UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Instituto de Ciência e Tecnologia
Câmpus de Sorocaba

CAIO ALENCAR MACIEL

PREDIÇÃO CLIMÁTICA UTILIZANDO REDES NEURAIS CONVOLUCIONAIS E REDES NEURAIS RECORRENTES

Sorocaba
2026

CAIO ALENCAR MACIEL

PREDIÇÃO CLIMÁTICA UTILIZANDO REDES NEURAIS CONVOLUCIONAIS E REDES NEURAIS RECORRENTES

Trabalho de Conclusão de Curso apresentado à Universidade Estadual Paulista (UNESP), Instituto de Ciência e Tecnologia, Sorocaba, como parte dos requisitos para obtenção do grau de Bacharel(a) em Engenharia de Controle e Automação.

Orientador(a): Prof. Dr. Leopoldo André Dutra Lusquino Filho

Sorocaba

2026

M152p

Maciel, Caio Alencar

Predição climática utilizando redes neurais convolucionais e redes neurais recorrentes / Caio Alencar Maciel. -- Sorocaba, 2026

174 p. : il., tabs.

Trabalho de conclusão de curso (Bacharelado - Engenharia de Controle e Automação) - Universidade Estadual Paulista (UNESP), Instituto de Ciência e Tecnologia, Sorocaba

Orientador: Leopoldo André Dutra Lusquino Filho

1. Redes neurais (Computação). 2. Previsão do tempo. 3. Aprendizado do computador. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

CAIO ALENCAR MACIEL

PREDIÇÃO CLIMÁTICA UTILIZANDO REDES NEURAIS CONVOLUCIONAIS E REDES NEURAIS RECORRENTES

Trabalho de Conclusão de Curso apresentado à Universidade Estadual Paulista (UNESP), Instituto de Ciência e Tecnologia, Sorocaba, para obtenção do título de Bacharel em Engenharia de Controle e Automação.

Data da defesa: 03/09/2026

Banca Examinadora:

_____

Prof. Dr. Leopoldo André Dutra Lusquino Filho
UNESP – Instituto de Ciência e Tecnologia- Campus de Sorocaba

_____

Prof. Dr. William Dantas Vichete
UNESP – Instituto de Ciência e Tecnologia- Campus de Sorocaba

_____

Prof. Dr. Luiz Fernando dos Reis de Oliveira
CEPEL – Centro de Pesquisas de Energia Elétrica

AGRADECIMENTOS

Agradeço aos meus pais, Vagner e Maria, pelo apoio emocional e financeiro incondicional, tanto antes do meu ingresso na universidade quanto ao longo de toda a graduação. Muito obrigado por sempre investirem em uma educação de qualidade para mim

Aos meus irmãos, Bianca e Maurício, por sempre servirem de inspiração para eu buscar e me dedicar aos estudos.

Também agradeço ao meu orientador, Prof. Dr. Leopoldo André Dutra Lusquino Filho, pelo apoio e paciência durante todo o desenvolvimento do trabalho.

Por fim, a todos os amigos que fiz e com quem pude compartilhar o ambiente universitário, por tornarem essa vivência mais fácil, em especial aos meus amigos André Silva, Levi Correia e Santiago Teixeira, que me deram um grande suporte durante toda a minha jornada.

RESUMO

A previsão de variáveis atmosféricas com aprendizado profundo tem usado redes híbridas convolucional-recorrentes com frequência. Essa prática parte do pressuposto de que a camada convolucional tende a melhorar o desempenho. Este trabalho verifica a validade desse pressuposto quando o problema é uma série temporal multivariada de ponto único, sem estrutura espacial. Os dados usados são horários da estação meteorológica automática de Sorocaba, SP, entre 2000 e 2025. Eles vieram do conjunto METBRA25Y. A partir de janelas de 336 horas, os modelos preveem ao mesmo tempo todas as variáveis da entrada em três horizontes: 1 hora, 24 horas e 168 horas. Foram comparadas as redes recorrentes puras LSTM e GRU com as híbridas CNN-LSTM, CNN-GRU, CNN-LSTM-GRU e suas variantes bidirecionais. O ARIMA serviu como linha de base estatística. Todas as redes usaram a mesma configuração de camadas, unidades e regularização. Assim, a única diferença entre os grupos foi a presença do bloco convolucional. O desempenho foi medido por MAE, R^2, SMAPE e DTW, com cinco execuções por modelo. O desvio padrão entre elas serviu para separar diferença real de empate técnico. Os resultados mostram que a convolução não trouxe ganho expressivo. Nos horizontes de 1 hora e 168 horas, dominou o empate técnico. No horizonte de 24 horas, as recorrentes puras venceram em dez das dezesseis variáveis, sem nenhum ganho expressivo das híbridas. O trabalho indica que, sem estrutura espacial, acrescentar convolução a uma rede recorrente tende a adicionar parâmetros sem trazer informação nova.

Palavras-chave: previsão do tempo; séries temporais; aprendizado profundo; redes neurais recorrentes; redes neurais convolucionais, redes híbridas.

ABSTRACT

The forecasting of atmospheric variables using deep learning has frequently employed hybrid convolutional-recurrent networks. This stems from the assumption that the convolutional layer tends to improve performance. This work investigates whether this idea holds true when the problem is a single-point multivariate time series with no spatial structure. The data used consists of hourly records from the automatic weather station of Sorocaba, SP, between 2000 and 2025, sourced from the METBRA25Y dataset. Using 336-hour windows, the models simultaneously predict all input variables over three horizons: 1 hour, 24 hours, and 168 hours. Pure recurrent networks (LSTM and GRU) were compared against hybrid networks (CNN-LSTM, CNN-GRU, CNN-LSTM-GRU) and their bidirectional variants. ARIMA served as the statistical baseline. All networks used the same configuration of layers, units, and regularization. Thus, the only difference between the groups was the presence of the convolutional block. Performance was measured using MAE, R^2, SMAPE, and DTW, with five runs per model. The standard deviation between them was used to distinguish actual differences from technical ties. The results show that convolution did not bring significant gains. In the 1-hour and 168-hour horizons, technical ties prevailed. In the 24-hour horizon, pure recurrent networks outperformed in ten out of the sixteen variables, with no significant gain from the hybrid models. This study indicates that, without spatial structure, adding convolution to a recurrent network tends to add parameters without bringing new information.

