



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Câmpus de Presidente Prudente

RAFAEL BARONI ALTRAN

**ANÁLISE ESTATÍSTICA DOS PADRÕES MENSAIS DE DIAS DE CHUVA EM SÃO
PAULO**

PRESIDENTE PRUDENTE

2025

RAFAEL BARONI ALTRAN

**ANÁLISE ESTATÍSTICA DOS PADRÕES MENSIS DE DIAS DE CHUVA EM SÃO
PAULO**

Relatório Final de Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Estatística da FCT/UNESP para aproveitamento na disciplina Trabalho de Conclusão de Curso.

Orientador: Prof. Dr. Fernando Antonio Moala.

PRESIDENTE PRUDENTE

2025

A469a

Altran, Rafael Baroni

Análise estatística dos padrões mensais de dias de chuva em São Paulo / Rafael Baroni Altran. -- Presidente Prudente, 2025

45 p.

Trabalho de conclusão de curso (Bacharelado - Estatística) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente

Orientador: Fernando Antonio Moala

1. Chuva. 2. ENSO. 3. Distribuição beta. 4. Cópulas. 5. Sazonalidade. I. Título.

TERMO DE APROVAÇÃO

Rafael Baroni Altran

ANÁLISE ESTATÍSTICA DOS PADRÕES MENSIS DE DIAS DE CHUVA EM SÃO PAULO

Relatório de Final de Trabalho de Conclusão de Curso aprovado como requisito para obtenção de créditos na disciplina Trabalho de Conclusão do curso de graduação em Estatística da Faculdade de Ciências e Tecnologia da Unesp, pela seguinte banca examinadora:

Orientador: _____

Prof. Dr. Fernando Antonio Moala
Departamento de Estatística

Prof. Dr. Klaus Schlunzen Junior
Departamento de Estatística

Presidente Prudente, 19 de dezembro de 2025.

RESUMO

Este trabalho analisou a frequência mensal de dias com chuva na cidade de São Paulo entre 1990 e 2019, considerando a proporção de dias com precipitação igual ou superior a 1 mm. Com base nas características da variável contínua e restrita ao intervalo (0,1) empregou-se a distribuição Beta, incluindo o ajuste mensal e global. Em seguida, foram ajustados modelos de regressão Beta para investigar efeitos temporais e sazonais, bem como possíveis interações entre ano e mês. Os resultados mostraram forte sazonalidade, com valores mais elevados nos meses de verão e menores no período seco, enquanto não foram observadas tendências temporais gerais; entretanto, a interação Ano $\times$ Mês indicou redução significativa da proporção de dias chuvosos em meses específicos, como fevereiro e julho.

Em uma segunda etapa, incorporou-se o índice Oceanic Niño Index (ONI), a fim de avaliar a influência do fenômeno El Niño–Oscilação Sul (ENSO) sobre a variabilidade pluviométrica da cidade. As análises descritivas mostraram maiores proporções médias de dias chuvosos durante eventos de El Niño e La Niña em comparação à fase neutra. O teste de Kruskal–Wallis confirmou diferenças estatísticas entre as fases, e os modelos de regressão Beta incluindo ENSO revelaram coeficiente significativo após controle dos efeitos sazonais, evidenciando que episódios de aquecimento ou resfriamento do Pacífico influenciam o regime de chuva local.

Por fim, avaliou-se a dependência conjunta entre proporção de dias chuvosos e precipitação total mensal utilizando cópulas, com destaque para a cópula de Clayton, aplicada separadamente para cada fase do ENSO. Observou-se forte dependência em todas as fases, com intensidade ligeiramente maior durante La Niña, indicando maior coerência entre frequência e volume total de chuva. Nos anos de El Niño, a dependência foi levemente menor, sugerindo maior ocorrência de eventos intensos concentrados em poucos dias.

A relevância deste estudo reside no fornecimento de evidências estatísticas que aprimoram a compreensão da variabilidade pluviométrica em São Paulo, contribuindo para o monitoramento climático, o planejamento urbano e ações de adaptação e prevenção voltadas à Defesa Civil. A integração entre ENSO, frequência de chuva e precipitação total permite identificar meses e condições climáticas mais suscetíveis a extremos, oferecendo subsídios técnicos para decisões estratégicas em gestão de risco meteorológico.

Palavras-chave: Chuva. ENSO. El Niño. La Niña. Distribuição Beta. Regressão Beta. Sazonalidade. Cópulas. Dependência.

ABSTRACT

This study analyzed the monthly frequency of rainy days in the city of São Paulo between 1990 and 2019, considering the proportion of days with precipitation equal to or greater than 1 mm. Given the continuous nature of the variable restricted to the (0,1) interval, the Beta distribution was employed, including both global and monthly fits. Subsequently, Beta regression models were used to investigate temporal and seasonal effects, as well as interactions between year and month. The results revealed strong seasonality, with higher values in summer and lower values during the dry season. No general temporal trends were detected, although the Year $\times$ Month interaction indicated significant reductions in specific months, such as February and July.

In the second stage, the Oceanic Niño Index (ONI) was incorporated to evaluate the role of the El Niño–Southern Oscillation (ENSO) in the rainfall variability of São Paulo. Descriptive analyses showed higher average proportions of rainy days during El Niño and La Niña events compared to neutral conditions. The Kruskal–Wallis test confirmed statistical differences among ENSO phases, and Beta regression models including ENSO revealed a significant effect once seasonality was controlled, demonstrating that Pacific Ocean temperature anomalies influence the local rainfall regime.

Finally, the joint dependence between the proportion of rainy days and total monthly precipitation was assessed using copulas, particularly the Clayton copula, applied separately for each ENSO phase. Strong dependence was observed across all phases, with slightly higher intensity during La Niña, and slightly weaker dependence during El Niño, suggesting a higher occurrence of intense rainfall concentrated in fewer days.

The relevance of this study lies in providing statistical evidence that enhances the understanding of rainfall variability in São Paulo, supporting climate monitoring, urban planning, and Civil Defense preparedness. The integration of ENSO dynamics, frequency of rainy days, and precipitation volume offers valuable insights for identifying months and climatic conditions more prone to extreme events, thus providing technical support for risk management decisions.

Keywords: Rainfall. ENSO. El Niño. La Niña. Beta Distribution. Beta Regression. Seasonality. Copulae. Dependence.

LISTA DE FIGURAS

Figura 1: Exemplo dos dados de precipitação diária obtidos a partir do INMET.....	15
Figura 2: Estrutura dos dados iniciais utilizados antes das transformações.	16
Figura 3 Exemplos de formas da distribuição Beta:.....	20
Figura 4:Boxplot da proporção mensal de dias chuvosos por mês.	26
Figura 5: Mapa de calor da proporção mensal de dias com chuva por mês e ano (1990-2019).	27
Figura 6: Histograma da proporção de dias com chuva e ajuste Beta.	28
Figura 7: Histogramas da proporção de dias com chuva com ajustes Beta por mês.	29
Figura 8: Proporções observadas e ajustadas pelo modelo Beta com efeito sazonal. ..	32
Figura 9:Resíduos padronizados versus valores ajustados do modelo Beta completo ($Proporção_{tr} \sim ENSO_{index} + Mês + Ano$).	36

LISTA DE TABELAS

Tabela 1: Estatísticas descritivas da proporção mensal de dias chuvosos.	26
Tabela 2: Estimativas dos parâmetros da distribuição Beta ajustada	28
Tabela 3: Estimativa dos parâmetros da distribuição Beta por mês	30
Tabela 4: Coeficientes estimados da regressão beta (Proporção de dias com chuva ~ Ano).	31
Tabela 5: Coeficientes estimados do modelo de regressão Beta com efeito sazonal.	31
Tabela 6: Comparação entre modelos de regressão Beta: simples (Ano) e completo (Ano + Mês).	32
Tabela 7: Coeficientes de interação significativos ou próximo do nível de significância (Ano x Mês).	33
Tabela 8: Estatísticas descritivas da proporção de dias chuvosos por fase do ENSO.	34
Tabela 9: Comparação entre modelos Beta com ENSO.	37
Tabela 10: Parâmetros da cópula de Clayton estratificados por ENSO.	38

SUMÁRIO

1 INTRODUÇÃO	9
1.1 OBJETIVOS.....	12
2 REFERENCIAL TEÓRICO.....	12
3 BASE DE DADOS.....	15
3.1 DADOS DE PRECIPITAÇÃO (INMET).....	15
3.2 DADOS DO FENÔMENO EL NIÑO–OSCILAÇÃO SUL (ENSO)	16
3.3 INTEGRAÇÃO DAS BASES	17
4 METODOLOGIA	17
5 FUNDAMENTAÇÃO TEÓRICA DA MODELAGEM E REGRESSÃO BETA	19
5.1 DISTRIBUIÇÕES PARA PROPORÇÕES E MOTIVAÇÃO PARA A DISTRIBUIÇÃO BETA	19
5.2 DISTRIBUIÇÃO BETA.....	20
5.3 TRATAMENTO DE PROPORÇÕES COM VALORES 0 E 1.....	21
5.4 MODELAGEM DE PROPORÇÕES VIA REGRESSÃO BETA.....	21
5.5 INTERPRETAÇÃO DOS PARÂMETROS NA REGRESSÃO BETA	22
6 TEORIA DAS CÓPULAS.....	22
6.1 TEOREMA DE SKLAR	23
6.2 TRANSFORMAÇÕES DAS MARGINAIS PARA O INTERVALO (0,1).....	23
6.3 CÓPULA DE CLAYTON.....	23
6.4 ESTIMAÇÃO DO PARÂMETRO	24
7 RESULTADOS E DISCUSSÃO	25
7.1 ANÁLISE DESCRITIVA DA PROPORÇÃO DE DIAS CHUVOSOS.....	25
7.2 AJUSTE DA DISTRIBUIÇÃO BETA.....	28
7.3 AJUSTE MENSAL DA DISTRIBUIÇÃO BETA	29
7.4 REGRESSÃO BETA	30
7.5 REGRESSÃO BETA COM INTERAÇÃO ANO X MÊS	32
7.6 INFLUÊNCIA DO ENSO NA PROPORÇÃO DE DIAS CHUVOSOS.....	33
7.7 MODELOS DE REGRESSÃO BETA COM ENSO	34
7.7.1 MODELO 1: ENSO CONTÍNUO.....	35
7.7.2 MODELO 2: ENSO + MÊS + ANO	35
7.7.3 MODELO 3: INTERAÇÃO ENSO × MÊS	36
7.7.4 COMPARAÇÃO ENTRE MODELOS.....	37
7.8 DEPENDÊNCIA ENTRE PROPORÇÃO E PRECIPITAÇÃO VIA CÓPULAS ESTRATIFICADAS POR ENSO	37
7.9 SÍNTESE INTEGRADA DOS RESULTADOS	38
8 CONCLUSÃO	40
9 REFERÊNCIAS.....	ERROR! BOOKMARK NOT DEFINED.

