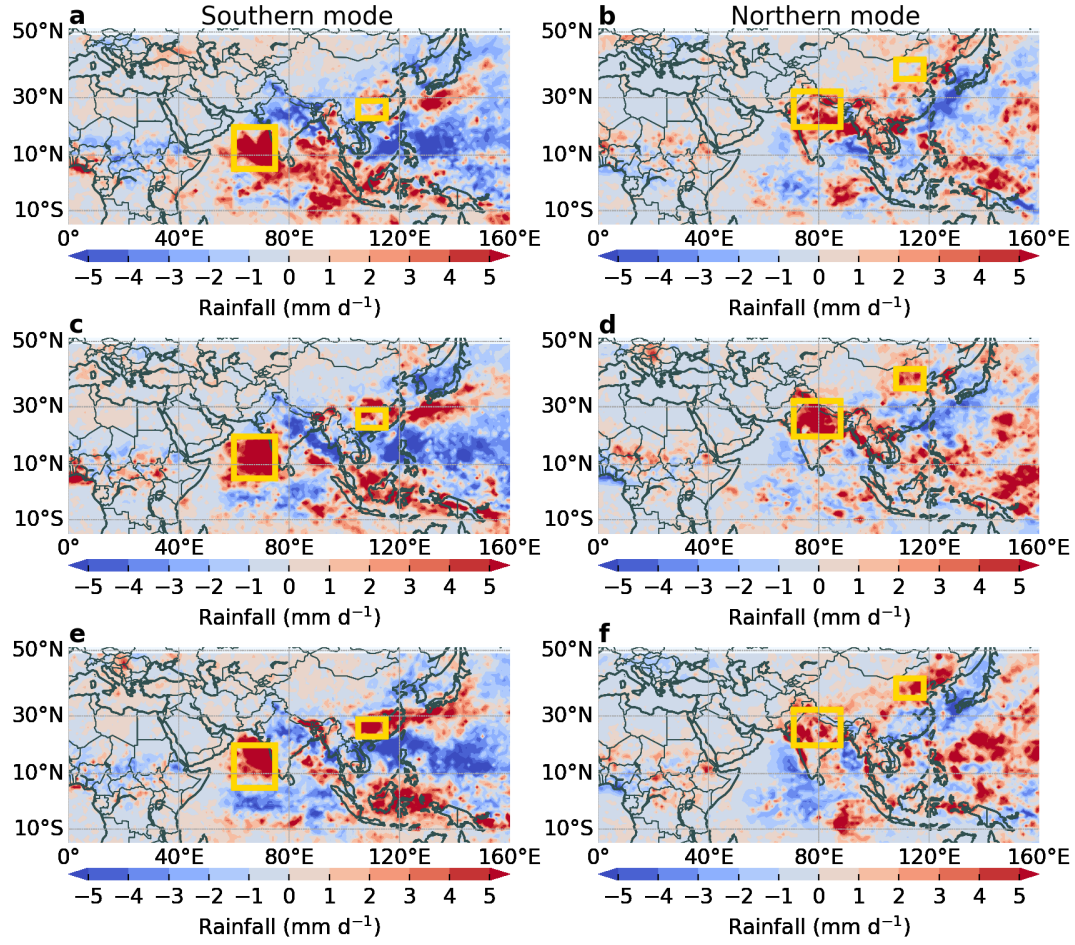
FIGURE A.5.: The composite anomalies of TRMM rainfall, with respect to JJA climatology, for Southern ((a), (c), and (e)) and Northern ((b), (d), and (f)) modes of ISM-EASM connection. Composites are computed for Day −2 ((a) and (b)), Day 0 ((c) and (d)), and Day +2 ((e) and (f)), where Day 0 corresponds to the day of high ERE synchronization for each ISM-EASM connection mode. Day −2 and +2 represent the days before and after the day of high synchronization for the Northern and Southern modes of the connection between ISM and EASM, respectively. For the Southern (Northern) connection mode, rainfall occurs first at ARB (CMZ) and then at SCN (NCN) after some lags. (Taken from Supporting Information of Gupta et al. [P4])
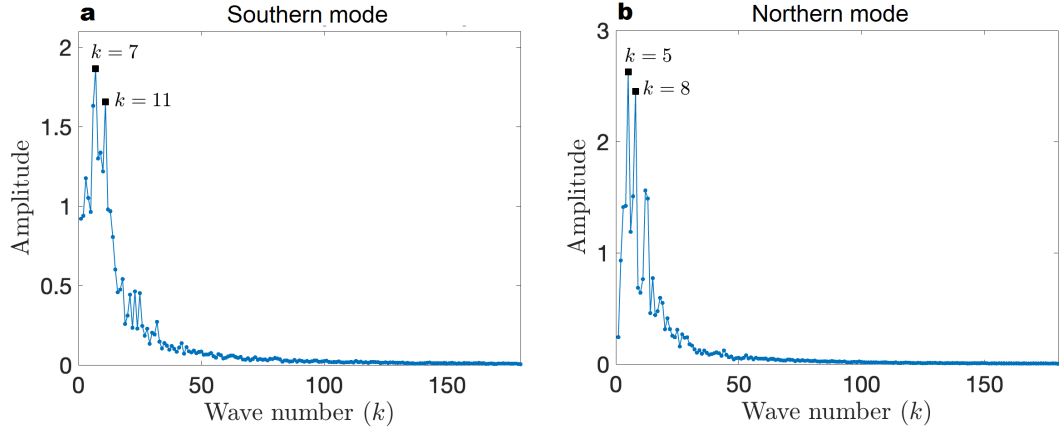
FIGURE A.6.: Spatial Fourier spectra showing the dominant wavenumbers associated with wave pattern in the composite anomalies of meridional wind component at 250 hPa shown in Figure 5.4 of the main text for (a) Southern mode determined from the latitude belt 25°N-35°N, and (b) Northern mode for the latitude belt 40°N-50°N. (Taken from Supporting Information of Gupta et al. [P4])
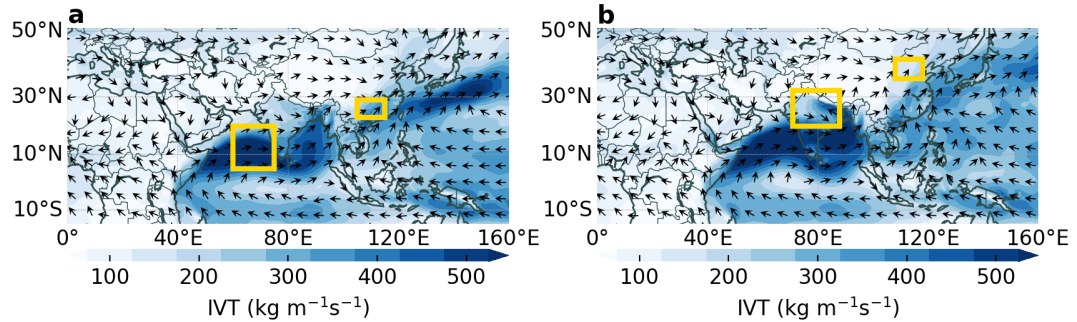
FIGURE A.7.: Composites of vertically integrated water vapour flux (IVT) for the (a) Southern and (b) Northern modes, based on the days of high ERE synchronization (Day 0). From (a), we see a high amount of water vapour transport along the Somalian LLJ to the ARB and then towards SCN via the Bay of Bengal. This gives a clear indication of the formation of the great moisture corridor. In case of (b), the general direction of flow is still south-westerly, from ARB towards China, however the amount of moisture transported from BOB over to China seems to be greatly reduced. (Taken from Supporting Information of Gupta et al. [P4])
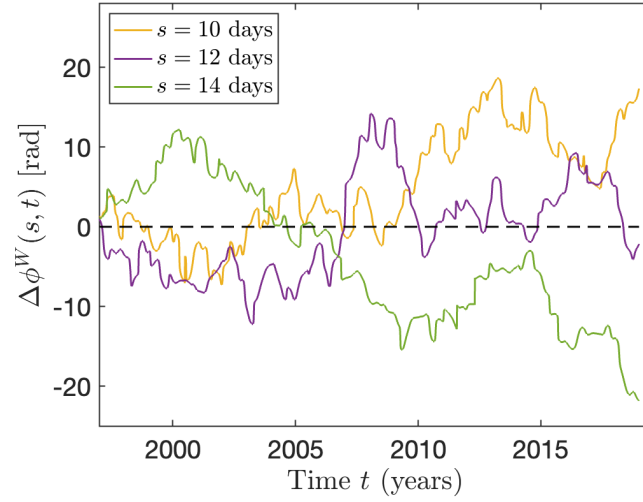
FIGURE A.8.: Unwrapped wavelet phase difference, $\phi^W(s,t)$, between the CMZ and the NCN precipitation time series versus time t during the JJA period, for scales $s = 10, 12$ and 14 days. $\phi^W(s,t)$ oscillates about zero and does not grow with time, implying that the CMZ and NCN time series are phase synchronized at these scales.
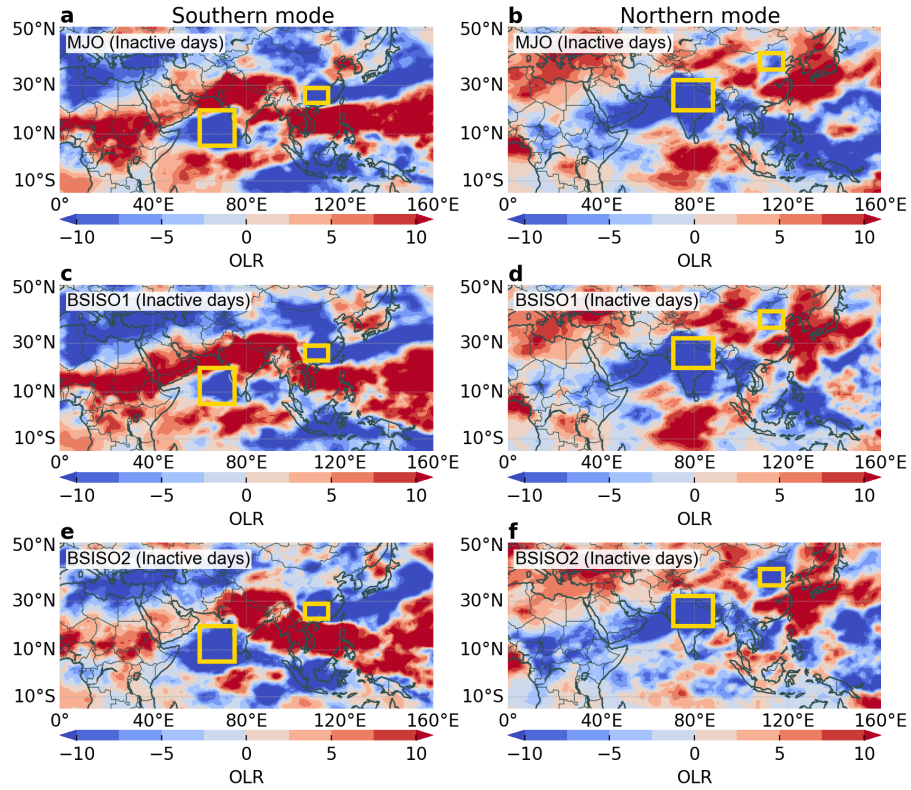
FIGURE A.9.: The composite anomalies of outgoing long-wave radiation, with respect to JJA climatology, based on the **inactive** days of high ERE synchronization on Day 0, for (a) Southern and (b) Northern modes of ISM-EASM connection. (Taken from Supporting Information of Gupta et al. [P4])

131
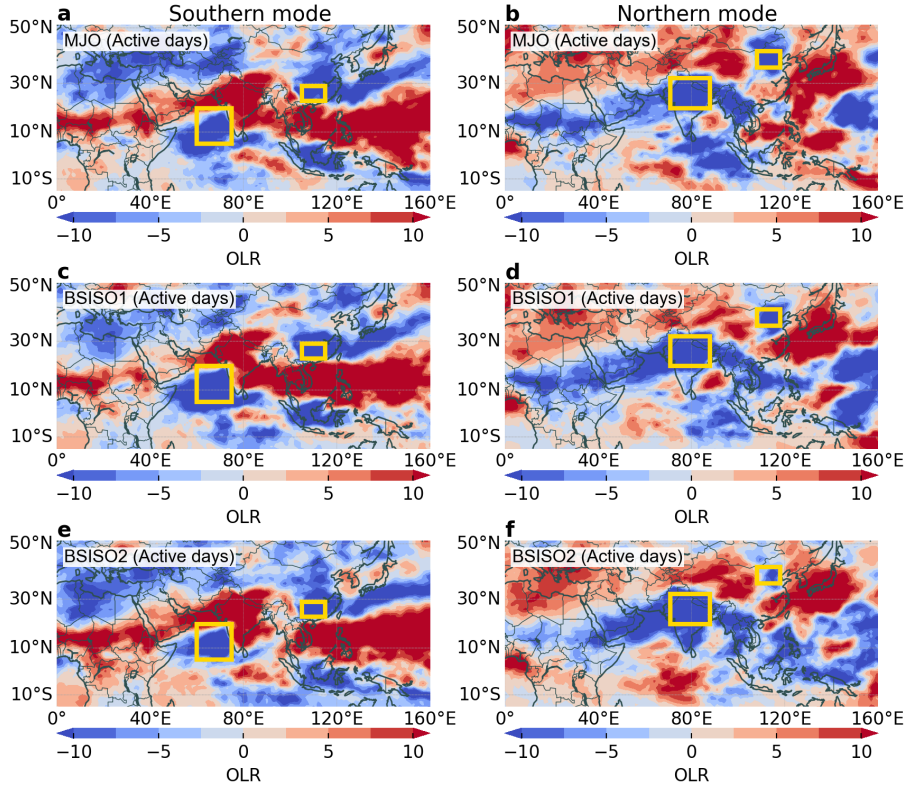
FIGURE A.10.: The composite anomalies of outgoing long-wave radiation, with respect to JJA climatology, based on the **active** days of high ERE synchronization on Day 0, for (a) Southern and (b) Northern modes of ISM-EASM connection. The negative OLR anomaly (positive convection) over the respective interacting regions of ISM and EASM for both modes is seen from the composites for both **inactive** (Figure A.9) and **active** (this figure) MJO/BSISO days of high ERE synchronization implies that MJO or BSISO alone are not responsible the interconnection between the two monsoon subsystems. However, it is also clear from the figure that the intraseasonal oscillations are an important ingredient of this linkage between ISM and EASM. (Taken from Supporting Information of Gupta et al. [P4])