Keywords: weather forecasting; time series; deep learning; recurrent neural networks; convolutional neural networks; hybrid networks.

LISTA DE ILUSTRAÇÕES

Figura 1 - Topologia de uma CNN 16
Figura 2 – Average Pooling 2x2 sendo aplicado em uma entrada 5x5 e captando os valores médios de cada quadrante. 17
Figura 3 - Célula LSTM 23
Figura 4 - Diagrama de uma célula GRU, destacando a ausência do estado da célula separado e as portas de Atualização e Redefinição 25
Figura 5 - Representação de uma LSTM BIDIRECIONAL 27
Figura 6 - Comparativo esquemático entre as duas topologias híbridas. À esquerda, a (blocos convolucionais em série alimentando a célula recorrente); à direita, a célula ConvLSTM 31
Figura 7 - Pipeline de treinamento e avaliação de Modelos que serão utilizados no trabalho 36
Figura 8 - Recorte de Precipitação, Radiação e Umidade no tempo em Sorocaba presentes no METBRA25Y 40
Figura 9 - Recorte de Pressão, Temperatura e Vento (separada em termos de COS e SEN) no tempo em Sorocaba presentes no METBRA25Y 41
Figura 10 - Janela deslizante com mascaramento 43
Figura 11- Topologia da LSTM utilizada nos experimentos 46
Figura 12 - Topologia da CNN - LSTM utilizada nos experimentos 49
Figura 13 - Curvas de predição em comparação com o real (INMET) - GRU - Radiação Global 57
Figura 14 - Curvas de predição em comparação com o real (INMET) - CNN-GRU - Radiação Global 58
Figura 15 - Curvas de predição em comparação com o real (INMET) - GRU - Temperatura do Bulbo 63
Figura 16 - Curvas de predição em comparação com o real (INMET) - CNN-GRU - Temperatura do Bulbo 64
Figura 17 - Curvas de predição em comparação com o real (INMET) - GRU – Velocidade do Vento 69
Figura 18 - Curvas de predição em comparação com o real (INMET) - CNN-GRU – Velocidade do Vento 70
Figura 19 - Curvas de predição em comparação com o real (INMET) - GRU - Precipitação 75
Figura 20 - Curvas de predição em comparação com o real (INMET) - CNN-GRU - Precipitação 76
Figura 21 - heatmap da comparação de R^2 da GRU com as redes híbridas 79
Figura 22 - heatmap da comparação de R^2 da GRU com as redes híbridas bidirecionais 79
Figura 23 - heatmap da comparação de MAE da GRU com as redes híbridas 80
Figura 24 - heatmap da comparação de MAE da GRU com as redes híbridas bidirecionais 81
Figura 25 - heatmap da comparação de R^2 e MAE da CNN-LSTM sem Average pooling com a CNN-LSTM com Average Pooling 83

LISTA DE TABELAS

Tabela 1 - Resultados das métricas para a Temperatura do Bulbo para o modelo ARIMA 51
Tabela 2- Resultados das métricas para a Radiação Global para o modelo ARIMA 52
Tabela 3 - Resultados das métricas para a Velocidade do vento para o modelo ARIMA.....52
Tabela 4 - Resultados das métricas para a Precipitação para o modelo ARIMA 53
Tabela 5 – Resultados das métricas entre arquiteturas para a variável de Radiação global no horizonte de 1 hora 54
Tabela 6 - Resultados das métricas entre arquiteturas para a variável de Radiação global no horizonte de 24 horas.....55
Tabela 7 - Resultados das métricas entre arquiteturas para a variável de Radiação global no horizonte de 168 horas.....56
Tabela 8 - Resultados das métricas entre arquiteturas para a variável de Temperatura do bulbo no horizonte de 1 hora 60
Tabela 9 - Resultados das métricas entre arquiteturas para a variável de Temperatura do bulbo no horizonte de 24 horas.....61
Tabela 10 - Resultados das métricas entre arquiteturas para a variável de Temperatura do bulbo no horizonte de 168 horas.....62
Tabela 11 - Resultados das métricas entre arquiteturas para a variável de Velocidade do vento no horizonte de 1 hora 66
Tabela 12 - Resultados das métricas entre arquiteturas para a variável de Velocidade do vento no horizonte de 24 horas.....67
Tabela 13 - Resultados das métricas entre arquiteturas para a variável de Velocidade do vento no horizonte de 168 horas.....68
Tabela 14 - Resultados das métricas entre arquiteturas para a variável de Precipitação no horizonte de 1 hora 72
Tabela 15 - Resultados das métricas entre arquiteturas para a variável de Precipitação no horizonte de 24 horas.....73
Tabela 16 - Resultados das métricas entre arquiteturas para a variável de Precipitação no horizonte de 168 horas.....74