1 INTRODUÇÃO

Durante toda a história da humanidade, o estudo e o conhecimento do clima foram fundamentais para o desenvolvimento humano, permitindo que as civilizações se adaptassem e progredissem em ambientes diversos. Isso pode ser evidenciado por diversos estudos, como IPCC (2014, 2021), Pielke e Niyogi (2010) e Rosenzweig et al. (2004), que mostram a influência climática em áreas como agricultura, infraestrutura urbana, gestão de recursos, política, entre outras.

A precipitação pluvial exerce um papel importantíssimo no equilíbrio ambiental e no desenvolvimento econômico das regiões. Em São Paulo, uma cidade caracterizada pelo clima subtropical úmido, segundo a classificação de Köppen-Geiger, aliado à intensa urbanização, observa-se uma variabilidade significativa nas distribuições das chuvas, o que pode causar impactos no abastecimento de água, na agricultura e na gestão de recursos hídricos (Saad, 1990). Considerando esses fatores, evidencia-se a necessidade de uma análise que busque compreender e, possivelmente, minimizar danos nessas áreas, especialmente por meio de um estudo relacionado ao número de dias de chuva, uma vez que essa variável está associada a eventos extremos, como secas prolongadas e inundações.

No caso da cidade de São Paulo, o elevado grau de urbanização, aliado à variabilidade climática, torna a região sensível a impactos hidrológicos negativos. A impermeabilização do solo e a canalização de rios alteram o escoamento da água, dificultando sua infiltração e contribuindo para a maior ocorrência de alagamentos. Dessa forma, torna-se fundamental uma compreensão mais ampla sobre os padrões mensais de dias de chuva, pois esses dados podem fornecer informações essenciais para otimizar sistemas de drenagem, aprimorar a gestão de recursos hídricos e orientar o desenvolvimento urbano sustentável.

A variabilidade das chuvas e seus impactos na cidade de São Paulo podem ser observados em eventos recentes, como o ocorrido em abril de 2025, quando, em um intervalo de apenas dois dias, registrou-se uma precipitação superior à média prevista para todo o mês (UOL, 2025). Diante dessa situação, a Defesa Civil orientou a população sobre como se proteger e identificar possíveis riscos, incluindo deslizamentos e estruturas comprometidas.

Outro aspecto relevante refere-se à segurança climática e à gestão de riscos. A análise da frequência de dias de chuva ao longo dos meses permite identificar

períodos historicamente mais sensíveis a incidentes como deslizamentos e alagamentos, fornecendo informações importantes para órgãos responsáveis, como a Defesa Civil.

Os impactos da variabilidade climática podem ser observados também em outras regiões do mundo. Estudos sobre a monção na Índia, por exemplo, demonstraram a relevância da quantidade e da regularidade da precipitação para o desenvolvimento agrícola e, conseqüentemente, para a economia do país (Kumar et al., 1992; Joshi et al., 2005). Esses casos reforçam a importância das análises climáticas, que permitem a adoção de estratégias de adaptação diante de incertezas ambientais. Embora as características climáticas dessas regiões sejam distintas das de São Paulo, as lições aprendidas sobre gestão de recursos hídricos e práticas agrícolas podem servir de referência para a elaboração de estratégias adaptativas que mitiguem os desafios impostos pela variabilidade climática.

No contexto da climatologia, estudos e análises sobre dados de chuva frequentemente se concentram no volume total de precipitação, em razão de sua relevância direta. Contudo, a frequência com que ocorrem os eventos de chuva também merece atenção, oferecendo uma perspectiva igualmente importante. Isso porque o volume mensal de precipitação pode ser influenciado por eventos extremos isolados, como citado anteriormente, o que não indica necessariamente uma mudança climática. Assim, analisar o número de dias de chuva possibilita avaliar a distribuição desses eventos, aspecto diretamente relacionado ao cotidiano urbano. No caso de São Paulo, uma cidade densamente urbanizada, a frequência de chuva afeta o trânsito, a prestação de serviços e a segurança da população frente a riscos como deslizamentos e alagamentos. Portanto, o aprofundamento estatístico sobre a frequência de dias chuvosos pode revelar padrões climáticos que não seriam identificados apenas por meio da análise do volume total de precipitação.

Esse tipo de estudo permite uma compreensão mais detalhada das mudanças no regime de chuvas e auxilia na tomada de decisões frente aos impactos e desafios ambientais e urbanos. Dessa forma, a análise estatística da frequência de dias de chuva mensal na cidade de São Paulo torna-se essencial, com o objetivo de identificar tendências e alterações no comportamento climático. Com base nesses resultados, torna-se possível aprimorar o entendimento climático da região, além de fornecer subsídios para a tomada de decisões e o desenvolvimento de estratégias de gestão urbana e ambiental, visando mitigar impactos observados.

O estudo da frequência de dias chuvosos, além de sua relevância local, é importante também no cenário internacional, especialmente diante das mudanças climáticas. Segundo o Painel Intergovernamental sobre Mudanças Climáticas (IPCC, 2014), existem evidências de que o aumento da temperatura média global influencia os padrões de precipitação, inclusive em áreas urbanas. Essa variabilidade climática também afeta a ocorrência de eventos extremos, como tempestades e longos períodos de seca (IPCC, 2021).

Além da perspectiva ambiental e científica, compreender essa variável é igualmente importante sob o ponto de vista social. A população de São Paulo enfrenta diariamente os impactos da variabilidade climática, como enchentes repentinas, congestionamentos e atrasos no transporte público. Esses efeitos atingem de forma mais intensa as populações vulneráveis, evidenciando que a análise desses dados ultrapassa questões ambientais e de planejamento urbano, envolvendo também aspectos relacionados ao bem-estar coletivo.

Apesar da ampla existência de estudos sobre o comportamento da precipitação em São Paulo, como Tarifa e Armani (2001) e Reboita et al. (2010), a maioria concentra-se no volume total acumulado mensalmente a maioria concentra-se no volume total acumulado mensalmente. No entanto, ainda são poucos os que tratam a frequência de dias de chuva como variável principal, o que representa uma lacuna na compreensão da regularidade das chuvas e de seu comportamento ao longo do tempo. Essa lacuna limita a atuação de órgãos públicos e planejadores urbanos, que frequentemente utilizam análises que não contemplam adequadamente essa dimensão. A frequência de eventos de chuva é crucial, pois está diretamente associada à percepção da população sobre a presença ou ausência de precipitação, além de afetar o manejo de espaços públicos, sistemas de drenagem e mecanismos de alerta.

Dessa forma, analisar a frequência de dias chuvosos por mês no período de 1990 a 2019 na cidade de São Paulo representa uma oportunidade de preencher essa lacuna e fornecer uma visão mais detalhada da variabilidade climática regional. Esse intervalo temporal foi selecionado devido à qualidade e à completude dos dados disponibilizados pelo Instituto Nacional de Meteorologia (INMET), uma vez que, a partir de 2019, observou-se um aumento expressivo no número de registros ausentes, o que comprometeria a consistência das séries históricas e a confiabilidade das análises estatísticas. Essa perspectiva, alinhada aos dados históricos mais robustos,

pode revelar tendências e fornece uma base sólida para decisões que aumentem a resiliência urbana diante dos desafios climáticos atuais e futuros.

Além disso, considerando o papel de fenômenos climáticos de grande escala na variabilidade das chuvas, este estudo incorpora uma análise complementar da influência do fenômeno El Niño–Oscilação Sul (ENSO) sobre a frequência de dias chuvosos ao longo do período analisado.

1.1 OBJETIVOS

O objetivo deste trabalho é, a partir dos dados de precipitação da cidade de São Paulo, realizar análises estatísticas voltadas para a frequência de dias de chuva mensais no período de 1990 a 2019. Pretende-se identificar padrões sazonais, variações entre os meses e possíveis indícios de mudanças no comportamento climático urbano.

Além disso, busca-se obter resultados por meio de métodos estatísticos, como análises descritivas, ajustes a distribuições teóricas e técnicas de regressão, com foco na distribuição Beta.

De forma complementar, será realizada uma análise conjunta entre a frequência de dias com chuva e a quantidade de precipitação mensal por meio das teorias das cópulas, com o objetivo de examinar a dependência entre essas variáveis.

Por fim, incorporou-se ao estudo uma análise adicional sobre a influência do fenômeno El Niño–Oscilação Sul (ENSO), avaliando de que forma suas diferentes fases podem afetar a frequência mensal de dias chuvosos e a estrutura de dependência entre as variáveis analisadas.

2 REFERENCIAL TEÓRICO

Este projeto utiliza como área de estudo a cidade de São Paulo, que apresenta um clima caracterizado como subtropical úmido, com chuvas irregulares e temperaturas elevadas ao longo do ano. Em Souza et al. (2017), observa-se a importância da variabilidade climática e da precipitação devido à sua influência direta sobre as atividades humanas e as dinâmicas ambientais. Os autores destacam que, no semiárido brasileiro, uma das principais causas dessa variabilidade são os fenômenos El Niño–Oscilação Sul (ENOS) e o Dipolo do Atlântico Tropical (DAT), que alteram significativamente o comportamento das chuvas. Embora o estudo tenha

como foco cidades do Nordeste, essa abordagem metodológica e os impactos observados servem como parâmetro para compreender o comportamento climático em áreas urbanas, como São Paulo, onde a distribuição e a dinâmica das chuvas afetam diretamente a gestão de riscos, a infraestrutura e o abastecimento hídrico.

