UNIVERSIDADE FEDERAL DE VIÇOSA

Spatio-Temporal Dynamics of Flying Rivers: Assessment of Precipitation Model Effectiveness and Its Influence on the SACZ

Arthur Amaral e Silva
Doctor Scientiae

VIÇOSA - MINAS GERAIS
2025

ARTHUR AMARAL E SILVA

Spatio-Temporal Dynamics of Flying Rivers: Assessment of Precipitation Model Effectiveness and Its Influence on the SACZ

Thesis submitted to the Civil Engineering Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

Adviser: Italo Oliveira Ferreira

Co-adviser: Leonardo Campos de Assis

VIÇOSA - MINAS GERAIS
2025

Ficha catalográfica elaborada pela Biblioteca Central da Universidade Federal de Viçosa - Campus Viçosa

T
S586d
2025

Silva, Arthur Amaral e, 1994-
Dinâmica espaço-temporal dos rios voadores: avaliação da eficácia dos modelos de precipitação e sua influência nas ZCAS / Arthur Amaral e Silva. – Viçosa, MG, 2025.
1 tese eletrônica (148 f.): il.

Inclui apêndice.
Orientador: Ítalo Oliveira Ferreira.
Tese (doutorado) - Universidade Federal de Viçosa, Departamento de Engenharia Civil, 2025.
Inclui bibliografia.
DOI: https://doi.org/10.47328/ufvbbt.2025.403
Modo de acesso: World Wide Web.

1. Precipitação (Meteorologia) - Variabilidade - Brasil. 2. Florestas tropicais - Amazonas. 3. Zonas climáticas - Brasil. 4. Mudanças climáticas. 5. Sensoriamento remoto. I. Ferreira, Ítalo Oliveira, 1988-. II. Universidade Federal de Viçosa. Departamento de Engenharia Civil. Programa de Pós-Graduação em Engenharia Civil. III. Título.

CDD 22. ed. 551.577281

Bibliotecário(a) responsável: Alice Regina Pinto Pires CRB-6/2523

ARTHUR AMARAL E SILVA

Spatio-Temporal Dynamics of Flying Rivers: Assessment of Precipitation Model Effectiveness and Its Influence on the SACZ

Thesis submitted to the Civil Engineering Graduate Program of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of *Doctor Scientiae*.

APPROVED: March 11, 2025.

Assent:

Arthur Amaral e Silva
Author

Italo Oliveira Ferreira
Adviser

À minha avó Maria Helena (*in memoriam*), que sempre esteve ao meu lado, mesmo após sua partida. Dedico.

ACKNOWLEDGMENTS

First and foremost, I express my deepest gratitude to my parents, Solange Amaral and Carlos Silva, for their unconditional support, companionship, love, and for always being by my side. Thank you.

To my brother, Matheus Amaral, I am immensely grateful for your unwavering support, for being my strength during challenging moments, and above all, for the unique bond we share. You are an extension of my heart. I love you. Thank you.

To Professor Italo Oliveira Ferreira, I am thankful for your guidance and supervision throughout these years. Having the opportunity to conduct research under such a remarkable mentor has provided me with personal growth and lifelong learning. Thank you.

To Professor Lúcia Calijuri, I extend my deep admiration and appreciation. From our very first interaction in your class at the university, I was welcomed, guided, and encouraged to explore new horizons. Your teaching and generosity have provided enriching experiences that have profoundly shaped my academic journey. This path would not have been the same without your presence. Thank you.

To Professor Leonardo Campos, my heartfelt thanks. Your collaboration was essential to the development of this work, making my academic journey even more enriching. The knowledge gained during my doctorate, a result of your dedication and teaching, will stay with me forever. I will be eternally grateful for your mentorship and commitment. Thank you.

To Professor Júlio César de Oliveira, I am grateful for your guidance throughout these years of research. It was a privilege to develop this project under your supervision. Thank you for sharing your knowledge. Thank you.
To my friends from Fortaleza-CE—Bruno Calixto, Mayron Freitas, George Façanha, Ivna Milério, and Gustavo Tavares—thank you for remaining present despite the distance and for helping me stay strong throughout this journey. Thank you.

To my brother-friend, Lucas Rodrigues, and to my cousin-sister and best friend, Luiza Helena, I am eternally grateful. You are my daily strength, my foundation, and an essential part of who I am. Thank you.

To my dear friends and chosen family in Viçosa-MG—Iara Magalhães, Alexia Aona, Thiago Abrantes, Laura Andrade, Daniel Louzada, Juliana Lorentz, and Lucas Borges—thank you for every comforting hug during hard times, every moment of celebration, and for all the support shared as we pursue our dreams together. Thank you.

To my beloved friends Hugo Reggiani, Matheus Chagas, and João Guilherme, my deepest gratitude for your friendship, affection, and unconditional support throughout this journey. Every shared moment, every memory built together, made this path even more special. You are an essential part of this achievement. I love you all. Thank you.

To the SIGEOnPA laboratory team, I recognize your support, camaraderie, and the shared learning that made this journey lighter and more rewarding. Thank you.

To the LUVE-UFV volleyball team, who welcomed me from the beginning and became my second home—a place where I could practice the sport I love alongside incredible people—my special thanks to Fernando Fiedler, Maycon Saraiva, Luís Xavier, Vinicius Rodrigues, and Pedro Rodrigues. I love you all. Thank you.

To the Federal University of Viçosa (UFV), for the opportunity to pursue my graduate studies.

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Finally, I extend my gratitude to everyone who contributed in any way to this journey, whether directly or indirectly. Thank you for being part of this achievement.

“Todas as vitórias ocultam
uma abdicação”

Simone de Beauvoir

ABSTRACT

SILVA, Arthur Amaral e, D.Sc., Universidade Federal de Viçosa, March, 2025. **Spatio-Temporal Dynamics of Flying Rivers: Assessment of Precipitation Model Effectiveness and Its Influence on the SACZ**. Adviser: Italo Oliveira Ferreira. Co-adviser: Leonardo Campos de Assis.

The Amazon Rainforest plays a fundamental role in regulating global and regional climate patterns, sustaining one of the most complex hydrological cycles on the planet. Its continuous evapotranspiration process generates atmospheric moisture, which is essential for the formation of both local and distant rainfall, contributing to moisture redistribution through the so-called "Flying Rivers." These large-scale atmospheric moisture transport systems carry vast amounts of water vapor from the Amazon Basin to other regions of South America, particularly southeastern Brazil, ensuring the continuity of precipitation and influencing the regional hydrological regime. However, increasing deforestation and climate change have significantly disrupted this process, compromising the intensity and consistency of the Flying Rivers and, consequently, altering precipitation patterns in southeastern Brazil. One of the primary meteorological systems influenced by the Flying Rivers is the South Atlantic Convergence Zone (SACZ), a key climatic phenomenon responsible for rainfall distribution in southeastern Brazil, especially during the austral summer. The SACZ is characterized by an extensive band of cloud cover and precipitation stretching from the Amazon to the South Atlantic Ocean, regulating water availability, agricultural productivity, and the frequency of extreme climate events such as droughts and floods. Given this climatic interconnectivity, this study aims to assess the variability of precipitation patterns associated with the SACZ and its relationship with the Flying Rivers using advanced remote sensing techniques, climate modeling, and statistical analysis. First, four widely used climate models were analyzed to forecast precipitation, applying statistical approaches to a 20-year time series to evaluate their accuracy against in situ measurements. Additionally, cluster analyses were conducted to examine seasonal rainfall patterns across Brazil and assess the influence of environmental and climatic factors—including temperature, evapotranspiration, NDVI, and topography—on precipitation distribution. Finally, the relationship between SACZ precipitation patterns and other atmospheric phenomena was analyzed over a 21-year period to better understand atmospheric moisture circulation dynamics and its hydrological impacts in Minas Gerais. The results highlight the increasing vulnerability of southeastern Brazil to changes in moisture transport patterns, with significant implications for

water resource management and climate adaptation strategies. The use of advanced Big Data techniques and machine learning improved the predictive analysis of precipitation patterns, reinforcing the importance of integrating satellite data with ground-based measurements for effective climate monitoring.

Keywords: Amazon Rainforest; Flying Rivers; South Atlantic Convergence Zone (SACZ); Climate Change; Remote Sensing; Precipitation Patterns; Climate Modeling.

RESUMO

SILVA, Arthur Amaral e, D.Sc., Universidade Federal de Viçosa, março de 2025. **Dinâmica Espaço-Temporal dos Rios Voadores: Avaliação da Eficácia dos Modelos de Precipitação e sua influência nas ZCAS**. Orientador: Italo Oliveira Ferreira. Coorientador: Leonardo Campos de Assis.

A Floresta Amazônica desempenha um papel fundamental na regulação dos padrões climáticos globais e regionais, sendo responsável por sustentar um dos ciclos hidrológicos mais complexos do planeta. Seu processo contínuo de evapotranspiração gera umidade atmosférica essencial para a formação de chuvas locais e distantes, contribuindo para a redistribuição de umidade através dos chamados "Rios Voadores". Esses sistemas atmosféricos de transporte de vapor d'água deslocam grandes volumes de umidade da Bacia Amazônica para outras regiões da América do Sul, especialmente para o Sudeste do Brasil, garantindo a manutenção da precipitação e influenciando o regime hidrológico regional. No entanto, o desmatamento, que tem levado à degradação da floresta e as mudanças climáticas têm afetado significativamente esse processo, comprometendo a intensidade e regularidade dos Rios Voadores e, consequentemente, alterando os padrões de precipitação no Sudeste brasileiro. Um dos principais sistemas meteorológicos influenciados pelos Rios Voadores é a Zona de Convergência do Atlântico Sul (ZCAS), um fenômeno climático de grande importância para a distribuição das chuvas no Sudeste, especialmente durante o verão austral. A ZCAS é caracterizada por uma extensa banda de nebulosidade e precipitação que se estende da Amazônia ao Oceano Atlântico Sul, modulando a disponibilidade hídrica, a produtividade agrícola e a ocorrência de eventos climáticos extremos, como secas e enchentes. Diante dessa interconectividade climática, o presente estudo teve como objetivo avaliar os padrões de precipitação associados à ZCAS e sua relação com os Rios Voadores, utilizando técnicas avançadas de sensoriamento remoto, modelagem climática e análise estatística. Primeiramente, foram analisados quatro modelos climáticos amplamente utilizados para prever precipitação, aplicando abordagens estatísticas a uma série temporal de 20 anos para avaliar sua precisão em comparação com medições in situ. Além disso, foram realizadas análises de agrupamento para investigar padrões sazonais de chuva no Brasil, bem como a influência de fatores ambientais e climáticos, como temperatura, evapotranspiração, NDVI e relevo, na distribuição das precipitações. Por fim, a relação entre os padrões de precipitação da ZCAS e outros fenômenos atmosféricos foi examinada em uma escala temporal de 21 anos, buscando compreender a dinâmica da circulação de umidade

atmosférica e seus impactos hidrológicos em Minas Gerais. Os resultados destacam a crescente vulnerabilidade do Sudeste do Brasil às mudanças nos padrões de transporte de umidade, com implicações significativas para a gestão de recursos hídricos e estratégias de adaptação climática. O uso de técnicas avançadas de Big Data e aprendizado de máquina permitiu aprimorar a análise preditiva dos padrões de precipitação, evidenciando a importância da integração de dados de satélite e medições terrestres para monitoramento climático. Este estudo contribui para o avanço do conhecimento sobre os impactos das mudanças na Floresta Amazônica na variabilidade climática regional, reforçando a necessidade de políticas de conservação e mitigação para garantir a sustentabilidade dos recursos hídricos e a resiliência climática no Brasil.

Palavras-chave: Floresta Amazônica; Rios Voadores; Zona de Convergência do Atlântico Sul (ZCAS); Mudanças Climáticas, Sensoriamento Remoto; Padrões de Precipitação; Modelagem Climática.

LIST OF TABLES

Table 4.1. Details of the precipitation gridded datasets……...….32
Table 4.2. Statistical Analysis of the Models……...36
Table 4.3. Analysis of models in response to the extreme events……..…...….42
Table 5.1. Details of the gridded datasets……...66
Table 6.1. Key characteristics and sources of the database ……...…...113

LIST OF ILLUSTRATIONS

Figure 4.1. Flowchart of the applied methodology……..……...….27
Figure 4.2. Location map of the study area……..……...….28
Figure 4.3. a) Location of rainfall stations; b) Elevation (m) model of the Brazilian relief; c) Biomes of Brazil. ……..……..…..31
Figure 4.4. Performance of climatic models in terms of the Coefficient of Determination - R^2 metric; A) CHIRPS Model; B) PERSEIANN-CDR Model; C) GLDAS Model; D) TerraClimate Model; E) Boxplot chart of data distribution. ……..……..39
Figure 4.5. Performance of climatic models in terms of the Root Mean Square Error - RMSE metric; A) CHIRPS Model; B) PERSEIANN-CDR Model; C) GLDAS Model; D) TerraClimate Model; E) Boxplot chart of data distribution…….40
Figure 4.6. Performance of climatic models in terms of the Mean Absolute Percentage Error - MAPE metric; A) CHIRPS Model; B) PERSEIANN-CDR Model; C) GLDAS Model; D) TerraClimate Model; E) Boxplot chart of data distribution……...41
Figure 5.1. Methodological flowchart……..…64
Figure 5.2. A) Rain gauge stations distributed in the study area; B) Focal study area..…65
Figure 5.3. Optimal number of clusters……72
Figure 5.4. Principal Component Analysis illustrating variable contributions by month: A) January; B) February; C) March; D) April; E) May; and F) June……74
Figure 5.5. Principal Component Analysis illustrating variable contributions by month: A) July; B) August; C) September; D) October; E) November; and F) December……...75
Figure 5.6. Correlation of variables through Principal Component Analysis (PCA) by month: A) January; B) February; C) March; D) April; E) May; and F) June……..78
Figure 5.7. Correlation of variables through Principal Component Analysis (PCA) by month: A) July; B) August; C) September; D) October; E) November; and F) December……79
Figure 5.8. Cluster analysis of climate and environmental variables. A) January; B) February; C) March; D) April; E) May; F) June; G) July; H) August; I) September; J) October; K) November; L) December; and M) Ribbon graph……...82

Figure 5.9. Cluster analysis focused on the seasonality of rainfall stations. A) January; B) February; C) March; D) April; E) May; F) June; G) July; H) August; I) September; J) October; K) November; L) December; and M) Ribbon graph……87
Figure 5.10. Behavior of Climatological Normals from the Years 1931–1960 (Green), 1961–1990 (Yellow), and 1991–2020 (Blue) in Brazilian Biomes……88
Figure 6.1. Methodological flowchart……110
Figure 6.2. Study area Location……111
Figure 6.3. Amazon Vegetation and Rainfall in MG (2003–2024)……118
Figure 6.4. Spatial distribution of extreme SACZ-related rainfall events across the mesoregions of Minas Gerais during the following hydrological years: A) 2004–2005; B) 2005–2006; C) 2006–2007; D) 2008–2009; E) 2009–2010; F) 2010–2011; G) 2011–2012; H) 2015–2016; I) 2017–2018; J) 2019–2020; K) 2021–2022; L) 2022–2023……121
Figure 6.5. Annual land cover evolution (2004–2023) by class for the mesoregions of Campo das Vertentes, Central Mineira, Jequitinhonha, Metropolitana de Belo Horizonte, Noroeste de Minas, and Norte de Minas. Bars represent the total area (ha) classified as 'Anthropic' (gray) and 'Natural' (green) each year……126
Figure 6.6. Annual land cover evolution (2004–2023) by class for the mesoregions of Oeste de Minas, Sul/Sudoeste de Minas, Triângulo Mineiro/Alto Paranaíba, Vale do Mucuri, Vale do Rio Doce, and Zona da Mata. Bars represent the total area (ha) classified as 'Anthropic' (gray) and 'Natural' (green) each year……127

SUMMARY

PRESENTATION 16

1. GENERAL INTRODUCTION 18

2. HYPOTHESES 21

3. OBJECTIVE 21

3.1. GENERAL OBJECTIVE 21

3.2. SPECIFIC OBJECTIVES 21

4. CAPÍTULO I. Rainfall from Brazilian Flying Rivers: Evaluating the Effectiveness of Precipitation Gridded Databases[1] 22

4.1. Introduction 22

4.2. Methodology 26

4.2.1. Study Area 27

4.2.2. Data Collection and Standardization 29

4.2.3. Mathematical Basis of Regression Metrics 34

4.2.4. Multiple Linear Regression (MLR) 35

4.2.5. Results Compilation 36

4.3. Results 36

4.3.1. Statistical Analysis of Rainfall Models 36

4.3.2. Analysis of models in the face of extreme weather events 42

4.4. Discussion 42

4.5. Conclusion 50

Acknowledgments 51

5. CAPÍTULO II. Multivariate Statistical Analysis of Rainfall Variability in Brazil: Assessing Climatic and Environmental Drivers of Precipitation 61

5.1. Introduction 62

5.2. Methodology 64

5.2.1. Study Area Definition 65

5.2.2. Data acquisition 66

5.2.3. Data preprocessing 68

5.2.4. Database processing 69

5.3. Results 73

5.3.1. The interrelationship between climatic and environmental variables and their impact on climate behavior 73
5.3.2. Cluster-based evaluation of rainfall stations: integrating climatic and environmental variables 81
5.4. Discussion 88
5.5. Conclusion 96
6. CAPÍTULO III. From Drought to Flood: The Role of SACZ and Land-Use Change in Shaping Extreme Rainfall Patterns in Minas Gerais 106
6.1. Introduction 107
6.2. Methodology 110
6.2.1. Study Area Definition 110
6.2.2. Database Acquisition 113
6.2.3. Data base pre-processing and processing 115
6.2.4. Results Interpretation 117
6.3. Results 117
6.3.1. Amazon Vegetation and Rainfall in MG (2003–2024) 118
6.3.2. Spatiotemporal analysis of the impacts of SACZ events in Minas Gerais 120
6.4. Discussion 128
6.5. Conclusion 131
References 132
7. GENERAL CONCLUSIONS 140
8. SUGGESTIONS FOR FUTURE RESEARCH 141
9. GENERAL REFERENCES 142
10. APPENDIX I 148

PRESENTATION

This proposal builds upon the master's research titled "Anthropic Actions and the Degradation of the Amazon Biome: Studies of Impacts and Soil Recovery", developed by the author. That study examined climate variability, changes in native vegetation, and the degree of anthropization in the Legal Amazon and the Amazon biome over time, associating these changes with the environmental impacts observed in the region. Based on the findings and considering the intensifying effects of climate change, the need to further investigate the consequences of changes in the Amazon biome beyond its physical boundaries became evident.

Recent studies reinforce the critical role of the Amazon Forest in climate regulation, particularly regarding rainfall patterns, through the atmospheric moisture transport process known as the Flying Rivers. These moisture fluxes are essential for recharging water resources such as the São Francisco River, which originates in the state of Minas Gerais and is one of Brazil's most important watercourses. The degradation of the Amazon's forest cover reduces the moisture available for transport, directly impacting rainfall regularity in the Southeast and jeopardizing the region's water availability.

In this context, the main objective of this proposal is to assess the influence of the Flying Rivers on precipitation variability in Brazil, with an emphasis on the relationship between Amazonian moisture transport and the dynamics of the South Atlantic Convergence Zone (SACZ), through the use of climate models, statistical analyses, and advanced meteorological monitoring techniques.

To achieve this, the research is structured into three chapters. Chapter one evaluates widely used climate models for precipitation forecasting, applying statistical methods to a 20-year monthly time series to analyze their performance in estimating the spatiotemporal variability of rainfall compared to in situ measurements. Chapter two conducts independent cluster analyses to investigate different aspects of rainfall variability in Brazil—one based solely on rainfall station data, and the other integrating biogeophysical and climatic variables to explore their influence on precipitation distribution. Finally, chapter three analyzes the relationship between SACZ precipitation patterns and other atmospheric phenomena over a 21-year daily time series, aiming to understand how atmospheric moisture transport processes influence regional hydrological behavior in Minas Gerais.

By deepening the understanding of these processes, this research aims to contribute to the development of more effective strategies for mitigating hydrological impacts and adapting to climate change, providing insights to support environmental conservation policies and the sustainable management of water resources in Minas Gerais.

1. GENERAL INTRODUCTION

The Amazon, the largest tropical rainforest in the world, plays a fundamental role in regulating both global and regional climate patterns. Its dense vegetation, extensive river systems, and continuous evapotranspiration processes sustain one of the most significant hydrological cycles on the planet. Over recent decades, scientific research has increasingly emphasized the critical relationship between forests and precipitation, demonstrating that tree transpiration is one of the primary sources of atmospheric moisture over continental areas (Jasechko et al., 2013).

The forest acts as a vast natural pump that absorbs groundwater, releases it into the atmosphere as water vapor, and maintains high humidity levels in the region. This process not only ensures local precipitation but also influences rainfall thousands of kilometers away. Makarieva and Gorshkov (2007) proposed the “biotic pump” mechanism, through which forests actively contribute to the formation of atmospheric pressure gradients, facilitating the transport of moisture from the Amazon to other parts of South America. Sheil and Murdiyarso (2009) expanded this concept by showing how forests generate low-pressure zones that attract clouds and promote rainfall, thereby creating their own climatic conditions.

Based on this understanding, the concept of “Flying Rivers” emerged, describing large-scale atmospheric moisture transport systems responsible for transferring vast amounts of water vapor from the Amazon Basin to southeastern Brazil and other regions. Initially introduced by Salati et al. (1979) and later refined by Salati and Vose (1984), Flying Rivers are analogous to terrestrial rivers, channeling moisture across great distances. Arraut et al. (2012) highlighted the role of low-level jets in driving these atmospheric moisture flows, emphasizing their importance as key hydro-meteorological pathways. The Amazon Forest, through its high rates of evapotranspiration, generates significant atmospheric moisture, which is transported westward by trade winds and then redirected southward by the Andes Mountains, forming the Flying Rivers (Pearce, 2019). The volume of water transported by this mechanism is estimated to be comparable to that of the Amazon River itself, underscoring the forest’s vital role in maintaining South America’s precipitation patterns (Pearce, 2020).

The significance of the Flying Rivers extends beyond the Amazon, as they provide a substantial portion of the rainfall that sustains Brazil’s agricultural regions and supplies

major urban centers such as São Paulo, Minas Gerais, and Rio de Janeiro (Monteiro & Campelo, 2022). However, deforestation and climate change have increasingly disrupted this delicate balance. The removal of large forested areas reduces the amount of moisture available for transport, weakening the intensity and consistency of the Flying Rivers (Nobre et al., 2016). As a result, southeastern Brazil has experienced increasing climate variability, with more frequent prolonged droughts and irregular rainfall patterns (Zemp et al., 2017). Weng et al. (2018) demonstrated a direct correlation between deforestation in the Amazon and decreased precipitation in distant regions, reinforcing the urgent need for conservation measures. Moreover, Nacur and Vartuli (2021) observed that changes in the Amazonian hydrological cycle—driven by land use change—affect climate patterns not only in Brazil but also in neighboring countries.

One of the meteorological systems most influenced by the Flying Rivers is the South Atlantic Convergence Zone (SACZ). This quasi-stationary convective system plays a crucial role in modulating rainfall distribution in southeastern Brazil, particularly during the austral summer (Silva et al., 2019). Characterized by an extensive band of cloudiness and precipitation stretching from the Amazon Basin to the South Atlantic Ocean, the SACZ is a key component in regulating regional hydrological cycles, influencing water availability, agricultural productivity, and the frequency of extreme weather events such as floods and droughts (Filho et al., 2022). Despite its importance, forecasting SACZ variability remains challenging due to complex interactions among atmospheric circulation patterns, large-scale climate oscillations, and moisture transport processes.

The interdependence between SACZ dynamics and moisture transport from the Amazon is a fundamental topic of study. Flying Rivers represent one of the main sources of moisture for SACZ formation, enhancing convective activity and sustaining prolonged rainfall episodes (Amaral e Silva et al., 2024). However, as deforestation continues to reduce the volume of water vapor transported to southeastern Brazil, the SACZ has become increasingly unstable. Studies by Costa et al. (2024) and Correia et al. (2024) showed that reduced moisture transport from the Amazon is correlated with changes in SACZ precipitation patterns, resulting in an increased frequency of hydro-meteorological extremes such as severe droughts, catastrophic floods, flash floods, and rainfall-induced landslides.

Given this climatic interconnectivity, there is an urgent need for high-resolution monitoring systems capable of capturing the complex factors that influence SACZ variability and its hydrological implications. Recent advances in remote sensing technologies and big data from satellite-based precipitation representations have significantly improved the ability to analyze SACZ dynamics and their associated impacts (Saddique et al., 2022). The integration of these datasets with ground-based meteorological observations allows for a more comprehensive understanding of rainfall variability and SACZ-driven precipitation patterns over long-term climatological scales.

Moreover, the growing availability of artificial intelligence (AI) techniques has revolutionized climate research, enabling the processing of vast volumes of meteorological data with high precision. Machine learning algorithms and cloud computing platforms facilitate the identification of complex spatiotemporal patterns in rainfall variability, enhancing real-time monitoring and predictive capabilities for extreme precipitation events associated with the SACZ (Sondermann et al., 2022; Elmahal & Musa, 2023). These technological innovations have direct implications for water resource management, disaster risk reduction, and climate adaptation planning in southeastern Brazil.

In this context, the objectives of this study were to: (i) analyze the spatiotemporal variability of rainfall using four precipitation models, applying statistical approaches to a 20-year monthly time series in order to evaluate their performance in estimating rainfall variability compared to in situ measurements; (ii) assess seasonal rainfall patterns and their relationship with the Flying Rivers, incorporating biogeophysical and climatic variables to investigate different aspects of rainfall variability in Brazil, while also presenting appropriate methodological tools and approaches that reinforce the influence of these factors on precipitation distribution; and (iii) examine the relationship between SACZ rainfall patterns and other atmospheric phenomena over a 21-year daily time series, aiming to understand how atmospheric moisture transport processes influence regional hydrological behavior in Minas Gerais.

This study contributes to a deeper understanding of rainfall variability and its implications for climate risk management, reinforcing the importance of data-driven approaches for monitoring and adapting to climate change in Brazil.

2. HYPOTHESES

- The transport of moisture from the Amazon Rainforest plays a fundamental role in sustaining the convective activity of the South Atlantic Convergence Zone (SACZ) and shaping precipitation patterns in southeastern Brazil. Consequently, the increasing anthropogenic degradation of the forest directly affects the spatiotemporal dynamics of the SACZ and alters the hydrological cycle in the state of Minas Gerais..

3. OBJECTIVE

3.1. GENERAL OBJECTIVE

To assess the influence of the Flying Rivers on precipitation variability in Brazil, with emphasis on the relationship between Amazonian moisture transport and SACZ dynamics, and their impact on southeastern Brazil, particularly the state of Minas Gerais.

3.2. SPECIFIC OBJECTIVES

- To evaluate the performance of the CHIRPS, GLDAS, TerraClimate, and PERSIANN precipitation datasets in representing rainfall patterns in regions influenced by Brazilian Flying Rivers, by comparing satellite estimates with in situ observations.

- To investigate the spatial and temporal variability of precipitation in relation to environmental and climatic factors, initially analyzing seasonal rainfall behavior within the context of the Flying Rivers, and subsequently incorporating climatic (temperature), hydrological (evapotranspiration), and biogeophysical (NDVI and topography) variables.

- To refine the characterization of SACZ variability and assess its socio-environmental impacts in Minas Gerais, focusing on extreme precipitation events and their implications for climate risk management, through the integration of satellite remote sensing, big data analytics, and land use/land cover change analyses over time.

4. CAPÍTULO I. Rainfall from Brazilian Flying Rivers: Evaluating the Effectiveness of Precipitation Gridded Databases[1]

Abstract: The uneven global distribution of rainfall significantly impacts water resources and environmental sustainability, emphasizing the need for reliable climate prediction models. Accurate predictions are vital for sectors such as food security, urban planning, and disaster management. Data from ground stations, radars, and satellites are essential, despite challenges like instrumental errors. Satellites, with their comprehensive sensors, are crucial for atmospheric observations, aiding in the prediction of large-scale climatic events. Climate models such as CHIRPS, GLDAS, TerraClimate, and PERSIANN use different approaches to analyze precipitation data, which is key to understanding its spatial and temporal variability. This study evaluated (rainfall data) from these four climate models over 20 years (within the Brazilian territory), focusing on the spatiotemporal behavior of rainfall using statistical metrics such as R^2, RMSE, and MAPE. The findings showed that CHIRPS had the best performance (R^2 = 0.843; RMSE = 42.83; MAPE = 0.09%), excelling in both overall database and extreme event analyses. TerraClimate, initially the lowest-performing model (R^2 = 0.413; RMSE = 91.56; MAPE = 0.23%), improved significantly when combined with elevation through multiple linear regression (MLR), achieving R^2 of 0.718, RMSE of 31.14, and MAPE of 9.56%. This made TerraClimate a viable model for studying the Flying Rivers. The study highlights that model selection should align with the specific characteristics of the area under consideration, with CHIRPS being particularly suitable for the studied region. This research enhances the understanding of the effectiveness of these models in estimating rainfall compared to in-situ measurements, which is crucial for various applications. The authors advocate for further studies to advance research on the Flying Rivers, their significance, and the impacts of climate change on them.

Keywords: Amazon Rainforest, Precipitation Models, Statistical Analysis, Flying Rivers, Remote Sensing.

[1] Versão publicada na revista *International Journal of Climatology* (Apêndice I)

4.1. Introduction

The link between forests and rainfall has become clearer in recent years. Initially, tree transpiration was identified as the primary source of water in continental areas

(Jasechko et al., 2013). Moreover, a mechanism by which forests pump water into the atmosphere and sustain the moisture necessary for their survival was proposed by Makarieva and Gorshkov (2007) and further elaborated by Sheil and Murdiyarso (2009). This mechanism involves a reduction in atmospheric pressure at lower levels, which draws clouds over forested regions. Additionally, forests release biogenic volatile organic compounds that act as nuclei for water condensation. These processes result in the production of vast amounts of atmospheric water in forested regions, essentially forming “flying rivers” (Nobre, 2014).

The concept of the "Flying River," also known as "Aerial Rivers," was first defined by Salati et al. (1979) and further by Salati and Vose (1984). It describes the aerial transfer of water vapor from the Amazon rainforest to other regions. The term "aerial river" refers to the preferential pathways of moisture flow, whether narrow or broad, due to the near-perfect analogy with surface rivers (Arraut et al., 2012). According to Pearce (2019), flying rivers travel across the globe, impacting rainfall over vast distances. Researches are increasingly recognizing that forests play a crucial role in sustaining these rivers, making the loss of moisture recycling due to deforestation an even more immediate threat than global warming in many parts of the world.

The term "Flying Rivers" gained popularity in Brazil through Dr. José Marengo, who used it to describe a meteorological phenomenon technically known as “low-level jets.” These air currents carry water vapor across the Amazon and down the eastern slope of the Andes, reaching various parts of Brazil and even as far as Northern Argentina. In some cases, these air currents develop a core of particularly high speed, known as the South American low-level jet (Arraut et al., 2012). In this context, flying rivers refer to a phenomenon composed of various climatological events. These, combined with the replenishment of air masses through the evapotranspiration of forests and their components, as well as the evaporation of water bodies, create a water-laden air current similar to terrestrial rivers. The Amazon flying river, for instance, is estimated to carry as much water as the giant terrestrial river flowing below it (Pearce, 2020).

The abundant vegetation in the Amazon acts like a massive sponge, absorbing water from the soil and transpiring it into the atmosphere through a process known as evapotranspiration (Satyamurty *et al.*, 2013). As water evaporates from the leaves, it forms water vapor that rises into the atmosphere. This water vapor can travel great

distances carried by prevailing winds. In the case of the Amazon, the moisture-laden air masses can travel thousands of kilometers, influencing weather patterns and precipitation in distant regions (Arraut *et al.*, 2012).

When this moisture-laden air encounters cooler temperatures or reaches higher altitudes, it condenses and forms clouds. These clouds eventually release the water as precipitation, which can fall as rain or snow depending on the local climate conditions. Thus, the water vapor transported from the Amazon rainforest contributes to rainfall in areas far beyond the forest itself (Fetter et al., 2018).

However, the Flying River faces threats due to deforestation and climate change. As large areas of the Amazon are cleared for agriculture, urbanization, and other purposes, the capacity of the rainforest to generate and release water vapor into the atmosphere is diminished. This not only affects local ecosystems but also disrupts weather patterns in distant regions that depend on the Flying River for precipitation.

The irregular distribution of rainfall around the globe, coupled with its direct influence on water resource availability and environmental sustainability, emphasizes the importance of developing reliable and comprehensive approaches for predicting and modeling these climatic events (Guofeng et al., 2017; Kumar et al., 2017). This type of modeling plays a fundamental role in understanding climatic systems and in decision-making across a wide range of sectors, essential for food and water security, natural disaster management, urban planning, environmental conservation, energy generation, and others (Rocha and Soares, 2015; Santos and Farias, 2017).

Among the primary data used to predict, understand, and analyze precipitation patterns are those collected through ground-based weather stations, weather radars, and satellites (Zhou et al., 2023). Ground-based weather stations are a fundamental source of local and regional precipitation data, while radars and satellites provide information with broader spatial coverage (Noh et al., 2009). Modern weather stations have employed pluviographs (automatic gauges) to ensure the data recording with a quality standard over the years. However, this poses a challenge for technicians as various pluviometric instruments exhibit inherent errors in their functional systems, in addition to the influence of meteorological variables such as wind, temperature, and variations in rainfall intensity, among others (Kidd et al., 2017).

Data collection from radars and satellites uses real-time information to anticipate future climate behavior. The satellites are equipped with sensors capable of capturing data in various parts of the electromagnetic spectrum, including visible light and infrared bands, which play a crucial role in atmospheric observation (Moraes, 2022). The sensors are especially useful for monitoring cloud cover and investigating complex atmospheric phenomena. In the visible light range, satellites capture images during the day, while in the near and mid-infrared, they detect thermal radiation, allowing observations even during the night or in low-light conditions (Arejano et al., 2023).

This information is essential for weather forecasting and climate research, assisting scientists in understanding climate patterns, such as sea surface temperature, moisture distribution, cloud presence and movement, and precipitation intensity (Givati et al., 2016; Ruggieri et al., 2021). It also helps measure the amount of precipitation in a specific area and the formation and movement of storms, among other relevant data. These observations are crucial for predicting large-scale weather systems, such as hurricanes, cold fronts, and other significant meteorological events (Urzagasti et al., 2020). Additionally, this data has applications in various other areas, including agriculture, environmental monitoring, and natural disaster management (Oliveira et al., 2022).

It is noteworthy that the systems mentioned above have distinct approaches to obtaining and analyzing precipitation data. The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) model, for example, stands out for combining satellite information with data from weather stations, providing real-time high-resolution estimates (Funk et al., 2015). GPM - Global Precipitation Measurement is a global mission that uses advanced sensors for precise measurements worldwide (Smith et al., 2007). The Tropical Rainfall Measuring Mission (TRMM) model, when active, focused on tropical regions and used radar and microwave technology. The Global Land Data Assimilation System (GLDAS) focuses on modeling land processes related to water (Kumar et al., 2017; Guofeng et al., 2017). The TerraClimate model provides high-resolution global climate data (Abatzoglou et al., 2018), while the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks- Climate Data Record (PERSIANN) uses neural networks to estimate precipitation from satellite data (Sorooshian et al., 2000).

Precipitation intensity exhibits significant fluctuation in both space and time. Point records obtained through rain gauges and remote sensing observations reveal that

rainfall data varies on a spatial scale from a few meters to hundreds of kilometers, generating a high degree of uncertainty in representing its variability (Trovati and Antonio, 2007). In this context, one way to analyze the quality of produced models is through the use of statistical metrics that assist in verifying the fit of these models when compared to observed in-situ values. An example of widely applied metrics in different studies is the Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE).

Studying Flying Rivers presents a significant challenge and remains a relatively underexplored topic in the literature. These air masses cover vast areas and exhibit varying behaviors depending on their location, influenced by multiple external factors, which complicates their study. The novelty of this work lies in the statistical analysis of which climate model is best suited to study the extent of Flying Rivers. The study focuses on the rainfall generated by this phenomenon, taking into account the diverse landscapes that influence its recharge and utilizing data from 937 weather stations along its pathway.

Furthermore, specific studies such as the assessment of extreme events and the association of rainfall occurrence with variables such as topography were also addressed in this research. So, this paper primarily evaluated four commonly employed climate models for predicting precipitation information using statistical approaches. Additionally, the research utilized a 20-year time series on a monthly scale to evaluate the models' performance in estimating the spatiotemporal behavior of rainfall compared to in-situ measured information.

4.2. Methodology

The methodology for developing this research was divided into three stages: study area, database collection and standardization, and mathematical background of regression metrics (Figure 4.1).

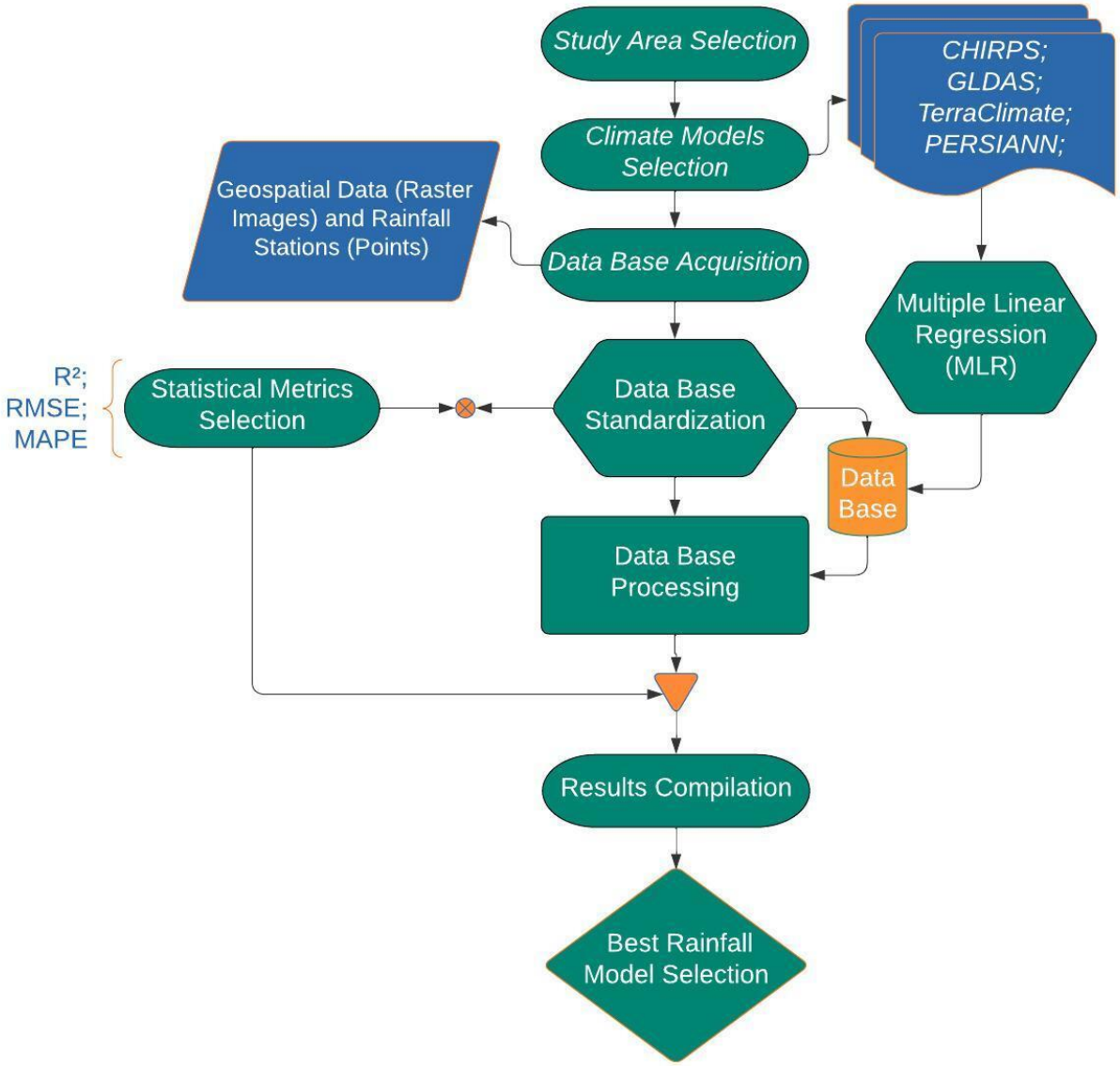

Figure 4.1. Flowchart of the applied methodology.

4.2.1. Study Area

The study area encompasses the trajectory of descending air masses from the Amazon towards Southeast Brazil, known as "Flying Rivers" (Figure 4.2). The figure displays the annual averages related to each data source used in this research. The curved line corresponds to the alignment of the natural path of Flying Rivers.