SUMÁRIO

1 INTRODUÇÃO 11
2 TRABALHOS RELACIONADOS 13
2.1 Uso de Machine Learning para predição de eventos meteorológicos extremos 13
2.2 Redes Neurais Convolucionais 15
2.2.1 Topologia CNN 16
2.2.2 Treinamento 18
2.3 Contribuições e Limitações no Escopo Meteorológico 19
2.4 Redes neurais Recorrentes 20
2.4.1 LSTM 23
2.4.2 Gated Recurrent Unit (GRU) 25
2.4.3 LSTM Bidirecional 27
2.4.3.1 Topologia e modelagem matemática 27
2.4.3.2 Treinamento 27
2.5 Redes híbridas convolucional-recorrente 29
2.5.1 Vantagens da hibridização 31
2.5.2 Limitações das redes híbridas para séries temporais de ponto único: ConvLSTM e CNNLSTM 32
3 ABORDAGEM PROPOSTA 35
3.1 Fluxograma da metodologia proposta 36
4 Experimentos 38
4.1 Ambiente de desenvolvimento e bibliotecas utilizadas 38
4.2 Dataset 39
4.3 Pré-Processamento 42
4.4 Métricas 44
4.4.1 Erro Absoluto Médio (MAE) 44
4.4.2 Erro Percentual Absoluto (SMAPE) 44
4.4.3 Coeficiente de determinação (R^2) 45
4.4.4 Distância por Deformação Temporal Dinâmica (DTW) 45
4.5 Topologias 46
5 Resultados e Discussões 51
5.1 Análise Redes Recorrentes vs Redes híbridas 77
6 CONCLUSÃO 86
REFERÊNCIAS 88
APÊNDICE A – Resultados completos 93

1 INTRODUÇÃO

As mudanças climáticas não são mais uma previsão distante. Seus efeitos já se manifestam no cotidiano urbano. Segundo o IPCC (2023), o aquecimento global de origem antrópica já chegou a cerca de 1,1 °C em relação ao período pré-industrial. Cada aumento de 0,5 °C faz crescer, de forma clara, a frequência e a força do calor extremo, das tempestades e das secas.

No Brasil, esse cenário se traduz em números. Só em 2025, o Centro Nacional de Monitoramento e Alertas de Desastres Naturais registrou 1.493 casos hidrológicos, incluindo secas fortes, alagamentos, enxurradas, cheias e deslizamentos de terra. Esses eventos afetaram mais de 336 mil pessoas e causaram cerca de R$ 3,9 bilhões em prejuízos (CEMADEN, 2026). Os impactos não se restringem ao prejuízo financeiro. Eles também atingem a produção agrícola, a geração de energia, o transporte nas cidades e a saúde pública. E, em última instância, resultam em perdas humanas. Em muitos desses casos, o que separa um evento grave de uma tragédia é o tempo de antecedência com que ele é identificado. Esse tempo permite emitir alertas e retirar pessoas de áreas de risco.

Prever esses eventos, no entanto, é uma tarefa difícil. A atmosfera é um sistema não linear, de modo que pequenas diferenças nas condições iniciais podem levar a resultados muito diferentes em poucos dias, o que limita o tempo útil de qualquer previsão. Os modelos numéricos tradicionais lidam com esse comportamento resolvendo equações físicas. Porém, eles exigem muito processamento, o que abriu espaço para o uso de técnicas de aprendizado de máquina. Tais técnicas, contudo, apresentam limitações próprias. Como os extremos são raros, eles aparecem pouco no histórico usado para treinar os modelos. Assim, as previsões tendem a se aproximar da média e suavizam justamente os picos de maior interesse operacional (Materia et al., 2024). Além disso, o clima está mudando, então as estatísticas da série não permanecem estáveis ao longo dos anos. Esse fenômeno se chama *concept drift* e faz com que padrões aprendidos no passado percam parte da validade. As redes profundas também têm baixa interpretabilidade, o que dificulta a confiança nas previsões para uso prático (Zhang et al., 2025). Há ainda um aspecto menos discutido na literatura. Nem toda previsão meteorológica usa uma grade espacial, como imagens de satélite ou de radar. Em muitos casos, o problema efetivo consiste em prever a evolução de uma série temporal medida por uma única estação. Nessa situação, é preciso revisar a escolha da arquitetura.

Diante disso, este trabalho investiga uma prática comum na literatura recente. Trata-se do uso de redes híbridas convolucional-recorrentes, que combinam camadas convolucionais,

responsáveis por extrair padrões locais, com camadas recorrentes, que modelam a dependência temporal. Essa combinação constitui o estado da arte em problemas com forte componente de imagem. No entanto, assume-se com frequência que ela traz benefícios mesmo quando não há estrutura espacial a ser explorada. O objetivo é verificar, com evidência experimental, se a adição do bloco convolucional realmente melhora a previsão em uma série temporal multivariada de ponto único. Para isso, foi usada a série horária da estação meteorológica automática de Sorocaba, SP, no período de 2000 a 2025, obtida do conjunto METBRA25Y, derivado dos registros públicos do Instituto Nacional de Meteorologia (Lima Castro et al., 2026). A partir de janelas de 336 horas, equivalentes a 14 dias, os modelos preveem ao mesmo tempo todas as variáveis da entrada em três horizontes: 1 hora, 24 horas e 168 horas. Foram comparadas as arquiteturas recorrentes puras LSTM e GRU com as híbridas CNN-LSTM, CNN-GRU, CNN-LSTM-GRU e suas variantes bidirecionais, usando também o modelo estatístico ARIMA como linha de base. Todas mantêm a mesma configuração de camadas, unidades e regularização, assim a única diferença entre os grupos é a presença do bloco convolucional. A avaliação usa quatro métricas complementares, MAE, R^2, SMAPE e DTW, calculadas por variável e por horizonte, com cinco execuções de cada modelo. O desvio padrão entre elas serve para distinguir uma diferença real de um empate técnico.

