



**UNIVERSIDADE FEDERAL DE ITAJUBÁ
INSTITUTO DE RECURSOS NATURAIS
PROGRAMA DE GRADUAÇÃO EM CIÊNCIAS ATMOSFÉRICAS**

**USO DE INTELIGÊNCIA ARTIFICIAL PARA
PREVISÃO DE PRECIPITAÇÃO SAZONAL NO
ESTADO DE MINAS GERAIS, BRASIL**

MONOGRAFIA DE GRADUAÇÃO

Matheus José Gomes

**Itajubá, MG, Brasil
2025**

USO DE INTELIGÊNCIA ARTIFICIAL PARA PREVISÃO DE PRECIPITAÇÃO SAZONAL NO ESTADO DE MINAS GERAIS, BRASIL

por

Matheus José Gomes

Monografia apresentada à comissão examinadora Programa de Graduação em Ciências Atmosféricas da Universidade Federal Itajubá (UNIFEI, MG), como requisito parcial para obtenção do grau de
Bacharel em Ciências Atmosféricas.

Orientadora: Maria Alice Machado Rodrigues
Coorientadora: Juliana Aparecida Anochi

**Itajubá, MG, Brasil
2025**

**Universidade Federal de Itajubá
Instituto de Recursos Naturais
Programa de Graduação em Ciências Atmosféricas**

A Comissão Examinadora, abaixo assinada, aprova a Monografia

**USO DE INTELIGÊNCIA ARTIFICIAL PARA PREVISÃO DE
PRECIPITAÇÃO SAZONAL NO ESTADO DE MINAS GERAIS,
BRASIL**

elaborada por
Matheus José Gomes

Como requisito parcial para a obtenção do grau de
Bacharel em Ciências Atmosféricas

Comissão Examinadora:

Maria Alice Machado Rodrigues, Ma. (UNIFEI)
(Presidente/Orientador)

Juliana Aparecida Anochi, Dra. (INPE)
(Coorientador)

Aline Araújo de Freitas, Ma. (UNIFEI)

Marilia Harumi Shimizu, Dra. (INPE)

Itajubá, 14 de novembro de 2025.

AGRADECIMENTOS

Agradeço à Universidade Federal de Itajubá (UNIFEI), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e ao Instituto Nacional de Pesquisas Espaciais (INPE), instituições que contribuíram de maneira fundamental para minha formação acadêmica e para o desenvolvimento deste trabalho.

Aos professores e pesquisadores que fizeram parte da minha trajetória, em especial à minha orientadora de Trabalho de Conclusão de Curso, Ma. Maria Alice Machado Rodrigues, e às minhas orientadoras de Iniciação Científica, Dras. Juliana Aparecida Anochi e Marília Harumi Shimizu, agradeço pela dedicação, paciência e por compartilharem conhecimentos que foram essenciais para meu crescimento acadêmico e profissional.

Aos colegas que fiz durante o curso, pela parceria, companheirismo e momentos compartilhados que tornaram essa jornada mais leve e enriquecedora. Aos meus amigos de república, com quem vivi momentos inesquecíveis que levarei para toda a vida.

E, acima de tudo, agradeço aos meus pais, pelo amor incondicional, pelo exemplo de dedicação e por nunca medirem esforços para que eu pudesse chegar até aqui. À minha namorada, pelo carinho, compreensão e apoio constante, que me deram forças para enfrentar os desafios e motivação para seguir em frente. Sem vocês, nada disso teria sido possível.

Dedico este trabalho à minha família, pois sou a continuação dos sonhos deles e fruto de todos que vieram antes de mim.

Quem é mais potente na vida social, na vida política, na vida econômica, sempre será menos afetado pela despotencialização da natureza.

(Gilberto Gil)

RESUMO

Monografia de Graduação
Programa de Graduação em Ciências Atmosféricas
Universidade Federal de Itajubá, MG, Brasil

USO DE INTELIGÊNCIA ARTIFICIAL PARA PREVISÃO DE PRECIPITAÇÃO SAZONAL NO ESTADO DE MINAS GERAIS, BRASIL

AUTOR(A): MATHEUS JOSÉ GOMES
ORIENTADOR(A): MARIA ALICE MACHADO RODRIGUES
Local e Data da Defesa: Itajubá, 14 de novembro de 2025.

Este trabalho investiga a aplicação de técnicas de Inteligência Artificial, com foco em redes neurais artificiais, na previsão da precipitação sazonal no estado de Minas Gerais, região marcada por forte variabilidade climática e suscetibilidade a eventos extremos como secas e chuvas intensas. A metodologia compreende a seleção de variáveis meteorológicas e de superfície, obtidas a partir da reanálise ERA5, do GPCP (*Global Precipitation Climatology Project*) v3.3, para servir como dados de entrada para o treinamento da rede neural. O processo incluiu a normalização dos dados, sua divisão em conjuntos de treino e teste, e uma otimização sistemática dos hiperparâmetros do modelo por meio de *Grid Search* para definir a arquitetura final. A performance foi avaliada por meio do raiz do erro quadrático médio (RMSE), erro quadrático médio (MSE), e erro relativo (RE) para verificar a capacidade da rede em representar padrões de variabilidade climática e em gerar previsões com acurácia satisfatória. Os resultados demonstraram que o modelo de IA superou a performance do NMME (*North American Multi-Model Ensemble*) nas estações de verão, outono e inverno, apresentando menor RMSE. Contudo, revelou limitações importantes, como uma baixa performance na primavera e a tendência a subestimar a precipitação nos períodos mais secos. A análise final sugere o alto potencial da abordagem, ao mesmo tempo que aponta para a complementaridade de erros entre o modelo de IA e o NMME, abrindo caminhos para futuros sistemas de previsão híbridos.

Palavras-chave: Inteligência Artificial. Previsão Sazonal. Redes Neurais Artificiais. Variabilidade Climática. Minas Gerais.

LISTA DE FIGURAS

FIGURA 1 - Topografia do estado de Minas Gerais.....	6
FIGURA 2 - Mesorregiões de Minas Gerais	8
FIGURA 3 - Etapas do desenvolvimento do modelo de previsão sazonal.....	10
FIGURA 4 - Modelo de Rede neural	12
FIGURA 5 - Climatologia sazonal em Minas Gerais (1992-2022)	15
FIGURA 6 – Anomalia de precipitação.	17
FIGURA 7 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas no Verão entre 1984 a 2023.....	20
FIGURA 8 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas no Outono entre 1984 a 2023.....	22
FIGURA 9 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas no Inverno entre 1984 a 2023.....	24
FIGURA 10 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas na Primavera entre 1984 a 2023.....	25
FIGURA 11 - Comparação entre o dado observado, a previsão, e o NMME para 2022.....	28

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

AIRS - *Atmospheric Infrared Sounder*

avg_tnlwrf – *Mean top net long-wave radiation flux*

avg_tnswrf – *Mean top net short-wave radiation flux*

CDS - *Climate Data Store*

cm – Centímetros

CNN - *Convolutional Neural Networks*

CORRA - *Combined Radar-Radiometer Analysis*

DJF – Dezembro, Janeiro e Fevereiro

ECMWF - *European Centre for Medium-Range Weather Forecasts*

ENSO - El Niño-Oscilação Sul

ERA5 - *ECMWF Reanalysis, 5th Generation*

ETOPO - *The Earth Topography*

GPCP - *Global Precipitation Climatology Project*

GPM - *Global Precipitation Measurement*

GPU - *Graphics Processing Unit*

IA - Inteligência artificial

IBGE - Instituto Brasileiro de Geografia e Estatística

IMERG - *Integrated Multi-satellitE Retrievals for GPM*

INPE - Instituto Nacional de Pesquisas Espaciais

JJA – Junho, Julho e Agosto

K – Kelvin

LSTM - *Long Short-Term Memory*

m – Metros

MAM – Março, Abril e Maio

mm dia⁻¹- Milímetro por dia

m³ m⁻³ – Metros cúbicos por metro cúbico

m s⁻¹ - Metros por segundo

MCGAs - Modelos de circulação geral da atmosfera

MONAN - *Model for Ocean-Land-Atmosphere Prediction*

MSE - *Mean squared error*

NMME - *North American Multi-Model Ensemble*

Pa - Pascal

r - Coeficiente de correlação de Pearson

RE – *Relative error*

ReLU- *Rectified Linear Unit*

RMSE - *Root mean square error*

SG - *Surface Gauge*

SON – Setembro, Outubro e Novembro

swvl1 – *Volumetric soil water layer 1 (0 - 7 cm)*

swvl2 – *Volumetric soil water layer 2 (7 - 28 cm)*

swvl3 – *Volumetric soil water layer 3 (28 - 100 cm)*

swvl4 - *Volumetric soil water layer 4 (100 - 289 cm)*

TI - Tecnologia da Informação

TOVS - *Tiros Operational Vertical Sounder*

TPU - *Tensor Processing Unit*

t2m – Temperatura a 2 metros

u - Componentes zonais do vento

u10 – Componentes zonais do vento a 10 metros

u100 – Componentes zonais do vento a 100 metros

v - Componentes meridionais do vento

v10 – Componentes meridionais do vento a 10 metros

v100 – Componentes meridionais do vento a 100 metros

V3.2 -Versão 3.2

V3.3 - Versão 3.3

W m⁻² - Watts por metro quadrado

ZCAS - Zona de Convergência do Atlântico Sul

SUMÁRIO

RESUMO.....	vi
LISTA DE FIGURAS.....	vii
LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS.....	vii
1. INTRODUÇÃO	1
1.1.OBJETIVO	5
1.1.1 Objetivo Geral	5
1.1.2 Objetivos Específicos	5
2. MATERIAL E MÉTODOS	6
2.1. DESCRIÇÃO DA ÁREA DE ESTUDO	6
2.2. DADOS.....	8
2.2.1 Precipitação.....	8
2.2.2 Variáveis Atmosféricas e de volume de água no solo	9
2.3. METODOLOGIA	9
2.3.1. Preparação dos dados e correlação.....	10
2.3.2. Treinamento e Validação dos Modelos de IA	13
2.3.3. Avaliação das Previsões pelos Modelos de IA	16
3. RESULTADOS E DISCUSSÃO	18
3.1 ANÁLISE CLIMATOLÓGICA E DEFINIÇÃO DAS VARIÁVEIS PREDITORAS.....	18
3.2. CONFIGURAÇÃO E OTIMIZAÇÃO E DA IA.....	28
3.3. AVALIAÇÃO DA PREVISÃO POR IA.....	28
4. CONCLUSÃO	31
5. REFERÊNCIAS	33

1. INTRODUÇÃO

Os modelos meteorológicos numéricos, assim como as previsões de tempo e clima evoluíram significativamente ao longo dos anos, impulsionados pelos avanços nas tecnologias computacionais, como, por exemplo, a inteligência artificial. Esses modelos são importantes para fornecer previsões que auxiliam na prevenção de eventos climáticos extremos, no planejamento agrícola, na gestão de recursos hídricos e em diversas outras aplicações (HOLTON, 2013). Entre os principais avanços científicos do último século, destaca-se a capacidade de simular sistemas físicos e complexos por meio de modelos numéricos, que representam matematicamente a atmosfera real (SAMPAIO, 2014).

Apesar dos avanços dos modelos, a precipitação é uma variável difícil de simular, mas que atualmente existem produtos de precipitação renomados como o GPCP (Global Precipitation Climatology Project) (HUFFMAN *et al.* 2025c). Em sua versão mais atualizada, a V3.3, o método de elaboração consiste em integrar e calibrar diversas fontes de estimativas de precipitação, principalmente dados de satélite, com destaque para o produto IMERG (*Integrated Multi-satellite Retrievals for GPM*) da missão GPM (*Global Precipitation Measurement*). Essas estimativas são ajustadas com dados mensais do GPCP V3.2 SG (*Surface Gauge*), assegurando coerência global, especialmente em regiões de alta latitude com menor cobertura de sensores como o CORRA (*Combined Radar-Radiometer Analysis*), TOVS (*Tiros Operational Vertical Sounder*) e AIRS (*Atmospheric Infrared Sounder*). O processo envolve calibração histórica, interpolação e correções proporcionais das versões anteriores, buscando gerar um campo diário de precipitação global e homogêneo (HUFFMAN *et al.*, 2025b).