A.2.2. Impact of El Niño-Southern Oscillation on spatial connectivity pattern of Convection during ASM
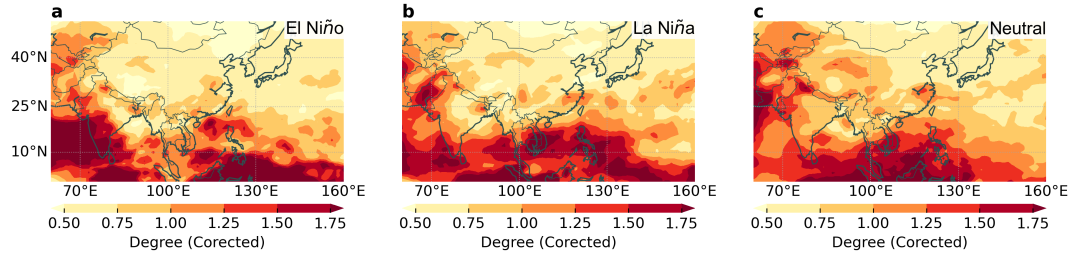
FIGURE A.11.: Spatial patterns of degree of OLR correlation network for the monsoon season for different ENSO phases after correcting the boundary effects using the procedure in Section 3.5.4.
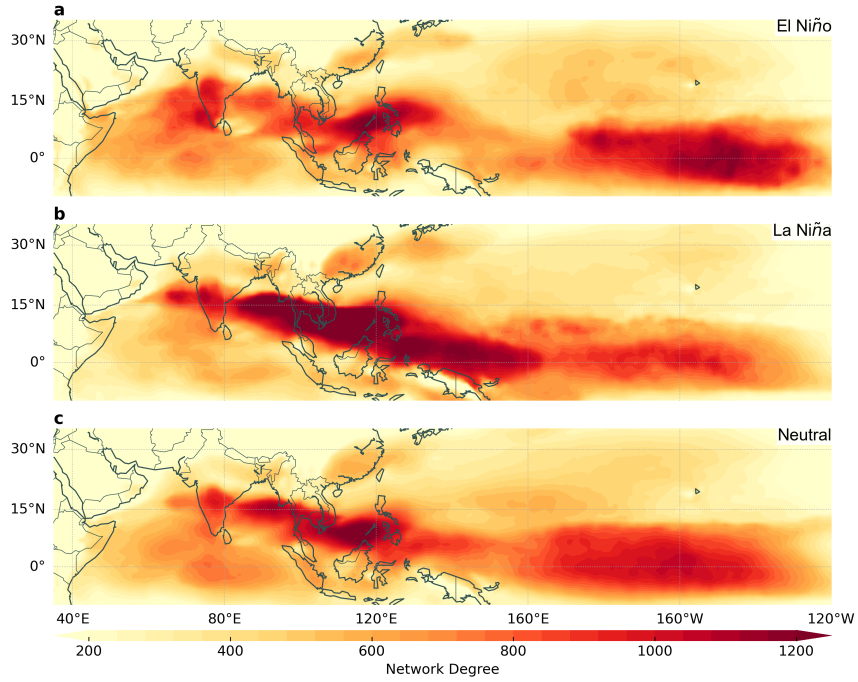
FIGURE A.12.: Spatial patterns of degree for functional networks of U850 constructed for the extended region of ASM which includes the adjacent oceans for JJA season when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral.
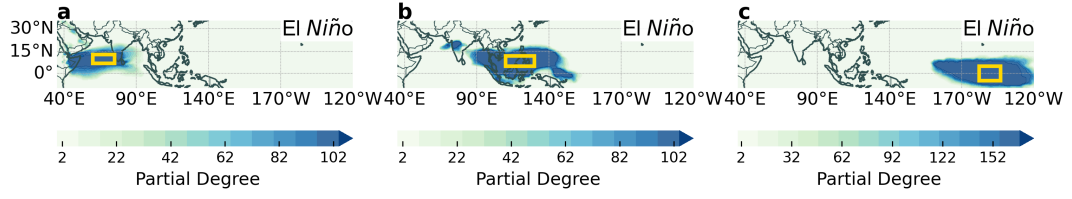
FIGURE A.13.: Partial degree plots corresponding to the high degree regions of the U850 network during El Niño period shown in Figure A.12a: (a) Arabian Sea, (b) Philippines and (c) equatorial eastern Pacific Ocean.
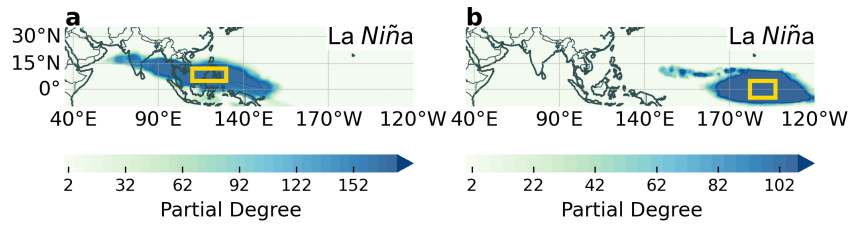
FIGURE A.14.: Partial degree plots corresponding to the high degree regions of the U850 network during La Niña period shown in Figure A.12b: (a) Philippines and (b) equatorial eastern Pacific Ocean.
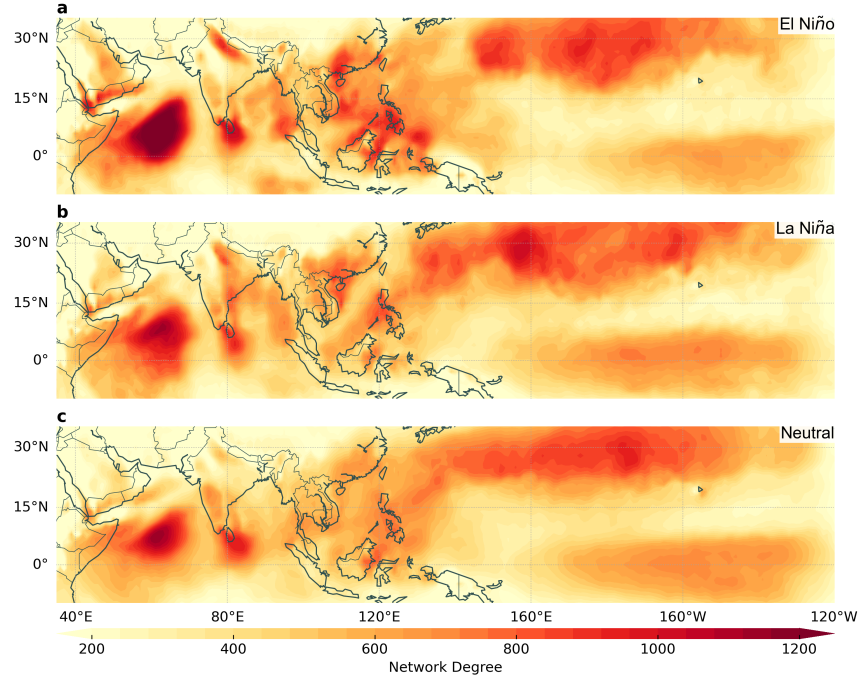
FIGURE A.15.: Spatial patterns of degree for functional networks of V850 constructed for the extended region of ASM which includes the adjacent oceans for JJA season when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral.

A.3. Spatial variability of error correlations

TABLE A.1.: Values of Degree Assortativity coefficient for Forecast error networks of U850, V850, Z850 and OLR. (Taken from Gupta et al. [P7])

Variable	Degree Assortativity coefficient (r)
U850	0.509
V850	0.556
Z850	0.858
OLR	0.434
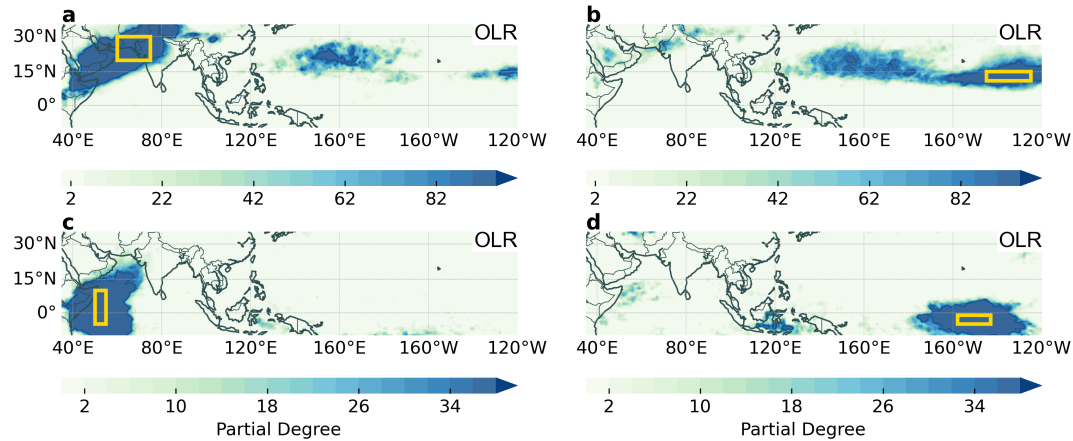
FIGURE A.16.: Partial degree associated with the regions in yellow boxes showing the areas connected to those regions in the OLR forecast error network (Figure 6.1d). (Taken from Gupta et al. [P7])
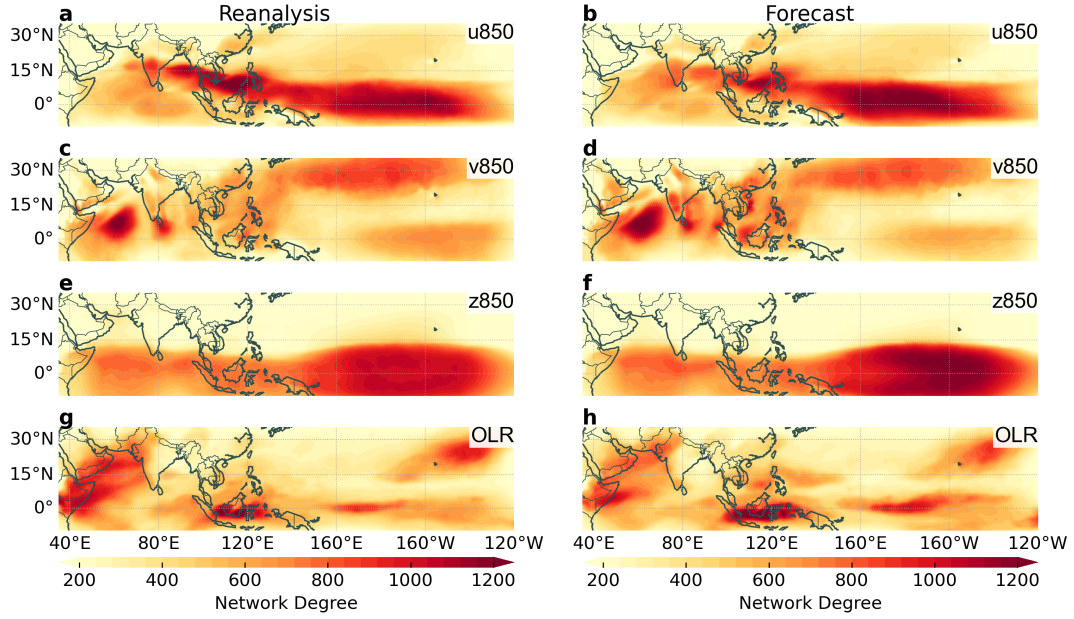
FIGURE A.17.: Comparison of spatial patterns of degree for networks of (a-b) U850, (c-d) V850, (e-f) Z850 and (g-h) OLR using reanalysis (a,c,e,g) and forecast (b,d,f,h) data. (Taken from Gupta et al. [P7])
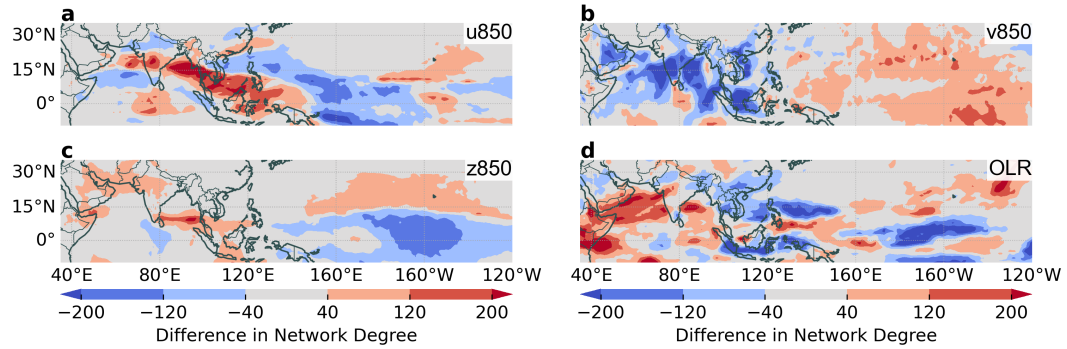
FIGURE A.18.: Difference in node degree between the reanalysis and the forecast networks shown in Figure A.17 for (a) U850, (b) V850, (c) Z850 and (d) OLR. (Taken from Gupta et al. [P7])
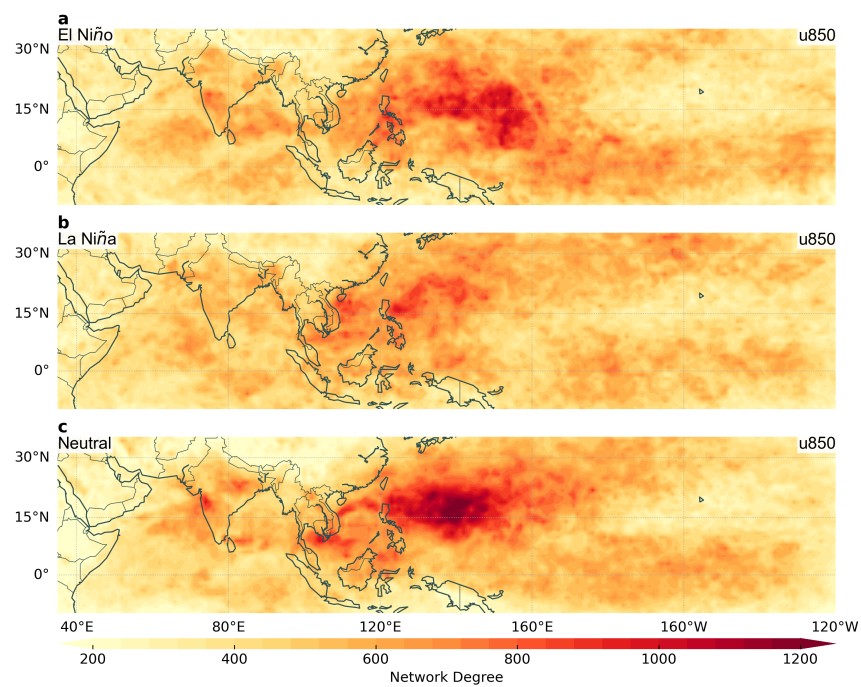
FIGURE A.19.: Spatial patterns of degree for functional networks of U850 forecast error when ENSO is (a) positive (El Niño), (b) negative (La Niña) and (c) neutral.

