

Universidade Tecnológica Federal do Paraná – UTFPR
Departamento Acadêmico de Eletrônica – DAELN
Bacharelado em Engenharia Eletrônica
Disciplina: ELTD4 – Reconhecimento de Padrões e Aprendizado de Máquina
Semestre: 2025.1
Prof: Gustavo Benvenutti Borba

RELATÓRIO

Análise Espaço-Temporal da Dinâmica da Cobertura Vegetal no Brasil (2019-2024) Utilizando Dados de Sensoriamento Remoto

Aluno:
Pedro Schuves Marodin / a2358522

06.2025

1 Introdução

O Brasil, uma nação megadiversa que abriga biomas de importância crítica para a estabilidade climática global — como a Floresta Amazônica, o Cerrado e a Mata Atlântica —, enfrenta pressões crescentes oriundas da expansão da fronteira agropecuária, da urbanização e da intensificação de eventos climáticos extremos. O monitoramento contínuo e em larga escala dessas transformações é logisticamente inviável por métodos tradicionais, tornando o sensoriamento remoto a ferramenta primária e indispensável para a análise ambiental em escala continental. Sua capacidade de aquisição de dados de forma sistemática e objetiva permite a detecção de padrões e tendências que seriam impossíveis de observar em campo.

Para traduzir os dados espectrais brutos em informações biofísicas úteis, a comunidade científica desenvolveu uma gama de índices e produtos. Este trabalho foca em três variáveis-chave. A primeira é o Índice de Vegetação por Diferença Normalizada (NDVI) [20], um indicador consolidado do vigor e saúde da vegetação. A ele, somam-se a Temperatura da Superfície Terrestre (LST), que mede a energia térmica irradiada pela superfície e atua como um proxy para o estresse hídrico e energético [28], e a Precipitação Pluviométrica, estimada por missões como a GPM para quantificar a disponibilidade de água, o principal motor dos ecossistemas terrestres [10].

A análise conjunta desta série temporal de dados, derivados dos sensores MODIS e GPM (*Global Precipitation Measurement*), abrange todo o território brasileiro no período de 2019 a 2024. A escolha do sensor MODIS para NDVI e LST justifica-se por sua alta frequência de revisita¹, essencial para mitigar a contaminação por nuvens em regiões tropicais. O escopo da análise, no entanto, avança para além da observação de tendências, incorporando técnicas de aprendizado de máquina e redes neurais para uma detecção mais sofisticada de padrões e anomalias, visando uma compreensão mais profunda dos vetores de mudança na paisagem brasileira.

¹Os satélites Terra e Aqua, que carregam o sensor MODIS, possuem uma órbita polar que lhes permite imagear a superfície terrestre inteira a cada 1 a 2 dias.

2 Objetivos

2.1 Objetivo Geral

O objetivo central deste trabalho é caracterizar, quantificar e interpretar a dinâmica espaço-temporal da cobertura vegetal em todo o território brasileiro no período de 2019 a 2024. Para tal, será utilizado um framework multi-método que integra dados de sensoriamento remoto (NDVI, LST e Precipitação) com análises estatísticas, de aprendizado de máquina e de redes neurais, a fim de identificar tendências, anomalias e padrões de mudança, e inferir seus possíveis vetores climáticos e antrópicos (tudo que é originado, causado ou modificado pela ação do ser humano no meio ambiente, como o desmatamento, a poluição e a construção de cidades).

2.2 Objetivos Específicos

Para alcançar o objetivo geral, os seguintes passos específicos foram definidos e executados:

1. **Aquisição e Pré-processamento de Dados:** Automatizar o processo de busca, download e preparação de dados de múltiplos sensores (MODIS e GPM) para os anos de interesse, gerando mosaicos anuais para as variáveis de NDVI, Temperatura da Superfície Terrestre (LST) e Precipitação Pluviométrica para todo o Brasil.
2. **Análise de Tendência Temporal em Escala Nacional:** Calcular a tendência geral da vegetação no país, aplicando um modelo de regressão linear sobre a série temporal do NDVI médio anual para determinar a magnitude e a significância estatística de processos de “esverdeamento” (*greening*) ou “escurecimento” (*browning*).
3. **Detecção de Anomalias Anuais:** Identificar anos com comportamento de vegetação atípico, calculando o desvio do NDVI médio de cada ano em relação à média de todo o período, destacando anos com potenciais anomalias climáticas.
4. **Análise de Mudança Espacial (Hotspots):** Gerar um mapa de diferença entre o início e o fim do período (2024 vs. 2019) para localizar espacialmente as áreas que sofreram as mais significativas perdas ou ganhos de vegetação.
5. **Análise Regional Estratificada:** Desagregar a análise em nível de unidade federativa e de grandes regiões, calculando as tendências temporais de NDVI para cada uma e comparando suas trajetórias para identificar padrões e dinâmicas distintas.
6. **Segmentação de Padrões por Clusterização:** Aplicar técnicas de aprendizado de máquina não supervisionado (K-Means) para segmentar o território nacional em zonas de comportamento de NDVI homogêneo (clusters), visando identificar e mapear macrozonas ecológicas de forma automatizada a partir dos dados.
7. **Detecção de Anomalias com Redes Neurais:** Implementar uma arquitetura de rede neural recorrente convolucional (ConvLSTM Autoencoder) para modelar a dinâmica espaço-temporal da vegetação e detectar anomalias, identificando focos de mudança abrupta que não são capturados por tendências lineares simples.
8. **Síntese e Visualização dos Resultados:** Consolidar todas as observações em gráficos e mapas, para explicitar os padrões e processos observados.

3 Fundamentação Teórica

Esta seção detalha os fundamentos do Índice de Vegetação por Diferença Normalizada (NDVI) e enquadra as análises subsequentes nos domínios do Processamento de Sinais e do Reconhecimento de Padrões.

3.1 Índice de Vegetação por Diferença Normalizada (NDVI)

3.1.1 A Assinatura Espectral da Vegetação

A capacidade de monitorar a vegetação a partir do espaço deriva de sua “assinatura espectral” característica, ou seja, a maneira como ela interage com a radiação eletromagnética em diferentes comprimentos de onda [6]. A chave para essa assinatura reside na fisiologia da planta:

- **Absorção na Faixa do Vermelho (Red):** Os pigmentos fotossintéticos, principalmente as clorofitas A e B, são altamente eficientes na absorção de energia na faixa do espectro visível, especialmente na banda do vermelho (aproximadamente 620-750 nm). Essa energia é utilizada no processo de fotossíntese. Uma planta saudável e com alta concentração de clorofila age como um forte absorvedor nesta faixa, resultando em baixa reflectância [9].
- **Reflectância na Faixa do Infravermelho Próximo (NIR):** A estrutura interna da folha, composta pelo parênquima lacunoso (mesófilo), não utiliza a energia do infravermelho próximo (aproximadamente 750-1300 nm). Essa radiação penetra na folha e é refletida e espalhada múltiplas vezes pelas interfaces entre as células e os espaços de ar, resultando em uma reflectância muito alta. Uma vegetação vigorosa, com estrutura celular intacta, é um forte refletor de NIR[9].

A baixa reflectância no vermelho e alta reflectância no NIR é a base física para a detecção da vegetação. Solos expostos, rochas e corpos d’água não exibem esse contraste acentuado [20].

3.1.2 Equação Matemática de NDVI

O NDVI não é uma medição direta de biomassa, mas sim um *proxy*, um índice numérico abstrato projetado para maximizar a sensibilidade a essa assinatura espectral. Ele é calculado pela equação (1):

$$\text{NDVI} = \frac{(\rho_{\text{NIR}} - \rho_{\text{Red}})}{(\rho_{\text{NIR}} + \rho_{\text{Red}})} \quad (1)$$

Onde ρ_{NIR} e ρ_{Red} representam os valores de reflectância de superfície nas respectivas bandas. A **normalização** exerce um papel fundamental, pois:

- **Minimização de Efeitos Externos:** Fatores como variações na iluminação solar (devido ao ângulo do sol ou à presença de nuvens finas) e efeitos topográficos (sombreamento em encostas) afetam a intensidade da radiação tanto no NIR quanto no vermelho. Ao usar uma razão, o NDVI minimiza esses efeitos, pois eles tendem a ser cancelados no quociente, tornando o índice mais robusto e comparável entre diferentes imagens e locais [17].

- **Escala Confinada:** A normalização confina os valores do índice a uma escala fixa de -1 a +1, facilitando a interpretação e a aplicação de limiares. Valores próximos a +1 indicam alta atividade fotossintética (vegetação densa); valores próximos a 0 indicam ausência de vegetação (solos expostos); e valores negativos tipicamente representam água ou neve [20].

3.1.3 Limitações e Considerações sobre NDVI

- **Saturação do Sinal:** Em ecossistemas de alta biomassa, como florestas tropicais densas, o NDVI pode atingir um ponto de saturação. A partir de um certo nível de densidade foliar, a absorção de vermelho se torna quase total e a reflectância de NIR se estabiliza. Consequentemente, o valor do NDVI para de aumentar, mesmo que a biomassa continue a crescer, tornando o índice menos sensível a variações em florestas muito densas [25].
- **Influência do Solo:** Em áreas de vegetação esparsa (e.g., Cerrado, Caatinga), a reflectância do solo subjacente pode influenciar significativamente o valor do pixel, mascarando a real contribuição da vegetação [8].
- **Efeitos Atmosféricos:** Embora a normalização reduza ruídos, aerossóis e vapor d'água na atmosfera ainda podem afetar as medições e requerer correções atmosféricas para análises de alta precisão [9].

3.2 Pluviosidade e Temperatura

3.2.1 Temperatura da Superfície Terrestre (LST)

A Temperatura da Superfície Terrestre, ou LST (do inglês, *Land Surface Temperature*), é definida como a temperatura radiativa da “pele” da superfície terrestre. É crucial distingui-la da temperatura do ar, que é tipicamente medida por estações meteorológicas a uma altura de 2 metros. A LST é uma variável-chave no balanço de energia da superfície, pois governa os fluxos de calor sensível e latente para a atmosfera e é diretamente influenciada pela radiação solar incidente, pelo albedo da superfície, pela cobertura vegetal e pela umidade do solo [29].

No contexto do sensoriamento remoto, a LST é estimada a partir de medições da radiação térmica emitida pela superfície terrestre, captada por sensores a bordo de satélites, como o MODIS (*Moderate Resolution Imaging Spectroradiometer*). Produtos como o MOD11A2 [28] fornecem estimativas globais de LST com alta frequência temporal (e.g., 8 dias), permitindo o monitoramento contínuo das condições térmicas. Anomalias na LST são indicadores de estresse ambiental. Um aumento anômalo da LST está frequentemente associado a condições de seca, onde a redução da umidade do solo e da transpiração da vegetação diminui o resfriamento evaporativo da superfície. Além disso, mudanças no uso e cobertura da terra, como o desmatamento, levam a um aumento significativo da LST, pois a superfície exposta do solo absorve e retém mais calor do que a cobertura florestal [23].

3.2.2 Precipitação Pluviométrica

A precipitação é a principal fonte de água para os ecossistemas terrestres e o principal motor do ciclo hidrológico. Sua alta variabilidade espacial e temporal torna o monitoramento por redes de pluviômetros em solo um desafio, especialmente em regiões

vastas e remotas como a Amazônia. O sensoriamento remoto orbital oferece uma solução para este problema, fornecendo estimativas de precipitação quase globais e contínuas.

Missões como a GPM (*Global Precipitation Measurement*) utilizam uma constelação de satélites equipados com radares de precipitação e radiômetros de micro-ondas passivos para detectar e quantificar partículas de chuva e gelo na atmosfera. Produtos como o IMERG (*Integrated Multi-satellite Retrievals for GPM*) integram essas múltiplas fontes de dados para gerar um produto de precipitação unificado e de alta qualidade [10].

A análise da precipitação e de suas anomalias é fundamental para a compreensão da dinâmica da vegetação. Déficits de chuva (anomalias negativas) prolongados caracterizam eventos de seca meteorológica, que levam ao estresse hídrico da vegetação e, consequentemente, à redução do NDVI. Por outro lado, excedentes de chuva (anomalias positivas) podem tanto favorecer “esverdeamento” intenso como, em casos extremos, causar inundações que também impactam negativamente certos tipos de ecossistemas.

3.3 Análise Integrada e Desacoplamento de Sinais Climáticos e Antrópicos

A análise isolada de um índice de vegetação como o NDVI é eficaz para quantificar e localizar mudanças na cobertura vegetal, revelando *o que* mudou e *onde*. Contudo, para determinar *por que* a mudança ocorreu, é imperativo integrar variáveis climáticas que atuam como os principais vetores de controle dos ecossistemas terrestres. A temperatura da superfície (LST) e a precipitação são os fatores primários que ditam a disponibilidade de energia e água, respectivamente. A análise conjunta e sinérgica dessas três variáveis (NDVI, LST e Precipitação) permite a construção de um *framework* de diagnóstico capaz de distinguir entre alterações impulsionadas pela variabilidade climática e aquelas resultantes de ações antrópicas diretas.

3.3.1 Acoplamento Físico e Respostas Fisiológicas

As variáveis LST e precipitação não são independentes; elas estão intrinsecamente acopladas pelo balanço de energia na superfície. A precipitação aumenta a umidade do solo, que, por sua vez, alimenta o processo de evapotranspiração da vegetação e a evaporação direta do solo. Este processo consome energia na forma de calor latente, resultando em um resfriamento da superfície e, consequentemente, em uma menor LST [21]. Em cenários de seca (baixa precipitação), a evapotranspiração é reduzida, e a maior parte da radiação solar incidente é convertida em calor sensível, elevando drasticamente a LST [21].