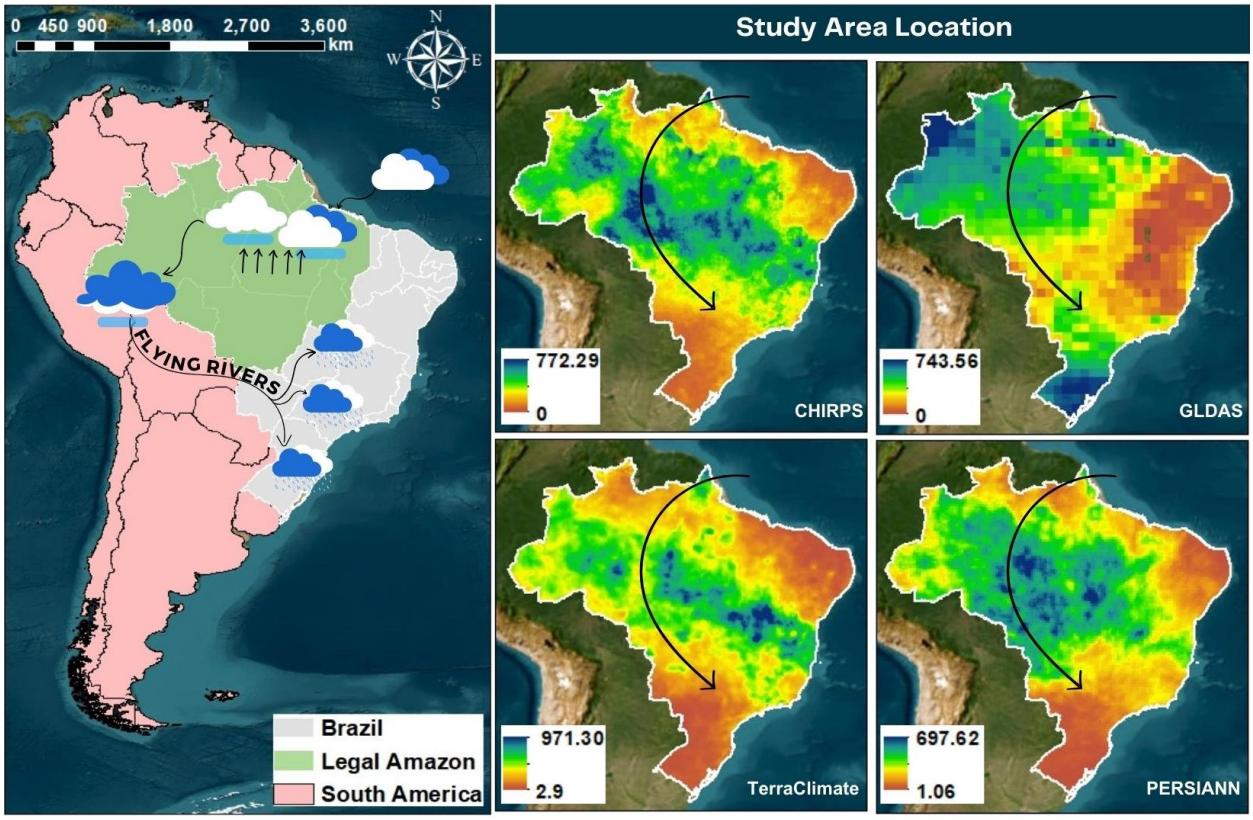

Figure 4.2. Location map of the study area.

The concept of " Tropospheric Rivers " was introduced in 1992 by Reginald Newell and Nicholas Newell to describe atmospheric flows that transport large amounts of water vapor, often surpassing the Amazon River's flow. The Amazon rainforest significantly contributes to this phenomenon by maintaining high air moisture content and exporting atmospheric rivers, leading to rain in distant regions during the southern hemisphere summer (Nobre, 2015). Its significance was first highlighted in 1979 by Brazilian agronomist Eneas Salati, who conducted studies on the isotopic composition of rainwater from the Amazon Basin, showing that half of the Amazon's rainfall comes from the forest's transpiration (Pearce, 2020). During meteorological investigations into the South American Low-Level Jet winds traveling east to west across the Amazon, Salati and his colleagues theorized that this jet carried much of the transpired moisture, which they termed as “Flying Rivers”. The term “Atmospheric River” was later introduced to describe filamentary structures in the vertically integrated moisture flow field responsible for intense water transport. Nowadays, it is understood that these moisture flows globally, affecting regions like the U.S., China, Pakistan, India, and the African Sahel (Newell et al., 1992; Newell and Zhu, 1994; Zhu and Newell, 1998; Pearce, 2020).

These Rivers in the Sky, present within Brazilian territory, cover an extensive area stretching from the northern region of the Amazon to the far north of the southern region, traversing the central-western, southeastern, and northeastern regions of the country (Pearce, 2019). These atmospheric transport systems connect diverse biomes, allowing interaction with vegetation of various characteristics, ranging from dense forests to wetlands, while also traversing different land use types (Weng *et al.*, 2018). Such a trajectory encompasses ecosystems ranging from vast forests, exemplified by the Amazon, to extensive areas dedicated to agriculture and pasture, predominantly found in the central-western region of the country, and also reaches urban centers of significant economic relevance, such as the metropolis of São Paulo (Weng *et al.*, 2018).

They are formed from the evaporation of water from the Atlantic Ocean brought into the Amazon by the trade winds, characterized by their high humidity, resulting in the high rainfall rate of the forest (Nacur and Vartuli, 2021). This air mass in Brazil covers the Amazon and Pantanal biomes, influencing the Cerrado and parts of the Atlantic Forest region (Monteiro and Campelo, 2022).

However, through their leaves, plants release this water back into the atmosphere through transpiration, replenishing the moisture-laden winds, which then move westward until they collide with the Andes Mountains. There, part of this water will replenish the Amazon Basin itself. The remainder will flow south, providing rain to a significant portion of the South American continent (Nacur and Vartuli, 2021).

Moreover, Flying Rivers can also affect the climate of other regions, such as North America. Some of the transported moisture can be carried by the jet stream to the western regions of the United States and Canada, influencing rainfall and snow patterns in these areas (Monteiro and Campelo, 2022).

The vitality of the Amazon Rainforest is crucial for maintaining the volume of Flying Rivers. Deforestation and climate change interfere with this process, negatively impacting rainfall patterns in various regions. Therefore, understanding and protecting them is essential for the Amazon region, global climate, and the world's water security.

4.2.2. Data Collection and Standardization

The database used in this work included point vector data from rainfall stations (Agência Nacional de Águas e Saneamento Básico - ANA), as well as rainfall models provided by different providers such as the National Aeronautics and Space

Administration (NASA), Climatology Lab, Climate Hazards Center, and National Centers for Environmental Information (NCEI).

4.2.2.1. Acquisition of the Rainfall Stations Database

The data related to rainfall stations were acquired through the National Water Agency - ANA. Its main objective is to implement and coordinate the shared and integrated management of water resources and regulate access to water, promoting its sustainable use (ANA, 2023).

With the assistance of the ANA Data Acquisition tool, an automatic download of data from various rainfall stations was performed (Petry et al., 2021). The tool was developed as part of the Large Basin Model (LBM) plugin for the Quantum GIS software.

The period of the studied time-series was defined according to the availability and quality of the data presented by the rainfall stations. Thus, the study comprised data existing between April 2000 and December 2019.

After acquiring rainfall data, they were filtered for subsequent insertion into the statistical analyses performed. Initially, the "Operators" of the rainfall stations were filtered. These organizations/companies/agencies are responsible for the station operation, maintenance, and data collection. In this context, some of the stations within the study area were operated by the Geological Research Company (CPRM), part of the Geological Survey of Brazil (SGB), a public company linked to the Ministry of Mines and Energy, and some from National Institute of Meteorology (INMET), an agency of the Ministry of Agriculture and Livestock, which aims to add value to production in Brazil through meteorological information, were used. CPRM and INMET were used as a limit due to the reliability and availability of historical data.

Subsequently, the availability of data from 1.629 stations was verified. Thus, all stations with a percentage greater than 20% of missing data were removed from the final analyzed dataset. Then, all stations that met the previously mentioned requirements were included in the final processing stage, totaling a sample of 998 rainfall stations. Finally, the number of stations present per pixel was checked to avoid bias in the studied data. Thus, a grid was constructed according to the cell size of the lowest resolution file, which was crossed with station data using the Delete Identical command in ArcMap 10.3 software, randomly removing stations present in the same pixel, leaving only one sample per pixel. Next, the presence of outliers was checked in the pluviometric models, mainly referring to data recorded as -9999. At the end of the processing, the studied sample

included 937 rainfall stations (Figure 4.3). The information on the rainfall stations is made available in the Supplementary Material, as a table, as well as in the repository, in vector format (.shp), available at the link provided in the “Data availability statement” section of this manuscript.

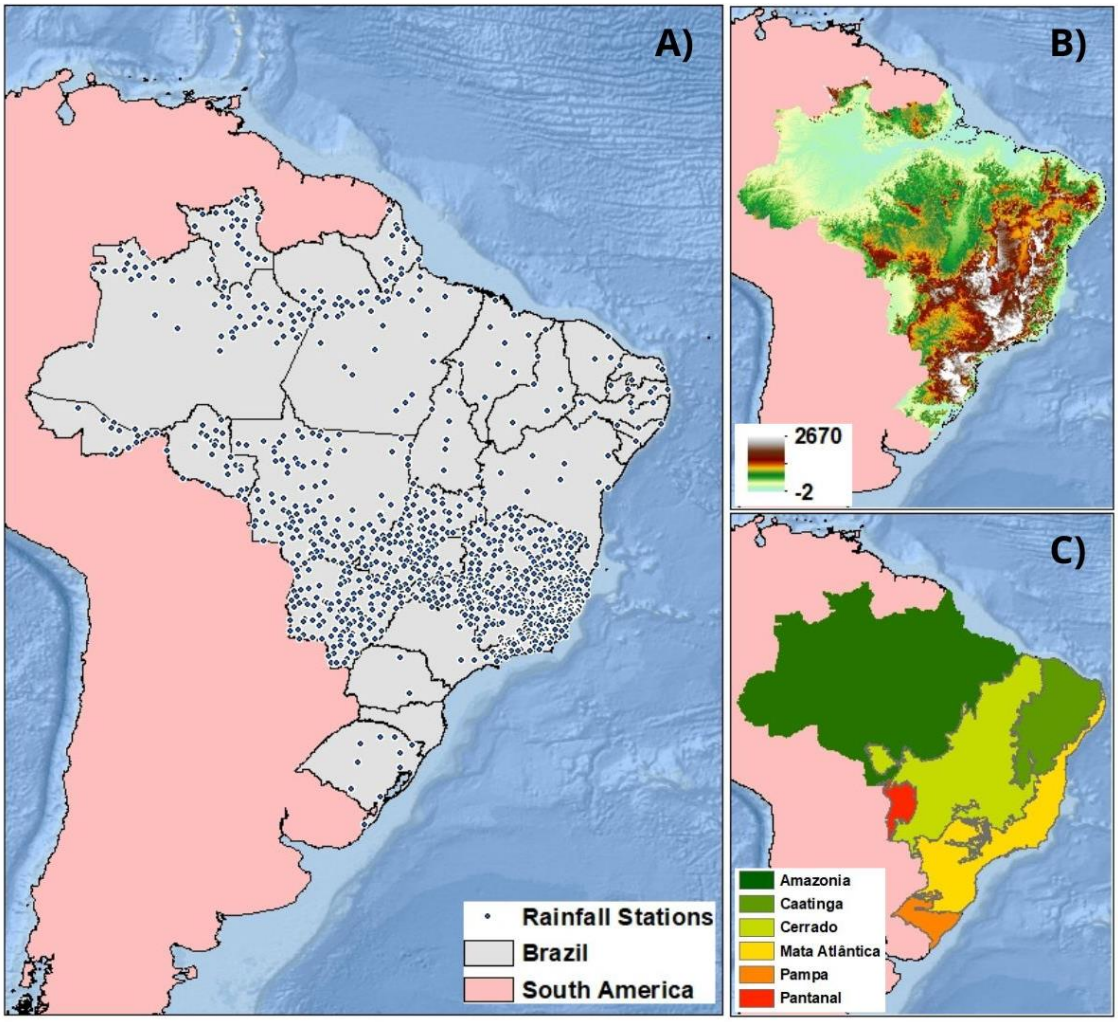

Figure 4.3. a) Location of rainfall stations; b) Elevation (m) model of the Brazilian relief; c) Biomes of Brazil.

As a way to fill in the missing data contained in the previously mentioned 20%, the station data was interpolated across the entire study area. To achieve this, the Kriging method was used, as previously done by Ananias et al. (2021) and Charles et al. (2022). Subsequently extracted using the Nearest Neighbor Search (NNS) identification technique.

The Near Neighborhood technique is a machine learning and data analysis method that relies on the proximity between data points in a dataset to make predictions or classifications (Roque et al., 2019). It is a simple and effective form of classification and

regression, especially useful when dealing with high-dimensional problems or when assumptions about data distribution are not made.

The two main approaches to building efficient NNS data structures are indexing and sketching. Indexing aims to build a data structure that, given a query point, produces a small subset of P (called the candidate set) that includes the desired neighbors (RezaAbbasifard et al., 2014). In contrast, the goal of sketching is to calculate compressed representations of points to allow for quickly calculating approximate distances. Indexing and sketching can be combined to maximize overall performance (Wang et al., 2016).

In this context, with the assistance of the Extract Values to Points tool in ArcMap 10.5 software, values were extracted based on the position of points associated with rainfall stations and stored in the attribute table of the generated layer. Thus, a database of 222,069 samples were generated.

4.2.2.2. Selection and Extraction of Gridded Data from Rainfall Models

Regarding rainfall models, as previously shown in Figure 4.1, the most widely applied models in the current literature were selected, namely: Climate Hazards Group InfraRed Precipitation with Station data – CHIRPS (Liu et al., 2019; Shen et al., 2020; GHOZAT et al., 2021; Hsu et al., 2020; López-Bermeo et al., 2022), TerraClimate (Neto et al., 2022; Araghi et al., 2023; Hanchane et al., 2023; Araghi and Adamowski, 2024), PERSIANN Climate Data Record (Santos et al., 2021; Saddique et al., 2021; Ray et al., 2022; Nadeem et al., 2022), and Global Land Data Assimilation System – GLDAS (Chen et al., 2020; Wang et al., 2021; Yan et al., 2022; Kuchak et al., 2023). Detailed information about the selected models is provided in Table 4.1.

Table 4.1. Details of the precipitation gridded datasets.

Source	Description	Spatial Resolution	References
GLDAS NOAH v2.1	Monthly data from April 2000 to December 2019 (237 months), represented by 237 images.	0.25º x 0.25º	GLDAS: Precipitation rate from satellite and ground-based observational data, using advanced land surface modeling and data assimilation techniques using infrared sensors and physiographic predictions. https://ldas.gsfc.nasa.gov/index.php/gldas
CHIRPS v2.0	Monthly data from April 2000 to December 2019 (237 months), represented by 237 images.	0.05º x 0.05º	CHIRPS: Precipitation estimates derived from rain gauges and satellite observations using field station data, infrared sensors, and physiographic predictions. https://www.chc.ucsb.edu/data/chirps/

Source	Description	Spatial Resolution	References
PERSIANN-CDR	Monthly data from April 2000 to December 2019 (237 months), represented by 237 images.	0.25º x 0.25º	PERSIANN: Precipitation estimation from remote sensing information using artificial neural networks and infrared sensors. https://chrsdata.eng.uci.edu/
TerraClimate	Monthly data from April 2000 to December 2019 (237 months), represented by 237 images.	~0.04º x 0.04º	TerraClimate: Precipitation data derived from the combination of high spatial resolution climatological normals from the WorldClim dataset, with lower spatial resolution, along with time-varying data from CRU Ts4.0 and the 55-year Japanese Reanalysis (JRA55). https://www.climatologylab.org/terraclimate.html

After selecting the models, information was extracted so that rainfall data could be individualized based on each of the 937 stations used. Thus, the point cloud related to rainfall stations intersected with the acquired images from the previously mentioned models, covering the entire study period. The data was extracted using the Nearest Neighbor Search (NNS) identification technique, as explained in section 4.4.1.

4.2.2.3. Database Standardization

The data standardization involved using ArcMap 10.5 software, where all acquired images were pre-processed so that each satellite was grouped chronologically, with a unique reference system covering the entire study area.

Regarding pixel dimensions, they were kept according to the characteristics of each satellite. Thus, the image resolution directly influenced data collection, allowing for comparison not only of the periodicity and processing method of satellite images but also of how the resolution of the images impacts the outcome of this research.

The database obtained after extracting the information was tabulated and entered into Excel 2019 software, where all mathematical and statistical analysis was conducted.

In addition, for conducting the analyses using Multiple Linear Regression, the data from the rain gauge stations and satellite models were normalized to generate the Provisional Normals, which covers a period of less than 30 but equal to or greater than 10 years database. The normalization was calculated using Equation 1 (INMET, 2022):

$$X_{ij} = \frac{\sum_k K_{ijk}}{N} \quad (1)$$

Where X_{ijk} is the observed value of variable X on day k, month i, year j, and N is the number of days in month i, year j, for which observations are available.

4.2.3. Mathematical Basis of Regression Metrics

The study employed three metrics to evaluate the analyzed datasets: the coefficient of determination, Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE).

The coefficient of determination, commonly denoted as R^2 or r^2, is the proportion of variation in a variable (dependent or response variable) explained by other variables (independent variables) in regression. It is a widely used measure to gauge the strength of the relationship between these variables in regression (KASUYA, 2018). It is a widely disseminated metric among authors and has been widely used over time. Miles (2005) showed that R-squared statistics are derived from analyses based on the general linear model (regression, ANOVA) and represent the proportion of variance between the explanatory and predictor variables in the sample. Chicco et al. (2021) suggested using R-squared statistics as a standard metric for evaluating regression analyses, which is popular in any scientific domain.

The Root Mean Square Error (RMSE) is a used metric in statistics and machine learning to measure the accuracy of a predictive model, typically a regression model. It quantifies the difference between predicted values and the actual values of a dataset (Hodson, 2022). RMSE has been applied as a statistical metric to measure the performance of models in areas such as meteorology, air quality, climate research studies, and geosciences (Chai and Draxler, 2014). This metric is commonly employed to assess the discrepancy between model-generated predictions and truly observed values in the real context. These values are evaluated in absolute units, corresponding to the same units used for reference values. Individual discrepancies are often referred to as residuals, and RMSE is used to summarize these residuals into a single measure of predictive capability. The closer the RMSE value is to 0, the better the model fits the data (Santos et al., 2021).

Finally, the Mean Absolute Percentage Error (MAPE) is a widely used error metric for assessing the accuracy of predictions in statistical analyses, particularly in time series forecasts and numerical data estimates (Kim and Kim, 2016). It quantifies the average percentage discrepancies between actual values and estimates relative to the actual

values. This metric is presented in percentage form, making it easy to interpret. A lower magnitude MAPE indicates greater accuracy of predictions, reflecting lower percentage discrepancies between estimates and actual values. This suggests that the forecasts are more aligned with actual values, which is desirable in forecasting and estimation analysis contexts (Mckenzi, 2011; Myttenaere et al., 2016).

4.2.4. Multiple Linear Regression (MLR)

The multiple linear regression (MLR) method relies on a regression equation that demonstrates the straightforward connection between dependent and independent variables. Widely regarded as a classic and popular approach, MLR serves as an effective method for addressing statistical regression issues (Huien and Peter, 1997; Malik and Kumar, 2015; Li et al., 2017; Tikhamarine et al. 2020). In multiple linear regression, a model is formulated through a linear equation. In this equation, the coefficients assigned to each independent variable measure the specific effect of that variable on the dependent one, while considering the effects of other independent variables. This technique is widely employed in various fields, such as social, economic, medical, and environmental sciences, for data analysis and prediction (Aiken et al., 2012).

The MLR equation in matrix notation is as follows (Equation 2):

$$y = X\beta + e \tag{2}$$

Where:

y: is the vector of dependent variable values to be estimated by the model.
β: is the vector of coefficients of the independent variables in matrix *X*.
e: is the vector of errors that cannot be explained by the model.

This linear system can be solved by the Least Squares Method (LSM), which in matrix notation is (Equation 3):

$$\hat{\beta} = (X^T X)^{-1}(X^T Y) \tag{3}$$

Where:

$\hat{\beta}$: is the vector of estimated coefficients for the independent variables.
X: is the matrix of independent variables.
Y: is the vector of dependent variable values.

In this study, MLR was used to associate the elevation of the terrain with models in which this variable is not present (TerraClimate and PERSIANN-CDR). Thus, it was

possible to identify how terrain features, such as the presence of mountains, valleys, among other landscapes, can modify rainfall patterns (orographic rainfall). Additionally, it allows us to observe how the presence or absence of this variabl may impact the produced model's quality.

4.2.5. Results Compilation

After obtaining the results from applying the pre-defined metrics, the final phase of the methodology involved evaluating and comparing these results for each of the analyzed models. The model that demonstrated the best performance according to the established metrics was selected, thereby identifying the most suitable climatic model for the research involving the data type under consideration.

4.3. Results

4.3.1. Statistical Analysis of Rainfall Models

Among all the evaluated models, based on the data obtained from the 937 precipitation stations used in this study, the precipitation model provided by the Climate Hazards Group InfraRed Precipitation with Station data – CHIRPS, showed the best results, as can be seen in Figures 4.4-4.6 and Table 4.2. Moreover, by applying multiple linear regression, it was possible to incorporate the effects of elevation into the models that did not originally include this information. The results obtained from the MLR are also available in Table 4.2.

Table 4.2. Statistical Analysis of the Models

Model	R^2	RMSE	MAPE
CHIRPS	0.843	42.83	0.09%
GLDAS	0.839	45.17	0.11%
PERSIANN-CDR	0.671	74.57	0.17%
TerraClimate	0.413	91.56	0.23%
PERSIANN-CDR with Elevation by MLR	0.726	29.93	9.21%
TerraClimate with Elevation by MLR	0.718	31.14	9.56%

Based on the applied statistical metrics from an isolated analysis, it can be observed that the model provided by the CHIRPS system stood out positively in all

analyses. Regarding the Coefficient of Determination R^2 (Figure 4.4), CHIRPS showed an average of 0.843, followed by the GLDAS, PERSIANN-CDR, and TerraClimate models, with values of 0.839, 0.671, and 0.413, respectively. This indicates the level of fit of the models to the data collected in situ by each station. In this context, it was found that CHIRPS has a higher explanatory power of the analyzed variable, providing a better model fit.

The Root Mean-Square Error (RMSE) (Figure 4.5) assessed the absolute error of the models. As it represents the standard deviation of the residuals (prediction errors), RMSE measured their dispersion. In other words, the analysis identified how concentrated the data were around the best-fit line. Thus, it was found that the CHIRPS model had the lowest RMSE value (42.83). Similar to the Coefficient of Determination R^2, the second-best model in RMSE evaluation was GLDAS (45.17), followed by the PERSIANN-CDR (74.57) and TerraClimate models (91.56).

Finally, based on the analysis of the Mean Absolute Percentage Error (MAPE), Figure 4.6, the average accuracy of the model in percentage terms was determined. In this metric, following the pattern of the others analyzed, the CHIRPS model exhibited higher quality, with an error of 23.10%, followed by the models provided by GLDAS (25.60%), PERSIANN-CDR (41.18%), and TerraClimate (53.41%).

The evaluation of observed data applied to the Brazilian region, specifically concerning the hydrological cycle effects of the Flying Rivers, shows that the Climate Hazards Group InfraRed Precipitation with Station data – CHIRPS stands out remarkably compared to the other evaluated models

In contrast to this highlighted performance of CHIRPS, the TerraClimate model presented the most unsatisfactory results within the study area. Additionally, the PERSIANN Climate Data Record demonstrated significant inadequacies, failing to capture relevant climate patterns.

On the other hand, it is noteworthy that GLDAS demonstrated better performance than the PERSIANN Climate Data Record and TerraClimate, approaching the values achieved by CHIRPS. However, CHIRPS remains the most appropriate and reliable choice for climatic studies in the area delimited in this research.

As for the analyses regarding the application of MLR, it was possible to verify that the PERSIANN-CDR and TerraClimate models showed considerable improvements. Especially the TerraClimate model, which started as an inadequate model for the study in question, and became a database with sufficient quality for its application. Thus, it can be observed that the region's physiography directly impacts the quality and accuracy of the produced model. According to Table 4.2, the R^2 and RMSE metrics improved; on the other hand, the MAPE values increased, demonstrating a greater amount of discrepant data in the generated model. However, even with more discrepancies present in the database, the data showed values closer to those provided by the rain gauge stations, which increased the level of correlation between the observed and predicted data.

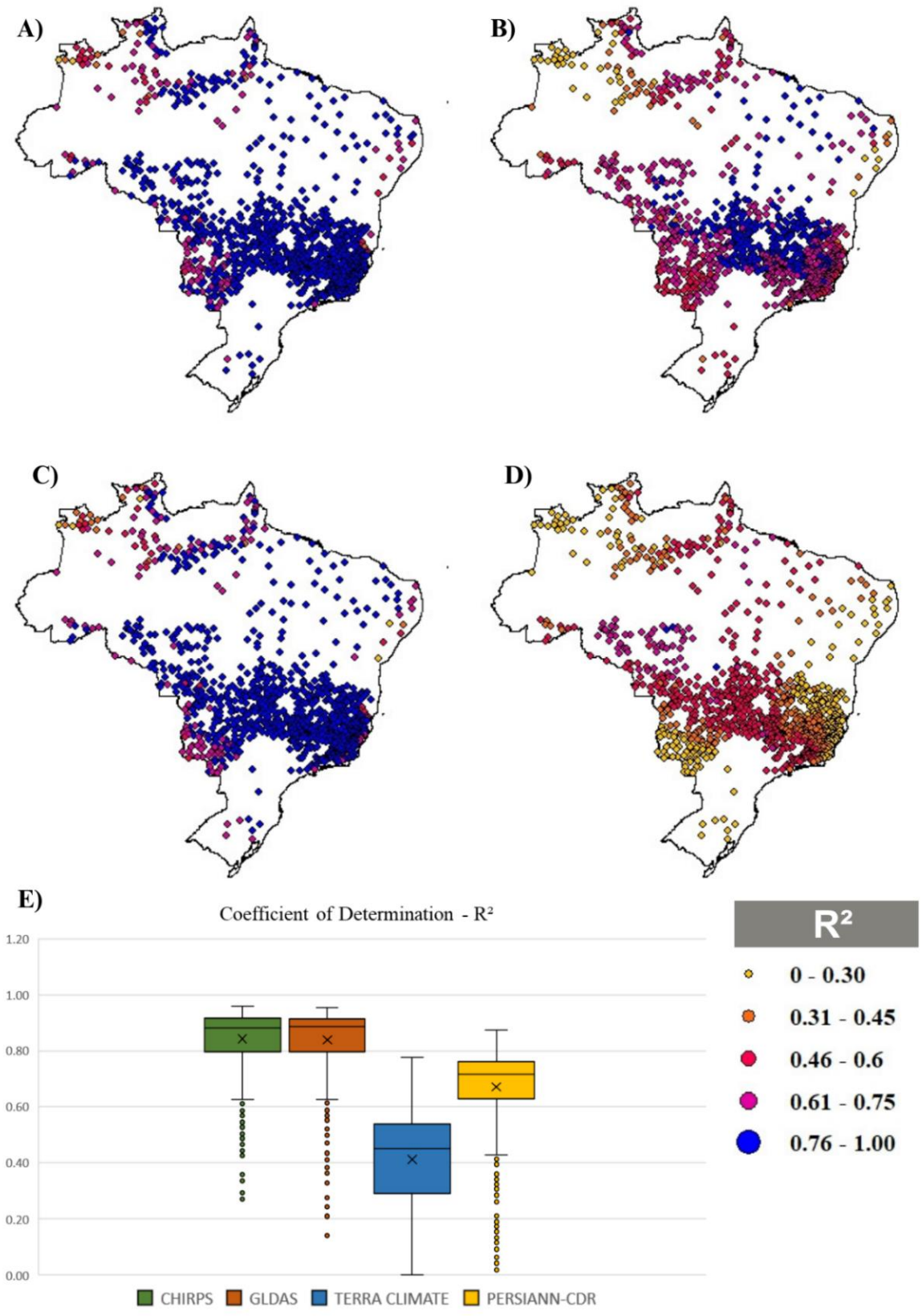

Figure 4.4. Performance of climatic models in terms of the Coefficient of Determination - R^2 metric; A) CHIRPS Model; B) PERSEIANN-CDR Model; C) GLDAS Model; D) TerraClimate Model; E) Boxplot chart of data distribution.

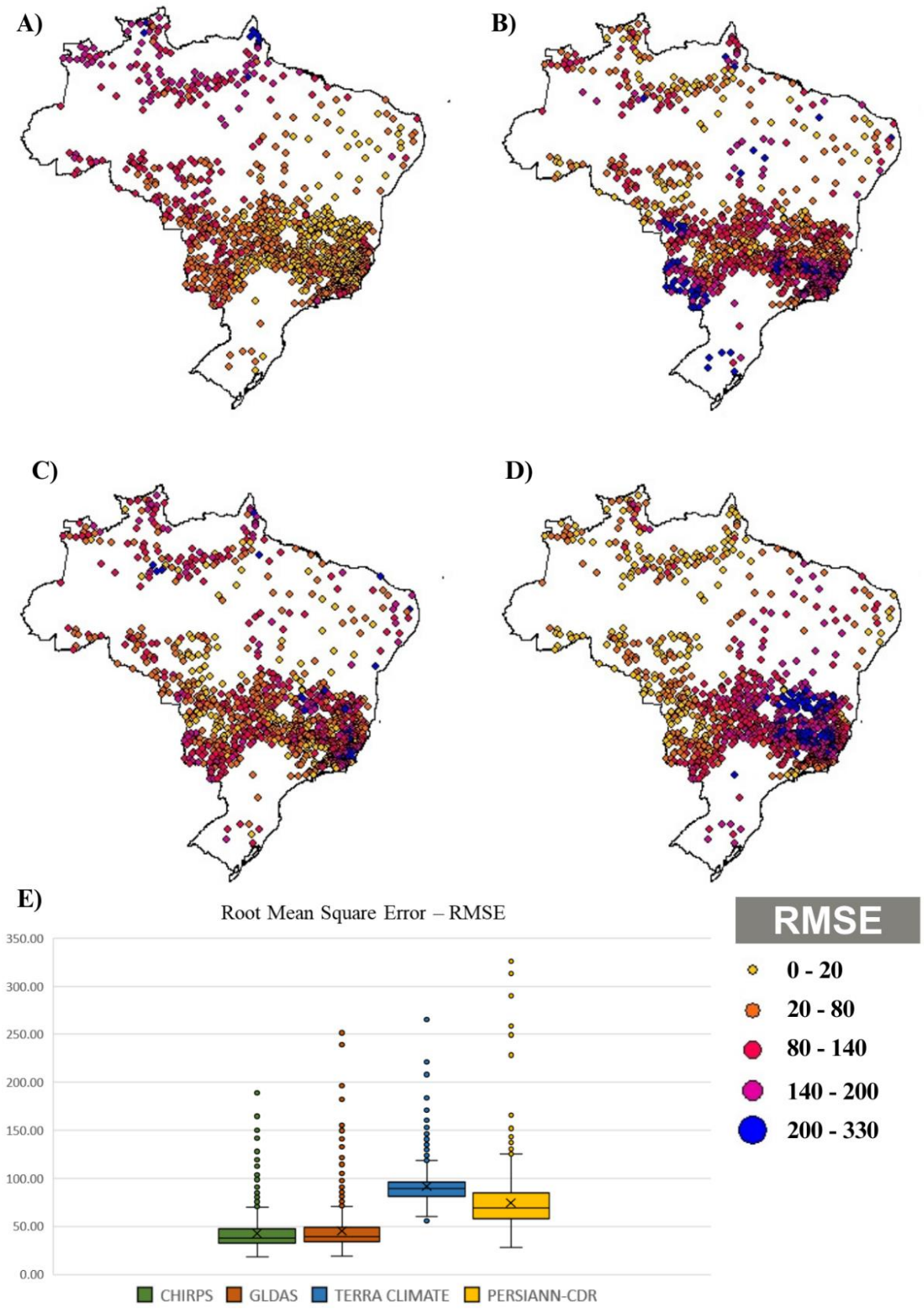

Figure 4.5. Performance of climatic models in terms of the Root Mean Square Error - RMSE metric; A) CHIRPS Model; B) PERSEIANN-CDR Model; C) GLDAS Model; D) TerraClimate Model; E) Boxplot chart of data distribution.

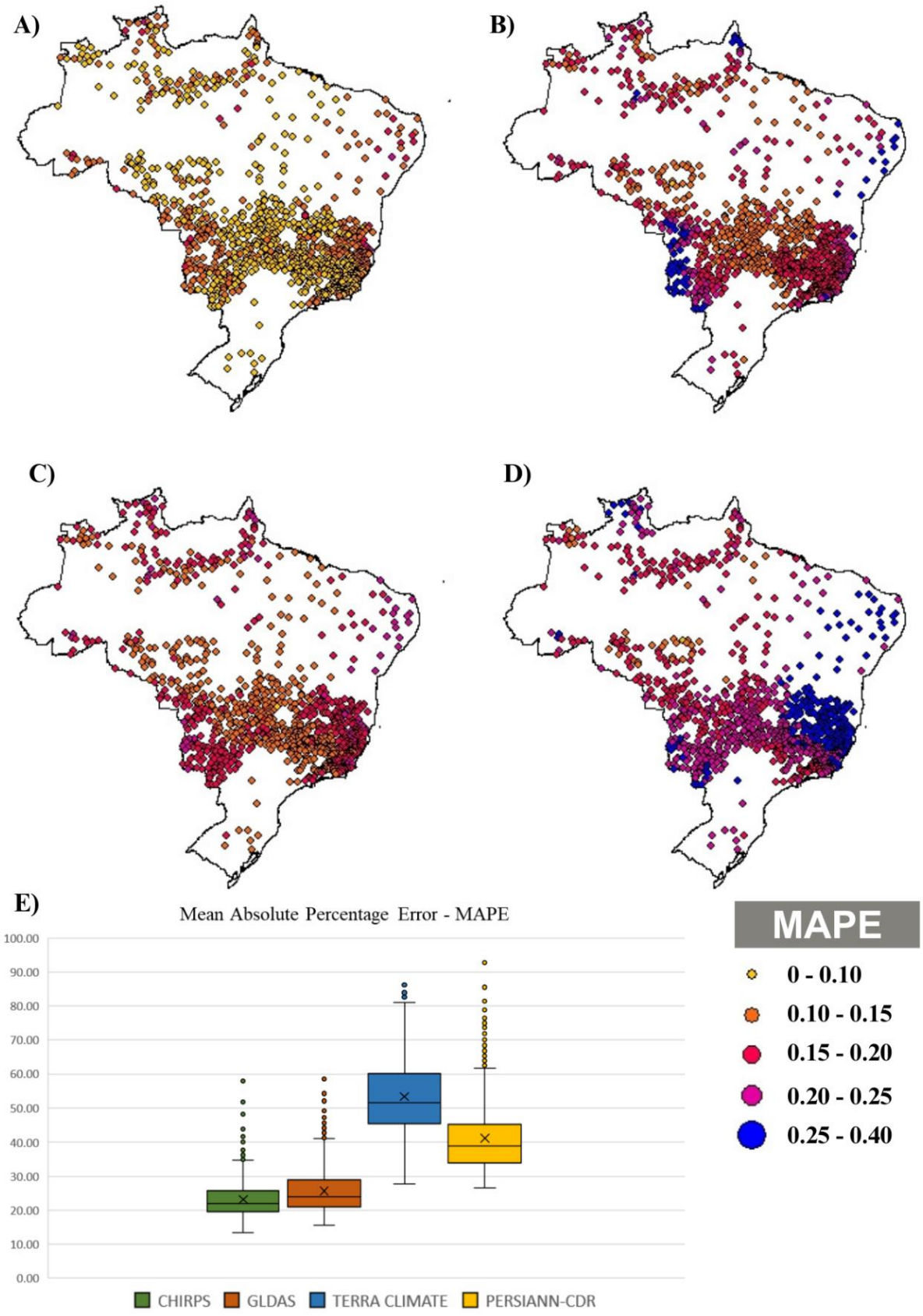

Figure 4.6. Performance of climatic models in terms of the Mean Absolute Percentage Error - MAPE metric; A) CHIRPS Model; B) PERSEIANN-CDR Model; C) GLDAS Model; D) TerraClimate Model; E) Boxplot chart of data distribution.

4.3.2. Analysis of models in the face of extreme weather events

To enhance the analysis of the conducted studies, the hydrological years of 2001-2002, 2005, 2007-2008, 2010 and 2012/2017 (Lima and Magalhães, 2018), marked by high aridity indices and the largest droughts occurrences in the database period, was examined. The main objective was to evaluate the model's predictive capacity in the face of extreme climatic events. Thus, when analyzing the metrics during the selected period, it was found that CHIRPS demonstrated the highest predictive quality, followed by the models GLDAS, PERSIANN-CDR, and TerraClimate (Table 4.3), following the same trend as the previous analysis.

Table 4.3. Analysis of models in response to the extreme events.

	Months												
Model	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Set**	**Oct**	**Nov**	**Dec**	**Coefficient of Determination - R^2**
CHIRPS	0.71	0.64	0.49	0.50	0.49	0.52	0.52	0.55	0.54	0.57	0.60	0.71	
GLDAS	0.72	0.64	0.53	0.43	0.41	0.51	0.42	0.43	0.54	0.52	0.60	0.71	
PERSIANN-CDR	0.43	0.45	0.32	0.36	0.33	0.40	0.34	0.38	0.44	0.32	0.31	0.46	
TerraClimate	0.06	0.07	0.10	0.10	0.08	0.10	0.07	0.07	0.06	0.10	0.06	0.07	

4.4. Discussion

According to the obtained results, it is possible to observe a pattern in the distribution of the evaluated metric values based on the position of the selected precipitation stations. The central-west region, the northwest of the southeast region, and the southern Bahia exhibit a more homogeneous distribution than other regions, such as the south of northern region and the coastal strip (Figures 4.4 - 4.6). These variations in distribution may be associated with various factors, ranging from physiographic issues to the production method of precipitation models.

As this research focuses on precipitation data, climatic and geographical factors significantly impact the observed data collected by precipitation stations. Orographic rains are examples of variables that can cause modifications in the precipitation regime. This type of rain occurs due to the uplift of moist air over a geographical barrier, such as

mountains. When moist air is forced to rise over a geographical elevation, it cools, condenses, and forms clouds, resulting in rainfall (Sanches et al., 2023).

The rainfall distribution in Brazil displays notable diversity influenced by the vast territorial expanse and varied geographical characteristics. The amplitude of these climatic variations is evidenced by distinct rainfall patterns in various country areas, each with its peculiarities (Fetter et al., 2018).

In the North, especially in the Amazon region, a regular and abundant rainfall distribution throughout the year is observed, vital for maintaining the tropical forest and local ecosystems (Barni et al., 2020). However, contrasting with this humidity, the Northeast faces a semiarid climate characterized by scarce and irregular rainfall, leading to frequent periods of drought, notably in the northeastern hinterlands (Rocha et al., 2021, Silva et al., 2021).

The Central-West Region, occupied by the Cerrado, experiences a seasonal rainfall pattern, with a rainy season during the summer and a dry season in the winter. These rains are of great importance for agriculture in the region, sustaining the production of essential crops (Confessor et al., 2022). In the Southeast, the presence of the Serra do Mar along the coast results in orographic rains, especially in areas of São Paulo, Rio de Janeiro, and parts of Minas Gerais, contributing to a regular rainfall regime and of great importance for water supply to large bodies like the São Francisco River and the Cantareira System (Brasiliense et al., 2020).

In regions like the Pantanal, located in the Central-West, a floodplain whose ecosystem is directly affected by seasonal rains, orographic rains are crucial for maintaining biodiversity (Maia et al., 2023). Additionally, in the Chapada Diamantina, located in Bahia, orographic rains are common, resulting from the uplift of moist air over the mountains, being a vital phenomenon for local water balance and, consequently, the region's ecological diversity (Oliveira et al., 2020).

The effects of orographic rains can be seen in various studies. For example, Aderaldo and Nery (2022) assessed the effects of relief configuration on local climate and its impacts on the desertification process in the Inhamuns nucleus, Ceará, Brazil. The authors concluded that this process in the region is amplified due to the influence of relief and climate interaction. In another study, Nascimento and Vale (2019) evaluated the

relationship between rainfall distribution and relief in an 18.8 km transect between Fundão and Santa Teresa municipalities in Espírito Santo during the hydrological year 2015/2016. From a topoclimatic perspective, the authors highlighted the study area's climatic individuality, especially regarding local orography. Thus, it was concluded that the location morphology and geographical characteristics impose an uneven distribution of precipitation.

