



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21d/2023/08.18.12.38-TDI

**CHARACTERIZATION OF PRECIPITATION
COMMUNITIES IN SOUTH AMERICA USING
COMPLEX NETWORKS**

César Arturo Sánchez Peña

Doctorate Thesis of the Graduate
Course in Applied Computing,
guided by Drs. Alan James Peixoto
Calheiros, and Elbert Einstein
Nehrer Macau, approved in August
16, 2023.

URL of the original document:

<<http://urlib.net/8JMKD3MGP3W34T/49LG54E>>

INPE
São José dos Campos
2023

PUBLISHED BY:

Instituto Nacional de Pesquisas Espaciais - INPE
Coordenação de Ensino, Pesquisa e Extensão (COEPE)
Divisão de Biblioteca (DIBIB)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

**BOARD OF PUBLISHING AND PRESERVATION OF INPE
INTELLECTUAL PRODUCTION - CEPPII (PORTARIA Nº
176/2018/SEI-INPE):****Chairperson:**

Dra. Marley Cavalcante de Lima Moscati - Coordenação-Geral de Ciências da Terra
(CGCT)

Members:

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação (CPG)
Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia, Tecnologia e
Ciência Espaciais (CGCE)
Dr. Rafael Duarte Coelho dos Santos - Coordenação-Geral de Infraestrutura e
Pesquisas Aplicadas (CGIP)
Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)

DIGITAL LIBRARY:

Dr. Gerald Jean Francis Banon
Clayton Martins Pereira - Divisão de Biblioteca (DIBIB)

DOCUMENT REVIEW:

Simone Angélica Del Ducca Barbedo - Divisão de Biblioteca (DIBIB)
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)

ELECTRONIC EDITING:

Ivone Martins - Divisão de Biblioteca (DIBIB)
André Luis Dias Fernandes - Divisão de Biblioteca (DIBIB)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21d/2023/08.18.12.38-TDI

CHARACTERIZATION OF PRECIPITATION COMMUNITIES IN SOUTH AMERICA USING COMPLEX NETWORKS

César Arturo Sánchez Peña

Doctorate Thesis of the Graduate
Course in Applied Computing,
guided by Drs. Alan James Peixoto
Calheiros, and Elbert Einstein
Nehrer Macau, approved in August
16, 2023.

URL of the original document:

<<http://urlib.net/8JMKD3MGP3W34T/49LG54E>>

INPE
São José dos Campos
2023

Cataloging in Publication Data

Sánchez Peña, César Arturo.

Sa55c Characterization of precipitation communities in South America using complex networks / César Arturo Sánchez Peña. – São José dos Campos : INPE, 2023.

xxiv + 149 p. ; (sid.inpe.br/mtc-m21d/2023/08.18.12.38-TDI)

Thesis (Doctorate in Applied Computing) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2023.

Guiding : Drs. Alan James Peixoto Calheiros, and Elbert Einstein Nehrer Macau.

1. Extreme rainfall. 2. Communities. 3. Complex networks. 4. South America. 5. Amazon basin. I.Title.

CDU 004.7:551.577



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA
E INOVAÇÃO



INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

DEFESA FINAL DE TESE CÉSAR ARTURO SÁNCHEZ PEÑA BANCA Nº 205/2023, REGISTRO 351986/2019

No dia 16 de agosto de 2023, por teleconferência, o(a) aluno(a) mencionado(a) acima defendeu seu trabalho final (apresentação oral seguida de arguição) perante uma Banca Examinadora, cujos membros estão listados abaixo. O(A) aluno(a) foi APROVADO(A) pela Banca Examinadora, por unanimidade, em cumprimento ao requisito exigido para obtenção do Título de Doutor em Computação Aplicada, com a exigência de que o trabalho final a ser publicado deverá incorporar as correções sugeridas pela Banca Examinadora, com revisão pelo(s) orientador(es).

Título: “Characterization of precipitation communities in South America using Complex Networks”.

Membros da Banca:

Dr. Reinaldo Roberto Rosa – Presidente – INPE
Dr. Alan James Peixoto Calheiros – Orientador – INPE
Dr. Elbert Einstein Nehrer Macau – Orientador – Unifesp
Dr. Marcos Gonçalves Quiles – Membro Interno – INPE
Dr. Henrique de Melo Jorge Barbosa – Membro externo – UMBC
Dr. Didier Augusto Vega Oliveros – Membro externo – Unicamp



Documento assinado eletronicamente por **Marcos Gonçalves Quiles (E), Usuário Externo**, em 21/08/2023, às 08:53 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Henrique de melo jorge barbosa (E), Usuário Externo**, em 21/08/2023, às 15:31 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Elbert Einstein nehrer macau (E), Usuário Externo**, em 21/08/2023, às 23:07 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Reinaldo Roberto Rosa, Pesquisador Titular**, em 22/08/2023, às 10:35 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Didier augusto vega oliveros (E), Usuário Externo**, em 22/08/2023, às 23:44 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alan James Peixoto Calheiros, Tecnologista**, em 25/08/2023, às 15:27 (horário oficial de Brasília), com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site <https://sei.mcti.gov.br/verifica.html>, informando o código verificador **11289941** e o código CRC **A8CBBE93**.

Referência: Processo nº 01340.006745/2023-99

SEI nº 11289941

To my grandmother

ACKNOWLEDGEMENTS

I express my heartfelt gratitude to my family, with a special mention to my parents, Lucy, and Raul. Their unwavering love, support, and encouragement have been instrumental in my journey. I am truly grateful for their guidance and support. I also want to express special thanks to my dear Uncle Miguel, who has always been an excellent role model, showing me the importance of dedication and effort.

I sincerely offer my deepest gratitude to my advisor, Alan Calheiros, for his invaluable guidance, mentorship, and unwavering support throughout the entire duration of my research. His expertise, patience, and dedication have played a crucial role in shaping my research. I am sincerely grateful for his constant encouragement, insightful feedback, and belief in mine. I am considerate fortunate to have had him as my advisor, and I am thankful for the valuable lessons I have learned under his supervision.

I also extend my heartfelt gratitude to Dr. Elbert Macau for his supervision and guidance throughout the development of this research. I am thankful for his dedicated time, valuable comments, and meticulous corrections, which have greatly contributed to the quality and refinement of this work.

I want to convey my appreciation to my dear friend, Isela, for her unwavering friendship and constant companionship throughout the ups and downs of this phase of my life. I am deeply grateful to Adriano and Helvecio for their invaluable company and support throughout the development of this research.

And to all those who are not specifically mentioned here, please know that I hold in high regard all the individuals who have been a part of my academic journey. Your presence and contributions have been valued and appreciated.

Finally, I would like to thank the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES) - Brasil for funding this research via the grant 88882.444490/2019-01.

ABSTRACT

The dynamics of precipitation systems are essential for understanding their origins, development, and potential impacts. This knowledge is crucial for predicting and mitigating the adverse effects of extreme rainfall events. These events pose significant threats to the safety and well-being of populations, as well as causing material damage and significant socio-economic impacts. Understanding precipitation is of utmost importance in increasing resilience, particularly in developing regions like South America. To achieve this, we utilize a complex network approach, which allows us to extract key information about the behavior of the precipitation system and identify and characterize highly connected regions, known as communities. By examining the details of these communities, we gain valuable insights into the dynamics and structure of precipitation, thereby enhancing our understanding and prediction of such events. The analysis starts with the network of extreme events in South America during the summer season. The objective is to extract information about the dynamics of this network and identify communities within it. Subsequently, a characterization of these communities is provided. To investigate variations at smaller scales, another network of extreme events was created specifically for the Amazon basin. In addition to these two extreme rainfall networks, three additional rainfall networks were created for the Amazon basin to compare the basin's response to both rainfall and extreme rainfall events. The outcomes revealed that the networks of extreme events in South America and the Amazon basin exhibit a similar spatial distribution of communities. These communities are primarily influenced by the transport of low-level moisture from the eastern Amazon towards the Andes, ultimately exiting the Amazon basin towards the extratropics. The central region of the Amazon basin acts as a source, while the southeast region acts as a sink. In the case of the rainfall networks in the Amazon basin, three important regions were identified: northwest, northeast, and southeast. The first two regions are related to the entry of moisture into the basin, while the third is related to the outflow. Finally, all networks highlight the significant role of the eastern Andes as a topographic barrier. Typically, this region consists of one or two communities, which serve as highly connected corridors for low-level moisture flow. The results showed that complex networks were able to define regions with similar characteristics in South America.

Palavras-chave: Extreme rainfall. Communities. Complex networks. South America. Amazon Basin.

CARACTERIZAÇÃO DE COMUNIDADES DE PRECIPITAÇÃO NA AMÉRICA DO SUL USANDO REDES COMPLEXAS

RESUMO

A dinâmica dos sistemas de precipitação é essencial para compreender suas origens, desenvolvimento e impactos potenciais. Esse conhecimento é crucial para prever e mitigar os efeitos adversos de eventos extremos de precipitação. Estes eventos representam ameaças significativas à segurança e bem-estar das populações, além de causarem danos materiais e impactos socioeconômicos significativos. Compreender a precipitação é de extrema importância para aumentar a resiliência, especialmente em regiões em desenvolvimento como a América do Sul. Para alcançar isso, utilizamos uma abordagem de rede complexa, que nos permite extrair informações-chave sobre o comportamento do sistema de precipitação e identificar e caracterizar regiões altamente conectadas, conhecidas como comunidades. Ao examinar os detalhes dessas comunidades, obtemos insights valiosos sobre a dinâmica e estrutura da precipitação, melhorando assim nossa compreensão e previsão de tais eventos. A análise começa com a rede de eventos extremos na América do Sul durante a estação de verão. O objetivo foi extrair informações sobre a dinâmica dessa rede e identificar as comunidades presentes, e posteriormente fornecer uma caracterização dessas comunidades. Para investigar variações em escalas menores, outra rede de eventos extremos foi criada especificamente para a bacia amazônica. Além dessas duas redes de precipitação extrema, foram criadas mais três redes de precipitação para a bacia amazônica, a fim de comparar a resposta da bacia tanto à precipitação quanto aos eventos extremos de precipitação. Os resultados revelaram que as redes de eventos extremos na América do Sul e na bacia amazônica apresentam uma distribuição espacial similar de comunidades. Essas comunidades são influenciadas principalmente pelo transporte de umidade em baixos níveis da Amazônia oriental em direção aos Andes, e posteriormente saindo da bacia amazônica em direção aos extratropicais. A região central da bacia amazônica atua como uma fonte, enquanto a região sudeste atua como um sumidouro. No caso das redes de precipitação na bacia amazônica, foram identificadas três regiões importantes: noroeste, nordeste e sudeste. As duas primeiras regiões estão relacionadas à entrada de umidade na bacia, enquanto a terceira está relacionada à saída. Por fim, todas as redes destacam o papel significativo dos Andes orientais como barreira topográfica. Geralmente, essa região consiste em uma ou duas comunidades, que atuam como corredores altamente conectados para o fluxo de umidade em baixos níveis. De um modo geral os resultados mostraram que as redes complexas foram hábeis para definir as regiões com características similares na América do Sul.

Keywords: Chuva extrema. Comunidades. Redes complexas. América do Sul. Bacia Amazônica.

LIST OF FIGURES

	<u>Page</u>
3.1 Advection mechanisms in South America.	8
3.2 Schematic representation of the stages that represent the life cycle of a storm.	12
3.3 Hourly frequency of deep convection cores.	15
3.4 Average lightning rates present over South America.	16
3.5 Regions of geographic interest in South America.	18
3.6 Brief review of Complex Networks in meteorology applications.	20
3.7 AWC for weather network using monthly geopotential height anomaly at 500 hPa.	21
3.8 Strong and long positive links greater than 5000 <i>km</i>	23
3.9 Topographic map and representation of the mechanisms of the South American monsoon system with the results of network divergence.	25
3.10 Schematic example of a network with three communities.	28
3.11 Community structures in a rain network in India.	30
3.12 Climate network communities of geopotential height (<i>hgt</i>) em 500 <i>hpa</i>	31
3.13 Change in community structures before and during the Australian summer monsoon.	32
4.1 Schematic of the scan geometries of the three primary TRMM rain sensors: TMI, PR e VIRS.	36
4.2 TRMM data processing scheme.	37
4.3 Representation of the satellite constellation of the GPM mission.	39
4.4 Scheme of DPR and GMI instrument scanning patterns onboard the GPM Core.	40
4.5 Vertical diagram of three types of PF.	44
5.1 Diagram of the two groups of networks that were used.	48
5.2 Average daily rainfall and 90th percentile (mm/day) over South America during the austral summer.	50
5.3 Schematic representation of the creation of a climate network.	52
5.4 Example of the process to define extreme events using the 90th percentile of a daily rainfall series for a specific point on the grid.	54
5.5 Schematic representation of the AND index for some nodes in a regular grid.	60
5.6 Clustering coefficient for the central node ($CC_i = 1/3$).	61

5.7	Schematic representation of the AWC index for some nodes in a regular grid.	62
5.8	Schematic representation of the GD index for some nodes in a regular grid.	63
5.9	Schematic representation of the BC index for some nodes in a regular grid.	64
5.10	Schematic representation of the DIL index for some nodes in a regular grid.	66
5.11	Schematic representation of the communities in a regular grid.	67
6.1	Density variation for different MLV values for South America network.	74
6.2	Scatter plot of Geographic Distance (km).	76
6.3	Communities in networks considering continent and ocean rainfall.	77
6.4	Communities in the extreme daily rainfall network.	79
6.5	Overview of important features of the low-level atmospheric circulation during the SAMS.	80
6.6	Results of the degrees for each community.	83
6.7	Node connections considering only three communities.	85
6.8	Normalized results for Community SA1.	87
6.9	Normalized results for Community SA2.	88
6.10	Normalized results for Community SA3.	90
6.11	Number of rainiest and deepest convection events for the two approaches.	94
6.12	Contoured frequency by Altitude Diagram (CFAD).	96
6.13	Contoured frequency by Altitude Diagram (CFAD).	97
6.14	The cumulative distribution function (CDF).	98
6.15	Normalized index for AMAZON network.	100
6.16	Normalized index for AMAZON network.	101
6.17	Communities in extreme rainfall AMAZON network.	101
6.18	Communities edges for the AMAZON network.	102
6.19	Communities in extreme rainfall networks.	103
6.20	Variations in the density and clustering for AMAZON Networks.	105
6.21	Relationship between density and clustering of AMAZON networks.	106
6.22	Degree distribution of the networks.	107
6.23	Normalized index for AMAZON networks.	108
6.24	Degree and important lines.	110
6.25	Normalized Betweenness centrality.	112
6.26	Betweenness centrality boxplot for nodes with/non-rivers.	114
6.27	Communities identified in each network.	115
6.28	Difference of normalized indices between Spearman and Kendall.	116
6.29	Scatter plot for the differences between Spearman and Kendall networks.	117
7.1	Possible regions to the location of new rain gauges.	122

LIST OF TABLES

	<u>Page</u>
4.1 Spatial resolution of TMI frequency bands.	38
4.2 New criteria used for defining Precipitation Features (LIU, 2007).	43
5.1 Characteristics of the networks used in this study.	48
6.1 The calculated values for different indices are presented for the South America network and each individual community.	81
6.2 Statistical information for extreme events in each community, with rain- fall values given in millimeters per day (mm/day).	92
A.1 List of links to data used.	149
A.2 Programming language and main packages used.	149

LIST OF ABBREVIATIONS

AMS	–	American Meteorology Society
AND	–	Average Neighbor Degree
ASM	–	Australian Summer Monsoon
AWC	–	Area weighted connectivity
BC	–	Betweenness Centrality
CC	–	Clustering Centrality
MCC	–	Mesoscale Convective Complex
MCS	–	Mesoscale Convective system
CDF	–	Cumulative Distribution Function
CERES	–	Clouds and Earth’s Radiant Energy System
CFAD	–	Contoured Frequency by Altitude Diagram
CN	–	Complex Networks
DC	–	Degree Centrality
DIL	–	Degree and the Importance of Lines
DPR	–	Dual-frequency Precipitation Radar
ECA	–	Eastern Central Andes
EOF	–	Empirical orthogonal functions
GD	–	Geographical Distance
GMI	–	Microwave Imager
GPM	–	Global Precipitation Measurement Mission
IMERG	–	Integrated Multi-satellitE Retrievals for GPM
IPCC	–	Intergovernmental Panel on Climate Change
IR	–	Infrared
ISM	–	Indian Summer Monsoon
JAXA	–	Japan Aerospace Exploration Agency
JBNAS	–	South American Low Level Jet
LLJ	–	Low level jet
LIS	–	Lightning Imaging Sensor
MLV	–	minimum link value
NASA	–	National Aeronautics and Space Administration
NAO	–	North Atlantic Oscillation Index
PDF	–	Probability Density Function
PF	–	Precipitation Features
PNA	–	Pacific North America Index
PSA	–	Pacific South America Index
PR	–	Precipitation radar
SESA	–	Southeastern South America
SOM	–	Self Organizing Map
TRMM	–	Tropical Rainfall Measuring Mission
TMI	–	Microwave Imager
VIRS	–	Visible and infrared scanners
WMO	–	World Meteorological Organization

LIST OF SYMBOLS

r	–	Pearson correlation coefficient
r_s	–	Spearman's rank correlation
τ_k	–	Kendall's rank correlation
k_i	–	Node degree
N	–	Total number of nodes
$\langle k \rangle$	–	Average degree of the network
ρ	–	Link density
$N(i)$	–	Number of neighbors of node i
E_i	–	Number of edges existing only between neighbor
λ_j	–	Latitude of node j
$dist(i, j)$	–	Geographical distance between points i and j
σ_{kl}	–	Number of shortest paths from node k to l
$\sigma_{kl}(i)$	–	number of shortest paths from k to l that pass through node i
Q	–	Modularity

CONTENTS

	<u>Page</u>
1 INTRODUCTION	1
2 OBJECTIVES	5
3 THEORETICAL BACKGROUND	7
3.1 South America summer precipitation	7
3.2 Extreme rainfall events	9
3.2.1 Convection	10
3.3 Complex networks in meteorology	19
3.4 Communities identification	27
4 DATA	35
4.1 Tropical Rainfall Measuring Mission (TRMM)	35
4.1.1 TRMM Microwave Imager (TMI)	36
4.1.2 Precipitation radar (PR)	38
4.1.3 Visible and infrared scanners (VIRS)	38
4.2 Global Precipitation Measurement Mission (GPM)	39
4.2.1 GPM Microwave Imager (GMI)	41
4.2.2 Dual-frequency Precipitation Radar (DPR)	41
4.3 Integrated Multi-satellitE Retrievals for GPM (IMERG)	42
4.4 Precipitation Features (PF)	42
4.4.1 Vertical Reflectivity Profile	45
5 METHODS	47
5.1 Study region	47
5.1.1 South America network	49
5.1.2 Amazon network	50
5.2 Climate networks	51
5.2.1 Extreme rainfall networks	53
5.2.1.1 Define extreme rainfall events	53
5.2.1.2 Event Synchronization	53
5.2.2 Rainfall networks	55
5.2.2.1 Pearson correlation coefficient	55

5.2.2.2	Spearman’s rank correlation	56
5.2.2.3	Kendall’s rank correlation	56
5.3	Network analysis	57
5.3.1	Adjacency matrix	57
5.3.2	Degree, distribution, and density	58
5.3.3	Degree Centrality (DC)	59
5.3.4	Average Neighbor Degree (AND)	59
5.3.5	Cluster Coefficient (CC)	60
5.3.6	Area Weighted Connectivity (AWC)	62
5.3.7	Geographical Distance (GD)	63
5.3.8	Betweenness Centrality (BC)	64
5.3.9	Degree value and the Importance of Lines (DIL)	65
5.4	Community identification	66
5.5	Spatial embedding	68
5.6	Community characterization	68
5.6.1	Network topology analysis	68
5.6.2	Distribution of extreme events	69
5.6.3	Deep convective precipitation	70
6	RESULTS	73
6.1	SOUTH AMERICA network	73
6.1.1	Geographical distance	75
6.1.2	Removing the ocean	77
6.1.3	Communities identification	78
6.1.4	Network analysis	79
6.1.5	Inter-communities	81
6.1.6	Intra-communities	86
6.1.7	Communities characterization	91
6.1.7.1	Extreme events distribution	91
6.1.7.2	Deep convective precipitation	93
6.2	AMAZON networks	99
6.2.1	Extreme rainfall network	99
6.2.2	Total rainfall networks	104
6.2.2.1	Networks analysis	106
6.2.2.2	Spearman and Kendall comparison	116
7	CONCLUSIONS	119
7.1	Extreme rainfall networks	119

7.2	Amazon Basin rainfall networks	120
7.3	Possible applications	122
7.4	Future works	123
REFERENCES		125
APPENDIX A		149
A.1	Data links	149
A.2	Software	149

1 INTRODUCTION

Rainfall, as a crucial component in the hydrological cycle, not only impacts various aspects of society but also influences the occurrence of extreme rainfall events. Extreme rainfall events are associated with risks such as flooding and landslides (BOERS et al., 2013; BOERS et al., 2015). To study and characterize rainfall and its extreme events, we can use various measurements, and variables from different sources. These include spatial dimensions of convective systems, cloud top temperature, radar reflectivity, microphysical variables of clouds, and surrounding meteorological conditions associated with convection (BOERS et al., 2014b; MUÑOZ et al., 2015). Selecting the appropriate variables is crucial to obtaining comprehensive information that cannot be captured by a single variable. Hence, combining multiple variables is often necessary, requiring careful consideration at the initial stages of the study. Understanding the spatial and temporal distribution of rainfall becomes even more critical when it comes to assessing the potential for extreme events and implementing effective monitoring and preparedness measures. Studies analyzing precipitation systems, like the present one, contribute to identifying areas with similar rainfall characteristics, atmospheric interactions, and topographic variations, which in turn aids in the development of strategies to mitigate the impacts of extreme events (KUMAR et al., 2019b). By comprehending the patterns and dynamics of rainfall, we can enhance our ability to manage water resources, reduce disaster risks, improve forecast models, and adapt to the challenges posed by climate change.

Understanding the dynamics of precipitation systems is essential for comprehending their origins, development, and potential impacts. This knowledge is crucial for predicting and mitigating the adverse effects of extreme rainfall events. Extreme events, serving as natural responses to atmospheric disturbances, play a significant role in regulating Earth's climate through various physical and thermodynamic mechanisms (LIU et al., 2020; YIN et al., 2023). While these mechanisms act as energy regulators, they also give rise to severe weather phenomena, including storms characterized by excessive rainfall, lightning, hail, and strong winds (GRIMM, 2011; BOERS et al., 2016).

A complex network is a powerful mathematical and computational tool used to study and analyze complex systems found in various disciplines (FREITAS et al., 2019; MATA, 2020). It is a representation of interconnected elements, often referred to as nodes or vertices, with links or edges that connect these nodes.

The nodes can represent individual entities, while the links describe the relationships, interactions, or connections between them. In meteorology, complex networks have yielded significant results, particularly in understanding atmospheric teleconnections (ZHOU et al., 2015; TSONIS et al., 2008; BOERS et al., 2019) and monsoon systems (MALIK et al., 2012; RODRIGUES et al., 2022; GUPTA et al., 2023; CHEUNG; OZTURK, 2020). By examining the system’s behavior and the transmission of information in response to disturbances, complex networks provide valuable insights associated with its connections (CONTICELLO et al., 2018; GELBRECHT et al., 2021). Additionally, the identification of groups that exhibit similar behaviors and possess strong internal links (SAHA; MITRA, 2019; FORTUNATO, 2010; TSONIS; SWANSON, 2011). There are also different studies of complex networks applied to obtain a better understanding of the precipitating system (BOERS et al., 2013; BOERS et al., 2019; NEWELL et al., 2022).

It is well known that the South American Monsoon System (SAMS) is the main modulator of rainfall in South America (VERA et al., 2006a). SAMS has the potential to cause significant impacts on agriculture, water resources, and human health (BERBERY; BARROS, 2002; MARTÍN-GÓMEZ et al., 2016). To enhance knowledge about these events, it is essential to adopt innovative approaches. In this context, the identification of regions that exhibit similar characteristics called communities emerges as a perspective to analyze the interactions among the mechanisms present. In this way, complex networks can reveal emergent structures and provide valuable insights into the dynamics of each specific region. By identifying these regions, it is possible to understand how different elements are grouped and interact with each other, allowing for a deeper understanding of interconnections and the propagation of precipitation systems. This approach provides a perspective that can help determine areas of greater importance or vulnerability.

The identification of communities can contribute to future research aimed at understanding interactions within the climate system, as these communities could potentially serve as climate indices (STEINHAEUSER et al., 2011). Rainfall communities provide valuable insights into the behavior and organization of precipitation patterns, aiding in the prediction and mitigation of extreme rainfall events that can have severe consequences for human safety, infrastructure, agriculture, and water resources. Different studies use complex networks to investigate and uncover patterns within the precipitation system, encompassing both regular rainfall and extreme rainfall events (BOERS et al., 2014a; FELDHOFF et al., 2015; CIEMER et al., 2018).

In these works the primary objective is to analyze the behavior of these networks and their characteristics. However, no studies focused on identifying communities on a precipitation network in South America. So to ensure the consistency and reliability of the findings, a comprehensive analysis is conducted, encompassing bibliographic, spatial, and dynamic perspectives.

South America possesses a variety of significant atmospheric mechanisms, complemented by its vast geographical diversity. Among these, the Andes mountain range plays a crucial role in influencing the continent's circulation (INSEL et al., 2010), while the *Sierras de Córdoba* region stands out for having some of the deepest convective cores globally (MULHOLLAND et al., 2018). The Brazilian semiarid region is considered the most populated semiarid region in the world (CUNHA et al., 2015; MENEZES et al., 2021). The Pisco-Ica desert is known for its characteristic dust storms locally referred to as *Paracas* (BRICEÑO-ZULUAGA et al., 2017). Meanwhile, the Atacama desert claims fame as the oldest continuously dry non-polar desert on Earth (BULL et al., 2018).

However, one of the most vital regions in South America is the Amazon basin, acclaimed as the largest rainforest and water system in the world (ALMAZROUI et al., 2021). The Amazon basin plays a pivotal role as a moisture source for the continent, impacting regional and global climate, and demonstrating a remarkable capacity for carbon dioxide absorption (FISCH et al., 1998; HAYLOCK et al., 2006; NOBRE et al., 2009). By identifying rainfall communities in the Amazon basin, we can better understand its role as a moisture source. This information can provide us with insights into localized climate drivers, study orographic rainfall processes, and explore the influence of river systems on precipitation patterns. Such knowledge is invaluable for improving weather forecasts, climate models, and sustainable development initiatives in the Amazon region. Identifying communities within the Amazon basin not only enables us to compare and contrast them with the communities in South America but also provides valuable information about the specific behavior and dynamics of precipitation systems in this important region. Furthermore, this comparison between the Amazon basin and South America allows us to verify the consistency and applicability of the complex network approach across different regions with specific weather and geographical characteristics. Understanding how the communities in the Amazon basin differ from those in the South American network provides insights into the factors that shape rainfall and extreme rainfall events. By doing so, we can examine the robustness of the technique and gain insights into the unique characteristics of each region.

Ultimately, this research contributes to a broader analysis of precipitation events, encompassing both regular rainfall and extreme events. The results of this study provide valuable information on the system's behavior, thus the identification of distinct communities. Therefore, through the identification and analysis of communities, a new perspective can be offered to understand the dynamics of precipitation events in South America, including the Amazon basin. This approach can provide valuable insights for various applications that can be applied in various fields, such as meteorological modeling, hydrology, weather/climate forecasting, or risk management.

The structure of this thesis consists of seven chapters and an appendix, with each serving a specific purpose. After this chapter, the main objectives of this study are presented in Chapter 2. Chapter 3 provides a concise theoretical framework on precipitation systems, including a literature review on the application of complex networks in meteorological applications and a brief review of communities. Chapter 4 focuses on describing the data used in this research, along with their relevant characteristics.

In Chapter 5, the study area and the methodology employed to achieve the thesis objectives are detailed. This chapter encompasses the creation of networks, selection of indices for analysis, identification of communities, and characterization of these communities. Chapter 6 presents the results obtained from the conducted analyses. These results are discussed in relation to the proposed objectives and research questions. The chapter covers the findings from the extreme rainfall network analysis and the characterization of the communities within the network. Notably, the comparison of different similarity functions is presented only for the rainfall networks in the Amazon basin. The final conclusions of the research are summarized in Chapter 7. Lastly, the Appendix contains additional relevant information. Section A.1 provides the links to access the data used in the research, while Section A.2 lists the software used along with their main packages.

2 OBJECTIVES

Within the study of rainfall and extreme rainfall events, it is of vital importance to recognize and characterize highly connected regions, known as communities, and analyze their representative features. The study of these systems is crucial due to their strong association with severe weather events that can cause significant damage to society. Understanding these systems is fundamental to mitigating the impacts of these events on the population and the economy.

In this context, the objective of this study is to identify highly connected regions called communities that exhibit similar characteristics in precipitation events during the summertime, using complex networks. This involves evaluating the topological, physical, and statistical properties of the communities and analyzing their relationships and connections. However, it is necessary to detail certain processes to accomplish this goal. Therefore, the specific objectives are presented below.

- a) Identify communities where extreme continental precipitation events in South America have similar characteristics during the summertime using complex networks and evaluate the results of topological, physical, and statistical properties of the communities;
- b) Analyze the relationships among the communities, as well as investigate possible influences and connections between them. Verifying the consistency of the identified communities with information from previous studies present in the literature;
- c) Define the physical properties of the precipitation communities. Characterizing the precipitation distribution and convective activity in each region of interest using physical information from vertical profiles of radar reflectivity;
- d) Examine the behavior of networks at smaller scales within each community to identify possible relationships or patterns with local characteristics, such as land cover, topography, and rainfall regime, among other aspects. Evaluating the sensitivity of communities to different approaches in defining the network for more restricted regions.

By achieving these objectives, the aim of this thesis is to gain a better understanding of precipitation systems and their related mechanisms.

3 THEORETICAL BACKGROUND

3.1 South America summer precipitation

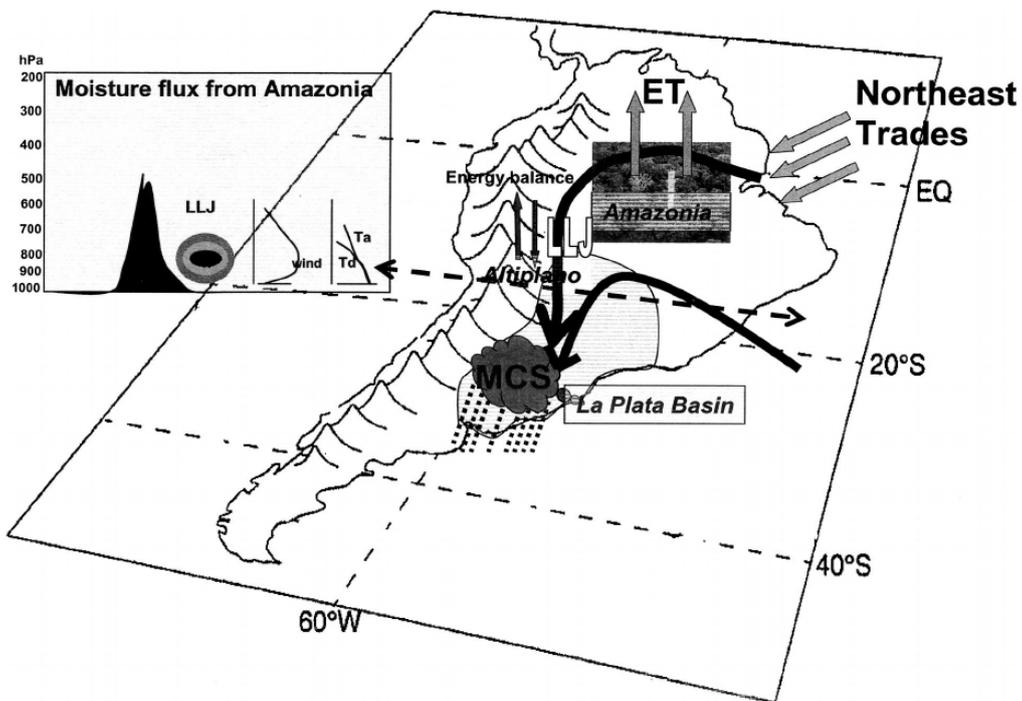
In South America, the precipitation regime is a result of the interaction between air masses brought from the North Atlantic by the trade winds and the orographic barrier of the Andes mountain range (MARENGO et al., 2012). When these air masses encounter the Andes, they ascend and change direction towards the south-east, leading to the formation of a moisture convergence band within the continent (WEBSTER et al., 1998; TRENBERTH et al., 2000; JORGETTI et al., 2014). This persistent and intense moisture transport during the austral summer significantly impacts the precipitation in South America, particularly the central and southeastern regions of Brazil (BOMBARDI; CARVALHO, 2011; MARENGO et al., 2012). This association with strong convective activity, precipitation, and atmospheric circulation is referred to as the South American Monsoon System (SAMS) (GAN et al., 2004; VERA et al., 2006b).

The intensity and configuration of the SAMS are primarily influenced by two key mechanisms: 1) The meridional shift of the Intertropical Convergence Zone (ITCZ), which controls the moisture flow from the North Atlantic to the South American continent (VUILLE et al., 2012). 2) The structure of the South Atlantic Convergence Zone (SACZ), which is a band of cloudiness and rainfall resulting from deep convection. This zone is formed by the convergence of moist air masses from the Amazon, the continental interior, and the South Atlantic, extending from northwest to southeast across the tropical region (KODAMA, 1992; LIEBMANN et al., 1999; CARVALHO et al., 2004; JORGETTI et al., 2014).

The Andes Mountains in South America act as a barrier to the moisture-laden air currents originating from the Amazon. These currents, known as the South American low-level jet (SALLJ), are deflected southward, serving as a moisture corridor between tropical South America and *La Plata* river basin (BOERS et al., 2013; BOERS et al., 2015; GELBRECHT et al., 2021). The SALLJ extends towards the Bolivian *Altiplano* and Paraguay (BERRI; INZUNZA, 1993) and continues across southeastern South America and subtropical latitudes, reaching northern Argentina, where it interacts with low-pressure systems (FERREIRA et al., 2010; BERRI; INZUNZA, 1993). This pattern of water vapor flux from the Amazon region to higher latitudes (RASMUSSEN; MO, 1996) establishes an exchange of air masses between the tropics and extratropical regions, creating favorable conditions for the development of deep convection processes in southeastern South America.

Various studies demonstrated significant progress in understanding the structure of the SALLJ and its temporal variability, showing that the SALLJ is much stronger during the late afternoon, thereby the SALLJ represents a corridor for moisture transport from the tropics (MARENGO et al., 2002; BERBERY; COLLINI, 2000). This moisture advection is more intense during the warm months (austral summer), while at latitudes near 20°S, it exhibits a higher frequency of events during the cold months (austral winter). During winter, the SALLJ shows greater hourly variability in frequency, which is consistent with the results shown by Velasco and Fritsch (1987). Marengo et al. (2004) developed a diagram illustrating the process of moisture transport from the Amazon to extratropical latitudes (Figure 3.1), and its contribution to precipitation systems over the *La Plata* basin.

Figure 3.1 - Advection mechanisms in South America.



Schematic representation of moisture transport from the Amazon region, driven by the SALLJ, and showing its contribution to MCS over Uruguay.
SOURCE: Marengo et al. (2004).

The studies mentioned above provide valuable insights into the mechanisms that influence precipitation patterns in South America. Now, let us delve into a concise review of extreme rainfall events and highlight the importance of comprehending their nature.

3.2 Extreme rainfall events

The adaptation and mitigation of adverse climate effects are of great importance. These extreme rainfall events pose an important threat to the safety and well-being of populations, as well as causing material damage and great socio-economic impacts. Therefore, understanding the patterns and mechanisms that lead to these events is crucial for the development of prevention and preparedness strategies. In this context, studies on the spatial and temporal distribution of these events, as well as the atmospheric influences and behavior of convective systems, are essential to improve the understanding and forecasting capacity of these extreme rainfall events in South America (BOERS et al., 2014b). For example, according to Marengo (2008), Brazil is subject to the effects of ongoing climate change and will be further affected by them in the future, especially regarding extreme weather events. To gather information, a literature review will be presented, discussing the points mentioned earlier and providing the theoretical context upon which this work is based.

The Intergovernmental Panel on Climate Change (IPCC), in its Climate Change 2014 report, provides substantial evidence of the increasing occurrence of extreme events in various regions worldwide, particularly in less developed countries with tropical climates (FIELD et al., 2014). The report highlights the intensification of extreme events in water resources as a result of global climate change. Additionally, the World Meteorological Organization (WMO) report confirms a global average temperature rise, which is closely linked to significant seasonal variations in precipitation across the globe (ALLAN; SODEN, 2008; MARENGO et al., 2020; STUART et al., 2022). According to The Global Annual to Decadal Climate Update report, the unprecedented concentrations of greenhouse gases and accumulated heat in the atmosphere have contributed to the past seven years being the warmest on record (HERMANSON et al., 2022). These changes have had a significant impact on the hydrological cycle (RANJAN et al., 2006; MADAKUMBURA et al., 2019; MA et al., 2020). The rise in temperatures has resulted in an intensification of extreme events in water resources, altering the frequency and intensity of extreme precipitation. This response is driven by physical mechanisms and feedback processes occurring within storms, primarily associated with the release of latent heat from increased thermodynamic moisture (LACKMANN, 2013; NIE et al., 2018; MIZUTA; ENDO, 2020). In this context, understanding the weather and climate systems plays a crucial role in alerting the population about imminent risks and assisting relevant authorities and civil defense in decision-making and planning aimed at mitigating the potential impacts of these events.

In the case of Brazil, it exhibits diverse meteorological characteristics and patterns. Extreme rainfall can affect different regions, but the magnitude of the impacts varies. In southern Brazil, more than 90% of extreme rainfall events are attributed to convective systems, while in the Northeast, this proportion ranges from 60% to 90% (PALHARINI et al., 2022). The floods that occurred in the states of *Bahia* and *Minas Gerais* in December 2021 are attributed to the presence of the South Atlantic Convergence Zone (SACZ) (MARENGO et al., 2023). The extreme rainfall events in the southwestern region of *Rio Grande do Sul* can be attributed to the formation of large convective systems known as Mesoscale Convective Systems (MCS). These systems form due to atmospheric convergence and low-level moisture transport, primarily originating from the Amazon region (SANCHES et al., 2019).

As mentioned above, extreme rainfall events are closely linked to convection. These events play a significant role in the variability of weather conditions across different regions of the world. These events are characterized by intense and localized precipitation, often occurring over short periods of time. Convection is the process responsible for the formation of storm clouds and the release of latent energy in the atmosphere, resulting in heavy rainfall and extreme weather events such as storms. Understanding extreme rainfall events, even if these are not associated with convection, is of utmost importance, as these events can have significant impacts on various areas, including agriculture, infrastructure, water resources, and public safety. Furthermore, with global climate change, it is expected that the frequency and intensity of these events may be affected, making the study of these phenomena even more crucial.

3.2.1 Convection

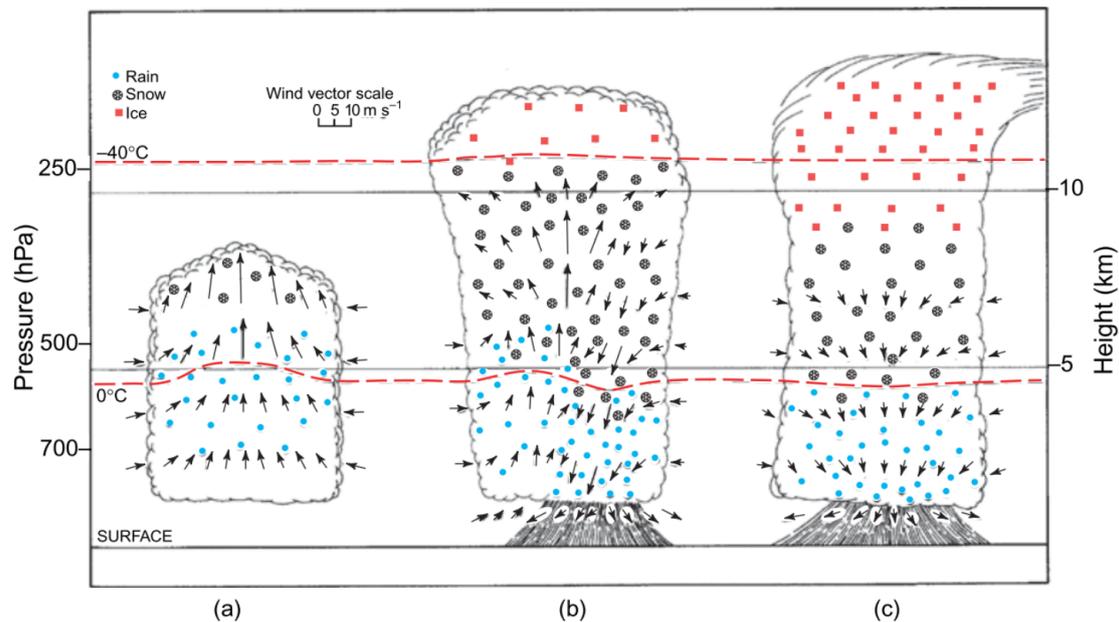
Convection is one of the three modes of heat transfer: conduction, convection, and radiation. In general terms, convection refers to the transfer of heat through the movement of a fluid. According to the American Meteorological Society (AMS), convection is predominantly vertical motion driven by buoyancy forces arising from static instability, which can be understood as thermally driven circulations (AMERICAN METEOROLOGICAL SOCIETY (AMS), 2020). Atmospheric convection can occur as free convection, where the transport is mainly driven by density differences within the fluid. Additionally, forced convection can also occur, where the transport is influenced by mechanical forces such as surface heterogeneity or turbulent processes. Turbulent convection is more common in the atmosphere. Furthermore, there is also the distinction between moist and dry convection.

For example, cloud formation is associated with moist convection, while convection that occurs without cloud formation is referred to as dry convection. The terms convection and storms are often used interchangeably, although storms are just one outcome of a specific type of convection. This association serves as an example of the potential events caused by convection. The study of deep convection and its relationship with severe weather phenomena is an area of research that receives considerable attention within meteorology. Since convection is a crucial process in atmospheric dynamics and due to the negative impacts it can have on society and the economy, such as heavy precipitation, lightning, strong gusts of wind, and hail, it is important to study it. These phenomena, in turn, can lead to other effects like flooding, landslides, power outages, tree falls, and most importantly, loss of human lives (PIELKE; CARBONE, 2002; NASCIMENTO, 2005). Therefore, it is evident that studying convection is of utmost importance as the information gathered can provide key insights for describing, characterizing, or predicting convective systems that may give rise to extreme events. As mentioned earlier, various techniques have been applied to the study of convective systems, often focusing on their morphological characteristics. However, in this study, complex networks will be used to analyze both rainfall events and their extreme episodes throughout South America.

Indeed, considering the relationship between atmospheric convection and extreme rainfall events, it is important to understand the formation and characteristics of *Cumulonimbus* clouds in order to advance in the study and comprehension of these significant events. *Cumulonimbus* is considered the final stage of development for a *cumulus* cloud and is characterized by its imposing and vertically extended structure. These clouds are capable of reaching great heights in the atmosphere, with altitudes exceeding 10 kilometers. In general terms, most *Cumulonimbus* clouds are associated with thunderstorms, and therefore, the study of the different characteristics of convective cells is of utmost importance (WEISMAN; KLEMP, 1986). This includes examining favorable thermodynamic conditions (WILLIAMS et al., 2005), the presence of updrafts (DOSWELL, 2001), and the life cycle of thunderstorms (BYERS; BRAHAM, 1948; CINTINEO et al., 2013). These factors play a crucial role in the formation, development, and intensity of convective storms and can provide valuable insights into the understanding of extreme rainfall events. These factors play a crucial role in the formation and intensification of convective clouds, as well as the interactions between different convective cells within the system. Moreover, understanding the evolution of a convective system also allows for the identification of key stages in this process, such as the development or cumulus stage, the mature stage, and the dissipating stage (BYERS; BRAHAM, 1948).

Figure 3.2 presents a schematic representation of the developmental phases. Each stage presents distinct characteristics and is associated with different types of events, such as the formation of updrafts and downdrafts, the release of latent heat, the production of precipitation, and the occurrence of strong winds. By recognizing the specific moment within the development process of a convective system in which certain events occur, it is possible to improve the prediction of adverse weather phenomena such as severe storms, tornadoes, and heavy rainfall. This contributes to the safety and planning of human activities that may be impacted by these events, such as agriculture, aviation, water resource management, and disaster response. Therefore, understanding a convective system is essential for the analysis and prediction of weather events, allowing for a better understanding of atmospheric processes and providing valuable information for management and decision-making regarding convective phenomena.

Figure 3.2 - Schematic representation of the stages that represent the life cycle of a storm.



(a) *cumulus* stage (b) mature stage (c) dissipation stage.
SOURCE: Wallace and Hobbs (2006).

The *cumulus* stage is characterized by towering storm clouds, primarily structured by strong upward movements of warm and moist air due to converging surface flow. Usually, the onset of precipitation in the lower layers indicates the transition to the next stage.

The mature stage is composed of more predominant upward movements compared to the previous stage, which causes divergence at upper levels of the troposphere, leading to the formation of the so-called cloud anvil (the tropopause acts as a natural boundary). The upward movement is fueled by the condensation process of water, as it releases latent heat, while the evaporation of water droplets within the cloud causes cooling and subsequent downward movement that triggers the gust front. Finally, the dissipation stage, unlike the previous stages, is predominantly characterized by downward vertical movements, presenting surface divergences and continuous cooling of the air through evaporation, which feeds the gust front moving away from the cloud. At this stage, the system no longer provides the necessary upward movements for the organization of the system, resulting in the dissipation of the cloud.

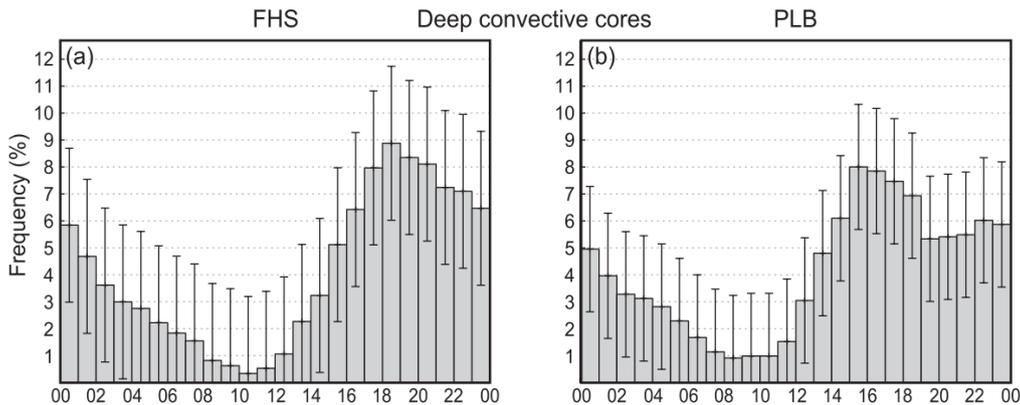
Convective systems can also be classified according to their organization and extent, and they can be grouped into four types. The first one is the *Single-cell* system, which consists of a single structure. It is the simplest convective storm system with a lifespan of approximately 30 to 60 minutes. On the other hand, Multicellular systems are more organized, presenting a group of storm families. Multicellular systems have much longer lifespans compared to single-cell systems, and they can generate adverse weather phenomena. The Supercells exhibit a higher degree of organization compared to the previous types, with a single, highly intense cell that makes it one of the most potentially damaging structures. Such systems are associated with the presence of hail, strong winds, and even tornadoes, and in some cases, their effects at the surface appear as very intense rainfall. These systems are characterized by having rotation in the updrafts. Lastly, Mesoscale Convective Systems (MCS) are structures with a higher degree of organization, typically composed of a cluster of convective clouds. During their mature phase, they often exhibit stratiform and continuous rainfall with the presence of very intense storm cores. Several studies have shown that for MCS, their lifespan is associated with their size (MADDOX, 1980; HOUZE et al., 1990; MACHADO et al., 2004; LIU et al., 2007). Among convective systems, there are some that have received particular interest, such as the Mesoscale Convective Complex (MCC), which is composed of a cluster of clouds with significant vertical development and symmetry in their shape (MADDOX, 1980). This criterion arose from the need to identify large convective systems based on their horizontal and vertical dimensions, without considering the internal structure and using infrared (IR) imagery. For some cases, smaller-scale MCS exhibit similar structures to mesoscale convective complexes, for example, Maddox (1980) used MCC criteria with certain modifications (BARTELS et al., 1984).