FIGURE A.20.: Probability density function (PDF) for link distances of error networks (refer to Figure 6.1a-d in Chapter 6) of (a) U850, (b) V850, (c) Z850 and (d) OLR. Distance distribution of U850 and V850 error networks fit well to power law. Z850 error network has large deviation from power law distribution. The distance distribution of Z850 network might be related to high assortativity of the Z850 network. The OLR error network shows a power law distribution for smaller distances but tends towards great-circle kernel density estimate (KDE) for larger distances implying long range connectivity along great-circle distances. (Taken from Gupta et al. [P7])
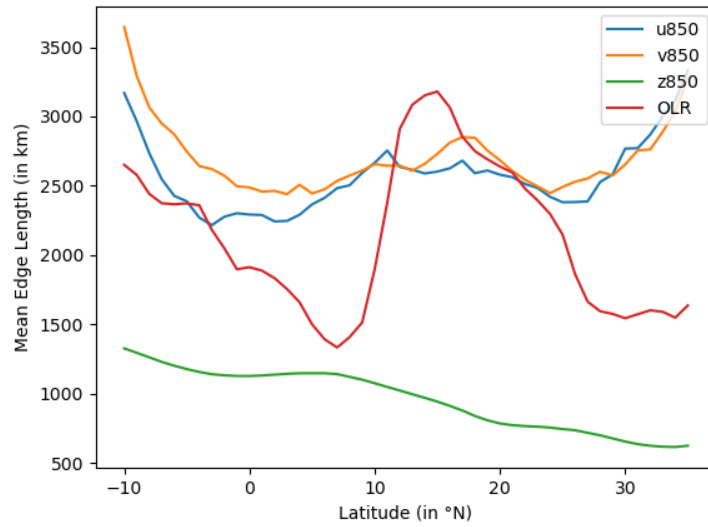
FIGURE A.21.: Mean geographic link distance of grid points averaged over all longitudes in the considered region computed from error network of U850 (blue), V850 (orange), Z850 (green) and OLR (red). It can be used to get an estimate of the correlation length scale. As the length here is calculated for a regional network without correcting for boundary effects, the end latitudes show high boundary errors. Despite that one can see a clear maxima close to the tropics for the U850, V850 and OLR networks. The measure needs to be computed from a global network to obtain a better picture of the latitude dependence of length scale unlike that computed from a regional network as done here. (Taken from Gupta et al. [P7])

Bibliography

[1]   *CRediT – Contributor Roles Taxonomy.* `https://credit.niso.org/`. (Accessed on 11/05/2022).

[2]   H. J. Schellnhuber. "'Earth system' analysis and the second Copernican revolution". In: *Nature* 402.6761 (1999), pp. C19–C23.

[3]   V. Bjerknes. "Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik". In: *Meteorolgische Zeitschrift* 21.1 (1904), pp. 1–7.

[4]   L. F. Richardson. *Weather prediction by numerical process.* Cambridge, England: Cambridge University Press, 1922.

[5]   J. G. Charney, R. Fjörtoft, and J. Von Neumann. "Numerical Integration of the Barotropic Vorticity Equation". In: *Tellus* 2.4 (1950), pp. 237–254.

[6]   Edward N. Lorenz. "Deterministic Nonperiodic Flow". In: *J. Atm. Sci.* 20.2 (1963), pp. 130–141.

[7]   S. Arrhenius. "On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground". In: *Philosophical Magazine and Journal of Science.* 5th ser. 41 (1896), pp. 237–276.

[8]   S. Manabe and R. T. Wetherald. "Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity". In: *J. Atm. Sci.* 24.3 (1967), pp. 241–259.

[9]   S. Manabe and R. F. Strickler. "Thermal Equilibrium of the Atmosphere with a Convective Adjustment". In: *J. Atm. Sci.* 21.4 (1964), pp. 361–385.

[10]   *UCAR E&O - Randy Russell - Climate Models - Components and Evolution.* `https://eo.ucar.edu/staff/rrussell/climate/modeling/climate_model_components_evolution.html`. (Last Accessed on 02/05/2022). 2022.

[11]   *European Centre for Medium-Range Weather Forecast (ECMWF).* `https://www.ecmwf.int/`. 2022.

[12]   *What is a GCM?* `https://www.ipcc-data.org/guidelines/pages/gcm_guide.html`. (Last Accessed on 02/21/2022).

[13]   *Climate Models, NOAA Climate.gov.* `https://www.climate.gov/maps-data/climate-data-primer/predicting-climate/climate-models`. (Last Accessed on 02/21/2022).

[14] P. Maher, E. P. Gerber, B. Medeirosa, T. M. Merlis, S. Sherwood, A. Sheshadri, A. H. Sobel, G. K. Vallis, A. Voigt, and P. Zurita-Gotor. "Model hierarchies for understanding atmospheric circulation". In: *Rev. Geophys.* 57.2 (2019), pp. 250–280.

[15] K. Hasselmann. "On the Problem of Multiple Time Scales in Climate Modeling". In: *Man's Impact on Climate.* Ed. by Wilfrid Bach, Jürgen Pankrath, and William Kellogg. Vol. 10. Developments in Atmospheric Science. Elsevier, 1979, pp. 43–55.

[16] K. Hasselmann. "Stochastic climate models Part I. Theory". In: *Tellus* 28.6 (1976), pp. 473–485.

[17] C. Frankignoul and K. Hasselmann. "Stochastic climate models Part II. Application to sea-surface temperature anomalies and thermocline variability". In: *Tellus* 29.4 (1977), pp. 289–305.

[18] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. "Stochastic resonance in climatic change". In: *Tellus* 34.1 (1982), pp. 10–15.

[19] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. "A Theory of Stochastic Resonance in Climatic Change". In: *SIAM J. Appl. Math.* 43 (1983), pp. 565–578.

[20] M. Milankovitch. "Mathematische Klimalehre und Astronomische Theorie der Klimaschwankungen". In: *Handbuch der Klimatologie.* Ed. by W. Köppen and R. Geiger. Vol. 1(A). [Nachdr. d. Ausg.] Berlin, 1930.

[21] V. Masson-Delmotte et al. "Climate Change 2021: The Physical Science Basis". In: *Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Sixth Assessment Report (AR5).* Cambridge University Press, 2021.

[22] K. Hasselmann. "On the signal-to-noise problem in atmospheric response studies". In: *Meteorology of Tropical Oceans.* Ed. by D. B. Shaw. Royal Meteorological Society, 1979, pp. 251–259.

[23] K. Hasselmann. "Optimal Fingerprints for the Detection of Time-dependent Climate Change". In: *J. Clim.* 6.10 (1993), pp. 1957–1971.

[24] K. Hasselmann. "Multi-pattern fingerprint method for detection and attribution of climate change". In: *Clim. Dyn.* 13 (1997), pp. 601–611.

[25] S. Seneviratne, N. Nicholls, D. Easterling, C. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. Mcinnes, M. Rahimi, M. Reichstein, A. Sorteberg, C. Vera, and X. Zhang. "Changes in climate extremes and their impacts on the natural physical environment". In: 2012.

[26] M. Ghil and V. Lucarini. "The physics of climate variability and climate change". In: *Rev. Mod. Phys.* 92.3 (2020), p. 035002.

[27] N. Wunderling, J. F. Donges, J. Kurths, and R. Winkelmann. "Interacting tipping elements increase risk of climate domino effects under global warming". In: *Earth System Dynamics* 12.2 (2021), pp. 601–619.

[28] M. Singh, R. Krishnan, B. Goswami, A. D. Choudhury, P. Swapna, R. Vellore, A. G. Prajeesh, N. Sandeep, C. Venkataraman, R. V. Donner, N. Marwan, and J. Kurths. "Fingerprint of volcanic forcing on the ENSO–Indian monsoon coupling". In: *Science Advances* 6.38 (2020), eaba8164.

[29] J. Fan, J. Meng, J. Ludescher, X. Chen, Y. Ashkenazy, J. Kurths, S. Havlin, and H. J. Schellnhuber. "Statistical physics approaches to the complex Earth system". In: *Physics Reports* 896 (2021), pp. 1–84.

[30] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "Complex networks in climate dynamics". In: *The European Physical Journal Special Topics* 174 (2009), pp. 157–179.

[31] J. Ludescher, M. Martin, N. Boers, A. Bunde, C. Ciemer, J. Fan, S. Havlin, M. Kretschmer, J. Kurths, J. Runge, V. Stolbova, E. Surovyatkina, and H. J. Schellnhuber. "Network-based forecasting of climate phenomena". In: *Proceedings of the National Academy of Sciences* 118.47 (2021).

[32] H. A. Dijkstra, E. Hernandez-Garcia, C. Masoller, and M. Barreiro. *Networks in Climate*. Cambridge University Press, 2019.

[33] P. Voosen. "Europe is building a 'digital twin' of Earth to revolutionize climate forecasts". In: *Science* (2020).

[34] P. Bauer, B. Stevens, and W. Hazeleger. "A digital twin of Earth for the green transition". In: *Nature Climate Change* 11.2 (2021), pp. 80–83.

[35] P. Coppola, J. Allanson, L. Naci, R. Adapa, P. Finoia, G. B. Williams, J. D. Pickard, A. M. Owen, D. K. Menon, and E. A. Stamatakis. "The complexity of the stream of consciousness". In: *Communications Biology* 5.1 (2022).

[36] S. Strogatz. "Exploring Complex Networks". In: *Nature* 410 (2001), pp. 268–276.

[37] A. Arenas, A. Diaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. "Synchronization in complex networks". In: *Physics Reports* 469.3 (2008), pp. 93–153.

[38] S. Boccaletti, V. Latora, Y. Moreno, N. Chavez, and D.-U. Hwang. "Complex networks: Structure and dynamics". In: *Physics Reports* 424.4 (2006), pp. 175–308.