O aprimoramento do conhecimento sobre os processos que controlam a evolução do sistema climático, aliado ao desenvolvimento de ferramentas observacionais, métodos de solução numérica, e o aumento da capacidade computacional, tem elevado significativamente a qualidade das previsões (LYNCH, 2007). Os modelos de circulação geral da atmosfera (MCGAs), por exemplo, são capazes de prever as condições do tempo com alto grau de confiabilidade por até 7 a 12 dias, dependendo da região e das condições iniciais (ZHANG, 2019). Além disso, tais modelos têm sido fundamentais para previsões climáticas sazonais e de longo prazo, considerando fatores como o aumento de gases de efeito estufa, aerossóis e

alterações no uso da terra, como o desmatamento e a urbanização (MELO; MARENGO, 2008).

Apesar desses avanços, a crescente complexidade dos modelos e a demanda por maior precisão impõem desafios computacionais relevantes, impulsionando a busca por abordagens alternativas, como o uso de técnicas de inteligência artificial na modelagem e previsão climática (ANOCHI, 2015). Nesse contexto, o aprendizado de máquina oferece um vasto campo de aplicações e novas perspectivas para aprimorar as análises climáticas e a previsão do tempo, buscando superar as limitações computacionais e de precisão dos modelos convencionais (BOCHENEK; USTRNUL, 2022).

As abordagens tradicionais de previsão climática, baseadas em modelos dinâmicos, utilizam equações físicas para descrever os processos atmosféricos (LAM, 2023). Embora esses modelos determinísticos tenham avançado ao longo do tempo, ainda apresentam limitações em termos de resolução espacial e temporal, e incertezas nos processos físicos parametrizados (VOLODINA; CHALLENGOR, 2021), além de exigirem alta capacidade computacional para processar grandes volumes de dados.

Nesse contexto, a inteligência artificial (IA) surge como uma alternativa promissora para melhorar a qualidade das previsões climáticas, ao ser incorporado em diferentes etapas de processamento das previsões numéricas (CINTRA; DE CAMPOS VELHO, 2018). Além disso, modelos de IA, como as redes neurais, têm mostrado que são capazes de lidar com grandes conjuntos de dados, identificar padrões complexos e gerar previsões empíricas de forma mais rápida (ANOCHI *et al.*, 2021).

No campo da inteligência artificial, que busca desenvolver sistemas capazes de realizar tarefas complexas, as Redes Neurais Artificiais destacam-se como uma abordagem computacional inspirada pela forma como o cérebro humano processa informações. Essas redes são construídas a partir de numerosas unidades de processamento simples, ou neurônios, que são densamente interconectadas e funcionam de maneira distribuída e paralela. Através de um processo de aprendizado adaptativo, no qual as conexões entre esses neurônios são ajustadas com base em dados de exemplo, as Redes Neurais Artificiais adquirem a capacidade de identificar padrões, classificar informações e realizar previsões, tornando-se ferramentas valiosas em uma vasta gama de aplicações (FURTADO, 2019).

Instituições meteorológicas renomadas, como o Centro Europeu de Previsão do Tempo de Médio Prazo (ECMWF) e o *Met Office* do Reino Unido, têm adotado técnicas de IA para melhorar a precisão das previsões climáticas, apresentando resultados encorajadores no tratamento de dados não lineares e no aumento da acurácia das previsões (CONTI, 2024). Já no Brasil, temos iniciativas como o projeto MONAN (*Model for Ocean-Land-Atmosphere Prediction*) do Instituto Nacional de Pesquisas Espaciais (INPE), que representa um esforço para desenvolver um modelo unificado e de estado-da-arte do Sistema Terrestre, ajustado às condições da América do Sul e focado no avanço da qualidade e confiabilidade das previsões de tempo, clima e ambiente. De forma crucial, o projeto MONAN estabelece como um de seus requisitos o aperfeiçoamento de suas previsões e rotinas de produção através da utilização de técnicas de inteligência artificial. Este novo sistema de modelagem visa cobrir um amplo espectro de escalas, desde previsões de curtíssimo prazo até projeções de mudanças climáticas de longo prazo (FREITAS, 2023).

Segundo Goodfellow *et al.* (2016), uma rede neural processa informações através de camadas de neurônios interconectados por pesos: inicialmente, as variáveis de entrada são inseridas na primeira camada; em seguida, o modelo é treinado, onde os pesos das conexões são ajustados iterativamente com base em dados de exemplo para que a rede aprenda a produzir as saídas desejadas. Por fim, o modelo é validado usando um conjunto de dados diferente (não visto no treinamento) para verificar sua precisão e capacidade de generalização a novas informações.

Entre as variáveis climáticas comumente empregadas como entradas em redes neurais destacam-se a temperatura do ar a 2m, a pressão ao nível da superfície e as componentes zonal (u) e meridional (v) do vento, conforme demonstrado por Monego, Anochi e De Campos Velho (2022). Ademais, Araújo e Souza (2010) evidenciaram a relevância da radiação solar para aperfeiçoar o desempenho dos modelos climáticos. Trabalhos mais recentes, como Zhu *et al.* (2021), também têm incorporado o volume de água no solo em múltiplas camadas, estabelecendo correlações entre esse parâmetro e o aumento de precipitação em determinadas regiões.

Uma das ferramentas gratuitas comumente utilizadas para facilitar a criação de sistemas de inteligência artificial é o TensorFlow, o qual foi desenvolvido pelo Google em 2015. Ele organiza operações complexas em forma de gráficos, permitindo que modelos de aprendizado sejam executados de forma eficiente em diferentes dispositivos, como GPUs (*Graphics Processing Unit*) e TPUs (*Tensor Processing*

Unit). Bastante utilizado com a linguagem *Python*, o *TensorFlow* é compatível com outras linguagens e se destaca por sua flexibilidade, sendo usado tanto em projetos acadêmicos quanto comerciais. A plataforma é aplicada em tarefas como reconhecimento de imagem, voz e análise de texto, e integra-se facilmente com os serviços de nuvem do Google, o que facilita o desenvolvimento e a distribuição de modelos (GOLDBROOUGH, 2019).

O objetivo deste trabalho, portanto, foi desenvolver um modelo aplicando uma rede neural de baixo custo computacional capaz de realizar previsões sazonais de precipitação para o estado de Minas Gerais, onde as variáveis de entrada são parâmetros atmosféricos observados. Para isso, o estudo inicia-se pela apresentação dos dados empregados. Em seguida, a metodologia foi dividida em três etapas: (1) análise climatológica, consistindo na preparação dos dados observados através de cálculos de médias espaciais, detecção de anomalias e análises de correlação; (2) treinamento e validação dos modelos de IA, e (3) avaliação das previsões a partir de métricas de desempenho (viés, RMSE e ER) com representação espacial dos resultados. Por fim, são apresentados os resultados obtidos em termos de acurácia e tempos de execução do modelo desenvolvido, seguido das conclusões, das contribuições do estudo, das limitações identificadas e das recomendações para pesquisas futuras.

1.1. Objetivo

1.1.1 Objetivo Geral

O objetivo geral deste trabalho foi desenvolver um modelo de inteligência artificial, baseado em redes neurais, para a previsão de precipitação sazonal no estado de Minas Gerais, utilizando parâmetros atmosféricos observados como variáveis de entrada.

1.1.2 Objetivos Específicos

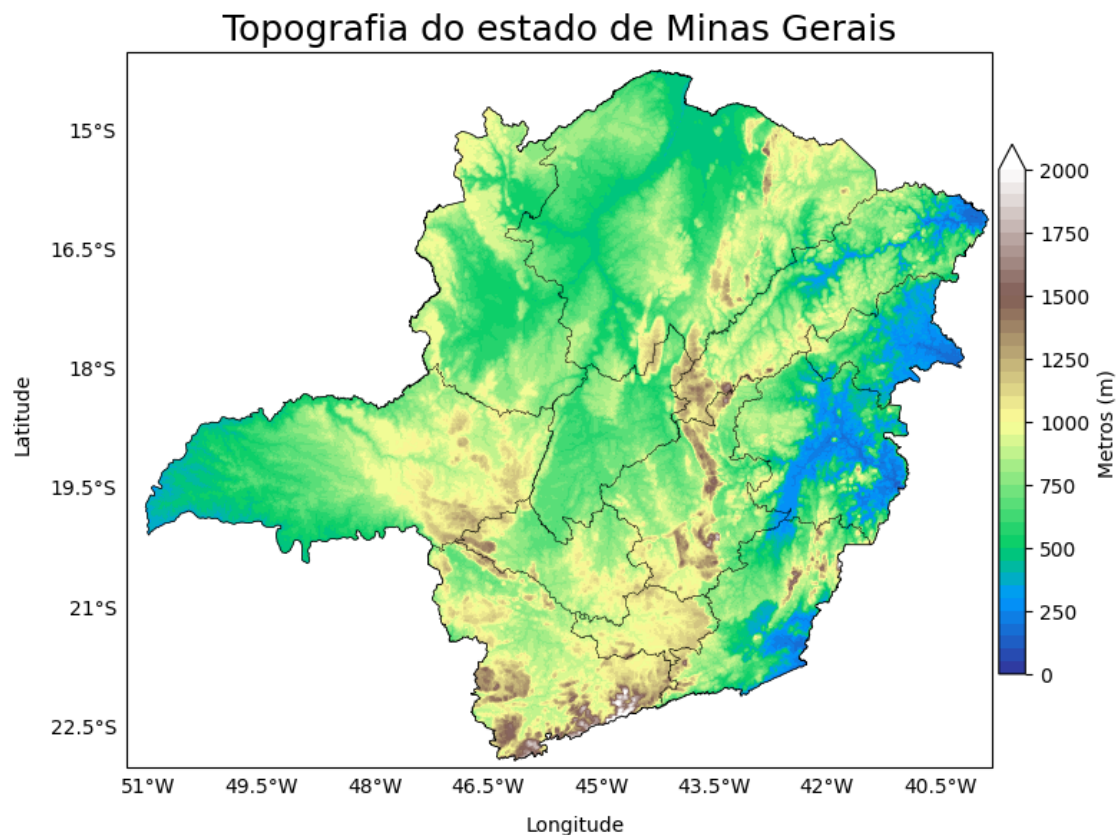
- Identificar as características climatológicas das variáveis observadas.
- Identificar as variáveis preditoras inseridas no modelo.
- Estruturar e treinar modelo de rede neural artificial.
- Avaliar a acurácia e o desempenho do modelo de inteligência artificial para a previsão da precipitação sazonal.
- Analisar a capacidade do modelo de rede neural otimizado em representar os padrões de variabilidade da precipitação sazonal e sua eficiência computacional.

2. MATERIAL E MÉTODOS

2.1. Descrição da área de estudo

O estado de Minas Gerais, localizado na região sudeste do Brasil, é conhecido por sua vasta extensão territorial e pela ampla diversidade de paisagens e climas. Com uma área de aproximadamente 586.521 km², Minas Gerais faz fronteira com os estados da Bahia, Goiás, Mato Grosso do Sul, Espírito Santo, São Paulo e Rio de Janeiro (IBGE, 2022). Essa diversidade geográfica resulta em variações significativas de altitude, que vão desde áreas mais baixas no nordeste e oeste do estado até regiões montanhosas ao sul, como a Serra da Mantiqueira, onde os picos podem atingir cerca de 2.800 metros de altitude (Figura 1).

Figura 1- Topografia do estado de Minas Gerais.