A vegetação responde fisiologicamente a essas condições. A disponibilidade hídrica oriunda da precipitação é um fator limitante primário para a fotossíntese. Um aumento na umidade do solo permite que as plantas abram seus estômatos para absorver CO₂, resultando em maior atividade fotossintética e biomassa foliar, o que é diretamente capturado por um aumento no NDVI. Inversamente, a LST elevada é um indicador de estresse térmico e hídrico. Sob altas temperaturas e baixa umidade, as plantas fecham seus estômatos para conservar água, reduzindo a fotossíntese e, em casos extremos, levando à senescência foliar (processo de envelhecimento e queda das folhas), o que causa uma queda no NDVI.

Essa inter-relação gerou o conceito do “espaço LST-NDVI”, uma metodologia para avaliar o estado de umidade da superfície [21]. Em um gráfico de dispersão de LST versus NDVI para uma dada região, os pixels tipicamente formam uma estrutura triangular ou trapezoidal. A “borda seca” (LST alta, NDVI baixo) representa solos expostos e vegetação sob estresse máximo, enquanto a “borda úmida” (LST baixa, NDVI alto) representa uma

vegetação saudável com umidade ilimitada. A posição de uma área dentro deste espaço vetorial é, portanto, um diagnóstico de sua condição ambiental.

3.3.2 A Dinâmica do Lapso Temporal (Time Lag)

A resposta da vegetação às variações de precipitação não é instantânea. Existe um lapso de tempo (*time lag*) entre o evento de chuva e a resposta máxima do NDVI. Esse atraso ocorre porque a água precisa primeiro infiltrar no solo e se tornar disponível na zona radicular antes que os processos fisiológicos da planta resultem em um aumento mensurável da área foliar e do conteúdo de clorofila. A duração deste *time lag* é ecossistema-dependente: pastagens e vegetações de ciclo curto em regiões áridas podem responder em questão de dias ou semanas, enquanto ecossistemas florestais, com maior biomassa e inércia, podem levar de um a três meses para exibir uma resposta completa a um pulso de precipitação [29]. Modelar e compreender este lapso temporal é fundamental para verificar a correlação entre pluviosidade(chuva) e efeito (resposta do NDVI) e evitar conclusões espúrias.

3.3.3 Diferenciação de Vetores de Mudança: O Desacoplamento do Sinal

O objetivo final da análise integrada é a atribuição de causalidade, ou seja, distinguir uma mudança antrópica de uma mudança climática. Isso é alcançado através da análise do acoplamento (ou desacoplamento) entre o sinal do NDVI e os sinais climáticos.

- **Mudança Impulsionada pelo Clima:** Uma alteração na vegetação é considerada de origem climática quando sua resposta está fortemente acoplada às variáveis meteorológicas. Por exemplo, uma queda acentuada e generalizada do NDVI em uma região que coincide com uma anomalia negativa de precipitação e uma anomalia positiva de LST (ou seja, uma seca) é claramente uma resposta a um estresse climático. A relação entre as variáveis é a esperada.
- **Mudança Impulsionada por Ação Antrópica:** O sinal de uma intervenção humana direta é o **desacoplamento** da resposta do NDVI em relação aos vetores climáticos. Se uma área exibe uma queda abrupta e severa no NDVI, mas os dados de precipitação e LST para aquele local e período estão dentro da normalidade histórica, a causa não pode ser climática. Este desacoplamento aponta para um evento localizado e não-natural, como desmatamento por corte raso, um incêndio provocado, expansão urbana ou atividade de mineração.

Portanto, esta abordagem multivariada transforma o sensoriamento remoto de uma ferramenta de monitoramento para um *framework* de diagnóstico, permitindo não apenas mapear as mudanças, mas também inferir suas causas subjacentes com um maior grau de confiança científica.

3.4 Abstração de Dados Geoespaciais como Sinais e Padrões

A essência do reconhecimento de padrões e do processamento de sinais é a transformação de um fenômeno do mundo real em um objeto matemático (um sinal, uma imagem, um vetor de características) que pode ser analisado por algoritmos. Neste trabalho, o conjunto de dados de NDVI é abstraído de duas formas principais: como um conjunto de sinais temporais e como uma sequência de padrões espaciais [9].

3.4.1 Domínio Temporal: A Vegetação como um Sinal Digital

Para um dado local (pixel) ou região (e.g., um estado), a sequência de valores de NDVI ao longo do tempo (2019, 2020, ..., 2024) pode ser modelada como um sinal discreto unidimensional, $s(t)$ [27]. A análise deste sinal permite decompor a dinâmica da vegetação em componentes interpretáveis:

- **Componente de Baixa Frequência (Tendência):** A tendência de longo prazo (aumento ou diminuição gradual da vegetação) representa a componente de mais baixa frequência do sinal. A **análise de tendência** via regressão linear é uma forma de aplicar um filtro passa-baixa para isolar essa componente, modelada como $NDVI(t) = \beta_0 + \beta_1 t + \epsilon$. O coeficiente β_1 quantifica essa tendência [9].
- **Componente Periódica (Sazonalidade):** A vegetação exibe um ciclo anual previsível (fenologia), que se manifesta como uma forte componente periódica no sinal de NDVI. Análises mais avançadas (não implementadas aqui, mas parte do campo) utilizam a Transformada de Fourier para decompor o sinal em suas frequências constituintes, permitindo identificar o início, o pico e o fim da estação de crescimento [17].
- **Componente Aperiódica (Anomalias e Ruído):** Eventos abruptos e não previsíveis, como secas, incêndios ou desmatamento, se manifestam como desvios aperiódicos do comportamento esperado (tendência + sazonalidade). A **deteção de anomalias**, ao calcular a diferença entre o sinal observado e um modelo de referência (neste caso, a média de longo prazo), atua como um filtro para realçar esses eventos de alta frequência ou impulsivos.

3.4.2 Domínio Espacial: O Mosaico como um Padrão de Imagem

Cada mosaico anual de NDVI pode ser tratado como uma imagem digital, um padrão bidimensional onde cada pixel possui uma intensidade (o valor do NDVI). As técnicas de **Reconhecimento de Padrões** são então utilizadas para extrair informações desta imagem:

- **Extração de Características (Feature Extraction):** O próprio valor de NDVI de um pixel é a característica mais fundamental. Características mais complexas podem ser derivadas, como a textura, ou agregadas, como o NDVI médio de uma região (um estado), que reduz a dimensionalidade do padrão.
- **Deteção de Mudanças (Change Detection):** A criação de um **mapa de diferença** é uma técnica clássica de processamento de imagens, onde a subtração de duas imagens (NDVI de 2024 - NDVI de 2019) gera um novo padrão. Neste novo padrão, pixels com valores próximos a zero indicam ausência de mudança, enquanto pixels com altos valores (positivos ou negativos) revelam padrões espaciais de mudança significativa.
- **Classificação e Segmentação:** A criação de um **mapa coroplético** [24] é uma forma de classificação supervisionada. As classes são predefinidas (os estados) e o algoritmo atribui a cada classe uma “etiqueta” visual (uma cor) baseada em uma característica agregada (o NDVI médio), segmentando o mapa em regiões de acordo com o comportamento da vegetação. Técnicas não supervisionadas, como a clusterização, poderiam ser usadas em trabalhos futuros para encontrar “regiões naturais” com comportamento de NDVI similar, sem usar as fronteiras políticas.

4 Materiais e Métodos

A metodologia deste estudo foi estruturada em um fluxo de trabalho computacional automatizado, concebido para garantir a reprodutibilidade desde a aquisição dos dados primários até a geração das análises finais. Esta seção detalha a fonte dos dados geoespaciais e o conjunto de ferramentas e bibliotecas que constituem a arquitetura do projeto.

4.1 Fonte de Dados Geoespaciais

4.1.1 Dados de Satélite: Produto MODIS NDVI (MOD13A2)

O dado primário para a análise da vegetação foi o produto **MOD13A2 Versão 6.1**, derivado do sensor MODIS (Moderate Resolution Imaging Spectroradiometer), operante a bordo da plataforma orbital **Terra** da NASA [13]. A escolha deste produto foi estratégica, baseada nas seguintes características:

- **Resolução Temporal Elevada:** O MOD13A2 é um produto de composição de 16 dias. Isso significa que ele não representa uma única imagem, mas sim o melhor valor de pixel selecionado a partir de múltiplas aquisições diárias ao longo de um período de 16 dias. Este processo de composição é fundamental para análises em regiões tropicais como o Brasil, pois minimiza a contaminação por nuvens, sombras e aerossóis atmosféricos, resultando em uma série temporal mais limpa e consistente [6].
- **Resolução Espacial Adequada:** Com uma resolução espacial de 1 quilômetro por pixel, o produto é ideal para análises em escala continental. Ele oferece um balanço ótimo entre o detalhe geográfico e a viabilidade computacional para processar um grande volume de dados cobrindo todo o território nacional [6].
- **Processamento de Nível 3:** Sendo um produto de Nível 3, os dados já vêm georreferenciados, em uma grade senoidal padronizada, e com correções atmosféricas aplicadas, o que aumenta a confiabilidade das medições de reflectância de superfície [13].

Os arquivos, no formato HDF-EOS (Hierarchical Data Format - Earth Observing System), contêm múltiplos subdatasets científicos, dos quais a camada de NDVI foi a principal utilizada. Os dados foram acessados através do portal **NASA Earthdata**, especificamente do arquivo do LP DAAC (Land Processes Distributed Active Archive Center), para o período temporal de 1 de Janeiro de 2019 a 31 de Dezembro de 2024.

4.1.2 Dados Vetoriais de Referência

Para o recorte e a estratificação das análises, foram utilizados dois conjuntos de dados vetoriais (shapefiles):

- **Limite Nacional:** Polígono do território brasileiro obtido da base de dados *Natural Earth* (escala 1:110m), utilizado para o recorte geral dos mosaicos de NDVI.
- **Limites Estaduais:** Polígonos das Unidades da Federação do Brasil, obtidos da malha territorial de 2022 disponibilizada pelo **Instituto Brasileiro de Geografia e Estatística (IBGE)**, utilizados para a análise regional.

4.2 Fluxo de Trabalho Computacional e Ferramentas

O núcleo metodológico do projeto consiste em um fluxo de trabalho implementado na linguagem de programação **Python** (versão 3.11). A escolha do Python deve-se ao seu robusto ecossistema de bibliotecas para ciência de dados e geoprocessamento. A arquitetura foi modularizada em scripts para aquisição e para análise, garantindo eficiência e reprodutibilidade.

4.2.1 Bibliotecas Python Centrais

O fluxo de trabalho foi sustentado por um conjunto de bibliotecas de código aberto [18], cada uma com um papel específico:

- **earthaccess**: Utilizada como interface de alto nível para o sistema da NASA. A biblioteca automatizou o processo de autenticação, busca no catálogo do Common Metadata Repository (CMR) com base em critérios espaciais e temporais, e o download paralelo dos arquivos HDF.
- **rasterio**: Peça fundamental para toda a manipulação de dados raster. Foi empregada para: (i) abrir os arquivos HDF e ler os subdatasets de NDVI; (ii) utilizar o método `rasterio.merge.merge` para combinar múltiplos tiles de satélite em um mosaico anual contínuo; (iii) usar a função `rasterio.mask.mask` para recortar os dados raster com base nos polígonos vetoriais; e (iv) escrever os resultados finais no formato GeoTIFF.
- **geopandas**: Utilizada para a manipulação dos dados vetoriais. Suas funções foram cruciais para ler os arquivos de shapefile do Brasil e dos estados e, principalmente, para realizar a reprojeção dos sistemas de coordenadas de referência (CRS) dos vetores para que coincidissem com o CRS dos rasters, um passo essencial para o recorte correto.
- **pandas**: Empregada para a estruturação e manipulação de todos os dados tabulares. Os resultados estatísticos e regionais foram organizados em DataFrames para facilitar os cálculos e, posteriormente, exportados para o formato CSV.
- **numpy**: A base para toda a computação numérica. Os dados de pixel dos rasters foram carregados como arrays NumPy, permitindo a execução de operações matemáticas vetorizadas de forma eficiente, como a aplicação do fator de escala, a substituição de valores nulos (*nodata*) e o cálculo de estatísticas.
- **scipy.stats**: Especificamente, a função `linregress` foi utilizada para realizar a análise de regressão linear ordinária sobre a série temporal do NDVI médio, calculando a inclinação da tendência, o coeficiente de determinação (R^2) e o p-valor.
- **matplotlib**: Utilizada como motor de visualização para gerar todos os gráficos estáticos (gráficos de linha, de barra e mapas temáticos) de forma programática, garantindo que as visualizações pudessem ser reproduzidas e salvas em alta resolução para inclusão neste relatório.

4.2.2 Software de Sistema de Informação Geográfica (SIG)

- **QGIS 3.x**: Utilizado como ferramenta de apoio essencial para a validação e análise exploratória dos dados. O software foi empregado para: (i) realizar a inspeção visual

dos mosaicos GeoTIFF anuais, verificando sua integridade e cobertura espacial; (ii) explorar interativamente os valores de pixel; e (iii) validar a sobreposição e o alinhamento entre os dados raster e os vetoriais, garantindo a precisão dos recortes e das análises regionais [19].

4.3 Descrição da Metodologia Computacional

O fluxo de trabalho foi implementado em um ambiente Python e modularizado em dois scripts principais, desacoplando o processo de aquisição de dados, que é computacionalmente intensivo, da etapa de análise e visualização, que é mais ágil. Esta abordagem garante a reprodutibilidade e permite a re-análise dos resultados sem a necessidade de reprocessar os dados brutos.

4.3.1 Script 1: coleta_dados_NDVI.py

Este script constitui o motor de processamento de dados do projeto, responsável por todas as operações pesadas, desde a busca até a geração dos produtos de base.