[39] R. Albert and A. L. Barabási. "Statistical mechanics of complex networks". In: *Rev. Mod. Phys.* 74.1 (2002), pp. 47–97.

[40] N. Boers, J. Kurths, and N. Marwan. "Complex systems approaches for Earth system data analysis". In: *Journal of Physics: Complexity* 2.1 (2021), p. 011001.

[41] M.E.J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.

[42]   Z. Lu, W. Dong, B. Lu, N. Yuan, Z. Ma, M. I. Bogachev, and J. Kurths. "Early warning of the Indian Ocean Dipole using climate network analysis". In: *Proceedings of the National Academy of Sciences* 119.11 (2022), e2109089119.

[43]   J. Ludescher, A. Gozolchiani, M. I. Bogachev, A. Bunde, S. Havlin, and H. J. Schellnhuber. "Improved El Niño forecasting by cooperativity detection". In: *Proceedings of the National Academy of Sciences* 110.29 (2013), pp. 11742–11745.

[44]   J. Ludescher, A. Gozolchiani, M. I. Bogachev, A. Bunde, S. Havlin, and H. J. Schellnhuber. "Very early warning of next El Niño". In: *Proceedings of the National Academy of Sciences* 111.6 (2014), pp. 2064–2066.

[45]   J. Meng, J. Fan, Y. Ashkenazy, and S. Havlin. "Percolation framework to describe El Niño conditions". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 035807.

[46]   O. Guez, A. Gozolchiani, Y. Berezin, S. Brenner, and S. Havlin. "Climate network structure evolves with North Atlantic Oscillation phases". In: *Europhysics Letters* 98.3 (2012), p. 38006.

[47]   O. Guez, A. Gozolchiani, Y. Berezin, Y. Wang, and S. Havlin. "Global climate network evolves with North Atlantic Oscillation phases: Coupling to Southern Pacific Ocean". In: *Europhysics Letters* 103.6 (2013), p. 68006.

[48]   N. Malik, B. Bookhagen, N. Marwan, and J. Kurths. "Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks". In: *Climate dynamics* 39.3 (2012), pp. 971–987.

[49]   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". In: *Nonlinear Processes in Geophysics* 21.4 (2014), pp. 901–917.

[50]   N. Boers, B. Bookhagen, N. Marwan, J. Kurths, and J. Marengo. "Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System". In: *Geophysical Research Letters* 40.16 (2013), pp. 4386–4392.

[51]   N. Boers, A. Rheinwalt, B. Bookhagen, H.M.J. Barbosa, N. Marwan, J. Marengo, and J. Kurths. "The South American rainfall dipole: A complex network analysis of extreme events". In: *Geophysical Research Letters* 41.20 (2014), pp. 7397–7405.

[52]   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". In: *Nature Communications* 5.1 (2014), p. 5199.

[53] V. Stolbova, E. Surovyatkina, B. Bookhagen, and J. Kurths. "Tipping elements of the Indian monsoon: Prediction of onset and withdrawal". In: *Geophysical Research Letters* 43.8 (2016), pp. 3982–3990.

[54] D. Traxl, N. Boers, A. Rheinwalt, B. Goswami, and J. Kurths. "The size distribution of spatiotemporal extreme rainfall clusters around the globe". In: *Geophysical Research Letters* 43.18 (2016), pp. 9939–9947.

[55] B. Goswami, P. Schultz, B. Heinze, N. Marwan, B. Bodirsky, H. Lotze-Campen, and J. Kurths. "Inferring interdependencies from short time series". In: *Proceedings of the Conference on Perspectives in Nonlinear Dynamics - 2016* 1.1 (2017), pp. 51–60.

[56] N. Boers, B. Bookhagen, J. Marengo, N. Marwan, J.-S. von Storch, and J. Kurths. "Extreme Rainfall of the South American Monsoon System: A Dataset Comparison Using Complex Networks". In: *Journal of Climate* 28.3 (2015), pp. 1031–1056.

[57] C. Dalelane, K. Winderlich, and A. Walter. "Evaluation of Global Teleconnections in CMIP6 Climate Projections using Complex Networks". In: *EGUsphere [preprints]* 2022 (2022), pp. 1–29.

[58] G. Di Capua, D. Coumou, B. van der Hurk, A. Weissheimer, A. G. Turner, and R. V. Donner. "Validation of boreal summer tropical-extratropical causal links in seasonal forecasts". In: *Weather and Climate Dynamics Discussions* 2022 (2022), pp. 1–40.

[59] W. Gregory, J. Stroeve, and M. Tsamados. "Network connectivity between the winter Arctic Oscillation and summer sea ice in CMIP6 models and observations". In: *The Cryosphere* 16.5 (2022), pp. 1653–1673.

[60] P. Nowack, J. Runge, V. Eyring, and J. D. Haigh. "Causal networks for climate model evaluation and constrained projections". In: *Nature Communications* 11.1 (2020), p. 1415.

[61] F. Li, Q. Zhu, W. J. Riley, K. Yuan, H. Wu, and Z. Gui. "Wetter California Projected by CMIP6 Models With Observational Constraints Under a High GHG Emission Scenario". In: *Earth's Future* 10.4 (2022).

[62] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger. "Performance of different synchronization measures in real data: A case study on electroencephalographic signals". In: *Phys. Rev. E* 65.4 (2002), p. 041903.

[63] A. A. Koronovskii and A. E. Hramov. "Wavelet transform analysis of the chaotic synchronization of dynamical systems". In: *Journal of Experimental and Theoretical Physics Letters* 79.7 (2004), pp. 316–319.

[64] Z.-Z. Hu and T. Nitta. "Wavelet Analysis of Summer Rainfall over North China and India and SOI Using 1891-1992 Data". In: *Journal of the Meteorological Society of Japan. Ser. II* 74.6 (1996), pp. 833–844.

Bibliography

[65] W. Jin, Z. Zu-qiang, H. Jin-hai, and Q. Li. "The analysis of the relationship between the subsystems of the Asian summer monsoon". In: *Journal of Tropical Meteorology* 20 (2014), pp. 342–348.

[66] D. Maraun and J. Kurths. "Cross wavelet analysis: significance testing and pitfalls". In: *Nonlinear Processes in Geophysics* 11.4 (2004), pp. 505–514.

[67] F. Lunkeit. "Synchronization experiments with an atmospheric global circulation model". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 11.1 (2001), pp. 47–51.

[68] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization - A Universal Concept in Nonlinear Sciences*. Vol. 12. Cambridge Nonlinear Science Series. Cambridge University Press, 2001, pp. 1–411.

[69] S. Boccaletti, J. Kurths, G. Osipov, D.L. Valladares, and C.S. Zhou. "The synchronization of chaotic systems". In: *Physics Reports* 366.1 (2002), pp. 1–101.

[70] D. Eroglu, J.S.W. Lamb, and T. Pereira. "Synchronisation of chaos and its applications". In: *Contemporary Physics* 58.3 (2017), pp. 207–243.

[71] C. Hugenni. "Horologium oscillatorium". In: *Apud F. Maguet, Paris* (1673).

[72] L. M. Pecora and T. L. Carroll. "Synchronization in chaotic systems". In: *Phys. Rev. Lett.* 64.8 (1990), pp. 821–824.

[73] Stephen G. Penny. "Mathematical foundations of hybrid data assimilation from a synchronization perspective". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.12 (2017), p. 126801.

[74] L. Kocarev and U. Parlitz. "General Approach for Chaotic Synchronization with Applications to Communication". In: *Phys. Rev. Lett.* 74.25 (1995), pp. 5028–5031.

[75] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths. "Phase Synchronization of Chaotic Oscillators". In: *Phys. Rev. Lett.* 76.11 (1996), pp. 1804–1807.

[76] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths. "From Phase to Lag Synchronization in Coupled Chaotic Oscillators". In: *Phys. Rev. Lett.* 78.22 (1997), pp. 4193–4196.

[77] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H.D.I. Abarbanel. "Generalized synchronization of chaos in directionally coupled chaotic systems". In: *Phys. Rev. E* 51.2 (1995), pp. 980–994.

[78] L. Kocarev and U. Parlitz. "Generalized Synchronization, Predictability, and Equivalence of Unidirectionally Coupled Dynamical Systems". In: *Phys. Rev. Lett.* 76.11 (1996), pp. 1816–1819.

[79] A. Hampton and D. H. Zanette. "Measure Synchronization in Coupled Hamiltonian Systems". In: *Phys. Rev. Lett.* 83.11 (1999), pp. 2179–2182.

[80] S. Gupta, S. De, M. S. Janaki, and A. N. Sekar Iyengar. "Exploring the route to measure synchronization in non-linearly coupled Hamiltonian systems". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.11 (2017), p. 113103.

[81] M. Carmen Romano Blasco. "Synchronization analysis by means of recurrences in phase space". doctoralthesis. Universität Potsdam, 2005.

[82] C. Torrence and G. P. Compo. "A practical guide to wavelet analysis". In: *Bulletin of the American Meteorological Society* 79.1 (1998), pp. 61–78.

[83] L. M. Pecora and T. L. Carroll. "Synchronization of chaotic systems". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.9 (2015), p. 097611.

[84] B. Ermentrout. "An adaptive model for synchrony in the firefly Pteroptyx malaccae". In: *Journal of Mathematical Biology* 29.6 (1991), pp. 571–585.

[85] F.A.S. Ferrari, R.L. Viana, S.R. Lopes, and R. Stoop. "Phase synchronization of coupled bursting neurons and the generalized Kuramoto model". In: *Neural Networks* 66 (2015), pp. 107–118.

[86] C. Schäfer, M. G. Rosenblum, J. Kurths, and H.-H. Abel. "Heartbeat synchronized with ventilation". In: *Nature* 392.6673 (1998), pp. 239–240.

[87] C. Schäfer, M. G. Rosenblum, H.-H. Abel, and J. Kurths. "Synchronization in the human cardiorespiratory system". In: *Phys. Rev. E* 60.1 (1999), pp. 857–870.

[88] D. Maraun and J. Kurths. "Epochs of phase coherence between El Niño/Southern Oscillation and Indian monsoon". In: *Geophysical Research Letters* 32.15 (2005).

[89] M. Gelbrecht, N. Boers, and J. Kurths. "Phase coherence between precipitation in South America and Rossby waves". In: *Science Advances* 4.12 (2018), eaau3191.

[90] J. Tian, H. Qiu, G. Wang, Y. Chen, and L. Fu. "Measure synchronization in a two-species bosonic Josephson junction". In: *Phys. Rev. E* 88.3 (2013), p. 032906.

[91] R. M. Gray. "Entropy and Information". In: *Entropy and Information Theory*. New York, NY: Springer New York, 1990, pp. 21–55.

[92] A. Kraskov, H. Stögbauer, and P. Grassberger. "Estimating mutual information". In: *Phys. Rev. E* 69.6 (2004), p. 066138.

[93] T. Schreiber. "Measuring Information Transfer". In: *Phys. Rev. Lett.* 85.2 (2000), pp. 461–464.

[94] R. Silini and C. Masoller. "Fast and effective pseudo transfer entropy for bivariate data-driven causal inference". In: *Scientific Reports* 11.1 (2021), p. 8423.

[95] C. W. J. Granger. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods". In: *Econometrica* 37.3 (1969), pp. 424–438.

Bibliography

[96]   I. I. Mokhov, D. A. Smirnov, P. I. Nakonechny, S. S. Kozlenko, E. P. Seleznev, and J. Kurths. "Alternating mutual influence of El-Niño/Southern Oscillation and Indian monsoon". In: *Geophysical Research Letters* 38.8 (2011).

[97]   J. Arnhold, P. Grassberger, K. Lehnertz, and C.E. Elger. "A robust method for detecting interdependences: application to intracranially recorded EEG". In: *Physica D: Nonlinear Phenomena* 134.4 (1999), pp. 419–430.

[98]   T. Ambrizzi, B. J. Hoskins, and H. Hsu. "Rossby Wave Propagation and Teleconnection Patterns in the Austral Winter". In: *Journal of Atmospheric Sciences* 52.21 (1995), pp. 3661–3672.

[99]   Daniel S. Wilks. "Chapter 3 - Empirical Distributions and Exploratory Data Analysis". In: *Statistical Methods in the Atmospheric Sciences (Fourth Edition)*. Ed. by D. S. Wilks. Fourth Edition. Elsevier, 2019, pp. 23–75.

[100]   M. Katz. *Study Design and Statistical Analysis: A Practical Guide for Clinicians*. Cambridge University Press, 2006.

[101]   J.D. Gibbons and S. Chakraborti. *Nonparametric statistical inference, fourth edition*. 4th ed. Statistics: A Series of Textbooks and Monographs. Boca Raton, FL: CRC Press, 2003.

[102]   M.G. Kendall. "A New Measure of Rank Correlation". In: *Biometrika* 30.1/2 (1938), pp. 81–93.

[103]   M.G. Kendall. "The Treatment of Ties in Ranking Problems". In: *Biometrika* 33.3 (1945), pp. 239–251.