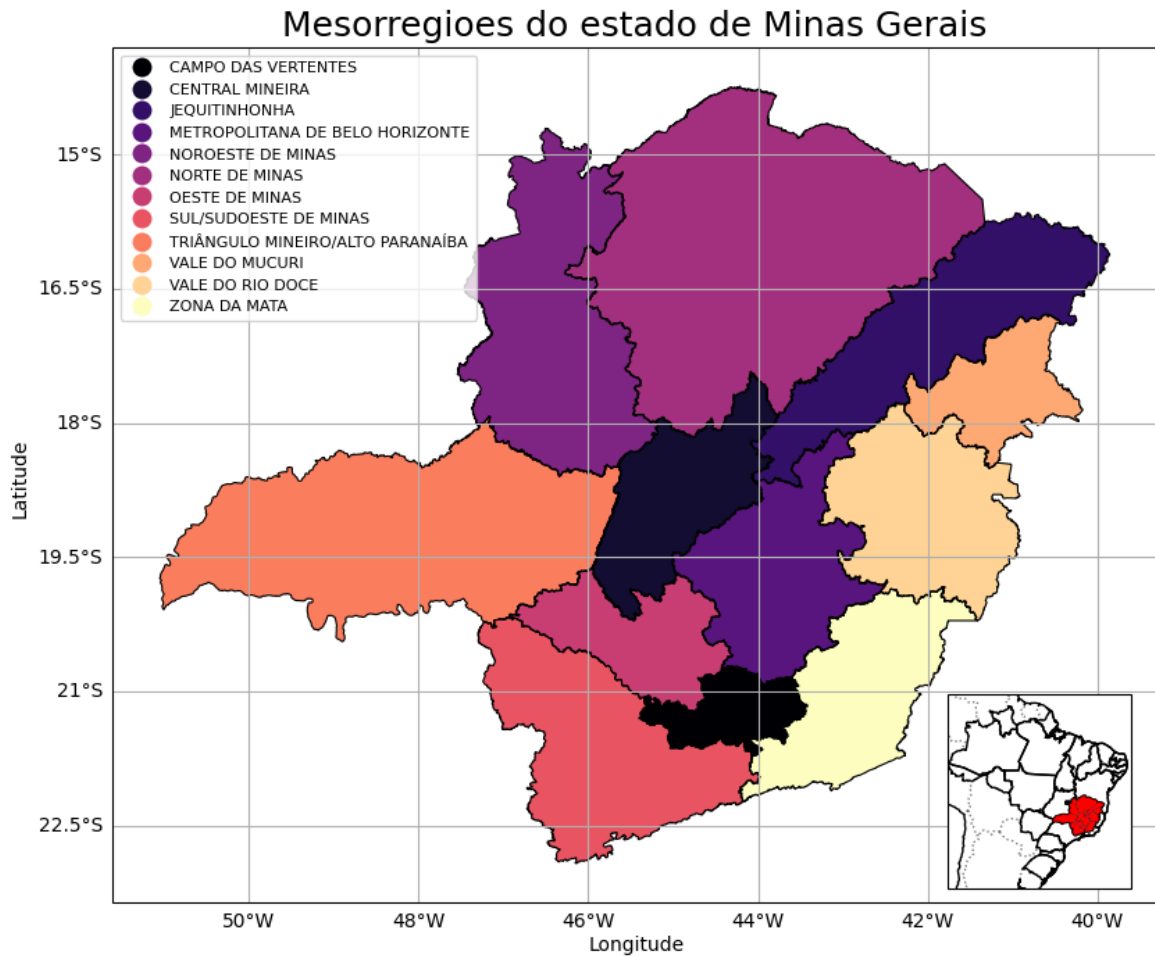
Topografia do estado de Minas Gerais em metros com base nos dados topográficos do ETOPO 2022 (MACFERRIN et al., 2022), com resolução espacial de 0,0167° × 0,0167°. Fonte: Elaboração própria.

O estado de Minas Gerais possui um clima predominantemente Tropical de Altitude e Tropical, com verões chuvosos e invernos secos (REBOITA et al., 2010; GRIMM, 2019; RODRIGUES et al., 2022). A topografia montanhosa desempenha um papel significativo no padrão de precipitação do estado. As regiões central e sul de Minas, com altitudes mais elevadas, tendem a presenciar a formação de chuvas orográficas devido ao relevo acidentado que força a ascensão do ar úmido (REBOITA et al., 2015). A maior parte da chuva concentra-se no verão (dezembro-fevereiro), influenciada principalmente pela Zona de Convergência do Atlântico Sul (ZCAS), frentes frias e outros sistemas como Complexos Convectivos de Mesoescala (CCM) e a circulação de umidade da Amazônia. As quantidades de chuva variam significativamente, sendo maiores no sul e Zona da Mata e menores no norte e nordeste do estado (REBOITA, et al., 2010).

Logo, por conta de toda a diversidade do estado, a destreza dos modelos numéricos se torna menor quando comparada às regiões do norte e sudeste da América do Sul (YNOUE *et al.*, [s.d.]). O El Niño-Oscilação Sul (ENOS) é um dos fenômenos climáticos globais que podem afetar o clima no estado, podendo causar períodos de seca prolongada ou chuvas intensas, relacionados ao posicionamento da ZCAS ou a passagem de sistemas transientes (DE ALBUQUERQUE CAVALCANTI; FERREIRA, 2021), impactando tanto a agricultura quanto os recursos hídricos.

Devido a essas grandes diferenças, são utilizadas divisões no estado chamadas de mesorregiões (Figura 2), definidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Esta divisão foi feita com base em semelhanças socioculturais e climáticas (IBGE, 1990).

Figura 2 - Mesorregiões de Minas Gerais.



12 Mesorregiões de Minas Gerais segundo o IBGE (1990). Fonte: Elaboração própria.

A previsão de precipitação é particularmente relevante em Minas Gerais, uma vez que atividades econômicas como a agricultura, mineração e a geração de energia hidrelétrica são altamente dependentes de condições climáticas. Analisar e prever a precipitação na região tem implicações diretas para a gestão dos recursos hídricos e para a mitigação de desastres naturais, como enchentes e secas severas (IGAM, 2020).

2.2. Dados

2.2.1 Precipitação

O dado de precipitação utilizado neste estudo foi o *Global Precipitation Climatology Project* (GPCP): dados diários com resolução espacial de $0,5^\circ \times 0,5^\circ$ (HUFFMAN et al., 2025a), no período de janeiro de 1983 a março de 2024, e os dados podem ser

acessados através de:

https://cmr.earthdata.nasa.gov/search/concepts/C3405928682-GES_DISC.html

2.2.2 Variáveis Atmosféricas e de volume de água no solo

Neste estudo foram utilizadas as seguintes variáveis mensais:

- Componentes horizontais do vento (zonal, u, e meridional, v) a 10 e 100 metros acima do nível do solo em metros por segundo (m s^{-1});
- Pressão atmosférica em superfície em Pascal (Pa);
- Temperatura do ar a 2 metros do solo em kelvin (K);
- Fluxo líquido médio de radiação de ondas longas e onda curta no topo da atmosfera em watts por metro quadrado (W m^{-2});
- Camada volumétrica de água do solo em quatro níveis (Camada 1: 0 - 7 cm, Camada 2: 7 - 28 cm, Camada 3: 28 - 100 cm, Camada 4: 100 - 289 cm) em metro cúbico de água por metro cúbico ($\text{m}^3 \text{ m}^{-3}$);

Todos os dados foram obtidos da reanálise ERA5 (Hersbach et al., 2020). O ERA5 possui uma resolução horizontal de $0,25^\circ \times 0,25^\circ$, e os dados podem ser acessados através do *Climate Data Store* (CDS): <https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means?tab=overview>.

2.3. Metodologia

A metodologia empregada neste trabalho foi estruturada em três etapas principais, que abrangem desde o tratamento dos dados brutos até a avaliação final do modelo de previsão. Este percurso metodológico foi concebido para garantir a robustez e a transparência do processo, permitindo a construção de um modelo de inteligência artificial otimizado para a previsão de precipitação sazonal no estado de Minas Gerais. A Figura 3 apresenta um fluxograma que sumariza visualmente todo o processo, ilustrando a sequência lógica e as principais atividades desenvolvidas em cada fase.

Figura 3 - Etapas do desenvolvimento do modelo de previsão sazonal.



Fluxograma das etapas metodológicas, desde a aquisição e análise de correlação dos dados (Etapa 1), passando pela otimização da arquitetura da rede neural via Grid Search (Etapa 2), até a geração da previsão por ensemble e sua avaliação final (Etapa 3). Fonte: Elaboração própria.

2.3.1. Preparação dos dados e correlação

Para escolha do ano que foi realizada a previsão, foi calculada a climatologia sazonal do conjunto de dados no período de 1992 a 2022, a fim de estabelecer uma referência para os cálculos das anomalias. Com base nesta etapa, determinaram-se as anomalias sazonais de precipitação, em que valores positivos e negativos indicam, respectivamente, períodos com precipitação acima e abaixo da média climatológica.

A escolha do ano para a previsão com o modelo não foi aleatória. O critério adotado consistiu em selecionar um período recente que apresentasse anomalias de precipitação positivas, caracterizando eventos extremos (chuvosos). Essa abordagem estratégica visa avaliar a robustez e a capacidade de generalização da rede neural

em cenários críticos, onde a variabilidade climática foge ao padrão médio, desafiando a capacidade preditiva do modelo frente a situações de alto impacto socioambiental.

Após a seleção do ano, foi realizada a conversão da série temporal de precipitação diária do produto GPCP para valores mensais, de modo a compatibilizar a resolução temporal com as demais variáveis analisadas. Essa transformação foi essencial para garantir coerência nas análises subsequentes e adequar os dados de entrada nos modelos de aprendizado de máquina.

Como forma de reduzir os efeitos das variações sazonais marcantes do regime pluviométrico de Minas Gerais, os dados de precipitação foram segmentados sazonalmente, em consonância com as estações do ano: verão (dezembro, janeiro e fevereiro), outono (março, abril e maio), inverno (junho, julho e agosto) e primavera (setembro, outubro e novembro). Essa abordagem contribui para a estabilização das séries temporais e evita que a modelagem estatística seja influenciada por oscilações cíclicas, preservando as variações interanuais.

A identificação das variáveis com maior influência sobre a precipitação, foi realizada por meio da análise de correlação linear de Pearson (equação 1) entre os dados de precipitação e cada possível variável preditora. A correlação foi aplicada entre a média espacial de precipitação sobre cada mesorregião e as variáveis atmosféricas para todo o estado, permitindo uma interpretação espacial detalhada. Segundo PARANHOS et al. (2014), o cálculo da correlação linear é dado por:

$$r = \frac{\sum_{t=1}^n (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{t=1}^n (x - \bar{x})^2 \sum_{t=1}^n (y - \bar{y})^2}} \quad (\text{Equação 1})$$

Onde: n é o número total de pares de observações temporais disponíveis; r é o coeficiente de correlação de Pearson; x e y são os valores temporais de duas variáveis comparadas; e $\bar{x}$ e $\bar{y}$ são suas respectivas médias.

Adotou-se um nível de confiança de 95% para a avaliação da correlação entre as séries. Isso significa que, quando a hipótese de ausência de correlação ($H_0: \rho = 0$) é testada, existe somente 5% de chance de rejeitá-la de forma equivocada, ou seja, de identificar uma correlação onde, na realidade, ela não existe. Conclui-se que a correlação é estatisticamente significativa. Em termos práticos, isso garante que os

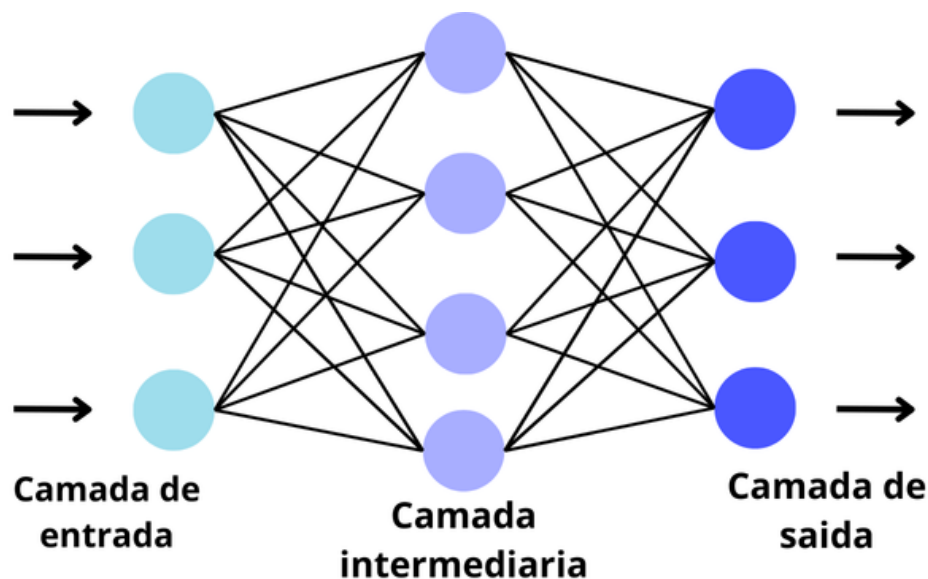
valores de correlação apresentam uma associação entre variáveis que dificilmente seria resultado de flutuações aleatórias dos dados.

Portanto as variáveis de entrada no modelo selecionadas foram aquelas que apresentaram maiores valores de correlação significativa e que demonstraram alguma relação física espacial.