Among the favorable conditions for the generation of convection are: high heat and humidity, convergence of winds at lower levels, upward vertical motion, and high values of specific humidity at the surface. Such conditions can be organized by larger-scale systems, such as cold fronts (SIQUEIRA; MACHADO, 2004; SIQUEIRA et al., 2005). One of the first studies on convective systems was conducted by Cavalcanti (1982), who observed the formation of MCS over Paraguay, suggesting that these systems were possibly linked to larger-scale circulation along with orographic forcing. Subsequently, several studies emerged, examining convective systems in South America and their possible origins (GUEDES, 1985). Fritsch et al. (1986) and colleagues investigated convective systems during the summers of 1982 and 1983, aiming to assess the frequency and contribution of convective events to precipitation. The results showed that during this season, precipitation increased by 30% to 50% due to MCC. Based on the criteria presented by Maddox (1980) for the definition of MCC, several studies have been conducted analyzing the presence of these systems (MCANELLY; COTTON, 1986; MERRITT; FRITSCH, 1984; BARTELS; ROCKWOOD, 1983). For instance, Kane et al. (1987) examined 106 MCS, of which 74 were identified as MCC, while the remaining systems did not meet the criteria of Maddox (1980) and were referred to as MCC-like systems at that time (FRITSCH et al., 1986). In the early 1990s, Houze et al. (1990) found no differences in the precipitation structure generated by smaller MCC and MCS. Mohr and Zipser (1996) defined characteristics to recognize intense MCS and showed that the tropical region of South America and Africa, along with the warm pool of the South Pacific, are areas with the highest number of MCS. They also demonstrated that MCS over continents tends to be colder than those over oceanic regions (higher cloud tops). Furthermore, they found that over continents, these systems are more frequent in the evening, while over oceans, they occur more often in the morning, confirming the findings of Velasco and Fritsch (1987).

Using reflectivity data from TRMM, convective events can be identified based on their development characteristics, such as the intensity of the convective core (higher reflectivity values), height, and horizontal extent (NESBITT et al., 2000). Based on this, convective events can be characterized as presented by Romatschke et al. (2010), who identified different structures of convective systems that were termed as *deep convective cores*, defined as young convective cells with strong updrafts and typically associated with severe weather conditions, reaching heights greater than 10 km. *Broad convective cores*, representing convective cores that extend horizontally over areas larger than 1000 km², often part of larger convective systems such as MCS (HOUZE JUNIOR, 2004).

Finally, *broad stratiform regions*, refers to systems with areas larger than 50000 km², which typically occur during the dissipation phase of an MCS. In addition to this categorization, the authors studied the hourly frequency, revealing that deep convective cores occur from late afternoon (Figure 3.3), suggesting that the diurnal cycle plays a crucial role in modulating convection and precipitation mechanisms in the continent.

Figure 3.3 - Hourly frequency of deep convection cores.



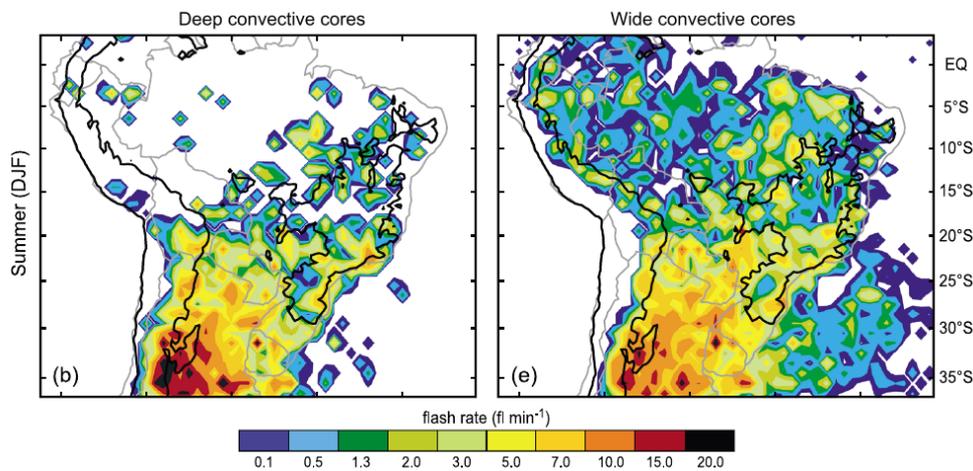
(a) Eastern Central Andes and (b) *La Plata* Basin - Argentina.

SOURCE: Romatschke et al. (2010).

According to Zipser et al. (2006), the regions with the most intense storms in the world are the eastern Andes in southeastern South America, the eastern Rocky Mountains in the United States, and near the western Himalayas in southern Asia. Furthermore, South America is an area of significant interest for the study of intense convective systems, particularly the subtropical region, which is the second-largest continental region in the world in terms of MCS extent, surpassed only by South Asia (MOHR; ZIPSER, 1996). Moreover, this part of the planet experiences the most extensive extreme precipitation events (KUMAR et al., 2019a) and the deepest convective storms in the world, occurring over Argentina (RASMUSSEN et al., 2014). The systems present similar characteristics to events observed in the Rocky Mountains of the United States, with the highest nocturnal activity (MARENGO et al., 2003). According to Machado et al. (1998), there is a direct relationship between the spatial dimensions and the life cycle of MCS (MACHADO et al., 2004; HOUZE et al., 1990). Sakamoto et al. (2009) examined the phase variations of MCS and the difference in diurnal cycles between continental and oceanic regions.

The authors noted a higher frequency over continental areas, similar to [Mohr and Zipser \(1996\)](#) and [Velasco and Fritsch \(1987\)](#). [Durkee et al. \(2009\)](#) obtained the climatology of MCC for South America and found that these complexes have larger precipitation areas compared to those observed in North America and Africa. Furthermore, MCC contributes between 15% and 21% to the total precipitation over northern Argentina and Paraguay during the period from 1998 to 2007. Other studies have focused on more specific regions, such as [Anabor et al. \(2008\)](#), which investigated the presence of MCS in the Plata basin and its local influence, as some of these systems have a life cycle of up to 69 hours, associated with strong moisture advection from the Amazon region.

Figure 3.4 - Average lightning rates present over South America.



Convective systems (deep and extensive) associated with extreme storms, showing a core of maximum activity over the *Sierras de Córdoba*, Argentina. The thick black line represents the contour of the 0.5km topography.

SOURCE: [Rasmussen et al. \(2014\)](#).

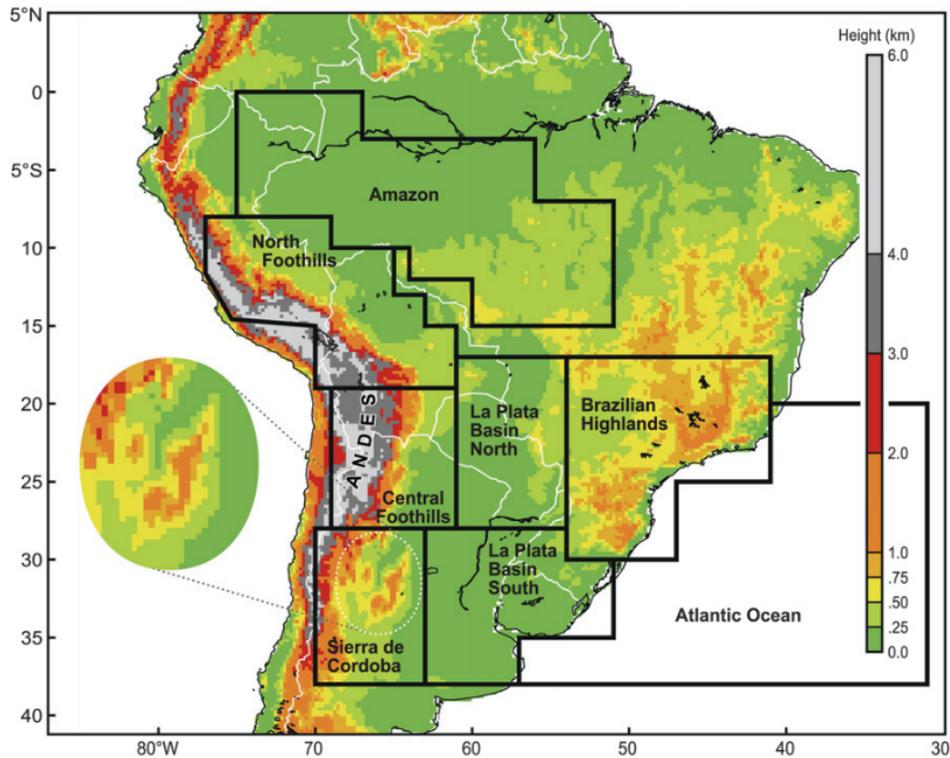
As mentioned, central Argentina is one of the regions that experience the widest convective cores ([DURKEE et al., 2009](#)) and the most intense convective storms with a high occurrence of lightning (Figure 3.4). Using weather radar data in the *Sierras de Córdoba* of Argentina, [Mulholland et al. \(2018\)](#) studied intense convection events and found that the majority of storms are multicellular (composed of multiple cells, each in a different stage of development). These events occurred from the early afternoon until late at night, with a higher number of occurrences in the austral summer. In the annual cycle, the months with the highest activity of deep convection were from November to February, while the lowest number of events occurred in the Austral winter.

The mechanisms or forcings associated with convective activity are diverse and include the interaction between the ocean and the land, the convergence of warm and moist air at lower atmospheric levels, the presence of mountains acting as orographic forcing, the advancement of cold fronts, among others. As mentioned earlier in Section 3.1, another significant mechanism is the low-level jet (LLJ), which has the capability to generate regions of moisture convergence at lower atmospheric levels. Given the importance of LLJ, numerous studies have examined nocturnal storms associated with them and how they facilitate the transport of warm and moist air in various regions around the world. For example in Africa (JURY; TOSEN, 1989), Asia (FINDLATER, 1969), Australia (KEENAN et al., 1989), India (KALAPUREDDY et al., 2007), South America (BERBERY; COLLINI, 2000; GELBRECHT et al., 2021), the Caribbean (WHYTE et al., 2008) and United States (KINCER, 1916; MEANS, 1952; WALLACE, 1975; ASTLING et al., 1985; CHEN; KPAEYEH, 1993; NICOLINI et al., 1993). These studies also highlight the strong relationship between the development of these jet streams and the presence of mountain ranges, such as the Rocky Mountains in the United States, the Southern Himalayas in South Asia, and the Andes in South America. This serves as an example of how significant geographical variations can influence the organization of convection.

In the case of South America, the South America Low-Level Jet (SALLJ) transports moisture to different regions, with the Amazon basin acting as a major source of water vapor. This is one of the reasons why the Amazon basin is highly significant, as it is known for its high complexity. Several studies have shown that precipitation systems over the Amazon basin region are more similar to those over the ocean rather than the continent. This indicates that these systems are more extensive in nature rather than intense (LIU et al., 2007; NESBITT et al., 2000; RASMUSSEN; HOUZE, 2011). For instance, during the summer, the Amazon region, along with the South Atlantic Ocean, exhibits few deep convective cores. Additionally, the Amazon region displays spatial variability in the occurrence of Mesoscale Convective Systems (MCS), with a higher concentration in the east compared to the west (JARAMILLO et al., 2017).

Another important region in South America is the central-eastern Andes, characterized by high activity of convective events with deep cores. This region exhibits some similarity to the seasonal pattern observed in the Cordoba region. While the highest activity of deep convective cores occurs in spring to the east of the central Andes, in the Cordoba region, this intense activity is observed during the summer, indicating an apparent connection with a lag between the two regions (ROMATSCHKE

Figure 3.5 - Regions of geographic interest in South America.



SOURCE: Rasmussen and Houze (2011).

et al., 2010; RASMUSSEN; HOUZE, 2011; BRUICK et al., 2019). The presence of these convective cores over the eastern Andes is associated with the occurrence of extreme rainfall events, especially over the region known as the Peruvian rainfall hotspot (ESPINOZA et al., 2015).

While identifying extreme rainfall events, these can be defined using the 90th percentile of rainfall. This method of definition is based on the statistical analysis of the distribution of precipitation values over a specific period of time in a particular region (ROY; BALLING JUNIOR, 2004; GRIMM; TEDESCHI, 2009; BOERS et al., 2013; ESPINOZA et al., 2015; SONEJA et al., 2016; ZUBIETA et al., 2019).

The 90th percentile represents the value below which 90% of the data lies. By calculating the 90th percentile for a time series of precipitation, we are finding a value that is greater than this percentile of the observed series at each point, this is because what is extreme in one place, might not be in another.

Precipitation values that exceed this threshold are considered extreme rainfall events. Indeed, this approach allows us to identify episodes of intense and concen-

trated rainfall within a specific time period. It is important to note that the value of the 90th percentile can vary depending on the region and the analysis period. In areas with a history of heavier rainfall, it is likely that the threshold for defining extreme events will be higher compared to regions with less intense rainfall. This variability takes into account the local climatic conditions and provides a relative measure of extreme rainfall events specific to a particular area.

In this section, we provided a brief introduction to studies related to extreme rainfall events, discussing their characteristics and different causes. The objective of this study is to apply new techniques in the analysis of extreme precipitation events in South America. In the following sections, we will conduct a literature review of the main technique that will be used to achieve these objectives, complex networks. This approach will be primarily employed to obtain key information about the system's dynamics and identify highly connected regions, known as communities.

3.3 Complex networks in meteorology

The mathematical definition of a graph is a set of objects called vertices or nodes connected by relationships known as edges, which represent binary relationships between elements of a set. A graph G consists of two sets. The first set, V , is the set of its nodes or vertices, and the second set, E , is the set of its edges or connections. Each edge is related to two elements of V , and this relationship defines the topology of the graph. When a graph is used to describe relationships between various discrete objects, it is called a network. The definition of a network is very general, allowing it to be applied to different systems.

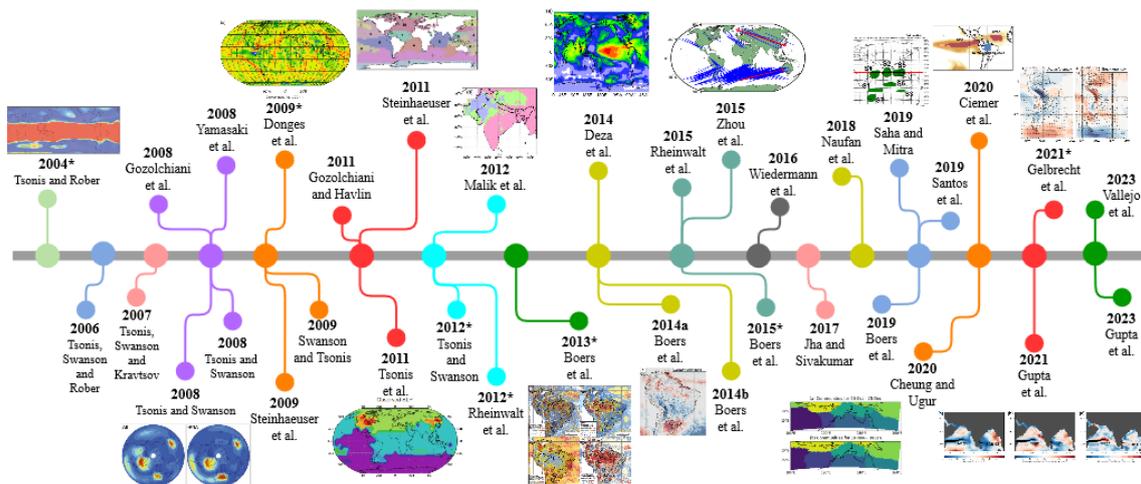
We could define a complex network as a graph that possesses certain non-trivial topological and statistical properties (BARABASI, 2003). Currently, a wide variety of applications can be found using complex networks, such as the World Wide Web (ALBERT et al., 1999), the collaboration network of film actors, protein folding, scientific collaboration network (EOM; JO, 2014), cellular networks, citation networks (ALBERT; BARABÁSI, 2002), gene expression data (AGRAWAL, 2002), the travel route map of Continental Airlines (TSONIS et al., 2006), ecological network (MONTROYA et al., 2006), airport network (BAGLER, 2008; ZANIN; LILLO, 2013; MALIK et al., 2019), public traffic network (AN et al., 2014), among others.

These examples demonstrate the great flexibility of the technique in various fields of study, and meteorology is no exception. Numerous studies have been conducted in meteorology using complex networks. It started with Tsonis and Roebber (2004),

who presented the first work on climate networks. The authors observed that the climate network exhibits small-world properties but is composed of two subnetworks: a regular tropical network and a scale-free extratropical network. Building on this, several studies have been published applying network theory concepts to climate data, yielding highly favorable results (TSONIS et al., 2006; TSONIS et al., 2007; TSONIS et al., 2008; TSONIS et al., 2011; TSONIS; SWANSON, 2008; YAMASAKI et al., 2008; GOZOLCHIANI et al., 2008; SWANSON; TSONIS, 2009).

Figure 3.6 presents a brief compilation of some of the principal studies in complex networks applied to meteorological data. In general terms, it can be observed that the early studies between the years 2004 and 2011 focused on identifying the dynamics of low-frequency disturbances on a global scale, with initial applications related to El Niño and teleconnections. Between the years 2012 and 2014, innovative work emerged concerning monsoonal systems, with a significant emphasis on understanding their dynamics. In recent years, a wide variety of applications has arisen, such as the use of networks before, during, and after the phenomenon of interest to study its evolution; improvements in teleconnection results; the use of complex networks as a comparison technique, to verify the differences between the dynamics of the system; the identification of key regions, and how they can serve as indicators for climate prediction; changes in the structure of communities to understand the response to anomalies; among others. Below, we present in greater detail the relevant studies related to this research.

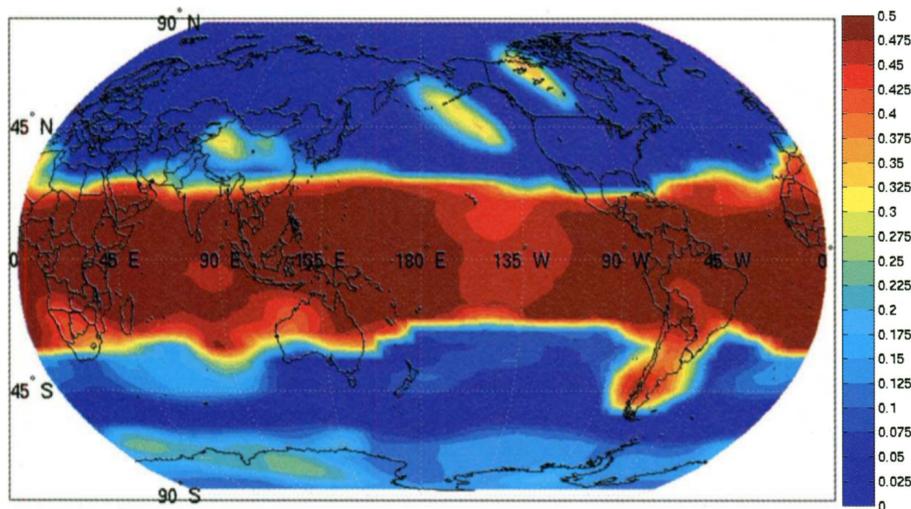
Figure 3.6 - Brief review of Complex Networks in meteorology applications.



SOURCE: Author's production.

Tsonis and Roebber (2004) obtained the Area Weighted Connectivity (AWC), which shows the fraction of the total area to which each vertex is connected (Figure 3.7). This network was constructed using geopotential height (*hgt*) data at 500 hPa. This variable was chosen because it provides a good representation of the general circulation of the atmosphere. The results were consistent with the original findings of Tsonis and Roebber (2004), where vertices in the tropics have almost the same number of connections, while highly connected vertices are found in the extratropics. Additionally, the authors used cross-correlation to define the links between the vertices. However, one of the advantages of complex networks is their versatility, and other methods of defining these links can be applied, such as using mutual information (YAMASAKI et al., 2008; MARWAN et al., 2009; DONGES et al., 2009b; DONGES et al., 2009a; DEZA et al., 2014).

Figure 3.7 - AWC for weather network using monthly geopotential height anomaly at 500 hPa.



Cross-correlations were used as measures of connection between vertices with zero lag. The time series used were monthly anomalies, and two vertices were considered connected only if the absolute value of their cross-correlation was $|r| \geq 0.5$, which is statistically significant above the 99% level.

SOURCE: Tsonis et al. (2006).

Donges et al. (2009b) presented the differences in the climate network using connections defined with Pearson's linear correlation and mutual information, where these differences mainly occur in regions of higher variability, which may be indicative of nonlinear physical processes in the climate system. Additionally, both networks exhibit small-world properties. Some of the results from these studies revealed im-

portant characteristics of the climate system, some of which were already known to the scientific community, while others were new. For example, in the Northern Hemisphere, the presence of highly connected vertices in North America, China, and the North Pacific was observed. In relation to the highly connected vertices mentioned earlier, they could be representing part of the atmospheric teleconnections associated with El Niño, specifically the Pacific-North America (PNA) pattern (WALLACE; GUTZLER, 1981). The PNA is an index that represents low-frequency variability over the Northern Hemisphere. This pattern is observed at the 500 hPa level and reflects the propagation of Rossby waves (planetary waves) generated by the exchange of energy between the atmosphere and the ocean. This mechanism is triggered by the anomalous warming of sea surface temperatures in the central equatorial Pacific. On the other hand, in the Southern Hemisphere, the presence of supervertices in southern South America, Antarctica, and the Indian Ocean is observed. In this case, these areas are associated with the Pacific-South America (PSA) pattern, which represents the propagation of waves towards the Southern Hemisphere (MO; HIGGINS, 1998).

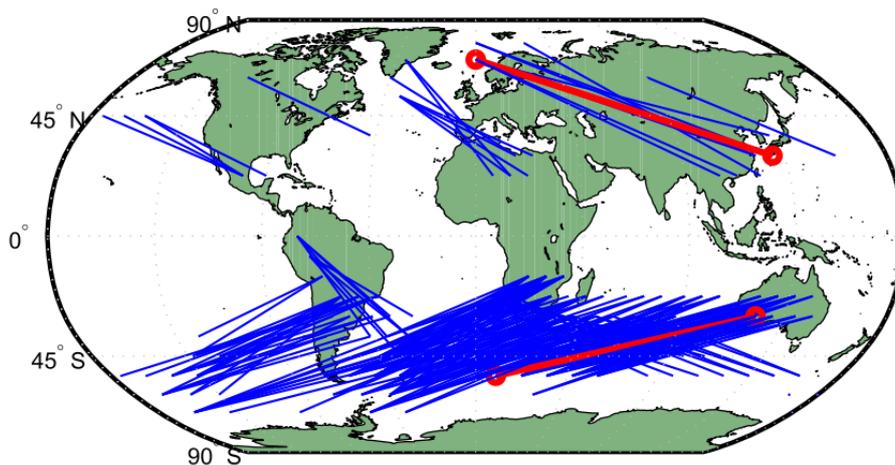
However, another interesting characteristic of the network presented by Donges et al. (2009b) is the absence of the North Atlantic Oscillation (NAO) Index. It should be noted that the NAO represents large-scale fluctuations in the atmosphere, specifically between the subtropical high-pressure zone (Northern Hemisphere) and the polar low-pressure zone over the North Atlantic. This index determines the winter climate variability in the North Atlantic region. The absence of the NAO in the network can be mainly attributed to the fact that it does not have strong connections with tropical regions. Additionally, its presence might be overshadowed by the high connectivity observed in other regions.

To verify the presence of the NAO, Tsonis et al. (2008) used Empirical Orthogonal Functions (EOF) to extract the influence of Pacific North America (PNA) pattern from the climate network, focusing on the Northern Hemisphere. It was observed that after removing the influence of the PNA, the structure of local correlations did not vary significantly. However, in the case of the NAO, the local correlations weakened in the absence of the index. This indicates that the PNA does not play a relevant role in the stability and efficiency of the climate network in extratropical regions. The results also show that the PNA is a linear response to forcing in the tropics, while the NAO has a more significant influence in the extratropical regions, establishing a stronger association.

The study conducted by Tsonis et al. (2011) identified that the structure of the geopotential height (*hgt*) network is associated with atmospheric teleconnection patterns, such as the Pacific North America (PNA) pattern in the Northern Hemisphere and the Pacific South America (PSA) pattern in the Southern Hemisphere. These findings highlight the influence of these teleconnection patterns on the organization and behavior of the climate system. Since the physical mechanism related to climate change is still not clear, complex networks have also been used to analyze this issue (WANG et al., 2009; TSONIS; SWANSON, 2011). The authors noted the important role of the North Atlantic Oscillation (NAO). The NAO, with its significant influence over the North Atlantic, affects the strength of the westerly flow across the mid-latitudes. Therefore, the NAO, which is the dominant mode of variability in the North Atlantic (TSONIS et al., 2008), impacts the variability of the El Niño-Southern Oscillation (ENSO) and global temperature variability.

At the end of the 1990s, the prevailing idea was that the global atmospheric response to extreme phases of El Niño-Southern Oscillation (ENSO) was linear. However, many studies have emphasized the nonlinearity of the atmosphere's response to tropical sea surface temperature (SST) forcing, particularly during the winter of the Northern Hemisphere (HOERLING et al., 1997; WU; HSIEH, 2004). Hence, there is interest in the nonlinear analysis of atmospheric teleconnections caused by El Niño. Due to the importance of these teleconnections, they have been analyzed using complex networks (YAMASAKI et al., 2008; GOZOLCHIANI et al., 2011; DEZA et al., 2014), as mentioned above.

Figure 3.8 - Strong and long positive links greater than 5000 km.



SOURCE: Zhou et al. (2015).

Some studies have focused on the use of other variables beyond those mentioned above, revealing unknown patterns that deserve attention. Notably, the work of [Zhou et al. \(2015\)](#) provided pathways of influence propagation between remote regions for an air temperature network, demonstrating that teleconnections are related to waves, winds, or currents (Figure 3.8). In this study, the authors emphasize the known teleconnection patterns. Particularly noteworthy are the results for South America, where strong relationships are observed between the Amazon Basin and the region between Uruguay and southern Brazil, an area characterized by high atmospheric electrical activity ([ROMATSCHKE et al., 2010](#)).

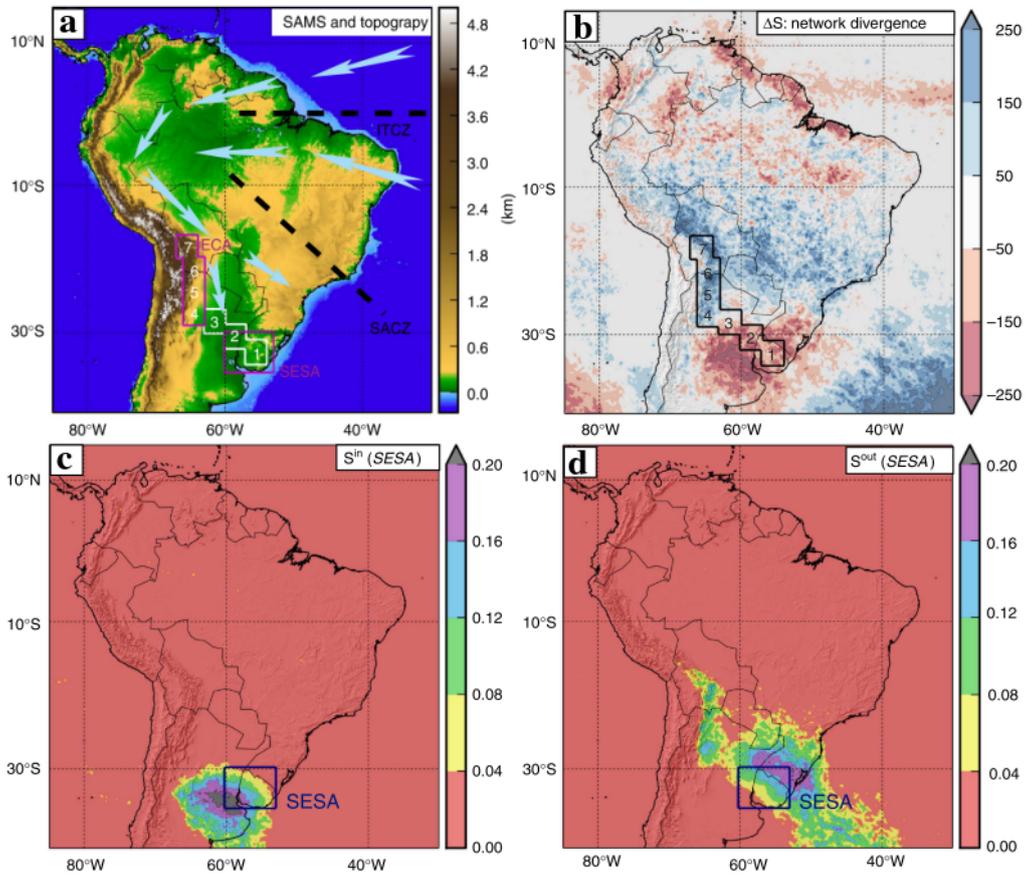
These findings highlight the importance of understanding the effects of teleconnections and their pathways. By recognizing these interconnected areas, we can not only monitor these events more effectively but also potentially make prognostic assessments.

Given the significance of this phenomenon and its associated impacts, various efforts have been made to forecast El Niño events ([BARNETT et al., 1994](#); [BARNSTON et al., 2012](#); [LUDESCHER et al., 2014](#)). In their study, [Tsonis and Swanson \(2008\)](#) investigated variations in the topology of the climate network during El Niño and La Niña years separately. They found that the El Niño network is less connected compared to the La Niña network. The ENSO phases induce fluctuations at different scales, which limits their predictability. Therefore, recognizing these effects could enhance the prediction of these events. Recently, [Broni-Bedaiko et al. \(2019\)](#) introduced a methodology for El Niño event prediction using recurrent neural networks in conjunction with complex networks. The network was constructed following the conventional procedure using surface temperature data and Pearson’s linear correlation ([TSONIS; ROEBBER, 2004](#); [TSONIS et al., 2006](#)). In this study, twelve metrics characterizing the network (such as betweenness, network diameter, centrality, AWC, clustering coefficient, etc.) were utilized along with El Niño3.4 index data as inputs for the neural networks. The neural networks were able to generate predictions of quarterly sea surface temperature anomalies for Niño3.4 ([BARNSTON et al., 2012](#)).

South America is constantly affected by various weather phenomena that cause high levels of precipitation. [Boers et al. \(2014b\)](#) aimed to provide predictions of extreme precipitation by introducing the concept of network divergence, which is based on the nonlinear measure of event synchronization. Two regions of interest were defined. The first region is a region of intense rainfall limited by 35°S-30°S and 60°W-53°W,

referred to as Southeastern South America (SESA). The second region is located east of the central Andes and is called ECA (Eastern Central Andes) (Figure 3.9). This analysis indicates that precipitation events in SESA are followed by extreme events along a narrow strip in the eastern Andes and western Bolivia. In these regions, synchronized extreme events that occur in ECA were finally found two days after they occurred in SESA. Using this method, it was possible to predict more than 60% of precipitation events in the central Andes (90% during El Niño conditions).

Figure 3.9 - Topographic map and representation of the mechanisms of the South American monsoon system with the results of network divergence.



(a) The white boxes labeled 1 to 7 indicate the climatological propagation path of extreme events, while the purple boxes represent the boundaries of the Eastern Central Andes (ECA) and Southeastern South America (SESA) regions. (b) Network divergence is defined as the difference in input and output strength for each point in the network. Positive values indicate sinks in the network, where synchronized extreme rainfall occurs 2 days after the occurrence in the source, while negative values indicate sources, where synchronized rainfall occurs within 2 days before it occurs in multiple other locations. (c) S^{in} represents extreme events preceding the events occurring in SESA. (d) S^{out} represents extreme events in SESA followed by extreme events along the band.

SOURCE: Boers et al. (2014b).

Different studies have explored the concept of complex networks in investigating the properties of precipitation, such as its spatial and temporal connectivity. [Jha and Sivakumar \(2017\)](#) studied precipitation in the Murray-Darling Basin in Australia, where they constructed three different pluviometric networks (N1, N2, and N3). N1 and N2 have the same data length but differ in the spatial distribution of the number of rain gauges. N2 and N3 have different data lengths but the same number of rain gauge stations. For the temporal analysis, six different time scales were used, ranging from daily to monthly, along with seven minimum correlation values (0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9). It was observed that the correlation between precipitation at different time scales depends on the correlation interval, and this correlation can change from purely linear to highly nonlinear. Therefore, determining the minimum degree of correlation is of utmost importance.

[Naufan et al. \(2018\)](#) used precipitation data from the Weather Research and Forecasting (WRF) model and employed a similar methodology as presented by [Jha and Sivakumar \(2017\)](#). The authors found that networks with longer time scales exhibit small-world properties. For shorter time scales, the precipitation network can be classified as either a small-world network or a scale-free network, depending on the correlation threshold used. Finally, the spatial connections of rainfall exhibit a transition phase at intermediate time scales, especially at higher thresholds. The results generally indicate a negative relationship between the clustering coefficient and the threshold level and a positive relationship between the clustering coefficient and the temporal scale.

These observations seem to suggest that spatial connections in rainfall are different at different time scales. Therefore, it is noted that for higher-resolution studies, the properties of networks in meteorology are more variable, which may indicate that for local studies, not only climatic factors but also regional characteristics need to be considered.

One important aspect in the study of networks is the identification of the participation or activity of vertices, as this information can provide key insights about the system. Whether it is to identify vertices that are more susceptible to disturbances or those that are notably more important in terms of information transmission. [Agarwal et al. \(2018\)](#) presented an approach based on understanding the role of each node in a community, specifically in the context of rainfall networks over India, using Z-P space information ([GUIMERA; AMARAL, 2005; GUIMERA et al., 2007](#)). In this way, they were able to comprehend the qualitative and quantitative

aspects of community members and identify regions of higher importance associated with node activity. This approach demonstrated how complex networks can be used as a technique to identify potential geographic locations for the installation of new weather stations.

3.4 Communities identification

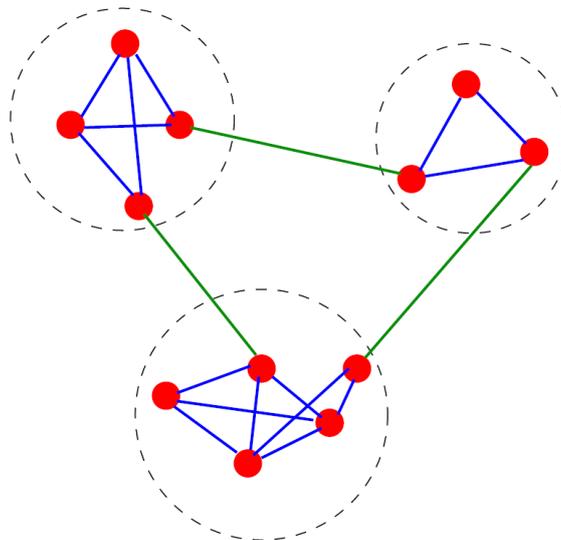
Usually, complex systems exhibit their own self-organized behaviors, depending on how information is transmitted and how vertices participate in this process. Typically, some vertices have a higher prevalence than others (BARABÁSI; ALBERT, 1999; NEWMAN, 2005; BARZEL; BARABÁSI, 2013). However, within the network, it is also possible to find specific groups that interact more intensively among themselves than with the rest of the network, characterizing a predominantly local interaction. These sub-graphs are known as communities (TSONIS; SWANSON, 2008; STEINHAEUSER et al., 2009; ABE; SUZUKI, 2012). Therefore, a community can be defined as a group of vertices that are more strongly interconnected among themselves than with the rest of the network. Vertices within the same community share similar characteristics.

Figure 3.10 illustrates in a simplified way a network with three communities. By analyzing the communities, it is possible to gain a better understanding of the structure and dynamics of the network, identifying groups of strongly interconnected nodes and exploring the relationships between them. This approach allows for a more detailed view and reveals important information about the organization and functioning of the network under study. Communities provide insights into natural groupings and can help identify patterns, processes, and significant interactions within the system represented by the network. This information can be used to understand the function of different parts of the network, identify influential nodes or important hubs, and even optimize processes or strategies based on the properties of the communities.

Therefore, the correct identification of communities is a fundamental step in the analysis of complex networks (MISSAOUI; SARR, 2015; SCHAUB et al., 2017).

There are several techniques developed for identifying communities in complex networks, including FastQ (NEWMAN; GIRVAN, 2004), WalkTrap (PONS; LATAPY, 2005; STEINHAEUSER et al., 2009), eigenvectors of the modularity matrix (NEWMAN, 2006), fuzzy clustering (SUN, 2015; NEPUSZ et al., 2008), IsoFdp (YOU et al., 2016), density peak model (DENG et al., 2019), and others (ZHANG et al., 2010;

Figure 3.10 - Schematic example of a network with three communities.



SOURCE: Fortunato (2010).

AGARWAL et al., 2018; TASGIN; BINGOL, 2019; ARTURO et al., 2022). However, the modularity function proposed by Newman and Girvan (2004) is widely used in the analysis of complex networks to estimate the quality of a network partition into communities. Modularity measures how different the original graph is compared to its randomizations. This measure is based on the idea that by randomizing the network structure, communities tend to break apart, making it possible to compare the real structure with its randomization and assess how non-random the community organization is (FORTUNATO; HRIC, 2016). In other words, modularity aims to identify clusters of nodes within the network that are more densely connected to each other than to the rest of the network. It quantifies the difference between the number of edges observed within the groups and the number of edges expected in a random network with the same degree distribution.

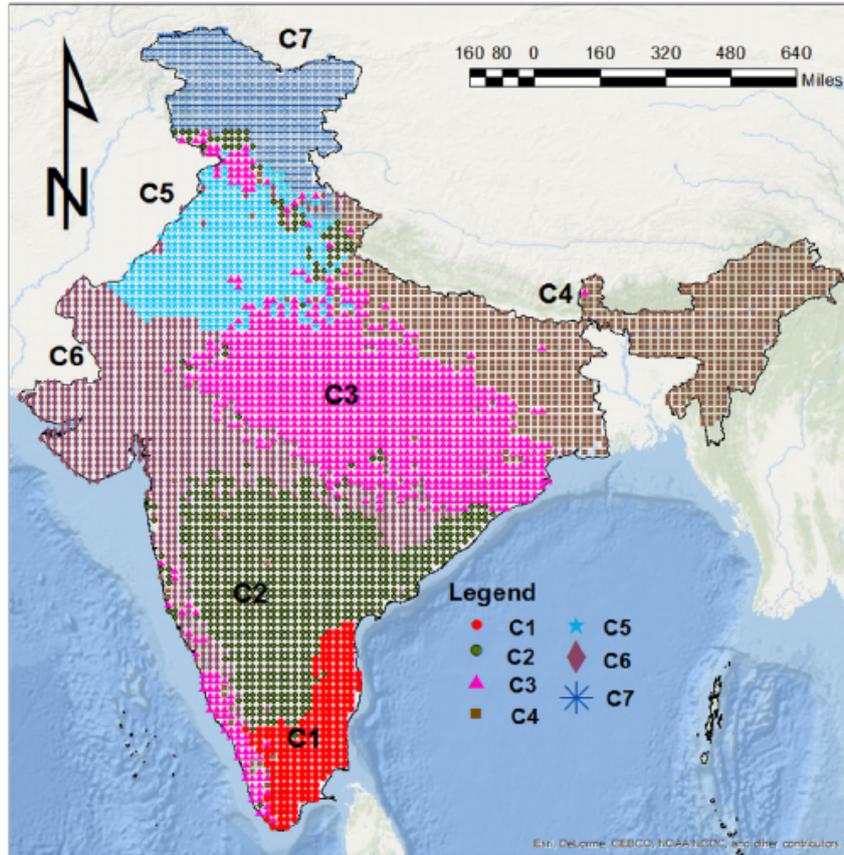
Several studies have focused on identifying groups of regions with similar patterns in the variable of interest using the concept of modularity. These studies have revealed the existence of communities and how they exhibit specific characteristics. The analysis of modularity allows for the grouping of regions that share similar behaviors, such as temperature patterns (STEINHAEUSER et al., 2009; TIRABASSI; MASOLLER, 2016), precipitation (AGARWAL et al., 2019; CERON et al., 2019), or other climatic parameters (STEINHAEUSER; TSONIS, 2014). These climate communities can encompass geographically close or distant areas that exhibit similar climate interactions.

The identification of these communities has been of great importance in understanding climatic processes and contributes to improving the modeling and prediction of weather events and phenomena in different regions of the world.

For example, the study conducted by [Steinhaeuser et al. \(2011\)](#) through the identification of communities investigated the structure of the climate network using sea surface temperature, sea level pressure, geopotential height, precipitable water, horizontal wind speed, and vertical wind speed. The results demonstrated that the identified communities were more effective as predictors compared to clusters obtained through traditional approaches. This finding suggests that changes in the distribution of communities can be used as a descriptive and predictive tool in the field of climatology. The authors emphasized the importance of communities in understanding climate patterns and developing more accurate models for climate predictions. [Tsonis et al. \(2011\)](#) explored the structure of communities in three climatic variables worldwide, comparing observed data with models. They identified a small number of relevant communities (between four and five) for these variables, as well as observed differences in the number of communities between the observed data and the models. This information is important for identifying regions where the models fail to provide an adequate representation. Another relevant study conducted by [Conticello et al. \(2018\)](#) investigated the relationship between intense precipitation events and atmospheric circulation in the central region of Italy, specifically in *Lazio*. They employed a combined approach of Self-Organizing Maps (SOM) and event synchronization to recognize atmospheric patterns that influence precipitation systems. Additionally, the authors used modularity to identify communities within the network of intense precipitation events, resulting in the definition of three rain gauge communities. This analysis revealed a clear representation of the region of interest, providing valuable insights for identifying large-scale atmospheric mechanisms. Meanwhile, [Agarwal et al. \(2018\)](#) demonstrated the importance of vertex participation within their respective communities in a rain network in India (Figure 3.11). They used the modularity criterion ([NEWMAN; GIRVAN, 2004](#)) to identify the communities, emphasizing the relevance of this measure in the context of the study.

The authors provided information about the connectivity among these regions identified as communities. It was concluded that the northern part of India exhibits disconnection from other communities, while the southern part shows a high degree of internal and external connections, and central India reveals long-range connections.

Figure 3.11 - Community structures in a rain network in India.



SOURCE: Agarwal et al. (2018).

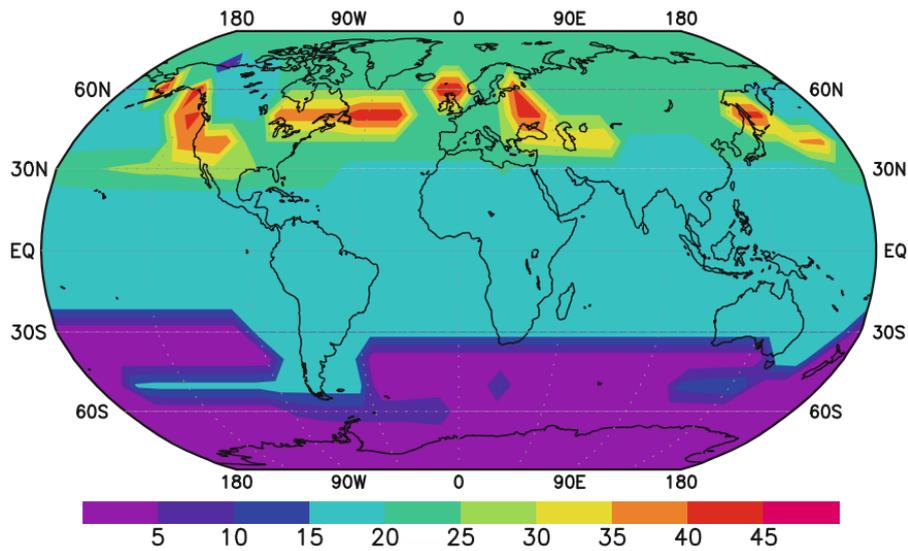
These analyses allowed for an understanding of the interconnection structure among the regions and provided information about the mechanisms or subsystems present, as well as the participation of nodes within these communities.

Complex networks have the ability to extract relevant information about the dynamics of complex systems, making them a valuable tool in various studies related to climatic phenomena. One such climatic phenomenon of interest is the monsoon system, a notable example of which is the Indian Summer Monsoon (ISM), which primarily occurs between June and September. During this period, the ISM experiences intense rainfall activity that plays a crucial role in agriculture, the economy, and the well-being of the population.

Therefore, different studies using complex networks have been conducted with the aim of extracting information associated with the dynamics of the ISM (MALIK et al., 2012; REHFELD et al., 2013; AGARWAL et al., 2018; GUPTA et al., 2023).

Tsonis et al. (2011) present an analysis of communities for the geopotential height 500 hPa network (Figure 3.12), and the identified communities are consistent with the presence of atmospheric teleconnection patterns for the Northern Hemisphere, including the North Atlantic Oscillation, as well as atmospheric teleconnections in the Southern Hemisphere.

Figure 3.12 - Climate network communities of geopotential height (*hgt*) em 500 *hpa*.



The colors represent the communities, which are regions associated with the Northern Hemisphere climate indices.

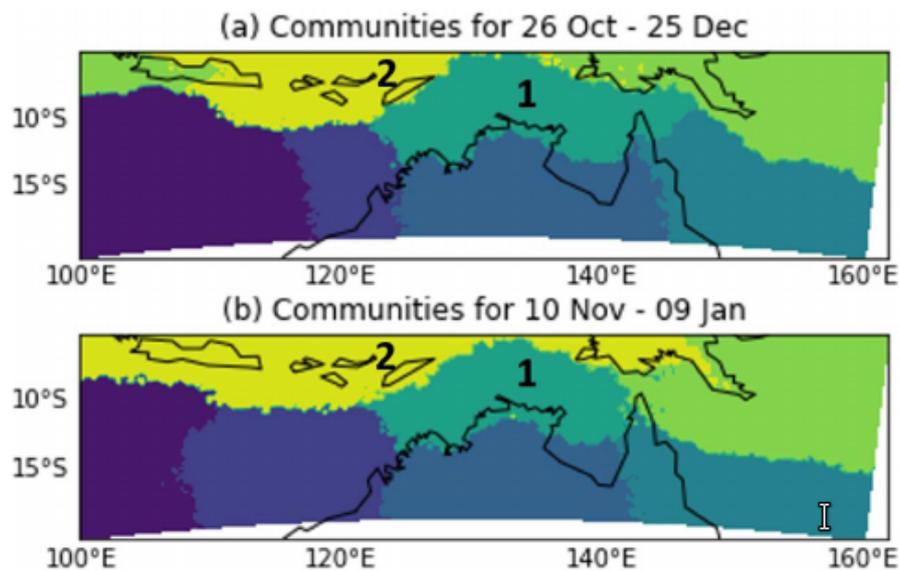
SOURCE: Tsonis et al. (2011).

The study conducted by Saha and Mitra (2019) used the fast-greedy method to identify communities during the ISM period. The fast-greedy method is a hierarchical agglomeration technique that aims to optimize the modularity of the network, meaning it identifies clusters of nodes that have a higher density of internal connections compared to connections with other clusters. The main motivation of this study was to identify community patterns that could serve as predictors and contribute to the prediction of ISM behavior.

By identifying and analyzing these communities, the researchers were able to obtain valuable information about the dynamics of the monsoon system and understand how different regions and atmospheric components interact during the monsoon period. They highlighted how the accurate interpretation of the identified communities allowed them to develop predictors that are capable of generating forecasts about the development of the ISM.