[104]   D. G. Bonett and T. A. Wright. "Sample size requirements for estimating pearson, kendall and spearman correlations". In: *Psychometrika* 65.1 (2000), pp. 23–28.

[105]   W. Xu, Y. Hou, Y.S. Hung, and Y. Zou. "A comparative analysis of Spearman's rho and Kendall's tau in normal and contaminated normal models". In: *Signal Processing* 93.1 (2013), pp. 261–276.

[106]   R. Quian Quiroga, T. Kreuz, and P. Grassberger. "Event synchronization: a simple and fast method to measure synchronicity and time delay patterns". In: *Physical review E* 66.4 (2002), p. 041904.

[107]   J. D. Victor and K. P. Purpura. "Metric-space analysis of spike trains: theory, algorithms and application". In: *Network: Computation in Neural Systems* 8.2 (1997), pp. 127–164.

[108]   I. Ozken, D. Eroglu, T. Stemler, N. Marwan, G. B. Bagci, and J. Kurths. "Transformation-cost time-series method for analyzing irregularly sampled data". In: *Physical Review E* 91 (2015), p. 062911.

[109]   I. Ozken, D. Eroglu, S. F. M. Breitenbach, N. Marwan, L. Tan, U. Tirnakli, and J. Kurths. "Recurrence plot analysis of irregularly sampled data". In: *Physical Review E* 98 (2018), p. 052215.

[110] A. Banerjee, B. Goswami, Y. Hirata, D. Eroglu, B. Merz, J. Kurths, and N. Marwan. "Recurrence analysis of extreme event-like data". In: *Nonlinear Processes in Geophysics* 28.2 (2021), pp. 213–229.

[111] N. Boers, B. Goswami, A. Rheinwalt, B. Bookhagen, B. Hoskins, and J. Kurths. "Complex networks reveal global pattern of extreme-rainfall teleconnections". In: *Nature* 566.7744 (2019), pp. 373–377.

[112] A. Agarwal, N. Marwan, M. Rathinasamy, B. Merz, and J. Kurths. "Multi-scale event synchronization analysis for unravelling climate processes: a wavelet-based approach". In: *Nonlinear Processes in Geophysics* 24.4 (2017), pp. 599–611.

[113] Y. Kuramoto and I. Nishikawa. "Statistical macrodynamics of large dynamical systems. Case of a phase transition in oscillator communities". In: *Journal of Statistical Physics* 49.3 (1987), pp. 569–605.

[114] A. A. Tsonis, K. L. Swanson, and P. J. Roebber. "What Do Networks Have to Do with Climate?." In: *Bulletin of the American Meteorological Society* 87.5 (2006), pp. 585–595.

[115] P. W. Anderson. "More Is Different". In: *Science* 177.4047 (1972), pp. 393–396.

[116] J. A. Dunne, R. J. Williams, and N. D. Martinez. "Food-web structure and network theory: The role of connectance and size". In: *Proceedings of the National Academy of Sciences* 99.20 (2002), pp. 12917–12922.

[117] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. "Epidemic processes in complex networks". In: *Rev. Mod. Phys.* 87.3 (2015), pp. 925–979.

[118] D. Witthaut, F. Hellmann, J. Kurths, S. Kettemann, H. Meyer-Ortmanns, and M. Timme. "Collective nonlinear dynamics and self-organization in decentralized power grids". In: *Rev. Mod. Phys.* 94.1 (2022), p. 015005.

[119] M. Golosovsky and S. Solomon. "Growing complex network of citations of scientific papers: Modeling and measurements". In: *Phys. Rev. E* 95.1 (2017), p. 012324.

[120] M. Faloutsos, P. Faloutsos, and C. Faloutsos. "On Power-Law Relationships of the Internet Topology". In: *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*. SIG-COMM '99. Cambridge, Massachusetts, USA: Association for Computing Machinery, 1999, pp. 251–262.

[121] F. Wolf, S. Lehmann, and P. Lorenz-Spreen. "Successive Cohorts of Twitter Users Show Increasing Activity and Shrinking Content Horizons". In: *Journal of Quantitative Description: Digital Media* 2 (2022).

[122] R. Shields. "Cultural Topology: The Seven Bridges of Königsburg, 1736". In: *Theory, Culture & Society* 29.4-5 (2012), pp. 43–57.

[123] X. Sun, S. Wandelt, C. Zheng, and A. Zhang. "COVID-19 pandemic and air transportation: Successfully navigating the paper hurricane". In: *Journal of Air Transport Management* 94 (2021), p. 102062.

[124] D. J. Watts and S. Strogatz. "Collective dynamics of 'small-world' networks". In: *Nature* 393 (1998), pp. 440–442.

[125] R. J. Wilson and J. J. Watkins. *Graphs: An Introductory Approach–A First Course in Discrete Mathematics.* en. Nashville, TN: John Wiley & Sons, 1990.

[126] M. E. J. Newman. "Mixing patterns in networks". In: *Phys. Rev. E* 67.2 (2003), p. 026126.

[127] N. R. Chopde and M. Nichat. "Landmark based shortest path detection by using A* and Haversine formula". In: *International Journal of Innovative Research in Computer and Communication Engineering* 1.2 (2013), pp. 298–302.

[128] L. Tupikina, K. Rehfeld, N. Molkenthin, V. Stolbova, N. Marwan, and J. Kurths. "Characterizing the evolution of climate networks". In: *Nonlinear Processes in Geophysics* 21.3 (2014), pp. 705–711.

[129] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. "Random graphs with arbitrary degree distributions and their applications". In: *Phys. Rev. E* 64.2 (2001), p. 026118.

[130] L. Barnett, E. Di Paolo, and S. Bullock. "Spatially embedded random networks". In: *Phys. Rev. E* 76.5 (2007), p. 056115.

[131] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley. "Classes of small-world networks". In: *Proceedings of the National Academy of Sciences* 97.21 (2000), pp. 11149–11152.

[132] S. N. Dorogovtsev and J. F. F. Mendes. "Language as an evolving word web". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268.1485 (2001), pp. 2603–2606.

[133] R. F. i Cancho and R. V. Solé. "The small world of human language". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268.1482 (2001), pp. 2261–2265.

[134] A.-L. Barabási and R. Albert. "Emergence of Scaling in Random Networks". In: *Science* 286.5439 (1999), pp. 509–512.

[135] E. Bullmore and O. Sporns. "Complex brain networks: graph theoretical analysis of structural and functional systems". In: *Nature Reviews Neuroscience* 10.3 (2009), pp. 186–198.

[136] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths. "Recurrence networks—a novel paradigm for nonlinear time series analysis". In: *New Journal of Physics* 12.3 (2010), p. 033025.

[137] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño. "From time series to complex networks: The visibility graph". In: *Proceedings of the National Academy of Sciences* 105.13 (2008), pp. 4972–4975.

[138] R. Albert and A. L. Barabási. "Topology of Evolving Networks: Local Events and Universality". In: *Phys. Rev. Lett.* 85.24 (2000), pp. 5234–5237.

[139] J. Hlinka, D. Hartman, N. Jajcay, M. Vejmelka, R. Donner, N. Marwan, J. Kurths, and M. Paluš. "Regional and inter-regional effects in evolving climate networks". In: *Nonlinear Processes in Geophysics* 21.2 (2014), pp. 451–462.

[140] N. Marwan, J. Donges, A. Radebach, J. Runge, and J. Kurths. "Evolving Climate Networks". In: *International Symposium on Nonlinear Theory and its Applications NOLTA2010*. Krakow, Poland, 2010, pp. 3–6.

[141] A. Barrat, M. Barthélemy, and A. Vespignani. "Weighted Evolving Networks: Coupling Topology and Weight Dynamics". In: *Phys. Rev. Lett.* 92.22 (2004), p. 228701.

[142] T. Gross and B. Blasius. "Adaptive coevolutionary networks: a review". In: *Journal of The Royal Society Interface* 5.20 (2008), pp. 259–271.

[143] W. M. Dines. "Correlation in Seasonal Variations of Weather. Memoirs of the Indian Meteorological Department. Vols. XX. and XXI. By Dr. Gilbert T. Walker, F.R.S. Simla, 1910 to 1915". In: *Quarterly Journal of the Royal Meteorological Society* 42.178 (1916), pp. 129–132.

[144] A.A. Tsonis and P.J. Roebber. "The architecture of the climate network". In: *Physica A: Statistical Mechanics and its Applications* 333 (2004), pp. 497–504.

[145] A. A. Tsonis and K. L. Swanson. "Topology and Predictability of El Niño and La Niña Networks". In: *Phys. Rev. Lett.* 100.22 (2008), p. 228502.

[146] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. "The backbone of the climate network". In: *Europhysics Letters* 87.4 (2009), p. 48007.

[147] J. F. Donges, H. C. H. Schultz, N. Marwan, Y. Zou, and J. Kurths. "Investigating the topology of interacting networks". In: *The European Physical Journal B* 84.4 (2011), pp. 635–651.

[148] A. Rheinwalt, N. Marwan, J. Kurths, P. Werner, and F. Gerstengarbe. "Boundary effects in network measures of spatially embedded networks". In: *EPL (Europhysics Letters)* 100.2 (2012), p. 28002.

[149] A. H. Fink and P. S. "Tropical cyclones". In: *Naturwissenschaften* 85.10 (1998), pp. 482–493.

[150] R. Emerton et al. "Emergency flood bulletins for Cyclones Idai and Kenneth: A critical evaluation of the use of global flood forecasts for international humanitarian preparedness and response". In: *International Journal of Disaster Risk Reduction* 50 (2020), p. 101811.

[151] T. Knutson, S. J. Camargo, J. C. L. Chan, K. Emanuel, C.-H. Ho, J. Kossin, M. Mohapatra, M. Satoh, M. Sugi, K. Walsh, et al. "Tropical cyclones and climate change assessment: Part I: Detection and attribution". In: *Bulletin of the American Meteorological Society* 100.10 (2019), pp. 1987–2007.

Bibliography

[152] C. Ciemer, N. Boers, H. Barbosa, J. Kurths, and A. Rammig. "Temporal evolution of the spatial covariability of rainfall in South America". In: *Climate Dynamics* 51 (2018).

[153] U. Ozturk, N. Marwan, O. Korup, H. Saito, A. Agarwal, M. J. Grossman, M. Zaiki, and J. Kurths. "Complex networks for tracking extreme rainfall during typhoons". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.7 (2018), p. 075301.

[154] U. Ozturk, N. Malik, K. Cheung, N. Marwan, and J. Kurths. "A network-based comparative study of extreme tropical and frontal storm rainfall over Japan". In: *Climate Dynamics* 53 (2019).

[155] R. Albert, I. Albert, and G. L. Nakarado. "Structural vulnerability of the North American power grid". In: *Phys. Rev. E* 69.2 (2004), p. 025103.

[156] L. Li, T. Zeng, L. Cai, Y. Wang, and Y. Li. "The local-world evolving network model for power grid considering different voltage levels". In: *The 11th IET International Conference on Advances in Power System Control, Operation and Management (APSCOM 2018)*. 2018, pp. 1–4.

[157] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, and J. Kurths. "Hierarchical Organization Unveiled by Functional Connectivity in Complex Brain Networks". In: *Phys. Rev. Lett.* 97.23 (2006), p. 238103.

[158] K. Lehnertz, G. Ansmann, S. Bialonski, H. Dickten, C. Geier, and S. Porz. "Evolving networks in the human epileptic brain". In: *Physica D: Nonlinear Phenomena* 267 (2014). Evolving Dynamical Networks, pp. 7–15.

[159] K. Cheung and U. Ozturk. "Synchronization of extreme rainfall during the Australian summer monsoon: Complex network perspectives". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30.6 (2020), p. 063117.

[160] K. Taira, A. G. Nair, and S. L. Brunton. "Network structure of two-dimensional decaying isotropic turbulence". In: *J. Fluid Mech.* 795 (2016).

[161] 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. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (Last accessed 02/12/2020). 2018.

[162] H. Hersbach et al. "The ERA5 global reanalysis". In: *Quarterly Journal of the Royal Meteorological Society* 146.730 (2020), pp. 1999–2049.

[163] J. Kouroutzoglou, H. Flocas, I. Simmonds, K. Keay, and M. Hatzaki. "Assessing characteristics of Mediterranean explosive cyclones for different data resolution". In: *Theoretical and Applied Climatology* 105 (2011), pp. 263–275.

[164] C. W. Landsea and J. L. Franklin. "Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format". In: *Monthly Weather Review* 141.10 (2013), pp. 3576–3592.

[165] K. Yamasaki, A. Gozolchiani, and S. Havlin. "Climate Networks around the Globe are Significantly Affected by El Niño". In: *Phys. Rev. Lett.* 100.22 (2008), p. 228501.

[166] W.M. Gray. "Global view of the origin of Tropical disturbances and storms". In: *Monthly Weather Review* 96.10 (1968), pp. 669–700.

[167] S. J. Camargo, K. A. Emanuel, and A. H. Sobel. "Use of a Genesis Potential Index to Diagnose ENSO Effects on Tropical Cyclone Genesis". In: *Journal of Climate* 20.19 (2007), pp. 4819–4834.