O Capítulo 2 mostra os trabalhos relacionados e a base teórica, em que há a discussão sobre o uso de aprendizado de máquina para prever eventos meteorológicos extremos e suas limitações. Também trata da topologia e do treinamento das redes convolucionais e recorrentes, incluindo LSTM, GRU e variantes bidirecionais, e das redes híbridas convolucional-recorrentes. O Capítulo 3 explica a abordagem proposta, com os requisitos considerados e o fluxo de treinamento, avaliação e comparação dos modelos. O Capítulo 4 detalha os experimentos, incluindo o ambiente de desenvolvimento, o conjunto de dados, o pré-processamento, as métricas e as topologias usadas. O Capítulo 5 mostra os resultados, com a comparação entre redes recorrentes puras e híbridas por variável e por horizonte. Por fim, o Capítulo 6 reúne as conclusões do trabalho.

2 TRABALHOS RELACIONADOS

Este capítulo apresenta a fundamentação teórica da pesquisa. A Seção 2.1 revisa os principais estudos dos últimos cinco anos sobre *Machine Learning* na predição de eventos meteorológicos extremos e suas limitações. As Seções 2.2 e 2.3 detalham a evolução, topologia e treinamento das Redes Convolucionais (CNNs) e Recorrentes (RNNs, LSTM, GRU e Bidirecionais). Por fim, a Seção 2.5 explora as redes híbridas convolucional-recorrentes.

2.1 Uso de Machine Learning para predição de eventos meteorológicos extremos

Nos últimos anos, a previsão de eventos meteorológicos extremos passou por uma transformação significativa com a adoção de técnicas de *Machine Learning* (ML) e *Deep Learning* (DL). Tradicionalmente, a meteorologia tem dependido de Modelos de Previsão Numérica de Tempo (NWP), baseados em complexas equações físicas e termodinâmicas. Contudo, devido ao alto custo computacional e a dificuldade de modelar dinâmicas não lineares extremas em tempo real, os algoritmos de aprendizado de máquina emergiram como uma alternativa nova. Trabalhos de revisão recentes destacam que modelos de DL já conseguem rivalizar ou até superar os modelos físicos em previsões de médio prazo e na extração de padrões climáticos complexos (Fang *et al.*, 2021; Verma *et al.*, 2023; Olivetti & Messori, 2024).

Uma das principais vantagens observadas na literatura recente é a capacidade do ML de superar a escassez de dados meteorológicos estruturados, um problema encontrado em várias partes do mundo. Esses estudos demonstram que algoritmos de aprendizado de máquina são altamente eficazes para melhorar a precisão da previsão do tempo em regiões remotas e com baixo monitoramento histórico. Kagabo *et al.* (2024), por exemplo, comprovaram a utilidade dessas ferramentas na predição de chuvas extremas em Ruanda, enfrentando desafios de dados esparsos de forma semelhante à abordagem de Yadav *et al.* (2025), que propuseram melhorias significativas nas previsões de regiões remotas da Índia. Em ambos os cenários, modelos de ML conseguiram mapear correlações nos dados disponíveis que modelos convencionais não capturavam.

Além disso, a evolução técnica permitiu que abordagens puramente estatísticas, como o modelo SARIMA (Kumari & Muthulakshmi, 2024), servissem de base para que *frameworks* mais avançados e sofisticados fossem desenvolvidos. Como os eventos extremos carregam alta incerteza, pesquisas atuais têm migrado de previsões determinísticas para abordagens probabilísticas. Um exemplo claro dessa transição é o *framework* proposto por Sattari et al.

(2025), que quantifica incertezas na previsão de eventos extremos diários, como inundações causadas por tempestades, oferecendo uma margem de confiabilidade para a tomada de decisão. Paralelamente, Jiang *et al.* (2022) ressaltam a importância de prever não apenas o quando, mas o onde esses eventos ocorrerão, demonstrando a necessidade de algoritmos capazes de lidar simultaneamente com dimensões espaciais e temporais.

A relevância da aplicação de ML nesses cenários transcende a própria meteorologia, tornando-se uma ferramenta importante para a mitigação de impactos em infraestruturas e na saúde pública. A previsão de eventos meteorológicos extremos tem sido acoplada a sistemas de monitoramento para garantir a segurança de navegações marítimas (Rawson *et al.*, 2021) e para estimar falhas e apagões em redes de energia sustentável durante tempestades (AlHaddad et al., 2023). No campo da saúde, ferramentas preditivas baseadas em dados climáticos são fundamentais para antecipar surtos de doenças e riscos sanitários atrelados a eventos extremos sensíveis ao clima (Ssebyala *et al.*, 2024).

Apesar do notável avanço e das diversas aplicações bem-sucedidas, o uso de ML na predição de eventos extremos ainda enfrenta desvantagens e desafios estruturais significativos. Uma das limitações mais críticas apontadas na literatura é o problema do desbalanceamento de dados e a dificuldade dos modelos em generalizar ocorrências raras. Como os extremos climáticos são, por definição, eventos de baixa probabilidade, os algoritmos tendem a subestimar sua intensidade, retornando previsões suavizadas e mais próximas da média histórica (Materia *et al.*, 2024).

Além disso, a ausência de restrições físicas no treinamento das redes gera o problema da "caixa-preta" (*black-box*). Conforme alertado por Zhang *et al.* (2025), a falta de interpretabilidade dos modelos de Deep Learning dificulta a confiança operacional por parte dos meteorologistas, uma vez que as redes podem acertar uma previsão sem respeitar princípios fundamentais de conservação de massa ou energia da atmosfera.