Além disso, os autores ressaltam que, mesmo em regiões como o Nordeste brasileiro, onde os índices pluviométricos são baixos, a distribuição dos dias de chuva é uma variável relevante, especialmente no caso dos veranicos períodos sem precipitação, com forte insolação e calor intenso, mesmo durante a estação chuvosa que impactam a agricultura e os recursos hídricos. Analisar como os dias de chuva se distribuem ao longo dos meses permite compreender a ocorrência de eventos extremos, como estiagens prolongadas e inundações. Essa lógica pode ser aplicada também a grandes centros urbanos, como São Paulo, onde a variabilidade das chuvas influencia diretamente a dinâmica urbana.

Em Silva et al. (2012), observa-se uma análise realizada no Nordeste sobre a variabilidade anual e intra-anual da precipitação, destacando a importância do número de dias de chuva como indicador climático complementar ao volume total. Os autores utilizaram séries históricas provenientes da Agência Nacional de Águas (ANA) e da Superintendência do Desenvolvimento do Nordeste (SUDENE) entre 1935 e 2000, realizando análises como coeficiente de variação, mapas de distribuição espacial e avaliação da variabilidade da frequência de dias de chuva nos períodos anuais, chuvosos e secos. Identificaram diferenças expressivas entre o litoral e o semiárido, sendo este último mais suscetível a oscilações intensas, especialmente em períodos de estiagem. Esse estudo demonstra que a frequência dos dias de chuva fornece informações relevantes sobre o clima de uma região, permitindo identificar períodos de maior insegurança hídrica. Embora o foco seja o Nordeste, suas metodologias e conclusões também são aplicáveis a áreas urbanas como São Paulo, nas quais a frequência de dias chuvosos impacta diretamente o cotidiano, como mencionado anteriormente.

Estudos sobre a frequência dos dias de chuva têm ganho visibilidade e importância até mesmo internacionalmente. Em Kumar et al. (2021), uma análise foi realizada no distrito de Navsari, na Índia, utilizando distribuições binomial e normal para investigar a distribuição semanal e mensal de dias chuvosos com base em dados diários de 36 anos. Os resultados demonstraram que tais distribuições foram eficazes para analisar a variabilidade da frequência de chuva e identificar tendências sazonais,

como o aumento da ocorrência de dias chuvosos durante o período de monções. Embora o estudo tenha foco agrícola, assim como outros trabalhos citados, a estratégia metodológica adotada é aplicável a análises voltadas a ambientes urbanos, oferecendo uma compreensão mais precisa da dinâmica climática.

A análise conjunta da distribuição dos dias de chuva e da quantidade total de precipitação tem se mostrado eficaz para aplicações em agricultura, climatologia e planejamento. No estudo de Haider e Hussain (1989), baseado em uma série diária de 43 anos no distrito de Faisalabad, no Paquistão, os autores investigaram a distribuição estatística da precipitação e as probabilidades associadas aos números de dias com chuva mensais e por décadas. Como metodologia, aplicaram testes de normalidade, como o qui-quadrado, e ajustaram distribuições teóricas para estimar probabilidades de ocorrência de diferentes quantidades de dias de chuva. Destacaram que, embora analisar a precipitação total seja fundamental, a frequência de dias chuvosos fornece informações complementares de grande relevância. A conclusão foi que apenas os meses de julho e agosto apresentaram distribuição normal, enquanto os demais exibiram distribuições assimétricas, reforçando a importância de avaliar essa variável separadamente.

Um estudo realizado por Silva et al. (2021) na bacia hidrográfica do rio Gurupi, em Cachoeira do Piriá (PA), teve como objetivo analisar séries históricas de precipitação ao longo de 30 anos, utilizando metodologias como tendência linear, Mann–Kendall, Spearman e Sen's Slope. Com base nesses métodos, os autores identificaram padrões sazonais bem definidos, com maior frequência de dias chuvosos entre janeiro e junho, além de avaliar possíveis tendências significativas ao longo do período. Embora a conclusão tenha indicado ausência de tendência significativa, o estudo reforça a utilidade dessas análises para apoio ao planejamento hídrico e ambiental.

Considerando os fenômenos de grande escala, como o ENOS, é importante destacar que suas diferentes fases El Niño e La Niña exercem influência reconhecida sobre o regime de chuvas no Brasil, ainda que com intensidades distintas entre regiões. O El Niño tende, em muitos casos, a modificar padrões de circulação atmosférica, o que pode resultar em alterações na distribuição temporal das chuvas, enquanto a La Niña frequentemente promove condições de maior regularidade pluviométrica em algumas regiões. Embora o presente estudo utilize São Paulo como área de referência região com padrões climáticos distintos dos citados a atuação do

ENOS constitui um componente relevante da variabilidade climática nacional e, portanto, fundamenta sua inclusão como variável complementar neste trabalho.

3 BASE DE DADOS

3.1 Dados de precipitação (INMET)

Para a realização deste trabalho, utilizou-se a base de dados de precipitação da cidade de São Paulo referente ao período de 1990 a 2019, obtida no site oficial do Instituto Nacional de Meteorologia (INMET), a partir da estação meteorológica Mirante de Santana. Na estrutura original da base, as informações disponíveis consistem nas datas das medições e na quantidade total de precipitação diária, expressa em milímetros (mm), conforme ilustrado na Figura 1.

Figura 1: Exemplo dos dados de precipitação diária obtidos a partir do INMET.

	Data	Precipitacao
1	Data Medicao	PRECIPITACAO TOTAL, DIARIO(mm)
2	01/01/1990	15.5
3	02/01/1990	35.9
4	03/01/1990	16.3
5	04/01/1990	5.2
6	05/01/1990	11.9
7	06/01/1990	9.4
8	07/01/1990	15.7
9	08/01/1990	5.9
10	09/01/1990	0
11	10/01/1990	9.3
12	11/01/1990	4.7

Fonte: bdmep.inmet.gov.br

A proposta central deste estudo envolve a análise da frequência de dias de chuva ao longo dos meses. Para isso, os dados diários foram organizados de forma a permitir sua posterior transformação em métricas mensais, incluindo a contagem de dias com chuva e a agregação da precipitação diária. Essa estrutura consiste na base inicial utilizada para a construção das variáveis analisadas ao longo do trabalho.

Figura 2: Estrutura dos dados iniciais utilizados antes das transformações.

	Ano	Mes	AnoMes	DiasTotais	DiasComChuva	TotalPrecipitacao	Proporcao
1	1990	1	1990-01	31	16	280.7	0.51612903
2	1990	2	1990-02	28	13	152.8	0.46428571
3	1990	3	1990-03	31	16	228.8	0.51612903
4	1990	4	1990-04	30	4	77.1	0.13333333
5	1990	5	1990-05	31	5	52.8	0.16129032
6	1990	6	1990-06	30	3	39.2	0.10000000
7	1990	7	1990-07	31	14	121.0	0.45161290
8	1990	8	1990-08	31	5	49.6	0.16129032
9	1990	9	1990-09	30	8	96.1	0.26666667
10	1990	10	1990-10	31	11	117.6	0.35483871
11	1990	11	1990-11	30	11	76.0	0.36666667
12	1990	12	1990-12	31	10	124.7	0.32258065
13	1991	1	1991-01	31	13	270.7	0.41935484
14	1991	2	1991-02	28	15	358.0	0.53571429
15	1991	3	1991-03	31	24	451.3	0.77419355

3.2 Dados do fenômeno El Niño–Oscilação Sul (ENSO)

Além da base de precipitação, este estudo incorporou informações referentes ao fenômeno El Niño–Oscilação Sul (ENSO), devido à sua influência reconhecida sobre o regime de chuvas no Brasil. Para isso, utilizou-se o Oceanic Niño Index (ONI), disponibilizado pela NOAA/Climate Prediction Center, amplamente empregado como medida operacional do ENSO.

O ONI é calculado como a média móvel trimestral das anomalias de temperatura da superfície do mar na região Niño 3.4 (5°N–5°S, 120°–170°W). Segundo a classificação da NOAA:

- Valores $\geq +0,5^{\circ}\text{C}$ indicam El Niño,
- Valores $\leq -0,5^{\circ}\text{C}$ indicam La Niña,
- Valores entre $-0,5^{\circ}\text{C}$ e $+0,5^{\circ}\text{C}$ representam fase neutra.

Para o presente estudo, foram utilizados os valores correspondentes ao período de 1990 a 2019, de forma a manter compatibilidade temporal com a série pluviométrica. Os dados foram organizados no formato largo (wide), contendo para cada ano as 12 médias trimestrais correspondentes aos períodos: DJF, JFM, FMA, MAM, AMJ, MJJ, JJA, JAS, ASO, SON, OND e NDJ.

Essas informações foram posteriormente transformadas em:

- ENSO_index: valor numérico contínuo da ONI associado a cada mês;
- ENSO_fase: categorização em El Niño, La Niña ou Neutro.

A utilização do ENSO como variável complementar permite investigar se suas fases influenciam a frequência de dias chuvosos na cidade de São Paulo, além de auxiliar na identificação de anos atípicos observados nas análises descritivas.

3.3 Integração das Bases

Para integrar as informações do ONI aos dados mensais de precipitação, cada mês da série histórica foi associado ao valor do ONI referente ao trimestre móvel no qual ele se encontra. Por exemplo:

- Janeiro → DJF
- Fevereiro → DJF
- Março → JFM
- Abril → MAM
- Etc.

Esse procedimento é amplamente utilizado em estudos climatológicos, uma vez que o ENSO atua de forma gradual, sendo melhor representado por médias móveis trimestrais do que por valores pontuais mensais. Após essa associação, a série passou a conter, para cada mês:

- proporção de dias chuvosos,
- Precipitação total mensal,
- Fase ENSO correspondente,
- Índice ENSO contínuo.

Essa estrutura permitiu a realização das análises complementares por meio de métodos descritivos, testes não paramétricos e modelos de regressão Beta, além de investigações envolvendo a teoria das cópulas, estratificando-se as dependências entre variáveis de acordo com a fase do ENSO.

4 METODOLOGIA

Nesta seção, descreve-se o procedimento adotado para a análise da proporção de dias com chuva na cidade de São Paulo durante o período de 1990 a 2019. A metodologia compreende desde a aquisição e preparação da base de dados até a aplicação das técnicas estatísticas de modelagem, bem como a interpretação dos resultados obtidos.