[168] M. Wahiduzzaman, E. Oliver, S. J. Wotherspoon, and N. J. Holbrook. "A climatological model of North Indian Ocean tropical cyclone genesis, tracks and landfall". In: *Climate Dynamics* 49 (2017), pp. 2585–2603.

[169] Z. Li, W. Yu, T. Li, V. S. N. Murty, and F. Tangang. "Bimodal Character of Cyclone Climatology in the Bay of Bengal Modulated by Monsoon Seasonal Cycle". In: *Journal of Climate* 26.3 (2013), pp. 1033–1046.

[170] N. Vissa, A.N.V. Satyanarayana, and B. Prasad Kumar. "Intensity of tropical cyclones during pre- and post-monsoon seasons in relation to accumulated tropical cyclone heat potential over Bay of Bengal". In: *Natural Hazards* 68 (2013).

[171] NASA Earth Observatory. *Seroja Slams Australia*. https://earthobservatory.nasa.gov/images/148180/seroja-slams-australia?src=eoa-iotd. (last accessed on 06/04/2022). 2021.

[172] Y.-A. Liou, J.-C. Liu, M.-X. Wu, Y.-J. Lee, C.-H. Cheng, C.-P. Kuei, and R.-M. Hong. "Generalized empirical formulas of threshold distance to characterize cyclone–cyclone interactions". In: *IEEE Transactions on Geoscience and Remote Sensing* 54.6 (2016), pp. 3502–3512.

[173] J.-C. Liu, Y.-A. Liou, M.-X. Wu, Y.-J. Lee, C.-H. Cheng, C.-P. Kuei, and R.-M. Hong. "Analysis of interactions among two tropical depressions and typhoons Tembin and Bolaven (2012) in Pacific Ocean by using satellite cloud images". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.3 (2014), pp. 1394–1402.

[174] E. W. Hoover. "Relative motion of hurricane pairs". In: *Monthly Weather Review* 89.7 (1961), pp. 251–255.

[175] S. Brand. "Interaction of binary tropical cyclones of the western North Pacific Ocean". In: *Journal of Applied Meteorology and Climatology* 9.3 (1970), pp. 433–441.

[176] L. E. Carr and R. L. Elsberry. "Dynamical Tropical Cyclone Track Forecast Errors. Part I: Tropical Region Error Sources". In: *Weather and Forecasting* 15.6 (2000), pp. 641–661.

[177] E.-J. Cha, S.-g. Yun, I.-J. Moon, and D.-H. Kim. "Binary interaction of typhoons Soulik and Cimaron in 2018 – Part I: Observational characteristics and forecast error". In: *Tropical Cyclone Research and Review* 10.1 (2021), pp. 32–42.

[178] A. D. Evans and R. J. Falvey. *Annual Tropical Cyclone Report, Joint Typhoon Warning Center (JTWC)*. `https://www.metoc.navy.mil/jtwc/products/atcr/2012atcr.pdf`. (Last Accessed on 02/22/2022). 2012.

[179] Y.-A. Liou and R. S. Pandey. "Interactions between typhoons Parma and Melor (2009) in North West Pacific Ocean". In: *Weather and Climate Extremes* 29 (2020), p. 100272.

[180] Y.-A. Liou, J.-C. Liu, C.-C. Liu, C.-H. Chen, K.-A. Nguyen, and J. P Terry. "Consecutive dual-vortex interactions between quadruple typhoons Noru, Kulap, Nesat and Haitang during the 2017 North Pacific typhoon season". In: *Remote Sensing* 11.16 (2019), p. 1843.

[181] J. Nott. "A 6000 year tropical cyclone record from Western Australia". In: *Quaternary Science Reviews* 30.5-6 (2011), pp. 713–722.

[182] K. Dong and C. J. Neumann. "On the relative motion of binary tropical cyclones". In: *Monthly Weather Review* 111.5 (1983), pp. 945–953.

[183] M. Lander and G. J. Holland. "On the interaction of tropical-cyclone-scale vortices. I: Observations". In: *Quarterly Journal of the Royal Meteorological Society* 119.514 (1993), pp. 1347–1361.

[184] C. K. Tang, J. C. L. Chan, and M. Yamaguchi. "Large tropical cyclone track forecast errors of global numerical weather prediction models in western North Pacific basin". In: *Tropical Cyclone Research and Review* 10.3 (2021), pp. 151–169.

[185] E. A. Ritchie and G. J. Holland. "On the interaction of tropical-cyclone-scale vortices. II: Discrete vortex patches". In: *Quarterly Journal of the Royal Meteorological Society* 119.514 (1993), pp. 1363–1379.

[186] J. Liang and L. Wu. "Sudden Track Changes of Tropical Cyclones in Monsoon Gyres: Full-Physics, Idealized Numerical Experiments". In: *Journal of the Atmospheric Sciences* 72.4 (2015), pp. 1307–1322.

[187] T. Leweke, S. Le Dizes, and C. H. K. Williamson. "Dynamics and instabilities of vortex pairs". In: *Annual Review of Fluid Mechanics* 48 (2016), pp. 507–541.

[188] P. Meunier, S. Le Dizes, and T. Leweke. "Physics of vortex merging". In: *Comptes Rendus Physique* 6.4-5 (2005), pp. 431–450.

[189] C. Cerretelli and C. H. K. Williamson. "The physical mechanism for vortex merging". In: *J. Fluid Mech.* 475 (2003), pp. 41–77.

[190] S. Wei-Jen Chang. "A numerical study of the interactions between two tropical cyclones". In: *Monthly Weather Review* 111.9 (1983), pp. 1806–1817.

[191] M. DeMaria and J. C. L. Chan. "Comments on "A numerical study of the interactions between two tropical cyclones"". In: *Monthly Weather Review* 112.8 (1984), pp. 1643–1645.

[192] K. V. Roberts and J. P. Christiansen. "Topics in computational fluid mechanics". In: *Computer Physics Communications* 3 (1972), pp. 14–32.

[193] P. G. Saffman and R. Szeto. "Equilibrium shapes of a pair of equal uniform vortices". In: *Physics of Fluids* 23.12 (1980), pp. 2339–2342.

[194] E. A. Overman and N. J. Zabusky. "Evolution and merger of isolated vortex structures". In: *Physics of Fluids* 25.8 (1982), pp. 1297–1305.

[195] D. G. Dritschel. "The nonlinear evolution of rotating configurations of uniform vorticity". In: *J. Fluid Mech.* 172 (1986), pp. 157–182.

[196] U. Meunier P.and Ehrenstein, T. Leweke, and M. Rossi. "A merging criterion for two-dimensional co-rotating vortices". In: *Physics of Fluids* 14.8 (2002), pp. 2757–2766.

[197] Ch. Josserand and M. Rossi. "The merging of two co-rotating vortices: a numerical study". In: *European Journal of Mechanics-B/Fluids* 26.6 (2007), pp. 779–794.

[198] L. Li and X. Ge. "Intensity Change of NORU (2017) During Binary Tropical Cyclones Interaction". In: *Asia-Pacific Journal of Atmospheric Sciences* 57.1 (2021), pp. 135–147.

[199] A. Benfield. *Global Catastrophe Recap August 2017.* http://thoughtleadership. aon.com. (Accessed on 05/11/2022).

[200] C. C. Lam and S. T. Lai. "Use of ECMWF 850-hPa vorticity fields in the forecasting of tropical cyclones and intense lows in June- July 1994". In: *Proc. 9th Guangdong- Hong Kong- Macau Joint Seminar on Hazardous Weather* (1994), pp. 137–156.

[201] E. Flaounas, V. Kotroni, K. Lagouvardos, and I. Flaounas. "CycloTRACK (v1. 0)–tracking winter extratropical cyclones based on relative vorticity: sensitivity to data filtering and other relevant parameters". In: *Geoscientific Model Development* 7.4 (2014), pp. 1841–1853.

[202] S. Chen, W. Li, Y. Lu, and Z. Wen. "Variations of latent heat flux during tropical cyclones over the South China Sea". In: *Meteorological Applications* 21.3 (2014), pp. 717–723.

[203] Y. Wu, S. Chen, W. Li, R. Fang, and H. Liu. "Relative vorticity is the major environmental factor controlling tropical cyclone intensification over the Western North Pacific". In: *Atmospheric Research* 237 (2020), p. 104874.

[204] W. M. Gray. "Global view of the origin of tropical disturbances and storms". In: *Monthly Weather Review* 96.10 (1968), pp. 669–700.

[205] C.-C. Wu, T.-S. Huang, W.-P. Huang, and K.-H. Chou. "A new look at the binary interaction: Potential vorticity diagnosis of the unusual southward movement of Tropical Storm Bopha (2000) and its interaction with Supertyphoon Saomai (2000)". In: *Monthly Weather Review* 131.7 (2003), pp. 1289–1300.

[206] S.-S. Ho and A. Talukder. "Cyclone Tracking Using Multiple Satellite Data Sources via Spatial-Temporal Knowledge Transfer". In: *AAAI-08 workshop, Transfer Learning for Complex Tasks* (2008).

[207] J. Molinari, D. Vollaro, and F. Robasky. "Use of ECMWF operational analyses for studies of the tropical cyclone environment". In: *Meteorl. Atmos. Phys.* 47 (1992), pp. 127–144.

[208] M. T. Pillay and J. M. Fitchett. "On the conditions of formation of Southern Hemisphere tropical cyclones". In: *Weather and Climate Extremes* 34 (2021), p. 100376.

[209] C. J. Neumann. "Use of deep-layer mean geopotential height fields in statistical prediction of Tropical Cyclone motion". In: *Bulletin of the American Meteorological Society* 60.5 (1979), pp. 584–584.

[210] T. B. Kimberlain and M. J. Breman. "Tropical cyclone motion". In: *Global Guide to Tropical Cyclone Forecasting* (2017), pp. 63–155.

[211] C. S. Velden and L. M. Leslie. "The Basic Relationship between Tropical Cyclone Intensity and the Depth of the Environmental Steering Layer in the Australian Region". In: *Weather and Forecasting* 6.2 (1991), pp. 244–253.

[212] A. Krishnan, R. I. Sujith, N. Marwan, and J. Kurths. "Suppression of thermoacoustic instability by targeting the hubs of the turbulent networks in a bluff body stabilized combustor". In: *J. Fluid Mech.* 916 (2021).

[213] M. G. Meena and K. Taira. "Identifying vortical network connectors for turbulent flow modification". In: *J. Fluid Mech.* 915 (2021).

[214] J. D. Jackson. *Classical electrodynamics.* 1999.

[215] R. Shankar. "Fundamentals of Physics II". In: Yale University Press, 2016.

[216] J. D. Anderson. In: *Fundamentals of Aerodynamics.* McGraw, 2009.

[217] Q. Deng, L. Smith, and A. Majda. "Tropical cyclogenesis and vertical shear in a moist Boussinesq model". In: *Journal of Fluid Mechanics* 706 (2012), pp. 384–412.

[218] Y. Moon, D. S. Nolan, and M. Iskandarani. "On the Use of Two-Dimensional Incompressible Flow to Study Secondary Eyewall Formation in Tropical Cyclones". In: *Journal of the Atmospheric Sciences* 67.12 (2010), pp. 3765–3773.

[219] G. J. Holland and G. S. Dietachmayer. "On the interaction of tropical-cyclone-scale vortices. III: Continuous barotropic vortices". In: *Quarterly Journal of the Royal Meteorological Society* 119.514 (1993), pp. 1381–1398.

[220] Weather Underground. *Hurricane & Tropical Cyclones*. `https://www.wunderground.com/hurricane`. (Last accessed on 11/28/2022).

[221] B. J. Hoskins and K. I. Hodges. "New perspectives on the Northern Hemisphere winter storm tracks". In: *Journal of the Atmospheric Sciences* 59.6 (2002), pp. 1041–1061.

[222] C. Cerretelli and C. H. K. Williamson. "A new family of uniform vortices related to vortex configurations before merging". In: *J. Fluid Mech.* 493 (2003), pp. 219–229.

[223] National Geographic Society. *The Coriolis Effect: Earth's Rotation and Its Effect on Weather*. `https://education.nationalgeographic.org/resource/coriolis-effect`. (Last accessed on 11/28/2022).

[224] Y. Couder and C. Basdevant. "Experimental and numerical study of vortex couples in two-dimensional flows". In: *Journal of Fluid Mechanics* 173 (1986), pp. 225–251.

[225] L. K. Brandt and K. K Nomura. "The physics of vortex merger and the effects of ambient stable stratification". In: *J. Fluid Mech.* 592 (2007), pp. 413–446.

[226] K. Song, L. Tao, and J. Gao. "Rapid Weakening of Tropical Cyclones in Monsoon Gyres Over the Western North Pacific: A Revisit". In: *Frontiers in Earth Science* 9 (2021), p. 507.

[227] B. N. Goswami and R.S.A. Mohan. "Intraseasonal oscillations and interannual variability of the Indian summer monsoon". In: *Journal of Climate* 14.6 (2001), pp. 1180–1198.

[228] Q. Ding and B. Wang. "Circumglobal teleconnection in the Northern Hemisphere summer". In: *Journal of Climate* 18.17 (2005), pp. 3483–3505.