Busca e Aquisição de Dados. A primeira etapa consiste na interação programática com a infraestrutura de dados da NASA. Utilizando a biblioteca earthaccess, o script realiza uma busca no Common Metadata Repository (CMR) por todos os grânulos do produto MOD13A2 que intersectam a caixa delimitadora (bounding box) do território brasileiro para cada ano do período de estudo. A busca é definida por parâmetros explícitos, como demonstrado no Listagem 1.

```
1 # Definição dos parâmetros para a busca de dados
2 results = earthaccess.search_data(
3     short_name='MOD13A2',
4     bounding_box=brasil_aoi.total_bounds,
5     temporal=(f'{year}-01-01', f'{year}-12-31'),
6 )
7
8 # Download dos arquivos encontrados
9 filepaths = earthaccess.download(results, local_path=local_dir)
```

Listing 1: Trecho de código para busca de dados na API da NASA.

Processamento e Mosaico Anual. Após o download, o script itera sobre cada arquivo HDF-EOS baixado. Para cada arquivo, a camada científica correspondente ao NDVI é extraída. Esta seleção identifica o subdataset cujo nome contém a string “NDVI”, como no Listagem 2.

```
1 # Identifica o caminho do subdataset de NDVI dentro do arquivo HDF
2 with rasterio.open(hdf_file) as src:
3     ndvi_sds_path = [s for s in src.subdatasets
4                       if 'NDVI' in s.split(':')[0]]
```

Listing 2: Extração do subdataset NDVI do arquivo HDF.

Cada camada de NDVI extraída é então recortada (*clipped*) usando a geometria do polígono do Brasil, previamente alinhada ao sistema de coordenadas do raster. Ao final do processamento de todos os arquivos de um ano, os múltiplos *tiles* recortados são combinados em um único mosaico anual. Para esta etapa, utilizou-se a função `rasterio.merge.merge` com o método ‘`max`’ (Listagem 3). A escolha deste método é uma técnica padrão em composição de NDVI, pois para cada pixel sobreposto, ele seleciona o valor máximo, o que tende a favorecer pixels com vegetação mais vigorosa e a minimizar a influência de nuvens residuais ou outras contaminações atmosféricas.

```
1 # Combina a lista de rasters em memória em um único mosaico
2 mosaic, out_trans = merge(srcs_to_mosaic, method='max')
3
4 # Aplica o fator de escala e trata valores nulos
5 ndvi_raw = mosaic[0].astype(float)
6 ndvi_raw[ndvi_raw == -3000] = np.nan
7 ndvi = ndvi_raw * 0.0001
```

Listing 3: Combinação dos tiles em um mosaico anual pelo método de máximo valor.

O mosaico resultante é então salvo como um arquivo GeoTIFF, que servirá de base para todas as análises subsequentes.

Extração de Dados Regionais. A etapa final deste script é a extração de estatísticas zonais. A função `gerar_dados_regionais` itera sobre cada mosaico GeoTIFF anual e, para cada um, calcula o valor médio de NDVI para cada estado brasileiro, utilizando o shapefile do IBGE como máscara. O resultado é consolidado e salvo em um único arquivo CSV, `ndvi_medio_por_estado.csv`, que servirá de entrada para as análises regionais.

4.3.2 Script 2: analise_NDVI.py

Este script é a camada de análise e visualização do projeto. Ele opera exclusivamente sobre os arquivos pré-processados pelo Script 1, o que o torna extremamente rápido e flexível.

Cálculo de Estatísticas e Análises Temporais. O script primeiramente carrega todos os GeoTIFFs anuais para calcular as estatísticas descritivas (média, mediana, desvio padrão, etc.) para o Brasil, os *outputs*, podem ser verificados na pasta `resultados` e `resultados_imagens`. Em seguida, este arquivo é utilizado para as análises temporais. A análise de tendência, por exemplo, é realizada aplicando a função `linregress` da biblioteca SciPy, como mostra a Listagem 4.

```
1 # Carrega os dados anuais e aplica a regressão
2 anos, ndvi_medio = df_stats['Ano'], df_stats['NDVI Médio']
3 slope, intercept, r_value, p_value, std_err = stats.linregress(
4     anos, ndvi_medio
5 )
```

Listing 4: Cálculo da regressão linear para a análise de tendência.

Análise de Mudança Espacial. Para identificar *hotspots* de mudança, o script `analise_NDVI.py` carrega o primeiro e o último GeoTIFF da série, usando a função `analise_mapa_de_diferenca`. A detecção de mudança é feita através de uma operação de álgebra de mapas: a subtração dos arrays NumPy correspondentes aos dados de pixel, como visto na Listagem 5.

```
1 # Carrega os rasters do primeiro e último ano e subtrai seus arrays
2 with rasterio.open(geotiff_files[0]) as src_primeiro, \
3     rasterio.open(geotiff_files[-1]) as src_ultimo:
4
5     diferenca_ndvi = src_ultimo.read(1) - src_primeiro.read(1)
```

Listing 5: Cálculo do mapa de diferença por subtração de arrays.

Geração de Gráficos e Mapas Temáticos. Finalmente, o script utiliza os arquivos CSV do folder `resultados` e o shapefile dos estados para gerar todas as figuras do relatório. A biblioteca `Matplotlib` é usada para os gráficos de linha e barra, enquanto a integração do `GeoPandas` com o `Matplotlib` é usada para criar o mapa coroplético que visualiza a distribuição do NDVI médio entre os estados.

4.3.3 Metodologia Computacional via Autoencoder ConvLSTM

O processo de detecção de anomalias, encapsulado no script `neural_analise.py`, foi executado de forma independente para cada uma das cinco grandes regiões do Brasil. Esta abordagem permite a criação de modelos especialistas, otimizados para os padrões de vegetação de cada macro-bioma. As etapas técnicas são as seguintes:

- 1. Carregamento e Preparo dos Dados:** O processo inicia com o carregamento dos mosaicos anuais de NDVI (em formato GeoTIFF) gerados na etapa de pré-processamento. Estes são empilhados para formar um tensor de dados espaço-temporal com dimensões (tempo, altura, largura).
- 2. Delimitação Regional (ROI):** A delimitação das Regiões de Interesse (ROI) foi realizada a partir de dados vetoriais de fontes oficiais. Foram utilizados os polígonos das Unidades da Federação do Brasil, obtidos da malha territorial de 2022 disponibilizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) [11]. Utilizando a biblioteca `GeoPandas`, o script primeiramente agrupa estas geometrias estaduais nas cinco grandes regiões através da operação de dissolução geoespacial (`dissolve(by='NM.REGIAO')`). Em seguida, para cada região, o tensor de dados nacional é recortado (`rasterio.mask.mask`) de acordo com sua geometria específica, garantindo que cada modelo seja treinado apenas com os dados de sua respectiva área. A reprojeção do sistema de coordenadas (CRS) do vetor para o do raster, via método `.to_crs()`, é um passo crítico executado nesta etapa para garantir a compatibilidade e precisão espacial.
- 3. Geração de Amostras (Patches):** O treinamento de redes convolucionais em imagens de alta resolução é computacionalmente intensivo. Para viabilizá-lo, a série temporal de cada região é segmentada em “patches” (recortes) menores, de dimensões 32x32 pixels. Esta técnica aumenta exponencialmente o número de amostras de treinamento e permite o processamento em lotes (batches), otimizando o uso

de memória. Os dados são encapsulados em uma classe `SpatiotemporalDataset` customizada para o PyTorch.

4. Arquitetura e Treinamento do Modelo: O núcleo da análise é um **Autoencoder ConvLSTM**.

- **Encoder:** Uma sequência de Células Convolucionais de Memória de Curto e Longo Prazo (ConvLSTM) [22] processa a série temporal de cada patch. Sua função é aprender a comprimir a informação espaço-temporal em um vetor de características latentes (o “estado oculto” final), que representa a essência da dinâmica daquele local.
- **Decoder:** Outra sequência de células ConvLSTM recebe esse vetor de características e tenta reconstruir a sequência original de imagens, passo a passo.
- **Aprendizado:** O modelo é treinado para minimizar o Erro Quadrático Médio (*Mean Squared Error* - MSE) entre os patches originais e os reconstruídos. Ao iterar por um número substancial de épocas (neste caso, 50), a rede se torna eficiente em reconstruir padrões de vegetação “normais” e previsíveis para aquela região.

5. Detecção de Anomalias por Erro de Reconstrução: Após o treinamento, o modelo especialista de cada região é utilizado para reconstruir todo o conjunto de dados regional. O erro de reconstrução é então calculado para cada pixel individualmente, como a média do erro quadrático ao longo de toda a série temporal. Um erro baixo significa que a evolução da vegetação naquele pixel foi “normal” e bem aprendida pelo modelo. Um erro alto, por sua vez, sinaliza uma anomalia: uma evolução atípica que o modelo não conseguiu prever, indicando uma provável mudança significativa.

6. Geração do Mapa Final: Para cada região, o mapa de erro contínuo gerado pela rede neural é submetido a um processo de pós-processamento para se obter o mapa binário de mudanças. A abordagem implementada no script `neural_analise.py` não utiliza operações morfológicas complexas, mas foca em uma limiarização (thresholding) estatística para isolar as anomalias mais extremas.

O método consiste em duas etapas principais, conforme detalhado no trecho de código da Listagem 6:

- (a) Primeiramente, a distribuição empírica de todos os valores de erro de reconstrução do mapa é analisada. Utilizando a função `numpy.percentile`, calcula-se o valor de erro que corresponde ao percentil 95 (p_{95}). Este valor, denotado como τ , representa o limiar acima do qual se encontram os 5% de pixels com os maiores erros de reconstrução em toda a região.
- (b) Em seguida, este limiar τ é usado para binarizar o mapa. Uma operação lógica é aplicada a cada pixel: se o seu valor de erro $e_{i,j}$ for maior ou igual a τ , ele é classificado como “mudança significativa” (recebendo o valor 1); caso contrário, é classificado como “não-mudança” (recebendo o valor 0).

Essa abordagem garante que apenas as anomalias mais estatisticamente improváveis, localizadas na cauda superior da distribuição de erros, sejam destacadas como pontos de mudança, resultando no mapa final.

```
1 # Trecho de código do script: neural_analise.py
2
3 # --- Geração do Mapa Binário de Mudanças ---
4
5 # O mapa 'error_map' contém os valores de erro de reconstrução para cada
6 # pixel.
7 # Filtramos apenas os erros maiores que zero para o cálculo do
8 # percentil.
9 threshold = np.percentile(error_map[error_map > 0], 95)
10
11 # Aplica o limiar: qualquer pixel com erro >= threshold se torna True
12 # → (1),
13 # e os demais se tornam False (0). O resultado é um array booleano
14 # que é convertido para um inteiro de 8 bits (uint8).
15 change_map = (error_map > threshold).astype(np.uint8)
16
17 # --- Salvando o resultado ---
18 # Define o perfil do novo GeoTIFF, especificando o tipo de dado como
19 # → uint8.
20 change_profile = current_profile.copy()
21 change_profile.update(dtype=rasterio.uint8, count=1, nodata=0)
22
23 # Define o caminho do arquivo de saída.
24 change_map_path_tif = os.path.join(OUTPUT_DIR,
25 → f'mapa_mudancas_{nome_regiao}.tif')
26
27 # Escreve o array do mapa de mudanças no arquivo GeoTIFF.
28 with rasterio.open(change_map_path_tif, 'w', **change_profile) as dst:
29     dst.write(change_map, 1)
30
31 print(f" -> Mapa de mudanças (GeoTIFF) salvo em: {change_map_path_tif}")
```

Listing 6: Implementação da limiarização estatística em Python para gerar o mapa de mudanças.

5 Análise dos Resultados Baseados em NDVI

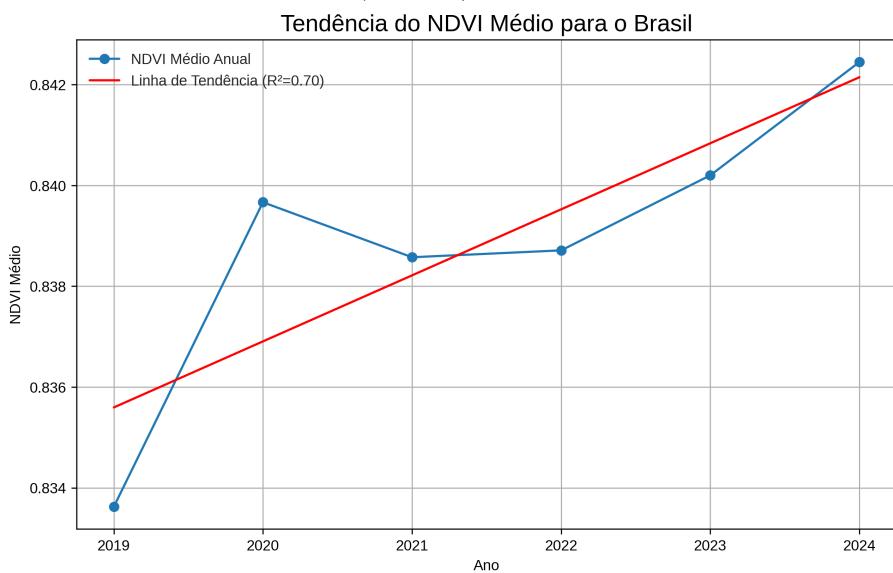
Nesta seção, são apresentados e discutidos os principais resultados obtidos a partir do processamento dos dados de NDVI para o período de 2019 a 2024. A análise foi estruturada em cinco eixos principais: (i) a caracterização estatística e a tendência temporal para o Brasil como um todo; (ii) a identificação de anomalias anuais; (iii) a análise espacial das áreas de mudança; (iv) a análise de agrupamento espacial (clusters) para identificação de macrozonas ecológicas; e (v) a análise detalhada das tendências regionais por estado.

5.1 Análise Estatística e de Tendência Temporal

A análise da série histórica do NDVI médio para o território brasileiro, ilustrada na Figura 1, revela uma clara e estatisticamente significativa tendência de “esverdeamento” (greening) ao longo do período de 2019 a 2024. Os valores médios anuais oscilaram, partindo de um vale em 2019, atingindo um pico intermediário em 2020 e culminando no valor mais alto da série em 2024.