2.3.2. Treinamento e Validação dos Modelos de IA

O treinamento do modelo de IA foi realizado utilizando a biblioteca *TensorFlow* em Python, uma ferramenta amplamente empregada para computação numérica em grande escala e aprendizagem de máquina. A escolha desta metodologia de *Machine Learning* baseou-se no estudo de Anochi (2021). O modelo de rede neural é estruturado em camadas, nas quais os dados de entrada são processados por múltiplos neurônios até atingirem a camada de saída como mostrado na Figura 4. Durante o treinamento, os dados são repetidamente propagados pela rede, ajustando-se os pesos internos com o objetivo de otimizar a acurácia do modelo. Cada ciclo completo desse processo contínuo de ajuste é denominado “época” (*epoch*).

Figura 4 – Modelo genérico de rede neural estruturado em camadas de entrada, ocultas e saída.



Fonte: Elaboração própria.

Para a etapa de treinamento e validação da rede neural, os dados foram particionados temporalmente de forma a respeitar a cronologia da série histórica e evitar o vazamento de dados futuros no treinamento. O conjunto de treinamento compreendeu o período de 1983 a 2020, totalizando 38 anos de observações das variáveis atmosféricas e de superfície. Para a etapa de teste, utilizada para avaliar a capacidade de generalização do modelo em dados não vistos, foram selecionados os

dados referentes ao ano de 2021. Adicionalmente, 10% dos dados do conjunto de treinamento foram separados aleatoriamente para validação interna do modelo a cada época.

A definição da arquitetura e dos hiperparâmetros ótimos do modelo de rede neural foi conduzida por um processo sistemático de busca em grade com validação cruzada, implementado com a função *GridSearchCV* (*Grid Search with Cross-Validation*) da biblioteca *Scikit-learn*.

O modelo base, construído com a biblioteca *Keras*, consiste em uma rede neural sequencial com uma camada de entrada para o número de variáveis preditoras de maiores correlações, um número variável de camadas ocultas e uma camada de saída com um único neurônio e ativação linear, adequada para problemas de regressão. A função de ativação *Rectified Linear Unit* (ReLU) foi mantida em todas as configurações, devido à sua consolidada eficácia (RIZZO et al., 2020).

Para encontrar a configuração de melhor desempenho, o modelo Keras foi encapsulado como um regressor compatível com o *Scikit-learn* usando *KerasRegressor*. A busca em grade foi então configurada para explorar o seguinte espaço de hiperparâmetros:

- Número de neurônios por camada oculta: [16, 32, 50, 64, 96]
- Número de camadas ocultas: [1, 2, 3]

O treinamento de cada combinação foi realizado com um tamanho de lote fixo de 10 e por 50 épocas.

Para a avaliação e seleção do melhor modelo, a *GridSearchCV* utilizou a métrica de pontuação '*neg_mean_squared_error*'. O objetivo é maximizar essa pontuação, o que é equivalente a minimizar o MSE. A combinação de hiperparâmetros que apresentou o melhor resultado médio na validação cruzada foi, portanto, selecionada para constituir o modelo final. Os dados de treinamento passaram por um processo de normalização antes de serem utilizados no ajuste do modelo.

Assim, obteve-se o modelo mais eficiente e para aumentar a precisão das previsões, aplicou-se a técnica de ensemble, que consiste em executar o mesmo modelo várias vezes e calcular a média dos resultados, reduzindo assim os erros aleatórios inerentes a cada execução. Após todas as rodadas de ensemble, calculou-se o MSE ao longo das iterações, conforme definido na Equação 3 (MARMOLIN, 1986).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - x)^2 \quad (\text{Equação 3})$$

Em que: y é o dado observado, x é o dado previsto, e n é o número de rodadas.

2.3.3. Avaliação das Previsões pelos Modelos de IA

Para contextualizar e avaliar o desempenho do modelo de IA desenvolvido, sua previsão foi comparada com a previsão sazonal média gerada por diversos modelos dinâmicos que compõem o sistema NMME. O NMME é um projeto operacional de ponta que reúne previsões sazonais de diversos modelos climáticos dinâmicos de centros de pesquisa da América do Norte. Diferente da técnica de ensemble de modelo único aplicada neste trabalho (múltiplas execuções do mesmo modelo de IA), ele utiliza uma abordagem de ensemble multimodelo. Esta técnica combina as saídas de diferentes modelos, cada um com sua própria representação das leis físicas da atmosfera e do oceano. A principal vantagem desta abordagem é a redução de erros e vieses sistemáticos, uma vez que as incertezas e deficiências de um modelo tendem a ser compensadas pelas de outros, resultando em uma previsão mais robusta e confiável. Portanto, a comparação com o NMME serve como um padrão de referência para quantificar a acurácia e o valor agregado do modelo de rede neural proposto, avaliando sua competitividade frente a uma das principais ferramentas de previsão climática sazonal utilizadas atualmente.

Após o treinamento do modelo de IA, gerou-se a previsão de precipitação para o ano de interesse de forma sazonal. A avaliação dessa previsão foi feita com base nos dados observados do mesmo período, utilizando-se as métricas de viés (Equação 4), MSE (Equação 3) e RMSE conforme CHAI *et al.* (2014) (Equação 5). Valores próximos de zero em ambas as métricas indicam um melhor desempenho do modelo em comparação com os dados observados. O viés (bias) aponta se o modelo tende a superestimar ou subestimar a precipitação, enquanto o RMSE e o MSE quantifica a magnitude média do erro, servindo como um indicador de precisão.

$$bias = \frac{1}{n} \sum_{i=1}^n (y - x) \quad (\text{Equação 4})$$

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y-x)^2}{n}} \quad (\text{Equação 5})$$

Em que nas equações 4 e 5: y é o dado observado; x é o dado previsto e n é o número de observações.

A comparação entre as previsões de cada rodada do *ensemble* também foi realizada por meio do cálculo do erro relativo (RE), conforme definido por GLECKLER et al. (2008) (Equação 6).

$$RE = \frac{E - \bar{E}}{\bar{E}} \quad (\text{Equação 6})$$

Em que: E é o valor de RMSE dos dados previstos e $\bar{E}$ é a mediana de todos os valores de RMSE obtidos nas previsões do modelo.

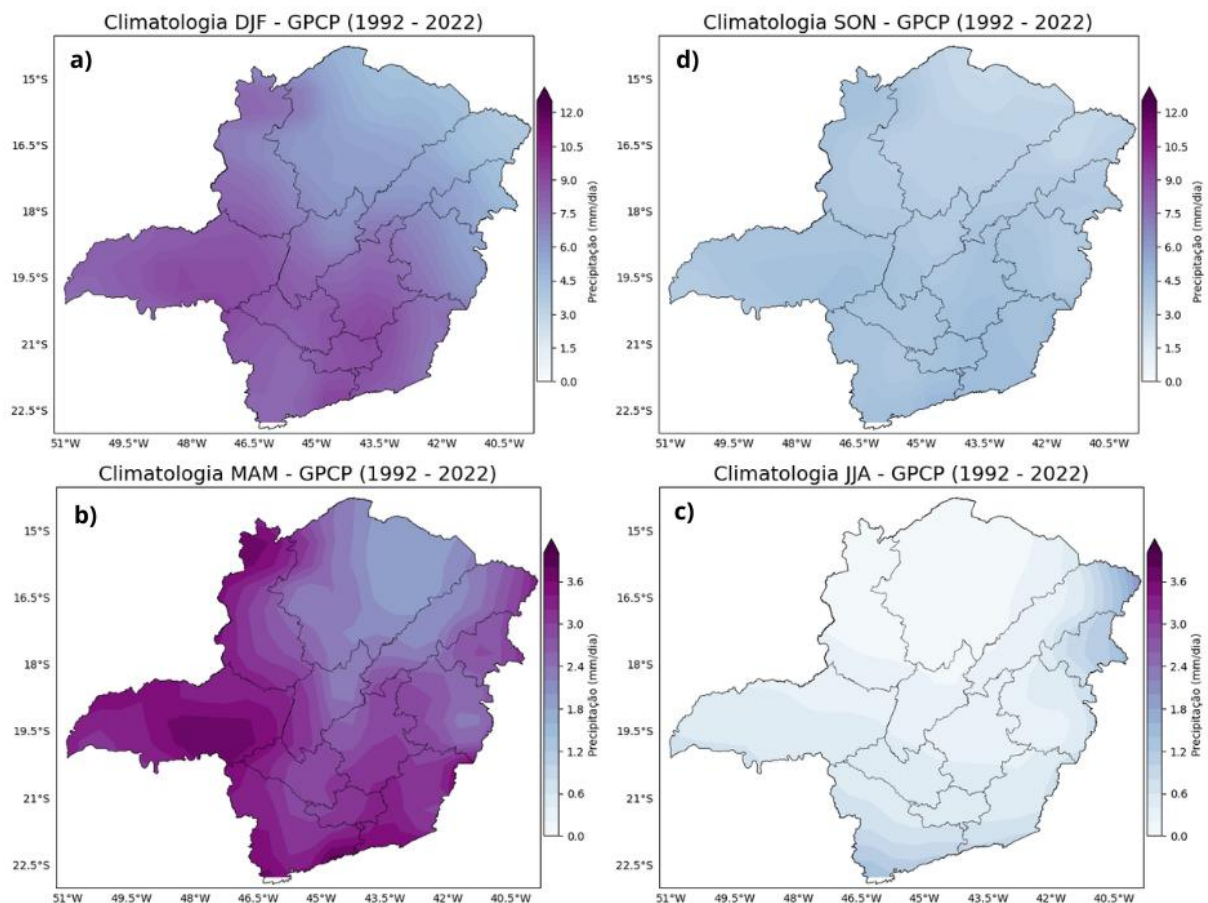
O modelo com melhor desempenho apresenta valores de viés e RMSE próximos de zero e de RE próximos de -1.

3. RESULTADOS E DISCUSSÃO

3.1 Análise Climatológica e Definição das Variáveis Predictoras

A Figura 5 mostra a climatologia sazonal da precipitação em Minas Gerais no período de 1992 a 2022, evidenciando a forte sazonalidade do regime de chuvas no estado. O intervalo de 30 anos adotado, garante uma representação estatisticamente robusta das condições médias recentes. Além disso, utilizar um período mais atualizado permite evitar a influência de décadas anteriores que podem apresentar padrões climáticos distintos, seja devido a limitações observacionais históricas ou a mudanças climáticas graduais.

Figura 5 – Climatologia sazonal de precipitação em Minas Gerais (1992-2022).



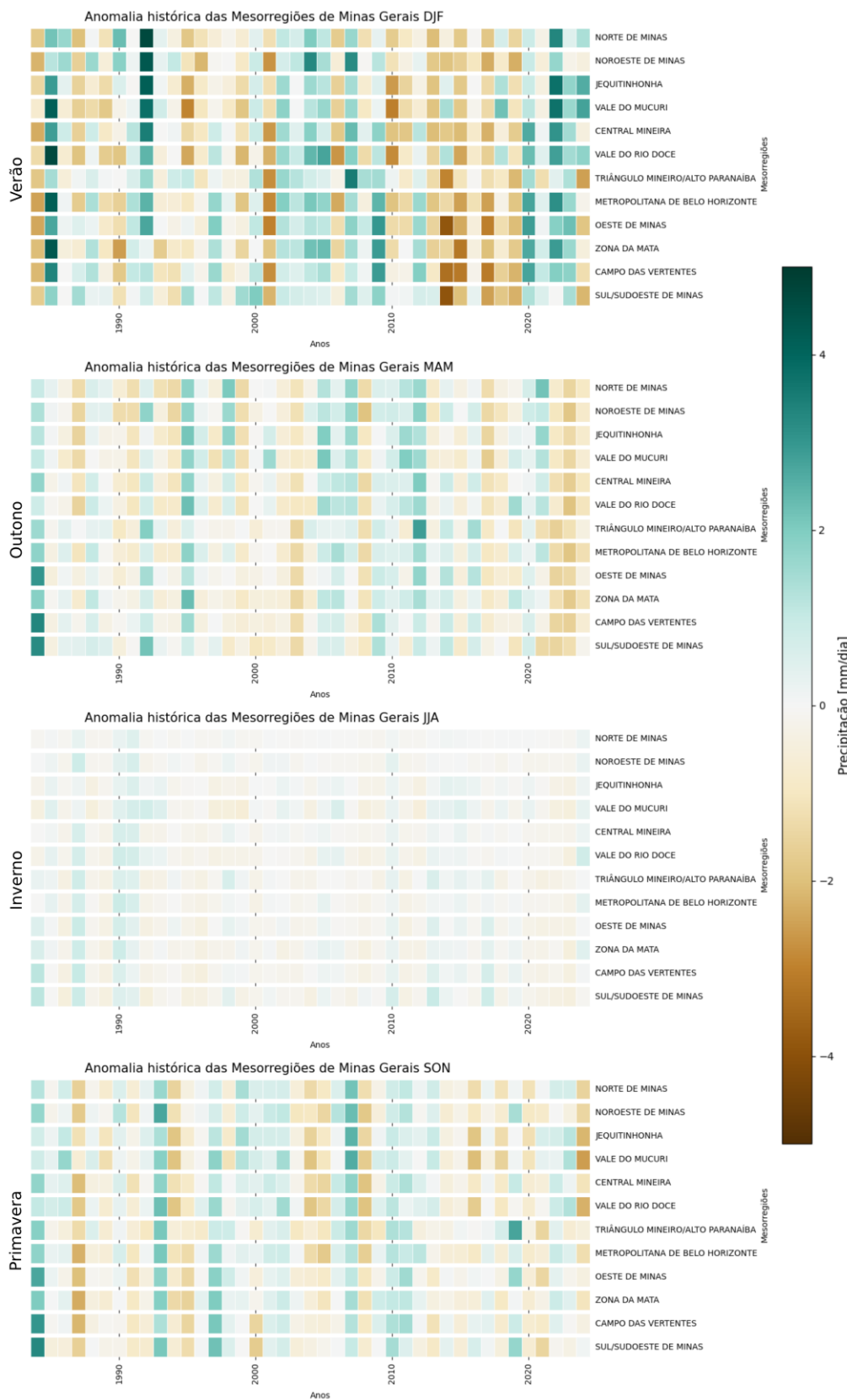
Climatologia sazonal da precipitação média diária (mm/dia) em Minas Gerais para o período de 1992–2022, obtida a partir do conjunto de dados GPCP. Os painéis representam: (a) verão (DJF), (b) outono (MAM), (c) inverno (JJA) e (d) primavera (SON). A escala de cores foram