Para contornar parte dessas limitações, especialmente no que tange à compreensão de padrões bidimensionais de sistemas meteorológicos (como o mapeamento de frentes frias ou furacões em imagens de satélite), a comunidade científica passou a explorar topologias especializadas na extração de características espaciais (Elabd *et al.*, 2025). É nesse contexto de necessidade de interpretação do "espaço" climático que as Redes Neurais Convolucionais (CNNs) ganham importância. Cabe ressalvar, todavia, que a estrutura do dado de entrada varia conforme a tarefa: nem toda predição meteorológica dispõe de uma grade espacial, e parte expressiva dos problemas, em especial a previsão a partir de séries temporais de estações isoladas, é de natureza essencialmente sequencial (Zhang *et al.*, 2025). Como se detalhará

adiante, isso condiciona quais arquiteturas, e em que configuração, são de fato adequadas a cada cenário.

2.2 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNNs) representam uma classe de arquiteturas de Deep Learning projetada especificamente para o processamento de dados dispostos em estruturas de grades bidimensionais ou multidimensionais, como imagens digitais ou matrizes de dados georreferenciados. Diferentemente das *Redes Neurais Perceptron Multicamadas* (MLP) tradicionais, as quais colapsam a estrutura espacial do dado ao transformá-lo em um vetor unidimensional, as CNNs preservam as relações de adjacência e a topologia local das variáveis de entrada por meio de conexões locais e compartilhamento de pesos.

Historicamente, as bases das CNNs fundamentam-se nos estudos biológicos do córtex visual de mamíferos conduzidos por Hubel e Wiesel nas décadas de 1950 e 1960, que identificaram células responsivas a pequenos campos receptivos locais. Contudo, o marco pioneiro computacional ocorreu com o desenvolvimento do *Neocognitron* por Fukushima em 1980, e posteriormente, com a consolidação da arquitetura LeNet-5 por Yann LeCun em 1998 (Jiang *et al.*, 2022). Projetada para o reconhecimento de caracteres manuscritos, a LeNet-5 introduziu formalmente a combinação de camadas de convolução, funções de ativação não lineares e subamostragem (*pooling*).

Apesar do sucesso inicial, o uso de redes convolucionais profundas permaneceu limitado pelo poder computacional e pela escassez de dados até o ano de 2012, quando a arquitetura AlexNet (Krizhevsky *et al.*, 2012) venceu a competição ImageNet. A AlexNet demonstrou a viabilidade de empilhar múltiplas camadas profundas utilizando unidades de ativação linear retificada (ReLU) e processamento paralelo massivo em Unidades de Processamento Gráfico (GPUs). Posteriormente, a evolução das CNNs foi marcada pela introdução de arquiteturas como a VGGNet, que padronizou filtros menores 3×3, e o divisor de águas ResNet (He *et al.*, 2015), que implementou conexões residuais para mitigar o problema do desvanecimento do gradiente em redes com mais de uma centena de camadas.

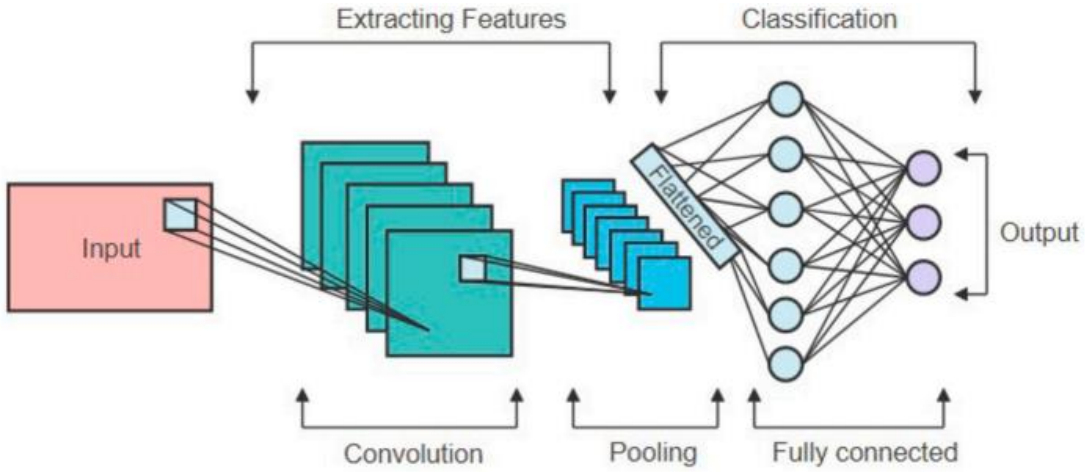
No escopo das ciências atmosféricas, a evolução das CNNs resultou na adaptação de redes de segmentação semântica, destacando-se a arquitetura U-Net (proposta por Ronneberger *et al.*, 2015). A estrutura simétrica em formato de "U", composta por uma sub-rede de contração (codificadora) e uma sub-rede de expansão (decodificadora), provou ser altamente eficaz para

tarefas de *nowcasting* e previsão meteorológica baseada em imagens de satélite e radares meteorológicos (Fang et al., 2021; Zhang et al., 2025).

2.2.1 Topologia CNN

A arquitetura típica de uma Rede Neural Convolucional baseia-se no empilhamento sequencial de três tipos principais de camadas: camadas convolucionais, camadas de subamostragem (*pooling*) e camadas totalmente conectadas (Fully Connected) (Ismail *et al., 2023*), como é mostrado na Figura 1 .

Figura 1 - Topologia de uma CNN



Fonte: *Ismail et al., 2023*

A Camada Convolucional consiste no núcleo computacional da rede. Nela, um conjunto de filtros convolucionais ajustáveis denominados kernels desliza sobre a matriz de entrada, realizando uma operação matemática de produto escalar local. Matematicamente, a operação de convolução discreta bidimensional para um elemento em uma posição (i, j) de um mapa de características (feature map) é expressa pela Equação (1):

$$S(i,j) = (I * K)(i,j) = \sum\sum I(i-m, j-n) \cdot K(m,n) \quad (1)$$

Onde I denota a matriz de entrada, K representa o filtro bidimensional de dimensões $m \times n$, e $S(i,j)$ constitui a saída convolucionada resultante (Rafi *et al.*, 2021).