Para quantificar a direção da mudança ao longo do tempo, foi realizada uma análise de regressão linear. O modelo de regressão aponta para uma forte tendência geral de crescimento da vegetação, com um coeficiente de determinação (R^2) de 0.70. Este valor robusto indica que 70% da variação no NDVI médio anual pode ser explicada pela tendência linear de crescimento ao longo dos seis anos. Diferente de uma análise preliminar que sugeria estagnação, os dados consolidados apontam para uma melhora geral no vigor da vegetação em escala nacional, embora, como veremos, essa tendência mascare dinâmicas regionais bastante distintas.

Figura 1: Série temporal e linha de tendência do NDVI médio para o Brasil (2019-2024). A linha vermelha representa a forte tendência linear positiva ($R^2=0.70$).

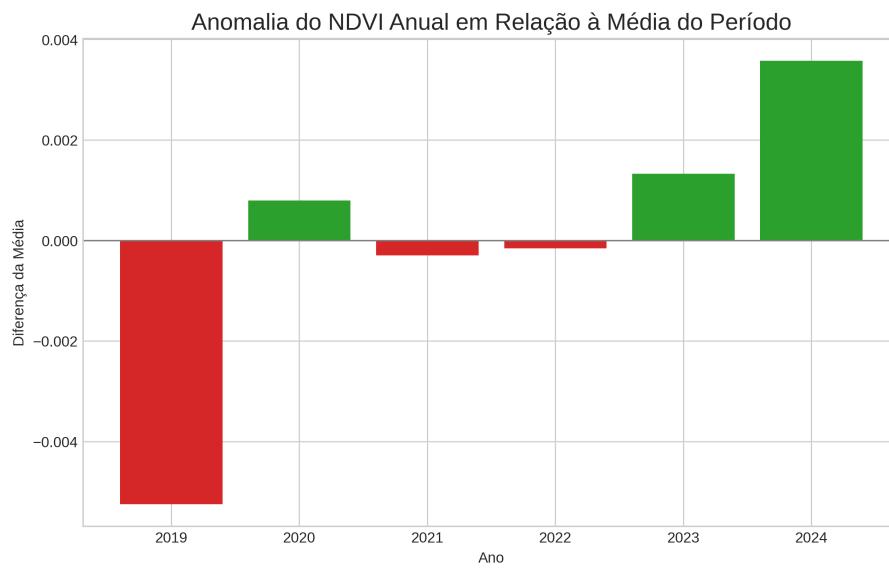


5.2 Análise de Anomalias Anuais

Para isolar os anos com comportamento atípico, foram calculadas as anomalias do NDVI médio anual em relação à média de todo o período. O ano de **2020** apresentou uma anomalia positiva significativa, indicando condições ambientais que foram particularmente

favoráveis ao desenvolvimento da vegetação em escala nacional. Em contrapartida, os anos de **2019** e **2023**, exibiram as maiores anomalias negativas. Estes resultados sugerem que, nestes anos, o Brasil pode ter enfrentado condições de estresse hídrico mais severas ou uma temporada de incêndios mais intensa do que a média, impactando negativamente a saúde da vegetação em grandes áreas. A análise de anomalias é, portanto, uma ferramenta eficaz para sinalizar anos de possíveis eventos climáticos extremos.

Figura 2: Anomalia do NDVI médio anual em relação à média do período 2019-2024. Barras verdes indicam anos com vegetação acima da média; barras vermelhas, abaixo da média.



5.3 Evolução Espacial Anual da Vegetação (2019-2024)

A análise da sequência de mapas anuais de NDVI permite uma investigação detalhada da distribuição espacial e da variação do vigor da vegetação para cada ano do período estudado. Esta perspectiva cronológica é fundamental para correlacionar os padrões observados com eventos climáticos e ambientais específicos de cada ano.

O mapa de NDVI para 2019 (Figura 3) revela um cenário de estresse vegetativo em importantes biomas brasileiros. Este foi um ano marcado por uma intensa temporada de queimadas na Amazônia e no Pantanal, além de condições de seca em porções do Centro-Oeste e Sudeste [4]. Visualmente, é possível notar valores de NDVI ligeiramente inferiores na borda sul da Amazônia (o “arco do desmatamento”) e uma menor intensidade do verde no Brasil Central. Estes padrões são consistentes com os registros de anomalias climáticas negativas para o período, que resultaram em um dos pontos mais baixos da série histórica do NDVI nacional, conforme mostrado nas figuras 1 e 2.

Figura 3: Mapa de NDVI para o Brasil em 2019.



Em 2020, o mapa (Figura 4) demonstra uma recuperação notável em parte do território, especialmente na Amazônia, que registrou um “esverdeamento” significativo, resultando em uma forte anomalia positiva no NDVI nacional. Este fenômeno pode estar associado a um regime de chuvas mais favorável em porções do bioma. Contudo, 2020 também foi o ano de queimadas históricas no Pantanal. Embora o NDVI seja uma média anual, a degradação severa neste bioma pode ser notada como uma mancha de verde menos intenso em sua localização geográfica.

Figura 4: Mapa de NDVI para o Brasil em 2020.



O ano de 2021 foi caracterizado por uma das piores crises hídricas no Centro-Sul do Brasil, sob a influência do fenômeno La Niña [1]. O mapa de NDVI (Figura 5) reflete este cenário, com uma visível redução do vigor da vegetação na Bacia do Paraná, afetando biomas como o Cerrado e a Mata Atlântica em áreas do Sudeste e Sul. A mancha de verde aparece mais “pálida” nessas regiões em comparação com 2020, corroborando os impactos da seca severa na agricultura e nos ecossistemas naturais.

Figura 5: Mapa de NDVI para o Brasil em 2021.



A continuidade do fenômeno La Niña em 2022 manteve a pressão sobre o Sul do Brasil, que enfrentou mais um ano de seca severa, especialmente no Rio Grande do Sul, com perdas significativas na produção agrícola [5]. Por outro lado, o mesmo fenômeno favoreceu um regime de chuvas acima da média em partes do Norte e Nordeste do país, causando inundações em alguns estados [30]. O mapa de 2022 (Figura 6) captura essa dualidade, mostrando valores de NDVI relativamente baixos no extremo sul do país, enquanto áreas do sertão nordestino apresentam uma condição vegetativa mais vigorosa que no ano anterior.

Figura 6: Mapa de NDVI para o Brasil em 2022.



O ano de 2023 foi marcado pela transição para um forte El Niño, que causou extremos climáticos opostos. A Bacia Amazônica enfrentou uma seca histórica e severa, com rios atingindo níveis mínimos e um aumento expressivo no risco de incêndios, evento amplamente documentado como um dos mais intensos já registrados [2]. Este impacto é visível no mapa de NDVI (Figura 7), onde a mancha verde da Amazônia se mostra notavelmente menos intensa do que nos anos anteriores. Simultaneamente, o Sul do país experimentou chuvas torrenciais e inundações, impulsadas por ciclones extratropicais associados ao mesmo fenômeno [14], que, paradoxalmente, podem levar a um aumento pontual do NDVI devido à alta disponibilidade de água para a vegetação que sobreviveu.

Figura 7: Mapa de NDVI para o Brasil em 2023.



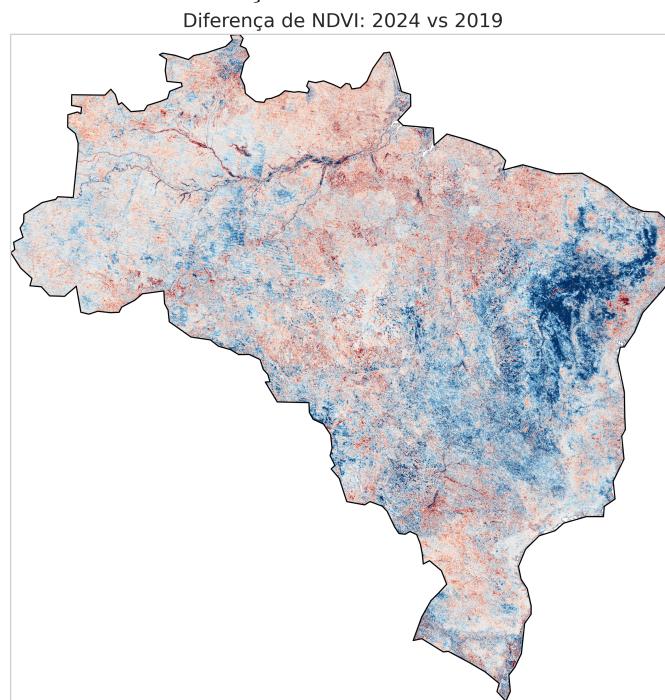
O mapa de 2024 (Figura 8) representa o pico de vigor vegetativo da série histórica em escala nacional. Este resultado é impulsionado por uma forte recuperação em áreas do Brasil Central e Sudeste, que saíram do período de estresse hídrico dos anos anteriores. No entanto, o mapa também captura os efeitos dos eventos extremos do final de 2023 e início de 2024. A recuperação na Amazônia ainda é parcial após a seca severa, e o Sul do país, embora verde, estava no limiar das inundações catastróficas que ocorreram no Rio Grande do Sul, cujos impactos totais no NDVI só serão plenamente visíveis em análises futuras.

Figura 8: Mapa de NDVI para o Brasil em 2024.



A comparação direta entre o início e o fim do período, através do mapa de diferença 9, sintetiza a evolução de longo prazo. As cores frias (azul) indicam áreas onde o NDVI em 2024 foi menor que em 2019, sugerindo perda líquida ou degradação da vegetação. Estas áreas se concentram no “arco do desmatamento” [31]. Por outro lado, as cores quentes (vermelho) indicam um “esverdeamento” líquido. Este fenômeno é visível em grandes extensões do Brasil Central, Sudeste e partes do Nordeste, e está alinhado com a tendência geral positiva observada para o país, sendo possivelmente um reflexo da intensificação agrícola e de mudanças no regime de chuvas.

Figura 9: Mapa de diferença de NDVI (2024 - 2019). Tons de azul indicam aumento no vigor da vegetação, enquanto tons de vermelho indicam diminuição.

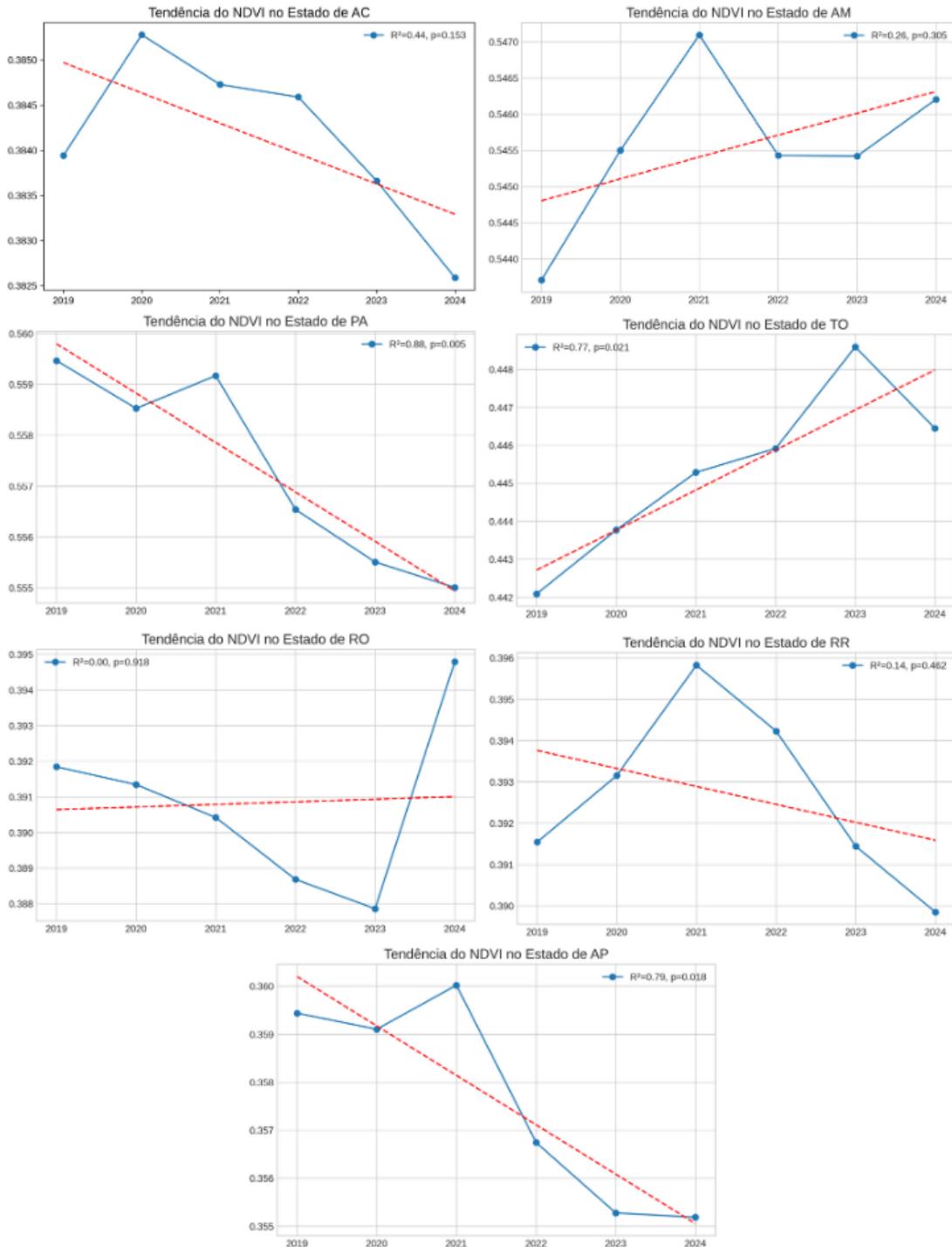


5.4 Análise das Tendências Regionais por Estado

A agregação dos dados por unidade federativa revela que a tendência nacional positiva é um mosaico de diferentes realidades regionais, com dinâmicas de degradação e regeneração ocorrendo simultaneamente.