Inicialmente, selecionou-se o software R, devido à sua ampla utilização em estudos estatísticos e à variedade de pacotes disponíveis para manipulação, transformação e modelagem de dados. A base de dados foi obtida a partir do Instituto Nacional de Meteorologia (INMET), referente à estação meteorológica do Mirante de Santana. Após a importação dos dados originais compostos exclusivamente pela data e pela precipitação total diária em milímetros foram realizadas as etapas de limpeza e organização da base.

Para possibilitar as análises propostas, os dados diários foram agregados em escala mensal. A partir da definição operacional de dia chuvoso, adotou-se o critério de considerar como chuva apenas dias com precipitação superior a 1 mm, valor amplamente utilizado pelo INMET, o que exclui ocorrências como garoa leve, orvalho ou medições muito pequenas. Utilizando os pacotes *dplyr* (para manipulação) e *lubridate* (para tratamento de datas), foram construídas as variáveis mensais:

- número de dias com chuva (≥ 1 mm);
- número total de dias do mês;
- precipitação total mensal;
- proporção de dias chuvosos.

Esse formato da base é fundamental para as etapas seguintes, que envolvem modelagem da proporção de dias chuvosos, ajustes de distribuições teóricas e avaliação de dependência entre variáveis.

A análise estatística iniciada utilizou, primeiramente, procedimentos descritivos, com o objetivo de identificar padrões gerais, diferenças sazonais e possíveis anos atípicos. Em seguida, a proporção de dias de chuva foi analisada como uma variável contínua no intervalo (0,1), justificando-se o uso da distribuição Beta. Foram ajustados modelos Beta via máxima verossimilhança, com o objetivo de estimar os parâmetros de forma α e β . A adequação dos modelos foi avaliada graficamente e por meio da análise de resíduos. Adicionalmente, com o intuito de investigar possíveis tendências temporais, aplicaram-se modelos de regressão Beta, permitindo estimar a probabilidade de ocorrência de meses com elevada frequência de dias chuvosos.

Para complementar o estudo, incorporou-se a teoria das cópulas, empregando especificamente a cópula de Clayton, com o objetivo de avaliar a dependência conjunta entre a proporção de dias chuvosos e a precipitação total mensal. Inicialmente, ajustaram-se separadamente as distribuições marginais: Beta para a

proporção de dias chuvosos e Gamma para a precipitação total. Após a transformação das marginais para o intervalo (0,1), realizou-se o ajuste da cópula por máxima verossimilhança. Tal abordagem permite investigar a intensidade da dependência entre as variáveis, especialmente quanto à ocorrência simultânea de meses secos ou de comportamento extremo no regime pluviométrico.

Por fim, considerando a influência de fenômenos climáticos de grande escala na variabilidade das chuvas, foi conduzida uma análise complementar envolvendo a relação entre a frequência de dias chuvosos e as fases do El Niño–Oscilação Sul (ENSO) ao longo do período estudado. Essa etapa auxilia na interpretação de eventuais anomalias e contribui para a compreensão mais ampla dos padrões observados.

5 FUNDAMENTAÇÃO TEÓRICA DA MODELAGEM E REGRESSÃO BETA

5.1 Distribuições para Proporções e Motivação para a Distribuição Beta

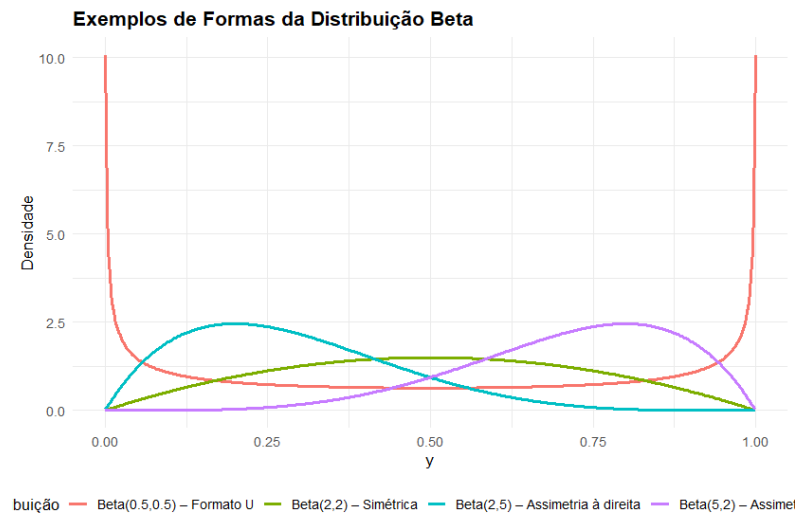
Variáveis do tipo proporção são representadas por valores contínuos contidos no intervalo (0,1). Elas aparecem com frequência em estudos ambientais, econômicos, biológicos e climáticos, como no presente trabalho, que inclui uma análise referente à proporção de dias de chuva.

Para uma modelagem adequada de uma variável dessa natureza, é necessário utilizar distribuições que respeitem o intervalo (0,1). Assim, modelos clássicos, como regressão linear ou normal, tornam-se inadequados pelos seguintes motivos:

- podem gerar previsões fora do intervalo (0,1);
- assumem variância constante, o que não é apropriado para proporções;
- não captam assimetrias, comuns quando há concentração de valores próximos de 0 ou de 1.

Diante dessas limitações, a distribuição Beta destaca-se como uma das mais utilizadas no contexto de modelagem de proporções. Trata-se de uma distribuição contínua com alta flexibilidade, capaz de assumir diferentes formatos, conforme ilustrado na Figura 3.

Figura 3 Exemplos de formas da distribuição Beta:



Fonte: Elaborado pelo Autor.

Essa versatilidade é especialmente útil em estudos climáticos, em que diferentes meses podem apresentar comportamentos distintos: alguns mais estáveis e outros mais dispersos ou assimétricos.

5.2 Distribuição Beta

A função densidade da distribuição Beta no intervalo (0,1) é dada por:

$$f(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 < x < 1$$

Onde:

- $\alpha > 0$ e $\beta > 0$ são os parâmetros de forma.
- Γ é a função gama.

As principais propriedades utilizadas para interpretação são:

- $E(Y) = \frac{\alpha}{\alpha + \beta}$
- $Var(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$

Essas expressões mostram que a variância depende da média, diferentemente do caso normal. Além disso, a distribuição é controlada por apenas dois parâmetros e permite capturar assimetrias, algo comum na proporção de dias com chuva. Essa flexibilidade contribui para a análise de dados pluviométricos, nos quais é recorrente a concentração de valores próximos de 0 em meses secos e próximos de 1 em meses mais chuvosos.

5.3 Tratamento de Proporções com Valores 0 e 1

Um aspecto fundamental para a utilização da distribuição Beta é que seus valores devem estar estritamente no intervalo (0,1). No entanto, dados climatológicos podem apresentar meses com proporção exatamente igual a 0 (por exemplo, meses de inverno sem chuva) ou igual a 1 (casos raros, mas possíveis).

Para lidar com esse problema, um procedimento amplamente aceito é a transformação proposta por Smithson e Verkuilen (2006), que garante que todos os valores estejam dentro do intervalo (0,1):

$$Y^* = \frac{Y(n - 1) + 0.5}{n}$$

Onde:

- Y é a proporção original;
- n é o número total de observações;
- Y^* é a nova variável utilizada no ajuste da distribuição Beta.

5.4 Modelagem de Proporções via Regressão Beta

A regressão Beta é uma extensão da distribuição Beta que permite incluir variáveis explicativas e modelar a média da proporção. O modelo assume:

$$Y_i \sim \text{Beta}(\mu_i, \phi),$$

onde:

- μ_i é a média da distribuição Beta para o mês/ano i ;
- ϕ é o parâmetro de precisão.

A ligação entre μ_i e os preditores ocorre por meio de uma função de ligação, geralmente o logit:

$$\text{logit}(\mu_i) = \eta_i = x_i \beta.$$

Esse modelo permite manter os valores no intervalo (0,1), além de possibilitar a representação de efeitos lineares e não lineares, inclusão de interações, consideração de sazonalidade e investigação de tendências temporais. Assim, torna-se possível explorar questões como:

- se a proporção média de dias chuvosos depende do mês;
- se há tendência temporal;
- se há variações significativas entre os meses.

5.5 Interpretação dos Parâmetros na Regressão Beta

A interpretação dos coeficientes estimados é feita na escala logit. Dessa forma:

- um coeficiente positivo indica aumento na média da proporção de dias com chuva;
- um coeficiente negativo indica redução.

Além dos coeficientes, a regressão Beta fornece métricas de avaliação do ajuste, como:

- Pseudo- R^2 , que expressa a qualidade do modelo;
- AIC, útil para comparar diferentes modelos ajustados.

6 TEORIA DAS CÓPULAS

Nesta etapa do trabalho, aborda-se a relação entre duas variáveis fundamentais em estudos climatológicos: a proporção mensal de dias chuvosos e a quantidade total de precipitação mensal. Comumente, análises desse tipo utilizam métricas como a correlação de Pearson; entretanto, tais métodos apresentam limitações importantes. A correlação linear captura apenas a dependência simétrica e linear entre variáveis, tornando-se insuficiente quando a associação é mais intensa em regiões específicas da distribuição, como valores simultaneamente baixos, ou quando a relação é nitidamente não linear.

Diante disso, as cópulas surgem como uma ferramenta apropriada para modelar dependências complexas. O ponto central dessa abordagem é a capacidade de separar as distribuições marginais da estrutura de dependência, permitindo analisar o comportamento conjunto sem impor suposições rígidas sobre as marginais. Com isso, obtém-se maior flexibilidade e aderência ao comportamento real dos dados climáticos.

6.1 TEOREMA DE SKLAR

O fundamento que sustenta toda a teoria das cópulas é o *Teorema de Sklar* (1959). Ele estabelece que, considerando duas variáveis aleatórias X e Y , com distribuições marginais F_X e F_Y , a distribuição conjunta pode ser representada por meio de uma cópula C definida no quadrado unitário $(0,1)^2$, tal que:

$$F_{X,Y}(x, y) = C(F_X(x), F_Y(y)).$$

Assim, as informações sobre a dependência entre as variáveis ficam completamente contidas na função C . Isso garante maior flexibilidade para as análises possíveis e permite ajustar separadamente as marginais e a estrutura de dependência.