O comportamento geométrico dessa operação é regulado por três hiperparâmetros essenciais:

***Stride* (passo):** define o deslocamento do filtro sobre a entrada.

***Padding* (preenchimento):** técnica de adicionar bordas artificiais, sendo geralmente zeros, para preservar as dimensões espaciais da matriz.

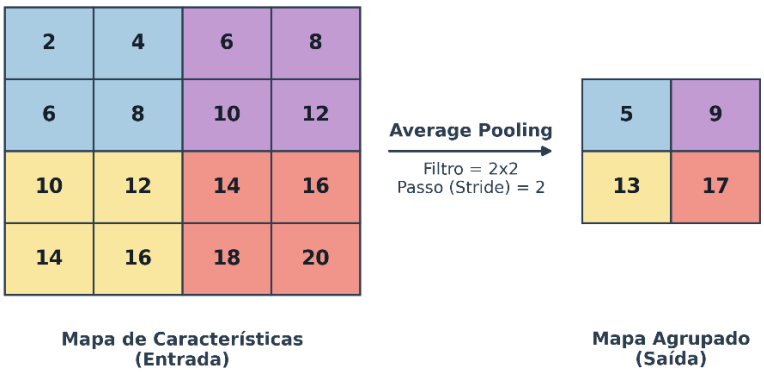
Número de filtros: determina a quantidade de características distintas extraídas simultaneamente.

A saída da convolução é submetida a uma função de ativação não linear. A unidade mais utilizada é a *Rectified Linear Unit* (ReLU), definida pela Equação (2), que zera valores negativos e introduz a não linearidade necessária para o aprendizado de padrões complexos sem saturar o gradiente (AlHaddad *et al.*, 2023):

$$f(x) = max(0, x) \quad (2)$$

A Camada de *Pooling*: Posicionada geralmente após as camadas convolucionais, tem como função reduzir as dimensões espaciais dos mapas de características, diminuindo a quantidade de parâmetros computacionais (Zha *et a*l., 2022) e conferindo invariância rotacional e translacional local à rede. A técnica mais difundida é o *Max-Pooling*, que extrai apenas o valor máximo contido em uma janela deslizante predefinida (Zha *et al*., 2022). Em contrapartida, o *Average Pooling* calcula a média dos valores presentes na região analisada, sendo uma alternativa quando o objetivo é suavizar as informações do mapa de características em vez de destacar apenas as ativações mais fortes (Zha *et al.*, 2022).

Figura 2 – *Average Pooling* 2x2 sendo aplicado em uma entrada 5x5 e captando os valores médios de cada quadrante.



Fonte: Autoria própria

Camadas Totalmente Conectadas: Localizadas na porção terminal da topologia, recebem os mapas de características achatados (*flattened*) vindos das convoluções anteriores (Rafi *et al.*, 2021) e realizam a combinação global dessas informações para computar o

resultado final do modelo, seja uma classificação ou um valor de regressão contínuo (Rafi *et al.*, 2021).

2.2.2 Treinamento

O ciclo operacional de uma CNN divide-se nas etapas complementares de treinamento (ajuste de pesos) e inferência.

O treinamento fundamenta-se no paradigma do aprendizado supervisionado e opera por meio do algoritmo de retropropagação do erro (*Backpropagation*) (Elabd *et al.*, 2025). Inicialmente, na etapa de propagação direta (*Forward Propagation*), os dados meteorológicos históricos atravessam a rede, gerando uma previsão (Elabd *et al.*, 2025). Essa previsão é confrontada com o dado real medido através de uma função de perda (*Loss Function*). Para problemas de modelagem climática contínua, adota-se predominantemente o Erro Quadrático Médio (*Mean Squared Error - MSE*) (Elabd *et al.*, 2025), conforme expresso na Equação (3):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(yi - \hat{y}i)^2 \quad (3)$$

Onde yi representa o valor meteorológico observado em uma determinada coordenada, $\hat{y}i$ denota a predição gerada pela CNN, e N constitui o número total de amostras ou pontos da grade espacial.

Uma vez calculado o erro, inicia-se a retropropagação. Utilizando a regra da cadeia do cálculo diferencial, os gradientes da função de perda em relação a cada peso e viés (*bias*) da rede são calculados de trás para frente (Elabd *et al.*, 2025). Os pesos dos filtros convolucionais são então atualizados iterativamente utilizando algoritmos de descida do gradiente, como o otimizador *Adam* (Adaptive Moment Estimation) (Elabd *et al.*, 2025).

A inferência, por sua vez, ocorre após a convergência do modelo e o congelamento definitivo dos parâmetros (pesos e vieses). Nesta fase, novos dados de entrada realizam uma única passagem direta (*forward*) pela arquitetura estabilizada. Como não há cálculo de derivadas nem atualizações de pesos nesta etapa, a inferência é capaz de prever novas instâncias em frações de segundo, a exemplo do *framework* de Elabd *et al.* (2025), que registra tempos de predição inferiores a 0,5 segundos por instância.

2.3 Contribuições e Limitações no Escopo Meteorológico

A inserção das CNNs no domínio da previsão de eventos meteorológicos extremos trouxe contribuições revolucionárias, mas também evidenciou limitações inerentes à sua própria formulação topológica.

A principal vantagem das CNNs no contexto meteorológico reside na sua capacidade de extração automatizada de características espaciais complexas. Fenômenos atmosféricos extremos, tais como complexos convectivos de mesoescala, frentes frias e ciclones tropicais, possuem assinaturas espaciais geométricas bem definidas em imagens de satélite e campos de refletividade de radar.