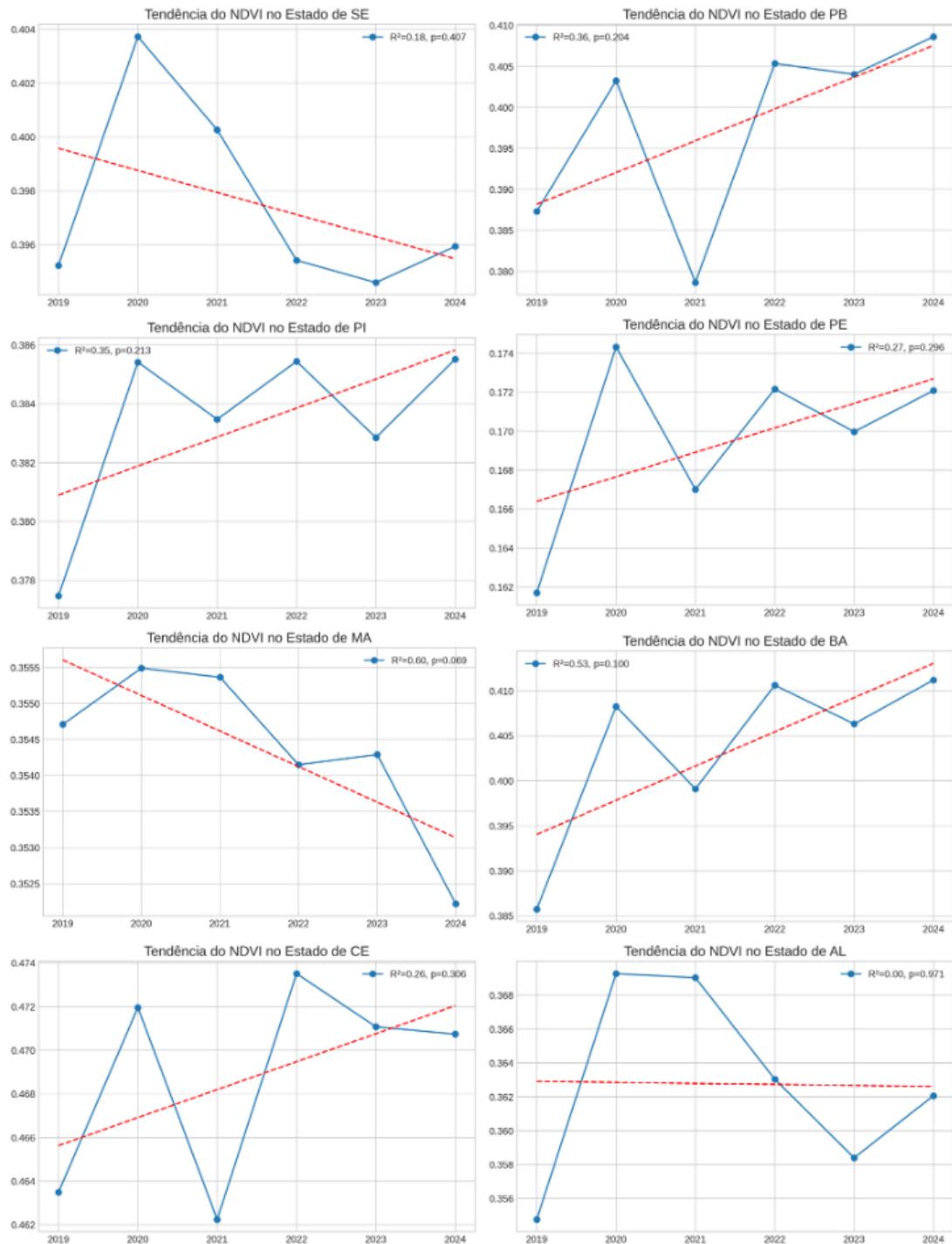
Região Norte (Figura 10): A Região Norte exibe um padrão preocupante e heterogêneo. Estados como **Acre (AC)**, **Pará (PA)**, **Rondônia (RO)** e **Amapá (AP)** mostram uma clara tendência de queda no NDVI, consistente com a pressão do “arco do desmatamento”. As tendências para Pará ($p<0.05$) e Amapá ($p<0.05$) são estatisticamente significativas. Em contraste, **Amazonas (AM)**, **Tocantins (TO)** e **Roraima (RR)** apresentam uma leve tendência de alta no período.

Figura 10: Tendências de NDVI para os estados da Região Norte.



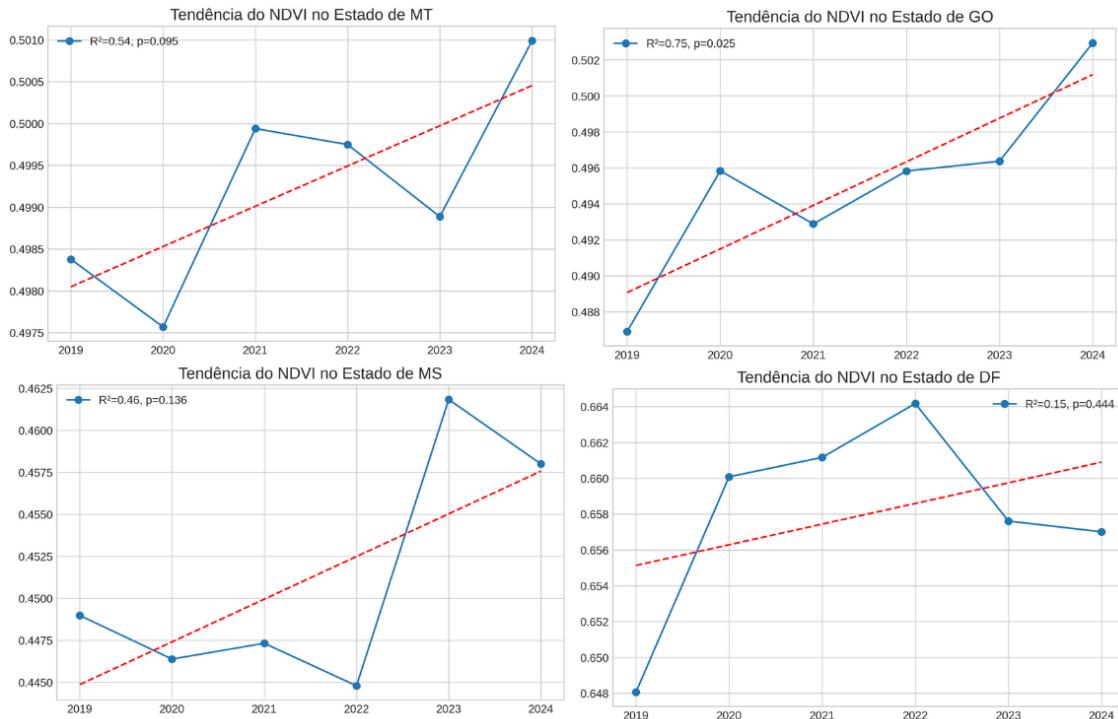
Região Nordeste (Figura 11): O Nordeste é caracterizado por alta volatilidade interanual. Apesar disso, a maioria dos estados, como **Piauí (PI)**, **Ceará (CE)**, **Bahia (BA)** e **Paraíba (PB)**, apresenta uma tendência de fundo positiva, embora sem significância estatística. **Maranhão (MA)** e **Sergipe (SE)** mostram uma leve tendência de queda, enquanto **Alagoas (AL)** se mantém praticamente estável.

Figura 11: Tendências de NDVI para os estados da Região Nordeste.



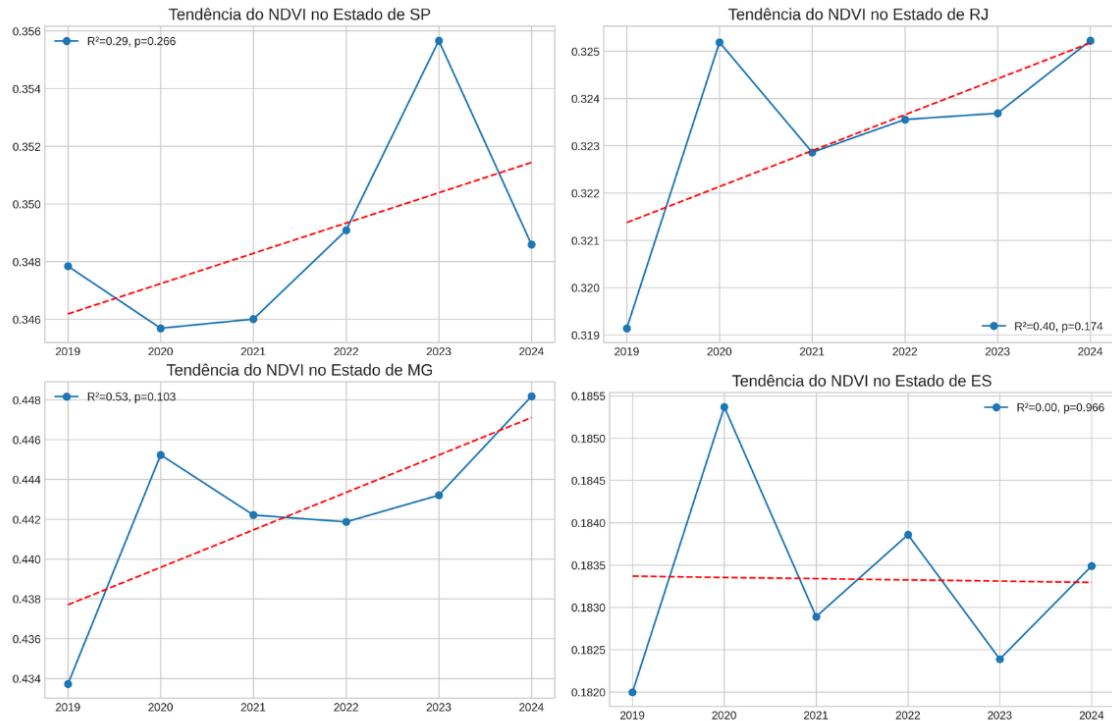
Região Centro-Oeste (Figura 12): O Centro-Oeste mostra uma tendência geral de “esverdeamento”. Goiás (GO) lidera com um aumento estatisticamente significativo ($p=0.025$). Mato Grosso (MT) e Mato Grosso do Sul (MS) também mostram tendências positivas, embora com maior variabilidade e sem significância estatística, refletindo a dinâmica complexa do agronegócio e a sazonalidade do Cerrado.

Figura 12: Tendências de NDVI para os estados da Região Centro-Oeste.



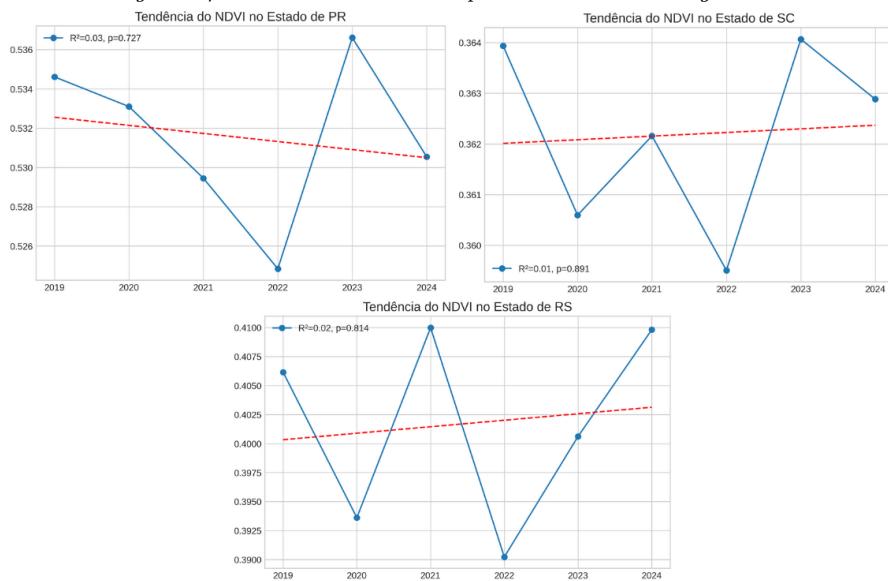
Região Sudeste (Figura 13): Todos os quatro estados do Sudeste apresentam uma tendência de aumento do NDVI, embora nenhuma seja estatisticamente significativa no período analisado. Esta tendência de “esverdeamento” pode estar associada a uma combinação de fatores, como a regeneração de áreas da Mata Atlântica, mudanças no uso do solo e variações climáticas favoráveis.

Figura 13: Tendências de NDVI para os estados da Região Sudeste.



Região Sul (Figura 14): A Região Sul apresenta uma dinâmica de alta variabilidade e ausência de uma tendência clara. **Paraná (PR)** mostra uma leve tendência de queda, enquanto **Santa Catarina (SC)** e **Rio Grande do Sul (RS)** demonstram uma leve tendência de alta, porém nenhuma dessas tendências é estatisticamente significativa. A grande oscilação anual pode estar ligada a eventos climáticos extremos, como as secas que afetaram a região no período.

Figura 14: Tendências de NDVI para os estados da Região Sul.