6.2 Transformações das marginais para o intervalo (0,1)

Para utilizar cópulas, é necessário transformar as variáveis originais por meio de suas distribuições acumuladas, garantindo que seus valores fiquem no intervalo $(0,1)$. Esse procedimento é chamado de uniformização, sendo dado por:

$$u_1 = F_X(x), \quad u_2 = F_Y(y).$$

No presente estudo, as variáveis utilizadas são:

- A proporção mensal de dias chuvosos, modelada por uma distribuição Beta;
- A precipitação total mensal, ajustada por uma distribuição Gamma.

Assim, a transformação ocorre da seguinte forma:

$$u_1 = F_{Beta}(Proporção), \quad u_2 = F_{Gamma}(Precipitação).$$

6.3 CÓPULA DE CLAYTON

Dentro da teoria das cópulas, existe uma ampla variedade de famílias, cada uma adequada a tipos específicos de dependência. As cópulas arquimedianas

incluem Clayton, Gumbel e Frank; já as cópulas elípticas incluem a Normal e a *t-Student*, entre outras.

A cópula escolhida neste trabalho é a Clayton, caracterizada por capturar de forma eficiente a dependência em eventos simultaneamente baixos — isto é, quando a associação é mais forte na cauda inferior. Esse comportamento justifica sua escolha, já que, no contexto pluviométrico, meses com poucos dias de chuva tendem a apresentar precipitação total baixa. A cópula de Clayton é definida por:

$$C(u_1, u_2; \theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}, \quad \theta > 0.$$

O parâmetro θ controla a intensidade da dependência: quanto maior seu valor, maior a associação entre as variáveis. Uma medida bastante utilizada é a tau de Kendall, que para a cópula de Clayton assume a forma:

$$\tau = \frac{\theta}{\theta + 2},$$

permitindo interpretar a dependência de forma direta e comparável entre estudos.

Outra medida importante, especialmente em pesquisas sobre escassez de chuvas, é a probabilidade de valores simultaneamente baixos nas duas variáveis, dada por:

$$\lambda_L = 2^{-\frac{1}{\theta}}.$$

6.4 Estimação do parâmetro

O método utilizado segue a abordagem tradicional em hidrologia, meteorologia e áreas afins, sendo amplamente empregado devido à eficiência assintótica e ao bom desempenho mesmo em amostras moderadas.

Após a transformação das marginais, o próximo passo é ajustar a cópula diretamente aos pares (μ_1, μ_2) . O parâmetro θ é estimado pelo método da máxima verossimilhança. A estimação fornece tanto o valor pontual de θ quanto a medida de

dependência correspondente via τ , permitindo interpretações mais coerentes sobre a relação entre as variáveis.

7 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados os resultados obtidos a partir da análise da frequência de dias chuvosos na cidade de São Paulo entre 1990 e 2019, bem como da investigação complementar envolvendo a precipitação total mensal, a dependência entre variáveis via cópulas e o papel do fenômeno El Niño–Oscilação Sul (ENSO).

Os resultados são discutidos considerando aspectos climáticos, estatísticos e aplicados, com destaque para a utilidade prática em contextos de planejamento urbano e gestão de riscos.

7.1 Análise descritiva da proporção de dias chuvosos

Inicialmente, foram calculadas medidas descritivas para a proporção mensal de dias com chuva entre 1990 e 2019. Os resultados indicam que a cidade apresenta uma média geral próxima de 0,30, correspondendo a cerca de 9 dias de chuva por mês, com variabilidade significativa entre as estações do ano. O coeficiente de variação indica heterogeneidade moderada, revelando que, apesar de haver sazonalidade, existe ampla variabilidade mês a mês.

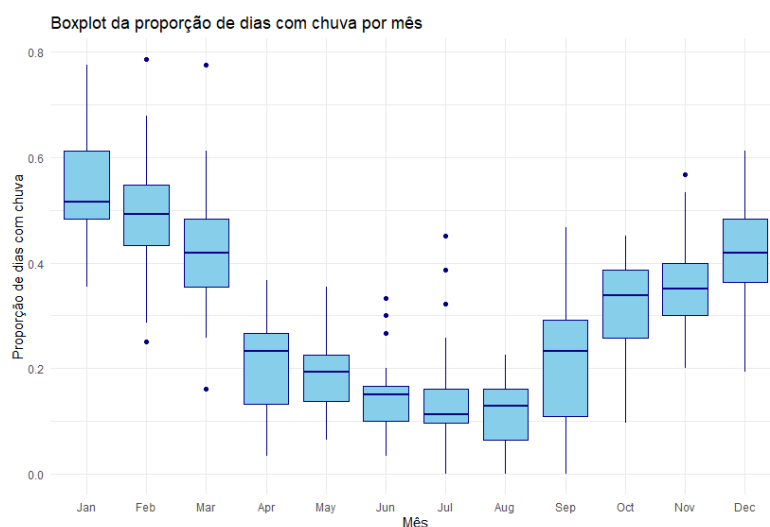
Tabela 1: Estatísticas descritivas da proporção mensal de dias chuvosos.

Mês	Média	Mediana	D.P	Mínimo	Máximo	C.V
Jan	0.537	0.516	0.098	0.355	0.774	0.182
Fev	0.488	0.491	0.124	0.250	0.786	0.254
Mar	0.429	0.419	0.122	0.161	0.774	0.284
Abr	0.212	0.233	0.093	0.033	0.367	0.439
Mai	0.191	0.194	0.069	0.065	0.355	0.362
Jun	0.153	0.150	0.086	0.033	0.333	0.561
Jul	0.141	0.113	0.109	0.000	0.452	0.774
Ago	0.119	0.129	0.071	0.000	0.226	0.594
Set	0.226	0.233	0.121	0.000	0.467	0.537
Out	0.322	0.339	0.088	0.097	0.452	0.274
Nov	0.353	0.350	0.100	0.20	0.567	0.284
Dez	0.430	0.419	0.0996	0.194	0.613	0.232

Fonte: Elaborado pelo Autor.

A distribuição apresenta assimetria moderada e cauda mais acentuada à direita, refletindo meses excepcionalmente chuvosos, especialmente entre dezembro e março. Esses padrões justificam o uso de modelos adequados para proporções com assimetria, como a distribuição Beta.

Uma análise preliminar concentra-se na observação do comportamento da proporção mensal de dias chuvosos. O boxplot apresentado a seguir, construído a partir de todos os meses entre 1990 e 2019, evidencia a forte variabilidade sazonal do regime pluviométrico da cidade.

Figura 4: Boxplot da proporção mensal de dias chuvosos por mês.

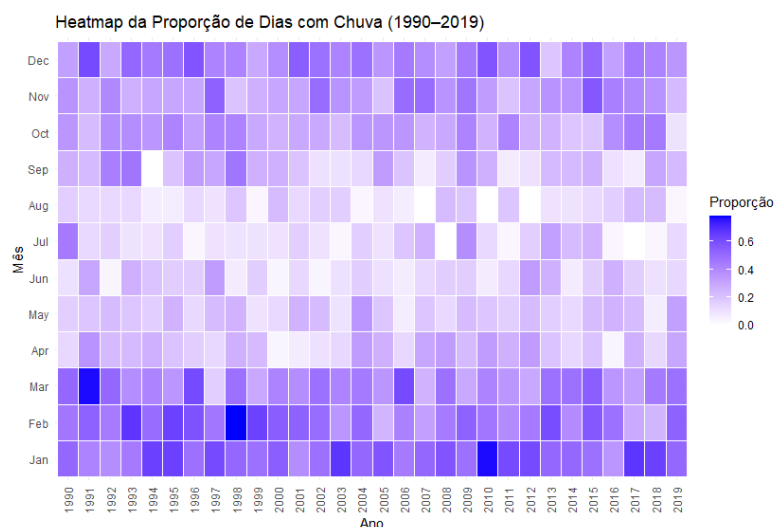
Fonte: Elaborado pelo Autor.

A Figura 4 deixa claro que os meses de verão apresentam proporções mais altas, com mediana elevada e maior dispersão, refletindo a maior instabilidade atmosférica típica dessa estação. Em contraste, os meses de inverno exibem valores sistematicamente baixos, com forte concentração próxima de zero. Essa assimetria entre períodos secos e chuvosos reforça a inadequação de modelos lineares clássicos e justifica o uso da distribuição Beta, cuja flexibilidade permite modelar proporções próximas de 0 e 1.

Além disso, observa-se a presença de valores extremos ao longo de todo o período, indicando meses com seca quase absoluta e meses com chuva praticamente diária. Essa variabilidade reforça a necessidade de técnicas de modelagem capazes de incorporar sazonalidade e eventuais tendências temporais.

Complementando a análise, apresenta-se a Figura 5, sendo um mapa de calor que sintetiza a evolução anual e mensal da proporção de dias chuvosos.

Figura 5: Mapa de calor da proporção mensal de dias com chuva por mês e ano (1990-2019).



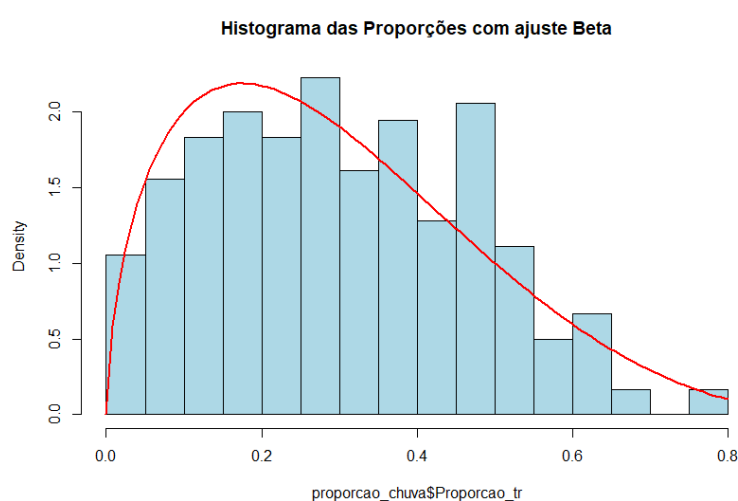
Fonte: Elaborado pelo Autor.