[229] M. Rajeevan, S. Gadgil, and J. Bhate. "Active and break spells of the Indian summer monsoon". In: *Journal of earth system science* 119.3 (2010), pp. 229–247.

[230] Y. Ding, P. Liang, Y. Liu, and Y. Zhang. "Multiscale variability of Meiyu and its prediction: A new review". In: *Journal of Geophysical Research: Atmospheres* 125.7 (2020), e2019JD031496.

[231] Y. Ding and J. Chan. "The East Asian summer monsoon: an overview". In: *Meteorology and Atmospheric Physics* 89.1 (2005), pp. 117–142.

[232] J. H. Feldhoff, R. V. Donner, J. F. Donges, N. Marwan, and J. Kurths. "Geometric detection of coupling directions by means of inter-system recurrence networks". In: *Physics Letters A* 376.46 (2012), pp. 3504–3513.

[233] K. Rehfeld, N. Marwan, S.F.M. Breitenbach, and J. Kurths. "Late Holocene Asian summer monsoon dynamics from small but complex networks of paleoclimate data". In: *Climate dynamics* 41.1 (2013), pp. 3–19.

[234] Y. Ding. *Monsoons over China*. Vol. 16. Springer Science & Business Media, 2013.

[235] K.-J. Ha, Y.-W. Seo, J.-Y. Lee, R. H. Kripalani, and K.-S. Yun. "Linkages between the South and East Asian summer monsoons: a review and revisit". In: *Climate Dynamics* 51.11 (2018), pp. 4207–4227.

[236] R. H. Kripalani and S. V. Singh. "Large scale aspects of India-China summer monsoon rainfall". In: *Advances in Atmospheric Sciences* 10.1 (1993), pp. 71–84.

[237] R. H. Kripalani and A. Kulkarni. "Monsoon rainfall variations and teleconnections over South and East Asia". In: *International Journal of Climatology: A Journal of the Royal Meteorological Society* 21.5 (2001), pp. 603–616.

[238] Y. Liu and Y. Ding. "Analysis and numerical simulations of the teleconnection between Indian summer monsoon and precipitation in North China". In: *Journal of Meteorological Research* 22.4 (2008), pp. 489–501.

[239] R. Wu. "Relationship between Indian and East Asian summer rainfall variations". In: *Advances in Atmospheric Sciences* 34.1 (2017), pp. 4–15.

[240] 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". In: *Climate Dynamics* 57.9 (2021), pp. 2615–2636.

[241] Y. Liu and Y. Ding. "Teleconnection between the Indian summer monsoon onset and the Meiyu over the Yangtze River Valley". In: *Science in China Series D: Earth Sciences* 51.7 (2008), pp. 1021–1035.

[242] Y. Liu, P. Liang, and Y. Sun. *The Asian summer monsoon: characteristics, variability, teleconnections and projection*. Elsevier, 2019.

[243] S. Woo, G. P. Singh, J.-H. Oh, and K.-M. Lee. "Possible teleconnections between East and South Asian summer monsoon precipitation in projected future climate change". In: *Meteorology and Atmospheric Physics* 131.3 (2019), pp. 375–387.

[244] Y. L. Serra, X. Jiang, B. Tian, J. Amador-Astua, E. D. Maloney, and G. N. Kiladis. "Tropical intraseasonal modes of the atmosphere". In: *Annual Review of Environment and Resources* 39 (2014), pp. 189–215.

[245] R. A. Madden and P. R. Julian. "Observations of the 40–50-day tropical oscillation—A review". In: *Monthly weather review* 122.5 (1994), pp. 814–837.

[246] M. C. Wheeler and H. H. Hendon. "An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction". In: *Monthly weather review* 132.8 (2004), pp. 1917–1932.

[247] T. Yasunari. "Cloudiness fluctuations associated with the Northern Hemisphere summer monsoon". In: *Journal of the Meteorological Society of Japan. Ser. II* 57.3 (1979), pp. 227–242.

[248] M. Murakami. "Analysis of the deep convective activity over the western Pacific and southeast Asia Part I: Diurnal variation". In: *Journal of the Meteorological Society of Japan. Ser. II* 61.1 (1983), pp. 60–76.

[249] B. Wang, P. J. Webster, and H. Teng. "Antecedents and self-induction of active-break south Asian monsoon unraveled by satellites". In: *Geophysical Research Letters* 32.4 (2005).

[250] K.-S. Yun, B. Ren, K.-J. Ha, J. Chan, and J.-G. Jhun. "The 30–60-day oscillation in the East Asian summer monsoon and its time-dependent association with the ENSO". In: *Tellus A: Dynamic Meteorology and Oceanography* 61.5 (2008), pp. 565–578.

[251] 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". In: *Climate Dynamics* 40.1 (2013), pp. 493–509.

[252] J. Ju and J. Slingo. "The Asian summer monsoon and ENSO". In: *Quarterly Journal of the Royal Meteorological Society* 121.525 (1995), pp. 1133–1168.

[253] B. Wang and Z. Fan. "Choice of South Asian Summer Monsoon Indices". In: *Bulletin of the American Meteorological Society* 80.4 (1999), pp. 629–638.

[254] S. N. Sandeep and F. Stordal. "Use of daily outgoing longwave radiation (OLR) data in detecting precipitation extremes in the tropics". In: *Remote Sensing Letters* 4.6 (2013), pp. 570–578.

[255] G. Di Capua, M. Kretschmer, R. V. Donner, B. Van Den Hurk, R. Vellore, R. Krishnan, and D. Coumou. "Tropical and mid-latitude teleconnections interacting with the Indian summer monsoon rainfall: a theory-guided causal effect network approach". In: *Earth System Dynamics* 11.1 (2020), pp. 17–34.

[256] J. Runge, V. Petoukhov, J. F. Donges, J.v Hlinka, N. Jajcay, M. Vejmelka, D. Hartman, N. Marwan, M. Paluš, and J. Kurths. "Identifying causal gateways and mediators in complex spatio-temporal systems". In: *Nature communications* 6.1 (2015), pp. 1–10.

[257] N. Marwan and J. Kurths. "Complex network based techniques to identify extreme events and (sudden) transitions in spatio-temporal systems". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.9 (2015), p. 097609.

[258] F. Wolf, U. Ozturk, K. Cheung, and R. V. Donner. "Spatiotemporal patterns of synchronous heavy rainfall events in East Asia during the Baiu season". In: *Earth System Dynamics* 12.1 (2021), pp. 295–312.

[259] G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker. "The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales". In: *Journal of hydrometeorology* 8.1 (2007), pp. 38–55.

[260] I. I. Mokhov, D. A. Smirnov, P. I. Nakonechny, S. S. Kozlenko, E. P. Seleznev, and J. Kurths. "Alternating mutual influence of El-Niño/Southern Oscillation and Indian monsoon". In: *Geophysical Research Letters* 38.8 (2011).

[261] S.-P. Xie, Y. Kosaka, Y. Du, K. Hu, J. S. Chowdary, and G. Huang. "Indo-western Pacific ocean capacitor and coherent climate anomalies in post-ENSO summer: A review". In: *Advances in Atmospheric Sciences* 33.4 (2016), pp. 411–432.

[262] A. Radebach, R. V. Donner, J. Runge, J. F. Donges, and J. Kurths. "Disentangling different types of El Niño episodes by evolving climate network analysis". In: *Phys. Rev. E* 88.5 (2013), p. 052807.

[263] M. Gelbrecht, N. Boers, and J. Kurths. "Variability of the low-level circulation of the South American Monsoon analysed with complex networks". In: *The European Physical Journal Special Topics* 230.14 (2021), pp. 3101–3120.

[264] A. Libertino, A. Sharma, V. Lakshmi, and P. Claps. "A global assessment of the timing of extreme rainfall from TRMM and GPM for improving hydrologic design". In: *Environmental Research Letters* 11.5 (2016), p. 054003.

[265] S. Gadgil and P. V. Joseph. "On breaks of the Indian monsoon". In: *Journal of Earth System Science* 112.4 (2003), pp. 529–558.

[266] K.-S. Choi, C.-C. Wu, and E.-J. Cha. "Change of tropical cyclone activity by Pacific-Japan teleconnection pattern in the western North Pacific". In: *Journal of Geophysical Research: Atmospheres* 115.D19 (2010).

[267] R.-Y. Lu, J.-H. Oh, and B.-J. Kim. "A teleconnection pattern in upper-level meridional wind over the North African and Eurasian continent in summer". In: *Tellus A: Dynamic Meteorology and Oceanography* 54.1 (2002), pp. 44–55.

[268] S.-S. Lee, Y.-W. Seo, K.-J. Ha, and J.-G. Jhun. "Impact of the western North Pacific subtropical high on the East Asian monsoon precipitation and the Indian Ocean precipitation in the boreal summertime". In: *Asia-Pacific Journal of Atmospheric Sciences* 49.2 (2013), pp. 171–182.

[269] R. Lu and B. Dong. "Westward extension of North Pacific subtropical high in summer". In: *Journal of the Meteorological Society of Japan. Ser. II* 79.6 (2001), pp. 1229–1241.

[270] A. Pathak, S. Ghosh, J. A. Martinez, F. Dominguez, and P. Kumar. "Role of oceanic and land moisture sources and transport in the seasonal and inter-annual variability of summer monsoon in India". In: *Journal of Climate* 30.5 (2017), pp. 1839–1859.

[271] T. F. Cheng and M. Lu. "Moisture source–receptor network of the East Asian summer monsoon land regions and the associated atmospheric steerings". In: *Journal of Climate* 33.21 (2020), pp. 9213–9231.

[272] T.-J. Zhou and R.-C. Yu. "Atmospheric water vapor transport associated with typical anomalous summer rainfall patterns in China". In: *Journal of Geophysical Research: Atmospheres* 110.D8 (2005).

[273] Y. Yang, T. Zhao, G. Ni, and T. Sun. "Atmospheric rivers over the Bay of Bengal lead to northern Indian extreme rainfall". In: *International Journal of Climatology* 38.2 (2018), pp. 1010–1021.

[274] M. Pan and M. Lu. "A novel atmospheric river identification algorithm". In: *Water Resources Research* 55.7 (2019), pp. 6069–6087.

[275] M. Pan and M. Lu. "East Asia atmospheric river catalog: Annual cycle, transition mechanism, and precipitation". In: *Geophysical Research Letters* 47.15 (2020), e2020GL089477.

[276] W. Wei, R. Zhang, M. Wen, X. Rong, and T. Li. "Impact of Indian summer monsoon on the South Asian High and its influence on summer rainfall over China". In: *Climate Dynamics* 43.5 (2014), pp. 1257–1269.

[277] W. Wei, M. Zhang R.and Wen, and S. Yang. "Relationship between the Asian westerly jet stream and summer rainfall over central Asia and North China: Roles of the Indian monsoon and the South Asian high". In: *Journal of Climate* 30.2 (2017), pp. 537–552.

[278] B. Wang, R. Wu, and K-M. Lau. "Interannual Variability of the Asian Summer Monsoon: Contrasts between the Indian and the Western North Pacific–East Asian Monsoons". In: *Journal of Climate* 14.20 (2001), pp. 4073–4090.

[279] T. Enomoto, B. J. Hoskins, and Y. Matsuda. "The formation mechanism of the Bonin high in August". In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 129.587 (2003), pp. 157–178.

[280] V. Krishnamurthy and R. S. Ajayamohan. "Composite structure of monsoon low pressure systems and its relation to Indian rainfall". In: *Journal of Climate* 23.16 (2010), pp. 4285–4305.

[281] K. M. R. Hunt and J. K. Fletcher. "The relationship between Indian monsoon rainfall and low-pressure systems". In: *Climate Dynamics* 53.3 (2019), pp. 1859–1871.

[282] D. M. Straus and V. Krishnamurthy. "The preferred structure of the interannual Indian monsoon variability". In: *Atmospheric and Oceanic.* Springer, 2007, pp. 1717–1732.

[283] K.-M. Lau and M.-T. Li. "The monsoon of East Asia and its global associations—A survey". In: *Bulletin of the American Meteorological Society* 65.2 (1984), pp. 114–125.

[284] T. Nitta and Z.-Z. Hu. "Summer Climate Variability in China and Its Association with 500 hPa Height and Tropical Convection". In: *Journal of the Meteorological Society of Japan. Ser. II* 74.4 (1996), pp. 425–445.

[285] C. P. Chang, Y. Zhang, and T. Li. "Interannual and interdecadal variations of the East Asian summer monsoon and tropical Pacific SSTs. Part I: Roles of the subtropical ridge". In: *Journal of Climate* 13.24 (2000), pp. 4310–4325.

[286] Y. Kosaka, H. Nakamura, M. Watanabe, and M. Kimoto. "Analysis on the dynamics of a wave-like teleconnection pattern along the summertime Asian jet based on a reanalysis dataset and climate model simulations". In: *Journal of the Meteorological Society of Japan. Ser. II* 87.3 (2009), pp. 561–580.

[287] M. Yanai and G.-X. Wu. "Effects of the Tibetan Plateau". In: *The Asian Monsoon.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 513–549.

[288] Q. Ding and B. Wang. "Intraseasonal Teleconnection between the Summer Eurasian Wave Train and the Indian Monsoon". In: *Journal of Climate* 20.15 (2007), pp. 3751–3767.