Antes do advento das CNNs, a identificação e o rastreamento desses sistemas dependiam de limiares manuais estipulados por meteorologistas ou de algoritmos de visão computacional clássica rígidos. As CNNs permitiram que os modelos computacionais aprendessem hierarquias de feições visuais de forma autônoma: as primeiras camadas aprendem bordas e gradientes de pressão; as camadas intermediárias identificam curvaturas e vorticidades; e as camadas profundas reconhecem a estrutura completa de uma tempestade severa (Jiang *et al.*, 2022). Isso reduziu drasticamente o tempo de processamento em comparação com os modelos físicos tradicionais e melhorou consideravelmente a acurácia na localização espacial do impacto do evento (Jiang *et al.*, 2022; Verma *et al.*, 2023).

Apesar do ganho expressivo na análise espacial, as CNNs tradicionais apresentam uma defasagem crítica: elas são inerentemente "cegas" para a dinâmica temporal e para a continuidade cronológica dos fluidos atmosféricos. A atmosfera não se comporta como um conjunto de quadros estáticos e isolados; ela constitui um sistema caótico contínuo onde o estado futuro é diretamente ligado ao histórico recente de evolução temporal.

Ao utilizar uma CNN pura para predição de variáveis atmosféricas, o modelo trata os dados de entrada de forma sequencialmente independente. Como consequência, a rede falha em modelar fenômenos dinâmicos como a aceleração de massas de ar, ciclos de dissipação de energia e a direção vetorial do deslocamento de células de tempestade ao longo do tempo (Verma *et al.*, 2023; Materia *et al.*, 2024). Tentar embutir a dimensão temporal empilhando imagens em canais extras resulta em uma perda rápida de correlação de longo prazo, gerando previsões que suavizam as intensidades e criam trajetórias fisicamente inconsistentes.

Essa limitação diz respeito à convolução aplicada ao espaço. A operação convolucional não é, em si, intrinsecamente espacial. Aplicada de forma unidimensional sobre o eixo temporal, ela pode extrair padrões locais da própria série (rampas, picos e periodicidades de curto prazo).

2.4 Redes neurais Recorrentes

Como discutido na seção anterior, embora as CNNs apresentem um bom desempenho na extração de características espaciais de sistemas meteorológicos, elas tratam os dados de entrada como matrizes estáticas independentes (Jiang *et al.*, 2022). Contudo, a atmosfera é descrita como um sistema dinâmico, fluido e caótico em constante evolução, onde o estado futuro depende da sequência cronológica de estados passados (Verma *et al.*, 2023; Olivetti & Messori, 2024). Para modelar essa continuidade temporal e superar a incapacidade das CNNs tradicionais em capturar dependências sequenciais, as Redes Neurais Recorrentes (RNNs) são utilizadas para meteorologia computacional (Fang *et al.*, 2021; Zhang *et al.*, 2025).

As RNNs compõem uma classe de redes neurais artificiais projetada especificamente para o processamento de dados sequenciais, textos ou séries temporais (Goodfellow *et al.*, 2016). Historicamente, os fundamentos da recorrência em redes artificiais baseiam-se nos sistemas dinâmicos com memória associativa propostos por Hopfield (1982) e nas arquiteturas com retroalimentação sequencial de Jordan (1986). No entanto, o avanço decisivo no processamento de sequências temporais foi o desenvolvimento da arquitetura criada por Jeffrey Elman (1990), conhecida como Elman Network ou RNN simples. A inovação de Elman (1990) foi a introdução de uma camada de contexto interna capaz de armazenar o estado oculto do passo temporal anterior, permitindo que a rede mantivesse uma representação ativa do histórico dinâmico dos dados.

A diferença fundamental entre uma rede feedforward e uma rede recorrente está na presença de laços de retroalimentação em suas camadas ocultas (Elman, 1990; Chollet, 2017). Em vez de processar uma sequência temporal inteira de forma simultânea ou independente, a RNN itera sobre os elementos da série passo a passo, mantendo um vetor de estado interno que funciona como uma memória volátil (Goodfellow *et al.*, 2016; Fang *et al.*, 2021). Conforme Chollet (2017), a topologia de uma RNN pode ser representada matematicamente de duas formas equivalentes: em seu formato compacto, com o laço de realimentação sobre o próprio nó oculto, ou em seu formato desenrolado no tempo. Essa segunda representação evidencia que, a cada instante t, a camada oculta recebe duas fontes de informação: o vetor de entrada atual

x_t, que na modelagem climática, corresponde a variáveis como temperatura, umidade ou pressão no instante corrente, e o vetor de estado oculto do instante anterior h_{t-1}, responsável por carregar a síntese dos padrões observados nos passos anteriores da série (Fang *et al.*, 2021; Yadav *et al.*, 2025). O funcionamento interno da camada oculta de uma RNN clássica é dado pela Equação (4), que descreve a atualização do estado de memória (Goodfellow et al., 2016):

$$ht = f(W_{hh}h_{t-1} + W_{xh}x_t + b_h) \quad (4)$$

Onde $h_t$representa o novo estado oculto no instante t, h_{t-1} denota o estado herdado do instante anterior $t-1$ e x_t é o vetor de entrada no passo atual. As matrizes W_{hh} e W_{xh} correspondem aos pesos associados à memória passada e a entrada atual, respectivamente, sendo compartilhadas ao longo de todos os passos de tempo da sequência (Elman, 1990). O termo b_h é o vetor de viés (*bias*) e $f(\cdot)$ representa uma função de ativação não linear, sendo a tangente hiperbólica ($tanh$) a mais adotada na literatura clássica, por mapear os valores no intervalo entre -1 e 1 (Chollet, 2017).