A Figura 5 revela padrões sazonais bem definidos e recorrentes ao longo dos 30 anos analisados: os meses de janeiro, fevereiro e março apresentam as maiores proporções, enquanto julho e agosto concentram os valores mais baixos. Nota-se, também, variação considerável entre diferentes anos, principalmente na estação chuvosa, em que a intensidade da chuva pode mudar substancialmente entre anos consecutivos. Esses padrões justificam a futura inclusão do mês como fator explicativo nos modelos de regressão Beta.

7.2 Ajuste da Distribuição Beta

Para avaliar a adequação da distribuição Beta à variável proporção mensal de dias chuvosos, foi aplicada a transformação proposta por Smithson e Verkuilen (2006), garantindo que todos os valores permanecessem estritamente dentro do intervalo (0,1). A Figura 6 apresenta o histograma da variável acompanhado da densidade estimada.

Figura 6: Histograma da proporção de dias com chuva e ajuste Beta.



Fonte: Elaborado pelo Autor.

Observa-se que a maior parte dos valores se encontra concentrada entre 0,1 e 0,4, indicando que, em grande parte dos meses, menos de 40% dos dias registram precipitação. A curva da distribuição Beta ajustada acompanha adequadamente o formato empírico, apresentando assimetria à direita, condizente com a predominância de meses menos chuvosos ao longo do ano. Essa coerência também pode ser verificada pelas estimativas dos parâmetros, apresentadas na Tabela 2.

Tabela 2: Estimativas dos parâmetros da distribuição Beta ajustada

Parâmetros	Estimativa	Erro Padrão
Alpha (α)	1.5948	0.1094
Beta (β)	3.8048	0.2852

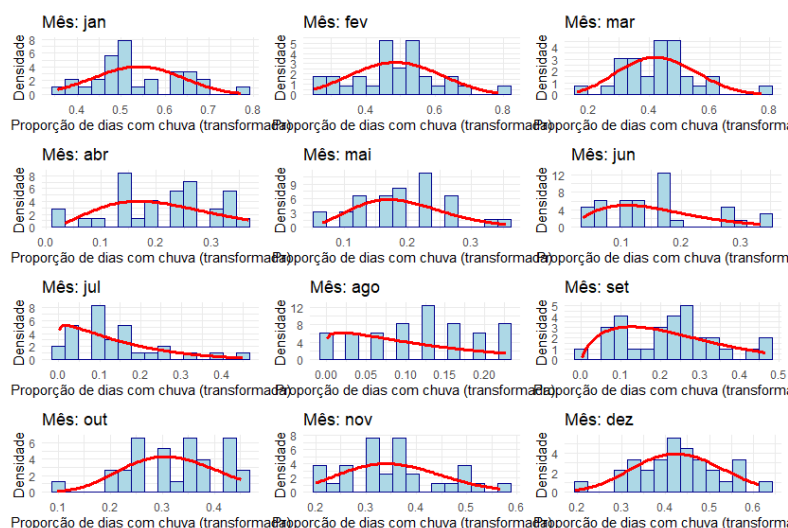
Fonte: Elaborado pelo Autor.

Os valores de α e β obtidos indicam que a distribuição apresenta cauda mais alongada à direita e maior densidade em proporções mais baixas, característica compatível com a série histórica de precipitação da cidade de São Paulo. Portanto, a distribuição Beta mostrou-se apropriada para representar a variabilidade da proporção mensal de dias chuvosos, justificando seu uso nas modelagens subsequentes.

7.3 Ajuste Mensal da Distribuição Beta

Com o objetivo de investigar possíveis diferenças sazonais no comportamento da variável, foram ajustadas distribuições Beta individualmente para cada mês do ano. A Figura 7 apresenta os histogramas mensais acompanhados das respectivas curvas Beta ajustadas.

Figura 7: Histogramas da proporção de dias com chuva com ajustes Beta por mês.



Fonte: Elaborado pelo Autor.

Os resultados evidenciam que a distribuição da proporção de dias chuvosos varia substancialmente ao longo do ano. Nos meses de verão (dezembro a março), observa-se uma concentração de valores mais altos, com distribuições próximas à simetria ou levemente assimétricas, refletindo maior frequência de dias com chuva. Já nos meses de inverno (junho a setembro), as distribuições tornam-se nitidamente concentradas próximas de zero, apresentando forte assimetria e, por vezes, elevada curtose, o que indica predominância de meses com poucos ou nenhum dia chuvoso.

Nos meses de transição (abril, maio e outubro) observa-se comportamento intermediário, com formatos mais heterogêneos entre os anos. Embora alguns desses meses apresentem maior variabilidade estrutural, a distribuição Beta foi capaz de representar satisfatoriamente a tendência central e o padrão geral dos dados.

As estimativas dos parâmetros α e β para cada mês encontram-se na Tabela 3.

Tabela 3: Estimativa dos parâmetros da distribuição Beta por mês.

Mês	Alpha(α)	E.P (α)	Beta (β)	E.P (β)
Jan	13.972	3.579	12.049	3.077
Fev	7.691	1.954	8.080	2.056
Mar	6.770	1.713	8.970	2.291
Abr	3.217	0.792	12.015	3.134
Mai	5.791	1.455	24.383	6.333
Jun	2.620	0.638	14.386	3.794
Jul	1.096	0.251	6.733	1.859
Ago	1.142	0.262	8.620	2.390
Set	1.728	0.411	6.086	1.606
Out	8.252	2.092	17.450	4.496
Nov	8.265	2.096	15.078	3.877
Dez	10.513	2.679	13.941	3.574

Fonte: Elaborado pelo Autor.

A análise desses parâmetros reforça a forte sazonalidade: valores mais altos de α e β nos meses de verão refletem distribuições mais concentradas em níveis moderados a elevados de proporção, enquanto valores mais baixos especialmente α nos meses secos explicam a concentração em proporções próximas de zero.

7.4 Regressão Beta

Com a finalidade de investigar possível tendência temporal na proporção de dias com chuva, foi realizada uma análise de regressão beta utilizando como variável explicativa o ano.

Considerando os valores obtidos na Tabela 4, os coeficientes indicam que não há evidências estatísticas significativas de tendência temporal na proporção mensal de dias com chuva ao longo do período estudado.

Tabela 4: Coeficientes estimados da regressão beta (Proporção de dias com chuva ~ Ano).

Parâmetro	Estimativa	E.P	p-valor
Intercepto (β_0)	12.749	10.059	0.205
Ano (β_1)	-0.007	0.005	0.176

Fonte: Elaborado pelo Autor.

Em seguida, foi ajustado um modelo mais completo, incluindo o fator mês tendo a seguinte forma:

$$\text{ProporçãoBeta} \sim \text{Ano} + \text{fator}(\text{Mês}).$$

A Tabela 5 mostra os coeficientes estimados desse novo modelo.

Tabela 5: Coeficientes estimados do modelo de regressão Beta com efeito sazonal.

Coeficiente	Estimativa	E.P	p-valor
Intercepto	13.232	6.819	0.052
Ano	-0.007	0.003	0.055
Fev	-0.194	0.128	0.131
Mar	-0.425	0.129	0.001
Abr	-1.464	0.141	< 0.001
Mai	-1.522	0.142	< 0.001
Jun	-1.833	0.149	< 0.001
Jul	-2.133	0.156	< 0.001
Ago	-2.289	0.161	< 0.001
Set	-1.506	0.142	< 0.001
Out	-0.870	0.132	< 0.001
Nov	-0.733	0.131	< 0.001
Dez	-0.419	0.129	0.001

Fonte: Elaborado pelo Autor.

Em seguida, a tabela 6 mostra que o modelo sazonal apresentou melhor ajuste, visto que o AIC é menor o pseudo- R^2 é maior. É possível observar que os meses de inverno apresentaram coeficientes menores, confirmando menor frequência de dias chuvosos, enquanto os meses de verão apresentaram valores maiores, reforçando o padrão sazonal.

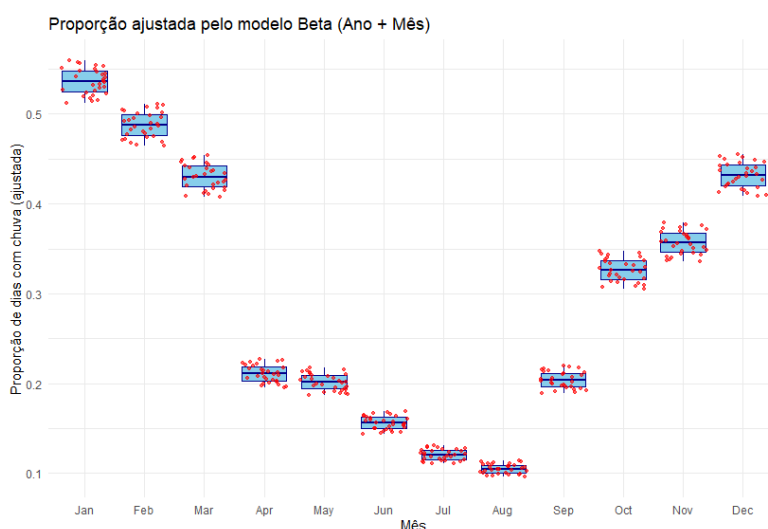
Tabela 6: Comparação entre modelos de regressão Beta: simples (Ano) e completo (Ano + Mês).

Modelo	AIC	Pseudo - R^2
Simples (Ano)	-273.370	0.004
Completo (Ano + Mês)	-612.026	0.5125

Fonte: Elaborado pelo Autor.

A seguir, a Figura 8 apresenta os valores ajustados.

Figura 8: Proporções observadas e ajustadas pelo modelo Beta com efeito sazonal.



Fonte: Elaborado pelo Autor.

Como uma análise complementar, foi avaliado a multicolineariedade através do VIF, com os resultados obtidos sendo inferiores a 1,9, indicando ausência de colinearidade problemática.

7.5 Regressão Beta com interação Ano x Mês

Com a finalidade de encontrar possíveis modificações ao longo do tempo de forma diferente entre os meses do ano referente a proporção de dias com chuva mensais, foi realizada um modelo de regressão Beta envolvendo interação entre ano e mês. Os coeficientes significativos são mostrados na Tabela 7.

Tabela 7: Coeficientes de interação significativos ou próximo do nível de significância (Ano x Mês).