The study conducted by Cheung and Ozturk (2020) investigated the Australian Summer Monsoon (ASM) through the application of complex networks and synchronization of extreme rainfall events. By utilizing complex networks, the researchers were able to analyze the structure of communities and identify patterns related to the stages of the monsoon. Figure 3.13 presents the changes in the structure of communities throughout the ASM. One of the identified communities in the network coincided with the monsoon region in northern Australia. Another community to the north extended during the early monsoon, indicating a high synchronization of this region with the Indonesian region. Subsequently, this community extended eastward during the onset of the monsoon. These results reveal that the processes occurring in these regions are critical to the development of the ASM.

Figure 3.13 - Change in community structures before and during the Australian summer monsoon.



SOURCE: Cheung and Ozturk (2020).

The analysis of changes in the structure of communities allowed for understanding the interactions between regions and patterns during the ASM. This approach provides valuable insights into the mechanisms and factors that influence the development of the Australian summer monsoon, contributing to the understanding of climatic processes. In summary, the study demonstrated how the application of complex networks and the analysis of communities can help identify patterns and processes crucial to the development of the Australian summer monsoon, providing useful information for understanding and predicting this climatic phenomenon.

The mentioned studies illustrate the application of community identification in complex networks and their relevance in understanding meteorological systems. Furthermore, they provide the necessary information to understand how different regions are connected, how nodes interact, and how climatic mechanisms operate at a regional scale. These findings have significant implications for the prediction and understanding of climate patterns, enabling a better assessment of local climate systems and their interactions with the global environment.

As noted by the cited works, the high versatility of complex networks allows for their application in meteorology through different approaches. These networks have proven effective in identifying various relationships existing among physical mechanisms in the atmosphere, providing key insights into the system under study. Such information can include patterns of interconnection between different regions, identification of communities with similar behavior, and analysis of how information is transmitted within the system. Therefore, complex networks have become a valuable tool for understanding and analyzing complex meteorological phenomena.

It is interesting to note that, although there have been several studies applying complex networks on the SAMS (BOERS et al., 2013; BOERS et al., 2014a; BOERS et al., 2014b; BOERS et al., 2015; GELBRECHT et al., 2021), no previous studies had been found that applied community identification to extreme rainfall events related to the SAMS. However, this particular study fills this gap and provides important information about the dynamics along with the identified communities.

4 DATA

In regions with low surface sensor density, the use of satellite data is essential. Optical sensors on satellites capture images at different wavelengths of the electromagnetic spectrum, with each wavelength associated with the target signal and corresponding radiative processes such as emission, absorption, and scattering of the emitted energy by the source. Typically, different sensors integrated into meteorological satellites read specific bands of the electromagnetic spectrum to estimate the desired information. Technologies developed for this purpose usually utilize visible (VIS) and infrared (IR) spectral channels, as well as microwave-based techniques such as radars. Satellite data provides a comprehensive and global view of atmospheric conditions, enabling monitoring and analysis of various meteorological phenomena, including precipitation.

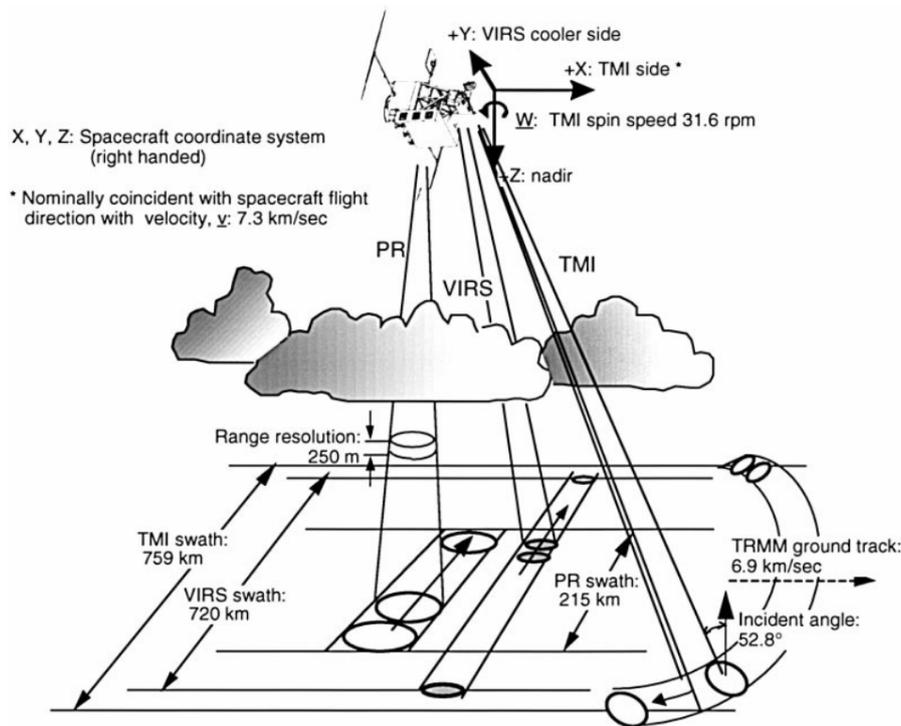
This chapter aims to compile the descriptions of the data that will be used in this research, as they play a fundamental role in the identification and study of precipitation systems. In this chapter, detailed information about the data sources, including their spatial and temporal resolution, will be presented. By providing a good understanding of the data used, this chapter establishes the foundation for the subsequent analyses and interpretations conducted in the study.

4.1 Tropical Rainfall Measuring Mission (TRMM)

TRMM (Tropical Rainfall Measuring Mission) was launched in November 1997 by NASA (National Aeronautics and Space Administration) in partnership with JAXA (Japan Aerospace Exploration Agency) and operated until June 2015, resulting in a measurement period of 17 years. During this time, TRMM played a crucial role in monitoring and understanding the distribution of tropical precipitation. The objective of TRMM was to monitor the temporal and spatial distribution of tropical precipitation (KUMMEROW et al., 1998; KUMMEROW et al., 2000). TRMM flew between the approximate latitudes of 35°N and 35°S and completed 16 orbits per day (PASSOW, 2010). The orbit of TRMM was polar and non-sun-synchronous, with an inclination of 35° and an orbit altitude of 350 km (until 2000) and 403 km (2001-2015).

TRMM is a low Earth orbit satellite and is composed of five instruments (Figure 4.1). Precipitation Radar (PR), Visible and Infrared Scanners (VIRS), Microwave Imager (TMI), Lightning Imaging Sensor (LIS), and Clouds and Earth's Radiant Energy System (CERES) (HUFFMAN et al., 2010; LIU et al., 2012).

Figure 4.1 - Schematic of the scan geometries of the three primary TRMM rain sensors: TMI, PR e VIRS.



SOURCE: Kummerow et al. (1998).

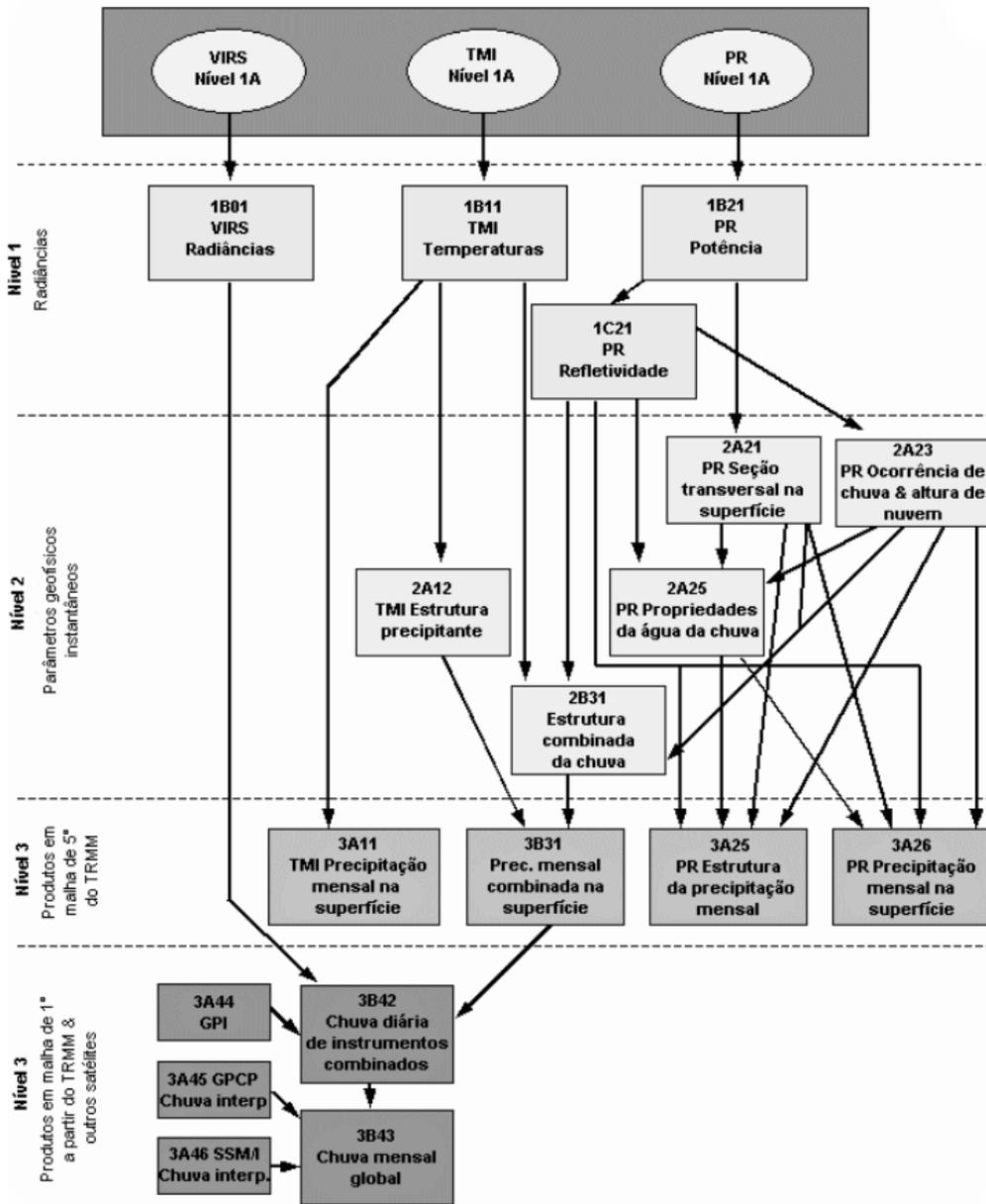
Figure 4.2 provides a more detailed overview of the processing scheme of TRMM products and the data used for their composition, along with their respective levels. All of these data have been compiled into a single repository by the University of Utah (Salt Lake City, Utah, USA). For more information about TRMM, it is recommended to visit the NASA-JAXA project page. The link to the project can be found in Table A.1 of the Appendix.

4.1.1 TRMM Microwave Imager (TMI)

A passive microwave sensor is an instrument designed to receive and measure the natural electromagnetic radiation emissions produced by the Earth and its atmosphere through emission and reflection. The TMI (Microwave Imager) was specifically designed to obtain quantitative information about precipitation. This sensor utilizes measurements of the energy emitted by the Earth and its atmosphere and has the capability to quantify rainfall intensity and water vapor.

The instrument consists of a parabolic antenna with an aperture size of 61 cm and a focal length of 50.8 cm. Based on this, the antenna traces a circle on the Earth's

Figure 4.2 - TRMM data processing scheme.



SOURCE: Collischonn (2006).

surface (Figure 4.1). The antenna has an angle with respect to the nadir of 49° , which provides an incidence angle of 52.8° with the Earth's surface. The TMI is capable of measuring the radiation intensity at five different frequencies: 10.7, 19.4, 21.3, 37, and 85.5 GHz. The 10.7 GHz frequency provides a more linear response to high rainfall rates. The TMI has a swath width of 878 km and different spatial resolutions (Table 4.1).

Table 4.1 - Spatial resolution of TMI frequency bands.

Band [Ghz]	Spatial resolution [km×km]
10.7	72×43
19.4	35×21
21.3	26×21
37	18×10
85.5	8×6

4.1.2 Precipitation radar (PR)

The Precipitation Radar (PR) provides vital information about the distribution and intensity of rainfall, precipitation type, storm depth, and other aspects associated with precipitation (KUMMEROW et al., 2000). The TRMM PR was the first spaceborne precipitation radar designed to provide three-dimensional maps of storm structures. Therefore, it is the most innovative instrument of the TRMM. The PR has an approximate horizontal resolution of 5 km, a swath width of 247 km, and a minimum detectable signal of approximately 15 dBZ. In terms of its ability to provide vertical profiles of rain and snow, the PR can offer profiles with heights of approximately 12 km. Additionally, it has the capability to detect light rain of about 0.7 *mm/h*. One of the main advantages of the PR is its ability to generate high-resolution three-dimensional maps of rainfall. The sensor has a vertical resolution of 250 *m* up to an altitude of 20 km. Beyond that, the data provides vertical information such as layer thickness and precipitation reaching the surface. The PR is capable of capturing precipitation systems larger than 16 km, primarily due to its resolution.

4.1.3 Visible and infrared scanners (VIRS)

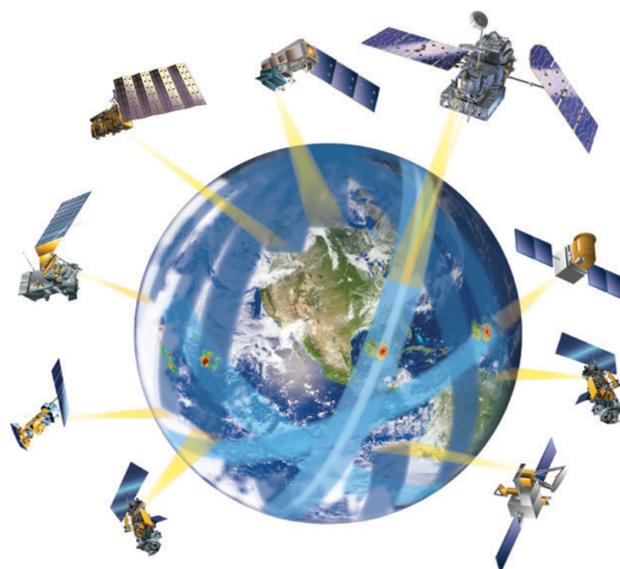
The VIRS is a sensor with the ability to detect radiation in five spectral bands, ranging from visible (0.63 μm) to infrared (12 μm). It has a spatial resolution of 2.4 km and an approximate swath width of 833 km. This sensor can estimate temperature using the received electromagnetic radiation. Under clear sky conditions, the infrared brightness temperature captured by the VIRS corresponds to the temperature of the Earth's surface. However, in the presence of clouds, the measured temperature will be that of the cloud tops. This is because clouds have distinct emission and reflection properties of radiation compared to the Earth's surface, resulting in a different temperature detected by the sensor. Therefore, lower temperatures are primarily associated with frozen surfaces, and higher intensities in shorter wavelength bands

are due to the reflection of sunlight. Meanwhile, higher temperatures will result in higher intensities in longer wavelength bands due to their emission. On the other hand, higher temperatures result in greater radiation emission in longer wavelength bands. Thus, in longer infrared spectral bands, the measured intensity will be higher in regions with higher temperatures because the emitted radiation will be greater. These differences in measured intensities in different spectral bands allow sensors like VIRS to obtain information about surface temperatures and identify features such as the presence of ice, snow, clouds, and other atmospheric elements, aiding in the understanding and monitoring of climatic processes.

4.2 Global Precipitation Measurement Mission (GPM)

GPM is an international network or constellation of satellites (Figure 4.3) aimed at providing observations of precipitation and snow. The GPM satellite constellation can observe precipitation globally every 2 to 3 hours. The central or core satellite, known as the Core Satellite, is considered the successor to TRMM and features two sensors for rain and snow measurements: the GPM Microwave Imager (GMI) and the Dual-frequency Precipitation Radar (DPR). The DPR provides three-dimensional measurements of precipitation structure and characteristics. The instrument consists of a Ka-band precipitation radar (KaPR) operating at a frequency of 35.5 GHz and a Ku-band precipitation radar (KuPR) operating at 13.6 GHz (HOU et al., 2014), the latter being equivalent to the PR of TRMM.

Figure 4.3 - Representation of the satellite constellation of the GPM mission.



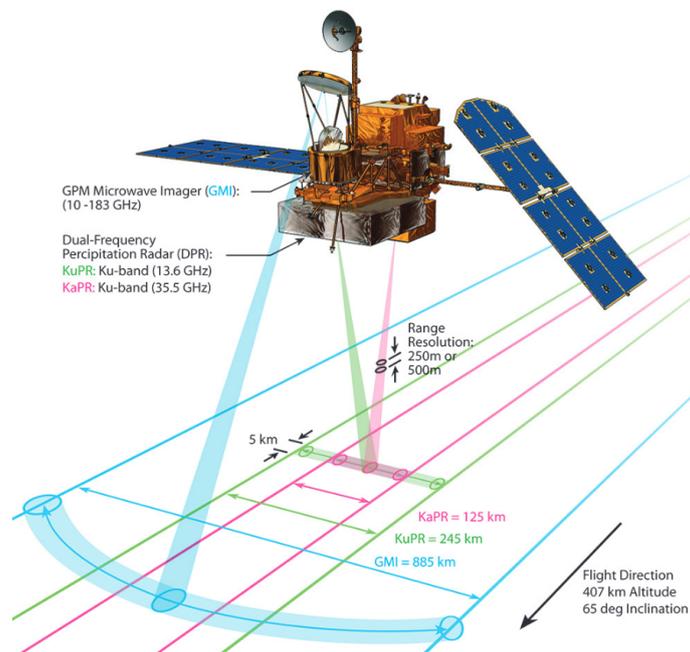
SOURCE: Hou et al. (2014).

The GPM core satellite has an altitude of 407 km and a non-sun-synchronous orbit, providing coverage from 65°N to 65°S with a spatial resolution of 0.1°x0.1° and a north-south inclination of 65°. The satellite was launched in February 2014 from the *Tanegashima* Space Center in Japan, marking the transition from the TRMM mission to the GPM era.

The GPM core satellite is equipped with two central instruments shown in Figure 4.4: the GPM Microwave Imager (GMI) and the Dual-frequency Precipitation Radar (DPR). The GMI is a microwave sensor with 13 channels ranging from 10 GHz to 183 GHz. These channels provide observations of the radiation emitted by different types of precipitation, from light to heavy rain and the presence of snow.

The second instrument, DPR, provides three-dimensional information about precipitation particles. This information is obtained from the energy reflected by these particles at different levels within the clouds. The DPR is more sensitive to light rain rates and the falling of precipitation compared to the PR/TRMM (Section 4.1.2).

Figure 4.4 - Scheme of DPR and GMI instrument scanning patterns onboard the GPM Core.



SOURCE: Hou et al. (2014).

4.2.1 GPM Microwave Imager (GMI)

The GMI is a passive microwave radiometer that features 13 frequency channels ranging from 10 GHz to 183 GHz. These different frequencies have been used in the past decades to detect rainfall of varying intensities, from light rain to heavy precipitation (HOU et al., 2014). Each frequency band is specifically used for different purposes. For example, the 10 GHz band is ideal for liquid precipitation, the bands between 19 GHz and 37 GHz are used to detect moderate rainfall, the 21 GHz band is related to water vapor, and the 183 GHz frequency band is used to detect signals of ice scattering and snowfall. The combination of GMI's different frequency channels allows for a comprehensive analysis of precipitation intensity in various regions of the planet. These data are crucial for continuous monitoring of global precipitation patterns, providing valuable information for climate studies and weather forecasts.

4.2.2 Dual-frequency Precipitation Radar (DPR)

The DPR consists of two precipitation radars: the KuPR (Ku-band Precipitation Radar) and the KaPR (Ka-band Precipitation Radar). The KuPR operates at a frequency of 13.6 GHz and is an enhanced version of the precipitation radar used in TRMM. It is suitable for measuring moderate to heavy rainfall, which is typical in tropical regions and mid-latitudes. On the other hand, the KaPR operates at a frequency of 35.5 GHz and aims to enhance the detection capability of weak rain and snow (HOU et al., 2014). The KaPR (Ka-band Precipitation Radar) has two scan modes: the dual-frequency mode (KaPR-MS) and the high-sensitivity mode (KaPR-HS). These scan modes offer different detection and resolution capabilities to capture various types of precipitation. Both sensors have a spatial resolution of 5 km, and in terms of swath width, the KuPR has 245 km while the KaPR has 125 km. Both radars have a vertical range resolution of 250 m, with a minimum detectable signal of 14.5 dBZ for KuPR, 16.47 dBZ for KaPR-MS, and 10.2 dBZ for KaPR-HS (TOYOSHIMA et al., 2015). The DPR provides three-dimensional observations of rainfall, meaning it is capable of mapping the vertical distribution of precipitation within storms.

This information is crucial for understanding storm characteristics such as intensity, vertical development, and the presence of regions with higher rainfall activity. Additionally, Iguchi et al. (2010) provides further details about the algorithm used to generate DPR data.

4.3 Integrated Multi-satellitE Retrievals for GPM (IMERG)

One of the most widely disseminated products derived from the combination of TRMM/GPM is provided by the IMERG algorithm. This algorithm provides an estimation of precipitation over a large part of the Earth's surface. As mentioned, it combines the information provided by the TRMM mission (2000-2015) and the GPM mission (2014-present), aiming to have a more robust precipitation record over a longer period of time (HOU et al., 2014). Therefore, this algorithm is designed to combine and interpolate estimates from all microwave precipitation satellites, along with infrared (IR) estimates, observed measurements, and other short-term and small-scale precipitation estimators. The IMERG data is available in three levels: "Early", the first multi-satellite product available 4 hours after observation (IMERG-E); "Late", the multi-satellite product available 12 hours after observation (IMERG-L); and "Final", which is released 3 months after the observation time (IMERG-F), as it is calibrated with monthly gauge data. The precipitation rate is represented in millimeters per hour (mm/h) with a spatial resolution of 0.1° and a temporal resolution of 30 minutes, covering an area from 60°N to 60°S . Therefore, IMERG is a powerful tool for obtaining precipitation estimates on a global scale, spanning a longer time period, and providing valuable information for climate, hydrological, and meteorological studies worldwide.

4.4 Precipitation Features (PF)

The PF data provides detailed information about the structure and characteristics of precipitation, serving as an important tool for studying and understanding the rainfall systems and their influence on climate. These data can include a variety of information, such as the area covered by precipitation, the temperature of the cloud tops, the stratiform and convective rainfall contribution, the precipitation depth (indicating the vertical thickness), the precipitation rate (indicating the amount of precipitation over a specific time interval), and the volumetric contribution (NESBITT et al., 2006; LIU et al., 2008). Furthermore, radar reflectivity plays a crucial role in identifying and evaluating rainfall, including its extreme events (ROMATSCHKE et al., 2010; BHAT; KUMAR, 2015; KUMAR; BHAT, 2016; KUMAR; BHAT, 2017; KUMAR; SILVA, 2019). Radar reflectivity data provides valuable insights into the structure and characteristics of precipitation systems, allowing for detailed analyses of rainfall patterns at both regional and vertical scales. By examining radar reflectivity, we can study the spatial distribution, intensity, and variability of rainfall, which is essential for understanding the dynamics of the precipitation system.

The Department of Atmospheric Sciences at the University of Utah (Salt Lake City, Utah, USA) has constructed a database of precipitation characteristics based on measurements from TRMM and GPM. This database provides output variables at three processing levels. The Level 1 data is composed of a combination of TRMM algorithms 1B01, 1B11, 2A12, 2A23, 2A25, 2B31, 2H25, and LIS, along with their corresponding GPM versions. A parallax correction was applied to TMI-PR, and nearest neighbor collocation was used for TMI-PR-LIS-VIRS. The Level 2 product consists of two groups: the old definition based on PR and TMI according to Nesbitt et al. (2000), and the new definition introduced in 2012 by Liu (2007). In addition, Table 4.2, shows other characteristics were aggregated as more information was considered in the calculations, such as latent heat and the 2B31 product, resulting in over 30 variables in Level 2.

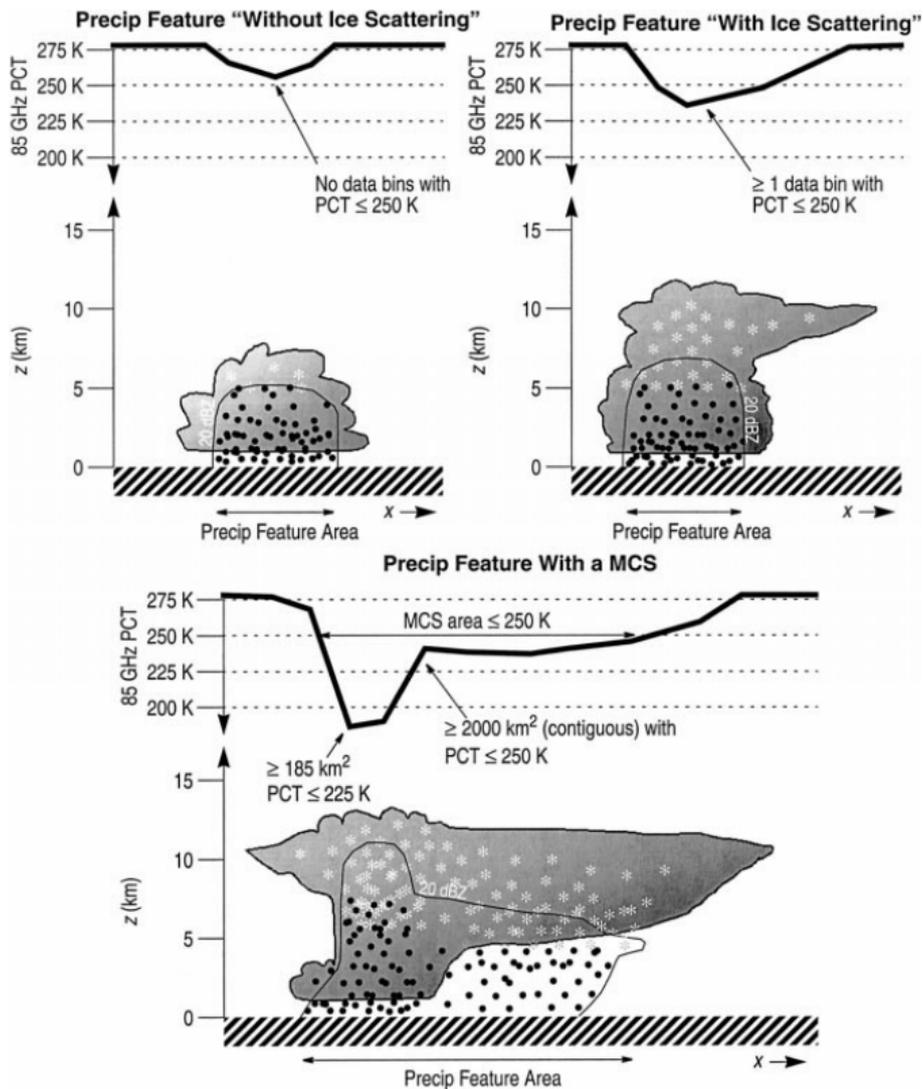
Table 4.2 - New criteria used for defining Precipitation Features (LIU, 2007).

Acronym	Definition	Criteria used
RTPF	Surface precipitation with PR and TMI	2A25 Pixels > 0 or TMI > 0
CLCONVF	Convective cells by convective pixels	Convective Rain Pixels 2A23
CL6KM30F	Convective cells by convective pixels > 30 dBz at 6 km	> 30 dBz at 6 km
CL40PF	Convective cells by convective pixels with 40 dBz column echo	> 40 dBz at column
CL12KM20F	Convective cells with 20 dBz at 12 km	> 20 dBz at 12 km
T200F	85 Ghz and PCT < 200 K	Pixels with pct85 < 200 K
TPCTF	85 Ghz and PCT < 250 K	Pixels with pct85 < 250 K

Finally, Level 3 was developed with the objective of creating climatologies, including convective precipitation, among others. This was done within an 80×360 grid. For the generation of Level 3 data, reanalysis data (KISTLER et al., 2001) were used, specifically the following variables: Temperature profile (K), Geopotential height (m), Omega (Pa/s), wind components (m/s), relative humidity at 8 levels from the surface (%), temperature (K), tropopause pressure (hPa), and precipitable water (kg/m²). More information can be found on the PF tab of the University of Utah’s website (LIU, 2007).

Figure 4.5 shows the scheme used by Nesbitt et al. (2000) to define three types of PF (Precipitation Features). This scheme is primarily based on top temperature data, reflectivity, area, and cold core height. Using these parameters, it is possible to identify the presence of these PFs in a specific region. Therefore, PFs are clusters of certain physical characteristics (associated with precipitation) with pre-defined boundaries that reflect the development of convective systems.

Figure 4.5 - Vertical diagram of three types of PF.



Where ice and hydrometeors are represented by asterisks and dots respectively.
SOURCE: Nesbitt et al. (2000).

4.4.1 Vertical Reflectivity Profile

Radar reflectivity plays a crucial role in identifying and assessing the intensity and physical characteristics of precipitation systems. Reflectivity is obtained from the radar signal's return when it encounters precipitation particles in the atmosphere. This reflectivity provides valuable information about the precipitation rate and the size of the precipitation particles.

In precipitation systems, areas of high reflectivity on radar often indicate extreme rainfall events. The Precipitation Features data provides information on radar reflectivity at 40 altitude levels, ranging from 0.5 km to 20 km, with a resolution of 0.1 km. This data allows for the creation of vertical reflectivity profiles, which are highly valuable for conducting detailed analyses and characterizing the vertical structure of precipitation.

By examining the reflectivity profiles, we can observe the distinct characteristics present in different precipitating systems. This information can be used to extract physical characteristics specific to each system, providing valuable insights into their behavior and dynamics. Analyzing the reflectivity profiles enhances our understanding of precipitation processes and aids in the interpretation and analysis of precipitation patterns and phenomena.

Different studies employ reflectivity profiles to investigate rainfall patterns and analyze extreme rainfall events, aiming to understand their intensity and behavior in the atmosphere (PETERSEN; RUTLEDGE, 1998; NESBITT et al., 2000; NESBITT et al., 2006; BOCCIPPIO et al., 2002; CAREY et al., 2003; JACOBSON; HEAVNER, 2005; LIU et al., 2007). When studying rainfall events, reflectivity profiles help in identifying and characterizing different precipitation types, such as convective or stratiform, and understanding their temporal and spatial variations. Reflectivity profiles are particularly useful for assessing the intensity and severity of precipitation. By utilizing reflectivity profiles, we can improve the information about rainfall patterns, the vertical structure of precipitation systems, and the characteristics of extreme rainfall events.

5 METHODS

In this chapter, the methods used to develop the study will be presented. It will start with the presentation of the study area and the steps for constructing complex networks. This will involve defining the nodes and edges of the network based on specific criteria, such as precipitation intensity. The analysis methods to identify the patterns associated with network metrics will be discussed, as well as the identification of communities. Furthermore, the characterization of the communities will be addressed, based on the physical and dynamic properties associated with extreme rainfall.

5.1 Study region

In this study, we will create a total of five networks to be analyzed. The first network, referred to as the SOUTH AMERICA Network (SA), will exclusively consist of extreme daily rainfall time series during the austral summer season in South America. The primary objective is to examine the current dynamics of extreme precipitation events, identify communities within this network, and provide a comprehensive characterization of these communities from both physical and statistical perspectives. Additionally, we will explore the relationship between these communities and the meteorological mean aspects, aiming to gain a deeper understanding of the patterns and processes associated with extreme rainfall in the region.

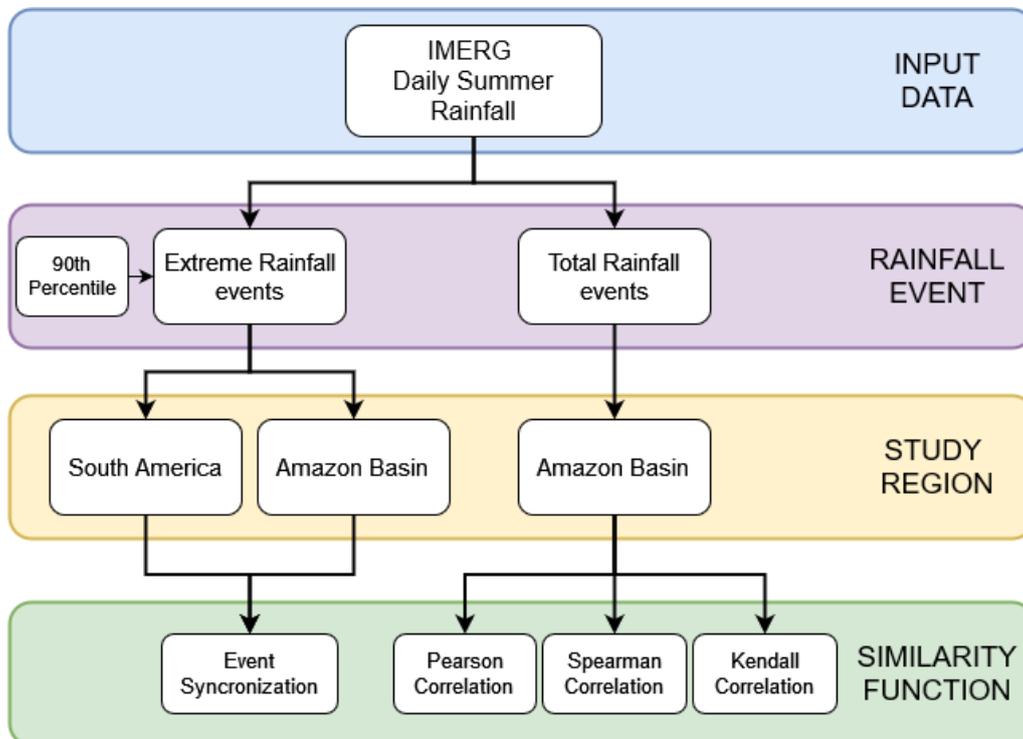
In order to investigate certain aspects of complex networks on a more local scale and the relationship between rainfall and the surface, more detailed analyses were conducted on the Amazon basin. Therefore, we will compare networks based on daily rainfall and extreme daily rainfall by creating four additional networks known as the AMAZON Networks (AMA). These networks are specifically designed for the Amazon Basin region, which is a significant and geographically diverse area within the continent. The main focus of these networks is to examine the similarities and differences between the two types of rainfall data (i.e. Extreme and total rain). The key distinction between these networks lies in the method used for the synchronization of extreme rainfall events and the correlation metrics (i.e. Pearson, Spearman, and Kendall) employed for daily rainfall analysis. By conducting this comparison, we aim to gain valuable insights into the behavior and characteristics of the network. Furthermore, it will be possible to assess the degree of sensitivity in defining communities based on correlation networks for this region. Table 5.1 presents a summary of all the network information.

Table 5.1 - Characteristics of the networks used in this study.

Network	Region	Event type	Similarity function
SA AMA 1	South America Amazon basin	Daily extreme rainfall	Event Synchronization
AMA 2 AMA 3 AMA 4	Amazon basin	Daily rainfall	Pearson linear correlation Spearman's rank correlation Kendall's rank correlation

The choice of the appropriate similarity function plays a fundamental role in analyzing climatic patterns in the network (Table 5.1). For this purpose, two groups of networks have been created (Figure 5.1). The first group was created using only extreme rainfall data for South America and the Amazon basin, while the second group will be a network of whole rainfall events only for the Amazon basin. For both networks, the study period is from December 2000 to February 2021.

Figure 5.1 - Diagram of the two groups of networks that were used.



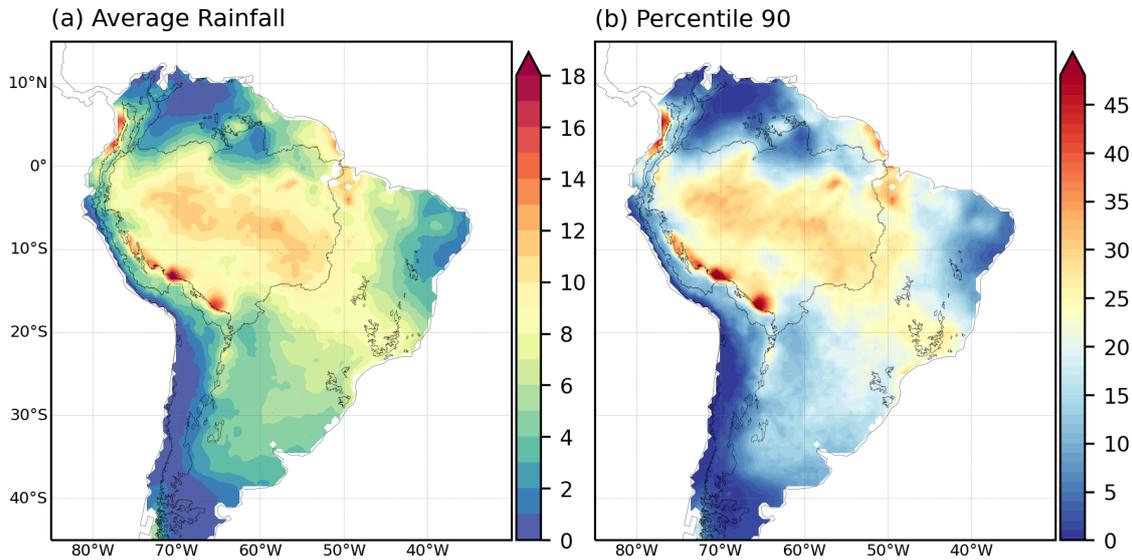
SOURCE: Author's production.

5.1.1 South America network

South America possesses a range of important atmospheric mechanisms and is characterized by significant geographical diversity. Several factors contribute to this diversity, including the Andes mountain range, which exerts a substantial influence on the continent's circulation patterns (INSEL et al., 2010). The Amazon basin is the largest tropical forest and water system in the world, playing a crucial role in the regional climate (ALMAZROUI et al., 2021). The *Sierras de Córdoba* region is known for its deep convective cores, making it an area of interest for convective studies (MULHOLLAND et al., 2018). The Brazilian semi-arid region is notable as the most populous semi-arid area worldwide (CUNHA et al., 2015; MENEZES et al., 2021). The Pisco-Ica desert experiences local dust storms known as *Paracas* (BRICEÑO-ZULUAGA et al., 2017). Furthermore, the Atacama Desert is renowned as the driest desert on Earth, characterized by extremely arid conditions (BULL et al., 2018). This geographical diversity in South America contributes in various ways to the complex climate variability and the occurrence of extreme weather events. The continent exhibits some notable features, such as being the second continental region with the presence of the largest mesoscale convective systems, leading to events of extreme precipitation (MOHR; ZIPSER, 1996; KUMAR et al., 2019a). Additionally, South America is known for hosting some of the deepest convective storms globally (ZIPSER et al., 2006; RASMUSSEN et al., 2014) and having the highest frequency of large hail (CECIL; BLANKENSHIP, 2012). These factors highlight the significance of South America in the study of extreme rainfall events. Based on the presented characteristics and the dynamics involving the entire region, South America stands out as an area with a great variety of active atmospheric systems. Figure 5.2 shows the exact extent of the region used for the SOUTH AMERICA network, which is defined by the box with latitudinal limits of 12°N - 45°S and longitudinal limits of 85°W - 31°W.

This figure also shows the mean and the 90th percentile of rainfall for the months corresponding to the Austral summer. It can be observed that the majority of the Amazon basin is presented as the region with the largest extension in terms of rainfall intensity. The two intense rainfall cores over the eastern Andes indicate the so-called Peruvian and Bolivian rainfall hotspots, which are associated with orographic precipitation (ESPINOZA et al., 2015).

Figure 5.2 - Average daily rainfall and 90th percentile (mm/day) over South America during the austral summer.



The black lines represent the topography and the Amazon basin limits.
SOURCE: Author's production.

5.1.2 Amazon network

The Amazon Basin is known as the largest tropical forest and hydrological system in the world (ALMAZROUI et al., 2021). It plays a crucial role in mitigating climate change by reducing the concentration of carbon dioxide in the atmosphere. Additionally, it is renowned for being an indispensable source of heat exchange with the atmosphere, particularly due to the high rate of evapotranspiration in the mid and upper troposphere. This characteristic creates favorable conditions for the formation of tropical and extratropical convective systems, playing a crucial role in regional circulation (FISCH et al., 1998; HAYLOCK et al., 2006; NOBRE et al., 2009). The Amazon Basin also exhibits similarities with oceanic convection (RASMUSSEN et al., 2016) and holds significant importance in various aspects, such as global/regional climate regulation, modulation of the hydrological cycle, and water purification (MARENGO, 2006; ALVES et al., 2017; PHILLIPS; BRIENEN, 2017). The study of rainfall in the Amazon is critical for understanding the functioning of the tropical forest ecosystem as a whole and its contribution to global climate regulation. The Amazon basin is the region highlighted in Figure 5.2.

The construction of the complex network in this study will be limited to the defined study areas, focusing solely on the specific region of interest. As a result, global

climate influences, known as teleconnections, will not be addressed in this research. The upcoming section will center on defining the networks of rainfall and extreme rainfall events, which is a crucial component of the analysis conducted in this study.

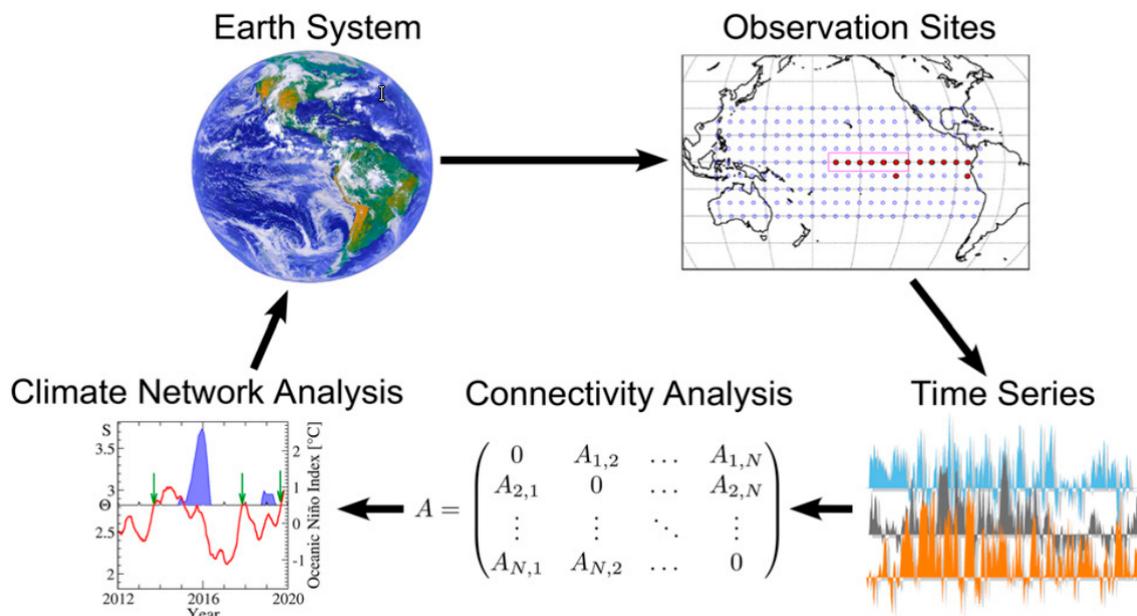
5.2 Climate networks

Complex networks have broad applicability. However, one of the main advantages of this technique is its ability to extract meaningful information from the system, which would often require more complex methodologies or the use of a set of variables together. An example of this potential is evident in the studies of atmospheric teleconnections associated with the El Niño phenomenon, which were among the earliest applications of complex networks to climate data (TSONIS; ROEBBER, 2004). This approach has allowed for a deeper understanding of the complex interactions between different regions of the global climate system. Due to their ability to describe how information propagates and relates within a system, complex networks have gained significant popularity in various fields of study (STROGATZ, 2001; COSTA et al., 2011; SALEH et al., 2018). This approach allows to identification of regions or notable elements within the system and provides information about the relationship or predominance of these in the dynamics of the system as a whole.

The methodology of climate networks is used for the creation of the network (TSONIS; ROEBBER, 2004; TSONIS et al., 2006; TSONIS et al., 2007; TSONIS et al., 2008; YAMASAKI et al., 2008; GOZOLCHIANI et al., 2008). In this approach, each vertex of the network is represented by a pair of latitude and longitude coordinates corresponding to each point on the spatial grid. This means that each vertex is defined at a specific geographic location, as illustrated in Figure 5.3. Therefore, the total number of vertices will be equal to the number of points in the grid ($N = \text{lat} \times \text{lon}$). Each vertex in the climate network contains a time series associated with a specific meteorological variable. This variable can be precipitation, atmospheric pressure, temperature, and humidity, among others. Each time series represents the measurements of the variable at a specific geographical point over time.

While the connections between the vertices, called edges, can be defined using different similarity functions, which may vary depending on the analysis objectives and the nature of the data. There have been different studies on climate networks using a variety of functions to define the edges. For example, Pearson correlation coefficient (NAUFAN et al., 2018), cross-correlation (TSO; YAU, 2007; GOZOLCHIANI et al., 2008), Spearman's rank correlation (BOERS et al., 2015), Kendall's rank correla-

Figure 5.3 - Schematic representation of the creation of a climate network.



SOURCE: Ludescher et al. (2021).

tion (GUPTA et al., 2021), mutual information (MARWAN et al., 2009), and event synchronization (BOERS et al., 2014b). Each of these similarity functions offers a unique perspective on the relationships between vertices in the climate network and can provide relevant information about climate dynamics and patterns of interaction between different geographical points.

Finally, in this study, undirected connections are utilized, indicating that the connection between two vertices does not possess a specific direction. This approach enables the exploration of relationships between geographical points and the analysis of climatic patterns associated with network behavior, without requiring specific information about the points of origin and destination. It is important to note that self-loops are excluded in this research.

Following the criterion of climatic networks, the next step is to define the data that will be used to create the network. Since the objective is to study the system of extreme precipitation events in South America, the variable of interest is rainfall. Therefore, it is also essential to choose a similarity function that captures the relevant relationships between geographical points based on rainfall characteristics.

5.2.1 Extreme rainfall networks

The analysis and understanding of extreme rainfall events are of vital importance due to the negative impacts they can have on society. To investigate the occurrence of these events, the networks in this study utilize data from extreme rainfall events. The objective is to understand the patterns of occurrence and the inter/intra connections between communities. This research aims to contribute to the prediction and mitigation of these events, as well as their associated impacts, defining areas with similar behavior.

5.2.1.1 Define extreme rainfall events

When using precipitation data, it is necessary to establish a specific rainfall threshold for identifying extreme events. In this work, the extreme daily rainfall events are identified by using the 90th percentile (P90) as a threshold, calculated for each grid point. This approach is commonly employed in various studies (BOERS et al., 2014b; ESPINOZA et al., 2015; ZUBIETA et al., 2019; BOERS et al., 2019). Thus, the 90th percentile represents a limit above which the events are considered extreme.