Based on the studied period, 2000 to 2019, it is important to highlight major drought events such as the ones in 2001-2002, 2005, 2007-2008, 2010 and 2012/2017 (Lima and Magalhães, 2018). The 2015 drought in Brazil was a climatic event that affected several country regions. Several states faced intense drought conditions, causing significant impacts on agriculture, water resources, and the lives of local communities. Among the most affected regions were the Northeast and part of the Southeast (Aquino and Nascimento, 2020). In other study, Cunha et al. (2019) evaluated drought events in Brazil from 2011 to 2019. During this period, the authors verified that Brazil experienced some of the most severe droughts in decades, impacting various regions, being most severe and widespread between the years 2011 and 2017. These droughts led to water crises that affected multiple economic sectors and the population, and also increased the number of fires due to the heightened risk under dry and warm conditions.

In the Northeast, the drought was particularly severe and prolonged, leading to a water crisis affecting millions. Water reservoirs reached critical levels, hindering water supply for various uses. Lack of rainfall also affected crops, contributing to economic and social problems in the region. In the Southeast, especially in São Paulo, a lack of rainfall led to a significant water crisis. Reservoir levels, such as the Cantareira System, plummeted, affecting the water supply in the metropolitan region (Braga and Molion, 2018).

Among the factors that contributed to exacerbating the drought situation in 2015, the El Niño climate event, climatic anomalies generating irregular atmospheric patterns, and physiographic differences between the affected regions, among others, can be highlighted (Marengo et al., 2016).

Based on the concepts and studies mentioned earlier, it can be observed that geographic conditions directly influence rainfall distribution, causing significant variation in the data observed by the stations according to their location. Thus, based on Figure

4.3, it can be verified that the stations used in this study are distributed within different geographic conditions, directly influencing the identified rainfall intensity. Based on the data behavior, it can be affirmed that the metrics showed homogeneity mainly in similar physiographic regions. When present in more geographically rugged areas or with pronounced relief differences, predicted data tended to show discrepancies compared to observed data.

The effects of terrain on the formation and distribution of rainfall, as well as on the quality of the analyzed models, become even clearer when we consider the application of Multiple Linear Regression (MLR) to rainfall models that do not originally account for this variable. Table 4.2 shows that the datasets from PERSIANN-CDR and TerraClimate significantly improve when MLR is applied, making both models applicable to the study of the region in analysis.

Regarding TerraClimate, which exhibited the lowest metrics values among those analyzed, other studies have similarly classified it as having low applicability. Saicharan and Rangaswamy (2023) investigated the suitability of CHIRPS, CRU, GLDAS, GPM, PERSIANN-CDR, SM2RAIN, and TerraClimate for rainfall measurement in India. Using skill metrics (γ, RMSE, NSE, RSR, and PBIAS), they assessed the suitability of each dataset at the pixel level and concluded that TerraClimate had the lowest performance among all datasets. Arias and Barriga (2022) analyzed the accuracy of CHIRPS and TerraClimate over 36 years of rainfall data at monthly, seasonal, and annual scales in a region with complex terrain. They found that the CHIRPS dataset demonstrated the best accuracy and performance across monthly, annual, and seasonal time scales.

Based on the above, it is clear that TerraClimate has low applicability in areas with significant terrain variation. However, when adjusted, the model becomes highly applicable, which is of great importance to the scientific community, especially considering the extensive data availability that this model provides.

In addition to the effects of relief, maritime influences also play a role. For this variable, regions near oceans and seas receive more rainfall than regions located inland due to higher oceanic water evaporation (Dantas, 2009). In these areas, moist air coming from the ocean encounters features that act as a natural barrier, being forced to rise and, consequently, cool, forming loaded clouds. This process is responsible for regular rainfall

in these areas, contributing significantly to the water supply for various cities and local ecosystems (Nascimento and Vale, 2019).

In a study conducted by Salton et al. (2021), which aimed to analyze various aspects of drought periods in the state of Paraná, it was observed that the lowest aridity indices are found in the vicinity of the Serra do Mar. This phenomenon can be attributed to a series of factors influencing precipitation patterns in the region. Local topography, maritime breeze, inland waters, and thermal contrasts between the land surface and the ocean stand out among these factors. The occurrence of orographic rains in the region results from air currents moving perpendicular to the relief on a synoptic scale, especially from the east and southeast directions, and the influence of maritime breeze manifesting in the late afternoon and early evening. These atmospheric phenomena play a crucial role in mitigating the effects of drought, ensuring a regular supply of moisture to the area near the Serra do Mar.

In more recent studies, Veiga et al. (2022) assessed the influence of three El Niño events on rainfall distribution in different physical spaces of Rio de Janeiro. The authors identified that rainfall concentration is due to favorable controlling factors such as relief, maritimity, and vegetation. The authors state that regions where precipitation stations are located and which are influenced by maritimity tend to present higher average rainfall values. Terassi and Galvani (2022) analyzed predominant atmospheric characteristics. They established climate types responsible for rainfall genesis by applying the Spatial Synoptic Classification (SCC) in Castro, Curitiba, and Paranaguá, in the eastern state of Paraná, Brazil. From the study, it was possible to identify that the climate types that act in these cities are related to maritimity, in the case of Paranaguá, and associated with extratropical atmospheric systems in Castro and Curitiba.

According to the above, and observing Figures 4.4-4.6, it is noted that stations closer to the coastal region of the Southeast, Northeast, and the Northeast of the Northern region show greater heterogeneity in the mean values of R^2, RMSE, and MAPE. This effect can be attributed to the region's climatic variation caused by the presence of oceanic masses, increasing the amplitude of observed data and interfering with the adjustment of predicted data. Additionally, such variations become even more disparate due to rainfall in these regions that are not measured by satellites due to their respective temporal resolutions. Thus, once brought by the oceans and affected by local relief, air

masses precipitate, and these data are captured by precipitation stations but not by satellites, causing discrepancies in observed and predicted monthly averages.

The farther from the coast, the greater the temperature range; with lower humidity concentration in the air, there tends to be drier weather in warmer months (Santos and Moraes, 2016). In this sense, air masses formed over the oceans are characterized by high humidity, which results in different precipitation levels when combined with topography and vegetation. Rain closer to the ocean does not exhibit the same behavior observed inland the continent.

In line with the previously mentioned factors, the vegetation cover of the study area also directly impacts rainfall distribution and, consequently, data collection. Studies demonstrate that rainfall is directly altered by latitudinal variation. Generally, areas near the Equator experience higher precipitation than other areas, such as poles and temperate regions. This is due to higher evaporation provided by the increased incidence of sunlight in the equatorial region. Regardless of the season, equatorial zones receive more solar heat. This establishes the following relationship between climate and latitudes: the lower the latitude, the higher the temperatures (Pavão, 2017).

According to Khan and Ul Hasan (2017), temperature directly affects evapotranspiration, increasing its rate with rising heat. On the other hand, high evapotranspiration rates impact rainfall, elevating this variable (Hanif et al., 2013). The Equator represents the hottest zone, where temperatures reach high values. In this area, evapotranspiration tends to increase due to heat; consequently, rainfall follows the same trend (Allen et al., 2016).

In this context, it can be affirmed that the different vegetation covers present in each biome, according to their robustness, contribute differently to water recharge (Britto et al., 2019). Thus, due to the stations' location within the study area, they exhibit variations in their data due to rainfall intensity according to evapotranspiration levels. As observed, various factors can cause changes in precipitation conditions. So, the combination of these factors can directly interfere with observed data, according to their characteristics, generating errors in predicted data.

Based on the analysis of Table 4.1, it is evident that the employed models show variations in the resolutions used, giving them different levels of detailed information

according to the specific characteristics of each model. Spatial resolution in satellite images refers to the ability to distinguish minutiae on the Earth's surface. This measure is quantified in terms of pixel size. The smaller the pixel size, the higher the spatial resolution, implying the ability to discern more precise and subtle details in the captured image (Zanotta et al., 2019).

Thus, in addition to the information extraction method discussed in section 4.4.2, the size of the studied sample and the extent of the study area, it was observed that the information detailing observed data directly impacted their quality in the study in question. It was found that higher resolutions, and consequently more detail richness, showed higher levels of fit between observed and predicted data.

In the study by Costa et al. (2019), the authors aimed to analyze monthly precipitation data from the CHIRPS product and verify its similarity to meteorological station data from the Weather Forecast and Climate Studies Center (CPTEC) of the National Institute for Space Research (INPE) and the National Institute of Meteorology (INMET) for the Brazilian territory. The study found that, CHIRPS estimates, linearly adjusted with INMET/CPTEC data, showed a more pronounced agreement. In a more recent study, Santos et al. (2020) evaluated seven rain estimation methods using different approaches with the aid of statistical analyses. The focus was on the long term, using 30 years of monthly scale data to verify the method's performance in estimating the spatial-temporal behavior of rainfall compared to *in situ* measured data from meteorological stations. The authors concluded that CHIRPS showed higher quality among all the analyzed models.

Regarding the study by Santos et al. (2020), the authors addressed a smaller study area with a sample of 560 precipitation stations, concluding that CHIRPS stands out among the other models. Although extensive, the Semiarid region does not present significant physiographic and atmospheric differences compared to the area covered by the Flying Rivers. Thus, the homogeneity of the region and its dimension contributed to a smaller disparity in the database and, consequently, a greater fit between observed and predicted data. In this context, it can be verified that the dimension of the areas and the sample size did not interfere with the quality of CHIRPS. So, demonstrating that its applicability maintained a high standard of quality.

Araghi and Adamwski (2024) analyzed 30 gridded precipitation datasets in a study conducted across Iran. The authors concluded that gridded precipitation datasets based solely on climate model outputs or satellite observations were not reliable alternatives to measured precipitation. However, incorporating weather station measurements into the production of gridded precipitation datasets significantly improved the accuracy of the final product, as seen with the multi-source datasets. So, they observed that TerraClimate was suggested over CHIRPS. The difference in results between this study and the one made by Araghi and Adamwski (2024) is attributed to the fact that TerraClimate was used in its original format, without incorporating any variable that would contribute to its improvement, causing CHIRPS to be identified as the superior gridded dataset, while TerraClimate was classified as the worse model. This is due to CHIRPS model incorporates corrections based on physiographic predictions, which provide greater accuracy in estimating rainfall data in the Flying Rivers context. The impact of physiographic characteristics on model quality becomes even clearer when we compare the data presented in Table 4.2.

Another point to be emphasized is the method of extracting information from satellite images. In the work of Santos et al. (2020), the authors used the "point-pixel" extraction method, where, according to the position of the point related to the studied precipitation station, the value of the pixel to which it is related is extracted. On the other hand, Costa et al. (2019) interpolated, through the Kriging technique, station data generating raster files that were compared to the studied hydrological models. However, as a way to expand the boundaries of knowledge of the present research, the authors used the Near Neighborhood technique, where a classification of the data is carried out so that each sample from a data set, evaluating its distance concerning the nearest neighbors, receives the value mostly present in its surroundings.

Based on the obtained results, it is possible to observe that, regardless of the extraction method, CHIRPS exhibited higher quality in this and the previously mentioned studies due to the resolution and the data production process.

Based on Table 4.1, CHIRPS has higher spatial resolution, with a smaller pixel size, consequently possessing greater information detailing. In addition, this dataset uses rainfall station data and includes terrain characterization, allowing for greater accuracy in information production. Regarding GLDAS, one of the lowest resolution datasets, due to

its processing, which also includes physiography, it achieves an acceptable quality compared to PERSIANN-CDR and TerraClimate. The contribution of correction through physiography during dataset processing is clear, as GLDAS and PERSIANN-CDR have the same spatial resolution. Still, the latter does not include terrain for adjusting the collected data.

In this sense, it is clear that the model's spatial resolution, combined with the sample size, its spatial distribution, and the method of extracting the required information, directly impacts the quality of the data due to the level of action of external factors and their effects on the information collection process.

While it is acknowledged that all four evaluated models demonstrate quality and find widespread use in global studies, their applicability is contingent upon the specific characteristics of the research. Nonetheless, the investigation of Flying Rivers presents a formidable challenge and remains relatively unexplored in the literature. The vast coverage of these air masses and their complex behaviors, influenced by location-specific exogenous factors, hinder our understanding of their dynamics. Therefore, CHIRPS emerges as the most suitable model due to its information extraction method and the sample size.

4.5. Conclusion

From the analyses conducted in this study, the results indicated that among the considered models, the Climate Hazards Group InfraRed Precipitation with Station data – CHIRPS exhibits the highest quality of fit to observed data, followed by the Global Land Data Assimilation System – GLDAS, PERSIANN Climate Data Record, and TerraClimate, respectively. The same results could be seen when analyzing extreme events of drought.

The authors could conclude that, even with lower resolution, the generalization of the GLDAS model leads to few errors in producing predicted information. This occurs because this dataset's data are processed and corrected alongside local physiography and terrestrial data. Thus, despite having lower quality than CHIRPS, GLDAS would not be a database to be excluded from use in studies of this nature. Additionally, it was observed that the area extent and the sample spatial distribution also directly impact the fit between observed and predicted data, as different variables causing modifications in precipitation data become more prominent.

Finally, it is understood that all four evaluated models exhibit quality and are widely used in studies around the world, with their applicability linked to the characteristics of the specific research. Thus, the novelty of this study lies in its statistical analysis aimed at identifying the optimal climatic model for studying the extent of Flying Rivers, considering the diverse landscapes affecting their recharge. Leveraging data from 937 weather stations along their trajectory further enhances the robustness of the analysis. Furthermore, by including the relief variable in models that did not originally use it, a great improvement in the quality of these models was observed. This caused PERSIANN-CDR to achieve more adequate values in the R^2 and RMSE metrics, and TerraClimate evolved from an inadequate base to a high-quality model for the study of Flying Rivers.

The authors acknowledge that this study faces limitations due to the database used. The extensive reach of the Flying Rivers leads to a scarcity of climate study-compatible data periods, as well as implications arising from the different biomes in which the rainfall stations are located. Furthermore, the varied terrain physiography directly influences local climatic conditions, adding complexity to the interpretation process. In light of these challenges, further research applying others gridded databases and variables that significantly influences the outcomes of the employed models, as well as studies focused on the seasonal zones to better understand the impacts of each region of Brazil on rainfall formation, we recommend to broaden the scope of findings from this study.

Acknowledgments

This work was carried out with the support of CAPES - Financing Code 001.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate is a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci Data 5, 170191. https://doi.org/10.1038/sdata.2017.191

Aderaldo, P.Í.C., Nery, J.T., 2022. Atuação do relevo e do clima no núcleo de desertificação dos Inhamuns, Ceará, Brasil. Ar@cne. Revista Electrónica de Recursos en Internet sobre Geografía y Ciencias Sociales 26. https://doi.org/10.1344/ara2022.268.37173

Aiken, L. S., West, S. G., Pitts, S. C., Baraldi, A. N., & Wurpts, I. C. (2012). Multiple Linear Regression. In Handbook of Psychology, Second Edition. Wiley. https://doi.org/10.1002/9781118133880.hop202018

Allen, J.L., Chown, S.L., Janion-Scheepers, C., Clusella-Trullas, S., 2016. Interactions between temperature change and acclimation rates affect latitudinal warming tolerance patterns. Conserv Physiol 4, 1–14. https://doi.org/10.1093/conphys/cow053

ANA - Agência Nacional de Águas e Saneamento Básico. Disponível em:< https://www.gov.br/ana/pt-br>. Acesso em: 12 de novembro de 2023.

Ananias, D.R.S., Liska, G.R., Beijo, L.A. et al. The assessment of annual rainfall field by applying different interpolation methods in the state of Rio Grande do Sul, Brazil. SN Appl. Sci. 3, 687 (2021). https://doi.org/10.1007/s42452-021-04679-1

Aquino, J.R. de, Nascimento, C.A. do, 2020. a Grande Seca E As Fontes De Ocupação E Renda Das Famílias Rurais No Nordeste Do Brasil (2011-2015). Rev Econ Nordeste 51, 81–97. https://doi.org/http://orcid.org/0000-0003-0772-7141

Araghi A, Adamowski JF. 2024. Assessment of 30 gridded precipitation datasets over different climates on a country scale. Earth Science Informatics. Springer Science and Business Media Deutschland GmbH, 17(2): 1301–1313. https://doi.org/10.1007/s12145-023-01215-0.

Araghi, A., Martinez, C.J., Adamowski, J.F., 2023. Evaluation of TerraClimate gridded data across diverse climates in Iran. Earth Sci Inform 16, 1347–1358. https://doi.org/10.1007/s12145-023-00967-z

Arejano, L.M., Ruppenthal, J.G., Nobre, F.L. de L., Pazuch, F.A., Santos, R.F., Pereira, I.R., Quadro, M.S., 2023. AVALIAÇÃO DAS RESOLUÇÕES ESPACIAL E TEMPORAL DO ÍNDICE VEGETATIVO GLI ATRAVÉS DE IMAGENS DE DRONE E SATÉLITE NA CIDADE DE CATANDUVAS - PR. International Journal of Environmental Resilience Research and Science 5, 1–12. https://doi.org/10.48075/ijerrs.v5i1.30574

Arias, E. C., & Barriga, J. C. 2022. Performance of high-resolution precipitation datasets CHIRPS and TerraClimate in a Colombian high Andean Basin. Geocarto International, 37(27), 17382–17402. https://doi.org/10.1080/10106049.2022.2129816

Arraut JM, Nobre C, Barbosa HMJ, Obregon G, Marengo J. 2012. Aerial Rivers and Lakes: Looking at Large-Scale Moisture Transport and Its Relation to Amazonia and to Subtropical Rainfall in South America. Journal of Climate, 25(2): 543–556. https://doi.org/10.1175/2011JCLI4189.1.

Barni, P.E., Barbosa, R.I., Xaud, H.A.M., Xaud, M.R., Fearnside, P.M., 2020. Precipitação no extremo norte da Amazônia: distribuição espacial no estado de Roraima, Brasil. Sociedade & Natureza 32, 439–456. https://doi.org/10.14393/SN-v32-2020-52769

Braga, H.A., Molion, L.C.B., 2018. The Droughts 2013/2014 in Southeast Brazil. Anuário do Instituto de Geociências - UFRJ 41, 100–107. https://doi.org/10.11137/2018_1_100_107

Brasiliense, C.S., Dereczynski, C.P., Satyamurty, P., Chou, S.C., Calado, R.N., 2020. Air Temperature and Precipitation Climatologies over Paraíba Sul River Basin, Southeast Brazil. Anuário do Instituto de Geociências - UFRJ 43, 355–365. https://doi.org/10.11137/2020_1_355_365

Britto, M., Baptista, G.M. de M., Lima, E.A. de, 2019. O estudo dos componentes do ciclo hidrológico desde métodos tradicionais até o uso de sensoriamento remoto: uma revisão. Paranoá Cad. arquitetura e Urban. 127–146. https://doi.org/10.18830/issn.1679-0944.n23.2019.11

Cepeda Arias, E., Cañon Barriga, J., 2022. Performance of high-resolution precipitation datasets CHIRPS and TerraClimate in a Colombian high Andean Basin. Geocarto Int 37, 17382–17402. https://doi.org/10.1080/10106049.2022.2129816

Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geosci Model Dev 7, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Charles T da S, Lopes TR, Duarte SN, Nascimento JG, Ricardo H de C, Pacheco AB. 2022. Estimating average annual rainfall by ordinary kriging and TRMM precipitation products in midwestern Brazil. Journal of South American Earth Sciences, 118: 103937. https://doi.org/10.1016/j.jsames.2022.103937.

Chen, Z., Zeng, Y., Shen, G., Xiao, C., Xu, L., Chen, N., 2021. Spatiotemporal characteristics and estimates of extreme precipitation in the Yangtze River Basin using GLDAS data. International Journal of Climatology 41, E1812–E1830. https://doi.org/10.1002/joc.6813

Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput. Sci. 7, e623. https://doi.org/10.7717/peerj-cs.623

Confessor, J.G., Silva, L.L., Araújo, P.M.S., 2022. Avaliação das Perdas de Água e Solo em Pastagem Inserida em Ambiente de Cerrado Brasileiro sob Chuva Simulada. Sociedade & Natureza 34. https://doi.org/10.14393/SN-v34-2022-65618

Costa, J.C., Pereira, G., Siqueira, M.E., Cardozo, F. da S., da Silva, V.V., 2019. Validação dos dados de precipitação estimados pelo CHIRPS para o Brasil. Revista Brasileira de Climatologia 24, 228–243. https://doi.org/http://dx.doi.org/10.5380/abclima.v24i0.60237

Cunha, Ana Paula M. A., Marcelo Zeri, Karinne Deusdará Leal, Lidiane Costa, Luz Adriana Cuartas, José Antônio Marengo, Javier Tomasella, Rita Marcia Vieira, Alexandre Augusto Barbosa, Christopher Cunningham, and et al. 2019. Extreme Drought Events over Brazil from 2011 to 2019. Atmosphere 10, no. 11: 642. https://doi.org/10.3390/atmos10110642

DANTAS, E.W.Correia., 2009. Maritimidade nos trópicos: por uma geografia do litoral.

Das, K., Jiang, J., Rao, J.N.K., 2004. Mean squared error of empirical predictor. The Annals of Statistics 32, 818–840. https://doi.org/10.1214/009053604000000201

Ferrante L, Getirana A, Baccaro FB, Schöngart J, Leonel ACM, Gaiga R, Garey MV, Fearnside PM. Effects of Amazonian flying rivers on frog biodiversity and populations in the Atlantic rainforest. Conserv Biol. 2023 Jun;37(3): e14033. doi: 10.1111/cobi.14033. Epub 2023 Mar 14. PMID: 36349503.

Fetter, R., Oliveira, C.H. de, Steinke, E.T., 2018. Um Índice para Avaliação da Variabilidade Espaço-Temporal das Chuvas no Brasil. Revista Brasileira de Meteorologia 33, 225–237. https://doi.org/10.1590/0102-7786332002

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci Data 2, 150066. https://doi.org/10.1038/sdata.2015.66

Ghozat, A., Sharafati, A., Hosseini, S.A., 2021. Long-term spatiotemporal evaluation of CHIRPS satellite precipitation product over different climatic regions of Iran. Theor Appl Climatol 143, 211–225. https://doi.org/10.1007/s00704-020-03428-5

Givati, A., Gochis, D., Rummler, T., Kunstmann, H., 2016. Comparing One-Way and Two-Way Coupled Hydrometeorological Forecasting Systems for Flood Forecasting in the Mediterranean Region. Hydrology 3, 19. https://doi.org/10.3390/hydrology3020019

Guofeng, Z., Dahe, Q., Yuanfeng, L., Fenli, C., Pengfei, H., Dongdong, C., Kai, W., 2017. Accuracy of TRMM precipitation data in the southwest monsoon region of China. Theor Appl Climatol 129, 353–362. https://doi.org/10.1007/s00704-016-1791-0

Hanchane, M., Kessabi, R., Krakauer, N.Y., Sadiki, A., El Kassioui, J., Aboubi, I., 2023. Performance Evaluation of TerraClimate Monthly Rainfall Data after Bias Correction in the Fes-Meknes Region (Morocco). Climate 11, 120. https://doi.org/10.3390/cli11060120

Hanif, M., Khan, A.H., Adnan, S., 2013. Latitudinal precipitation characteristics and trends in Pakistan. J Hydrol (Amst) 492, 266–272. https://doi.org/10.1016/j.jhydrol.2013.03.040

Hodson, T.O., 2022. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci Model Dev 15, 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Hsu, J., Huang, W.-R., Liu, P.-Y., Li, X., 2021. Validation of CHIRPS Precipitation Estimates over Taiwan at Multiple Timescales. Remote Sens (Basel) 13, 254. https://doi.org/10.3390/rs13020254

Huien, H., Peter, F., 1997. Estimation of daily soil water evaporation using an artificial neural network. J. Arid Environ.

Instituto Nacional de Meteorologia (INMET). Normais Climatológicas do Brasil 1991-2020. Brasília/DF, Brasil, 2022.

Jasechko, S., Sharp, Z.D., Gibson, J.J., Birks, S.J., Yi, Y., Fawcett, P.J.,2013. Terrestrial water fluxes dominated by transpiration. Nature 496, 347–350.

Kasuya, E., 2019. On the use of r and r squared in correlation and regression. Ecol Res 34, 235–236. https://doi.org/10.1111/1440-1703.1011

Khan, S., Ul Hasan, M., 2017. Evapotranspiration Distribution and Variation of Pakistan (1931-2015). Annals of Valahia University of Targoviste, Geographical Series 17, 184–197. https://doi.org/10.1515/avutgs-2017-0017

Kidd, C., Becker, A., Huffman, G.J., Muller, C.L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D.B., 2017. So, How Much of the Earth's Surface Is Covered by Rain Gauges? Bull. Am. Meteorol. Soc. 98, 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1

Kim, S., Kim, H., 2016. A new metric of absolute percentage error for intermittent demand forecasts. Int J Forecast 32, 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003

Kuchak, V.S., Morid, S., Delavar, M., 2023. Evaluation of widespread flooding of the Karkheh Basin in Iran using SWAT model and GLDAS database. Natural Hazards 117, 2165–2185. https://doi.org/10.1007/s11069-023-05881-7

Kumar, D., Gautam, A.K., Palmate, S.S., Pandey, A., Suryavanshi, S., Rathore, N., Sharma, N., 2017. Evaluation of TRMM multi-satellite precipitation analysis (TMPA) against terrestrial measurement over a humid sub-tropical basin, India. Theor Appl Climatol 129, 783–799. https://doi.org/10.1007/s00704-016-1807-9

Kvålseth, T.O., 1985. Cautionary Note about R 2. Am. Stat. 39, 279–285. https://doi.org/10.1080/00031305.1985.10479448

Li, X., Sha, J., Wang, Z., 2017. A comparative study of multiple linear regression, artificial neural network and support vector machine for the prediction of dissolved oxygen. Hydrol. Res. 48, 1214–1225. https://doi.org/10.2166/nh.2016.149.

LIMA, J. R. de; MAGALHÃES, A. R. Secas no Nordeste: registros históricos das catástrofes econômicas e humanas do século 16 ao século 21, Parc. Estrat., v. 23 (46), p. 191-212, 2018.

Liu, J., Shangguan, D., Liu, S., Ding, Y., Wang, S., Wang, X., 2019. Evaluation and comparison of CHIRPS and MSWEP daily-precipitation products in the Qinghai-Tibet Plateau during the period of 1981–2015. Atmos Res 230, 104634. https://doi.org/10.1016/j.atmosres.2019.104634

López-Bermeo, C., Montoya, R.D., Caro-Lopera, F.J., Díaz-García, J.A., 2022. Validation of the accuracy of the CHIRPS precipitation dataset at representing climate variability in a tropical mountainous region of South America. Physics and Chemistry of the Earth, Parts A/B/C 127, 103184. https://doi.org/10.1016/j.pce.2022.103184

Maia, F.C., Almeida, T., Oliveira, P.L.G. de, Ferreira, D.V.P., Má, J.C., Cicerelli, R.E., 2023. Avaliação Temporal da Dinâmica de Regeneração da Vegetação em Áreas Queimadas no Pantanal. Rev. Bras. Cartogr. 75, 1–12. https://doi.org/10.14393/rbcv75n0a-66772

Makarieva, A.M., Gorshkov, V.G., 2007. Biotic pump of atmospheric moisture as driver of the hydrological cycle on land. Hydrol. Earth Syst. Sci. 11, 1013–1033.

Malik, A., Kumar, A., 2015. Pan Evaporation Simulation Based on Daily Meteorological Data Using Soft Computing Techniques and Multiple Linear Regression. Water Resour. Manage. 29, 1859–1872. https://doi.org/10.1007/s11269-015-0915-0.

Marengo, J.A., Cunha, A.P., Alves, L.M., 2016. A seca de 2012-15 no semiárido do Nordeste do Brasil no contexto histórico. Revista Climanálise 4, 49–54.

Marielos Peña-Claros, Carlos Nobre, A regional approach to save the Amazon.Science381,1261-1261(2023). DOI:10.1126/science.adk8794

McKenzie, J., 2011. Mean absolute percentage error and bias in economic forecasting. Econ Lett 113, 259–262. https://doi.org/10.1016/j.econlet.2011.08.010

Miles, J., 2005. R -Squared, Adjusted R -Squared, in: Encyclopedia of Statistics in Behavioral Science. Wiley. https://doi.org/10.1002/0470013192.bsa526

Monteiro, A., Campelo, M.M., 2022. O Pará no foco dos estudos sobre populações negras da Amazônia Oriental. Afros & Amazônia 1.

Moraes, E.C. De, 2002. Fundamentos de Sensoriamento Remoto, Inpe-8984-Pud/62. São José dos Campos.

Myttenaere, A., Golden, B., Le Grand, B., Rossi, F., 2016. Mean Absolute Percentage Error for regression models. Neurocomputing 192, 38–48. https://doi.org/10.1016/j.neucom.2015.12.114

Nacur, E.R., Vartuli, V., 2021. FLYING RIVERS AND CLIMATE CHANGE CAUSED BY THE DEFORESTATION OF THE AMAZON RAINFOREST: A PERSPECTIVE BASED ON LATIN AMERICAN CONSTITUTIONALISM. Revista Brasileira de Direito Animal 16, 100–115.

Nadeem, M.U., Anjum, M.N., Afzal, A., Azam, M., Hussain, F., Usman, M., Javaid, M.M., Mukhtar, M.A., Majeed, F., 2022. Assessment of Multi-Satellite Precipitation Products over the Himalayan Mountains of Pakistan, South Asia. Sustainability 14, 8490. https://doi.org/10.3390/su14148490

Nascimento, F.H., Vale, C.C., 2019. Orographic effect in a transparent between Fundão and Santa Teresa, Espírito Santo in the hydrological year 2015/2016. Geografares 29.

Neto, A.K., Ribeiro, R.B., Pruski, F.F., 2022. Assessment water balance through different sources of precipitation and actual evapotranspiration. Res Sq 1–30. https://doi.org/https://orcid.org/0000-0001-6552-7282

Newell, R.E., Newell, N.E., Zhu, Y., Scott, C., 1992. Tropospheric rivers? – A pilot study. Geophys. Res. Lett. 19, 2401–2404. https://doi.org/10.1029/92GL02916

Newell, R.E., Zhu, Y., 1994. Tropospheric rivers: A one-year record and a possible application to ice core data. Geophys. Res. Lett. 21, 113–116. https://doi.org/10.1029/93GL03113

Ngaruye, I., Rosen, D. Von, Singull, M., 2019. Mean-Squared errors of small area estimators under a multivariate linear model for repeated measures data. Commun Stat Theory Methods 48, 2060–2073. https://doi.org/10.1080/03610926.2018.1444178

Nobre, A.D., 2014. The Future Climate of Amazonia, Scientific Assessment Report. Sponsored by CCST-INPE, INPA and ARA. São José dos Campos, Brazil, 42p

Nobre, A.D., 2015. O futuro climático da Amazônia. Instituto Nacional de Pesquisa Espaciais – INPE.

Noh, Y., Liu, G., Jones, A.S., Vonder Haar, T.H., 2009. Toward snowfall retrieval over land by combining satellite and in situ measurements. Journal of Geophysical Research: Atmospheres 114, 1–15. https://doi.org/10.1029/2009JD012307

Oliveira, A.C.L. de, Tótola, L.A., Lorentz, J.F., Amaral e Silva, A., Assis, L.R. de, dos Santos, V.J., Calijuri, M.L., 2022. Spatial analysis of energy indicators and proposition of alternative generation sources for the Brazilian territory. J Clean Prod 356, 131894. https://doi.org/10.1016/j.jclepro.2022.131894

Pavão, L.L., Querino, C.A.S., Biudes, M.S., Pavão, V.M., Querino, J.K.A. da S., Machado, N.G., Beneditti, C.A., Peixoto, K.L.G., 2017. DISTRIBUIÇÃO ESPAÇO-TEMPORAL DA TEMPERATURA SUPERFICIAL URBANA NO SUL DO AMAZONAS. Raega - O Espaço Geográfico em Análise 42, 210. https://doi.org/10.5380/raega.v42i0.47440

Pearce F. 2019. Rivers in the sky. New Scientist, 244(3254): 40–43. https://doi.org/10.1016/S0262-4079(19)32070-6.

Pearce, F., 2020. Weather makers. Science 368, 1302–1305. https://doi.org/10.1126/science.368.6497.1302

Petry, I., Jardim, P., Fan, F.M., 2021. Manual de aplicação ANA Data Acquisition V 1.0. Universidade Federal do Rio Grande do Sul 12.

Quintella Veiga, R., José de Lucena, A., Silva Wanderley, H., 2021. INFLUÊNCIA DOS EVENTOS EL NIÑO NA DISTRIBUIÇÃO DAS CHUVAS NA CIDADE DO RIO DE JANEIRO. Raega - O Espaço Geográfico em Análise 53, 22. https://doi.org/10.5380/raega.v53i0.73975

Ray, R.L., Sishodia, R.P., Tefera, G.W., 2022. Evaluation of Gridded Precipitation Data for Hydrologic Modeling in North-Central Texas. Remote Sens (Basel) 14, 3860. https://doi.org/10.3390/rs14163860

RezaAbbasifard, M., Ghahremani, B., Naderi, H., 2014. A Survey on Nearest Neighbor Search Methods. Int. J. Comput. Appl. 95, 39–52. https://doi.org/10.5120/16754-7073

Rocha, R., Soares, R.R., 2015. Water scarcity and birth outcomes in the Brazilian semiarid. J Dev Econ 112, 72–91. https://doi.org/10.1016/j.jdeveco.2014.10.003

Rocha, T.B.C., Vasconcelos Júnior, F. das C., Silveira, C. da S., Martins, E.S.P.R., Gonçalves, S.T.N., Silva, E.M. da, Alves, J.M.B., Sakamoto, M.S., 2021. Indicadores de Veranicos e de Distribuição de Chuva no Ceará e os Impactos na Agricultura de Sequeiro. Revista Brasileira de Meteorologia 36, 579–589. https://doi.org/10.1590/0102-77863630041

Roque, F.V., Macarini, L.A., Crotti, Y., OliveiraWeber, T., Cechinel, C., 2019. Detecção de defeitos visuais em tecidos utilizando Wavelets e algoritmos de aprendizado de máquina, in: Anais Do Computer on the Beach. pp. 619–628.

Ruggieri, G., Allocca, V., Borfecchia, F., Cusano, D., Marsiglia, P., De Vita, P., 2021. Testing Evapotranspiration Estimates Based on MODIS Satellite Data in the Assessment of the Groundwater Recharge of Karst Aquifers in Southern Italy. Water (Basel) 13, 118. https://doi.org/10.3390/w13020118

Saddique, N., Muzammil, M., Jahangir, I., Sarwar, A., Ahmed, E., Aslam, R.A., Bernhofer, C., 2022. Hydrological evaluation of 14 satellite-based, gauge-based and reanalysis precipitation products in a data-scarce mountainous catchment. Hydrological Sciences Journal 67, 436–450. https://doi.org/10.1080/02626667.2021.2022152

Saicharan, Vasala, and Shwetha Hassan Rangaswamy. 2023. "A Comparison and Ranking Study of Monthly Average Rainfall Datasets with IMD Gridded Data in India" Sustainability 15, no. 7: 5758. https://doi.org/10.3390/su15075758

Salati, E. & Vose, P. B. (1984). Amazon Basin: A System in Equilibrium. Science, 225, 129-138.

Salati, E., Dall'Olio, A., Matsui, E., Gat, J. R. (1979). Recycling of water in the Amazon Basin: An isotopic study. Water Resources Research, 15, 1250–1258.

Salton, F.G., Morais, H., Lohmann, M., 2021. Períodos Secos no Estado do Paraná. Revista Brasileira de Meteorologia 36, 295–303. https://doi.org/10.1590/0102-77863620163

Sanches, R.G., Santos, B.C. dos, Moreira, R.M., Bourscheidt, V., Souza, P.H. de, 2023. AS CHUVAS DE RELEVO: UM LEVANTAMENTO BIBLIOMÉTRICO SOBRE O TEMA NO MUNDO E NO BRASIL. Caderno de Geografia 33, 1148. https://doi.org/10.5752/P.2318-2962.2023v33n75p1148

Santos, D., Moraes, S., 2016. Variação Da Temperatura Do Ar Média, Mínima E Máxima No Perfil Topoclimático Da Trilha Caminhos Do Mar (Sp). Revista Equador 7, 01–09.

Santos, S.M. dos, de Farias, M.M.M.W.E.C., 2017. Potential for rainwater harvesting in a dry climate: Assessments in a semiarid region in northeast Brazil. J Clean Prod 164, 1007–1015. https://doi.org/10.1016/j.jclepro.2017.06.251

Santos, V.J., Calijuri, M.L., Ribeiro Júnior, J.I., de Assis, L.C., 2021. Rainfall estimation methods in the Brazilian semiarid region: 30-year evaluation on a monthly scale. International Journal of Climatology 41, E752–E767. https://doi.org/10.1002/joc.6723

Satyamurty P, da Costa CPW, Manzi AO. 2013. Moisture source for the Amazon Basin: a study of contrasting years. Theoretical and Applied Climatology. Springer-Verlag Wien, 111(1–2): 195–209. https://doi.org/10.1007/s00704-012-0637-7.

Sheil, D., Murdiyarso, D., 2009. How forests attract rain: an examination of a new hypothesis. Bioscience 59, 341–347.

Shen, Z., Yong, B., Gourley, J.J., Qi, W., Lu, D., Liu, J., Ren, L., Hong, Y., Zhang, J., 2020. Recent global performance of the Climate Hazards group Infrared Precipitation (CHIRP) with Stations (CHIRPS). J Hydrol (Amst) 591, 125284. https://doi.org/10.1016/j.jhydrol.2020.125284

Silva, T.A., Ferreira, J., Calijuri, M.L., dos Santos, V.J., Alves, S. do C., Castro, J. de S., 2021. Efficiency of technologies to live with drought in agricultural development in Brazil's semi-arid regions. J. Arid Environ. 192, 104538. https://doi.org/10.1016/j.jaridenv.2021.104538

Smith, E.A., Asrar, G., Furuhama, Y., Ginati, A., Mugnai, A., Nakamura, K., Adler, R.F., Chou, M.-D., Desbois, M., Durning, J.F., Entin, J.K., Einaudi, F., Ferraro, R.R., Guzzi,

R., Houser, P.R., Hwang, P.H., Iguchi, T., Joe, P., Kakar, R., Kaye, J.A., Kojima, M., Kummerow, C., Kuo, K.-S., Lettenmaier, D.P., Levizzani, V., Lu, N., Mehta, A. V., Morales, C., Morel, P., Nakazawa, T., Neeck, S.P., Okamoto, K., Oki, R., Raju, G., Shepherd, J.M., Simpson, J., Sohn, B.-J., Stocker, E.F., Tao, W.-K., Testud, J., Tripoli, G.J., Wood, E.F., Yang, S., Zhang, W., 2007. International Global Precipitation Measurement (GPM) Program and Mission: An Overview, in: Measuring Precipitation From Space. Springer Netherlands, Dordrecht, pp. 611–653. https://doi.org/10.1007/978-1-4020-5835-6_48

Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H. V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN System Satellite–Based Estimates of Tropical Rainfall. Bull Am Meteorol Soc 81, 2035–2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2

Terassi, P.M. de B., Galvani, E., 2022. Análise da gênese pluvial a partir da aplicação da técnica Spatial Synoptic Classification na região leste do estado do Paraná. Geography Department University of Sao Paulo 42, e185541. https://doi.org/10.11606/eISSN.2236-2878.rdg.2022.185541

Tikhamarine, Y., Souag-Gamane, D., Ahmed, A. N., Sammen, S. Sh., Kisi, O., Huang, Y. F., & El-Shafie, A. (2020). Rainfall-runoff modelling using improved machine learning methods: Harris hawks optimizer vs. particle swarm optimization. Journal of Hydrology, 589, 125133. https://doi.org/10.1016/j.jhydrol.2020.125133

Urzagasti, C.A., Maciel, J.N., Wentz, V.H., Ledesma, J.J.G., Ando Junior, O.H., 2020. Captura Automatizada de Informações Meteorológicas e Imagens de Satélite para a Predição de Geração de Energia Solar Fotovoltaica, in: Anais Do XVII Congresso Latino-Americano de Software Livre e Tecnologias Abertas (Latinoware 2020). Sociedade Brasileira de Computação - SBC, pp. 160–163. https://doi.org/10.5753/latinoware.2020.18625

Wang, D., Zhang, M., Fu, M., Cai, Z., Li, Z., Han, H., Cui, Y., Luo, B., 2016. Nonlinearity Mitigation Using a Machine Learning Detector Based on k-Nearest Neighbors. IEEE Photonics Technol. Lett. 28, 2102–2105. https://doi.org/10.1109/LPT.2016.2555857

Wang, W., Lin, H., Chen, N., Chen, Z., 2021. Evaluation of multi-source precipitation products over the Yangtze River Basin. Atmos Res 249, 105287. https://doi.org/10.1016/j.atmosres.2020.105287

Weng W, Luedeke MKB, Zemp DC, Lakes T, Kropp JP. 2018. Aerial and surface rivers: downwind impacts on water availability from land use changes in Amazonia. Hydrology and Earth System Sciences, 22(1): 911–927. https://doi.org/10.5194/hess-22-911-2018.