ajustadas com maiores valores para a primavera-verão, e com menores para outono-inverno. Fonte: Elaboração própria.

O verão (DJF) é marcado pelos maiores acumulados médios diários, que chegam 12 mm/dia em algumas regiões, especialmente no centro e sul do estado, refletindo a atuação da ZCAS e a maior disponibilidade de umidade atmosférica. No outono (MAM), observa-se uma redução expressiva nos totais pluviométricos, com valores inferiores a 4 mm/dia. O inverno (JJA) representa a estação mais seca, com precipitação média geralmente abaixo de 1 mm/dia, exceto em áreas do leste do estado, onde a circulação marítima e a passagem de frentes frias ocasionais ainda contribuem para alguns eventos de chuva. A primavera (SON), por sua vez, marca o retorno gradual das chuvas, configurando a transição para o regime úmido do verão seguinte. Esses padrões refletem a dinâmica do clima tropical de monções na região Sudeste do Brasil, caracterizado por contrastes acentuados entre o período chuvoso e o período seco (REBOITA et al., 2010; GRIMM, 2019; RODRIGUES et al., 2022).

Já a análise do histórico de anomalias de precipitação nas mesorregiões de Minas Gerais, compreende o período de 1984 a 2024 para avaliar toda a extensão do dado pode ser visualizado em mapas de calor sazonais (Figura 6), e fornecem uma representação gráfica contundente das flutuações entre períodos excessivamente úmidos e severamente secos.

Figura 6 - Anomalia de Precipitação por mesorregião.



Anomalias de precipitação (mm dia^{-1}) nas mesorregiões de Minas Gerais, com base na climatologia de 1984 a 2024 para as estações: verão (DJF), outono (MAM), inverno (JJA) e

primavera (SON). Valores positivos (tons verdes) indicam precipitação acima da média climatológica, enquanto valores negativos (tons marrons) indicam precipitação abaixo da média. As linhas horizontais representam as diferentes mesorregiões, e as colunas representam os anos. Fonte: Elaboração própria.

As variações mais intensas de precipitação ocorrem, naturalmente, durante o verão (DJF), por ser este o período mais chuvoso do estado. Este efeito é visualmente acentuado na análise das anomalias, uma vez que a utilização de uma paleta de cores padronizada para as quatro estações faz com que os maiores desvios absolutos, típicos do verão, se destaquem com cores mais vivas em comparação às estações mais secas. Foi no verão que se registraram as anomalias positivas mais significativas, com anos como 1985, 1992 e, mais recentemente, 2022, destacando-se por suas chuvas torrenciais.

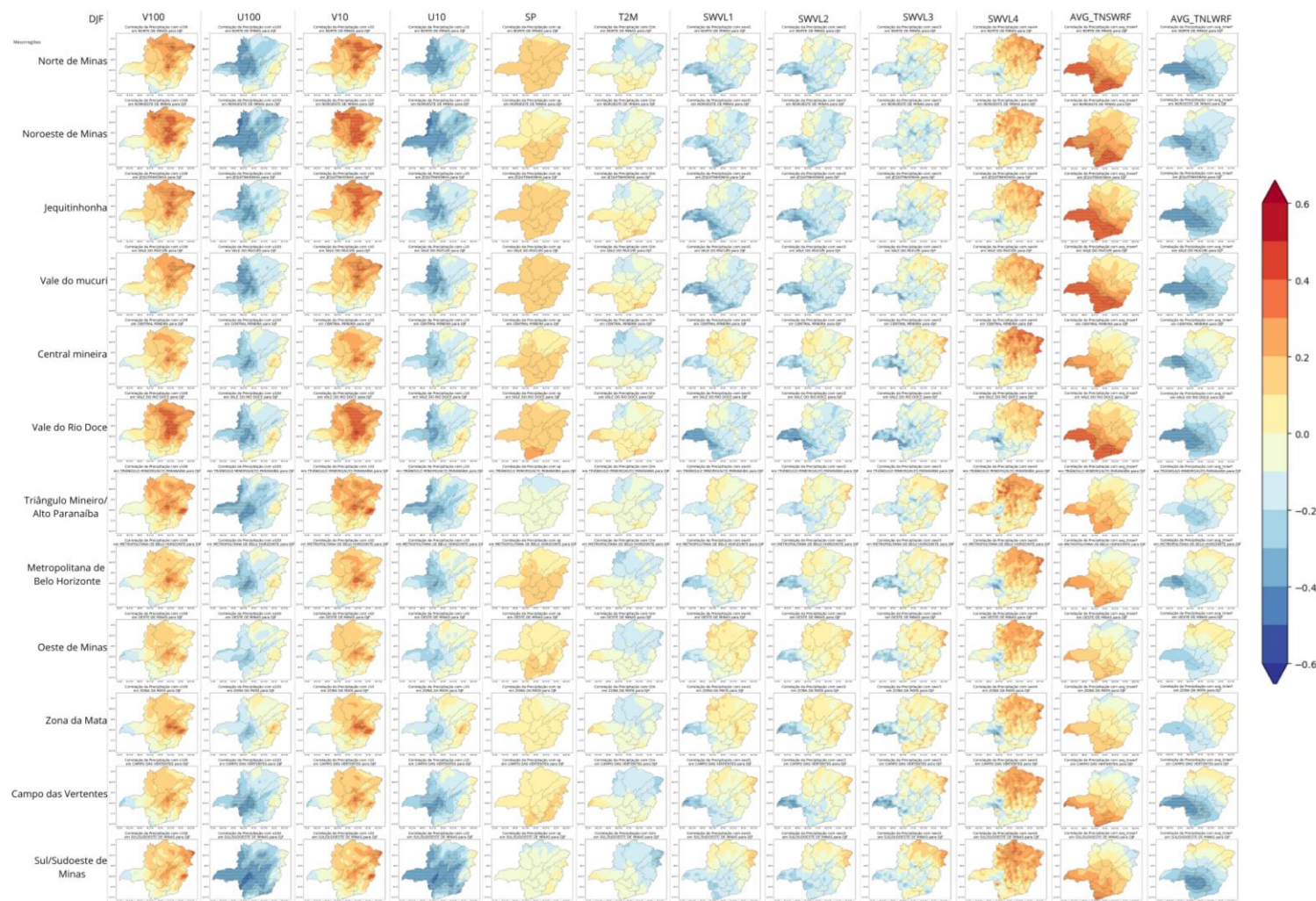
O evento de 1992 foi particularmente severo nas mesorregiões do Norte de Minas, Jequitinhonha e Vale do Mucuri, enquanto as chuvas de 2022 causaram grande impacto na porção central e leste do estado, incluindo a Região Metropolitana de Belo Horizonte e a Zona da Mata, caracterizada por volumes superiores à média climatológica. De acordo com Pezzi et al. (2023), esse fenômeno esteve associado à intensificação de situações atmosféricas favoráveis à formação de grandes sistemas de precipitação, como frentes frias e Zonas de Convergência do Atlântico Sul.

A análise sinótica realizada por Da Rocha Bartolomei et al. (2023) na Secretaria de Estado de Meteorologia aponta que, nesse período, a passagem de frentes frias próximas à costa e a atuação de sistemas de baixa pressão contribuíram para a elevada umidificação e o aumento da precipitação, culminando em eventos extremos com graves consequências socioeconômicas.

O ano de 2022 foi selecionado para a previsão por ser um evento de chuva extrema, o que representa um teste ideal para a capacidade da rede neural. O longo período anterior (1983-2021) fornece uma base de dados rica e variada para o treinamento, incluindo tanto secas severas quanto anos mais úmidos, permitindo que o modelo aprenda os complexos padrões que antecedem uma anomalia tão significativa. Embora as anomalias nas outras estações tenham sido próximas da neutralidade ou até negativas, a anomalia positiva durante o verão foi expressiva e dominante, tornando-o um caso de estudo ideal para testar a capacidade do modelo em prever um evento sazonal extremo.

A seleção de variáveis de entrada concentra-se nos processos de transporte atmosférico (ventos) e nos efeitos da nebulosidade (radiação). Na figura 7 é possível ver a correlação da precipitação no verão com as componentes zonal e meridional do vento a 10 e 100 metros (u10, v10, u100, v100,) radiação de onda curta (avg_tnswrf) e radiação de onda longa (avg_tnlwrf).

Figura 7 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas no Verão entre 1984 a 2023.



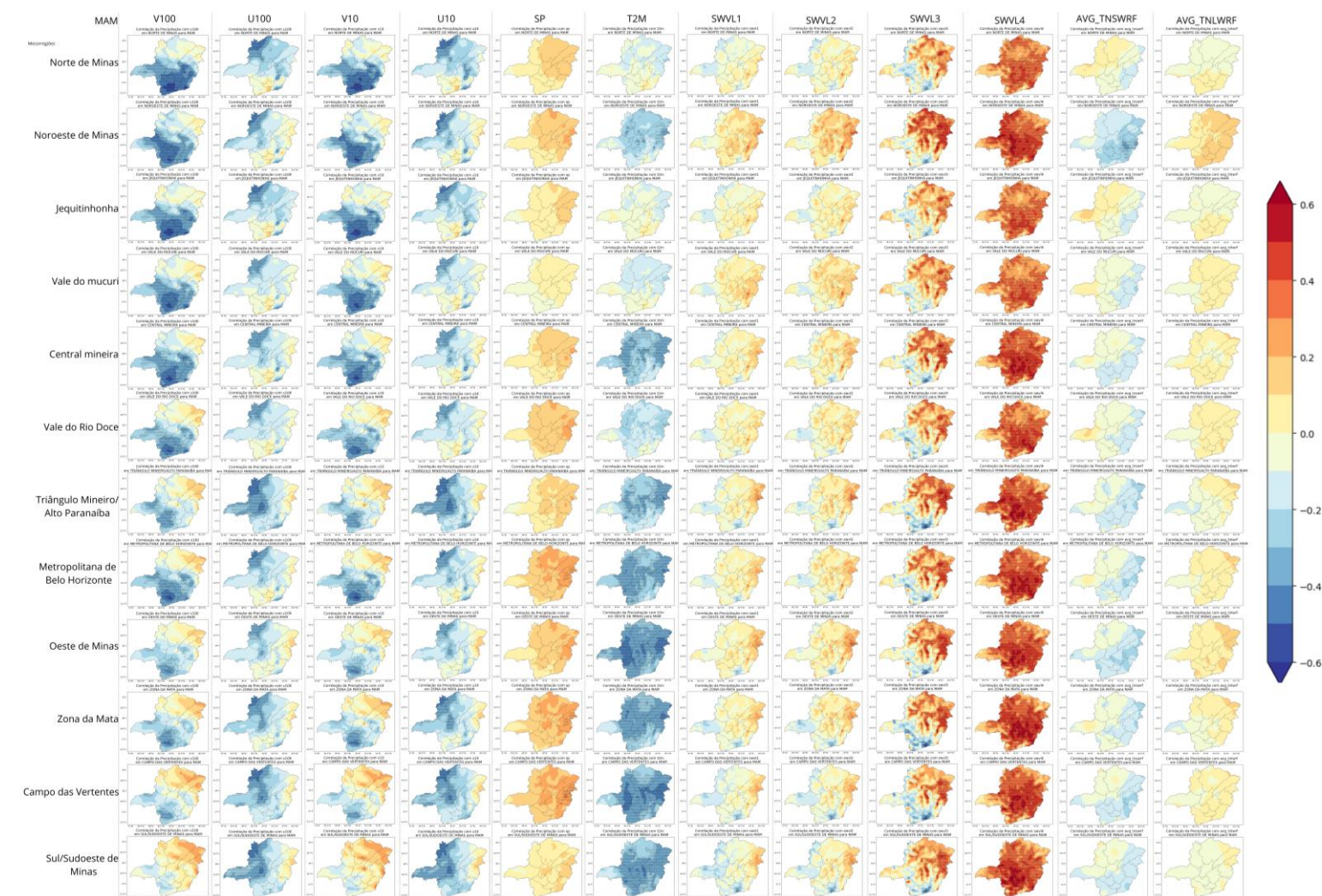
Correlação entre a precipitação média por mesorregião, e possíveis variáveis preditoras em todo o estado para o inverno. Em pontilhado a significância estatística de 95%. Fonte: Elaboração própria. Fonte: Elaboração própria.