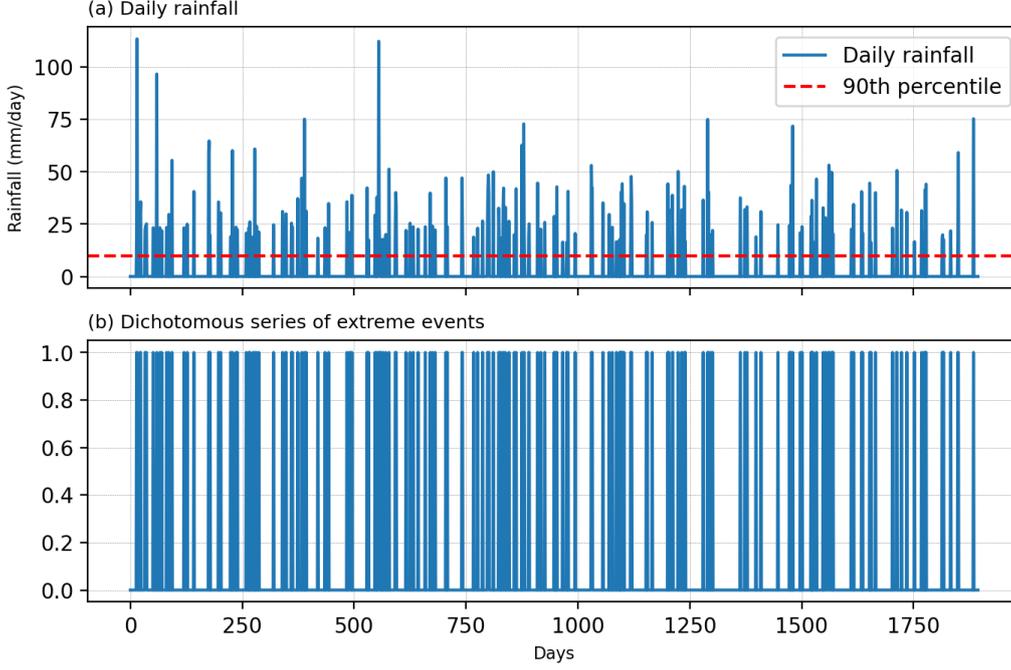
Therefore, extreme events will be determined by considering the top 10% of the most intense rainfall events. This approach allows us to focus on the most significant occurrences in terms of rainfall magnitude. The procedure for identifying extreme events in a rainfall series begins by obtaining the value of the P90. Then, all rainfall events with magnitudes lower than the P90 value are classified as zeros, indicating the absence of extreme events. On the other hand, rainfall events with magnitudes equal to or greater than the P90 value are classified as ones, that represent the occurrence of extreme events. Figure 5.4 illustrates the process of converting a rainfall series into an extreme rainfall series.

In this way, a dichotomous (or binary) series is created that allows us to distinguish the presence of extreme rainfall events from the rest. This series of extreme events is used as an input condition for the similarity function.

5.2.1.2 Event Synchronization

For the extreme rainfall networks (South America and Amazon basin), the links are defined using Event Synchronization (ES) (QUIROGA et al., 2002; MALIK et al., 2012; BOERS et al., 2019). ES is a technique that measures the similarity between time series by identifying patterns of event synchronization or co-occurrence. By using ES as the similarity function, the network will be able to capture the spatial

Figure 5.4 - Example of the process to define extreme events using the 90th percentile of a daily rainfall series for a specific point on the grid.



Where the red line represent the 90th percentile threshold.

SOURCE: Author's production.

patterns associated with extreme events and identify regions that exhibit similar patterns of occurrence for these events. One of the advantages of this metric is that it is designed to calculate nonlinear relationships between time series of defined events (AGARWAL et al., 2018). Event synchronization is defined based on two series of extreme events, denoted as e_i and e_j , occurring at grid points i and j , respectively. A specific extreme event in time i and j are, represented by l_i and l_j . Consider two extreme events, denoted as e_i^u and e_j^v , where $0 \leq u \leq l_i$ and $0 \leq v \leq l_j$. To check if two events can be uniquely linked to each other ($d_{i,j}^{u,v} := e_i^u - e_j^v$), a dynamic time lag is defined using:

$$\tau = \frac{\min\{d_{ii}^{u,u-1}, d_{ii}^{u,u+1}, d_{jj}^{v,v-1}, d_{jj}^{v,v+1}\}}{2} \quad (5.1)$$

A filter is defined using a minimum (τ_{min}) and maximum (τ_{max}) time lag, which allows us to analyze processes at different time scales. Thus, synchronization can be defined as:

$$S_{ij}^{u,v} = \begin{cases} 1 & 0 < d_{ij}^{u,v} \leq \tau \\ 0.5 & \tau_{min} \leq \tau \leq \tau_{max} \\ 0 & \text{otherwise} \end{cases} \quad (5.2)$$

The synchronization of events directed from e_j to e_i is defined by a normalized sum:

$$q_{ij} = \frac{\sum_{u,v} S_{ij}^{u,v}}{\sqrt{l_i l_j}} \quad (5.3)$$

Finally, undirected event synchronization is determined by summing the directed event synchronization between the two vertices, given by: $ES_{ij} = q_{ij} + q_{ji}$.

5.2.2 Rainfall networks

The rainfall networks will be constructed using precipitation time series as vertices. These time series will be obtained from the highlighted region shown in Figure 5.2. The objective is to verify the sensibility of the complex network and the additional information shown for a specific region (Amazon Basin) in comparison with the entire region (South America). To achieve this, we will apply three distinct similarity functions to analyze the interconnections and patterns of rainfall events, resulting in the creation of three separate networks. In the following sections, we will provide a detailed presentation of these functions.

5.2.2.1 Pearson correlation coefficient

Also known simply as Pearson correlation, it is a statistical measure that evaluates the linear relationship between two continuous variables and is widely used to assess the strength and direction of this relationship. Pearson correlation is defined as follows:

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2} \sqrt{\sum(Y_i - \bar{Y})^2}} \quad (5.4)$$

Where X_i and Y_i are the individual values of variables X and Y , respectively, while \bar{X} e \bar{Y} are the means of each series. It should be noted that Pearson correlation is a parametric technique. This means that it assumes a specific distribution for the variables involved, in this case, a normal distribution. Applying this correlation to data that does not follow a normal distribution can compromise the interpretation and introduce artificial bias into the results.

Rainfall often does not follow a normal distribution, which can lead to an inadequate interpretation of Pearson linear correlation when applied directly to this data. To address this issue, the methodology proposed by Ciemer et al. (2018) was used. It allows for reducing artificial bias and obtaining a more accurate estimate of the Pearson correlation between variables such as precipitation. It is important to note that this specific methodology can be a useful approach for dealing with the non-normal distribution of rainfall when applying Pearson linear correlation.

5.2.2.2 Spearman's rank correlation

Also known as Spearman's correlation or Spearman's rho, it is a non-parametric statistical measure for assessing the linear relationship between two variables. Spearman's correlation considers the monotonic relationship between variables, meaning that one variable consistently increases or decreases with the other regardless of the shape of the relationship. Similar to Pearson's correlation, Spearman's correlation ranges from -1 to 1. Spearman's correlation is defined as follows:

$$r_s = 1 - \frac{\sum_{i=1}^n (a^i - b^i)^2}{n(n^2 - 1)} \quad (5.5)$$

Where a^i and b^i are the ranks of x e y respectively, while n is the number of observations. This correlation is based on the ranking of the variables, as each variable is individually ranked and then the differences between these rankings are calculated to capture non-linear relationships. Spearman's correlation is a useful alternative when the data does not follow a normal distribution or when the relationship between variables is not linear.

5.2.2.3 Kendall's rank correlation

Kendall's correlation is also a non-parametric statistical measure used to assess the relationship between two ordinal variables. Similar to Spearman's correlation,

Kendall's correlation is based on the ranks of the observations for each variable. This correlation measures the agreement between the rankings of the two variables, indicating whether the observations tend to have a consistent order between the variables. Kendall's correlation does not assume any specific functional relationship between the variables, making it suitable for capturing non-linear and monotonic relationships. Kendall's correlation is computed based on the number of concordant and discordant pairs, which are defined according to the following conditions. A pair is considered concordant if $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$. A pair is considered discordant if $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$. Otherwise, the pair is neither concordant nor discordant. Kendall's correlation coefficient is defined as:

$$\tau_k = \frac{NC - ND}{n(n-1)/2} \quad (5.6)$$

where NC and ND are the numbers of concordant and discordant pairs respectively, while n is the number of observations. Like Pearson and Spearman, Kendall's correlation ranges from -1 to 1, where +1 indicates a perfect concordance between the rankings, -1 indicates a perfect discordance, and 0 represents no concordance.

5.3 Network analysis

After defining the similarity functions to be used in the networks, the methods for their creation and analysis are presented. This chapter discusses the criteria used to obtain the adjacency matrix and then describes the main metrics for analysis.

5.3.1 Adjacency matrix

Once the similarity functions for the rainfall and extreme events networks have been defined (sections 5.2.2 and 5.2.1), the next step is to calculate the connections between all pairs of vertices, resulting in an $N \times N$ connectivity matrix. From the statistical distribution of all existing connections, ES values greater than 98% were identified to verify statistical significance (0.02) (BOERS et al., 2013). This allows us to create an adjacency matrix based on this threshold. Values equal to or above this threshold are classified as 1, while values below are classified as 0. This approach allows us to focus on the most relevant connections in the network. Thus, the resulting network contains only the strongest connections, which contributes to identifying patterns and structures present in rainfall and extreme rainfall events.

5.3.2 Degree, distribution, and density

The degree of a vertex indicates the number of connections that the vertex v_i has. This index is useful for evaluating the connectivity of vertices in the network. The degree can be obtained from the adjacency matrix as follows:

$$k_i = \sum_{j=1}^N A_{ij} \quad (5.7)$$

Where j represents the vertices v_j that are connected to vertex v_i , and N is the total number of vertices in the network. In unweighted networks, each edge or connection in the adjacency matrix is represented by the value 1, indicating the existence of a connection between two nodes. The degree of a node is the number of edges connected to it, i.e., the number of neighbors it has in the network. For example, if a node has 5 connected neighbors, it will have a degree equal to 5. The degree of a node is an individual measure as it describes the specific connectivity of that node in the network.

The average degree, or mean degree, provides an estimate of the average connectivity of nodes in the network, indicating how many connections, on average, each node has. The average degree is useful for comparing the connectivity of different networks and understanding the overall distribution of connections in the network. It is calculated as the average of the degrees of all nodes in the network and can be defined as follows:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i \quad (5.8)$$

The degree distribution of a network, denoted as $P(k)$, is a statistical measure that describes the frequency at which different degrees are present. The degree distribution of a network can take different forms depending on the network's structure. In random networks, the degree distribution follows a Poisson distribution, where most nodes have an average degree and the number of nodes with higher degrees decreases rapidly. In scale-free networks, on the other hand, the degree distribution follows a power-law distribution, where some nodes have extremely high degrees while the majority of nodes have low degrees.

The link density of a network is a measure that describes the proportion of existing links in relation to the maximum possible number of links. It indicates how densely interconnected the vertices are in the network. Link density ranges from 0 to 1, where 0 represents a completely disconnected or fragmented network, and 1 indicates a fully connected network. To calculate the network's link density, it is necessary to first obtain the total number of possible links. For a network with N vertices, the maximum possible number of connections is $N(N - 1)/2$ (for an undirected network). The link density of a network can be expressed as:

$$\rho = \frac{m}{N(N - 1)/2} \quad (5.9)$$

5.3.3 Degree Centrality (DC)

The degree centrality (DC) is a measure that indicates the importance of a node based on the number of links it has (SRINIVASAN et al., 2020). In other words, DC measures how central or influential a node is within the network, taking into account its degree, which is the number of links connected to it. The higher the DC value of a node, the more important it is in terms of connections within the network. A node with a high DC has a large number of links, indicating that it plays a crucial role in communication and information transfer within the network. DC is obtained by dividing the number of edges of a node by the maximum possible number of edges the node could have. In other words, it is the ratio between the node's degree and the maximum possible degree the node could have.

$$DC = \frac{1}{L - 1} \sum_j^N A_{ij} \quad (5.10)$$

Where L represents the total of vertices in the network, N is the total number of nodes and A_{ij} is the ij -th element of the adjacency matrix A .

5.3.4 Average Neighbor Degree (AND)

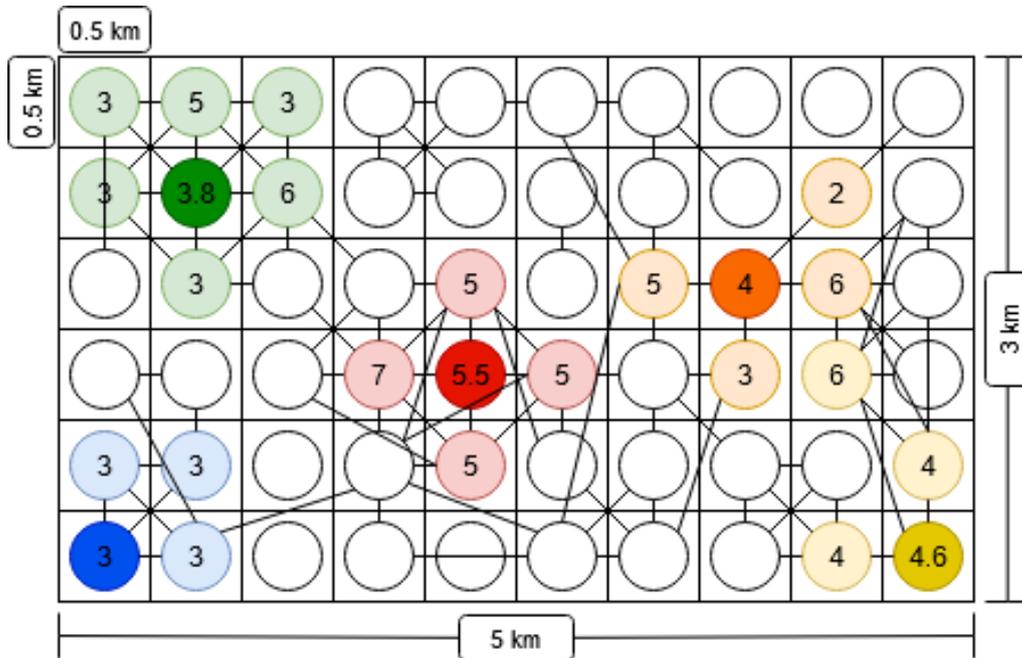
The average neighbor degree (AND) is a measure that indicates the average degree of the neighbors of a given vertex (BARRAT et al., 2004). This measure provides information about the connectivity of a node's neighbors and helps to understand the neighborhood structure (Figure 5.5). The AND of a node is calculated by finding

all the neighbors of that node in the network and then calculating the average degree of those neighbors. This measure gives an idea of how the neighboring nodes of a particular node are connected to each other. AND is expressed as:

$$AND_i = \frac{1}{N(i)} \sum_{j \in N(i)} k_j \quad (5.11)$$

Where $N(i)$ represents the neighbors of node i , and k_j is the degree of neighbor node j belonging to the set $N(i)$.

Figure 5.5 - Schematic representation of the AND index for some nodes in a regular grid.



The numbers in light-colored nodes indicate the degree, while the number in dark nodes shows the AND value.

SOURCE: Author's production.

5.3.5 Cluster Coefficient (CC)

The clustering coefficient (CC) indicates the degree of clustering or the tendency of vertices in a network to form groups or clusters. This index provides information about the local structure of the network, revealing how interconnected the neighbors of a specific node are (KAISER, 2008).

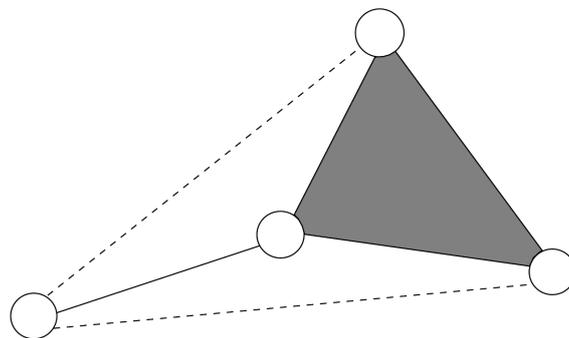
This measure is useful for understanding the clustering structure of the network and identifying regions where interconnectivity is high, providing insights into the formation of potential communities. The clustering coefficient ranges from 0 to 1. A value of 0 indicates that the neighbors of a node are not connected to each other. A value of 1 indicates that all neighbors are fully connected to each other.

For a given node i , the clustering coefficient (CC_i) is a measure of the proportion of connectivity among its neighbors, i.e., it is the average fraction of pairs of neighbors (belonging to the same node) that are also neighbors of each other. The clustering coefficient is defined as:

$$CC_i = \frac{2E_i}{k_i(k_i - 1)} \quad (5.12)$$

Where, for a given node i , k_i is the number of neighbors and E_i represents the number of edges existing only between these neighbors. Figure 5.6 illustrates a network in which we calculate the clustering coefficient for the central vertex. The central vertex has three neighbors, which implies that these neighbors can have a maximum of three edges between them (calculated as $3 \cdot (3-1)/2$). However, in this case, the neighbors have only one edge between them. Therefore, the clustering coefficient for the central vertex is equal to $1/3$, indicating that only one-third of the possible connections between the neighbors are present. This value suggests a low degree of local clustering for the central vertex in the network.

Figure 5.6 - Clustering coefficient for the central node ($CC_i = 1/3$).



The continuous and dashed lines represent the presence and absence of connection between nodes respectively.

SOURCE: Caldarelli (2007).

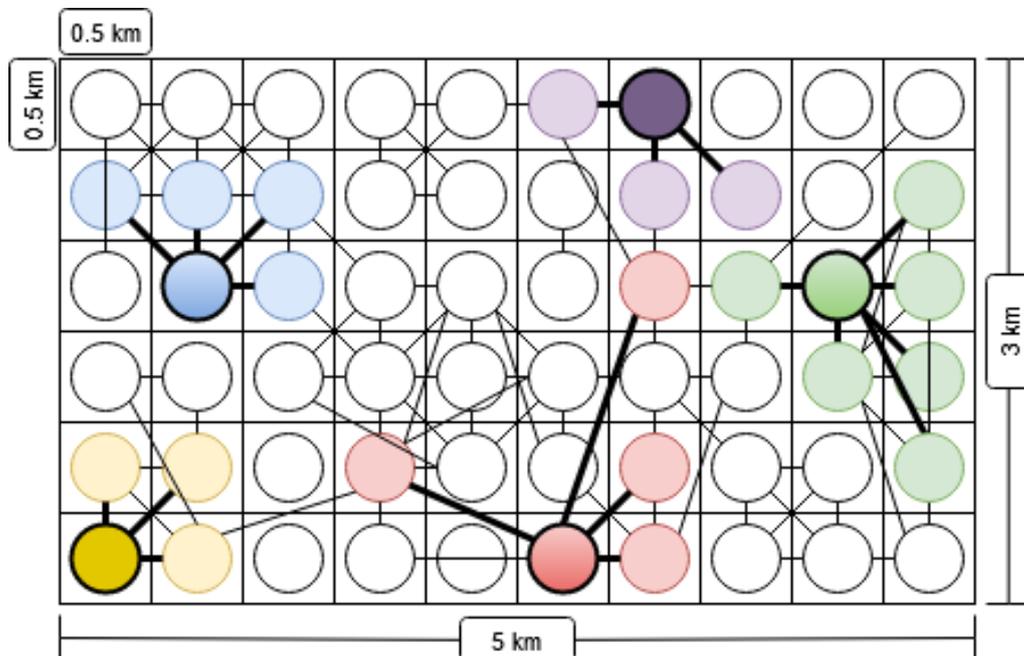
Where λ_j is the latitude of node j and A_{ij} is equal to 1 if nodes i and j are connected, otherwise A_{ij} is equal to 0. The term $\cos(\lambda_j)$ takes into account the decreasing area associated with higher latitude grid cells.

Therefore, AWC can be interpreted as the fraction of the study region area that is connected to a given node. The CC and the AWC provide information about the degree of connection between neighbors and the proportion of connectivity relative to the total area, both for a given vertex. This allows us to identify which vertices have a higher degree of local connectivity and which vertices are more connected in relation to the area.

5.3.7 Geographical Distance (GD)

Geographical distance (GD) refers to the physical distance between nodes or vertices in the network, taking into account their geographic locations (Figure 5.8). In some cases, geographic distance can play a significant role in network connectivity or influence the relationships between nodes (BOERS et al., 2013; BOERS et al., 2014b).

Figure 5.8 - Schematic representation of the GD index for some nodes in a regular grid.



The wide black lines, representing the links of the neighboring nodes, will be used to calculate the respective geographical distances for the dark-colored nodes.
SOURCE: Author's production.

GD can be obtained as follows:

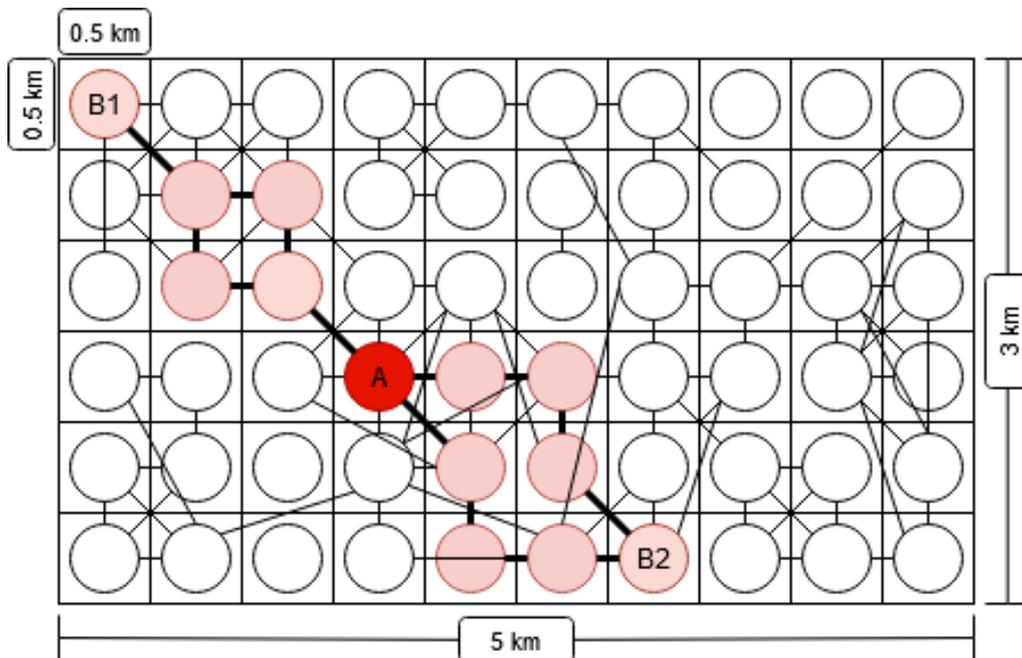
$$GD_i = \frac{1}{D_i} \sum_{j=1}^{D_i} dist(i, j) \quad (5.15)$$

Where $dist(i, j)$ is the distance between points i and j , and $D_i = \sum_{i=j}^N A_{ij}$. GD can be used to analyze spatial patterns or identify potential regions of information flow within the network, indicating areas with local or regional influence.

5.3.8 Betweenness Centrality (BC)

Betweenness centrality (BC) is an important centrality metric due to the information it provides. BC is based on the concept of shortest paths and is defined as the proportion of shortest paths that pass through a particular node compared to the total number of shortest paths in the network (BOERS et al., 2013). BC for a node i is calculated as the sum of the fraction of shortest paths between all pairs of nodes in the network that pass through node i (Figure 5.9).

Figure 5.9 - Schematic representation of the BC index for some nodes in a regular grid.



The four shortest paths between B1 and B2 passing through node A.
SOURCE: Author's production.

Mathematically, it can be defined as:

$$BC_i = \frac{\sum_{l < k \neq i} \sigma_{kl}(i)}{\sum_{l < k \neq i} \sigma_{kl}} \quad (5.16)$$

Where σ_{kl} represents the number of shortest paths from node k to l , and $\sigma_{kl}(i)$ is the number of shortest paths from k to l that pass through node i . BC indicates the ability of a node to act as an intermediary in the transmission of information or flow of resources through the network. This metric is valuable for identifying the most participatory paths and bridges that connect different parts of the network. By calculating the betweenness centrality of each node, we can identify those nodes that act as key intermediaries in the flow of information between other nodes in the network.

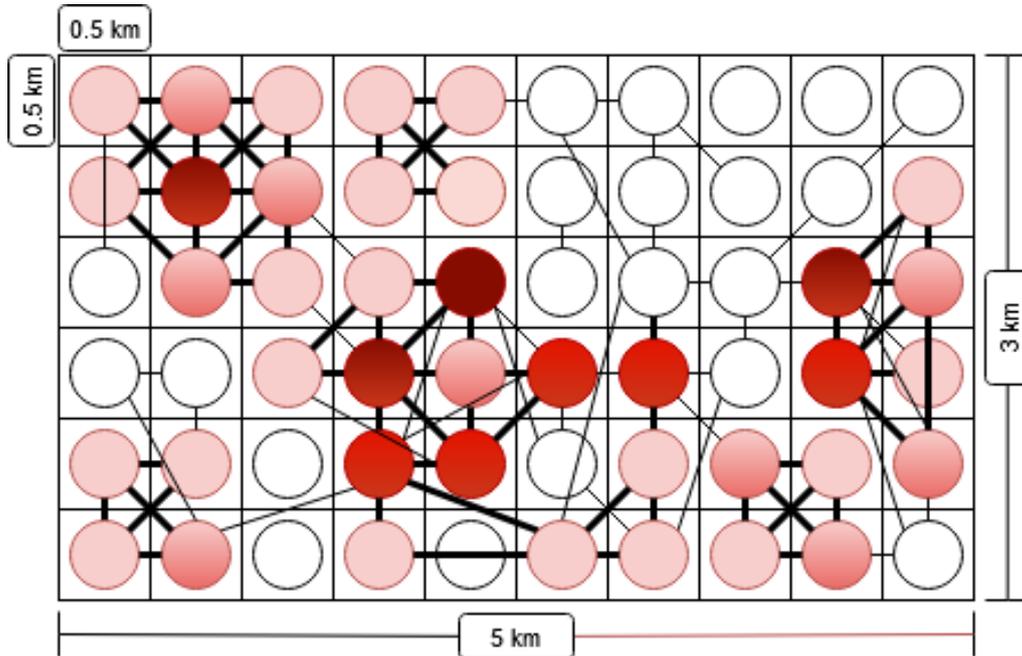
5.3.9 Degree value and the Importance of Lines (DIL)

The Degree value and Importance of Lines (DIL) provide a measure of node importance (LIU et al., 2016). This method provides information about the participation of vertices, taking into account their degree and the importance of the links they participate in. It identifies highly participatory vertices in terms of network activity. The importance of the line (link) between two vertices i and j with link e_{ij} is defined as:

$$L_{ij} = \frac{U}{\lambda} = \frac{(k_i - p - 1)(k_j - p - 1)}{1 + p/2} \quad (5.17)$$

Where p is the number of triangles where the edge e_{ij} participates, k_i and k_j are the degrees of vertices i and j respectively. The DIL allows us to understand the vertex participation in terms of its degree and the importance of its links, providing information about the overall structure and dynamics of the network, including the identification of influential nodes and critical connections for information propagation. Figure 5.10 shows the DIL calculation based on degree and link importance.

Figure 5.10 - Schematic representation of the DIL index for some nodes in a regular grid.



The red colors gradient are the DIL values, while the wide black lines are the vertices that belong to a triangle.

SOURCE: Author's production.

5.4 Community identification

Typically, a network can be composed of subgroups with a high degree of internal connectivity, which exhibit distinct behaviors compared to the rest of the nodes in the network. These subgroups are known as communities.

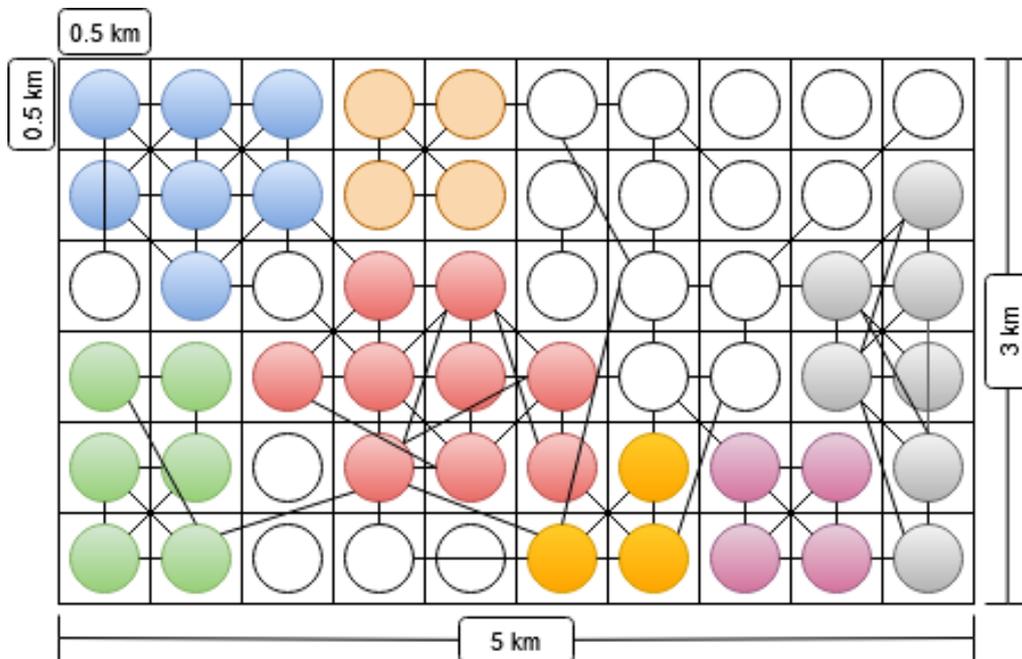
Communities often have different characteristics and provide important information about the dynamics of the network. There are different methods to identify communities (FORTUNATO, 2010). However, in this study, we applied one of the most commonly used methods, the Greedy Modularity Maximization (CLAUSET et al., 2004). This method is based on the application of modularity (Q) (NEWMAN; GIRVAN, 2004; COSTA et al., 2007). Modularity is a measure that evaluates the community structure in the network and quantifies how well the network is divided into distinct communities. This means that vertices within each community are densely connected to each other, while there are fewer connections between communities. Higher values of modularity indicate a more pronounced community structure. Modularity is defined as follows:

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - \gamma \frac{k_i k_j}{2m}) \delta(c_i, c_j) \quad (5.18)$$

Where A is the adjacency matrix, k_i and k_j are the degrees of vertices i and j respectively, m is the number of edges, γ is a resolution parameter, and $\delta(c_i, c_j)$ is 1 if vertices i and j belong to the same community, and 0 otherwise.

Greedy modularity maximization works hierarchically, starting with each node as an individual community and then repeatedly merging with other communities to increase modularity. The goal is to find the community configuration that maximizes the network's modularity. This process is repeated several times until there are no significant gains in modularity. Proper identification of communities can improve the accuracy and reliability of the results, providing more robust insights into the behavior of precipitation events and their implications for the studied region. This information can be valuable in gaining a better understanding and developing more accurate and efficient local weather forecasting.

Figure 5.11 - Schematic representation of the communities in a regular grid.



SOURCE: Author's production.

5.5 Spatial embedding

The mechanisms that influence climate and weather are not limited by political boundaries or artificial divisions of study areas (RHEINWALT et al., 2012). These restrictions can affect connections and have an influence on centrality measures, also known as spatial embedding. To correct this effect, we create 400 sets of artificial rainfall data, called surrogates. These surrogates were obtained from the Fourier transform. These surrogates are artificially generated data that preserve certain properties or statistical characteristics of the original data series. They are used as null or control models, allowing the assessment of the significance or randomness of the observed network. From these surrogates, 400 networks will be created. In these networks, indices that require spatial correction (CC, BC, and CL) will be calculated, and the average values of these indices will be obtained. Subsequently, these values will be divided by the respective values from the original measurements, thus obtaining the corrected indices.

5.6 Community characterization

Once the communities had been identified, a study was conducted to investigate the similarities between them. The analysis consisted of studying the topology and specific characteristics of the networks, in other words, the information that complex network techniques provide. The second part focuses on the statistical analysis of both rainfall and extreme events within each community. In this analysis, spatial and temporal information was obtained regarding average behaviors, frequency, intensity, trends, and others. Lastly, the analysis of precipitation characteristics provides a physical approach to precipitating systems, offering key insights into how these events unfold. This information is crucial for understanding the differences between events occurring in different communities.

5.6.1 Network topology analysis

This analysis begins with the results of the degree distribution for each of the communities, from these results, we can observe which communities exhibit more homogeneous behaviors or stand out as highly connected sub-regions. Additionally, we calculated the connectivity density to identify which communities are the most densely interconnected. These results of degree distribution and density are complemented by the spatial results of the CC, AWC, GD, BC, and DIL indices. These indices helped us understand the dynamics present in the communities and obtain characteristic information about each of them.

The CC provided the degree of internal connectivity within each community and the presence of local clusters. The AWC gave us insights into the spatial influence of vertices within the communities. The GD allowed us to identify regions of information exchange and local or regional influences. The BC and CL revealed the most important nodes and central vertices within the communities. Lastly, the DIL helped us identify crucial nodes for communication between different parts of the network.

By combining these different results, we were able to obtain a comprehensive view of the dynamics of the communities and gather representative and distinctive information about each of them. This will allow us to better understand the structure and functioning of the communities within the complex network under study.

5.6.2 Distribution of extreme events

This stage consists of statistical analyses of the communities. Initially, rainfall events and extreme rainfall were examined in statistical terms, taking into account the structure of the network under consideration. Characteristic information was extracted for each community, such as box plots, mean, median, maximum value, minimum value, standard deviation, and variance. Additionally, the probability density function (PDF) was analyzed to identify magnitudes with the highest probability of occurrence within the community. Furthermore, skewness and kurtosis values were calculated, which was useful for comparing distributions among different communities.

To verify the statistical significance of two datasets representing rainfall in different communities, we employ two non-parametric statistical methods, both within a 95% confidence interval. First, we use the Mann-Whitney U test to assess the differences in averages between the two datasets. This test helps us determine whether one group tends to have higher or lower rainfall values compared to the other. Secondly, we utilize the Kolmogorov-Smirnov test to compare the quantified distance between the empirical distribution functions of the two datasets. This test helps us assess how the overall distribution of rainfall values differs between the two communities. These methods are valuable because they don't make any assumptions about the underlying data distribution, making them suitable for comparing datasets with potentially different characteristics. The statistical analysis allowed us to identify which communities exhibit similarities in terms of magnitude and frequency of precipitation events (rainfall or extreme rainfall).

By grouping communities with similar patterns, we gained a deeper understanding of the shared characteristics among them. This contributes to a more detailed understanding of the spatial and temporal distribution of precipitation events in the studied region.

5.6.3 Deep convective precipitation

An additional part of the characterization of communities in SA involves the analysis of deep convective profiles. This analysis provides valuable insights into the physics underlying extreme rainfall events likely associated with convection. Reflectivity profiles, which depict the vertical structure of precipitation, are used as a characteristic signature of each community. These profiles offer diverse information about the presence of hydrometeors, which can indicate various processes associated with precipitation, such as light rain or deep convection.

In addition to reflectivity profiles, other variables from the precipitation features dataset that provide more detailed information about rain microphysics are also considered. These variables include ice content, water distribution, 85 GHz temperature with corrected polarization, and maximum height of 30 and 40 dBZ echo-top. By incorporating these variables, a more comprehensive and accurate analysis of convective events is possible, aiding in the understanding of atmospheric processes and the identification of specific characteristics of the communities under study.

For the case of the communities of extreme rainfall networks, in addition to the aforementioned analysis in previous sections, it includes a complementary analysis focused only on deep convective events. These events were identified based on specific criteria, targeting pixels that exhibited the highest rainfall rate with the deepest convection (HAMADA et al., 2015). The identification process involved setting thresholds based on the 95th percentile of maximum near-surface precipitation (MAXNSRAIN) and the maximum height of 40 dBZ echo-top (MAXHT40).

Two approaches were defined based on the spatial distribution used. The "homogeneous" approach applies the 95th percentile to each grid point (with a resolution of 0.5°) within each community, allowing for an analysis of the overall behavior of the community. On the other hand, the "non-homogeneous" approach considers the sample for the community as a whole, leading to an imbalance in the spatial distribution of the data. This approach favors regions within the communities where the chosen criteria for deep convection are more frequent. This allows for a focused analysis of the importance of these areas within the communities.

Any pixel that exceeded these limits was considered to exhibit deep convective precipitation. By applying these criteria, the analysis specifically focused on events that displayed intense rainfall and significant vertical development, indicating the presence of deep convective processes. Note that the 95th percentile is applied to select the deep convective precipitation. Once this information is extracted, it will be used to filter the events in each of the approaches. In the "homogeneous" approach, these values vary according to the grid point, while for the "non-homogeneous" approach, they are defined for the entire community.

6 RESULTS

This chapter presents the results obtained from the analysis of the networks used in this study and is divided into two sections.

The first section focuses on the results of the extreme rainfall events networks, specifically the South America Network. In this analysis, event synchronization was employed to create these networks, facilitating the identification of patterns of extreme events and communities. In the case of the communities within the South America Network, they will be characterized through a combination of physical and statistical analysis. This characterization will include studying their structure, connectivity, and specific attributes. By conducting this analysis, we aim to gain a better understanding of how extreme rainfall events are organized within highly interconnected groups. These findings will provide valuable insights into the dynamics and underlying mechanisms of these events within South America.

The second section focuses on the results of the extreme and total rainfall networks over the Amazon Basin. Based on the results of the South American network, as we shall see, it was evident that the Amazon basin itself presents a specific dynamic. In order to understand what happens in this specific region, deeper analyses were carried out with different networks. By the analysis of this specific region, we gain a deeper understanding of how they affect the structure and properties of the network, providing valuable insights into the dynamics of rainfall in this important region.

We will discuss the identified patterns and their relationship to regional climate phenomena, offering valuable insights into these events. Overall, this chapter will provide a comprehensive view of the studied networks and present key analysis results.

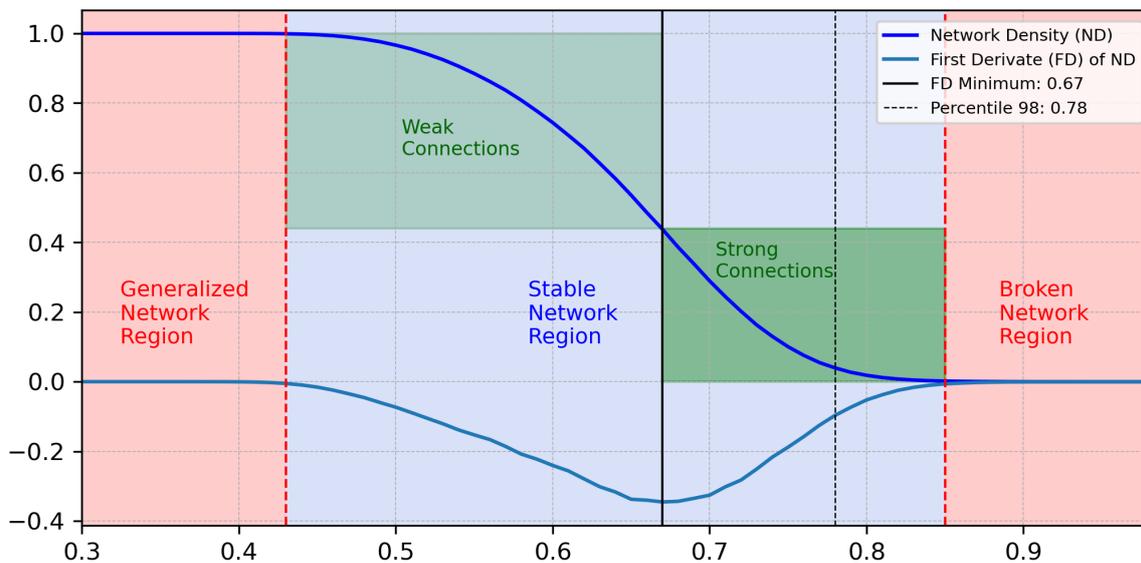
6.1 SOUTH AMERICA network

An undirected and unweighted network of extreme daily rainfall events was constructed for the Austral summer, this network is called "South America". The network is composed of a regular grid with 120 latitude points and 110 longitude points with a total of 13200 nodes, and a spatial resolution of $0.5^\circ \times 0.5^\circ$. Each node corresponds to a time series of dichotomous extreme rainfall events. To define the link between two nodes, we use the event synchronization with a maximum delay of 3 days (τ_{max}), as described in previous studies (QUIROGA et al., 2002; MALIK et al., 2012; BOERS et al., 2013).

Using this information, the synchronization values between all possible pairs of nodes were computed. Since the network is undirected, the synchronization from node i to node j and vice versa were summed. To define the links in the network, only the values greater than or equal to a threshold, known as the Minimum Link Value (MLV), were considered.

Figure 6.1 visually illustrates the variation of link density for different MLV values. The link density provides an indication of the network's connectivity level (ZHANG et al., 2010). It is calculated as the ratio of the actual number of links to the total number of possible links in the network. The link density ranges between 0 and 1, where a value of 1 represents a fully connected network, while a value of 0 indicates a disconnected or fragmented network. The link density was then calculated for various MLV values, ranging from 0.3 to 0.98 with an increment of 0.01 (DONGES et al., 2009b).

Figure 6.1 - Density variation for different MLV values for South America network.



The colored regions indicate the degree of connectivity of the network, while the boundaries of the regions (red and black dotted lines) were defined from the variation of the first derivative of the network density (blue line).

SOURCE: Author's production.

By examining the link density across MLV values, we can identify the threshold at which the network achieves a good balance between connectivity and sparsity, allowing us to obtain a reference of the connection intensity of the MLV used.

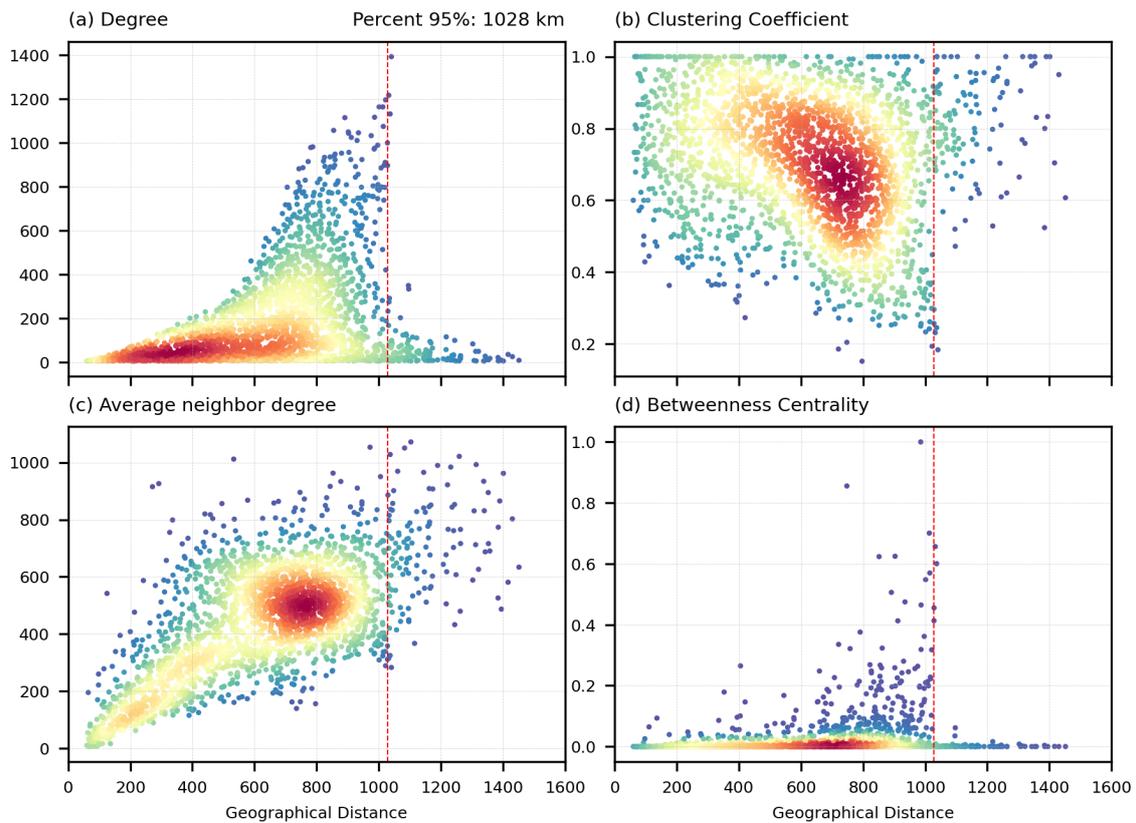
The analysis of the network density variation reveals distinct connectivity regimes based on the MLV values. For low MLV, the network exhibits high connectivity, with numerous links between nodes. As the MLV increases, the network gradually transitions into a disconnected or broken state, as indicated by the red region. By examining the first derivative of the density curve (light blue line), an inflection point is identified as the minimum value of the derivative (vertical black line). This inflection point separates two sub-regions within the blue region. These sub-regions are characterized as weak and strong connections, representing different levels of connectivity within the stable or regularly connected network. Overall, this analysis provides information into the network connectivity at different MLV thresholds. To identify highly connected communities, the MLV threshold was set to include only the 2% of the strongest links in the network. This threshold selects links with values greater than or equal to the 98th percentile of the synchronization values (BOERS et al., 2013). From Figure 6.1, it can be observed that the MLV defined by the 98th percentile (represented by the dotted black line) falls within the region of strong connections. This indicates that the network exclusively incorporates extreme rainfall events with a robust synchronization of up to 3 days. By considering only these strong connections, the network analysis focuses on capturing the most significant and tightly linked extreme events.

6.1.1 Geographical distance

The relationship between geographic distance (GD) with several indices of the network was examined in order to identify the window of greatest activity in the network. These indicators include Node Degree (DG), which measures the number of links connected to a node (neighbors); Clustering Coefficient (CC), which quantifies the degree of clustering in the node; Average Neighbor degree (AND), representing the average degree of the neighboring nodes and Betweenness Centrality (BC), which indicates the importance and participation of a node in the transmission of information throughout the network, displaying the bridge role of a certain node.

By analyzing the relationship between GD and these network indexes, we can identify the spatial distance regions or windows of greatest activity inside the network, associated with the response or connectivity patterns of the extreme rainfall events. Figure 6.2 displays a scatter plot of the GD and these indexes. Notably, it reveals that as the geographical distance approaches 1000 km, the network presents an important decrease in the number of nodes. This indicates a clear limitation in the connectivity or interaction between nodes within this distance range.

Figure 6.2 - Scatter plot of Geographic Distance (km).



(a) Degree, (b) Clustering Coefficient, (c) Average Neighbor Degree, and (d) Betweenness Centrality. The colors indicate the spatial density of the nodes, while the red line represents the 95th percentile of the geographical distances (1028 km).

SOURCE: Author's production.

The analysis of network metrics revealed important observations. Nodes with higher degrees are concentrated within a geographical distance range of approximately 600 to 1100 km (Figure 6.2a), indicating the presence of localized regions with a higher number of connections. Additionally, the metrics of clustering coefficient, average neighbor degree, and betweenness centrality demonstrate better performance within a distance range of 500 to 1100 km (Figures 6.2b, 6.2c and 6.2d). This suggests that extreme rainfall events within this window exhibit stronger network activity and connectivity.

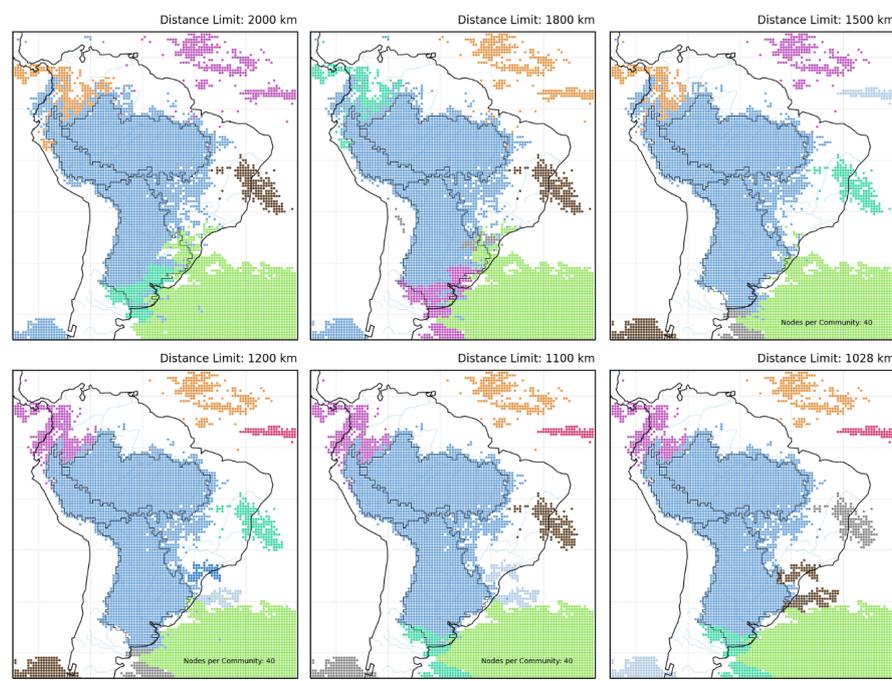
The analysis suggests that while links outside the distance range are less frequent, they can still provide valuable information depending on the research objective. However, this study specifically focuses on regional connections rather than very distant ones, as they tend to favor larger or more extensive communities.