Yan, X., Zhang, B., Yao, Y., Yin, J., Wang, H., Ran, Q., 2022. Jointly using the GLDAS 2.2 model and GRACE to study the severe Yangtze flooding of 2020. J Hydrol (Amst) 610, 127927. https://doi.org/10.1016/j.jhydrol.2022.127927

Zanotta, D.C., Ferreira, M.P., Zortea, M., 2019. Processamento de satélite de imagens. Oficina de Textos, São Paulo.

Zhou, H., Ning, S., Li, D., Pan, X., Li, Q., Zhao, M., Tang, X., 2023. Assessing the Applicability of Three Precipitation Products, IMERG, GSMaP, and ERA5, in China

over the Last Two Decades. Remote Sens (Basel) 15, 4154. https://doi.org/10.3390/rs15174154

Zhu, Y., Newell, R.E., 1998. A Proposed Algorithm for Moisture Fluxes from Atmospheric Rivers. Mon. Weather Rev. 126, 725–735. https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2

5. CAPÍTULO II. Multivariate Statistical Analysis of Rainfall Variability in Brazil: Assessing Climatic and Environmental Drivers of Precipitation

Abstract: This study combines Principal Component Analysis (PCA) and Cluster Analysis (CLARA) to analyze rainfall dynamics and its interactions with environmental and climatic variables. Understanding these relationships is crucial for monitoring hydrological patterns, predicting weather variability, and assessing climate change impacts. Rainfall distribution is shaped by atmospheric moisture transport, land surface characteristics, and broader climate systems, highlighting the complexity of precipitation patterns at regional and global scales. PCA was applied to reduce dimensionality and identify dominant climatic variables influencing precipitation patterns, including temperature, evapotranspiration, NDVI, and relief. The results indicate that rainfall dominates seasonal variability, explaining 43.93% of the variance in January and 41.97% in February, highlighting its role in wet-season dynamics. NDVI and evapotranspiration account for 26–30% of the variance in Dim.2, highlighting their influence on vegetation and moisture recycling. During the dry season, temperature emerges as a primary controlling factor, peaking at 34.13% variance in March and becoming inversely correlated with precipitation. The CLARA cluster analysis classified 921 rainfall stations into distinct seasonal groupings, revealing dynamic monthly transitions. Group 1, characterized by low/moderate rainfall and NDVI, remains stable throughout the year, whereas Group 2 dominates wet periods, exhibiting high NDVI (>0.7) and evapotranspiration (>0.4 mm/day). Group 3 represents arid zones with low precipitation (30–75 mm) and high temperatures (>28°C), particularly prominent in July and August. The findings highlight the influence of topography, with orographic effects enhancing precipitation in elevated areas. This study also emphasizes the impact of Amazon deforestation, disrupting moisture transport and reducing precipitation in southeastern Brazil. The weakening of the Flying Rivers threatens regional water security, reinforcing the need for sustainable land-use policies. By integrating multivariate statistical methods, this research provides quantitative insights into rainfall variability, aiding climate resilience strategies for agriculture and water management in South America.

Keywords: Principal Component Analysis (PCA), Cluster Analysis, Flying Rivers, Climate Change, Environmental Variables, Precipitation Patterns.

5.1. Introduction

Understanding the dynamics of rainfall and its interaction with environmental and climatic variables is essential for monitoring hydrological patterns, predicting weather variability, and assessing the impacts of climate change. Rainfall distribution is influenced by multiple factors, including atmospheric moisture transport, land surface characteristics, and regional and global climate systems.

In South America, the Flying Rivers play a crucial role in shaping precipitation patterns, acting as major conduits for atmospheric moisture transported from the Amazon Basin to other regions, such as the Brazilian Southeast (Pearce, 2019; Monteiro & Campelo, 2022). This phenomenon begins with the evaporation of Atlantic Ocean waters, which, upon reaching the Amazon, is augmented by evapotranspiration from the dense rainforest canopy. This continuous recycling of moisture generates humid air masses that are transported westward by trade winds and then redirected southward by the Andes Mountains (Nobre et al., 2016).

These atmospheric rivers contribute significantly to the annual rainfall received in Brazil's agricultural heartlands and major urban centers, such as São Paulo and Rio de Janeiro. However, deforestation and climate change have disrupted this delicate balance, leading to reduced rainfall, prolonged droughts, and increased climatic variability (Weng et al., 2018; Nacur & Vartuli, 2021). The weakening of the Amazon's hydrological cycle has been directly linked to declining precipitation in southeastern Brazil, highlighting the urgency of understanding the spatial and temporal patterns of rainfall formation (Zemp et al., 2017).

To understand these spatial and temporal rainfall patterns, Cluster Analysis has emerged as a powerful tool in climate and environmental studies, allowing the classification of rainfall stations based on precipitation trends and their interaction with environmental factors (Gupta et al., 2021). By grouping similar meteorological stations, cluster analysis enables the identification of underlying patterns, seasonality, and regional differences in precipitation regimes. Among the various clustering algorithms available, CLARA (Clustering Large Applications) has been widely recognized for its efficiency in handling large datasets. CLARA is an extension of the Partitioning Around Medoids (PAM) algorithm and is specifically designed to improve computational efficiency in large-scale clustering tasks (Kaufman and Rousseeuw, 1991). Unlike traditional clustering methods,

which may struggle with high-dimensional data, CLARA selects representative subsets of the dataset, applies PAM clustering, and then generalizes the results to the entire dataset. This approach ensures robust clustering performance while maintaining computational feasibility, making CLARA particularly useful for environmental and meteorological applications (Schubert & Rousseeaw, 2019; Harb et al., 2024).

Following Cluster, the Principal Component Analysis (PCA) has become a fundamental tool in meteorology and environmental sciences. PCA is a widely used technique for reducing the dimensionality of large datasets while retaining the most significant variance in the data (Jollife and Cadima, 2016). By transforming correlated variables into a smaller set of uncorrelated principal components (PCs), PCA helps identify dominant climatic and environmental factors that influence rainfall variability (Lever et al., 2017). This method is particularly valuable in climate research, where multiple variables, such as temperature, evapotranspiration, vegetation indices (NDVI), and relief, interact in complex ways to shape precipitation patterns. PCA enables researchers to detect relationships among climatic drivers, assess seasonal fluctuations, and improve clustering accuracy by reducing noise and redundancy in the dataset (Altman and Krzywinski, 2015).

By integrating PCA with CLARA, this study aims to provide a comprehensive understanding of the spatial and temporal variability of rainfall in relation to environmental and climatic drivers, thereby improving climate monitoring efforts. PCA enables the extraction of key components from large climate datasets, which are then used to optimize clustering accuracy. This integration ensures that the most relevant variables contribute to the clustering process, reducing dimensionality while preserving essential information about rainfall patterns.

So, this research conducted two independent cluster analyses to examine different aspects of rainfall variability in Brazil. The first analysis focused exclusively on rainfall data, aiming to classify precipitation stations based on seasonal behavior and their alignment with Flying Rivers pathways. The second analysis incorporated additional environmental and climatic variables, including temperature, evapotranspiration, NDVI, and relief, to explore their influence on rainfall formation and distribution. To improve the accuracy and interpretability of the results, Principal Component Analysis (PCA) was applied. This step reduced data dimensionality while preserving the most significant

variance, ensuring that only the most influential variables were used, optimizing the identification of distinct rainfall regimes and their environmental drivers.

By adopting a multivariate statistical approach, this study provides new insights into the complex interactions between precipitation, temperature, vegetation dynamics, and water balance processes across Brazil. These findings contribute to climate monitoring, water resource management, and sustainable land-use planning, offering valuable data-driven strategies to mitigate the impacts of climate change on regional hydrological cycles.

5.2. Methodology

The methodology for developing the research in question was divided into four stages: study area definition, database collection, data pre-processing and data base processing (Cluster and PCA Analysis) (Figure 5.1).

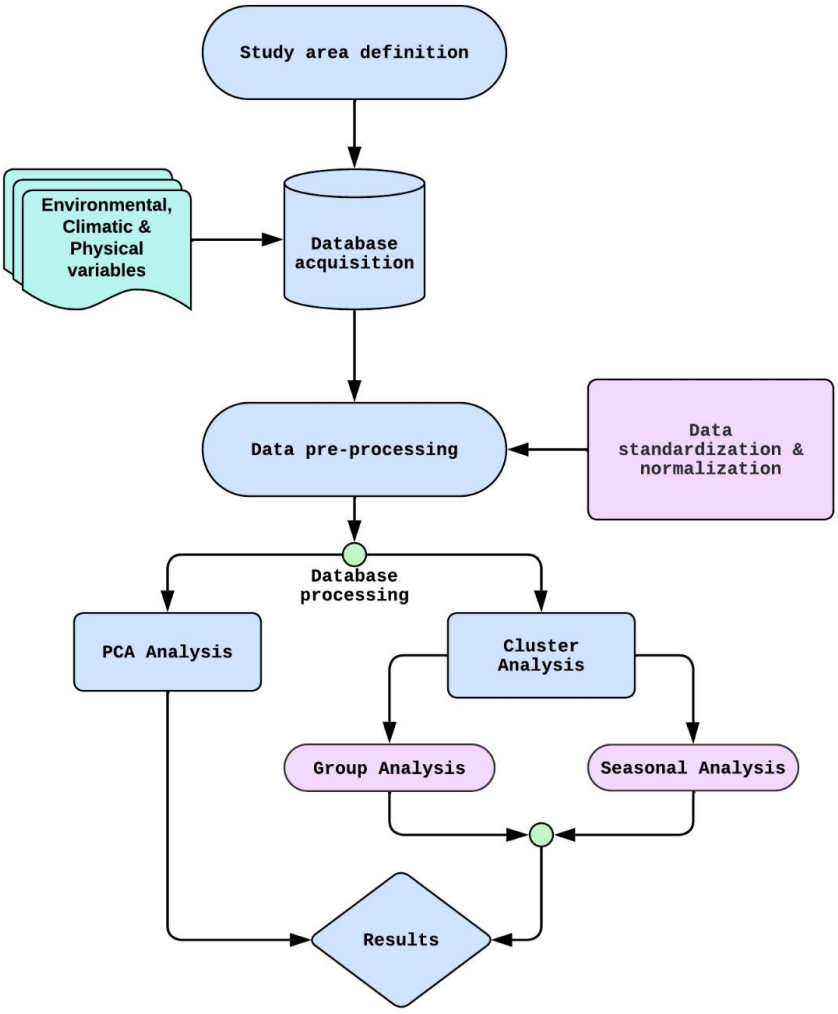

Figure 5.1. Methodological flowchart.

5.2.1. Study Area Definition

The study area covers the Brazilian territory, focusing on the regions monitored by the rainfall stations selected for this research (Figure 5.2). It is worth noting that, in order to more accurately analyze the interconnection of rainfall events between the Amazon region and the Southeast, mediated by the Flying Rivers, the main area of interest in this study is properly highlighted in Figure 5.2B.

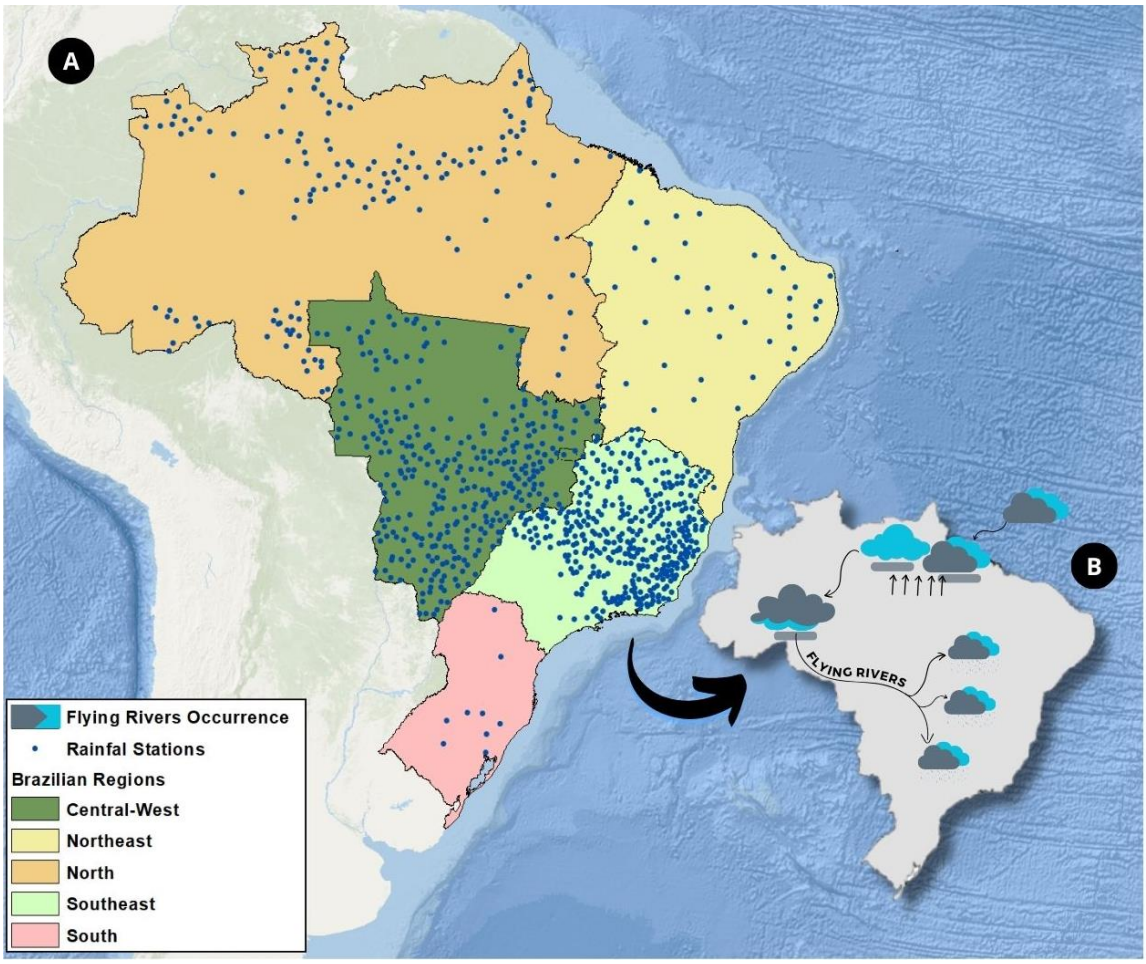

Figure 5.2. A) Rain gauge stations distributed in the study area; B) Focal study area.

The focal study area encompasses the region influenced by the Flying Rivers, atmospheric moisture transport systems spanning Brazil. These "Rivers in the Sky" extend from the northern Amazon to the southernmost regions, crossing the central-western, southeastern, and northeastern parts of the country (Pearce, 2019). These systems connect diverse biomes, facilitating interactions between various vegetation types, from dense forests to wetlands, and traversing multiple land uses (Weng et al., 2018). Their trajectory includes ecosystems ranging from the vast Amazon rainforest to

extensive agricultural and pasturelands, particularly in the central-western region, and urban centers of economic significance, such as São Paulo (Weng et al., 2018).

Flying Rivers originate from the evaporation of Atlantic Ocean waters, transported into the Amazon by trade winds, resulting in the high humidity and substantial rainfall characteristic of the forest (Nacur and Vartuli, 2021). These air masses influence the Amazon and Pantanal biomes, impacting the Cerrado and sections of the Atlantic Forest (Monteiro and Campelo, 2022).

Additionally, Flying Rivers contribute to climatic patterns beyond South America. Some of the transported moisture reaches North America via the jet stream, affecting precipitation and snowfall in the western United States and Canada (Amaral e Silva et al., 2024).

The integrity of the Amazon Rainforest is crucial for sustaining the volume of Flying Rivers. Deforestation and climate change disrupt this system, altering rainfall patterns across multiple regions. Consequently, preserving these atmospheric rivers is vital for regional ecosystems, global climate stability, and water security.

5.2.2. Data acquisition

The detailed database applied in this research can be found in Table 5.1.

Table 5.1. Details of the gridded datasets.

Database	Resolution	Description	Source
Evapotranspiration	0.25º x 0.25º	Monthly data from April 2000 to December 2019 (237 months), represented by 237 images.	Global Land Data Assimilation System - GLDAS: Evapotranspiration rate that expresses the amount of water lost from a unit area of surface per unit time. http://ldas.gsfc.nasa.gov/faq/#NLDAS_evap
Temperature			Global Land Data Assimilation System - GLDAS: Average temperature of the Earth's surface. The actual temperature of the surface, and it can be considerably different than the air temperature above the surface, particularly in warm, sunny conditions. https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH025_M_2.1/summary?keywords=GLDAS
Rainfall	0.05º x 0.05º	Monthly data from April 2000 to December 2019 (237 months), represented by 237 images.	CHIRPS: Precipitation estimates derived from rain gauges and satellite observations using field station data, infrared sensors, and physiographic predictions.

Database	Resolution	Description	Source
			https://www.chc.ucsb.edu/data/chirps/
Normalized Difference Vegetation Index (NDVI)			MOD13C2 Version 6: The Normalized Difference Vegetation Index (NDVI) is referred to as the global monitoring of vegetation conditions and are used in products displaying land cover and land cover changes. https://doi.org/10.5067/MODIS/MOD13C2.006
Relief	30 x 30 m	High-resolution global topographic database, from February 2011, created by the Space Shuttle Endeavour mission.	Shuttle Radar Topography Mission - SRTM: digital elevation dataset obtained through radar interferometry conducted by NASA and the National Geospatial-Intelligence Agency (NGA) in collaboration with the German and Italian space agencies. https://www.dsr.inpe.br/topodata/acesso.php

Regarding rainfall data, it was obtained from the Climate Hazards Center through the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), one of the most widely applied models in climate studies (Liu et al., 2019; Shen et al., 2020; Ghozat et al., 2021; Hsu et al., 2021; López-Bermeo et al., 2022). Furthermore, a study by Amaral e Silva et al. (2024) highlighted that CHIRPS stands out among freely available precipitation datasets due to its robust preprocessing methodology.

Temperature and evapotranspiration data were obtained from the GLDAS. Developed by NASA's Goddard Space Flight Center (GSFC) in collaboration with the National Oceanic and Atmospheric Administration (NOAA), GLDAS is a land surface modeling system designed to assimilate satellite and ground-based observational data into land surface models (LSMs). This system enables the generation of global-scale simulations of land surface conditions with high spatial and temporal resolution in near real-time.

The NDVI and relief data were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Shuttle Radar Topography Mission (SRTM), respectively, both made available by the National Aeronautics and Space Administration (NASA). The MODIS vegetation index, produced at 16-day intervals and at multiple spatial resolutions, enables consistent spatial and temporal monitoring of vegetation canopy greenness. This index reflects a combination of leaf area, chlorophyll content, and canopy structure, making it a valuable tool for analyzing vegetation dynamics (Didan, 2021). The SRTM mission, conducted in February 2000 aboard the Space Shuttle

Endeavour, aimed to generate a high-resolution global digital elevation model (DEM). Using radar interferometry, the mission mapped over 80% of Earth's land surface, covering latitudes from 60°N to 56°S, with unprecedented accuracy in topographic representation.

Finally, rainfall station data were obtained from the National Water Agency (ANA), whose primary mission is to implement and coordinate the integrated and shared management of water resources, regulate access to water, and promote its sustainable use (ANA, 2023). To facilitate data collection, the ANA Data Acquisition tool was used to automatically download precipitation records from multiple rainfall stations (Petry et al., 2021). This tool was developed as part of the Large Basin Model (LBM) plugin for Quantum GIS (QGIS), enhancing the efficiency of hydrological data retrieval and analysis. So, 921 stations were used.

5.2.3. Data preprocessing

The data standardization process was carried out using ArcMap 10.5, where all acquired images were preprocessed to ensure that each satellite dataset was chronologically organized and aligned within a single geographical reference system covering the entire study area.

Pixel dimensions were preserved according to the specifications of each satellite, ensuring that image resolution remained consistent with the original data source. This approach allowed for a comparative analysis of satellite image periodicity, processing methods, and the impact of resolution on data collection and research outcomes.

The Nearest Neighbor Search (NNS) technique was employed to extract information from satellite data. NNS is a widely used machine learning and data analysis method that identifies patterns based on the proximity of data points, such as rainfall stations, within a dataset. This technique is particularly effective for classification and regression tasks, especially in high-dimensional spaces or when no prior assumptions about data distribution are made (Roque et al., 2019).

Efficient NNS data structures are typically constructed using two primary approaches: indexing and sketching. Indexing involves creating a structured dataset that, for any given query point, generates a small candidate set P containing potential nearest neighbors (RezaAbbasifard et al., 2014). Conversely, sketching focuses on computing

compressed representations of data points, allowing for rapid approximation of distances. These two methods can be integrated to optimize computational performance (Wang et al., 2016).

In this study, the Extract Values to Points tool in ArcMap 10.5 was utilized to extract values based on the spatial positions of rainfall stations. The extracted data was subsequently stored in the attribute table of the generated layer, facilitating further analysis.

Regarding data normalization, it used the *data.Normalization* function, part of the *clusterSim* package in R, rescaling the values to a range from 0 to 1. This function offers various methods to normalize data, which is essential for preparing datasets for analysis. The normalization adjusts the values of numeric variables to a common scale, without distorting differences in the ranges of values (Walesiak, 2016). This process ensures that each variable contributes equally to the analysis and prevents variables with larger ranges from dominating the results. This research used the *Log Transformation* method, reducing the impact of outliers.

Furthermore, to facilitate the analysis, the Provisional Normals for each variable was calculated. This process applies to datasets spanning at least 10 years but fewer than 30 years. The process was performed using Equation 1 (INMET, 2022):

$$X_{ij} = \frac{\sum_k K_{ijk}}{N} \quad (1)$$

Where X_{ijk} is the observed value of variable X on day k, month i, year j, and N is the number of days in month i, year j, for which observations are available.

5.2.4. Database processing

To achieve the study's objectives, data processing was carried out using Principal Component Analysis (PCA) and Cluster Analysis.

5.2.4.1. Principal Components Analysis (PCA)

According to Lever et al. (2017), Principal Component Analysis (PCA) reduces data by projecting it into lower dimensions, called principal components (PCs), to provide the best summary using a limited number of components. The first PC is selected to minimize the distance between the data and its projection, maximizing variance. The second PC is chosen similarly but must be uncorrelated with the first. This process

maximizes the correlation between the data and its projection and is akin to performing linear regression (Altman and Krzywinski, 2015).

However, PCA has limitations, such as assuming linear data structures, which may overlook highly correlated patterns, as all PCs are uncorrelated (Altman and Krzywinski, 2015). The scale of variables can also affect results, making it crucial to adjust the scale based on prior knowledge. When applied correctly, PCA is a powerful tool for identifying key variables and outliers in data (Lever et al., 2017).

Before conducting PCA, several R packages are essential for analysis and visualization. The *ggplot2* package is used to create various data visualizations, including scatter plots and bar charts. The *factoextra* package offers specialized functions for visualizing PCA results, such as scree plots, biplots, and variable contribution plots. Additionally, the *dplyr* package can be utilized for data manipulation, though its use is optional.

To execute PCA, these packages were installed and loaded in RStudio, providing the necessary tools for both computing PCA and generating insightful visual representations of the results. The *prcomp()* function in R is the most commonly used method for performing PCA, as it computes principal components while automatically centering and scaling the data when specified. The output provides key information, including the standard deviation of each principal component, the proportion of variance explained by each component, and the cumulative variance retained (Narvaez-Montoya et al., 2023). In addition to the coefficients, the original variables are linked to the principal components (PCs) through loadings, which range from -1 to 1 and represent each variable's contribution to a given component. Loadings close to -1 or 1 indicate a strong negative or positive influence, respectively, while values near 0 suggest minimal impact (Jollife and Cadima, 2016).

Determining the optimal number of principal components to retain requires analyzing the cumulative variance explained. A general guideline is to keep components that collectively account for at least 70-80% of the total variance (Jolliffe & Cadima, 2016). By examining the cumulative variance, users can decide whether to retain only the first few principal components or include additional ones based on their contribution to variance explanation. In this study, the analysis focused on the first four principal components derived from the PCA processing.

5.2.4.2. Cluster Analysis

Cluster analysis is a widely used technique in data science and statistics to group similar observations based on certain characteristics. Among the various clustering algorithms available, CLARA (Clustering Large Applications) is particularly effective for handling large datasets (Kaufman and Rousseeuw, 1991). It is an extension of the Partitioning Around Medoids (PAM) algorithm and is designed to improve computational efficiency when clustering large datasets in RStudio (Kaufman and Rousseeuw, 1987).

CLARA works by selecting a sample of the data, applying the PAM algorithm, and using the medoids (central points of clusters) to classify the rest of the dataset (Schubert and Rousseeuw, 2019). The key idea behind this method is that instead of clustering the entire dataset at once, it randomly selects a subset (sample) of observations, applies clustering on the sample, and then assigns the remaining data points based on their similarity to the obtained clusters (Harb et al., 2024). By repeating this process multiple times with different samples, CLARA improves clustering quality while maintaining computational efficiency.

To perform CLARA clustering in R, several packages are required to provide the necessary functions for implementation and visualization. Each package plays a specific role in executing and interpreting the clustering process. The *cluster* package is the primary package for CLARA and other clustering algorithms. It contains the *clara()* function, which executes the CLARA algorithm and provides clustering results based on medoids rather than centroids (Kassambara, 2017).

When performing clustering analysis in RStudio, two key challenges arise: determining whether the dataset has a natural cluster structure and identifying the optimal number of clusters. To address these issues, the Hopkins statistic and *fviz_nbclust* function from the *factoextra* package are commonly used.

The Hopkins statistic measures the clustering tendency of a dataset by comparing the distribution of actual data points with randomly generated points (Wright, 2022). The value of the statistic ranges between 0 and 1: values close to 1 indicate that the data has a strong cluster structure, values close to 0 suggest that the data is randomly distributed, making clustering less meaningful, while values around 0.5 indicate no clear clustering tendency (Hopkins and Skellam, 1954). The dataset used in this research obtained a Hopkins statistic value of 0.802, indicating a strong clustering tendency.

Concerning to the optimal number of clusters. The *fviz_nbclust* function assists in this process by offering different statistical methods for estimating the ideal number of clusters. Among the most commonly used methods is the Elbow Method, Silhouette Method and the Gap Statistic, the last one compares the performance of the clustering model with a randomly generated dataset, helping to determine the optimal number of clusters (Amato et al., 2019). This study applied the Gap Statistic method to grouping database containing climatic (Rainfall, Temperature and Evapotranspiration) and biogeophysical (NDVI and Relief) variables, identifying an optimal clustering solution with six groups (Figure 5.3).

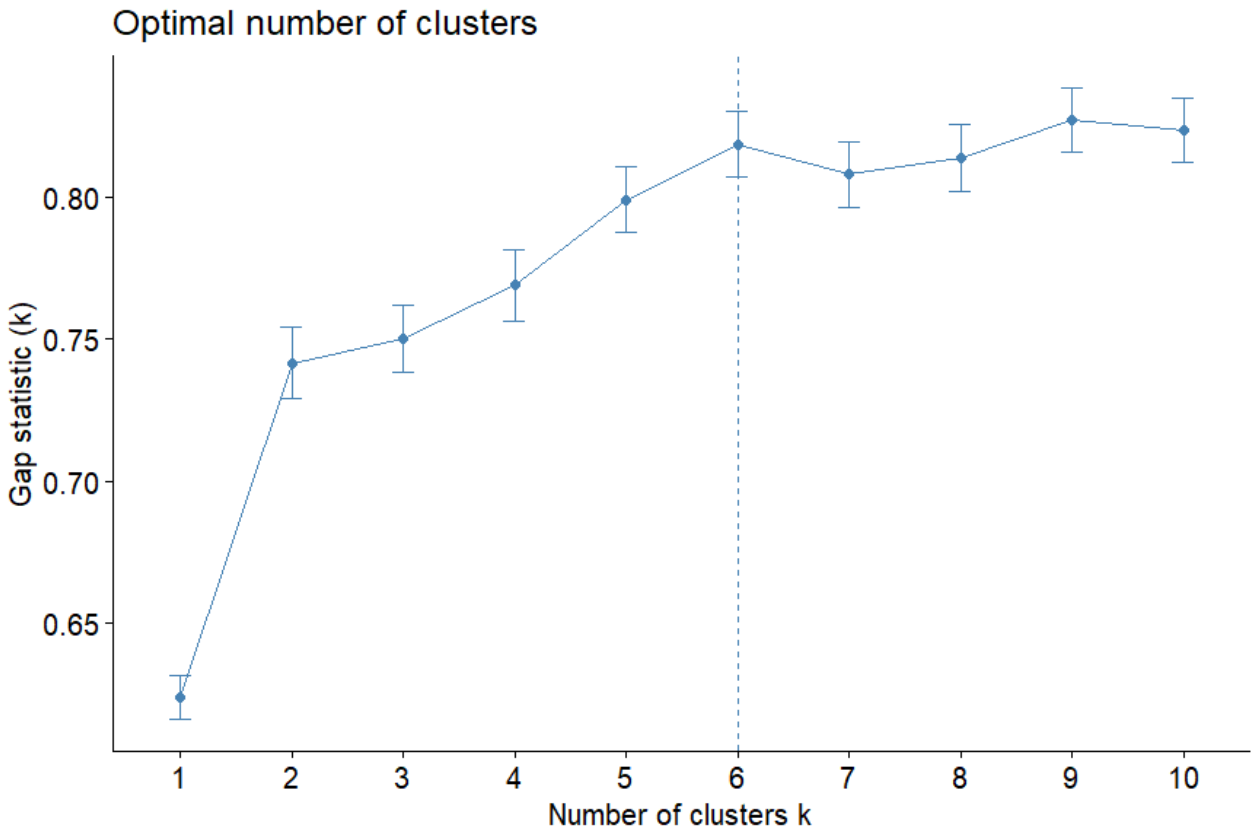

Figure 5.3. Optimal number of clusters.

This study included two cluster analyses. The first considered only rainfall data, aiming to identify the seasonal behavior of the rainfall stations, with a focus on the path covered by Flying Rivers. The second analysis incorporated environmental, climatic, and physical variables to understand the interaction between these factors and rainfall formation.

5.3. Results

5.3.1. The interrelationship between climatic and environmental variables and their impact on climate behavior

The Principal Component Analysis (PCA) illustrated in the figures elucidates the relationships among the variables Rainfall, Temperature, Evapotranspiration, and NDVI throughout the year. Each principal component captures a proportion of the variance in the dataset, with the contributions of each variable varying across months. This analysis reveals seasonal patterns and interactions among climatic and environmental factors. Figure 5.4 presents the PCA analysis for the months of January to June, whereas Figure 5.5 illustrates the analysis for the months of July to December.

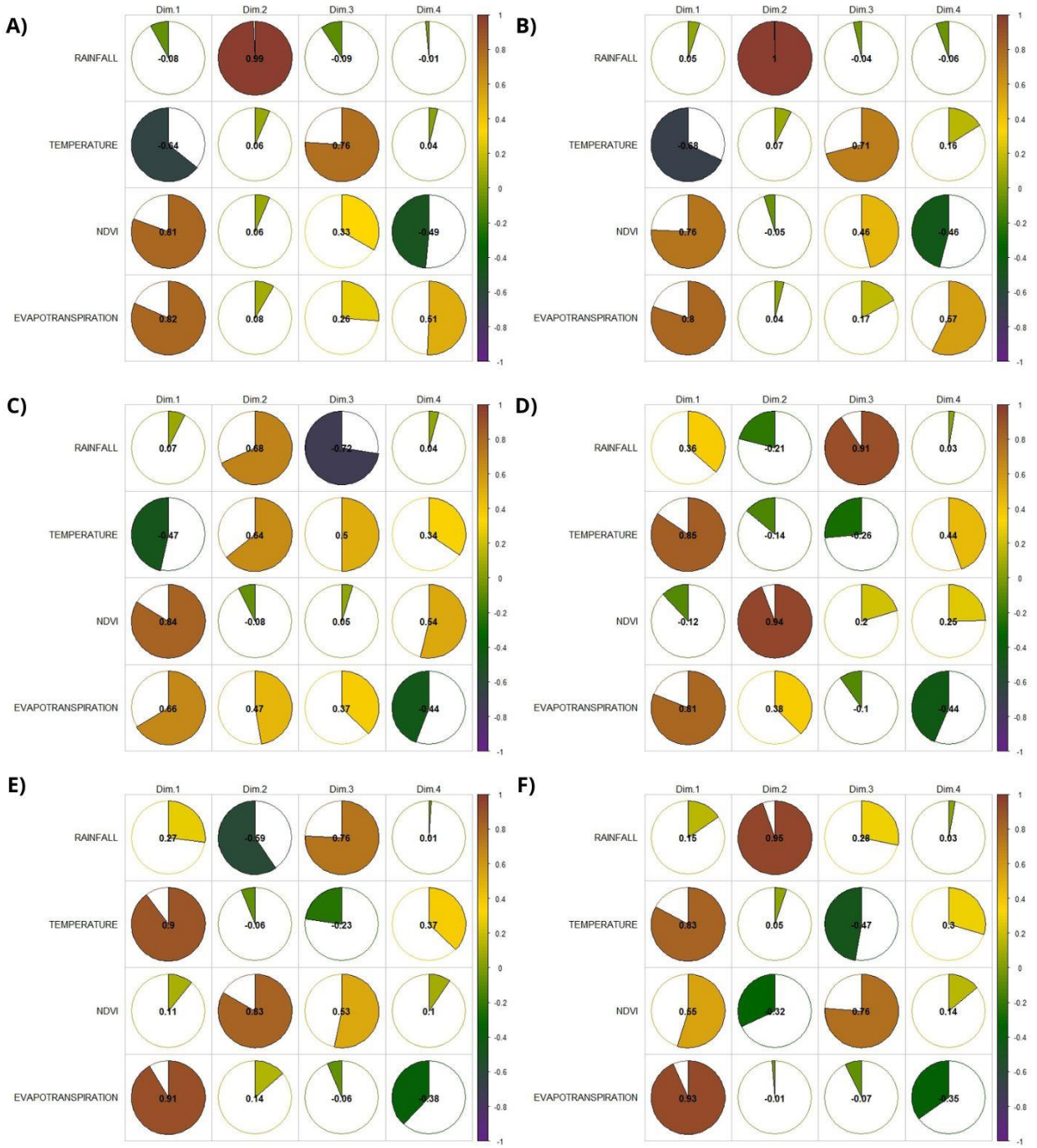

Figure 5.4. Principal Component Analysis illustrating variable contributions by month: A) January; B) February; C) March; D) April; E) May; and F) June.

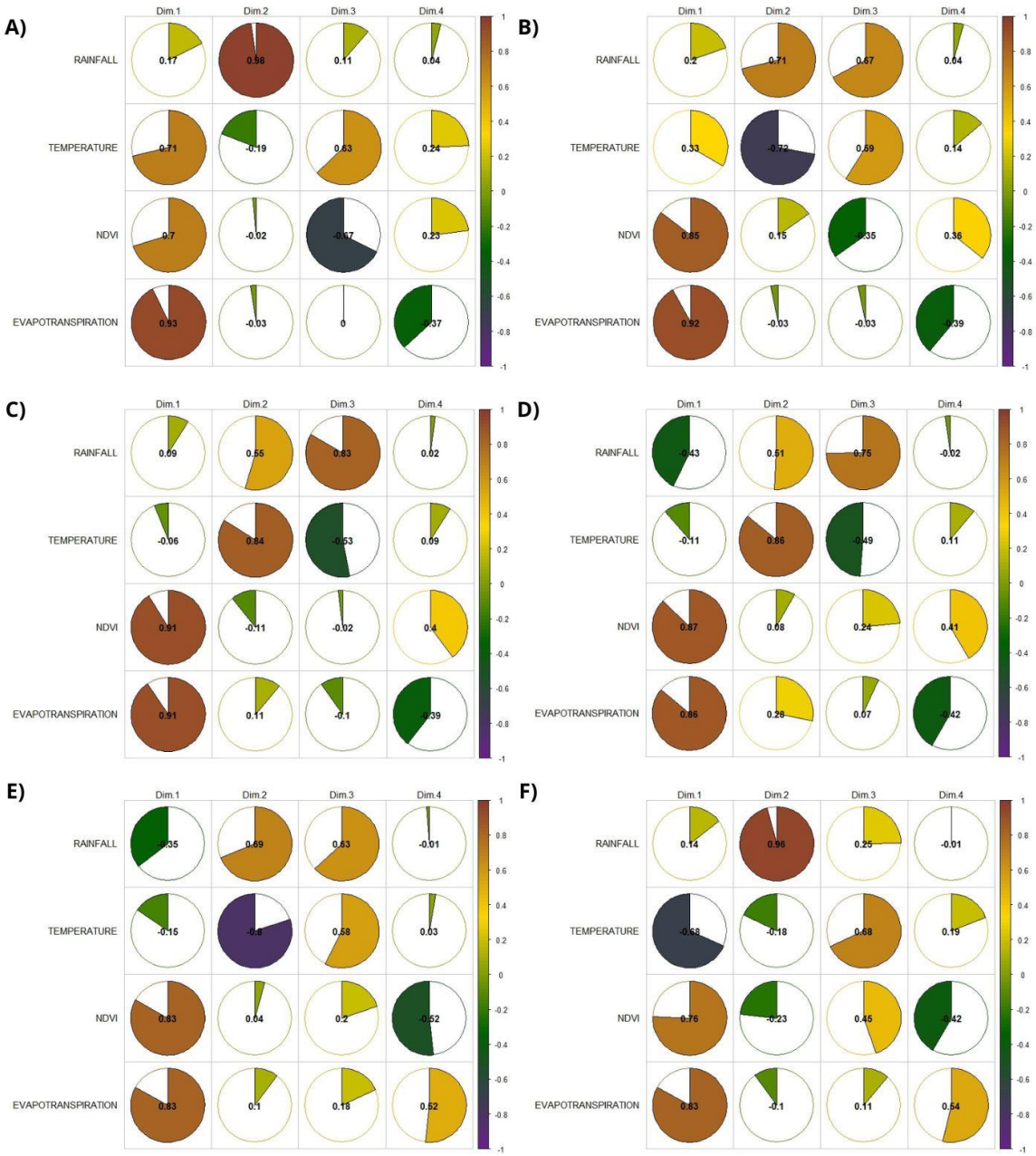

Figure 5.5. Principal Component Analysis illustrating variable contributions by month: A) July; B) August; C) September; D) October; E) November; and F) December.

The Principal Component Analysis (PCA) revealed key seasonal interactions among vegetation activity (NDVI), evapotranspiration, temperature, and rainfall, which collectively shape atmospheric moisture flux and the formation of Flying Rivers—airborne moisture flows that redistribute water across vast distances. These moisture transport systems are primarily sustained by evapotranspiration from forested regions, which recharge the atmosphere with water vapor and influence precipitation patterns.

During January and February (Figures 5.4A and 5.4B), NDVI (0.81, 0.76) and Evapotranspiration (0.82, 0.80) dominate Dim.1, highlighting strong vegetation-driven

moisture flux. Temperature (-0.64, -0.68) negatively correlates, indicating that higher temperatures reduce soil moisture retention and vegetation cover. Meanwhile, Rainfall (0.99, 1.00) emerges as an independent climatic driver in Dim.2, showing its seasonal variability. This period marks the early accumulation of atmospheric moisture, as forests and other vegetated landscapes release large amounts of water vapor into the lower atmosphere, fueling the initial phase of Flying River development.

As the wet season progresses into March and April (Figures 5.4C and 5.4D), NDVI (0.84, -0.12) and Evapotranspiration (0.66, 0.81) remain dominant on Dim.1, while Temperature (-0.47, 0.85) transitions from a weak negative to a strong positive influence. This shift suggests that rising temperatures now enhance evapotranspiration, increasing atmospheric water vapor content. Rainfall (0.68, 0.91) in Dim.2 gains importance, reflecting stronger seasonal interactions between precipitation and warming. These conditions indicate a peak phase of moisture cycling, where water vapor is continuously replenished and transported by prevailing winds, intensifying the long-distance movement of the air masses.