A análise da correlação mostra que as variáveis de radiação possuem uma relação forte com a chuva: a radiação de onda curta (avg_tnsurf) tem uma correlação fortemente positiva, indicando que a presença de nuvens de chuva aumenta o albedo, ou seja, refletindo mais a radiação solar, enquanto a radiação de onda longa (avg_tnlwrf) tem uma correlação negativa, pois onde a precipitação é mais intensa, maior é a nebulosidade típica de verão com grande desenvolvimento vertical com isso menor é o fluxo radiativo líquido de onda longa. O conjunto de ventos, por sua vez, captura a dinâmica do transporte de umidade. Embora os componentes u10, u100 e v10 tenham correlações mais fracas, o vento meridional em 100 metros (v100) exhibe um padrão espacial claro, com correlação negativa no norte do estado. Este padrão é crucial, pois representa o fluxo de umidade vindo da região amazônica, um mecanismo essencial para a ocorrência de chuva na ZCAS, que atinge seu pico no verão.

A pressão apresentou correlações fracas, pois a precipitação de verão na região é mais influenciada por fatores termodinâmicos locais do que por variações no campo de pressão sinótico. Já a temperatura demonstrou uma correlação inconsistente, resultado de processos antagônicos: enquanto altas temperaturas favorecem a convecção, a presença de nuvens e chuva causa um resfriamento da superfície, resultando em um sinal de correlação sazonal fraco. Portanto, este conjunto de variáveis foca em modelar a precipitação com base na dinâmica da cobertura de nuvens e nos padrões de larga escala de transporte de umidade que definem a estação.

Na Figura 8, para o Outono as variáveis selecionadas foram u10, u100, v10, v100, t2m, swvl3, swvl4. Tendo assim, um modelo híbrido que combina indicadores de superfície e de dinâmica atmosférica para caracterizar esta estação de transição. A inclusão da temperatura a 2 metros (t2m) e da quantidade de água no solo (swvl3, swvl4) captura as condições termodinâmicas e hídricas da superfície. A t2m apresenta uma forte correlação negativa, indicando o efeito de resfriamento causado pela nebulosidade e pela chuva, enquanto os níveis de swvl mostram uma correlação positiva muito forte, refletindo a umidade residual do verão e a resposta do solo às últimas chuvas da estação.

Figura 8 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas no Outono entre 1984 a 2023.

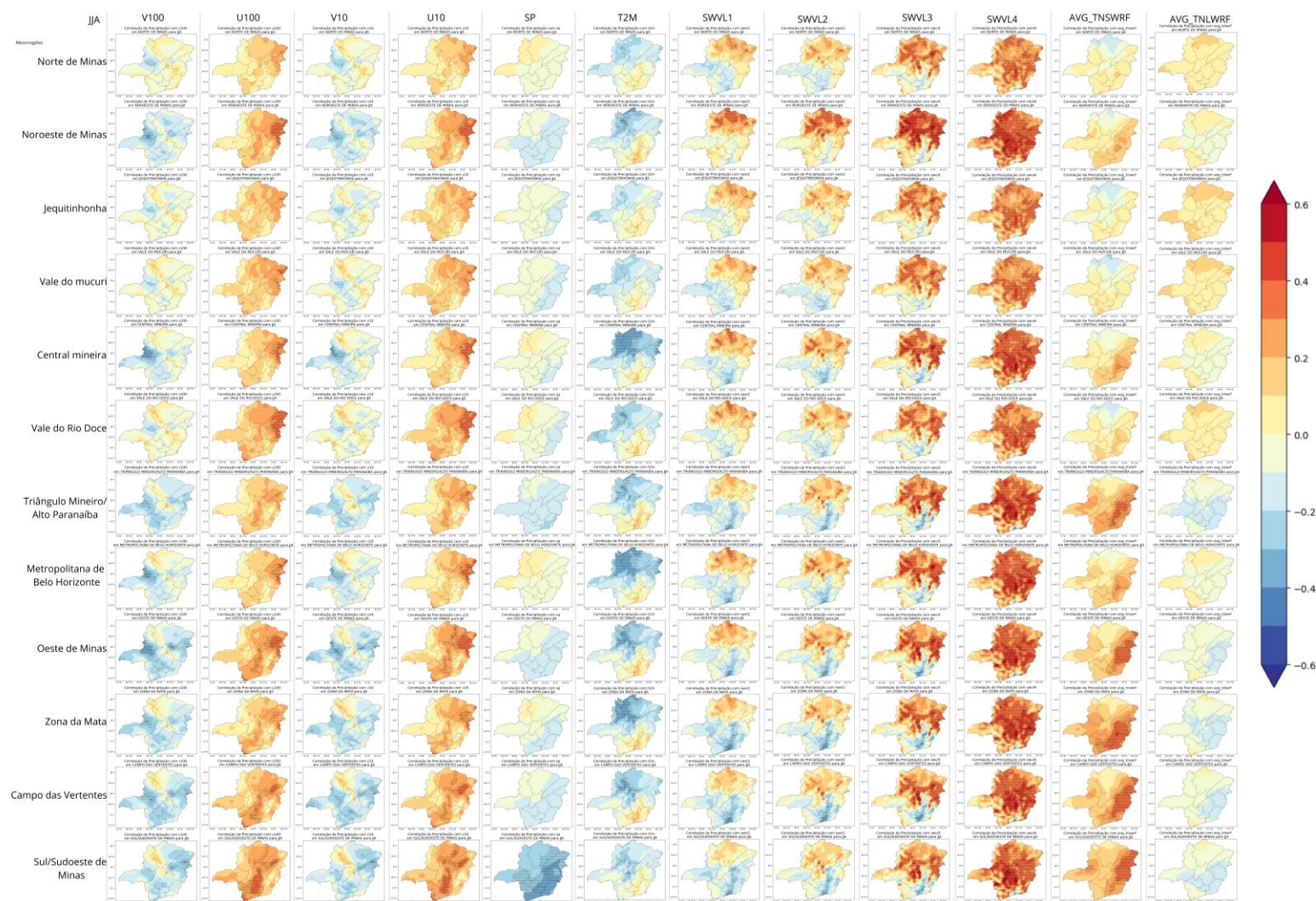


Correlação entre a precipitação média por mesorregião, e possíveis variáveis preditoras em todo o estado para o Outono. Em pontilhado a significância estatística de 95%. Fonte: Elaboração própria. Fonte: Elaboração própria.

O conjunto de todos os componentes do vento é incluído para capturar os padrões de circulação durante a retirada da estação chuvosa. Embora suas correlações sejam mais fracas e difusas que em outras estações, sua presença no modelo visa a identificar os sinais sutis da mudança no regime de ventos, que marcam a transição de um padrão de umidade de verão para as condições mais secas e estáveis que se instalarão no inverno. Em contrapartida, as camadas superficiais de umidade do solo foram descartadas por apresentarem correlações mais fracas. Elas respondem muito rapidamente a eventos isolados de chuva e evaporação, representando uma memória de curto prazo, enquanto as camadas profundas indicam a condição hídrica de longo prazo, mais relevante para a escala sazonal.

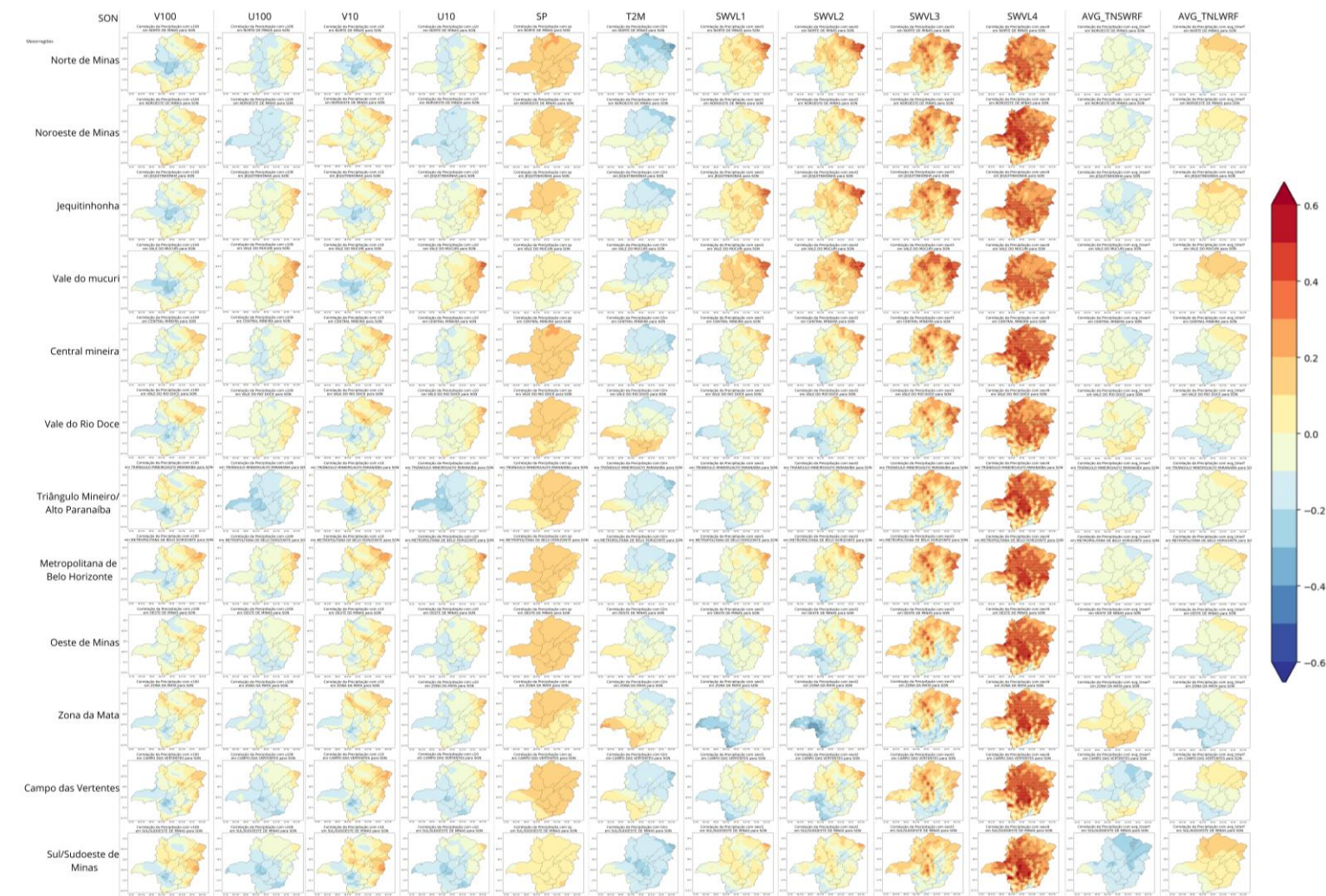
Para as estações da Primavera e do Inverno, a seleção de apenas duas variáveis (swvl3 e swvl4) indica um modelo focado exclusivamente na condição hídrica do solo. A análise mostra que, em ambas as estações, estes níveis mais profundos de umidade do solo possuem uma forte e consistente correlação positiva com a precipitação podendo ser visto na Figura 9 e Figura 10. As demais variáveis, como ventos e radiação, foram descartadas nestes períodos por apresentarem correlações muito fracas, esparsas e sem um padrão físico claro, uma vez que os processos dinâmicos que geram chuva são menos frequentes e intensos, tornando-as preditores sazonais pouco eficazes.