By excluding these longer connections, the analysis can be more focused on understanding the dynamics and interactions within smaller, geographically localized communities. To establish a threshold for connections, the 95th percentile of GD was used. Connections with GD greater than 1028 km (red line in Figure 6.2) were excluded, resulting in a network without longer connections.

6.1.2 Removing the ocean

Tropical convective events are typically classified into two groups: marine and continental (MOHR; ZIPSER, 1996; XU; ZIPSER, 2012). Rainy events occurring over the ocean and the continent exhibit distinct characteristics. For instance, convective events over the continent tend to have strong updrafts that promote deep convection, resulting in increased ice and electrical discharges. On the other hand, over the ocean, convection is not as intense, being weak to moderate, and it has a higher contribution from low-level processes (warm rain). Therefore, one of the main differences between these two types of events is the intensity of convection (ZIPSER et al., 2006). As this study focuses on identifying extreme rainfall communities exclusively on the continent, precipitation occurring over the ocean was ignored (Figure 6.3).

Figure 6.3 - Communities in networks considering continent and ocean rainfall.



The geographical distance limits are indicated at the top of the figures. The black line represents the communities that were utilized in this study.

SOURCE: Author's production.

Therefore, the network only includes data from the continent, as shown in Figure 5.2. The exclusion of oceanic precipitation aims to highlight the predominance of naturally occurring extreme events over the continent. However, to assess the difference in continental communities with and without oceanic data, another network was constructed, encompassing the study area including the ocean (Figure 6.3).

It is worth noting that the inclusion of ocean data in the analysis revealed regions where transient systems such as cold fronts and the Intertropical Convergence Zone (ITCZ) are more frequent. However, these regions formed relatively small communities compared to the continental ones. This is likely due to the high space-time variability of extreme events, which can lead to weak linkages through synchronization. Nevertheless, the technique employed in this study was robust enough to distinguish between predominantly continental communities and those associated with the ocean.

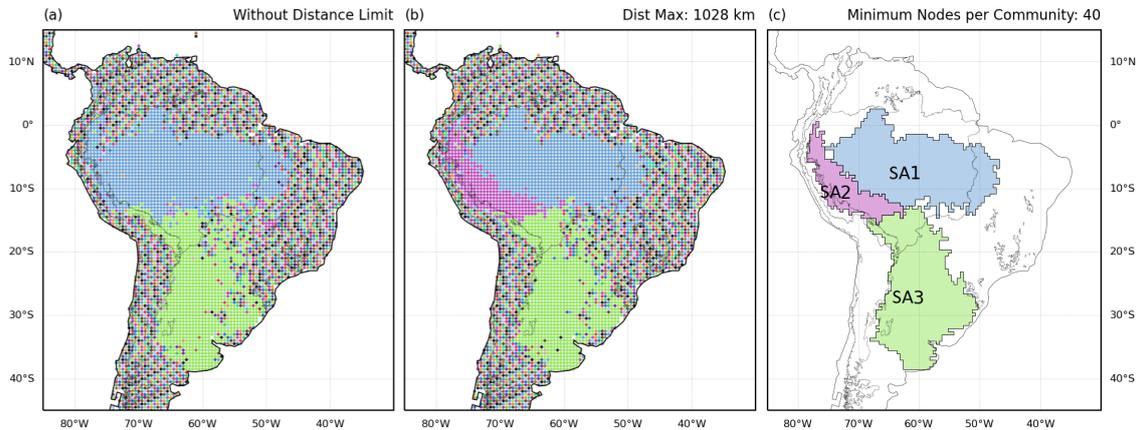
6.1.3 Communities identification

For the community identification, we employ the widely used Greedy Modularity Maximization method (CLAUSET *et al.*, 2004). This method relies on optimizing modularity (M) (NEWMAN; GIRVAN, 2004; COSTA *et al.*, 2007). It iteratively clusters nodes together to maximize modularity, starting with each node as a separate community and merging them until modularity stops improving. The geographical distance limit of 1028 km was chosen based on the objective of studying regional communities. The exclusion of connections beyond this distance helps focus on more localized interactions and smaller-scale communities. In Figure 6.4a, the network without the distance filter exhibits two large communities.

However, after applying the filter (Figure 6.4b), the network is divided into three communities, highlighting the splitting of the community located between 2°N and 10°S (blue community in Figure 6.4a). It suggests that the chosen distance limit removes links greater than 1028 km, probably from the eastern Amazon basin. In this way, by removing these links, the blue community cannot be grouped. Finally, a predominant community associated with the topographical barrier of the Andes (pink community) is identified.

Three communities were identified, and they comprise the biggest component of the network. These communities exhibit strong link values above the MLV threshold, indicating the occurrence of extreme events that are synchronized within a maximum time displacement of three days and a geographical range of up to 1028 km.

Figure 6.4 - Communities in the extreme daily rainfall network.



(a) all nodes in the network are displayed, indicating the composition of each community without applying a distance limit on the links. (b) The same communities are depicted but with a link distance limit of 1028 km. (c) Polygons represent the boundaries. The communities compose the biggest component of the network. SOURCE: Author's production.

Furthermore, these communities were clearly delineated, there was no necessity to set a minimum number of nodes for their identification. The extraction of polygons from each community is depicted in Figure 6.4c, employing a Von Neumann neighborhood with a radius of 1 pixel (ZAITSEV, 2017). In general, these three communities demonstrate a similar spatial pattern observed in previous studies (RASMUSSEN; HOUZE, 2011; ROMATSCHKE; JR, 2010; RASMUSSEN et al., 2016; KUMAR; SILVA, 2020).

6.1.4 Network analysis

The spatial distribution of the communities shown in Figure 6.4 exhibits a remarkable association with moisture transport from the central Amazon basin through the eastern Andes to southern South America. Where, this is a notable characteristic of the dynamics of the SAMS during the austral summer (VERA et al., 2006a; MARENGO et al., 2012; GELBRECHT et al., 2021). Figure 6.5 provides a schematic representation of the SAMS mechanisms. The three communities consist of important regions in South America.

Community 1 encompasses a large part of the Amazon basin, which is characterized by its contribution of moisture to the atmosphere through evapotranspiration (FISCH et al., 1998; HAYLOCK et al., 2006; NOBRE et al., 2009).

Figure 6.5 - Overview of important features of the low-level atmospheric circulation during the SAMS.



The red arrows represent the winds, and the black circles represent the South Atlantic Converge Zone (SACZ) and southeastern South America (SESA). These regions are associated with the phases of the SAMS.

SOURCE: Gelbrecht et al. (2021).

Community 2 is defined as the region east of the tropical Andes. In this community, the South American Low-Level Jets (SALLJ) change their direction from eastward to southward due to the presence of the Andes, which act as a topographic barrier (BOOKHAGEN; STRECKER, 2008; DRUMOND et al., 2008; SILVA; KOUSKY, 2012). The interaction between the SALLJ, which transports moisture from the Amazon basin, and the Andes leads to the development of orographic rainfall. The moisture transport to the extratropics of South America provides favorable conditions for the development of convective systems over the *La Plata* basin (MONTINI et al., 2019; BERBERY; BARROS, 2002; SALIO et al., 2007; PIERSANTE et al., 2021). Finally, community 3 extends from the central Andes to southern Uruguay, including the *Sierras de Córdoba*. This region is characterized by the presence of the deepest convective systems in South America (RASMUSSEN et al., 2014; MULHOLLAND et al., 2018), as well as the occurrence of Mesoscale Convective Systems (MCS) and frontal systems (SALIO et al., 2007; PIERSANTE et al., 2021).

The results of the extreme events networks are divided into two parts. The first part, presented in Section 6.1.5, is related to the communication between communities and is referred to as the inter-communities results. It includes an analysis of the degree of connection that exists beyond their boundaries, as well as a comparison with the network composed of all of South America. The second part of the analysis focuses on the individual participation of the vertices within their own community and is referred to as the intra-communities results. Additionally, it involves the identification of sub-regions with the highest activity. These results can be found in Section 6.1.6.

6.1.5 Inter-communities

The network composed of all the vertices in South America (Figure 6.4b) was referred to as the South America Network. The statistical results of the South America Network and each community can be found in Table 6.1. The South America Network exhibits sparse connectivity, with a density of 0.002 and low entropy of 0.179. This indicates that the network is characterized by a lack of strong connections, and this characteristic is even more pronounced when compared to individual communities.

Table 6.1 - The calculated values for different indices are presented for the South America network and each individual community.

Index	South America Network	Community ID		
		1	2	3
Area (km^2)	17 840 000	3 248 558	919 971	2 876 030
Number of Nodes	13 200	1 074	307	1 028
Mean Degree	109.425	130.78	80.847	68.58
Average Clustering	0.133	0.652	0.749	0.686
Shannon Entropy	0.179	0.798	0.858	0.716
Network Efficiency	0.017	0.519	0.616	0.445
Average Shortest Path Length	3.248	2.14	1.834	2.538
Link Density	0.002	0.122	0.264	0.067
Diameter	11	5	4	7

This observation aligns with the findings of Freitas et al. (2019), who analyzed thirteen real-world networks and found similar trends, with link densities below 0.009 and low entropy values (less than 0.25). The low density of the South America network results in lower values of Average Clustering and Network Efficiency, indicating a reduced ability to efficiently transmit information across the network.

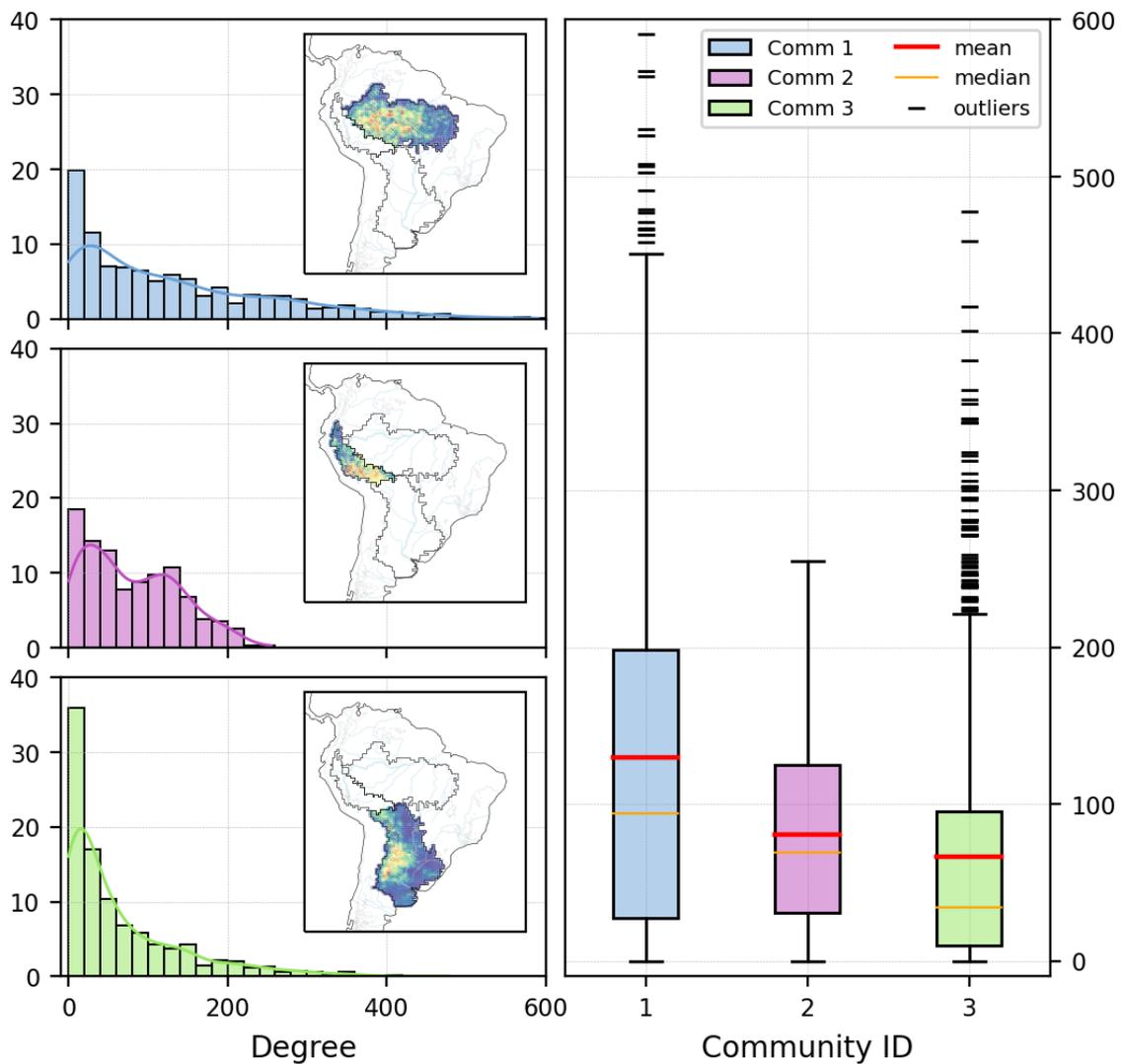
Table 6.1 indicates that Community SA1 is the largest among the three communities, with a size that is 71% larger than Community SA2 and 11% larger than Community SA3. It is characterized by a significant internal area and a high number of interconnected nodes, implying a potentially higher occurrence of extensive extreme events. This behavior of Community SA1 could be attributed to the characteristics of the Amazon region, which is known for its dense vegetation and its influence on local climate dynamics through processes such as moisture recycling (ZEMP et al., 2014). These processes can in turn affect the frequency, intensity, and spatial distribution of rainfall. Some studies suggest that these mechanisms may result in convective patterns that resemble those observed over oceans rather than continents (RASMUSSEN et al., 2016). As a consequence, Community SA1 exhibits a relatively higher Network Efficiency (0.519).

Community SA2 exhibits the highest link density among the three communities, reaching approximately 26%. It is characterized by the highest Average Clustering coefficient of 0.749, indicating a densely interconnected network. With a diameter of 4, it has the shortest maximum distance between nodes within the community. The presence of the central Andes acts as a topographic barrier, restricting the expansion of this community. The increased connectivity within this community can be attributed to the influence of the topographic barrier, where the Andes cause a redirection of wind directions towards the south (BYERLE; PAEGLE, 2002). This alteration in wind patterns contributes to similar behaviors among the vertices within Community SA2 due to orographic lifting (ROMATSCHKE; HOUZE, 2013). The high connectivity within this community is associated with the frequency and spatial limitation of extreme events in the area, which is also reflected in its elevated Network Efficiency (0.616), making it the most efficient community. Additionally, Community SA2 demonstrates the highest entropy value (0.858), indicating greater variability in the distribution of extreme events compared to the other communities, particularly Community SA1. This observation is further supported by examining the degree distributions of each community.

Community SA3 displays the lowest connectivity among the three communities, characterized by a link density of 0.067, a diameter of 7, and an efficiency of 0.445, indicating lower efficiency compared to the other communities. However, it exhibits a slightly higher average clustering coefficient of 0.686 compared to Community SA1, which has a value of 0.652. This observation suggests the presence of highly interconnected subregions within Community SA3.

Figure 6.6 presents the results of degree maps and their distribution for each community. The node degree maps reveal the subregions with a higher degree. For Community SA1, the central part of the Amazon region exhibits a dense network of higher degrees. In Community SA2, the high region is predominantly located to the east of the central Andes. Community 3 extends from the northern region of the community towards the *Sierras de Córdoba*, showcasing its own distinct area.

Figure 6.6 - Results of the degrees for each community.



The left and right column shows the degree frequency (%) and degree boxplot for each community respectively. The orange, red, and black lines indicate the median, mean, and outliers. The little maps represent the node degree in each community.

SOURCE: Author's production.

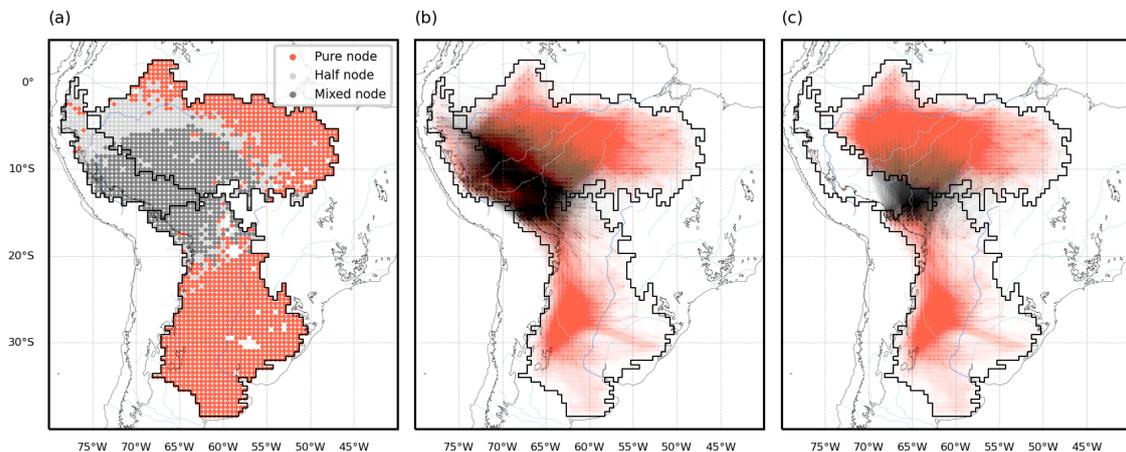
Community SA2 demonstrates a relatively uniform distribution of node degrees, with a gradual decline in high-degree nodes, indicating low variability. The community exhibits a lower frequency of nodes with higher degrees, with a maximum of 200 links. The boxplot confirms the homogeneous nature of Community SA2, as it shows low variability and the absence of outliers. In comparison to communities SA1 and SA3, Community SA2 has a smaller difference between the mean and median (11.2), while Communities SA1 and SA3 have differences of 35.69 and 32.3, respectively. These findings suggest that Community SA2 displays less heterogeneity compared to the other communities. Moreover, Community SA2 has the highest Average Clustering value (0.749), indicating a stronger tendency to form dense and interconnected clusters of extreme precipitation events. The degree distribution of Community SA1 exhibits a skewness of 1.07 and a variability of 121.66, with a long tail and a smooth drop. This indicates a higher probability of the presence of potential hubs within Community SA1. In contrast, Community SA2 shows low variability and skewness values among the communities (0.5 and 58.7, respectively). On the other hand, Community SA3 displays the highest skewness value (1.79) and a variability of 78.34. This suggests that Community SA3 experiences less extensive events or events with less spatial displacement but with a higher probability of occurrence, resulting in a longer right tail with a steep drop in the distribution. The connectivity within Community SA2 exhibits a more balanced distribution, indicating a higher proportion of nodes with a greater number of connections compared to its area (44%). This suggests that Community SA2 plays a significant role in facilitating intercommunication between Communities SA1 and SA3 or has a higher level of internal connectivity.

Community SA2 serves as a crucial region, enabling the transport of moisture and influencing convection over the *La Plata* basin (MONTINI et al., 2019; BERBERY; BARROS, 2002). These conditions create favorable environments for the formation of MCS (SALIO et al., 2007; PIERSANTE et al., 2021). To further investigate this, a focused analysis was conducted exclusively on the links connecting the three communities. This analysis excluded nodes outside of these three communities and solely examined the connections within the three communities (inter and intra links). Consequently, the network comprises only the nodes from these communities.

This analysis allowed us to examine the level of communication between the communities. Figure 6.7a illustrates the quantification of node connectivity with respect to communities, considering three defined conditions: "Pure nodes", "Half nodes", and "Mixed nodes".

Pure nodes are those that have connections exclusively within their own community. Half nodes are those that have connections to one external community in addition to their own. Mixed nodes have connections to two other communities along with their own. No nodes were found to have links solely outside their own community. The results indicate that Community SA2 has a very small percentage of pure nodes, with only 2.61% of nodes having connections exclusively within their own community. The remaining 97.39% of nodes have connections to at least one other community, making Community SA2 the community with the highest level of external communication. In terms of the percentage of half and mixed nodes, Community SA1 and Community SA3 have 54.38% and 20.64%, respectively.

Figure 6.7 - Node connections considering only three communities.



a) Nodes classification in relation to communities connections, b) inter-communities and intra-communities vertices represented by black and red lines respectively c) same as b but without Community 2.

SOURCE: Author's production.

Figure 6.7b illustrates the inter and intra links between communities, highlighting the dominant role of Community SA2 in communication, particularly with Community SA1. This suggests a significant relationship between extreme events in Community SA2 and the mid-western Amazon basin (ROCHA et al., 2009; MONTINI et al., 2019). In contrast, connections between communities SA2 and SA3 are primarily concentrated in the northern region of Community SA3. These connections are limited by the GD limit (1028 km) and the steep slope of the central Andes. This results in a substantial contribution from communities SA1 and SA2 to the characteristic region known as the Bolivian hotspot (ESPINOZA et al., 2015).

Figure 6.7c displays the links between communities SA1 and SA3, excluding nodes belonging to Community SA2. A comparison between this network (without Community SA2) and the network in Figure 6.7b reveals a loss of approximately 83% of inter-community links. This underscores the significant intercommunication role of Community SA2, primarily attributed to the presence of the Andes. The links also indicate the contribution of Community SA1 to the northern region of Community SA2. Notably, nodes surrounding the Bolivian hotspot exhibit higher degrees compared to other nodes in Community SA3, particularly in the northeast. This suggests a direct and influential connection between the south-central Amazon basin and the Bolivian hotspot.

The lower number of links observed in the eastern region of Community SA1 appears could be associated with the genesis of extreme events. MCS tend to propagate following a westward flow (ANSELMO et al., 2021), developing and expanding over the central-western Amazon basin, which explains the higher density of links in that area. It is important to note that the connections are constrained by a maximum geographical distance of 1028 km. Consequently, the communities do not have nodes with connections that exceed this distance. However, this limitation does not undermine the existing connections or diminish the significance of Community SA2. But in turn, this does not allow us to realize the synchronization of events with distances greater than this maximum link distance.

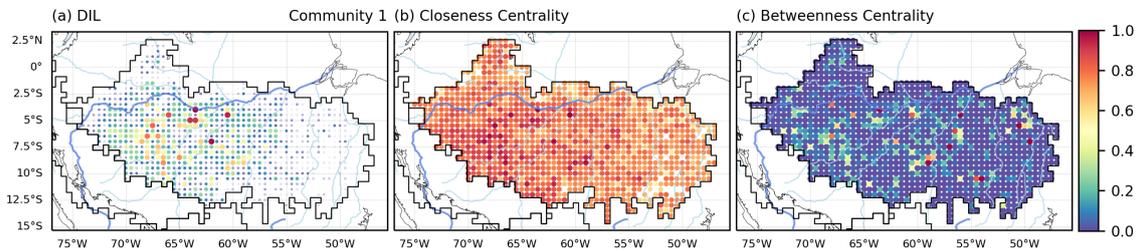
6.1.6 Intra-communities

However, the degree alone does not offer a comprehensive understanding of the network. To gain more insightful information, we calculated additional metrics, including DIL, CC, and BC, all of which were normalized. These metrics were computed separately for each community to extract characteristic insights (refer to Figures 6.8, 6.9, and 6.10). Consequently, our analysis of the communities is presented individually, and it's important to note that the results have been normalized to a range of 0 to 1. As a result, the values and colors do not represent identical measurements across the different communities.

Within Community SA1, the area of highest activity is primarily situated in the mid-western region, aligning with previous research highlighting the prevalence of precipitation events propagating westward in the Amazon basin due to the easterly trade winds (MACHADO et al., 1998). The Amazon basin is renowned as one of the world's wettest regions, with a notable frequency of MCS events (NESBITT et al., 2006; REHBEIN et al., 2018).

Figure 6.8 illustrates that Community SA1 exhibits a more active network relative to its size, with the largest participation area of the community positioned south of the Amazon River. This region may indicate a higher occurrence of extreme events or serve as a location where such events mature and/or intensify. Notably, Marengo et al. (2004) reported that the westerly regime is characterized by baroclinic waves extending towards the central part of the Amazon basin, facilitating moisture transport towards the eastern side of the Andes. This may serve as an explanation for the increased activity predominantly situated in the west-central part of Community SA1.

Figure 6.8 - Normalized results for Community SA1.



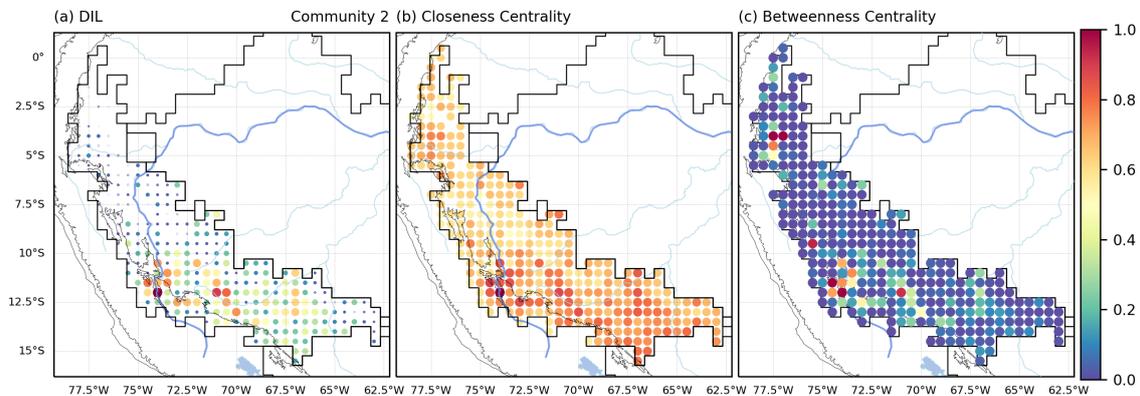
(a) Degree value and the Importance of Lines, (b) Closeness Centrality, and (c) Betweenness Centrality, while the blue line contained in the community represents the Amazon River.

SOURCE: Author's production.

In this community, the distribution of high BC values exhibits a spatial pattern that is prevalent throughout the community, indicating that this community has a higher level of engagement in communication compared to other communities within the network. This suggests that Community SA1 holds a central position in the network's overall functioning, and any disruption within it could significantly impact network performance (Figure 6.8c). This emphasizes the importance of the Amazon region for rainfall across South America. Additionally, Community SA1 displays a group of actively participating nodes in the eastern part of the community, approximately around 50°W and 5°S. In contrast, the patterns observed for Degree Importance and Closeness Centrality are less pronounced. The elevated BC values along the eastern edge might be linked to extreme events originating from the east, possibly influenced by coastal systems like instability lines that propagate westward into the Amazon basin (GRECO et al., 1990; GARSTANG et al., 1994).

The central-southern region of Community SA2 exhibits the highest level of activity, as depicted in Figure 6.9a. This distribution is directly linked to the occurrence of extreme events originating from Community SA1. This association is supported by the numerous connections between the western-central area of Community SA1 and the central-southern region of Community SA2, as shown in Figure 6.7b. Community SA2 displays a more concentrated pattern, with indicators such as DIL, CC, and BC highlighting the prevalence of extreme event activity primarily in the south-central part of the community. This could potentially be influenced by the mature phase of the SAMS (ZHOU; LAU, 1998; VERA et al., 2006a) and the presence of the Bolivia high (ACEITUNO; MONTECINOS, 1993; THIBEAULT et al., 2012).

Figure 6.9 - Normalized results for Community SA2.



(a) Degree value and the Importance of Lines, (b) Closeness Centrality, and (c) Betweenness Centrality. The fraction of the blue line contained in the community represents the *Ucajali* river.

SOURCE: Author's production.

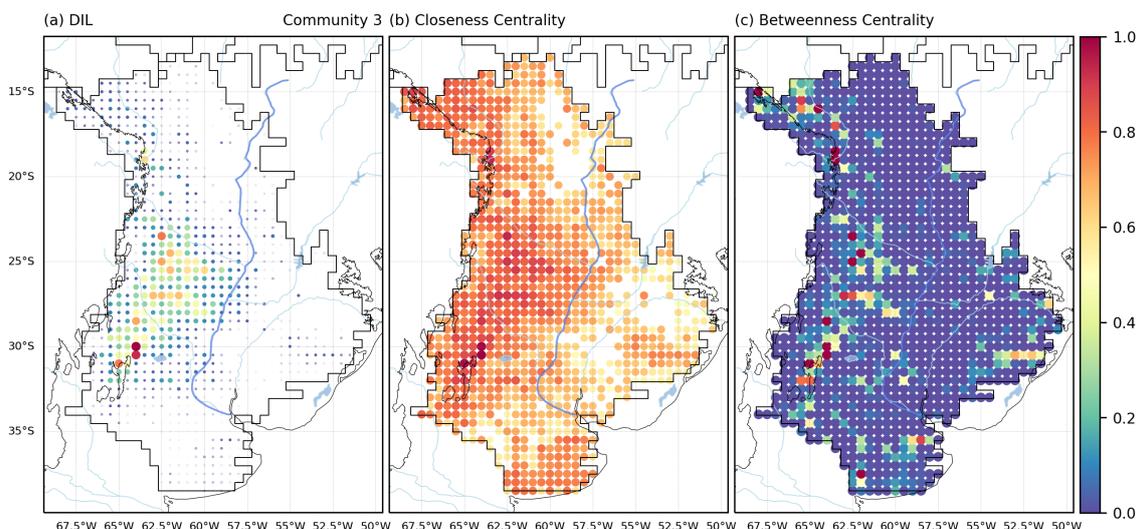
In the study by Anselmo et al. (2021), the eastern central Andes, specifically the Peruvian Hotspot between 12°SS and 71°SW, was identified as a region with a high probability density of MCS. This region exhibits elevated values of DIL, CC, and BC. Furthermore, Espinoza et al. (2015) highlighted the significant relationship between large-scale wind circulation and extreme rainfall events in the Peruvian Hotspot. These events are associated with wind patterns at 850 hPa, where the winds are mostly perpendicular (around 90°S) to the main axis of the Andes. Community SA2 is primarily characterized by the influence of orographic and convective precipitation (BOOKHAGEN; STRECKER, 2008). Specifically, in the eastern Andes, the moisture stream at low levels follows the slopes of the mountains, resulting in increased rainfall and convective activity in this region.

This phenomenon significantly contributes to the spatial and temporal distribution of precipitation in the area (FIGUEROA; NOBRE, 1990; LENTERS; COOK, 1995; MONTINI et al., 2019; JONES, 2019). Relief features play a crucial role in modulating extreme rainfall events within Community SA2. In a study conducted by Bookhagen and Strecker (2008), it was observed that relief features have three main characteristics that influence the rainfall-orography relationship: (i) similarity between elevation and relief, (ii) base level or minimum elevation conditions, and (iii) moisture flux. These characteristics help explain the more active response of the network to the west of Community SA2.

The dominant participation of vertices within Community SA3 (Figure 6.10) is concentrated on the eastern side of the Andes (i.e., the western portion of the community). In this community, the high activity could represent warm moisture from the Amazon basin meeting the frontal system with dry and cold air coming from the south of South America. Consequently, this flow significantly influences the circulation patterns and the distribution of extreme events in the region, leading to the emergence of two distinct climatic regimes: a relatively cold and dry regime in the western region and a warm and humid regime in the eastern region (SELUCHI; MARENGO, 2000). Community 3 exhibits the highest level of activity along the subtropical Andes, particularly between the *Sierras de Córdoba* and just below 20°S (as depicted in Figure 6.10).

The occurrence of intense precipitation events related to deep convection can result in elevated values of DIL, clustering coefficient, and betweenness centrality. This phenomenon is particularly prominent in regions like the *Sierras de Córdoba*, known for experiencing some of the world's most severe deep convective storms (MULHOLLAND et al., 2018). These extreme events contribute to a higher concentration of connections between nodes in the network, thereby enhancing their connectivity and centrality measures. The atmospheric conditions conducive to convective activity, coupled with the orographic features of the *Sierras de Córdoba*, play a crucial role in the initiation and progression of intense convective storms, which can rapidly evolve into supercells and subsequently develop into MCSs (NESBITT et al., 2021). Nodes with higher DIL values are predominantly located within the region identified by Seluchi et al. (2003) as being associated with the eastern sector of the low-pressure system referred to as the northwestern Argentinean low (NAL). The NAL typically resides over northeastern Argentina and exerts significant influence on wind patterns and summer precipitation.

Figure 6.10 - Normalized results for Community SA3.



(a) Degree value and the Importance of Lines, (b) Closeness Centrality, and (c) Betweenness Centrality. The fraction of the blue line contained in the community represents the *Paraná* River.

SOURCE: Author's production.

Nonetheless, there is also higher activity observed in the northern portion of the community, as indicated by CC and BC. This suggests that the northern part plays an important role in the overall dynamics of the community. The heightened activity in this region can be attributed to the Bolivian hotspot. Gelbrecht et al. (2021) defined a reference area in the northern section of Community SA3 (13-18°S, 59-65°W) that serves as an indicator of SAMS dynamics. By analyzing wind anomalies in this region, they were able to identify the active or break phase of SAMS. During the active phase, the reference area exhibits strong positive wind anomalies, characterized by intensified winds along the eastern Andes. The authors also observed reversed sign (negative) in wind anomalies over the *Sierras de Córdoba* during the active phase of SAMS. Although DIL values in the north of the community are moderate, its contribution to the community's dynamics becomes evident through the CC and BC values, which show heightened values in the northeastern region. However, for this community, the DIL values are more predominant near the *Sierras de Córdoba* compared to the north of the community, indicating a more predominant activity over this region. The BC values highlight the highest activity in the western section of the community, roughly between 20°S and 30°S (east of the subtropical Andes), corresponding to the region referred to as the Eastern Central Andes (ECA) (BOERS et al., 2014b; BOERS et al., 2015).

Figure 6.10 also reveals the existence of a region situated in the southern part of the community, specifically south of Uruguay, which exhibits strong connectivity or pathways connecting this area to the eastern Andes. This connectivity is evident from the CC and BC values in this region. This pattern is associated with the cyclonic circulation depicted by anomalous wind fields during the active phase of the SAMS (GELBRECHT et al., 2021). The heightened activity in the western section of Community SA3 can be attributed to the influence of cold fronts originating from the south, which are facilitated by cyclone development in the southwest Atlantic. The presence of the Andes enhances this influence by obstructing the geostrophic balance and accelerating southerly winds, thereby promoting the movement of cold fronts across the continent. These cold fronts typically initiate their incursion from Patagonia towards northern Argentina and sometimes extend into the southwestern Amazon basin (MARENGO et al., 2004; SELUCHI et al., 2006).

As can be noticed in the previous analyses, the distribution of rainfall within these communities has specific characteristics. In order to better understand this distribution, analyses of rainfall within each community were performed in the next section.

6.1.7 Communities characterization

In this section, we present the analysis of precipitation behavior in each community, aiming to identify significant differences that contribute to their characterization. The section starts by presenting the statistical results of extreme events, followed by an exploration of the characteristics of deep convective precipitation events. This comprehensive examination will provide valuable insights into the specific features of these regions and their spatial precipitation patterns.

6.1.7.1 Extreme events distribution

In the previous section, the topological and dynamic characteristics of the communities were presented, highlighting the most active subregions. However, it is also possible to characterize each community through statistical analysis of the extreme events present within them. This provides us with information on the differences or similarities in terms of the frequency and intensity of these events. Table 6.2 was computed using extreme event series from each community, providing different statistical values in order to find the differences or similarities between them. According to Table 6.2, as anticipated, Community SA1 exhibits the highest average rainfall characteristics (mean and median), followed by communities SA2 and SA3.

Table 6.2 - Statistical information for extreme events in each community, with rainfall values given in millimeters per day (mm/day).

Metric	Community ID		
	1	2	3
Mean	43.17	42.62	33.52
Median	38.28	36.32	28.17
Mean - Median	4.89	6.3	5.35
Interquartile distance	19.67	24.83	21.83
Standard Deviation	16.48	21.87	18.23
Variance	271.63	478.30	332.35
21-years trend	-0.009	-0.013	0.002
Skewness	1.30	1.35	1.29

The skewness values for rainfall distribution are 1.91, 2.26, and 2.38 for communities SA1, SA2, and SA3, respectively. Community SA1 has the lowest skewness, indicating a less skewed rainfall distribution with a slightly more uniform distribution of rainfall values between communities. Community SA3 has the highest skewness. Concerning extreme events, Community SA1 shows the least variability. This is evidenced by the smaller differences between the mean and median (4.89 mm/day), interquartile range (19.67 mm/day), standard deviation (16.48 mm/day), and variance (271.63) compared to the other communities, suggesting that extreme events in Community SA1 have less variability in magnitude. Community SA3 ranks as the second community with relatively low variability, while Community SA2 exhibits the highest variability among the three communities. The greater diversity observed in communities SA2 and SA3 can be attributed to the spatial variations in the occurrence of various precipitation systems, such as shallow convection or deep convection. In contrast, community SA1 displays a more homogeneous spatial distribution, suggesting a dominant or more consistent type of precipitation system prevalent in that region.

In relation to the linear trends, the communities present in general terms extreme precipitation trends close to zero, which indicates a low change during the study period. However, trends for extreme precipitation events are similar to SA1 and SA2, but different for SA3. Community SA1 exhibited a weak negative trend, indicating a slight weakening of daily extreme events every 21 years. Community SA2 displayed a slightly more pronounced negative trend of -0.013 mm/day. This suggests that the weakening of extreme events in Community SA1 may influence extreme events in

Community SA2, possibly due to atmospheric circulation, but this needs a deeper investigation. The strong connectivity between communities 1 and 2 may contribute to the shared negative trends in extreme events. However, further analysis is necessary to confirm this hypothesis.

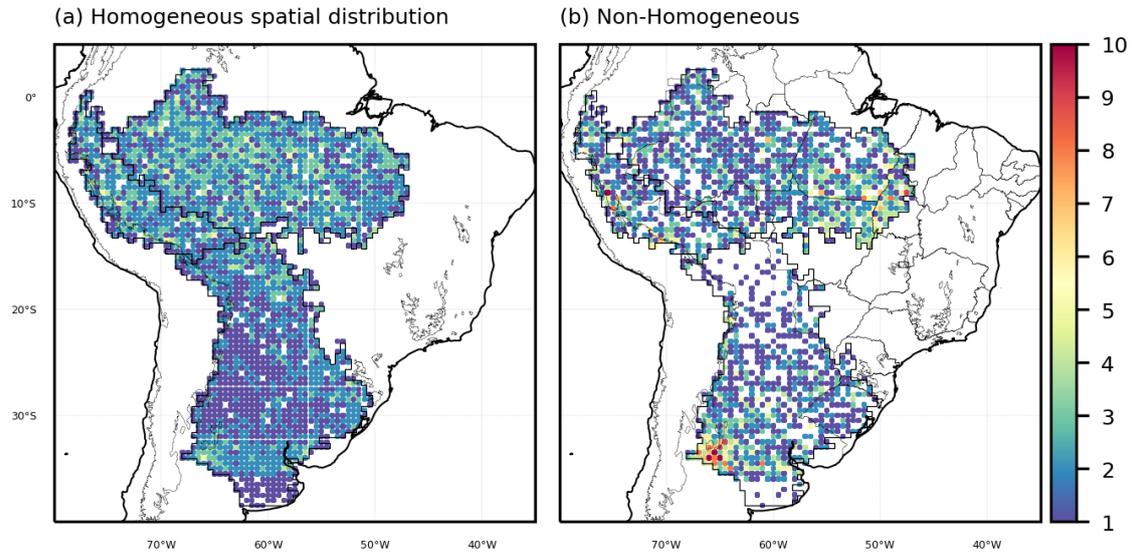
There are various studies examining these trends. For instance, [Marengo \(2004\)](#) identified negative rainfall trends observed across the entire Amazon basin using meteorological stations, with a negative trend in the northern and a positive trend in the southern at the regional level. Other studies have observed a negative trend specifically in the central part of the Amazon basin, particularly in the deforested areas of the Brazilian Amazon region. These negative trends could be linked to the removal of native vegetation cover ([FU et al., 2013](#); [DEBORTOLI et al., 2015](#); [LAWRENCE; VANDECAR, 2015](#); [SANTOS et al., 2015](#); [NOBRE et al., 2016](#); [SILVA JUNIOR et al., 2018](#)). It is believed that these deforested areas in the Amazon experience reduced precipitation due to changes in local atmospheric circulation patterns ([LIEBMANN; MARENGO, 2001](#); [HAGHTALAB et al., 2020](#)).

While, Community SA3 displays a weak positive trend of 0.002 mm/day, suggesting a slight increase in the intensity of extreme events every 21 years. Community SA3 shows strong local connectivity mainly in *Sierras de Córdoba*, with meaningful connections to the more subtropical region of South America. Although not considered in this analysis, it is acknowledged that the South Atlantic exerts significant influence in this area, as well as events originating from the southern part of the continent. However, a more comprehensive analysis would provide a more precise understanding of the dynamics in this area. A similar trend is observed in the degree boxplot (Figure 6.6). These findings indicate that Community SA1 and SA2 share similar characteristics in terms of extreme rainfall events and connectivity, whereas Community SA3 exhibits less intense extreme events compared to the other communities.

6.1.7.2 Deep convective precipitation

The following analyses aim to investigate intense rainfall and the tallest storm cells, which were identified by focusing on specific pixels with heavy rainfall and deep convection. More details on this approach can be found in Section 5.6.3. The results were derived by extracting data based on two spatial conditions: (a) a "homogeneous" spatial distribution, where we considered a local 95% threshold in each grid box (Figure 6.11a), and (b) a "non-homogeneous" spatial distribution, where we used a global 95% threshold (Figure 6.11b).

Figure 6.11 - Number of rainiest and deepest convection events for the two approaches.



SOURCE: Author's production.

Figure 6.11a provides the number of rainiest and deepest convection events identified using a homogeneous spatial distribution. A homogeneous distribution does not yield the same number of events at each grid point due to variations in event occurrences. This disparity can be attributed to factors such as satellite scan patterns (which follow a non-sun-synchronous circular orbit) and the inherent non-uniformity in the data distribution. This illustrates how the communities highlight specific regions with increased activity. Particularly, communities SA1 and SA2 show a more spatially homogeneous occurrence of the number of events compared to community SA3. This indicates a higher activity of rainiest and deepest convection events over communities SA1 and SA2, while community SA3 shows certain regions with higher activity.

Figure 6.11b illustrates the distribution of events based on the criteria of intense rainiest and deepest convection events for the non-homogeneous condition. Where, Community SA1 exhibits a widespread distribution across its entire area, with a noticeable concentration in the eastern section. This eastern region corresponds to the eastern part of the State of *Pará* and the northern part of the State of *Tocantins* in Brazil. This region is renowned for its elevated atmospheric electrical activity, which serves as an indicator of intense and well-developed storms (RASMUSSEN et al., 2014). Some studies have centered their attention on deep convection and lightning activity in this region (PETERSEN et al., 2006; ALMEIDA et al., 2012).

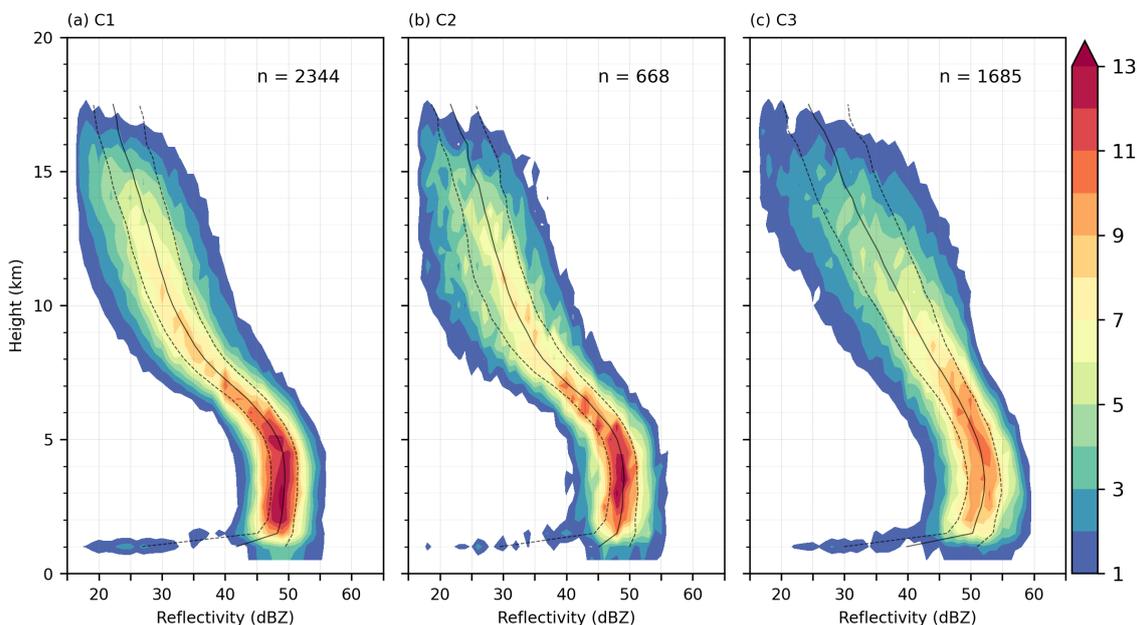
From Figure 6.8, it is evident that the eastern region of Community SA1 exhibits a low activity of Betweenness Centrality. This observation could potentially be linked to the occurrence of deep convection in that area. The low Betweenness Centrality values suggest that this region primarily represents localized activity and does not significantly contribute to the overall connectivity of the network. Similarly, Community SA2 shows a well-organized distribution that is similar to Community SA1, but with a slightly higher concentration towards the eastern side of the central Andes. This similarity indicates significant connectivity between these two communities in terms of the development of convective cloud processes, particularly for the rainiest deep convection events. Furthermore, the influence of the Andes on the precipitating systems can be observed once again. On the other hand, Community SA3 exhibits a more scattered distribution, with lower activity in the northern part of the community and a higher concentration in the southern part (6.11b), specifically in the southern region of the *Sierras de Córdoba*. This distribution of deep convective events aligns with the findings presented by Rasmussen et al. (2014).

The following analysis focuses on the vertical distribution of the deepest convective systems. This type of approach aims to identify some common characteristics or differences that can help to understand the dominant physical processes in each region. It should be noted that this type of study is important in defining the predominant physical processes in generating rainfall in numerical weather forecast models and in characterizing severe events for real-time monitoring.

Community SA1 exhibits a vertical distribution of precipitation, with the highest occurrence observed up to approximately 6 km, and the most intense reflectivity observed just below 5 km (Figure 6.12a). This specific structure, known as the "low-echo centroid" is commonly associated with intense precipitation events (HAMADA et al., 2015). The presence of the low-echo centroid indicates the development and continuity of high rainfall rates, highlighting the significance of warm rainfall processes within this community (HUANG et al., 2019). The low-echo centroid serves as an indicator of persistent and substantial rainfall activity in the area.

For Community SA2 (Figure 6.12b), the vertical profile of reflectivity exhibits similarities to Community SA1, albeit with slightly lower reflectivity occurrence values in the lower levels. This disparity in occurrence could potentially be attributed to the influence of orographic forcing, which impacts precipitation processes. The orographic features in the region may play an important role in modifying the vertical distribution of precipitation and influencing the frequency of intense rainfall events

Figure 6.12 - Contoured frequency by Altitude Diagram (CFAD).



The frequency of reflectivity values at different heights in percent for deep convection rainfall is shown in a homogeneous spatial distribution. The value "n" indicates the number of samples for the corresponding event. The three black lines represent the reflectivity at the 25th, 50th, and 75th percentiles at each level, respectively.