During May and June (Figures 5.4E and 5.4F), the peak evapotranspiration period is evident as Temperature (0.90, 0.83) and Evapotranspiration (0.91, 0.93) dominate Dim.1, reinforcing warming-driven moisture release. NDVI (0.11, 0.76) remains secondary, suggesting a stabilization of vegetation activity. Meanwhile, Rainfall (0.76, 0.95) continues as an independent driver in Dim.2, showing that precipitation responds more to seasonal climatic conditions than to local vegetation cycles during this time of the year. This period marks the maximum recharge of Flying Rivers, where extensive evapotranspiration ensures a steady flow of moisture into the atmosphere. The transported water vapor sustains precipitation in remote regions, including agricultural areas and semi-arid zones that rely on these atmospheric moisture pathways.

As the dry season approaches in July and August (Figures 5.5A and 5.5B), Evapotranspiration (0.93, 0.92) remains the strongest factor in Dim.1, with NDVI (0.70, 0.85) and Temperature (0.71, 0.33) indicating a sustained but slightly reduced moisture contribution. Rainfall (0.98, 0.71) in Dim.2 continues its role as an independent driver, while Dim.3, for July, captures an inverse relationship between Temperature (0.63) and NDVI (-0.67), suggesting that warming begins to stress vegetation, gradually reducing its water cycling efficiency. During this period, Flying Rivers reach their farthest extent,

delivering precipitation to areas far from their moisture source, maintaining critical hydrological balance in downwind regions.

Regarding September and October (Figures 5.5C and 5.5D), NDVI (0.91, 0.87) and Evapotranspiration (0.91, 0.86) remain dominant in Dim.1, indicating that residual moisture cycling persists despite the approaching dry conditions. However, a notable shift occurs in October, as Rainfall (-0.43) becomes negatively correlated with Dim.1, marking a decoupling of precipitation from vegetation-driven moisture flux. Temperature (0.84, 0.86) gains dominance in Dim.2, reinforcing its role in drying out landscapes. Dim.3 captures Rainfall (0.75) and Temperature (-0.49), illustrating the seasonal redistribution of atmospheric moisture. This shift signals the weakening of the Flying River system, as reduced evapotranspiration leads to decreased atmospheric moisture availability.

By November and December (Figures 5.5E and 5.5F), the effects of declining evapotranspiration and increasing temperature-induced drying become more evident. NDVI (0.83, 0.76) and Evapotranspiration (0.83, 0.83) continue to define Dim.1, but Temperature (-0.68) in December confirms its inverse relationship with vegetation indices. Rainfall (0.69, 0.96) remains independent in Dim.2, indicating that precipitation is now less reliant on vegetation-driven moisture input. This period marks the final stage of the Flying River cycle, where water vapor transport weakens, and precipitation patterns become more irregular. By the end of the year, atmospheric moisture flux stabilizes at lower levels, waiting for the next cycle of vegetation-driven recharge.

To further explore these interrelationships, we incorporated relief into the PCA as a physical environmental variable (Figures 5.6 and 5.7). This enhancement enabled a comprehensive evaluation of the extent to which other variables contribute to rainfall formation and their proportional relationships across months. The inclusion of relief is particularly relevant due to its influence on orographic rain, where moist air is forced to rise over elevated terrain, leading to cooling, condensation, and increased precipitation on windward slopes, while often causing drier conditions on leeward sides.

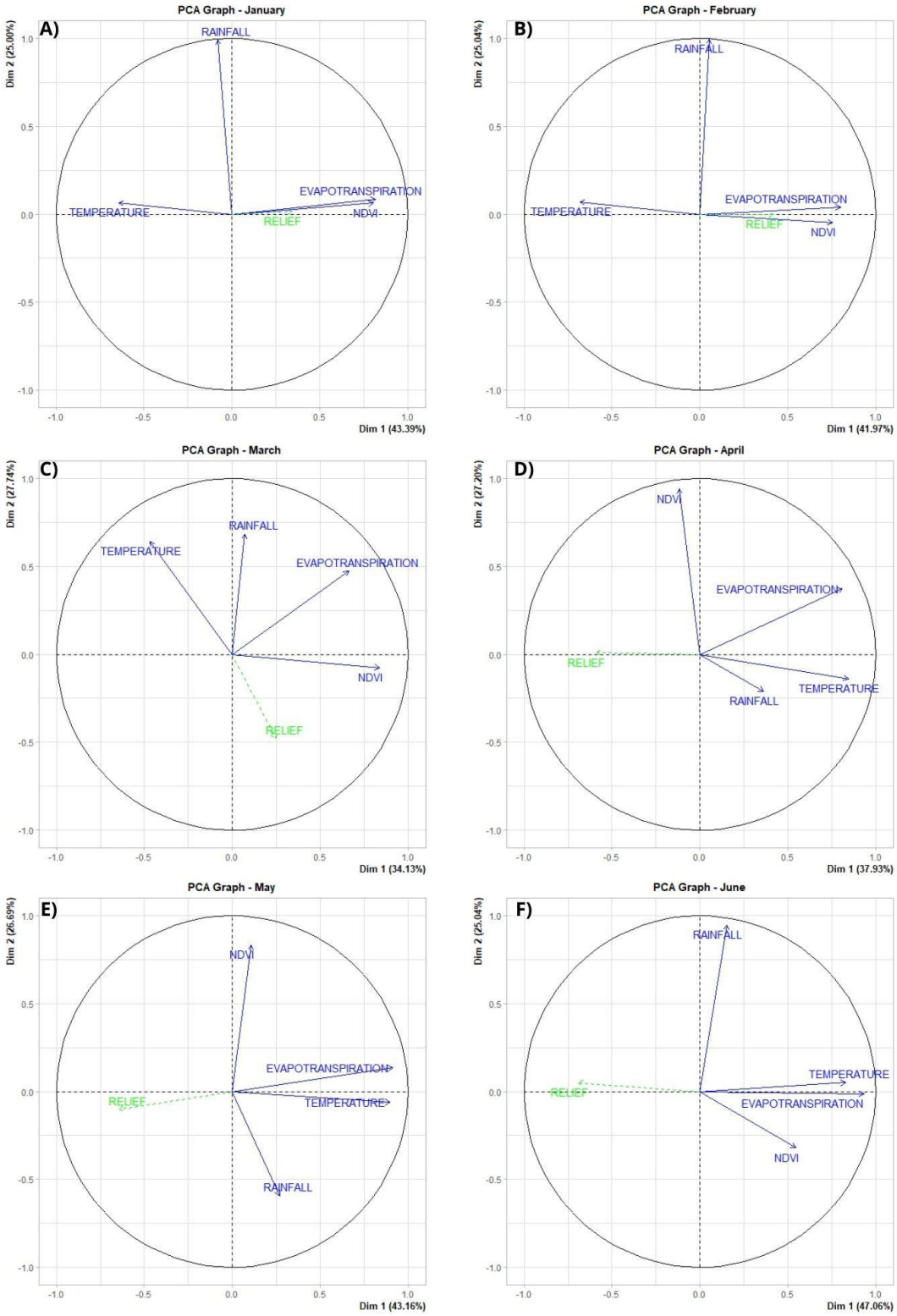

Figure 5.6. Correlation of variables through Principal Component Analysis (PCA) by month: A) January; B) February; C) March; D) April; E) May; and F) June.

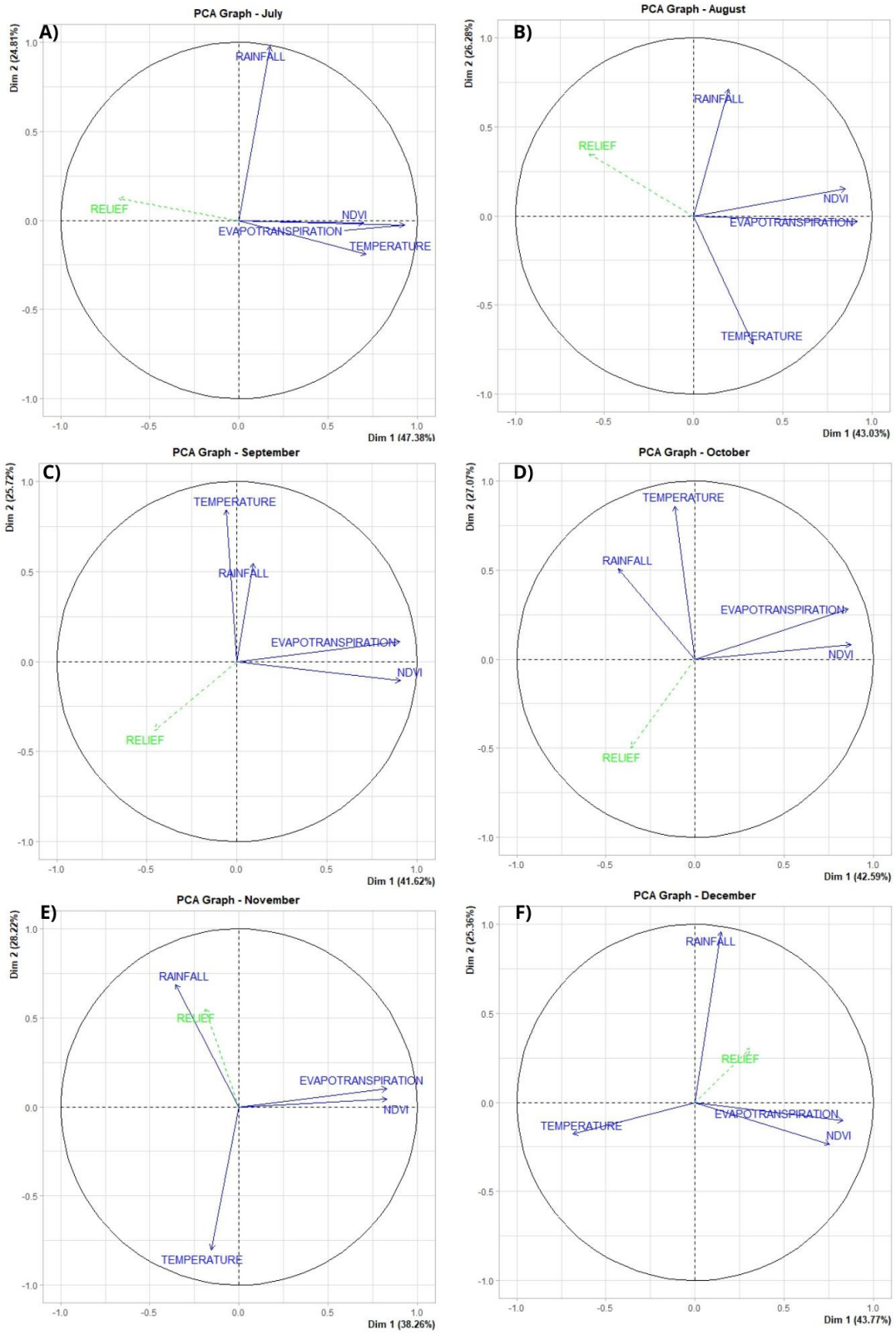

Figure 5.7. Correlation of variables through Principal Component Analysis (PCA) by month: A) July; B) August; C) September; D) October; E) November; and F) December.

The Principal Component Analysis (PCA) correlation circle plots for each month provided valuable insights into the relationships among key environmental variables:

Rainfall, Temperature, NDVI (Normalized Difference Vegetation Index), Evapotranspiration, and Relief. The direction and length of the arrows in the PCA plots indicate the strength and correlation of each variable with the principal components (Dim.1 and Dim.2), revealing seasonal trends and underlying climatic and ecological interactions.

Concerning to January and February (Figures 5.6A and 5.6B), during these months, Dim.1 (explaining over 40% of the variance) groups NDVI, Evapotranspiration, and Relief, indicating that topography strongly shapes vegetation and moisture availability. Rainfall aligns with Dim.2 (around 25%), acting as an independent driver, while Temperature negatively correlates with NDVI and Evapotranspiration, reducing moisture retention and affecting the atmospheric moisture flows (“flying rivers”).

By March (Figure 5.6C), Dim.1 explains less variance, suggesting higher environmental variability, indicating a certain independence between the variables, which characterizes a transitional period in the rainfall regime. Rainfall and Temperature shift to Dim.2, becoming more influential as seasonal climatic drivers, while NDVI and Evapotranspiration remain correlated on Dim.1. This shift often reflects more frequent or variable rainfall events, as well as changing temperature patterns that can alter the direction and strength of moisture-laden air masses, such as the Flying Rivers. April (Figure 5.6D) sees a resurgence in Dim.1’s explanatory power, with NDVI and Evapotranspiration still tightly linked, and Rainfall and Temperature now more evenly spread along Dim.2. This balance underscores the interplay between precipitation and warming trends in shaping large-scale airflows and the atmospheric transport of water vapor.

In May (Figure 5.6E), Temperature begins to correlate positively with NDVI and Evapotranspiration on Dim.1, revealing a warming-driven increase in vegetation and moisture transfer to the atmosphere. Rainfall shifts slightly further along Dim.2, underscoring its seasonal independence. Regarding June (Figure 5.6F), Dim.1 reaches its peak explanatory power as NDVI, Evapotranspiration, and Temperature form a robust cluster, emphasizing the major contribution of vegetation-driven water vapor to Flying Rivers recharge.

July (Figure 5.7A) maintains a strong NDVI–Evapotranspiration loading on Dim.1, with Rainfall and Temperature dominating Dim.2, reflecting temperature’s role in

modulating precipitation patterns. In August (Figure 5.7B), this split remains consistent: vegetation-driven moisture continues on Dim.1, while rainfall persists as an independent driver in Dim.2, supporting stable seasonal moisture flows across the landscape.

For September (Figure 5.7C), Dim.1 and Dim.2 become more balanced, as Temperature joins Rainfall in shaping seasonal dynamics on Dim.2. Meanwhile, NDVI and Evapotranspiration stay on Dim.1, emphasizing their steady response to climatic conditions that support atmospheric moisture transport. By October (Figure 5.7D), the correlation patterns persist: NDVI and Evapotranspiration remain primary on Dim.1, while Rainfall and Temperature drive changes on Dim.2.

Throughout November (Figure 5.7E), NDVI and Evapotranspiration, continue their positive correlation on Dim.1, highlighting vegetation's persistent role in water cycling, while Temperature presents a negative correlation with those variables. Rainfall strengthens its position on Dim.2, reflecting seasonal independence from relief characteristics. In December (Figure 5.7F), NDVI and Evapotranspiration remain dominant on Dim.1, and Rainfall stays the most independent factor, with Temperature inversely related to vegetation activity, a pattern that closes the annual cycle of moisture flows powering the Flying Rivers.

5.3.2. Cluster-based evaluation of rainfall stations: integrating climatic and environmental variables

The cluster analysis provided a comprehensive overview of the interaction between climatic and environmental variables across months, highlighting distinct patterns within each cluster (Figure 5.8).

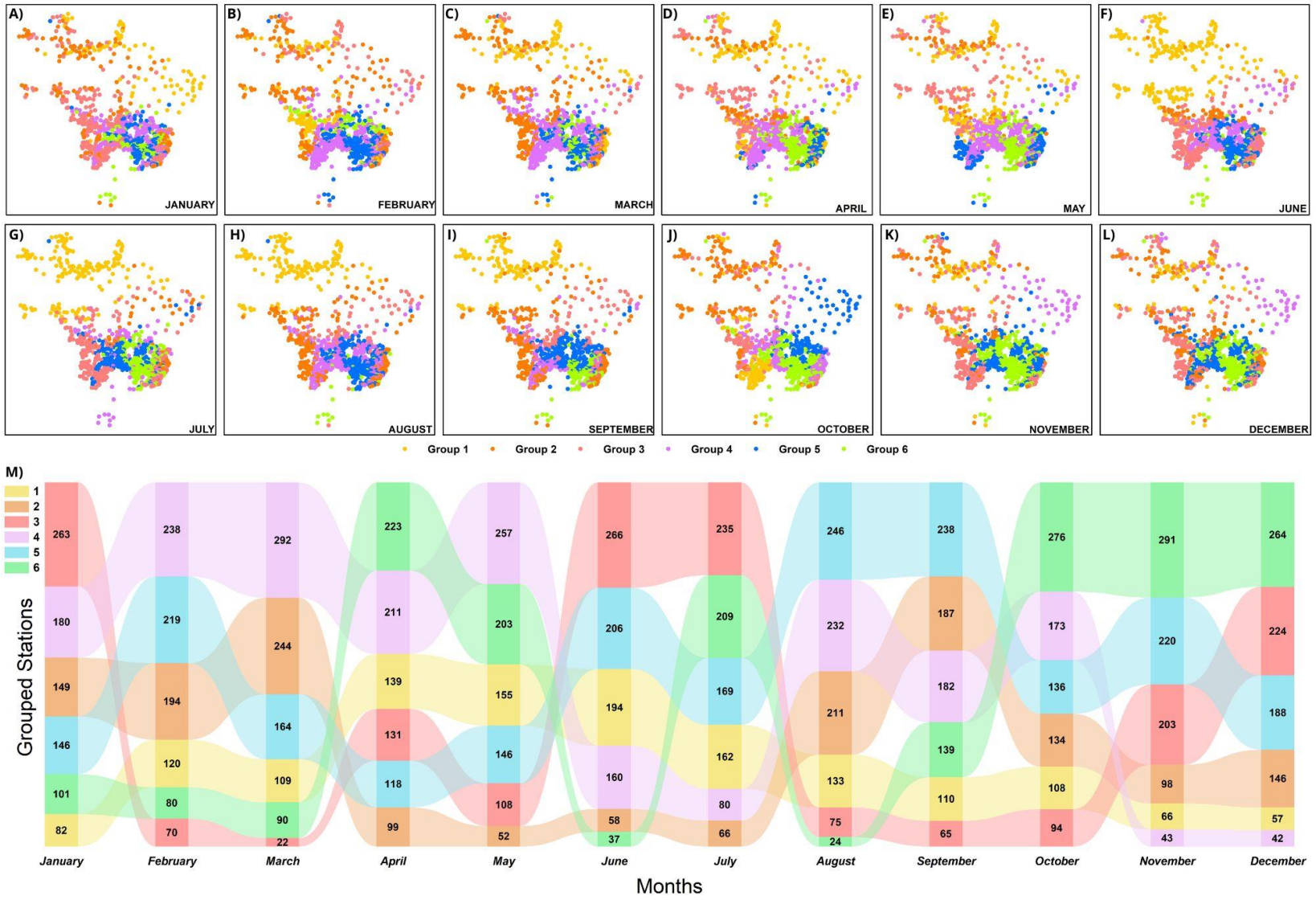

Figure 5.8. Cluster analysis of climate and environmental variables. A) January; B) February; C) March; D) April; E) May; F) June; G) July; H) August; I) September; J) October; K) November; L) December; and M) Ribbon graph.

The yearly dynamics of rainfall station groups, as illustrated in Figures 5.8A to 5.8M, reveal a continuous interplay of climatic and environmental factors, with notable variations in rainfall, NDVI, temperature, and evapotranspiration shaping the characteristics of the clusters. It is important to highlight that the groups, based on the monthly characteristics of the variables, exhibit different distributions within the biomes, thus presenting distinctive hydrological characteristics.

Cluster 1 consistently represents regions with moderate rainfall, temperature, and NDVI, maintaining stability throughout the year and reflecting areas of steady vegetation and mild climatic conditions. In contrast, Cluster 2 emerges as the dominant group during wetter periods, characterized by high NDVI and evapotranspiration, which signify dense vegetation and active ecological processes. Cluster 3, on the other hand, consistently represents arid regions with low rainfall and NDVI, coupled with higher temperatures and sparse vegetation.

In January (Figure 5.8A), Group 3 dominates, indicating a prevalence of stations with moderate rainfall and vegetation, while Group 2, with its high NDVI and evapotranspiration, captures densely vegetated areas. As February progresses (Figure 5.8B), Group 5 expands significantly, and Group 3 continues to reflect moderate rainfall regions, though Groups 1 and 6 decline, showing less vegetation and climatic activity. By March (Figure 5.8C), transitional zones emerge with Group 5 maintaining prominence, but a decline in Groups 3 and 6 signals a reduction in the stations. It is important to highlight that in the months of February and March, Group 4, which is predominant in the Atlantic Forest biome, experiences a significant increase in the number of stations. This phenomenon demonstrates the close interrelationship between climatic, biogeophysical factors and the hydrological regime of this biome.

The patterns shift in April (Figure 5.8D), with a decline in Groups 2 and 5, reflecting areas with decreasing NDVI and rainfall. Group 1, aligning with Group 6, increasing values, while Cluster 3 continues to reflect arid areas. May (Figure 5.8E) marks the peak of ecological activity, with Groups 4 and 6 dominating achieving high rainfall and NDVI, signifying vibrant vegetation. Group 3 regains prominence by June (Figure 5.8F), reflecting seasonal shifts toward moderate rainfall and steady vegetation.

During the second half of the year, the dominance of different groups highlights changing environmental conditions. In July (Figure 5.8G), Groups 3 and 6 dominate,

reflecting stable vegetation and regions with significant NDVI and evapotranspiration. By August (Figure 5.8H), Group 5 gains prominence, showcasing high NDVI and rainfall, while Group 3 remains associated with sparse vegetation and drier conditions, typical for this latitude at this time of year. In September (Figure 5.8I), Group 5 sustains its dominance, aligning with high NDVI and rainfall, while Group 1 resurges slightly, reflecting consistent vegetation and climate.

The final quarter of the year sees Group 6, the Atlantic Forest region, emerge as the dominant group in October (Figure 5.8J), continuing through November and December (Figure 5.8K), with Group 5, covering mainly the Pantanal, maintaining significant contributions. By December (Figure 5.8L), Groups 3 regain prominence, reflecting the return of moderate and wetter conditions, while Group 5 slightly declines. Throughout this period, Cluster 2 remains stable with moderate rainfall and NDVI, Cluster 3 highlights dense vegetation and ecological activity, Cluster 4 presents a drastic decline representing a decrease in precipitation and reflecting drier conditions, and Cluster 1 persists in representing arid areas with sparse vegetation.

The ribbon graph (Figure 5.8M) encapsulates these monthly transitions, illustrating the dynamic interplay between rainfall, NDVI, temperature, and evapotranspiration. Based on the hydrological patterns (Figure 5.10) in the southeast and northern regions, Groups 3 and 5 prevail during drier, vegetative periods, while Group 6 becomes more prominent in wetter months. This dynamic underscore the resilience and adaptability of vegetation and water-use processes to seasonal climatic variations. These patterns underscore the temporal variability of rainfall station characteristics and the complex interactions among climatic and environmental factors, offering valuable insights into regional and seasonal dynamics.

The cluster analysis focused on the seasonality of rainfall stations (Figure 5.9) reveals distinct seasonal patterns of precipitation, which directly influence atmospheric air mass behavior and the recharge of Flying Rivers. The results show two major groups: Cluster 1 (yellow), representing stations with lower precipitation (≤ 455 mm), and Cluster 2 (blue), indicating regions with higher rainfall (> 455 mm). These classifications highlight the shifting balance between moisture-laden air masses and drier conditions, which govern the seasonal cycle of atmospheric moisture transport.

During the rainy season from January to March (Figure 5.9 A-C), most stations fall into Cluster 2, with higher precipitation concentrated in southern and central regions. This period marks the peak of moisture availability, where vegetation-driven evapotranspiration enriches the atmosphere with water vapor, fueling the formation of Flying Rivers. The prevalence of high-rainfall zones reflects intense convective activity and large-scale atmospheric circulation, which transport moisture-laden air to distant regions, supporting rainfall far from the original sources.

By April (Figure 5.9D), a transitional period begins, as the number of Cluster 1 (low rainfall) stations increases, particularly in northern regions. This shift indicates the initial weakening of moisture recycling, as rainfall declines and atmospheric moisture transport starts to diminish. Despite this, Cluster 2 stations still experience significant precipitation, with an average rainfall of 330 mm, compared to only 91 mm in Cluster 1, emphasizing the continued role of remaining moisture sources in maintaining the water vapor pathways.

From May to July (Figure 5.9 E-G), dry season conditions intensify, leading to the expansion of Cluster 1 across most regions. The reduction in high-precipitation areas signifies lower evapotranspiration rates, decreasing the amount of atmospheric moisture available for transport. Although localized pockets of Cluster 2 persist, particularly in the northern and southern regions, the overall decline in rainfall suggests that moisture transport processes have weakened, limiting the Flying River's ability to sustain long-distance precipitation. This seasonal drying trend corresponds to a reduction in convection and cloud formation, further restricting atmospheric moisture flux.

In August and September (Figures 5.9H and 5.9I), the atmospheric moisture cycle begins to recover, as Cluster 2 gradually expands. By September, a significant number of stations return to the high-precipitation category, indicating the progressive reactivation of the Flying River system. This shift corresponds to strengthening convective activity and renewed moisture recycling, allowing for the redistribution of atmospheric water vapor as the next rainy season approaches.

As the wet season progresses into October in the southeast (Figure 5.9J), Cluster 2 remains dominant, with a slight increase in Cluster 1 stations. This shift occurs as the dry season sets in across the northeast and parts of the northern region. The strengthening of high-precipitation zones signals the seasonal growth of moisture

transport networks, as vegetation enters a phase of high evapotranspiration, increasing the replenishment of airborne moisture.

In November (Figure 5.9K), Cluster 2 expands significantly, reaching 779 stations, while Cluster 1 decreases to 142 stations. This shift highlights the intensification of moisture transport, as more regions experience high precipitation levels. The dominance of Cluster 2 indicates that evapotranspiration-driven moisture cycling has regained strength, reinforcing the role of atmospheric circulation in distributing rainfall across broader areas.

By December (Figure 5.9L), Cluster 2 remains predominant, with 586 stations, compared to 335 in Cluster 1. This pattern confirms that the Flying River system has fully recharged, supplying abundant moisture to support widespread rainfall across the region. The strong presence of high-rainfall stations suggests that vegetation-driven evapotranspiration is at its seasonal peak, sustaining the atmospheric moisture flux necessary for long-distance transport.

The cluster analysis highlights the seasonal cycle of atmospheric moisture transport, shaping rainfall distribution throughout the year. In summary, from January to March, Cluster 2 dominates, reflecting peak of activity, where evapotranspiration-driven moisture transport sustains widespread rainfall. Between April and July, Cluster 1 expands, indicating weakening moisture transport as evapotranspiration declines and atmospheric moisture diminishes. From August to December, Cluster 2 resurges, marking the reactivation of Flying Rivers, restoring atmospheric moisture cycling, and ensuring widespread precipitation. These findings underscore the vital role of forests in sustaining atmospheric moisture flux, essential for hydrological stability, climate resilience, and water resource management.

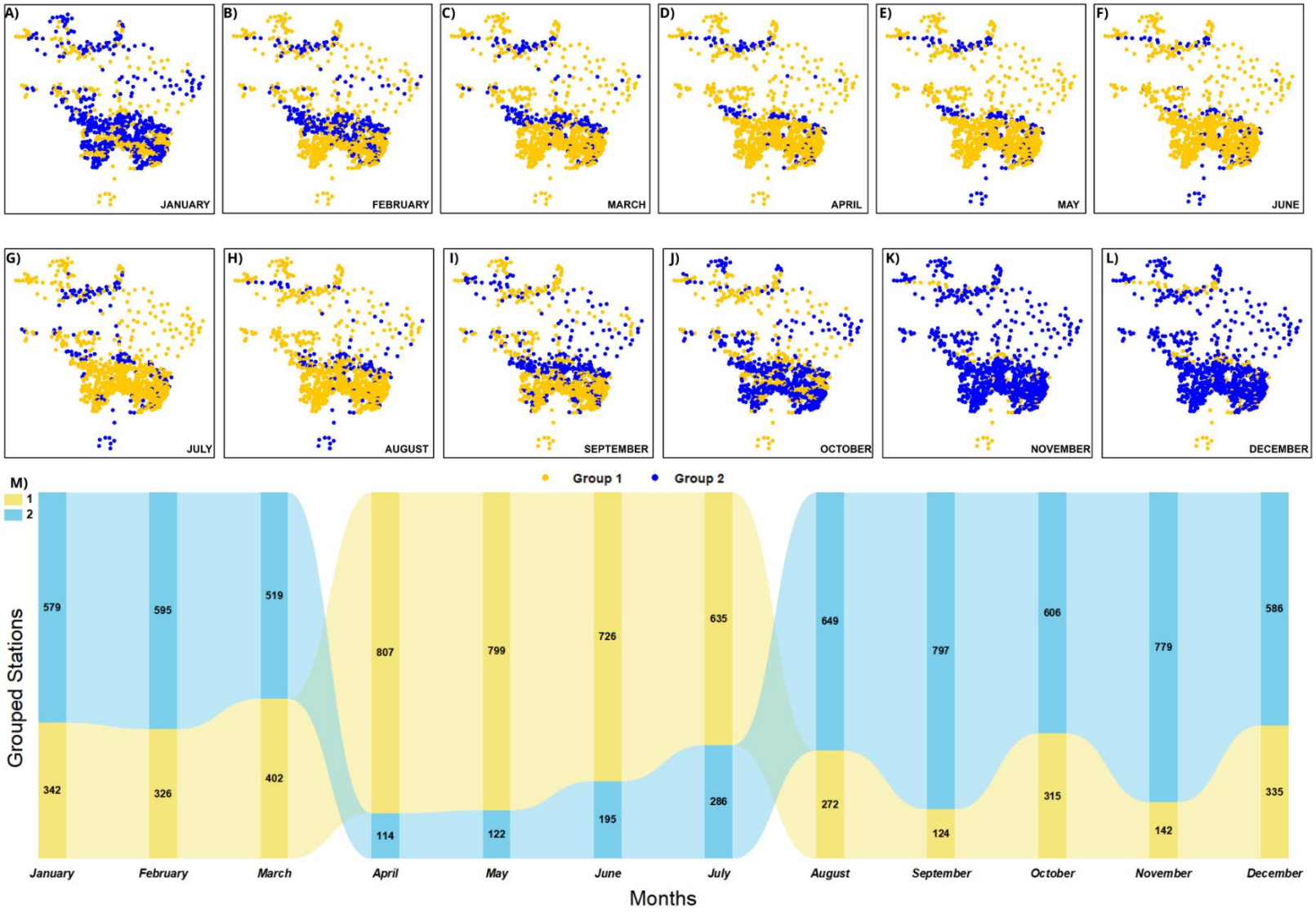

Figure 5.9. Cluster analysis focused on the seasonality of rainfall stations. A) January; B) February; C) March; D) April; E) May; F) June; G) July; H) August; I) September; J) October; K) November; L) December; and M) Ribbon graph.

5.4. Discussion

To facilitate the understanding of the hydrological behavior of Brazilian biomes, Figure 5.10 has been included to support the analysis and discussion of the results obtained.

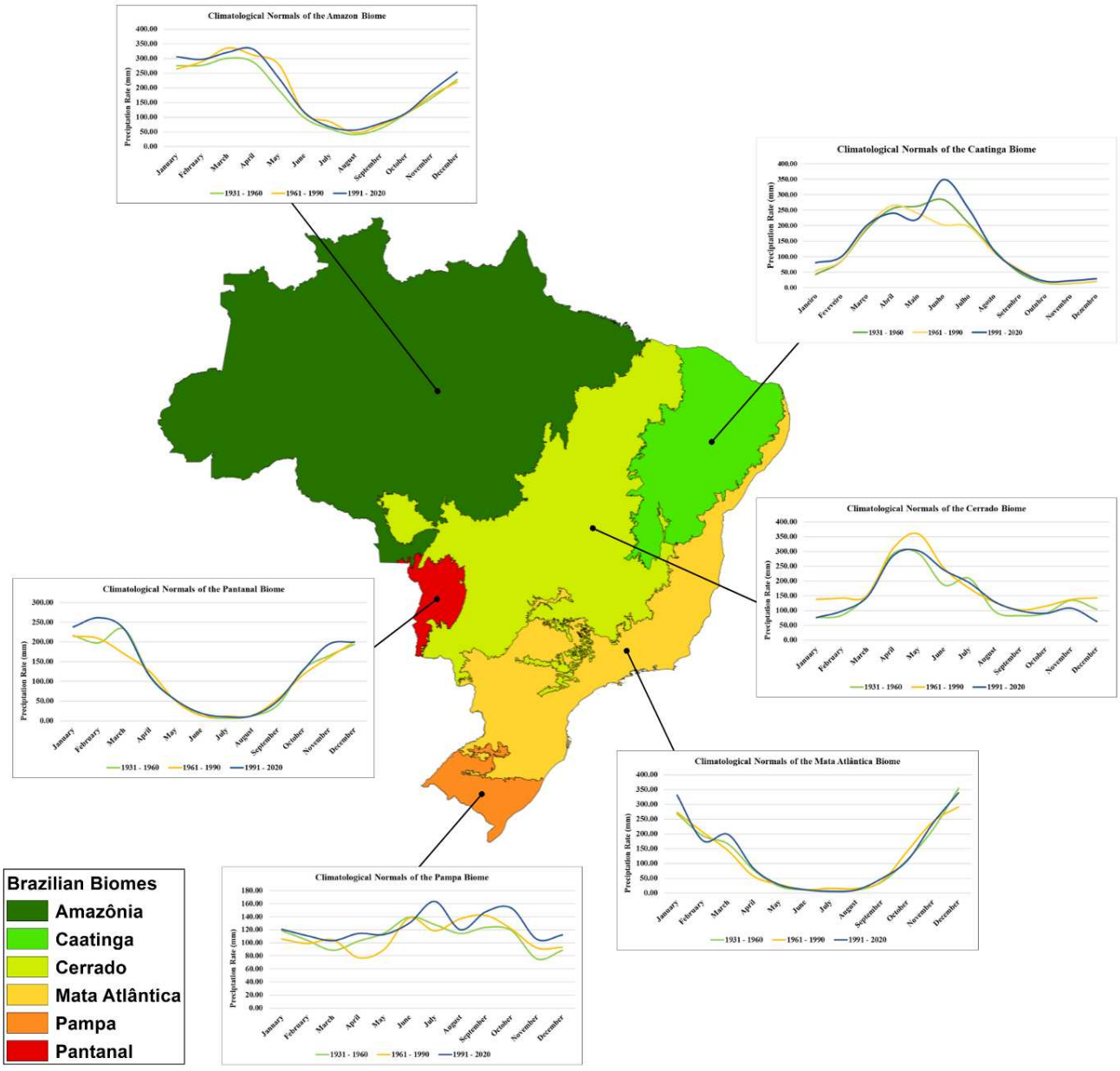

Figure 5.10. Behavior of Climatological Normals from the Years 1931–1960 (Green), 1961–1990 (Yellow), and 1991–2020 (Blue) in Brazilian Biomes.

The Principal Component Analysis (PCA) results provide valuable insights into the interrelationships among key environmental variables—NDVI, Rainfall, Temperature, Evapotranspiration, and Relief—and their seasonal impacts on climate behavior (Jolliffe & Cadima, 2016). These interactions play a critical role in shaping atmospheric moisture flux and the development of large-scale moisture transport systems, such as the Flying Rivers, which redistribute water across vast distances.

During the early months of the year (January and February), NDVI and Evapotranspiration exhibit strong loadings on the first principal component (Dim.1), indicating the dominance of vegetation-driven moisture flux in the atmospheric water cycle. This is consistent with previous studies demonstrating that evapotranspiration from forested regions significantly contributes to atmospheric humidity, which in turn influences precipitation patterns (Salati & Vose, 1984; Amaral e Silva et al., 2020). Rainfall, on the other hand, appears as an independent driver (in Dim.2), emphasizing its seasonal variability and relatively autonomous behavior during this period. Concurrently, temperature negatively correlates with both NDVI and Evapotranspiration, suggesting that higher temperatures lead to increased soil moisture loss and reduced vegetation water retention (Poveda et al., 2014)

As the wet season progresses into March and April, in biomes such as the Cerrado and Caatinga, the interaction between NDVI, Evapotranspiration, and Rainfall becomes more pronounced. The positive correlation of temperature with these variables suggests that warming enhances evapotranspiration, increasing atmospheric moisture availability and precipitation potential. This aligns with findings that temperature-driven evapotranspiration plays a crucial role in sustaining the hydrological cycle, particularly in tropical and subtropical regions (van der Ent et al., 2010). Rainfall's growing influence (on Dim.2) and further reinforces the dynamic relationship between temperature and precipitation during these transitional months, when moisture cycling reaches its peak intensity.

By May and June, the dominance of Temperature and Evapotranspiration (in Dim.1) reflects the peak period of atmospheric moisture recharge, particularly in the Amazon biome as the dry season begins. This phase is characterized by high temperatures and

increased evapotranspiration (D'Acunha et al., 2024), which play a crucial role in replenishing the Flying Rivers. NDVI, while still relevant, appears to stabilize, suggesting that vegetation activity has reached a steady state. Rainfall continues as an independent factor (in Dim.2), reflecting its responsiveness to seasonal climatic conditions rather than direct vegetation cycles. This period coincides with the maximum development of the Flying Rivers, as evapotranspiration-driven moisture input fuels long-distance atmospheric water transport, benefiting downwind agricultural and semiarid regions (Zemp et al., 2014).

As the dry season intensifies in July and August in Amazonia, Evapotranspiration remains a dominant factor, though NDVI and Temperature suggest a gradual decline in vegetation-driven moisture cycling. Rainfall remains independent (in Dim.2), while a new inverse relationship between Temperature and NDVI emerges (in Dim.3). This finding suggests that increased temperatures begin to stress vegetation, reducing its water cycling efficiency (Emmerichs et al., 2024). These observations are consistent with studies indicating that prolonged dry-season warming can weaken vegetation resilience, limiting its role in atmospheric moisture recycling (Malhi et al., 2008). The farthest extent of the Flying Rivers is reached during this time, sustaining precipitation in remote regions, such as the interior of the Cerrado, in the westernmost part of the biome, before atmospheric moisture flux begins to decline.

During September and October, NDVI and evapotranspiration remain dominant in Dimension 1. However, the emergence of a negative correlation between rainfall and these variables in October suggests a decoupling of precipitation from vegetation-driven moisture flux. This shift marks the weakening of the Flying River system, as evapotranspiration rates decline and atmospheric moisture transport slows down (Gimeno-Sotelo et al., 2024). Meanwhile, Temperature gains dominance (Dim.2), reinforcing its role in accelerating landscape drying. This seasonal redistribution of atmospheric moisture aligns with researches showing that temperature increases during the late dry season contributing to delayed rainfall onset and intensifying drought conditions (Fu et al., 2013; Zhou and Shi, 2024).

By November and December, declining Evapotranspiration and increasing Temperature-induced drying effects become more evident. NDVI and Evapotranspiration still define these months (Dim.1), but a strong inverse relationship between Temperature and vegetation indices in December confirms the seasonal impact of warming on plant water availability. Rainfall, meanwhile, remains independent (in Dim.2), suggesting that precipitation events are less dependent on local vegetation-driven moisture input during this period (Theeuwen et al., 2023). These findings align with literature indicating that seasonal shifts in evapotranspiration and precipitation regimes regulate the annual cycle of atmospheric moisture flux, with critical implications for regional water security (Ellison et al., 2017).

The inclusion of Relief in the PCA provided further insights into the role of topography in shaping climate interactions. The strong clustering of NDVI, Evapotranspiration, and Relief (in Dim.1) during January and February highlights how topographic variation influences vegetation distribution and moisture availability. This pattern is particularly relevant in montane and hilly landscapes, where elevation-driven climatic gradients modulate temperature and precipitation, thereby affecting evapotranspiration rates and water fluxes (Li et al., 2024). The shifting influence of Rainfall and Temperature (on Dim.2) throughout the year reflects the seasonal variability in atmospheric moisture sources and transport mechanisms, reinforcing the importance of considering elevation in hydrological and climatic models.

Overall, the PCA analysis underscores the dynamic nature of climate-vegetation interactions and the pivotal role of evapotranspiration in sustaining atmospheric moisture transport. The findings corroborate with existing research on the influence of land cover on regional precipitation patterns, demonstrating that changes in vegetation structure and temperature regimes can significantly impact the hydrological cycle (Amaral e Silva et al., 2020; Pratap and Markonis, 2022). Understanding these relationships is crucial for predicting future climate variability and developing strategies to mitigate the impacts of deforestation and land-use change on regional water balance.

Regarding cluster analysis, it was conducted to integrate climatic and environmental variables (e.g., rainfall, NDVI, temperature, and evapotranspiration),

facilitating a nuanced examination of the temporal and spatial dynamics across rainfall stations. Clustering techniques are extensively used in climatology to classify atmospheric and oceanic states, thus revealing underlying structures in the climate system (Wilks, 2011). In this study, clustering enabled the identification of coherent station groupings based on precipitation seasonality, representing a geographically defined area, offering insights into how climatic-environmental-physical interactions evolve throughout the year.

Group 1 consistently represents regions with moderate rainfall, temperature, and NDVI, indicative of stable vegetation and relatively mild climatic conditions, this group is more prevalent in the Amazon biome, the southwestern Cerrado, and the northern region of the Caatinga. Such stability often reflects ecosystems capable of maintaining hydrological and thermal balance. Regarding the rain gauge stations located in the Amazon and Pantanal biomes, which are predominantly found in Group 2, these biomes experience wetter periods marked by high NDVI values and elevated evapotranspiration rates. This pattern highlights robust plant growth driven by plentiful water resources, consistent with findings that link elevated evapotranspiration fluxes to dense vegetation cover (Zhang et al., 2001; Zheng et al., 2024). Group 3, in contrast, characterizes more arid regions or seasons featuring low rainfall, high temperatures, and sparse vegetation, posing challenges related to water scarcity and thermal stress on plant communities (Huang et al., 2022).