Figura 9 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas no Inverno entre 1984 a 2023.



Correlação entre a precipitação média por mesorregião, e possíveis variáveis preditoras em todo o estado para o inverno. Em pontilhado a significância estatística de 95%. Fonte: Elaboração própria. Fonte: Elaboração própria.

Figura 10 - Correlação no estado de Minas Gerais entre a precipitação em cada mesorregião e variáveis atmosféricas na Primavera entre 1984 a 2023.



Correlação entre a precipitação média por mesorregião, e possíveis variáveis preditoras em todo o estado para a primavera. Em pontilhado a significância estatística de 95%. Fonte: Elaboração própria.

No Inverno, a estação seca, a correlação positiva se mantém forte porque qualquer evento de chuva, mesmo que raro, causa um aumento imediato e detectável na umidade do solo, tornando swvl3 e swvl4 indicadores sensíveis da ocorrência de precipitação. Nesse caso, eles não representam um feedback para gerar mais chuva, mas sim uma consequência direta e um registro fiel de que a precipitação de fato ocorreu.

Na Primavera (Figura 10), a umidade armazenada nas camadas mais profundas do solo (swvl3, swvl4) funciona como uma memória do período seco e fonte de umidade para a atmosfera, através da evapotranspiração, no início da estação chuvosa. A forte correlação positiva sugere que um solo com maior capacidade de armazenamento ou que reteve mais umidade está diretamente associado a um maior volume de chuva na estação, servindo como um indicador fundamental da resposta hidrológica da bacia.

Com base nas análises apresentadas anteriormente para cada estação do ano, é possível observar na Tabela 1 as variáveis selecionadas que serão utilizadas como entrada na rede neural, representando os principais mecanismos atmosféricos e de superfície associados à precipitação em Minas Gerais.

Tabela 1 - Variáveis de entrada para cada estação

Estação	Variáveis Preditoras Selecionadas
Verão	Precipitação nos verões, u10, u100, v10, v100, avg_tnswrf, avg_tnlwrf
Outono	Precipitação nos outonos, u10, u100, v10, v100, t2m, swvl3, swvl4
Inverno	Precipitação nos invernos, swvl3, swvl4
Primavera	Precipitação nas primaveras, swvl3, swvl4

Variáveis meteorológicas selecionadas como entradas para o modelo em cada estação do ano. As variáveis incluem componentes do vento em diferentes níveis (u e v a 10 m e 100 m), temperatura do ar a 2 m (t2m), fluxo de radiação de onda curta e longa na superfície (avg_tnswrf e avg_tnlwrf) e umidade do solo em diferentes camadas (swvl3 e swvl4). Fonte: Elaboração própria.

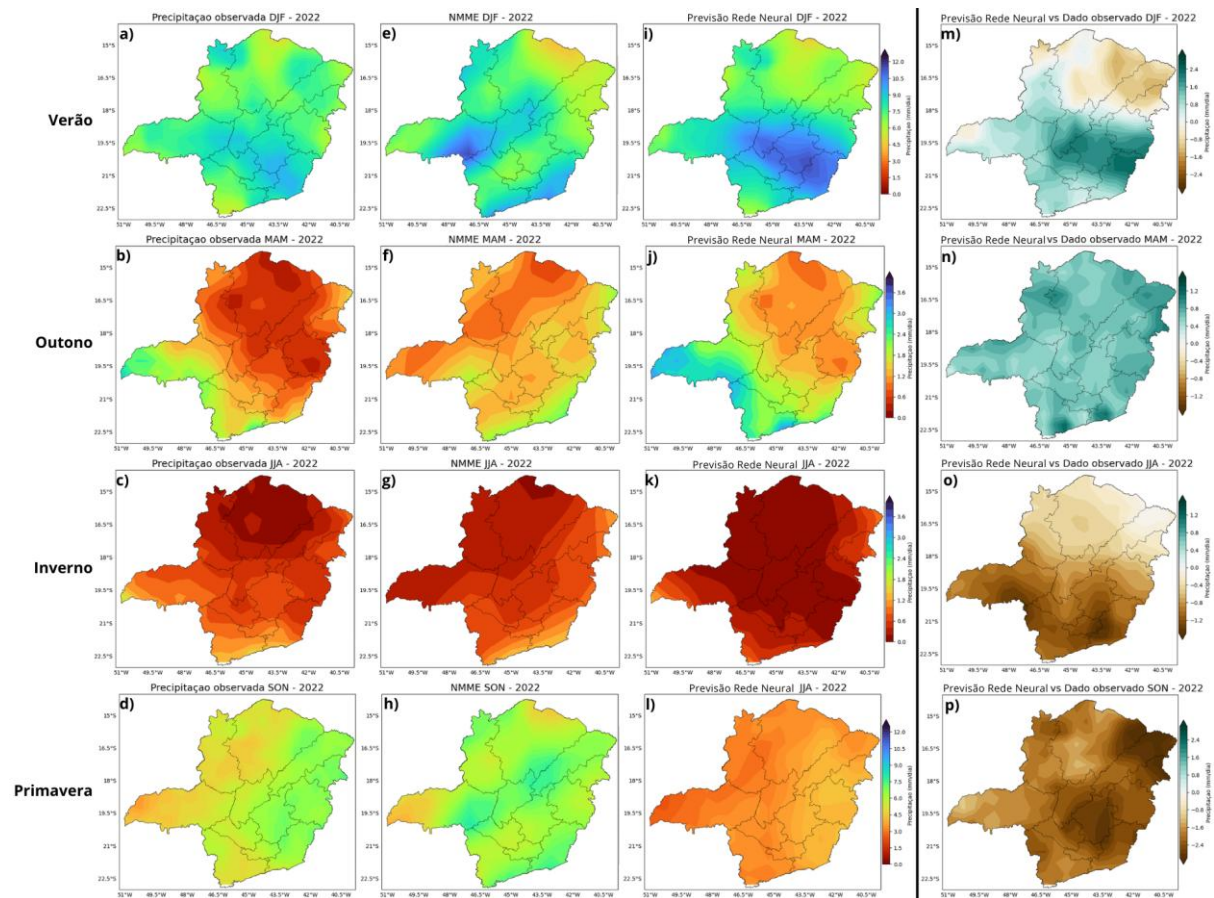
3.2. Configuração e Otimização e da IA

A configuração da rede neural que apresentou o melhor desempenho consistiu em uma arquitetura sequencial com três camadas ocultas, cada uma contendo 96 neurônios, função de ativação ReLU, tamanho de lote (batch) de 10 e treinamento por 50 épocas. Essa combinação foi selecionada por apresentar o menor MSE médio na validação cruzada (`neg_mean_squared_error`), superando as demais configurações testadas no espaço de hiperparâmetros. Com o objetivo de aumentar a robustez das previsões, aplicou-se a técnica de ensemble, essa abordagem reduziu a variabilidade entre execuções e melhorou o desempenho médio observado na validação, consolidando sua escolha como versão final para a etapa de previsão. Os testes foram conduzidos separadamente para as quatro estações do ano (verão, outono, inverno e primavera) permitindo avaliar o comportamento do modelo em diferentes condições sazonais.

3.3. Avaliação da previsão por IA

Na Figura 11, é possível ver a comparação entre a média do ensemble e os dados observados para 2022. Observa-se que as paletas de cor de SON e JJA foram definidas com limites diferentes daqueles aplicados a MAM e DJF, para evidenciar a maior amplitude de vieses observada em certas estações. No verão (DJF) o ensemble apresenta superestimação localizada no centro e sudeste de Minas Gerais, com magnitudes próximas a 2.4 mm dia^{-1} , enquanto o extremo norte do estado mostra subestimação. No outono (MAM) os desvios são menores (ordem de 1.2 mm dia^{-1}) e tendem a leve superestimação. Durante o inverno (JJA) e primavera (SON) observa-se uma subestimação ampla e consistente em grande parte do território (valores negativos até -1.2 a -2.4 mm dia^{-1}), indicando um viés seco sazonal do sistema preditivo. Esses padrões sugerem deficiências na representação de processos convectivos e/ou efeitos orográficos. Apontam para a necessidade de correções de viés sazonais, investigação da variabilidade do ensemble e também verificação de variáveis de entrada que melhor correlacionam com a precipitação.

Figura 11 - Comparação entre o dado observado, o NMME, e a previsão pela Rede Neural para 2022.



Comparação da precipitação sazonal observada (GPCC) (a-d), o modelo NMME (e-h) e estimada pela rede neural com método de ensemble (i-l). Os painéis mostram os resultados para o verão, outono, inverno, e primavera com escalas de cores ajustadas com maiores valores para a primavera-verão, e com menores para outono-inverno. Diferença de precipitação prevista pela rede neural e observada (m-p). Fonte: Elaboração própria.

Enquanto o ensemble de IA mostra um viés seco predominante em SON e JJA e superestimções localizadas no DJF (centro-sudeste), o NMME exibe padrões de erro deslocados espacialmente com superestimções mais pronunciadas no nordeste do domínio em DJF e maior heterogeneidade em MAM. Em outras palavras, os erros são frequentemente complementares: regiões onde a IA subestima podem ser zonas onde o NMME superestima, e vice-versa. Essa complementaridade sugere que estratégias de combinação multi-modelo ou de recalibração para reduzir o viés espacial e melhorar a skill sazonal.

A Tabela 2 resume as métricas de desempenho (RMSE, MSE e RE) da rede neural e do produto NMME para as quatro estações de 2022. A rede neural apresenta

menor RMSE em Verão (0.9993 vs 1.3015), Outono (0.7037 vs 0.7980) e Inverno (0.8210 vs 0.8417), indicando desempenho médio superior nessas estações. Contudo, na Primavera o comportamento inverte-se: a rede neural tem RMSE substancialmente maior que o NMME (2.1086 vs 1.2339), apontando limitação clara do modelo de aprendizagem nessa estação. Os valores de RE (próximos de zero em todas as estações) indicam que as rodadas avaliadas estão muito próximas das medianas internas do experimento, mas não mitigam as diferenças absolutas de RMSE observadas, especialmente na primavera.

Tabela 2 - Avaliação dos Erros da rede neural, e do NMME nas 4 estações

	Avaliação do modelo da rede neural			Avaliação do NMME	
Estação	RMSE	MSE	RE	RMSE NMME	MSE NMME
Verão	0.9993	0.9986	0.00294	1.3015	1.6939
Outono	0.7037	0.4952	-0.00728	0.7980	0.6369
Inverno	0.8210	0.6740	-0.00086	0.8417	0.7085
Primavera	2.1086	4.4461	0.00013	1.2339	1.5225

Métricas de avaliação do desempenho da rede neural e do modelo NMME para cada estação do ano. São apresentadas a raiz do erro quadrático médio (RMSE), o erro quadrático médio (MSE) e o erro relativo (RE) para a rede neural, além do RMSE e MSE para o NMME. Valores em azul indicam métricas nas quais o método apresentou melhor desempenho, enquanto valores em vermelho destacam aquelas em que houve pior desempenho. Fonte: Elaboração própria.

4. CONCLUSÃO

O presente trabalho demonstrou com sucesso a viabilidade do desenvolvimento de um modelo de rede neural de baixo custo computacional para a previsão de precipitação sazonal em Minas Gerais, apresentando tempo de processamento significativamente inferior ao de grandes modelos climáticos tradicionais. A metodologia, que partiu da análise de correlação para selecionar as variáveis preditoras mais influentes em cada estação do ano, mostrou-se eficaz mesmo com um conjunto enxuto de informações, resultando em previsões consistentes. Além disso, observou-se que a rede neural e o modelo dinâmico NMME apresentam padrões de erro espacialmente distintos, o que indica que a IA não substitui, mas complementa os modelos físicos, podendo melhorar suas limitações em diferentes regiões e períodos sazonais.