5.5 Análise de Agrupamento Espacial (Clusters)

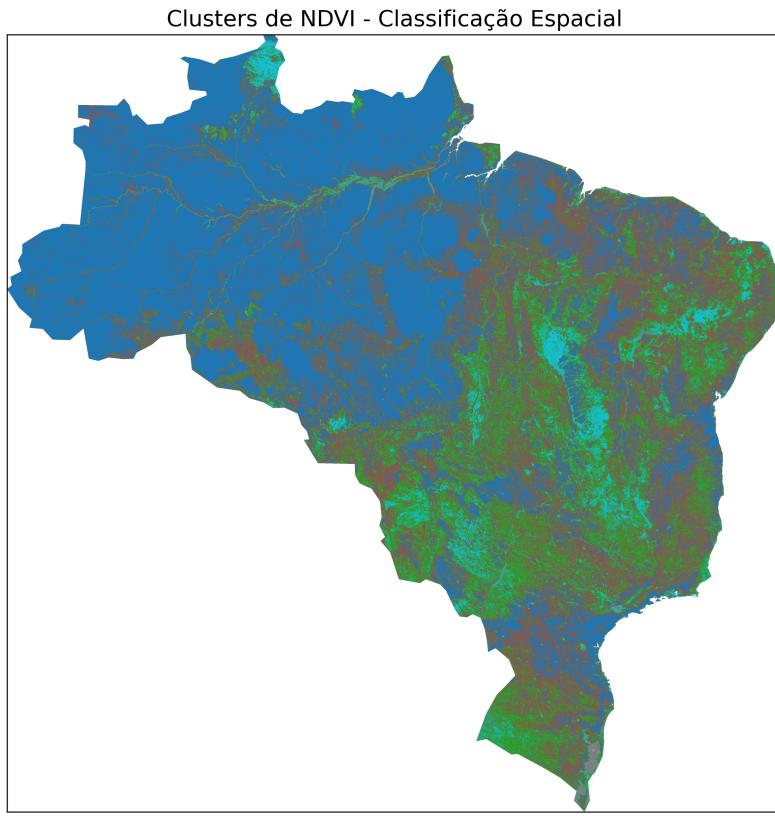
Para identificar macrozonas com características de vegetação similares de forma automatizada, foi aplicada uma técnica de aprendizado de máquina não supervisionado (K-Means Clustering). Este método agrupa todos os pixels do território em um número pré-definido de clusters (neste caso, 5), com base unicamente em seus valores de NDVI, sem conhecimento prévio sobre biomas ou limites estaduais. O resultado é um mapa de zoneamento ecológico (Figura 15).

Interpretação do Mapa de Clusters: Cada cor no mapa representa um cluster, ou seja, uma região com um comportamento de NDVI distinto e internamente consistente. A interpretação de cada cluster é feita por correlação com a geografia conhecida do Brasil:

- **Cluster Azul Escuro:** Dominante na Bacia Amazônica, este cluster representa as áreas de maior densidade de vegetação, correspondendo à floresta tropical úmida e contínua. São as regiões com os valores de NDVI mais elevados.
- **Cluster Verde:** Este grupo aparece de forma proeminente no bioma Cerrado, em áreas da Mata Atlântica e em zonas de transição. Representa um NDVI de nível intermediário a alto, característico de savanas mais densas, florestas secundárias e áreas agrícolas intensivas.
- **Cluster Ciano (Azul Claro):** Observado em manchas dentro da Amazônia, este cluster provavelmente corresponde a áreas de vegetação de várzea ou áreas com um dossel florestal ligeiramente diferente, que possuem uma assinatura de NDVI distinta da floresta de terra firme.
- **Cluster Marrom/Vermelho:** Este cluster concentra-se na região Nordeste, alinhado com o bioma Caatinga, e em áreas do Pantanal. Representa as zonas com os menores valores de NDVI, características de vegetação mais esparsa, adaptada a climas mais secos.

A análise de clusters é uma ferramenta para o reconhecimento de padrões, pois segmenta o país em domínios ecológicos funcionais de forma objetiva, baseada apenas nos dados de satélite.

Figura 15: Mapa de clusters de NDVI gerado por K-Means. Cada cor representa uma zona com características de vegetação estatisticamente similares.



5.6 Detecção de Anomalias Temporais com Redes Neurais

Para aprofundar a identificação de áreas com alterações significativas, foi implementada uma análise avançada utilizando uma arquitetura de rede neural profunda. O objetivo desta técnica é modelar a complexa dinâmica espaço-temporal da vegetação e, a partir dela, detectar anomalias que podem corresponder a eventos de desmatamento, regeneração, ou outras transformações abruptas no uso do solo.

5.6.1 Interpretação do Mapa de Anomalias Regionais

O mapa consolidado na Figura 16 exibe as áreas de anomalia detectadas por modelos especialistas em cada uma das cinco grandes regiões do Brasil, com uma cor designada para cada uma. Os pontos coloridos não representam ruído aleatório; eles indicam os locais onde a evolução do NDVI entre 2019 e 2024 foi mais “inesperada” ou anômala segundo a rede neural treinada para os padrões daquela região específica.

Figura 16: Mapa consolidado de anomalias temporais detectadas pela rede neural ConvLSTM. Cada cor representa os 5% de pixels com maior erro de reconstrução em sua respectiva região, sinalizando mudanças significativas na cobertura vegetal entre 2019 e 2024.



A interpretação das anomalias deve ser feita considerando o contexto de cada bioma:

- **Norte (Verde):** Os pontos verdes destacam anomalias na Bacia Amazônica. Sua distribuição espacial alinha-se de forma consistente com as fronteiras do “arco do desmatamento”.
- **Nordeste (Vermelho):** As anomalias em vermelho na região Nordeste ocorrem em um ambiente de alta variabilidade climática (Caatinga e Cerrado). Os pontos podem indicar tanto áreas de degradação e desertificação acelerada, impulsionadas por secas severas, quanto eventos de “esverdeamento” muito rápidos e atípicos após anos excepcionalmente chuvosos, que também fogem à norma temporal.
- **Centro-Oeste (Lilás):** Os pontos lilás estão imersos na principal região do agro-negócio brasileiro e do bioma Cerrado. As anomalias aqui podem ter múltiplas

causas: a conversão de novas áreas de vegetação nativa para agricultura, a rotação intensa de culturas com diferentes assinaturas de NDVI, ou a implementação de grandes projetos de irrigação que alteram drasticamente o ciclo anual da vegetação.

- **Sudeste (Azul):** As anomalias em azul, na região mais urbanizada do país, tendem a ser mais localizadas. Podem representar a expansão de perímetros urbanos sobre áreas verdes, mudanças significativas no manejo agrícola (ex: conversão de pastagens para culturas anuais como soja ou cana-de-açúcar) ou, inversamente, projetos de reflorestamento e recuperação de áreas degradadas da Mata Atlântica.
- **Sul (Preto):** Os pontos pretos na Região Sul ocorrem em um contexto de clima temperado e agricultura bem estabelecida. As anomalias são fortemente influenciadas por eventos climáticos extremos. Anos de seca severa (associados ao fenômeno La Niña) ou de inundações catastróficas (como as ocorridas em 2024) causam quedas ou picos abruptos no NDVI que são corretamente identificados como desvios do padrão temporal médio da região.

6 Análise dos Resultados Baseados em Dados Climáticos

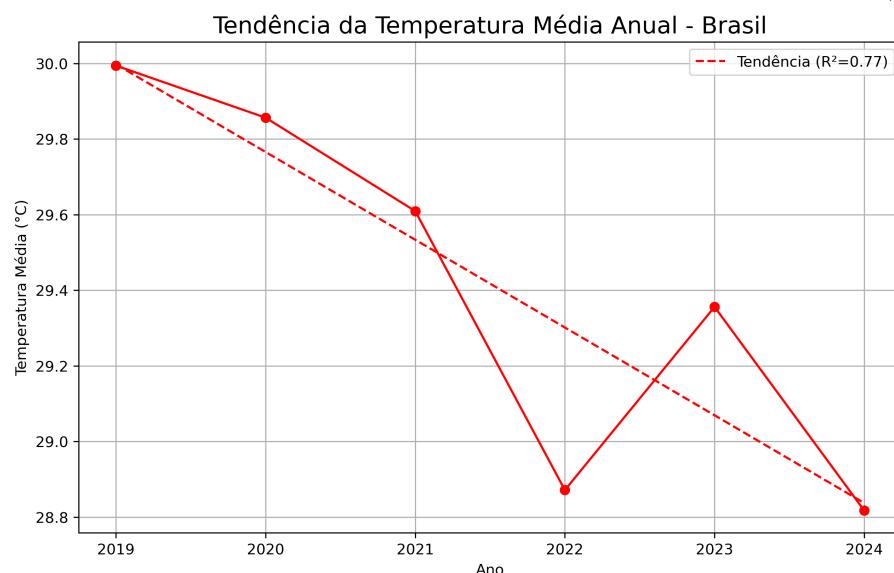
Nesta seção, são apresentados os resultados derivados da análise das variáveis climáticas — Temperatura da Superfície Terrestre (LST) e Precipitação Pluviométrica — para o período de 2019 a 2024. A análise foca nas tendências temporais, na identificação de anomalias anuais e na caracterização de zonas climáticas através de clusterização.

6.1 Tendências Temporais de Temperatura e Pluviosidade

A análise da série histórica das médias nacionais revela dinâmicas distintas para a temperatura e a pluviosidade no período.

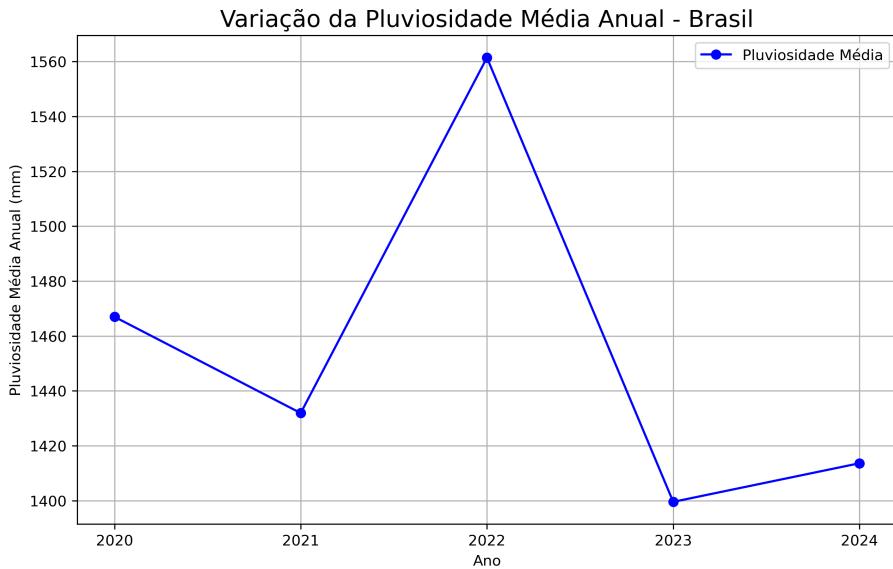
A Figura 17 ilustra a tendência da temperatura média anual para o Brasil. Observa-se uma clara e estatisticamente significativa tendência de resfriamento, com um coeficiente de determinação (R^2) de 0.77.

Figura 17: Série temporal e linha de tendência da Temperatura Média Anual para o Brasil (2019-2024).



Por outro lado, a análise da pluviosidade (Figura 18) exibe uma alta variabilidade interanual, sem uma tendência linear definida. O período é marcado por um pico de precipitação excepcional no ano de 2022, contrastando fortemente com os anos subsequentes, 2023 e 2024, que registraram os menores volumes de chuva da série. Esta ausência de tendência linear sugere que a dinâmica pluviométrica no Brasil, no período analisado, foi dominada por eventos anuais discretos em vez de uma mudança direcional gradual.

Figura 18: Série temporal da Pluviosidade Média Anual para o Brasil (2019-2024).

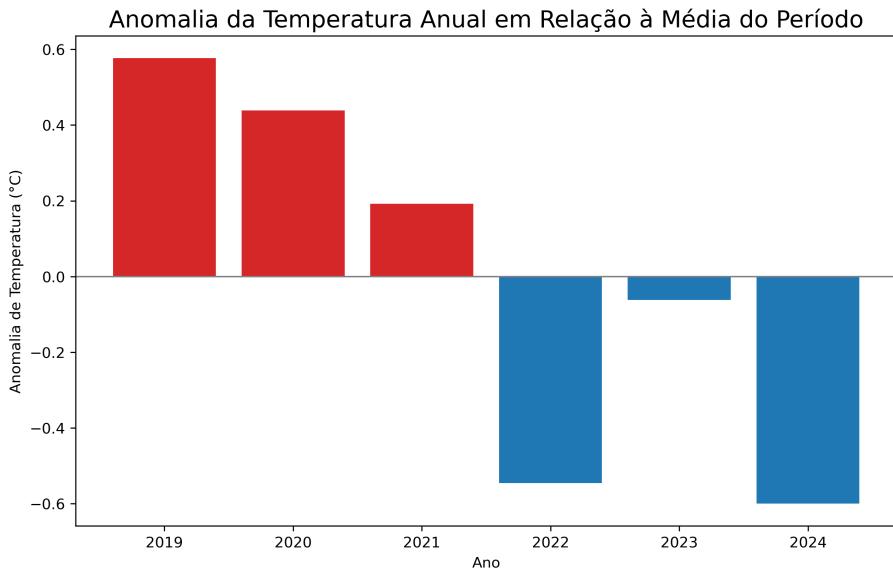


6.2 Análise de Anomalias Climáticas Anuais

A análise de anomalias, que compara cada ano com a média de todo o período, permite identificar os anos de eventos climáticos mais extremos.

O gráfico de anomalias de temperatura (Figura 19) corrobora a análise de tendência, destacando os anos de 2019, 2020 e 2021 como anomalamamente quentes (barras vermelhas). Em contrapartida, 2022 e 2024 foram significativamente mais frios que a média do período, com as maiores anomalias negativas. O ano de 2023 apresentou uma anomalia negativa mais branda.