Mês	Coef. Interação	E.P	z-valor	p-valor	Tendência
Fev	-0.0327	0.0146	-2.232	0.0256	Diminuição significativa
Mar	-0.0195	0.0147	-1.327	0.1844	Sem evidência
Abr	-0.0097	0.0160	-0.604	0.5459	Sem evidência
Mai	-0.0100	0.0161	-0.618	0.5363	Sem evidência
Jun	-0.0045	0.0169	-0.267	0.7893	Sem evidência
Jul	-0.0393	0.0178	-2.205	0.0275	Diminuição significativa
Ago	-0.0321	0.0183	-1.756	0.0791	Indício de diminuição
Set	-0.0169	0.0161	-1.047	0.2952	Sem evidência
Out	-0.0217	0.0150	-1.445	0.1486	Sem evidência
Nov	-0.00039	0.0149	-0.026	0.9792	Sem evidência
Dez	-0.0159	0.0146	-1.088	0.2764	Sem evidência

Fonte: Elaborado pelo Autor.

Com base nos resultados obtidos, o nível de significância dos coeficientes dos meses de fevereiro e julho revelaram uma tendência significativa de diminuição da proporção de dias com chuva durante esses anos analisados, enquanto agosto mostrou indício de diminuição. O pseudo- R^2 foi 0,5236, indicando boa capacidade explicativa. O AIC foi próximo ao do modelo sazonal sem interação. Esses resultados consolidam a ideia de mudanças nos regimes anuais de chuva, mostrando que os impactos das alterações climáticas podem se apresentar de forma sazonalmente diferente.

7.6 Influência do ENSO na Proporção de Dias Chuvosos

A fim de investigar uma possível relação entre a variabilidade pluviométrica mensal e fenômenos climáticos de grande escala, incorporou-se ao estudo o índice ONI, que caracteriza as fases do El Niño–Oscilação Sul (ENSO).

Inicialmente, foi realizada uma análise descritiva da proporção de dias chuvosos de acordo com cada fase do ENSO. A Tabela 8 apresenta medidas resumo da série estratificada em três categorias: El Niño, La Niña e fase neutra.

Tabela 8: Estatísticas descritivas da proporção de dias chuvosos por fase do ENSO.

ENSO	<i>n</i>	Média	Mediana	DP
La Niña	100	0.322	0.323	0.174
Neutro	160	0.260	0.226	0.161
El Niño	100	0.342	0.355	0.170

Fonte: Elaborado pelo Autor.

Observa-se que tanto El Niño quanto La Niña apresentam proporções médias mais elevadas que os meses classificados como fase neutra. A fase El Niño registra o maior valor médio, indicando que condições de aquecimento no Pacífico tendem a intensificar a ocorrência de dias chuvosos em São Paulo. Por outro lado, a fase neutra apresenta a menor média e menor mediana, refletindo uma distribuição mais concentrada em meses relativamente secos.

Para verificar se essas diferenças são estatisticamente significativas, foi realizado o teste não paramétrico de Kruskal–Wallis, adequado devido à assimetria evidente da distribuição da variável.

O resultado indicou diferenças significativas entre as três fases ($X^2 = 17,932; p < 0,001$), sugerindo que o comportamento da frequência de dias chuvosos é influenciado pelo ENSO.

7.7 Modelos de Regressão Beta com ENSO

A fim de avaliar o papel do ENSO na dinâmica da proporção de dias chuvosos, foram ajustados três modelos de regressão Beta:

- modelo simples com ENSO como variável contínua;
- modelo completo incluindo efeitos sazonais e temporais;
- modelo com interação entre ENSO e mês,

Com a finalidade de verificar se o efeito do ENSO difere ao longo do ano.

7.7.1 Modelo 1: ENSO contínuo

O primeiro modelo considerou apenas o índice ENSO contínuo como variável explicativa:

$$Proporção_{tr} \sim ENSO_{index}$$

Os resultados indicaram coeficiente não significativo ($p \approx 0,102$) e pseudo- R^2 muito baixo (0,0049). Isso mostra que, isoladamente, o ENSO não explica adequadamente a variabilidade mensal da proporção de dias chuvosos, reforçando a forte dominância do padrão sazonal observado.

7.7.2 Modelo 2: ENSO + Mês + Ano

O segundo modelo incorporou os efeitos de mês (como fator) e ano, além do ENSO:

$$Proporção_{tr} \sim ENSO_{index} + Mês + Ano.$$

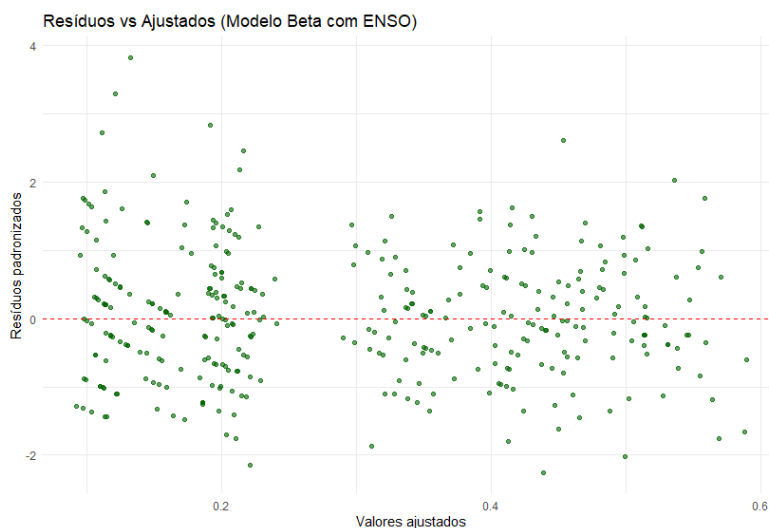
Os resultados mostraram:

- ENSO_index significativo ($p \approx 0,015$)
- pseudo- $R^2 = 0,5186$
- AIC muito inferior ao modelo simples.

Esse resultado evidencia que, ao controlar a sazonalidade, o ENSO passa a apresentar influência significativa. O coeficiente positivo sugere que aumento no índice ENSO (tendência a El Niño) está associado a maior proporção de dias chuvosos, reforçando os padrões encontrados na análise descritiva.

Abaixo, a Figura 9 apresenta os resíduos padronizados em função dos valores ajustados para o modelo Beta completo.

Figura 9: Resíduos padronizados versus valores ajustados do modelo Beta completo ($Proporção_{tr} \sim ENSO_{index} + Mês + Ano$).



Fonte: Elaborado pelo Autor.

A distribuição dos resíduos não revela padrão sistemático em função dos ajustados: a nuvem de pontos encontra-se, de forma geral, centrada em torno de zero e sem tendência marcada, o que sugere adequação do modelo à estrutura média dos dados. Observa-se, contudo, a presença de alguns resíduos extremos (valores padronizados acima de 2 ou abaixo de -2), que merecem investigação pontual (meses/anos com comportamento atípico). De modo geral, o gráfico indica que a suposição de relação média modelada pelo logit é razoável e que não existem evidências claras de heterocedasticidade severa não capturada pelo modelo.

7.7.3 Modelo 3: Interação ENSO × Mês

O terceiro modelo introduziu termos de interação:

$$Proporção_{tr} \sim ENSO_{index} \times Mês + Ano.$$

Esse modelo busca identificar se a influência do ENSO é diferente ao longo do ano. Os resultados indicaram que:

- algumas interações foram significativas, como ENSO × Novembro ($p \approx 0,044$);
- outras apresentaram indícios relevantes, como ENSO × Maio e ENSO × Julho ($p \approx 0,06$).

A interpretação sugere que o impacto do ENSO varia entre os meses, sendo mais relevante em períodos de transição ou início da estação chuvosa.

7.7.4 Comparação entre modelos

A Tabela 9 sintetiza os resultados de ajuste:

Tabela 9: Comparação entre modelos Beta com ENSO.

Modelo	df	AIC	pseudo-R ²
Modelo simples (ENSO)	3	-273.94	0.0049
Modelo completo (ENSO + Mês + Ano)	15	-615.77	0.5186
Modelo com interação ENSO×Mês	26	-606.67	0.5321

Fonte: Elaborado pelo Autor.

O modelo completo apresentou o melhor AIC, enquanto o modelo com interação obteve o maior pseudo-R². Ambos mostraram desempenho muito superior ao modelo simples. Do ponto de vista interpretativo:

- o ENSO realmente modula a proporção de dias chuvosos;
- esse efeito é mais bem observado quando a sazonalidade é controlada;
- há meses específicos em que o ENSO exerce influência mais forte.

7.8 Dependência entre Proporção e Precipitação via Cópulas Estratificadas por ENSO

Com o objetivo de avaliar se a relação entre proporção de dias chuvosos e precipitação total varia conforme a fase do ENSO, ajustou-se uma cópula de Clayton separadamente para cada categoria.

A Tabela 10 apresenta os parâmetros θ e a medida de dependência τ de Kendall.

Tabela 10: Parâmetros da cópula de Clayton estratificados por ENSO.

Fase	θ	τ
La Niña	5.984	0.75
Neutro	5.859	0.746
El Niño	5.031	0.716

Fonte: Elaborado pelo Autor.

Os resultados indicam dependência forte em todas as fases. Isso significa que meses com alta proporção de dias chuvosos também tendem a apresentar alta precipitação total relação esperada, mas agora estatisticamente quantificada.

Observa-se dependência ligeiramente maior em anos de La Niña. Já anos de El Niño apresentam dependência um pouco menor, sugerindo maior ocorrência de eventos volumosos em poucos dias, que aumentam a precipitação total sem elevar proporcionalmente o número de dias chuvosos.

Esse resultado é altamente relevante, pois evidencia que a estrutura conjunta das variáveis também responde às fases do ENSO.

7.9 Síntese Integrada dos Resultados

A análise conduzida ao longo deste capítulo permitiu compreender de forma abrangente os padrões de ocorrência de dias chuvosos em São Paulo entre 1990 e 2019, bem como sua relação com a precipitação total, a sazonalidade e a variabilidade climática associada ao ENSO. Os resultados revelaram, de maneira consistente, que o regime pluviométrico da cidade é fortemente marcado pela sazonalidade, com meses de verão exibindo proporções mais elevadas de dias chuvosos e maior variabilidade, enquanto os meses de inverno apresentam valores sistematicamente baixos.