The hydrological cycles in Brazil exhibit considerable regional variation due to differences in biome characteristics, atmospheric circulation patterns, and seasonal climatic conditions (Souza and Medeiros, 2024). A key driver of seasonal moisture conveyance in South America is the Flying Rivers phenomenon, wherein vast air currents transport water vapor from the Amazon Basin to remote parts of the continent (Salati & Vose, 1984; World Economic Forum, 2024). The Amazon rainforest, via substantial transpiration, plays a critical role in sustaining this aerial moisture corridor. However, deforestation disrupts these processes, diminishing transpiration and altering atmospheric circulation pathways—a development that can lead to reduced precipitation in regions historically reliant on these moisture flows (Nóbrega, 2014).

Monthly analyses reveal dynamic shifts in cluster memberships. For instance, in January, Group 3 prevails, suggesting moderate rainfall conditions in the Southeast and Central-West regions of Brazil, which characterize the rainy season in the Pantanal and Atlantic Forest biomes. By March (the end of the wet season), the emergence of Group 4, predominantly in the southwestern Cerrado biome and southern Amazon, signifies accelerated vegetation growth driven by early seasonal precipitation in these biomes. (Smith and Boers, 2023). Mid-year observations (around July) indicate climatic stabilization, dominated by Groups 3 and 6, which correspond to the Atlantic Forest, southern Cerrado, and Pantanal biomes, although significant temperature fluctuations persist. (Miranda et al., 2024), This period is hydrologically characterized by dry conditions with sporadic rainfall in the southeastern region.

By September, renewed precipitation fosters a temporary rise in Group 5, which corresponds to the core regions of the Cerrado and Atlantic Forest biomes, where the rainy season begins. In contrast, vegetation responses in October appear dampened, as this group becomes more representative of the Caatinga, possibly due to prolonged dryness (Figure 5.10) or reduced soil moisture retention (Zheng et al., 2024). These findings, visually consolidated in ribbon graph analyses, underscore the intricate interplay among rainfall, NDVI, temperature, and evapotranspiration in shaping the seasonal climate mosaic.

Focusing on the seasonality of rainfall stations, the cluster analysis identified stations with lower precipitation (Group 1) versus those with higher precipitation (Group 2), reflecting regional and temporal climatic discrepancies. In March, for example, stations in Group 1 register an average precipitation of approximately 75 mm (ranging between 30 mm and 200 mm), whereas stations in Group 2 exhibit substantially higher rainfall, mainly in Cerrado, Pampa and Caatinga biomes (Figure 5.10), averaging 290 mm and peaking at 650 mm (Miller et al., 2021). These fluctuations underscore the need for adaptive water resource management strategies that anticipate short-term climatic shifts and their impact on water availability.

Brazil’s climatic regimes, heavily influenced by both atmospheric circulation patterns and geographic factors, manifest pronounced wet and dry seasons. During the

wet season (November to May) in the Amazon, underpinned by the South American Monsoon System (SAMS), during which Flying Rivers enhance precipitation in southeastern Brazil (Marengo et al., 2024), elevated evapotranspiration rates contribute significantly to atmospheric moisture. This moisture is transported southward, enhancing rainfall in the Cerrado and southeastern Brazil. The increased precipitation during this period supports agricultural productivity and ecosystem resilience (Flach et al., 2021). However, as the dry season begins in June, evapotranspiration continues to play a role in atmospheric moisture flux despite reduced rainfall. This persistent moisture export supports precipitation in more distant regions, underscoring the biome's hydrological influence beyond its geographic boundaries (Versieux and Costa, 2024).

The Cerrado biome, characterized by its pronounced wet and dry seasons, serves as both a recipient and contributor to atmospheric moisture transport. During the rainy season (March to August), moisture input from the Amazon enhances precipitation, promoting vegetation growth and evapotranspiration (Hofmann et al., 2021). In turn, the evapotranspiration process contributes to localized atmospheric moisture recycling. As the dry season intensifies, soil moisture depletion leads to reduced vegetation activity and evapotranspiration, contributing to a decline in atmospheric moisture flux and influencing rainfall patterns in adjacent regions (Xu et al., 2022).

Interactions between the Amazon Rainforest and Brazil’s southeastern and southern regions underscore the regional importance of moisture transport and atmospheric feedbacks. Studies indicate that while moderate forest clearing (up to ~55–60% within 28-km grid cells) can initially augment local rainfall, extensive deforestation ultimately depletes precipitation levels through reduced moisture recycling and altered circulation (Zemp et al., 2017). This relationship can be observed when analyzing the hydrological behavior of the biomes that cover the Amazon and southeastern Brazil in Figure 5.10. Such disruptions to the Flying Rivers have profound socio-economic implications, including threats to agricultural productivity, hydropower generation, and urban water security (Lovejoy and Nobre, 2019).

The Pantanal and Caatinga biomes also experience hydrological cycles influenced by broader climatic processes. This biome, as a lowland wetland system, depends on

rainfall patterns from the surrounding highland areas and moisture inputs from the Amazon (Lázaro et al., 2020; Marques et al., 2021). Seasonal flooding during the wet season supports diverse ecosystems, while dry season moisture deficits expose the biome to increased drought risks (Furtak and Wolińska, 2023). The Caatinga, characterized by its semiarid climate, receives moisture primarily through episodic rainfall events during the wet season, driven by atmospheric dynamics linked to the Intertropical Convergence Zone (ZCIT) (Utida et al., 2023). Prolonged dry periods in the Caatinga highlight the biome's vulnerability to climate variability and the importance of moisture inputs from adjacent regions.

Another critical system shaping rainfall variability is the South Atlantic Convergence Zone (ZCAS), generated by moisture convergence between the South Atlantic High and the continental thermal low-pressure center (Carvalho et al., 2004). The intensity and position of the ZCAS exhibit strong seasonality, particularly in austral summer, modulated by oceanic and terrestrial thermal gradients. Variations in ZCAS intensity can differentially influence rainfall in northern Argentina, southern Brazil, and central-east Brazil, highlighting the need for integrated regional-scale analyses.

Climate change is projected to exacerbate existing vulnerabilities within these interconnected systems. Rising temperatures and shifting precipitation patterns may push portions of the Amazon rainforest toward a drier savanna-like state, with cascading repercussions for both local ecosystems and distant regions reliant on Amazon-derived moisture (Strand et al., 2018). Consequently, land management practices, climate adaptation policies, and global mitigation efforts are paramount to sustaining the hydrological balance and ecosystem services provided by the Amazon Basin (Nobre, 2014).

In summary, the rainfall patterns of the Amazon and southeastern/southern Brazil are intricately linked through atmospheric moisture transport and regional climate systems, as can be seen in Figures 5.9 and 5.10. Deforestation and climate change pose significant threats to these patterns, underscoring the need for sustainable land management and climate mitigation efforts to preserve the hydrological balance across these interconnected regions.

5.5. Conclusion

This study provides a comprehensive assessment of rainfall variability in Brazil, incorporating multivariate statistical techniques to analyze its interactions with climatic and environmental factors. The application of PCA identified dominant variables influencing precipitation patterns, revealing a strong correlation between rainfall, NDVI, and evapotranspiration during wet months, while temperature emerged as a key driver in drier periods. The integration of relief data further demonstrated the role of orographic effects in modulating precipitation distribution, reinforcing previous studies on topographic influence on rainfall.

The CLARA cluster analysis effectively categorized rainfall stations into distinct seasonal and environmental groupings, illustrating the dynamic nature of rainfall seasonality. Areas with high NDVI and evapotranspiration were linked to elevated precipitation levels, highlighting the essential role of vegetation in moisture recycling and climate regulation. Conversely, arid zones, characterized by lower NDVI and higher temperatures, exhibited reduced precipitation and increased vulnerability to climatic stress.

One of the most significant findings of this research is the influence of the Flying Rivers phenomenon in transporting Amazonian moisture to southeastern Brazil. The study underscores how deforestation and land-use changes are disrupting this hydrological cycle, leading to decreased precipitation and heightened climatic instability. These disruptions pose severe risks to agriculture, water security, and energy production, necessitating urgent mitigation efforts.

By leveraging PCA and CLARA clustering, this study contributes valuable insights for climate monitoring, sustainable water resource management, and policy formulation, while also providing a methodological framework for spatiotemporal analysis of climatic variables (rainfall, temperature, and evapotranspiration) and biogeophysical factors (NDVI and topography). The findings emphasize the importance of preserving the Amazon Rainforest to sustain regional and continental hydrological balance. Future

research should focus on long-term climate modeling and predictive analytics to enhance adaptive strategies against climate change-induced rainfall variability.

References

Altman, N., Krzywinski, M., 2015. Points of Significance: Association, correlation and causation. Nat Methods 12, 899–900. https://doi.org/10.1038/nmeth.3587

Amaral e Silva, A., Braga, M.Q., Ferreira, J., Juste dos Santos, V., do Carmo Alves, S., de Oliveira, J.C., Calijuri, M.L., 2020. Anthropic activities and the Legal Amazon: Estimative of impacts on forest and regional climate for 2030. Remote Sens Appl 18, 100304. https://doi.org/10.1016/j.rsase.2020.100304

Amaral e Silva, A., de Assis, L.C., dos Santos, V.J., de Andrade, L.C., Lorentz, J.F., Henriques, B.S., Calijuri, M.L., Ferreira, I.O., 2024. Rainfall From Brazilian Flying Rivers: Evaluating the Effectiveness of Precipitation Gridded Databases. International Journal of Climatology. https://doi.org/10.1002/joc.8707

Amato, G., Gennaro, C., Oria, V., Radovanovic, M., 2019. Similarity Search and Applications, Lecture Notes in Computer Science. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-32047-8

ANA - Agência Nacional de Águas e Saneamento Básico. Disponível em:< https://www.gov.br/ana/pt-br>. Acesso em: 12 de novembro de 2024.

Carvalho, L. M. V; 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, p. 88–108, 2004.

D'Acunha, B., Dalmagro, H. J., Zanella de Arruda, P. H., Biudes, M. S., Lathuillière, M. J., Uribe, M., Couto, E. G., Brando, P. M., Vourlitis, G., & Johnson, M. S. (2024). Changes in evapotranspiration, transpiration and evaporation across natural and managed landscapes in the Amazon, Cerrado and Pantanal biomes. Agricultural and Forest Meteorology, 346, 109875. https://doi.org/10.1016/j.agrformet.2023.109875

Didan, K. MODIS/Terra Vegetation Indices Monthly L3 Global 0.05Deg CMG V061. 2021, distributed by NASA EOSDIS Land Processes Distributed Active Archive Center, https://doi.org/10.5067/MODIS/MOD13C2.061. Accessed 2025-02-18.

Ellison, D., Morris, C.E., Locatelli, B., Sheil, D., Cohen, J., Murdiyarso, D., Gutierrez, V., Noordwijk, M. van, Creed, I.F., Pokorny, J., Gaveau, D., Spracklen, D. V., Tobella, A.B., Ilstedt, U., Teuling, A.J., Gebrehiwot, S.G., Sands, D.C., Muys, B., Verbist, B., Springgay, E., Sugandi, Y., Sullivan, C.A., 2017. Trees, forests and water: Cool insights for a hot world. Global Environmental Change 43, 51–61. https://doi.org/10.1016/j.gloenvcha.2017.01.002

Emmerichs, T., Lu, Y.-S., Taraborrelli, D., 2024. The influence of plant water stress on vegetation–atmosphere exchanges: implications for ozone modelling. Biogeosciences 21, 3251–3269. https://doi.org/10.5194/bg-21-3251-2024

Flach, R., Abrahão, G., Bryant, B., Scarabello, M., Soterroni, A. C., Ramos, F. M., Valin, H., Obersteiner, M., & Cohn, A. S. (2021). Conserving the Cerrado and Amazon biomes of Brazil protects the soy economy from damaging warming. World Development, 146, 105582. https://doi.org/10.1016/j.worlddev.2021.105582

Fu, R., Yin, L., Li, W., Arias, P.A., Dickinson, R.E., Huang, L., Chakraborty, S., Fernandes, K., Liebmann, B., Fisher, R., Myneni, R.B., 2013. Increased dry-season length over southern Amazonia in recent decades and its implication for future climate projection. Proceedings of the National Academy of Sciences 110, 18110–18115. https://doi.org/10.1073/pnas.1302584110

Furtak, K., & Wolińska, A. (2023). The impact of extreme weather events as a consequence of climate change on the soil moisture and on the quality of the soil environment and agriculture – A review. CATENA, 231, 107378. https://doi.org/10.1016/j.catena.2023.107378

Ghozat, A., Sharafati, A., Hosseini, S.A., 2021. Long-term spatiotemporal evaluation of CHIRPS satellite precipitation product over different climatic regions of Iran. Theor Appl Climatol 143, 211–225. https://doi.org/10.1007/s00704-020-03428-5

Gimeno-Sotelo, L., Fernández-Alvarez, J.C., Nieto, R., Vicente-Serrano, S.M., Gimeno, L., 2024. The increasing influence of atmospheric moisture transport on hydrometeorological extremes in the Euromediterranean region with global warming. Commun Earth Environ 5. https://doi.org/10.1038/s43247-024-01787-9

Gupta, S.K., Gupta, N., Singh, V.P., 2021. Variable-Sized Cluster Analysis for 3D Pattern Characterization of Trends in Precipitation and Change-Point Detection. J Hydrol Eng 26. https://doi.org/10.1061/(asce)he.1943-5584.0002010

Harb, H., Nader, C.A., Jaber, A., Hakem, M., Charr, J.C., Jaoude, C.A., Zaki, C., 2024. CLARA: A cluster-based node correlation for sampling rate adaptation and fault tolerance in sensor networks. Internet of Things (Netherlands) 28. https://doi.org/10.1016/j.iot.2024.101345

Hofmann, G. S., Cardoso, M. F., Alves, R. J. V., Weber, E. J., Barbosa, A. A., de Toledo, P. M., Pontual, F. B., Salles, L. de O., Hasenack, H., Cordeiro, J. L. P., Aquino, F. E., & de Oliveira, L. F. B. (2021). The Brazilian Cerrado is becoming hotter and drier. Global Change Biology, 27(17), 4060–4073. https://doi.org/10.1111/gcb.15712

Hopkins, B., Skellam, J.G., 1954. A New Method for determining the Type of Distribution of Plant Individuals WITH AN APPENDIX BY With one Figure in the Text. Ann Bot 18.

Hsu, J., Huang, W.-R., Liu, P.-Y., Li, X., 2021. Validation of CHIRPS Precipitation Estimates over Taiwan at Multiple Timescales. Remote Sens (Basel) 13, 254. https://doi.org/10.3390/rs13020254

Huang, X., Han, S., Shi, C., 2022. Evaluation of Three Air Temperature Reanalysis Datasets in the Alpine Region of the Qinghai–Tibet Plateau. Remote Sens (Basel) 14. https://doi.org/10.3390/rs14184447

INMET, 2022. Normais Climatológicas do Brasil 1991 - 2020, Instituto Nacional de Meteorologia. Brasilia.

Jollife, I.T., Cadima, J., 2016. Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. https://doi.org/10.1098/rsta.2015.0202

Kassambara, A., 2017. Multivariate Analysis I - Practical Guide To Cluster Analysis in R Unsupervised Machine Learning.

Kaufman, L., Rousseeuw, P.J., 1987. Clustering by means of medoids, in: Statistical Data Analysis Based on the L1-Norm and RelatedMethods, North-Holland, Dordrecht. pp. 405–416.

Kaufman, L., Rousseeuw, P.J., 1991. Clustering Large Applications (Program CLARA), in: Finding Groups in Data: An Int6roduction to Cluster Analysis.

Lázaro, W. L., Oliveira-Júnior, E. S., Silva, C. J. da, Castrillon, S. K. I., & Muniz, C. C. (2020). Climate change reflected in one of the largest wetlands in the world: an overview of the Northern Pantanal water regime. Acta Limnologica Brasiliensia, 32. https://doi.org/10.1590/s2179-975x7619

Lever, J., Krzywinski, M., Altman, N., 2017. Points of Significance: Principal component analysis. Nat Methods 14, 641–642. https://doi.org/10.1038/nmeth.4346

Li, H., Liu, X., Zhang, W., Zhu, H., Zhao, X., Liu, J., Luo, X., Wang, R., Zhao, H., Wang, C., 2024. Elevational Patterns of Forest Evapotranspiration and Its Sensitivity to Climatic Variation in Dryland Mountains. Water (Switzerland) 16. https://doi.org/10.3390/w16091252

Liu, J., Shangguan, D., Liu, S., Ding, Y., Wang, S., Wang, X., 2019. Evaluation and comparison of CHIRPS and MSWEP daily-precipitation products in the Qinghai-Tibet Plateau during the period of 1981–2015. Atmos Res 230, 104634. https://doi.org/10.1016/j.atmosres.2019.104634

López-Bermeo, C., Montoya, R.D., Caro-Lopera, F.J., Díaz-García, J.A., 2022. Validation of the accuracy of the CHIRPS precipitation dataset at representing climate variability in a

tropical mountainous region of South America. Physics and Chemistry of the Earth, Parts A/B/C 127, 103184. https://doi.org/10.1016/j.pce.2022.103184

Lovejoy, T.E., Nobre, C., 2019. Amazon tipping point: Last chance for action. Sci Adv 5, 4–6. https://doi.org/10.1126/sciadv.aba2949

Malhi, Y., Roberts, J.T., Betts, R.A., Killeen, T.J., Li, W., Nobre, C.A., 2008. Climate Change, Deforestation, and the Fate of the Amazon. Science (1979) 319, 169–172. https://doi.org/10.1126/science.1146961

Marengo, J.A., Espinoza, J.-C., Fu, R., Jimenez Muñoz, J.C., Alves, L.M., Da Rocha, H.R., Schöngart, J., 2024. Long-term variability, extremes and changes in temperature and hydrometeorology in the Amazon region: A review. Acta Amazon 54. https://doi.org/10.1590/1809-4392202200980

Marques, J. F., Alves, M. B., Silveira, C. F., Amaral e Silva, A., Silva, T. A., dos Santos, V. J., & Calijuri, M. L. (2021). Fires dynamics in the Pantanal: Impacts of anthropogenic activities and climate change. Journal of Environmental Management, 299, 113586. https://doi.org/10.1016/j.jenvman.2021.113586

Miller, R.L., Schmidt, G.A., Nazarenko, L.S., Bauer, S.E., Kelley, M., Ruedy, R., Russell, G.L., Ackerman, A.S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T.L., Cook, B.I., Cruz, C.A., Del Genio, A.D., Elsaesser, G.S., Faluvegi, G., Kiang, N.Y., Kim, D., Lacis, A.A., Leboissetier, A., LeGrande, A.N., Lo, K.K., Marshall, J., Matthews, E.E., McDermid, S., Mezuman, K., Murray, L.T., Oinas, V., Orbe, C., Pérez García-Pando, C., Perlwitz, J.P., Puma, M.J., Rind, D., Romanou, A., Shindell, D.T., Sun, S., Tausnev, N., Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., Yao, M., 2021. CMIP6 Historical Simulations (1850–2014) With GISS-E2.1. J Adv Model Earth Syst 13. https://doi.org/10.1029/2019MS002034

Miranda, V.F.V. V., dos Santos, D.M., Peres, L.F., Salvador, C., Nieto, R., Müller, G. V., Thielen, D., Libonati, R., 2024. Heat stress in South America over the last four decades: a bioclimatic analysis. Theor Appl Climatol 155, 911–928. https://doi.org/10.1007/s00704-023-04668-x

Monteiro, A., Campelo, M.M., 2022. O Pará no foco dos estudos sobre populações negras da Amazônia Oriental. Afros & Amazônia 1.

Nacur, E.R., Vartuli, V., 2021. Flying rivers and climate change caused by the deforestation of the amazon rainforest: a perspective based on Latin American constitutionalism. Revista Brasileira de Direito Animal 16, 100–115.

Narvaez-Montoya, C., Mahlknecht, J., Torres-Martínez, J.A., Mora, A., Bertrand, G., 2023. Seawater intrusion pattern recognition supported by unsupervised learning: A systematic review and application. Science of the Total Environment. https://doi.org/10.1016/j.scitotenv.2022.160933

Nobre, A.D., 2014. The Future Climate of Amazonia Scientific Assessment Report, 1st ed. Articulación Regional Amazônica, São José dos Campos.

Nobre, C.A., Sampaio, G., Borma, L.S., Castilla-Rubio, J.C., Silva, J.S., Cardoso, M., 2016. 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 113, 10759–10768. https://doi.org/10.1073/pnas.1605516113

Nóbrega, R.S., 2014. Impactos do desmatamento e de mudanças climáticas nos recursos hídricos na Amazônia ocidental utilizando o modelo SLURP. Revista Brasileira de Meteorologia 29, 111–120. https://doi.org/10.1590/0102-778620130024

Pearce, F., 2019. Rivers in the sky. New Sci (1956) 244, 40–43. https://doi.org/10.1016/S0262-4079(19)32070-6

Petry, I., Jardim, P., Fan, F.M., 2021. Manual de aplicação ANA Data Acquisition V 1.0. Universidade Federal do Rio Grande do Sul 12.

Poveda, G., Jaramillo, L., Vallejo, L.F., 2014. Seasonal precipitation patterns along pathways of South American low-level jets and aerial rivers. Water Resour Res 50, 98–118. https://doi.org/10.1002/2013WR014087

Pratap, S., Markonis, Y., 2022. The response of the hydrological cycle to temperature changes in recent and distant climatic history. Prog Earth Planet Sci. https://doi.org/10.1186/s40645-022-00489-0

RezaAbbasifard, M., Ghahremani, B., Naderi, H., 2014. A Survey on Nearest Neighbor Search Methods. Int J Comput Appl 95, 39–52. https://doi.org/10.5120/16754-7073

Roque, F.V., Macarini, L.A., Crotti, Y., OliveiraWeber, T., Cechinel, C., 2019. Detecção de defeitos visuais em tecidos utilizando Wavelets e algoritmos de aprendizado de máquina, in: Anais Do Computer on the Beach. pp. 619–628.

Salati, E., Vose, P.B., 1984. Amazon Basin: A System in Equilibrium. Science (1979) 225, 129–138. https://doi.org/10.1126/science.225.4658.129

Schubert, E., Rousseeuw, P.J., 2019. Faster k-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms, in: Amato, G., Gennaro, C., Oria, V., Radovanović, M. (Eds.), Lecture Notes in Computer Science. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-32047-8

Shen, Z., Yong, B., Gourley, J.J., Qi, W., Lu, D., Liu, J., Ren, L., Hong, Y., Zhang, J., 2020. Recent global performance of the Climate Hazards group Infrared Precipitation (CHIRP) with Stations (CHIRPS). J Hydrol (Amst) 591, 125284. https://doi.org/10.1016/j.jhydrol.2020.125284

Smith, T., Boers, N., 2023. Global vegetation resilience linked to water availability and variability. Nat Commun 14, 498. https://doi.org/10.1038/s41467-023-36207-7

Souza, A. de, & Medeiros, E. S. de. (2024). Exploring climatic patterns among Brazilian capitals. Mercator, 23(2024), 1–15. https://doi.org/10.4215/rm2024.e23028

Strand, J., Soares-Filho, B., Costa, M.H., Oliveira, U., Ribeiro, S.C., Pires, G.F., Oliveira, A., Rajão, R., May, P., van der Hoff, R., Siikamäki, J., da Motta, R.S., Toman, M., 2018. Spatially explicit valuation of the Brazilian Amazon Forest's Ecosystem Services. Nat Sustain 1, 657–664. https://doi.org/10.1038/s41893-018-0175-0

Theeuwen, J.J.E., Staal, A., Tuinenburg, O.A., Hamelers, B.V.M., Dekker, S.C., 2023. Local moisture recycling across the globe. Hydrol Earth Syst Sci 27, 1457–1476. https://doi.org/10.5194/hess-27-1457-2023

Utida, G., Cruz, F. W., Vuille, M., Ampuero, A., Novello, V. F., Maksic, J., Sampaio, G., Cheng, H., Zhang, H., Dias de Andrade, F. R., & Edwards, R. L. (2023). Spatiotemporal Intertropical Convergence Zone dynamics during the last 3 millennia in northeastern Brazil and related impacts in modern human history. Climate of the Past, 19(10), 1975–1992. https://doi.org/10.5194/cp-19-1975-2023

van der Ent, R.J., Savenije, H.H.G., Schaefli, B., Steele-Dunne, S.C., 2010. Origin and fate of atmospheric moisture over continents. Water Resour Res 46. https://doi.org/10.1029/2010WR009127

Versieux, V., & Costa, M. H. (2024). Local Evapotranspiration Is the Only Relevant Source of Moisture at the Onset of the Rainy Season in South America. Atmosphere, 15(8), 932. https://doi.org/10.3390/atmos15080932

Walesiak, M., 2016. The Choice of Groups of Variable Normalization Methods in Multidimensional Scaling. Przegląd Statystyczny 63, 7–18. https://doi.org/10.5604/01.3001.0014.1145

Wang, D., Zhang, M., Fu, M., Cai, Z., Li, Z., Han, H., Cui, Y., Luo, B., 2016. Nonlinearity Mitigation Using a Machine Learning Detector Based on k-Nearest Neighbors. IEEE Photonics Technology Letters 28, 2102–2105. https://doi.org/10.1109/LPT.2016.2555857

Weng, W., Luedeke, M.K.B., Zemp, D.C., Lakes, T., Kropp, J.P., 2018. Aerial and surface rivers: downwind impacts on water availability from land use changes in Amazonia. Hydrol Earth Syst Sci 22, 911–927. https://doi.org/10.5194/hess-22-911-2018

Wilks, D.S., 2011. Cluster Analysis. pp. 603–616. https://doi.org/10.1016/B978-0-12-385022-5.00015-4

Wright, K., 2022. Will the Real Hopkins Statistic Please Stand Up? R J 14.

Xu, H., Lian, X., Slette, I. J., Yang, H., Zhang, Y., Chen, A., & Piao, S. (2022). Rising ecosystem water demand exacerbates the lengthening of tropical dry seasons. Nature Communications, 13(1), 4093. https://doi.org/10.1038/s41467-022-31826-y

Zemp, D.C., Schleussner, C.F., Barbosa, H.M.J., Hirota, M., Montade, V., Sampaio, G., Staal, A., Wang-Erlandsson, L., Rammig, A., 2017. Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks. Nat Commun 8. https://doi.org/10.1038/ncomms14681

Zemp, D.C., Schleussner, C.F., Barbosa, H.M.J., Rammig, A., 2014. Exploring the complex network of the Amazon's water pump and flying rivers.

Zhang, L., Dawes, W.R., Walker, G.R., 2001. Response of mean annual evapotranspiration to vegetation changes at catchment scale. Water Resour Res 37, 701–708. https://doi.org/10.1029/2000WR900325

Zheng, J., Jin, X., Li, Q., Lang, J., Yin, X., 2024. Soil moisture variation and affecting factors analysis in the Zhangjiakou–Chengde district based on modified temperature vegetation dryness index. Ecol Indic 168, 112775. https://doi.org/10.1016/j.ecolind.2024.112775

Zhou, K., Shi, X., 2024. Understanding Precipitation Moisture Sources and Their Dominant Factors During Droughts in the Vietnamese Mekong Delta. Water Resour Res 60. https://doi.org/10.1029/2023WR035920

6. CAPÍTULO III. From Drought to Flood: The Role of SACZ and Land-Use Change in Shaping Extreme Rainfall Patterns in Minas Gerais

Abstract: This study analyzes the spatial-temporal variability of extreme precipitation events associated with the South Atlantic Convergence Zone (SACZ) across the state of Minas Gerais, Brazil, over the 2004–2024 period. By integrating satellite-based rainfall datasets (CHIRPS), historical precipitation records, and land cover classifications, the research provides a comprehensive assessment of how SACZ activity has influenced regional hydroclimatic dynamics, with particular emphasis on the interplay between extreme rainfall, land-use changes, and socio-environmental vulnerability. The results reveal a pronounced spatial concentration of SACZ-related anomalies along a northeast–southwest corridor, notably affecting mesoregions such as Zona da Mata, Vale do Rio Doce, Jequitinhonha, Triângulo Mineiro/Alto Paranaíba, and the Metropolitan Region of Belo Horizonte. These regions are characterized by intense anthropogenic land use, including urban expansion, deforestation, and agricultural intensification, which have diminished hydrological buffering capacity and increased exposure to flood events, landslides, and infrastructure collapse. The study also identifies strong interannual variability in SACZ behavior linked to large-scale climate modes, especially ENSO phases. La Niña episodes correspond to a reduction in SACZ intensity and prolonged droughts (e.g., 2007, 2012, 2017), while El Niño phases (e.g., 2008, 2011, 2016) enhance convective rainfall and extreme precipitation. Additionally, high-pressure anomalies and atmospheric blocking patterns exacerbate these conditions by modulating moisture transport. Notably, mesoregions with greater preservation of natural vegetation, such as portions of Norte de Minas, demonstrated more resilient hydrological responses, highlighting the role of land cover in mitigating climate extremes. The findings underscore the urgent need for integrated climate adaptation strategies that combine early warning systems, sustainable land management, and investments in resilient infrastructure. This study contributes to a growing body of knowledge supporting regional planning and disaster risk reduction in the face of escalating climate variability across southeastern Brazil.

Keywords: South Atlantic Convergence Zone (SACZ), Extreme rainfall events, Anthropogenic land use, Hydroclimatic impacts, Remote Sensing, Hydrometeorological Monitoring, Water Resource Management.

6.1. Introduction

The South Atlantic Convergence Zone (SACZ) is a fundamental meteorological system that exerts a significant influence on the climate of South America, particularly southeastern Brazil. It is characterized as a quasi-stationary convective band associated with the convergence of moisture-laden air masses from the Amazon Basin and the South Atlantic Ocean, resulting in prolonged and intense rainfall during the austral summer (Silva et al., 2019). The variability of SACZ activity plays a crucial role in modulating regional hydrological regimes, impacting water resource availability, agricultural productivity, and the occurrence of extreme weather events such as floods and droughts (Filho et al., 2022). However, despite its importance, the spatiotemporal characterization of SACZ remains a challenge due to the dynamic interactions between atmospheric circulation patterns, large-scale climate variability, and moisture transport mechanisms.

A critical component of SACZ-related precipitation is the "Flying Rivers" phenomenon, which describes the transport of Amazonian moisture to southeastern Brazil through low-level jet streams, particularly the South American Low-Level Jet (SALLJ) (Amaral e Silva et al., 2024). This moisture transport process is vital for sustaining precipitation patterns in Minas Gerais and surrounding areas, as it enhances convective activity and reinforces the persistence of SACZ events (Silva et al., 2019; Amaral e Silva et al., 2024).

However, deforestation and land-use changes in the Amazon have been linked to disruptions in this moisture transport mechanism, leading to modifications in rainfall distribution and an increased frequency of extreme hydrometeorological events, including severe droughts and catastrophic flooding (Costa et al., 2024; Correia et al., 2024). Given these concerns, an improved understanding of the relationship between SACZ activity, moisture transport processes, and external anthropogenic influences is essential for refining climate models and informing sustainable adaptation strategies.

The El Niño-Southern Oscillation (ENSO) also plays a significant role in SACZ variability, modulating seasonal precipitation trends across southeastern Brazil. Previous studies have shown that El Niño conditions tend to enhance SACZ activity, leading to increased precipitation and a higher frequency of extreme rainfall events (Barreiro, 2024; Pedreira Junior et al., 2020). Conversely, La Niña phases are typically associated with weakened SACZ intensity and below-average precipitation, exacerbating drought conditions (Verdan & Silva, 2022). Additionally, the interaction between ENSO and other large-scale climate oscillations, such as the Atlantic Multidecadal Oscillation (AMO) and the South Atlantic Subtropical High (SASH), further complicates precipitation predictability in the region (Marengo et al., 2020; Maia et al., 2022). These climatic interactions underscore the necessity of high-resolution monitoring systems that can capture the complex drivers of SACZ behavior and their hydrological implications.

Recent advancements in remote sensing and satellite-based precipitation datasets have provided new opportunities for analyzing SACZ dynamics and their associated impacts (Saddique et al., 2022). Studies have demonstrated the efficacy of gridded precipitation products, such as CHIRPS (Climate Hazards Group InfraRed Precipitation with Stations), PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks), and TerraClimate, in capturing precipitation variability across South America (Arias & Barriga, 2022; Du et al., 2024).

These datasets, when integrated with ground-based rainfall measurements, allow for a more comprehensive spatiotemporal assessment of SACZ-driven precipitation patterns and their long-term climatological trends. However, challenges remain in the validation and calibration of satellite-derived precipitation estimates, particularly in regions with complex topography and highly variable rainfall regimes (López-Bermeo et al., 2022; Nadeem et al., 2022).

While CHIRPS has been found to perform well in humid environments, it has shown limitations in accurately representing rainfall totals in semi-arid regions, necessitating bias correction techniques to improve its reliability (Araghi & Adamowski, 2024). Similarly, PERSIANN datasets, despite their ability to detect extreme rainfall events, require post-processing adjustments to align with in situ observations (Freitas et

al., 2022). Integrating multiple precipitation sources through advanced data assimilation techniques offers a promising avenue for enhancing the accuracy and applicability of SACZ rainfall estimates in climate risk assessments.

The growing availability of Big Data analytics in climatology has further transformed the study of atmospheric processes, enabling the processing of vast meteorological datasets with unprecedented precision. Machine learning algorithms and cloud computing frameworks facilitate the identification of complex spatiotemporal patterns in precipitation trends, improving real-time monitoring and predictive capabilities for SACZ-related events (Sondermann et al., 2022; Elmahal & Musa, 2023). This technological advancement has significant implications for water resource management, disaster risk reduction, and climate adaptation planning in southeastern Brazil. By leveraging large-scale satellite datasets and advanced computational techniques, researchers can enhance the forecasting of extreme rainfall and drought events, ultimately strengthening regional climate resilience (Funk et al., 2015; RezaAbbasifard et al., 2014).

In this context, this study aims to bridge a critical knowledge gap by integrating satellite remote sensing, Big Data analytics, and land coverage observations to refine the characterization of SACZ variability in Minas Gerais. By investigating the socio-environmental impacts of SACZ dynamics, this research will contribute to a deeper understanding of precipitation variability and its implications for climate risk management.

Furthermore, the study explored the interrelationship between SACZ precipitation patterns and other atmospheric events, in a 21-years' time series, shedding light on how atmospheric moisture transport processes influence regional hydrology. As climate change continues to disrupt atmospheric circulation patterns, developing robust data-driven approaches for climate monitoring and adaptation becomes increasingly imperative. This research represents a significant step toward improving the predictive capabilities of SACZ-related precipitation models and fostering an evidence-based approach to climate resilience in southeastern Brazil.

6.2. Methodology

The methodology for developing the research in question was divided into five stages: study area definition, database acquisition, data pre-processing, database processing (rainfall data evaluation) and results interpretation (Figure 6.1).

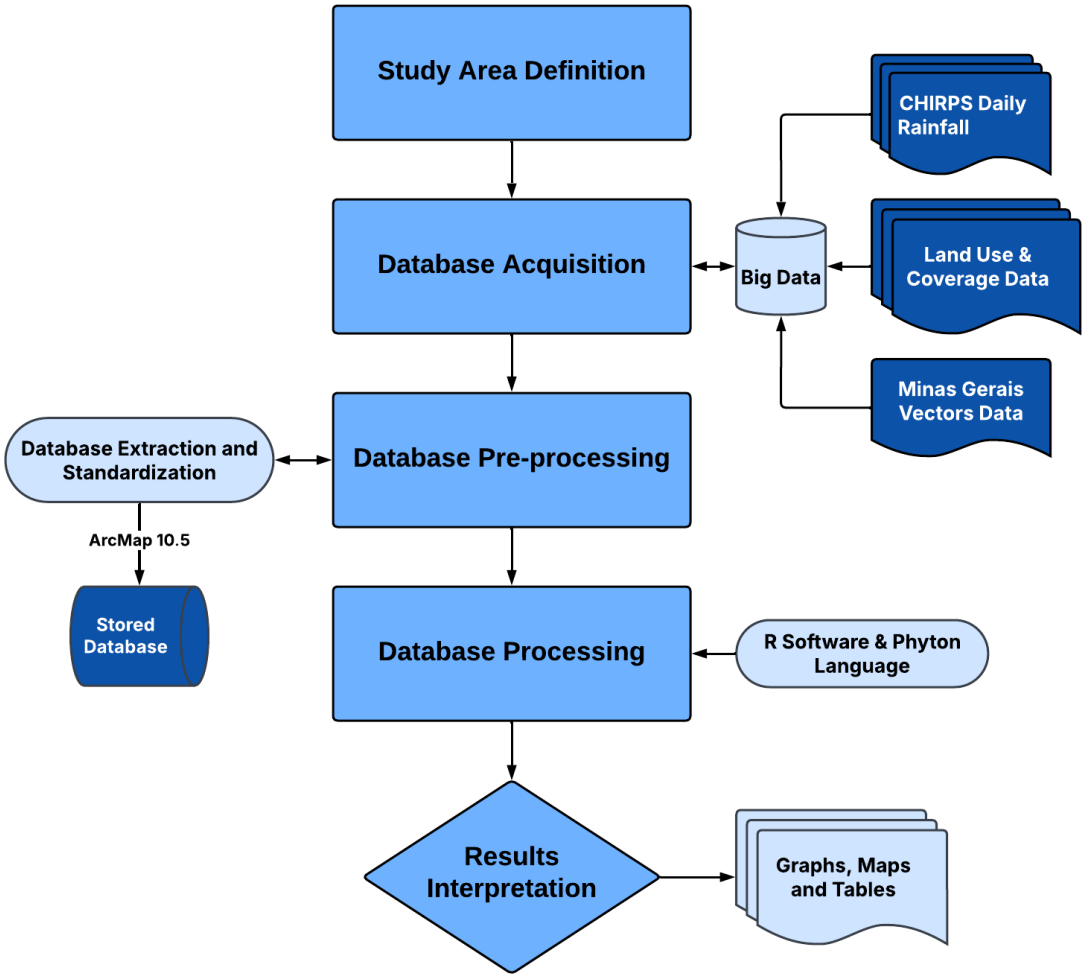

Figure 6.1. Methodological flowchart.

6.2.1. Study Area Definition

The study area encompasses Minas Gerais, a vast and ecologically diverse state in southeastern Brazil (Figure 6.2). This area serves as a crucial region for studying rainfall due to its varied physiography, hydrological significance, and substantial population. Its topography is characterized by mountain ranges, plateaus, and valleys, which significantly influence local and regional climate patterns (Aquino et al., 2012). One of its most important

hydrological features is the São Francisco River, which originates in the Serra da Canastra Mountain range at an elevation of approximately 730 meters above sea level (Correia Filho et al., 2022). This river extends over 2,914 kilometers and plays a vital role in supplying water for agriculture, industry, and hydroelectric power generation. The São Francisco River Basin, covers about 639,219 square kilometers, encompasses 37% of Minas Gerais' territory, further emphasizing the state's importance in Brazil's water cycle (Oliveira et al., 2023).