[289] T. Enomoto. "Interannual Variability of the Bonin High Associated with the Propagation of Rossby Waves along the Asian Jet". In: *Journal of the Meteorological Society of Japan. Ser. II* 82.4 (2004), pp. 1019–1034.

[290] A. Dethof, A. O'neill, J. M. Slingo, and H. G. J. Smit. "A mechanism for moistening the lower stratosphere involving the Asian summer monsoon". In: *Quarterly Journal of the Royal Meteorological Society* 125.556 (1999), pp. 1079–1106.

[291] C. Zhang. "Madden-Julian Oscillation". In: *Reviews of Geophysics* 43.2 (2005).

[292] D. S. Pai, J. Bhate, O. P. Sreejith, and H. R. Hatwar. "Impact of MJO on the intraseasonal variation of summer monsoon rainfall over India". In: *Climate Dynamics* 36.1 (2011), pp. 41–55.

[293] R. Bhatla, M. Singh, and D. R. Pattanaik. "Impact of Madden-Julian oscillation on onset of summer monsoon over India". In: *Theoretical and Applied Climatology* 128.1 (2017), pp. 381–391.

[294] X. Li, G. Gollan, R. J. Greatbatch, and R. Lu. "Impact of the MJO on the interannual variation of the Pacific–Japan mode of the East Asian summer monsoon". In: *Climate Dynamics* 52.5 (2019), pp. 3489–3501.

[295] B. Wang and X. Xie. "A model for the boreal summer intraseasonal oscillation". In: *Journal of the Atmospheric Sciences* 54.1 (1997), pp. 72–86.

[296] G. Chen and B. Wang. "Diversity of the boreal summer intraseasonal oscillation". In: *Journal of Geophysical Research: Atmospheres* 126.8 (2021).

[297] A. Kulkarni, R. Kripalani, S. Sabade, and M. Rajeevan. "Role of intra-seasonal oscillations in modulating Indian summer monsoon rainfall". In: *Climate dynamics* 36.5 (2011), pp. 1005–1021.

[298] J. Li, J. Mao, and G. Wu. "A case study of the impact of boreal summer intraseasonal oscillations on Yangtze rainfall". In: *Climate Dynamics* 44.9 (2015), pp. 2683–2702.

[299] Y. Xia, Q. Huang, S. Yao, and T. Sun. "Multiscale Analysis of Persistent Heavy Rainfall in the Meiyu Period over the Middle and Lower Reaches of the Yangtze River". In: *Frontiers in Earth Science* 9 (2021), p. 668.

[300] I.-S. Kang, C.-H. Ho, Y.-K. Lim, and K. M. Lau. "Principal modes of climatological seasonal and intraseasonal variations of the Asian summer monsoon". In: *Monthly weather review* 127.3 (1999), pp. 322–340.

[301] H. Annamalai and J. M. Slingo. "Active/break cycles: Diagnosis of the intraseasonal variability of the Asian summer monsoon". In: *Climate Dynamics* 18.1 (2001), pp. 85–102.

[302] Q. Ding and B. Wang. "Predicting extreme phases of the Indian summer monsoon". In: *Journal of Climate* 22.2 (2009), pp. 346–363.

[303] W. K.-M. Lau and D. E. Waliser. *Intraseasonal variability in the atmosphere-ocean climate system.* Springer Science & Business Media, 2011.

[304] V. Krishnamurthy and J. Shukla. "Intraseasonal and seasonally persisting patterns of Indian monsoon rainfall". In: *Journal of climate* 20.1 (2007), pp. 3–20.

[305] V. Krishnamurthy and J. Shukla. "Seasonal persistence and propagation of intraseasonal patterns over the Indian monsoon region". In: *Climate Dynamics* 30.4 (2008), pp. 353–369.

[306] W. Qian, H.-S Kang, and D.-K Lee. "Distribution of seasonal rainfall in the East Asian monsoon region". In: *Theoretical and Applied Climatology* 73 (2002), pp. 151–168.

[307] M. Rathinasamy, A. Agarwal, B. Sivakumar, N. Marwan, and J. Kurths. "Wavelet analysis of precipitation extremes over India and teleconnections to climate indices". In: *Stochastic Environmental Research and Risk Assessment* 33.11 (2019), pp. 2053–2069.

[308] NOAA. *Climate Prediction Center – Cold and warm episodes by season.* `https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php`. Last accessed on 10/31/2022. 2022.

[309] T. N. Palmer, Č Branković, P. Viterbo, and M. J. Miller. "Modeling Interannual Variations of Summer Monsoons". In: *Journal of Climate* 5.5 (1992), pp. 399–417.

[310] B. Wang, R. Wu, and X. Fu. "Pacific–East Asian Teleconnection: How Does ENSO Affect East Asian Climate?" In: *Journal of Climate* 13.9 (2000), pp. 1517–1536.

[311] A. M. Herzberg. "The design of experiments for correlated error structures: Layout and robustness". In: *Canadian Journal of Statistics* 10.2 (1982), pp. 133–138.

[312] G. J. Boer. "Systematic and Random Error in an Extended-Range Forecasting Experiment". In: *Monthly Weather Review* 121.1 (1993), pp. 173–188.

Bibliography

[313] J. Derber and F. Bouttier. "A reformulation of the background error covariance in the ECMWF global data assimilation system". In: *Tellus A: Dynamic Meteorology and Oceanography* 51.2 (1999), pp. 195–221.

[314] D. F. Parrish and J. C. Derber. "The National Meteorological Center's Spectral Statistical-Interpolation Analysis System". In: *Monthly Weather Review* 120.8 (1992), pp. 1747–1763.

[315] W.-S. Wu, R. J. Purser, and D. F. Parrish. "Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances". In: *Monthly Weather Review* 130.12 (2002), pp. 2905–2916.

[316] D. T. Kleist, D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord. "Introduction of the GSI into the NCEP Global Data Assimilation System". In: *Weather and Forecasting* 24.6 (2009), pp. 1691–1705.

[317] M. B. Pereira and L. Berre. "The Use of an Ensemble Approach to Study the Background Error Covariances in a Global NWP Model". In: *Monthly Weather Review* 134.9 (2006), pp. 2466–2489.

[318] R. Daley and E. Barker. "NAVDAS: Formulation and Diagnostics". In: *Monthly Weather Review* 129.4 (2001), pp. 869–883.

[319] Ingleby, N. Bruce.al structure of forecast errors and its representation in The Met. Office Global 3-D Variational Data Assimilation Scheme. In: *Quarterly Journal of the Royal Meteorological Society* 127.571 (2001), pp. 209–231.

[320] M. Gelbrecht, N. Boers, and J. Kurths. "A complex network representation of wind flows". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 035808.

[321] K. Steinhaeuser and A. A. Tsonis. "A climate model intercomparison at the dynamics level". In: *Climate Dynamics* 42.5 (2014), pp. 1665–1670.

[322] J. H. Feldhoff, S. Lange, J. Volkholz, J. F. Donges, J. Kurths, and F.-W. Gerstengarbe. "Complex networks for climate model evaluation with application to statistical versus dynamical modeling of South American climate". In: *Climate Dynamics* 44.5 (2015), pp. 1567–1581.

[323] W. Gregory, J. Stroeve, and M. Tsamados. "Network connectivity between the winter Arctic Oscillation and summer sea ice in CMIP6 models and observations". In: *The Cryosphere* 16.5 (2022), pp. 1653–1673.

[324] L. Magnusson. "Diagnostic methods for understanding the origin of forecast errors". In: *Quarterly Journal of the Royal Meteorological Society* 143.706 (2017), pp. 2129–2142.

[325] C. Qian, F. Zhang, B. W. Green, J. Zhang, and X. Zhou. "Probabilistic Evaluation of the Dynamics and Prediction of Supertyphoon Megi (2010)". In: *Weather and Forecasting* 28.6 (2013), pp. 1562–1577.

[326] L. Magnusson, J. D. Doyle, W. A. Komaromi, R. D. Torn, C. K. Tang, J. C. L. Chan, M. Yamaguchi, and F. Zhang. "Advances in understanding difficult cases of tropical cyclone track forecasts". In: *Tropical Cyclone Research and Review* 8.3 (2019). Special Issue for the 9th WMO International Workshop on Tropical Cyclones, pp. 109–122.

[327] L. Gao, P. Ren, and J. Zheng. "Medium-Range Predictability of Boreal Summer Western North Pacific Subtropical High and Its ENSO Modulation". In: *Frontiers in Earth Science* 10 (2022).

[328] B. Wang, B. Xiang, and J.-Yi Lee. "Subtropical High predictability establishes a promising way for monsoon and tropical storm predictions". In: *Proceedings of the National Academy of Sciences* 110.8 (2013), pp. 2718–2722.

[329] N. Žagar, M. Blaauw, B. Jesenko, and L. Magnusson. "Diagnosing model performance in the tropics". In: *ECMWF Newsletter No. 147 – Spring 2016* (2016), pp. 26–33.

[330] N. Žagar, A. Kasahara, K. Terasaki, J. Tribbia, and H. Tanaka. "Normal-mode function representation of global 3-D data sets: open-access software for the atmospheric research community". In: *Geoscientific Model Development* 8.4 (2015), pp. 1169–1195.

[331] J. C. K. Lee and X.-Y. Huang. "Background error statistics in the Tropics: Structures and impact in a convective-scale numerical weather prediction system". In: *Quarterly Journal of the Royal Meteorological Society* 146.730 (2020), pp. 2154–2173.

[332] K. Kikuchi, G. N. Kiladis, J. Dias, and T. Nasuno. "Convectively coupled equatorial waves within the MJO during CINDY/DYNAMO: slow Kelvin waves as building blocks". In: *Climate Dynamics* 50.11 (2018), pp. 4211–4230.

[333] B. Raphaldini, A. S. W. Teruya, C. F. M. Raupp, P. L. Silva-Dias, and D. Y. Takahashi. "Information flow between MJO-related waves: a network approach on the wave space". In: *The European Physical Journal Special Topics* 230.14 (2021), pp. 3009–3017.

[334] G. Li and S.-P. Xie. "Tropical Biases in CMIP5 Multimodel Ensemble: The Excessive Equatorial Pacific Cold Tongue and Double ITCZ Problems". In: *Journal of Climate* 27.4 (2014), pp. 1765–1780.

[335] Y. Y. Toh, A. G. Turner, S. J. Johnson, and C. E. Holloway. "Maritime Continent seasonal climate biases in AMIP experiments of the CMIP5 multimodel ensemble". In: *Climate Dynamics* 50.3 (2018), pp. 777–800.

[336] P.-H. Chung, C.-H. Sui, and T. Li. "Interannual relationships between the tropical sea surface temperature and summertime subtropical anticyclone over the western North Pacific". In: *Journal of Geophysical Research: Atmospheres* 116.D13 (2011).

[337] N. Molkenthin, H. Kutza, L. Tupikina, N. Marwan, J. F. Donges, U. Feudel, J. Kurths, and R. V. Donner. "Edge anisotropy and the geometric perspective on flow networks". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27.3 (2017), p. 035802.

[338] A. Rheinwalt, N. Boers, N. Marwan, J. Kurths, P. Hoffmann, F.-W. Gerstengarbe, and P. Werner. "Non-linear time series analysis of precipitation events using regional climate networks for Germany". In: *Climate Dynamics* 46.3 (2016), pp. 1065–1074.

[339] L. Lü, L. Pan, T. Zhou, Y.-C. Zhang, and H. E. Stanley. "Toward link predictability of complex networks". In: *Proceedings of the National Academy of Sciences* 112.8 (2015), pp. 2325–2330.

[340] G. Garcıéa-Pérez, R. Aliakbarisani, A. Ghasemi, and M. Á. Serrano. "Precision as a measure of predictability of missing links in real networks". In: *Phys. Rev. E* 101.5 (2020), p. 052318.

[341] M. Haas, B. Goswami, and U. Luxburg. "Pitfalls of Climate Network Construction: A Statistical Perspective". In: *arXiv preprint* (2022).

[342] C. T. Butts. "Revisiting the Foundations of Network Analysis". In: *Science* 325.5939 (2009), pp. 414–416.

[343] L. Peel, T. P. Peixoto, and M. De Domenico. "Statistical inference links data and theory in network science". In: *Nature Communications* 13.1 (2022), p. 6794.

[344] N. Molkenthin, K. Rehfeld, N. Marwan, and J. Kurths. "Networks from Flows - From Dynamics to Topology". In: *Scientific Reports* 4.1 (2014), p. 4119.

[345] N. Ekhtiari, A. Agarwal, N. Marwan, and R. V. Donner. "Disentangling the multi-scale effects of sea-surface temperatures on global precipitation: A coupled networks approach". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29.6 (2019), p. 063116.

Declaration

I declare that I have completed the thesis independently using only the aids and tools specified. I have not applied for a doctor's degree in Physics elsewhere and do not hold a corresponding doctor's degree. I have taken due note of the Faculty of Mathematics and Natural Sciences PhD Regulations, published in the Official Gazette of Humboldt-Universität zu Berlin no. 42/2018 on 11/07/2018.

Berlin, on 27.04.2023   _____

Shraddha Gupta