A modelagem Beta aplicada tanto de forma global quanto mensal mostrou-se adequada para capturar a assimetria característica da distribuição da variável de interesse. O ajuste por mês evidenciou, de forma clara, as diferenças estruturais no comportamento da proporção de dias com chuva ao longo do ano, reforçando a necessidade do tratamento sazonal nas modelagens subsequentes.

A análise de regressão sem o ENSO demonstrou ausência de tendência temporal global, embora tenha revelado, na modelagem com interação Ano $\times$ Mês,

sinais de mudança em meses específicos (principalmente fevereiro, julho e, com menor intensidade, agosto). Esse resultado sugere que possíveis alterações climáticas que afetam a variabilidade intra-anual podem não se manifestar de forma homogênea ao longo do ano.

Ao incorporar o ENSO, observou-se que as fases do fenômeno exercem influência mensurável sobre a frequência de dias chuvosos. A fase El Niño associou-se às maiores médias, enquanto a fase neutra apresentou os valores mais baixos. O efeito contínuo do índice ONI torna-se significativo somente quando a sazonalidade é controlada, reforçando que a variabilidade interanual do Pacífico modula o regime de chuvas, mas não substitui o papel dominante da sazonalidade local.

Os modelos de regressão Beta com ENSO revelaram desempenho robusto, mostrando que o fenômeno atua como modulador adicional no comportamento pluviométrico da cidade. Ademais, a interação ENSO $\times$ Mês evidenciou que a influência do ENSO não é uniforme ao longo do ano, indicando sensibilidade diferenciada entre meses — achado relevante para previsões sazonais.

Por fim, a análise com cópulas mostrou dependência muito forte entre a proporção de dias chuvosos e a precipitação total mensal, mas também revelou distinções sutis entre fases do ENSO, particularmente maior dependência em anos de La Niña e menor em anos de El Niño. Esse resultado é importante, pois sugere que eventos extremos de precipitação podem ocorrer de forma diferenciada entre as fases, reforçando a utilidade das cópulas como ferramenta para compreender dinâmicas conjuntas.

Em síntese, o conjunto de resultados mostra que a variabilidade mensal de dias chuvosos em São Paulo é governada por três eixos principais: sazonalidade marcante, modificações em meses específicos ao longo do período analisado e influência adicional do ENSO, que afeta tanto a proporção quanto a estrutura conjunta de precipitação e frequência de dias chuvosos. Tais achados fortalecem a compreensão climatológica da cidade e servem de subsídio para estratégias de monitoramento e planejamento hídrico.

8 CONCLUSÃO

O presente trabalho analisou de maneira detalhada a proporção mensal de dias com chuva na cidade de São Paulo entre 1990 e 2019, empregando métodos estatísticos avançados, incluindo modelagem Beta, análise de tendência temporal, investigação de efeitos sazonais, incorporação do fenômeno ENSO como covariável climática e avaliação de dependência bivariada via cópulas. Os resultados obtidos permitem compreender com maior profundidade tanto a dinâmica intrínseca do regime pluviométrico da cidade quanto os fatores externos que modulam essa variabilidade.

A investigação revelou inicialmente que a proporção mensal de dias chuvosos apresenta forte padrão sazonal. Os meses de verão concentram as proporções mais elevadas, enquanto o inverno é marcado por valores próximos de zero em diversos anos. Essa sazonalidade, evidente nas análises descritivas e confirmada nos ajustes mensais da distribuição Beta, representa o componente dominante da variabilidade pluviométrica paulistana.

O estudo da tendência temporal indicou ausência de mudanças significativas no comportamento médio da proporção de dias chuvosos ao longo das três décadas analisadas. Entretanto, a inclusão da interação entre ano e mês revelou que possíveis alterações climáticas podem afetar de maneira diferenciada meses específicos, com destaque para fevereiro e julho, sugerindo que mudanças intra-anuais merecem atenção em estudos de monitoramento de longo prazo.

A análise da influência do ENSO mostrou que o fenômeno exerce papel modulador importante na frequência de dias chuvosos em São Paulo. As fases El Niño e La Niña apresentaram maiores proporções médias de dias chuvosos do que a fase neutra, confirmando a relevância climática dessa oscilação oceânica-atmosférica. Os modelos de regressão Beta indicaram que o ENSO, embora não explique a variabilidade sozinho, torna-se significativo quando ajustado em conjunto com os efeitos sazonais e temporais, evidenciando que sua influência é complementar e não substitutiva à sazonalidade local. A interação ENSO $\times$ Mês reforçou que o impacto do ENSO varia ao longo do ano, afetando alguns períodos de forma mais intensa.

Por sua vez, a análise com cópulas trouxe uma contribuição adicional ao revelar que a proporção de dias chuvosos e a precipitação total mensal apresentam dependência muito forte, porém não idêntica entre as fases do ENSO. Em anos de La Niña, a dependência é ligeiramente maior, enquanto em El Niño é um pouco reduzida,

sugerindo padrões de chuva diferenciados entre as fases. Esses resultados podem ter implicações práticas para modelagem de extremos e previsão sazonal.

Do ponto de vista aplicado, os achados deste estudo oferecem subsídios importantes para órgãos de planejamento urbano, defesa civil e entidades responsáveis pela gestão de recursos hídricos. A compreensão detalhada da sazonalidade e da influência do ENSO pode auxiliar na elaboração de estratégias voltadas à prevenção de alagamentos, otimização de reservatórios, planejamento de infraestrutura e manejo de riscos hidro meteorológicos. Além disso, o uso de cópulas permite caracterizar de forma mais realista a relação entre frequência de dias chuvosos e volume total precipitado, o que é essencial para estudos de impacto, modelagem de cheias e avaliação de vulnerabilidades.

Por fim, destaca-se que a abordagem metodológica empregada combinando distribuição Beta, regressão com efeitos sazonais, interação temporal, ENSO e cópulas fornece uma estrutura analítica robusta e replicável para outros municípios ou regiões brasileiras. Trabalhos futuros podem incorporar variáveis adicionais (como temperatura, pressão atmosférica e umidade), explorar diferentes famílias de cópulas, utilizar modelos hierárquicos bayesianos ou integrar previsões climáticas sazonais operacionais, ampliando o alcance e a utilidade prática das análises aqui desenvolvidas.

Assim, conclui-se que a variabilidade pluviométrica de São Paulo resulta da combinação entre sazonalidade forte, modulações interanuais associadas ao ENSO e dependência complexa entre frequência e volume de chuva. O entendimento desses elementos contribui para aprimorar o monitoramento climático e fundamentar ações preventivas, consolidando a importância de análises estatísticas avançadas para a gestão urbana e ambiental.

9 REFERÊNCIAS

HAIDER, S. F.; HUSSAIN, H. Statistical distribution of rainfall data and probabilities of the number of rainy days. *Journal of Islamic Academy of Sciences*, v. 2, n. 4, p. 299-302, 1989.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Summary for Policymakers*. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva: IPCC, 2014. Disponível em: <https://www.ipcc.ch/report/ar5/wg2/summary-for-policymaker> . Acesso em: 17 maio 2025.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, 2022. Disponível em: <https://www.ipcc.ch/report/ar6/wg2/>. Acesso em: 25 out. 2024.

INSTITUTO NACIONAL DE METEOROLOGIA – INMET. *Normais climatológicas do Brasil 1961-1990*. Brasília, DF: INMET, 1996 (arquivo atualizado em 2022). Disponível em:

Acesso em: 29 maio 2025.

KUMAR, Neeraj; MOTE, B. M.; SRIVASTAVA, Alok; PATEL, D. D.; CHALODIA, A. L. Rainy days analysis by using normal, binomial distribution and discrete probability for Navsari district of south Gujarat. *International Journal of Chemical Studies*, v. 9, n. 1, p. 11712, 2021. Disponível em: <https://doi.org/10.22271/chemi.2021.v9.i1ar.11712>.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to Linear Regression Analysis*. 5. ed. Hoboken: John Wiley & Sons, 2012.

NAÇÕES UNIDAS BRASIL: Objetivos de Desenvolvimento Sustentável. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 18 set. 2024.

Pielke, R.A. & Niyogi, D. (2010). The Role of Landscape Processes within the Climate System. *Nature*, 465: 1089–1094.

Reboita, M.S.; Gan, M.A.; Rocha, R.P.; Ambrizzi, T. (2010). Regimes de Precipitação no Brasil. *Revista Brasileira de Meteorologia*, 25(2): 185–204.

REBOITA, M. S. et al. Regimes de precipitação na América do Sul: uma revisão bibliográfica. *Revista Brasileira de Climatologia*, v. 6, p. 1-17, 2010.

Rosenzweig, C.; Iglesias, A.; Yang, X.B.; Epstein, P.R.; Chivian, E. (2004). Climate Change and Extreme Weather Events: Implications for Food Production, Plant Diseases, and Pests. *Global Change & Human Health*, 2: 90–104.

SAAD, João Carlos Cury. *Estudo das distribuições de frequências da evapotranspiração de referência e da precipitação pluvial para fins de dimensionamento de sistemas de irrigação*. Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, Brasil, 1990.

SILVA, V. P. R. da; PEREIRA, E. R. R.; ALMEIDA, R. S. R. Estudo da variabilidade anual e intra-anual da precipitação na região Nordeste do Brasil. *Revista Brasileira de Meteorologia*, v. 27, n. 2, p. 163-172, 2012.

SMITHSON, M.; VERKUILEN, J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, v. 11, n. 1, p. 54-71, 2006.

SOUZA, C. L. O. de; NOGUEIRA, V. de F. B.; NOGUEIRA, V. da S. Variabilidade interanual da precipitação em cidades do semiárido brasileiro entre os anos de 1984 e 2015. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*, v. 12, n. 4, p. 740-747, 2017.

Tarifa, J.R.; Armani, G. (2001). *O Clima da Cidade de São Paulo*. São Paulo: USP.

UOL. Virada de tempo em SP: capital deve ter chuva prevista para um mês em 48h. São Paulo, 3 abr. 2025. Disponível em: <https://noticias.uol.com.br/cotidiano/ultimas-noticias/2025/04/03/tempo-frente-fria-sao-paulo.htm>. Acesso em: 10 abr. 2025.