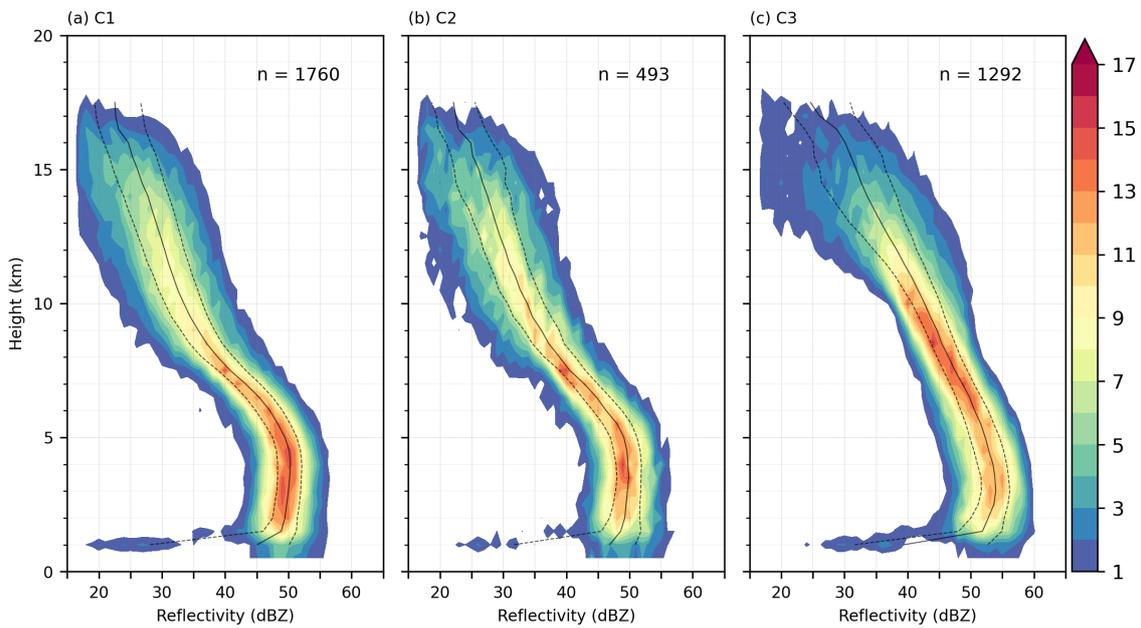
SOURCE: Author's production.

(KUMAR et al., 2019b; CHAVEZ et al., 2020). Moving on to Community SA3 (Figure 6.12c), there is a higher occurrence of maximum reflectivity at middle levels, specifically around 5 km. This indicates a more prominent presence of mixed-phase clouds in this community. Furthermore, Community SA3 exhibits the smallest vertical gradient intensity in maximum reflectivity.

Moreover, when comparing the results with the non-homogeneous distribution, all communities (Figure 6.13a-c) exhibit more pronounced values at higher levels. These values, reaching 45 dBZ echo top above the freezing level, serve as a reliable indicator of hail occurrence (WALDVOGEL et al., 1979; NI et al., 2016; MROZ et al., 2017). This suggests that the lifting processes in these regions with deeper convection are more efficient compared to the entire community (homogeneous distribution in Figure 6.12a-c). Particularly in Community SA3, the profiles demonstrate a greater vertical development. These findings also imply that in these regions, with greater vertical development, there may be a higher frequency of electrical discharges and hail.

The occurrence of reflectivity values around 50 dBZ between 6 and 7 km further suggests the potential for larger hail due to intense updrafts. The relationship between the depth of the melting level and the height of the 50 dBZ echo has been investigated by Donavan and Jungbluth (2007), revealing a strong linear correlation. This emphasizes the importance of gaining a better understanding of the microphysical processes of ice formation in these areas.

Figure 6.13 - Contoured frequency by Altitude Diagram (CFAD).

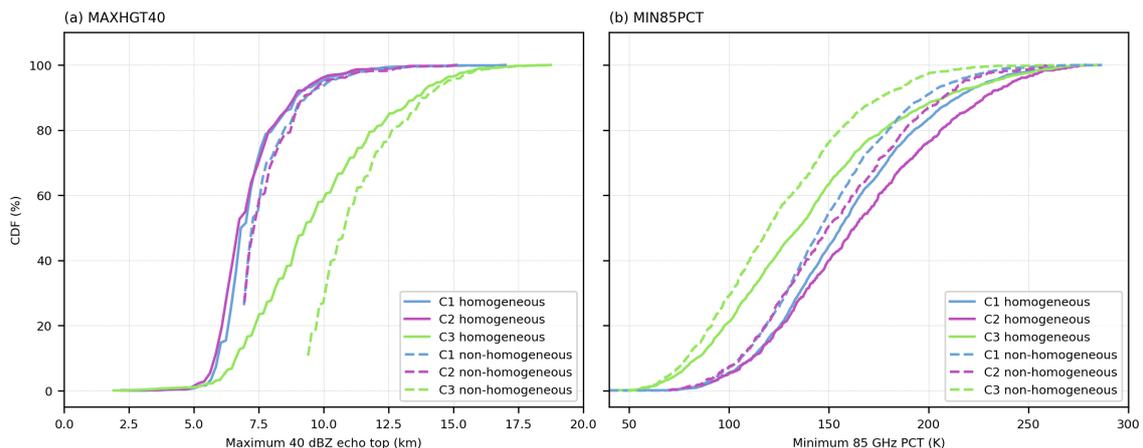


Same that Figure 6.12, but for non-homogeneous condition.
SOURCE: Author's production.

The following analyses confirm the hypotheses mentioned earlier. The presence of hail and the intensity of convection were examined using the cumulative density function (CDF) of MAXHT40 (maximum altitude of echo top 40 dBZ) and MIN85PCT (minimum polarization-corrected temperature at 85 GHz) for both homogeneous and non-homogeneous conditions (Figure 6.14). The results reveal that Community SA3 has a higher probability of encountering high storm heights and a greater likelihood of hail occurrence. In contrast, Communities SA1 and SA2 exhibit lower storm intensity and less frequent hail events. These differences between Community SA3 and the other two communities become even more pronounced in regions with higher occurrences of rainiest and deepest convective events (non-homogeneous conditions). Specifically, Community SA3 displays significantly higher echo top heights

at 40 dBZ compared to the other communities (Figure 6.14a). This characteristic is noteworthy, as even the uniform events in Community SA3 have a higher probability of reaching greater heights compared to the non-uniform events in the other two communities. Moreover, these higher storm heights in Community SA3 are closely associated with a greater presence of hail, indicated by the highest cumulative probability of MIN85PCT being less than 130 K. Previous studies have identified this temperature threshold (130 K) as an indicator of hail presence (CECIL; BLANKENSHIP, 2012). Overall, these results highlight the distinctive behavior of Community SA3, characterized by higher storm heights and a higher likelihood of hail compared to the other communities. This information provides valuable insights into the intensity of convection and the associated hail risk, contributing to a better understanding of severe precipitation events within each community.

Figure 6.14 - The cumulative distribution function (CDF).



(a) maximum altitude of 40 dBZ echo top (MAXHT40) and (b) minimum 85 GHz polarization-corrected temperature (MIN85PCT) is shown for both homogeneous and non-homogeneous conditions.

SOURCE: Author's production.

These findings indicate notable distinctions in convective systems among the communities, particularly above the 5 km height level. When considering the region with the highest occurrence of deep convection, communities SA1 and SA2 demonstrate nearly indistinguishable reflectivity profiles. This suggests that the convective events in these communities share similar attributes in terms of reflectivity intensity and vertical growth. Conversely, community SA3 exhibits distinctive characteristics in both scenarios, indicating significant disparities compared to the other communities.

6.2 AMAZON networks

As shown in previous results, communities SA1 and SA2 are located over the Amazon Basin. They share many similar characteristics, most notably in their rainfall distribution. The Amazon Basin is the largest river basin in the world and a region of great importance for the climate of South America. Therefore, it is possible that this region has great importance in defining the dynamics of the South America network and, consequently, in the communities. Therefore, deeper analyses were carried out in this specific region. The goal was to clarify several aspects, such as the influence of the basin in the definition of previous communities, the capacity of even more localized communities, and the influence of other techniques in the definition of communities. Therefore, two approaches were performed to better understand the characteristics of the network and its regional dynamics. The first (Section 6.2.1) was based on the identification of communities with local extreme events in that region, following the same principles adopted in the analysis of extreme events by the SOUTH AMERICA network. Understanding if there are even more communities located in this region and the influence of the basin in the previous results. The second (Section 6.2.2) aims to understand how different techniques can influence the definition of communities.

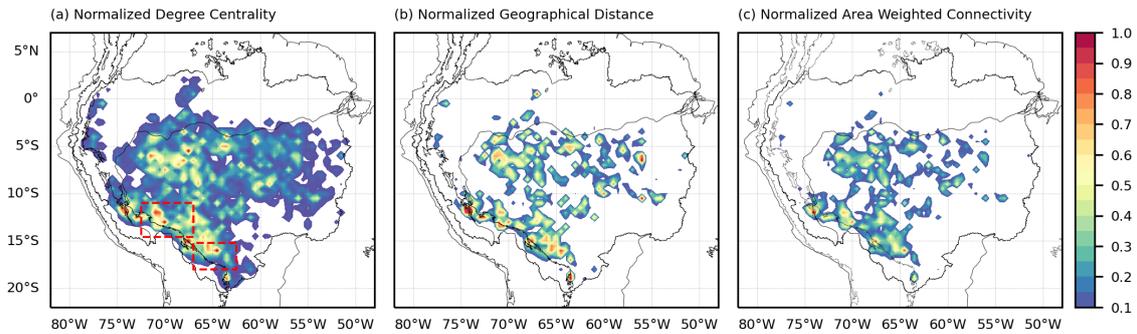
6.2.1 Extreme rainfall network

This section follows the same line of reasoning as the previous one, which presented the results of the extreme event network in South America. Here, the focus is the Amazon basin, as mentioned in Section 5.1.2. The analysis involves the extraction of essential information from the network related to the precipitation system and its prevailing mechanisms. Furthermore, the use of the AMAZON network allows us to gain insights into the specific dynamics within the Amazon basin. This comparison provides an opportunity to identify potential differences between studying the isolated region of the AMAZON network, and considering it as part of the broader SOUTH AMERICA network, as well as to verify the sensitivity of the network when analyzing this individual region.

This network was created using rainfall daily data (IMERG) for the summer months from December 2000 to February 2021. The extreme events over the Amazon basin were defined as events greater than or equal to the 90th percentile. The similarity function used is the event synchronization with 3 days. While the threshold was determined by the 98th percentile of the connections (BOERS *et al.*, 2013). Figure 6.15 presents the normalized results of Degree centrality, Geographical distance, and

Area Weighted Connectivity (AWC). These results show how the greatest activity of the extreme event network is concentrated to the south of the Amazon River, mainly over the central-western region, as indicated by [Anselmo et al. \(2021\)](#).

Figure 6.15 - Normalized index for AMAZON network.



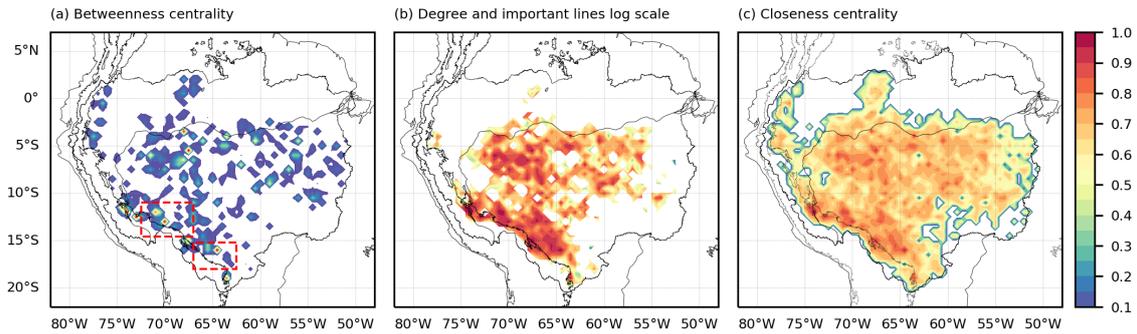
(a) Degree centrality, (b) Geographical distance, and (c) Area weighted connectivity. Red boxes represent the Peruvian and Bolivian rainfall hotspots.

SOURCE: Author's production.

Figure 6.16 provides the results of BC (a), DIL in log scale (b), and CC (c). Figure 6.16a shows that the BC exhibits high values to the east, as well as in two regions to the north of the basin. This indicates the displacement of extreme events from these areas. It should be noted that the DIL results (Figure 6.16b) show high values to the south of the Amazon River ([ANSELMO, 2015](#)) and slightly more intense to the west of the basin, regions associated with the Peruvian and Bolivian precipitation hotspots ([ESPINOZA et al., 2015](#)). This indicates the high connectivity between the center and west of the basin. Figure 6.16c depicts how CC presents a similar spatial pattern exhibiting higher values to the east of the central Andes, indicating greater efficiency in information transmission in that region. This increased efficiency and connectivity over the eastern Andes is associated with a change in the direction of low-level moisture flow, creating a corridor with predominant southward displacements.

The results of Figures 6.15 and 6.16 are consistent with the results obtained from the SOUTH AMERICA network (from Figure 6.6 to Figure 6.10), indicating the presence of higher network activity in the center-west of the Amazon basin and the center-south of the central Andes. This demonstrates the similarity in the highest network activity between both networks. However, when focusing on a specific region, many links that are outside this area are eliminated. Consequently, while the

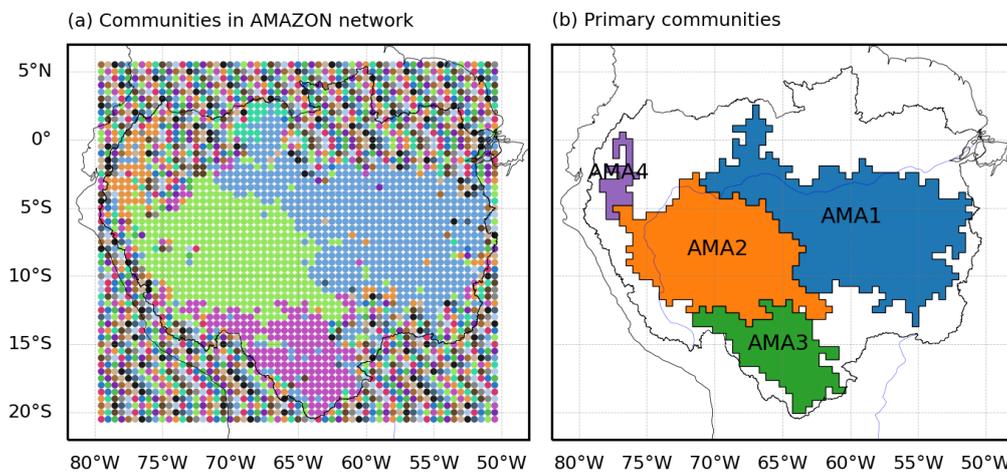
Figure 6.16 - Normalized index for AMAZON network.



(a) Betweenness centrality, (b) degree and important lines in log scale, and (c) Closeness centrality. Red boxes represent the Peruvian and Bolivian rainfall hotspots. SOURCE: Author's production.

overall dominant dynamics may remain unchanged, it is expected that removing these connections will alter the subgroups. To verify these differences, Figure 6.17 displays the communities in the AMAZON network.

Figure 6.17 - Communities in extreme rainfall AMAZON network.

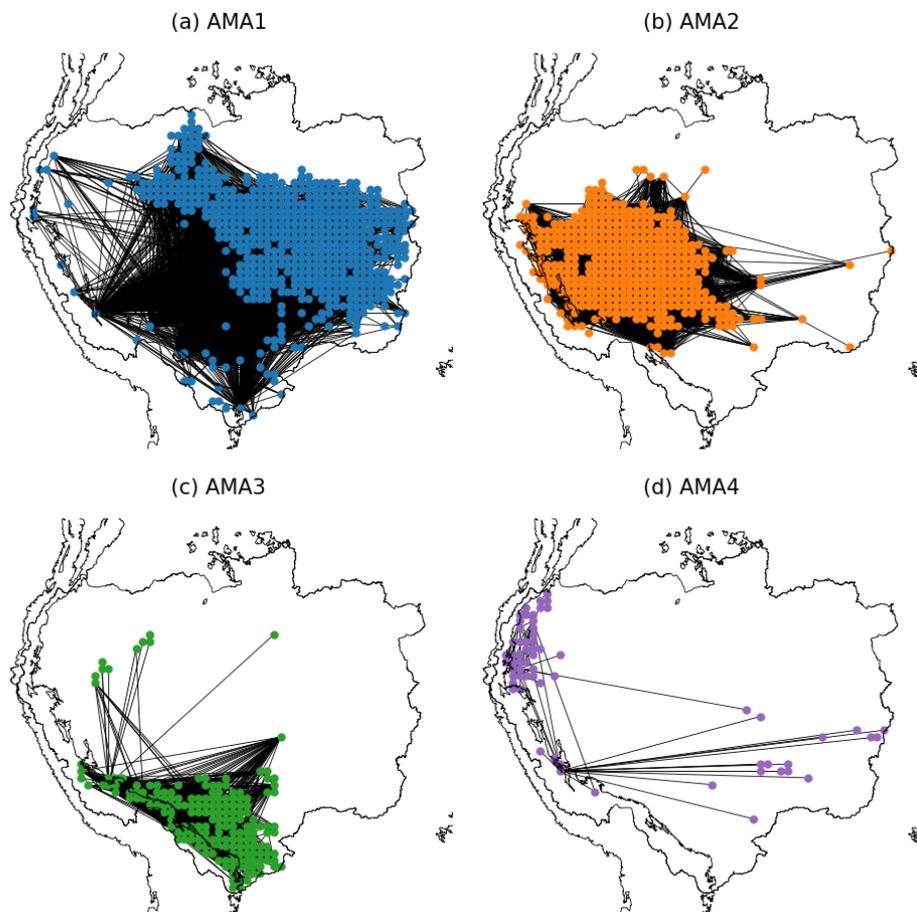


(a) Communities for all nodes while (b) represents the major communities and their IDs. SOURCE: Author's production.

Figure 6.18 provides information about the connections within the communities. The AMA1 community (Blue in Figure 6.18) exhibits the highest connections within the Amazon basin, with connections extending over the eastern Andes. This community is associated with the eastward spread of extreme winds (MACHADO et al., 1998).

The AMA2 community (Orange in Figure 6.18) includes a portion of the Amazon basin and the Amazon-Andes transition region. It encompasses well-developed extreme events, and topographical characteristics that facilitate precipitation over the central Andes (GARSTANG et al., 1994). The AMA3 community (Green in Figure 6.18) is primarily centered around the Peruvian and Bolivian hotspot regions (ESPINOZA et al., 2015). This region acts as a moisture sink from the Amazon basin. In addition to presenting a high percentage of precipitation related to MCS (FENG et al., 2021). Lastly, the AMA4 community highlights the influence of the northern part of the community on the Peruvian hotspot.

Figure 6.18 - Communities edges for the AMAZON network.

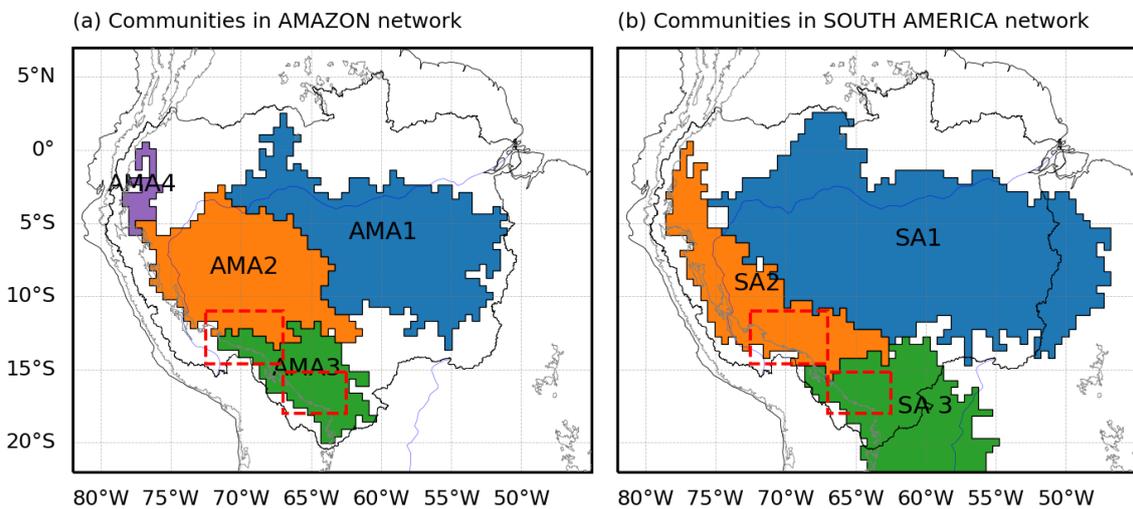


SOURCE: Author's production.

As previously mentioned, one of the objectives of this Amazon basin network is to verify whether there are differences between the communities in the South America network originating in this area.

To compare the differences between the extreme event networks, Figure 6.19 illustrates the communities within these networks. It can be observed that the AMAZON network consists of 4 communities, while the SOUTH AMERICA network has 3 communities. Where the main differences are between the AMA1, AMA2, and AMA4 communities with SA1 and SA2. Since SA1 covers a larger area of the Amazon basin compared to AMA1. While SA2 contains AMA4 and part of AMA2. Therefore, it is evident that even with the observed differences, the locations of the communities are similar, with AMA2 expanding its area towards SA1. This shows that internally, there are connections that favor links between the central west of the Amazon and the east of the Andes.

Figure 6.19 - Communities in extreme rainfall networks.



(a) AMAZON network and (b) SOUTH AMERICA network. Red boxes represent the Peruvian and Bolivian rainfall hotspots.

SOURCE: Author's production.

While AMA3 presents an interesting characteristic since this community contains the Bolivian and Peruvian hotspots. This does not happen in SA3 which only contains the Bolivian hotspot. This difference may be associated with the importance of the Bolivian hotspot on the dynamics in the extratropics of South America. This is because this region is influenced by the Chaco low and the northwestern Argentinean low (NAL) (SELUCHI *et al.*, 2003), which in turn is associated with the different phases of SAMS (GELBRECHT *et al.*, 2021). The communities in both networks present a similar spatial pattern, except that the AMAZON network provides more specialized communities, as can be seen in community AMA4.

Such results show that more local communities corroborated the main characteristics of the communities in South America, bringing more details about the internal structure of the region. This shows the importance of the Amazon basin in defining extreme event communities. In the following section, the total rainfall within this region will be analyzed in the definition of communities, since it is expected that they have more importance in the characteristics of the basin.

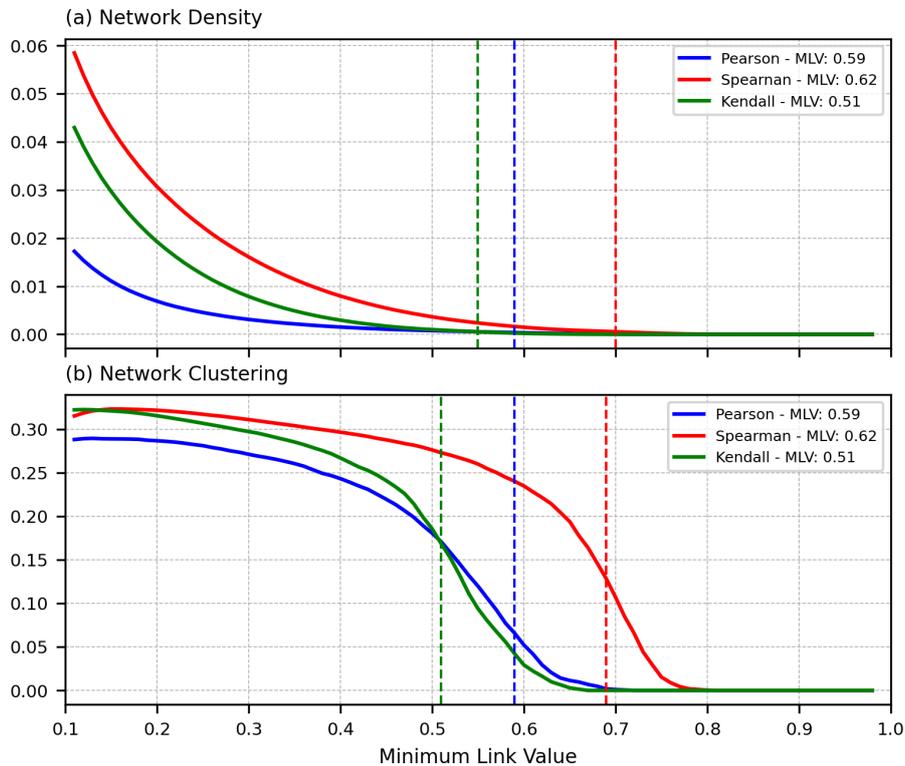
6.2.2 Total rainfall networks

The information about rainfall can help us better understand the mechanisms governing the region and the potential relation between precipitation and the rivers. This section presents the results of the differences between three correlation metrics (Pearson, Spearman, and Kendall) used as the similarity function for the rainfall network over the Amazon basin during the austral summer. As mentioned in Section 5.2.2, three rainfall networks were created using IMERG data from December 2000 to February 2021. After calculating the correlations between all possible vertex combinations, a threshold was set to define the strongest connections. This threshold was determined by the 98th percentile of the connections (BOERS et al., 2013), and this criterion was applied to all three networks. Therefore, only the strongest connections, corresponding to 2% of the total, are considered in the subsequent analyses.

It is known that the threshold used to define connections in a network can result in excessively connected or fragmented networks. These changes also depend on the metric used to calculate the links. In order to assess sensitivity, Figure 6.20 presents the variability of the three rainfall networks with respect to different connection thresholds. This analysis was performed using network density and average clustering coefficient. Figure 6.20a shows that the Spearman network exhibits the highest density values across the connection thresholds, followed by the Kendall network, and finally, with significantly lower values, the Pearson network. This indicates that the Spearman network has a greater number and intensity of connections compared to the other networks.

In Figure 6.20b we can see again how Spearman provides the highest connectivity values. It is important to note that the difference between the minimum link values between 0.5 and 0.75 shows the most notable differences between Spearman with Pearson and Kendall.

Figure 6.20 - Variations in the density and clustering for AMAZON Networks.



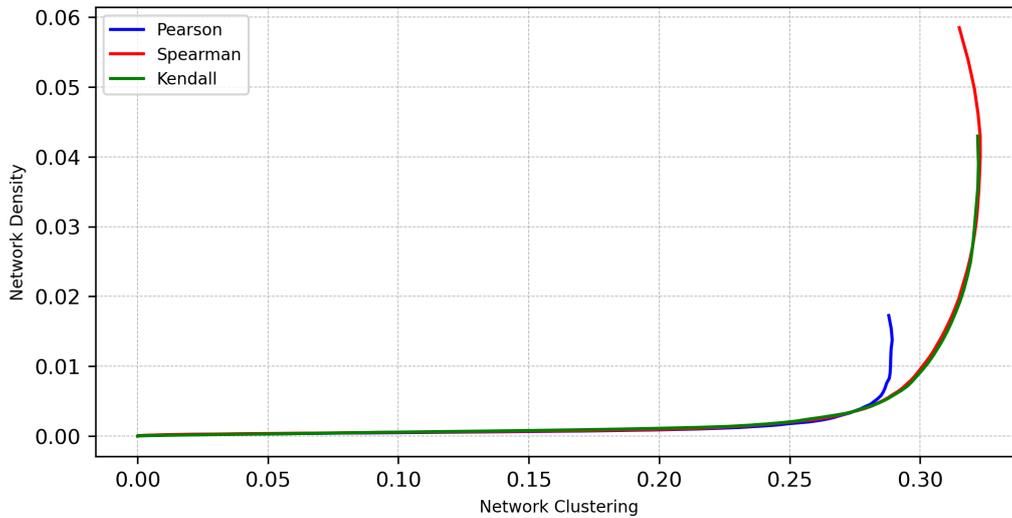
The dotted lines represent the minimum values to define connections in each network obtained from the 90th percentile.

SOURCE: Author's production.

Although these results demonstrate the behavior of these two indices with respect to the thresholds, Figure 6.21 depicts the changes in these networks concerning the mutual relationship between density and clustering coefficient. This figure shows how the networks exhibit a low response of clustering for very low-density values, indicating virtually fragmented networks. Then, higher activity is observed for clustering values above approximately 0.27. Overall, the Spearman and Kendall networks exhibit more similar behaviors, but the Spearman covers a wider range of values in the relationship between clustering and density. This indicates that the network will become fragmented for very high connection thresholds, which is the main difference compared to Pearson, which has a more broken structure.

The rapid fragmentation of the Pearson network may be associated with the nature of the linear relationship of this function, resulting in a low number of vertices with high degrees.

Figure 6.21 - Relationship between density and clustering of AMAZON networks.

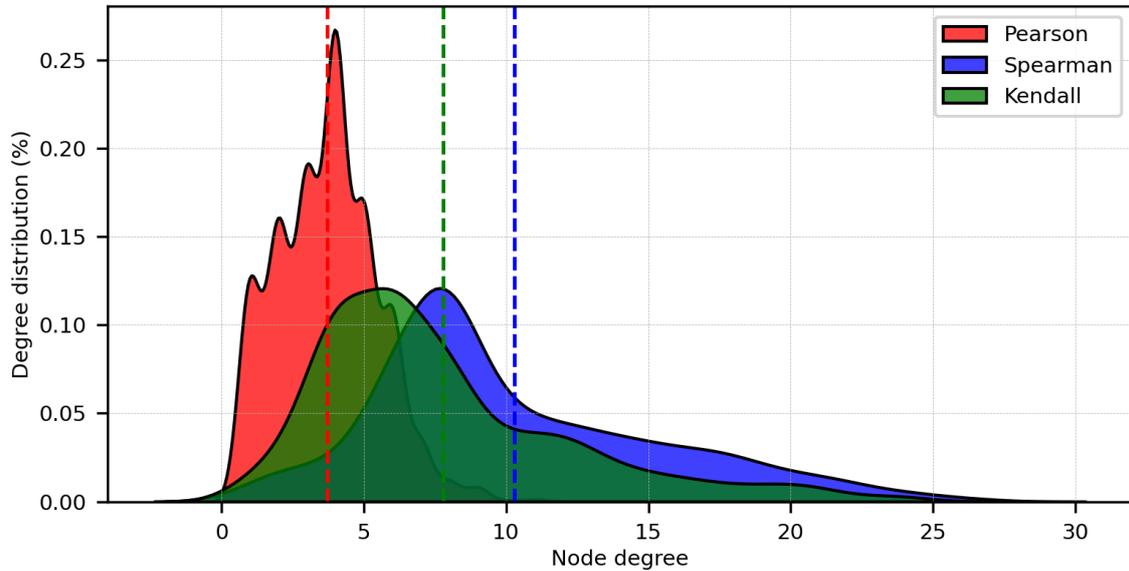


SOURCE: Author's production.

6.2.2.1 Networks analysis

This analysis is based on the variations of the degree distribution. The degree of a node represents the number of edges that are connected to that node and also provides information regarding its importance or possible influence on the network. To analyze the distribution of degrees, Figure 6.22 presents the degree distribution of the networks. In this figure, it can be observed the significant difference between the Pearson network compared to the others. The Pearson network exhibits a scarcity of vertices with high degrees, where its maximum degree is 11, while the Spearman and Kendall networks are 27 and 25, respectively. Additionally, the Pearson network also shows a median (4) higher than the mean (3.7), indicating a left-skewed distribution, which means that the most frequent degree occurrences are concentrated in values smaller than the median. This difference is noticeable when comparing it to Spearman and Kendall. In addition to having a longer tail, these networks show vertices with high degrees and means that are higher than the medians. Based on these topological results, it can be observed that the Pearson network is providing suspicious information associated with sub/under-estimations, as it exhibits a lower connected network. This difference can be attributed to the low presence or intensity of connections, as evidenced by the low maximum degree of this network. While Spearman and Kendall, they are present quite similar results, with the difference that Kendall provides lower grade values compared to Spearman.

Figure 6.22 - Degree distribution of the networks.



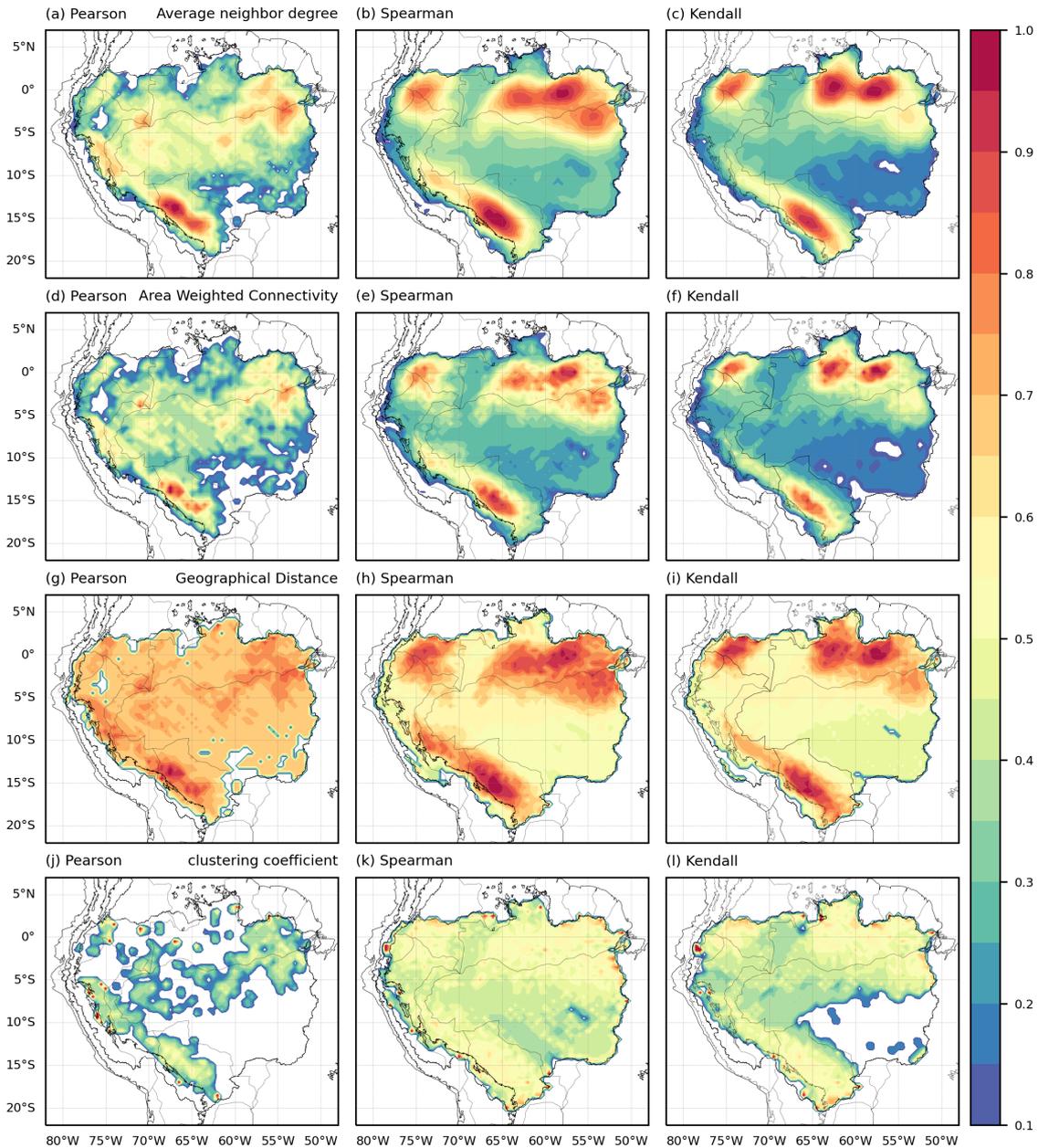
SOURCE: Author's production.

Figure 6.23 presents the results of different indexes for the Amazon networks. The results are components for the normalized AND, AWC, GD, and CC. These indexes allow us to identify central and influential regions in terms of their connections.

We can observe that the three networks exhibit similar spatial structures, with the biggest difference being presented in the Pearson network. Specifically, in the northern regions, both Spearman and Kendall networks show well-defined regions, while Pearson exhibits lower connectivity and homogeneity, which can be easily seen in the CC (Figure 6.23j). This figure also shows the lower degree of connectivity present in the Pearson network.

The AND, AWC, and GD in Figure 6.23 show a more similar spatial distribution between Spearman and Kendall networks. In both networks, we can observe well-defined regions with densely connected neighborhoods. These indexes reveal the presence of two regions in the north of the Amazon basin and one region to the southeast of the central Andes. These regions are identified as areas where nodes have a high number of connections with their neighbors. These spatial patterns can be interpreted as indicators of regions with greater intensity of interaction or influence within the rainfall network.

Figure 6.23 - Normalized index for AMAZON networks.



SOURCE: Author's production.

In fact, when comparing the structure of the Pearson network with Spearman and Kendall, we can observe that Pearson fails to identify the importance of the northern region of the Amazon basin. Pearson shows lower connectivity values in the northern region, indicating weaker connections between neighboring nodes. However, it is worth noting that the Pearson network does identify the importance of the region near the mouth of the Amazon River, although to a lesser extent.

The maximum distances recorded for GD were 90 km, 95 km, and 105 km for Pearson, Spearman, and Kendall respectively. These results reflect another fact that the Pearson network has shorter connections, indicating a bit lower distance connections compared to the other two networks.

On the other hand, the Spearman network has longer connections, which may be related to its identification of active regions with larger areas. As observed in the results of CC (Figure 6.23j, k and l). The Spearman and Kendall networks exhibit similar spatial patterns, indicating stronger cohesion in the same regions of the Amazon basin. However, the Pearson network shows values in smaller and slightly lower clusters, suggesting lower cohesion in this network.

There is a particularly noticeable difference in the central-east region of the basin, where clustering is better defined. This result is due to the linear nature of Pearson's correlation. The common region between the three networks is the eastern Andes. Where all networks show a similar spatial pattern, particularly above the Bolivian hotspot. This similarity in response can be attributed to the intense precipitation in the region and the predominantly southward direction of low-level flow caused by the presence of the Andes. As a result, a well-defined behavior is observed in this region across all networks.

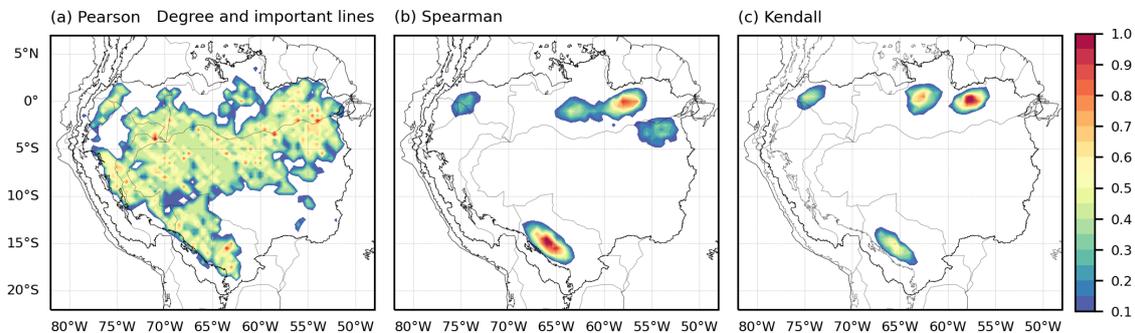
Therefore, when identifying communities or clusters in the network, it is likely that the Spearman and Kendall networks will produce more similar results, while the Pearson network may identify smaller communities. It is important to note that although Spearman and Kendall show a high degree of similarity, they exhibit a noticeable meridional difference (north-south), particularly over the northern part of the Amazon basin.

To obtain better information about these important regions, we will use an indicator that directly quantifies the importance or participation of the nodes, for instance, DIL. This index provides the node's participation based on its degree and the importance of its connections. The importance of connections is computed by the number of triangles in which the node participates¹. Therefore, from the number of connections together with the participation in triangles, the importance of a node is quantified. The results of DIL for each correlation method, provide information about the node participation in the network.

¹A triangle is a configuration in which three nodes are interconnected, forming a closed cycle. Triangles are especially relevant in graph theory as they can be used to measure the connectivity and transitivity of a network.

In Figure 6.24 we can observe that the Pearson network highlights areas with high DIL values in the middle of the basin and to the east of the Andes. In contrast, the results of Spearman and Kendall show a notable difference as they do not exhibit a spatially homogeneous DIL distribution, where both networks indicate spatially well-defined regions. Instead, they show the presence of distinct regions in the north and southeast of the basin. However, all three networks reveal areas related to the spatial structure of moisture transport. Spearman and Kendall's networks particularly emphasize the entry and exit regions of the moisture flow (VERA et al., 2006a).

Figure 6.24 - Degree and important lines.



SOURCE: Author's production.

For Spearman and Kendall Three main regions are identified. The first region, located in the northeast of the basin, extends from the mouth of the Amazon River and covers a larger spatial extent. While the second region, located in the northwest, is the smallest. These regions are associated with the influx of warm and humid air from the Tropical Atlantic Ocean, which contributes significantly to positive moisture divergence (ZHOU; LAU, 1998; MARENGO, 1992; MARENGO et al., 2003).

These regions to the north of the basin are also related to an intensification of the ascending branches of the Hadley and Walker circulations (ESPINOZA et al., 2022). The third region is situated in the southwest of the basin, specifically east of the Andes. This region can be considered an outlet or sink of the Amazon basin, as winds travel from the eastern part of the central Andes towards the extratropics of the continent (MARENGO et al., 2012; MONTINI et al., 2019; JONES, 2019). The results suggest that these three regions have a concentration of nodes with higher degrees, indicating that they are more central and influential areas within the rainfall network.

Therefore, the two northern areas would be related to rainfall over the central region of the basin, while the third region in the southeast could be interpreted as a response to the flow coming from the central region. The spatial distribution of rainfall displayed in previous results is closely related to the presence of the SAMS. This distribution pattern is more evidenced by the Spearman and Kendall networks, this is consistent with other studies that have investigated the dynamics of the SAMS in the region (MARENGO et al., 2012; MONTINI et al., 2019; JONES, 2019; GELBRECHT et al., 2021). These studies have also observed a significant influence of the SAMS on rainfall distribution and the formation of spatial patterns in the Amazon basin.

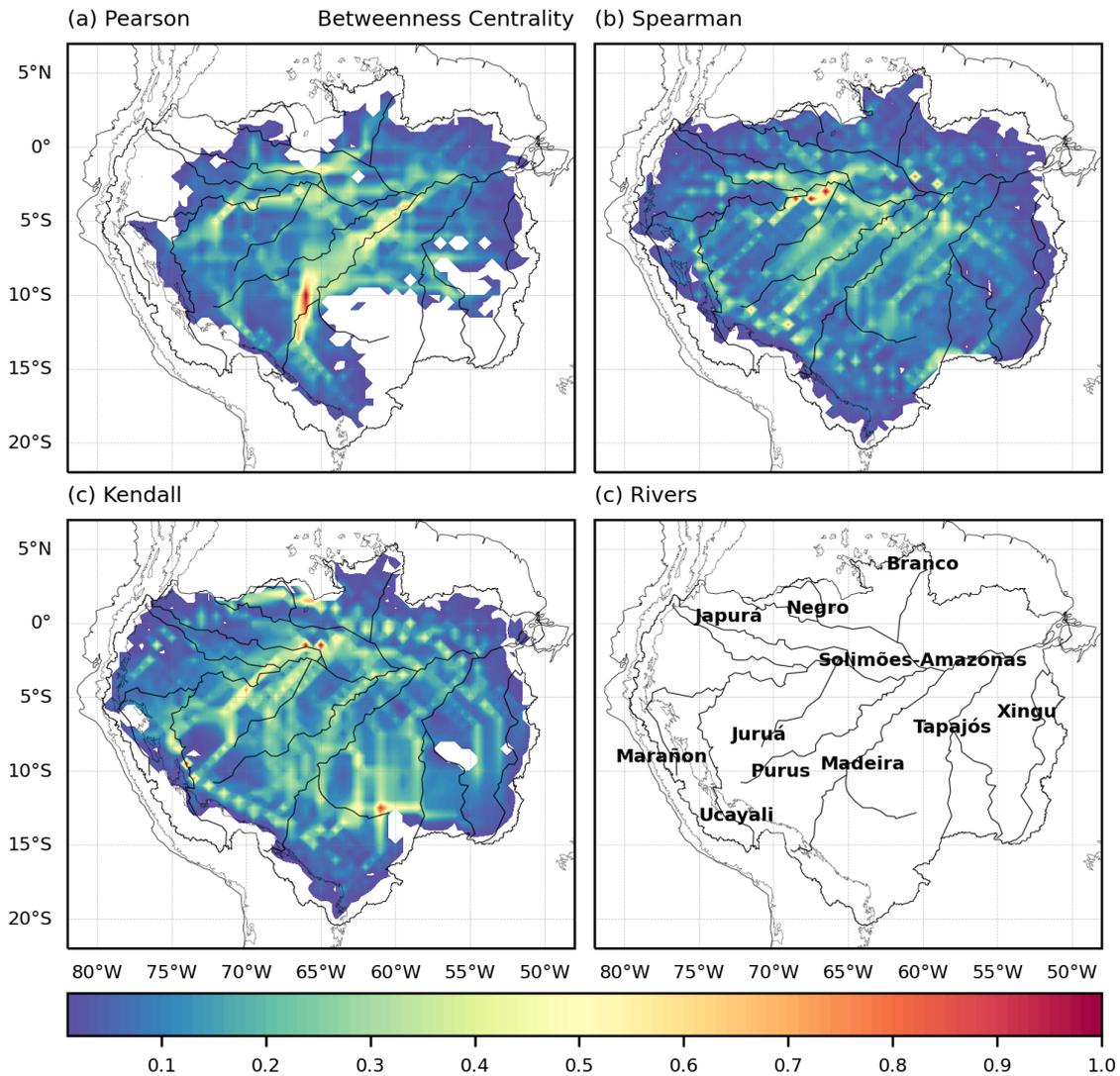
Therefore, the results of the correlation networks are consistent with the current understanding of the SAMS influence in the region. In addition to analyzing the importance of individual vertices, it is also relevant to investigate the existence of predominant paths that are frequently used in information transmission. This analysis is conducted using the Betweenness Centrality.

Figure 6.25 presents the results of BC, where the three networks exhibit the pattern associated with the moisture advection. These systems begin with the arrival of east winds at lower levels, which provide moisture and promote convective activity, resulting in increased rainfall in the region (COHEN et al., 1995). As convection produces more precipitation, evapotranspiration increases the humidity of the atmosphere at low levels, while easterly winds carry moisture towards the Andes (ELTAHIR; BRAS, 1996; SATYAMURTY et al., 2013). This phenomenon explains the high BC values observed in the central part of the Amazon basin. Additionally, when these precipitating systems from the North Atlantic Ocean reach land, they provide high BC values near the mouth of the Amazon River in the northeast region of the basin (BOERS et al., 2013). This observation is directly related to the previous results of DIL, as shown in Figure 6.24. The BC could be interpreted as the more common path for information transmission in the rainfall networks.

It is evident that the Pearson network exhibits higher values of BC focused in the central region, mainly between the Purus and Madeira rivers, while Spearman and Kendall show a more distributed response. This indicates that Spearman and Kendall have paths that cover a larger area of the Amazon basin, displaying a broader distribution. The three networks show a significant response of BC in the south of the Amazon River, indicating a major activity of rainfall. Similar to mentioned by (ANSELMO et al., 2021), the authors suggest that during the austral

summer, the most intense rainfall activity is primarily observed south of the Amazon River. This is associated with high levels of surface water elevation of the rivers over the center and center-west of the basin (FASSONI-ANDRADE et al., 2021).

Figure 6.25 - Normalized Betweenness centrality.



SOURCE: Author's production.

The networks also exhibit a good response to the presence of the Andes, displaying a well-defined path, mainly in the central part of this region. This is evident from the high BC values observed to the east of the Andes, which indicate the strong influence of orographic rainfall processes in the region (BOOKHAGEN; STRECKER, 2008; MONTINI et al., 2019; JONES, 2019).