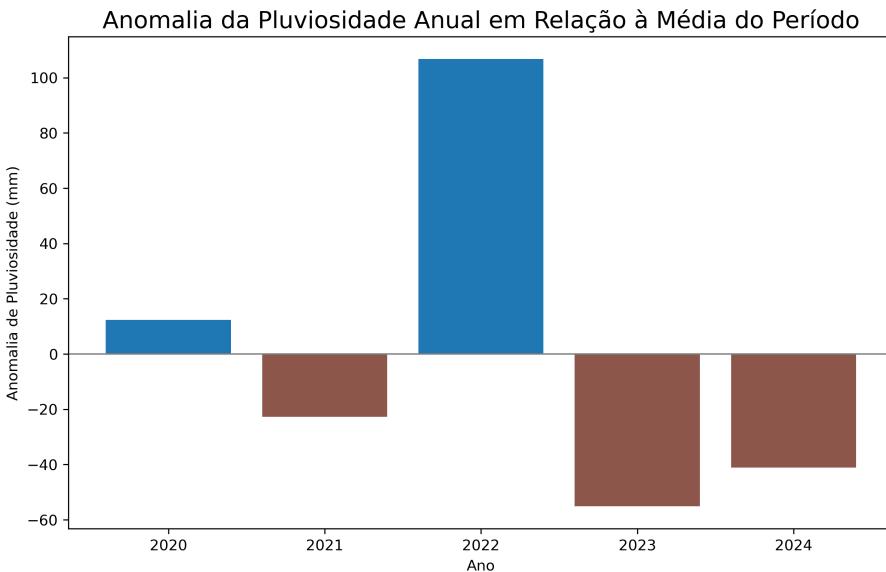
Figura 19: Anomalia da Temperatura Média Anual em relação à média do período 2019-2024.



Para a pluviosidade (Figura 20), o ano de **2022** emerge como um evento extremo, com uma anomalia positiva superior a 100 mm em relação à média, indicando um ano

excepcionalmente úmido em escala nacional. Em oposição, **2023** destaca-se como o ano mais seco da série, com uma anomalia negativa de quase -60 mm, seguido por 2024, também um ano mais seco que a média.

Figura 20: Anomalia da Pluviosidade Média Anual em relação à média do período 2019-2024.



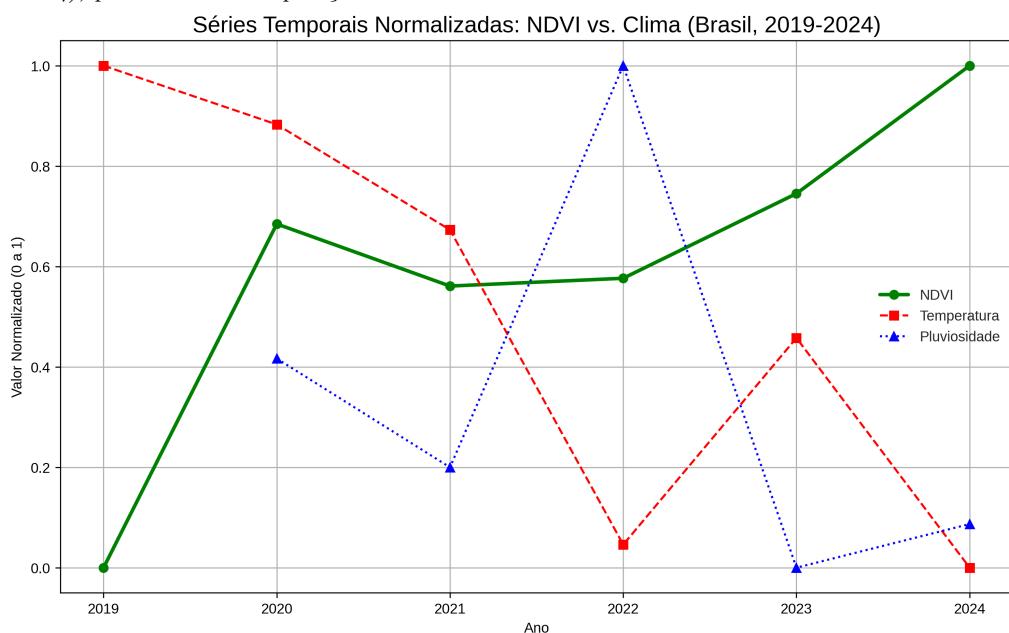
7 Correlação entre NDVI e Dados Climáticos

A etapa final da análise consiste em integrar os resultados da dinâmica da vegetação (NDVI) com as variáveis climáticas de Temperatura da Superfície (LST) e Pluviosidade. O objetivo é quantificar a relação entre esses fatores e inferir como o clima tem modulado a resposta da vegetação no período de 2019 a 2024, tanto em escala nacional quanto regional.

7.1 Análise Integrada em Escala Nacional

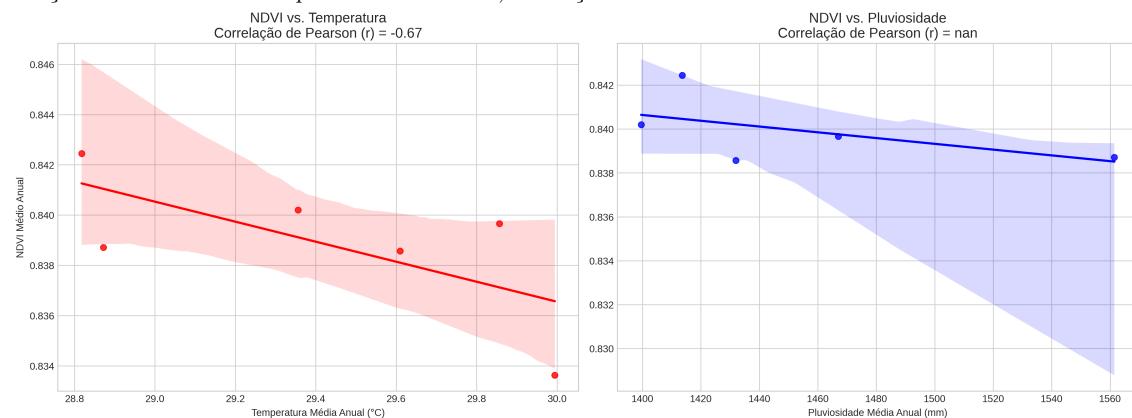
A Figura 21 apresenta as séries temporais normalizadas para o NDVI, temperatura e pluviosidade médios do Brasil. A normalização dos dados para uma escala de 0 a 1 permite a comparação direta da forma e da fase das tendências. Observa-se uma clara relação inversa entre a temperatura (linha vermelha) e o NDVI (linha verde); picos de temperatura, como em 2019, coincidem com vales de NDVI, enquanto o ano mais frio, 2022, corresponde a um platô de alta no vigor da vegetação. A relação com a pluviosidade (linha azul) é menos direta em escala nacional, embora o pico de chuvas em 2022 preceda um período de NDVI elevado e a queda abrupta da pluviosidade em 2023 coincida com a maior anomalia negativa, como visto na figura 20.

Figura 21: Séries temporais normalizadas (0 a 1) para NDVI, Temperatura e Pluviosidade para o Brasil (2019-2024), permitindo a comparação visual das tendências.



Para quantificar essas relações, foram gerados gráficos de dispersão com o cálculo do coeficiente de correlação de Pearson (r), conforme a Figura 22. A análise revela uma correlação negativa de moderada a forte ($r = -0.67$) entre o NDVI e a temperatura. Este resultado é estatisticamente robusto e indica que, em escala nacional, anos mais quentes estão associados a uma menor vitalidade da vegetação. Em contrapartida, a correlação entre o NDVI e a pluviosidade média anual é inconclusiva em escala nacional, com um coeficiente próximo de zero, sugerindo que a relação é mais complexa e provavelmente mascarada por dinâmicas regionais distintas.

Figura 22: Gráficos de dispersão e regressão linear para as variáveis em escala nacional. À esquerda, a relação entre NDVI e Temperatura. À direita, a relação entre NDVI e Pluviosidade.

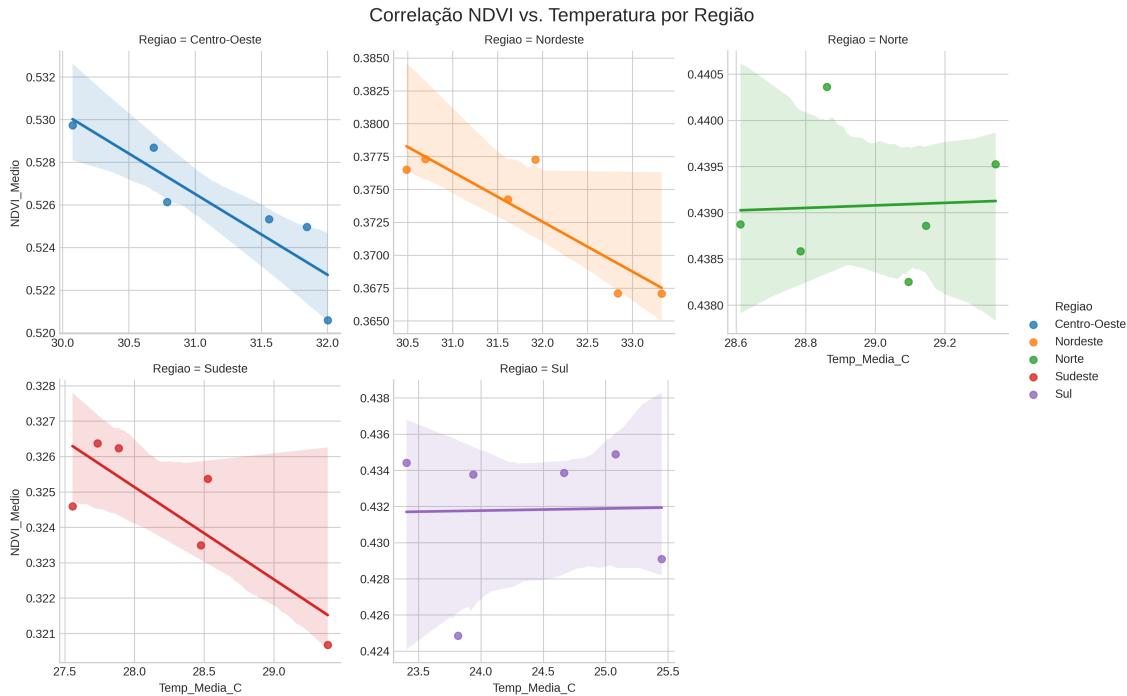


7.2 Dinâmicas de Correlação por Região

A análise desagregada por região confirma que a fraca correlação nacional com a pluviosidade esconde padrões regionais muito distintos e até opostos.

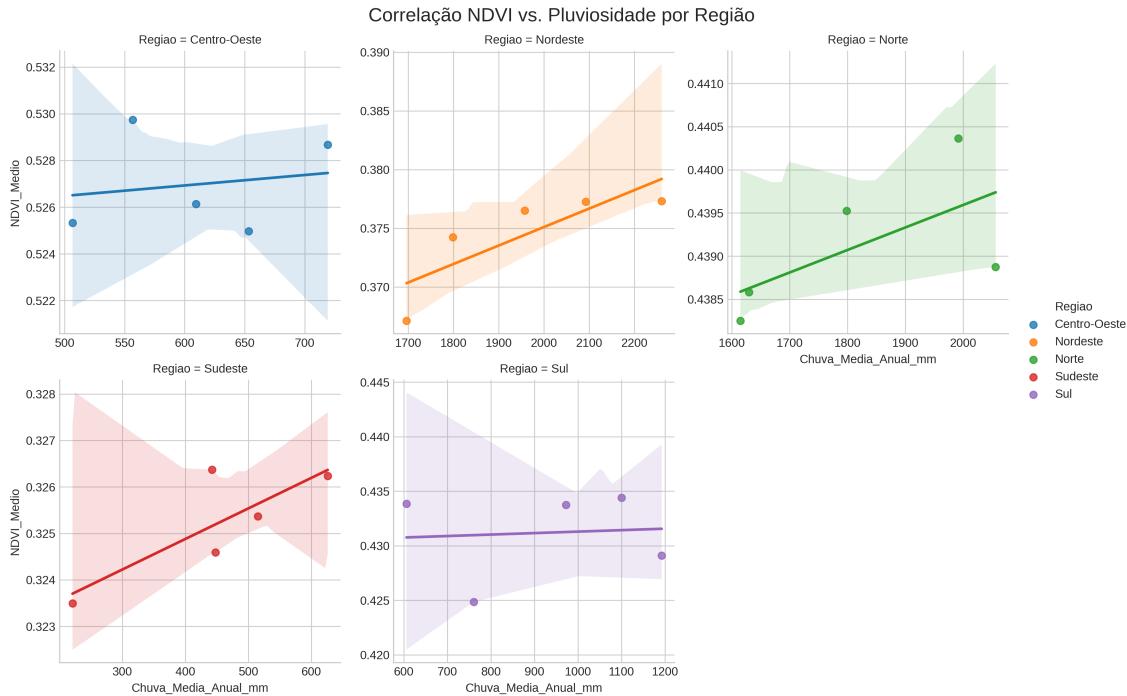
A Figura 23 mostra a correlação entre NDVI e temperatura para cada região. Observa-se uma forte correlação negativa nas regiões **Centro-Oeste**, **Nordeste** e **Sudeste**, indicando que a vegetação nestas áreas é altamente sensível ao estresse térmico. Notavelmente, na **Região Norte**, a correlação é praticamente nula. Isso sugere que, para o ecossistema amazônico, que opera em um regime de alta energia, as flutuações de temperatura observadas no período não foram o principal fator limitante para a vegetação. A **Região Sul** também apresenta uma correlação fraca, indicando que outros fatores podem ter maior influência.

Figura 23: Correlação entre NDVI e Temperatura Média Anual, por região.