Os resultados obtidos foram promissores, evidenciando que a arquitetura de rede neural otimizada, composta por três camadas ocultas com 96 neurônios cada, foi capaz de reproduzir os padrões de precipitação sazonal com boa acurácia, especialmente para o evento extremo de 2022. A validação demonstrou que o modelo de inteligência artificial superou o desempenho do NMME nas estações de verão, outono e inverno, apresentando um menor RMSE, o que reforça o potencial da abordagem.

No entanto, o modelo apresentou limitações significativas, como uma subestimação generalizada nas estações de inverno e primavera e uma superestimação localizada no centro-sul durante o verão. A baixa performance na primavera, onde o RMSE foi consideravelmente superior ao do NMME, sugere que as variáveis selecionadas (umidade do solo) podem não ser suficientes para capturar a complexidade da transição entre a estação seca e a úmida, período em que os mecanismos convectivos se intensificam. A complementaridade observada entre os erros do modelo de IA e os do NMME aponta para a oportunidade de desenvolver sistemas híbridos que combinem as fortalezas de ambas as abordagens para mitigar vieses e aprimorar a precisão das previsões.

Para trabalhos futuros, recomenda-se a exploração de um conjunto mais amplo de variáveis preditoras, incluindo parâmetros de teleconexões climáticas que influenciam a variabilidade da precipitação em Minas Gerais. Além disso, a aplicação de técnicas de pós-processamento para correção de viés e a investigação de outras

arquiteturas de redes neurais, como as redes neurais convulsionais (CNNs) ou de memória de curto e longo prazo (LSTMs), podem aprimorar a representação dos padrões espaço-temporais e a captura de dependências de longo prazo. A continuidade desta pesquisa é de grande relevância para a consolidação de ferramentas de previsão climática eficientes e de baixo custo, com potencial para auxiliar na gestão de recursos hídricos e na mitigação dos impactos de eventos climáticos extremos no estado.

5. REFERÊNCIAS

- ANOCHI, Juliana Aparecida. **Previsão climática de precipitação por redes neurais autoconfiguradas**. 2015. 2015. Tese de Doutorado. Tese (Doutorado). Curso de Pós-Graduação em Computação Aplicada, Instituto Nacional de Pesquisas Espaciais–INPE. São José dos Campos.
- ANOCHI, Juliana Aparecida; DE ALMEIDA, Vinícius Albuquerque; DE CAMPOS VELHO, Haroldo Fraga. Machine learning for climate precipitation prediction modeling over South America. **Remote Sensing**, v. 13, n. 13, p. 2468, 2021.
- ANOCHI, Juliana Aparecida; HERNÁNDEZ TORRES, Reynier; DE CAMPOS VELHO, Haroldo Fraga. Two geoscience applications by optimal neural network architecture. **Pure and Applied Geophysics**, v. 177, n. 6, p. 2663-2683, 2020.
- ARAUJO, Taciana Lima; SOUZA, Ê. P. Influência da Superfície no Balanço de Energia e na Intensidade da Precipitação Urbana. **Revista Brasileira de Geografia Física**, v. 3, p. 204-217, 2010.
- BENISTON, Martin; STEPHENSON, David B. Extreme climatic events and their evolution under changing climatic conditions. **Global and planetary change**, v. 44, n. 1-4, p. 1-9, 2004.
- BOCHENEK, Bogdan; USTRNUL, Zbigniew. Machine learning in weather prediction and climate analyses—applications and perspectives. **Atmosphere**, v. 13, n. 2, p. 180, 2022.
- BOMVENTI, Tiago Nicolosi; WAINER, Ilana Elazari Klein Coaracy; TASCHETTO, Andrea Sardinha. Relação entre a radiação de onda longa, precipitação e temperatura da superfície do mar no oceano atlântico tropical. **Revista Brasileira de Geofísica**, v. 24, p. 513-524, 2006.
- CHAI, Tianfeng et al. Root mean square error (RMSE) or mean absolute error (MAE). **Geoscientific model development discussions**, v. 7, n. 1, p. 1525-1534, 2014.
- CONTI, Silvia. Artificial intelligence for weather forecasting. **Nature Reviews Electrical Engineering**, v. 1, n. 1, p. 8-8, 2024.
- CINTRA, Rosangela Saher; DE CAMPOS VELHO, Haroldo F. Data assimilation by artificial neural networks for an atmospheric general circulation model. **Advanced applications for artificial neural networks**, v. 265, 2018.

DA ROCHA BARTOLOMEI, Fabiana; RIBEIRO, João Gabriel Martins; REBOITA, Michelle Simões. Eventos extremos de precipitação no sudeste do Brasil: Verão 2021/2022. **Revista Brasileira De Geografia Física**, v. 16, n. 05, p. 2658-2676, 2023.

DE ALBUQUERQUE CAVALCANTI, Iracema Fonseca; FERREIRA, Nelson Jesus. **Clima das regiões brasileiras e variabilidade climática**. Oficina de Textos, 2021.
GLECKLER, Peter J.; TAYLOR, Karl E.; DOUTRIAUX, Charles. Performance metrics for climate models. **Journal of Geophysical Research: Atmospheres**, v. 113, n. D6, 2008.

FREITAS, Saulo Ribeiro de. **MODEL FOR OCEAN-LAND-ATMOSPHERE PREDICTION (MONAN): AVALIAÇÃO DOS CANDIDATOS AO NÚCLEO DINÂMICO DO COMPONENTE ATMOSFÉRICO (MONAN-ATM) RECOMENDAÇÃO AO COMITÊ CIENTÍFICO DO MONAN**. São José dos Campos: INPE, 2023. (Nota Técnico-Científica sid.inpe.br/mtc-m21d/2023/08.25.12.01-NTC). Disponível em: <http://urlib.net/8JMKD3MGP3W34T/49MKML8>. Acesso em: 20 maio 2025

FURTADO, Maria Inês Vasconcellos. Redes neurais artificiais: uma abordagem para sala de aula. **Ponta Grossa, PR. Atena Editora**, v. 19, 2019.

GOLDSBOROUGH, Peter. A tour of tensorflow. **arXiv preprint arXiv:1610.01178**, 2016.

GOODFELLOW, Ian et al. **Deep learning**. Cambridge: MIT press, 2016.

GRIMM, Alice M. South American monsoon and its extremes. In: **Tropical extremes**. Elsevier, 2019. p. 51-93.

HERSBACH, H. et al. *ERA5 monthly averaged data on single levels from 1940 to present* [recurso eletrônico]. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2019. Disponível em: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means>. Acesso em: 2 maio 2025. DOI: 10.24381/cds.f17050d7.

HOLTON, James R.; HAKIM, Gregory J. **An introduction to dynamic meteorology**. Academic press, 2013.

HUFFMAN, G. J.; BEHRANGI, A.; BOLVIN, D. T.; NELKIN, E. J. *GPCP Version 3.3 Daily Precipitation Data Set* [recurso eletrônico]. Greenbelt, Maryland, USA: Goddard Earth Sciences Data and Information Services Center (GES DISC), 2025a. Disponível em: https://disc.gsfc.nasa.gov/datacollection/GPCPDAY_3.3.html. Acesso em: 2 maio 2025. DOI: 10.5067/MEASURES/GPCP/DATA307.

HUFFMAN, George J. et al. **Algorithm Theoretical Basis Document (ATBD) for Global Precipitation Climatology Project Version 3.3 Daily Precipitation Data**. Greenbelt, MD: NASA Goddard Space Flight Center, MEaSUREs Project, 11 fev. 2025b. Disponível em: https://docserver.gesdisc.eosdis.nasa.gov/public/project/MEaSUREs/GPCP/GPCP_ATBD_V3.3_Daily.pdf. Acesso em: 16 maio 2025.

HUFFMAN, George J. et al. **README document for Global Precipitation Climatology Project Version 3.3 Daily Precipitation Data**. Greenbelt, MD: NASA Goddard Space Flight Center, Goddard Earth Sciences Data and Information Services Center (GES DISC), 18 fev. 2025c. Disponível em: <https://doi.org/10.5067/MEASURES/GPCP/DATA307>. Acesso em: 16 maio 2025.

IBGE. **Área territorial brasileira 2022**. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, 2023. Acesso em: <https://cidades.ibge.gov.br/brasil/mg/panorama>. Acesso em: 05 nov. 2024.

INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS (IGAM). **Relatório de Gestão e Situação das Águas de Minas Gerais 2020**. Belo Horizonte: IGAM, 2020. Disponível em: <https://www.google.com/search?q=http://www.igam.mg.gov.br>. Acesso em: 16 novembro 2025.

JARDIM, Carlos Henrique. A "crise hídrica" no sudeste do Brasil: aspectos climáticos e repercussões ambientais. **Revista Tamoios**, v. 11, n. 2, 2015.

LAM, Remi et al. Learning skillful medium-range global weather forecasting. **Science**, v. 382, n. 6677, p. 1416-1421, 2023.

LYNCH, Peter. The origins of computer weather prediction and climate modeling. **Journal of computational physics**, v. 227, n. 7, p. 3431-3444, 2008.

MACFERRIN, Michael J. et al. ETOPO 2022: NOAA's new seamless topography-and-bathymetry bare earth surface elevation dataset. In: **AGU Fall Meeting Abstracts**. 2022. p. NH22C-0433.

MARMOLIN, Hans. Subjective MSE measures. **IEEE transactions on systems, man, and cybernetics**, v. 16, n. 3, p. 486-489, 1986.

MELO, Maria Luciene Dias de; MARENGO, José Antonio. Simulações do clima do Holoceno médio na América do Sul com o modelo de Circulação Geral da Atmosfera do CPTEC. **Revista Brasileira de Meteorologia**, v. 23, p. 191-205, 2008.

OBILO, Ezezi Isaac; AMADI, Eric Chikweru. Test for significance of Pearson's correlation coefficient. **International Journal of Innovative Mathematics, Statistics & Energy Policies**, v. 6, n. 1, p. 11-23, 2018.

PARANHOS, Ranulfo et al. Desvendando os mistérios do coeficiente de correlação de Pearson: o retorno. **Leviathan (São Paulo)**, n. 8, p. 66-95, 2014.
pressure-levels-monthly-means. Acesso em: 2 maio 2025. DOI:
10.24381/cds.6860a573.

REBOITA, Michelle Simões et al. Aspectos climáticos do estado de Minas Gerais. **Revista brasileira de Climatologia**, v. 17, 2015.

REBOITA, Michelle Simões et al. Regimes de precipitação na América do Sul: uma revisão bibliográfica. **Revista brasileira de meteorologia**, v. 25, 2010.

RIZZO, Igor Vilela; CANATO, Robson Leandro Carvalho. Inteligência artificial: funções de ativação. **Prospectus (ISSN: 2674-8576)**, v. 2, n. 2, 2020.

RODRIGUES, Maria AM et al. Onset and demise dates of the rainy season in the South American monsoon region: A cluster analysis result. **International Journal of Climatology**, v. 42, n. 3, p. 1354-1368, 2022.

SAMPAIO, Gilvan; DA SILVA DIAS, Pedro Leite. Evolução dos Modelos Climáticos e de Previsão de Tempo e Clima. **Revista USP**, n. 103, p. 41-54, 2014.

VOLODINA, Victoria; CHALLENGOR, Peter. The importance of uncertainty quantification in model reproducibility. **Philosophical Transactions of the Royal Society A**, v. 379, n. 2197, p. 20200071, 2021.

YNOUE, Rita Yuri; REBOITA, Michelle S.; AMBRIZZI, Tércio; SILVA, Gyrlene A. M. da; BOIASKI, Nathalie T. Previsão de tempo e clima. In: UNIVERSIDADE DE SÃO PAULO; UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO. **Meteorologia: Licenciatura em Ciências: Módulo 2**. São Paulo: Universidade de São Paulo: UNIVESP, [s.d.]. cap. 12, p. 257-275. Disponível em:
https://midia.atp.usp.br/plc/plc0009/impressos/plc0009_12.pdf. Acesso em: 22 maio 2025.

ZHANG, Fuqing et al. What is the predictability limit of midlatitude weather?. **Journal of the Atmospheric Sciences**, v. 76, n. 4, p. 1077-1091, 2019.

ZHU, Xinrong et al. Quantificando o papel do solo na redistribuição da precipitação local para o crescimento da vegetação. **Indicadores Ecológicos**, v. 124, p. 107355, 2021.