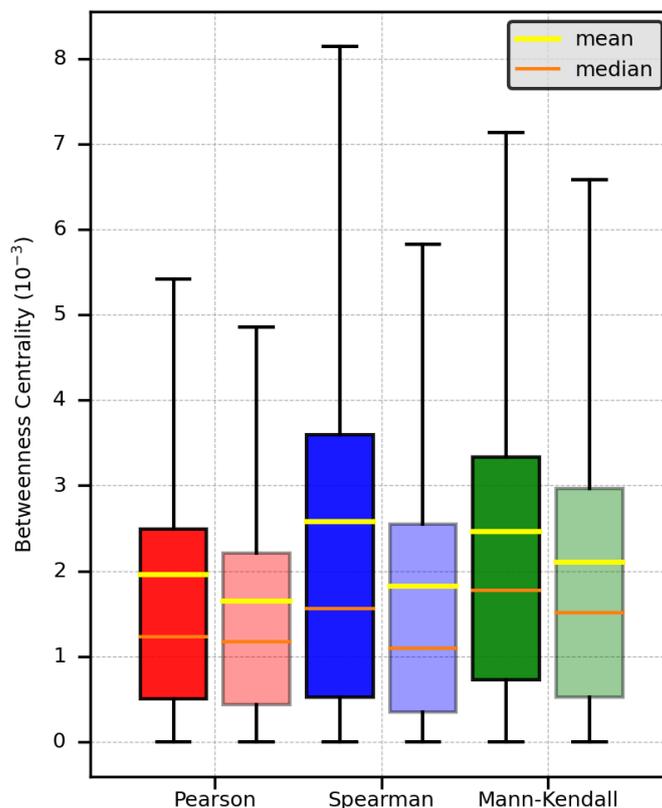
Pearson exhibits low BC values in the eastern part of the basin localized in the northeast of *Mato Grosso* (52.5°W - 57.5°W and 10°S - 15°S), while Spearman and Kendall show BC activity over this region. This region is known for the presence of mesoscale convective systems, where these events are propagated westward (ANSELMO et al., 2021).

It can be observed how some BC values are higher around rivers (represented by black lines inside of the basin map), indicating a response to the possible influence of land cover on precipitation. This relationship between river and land is known as the breeze effect, and several studies have provided insights into this phenomenon.

The land-river breeze effect is strongly linked to an increased occurrence of rain events during the morning and nighttime hours in the rivers and vicinity (LU et al., 2005; RAMOS DA SILVA et al., 2011; TANAKA et al., 2014), while a decrease is observed in the afternoon (FITZJARRALD et al., 2008; PAIVA et al., 2011). This effect is influenced by temperature differences between land and water, which create air circulation patterns that result in cloud formation and subsequent rainfall. Essentially, during the early morning hours, the temperature of the river surface tends to be warmer than the temperature of the land surface, leading to upward movement of air over the water and subsidence over the land. This convective effect enhances rainfall over the river surface. The three networks show different regions with better responses between the BC and the presence of rivers. For example, on the Amazon River, the networks exhibit high values of BC. Pearson and Spearman demonstrate a strong relationship of BC with the *Purus*, *Madeira*, and *Juruá* rivers, while Spearman and Kendall show this relationship on the *Marañón* river. Finally, Spearman shows the behavior of BC is more parallel to the rivers. It is also observed that Kendall presents greater BC activity on the north of the *Negro* river, indicating how this network assigns greater importance to this region.

To assess the degree of BC activity over rivers in the networks, two groups of nodes were defined based on whether they were located on a river or not, considering only BC values greater than 0 (blue regions are ignored). The results of these groupings are presented in Figure 6.26. It is evident that all three networks show how nodes with river presence (represented by boxes with dark colors) exhibit higher BC mean and median values compared to nodes without river presence (represented by boxes with light colors). These results demonstrate how all three networks are able to provide results associated with the land-river breeze effect, with more pronounced outcomes in the Spearman and Kendall networks.

Figure 6.26 - Betweenness centrality boxplot for nodes with/non-rivers.



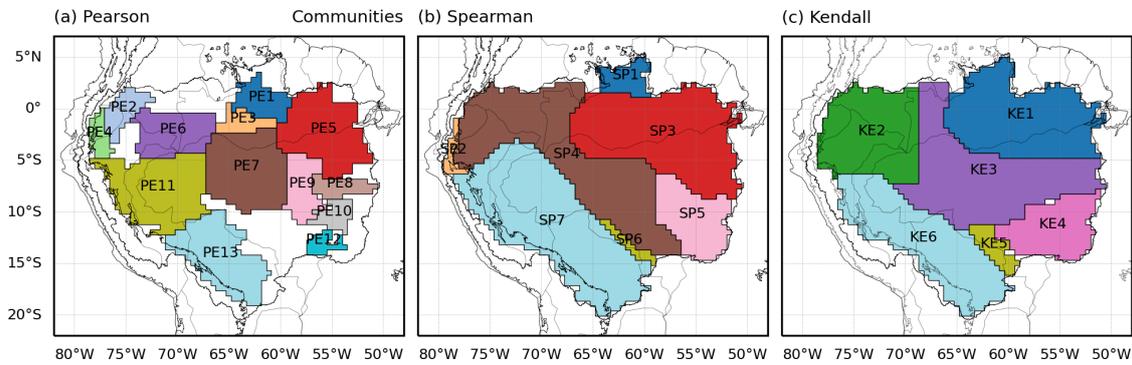
The dark and light boxplot colors represent the vertices groups with and without the presence of rivers, respectively.
 SOURCE: Author's production.

The BC in the networks show the more common paths of the precipitating events, which supports the east-west circulation pattern. However, Spearman provides a better relationship between the BC path with rivers, while Pearson focuses more on the center of the basin. Kendall also showed parallel-river behavior but tends to have a more traverse-river path. In addition to the boxplot results, it can be understood how Spearman better represents the association with river locations, related to the land-river breeze effect.

Figure 6.27 illustrates the results of community identification. It is evident that in the northeastern part of the Amazon basin, the Pearson network identified three distinct communities (PE1, PE3, and PE5), the Spearman network identified only one community (SP3), and the Kendall network presents one community (KE1). For the northwest region, the communities are PE2 and PE4 for Pearson, SP2, and SP4 for Spearman, and KE2 for Kendall.

In the case of the eastern Andes, the Pearson network divided this region into two distinct communities. One community for the north of central Andes (PE11), while the other community contains the Peruvian and Bolivian hotspots (PE13). This division into two communities may be attributed to the linear response of Pearson or also to the artificial bias (CIEMER et al., 2018), which is influenced by the high rainfall values characteristic of these two regions and in this way causing a greater degree of local connectivity. On the other hand, the Spearman and Kendall networks identified only one community in this region (SP7 and KE6).

Figure 6.27 - Communities identified in each network.



SOURCE: Author's production.

These differences indicate that the Spearman and Kendall networks provide a more comprehensive view of rainfall relations in this area by considering nonlinear interactions and correlations between variables.

Spearman and Kendall's networks exhibit a similar spatial distribution of communities, particularly in the northeast (SP1, SP3, and KE1), southeast (SP5 and KE4), and east (SP7 and KE6) of the Amazon basin. It is interesting to note that SP1 is included within the KE1 community. This is because the region defined by SP1 exhibits very low mean rainfall values. This indicates that Spearman has a greater sensitivity to the precipitation regime (intensity) and is able to identify this region as a separate community within the network. However, differences arise in the north-west region, where Spearman presents a large community to the southeast (SP4), while Kendall divides this region into two communities, one directly representative of the north (KE2) and the second defines much of the center of the Amazon basin (KE3).

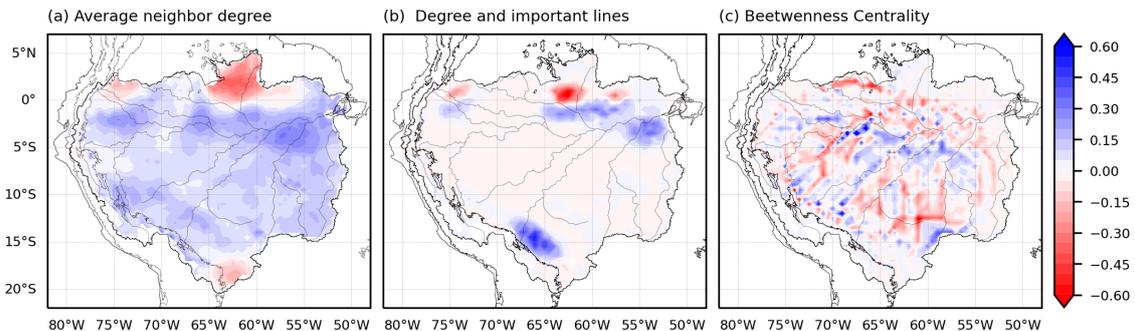
Finally, the rainfall communities analysis highlights the differences between Pearson, Spearman, and Kendall networks. As anticipated, Pearson tends to generate smaller communities, due to the network’s low degree of connectivity. This is particularly evident in the eastern Andes. In contrast, Spearman and Kendall identify a well-known corridor of the low-level jet stream. Thus providing a community that is more representative of the local dynamics.

6.2.2.2 Spearman and Kendall comparison

The results presented in the previous section indicate that all three networks provide findings that are linked to the pattern of the SAMS, with the Spearman and Kendall networks demonstrating the most consistent outcomes. These networks share a significant degree of similarity, and thus, this section will focus on analyzing their primary differences.

Additionally, it is worth noting that the similarities observed in their results provide further support for the relationship between the networks and the SAMS pattern. Figure 6.28 presents the differences in the normalized indices between the Spearman and Kendall networks. Positive values indicate a predominance of the Spearman network, while negative values indicate stronger indices from the Kendall network.

Figure 6.28 - Difference of normalized indices between Spearman and Kendall.



SOURCE: Author’s production.

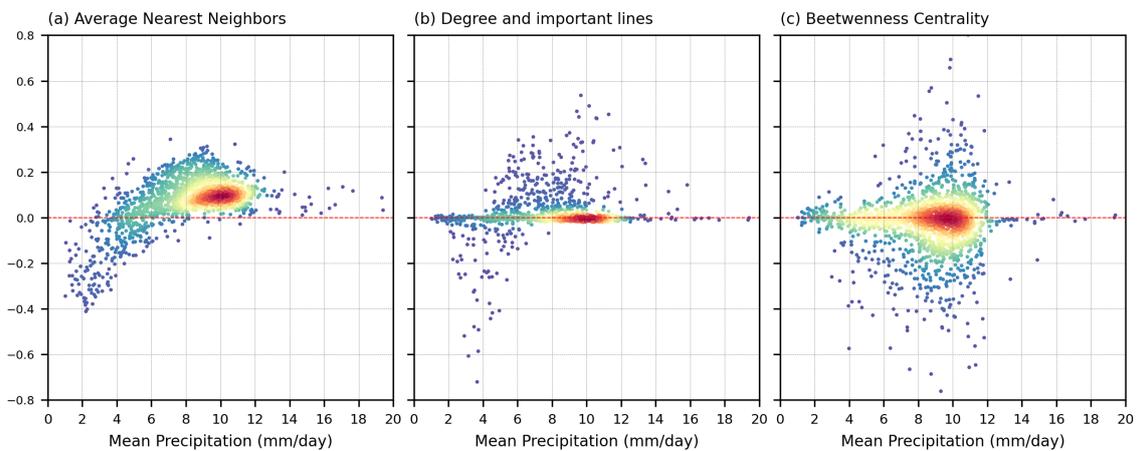
Figure 6.28a shows that the Kendall displays higher values at the extremes in the north and south of the basin, while Figure and 6.28b only in the north. This observation may be attributed to the specific characteristics of these regions, such as the low average rainfall amounts highlighted in the Amazon basin, shown in Figure 5.2a.

As mentioned earlier, Spearman exhibits higher sensitivity to rainfall regimes compared to Kendall. This difference can be observed in the region defined by 60°W - 65°W and 5°N - 0° (north of *Roraima* State). In Figure 6.27, this region is identified as part of the blue community for Kendall (KE1), indicating a low rainfall rate within the community associated with moisture influx from the North Atlantic. However, Spearman excludes this region from the community associated with moisture influx (SP3). This difference in community assignment highlights the contrasting response of the networks, with Spearman showing a response of AND and DIL more towards the south.

Finally, Figure 6.28c illustrates the differences in BC, with Kendall displaying higher values over most of the Amazon basin. Spearman, on the other hand, shows higher values with more predominance in the south of the Amazon River. Both networks show a well-defined path towards the central Andes, indicating moisture transport from the east. It is evident that the Spearman network exhibits higher AND values over most of the Amazon basin, showing superior connectivity in this region, and providing a good interpretation of the inflow and outflow of moisture. These results suggest that Spearman tends to generate a more connected and integrated network.

Figure 6.29 presents the scatter plot comparing the indices of the Spearman and Kendall networks. From Figure 6.29a, it can be observed that the largest differences in the AND occur at low rainfall values. Shows how Kendall provides a notable difference with Spearman for nodes with low average rainfall.

Figure 6.29 - Scatter plot for the differences between Spearman and Kendall networks.



SOURCE: Author's production.

This indicates the main difference between Spearman and Kendall, showing how Kendall overestimated the regions with low rainfall values. This pattern is also slightly observed in the DIL, as shown in Figures 6.29b. Where Kendall presents a strong response for series with low rainfall means. Spearman network consistently exhibits significantly higher values in the AND and DIL indices, suggesting a predominantly more connected network.

Lastly, the BC metric shows symmetric differences compared to the previous indices, indicating a similar response for the mean precipitation (Figure 6.29c). Both networks demonstrate that the most intense values of BC occur between 8 and 12 mm/day of rain. Additionally, nodes with low and very high mean rainfall values do not exhibit a strong presence of BC.

7 CONCLUSIONS

It is important to analyze the behavior of precipitation in each community to understand the distinct characteristics of different regions. Through the examination of temporal and spatial precipitation patterns, potential drivers of rainfall and its extreme events, such as convective systems (e.g. squall lines), orographic features, and atmospheric conditions, could be identified. An adequate identification or definition of the communities can improve the accuracy and reliability of their findings, leading to more robust insights into the behavior of precipitation systems and their implications for the studied region. This section presents the main conclusions of the networks for South America and the Amazon basin, thus the possible community applications and future works.

7.1 Extreme rainfall networks

The spatial pattern of the communities reflects the southward moisture transport from the central Amazon basin to southern South America during the austral summer, aligning with the active phase of the SAMS (GELBRECHT et al., 2021). This observation suggests a prevailing influence of the active phase of SAMS throughout the network. These communities are mainly represented by notable regions. Community SA1 is defined by most of the Amazon basin, Community SA2 contains the eastern central Andes, and finally Community SA3 from the eastern central Andes to the south of the *Sierras de Córdoba*.

Each of these communities exhibits distinct characteristics. Community SA1 demonstrates high activity, with nodes having the highest degrees and greater efficiency, showing the spatially more homogeneous and participatory region compared to the other two communities. The higher level of local interaction and cooperation among nodes within this community indicates its significance in maintaining connections between different parts of the network and facilitating the flow of information. Recognizing the source role of Community SA1 allows for a better understanding of the origins and potential pathways of information flow within the network of extreme events.

Community SA2 has the highest link density, indicating that it is the most interconnected region among the communities SA1 and SA3. This community is presented as an important bridge region between communities SA1 and SA3, specifically in the center-south. However, the connection between communities SA1 and SA2 is much stronger than between SA2 and SA3. This is due to various reasons such as

the predominant low-level wind direction, the geographical distance between nodes, the phases of the SAMS, and the presence of Andes. This demonstrates that the dynamics of the network of extreme events and the spatial distribution of communities align with the current literature, validating the effectiveness of this approach in identifying areas with similar characteristics and their interrelations.

Community SA3 has the highest probability for very low degrees and a rapid drop in the distribution of node degrees. This community also displays the highest values for skewness and outliers. These characteristics indicate a more heterogeneous behavior, which is associated with regions of both high and low connectivity.

The vertical reflectivity profiles revealed that Communities SA1 and SA2 have similar development and a moderate vertical gradient, with a slight difference at low levels. Community SA1 seems to have a greater presence of warm rain process compared to Community SA2, while the higher occurrence in the middle levels of Community SA2 may be due to the acceleration of rainfall processes related to orographic forcing. On the other hand, Community SA3 is characterized by intense vertical development, featuring deep convection. This can be associated with strong updrafts and vertical wind shear, making it the community with the highest convective cores. In turn, as a consequence, this community presents the greatest amount of ice.

7.2 Amazon Basin rainfall networks

The communities of extreme events in the Amazon basin showed a similar spatial pattern to the communities of extreme events in South America. The main distinctions are observed in the community to the east of the Andes. In the South America network, the Peruvian and Bolivian hotspots are located in two distinct communities, highlighting the importance of the Bolivian hotspot for extreme events in the extratropics of South America. In contrast, in the Amazon network, both hotspots are contained within the same community, clearly identifying this community as an exit corridor (sink community).

The differences in the spatial distribution of the indices among the Pearson, Spearman, and Kendall networks can be attributed to their sensitivities to rainfall patterns. While Spearman and Kendall exhibit similar results, the same cannot be said for Pearson. In order to obtain accurate results using Pearson, the rainfall series must follow a normal distribution. However, it is well-known that rainfall data does not often exhibit this characteristic. To address this issue, the criterion proposed by Ciemer et al. (2018) was applied. The results show how Pearson correlation has

certain limitations due to its linear nature. The linear relationship measured by Pearson may not fully capture the complexity of the interactions between nodes in the network. As a result, areas with high connectivity and strong interactions between regions may be underestimated or not fully identified, leading to a more fragmented representation of the network.

The results obtained from the rainfall networks over the Amazon basin, it is worth highlighting the identification of three important regions located in the northwest, northeast, and south of the basin. The northwest region is an area where moisture enters from the northern part of South America, as well as contributions from the Pacific Ocean. The northeast region of the basin is mainly influenced by the North Tropical Atlantic, which is the main source of warm humid air over the Amazon basin. The third region identified in the southeast of the basin represents the outflow of low-level flow towards the extratropics. This region is characterized by the presence of Peruvian and Bolivian precipitation hotspots.

The results of betweenness centrality (BC) demonstrate how the three rainfall networks exhibit high values following the spatial pattern of moisture transport from the central Amazon basin to the eastern Andes. While Pearson shows the most intense values in the central basin, Kendall is slightly more focused on the eastern center. In the case of Spearman, this network presents a bit more homogeneous distribution. The three networks also can capture the land-river breeze effect, since the results of networks exhibited higher means and medians of BC on nodes in the presence of rivers. With Spearman showing the most notable differences. This indicates a more accurate response in relation to this phenomenon.

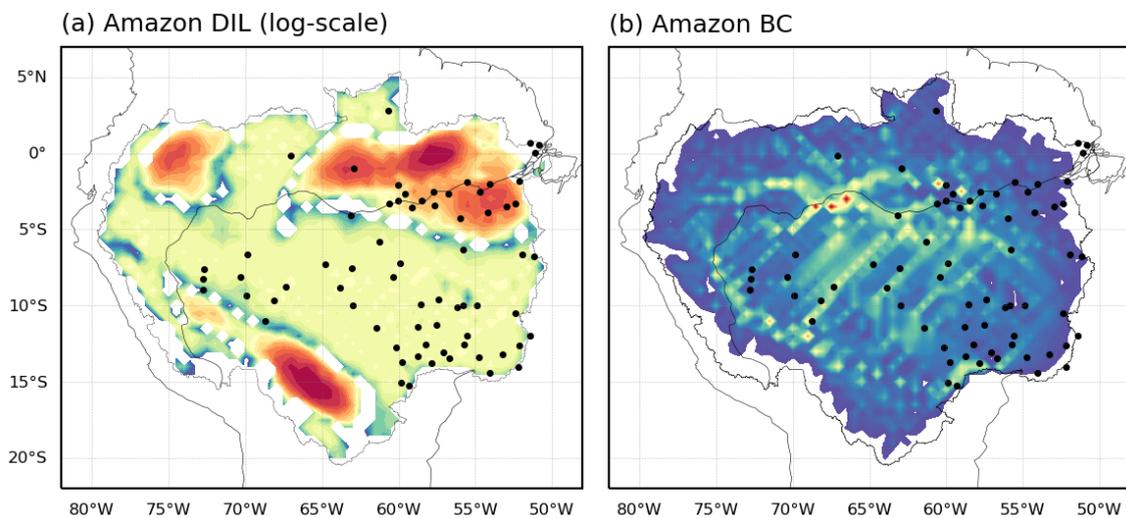
Finally, Spearman and Kendall networks exhibit similar results, identifying a more pronounced spatial influence of the entrance and exit regions. Spearman showed a better response to the importance of the northeast of the Amazon basin and a better identification of the land-river breeze effect.

These more adequate results shown by Spearman indicate better sensitivity to different rainfall regimes. By considering the ranks, Spearman can capture the underlying relationships and dependencies between variables more effectively, resulting in a higher density of connections in the network. In general, the Spearman correlation proved to be a more robust measure than the Kendall correlation.

7.3 Possible applications

An interesting application was demonstrated by [Agarwal et al. \(2018\)](#), who identified the most important nodes to identify the most suitable locations for installing new rain gauges. This method is of particular interest for regions of difficult access, where the number of rain gauges may be reduced. For example Figure 7.1 shows the DIL and BC as possible regions for installing new rain gauges. Also, this figure shows us the relationship between the rain gauge stations installed by INEMET and these indices.

Figure 7.1 - Possible regions to the location of new rain gauges.



DIL and BC for the Spearman Rainfall network. Black points represent the actually INEMET rain gauge locations.
SOURCE: Author's production.

We can observe that in the state of *Pará* near the mouth of the Amazon River, this region has several rain gauge stations that are well correlated with the DIL results (Figure 7.1a). However, to the north of the Amazon River, which also shows a high DIL, it is characterized by a lack of stations. This suggests a potential region for the installation of new stations. Similar information is provided by BC, where some stations are related to this metric. However, a region of high BC is noticeable over the Amazon River, which could be considered for the possible installation of another rain gauge station.

As previously mentioned, communities offer valuable information that can be utilized in various ways. For example, several studies examine the behavior of specific regions to anticipate different responses in other reference areas (ESPINOZA et al., 2015; CIEMER et al., 2020; GELBRECHT et al., 2021). However, these studies often identify important regions using rectangular boundaries, which can lead to underestimated relationships and impact the expected results. Therefore, using the communities or network indexes with appropriate boundaries will yield more accurate results.

An important application of communities lies in the development of specialized models for the region, be a statistical, physically based, or using artificial intelligence. A model specifically tailored to the community will provide a more detailed and comprehensive analysis of the specific processes and interactions occurring in that region. This approach also offers other benefits, such as improving resolution in identifying local features, topography, and land-cover characteristics.

Communities can serve as geographical boundaries for extracting specific weather-related physical parameters. For instance, analyzing extreme rainfall events within a community can provide insights into factors like convection depth, ice presence, and contributions from stratiform and convective processes. These findings can contribute to the development of tailored physical parameterizations for numerical weather prediction (NWP) models in that particular region.

Finally, using a specific community reduces computational complexity compared to global or regional models, which often cover extensive geographical areas and require significant computational resources and time. This approach is more computationally efficient and suitable for conventional computers.

7.4 Future works

In future research, it would be beneficial to identify the most active nodes or sub-regions within each community using metrics such as DIL or BC. This approach would allow for a more comprehensive characterization of the active regions and enhance the accuracy of defining extreme events. By focusing on these critical nodes, a deeper understanding of each community's distinct characteristics and behavior can be obtained, resulting in more precise and representative results. However, further analysis incorporates auxiliary data, such as instability indices or atmosphere electrical activity, to arrive at more robust conclusions.

Regarding the rainfall network in the Amazon basin, Spearman presented the most favorable results. In this way, it is intended to use rainfall data with higher resolutions (space and time), to verify the best response of the network. For example, The relationship between BC and the presence of rivers could be further enhanced with higher resolutions. In other words, a better spatial resolution can provide a noticeable difference between neighboring nodes. Similarly, an hourly network could also yield better results, as the land-river breeze effect occurs during mornings and nights.

Although the communities yielded favorable results for the dynamics of both the rainfall networks and the extreme rainfall networks, it is important to highlight the high sensitivity to the threshold used to define a link between two nodes. Even slight variations in the threshold can result in changes in the spatial structures of the communities. Therefore, it is important to analyze different methods of community identification in order to assess their sensitivity and identify more robust methods for this task.

REFERENCES

- ABE, S.; SUZUKI, N. Dynamical evolution of the community structure of complex earthquake network. **EPL (Europhysics Letters)**, v. 99, n. 3, p. 39001, 2012. 27
- ACEITUNO, P.; MONTECINOS, A. Análisis de la estabilidad de la relación entre la oscilación del sur y la precipitación en América del Sur. **Bulletin de Institut Français d'Études Andines**, v. 22, n. 1, p. 53–64, 1993. 88
- AGARWAL, A.; MARWAN, N.; MAHESWARAN, R.; MERZ, B.; KURTHS, J. Quantifying the roles of single stations within homogeneous regions using complex network analysis. **Journal of Hydrology**, v. 563, p. 802–810, 2018. 26, 27, 28, 29, 30, 54, 122
- AGARWAL, A.; MARWAN, N.; OZTURK, U.; MAHESWARAN, R. Unfolding community structure in rainfall network of Germany using complex network-based approach. In: RATHINASAMY, M.; CHANDRAMOULI, S.; PHANINDRA, K. B. V. N.; MAHESH, U. (Ed.). **Water resources and environmental engineering II: climate and environment**. Berlin: Springer: [s.n.], 2019. p. 179–193. 28
- AGRAWAL, H. Extreme self-organization in networks constructed from gene expression data. **Physical Review Letters**, v. 89, n. 26, p. 268702, 2002. 19
- ALBERT, R.; BARABÁSI, A.-L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, v. 74, n. 1, p. 47, 2002. 19
- ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world-wide web. **Nature**, v. 401, n. 6749, p. 130–131, 1999. 19
- ALLAN, R. P.; SODEN, B. J. Atmospheric warming and the amplification of precipitation extremes. **Science**, v. 321, n. 5895, p. 1481–1484, 2008. 9
- ALMAZROUI, M. et al. Assessment of cmip6 performance and projected temperature and precipitation changes over South America. **Earth Systems and Environment**, v. 5, n. 2, p. 155–183, 2021. 3, 49, 50
- ALMEIDA, A. C.; ROCHA, B. R.; SOUZA, J. R. S.; SÁ, J. A. S.; PISSOLATO FILHO, J. A. Cloud-to-ground lightning observations over the eastern Amazon region. **Atmospheric Research**, v. 117, p. 86–90, 2012. 94
- ALVES, L. M.; MARENGO, J. A.; FU, R.; BOMBARDI, R. J. Sensitivity of Amazon regional climate to deforestation. **American Journal of Climate Change**, v. 6, n. 1, p. 75–98, 2017. 50
- AMERICAN METEOROLOGICAL SOCIETY (AMS). **Convection definition by American Meteorological Society**. 2020. Available from: <<https://glossary.ametsoc.org/wiki/Convection>>. Access on: 03 Apr. 2020. 10

- AN, X.-l.; ZHANG, L.; LI, Y.-z.; ZHANG, J.-g. Synchronization analysis of complex networks with multi-weights and its application in public traffic network. **Physica A: Statistical Mechanics and its Applications**, v. 412, p. 149–156, 2014. 19
- ANABOR, V.; STENSRUD, D. J.; MORAES, O. L. D. Serial upstream-propagating mesoscale convective system events over southeastern South America. **Monthly Weather Review**, v. 136, n. 8, p. 3087–3105, 2008. 16
- ANSELMO, E. M. **Morfologia das tempestades elétricas na América do Sul**. 129 p. Tese (Doutorado em Ciências Atmosféricas) — Universidade de São Paulo, São Paulo, 2015. 100
- ANSELMO, E. M.; MACHADO, L. A.; SCHUMACHER, C.; KILADIS, G. N. Amazonian mesoscale convective systems: life cycle and propagation characteristics. **International Journal of Climatology**, v. 41, n. 7, p. 3968–3981, 2021. 86, 88, 100, 111, 113
- ARTURO, S. P. C.; JAMES, P. C. A.; ADRIANO, P. A.; HELVECIO, B. L. N. Spatial sensitivity of complex network communities in the Amazon basin in relation to the minimum link value. **Authorea Preprints**, 2022. 27, 28
- ASTLING, E.; PAEGLE, J.; MILLER, E.; O'BRIEN, C. Boundary layer control of nocturnal convection associated with a synoptic scale system. **Monthly Weather Review**, v. 113, n. 4, p. 540–552, 1985. 17
- BAGLER, G. Analysis of the airport network of India as a complex weighted network. **Physica A: Statistical Mechanics and its Applications**, v. 387, n. 12, p. 2972–2980, 2008. 19
- BARABASI, A.-L. **Linked: how everything is connected to everything else and what it means**. 2003. 19
- BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, n. 5439, p. 509–512, 1999. 27
- BARNETT, T. et al. Forecasting global enso-related climate anomalies. **Tellus A**, v. 46, n. 4, p. 381–397, 1994. 24
- BARNSTON, A. G.; TIPPETT, M. K.; L'HEUREUX, M. L.; LI, S.; DEWITT, D. G. Skill of real-time seasonal ENSO model predictions during 2002–11: is our capability increasing? **Bulletin of the American Meteorological Society**, v. 93, n. 5, p. 631–651, 2012. 24
- BARRAT, A.; BARTHELEMY, M.; PASTOR-SATORRAS, R.; VESPIGNANI, A. The architecture of complex weighted networks. **Proceedings of the National Academy of Sciences**, v. 101, n. 11, p. 3747–3752, 2004. 59
- BARTELS, D.; ROCKWOOD, A. Internal structure and evolution of a dual mesoscale convective complex. In: CONFERENCE ON HYDROMETEOROLOGY, 1983, USA. **Proceedings...** Tulsa: AMS, 1983. p. 97–102. 14

- BARTELS, D. L.; SKRADSKI, J. M.; MENARD, R. D. **Mesoscale convective systems: a satellite-data-based climatology**. [S.l.: s.n.], 1984. 13
- BARZEL, B.; BARABÁSI, A.-L. Universality in network dynamics. **Nature Physics**, v. 9, n. 10, p. 673–681, 2013. 27
- BERBERY, E. H.; BARROS, V. R. The hydrologic cycle of the La Plata basin in South America. **Journal of Hydrometeorology**, v. 3, n. 6, p. 630–645, 2002. 2, 80, 84
- BERBERY, E. H.; COLLINI, E. A. Springtime precipitation and water vapor flux over southeastern South America. **Monthly Weather Review**, v. 128, n. 5, p. 1328–1346, 2000. 8, 17
- BERRI, G. J.; INZUNZA, J. B. The effect of the low-level jet on the poleward water vapor transport in the central region of South America. **Atmospheric Environment. Part A. General Topics**, v. 27, n. 3, p. 335–341, 1993. 7
- BHAT, G.; KUMAR, S. Vertical structure of cumulonimbus towers and intense convective clouds over the South Asian region during the summer monsoon season. **Journal of Geophysical Research: Atmospheres**, v. 120, n. 5, p. 1710–1722, 2015. 42
- BOCCIPPIO, D. J.; KOSHAK, W. J.; BLAKESLEE, R. J. Performance assessment of the optical transient detector and lightning imaging sensor. part I: predicted diurnal variability. **Journal of Atmospheric and Oceanic Technology**, v. 19, n. 9, p. 1318–1332, 2002. 45
- BOERS, N.; BOOKHAGEN, B.; BARBOSA, H. M.; MARWAN, N.; KURTHS, J.; MARENGO, J. Prediction of extreme floods in the eastern central Andes based on a complex networks approach. **Nature Communications**, v. 5, n. 1, p. 1–7, 2014b. 1, 9, 24, 25, 33, 52, 53, 63, 90
- BOERS, N.; BOOKHAGEN, B.; MARENGO, J.; MARWAN, N.; STORCH, J.-S. von; KURTHS, J. Extreme rainfall of the south american monsoon system: a dataset comparison using complex networks. **Journal of Climate**, v. 28, n. 3, p. 1031–1056, 2015. 1, 7, 90
- BOERS, N.; BOOKHAGEN, B.; MARWAN, N.; KURTHS, J.; MARENGO, J. Complex networks identify spatial patterns of extreme rainfall events of the south american monsoon system. **Geophysical Research Letters**, v. 40, n. 16, p. 4386–4392, 2013. 1, 2, 7, 18, 33, 57, 63, 64, 73, 75, 99, 104, 111
- BOERS, N.; BOOKHAGEN, B.; MARWAN, N.; KURTHS, J. Spatiotemporal characteristics and synchronization of extreme rainfall in South America with focus on the Andes mountain range. **Climate Dynamics**, v. 46, p. 601–617, 2016. 1
- BOERS, N.; DONNER, R. V.; BOOKHAGEN, B.; KURTHS, J. Complex network analysis helps to identify impacts of the El Niño southern oscillation on moisture divergence in South America. **Climate Dynamics**, v. 45, p. 619–632, 2015. 33, 51

BOERS, N.; GOSWAMI, B.; RHEINWALT, A.; BOOKHAGEN, B.; HOSKINS, B.; KURTHS, J. Complex networks reveal global pattern of extreme-rainfall teleconnections. **Nature**, v. 566, n. 7744, p. 373–377, 2019. 2, 53

BOERS, N.; RHEINWALT, A.; BOOKHAGEN, B.; BARBOSA, H. M.; MARWAN, N.; MARENGO, J.; KURTHS, J. The south american rainfall dipole: a complex network analysis of extreme events. **Geophysical Research Letters**, v. 41, n. 20, p. 7397–7405, 2014a. 2, 33

BOMBARDI, R. J.; CARVALHO, L. M. The South Atlantic dipole and variations in the characteristics of the south american monsoon in the wcrp-cmip3 multi-model simulations. **Climate Dynamics**, v. 36, p. 2091–2102, 2011. 7

BOOKHAGEN, B.; STRECKER, M. R. Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes. **Geophysical Research Letters**, v. 35, n. 6, 2008. 80, 88, 89, 112

BRICEÑO-ZULUAGA, F.; CASTAGNA, A.; RUTLLANT, J. A.; FLORES-AQUEVEQUE, V.; CAQUINEAU, S.; SIFEDDINE, A.; VELAZCO, F.; GUTIERREZ, D.; CARDICH, J. Paracas dust storms: sources, trajectories and associated meteorological conditions. **Atmospheric Environment**, v. 165, p. 99–110, 2017. 3, 49

BRONI-BEDAİKO, C.; KATSRIKU, F. A.; UNEMI, T.; ATSUMI, M.; ABDULAI, J.-D.; SHINOMIYA, N.; OWUSU, E. El Niño-southern oscillation forecasting using complex networks analysis of lstm neural networks. **Artificial Life and Robotics**, v. 24, n. 4, p. 445–451, 2019. 24

BRUICK, Z. S.; RASMUSSEN, K. L.; ROWE, A. K.; MCMURDIE, L. A. Characteristics of intense convection in subtropical South America as influenced by El Niño–southern oscillation. **Monthly Weather Review**, v. 147, n. 6, p. 1947–1966, 2019. 18

BULL, A. T.; ANDREWS, B. A.; DORADOR, C.; GOODFELLOW, M. Introducing the atacama desert. **Antonie Van Leeuwenhoek**, v. 111, p. 1269–1272, 2018. 3, 49

BYERLE, L. A.; PAEGLE, J. Description of the seasonal cycle of low-level flows flanking the Andes and their interannual variability. **Meteorologica**, v. 27, p. 71–88, 2002. 82

BYERS, H. R.; BRAHAM, R. R. Thunderstorm structure and circulation. **Journal of Atmospheric Sciences**, v. 5, n. 3, p. 71–86, 1948. 11

CALDARELLI, G. **Scale-free networks: complex webs in nature and technology**. Illustrated. New York: [s.n.], 2007. 309 p. ISBN 978-0199665174. 61

CAREY, L. D.; RUTLEDGE, S. A.; PETERSEN, W. A. The relationship between severe storm reports and cloud-to-ground lightning polarity in the contiguous united states from 1989 to 1998. **Monthly Weather Review**, v. 131, n. 7, p. 1211–1228, 2003. 45

CARVALHO, L. M.; JONES, C.; LIEBMANN, B. The South Atlantic convergence zone: intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. **Journal of Climate**, v. 17, n. 1, p. 88–108, 2004. 7

CAVALCANTI, I. d. A. **Um estudo sobre interações entre sistemas de circulação de escala sinótica e circulações locais**. 100 p. Dissertação (Mestrado em Meteorologia) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 1982. 14

CECIL, D. J.; BLANKENSHIP, C. B. Toward a global climatology of severe hailstorms as estimated by satellite passive microwave imagers. **Journal of Climate**, v. 25, n. 2, p. 687–703, 2012. 49, 98

CERON, W.; SANTOS, L. B.; DOLIF NETO, G.; QUILES, M. G.; CANDIDO, O. A. Community detection in very high-resolution meteorological networks. **IEEE Geoscience and Remote Sensing Letters**, v. 17, n. 11, p. 2007–2010, 2019. 28

CHAVEZ, S. P.; SILVA, Y.; BARROS, A. P. High-elevation monsoon precipitation processes in the central Andes of Peru. **Journal of Geophysical Research: Atmospheres**, v. 125, n. 24, p. e2020JD032947, 2020. 96

CHEN, T.-C.; KPAEYEH, J. A. The synoptic-scale environment associated with the low-level jet of the great plains. **Monthly Weather Review**, v. 121, n. 2, p. 416–420, 1993. 17

CHEUNG, K. K.; OZTURK, U. Synchronization of extreme rainfall during the australian summer monsoon: complex network perspectives. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, v. 30, n. 6, p. 063117, 2020. 2, 32

CIEMER, C.; BOERS, N.; BARBOSA, H. M.; KURTHS, J.; RAMMIG, A. Temporal evolution of the spatial covariability of rainfall in South America. **Climate Dynamics**, v. 51, p. 371–382, 2018. 2, 56, 115, 120

CIEMER, C.; REHM, L.; KURTHS, J.; DONNER, R. V.; WINKELMANN, R.; BOERS, N. An early-warning indicator for Amazon droughts exclusively based on tropical Atlantic sea surface temperatures. **Environmental Research Letters**, v. 15, n. 9, p. 094087, 2020. 123

CINTINEO, J. L.; PAVOLONIS, M. J.; SIEGLAFF, J. M.; HEIDINGER, A. K. Evolution of severe and nonsevere convection inferred from goes-derived cloud properties. **Journal of Applied Meteorology and Climatology**, v. 52, n. 9, p. 2009–2023, 2013. 11

CLAUSET, A.; NEWMAN, M. E.; MOORE, C. Finding community structure in very large networks. **Physical Review E**, v. 70, n. 6, p. 066111, 2004. 66, 78

COHEN, J. C.; DIAS, M. A. S.; NOBRE, C. A. Environmental conditions associated with amazonian squall lines: a case study. **Monthly Weather Review**, v. 123, n. 11, p. 3163–3174, 1995. 111

COLLISCHONN, B. **Uso de precipitação estimada pelo satélite TRMM em modelo hidrológico distribuído**. 193 p. Dissertação (Mestrado em Recursos Hídricos e Saneamento Ambiental) — Universidade Federal de Rio Grande do Sul, Porto Alegre, 2006. 37

CONTICELLO, F.; CIOFFI, F.; MERZ, B.; LALL, U. An event synchronization method to link heavy rainfall events and large-scale atmospheric circulation features. **International Journal of Climatology**, v. 38, n. 3, p. 1421–1437, 2018. 2, 29

COSTA, L. d. F.; JR, O. N. O.; TRAVIESO, G.; RODRIGUES, F. A.; BOAS, P. R. V.; ANTIQUEIRA, L.; VIANA, M. P.; ROCHA, L. E. C. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, v. 60, n. 3, p. 329–412, 2011. 51

COSTA, L. d. F.; RODRIGUES, F. A.; TRAVIESO, G.; BOAS, P. R. V. Characterization of complex networks: a survey of measurements. **Advances in Physics**, v. 56, n. 1, p. 167–242, 2007. 66, 78

CUNHA, A.; ALVALÁ, R. C.; NOBRE, C. A.; CARVALHO, M. A. Monitoring vegetative drought dynamics in the brazilian semiarid region. **Agricultural and Forest Meteorology**, v. 214, p. 494–505, 2015. 3, 49

DEBORTOLI, N. S.; DUBREUIL, V.; FUNATSU, B.; DELAHAYE, F.; OLIVEIRA, C. H. D.; RODRIGUES-FILHO, S.; SAITO, C. H.; FETTER, R. Rainfall patterns in the southern Amazon: a chronological perspective (1971–2010). **Climatic Change**, v. 132, p. 251–264, 2015. 93

DENG, Z.-H.; QIAO, H.-H.; GAO, M.-Y.; SONG, Q.; GAO, L. Complex network community detection method by improved density peaks model. **Physica A: Statistical Mechanics and its Applications**, v. 526, p. 121070, 2019. 27

DEZA, J. I.; MASOLLER, C.; BARREIRO, M. Distinguishing the effects of internal and forced atmospheric variability in climate networks. **Nonlinear Processes in Geophysics**, v. 21, n. 3, p. 617–631, 2014. 21, 23, 62

DONAVON, R. A.; JUNGBLUTH, K. A. Evaluation of a technique for radar identification of large hail across the upper midwest and central plains of the United States. **Weather and Forecasting**, v. 22, n. 2, p. 244–254, 2007. 97

DONGES, J. F.; ZOU, Y.; MARWAN, N.; KURTHS, J. The backbone of the climate network. **EPL (Europhysics Letters)**, v. 87, n. 4, p. 48007, 2009. 21

_____. Complex networks in climate dynamics. **The European Physical Journal Special Topics**, v. 174, n. 1, p. 157–179, 2009. 21, 22, 74

DOSWELL, C. A. Severe convective storms—an overview. **Severe Convective Storms**, p. 1–26, 2001. 11

DRUMOND, A.; NIETO, R.; GIMENO, L.; AMBRIZZI, T. A lagrangian identification of major sources of moisture over central Brazil and La Plata basin. **Journal of Geophysical Research: Atmospheres**, v. 113, n. D14, 2008. 80

- DURKEE, J. D.; MOTE, T. L.; SHEPHERD, J. M. The contribution of mesoscale convective complexes to rainfall across subtropical South America. **Journal of Climate**, v. 22, n. 17, p. 4590–4605, 2009. 16
- ELTAHIR, E. A.; BRAS, R. L. Precipitation recycling. **Reviews of Geophysics**, v. 34, n. 3, p. 367–378, 1996. 111
- EOM, Y.-H.; JO, H.-H. Generalized friendship paradox in complex networks: the case of scientific collaboration. **Scientific Reports**, v. 4, n. 1, p. 1–6, 2014. 19
- ESPINOZA, J. C.; CHAVEZ, S.; RONCHAIL, J.; JUNQUAS, C.; TAKAHASHI, K.; LAVADO, W. Rainfall hotspots over the southern tropical Andes: spatial distribution, rainfall intensity, and relations with large-scale atmospheric circulation. **Water Resources Research**, v. 51, n. 5, p. 3459–3475, 2015. 18, 49, 53, 85, 88, 100, 102, 123
- ESPINOZA, J.-C.; MARENGO, J. A.; SCHONGART, J.; JIMENEZ, J. C. The new historical flood of 2021 in the amazon river compared to major floods of the 21st century: atmospheric features in the context of the intensification of floods. **Weather and Climate Extremes**, v. 35, p. 100406, 2022. 110
- FASSONI-ANDRADE, A. C. et al. Amazon hydrology from space: scientific advances and future challenges. **Reviews of Geophysics**, v. 59, n. 4, p. e2020RG000728, 2021. 112
- FELDHOFF, J. H.; LANGE, S.; VOLKHOLZ, J.; DONGES, J. F.; KURTHS, J.; GERSTENGARBE, F.-W. Complex networks for climate model evaluation with application to statistical versus dynamical modeling of south american climate. **Climate Dynamics**, v. 44, p. 1567–1581, 2015. 2
- FENG, Z.; LEUNG, L. R.; LIU, N.; WANG, J.; JR, R. A. H.; LI, J.; HARDIN, J. C.; CHEN, D.; GUO, J. A global high-resolution mesoscale convective system database using satellite-derived cloud tops, surface precipitation, and tracking. **Journal of Geophysical Research: Atmospheres**, v. 126, n. 8, p. e2020JD034202, 2021. 102
- FERREIRA, L. J.; SAULO, A. C.; SELUCHI, M. E. Características de la depresión del noroeste argentino en el período 1997-2003: criterios de selección y análisis estadístico. **Meteorologica**, 2010. 7
- FIELD, C. B. et al. Climate change 2014: impacts, adaptation, and vulnerability. In: CHANGE, I. P. O. C. (Ed.). **Fifth assessment report: global and sectorial aspects**. USA: [s.n.], 2014. p. 1–1101. 9
- FIGUEROA, S. N.; NOBRE, C. A. Precipitation distribution over central and western tropical South America. **Climanálise**, v. 5, n. 6, p. 36–45, 1990. 89
- FINDLATER, J. A major low-level air current near the indian ocean during the northern summer. **Quarterly Journal of the Royal Meteorological Society**, v. 95, n. 404, p. 362–380, 1969. 17

FISCH, G.; MARENGO, J. A.; NOBRE, C. A. Uma revisão geral sobre o clima da Amazônia. **Acta Amazônica**, p. 101–101, 1998. **3**, **50**, **79**

FITZJARRALD, D. R.; SAKAI, R. K.; MORAES, O. L.; OLIVEIRA, R. Cosme de; ACEVEDO, O. C.; CZIKOWSKY, M. J.; BELDINI, T. Spatial and temporal rainfall variability near the Amazon-Tapajós confluence. **Journal of Geophysical Research: Biogeosciences**, v. 113, n. G1, 2008. **113**

FORTUNATO, S. Community detection in graphs. **Physics Reports**, v. 486, n. 3–5, p. 75–174, 2010. **2**, **28**, **66**

FORTUNATO, S.; HRIC, D. Community detection in networks: a user guide. **Physics Reports**, v. 659, p. 1–44, 2016. **28**

FREITAS, C. G.; AQUINO, A. L.; RAMOS, H. S.; FRERY, A. C.; ROSSO, O. A. A detailed characterization of complex networks using information theory. **Scientific Reports**, v. 9, n. 1, p. 1–12, 2019. **1**, **81**

FRITSCH, J.; KANE, R.; CHELIUS, C. The contribution of mesoscale convective weather systems to the warm-season precipitation in the United States. **Journal of Applied Meteorology and Climatology**, v. 25, n. 10, p. 1333–1345, 1986. **14**

FU, G.; CHARLES, S. P.; KIRSHNER, S. Daily rainfall projections from general circulation models with a downscaling nonhomogeneous hidden markov model (nhmm) for south-eastern Australia. **Hydrological Processes**, v. 27, n. 25, p. 3663–3673, 2013. **93**

GAN, M. A.; KOUSKY, V. E.; ROPELEWSKI, C. F. The South America monsoon circulation and its relationship to rainfall over west-central Brazil. **Journal of Climate**, v. 17, n. 1, p. 47–66, 2004. **7**

GARSTANG, M.; MASSIE, H. L.; HALVERSON, J.; GRECO, S.; SCALA, J. Amazon coastal squall lines. part I: structure and kinematics. **Monthly Weather Review**, v. 122, n. 4, p. 608–622, 1994. **87**, **102**

GELBRECHT, M.; BOERS, N.; KURTHS, J. Variability of the low-level circulation of the south american monsoon analysed with complex networks. **The European Physical Journal Special Topics**, v. 230, n. 14–15, p. 3101–3120, 2021. **2**, **7**, **17**, **33**, **79**, **80**, **90**, **91**, **103**, **111**, **119**, **123**

GOZOLCHIANI, A.; HAVLIN, S.; YAMASAKI, K. Emergence of El Niño as an autonomous component in the climate network. **Physical Review Letters**, v. 107, n. 14, p. 148501, 2011. **23**

GOZOLCHIANI, A.; YAMASAKI, K.; GAZIT, O.; HAVLIN, S. Pattern of climate network blinking links follows El Niño events. **EPL (Europhysics Letters)**, v. 83, n. 2, p. 28005, 2008. **20**, **51**