A análise da pluviosidade (Figura 24) é ainda mais reveladora. As regiões **Norte**, **Nordeste** e **Sudeste** exibem uma forte correlação positiva, demonstrando que a disponibilidade de água é um vetor crucial para o vigor da vegetação nessas áreas. A correlação é particularmente acentuada no Nordeste, como esperado para um bioma semiárido. Em contraste, as regiões **Centro-Oeste** e **Sul** mostram uma correlação muito fraca ou inexiste, sugerindo que, para a agricultura de alta tecnologia e os ecossistemas locais, o total anual de chuva pode não ser tão determinante quanto a sua distribuição ao longo do ano ou outros fatores.

Figura 24: Correlação entre NDVI e Pluviosidade Média Anual, por região.



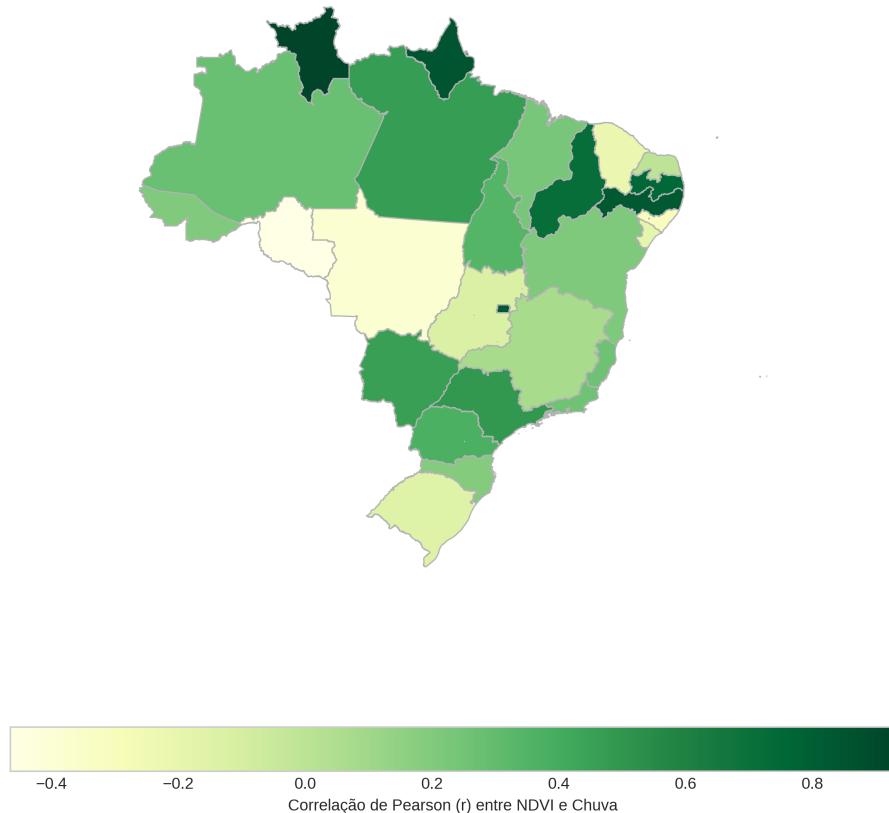
7.3 Análise Espacial da Dependência Pluviométrica por Estado

Para visualizar a dependência da vegetação à chuva de forma espacialmente explícita, foi gerado um mapa coroplético (Figura 25), onde cada estado é colorido pela intensidade de sua correlação NDVI-chuva.

O mapa ilustra de forma clara e intuitiva os padrões regionais. Os estados com os tons de verde mais escuros, como **Piauí (PI)**, **Ceará (CE)**, **Pernambuco (PE)** e outros no Nordeste, são aqueles cuja vegetação é mais fortemente dependente do volume de chuva anual. A mesma forte dependência é vista em estados da Amazônia Oriental e Setentrional, como **Pará (PA)**, **Amapá (AP)** e **Roraima (RR)**. Em contrapartida, os estados da Região Sul (**PR**, **SC**, **RS**) e o estado do **Amazonas (AM)** apresentam as cores mais claras, indicando uma correlação fraca ou negativa. No caso do Sul, isso pode ser explicado por eventos extremos (secas e enchentes) que quebram a relação linear simples. Para o Amazonas, a pluviosidade é tão abundante que variações no total anual não são o fator limitante para uma floresta que já se encontra em seu vigor máximo, e o excesso de nuvens pode até mesmo reduzir a radiação solar disponível.

Figura 25: Mapa de correlação de Pearson (r) entre NDVI e pluviosidade anual para cada estado. Tons de verde mais escuros indicam maior dependência da vegetação em relação à chuva.

Dependência da Vegetação à Chuva por Estado (2019-2024)



8 Conclusão

Este trabalho realizou com sucesso uma análise integrada e multi-variada da dinâmica ambiental do Brasil para o período de 2019 a 2024. Foi desenvolvido e implementado um *framework* computacional robusto que automatiza a aquisição, o pré-processamento e a análise de grandes volumes de dados de sensoriamento remoto, incluindo o Índice de Vegetação por Diferença Normalizada (NDVI), a Temperatura da Superfície Terrestre (LST) e a Precipitação Pluviométrica. A metodologia incluiu a análise estatística tradicional, incorporando técnicas de aprendizado de máquina e redes neurais profundas para uma caracterização mais sofisticada dos padrões de mudança.

As principais conclusões apontam para um cenário complexo, onde tendências nacionais mascaram realidades regionais distintas e fortemente acopladas às variáveis climáticas. Em escala nacional, foi identificada uma tendência estatisticamente significativa de “esverdeamento” (aumento do NDVI com $R^2=0.70$), que se mostrou fortemente correlacionada com uma tendência de resfriamento da superfície (queda da LST, com $R^2=0.77$ e correlação de Pearson $r=-0.67$). A pluviosidade, por sua vez, não apresentou uma tendência linear clara em nível nacional, sendo sua influência mais bem compreendida através da análise regional.

A análise estratificada revelou a heterogeneidade desses processos. Regiões como o Nordeste e o Sudeste demonstraram uma forte correlação positiva entre a saúde da vegetação e a chuva, enquanto a Região Norte, apesar da alta disponibilidade hídrica, apresentou tendências de degradação em estados do “arco do desmatamento”. A aplicação de clusterização por K-Means conseguiu segmentar o território em zonas climaticamente consistentes, alinhadas aos grandes biomas. Adicionalmente, a implementação de um autoencoder ConvLSTM validou sua capacidade de detectar anomalias espaço-temporais, identificando hotspots de mudança abrupta, especialmente em áreas de conhecida pressão antrópica.

9 Limitações e Recomendações para Trabalhos Futuros

Apesar da abrangência do *framework* desenvolvido, existem limitações inerentes que abrem novos caminhos para futuras investigações. As seguintes direções são propostas para aprofundar e refinar os resultados obtidos:

- **Análise em Alta Resolução Espacial:** A resolução de 1 km dos dados utilizados é ideal para tendências macro, mas não captura mudanças de uso do solo em escala fina. Recomenda-se o uso dos mapas de anomalia gerados como um guia para selecionar *hotspots*. Nesses locais, uma análise de *zoom-in* com dados de maior resolução, como os dos satélites da série **Landsat (30 m)** [26] ou **Sentinel-2 (10-20 m)** [3], permitiria uma caracterização precisa dos vetores de mudança (e.g., desmatamento seletivo, cicatrizes de incêndio).
- **Validação e Classificação de Anomalias com Modelos Supervisionados:** O modelo ConvLSTM operou de forma não supervisionada para detectar *onde* ocorreram mudanças. Um avanço significativo seria a criação de um dataset de treinamento rotulado, utilizando polígonos de desmatamento e regeneração de fontes como o projeto **PRODES/DETER do INPE** [12]. Com esses dados, seria possível treinar uma rede de segmentação semântica (e.g., U-Net) para não apenas detectar a mudança, mas também **classificá-la automaticamente** (ex: “desmatamento”, “incêndio”, “regeneração”).
- **Uso de Índices de Vegetação Complementares:** O NDVI é suscetível à saturação em áreas de alta biomassa. Uma investigação futura valiosa seria refazer as análises utilizando o **Índice de Vegetação Melhorado (EVI - Enhanced Vegetation Index)** [7], que também está disponível no produto MODIS. O EVI foi projetado para ser mais sensível a variações em florestas densas e pode revelar tendências de degradação ou crescimento sutil que o NDVI não captura.
- **Análise de Causalidade e Lapsos Temporais (Time Lag):** Este trabalho estabeleceu correlações entre NDVI e clima. O próximo passo seria aplicar métodos estatísticos mais avançados para inferência de causalidade, como a Causalidade de Granger ou o Mapeamento Cruzado Convergente, para determinar mais formalmente a direção da influência. Adicionalmente, modelar o lapso de tempo (*time lag*) específico da resposta do NDVI à precipitação em cada bioma traria um refinamento crucial à análise.
- **Fusão de Múltiplos Dados (Data Fusion):** A análise pode ser ainda mais enriquecida pela fusão com outros tipos de dados de satélite. Recomenda-se a incorporação de dados de focos de calor, como os do sensor **VIIRS** [15], e de umidade do solo, como os da missão **SMAP** [16], para criar um modelo preditivo e diagnóstico ainda mais completo e preciso sobre a saúde e a resiliência dos ecossistemas brasileiros.

Referências

- [1] Agência Brasil. Crise hídrica de 2021 é a pior dos últimos 91 anos, sep 2021. Acessado em 24 de Junho de 2025.
- [2] Agência Brasil. Seca na amazônia em 2023 foi a pior em 120 anos, mostra estudo, jan 2024. Acessado em 24 de Junho de 2025.
- [3] European Space Agency (ESA). Sentinel-2 mission, 2024. Acessado em 24 de Junho de 2025.
- [4] G1. Queimadas no brasil aumentam 82% em relação ao mesmo período de 2018, ago 2019.
- [5] G1 RS. Estiagem no RS: Com 418 cidades em situação de emergência, perdas no campo chegam a R\$ 36 bilhões, mar 2022. Acessado em 24 de Junho de 2025.
- [6] Sean Gillies. *Rasterio: Geospatial data I/O for Python programmers*, 2013.
- [7] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2):195–213, 2002.
- [8] Alfredo R. Huete. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3):295–309, 1988.
- [9] Alfredo R. Huete, Kamel Didan, Tomoaki Miura, E. P. Rodriguez, Xiang Gao, and Laerte G. Ferreira. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2):195–213, 2002.
- [10] G. J. Huffman, E. F. Stocker, D. T. Bolvin, E. J. Nelkin, and J. Tan. GPM IMERG final precipitation L3 1 month 0.1 degree x 0.1 degree V07. Technical report, NASA Goddard Earth Sciences Data and Information Services Center (GES DISC), 2020.
- [11] Instituto Brasileiro de Geografia e Estatística (IBGE). Malhas digitais - malha de unidades da federação 2022. <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/15774-malhas.html>, 2022. Acessado em 24 de Junho de 2025.
- [12] Instituto Nacional de Pesquisas Espaciais (INPE). Projeto PRODES - monitoramento do desmatamento da floresta amazônica brasileira por satélite. <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>, 2024. Acessado em 24 de Junho de 2025.
- [13] Kelsey Jordahl, Joris van den Bossche, Martin Fleischmann, et al. geopandas: Python tools for geographic data. 2014.
- [14] MetSul Meteorologia. O ano de extremos no sul com ciclones e enchentes sob forte El Niño, dez 2023.
- [15] NASA. Fire information for resource management system (FIRMS), 2024.
- [16] NASA Jet Propulsion Laboratory. SMAP - soil moisture active passive, 2024.

- [17] Nathalie Pettorelli, Jon Olav Vik, Atle Mysterud, Jean-Michel Gaillard, Compton J. Tucker, and Nils Chr. Stenseth. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution*, 20(9):503–510, 2005.
- [18] Python Software Foundation. *Python Language Reference, version 3.11*, 2023.
- [19] QGIS Development Team. *QGIS Geographic Information System*. QGIS Association, 2024.
- [20] John W Rouse, Robert H Haas, J. A. Schell, and D. W. Deering. Monitoring vegetation systems in the great plains with ERTS. *NASA Special Publication*, 351:309, 1974.
- [21] Inge Sandholt, Kjeld Rasmussen, and Jens Andersen. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of Environment*, 79(2–3):213–224, 2002.
- [22] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 802–810. Curran Associates, Inc., 2015.
- [23] Y. E. Shimabukuro, E. Arai, A. C. Dutra, V. Duarte, L. O. Anderson, J. R. dos Santos, A. P. D. de Aguiar, and L. E. O. C. Aragão. Drought and fire impacts on the amazon forest in the state of acre, brazil. *Acta Amazonica*, 49(4):336–345, 2019.
- [24] Terry A. Slocum, Robert B. McMaster, Fritz C. Kessler, and Hugh H. Howard. *Thematic Cartography and Geovisualization*. Pearson Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2009.
- [25] Compton J. Tucker. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2):127–150, 1979.
- [26] U.S. Geological Survey. Landsat missions, 2024. Acessado em 24 de Junho de 2025.
- [27] Jan Verbesselt, Rob Hyndman, Glenn Newnham, and Darius Culvenor. Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, 114(1):106–115, 2010.
- [28] Zhengming Wan. New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sensing of Environment*, 140:1–13, 2013.
- [29] J. Wang, P. M. Rich, and K. P. Price. Temporal responses of NDVI to precipitation and temperature in the central great plains, USA. *International Journal of Remote Sensing*, 24(11):2345–2364, 2003.
- [30] World Meteorological Organization (WMO). WMO update: La Niña persists and is likely to continue, jun 2022. Reportando a persistência do La Niña e seus impactos, incluindo chuvas acima da média no norte da América do Sul. Acessado em 24 de Junho de 2025.
- [31] WWF-Brasil. Arco do desmatamento: o que é e quais as principais ameaças, 2019.