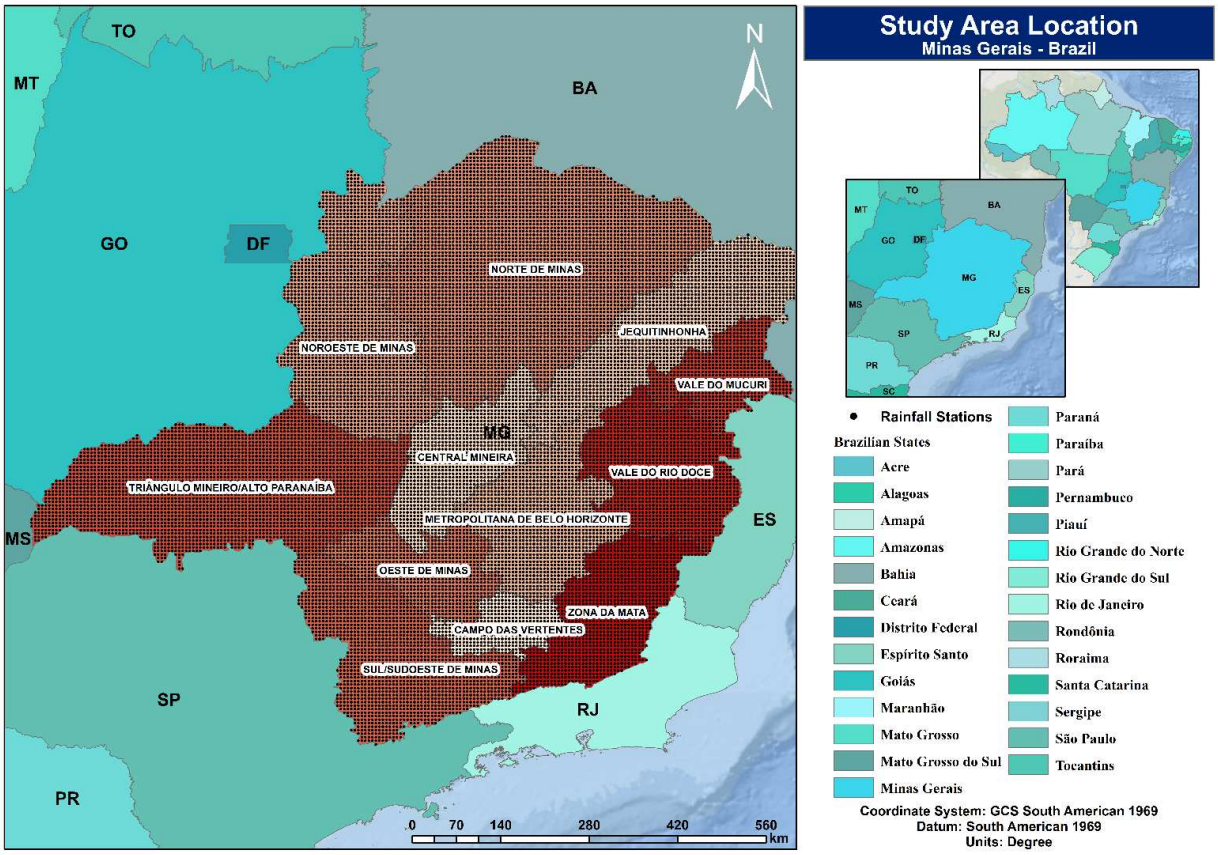

Figure 6.2. Study area Location.

Minas Gerais exhibits significant climatic variation due to its diverse topography. The state experiences a range of climatic conditions, from tropical in lower elevations to temperate in higher altitudes (Sondermann et al., 2022). Annual rainfall varies considerably, from approximately 600 millimeters in semi-arid regions to around 1,400 millimeters in wetter areas such as the Serra da Canastra. The rainy season generally extends from October to March and is heavily influenced by the South Atlantic Convergence Zone (SACZ), which plays a central role in modulating precipitation patterns

(Freitas et al., 2022). Studies have revealed that rainfall distribution in Minas Gerais is highly heterogeneous, with precipitation zones forming distinct clusters across the region. This spatial variability is crucial for understanding water availability, managing agriculture, and predicting hydrological responses to climate change and extreme weather events.

The vegetation of Minas Gerais is diverse, reflecting the state's varied climate and topography. The Cerrado biome dominates the central and western regions, characterized by savanna-like vegetation with deep-rooted plants adapted to seasonal rainfall variations (Nunes et al., 2022). In contrast, the Atlantic Forest biome, which once covered vast areas of the state, is now fragmented due to urban expansion, agriculture, and deforestation. This biome is one of the most biodiverse in the world and plays a crucial role in water regulation, as its dense vegetation helps maintain soil moisture and reduce runoff (Lembi et al., 2020). The northern parts of the state transition into the Caatinga biome, a semi-arid ecosystem with drought-resistant vegetation adapted to prolonged dry periods.

Water consumption in Minas Gerais is highly dependent on rainfall distribution and the availability of surface and groundwater resources. The state's water demand is driven primarily by agriculture, which accounts for approximately 70% of total water usage, followed by domestic consumption (15%) and industrial activities (10%) (Carvalho et al., 2024). Irrigation is particularly intensive in the western and northern regions of the state, where rainfall is lower, and water resources must be carefully managed. The São Francisco River and its tributaries provide crucial water for agricultural production, which includes crops such as coffee, soybeans, and sugarcane. However, prolonged droughts and increasing water demand have led to periodic water shortages, highlighting the need for sustainable water management practices, improved irrigation techniques, and investments in water storage infrastructure (Silva et al., 2018). Groundwater resources also play a significant role in supplying water for both urban and rural areas, but over-extraction has led to concerns about aquifer depletion and water quality deterioration (Sajjad et al., 2023).

Minas Gerais' diverse landscape, climatic variability, and hydrological importance make it a vital region for rainfall studies in Brazil. The state's role in national water security

and energy production underscores the importance of continuous research to better understand precipitation patterns and mitigate the impacts of climate change. By integrating scientific knowledge with policy and management strategies, Minas Gerais can serve as a model for sustainable water resource management in regions facing similar environmental and socio-economic challenges.

6.2.2. Database Acquisition

To carry out this research, the database shown in Table 6.1 was used.

Table 6.1. Key characteristics and sources of the database

Data	Source	Period	Description
Daily Rainfall	Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS)	2004 - 2024	CHIRPS v2.0 is a high-resolution, quasi-global rainfall dataset covering latitudes 50°S to 50°N from 1981 to the present. It combines 0.05° resolution satellite imagery with in-situ station data to create gridded rainfall time series, aiding in trend analysis and drought monitoring.
Minas Gerais Mesoregions	Brazilian Institute of Geography and Statistics (IBGE)	2023	Subdivision of the State of Minas Gerais that brings together several municipalities in a geographic area with economic and social similarities.
Land Use and Occupation	MapBiomas Brazil	2003 - 2023	MapBiomas operates collaboratively, bringing together institutions specialized in different biomes and cross-cutting themes to generate data on land use.
Climatic Data	National Institute of Meteorology	-	Monitoring and forecasting of weather and climate, issuance of severe weather warnings, and the development of various products to support agriculture, including research and innovation.

Regarding rainfall data, it was sourced from the Climate Hazards Center through the Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS), a widely utilized dataset in climate studies (Funk et al., 2015; Paredes-Trejo et al., 2017; Du et al., 2024). The study utilized daily rainfall data spanning from 2004 to 2024, covering a total of 21 years and comprising information from 7,670 days.

The vector data for the study area encompassed both the region's boundaries and the subdivision of its mesoregions. Minas Gerais is traditionally divided into 12 mesoregions, which group municipalities based on similar socio-economic and physical characteristics (Caetano et al., 2017). The study of mesoregions is particularly important

for assessing water availability, agricultural conditions, and economic opportunities across Minas Gerais. Different regions experience distinct challenges, with some areas benefiting from abundant water resources while others face persistent scarcity (Rodrigues et al., 2020). For example, the Triângulo Mineiro region enjoys ample water supply, facilitating large-scale farming, whereas the Jequitinhonha and Norte de Minas regions struggle with drought and limited water access (Correia et al., 2024). This variability underscores the importance of mesoregional analysis as a tool for researchers and policymakers aiming to enhance environmental sustainability, optimize water management, and promote balanced regional development strategies throughout the state.

Regarding land use and land cover data, it was obtained from the MapBiomas platform (https://brasil.mapbiomas.org/). This platform is a crucial tool for environmental monitoring and land-use analysis in Brazil. By integrating big data, artificial intelligence, and remote sensing technologies, MapBiomas enables comprehensive tracking of ecological transformations, supporting conservation efforts and informing public policies (Souza et al., 2020). For this study, land use and cover transition data from 2004 to 2023, the latest year available on the platform, were analyzed. This dataset was used to identify land-use variations across the state of Minas Gerais and the Amazon Rainforest, and to assess their correlation with the impacts of extreme rainfall events observed in this research.

Finally, climate data from the National Institute of Meteorology (INMET), specifically the Standardized Precipitation Index (SPI), were used not only to validate the information obtained after processing but also to enable a visual analysis of climate behavior in the region throughout the study period. INMET plays a crucial role in promoting and coordinating activities focused on generating accurate and relevant meteorological data. Its mission is to mitigate climate risks, support the sustainable development of the agricultural sector, enhance environmental conservation, and ensure the safety of Brazilian society (Musah et al., 2022).

6.2.3. Data base pre-processing and processing

Once the database was acquired, the next step involved pre-processing, which consists of organizing and standardizing the dataset before proceeding with further analyses. This phase was carried out using ArcMap 10.5, RStudio, and Excel 2019, each serving specific functions in data processing and refinement.

The pre-processing stage began with the development of a routine algorithm in RStudio and Phyton to automate key procedures, including data acquisition, filtering of relevant information, and graphical representation of results. The first step involved designing an algorithm to access the CHIRPS v2.0 online database through the links provided by the platform. This algorithm facilitated the download of 7,670 files, corresponding to each day in the studied time series. Consequently, satellite imagery was obtained for the period spanning January 1, 2004, to December 31, 2024, marking the beginning of the Big Data construction process.

Following data acquisition, a second script was implemented to extract the study area using a vector mask and reproject the dataset. For this step, the terra and sf packages in R were employed, ensuring that the raster datasets were clipped to the region of interest and aligned with the appropriate coordinate reference system (CRS) (Bivand, 2022). This transformation was performed using the project() function from the terra package, which converted the raster to a well-defined spatial reference system (Elmahal and Musa, 2023). Reprojection is a critical step, as it ensures spatial consistency, precise distance measurements, and seamless integration of datasets from various sources, thus improving the reliability of subsequent analyses (Nogueira, 2024).

The next step involved defining the precipitation data collection points. Since rain gauge stations do not provide full coverage across the state of Minas Gerais, a grid of points was generated based on the centroids of the pixels encompassing the state, resulting in a total of 18,651 data collection points. This approach enabled a spatial-temporal assessment of the occurrence patterns of SACZ events. This process is particularly crucial for regional studies and hydrological modeling, where accurate spatial representation is essential.

The following phase involved extracting information from the rainfall raster datasets. To achieve this, the Nearest Neighbor Search (NNS) technique was applied to extract satellite-derived values. NNS is a widely used machine learning and data analysis method that identifies patterns based on the proximity of data points within a dataset, such as rainfall stations. This method is particularly effective for classification and regression tasks, especially when dealing with high-dimensional spaces or datasets where prior assumptions about data distribution are not feasible (Roque et al., 2019).

Efficient NNS data structures are typically constructed using two primary approaches: indexing and sketching. Indexing involves structuring the dataset to generate a small candidate set (P) of potential nearest neighbors for any given query point (RezaAbbasifard et al., 2014). On the other hand, sketching focuses on computing compressed representations of data points, allowing for the rapid approximation of distances. These two methods can be integrated to optimize computational performance and enhance data retrieval efficiency (Wang et al., 2016).

In this study, the Extract Values to Points tool in ArcMap 10.5 was employed to extract raster values based on the precipitation data collection points. The extracted data was then stored in csv. archives, enabling further analyses. This approach ensured a robust and methodologically extraction process, enhancing the reliability of rainfall data for subsequent hydrological and climatological studies.

With the dataset compiled, the next step involved identifying South Atlantic Convergence Zone (SACZ) events. According to Carvalho et al. (2004), SACZ is characterized by a persistent band of cloudiness and precipitation oriented in a northwest-southeast direction, extending from the southern and southeastern Amazon region to the southwestern South Atlantic Ocean. This meteorological system is a major driver of prolonged rainfall periods, often leading to extreme precipitation events that can trigger floods and landslides (Seluchi & Chou, 2009; Reboita et al., 2017).

To systematically identify significant SACZ extreme events, the criteria established by Cardoso et al. (2020) regarding intensity and persistence were applied. According to their definition, Persistent Rainfall (PR) is characterized by the occurrence of daily

accumulated exceeding 100 mm for at least four consecutive days within the dataset. By applying this criterion, it was possible to detect and analyze the most significant SACZ-related extreme rainfall events during the study period.

6.2.4. Results Interpretation

After obtaining all post-processed data, the results were analyzed and interpreted through the construction of graphs, maps, and tables to facilitate visualization and comprehension. The graphical representations were developed in Phyton environment and RStudio using the ggplot2 package, which is widely recognized for its efficiency in data visualization. The ggplot2 package, developed by Hadley Wickham, is based on the Grammar of Graphics framework, which allows for a structured and layered approach to plot construction (Wickham, 2016). This methodology enables the systematic addition of graphical elements, improving the clarity and interpretability of complex datasets.

The ArcMap 10.5 software was also employed for mapping the obtained results, enabling a spatial analysis of the information. This geospatial approach allowed for a more precise interpretation of the dataset, particularly in identifying regional patterns and variations in precipitation events.

Additionally, tables constructed in Excel 2019, in conjunction with data processing in RStudio, facilitated data filtering and the identification of extreme events. The integration of these tools ensured a comprehensive assessment of the dataset, improving the accuracy of event classification and trend analysis.

In this context, all results were analyzed collectively, allowing for the identification of all SACZ events that occurred within the studied time series, as well as those that exhibited extreme characteristics. This combined approach provided a robust framework for evaluating the impact and persistence of SACZ-related precipitation patterns.

6.3. Results

Based on the applied methodology, it was possible to obtain relevant information on the interrelationship between deforestation in the Amazon biome and rainfall formation in the state of Minas Gerais. In addition, the approach enabled the analysis of the

spatiotemporal dynamics of the South Atlantic Convergence Zone (SACZ) in the state, as well as its socio-economic and environmental impacts.

6.3.1. Amazon Vegetation and Rainfall in MG (2003–2024)

Figure 6.3 offers a comprehensive visualization of the interdependent dynamics between deforestation in the Amazon biome and annual precipitation patterns in the state of Minas Gerais.

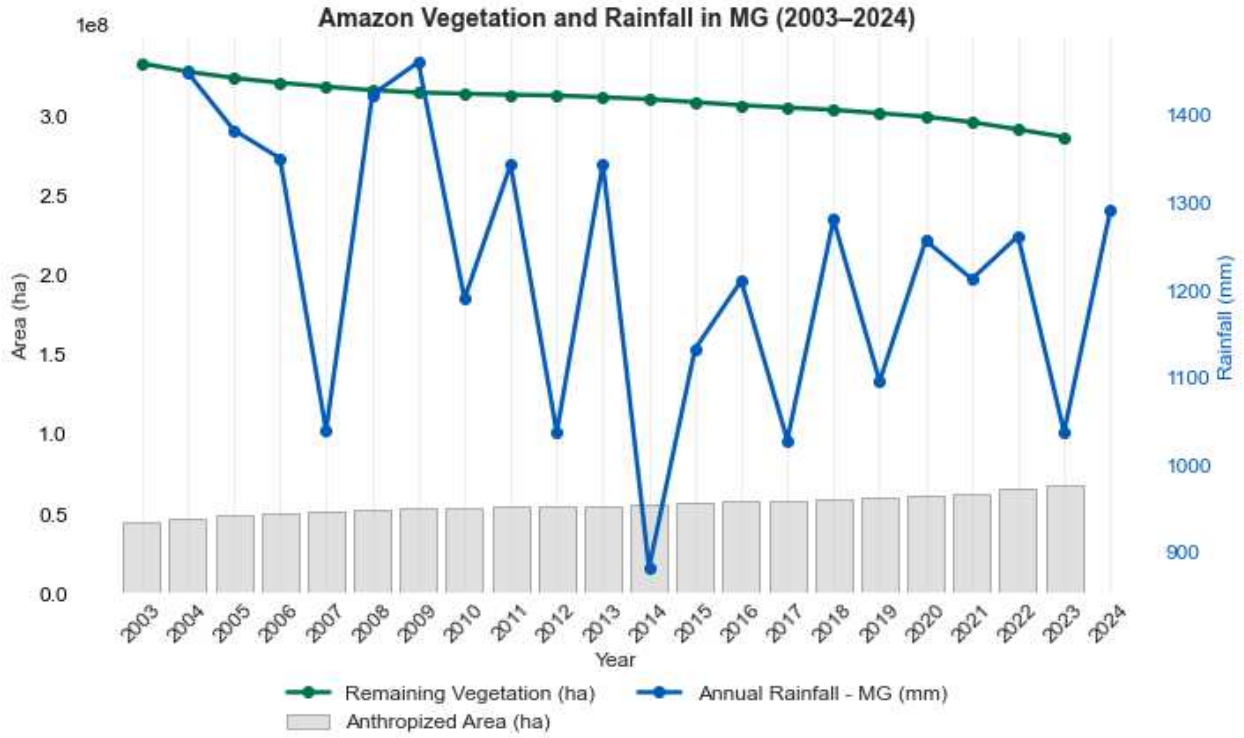

Figure 6.3. Amazon Vegetation and Rainfall in MG (2003–2024).

Three key variables are represented: the remaining native vegetation in the Amazon (green line), the annual rainfall observed in Minas Gerais (blue line), and the extent of anthropized—i.e., deforested—areas within the Amazon biome (grey bars). The temporal alignment of these variables enables an exploratory assessment of how land-use change in the Amazon influences hydrometeorological conditions in southeastern Brazil.

Throughout the two-decade timespan analyzed, there is a clear and progressive decline in the remaining Amazonian vegetation cover, highlighting the persistent advance

of deforestation. This decline is mirrored by a concomitant increase in the anthropized area, with more pronounced surges occurring after 2010. The annual rainfall series in Minas Gerais, in turn, exhibits high interannual variability, yet certain patterns suggest a delayed response to forest loss in the Amazon. Specifically, years of intensified deforestation are frequently followed by significant reductions in rainfall during the subsequent year. Notable examples include the periods 2006–2007, 2013–2014, 2020–2021 and 2022–2023. The year 2014, which registers the lowest rainfall in the entire series, follows a peak in deforestation in 2013, coinciding with a severe drought historically recorded in southeastern Brazil.

This temporal lag reinforces the hypothesis of a causative link mediated by atmospheric moisture recycling mechanisms. The removal of forest cover reduces evapotranspiration and disrupts the flow of humid air masses—phenomena often referred to as “flying rivers”—which are responsible for transporting moisture from the Amazon Basin to central and southeastern Brazil. The data suggest that large-scale deforestation compromises this mechanism, leading to decreased rainfall and potentially exacerbating water scarcity and socio-environmental vulnerabilities in downstream regions.

Interestingly, periods of relative stabilization or deceleration in deforestation—such as those preceding 2010 and 2024—appear to be associated with partial recovery in precipitation levels. Although this does not establish causality, it points toward a potential for hydrological resilience when forest degradation is mitigated, further underscoring the role of the Amazon biome in regulating regional climate systems.

In conclusion, the graph presents strong indicative evidence of a spatiotemporal teleconnection between Amazonian deforestation and rainfall variability in Minas Gerais. The observed lagged effects align with theoretical and empirical studies on biosphere-atmosphere interactions, reinforcing the urgency of forest conservation as a strategy not only for biodiversity protection but also for maintaining climatic and hydrological stability across Brazil. These findings highlight the need for integrated, cross-regional environmental policies that address the systemic impacts of deforestation beyond local or biome-specific scales.

6.3.2. Spatiotemporal analysis of the impacts of SACZ events in Minas Gerais

The spatial distribution of extreme rainfall events associated with the South Atlantic Convergence Zone across the state of Minas Gerais, based on a consolidated rainfall database from the hydrological years 2004 to 2023, reveals a highly uneven pattern of occurrence (Figure 6.4). These events were concentrated in specific mesoregions, particularly where geomorphological vulnerability intersects with intense land use and occupation.

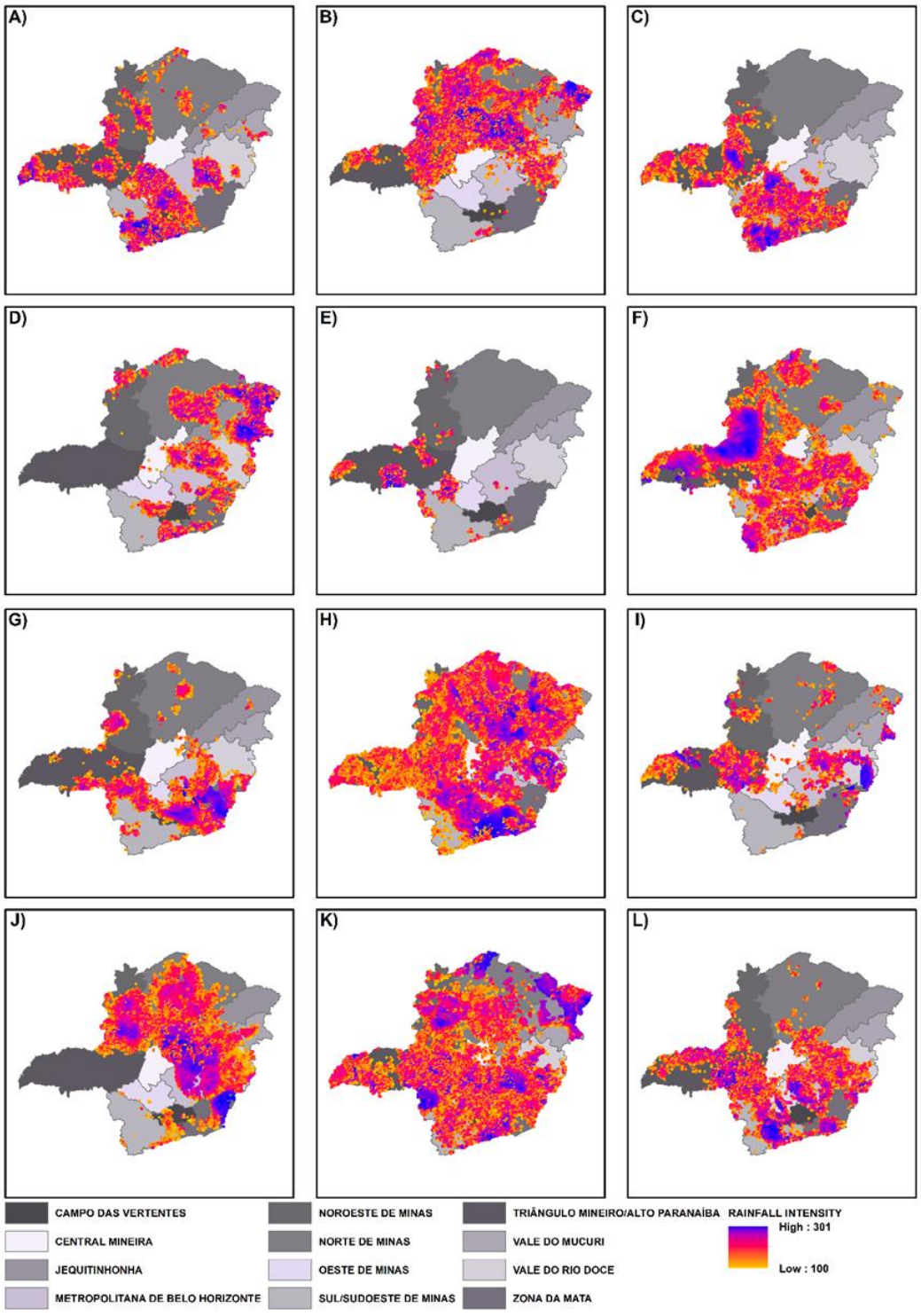

Figure 6.4. Spatial distribution of extreme SACZ-related rainfall events across the mesoregions of Minas Gerais during the following hydrological years: A) 2004–2005; B) 2005–2006; C) 2006–2007; D) 2008–2009; E) 2009–2010; F) 2010–2011; G) 2011–2012; H) 2015–2016; I) 2017–2018; J) 2019–2020; K) 2021–2022; L) 2022–2023.

The regions most affected in terms of frequency and rainfall magnitude were the Zona da Mata, Vale do Rio Doce, Jequitinhonha, Triângulo Mineiro/Alto Paranaíba, and the Metropolitan Region of Belo Horizonte. Most events occurred between November and March, following the seasonal activity of the SACZ. Spatially, the concentration of extreme events forms a northeast–southwest arc across the state, particularly along mountainous corridors and densely inhabited urban zones.

In the Zona da Mata, the convergence of topographic complexity and dense human occupation has amplified the impacts of SACZ events over the decades. Historical records confirm that cities such as Cataguases, Muriaé, Ponte Nova, and Juiz de Fora experienced some of the most catastrophic flood events in Minas Gerais. For instance, in January 2005, the Pomba River overflowed in Cataguases, displacing over 400 residents. These impacts align with rainfall events in the dataset exceeding 120 mm in a single day, confirming the SACZ's role in triggering such disasters. The presence of urban infrastructure along riverbanks and steep slopes has significantly increased hydrological sensitivity to these convective systems.

In the Vale do Rio Doce, several SACZ events caused severe socio-environmental damage, particularly in Coronel Fabriciano and Ipatinga. In January 2020, continuous rainfall led to a major flood of the Piracicaba River, resulting in over 500 people being displaced. The rainfall records for this region reveal multiple occurrences with accumulations above 150 mm, often within urbanized catchments. The classification of this region as predominantly "antropic" in land-use mapping reflects widespread urbanization and deforestation, which reduce infiltration and exacerbate surface runoff and landslide risk.

The Jequitinhonha mesoregion exhibited the highest cumulative rainfall values in the dataset, surpassing one million millimeters over the 20-year period. Towns like Padre Paraíso, Araçuaí, and Monte Formoso experienced floods, infrastructure collapse, and mass displacements. In January 2022 alone, rainfall in the region led to the overflow of the Jucuruçu River and the collapse of a dam in Crisolita. Several of these municipalities appear in the historical reports of emergency declarations and federal assistance. Notably,

this region faces a dual vulnerability: frequent SACZ activity and high socio-economic fragility, which limits disaster response capacity.

The Metropolitan Region of Belo Horizonte is a critical hotspot for SACZ-induced disasters. In January 2020, the city recorded over 935 mm of rainfall—its highest monthly total since the beginning of monitoring—resulting in over 70 fatalities and thousands of displaced people. In just 24 hours (23–24 January), 171.8 mm of rain fell, overwhelming the city's drainage systems. These volumes are well reflected in the rainfall database, confirming the role of SACZ events in overwhelming infrastructure in highly urbanized settings.

The Triângulo Mineiro/Alto Paranaíba mesoregion exhibited substantial exposure to SACZ-driven rainfall extremes throughout the analyzed period. Municipalities such as Uberaba, Uberlândia, Patos de Minas, and Monte Carmelo were repeatedly affected by high-impact hydrometeorological events. A notable example occurred in March 2024, when Monte Carmelo recorded over 90 mm of rainfall within just 12 hours, triggering flash floods and disrupting urban mobility. These intense precipitation episodes align with pronounced spikes in the SACZ event database, reflecting the mesoregion’s positioning along moisture convergence corridors. While the region does not exhibit the same topographic steepness as others like Zona da Mata or Jequitinhonha, its rapid urban expansion, soil compaction from agriculture, and significant loss of vegetative cover have critically diminished its hydrological buffering capacity.

Land cover analysis reinforces this trend. The natural vegetation area in Triângulo Mineiro/Alto Paranaíba has declined steadily over the last two decades, while anthropic land use - primarily linked to intensive agriculture, pasture, and urban development - has increased. This shift amplifies surface runoff, accelerates drainage system saturation, and increases the vulnerability of both urban and rural communities to SACZ-triggered flash flooding. Moreover, many urban centers in the region have expanded over former floodplains or in proximity to drainage channels, compounding the risk of urban inundation even under moderate rainfall intensities. The recurrence of these events highlights the need for stronger integration between urban planning, watershed management, and

climate risk mitigation strategies in this economically strategic yet environmentally sensitive mesoregion.

The integration of the rainfall database with the mesoregion reference and the historical disaster narratives makes evident the spatial signature of the SACZ in Minas Gerais. Not only does the spatial distribution of SACZ events align with areas of high rainfall accumulation and recurring flood impacts, but there is also a clear relationship with anthropogenic land use. Mesoregions classified as predominantly anthropic consistently show both higher frequency of extreme events and significantly worse impacts. For example, the Central Mineira region recorded events with more than 200 mm, and recurrent overflow of urban creeks in Curvelo and Bom Despacho, despite moderate rainfall levels.

This relationship is further supported by the analysis of land cover dynamics illustrated in Figures 6.5 and 6.6, which depict the annual variation in anthropic and natural land use across all mesoregions from 2004 to 2023. The figures reveal a consistent trend of increasing anthropic area—characterized by urban expansion and agricultural use—at the expense of natural vegetation. In mesoregions such as Central Mineira, Metropolitana de Belo Horizonte, Zona da Mata, and Vale do Rio Doce, the reduction in natural cover was particularly pronounced, with the anthropic area exceeding natural vegetation from the earliest years in the series and continuing to grow steadily. This land transformation trend aligns with the spatial clusters of extreme SACZ events and correlates with greater socio-environmental impacts observed in these regions. Conversely, mesoregions like Vale do Mucuri, Noroeste de Minas, and Norte de Minas, where natural cover remains proportionally higher and relatively stable, exhibit a lower frequency and intensity of SACZ-related disasters. These areas demonstrate greater hydrological resilience, even when exposed to high rainfall, highlighting the buffering effect of preserved ecosystems.

Taken together, these results emphasize that the intensity and recurrence of SACZ impacts in Minas Gerais are not solely determined by atmospheric forcing but are strongly modulated by land cover patterns and territorial occupation. The progressive anthropization of landscapes not only reduces natural water absorption and amplifies

surface runoff but also places human settlements in increasingly vulnerable positions. This reinforces the need for integrated land use planning, ecosystem preservation, and climate adaptation policies that consider both climatic and territorial drivers of risk.

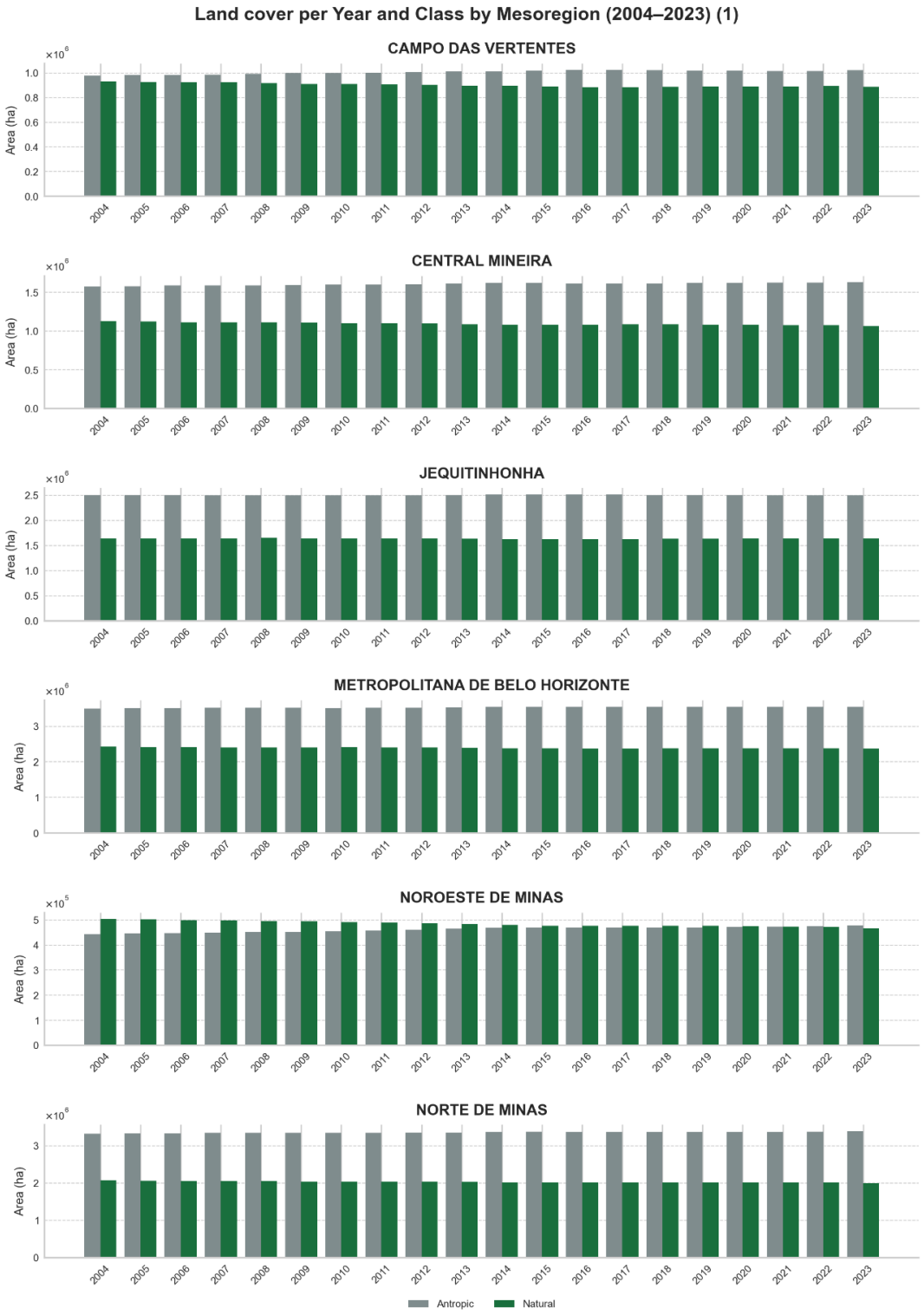

Figure 6.5. Annual land cover evolution (2004–2023) by class for the mesoregions of Campo das Vertentes, Central Mineira, Jequitinhonha, Metropolitana de Belo Horizonte, Noroeste de Minas, and Norte de Minas. Bars represent the total area (ha) classified as 'Anthropic' (gray) and 'Natural' (green) each year.

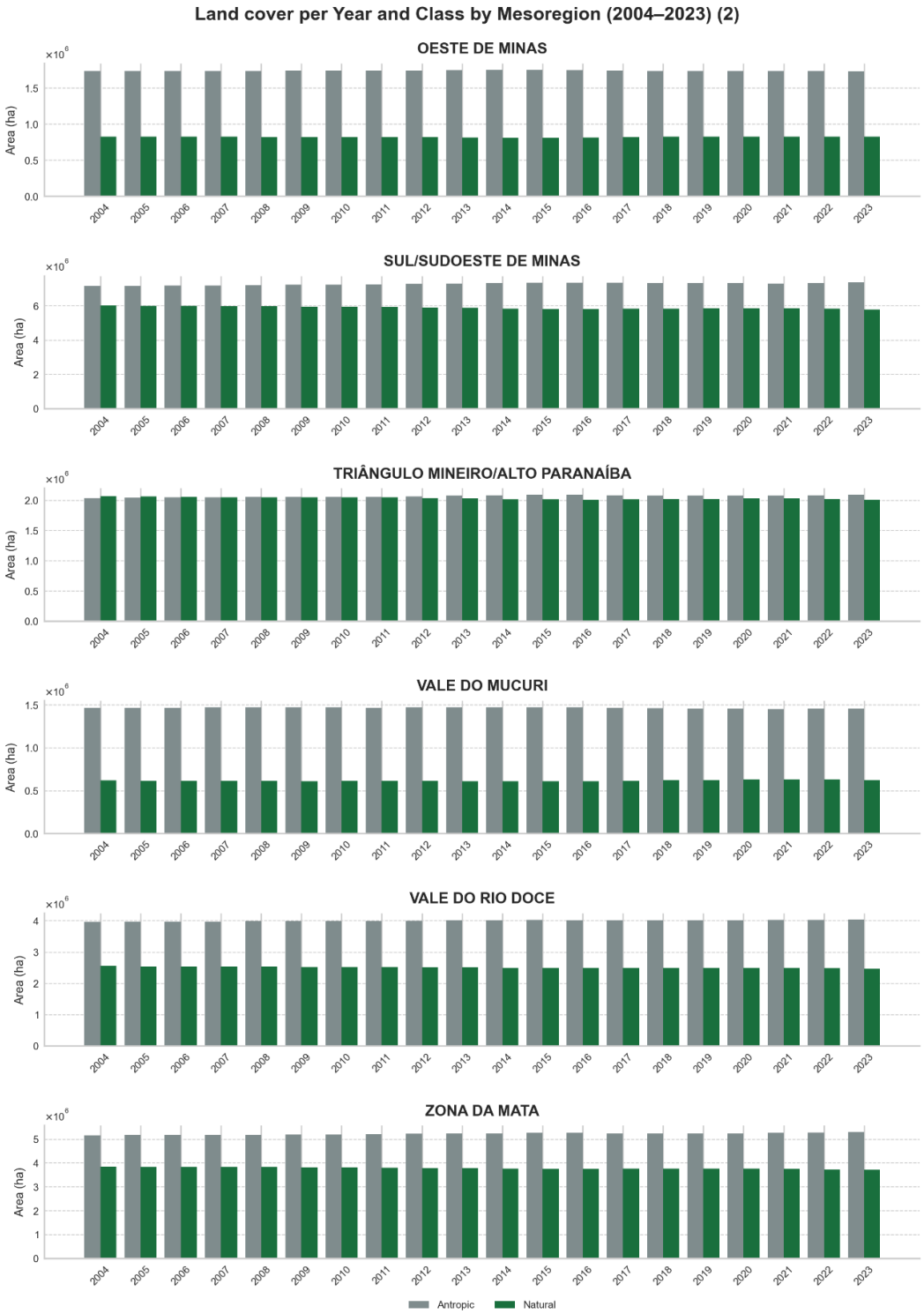

Figure 6.6. Annual land cover evolution (2004–2023) by class for the mesoregions of Oeste de Minas, Sul/Sudoeste de Minas, Triângulo Mineiro/Alto Paranaíba, Vale do Mucuri, Vale do Rio Doce, and Zona da Mata. Bars represent the total area (ha) classified as 'Anthropic' (gray) and 'Natural' (green) each year.

6.4. Discussion

The precipitation patterns observed in southeastern Brazil between 2004 and 2024 reveal significant interannual variability driven by multiple large-scale climatic mechanisms, including the South Atlantic Convergence Zone (SACZ), the El Niño-Southern Oscillation (ENSO), and subtropical high-pressure systems. These climatic drivers shape both seasonal and long-term hydrological trends, influencing regional water availability, agricultural productivity, and ecosystem stability. The analysis of precipitation levels over this period reveals a succession of alternating wet and dry years, with notable declines in annual precipitation during certain periods, leading to hydrological stress and drought conditions. These fluctuations underscore the region's increasing hydroclimatic sensitivity to both natural variability and anthropogenic pressures (Marengo et al., 2020).

The SACZ, a primary modulator of summer rainfall in southeastern Brazil, played a crucial role throughout the study period. It sustains prolonged convective activity during the austral summer by channeling moisture from the Amazon and the South Atlantic. The variability of the SACZ is closely linked to ENSO phases—La Niña years tend to weaken the SACZ, while El Niño years can intensify it (Carvalho et al., 2004; Verdan & Silva, 2022). This relationship was evident in the exceptionally dry years of 2007, 2012, and 2017, all of which coincided with La Niña phases and featured suppressed convective activity (Gushchina et al., 2020). Conversely, SACZ-intensified years such as 2008, 2011, and 2020 were marked by above-average precipitation, often leading to floods and landslides.

The regionalized analysis of SACZ impacts in Minas Gerais shows pronounced spatial disparities in precipitation extremes and their socio-environmental consequences. The mesoregions of Zona da Mata, Vale do Rio Doce, Jequitinhonha, Triângulo Mineiro/Alto Paranaíba, and the Metropolitan Region of Belo Horizonte exhibited the highest concentration of extreme rainfall events. These areas were also among those with the most intense human land use, according to Figures 5 and 6. Notably, regions with high anthropogenic land cover experienced more severe hydrological impacts, such as urban floods, landslides, and disruptions to infrastructure, reinforcing the connection between land-use change and vulnerability to extreme weather (Banjara et al., 2024).

In the Zona da Mata, flood-prone cities like Cataguases and Muriaé suffered severe damage during major SACZ events, such as in January 2005. The concurrence of steep topography, dense urbanization, and SACZ-induced downpours created conditions conducive to riverine flooding (Li et al., 2024). Likewise, in Vale do Rio Doce, urbanized municipalities such as Ipatinga and Coronel Fabriciano were repeatedly affected by flooding during SACZ-active years. The bar plots of land cover reveal that these mesoregions are predominantly classified as "antropic," which reflects deforestation and impermeabilization that exacerbate runoff and reduce infiltration capacity.

Concerning to the Jequitinhonha mesoregion, despite having a lower density of urban development, exhibited the highest cumulative rainfall values and one of the most severe SACZ impacts. Its dual vulnerability, frequent SACZ episodes and socio-economic fragility, was evident in the catastrophic floods and infrastructure failures recorded in 2022. Events such as the overflow of the Jucuruçu River and the collapse of a dam in Crisolita underscore the region's limited capacity for disaster response and highlight the role of climatic extremes in magnifying existing vulnerabilities.

The Metropolitan Region of Belo Horizonte presented some of the most severe cases of SACZ-related flooding. The historic rainfall in January 2020, which exceeded 935 mm in a single month, led to widespread fatalities and infrastructural collapse. Given this region’s extensive urban sprawl and impervious surface coverage, even moderate rainfall accumulations can overwhelm drainage systems (Dalagnol et al., 2022). Figures 5 and 6 further corroborate that this mesoregion is dominated by anthropogenic land cover, making it highly susceptible to SACZ-induced hydrometeorological hazards.