GRECO, S.; SWAP, R.; GARSTANG, M.; ULANSKI, S.; SHIPHAM, M.; HARRISS, R.; TALBOT, R.; ANDREAE, M.; ARTAXO, P. Rainfall and surface kinematic conditions over central Amazonia during able 2b. **Journal of Geophysical Research: Atmospheres**, v. 95, n. D10, p. 17001–17014, 1990. 87

GRIMM, A. M. Interannual climate variability in South America: impacts on seasonal precipitation, extreme events, and possible effects of climate change. **Stochastic Environmental Research and Risk Assessment**, v. 25, p. 537–554, 2011. 1

GRIMM, A. M.; TEDESCHI, R. G. Enso and extreme rainfall events in South America. **Journal of Climate**, v. 22, n. 7, p. 1589–1609, 2009. 18

GUEDES, R. **Condições de grande escala associadas a sistemas convectivos de mesoescala sobre a região central da América do Sul**. Dissertação (Mestrado em Meteorologia) — Universidade de São Paulo, São Paulo, São Paulo, 1985. 14

GUIMERA, R.; AMARAL, L. A. N. Cartography of complex networks: modules and universal roles. **Journal of Statistical Mechanics: Theory and Experiment**, v. 2005, n. 02, p. P02001, 2005. 26

GUIMERA, R.; SALES-PARDO, M.; AMARAL, L. A. Classes of complex networks defined by role-to-role connectivity profiles. **Nature Physics**, v. 3, n. 1, p. 63–69, 2007. 26

GUPTA, S.; BOERS, N.; PAPPENBERGER, F.; KURTHS, J. Complex network approach for detecting tropical cyclones. **Climate Dynamics**, v. 57, p. 3355–3364, 2021. 52

GUPTA, S.; SU, Z.; BOERS, N.; KURTHS, J.; MARWAN, N.; PAPPENBERGER, F. Interconnection between the indian and the east asian summer monsoon: spatial synchronization patterns of extreme rainfall events. **International Journal of Climatology**, v. 43, n. 2, p. 1034–1049, 2023. 2, 30

HAGHTALAB, N.; MOORE, N.; HEERSPINK, B. P.; HYNDMAN, D. W. Evaluating spatial patterns in precipitation trends across the Amazon basin driven by land cover and global scale forcings. **Theoretical and Applied Climatology**, v. 140, p. 411–427, 2020. 93

HAMADA, A.; TAKAYABU, Y. N.; LIU, C.; ZIPSER, E. J. Weak linkage between the heaviest rainfall and tallest storms. **Nature Communications**, v. 6, n. 1, p. 6213, 2015. 70, 95

HAYLOCK, M. R. et al. Trends in total and extreme South American rainfall in 1960–2000 and links with sea surface temperature. **Journal of Climate**, v. 19, n. 8, p. 1490–1512, 2006. 3, 50, 79

HERMANSON, L. et al. Wmo global annual to decadal climate update: a prediction for 2021–25. **Bulletin of the American Meteorological Society**, v. 103, n. 4, p. E1117–E1129, 2022. 9

HOERLING, M. P.; KUMAR, A.; ZHONG, M. El Niño, La Niña, and the nonlinearity of their teleconnections. **Journal of Climate**, v. 10, n. 8, p. 1769–1786, 1997. 23

HOU, A. Y.; KAKAR, R. K.; NEECK, S.; AZARBARZIN, A. A.; KUMMEROW, C. D.; KOJIMA, M.; OKI, R.; NAKAMURA, K.; IGUCHI, T. The global precipitation measurement mission. **Bulletin of the American Meteorological Society**, v. 95, n. 5, p. 701–722, 2014. 39, 40, 41, 42

HOUZE JUNIOR, R. A. Mesoscale convective systems. **Reviews of Geophysics**, v. 42, n. 4, 2004. 14

HOUZE, R. A.; SMULL, B. F.; DODGE, P. Mesoscale organization of springtime rainstorms in Oklahoma. **Monthly Weather Review**, v. 118, n. 3, p. 613–654, 1990. 13, 14, 15

HUANG, Y.; LIU, Y.; LIU, Y.; KNIEVEL, J. C. Budget analyses of a record-breaking rainfall event in the coastal metropolitan city of Guangzhou, China. **Journal of Geophysical Research: Atmospheres**, v. 124, n. 16, p. 9391–9406, 2019. 95

HUFFMAN, G. J.; ADLER, R. F.; BOLVIN, D. T.; NELKIN, E. J. The TRMM multi-satellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. In: GEBREMICHAEL, M.; HOSSAIN, F. (Ed.). **Satellite rainfall applications for surface hydrology**. Berlin: Springer: [s.n.], 2010. p. 3–22. 35

IGUCHI, T.; SETO, S.; MENEGHINI, R.; YOSHIDA, N.; AWAKA, J.; LE, M.; CHANDRASEKAR, V.; KUBOTA, T. **GPM/DPR level-2 algorithm theoretical basis document**. 6. ed. Washington: [s.n.], 2010. 127 p. NASA Goddard Space Flight Center. 41

INSEL, N.; POULSEN, C. J.; EHLERS, T. A. Influence of the Andes mountains on South American moisture transport, convection, and precipitation. **Climate Dynamics**, v. 35, n. 7, p. 1477–1492, 2010. 3, 49

JACOBSON, A. R.; HEAVNER, M. J. Comparison of narrow bipolar events with ordinary lightning as proxies for severe convection. **Monthly Weather Review**, v. 133, n. 5, p. 1144–1154, 2005. 45

JARAMILLO, L.; POVEDA, G.; MEJÍA, J. et al. Mesoscale convective systems and other precipitation features over the tropical Americas and surrounding seas as seen by TRMM. **International Journal of Climatology**, v. 37, n. s1, p. 380–397, 2017. 17

JHA, S. K.; SIVAKUMAR, B. Complex networks for rainfall modeling: spatial connections, temporal scale, and network size. **Journal of Hydrology**, v. 554, p. 482–489, 2017. 26

- JONES, C. Recent changes in the South America low-level jet. **Npj Climate and Atmospheric Science**, v. 2, n. 1, p. 20, 2019. 89, 110, 111, 112
- JORGETTI, T.; DIAS, P. L. da S.; FREITAS, E. D. de. The relationship between South Atlantic SST and SACZ intensity and positioning. **Climate Dynamics**, v. 42, p. 3077–3086, 2014. 7
- JURY, M. R.; TOSEN, G. R. Characteristics of the winter boundary layer over the african plateau: 26 s. **Boundary-Layer Meteorology**, v. 49, n. 1, p. 53–76, 1989. 17
- KAISER, M. Mean clustering coefficients: the role of isolated nodes and leafs on clustering measures for small-world networks. **New Journal of Physics**, v. 10, n. 8, p. 083042, 2008. 60
- KALAPUREDDY, M.; RAO, D.; JAIN, A.; OHNO, Y. Wind profiler observations of a monsoon low-level jet over a tropical indian station. **Annales Geophysicae**, v. 25, n. 10, p. 2125–2137, 2007. 17
- KANE, R.; CHELIUS, C.; FRITSCH, J. Precipitation characteristics of mesoscale convective weather systems. **Journal of Applied Meteorology and Climatology**, v. 26, n. 10, p. 1345–1357, 1987. 14
- KEENAN, T.; MCBRIDE, J.; HOLLAND, G.; DAVIDSON, N.; GUNN, B. Diurnal variations during the australian monsoon experiment (amex) phase ii. **Monthly Weather Review**, v. 117, n. 11, p. 2535–2553, 1989. 17
- KINCER, J. B. Daytime and nighttime precipitation and their economic significance. **Monthly Weather Review**, v. 44, n. 11, p. 628–633, 1916. 17
- KISTLER, R. et al. The ncep–ncar 50-year reanalysis: monthly means cd-rom and documentation. **Bulletin of the American Meteorological Society**, v. 82, n. 2, p. 247–268, 2001. 43
- KODAMA, Y. Large-scale common features of subtropical precipitation zones (the baiu frontal zone, the spcz, and the sacz) part i: characteristics of subtropical frontal zones. **Journal of the Meteorological Society of Japan. Ser. II**, v. 70, n. 4, p. 813–836, 1992. 7
- KUMAR, S.; BHAT, G. Vertical profiles of radar reflectivity factor in intense convective clouds in the tropics. **Journal of Applied Meteorology and Climatology**, v. 55, n. 5, p. 1277–1286, 2016. 42
- _____. Vertical structure of orographic precipitating clouds observed over south Asia during summer monsoon season. **Journal of Earth System Science**, v. 126, n. 8, p. 1–12, 2017. 42
- KUMAR, S.; SILVA, Y. Vertical characteristics of radar reflectivity and dsd parameters in intense convective clouds over south east south Asia during the indian summer monsoon: GPM observations. **International Journal of Remote Sensing**, v. 40, n. 24, p. 9604–9628, 2019. 42

_____. Distribution of hydrometeors in monsoonal clouds over the south american continent during the austral summer monsoon: GPM observations. **International Journal of Remote Sensing**, v. 41, n. 10, p. 3677–3707, 2020. 79

KUMAR, S.; SILVA, Y.; MOYA-ALVAREZ, A. S.; MARTÍNEZ-CASTRO, D. Seasonal and regional differences in extreme rainfall events and their contribution to the world's precipitation: GPM observations. **Advances in Meteorology**, v. 2019, 2019. 15, 49

KUMAR, S.; VIDAL, Y.-S.; MOYA-ÁLVAREZ, A. S.; MARTÍNEZ-CASTRO, D. Effect of the surface wind flow and topography on precipitating cloud systems over the Andes and associated Amazon basin: GPM observations. **Atmospheric Research**, v. 225, p. 193–208, 2019. 1, 96

KUMMEROW, C.; BARNES, W.; KOZU, T.; SHIUE, J.; SIMPSON, J. The tropical rainfall measuring mission (TRMM) sensor package. **Journal of Atmospheric and Oceanic Technology**, v. 15, n. 3, p. 809–817, 1998. 35, 36

KUMMEROW, C. et al. The status of the tropical rainfall measuring mission (TRMM) after two years in orbit. **Journal of Applied Meteorology**, v. 39, n. 12, p. 1965–1982, 2000. 35, 38

LACKMANN, G. M. The south-central us flood of may 2010: present and future. **Journal of Climate**, v. 26, n. 13, p. 4688–4709, 2013. 9

LAWRENCE, D.; VANDECAR, K. Effects of tropical deforestation on climate and agriculture. **Nature Climate Change**, v. 5, n. 1, p. 27–36, 2015. 93

LENTERS, J. D.; COOK, K. Simulation and diagnosis of the regional summertime precipitation climatology of South America. **Journal of Climate**, v. 8, n. 12, p. 2988–3005, 1995. 89

LIEBMANN, B.; KILADIS, G. N.; MARENGO, J.; AMBRIZZI, T.; GLICK, J. D. Submonthly convective variability over South America and the South Atlantic Convergence Zone. **Journal of Climate**, v. 12, n. 7, p. 1877–1891, 1999. 7

LIEBMANN, B.; MARENGO, J. Interannual variability of the rainy season and rainfall in the brazilian Amazon basin. **Journal of Climate**, v. 14, n. 22, p. 4308–4318, 2001. 93

LIU, B.; TAN, X.; GAN, T. Y.; CHEN, X.; LIN, K.; LU, M.; LIU, Z. Global atmospheric moisture transport associated with precipitation extremes: mechanisms and climate change impacts. **Wiley Interdisciplinary Reviews: Water**, v. 7, n. 2, p. e1412, 2020. 1

LIU, C. **TRMM precipitation and cloud feature database: description version 1.0**. 1. ed. Utah, USA: [s.n.], 2007. 780 p. University of Utah. xv, 43

LIU, C.; ZIPSER, E. J.; CECIL, D. J.; NESBITT, S. W.; SHERWOOD, S. A cloud and precipitation feature database from nine years of TRMM observations. **Journal of Applied Meteorology and Climatology**, v. 47, n. 10, p. 2712–2728, 2008. 42

LIU, C.; ZIPSER, E. J.; NESBITT, S. W. Global distribution of tropical deep convection: different perspectives from TRMM infrared and radar data. **Journal of Climate**, v. 20, n. 3, p. 489–503, 2007. 13, 17, 45

LIU, J.; XIONG, Q.; SHI, W.; SHI, X.; WANG, K. Evaluating the importance of nodes in complex networks. **Physica A: Statistical Mechanics and its Applications**, v. 452, p. 209–219, 2016. 65

LIU, Z.; OSTRENGA, D.; TENG, W.; KEMPLER, S. Tropical rainfall measuring mission (TRMM) precipitation data and services for research and applications. **Bulletin of the American Meteorological Society**, v. 93, n. 9, p. 1317–1325, 2012. 35

LU, L.; DENNING, A. S.; SILVA-DIAS, M. A. da; SILVA-DIAS, P. da; LONGO, M.; FREITAS, S. R.; SAATCHI, S. Mesoscale circulations and atmospheric CO₂ variations in the Tapajós region, Pará, Brazil. **Journal of Geophysical Research: Atmospheres**, v. 110, n. D21, 2005. 113

LUDESCHER, J.; GOZOLCHIANI, A.; BOGACHEV, M. I.; BUNDE, A.; HAVLIN, S.; SCHELLNHUBER, H. J. Very early warning of next El Niño. **Proceedings of the National Academy of Sciences**, v. 111, n. 6, p. 2064–2066, 2014. 24

LUDESCHER, J. et al. Network-based forecasting of climate phenomena. **Proceedings of the National Academy of Sciences**, v. 118, n. 47, p. e1922872118, 2021. 52

MA, J.; ZHOU, L.; FOLTZ, G. R.; QU, X.; YING, J.; TOKINAGA, H.; MECHOSO, C. R.; LI, J.; GU, X. Hydrological cycle changes under global warming and their effects on multiscale climate variability. **Annals of the New York Academy of Sciences**, v. 1472, n. 1, p. 21–48, 2020. 9

MACHADO, L.; LAURENT, H.; DESSAY, N.; MIRANDA, I. Seasonal and diurnal variability of convection over the Amazonia: a comparison of different vegetation types and large scale forcing. **Theoretical and Applied Climatology**, v. 78, n. 1, p. 61–77, 2004. 13, 15

MACHADO, L.; ROSSOW, W.; GUEDES, R.; WALKER, A. Life cycle variations of mesoscale convective systems over the Americas. **Monthly Weather Review**, v. 126, n. 6, p. 1630–1654, 1998. 15, 86, 101

MADAKUMBURA, G. D.; KIM, H.; UTSUMI, N.; SHIOGAMA, H.; FISCHER, E. M.; SELAND, Ø.; SCINOCCHA, J. F.; MITCHELL, D. M.; HIRABAYASHI, Y.; OKI, T. Event-to-event intensification of the hydrologic cycle from 1.5 °C to a 2 °C warmer world. **Scientific Reports**, v. 9, n. 1, p. 3483, 2019. 9

MADDOX, R. A. Mesoscale convective complexes. **Bulletin of the American Meteorological Society**, p. 1374–1387, 1980. 13, 14

- MALIK, H. A. M.; MAHMOOD, N.; USMAN, M. H.; ABID, F. Un-weighted network study of pakistani airports. In: INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND ENGINEERING TECHNOLOGIES, 2., 2019, Pakistan. **Proceedings...** Sukkur, 2019. p. 1–6. 19
- MALIK, N.; BOOKHAGEN, B.; MARWAN, N.; KURTHS, J. Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. **Climate Dynamics**, v. 39, n. 3, p. 971–987, 2012. 2, 30, 53, 73
- MARENGO, J. Interdecadal variability and trends of rainfall across the Amazon basin. **Theoretical and Applied Climatology**, v. 78, n. 1-3, p. 79, 2004. 93
- MARENGO, J. et al. Recent developments on the south american monsoon system. **International Journal of Climatology**, v. 32, n. 1, p. 1–21, 2012. 7, 79, 110, 111
- MARENGO, J. A. Interannual variability of surface climate in the Amazon basin. **International Journal of Climatology**, v. 12, n. 8, p. 853–863, 1992. 110
- _____. On the hydrological cycle of the Amazon basin: a historical review and current state-of-the-art. **Revista Brasileira de Meteorologia**, v. 21, n. 3, p. 1–19, 2006. 50
- _____. Água e mudanças climáticas. **Estudos Avançados**, v. 22, p. 83–96, 2008. 9
- MARENGO, J. A.; DOUGLAS, M. W.; DIAS, P. L. S. The south american low-level jet east of the Andes during the 1999 LBA-TRMM and lba-wet amc campaign. **Journal of Geophysical Research: Atmospheres**, v. 107, n. D20, p. LBA–47, 2002. 8
- MARENGO, J. A.; FISCH, G.; MORALES, C.; VENDRAME, I.; DIAS, P. C. Diurnal variability of rainfall in southwest Amazonia during the LBA-TRMM field campaign of the austral summer of 1999. **Acta Amazonica**, v. 34, p. 593–603, 2004. 87, 91
- MARENGO, J. A. et al. Assessing drought in the drylands of northeast Brazil under regional warming exceeding 4 c. **Natural Hazards**, v. 103, p. 2589–2611, 2020. 9
- _____. Heavy rainfall associated with floods in southeastern Brazil in november–december 2021. **Natural Hazards**, p. 1–28, 2023. 10
- MARENGO, J. A.; SOARES, W. R.; SAULO, C.; NICOLINI, M. Climatology of low-level jet east of the Andes as derived from the ncep-ncar reanalyses temporal and spatial variability. **American Meteorological Society**, 2003. 15, 110
- _____. Climatology of the low-level jet east of the Andes as derived from the ncep-ncar reanalyses: characteristics and temporal variability. **Journal of Climate**, v. 17, n. 12, p. 2261–2280, 2004. 8
- MARTÍN-GÓMEZ, V.; HERNÁNDEZ-GARCIA, E.; BARREIRO, M.; LÓPEZ, C. Interdecadal variability of southeastern South America rainfall and moisture sources during the austral summertime. **Journal of Climate**, v. 29, n. 18, p. 6751–6763, 2016. 2

- MARWAN, N.; DONGES, J. F.; ZOU, Y.; DONNER, R. V.; KURTHS, J. Complex network approach for recurrence analysis of time series. **Physics Letters A**, v. 373, n. 46, p. 4246–4254, 2009. 21, 52
- MATA, A. S. d. Complex networks: a mini-review. **Brazilian Journal of Physics**, v. 50, p. 658–672, 2020. 1
- MCANELLY, R. L.; COTTON, W. R. Meso- β -scale characteristics of an episode of meso- α -scale convective complexes. **Monthly Weather Review**, v. 114, n. 9, p. 1740–1770, 1986. 14
- MEANS, L. L. On thunderstorm forecasting in the central United States. **Monthly Weather Review**, v. 80, n. 10, p. 165–189, 1952. 17
- MENEZES, J. A.; MADUREIRA, A. P.; SANTOS, R. B. d.; DUVAL, I. d. B.; REGOTO, P.; MARGONARI, C.; BARATA, M. M. d. L.; CONFALONIERI, U. Analyzing spatial patterns of health vulnerability to drought in the brazilian semi-arid region. **International Journal of Environmental Research and Public Health**, v. 18, n. 12, p. 6262, 2021. 3, 49
- MERRITT, J.; FRITSCH, J. M. On the movement of the heavy precipitation areas of mid-latitude mesoscale convective complexes. In: CONFERENCE ON WEATHER FORECASTING AND ANALYSIS, 10., 1984, USA. **Proceedings...** Florida, 1984. p. 529–536. 14
- MISSAOUI, R.; SARR, I. **Social network analysis-community detection and evolution**. Berlin: Springer: [s.n.], 2015. 27
- MIZUTA, R.; ENDO, H. Projected changes in extreme precipitation in a 60-km agcm large ensemble and their dependence on return periods. **Geophysical Research Letters**, v. 47, n. 13, p. e2019GL086855, 2020. 9
- MO, K. C.; HIGGINS, R. W. The pacific–south american modes and tropical convection during the southern hemisphere winter. **Monthly Weather Review**, v. 126, n. 6, p. 1581–1596, 1998. 22
- MOHR, K. I.; ZIPSER, E. J. Mesoscale convective systems defined by their 85-ghz ice scattering signature: size and intensity comparison over tropical oceans and continents. **Monthly Weather Review**, v. 124, n. 11, p. 2417–2437, 1996. 14, 15, 16, 49, 77
- MONTINI, T. L.; JONES, C.; CARVALHO, L. M. The south american low-level jet: a new climatology, variability, and changes. **Journal of Geophysical Research: Atmospheres**, v. 124, n. 3, p. 1200–1218, 2019. 80, 84, 85, 89, 110, 111, 112
- MONTOYA, J. M.; PIMM, S. L.; SOLÉ, R. V. Ecological networks and their fragility. **Nature**, v. 442, n. 7100, p. 259–264, 2006. 19
- MROZ, K.; BATTAGLIA, A.; LANG, T. J.; CECIL, D. J.; TANELLI, S.; TRIDON, F. Hail-detection algorithm for the GPM core observatory satellite sensors. **Journal of Applied Meteorology and Climatology**, v. 56, n. 7, p. 1939–1957, 2017. 96

MULHOLLAND, J. P.; NESBITT, S. W.; TRAPP, R. J.; RASMUSSEN, K. L.; SALIO, P. V. Convective storm life cycle and environments near the Sierras de Córdoba, Argentina. **Monthly Weather Review**, v. 146, n. 8, p. 2541–2557, 2018. [3](#), [16](#), [49](#), [80](#), [89](#)

MUÑOZ, Á. G.; GODDARD, L.; ROBERTSON, A. W.; KUSHNIR, Y.; BAETHGEN, W. Cross-time scale interactions and rainfall extreme events in southeastern South America for the austral summer. part I: potential predictors. **Journal of Climate**, v. 28, n. 19, p. 7894–7913, 2015. [1](#)

NASCIMENTO, E. d. L. Previsão de tempestades severas utilizando-se parâmetros convectivos e modelos de mesoescala: uma estratégia operacional adotável no Brasil. **Revista Brasileira de Meteorologia**, v. 20, n. 1, p. 121–140, 2005. [11](#)

NAUFAN, I.; SIVAKUMAR, B.; WOLDEMESKEL, F. M.; RAGHAVAN, S. V.; VU, M. T.; LIONG, S.-Y. Spatial connections in regional climate model rainfall outputs at different temporal scales: application of network theory. **Journal of Hydrology**, v. 556, p. 1232–1243, 2018. [26](#), [51](#)

NEPUSZ, T.; PETRÓCZI, A.; NÉGYESSY, L.; BAZSÓ, F. Fuzzy communities and the concept of bridgeness in complex networks. **Physical Review E**, v. 77, n. 1, p. 016107, 2008. [27](#)

NESBITT, S. W.; CIFELLI, R.; RUTLEDGE, S. A. Storm morphology and rainfall characteristics of TRMM precipitation features. **Monthly Weather Review**, v. 134, n. 10, p. 2702–2721, 2006. [42](#), [45](#), [86](#)

NESBITT, S. W. et al. A storm safari in subtropical South America: proyecto RELAMPAGO. **Bulletin of the American Meteorological Society**, v. 102, n. 8, p. E1621–E1644, 2021. [89](#)

NESBITT, S. W.; ZIPSER, E. J.; CECIL, D. J. A census of precipitation features in the tropics using TRMM: radar, ice scattering, and lightning observations. **Journal of Climate**, v. 13, n. 23, p. 4087–4106, 2000. [14](#), [17](#), [43](#), [44](#), [45](#)

NEWELL, F. L.; AUSPREY, I. J.; ROBINSON, S. K. Spatiotemporal climate variability in the Andes of northern Peru: evaluation of gridded datasets to describe cloud forest microclimate and local rainfall. **International Journal of Climatology**, v. 42, n. 11, p. 5892–5915, 2022. [2](#)

NEWMAN, M. E. Power laws, pareto distributions and zipf's law. **Contemporary Physics**, v. 46, n. 5, p. 323–351, 2005. [27](#)

_____. Finding community structure in networks using the eigenvectors of matrices. **Physical Review E**, v. 74, n. 3, p. 036104, 2006. [27](#)

NEWMAN, M. E.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E**, v. 69, n. 2, p. 026113, 2004. [27](#), [28](#), [29](#), [66](#), [78](#)

- NI, X.; LIU, C.; ZHANG, Q.; CECIL, D. J. Properties of hail storms over China and the United States from the tropical rainfall measuring mission. **Journal of Geophysical Research: Atmospheres**, v. 121, n. 20, p. 12–031, 2016. 96
- NICOLINI, M.; WALDRON, K. M.; PAEGLE, J. Diurnal oscillations of low-level jets, vertical motion, and precipitation: a model case study. **Monthly Weather Review**, v. 121, n. 9, p. 2588–2610, 1993. 17
- NIE, J.; SOBEL, A. H.; SHAEVITZ, D. A.; WANG, S. Dynamic amplification of extreme precipitation sensitivity. **Proceedings of the National Academy of Sciences**, v. 115, n. 38, p. 9467–9472, 2018. 9
- NOBRE, C. A.; SAMPAIO, G.; BORMA, L. S.; CASTILLA-RUBIO, J. C.; SILVA, J. S.; CARDOSO, M. Land-use and climate change risks in the Amazon and the need of a novel sustainable development paradigm. **Proceedings of the National Academy of Sciences**, v. 113, n. 39, p. 10759–10768, 2016. 93
- NOBRE, F. S. S.; COSTA, C. L. A.; OLIVEIRA, D. L. de; CABRAL, D. A.; NOBRE, G. C.; CAÇOLA, P. Análise das oportunidades para o desenvolvimento motor (affordances) em ambientes domésticos no Ceará-Brasil. **Journal of Human Growth and Development**, v. 19, n. 1, p. 9–18, 2009. 3, 50, 79
- PAIVA, R. C. D.; BUARQUE, D. C.; CLARKE, R. T.; COLLISCHONN, W.; ALLASIA, D. G. Reduced precipitation over large water bodies in the Brazilian Amazon shown from TRMM data. **Geophysical Research Letters**, v. 38, n. 4, 2011. 113
- PALHARINI, R.; VILA, D.; RODRIGUES, D.; PALHARINI, R.; MATTOS, E.; UNDURRAGA, E. Analysis of extreme rainfall and natural disasters events using satellite precipitation products in different regions of Brazil. **Atmosphere**, v. 13, n. 10, p. 1680, 2022. 10
- PASSOW, M. J. TRMM-tropical rainfall measuring mission: bringing remote sensing of precipitation into your classroom. **Terræ Didactica**, v. 6, n. 1, p. 03–08, 2010. 35
- PETERSEN, W. A.; FU, R.; CHEN, M.; BLAKESLEE, R. Intraseasonal forcing of convection and lightning activity in the southern Amazon as a function of cross-equatorial flow. **Journal of Climate**, v. 19, n. 13, p. 3180–3196, 2006. 94
- PETERSEN, W. A.; RUTLEDGE, S. A. On the relationship between cloud-to-ground lightning and convective rainfall. **Journal of Geophysical Research: Atmospheres**, v. 103, n. D12, p. 14025–14040, 1998. 45
- PHILLIPS, O. L.; BRIENEN, R. J. Carbon uptake by mature Amazon forests has mitigated Amazon nations' carbon emissions. **Carbon Balance and Management**, v. 12, n. 1, p. 1–9, 2017. 50
- PIELKE, R.; CARBONE, R. E. Weather impacts, forecasts, and policy: an integrated perspective. **Bulletin of the American Meteorological Society**, v. 83, n. 3, p. 393–406, 2002. 11

PIERSANTE, J. O.; RASMUSSEN, K. L.; SCHUMACHER, R. S.; ROWE, A. K.; MCMURDIE, L. A. A synoptic evolution comparison of the smallest and largest mcs in subtropical South America between spring and summer. **Monthly Weather Review**, v. 149, n. 6, p. 1943–1966, 2021. 80, 84

PONS, P.; LATAPY, M. Computing communities in large networks using random walks. In: INTERNATIONAL SYMPOSIUM ON COMPUTER AND INFORMATION SCIENCES, 20., 2005, Turkey. **Proceedings...** Istanbul, 2005. p. 284–293. 27

QUIROGA, R. Q.; KREUZ, T.; GRASSBERGER, P. Event synchronization: a simple and fast method to measure synchronicity and time delay patterns. **Physical Review E**, v. 66, n. 4, p. 041904, 2002. 53, 73

RAMOS DA SILVA, R.; GANDU, A. W.; SÁ, L. D.; DIAS, M. A. S. Cloud streets and land–water interactions in the Amazon. **Biogeochemistry**, v. 105, p. 201–211, 2011. 113

RANJAN, P.; KAZAMA, S.; SAWAMOTO, M. Effects of climate change on coastal fresh groundwater resources. **Global Environmental Change**, v. 16, n. 4, p. 388–399, 2006. 9

RASMUSSEN, K. L.; CHAPLIN, M.; ZULUAGA, M.; JR, R. H. Contribution of extreme convective storms to rainfall in South America. **Journal of Hydrometeorology**, v. 17, n. 1, p. 353–367, 2016. 50, 79, 82

RASMUSSEN, K. L.; HOUZE, R. A. Orographic convection in subtropical South America as seen by the TRMM satellite. **Monthly Weather Review**, v. 139, n. 8, p. 2399–2420, 2011. 17, 18, 79

RASMUSSEN, K. L.; ZULUAGA, M. D.; HOUZE, R. A. Severe convection and lightning in subtropical South America. **Geophysical Research Letters**, v. 41, n. 20, p. 7359–7366, 2014. 15, 16, 49, 80, 94, 95

RASMUSSEN, E. M.; MO, K. C. Large-scale atmospheric moisture cycling as evaluated from nmc global analysis and forecast products. **Journal of Climate**, v. 9, n. 12, p. 3276–3297, 1996. 7

REHBEIN, A.; AMBRIZZI, T.; MECHOSO, C. R. Mesoscale convective systems over the Amazon basin. part I: climatological aspects. **International Journal of Climatology**, v. 38, n. 1, p. 215–229, 2018. 86

REHFELD, K.; MARWAN, N.; BREITENBACH, S. F.; KURTHS, J. Late holocene asian summer monsoon dynamics from small but complex networks of paleoclimate data. **Climate Dynamics**, v. 41, p. 3–19, 2013. 30

RHEINWALT, A.; MARWAN, N.; KURTHS, J.; WERNER, P.; GERSTENGARBE, F.-W. Boundary effects in network measures of spatially embedded networks. In: **2012 SC Companion: High Performance Computing, Networking Storage and Analysis**. [S.l.: s.n.], 2012. p. 500–505. 68

- ROCHA, R. P. D.; MORALES, C. A.; CUADRA, S. V.; AMBRIZZI, T. Precipitation diurnal cycle and summer climatology assessment over South America: an evaluation of regional climate model version 3 simulations. **Journal of Geophysical Research: Atmospheres**, v. 114, n. D10, 2009. 85
- RODRIGUES, M. A.; GARCIA, S. R.; KAYANO, M. T.; CALHEIROS, A. J.; ANDREOLI, R. V. Onset and demise dates of the rainy season in the south american monsoon region: a cluster analysis result. **International Journal of Climatology**, v. 42, n. 3, p. 1354–1368, 2022. 2
- ROMATSCHKE, U.; HOUZE, R. A. Characteristics of precipitating convective systems accounting for the summer rainfall of tropical and subtropical South America. **Journal of Hydrometeorology**, v. 14, n. 1, p. 25–46, 2013. 82
- ROMATSCHKE, U.; JR, R. A. H. Extreme summer convection in South America. **Journal of Climate**, v. 23, n. 14, p. 3761–3791, 2010. 79
- ROMATSCHKE, U.; MEDINA, S.; HOUZE, R. A. Regional, seasonal, and diurnal variations of extreme convection in the South Asian region. **Journal of Climate**, v. 23, n. 2, p. 419–439, 2010. 14, 15, 18, 24, 42
- ROY, S. S.; BALLING JUNIOR, R. C. Trends in extreme daily precipitation indices in India. **International Journal of Climatology**, v. 24, n. 4, p. 457–466, 2004. 18
- SAHA, M.; MITRA, P. Identification of indian monsoon predictors using climate network and density-based spatial clustering. **Meteorology and Atmospheric Physics**, v. 131, p. 1301–1314, 2019. 2, 31
- SAKAMOTO, M.; AMBRIZZI, T.; POVEDA, G. Life cycle of convective systems over western Colombia. In: MOCA, IAMAS, IAPSO AND IACS JOINT ASSEMBLY, 2009, Canada. **Proceedings...** Montreal, 2009. p. 15. 15
- SALEH, M.; ESA, Y.; MOHAMED, A. Applications of complex network analysis in electric power systems. **Energies**, v. 11, n. 6, p. 1381, 2018. 51
- SALIO, P.; NICOLINI, M.; ZIPSER, E. J. Mesoscale convective systems over southeastern South America and their relationship with the South American low-level jet. **Monthly Weather Review**, v. 135, n. 4, p. 1290–1309, 2007. 80, 84
- SANCHES, F. et al. Extreme rainfall events in the southwest of Rio Grande do Sul (Brazil) and its association with the sandization process. **American Journal of Climate Change**, v. 8, n. 04, p. 441, 2019. 10
- SANTOS, E. B.; LUCIO, P. S.; SILVA, C. M. S. e. Precipitation regionalization of the brazilian Amazon. **Atmospheric Science Letters**, v. 16, n. 3, p. 185–192, 2015. 93
- SATYAMURTY, P.; COSTA, C. P. W. da; MANZI, A. O. Moisture source for the Amazon basin: a study of contrasting years. **Theoretical and Applied Climatology**, v. 111, p. 195–209, 2013. 111

- SCHAUB, M. T.; DELVENNE, J.-C.; ROSVALL, M.; LAMBIOTTE, R. The many facets of community detection in complex networks. **Applied Network Science**, v. 2, n. 1, p. 1–13, 2017. 27
- SELUCHI, M. E.; GARREAUD, R.; NORTE, F. A.; SAULO, A. C. Influence of the subtropical Andes on baroclinic disturbances: a cold front case study. **Monthly Weather Review**, v. 134, n. 11, p. 3317–3335, 2006. 91
- SELUCHI, M. E.; MARENGO, J. A. Tropical–midlatitude exchange of air masses during summer and winter in South America: climatic aspects and examples of intense events. **International Journal of Climatology**, v. 20, n. 10, p. 1167–1190, 2000. 89
- SELUCHI, M. E.; SAULO, A. C.; NICOLINI, M.; SATYAMURTY, P. The north-western argentinian low: a study of two typical events. **Monthly Weather Review**, v. 131, n. 10, p. 2361–2378, 2003. 89, 103
- SILVA JUNIOR, C. H.; ALMEIDA, C. T.; SANTOS, J. R.; ANDERSON, L. O.; ARAGÃO, L. E.; SILVA, F. B. Spatiotemporal rainfall trends in the Brazilian legal Amazon between the years 1998 and 2015. **Water**, v. 10, n. 9, p. 1220, 2018. 93
- SILVA, V. B.; KOUSKY, V. E. The south american monsoon system: climatology and variability. **Modern Climatology**, v. 123, p. 152, 2012. 80
- SIQUEIRA, J. R.; MACHADO, L. A. T. Influence of the frontal systems on the day-to-day convection variability over South America. **Journal of Climate**, v. 17, n. 9, p. 1754–1766, 2004. 14
- SIQUEIRA, J. R.; ROSSOW, W. B.; MACHADO, L. A. T.; PEARL, C. Structural characteristics of convective systems over South America related to cold-frontal incursions. **Monthly Weather Review**, v. 133, n. 5, p. 1045–1064, 2005. 14
- SONEJA, S.; JIANG, C.; UPPERMAN, C. R.; MURTUGUDDE, R.; MITCHELL, C. S.; BLYTHE, D.; SAPKOTA, A. R.; SAPKOTA, A. Extreme precipitation events and increased risk of campylobacteriosis in Maryland, USA. **Environmental Research**, v. 149, p. 216–221, 2016. 18
- SRINIVASAN, S.; HYMAN, J. D.; O’MALLEY, D.; KARRA, S.; VISWANATHAN, H. S.; SRINIVASAN, G. **Machine learning techniques for fractured media**. United States: [s.n.], 2020. 109–150 p. ISBN 978-0-12-821669-9. 59
- STEINHAEUSER, K.; CHAWLA, N. V.; GANGULY, A. R. An exploration of climate data using complex networks. In: INTERNATIONAL WORKSHOP ON KNOWLEDGE DISCOVERY FROM SENSOR DATA, 3., 2009. **Proceedings...** Paris, France, 2009. p. 23–31. 27, 28
- _____. Complex networks as a unified framework for descriptive analysis and predictive modeling in climate science. **Statistical Analysis and Data Mining: The ASA Data Science Journal**, v. 4, n. 5, p. 497–511, 2011. 2, 29

- STEINHAEUSER, K.; TSONIS, A. A. A climate model intercomparison at the dynamics level. **Climate Dynamics**, v. 42, p. 1665–1670, 2014. 28
- STROGATZ, S. H. Exploring complex networks. **Nature**, v. 410, n. 6825, p. 268–276, 2001. 51
- STUART, L.; LÜTERBACHER, J.; PATERSON, L.; DEVILLER, R.; CASTONGUAY, S. **United In Science 2022: a multi-organization high-level compilation of the most recent science related to climate change, impacts and responses**. WMO: [s.n.], 2022. 40 p. 9
- SUN, P. G. Community detection by fuzzy clustering. **Physica A: Statistical Mechanics and its Applications**, v. 419, p. 408–416, 2015. 27
- SWANSON, K. L.; TSONIS, A. A. Has the climate recently shifted? **Geophysical Research Letters**, v. 36, n. 6, 2009. 20
- TANAKA, L. d. S.; SATYAMURTY, P.; MACHADO, L. A. Diurnal variation of precipitation in central Amazon basin. **International Journal of Climatology**, v. 34, n. 13, p. 3574–3584, 2014. 113
- TASGIN, M.; BINGOL, H. O. Community detection using boundary nodes in complex networks. **Physica A: Statistical Mechanics and its Applications**, v. 513, p. 315–324, 2019. 27, 28
- THIBEAULT, J.; SETH, A.; WANG, G. Mechanisms of summertime precipitation variability in the bolivian altiplano: present and future. **International Journal of Climatology**, v. 32, n. 13, p. 2033–2041, 2012. 88
- TIRABASSI, G.; MASOLLER, C. Unravelling the community structure of the climate system by using lags and symbolic time-series analysis. **Scientific Reports**, v. 6, p. 29804, 2016. 28
- TOYOSHIMA, K.; MASUNAGA, H.; FURUZAWA, F. A. Early evaluation of ku- and ka-band sensitivities for the global precipitation measurement (GPM) dual-frequency precipitation radar (dpr). **Sola**, v. 11, p. 14–17, 2015. 41
- TRENBERTH, K. E.; STEPANIAK, D. P.; CARON, J. M. The global monsoon as seen through the divergent atmospheric circulation. **Journal of Climate**, v. 13, n. 22, p. 3969–3993, 2000. 7
- TSO, G. K.; YAU, K. K. Predicting electricity energy consumption: a comparison of regression analysis, decision tree and neural networks. **Energy**, v. 32, n. 9, p. 1761–1768, 2007. 51
- TSONIS, A.; SWANSON, K. "On the origins of decadal climate variability: a network perspective". **Nonlinear Processes in Geophysics**, v. 19, n. 5, 2012. 62
- TSONIS, A. A.; ROEBBER, P. J. The architecture of the climate network. **Physica A: Statistical Mechanics and its Applications**, v. 333, p. 497–504, 2004. 19, 21, 24, 51

- TSONIS, A. A.; SWANSON, K.; KRAVTSOV, S. A new dynamical mechanism for major climate shifts. **Geophysical Research Letters**, v. 34, n. 13, 2007. 20, 51
- TSONIS, A. A.; SWANSON, K. L. Topology and predictability of El Niño and La Niña networks. **Physical Review Letters**, v. 100, n. 22, p. 228502, 2008. 20, 24, 27
- _____. Climate mode covariability and climate shifts. **International Journal of Bifurcation and Chaos**, v. 21, n. 12, p. 3549–3556, 2011. 2, 23
- TSONIS, A. A.; SWANSON, K. L.; ROEBBER, P. J. What do networks have to do with climate? **Bulletin of the American Meteorological Society**, v. 87, n. 5, p. 585–596, 2006. 19, 20, 21, 24, 51
- TSONIS, A. A.; SWANSON, K. L.; WANG, G. On the role of atmospheric teleconnections in climate. **Journal of Climate**, v. 21, n. 12, p. 2990–3001, 2008. 2, 20, 22, 23, 51
- TSONIS, A. A.; WANG, G.; SWANSON, K. L.; RODRIGUES, F. A.; COSTA, L. da F. Community structure and dynamics in climate networks. **Climate Dynamics**, v. 37, n. 5-6, p. 933–940, 2011. 20, 23, 29, 31
- VELASCO, I.; FRITSCH, J. M. Mesoscale convective complexes in the Americas. **Journal of Geophysical Research: Atmospheres**, v. 92, n. D8, p. 9591–9613, 1987. 8, 14, 16
- VERA, C. et al. The South American low-level jet experiment. **Bulletin of the American Meteorological Society**, v. 87, n. 1, p. 63–78, 2006. 2, 79, 88, 110
- _____. Toward a unified view of the american monsoon systems. **Journal of Climate**, v. 19, n. 20, p. 4977–5000, 2006. 7
- VUILLE, M.; BURNS, S. J.; TAYLOR, B. L.; CRUZ, F. W.; BIRD, B. W.; ABBOTT, M. B.; KANNER, L. C.; CHENG, H.; NOVELLO, V. F. A review of the south american monsoon history as recorded in stable isotopic proxies over the past two millennia. **Climate of the Past**, v. 8, n. 4, p. 1309–1321, 2012. 7
- WALDVOGEL, A.; FEDERER, B.; GRIMM, P. Criteria for the detection of hail cells. **Journal of Applied Meteorology and Climatology**, v. 18, n. 12, p. 1521–1525, 1979. 96
- WALLACE, J. M. Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. **Monthly Weather Review**, v. 103, n. 5, p. 406–419, 1975. 17
- WALLACE, J. M.; GUTZLER, D. S. Teleconnections in the geopotential height field during the northern hemisphere winter. **Monthly Weather Review**, v. 109, n. 4, p. 784–812, 1981. 22
- WALLACE, J. M.; HOBBS, P. V. **Atmospheric science: an introductory survey**. Academic Press: [s.n.], 2006. 12

WANG, G.; SWANSON, K. L.; TSONIS, A. A. The pacemaker of major climate shifts. **Geophysical Research Letters**, v. 36, n. 7, 2009. 23

WEBSTER, P. J.; MAGANA, V. O.; PALMER, T.; SHUKLA, J.; TOMAS, R.; YANAI, M.; YASUNARI, T. Monsoons: processes, predictability, and the prospects for prediction. **Journal of Geophysical Research: Oceans**, v. 103, n. C7, p. 14451–14510, 1998. 7

WEISMAN, M. L.; KLEMP, J. B. Characteristics of isolated convective storms. In: RAY, P. S. (Ed.). **Mesoscale meteorology and forecasting**. Berlin: Springer: [s.n.], 1986. p. 331–358. 11

WHYTE, F. S.; TAYLOR, M. A.; STEPHENSON, T. S.; CAMPBELL, J. D. Features of the caribbean low level jet. **International Journal of Climatology**, v. 28, n. 1, p. 119–128, 2008. 17

WILLIAMS, E.; MUSHTAK, V.; ROSENFELD, D.; GOODMAN, S.; BOCCIPPIO, D. Thermodynamic conditions favorable to superlative thunderstorm updraft, mixed phase microphysics and lightning flash rate. **Atmospheric Research**, v. 76, n. 1-4, p. 288–306, 2005. 11

WU, A.; HSIEH, W. W. The nonlinear northern hemisphere winter atmospheric response to enso. **Geophysical Research Letters**, v. 31, n. 2, 2004. 23

XU, W.; ZIPSER, E. J. Properties of deep convection in tropical continental, monsoon, and oceanic rainfall regimes. **Geophysical Research Letters**, v. 39, n. 7, 2012. 77

YAMASAKI, K.; GOZOLCHIANI, A.; HAVLIN, S. Climate networks around the globe are significantly affected by El Niño. **Physical Review Letters**, v. 100, n. 22, p. 228501, 2008. 20, 21, 23, 51

YIN, J.; GUO, S.; WANG, J.; CHEN, J.; ZHANG, Q.; GU, L.; YANG, Y.; TIAN, J.; XIONG, L.; ZHANG, Y. Thermodynamic driving mechanisms for the formation of global precipitation extremes and ecohydrological effects. **Science China Earth Sciences**, v. 66, n. 1, p. 92–110, 2023. 1

YOU, T.; CHENG, H.-M.; NING, Y.-Z.; SHIA, B.-C.; ZHANG, Z.-Y. Community detection in complex networks using density-based clustering algorithm and manifold learning. **Physica A: Statistical Mechanics and its Applications**, v. 464, p. 221–230, 2016. 27

ZAITSEV, D. A. A generalized neighborhood for cellular automata. **Theoretical Computer Science**, v. 666, p. 21–35, 2017. 79

ZANIN, M.; LILLO, F. Modelling the air transport with complex networks: a short review. **The European Physical Journal Special Topics**, v. 215, n. 1, p. 5–21, 2013. 19

ZEMP, D.; SCHLEUSSNER, C.-F.; BARBOSA, H.; ENT, R. Van der; DONGES, J. F.; HEINKE, J.; SAMPAIO, G.; RAMMIG, A. On the importance of cascading moisture recycling in South America. **Atmospheric Chemistry and Physics**, v. 14, n. 23, p. 13337–13359, 2014. 82

ZHANG, H.; ZAN, X.; HUANG, C.; ZHU, X.; WU, C.; WANG, S.; LIU, W. Detecting dense subgraphs in complex networks based on edge density coefficient. In: INTERNATIONAL CONFERENCE ON BIO-INSPIRED COMPUTING: THEORIES AND APPLICATIONS, 5., 2010. **Proceedings...** Liver pool, UK, 2010. p. 51–53. 27, 28, 74

ZHOU, D.; GOZOLCHIANI, A.; ASHKENAZY, Y.; HAVLIN, S. Teleconnection paths via climate network direct link detection. **Physical Review Letters**, v. 115, n. 26, p. 268501, 2015. 2, 23, 24

ZHOU, J.; LAU, K. Does a monsoon climate exist over South America? **Journal of Climate**, v. 11, n. 5, p. 1020–1040, 1998. 88, 110

ZIPSER, E. J.; CECIL, D. J.; LIU, C.; NESBITT, S. W.; YORTY, D. P. Where are the most intense thunderstorms on earth? **Bulletin of the American Meteorological Society**, v. 87, n. 8, p. 1057–1072, 2006. 15, 49, 77

ZUBIETA, R.; SAAVEDRA, M.; ESPINOZA, J. C.; RONCHAIL, J.; SULCA, J.; DRAPEAU, G.; MARTIN-VIDE, J. Assessing precipitation concentration in the Amazon basin from different satellite-based data sets. **International Journal of Climatology**, v. 39, n. 7, p. 3171–3187, 2019. 18, 53

APPENDIX A

A.1 Data links

Table A.1 - List of links to data used.

Base	url
GPM	< gpm.nasa.gov/data/directory >
TRMM	< trmm.gsfc.nasa.gov/images_dir/images.html >
IMERG	< disc.gsfc.nasa.gov/datasets/GPM_3IMERGDF_06/summary >
PF - Utah	< atmos.tamucc.edu/trmm/ >

A.2 Software

Table A.2 - Programming language and main packages used.

Software/Package	Version	Description
cartopy	0.21.0	Geospatial data processing and plotting
clima_anom	0.7.5	Extract the rainfall data from shapefile
geopandas	0.10.2	Extension for spatial and geometric operations for the data type used by pandas
matplotlib	3.5.0	Creating static and animated visualizations
netcdf4	1.5.7	netCDF read and write
networkx	2.6.3	Creation, manipulation and analysis of complex networks
numpy	1.21.2	Creation is calculations of multidimensional arrays
pandas	1.3.4	Data analysis and manipulation
pip	22.2.2	Official third-party software repository for Python
Python	3.9.13	Programming language used for all study
scipy	1.7.3	Efficient array operations
shapely	1.8.4	Manipulation and analysis of flat geometric objects