In the Triângulo Mineiro/Alto Paranaíba region, cities such as Uberaba, Uberlândia, and Patos de Minas experienced recurrent flood events. The March 2024 storm in Monte Carmelo, where over 90 mm of rain fell in 12 hours, exemplifies this trend. While this region is partly agricultural and less urbanized than Belo Horizonte, land cover analyses show a predominance of antropic classification, indicating significant deforestation and cropland expansion (Sanches et al., 2017). These land-use changes

have degraded the region's hydrological buffering capacity, thereby intensifying the impact of SACZ events.

Climatic teleconnections continue to play a dominant role in modulating SACZ behavior. ENSO phases have induced both wet and dry anomalies across southeastern Brazil, often mediated by the Pacific South American (PSA) pattern, the Southern Annular Mode (SAM), and the Atlantic Multidecadal Oscillation (AMO). For instance, the droughts of 2012 and 2017 were associated with La Niña and intensified SASH activity, which inhibited moisture transport and deep convection (Marengo et al., 2013; Correia et al., 2024). On the other hand, El Niño events in 2011 and 2016 coincided with enhanced SACZ persistence and widespread flooding, particularly in North and Northeast Minas Gerais.

The severe droughts of 2012 and 2015 also underscore the cumulative impacts of atmospheric blocking and high-pressure anomalies over the South Atlantic. The SASH's intensification during these years curtailed precipitation over southeastern Brazil, affecting not only surface water resources but also groundwater recharge and agricultural productivity (Silva et al., 2018). In contrast, years with strong SACZ activity, such as 2008 and 2020, provided hydrological relief but triggered destructive floods, landslides, and urban disruption.

As noted in Figure 4, the spatial recurrence of SACZ-related events follows a northeast-southwest corridor of precipitation anomalies, which coincides with densely populated and environmentally degraded regions. This spatial signature is critical for understanding regional vulnerability. In particular, mesoregions with a higher proportion of natural land cover, such as parts of Norte de Minas, demonstrated a relatively buffered hydrological response, suggesting that natural vegetation plays a key role in mitigating extreme rainfall impacts (Alavipanah et al., 2015).

An apparent decline in SACZ activity in 2023 may indicate the onset of a new dry phase, though longer-term monitoring is needed to confirm this trend. However, projections suggest that climate change will likely amplify hydrometeorological extremes, leading to both more frequent droughts and more intense rainfall events (Sondermann et

al., 2022). These findings emphasize the need for adaptive water governance, regional planning that accounts for climate variability, and the integration of land-use policy with hydrological risk management.

Future efforts should prioritize investments in early warning systems, restoration of natural vegetation buffers, and climate-resilient infrastructure—particularly in SACZ-prone mesoregions. Advances in remote sensing and high-resolution climate modeling will be critical for forecasting SACZ dynamics and supporting real-time decision-making (Amaral e Silva et al., 2024). The convergence of hydroclimatic data, land use classification, and disaster records offers a robust foundation for targeted interventions aimed at reducing vulnerability to SACZ-related extremes in Minas Gerais and beyond.

6.5. Conclusion

The analysis of precipitation patterns and SACZ-related extreme rainfall events from 2004 to 2024 reveals a complex and spatially uneven hydroclimatic regime across southeastern Brazil, with Minas Gerais standing out as a hotspot for both climatic variability and socio-environmental vulnerability. The interplay between large-scale atmospheric drivers, particularly the SACZ, ENSO phases, and the South Atlantic Subtropical High, and local land cover configurations critically shapes regional precipitation dynamics, drought frequency, and flood risk.

Findings from this study underscore that SACZ variability is not only influenced by ocean-atmosphere interactions, such as El Niño and La Niña events, but also by regional feedbacks including land-use change and geomorphological exposure. Mesoregions with high proportions of anthropogenic land cover, such as the Zona da Mata, Vale do Rio Doce, and the Metropolitan Region of Belo Horizonte, experienced both higher frequency and severity of SACZ-related impacts, including recurrent urban flooding, landslides, and infrastructure collapse. In contrast, regions with a predominance of natural vegetation, such as parts of Norte de Minas, demonstrated greater hydrological resilience for equivalent precipitation totals, highlighting the buffering capacity of preserved ecosystems.

The regionalization of extreme events further revealed the northeast-southwest corridor as a spatial axis of SACZ influence, where convective systems persistently interact with topographic and urban features. The Triângulo Mineiro/Alto Paranaíba region, for example, exemplifies how agricultural expansion and loss of vegetative cover diminish hydrological buffering, rendering the area increasingly susceptible to flash floods during SACZ peaks. Similarly, the Jequitinhonha Valley demonstrates how high rainfall accumulation, combined with socio-economic fragility, amplifies disaster impacts even in less urbanized settings.

These results point to the urgent need to integrate climate diagnostics with land management policies. Satellite-derived datasets, such as CHIRPS, proved instrumental in identifying spatial patterns of precipitation variability and filling observational gaps across complex terrains. Combined with land use classification and disaster reporting, they form a powerful toolset for targeting adaptive interventions.

Looking forward, climate change projections indicate an increase in the intensity and irregularity of SACZ behavior, with potential delays in onset, shifts in duration, and heightened rainfall extremes. Therefore, enhancing climate resilience in Minas Gerais requires a multifaceted strategy: investing in real-time monitoring systems, restoring ecological buffers, strengthening urban drainage infrastructure, and aligning territorial planning with hydroclimatic risk assessments. This study reinforces that managing SACZ-related hydrological extremes demands not only atmospheric science but also coordinated land governance and infrastructure adaptation.

References

Alavipanah, S.; Wegmann, M.; Qureshi, S.; Weng, Q.; Koellner, T. The Role of Vegetation in Mitigating Urban Land Surface Temperatures: A Case Study of Munich, Germany during the Warm Season. Sustainability 2015, 7, 4689-4706. https://doi.org/10.3390/su7044689

Amaral e Silva, A., de Assis, L. C., dos Santos, V. J., de Andrade, L. C., Lorentz, J. F., Henriques, B. S., Calijuri, M. L., & Ferreira, I. O. (2024). Rainfall From Brazilian

Flying Rivers: Evaluating the Effectiveness of Precipitation Gridded Databases. International Journal of Climatology. https://doi.org/10.1002/joc.8707

Aquino, R. F., Silva, M. L. N., Freitas, D. A. F. de, Curi, N., Mello, C. R. de, & Avanzi, J. C. (2012). Spatial variability of the rainfall erosivity in southern region of Minas Gerais state, Brazil. Ciência e Agrotecnologia, 36(5), 533–542. https://doi.org/10.1590/S1413-70542012000500006

Araghi, A., & Adamowski, J. F. (2024). Assessment of 30 gridded precipitation datasets over different climates on a country scale. Earth Science Informatics, 17(2), 1301–1313. https://doi.org/10.1007/s12145-023-01215-0

Arias, E. C., & Barriga, J. C. (2022). Performance of high-resolution precipitation datasets CHIRPS and TerraClimate in a Colombian high Andean Basin. Geocarto International, 37(27), 17382–17402. https://doi.org/10.1080/10106049.2022.2129816

Banjara, M.; Bhusal, A.; Ghimire, A.B.; Kalra, A. Impact of Land Use and Land Cover Change on Hydrological Processes in Urban Watersheds: Analysis and Forecasting for Flood Risk Management. Geosciences 2024, 14, 40. https://doi.org/10.3390/geosciences14020040

Barreiro, M. (2024). Combined Effects of the Tropical and Extratropical Modes of Variability on Precipitation in Southeastern South America. In Oxford Research Encyclopedia of Climate Science. Oxford University Press. https://doi.org/10.1093/acrefore/9780190228620.013.960

Bivand, R. (2022). R Packages for Analyzing Spatial Data: A Comparative Case Study with Areal Data. Geographical Analysis, 54(3), 488–518. https://doi.org/10.1111/gean.12319

Caetano, C. C. R., Ávila, L. A. C. de, & Tavares, M. (2017). A relação entre as transferências governamentais, a arrecadação tributária própria e o índice de educação dos municípios do estado de Minas Gerais. Revista de Administração Pública, 51(5), 897–916. https://doi.org/10.1590/0034-7612174433

Cardoso, C. de S., Quadro, M. F. L. de, & Bonetti, C. (2020). Persistência e Abrangência dos Eventos Extremos de Precipitação no Sul do Brasil: Variabilidade Espacial e Padrões Atmosféricos. Revista Brasileira de Meteorologia, 35(2), 219–231. https://doi.org/10.1590/0102-7786352031

Carvalho, J. D. de, Melo, M. C. de, Galvão, P. G. P., Miranda, W. L., & Lasmar, B. E. (2024). Groundwater in water scarcity context in the São Francisco River Basin (MG). Ambiente & Sociedade, 27. https://doi.org/10.1590/1809-4422asoc0177r5vu27l2oa

Carvalho, L. M. V, Jones, C., & Liebmann, B. (2004). The South Atlantic Convergence Zone: Intensity, Form, Persistence, and Relationships with Intraseasonal to Interannual Activity and Extreme Rainfall. Journal of Climate, 17(1), 88–108. https://doi.org/10.1175/1520-0442(2004)017<0088:TSACZI>2.0.CO;2

Correia Filho, W. L. F., Souza, P. H. de A., Oliveira-Júnior, J. F. de, Santiago, D. de B., Lyra, G. B., Zeri, M., & Cunha-Zeri, G. (2022). The wind regime over the Brazilian Southeast: Spatial and temporal characterization using multivariate analysis. International Journal of Climatology, 42(3), 1767–1788. https://doi.org/10.1002/joc.7334

Correia, P. O., Ribeiro, E. M., Galizoni, F. M., Simão, E. J. de P., & Santos, L. R. (2024). Águas, energias e alimentos na agricultura familiar do Alto Vale do Jequitinhonha, Minas Gerais. Revista Brasileira de Estudos Urbanos e Regionais, 26(1). https://doi.org/10.22296/2317-1529.rbeur.202431

Dalagnol, R., C. B. Gamcianinov, N. M. Crespo, et al. 2022. “Extreme Precipitation and Its Impacts in the Brazilian Minas Gerais State in January 2020: Can We Blame Climate Change?” Climate Resilience and Sustainability 1, no. 1: 15. https://doi.org/10.1002/cli2.15.

dos Santos, V., Dias de Oliveira, M., Boll, J., Sánchez-Murillo, R., Menegário, A. A., Gozzo, L. F., & Gastmans, D. (2019). Isotopic composition of precipitation during

strong El Niño–Southern Oscillation events in the Southeast Region of Brazil. Hydrological Processes, 33(4), 647–660. https://doi.org/10.1002/hyp.13351

Du, H., Tan, M. L., Zhang, F., Chun, K. P., Li, L., & Kabir, M. H. (2024). Evaluating the effectiveness of CHIRPS data for hydroclimatic studies. Theoretical and Applied Climatology, 155(3), 1519–1539. https://doi.org/10.1007/s00704-023-04721-9

Elmahal, A. E., & Musa, M. M. I. (2023). Spatial Cloud Computing Using Google Earth Engine and R Packages. In Geographic Information Systems - Data Science Approach. IntechOpen. https://doi.org/10.5772/intechopen.1002686

Filho, W. L. F. C., de Almeida Souza, P. H., de Oliveira-Júnior, J. F., de Barros Santiago, D., Lyra, G. B., Zeri, M., & Cunha-Zeri, G. (2022). The wind regime over the Brazilian Southeast: Spatial and temporal characterization using multivariate analysis. International Journal of Climatology, 42(3), 1767–1788. https://doi.org/10.1002/joc.7334

Freitas, A. A., Drumond, A., Carvalho, V. S. B., Reboita, M. S., Silva, B. C., & Uvo, C. B. (2022). Drought Assessment in São Francisco River Basin, Brazil: Characterization through SPI and Associated Anomalous Climate Patterns. Atmosphere, 13(1), 41. https://doi.org/10.3390/atmos13010041

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific Data, 2(1), 150066. https://doi.org/10.1038/sdata.2015.66

Gushchina, D., Zheleznova, I., Osipov, A., & Olchev, A. (2020). Effect of Various Types of ENSO Events on Moisture Conditions in the Humid and Subhumid Tropics. Atmosphere, 11(12), 1354. https://doi.org/10.3390/atmos11121354

Lembi, R. C., Cronemberger, C., Picharillo, C., Koffler, S., Sena, P. H. A., Felappi, J. F., Moraes, A. R. de, Arshad, A., Santos, J. P. dos, & Mansur, A. V. (2020). Urban expansion in the Atlantic Forest: applying the Nature Futures Framework to

develop a conceptual model and future scenarios. Biota Neotropica, 20(suppl 1). https://doi.org/10.1590/1676-0611-bn-2019-0904

Li, J.; Zhou, W.; Tao, C. The Impact of Urbanization on Surface Runoff and Flood Prevention Strategies: A Case Study of a Traditional Village. Land 2024, 13, 1528. https://doi.org/10.3390/land13091528

López-Bermeo, C., Montoya, R. D., Caro-Lopera, F. J., & Díaz-García, J. A. (2022). Validation of the accuracy of the CHIRPS precipitation dataset at representing climate variability in a tropical mountainous region of South America. Physics and Chemistry of the Earth, Parts A/B/C, 127(July 2021), 103184. https://doi.org/10.1016/j.pce.2022.103184

Maia, N. Z., Almeida, L. P., Emmendorfer, L., Nicolodi, J. L., & Calliari, L. (2022). Wave climate trends and breakpoints during the Atlantic Multidecadal Oscillation (AMO) in southern Brazil. Ocean and Coastal Research, 70. https://doi.org/10.1590/2675-2824070.21086nzm

Marengo, J. A., Alves, L. M., Soares, W. R., Rodriguez, D. A., Camargo, H., Riveros, M. P., & Pabló, A. D. (2013). Two Contrasting Severe Seasonal Extremes in Tropical South America in 2012: Flood in Amazonia and Drought in Northeast Brazil. Journal of Climate, 26(22), 9137–9154. https://doi.org/10.1175/JCLI-D-12-00642.1

Marengo, J. A., Cunha, A. P. M. A., Nobre, C. A., Ribeiro Neto, G. G., Magalhaes, A. R., Torres, R. R., Sampaio, G., Alexandre, F., Alves, L. M., Cuartas, L. A., Deusdará, K. R. L., & Álvala, R. C. S. (2020). Assessing drought in the drylands of northeast Brazil under regional warming exceeding 4 °C. Natural Hazards, 103(2), 2589–2611. https://doi.org/10.1007/s11069-020-04097-3

Musah, A., Dutra, L. M. M., Aldosery, A., Browning, E., Ambrizzi, T., Borges, I. V. G., Tunali, M., Başibüyük, S., Yenigün, O., Moreno, G. M. M., da Silva, A. C. G., dos Santos, W. P., de Lima, C. L., Massoni, T., Jones, K. E., Campos, L. C., & Kostkova, P. (2022). An Evaluation of the OpenWeatherMap API versus INMET Using

Weather Data from Two Brazilian Cities: Recife and Campina Grande. Data, 7(8), 106. https://doi.org/10.3390/data7080106

Nadeem, M. U., Anjum, M. N., Afzal, A., Azam, M., Hussain, F., Usman, M., Javaid, M. M., Mukhtar, M. A., & Majeed, F. (2022). Assessment of Multi-Satellite Precipitation Products over the Himalayan Mountains of Pakistan, South Asia. Sustainability, 14(14), 8490. https://doi.org/10.3390/su14148490

Nogueira, P. M. (2024). Spatial Analysis in Geology Using R. Chapman and Hall/CRC. https://doi.org/10.1201/9781032651880

Nunes, Y. R. F., Souza, C. S., Azevedo, I. F. P. de, Oliveira, O. S. de, Frazão, L. A., Fonseca, R. S., Santos, R. M. dos, & Neves, W. V. (2022). Vegetation structure and edaphic factors in veredas reflect different conservation status in these threatened areas. Forest Ecosystems, 9, 100036. https://doi.org/10.1016/j.fecs.2022.100036

Oliveira, e. C., barboza, r. D., silva, b. G. G., & diaz filho, m. C. (2023). Erosion of four Brazilian coastal deltas: how dam construction is changing the natural pattern of coastal sedimentary systems. Anais Da Academia Brasileira de Ciências, 95(suppl 1). https://doi.org/10.1590/0001-3765202320220576

Paredes-Trejo, F. J., Barbosa, H. A., & Lakshmi Kumar, T. V. (2017). Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. Journal of Arid Environments, 139, 26–40. https://doi.org/10.1016/j.jaridenv.2016.12.009

Pedreira Junior, A. L., Querino, C. A. S., Biudes, M. S., Machado, N. G., Santos, L. O. F. dos, & Ivo, I. O. (2020). Influence of El Niño and La Niña phenomena on seasonality of the relative frequency of rainfall in southern Amazonas mesoregion. RBRH, 25. https://doi.org/10.1590/2318-0331.252020190152

RezaAbbasifard, M., Ghahremani, B., & Naderi, H. (2014). A Survey on Nearest Neighbor Search Methods. International Journal of Computer Applications, 95(25), 39–52. https://doi.org/10.5120/16754-7073

Rodrigues, E. C., Tavares, R., & Meireles, A. L. (2020). Inequalities in the spatial concentration of agricultural crops among the micro- and mesoregions of Minas Gerais, Brazil. Ciência Rural, 50(5). https://doi.org/10.1590/0103-8478cr20190116

Roque, F. V., Macarini, L. A., Crotti, Y., OliveiraWeber, T., & Cechinel, C. (2019). Detecção de defeitos visuais em tecidos utilizando Wavelets e algoritmos de aprendizado de máquina. Anais Do Computer on the Beach, 619–628. https://periodicos.univali.br/index.php/acotb/article/view/14359

Saddique, N., Muzammil, M., Jahangir, I., Sarwar, A., Ahmed, E., Aslam, R. A., & Bernhofer, C. (2022). Hydrological evaluation of 14 satellite-based, gauge-based and reanalysis precipitation products in a data-scarce mountainous catchment. Hydrological Sciences Journal, 67(3), 436–450. https://doi.org/10.1080/02626667.2021.2022152

Sajjad, M. M., Wang, J., Afzal, Z., Hussain, S., Siddique, A., Khan, R., Ali, M., & Iqbal, J. (2023). Assessing the Impacts of Groundwater Depletion and Aquifer Degradation on Land Subsidence in Lahore, Pakistan: A PS-InSAR Approach for Sustainable Urban Development. Remote Sensing, 15(22), 5418. https://doi.org/10.3390/rs15225418

Sanches, F. de O., Silva, R. V. da, Ferreira, R. V., & Campos, C. A. A. (2017). Climate change in the Triângulo Mineiro region – Brazill. Revista Brasileira De Climatologia, 21. https://doi.org/10.5380/abclima.v21i0.51867

Satyamurty, P., Nobre, C. A., & Silva Dias, P. L. (1998). South America. In Meteorology of the Southern Hemisphere (pp. 119–139). American Meteorological Society. https://doi.org/10.1007/978-1-935704-10-2_5

Silva, F. B., Pereira, S. B., Martinez, M. A., Silva, D. D. da, & Vieira, N. P. A. (2018). Water needs and equivalence relations for different irrigated crops in the São Francisco basin. REVISTA CIÊNCIA AGRONÔMICA, 49(3). https://doi.org/10.5935/1806-6690.20180046

Silva, J. P. R., Reboita, M. S., & Escobar, G. C. J. (2019). Caracterização da zona de convergência do atlântico sul em campos atmosféricos recentes. Revista Brasileira de Climatologia, 25. https://doi.org/10.5380/abclima.v25i0.64101

Sondermann, M., Chou, S. C., Lyra, A., Latinovic, D., Siqueira, G. C., Junior, W. C., Giornes, E., & Leite, F. P. (2022). Climate change projections and impacts on the eucalyptus plantation around the Doce River basin, in Minas Gerais, Brazil. Climate Services, 28, 100327. https://doi.org/10.1016/j.cliser.2022.100327

Souza, C. M., Z. Shimbo, J., Rosa, M. R., Parente, L. L., A. Alencar, A., Rudorff, B. F. T., Hasenack, H., Matsumoto, M., G. Ferreira, L., Souza-Filho, P. W. M., de Oliveira, S. W., Rocha, W. F., Fonseca, A. V., Marques, C. B., Diniz, C. G., Costa, D., Monteiro, D., Rosa, E. R., Vélez-Martin, E., … Azevedo, T. (2020). Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. Remote Sensing, 12(17), 2735. https://doi.org/10.3390/rs12172735

Verdan, I., & Silva, M. E. S. (2022). Variabilidade da Zona de Convergência do Atlântico Sul em relação a eventos ENOS de 2000 a 2021. Geography Department University of Sao Paulo, 42, e193110. https://doi.org/10.11606/eISSN.2236-2878.rdg.2022.193110

Wang, D., Zhang, M., Fu, M., Cai, Z., Li, Z., Han, H., Cui, Y., & Luo, B. (2016). Nonlinearity Mitigation Using a Machine Learning Detector Based on k-Nearest Neighbors. IEEE Photonics Technology Letters, 28(19), 2102–2105. https://doi.org/10.1109/LPT.2016.2555857

Wickham, H. (2016). ggplot2. Springer International Publishing. https://doi.org/10.1007/978-3-319-24277-4

7. GENERAL CONCLUSIONS

This research represents a significant advance in understanding climate variability in Brazil, with a focus on the Flying Rivers and the South Atlantic Convergence Zone (SACZ). Through the use of statistical methodologies, remote sensing, and artificial intelligence, the study analyzed the interaction between atmospheric and hydrological systems, with an emphasis on precipitation in southeastern Brazil, particularly in the state of Minas Gerais. Multivariate analysis of precipitation data revealed that factors such as temperature, evapotranspiration, and vegetation cover significantly influence rainfall distribution in the Flying Rivers region. The Amazon Forest plays a crucial role in moisture recycling and water vapor transport, and its degradation is directly associated with reduced rainfall in agricultural and urban areas of the Southeast, thereby increasing water vulnerability and the frequency of extreme events. A strong correlation was also identified between SACZ variability and global climate phenomena such as El Niño and La Niña. During El Niño, SACZ activity intensifies, resulting in heavier rainfall, whereas La Niña weakens its activity, prolonging droughts and raising the risk of water crises. These findings underscore the need for accurate predictive models to support water resource management and climate adaptation policies. The comparison between rain gauge measurements and satellite-based precipitation estimates showed that models such as CHIRPS, PERSIANN, and TerraClimate provide broader spatial coverage and stronger predictive capability than traditional monitoring networks. Integrating these technologies enhances the monitoring of extreme climate events and supports effective disaster mitigation strategies. Finally, the study emphasizes the importance of preserving the Amazon Forest for Brazil's climate stability and the need for public policies that strengthen water resilience. Continuous improvement in climate modeling techniques and the expansion of meteorological monitoring networks are essential for bolstering climate resilience and ensuring the country's water security in the coming decades.

8. SUGGESTIONS FOR FUTURE RESEARCH

Future studies may enhance the understanding of climate variability in Brazil by refining predictive models for the South Atlantic Convergence Zone (SACZ) and extreme precipitation events. The use of machine learning and artificial intelligence can improve forecast accuracy, while the integration of satellite data and statistical modeling may increase the reliability of predictions concerning the intensity and duration of such events. Furthermore, analyzing the impacts of deforestation on moisture recycling and atmospheric transport by the Flying Rivers is crucial, as deforestation has been shown to reduce precipitation in distant regions. Numerical simulations can be employed to evaluate scenarios of forest conservation and restoration, providing guidance for more effective environmental policies.

Additionally, future research could focus on monitoring extreme hydrometeorological events—such as floods and droughts—through remote sensing and geospatial modeling, aiming to identify vulnerable areas and support climate impact mitigation. The relationship between SACZ variability and agricultural productivity also warrants investigation, given that shifts in rainfall patterns may jeopardize food security. Another important research direction involves examining the role of public policies in managing water and climate resources, assessing strategies such as resilient infrastructure, water reserve systems, and adaptive urban planning. Moreover, the correlation between SACZ frequency and public health could be explored by analyzing how extreme weather events influence the spread of tropical and respiratory diseases.

The continuation of these research efforts will contribute to advancing scientific knowledge and developing effective strategies for environmental, water, and agricultural management, thereby strengthening Brazil’s climate resilience in the face of global change.

9. GENERAL REFERENCES

AMARAL E SILVA, A.; BRAGA, M. Q.; FERREIRA, J.; JUSTE DOS SANTOS, V.; DO CARMO ALVES, S.; DE OLIVEIRA, J. C.; CALIJURI, M. L. Anthropic activities and the Legal Amazon: Estimative of impacts on forest and regional climate for 2030. **Remote Sensing Applications**, v. 18, p. 100304, 2020.

AMARAL E SILVA, A.; DE ASSIS, L. C.; DOS SANTOS, V. J.; DE ANDRADE, L. C.; LORENTZ, J. F.; HENRIQUES, B. S.; CALIJURI, M. L.; FERREIRA, I. O. Rainfall From Brazilian Flying Rivers: Evaluating the Effectiveness of Precipitation Gridded Databases. **International Journal of Climatology**, 2024.

ARRAUT, J. M.; NOBRE, C.; BARBOSA, H. M. J.; OBREGON, G.; MARENGO, J. Aerial Rivers and Lakes: Looking at Large-Scale Moisture Transport and Its Relation to Amazonia and to Subtropical Rainfall in South America. **Journal of Climate**, v. 25, n. 2, p. 543–556, 2012.

CARVALHO, J. D. DE; MELO, M. C. DE; GALVÃO, P. G. P.; MIRANDA, W. L.; LASMAR, B. E. Groundwater in water scarcity context in the São Francisco River Basin (MG). **Ambiente & Sociedade**, v. 27, 2024.

CEPEDA ARIAS, E.; CAÑON BARRIGA, J. Performance of high-resolution precipitation datasets CHIRPS and TerraClimate in a Colombian high Andean Basin. **Geocarto International**, v. 37, p. 17382–17402, 2022.

CONFESSOR, J. G.; SILVA, L. L.; ARAÚJO, P. M. S. Avaliação das Perdas de Água e Solo em Pastagem Inserida em Ambiente de Cerrado Brasileiro sob Chuva Simulada. **Sociedade & Natureza**, v. 34, 2022.

CORREIA FILHO, W. L. F.; SOUZA, P. H. DE A.; OLIVEIRA-JÚNIOR, J. F. DE; SANTIAGO, D. DE B.; LYRA, G. B.; ZERI, M.; CUNHA-ZERI, G. The wind regime over the Brazilian Southeast: Spatial and temporal characterization using multivariate analysis. **International Journal of Climatology**, v. 42, n. 3, p. 1767–1788, 2022.

CORREIA, P. O.; RIBEIRO, E. M.; GALIZONI, F. M.; SIMÃO, E. J. DE P.; SANTOS, L. R. Águas, energias e alimentos na agricultura familiar do Alto Vale do Jequitinhonha, Minas Gerais. **Revista Brasileira de Estudos Urbanos e Regionais**, v. 26, n. 1, 2024.

COSTA, J. C.; PEREIRA, G.; SIQUEIRA, M. E.; CARDOZO, F. DA S.; DA SILVA, V. V. Validação dos dados de precipitação estimados pelo CHIRPS para o Brasil. **Revista Brasileira de Climatologia**, v. 24, p. 228–243, 2024.

CUNHA, A. P. M. A.; ZERI, M.; LEAL, K. D.; COSTA, L.; CUARTAS, L. A.; MARENGO, J. A.; TOMASELLA, J.; VIEIRA, R. M.; BARBOSA, A. A.; CUNNINGHAM, C.; et al. Extreme Drought Events over Brazil from 2011 to 2019. **Atmosphere**, v. 10, n. 11, p. 642, 2019.

D'ACUNHA, B.; DALMAGRO, H. J.; ZANELLA DE ARRUDA, P. H.; BIUDES, M. S.; LATHUILLIÈRE, M. J.; URIBE, M.; COUTO, E. G.; BRANDO, P. M.; VOURLITIS, G.; JOHNSON, M. S. Changes in evapotranspiration, transpiration and evaporation across natural and managed landscapes in the Amazon, Cerrado and Pantanal biomes. **Agricultural and Forest Meteorology**, v. 346, p. 109875, 2024.

ELMAHAL, A. E.; MUSA, M. M. I. Spatial Cloud Computing Using Google Earth Engine and R Packages. In: Geographic Information Systems - Data Science Approach. **IntechOpen**, 2023.

EMMERICHS, T.; LU, Y.-S.; TARABORRELLI, D. The influence of plant water stress on vegetation–atmosphere exchanges: implications for ozone modelling. **Biogeosciences**, v. 21, p. 3251–3269, 2024.

FILHO, W. L. F. C.; ALMEIDA SOUZA, P. H.; OLIVEIRA-JÚNIOR, J. F.; BARROS SANTIAGO, D.; LYRA, G. B.; ZERI, M.; CUNHA-ZERI, G. The wind regime over the Brazilian Southeast: Spatial and temporal characterization using multivariate analysis. **International Journal of Climatology**, v. 42, n. 3, p. 1767–1788, 2022.

FREITAS, A. A.; DRUMOND, A.; CARVALHO, V. S. B.; REBOITA, M. S.; SILVA, B. C.; UVO, C. B. Drought Assessment in São Francisco River Basin, Brazil: Characterization

through SPI and Associated Anomalous Climate Patterns. **Atmosphere**, v. 13, n. 1, p. 41, 2022.

FU, R.; YIN, L.; LI, W.; ARIAS, P. A.; DICKINSON, R. E.; HUANG, L.; CHAKRABORTY, S.; FERNANDES, K.; LIEBMANN, B.; FISHER, R.; MYNENI, R. B. Increased dry-season length over southern Amazonia in recent decades and its implication for future climate projection. **Proceedings of the National Academy of Sciences**, v. 110, p. 18110–18115, 2013.

GHOZAT, A.; SHARAFATI, A.; HOSSEINI, S. A. Long-term spatiotemporal evaluation of CHIRPS satellite precipitation product over different climatic regions of Iran. Theoretical and **Applied Climatology**, v. 143, p. 211–225, 2021.

HSU, J.; HUANG, W.-R.; LIU, P.-Y.; LI, X. Validation of CHIRPS Precipitation Estimates over Taiwan at Multiple Timescales. **Remote Sensing**, v. 13, p. 254, 2021.

INSTITUTO NACIONAL DE METEOROLOGIA (INMET). Normais Climatológicas do Brasil 1991-2020. Brasília/DF, Brasil, 2022.

JASECHKO, S.; SHARP, Z. D.; GIBSON, J. J.; BIRKS, S. J.; YI, Y.; FAWCETT, P. J. Terrestrial water fluxes dominated by transpiration. **Nature**, v. 496, p. 347–350, 2013.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, **Physical and Engineering Sciences**, 2016.

LÁZARO, W. L.; OLIVEIRA-JÚNIOR, E. S.; SILVA, C. J. DA; CASTRILLON, S. K. I.; MUNIZ, C. C. Climate change reflected in one of the largest wetlands in the world: an overview of the Northern Pantanal water regime. **Acta Limnologica Brasiliensia**, v. 32, 2020.

LEVER, J.; KRZYWINSKI, M.; ALTMAN, N. Points of Significance: Principal component analysis. **Nature Methods**, v. 14, p. 641–642, 2017.

LÓPEZ-BERMEO, C.; MONTOYA, R. D.; CARO-LOPERA, F. J.; DÍAZ-GARCÍA, J. A. Validation of the accuracy of the CHIRPS precipitation dataset at representing climate

variability in a tropical mountainous region of South America. **Physics and Chemistry of the Earth, Parts A/B/C**, v. 127, p. 103184, 2022.

LOVEJOY, T. E.; NOBRE, C. Amazon tipping point: Last chance for action. **Science Advances**, v. 5, p. 4–6, 2019.

MAKARIEVA, A. M.; GORSHKOV, V. G. Biotic pump of atmospheric moisture as driver of the hydrological cycle on land. **Hydrology and Earth System Sciences**, v. 11, p. 1013–1033, 2007.

MARENGO, J. A.; ALVES, L. M.; SOARES, W. R.; RODRIGUEZ, D. A.; CAMARGO, H.; RIVEROS, M. P.; PABLÓ, A. D. Two Contrasting Severe Seasonal Extremes in Tropical South America in 2012: Flood in Amazonia and Drought in Northeast Brazil. **Journal of Climate**, v. 26, n. 22, p. 9137–9154, 2013.

MARENGO, J. A.; CUNHA, A. P. M. A.; NOBRE, C. A.; RIBEIRO NETO, G. G.; MAGALHÃES, A. R.; TORRES, R. R.; SAMPAIO, G.; ALEXANDRE, F.; ALVES, L. M.; CUARTAS, L. A.; DEUSDARÁ, K. R. L.; ÁLVALA, R. C. S. Assessing drought in the drylands of northeast Brazil under regional warming exceeding 4 °C. **Natural Hazards**, v. 103, n. 2, p. 2589–2611, 2020.

MARENGO, J. A.; CUNHA, A. P.; ALVES, L. M. A seca de 2012-15 no semiárido do Nordeste do Brasil no contexto histórico. **Revista Climanálise**, v. 4, p. 49–54, 2016.

MARENGO, J. A.; ESPINOZA, J.-C.; FU, R.; JIMENEZ MUÑOZ, J. C.; ALVES, L. M.; DA ROCHA, H. R.; SCHÖNGART, J. Long-term variability, extremes and changes in temperature and hydrometeorology in the Amazon region: A review. **Acta Amazonica**, v. 54, 2024.

MONTEIRO, A.; CAMPELO, M. M. O Pará no foco dos estudos sobre populações negras da Amazônia Oriental. **Afros & Amazônia**, v. 1, 2022.

MORAES, E. C. DE. Fundamentos de Sensoriamento Remoto. INPE-8984-PUD/62, São José dos Campos, 2002.

MUSAH, A.; DUTRA, L. M. M.; ALDOSERY, A.; BROWNING, E.; AMBRIZZI, T.; BORGES, I. V. G.; TUNALI, M.; BAŞIBÜYÜK, S.; YENIGÜN, O.; MORENO, G. M. M.; DA SILVA, A. C. G.; DOS SANTOS, W. P.; DE LIMA, C. L.; MASSONI, T.; JONES, K. E.; CAMPOS, L. C.; KOSTKOVA, P. An Evaluation of the OpenWeatherMap API versus INMET Using Weather Data from Two Brazilian Cities: Recife and Campina Grande. **Data**, v. 7, n. 8, p. 106, 2022.

MYTTENAERE, A.; GOLDEN, B.; LE GRAND, B.; ROSSI, F. Mean Absolute Percentage Error for regression models. **Neurocomputing**, v. 192, p. 38–48, 2016.

NACUR, E. R.; VARTULI, V. Flying rivers and climate change caused by the deforestation of the Amazon rainforest: A perspective based on Latin American constitutionalism. **Revista Brasileira de Direito Animal**, v. 16, p. 100–115, 2021.

NOBRE, A. D. The Future Climate of Amazonia Scientific Assessment Report. **Articulación Regional Amazônica**, São José dos Campos, 2014.

NOBRE, A. D. O futuro climático da Amazônia. **Instituto Nacional de Pesquisas Espaciais – INPE**, 2015.

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, p. 10759–10768, 2016.

NÓBREGA, R. S. Impactos do desmatamento e de mudanças climáticas nos recursos hídricos na Amazônia Ocidental utilizando o modelo SLURP. **Revista Brasileira de Meteorologia**, v. 29, p. 111–120, 2014.

PEARCE, F. Rivers in the sky. **New Scientist**, v. 244, n. 3254, p. 40–43, 2019.

PEARCE, F. Weather makers. **Science**, v. 368, p. 1302–1305, 2020.

SADDIQUE, N.; MUZAMMIL, M.; JAHANGIR, I.; SARWAR, A.; AHMED, E.; ASLAM, R. A.; BERNHOFER, C. Hydrological evaluation of 14 satellite-based, gauge-based and

reanalysis precipitation products in a data-scarce mountainous catchment. **Hydrological Sciences Journal**, v. 67, n. 3, p. 436–450, 2022.

SALATI, E.; DALL'OLIO, A.; MATSUI, E.; GAT, J. R. Recycling of water in the Amazon Basin: An isotopic study. **Water Resources Research**, v. 15, p. 1250–1258, 1979.

SALATI, E.; VOSE, P. B. Amazon Basin: A System in Equilibrium. **Science**, v. 225, p. 129–138, 1984.

SHEIL, D.; MURDIYARSO, D. How forests attract rain: an examination of a new hypothesis. **Bioscience**, v. 59, p. 341–347, 2009.

SILVA, J. P. R.; REBOITA, M. S.; ESCOBAR, G. C. J. Caracterização da zona de convergência do Atlântico Sul em campos atmosféricos recentes. **Revista Brasileira de Climatologia**, v. 25, 2019.

SONDERMANN, M.; CHOU, S. C.; LYRA, A.; LATINOVIC, D.; SIQUEIRA, G. C.; JUNIOR, W. C.; GIORNES, E.; LEITE, F. P. Climate change projections and impacts on the eucalyptus plantation around the Doce River basin, in Minas Gerais, Brazil. **Climate Services**, v. 28, p. 100327, 2022.

WENG, W.; LUEDEKE, M. K. B.; ZEMP, D. C.; LAKES, T.; KROPP, J. P. Aerial and surface rivers: downwind impacts on water availability from land use changes in Amazonia. **Hydrology and Earth System Sciences**, v. 22, n. 1, p. 911–927, 2018.

ZEMP, D. C.; SCHLEUSSNER, C. F.; BARBOSA, H. M. J.; HIROTA, M.; MONTADE, V.; SAMPAIO, G.; STAAL, A.; WANG-ERLANDSSON, L.; RAMMIG, A. Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks. **Nature Communications**, v. 8, 2017.

10. APPENDIX I

International Journal of Climatology

International Journal of Climatology | RMetS

RESEARCH ARTICLE

Rainfall From Brazilian Flying Rivers: Evaluating the Effectiveness of Precipitation Gridded Databases

Arthur Amaral e Silva[1] | Leonardo Campos de Assis[2] | Vitor Juste dos Santos[1] | Laura Coelho de Andrade[1] | Juliana Ferreira Lorentz[1] | Bruno Silva Henriques[1] | Maria Lucia Calijuri[1] | Italo Oliveira Ferreira[1]

[1]Department of Civil Engineering s/n, Federal University of Viçosa, Viçosa, Brazil | [2]Pro-Rectorate for Research, Postgraduate Studies, and Extension, University of Uberaba, Uberaba, Brazil

Correspondence: Arthur Amaral e Silva (arthur.amaral@ufv.br)

Received: 25 January 2024 | **Revised:** 19 November 2024 | **Accepted:** 20 November 2024

Funding: This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

Keywords: Amazon rainforest | flying Rivers | precipitation models | remote sensing | statistical analysis

ABSTRACT

The uneven global distribution of rainfall significantly impacts water resources and environmental sustainability, emphasising the need for reliable climate prediction models. Accurate predictions are vital for sectors such as food security, urban planning and disaster management. Data from ground stations, radars and satellites are essential, despite challenges like instrumental errors. Satellites, with their comprehensive sensors, are crucial for atmospheric observations, aiding in the prediction of large-scale climatic events. Climate models such as CHIRPS, GLDAS, TerraClimate, and PERSIANN use different approaches to analyse precipitation data, which is key to understanding its spatial and temporal variability. This study evaluated (rainfall data) from these four climate models over 20 years (within the Brazilian territory), focusing on the spatiotemporal behaviour of rainfall using statistical metrics such as R^2, RMSE, and MAPE. The findings showed that CHIRPS had the best performance ($R^2 = 0.843$; RMSE = 42.83; MAPE = 0.09%), excelling in both overall database and extreme event analyses. TerraClimate, initially the lowest-performing model ($R^2 = 0.413$; RMSE = 91.56; MAPE = 0.23%), improved significantly when combined with elevation through multiple linear regression (MLR), achieving R^2 of 0.718, RMSE of 31.14, and MAPE of 9.56%. This made TerraClimate a viable model for studying the Flying Rivers. The study highlights that model selection should align with the specific characteristics of the area under consideration, with CHIRPS being particularly suitable for the studied region. This research enhances the understanding of the effectiveness of these models in estimating rainfall compared to in situ measurements, which is crucial for various applications. The authors advocate for further studies to advance research on the Flying Rivers, their significance, and the impacts of climate change on them.

1 | Introduction

The link between forests and rainfall has become clearer in recent years. Initially, tree transpiration was identified as the primary source of water in continental areas (Jasechko et al. 2013). Moreover, a mechanism by which forests pump water into the atmosphere and sustain the moisture necessary for their survival was proposed by Makarieva and Gorshkov (2007) and further elaborated by Sheil and Murdiyarso (2009). This mechanism involves a reduction in atmospheric pressure at lower levels, which draws clouds over forested regions. Additionally, forests release biogenic volatile organic compounds that act as nuclei for water condensation. These processes result in the production of vast amounts of atmospheric water in forested regions, essentially forming 'flying rivers' (Nobre 2014a).