A partir do mapeamento do estado oculto atual h_t, a rede projeta a sua saída preditiva y_t para o instante corrente, baseando-se na transformação linear descrita pela Equação (5).

$$yt = g(W_{hy}h_t + b_y) \quad (5)$$

Onde W_{hy} e b_y são os pesos e vieses da camada de saída, e $g(\cdot)$ é a função de ativação final, geralmente linear ou identidade em problemas de regressão contínua, como a estimativa de volume de precipitação (Kagabo *et al.*, 2024; Kumari & Muthulakshmi, 2024).

O ajuste das matrizes de pesos compartilhados (W_{xh}, W_{hh}, W_{hy}) é realizado pelo algoritmo BPTT.

No BPTT, a rede é virtualmente desenrolada ao longo de toda a janela temporal de treinamento (Goodfellow *et al.*, 2016). Durante o passo direto (*Forward Pass*), os dados meteorológicos históricos são processados sequencialmente, gerando previsões parciais e computando uma função de perda global, como o *Erro Quadrático Médio (MSE)*, avaliada contra as observações reais (Kagabo *et al.*, 2024; Sattari *et al.*, 2025). No passo de retropropagação (*Backward Pass*), os gradientes do erro são calculados pela regra da cadeia, propagando-se do último passo temporal ($t = T$) até o instante inicial ($t = 0$), acumulando as derivadas parciais para atualizar os parâmetros via otimizadores estocásticos como o Adam (Kingma & Ba, 2015; Alhadadd *et al.*, 2023).

A etapa de inferência ocorre de forma autorregressiva após o congelamento dos pesos (Chollet, 2017). Em sistemas de alerta meteorológico precoce, a rede processa as variáveis do instante atual (x_t) e gera uma previsão imediata (y_{t+1}); essa saída e o estado oculto resultante (h_t) são então realimentados como entradas para o próximo passo (x_{t+1}), viabilizando previsões contínuas de trajetórias meteorológicas em escala de milissegundos (Zhang et al., 2025; Olivetti & Messori, 2024).

As RNNs clássicas apresentam limitações quando aplicadas a séries temporais meteorológicas longas. Hochreiter (1991) e Bengio et al. (1994) documentaram extensivamente dois fenômenos que comprometem seu uso prático: o desvanecimento do gradiente (*vanishing gradient problem*) e a explosão do gradiente (*exploding gradient problem*).

Durante a execução do algoritmo BPTT em sequências longas, o cálculo do gradiente em relação aos primeiros passos temporais exige multiplicações sucessivas da matriz de pesos recorrentes W_{hh}. Quando os autovalores dessa matriz são numericamente inferiores a unidade, o gradiente decresce exponencialmente a medida que a janela temporal se expande, tendendo a zero e impedindo a rede de atualizar seus pesos para horizontes passados distantes (Bengio *et al.*, 1994; Hochreiter, 1991; Goodfellow *et al.*, 2016). O resultado prático é uma perda crônica de memória de longo prazo (Materia *et al.*, 2024).

No contexto das ciências climáticas, essa limitação inviabiliza o uso isolado de RNNs simples. Fenômenos extremos como secas prolongadas, ondas de calor ou a ciclogênese de furacões são governados por teleconexões atmosféricas e condições de contorno de grande escala estabelecidas dias ou semanas antes do surgimento do evento (Materia *et al.*, 2024; Sattari *et al.*, 2025). Uma RNN clássica tende a "esquecer" esses gatilhos históricos remotos, tornando-se excessivamente dependente das flutuações recentes da série (Verma *et al.*, 2023).

A superação dessa falha estrutural levou ao desenvolvimento de arquiteturas baseadas em mecanismos de controle de fluxo de informação como as redes de memória com comportas (*gated RNNs*), capazes de reter dependências de longo prazo de forma estável (Goodfellow *et al.*, 2016). Essa evolução justifica a introdução das arquiteturas de Longa Memória de Curto Prazo (LSTM), Unidades Recorrentes Fechadas (GRU) e suas variantes bidirecionais, detalhadas a seguir.

Para resolver o problema do desvanecimento do gradiente e permitir que a rede neural lembrasse de informações por longos períodos de tempo, os pesquisadores Hochreiter e Schmidhuber (1997) criaram uma nova arquitetura chamada *Long Short-Term Memory* (LSTM).

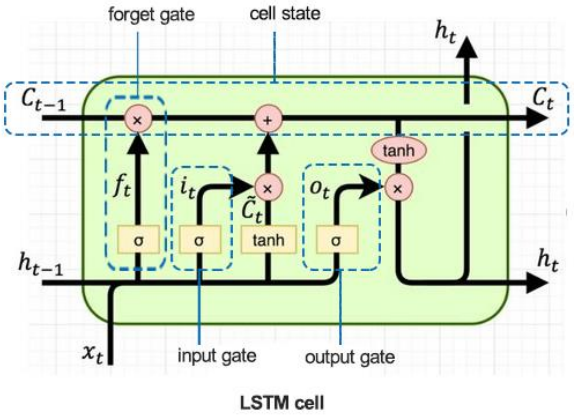
2.4.1 LSTM

A principal diferença entre a LSTM e uma RNN simples está em sua estrutura interna. Enquanto a RNN conta com apenas uma camada de processamento, a LSTM possui um mecanismo composto por quatro camadas que interagem entre si, formando o que se denomina célula de memória (Rafi *et al.*, 2021).

A diferença central da topologia LSTM é o estado da célula (C_t). Conforme Chollet (2017), o estado da célula funciona como uma correia transportadora que percorre toda a sequência de dados sofrendo apenas modificações pontuais, o que permite que informações climáticas de passos anteriores fluam pela rede sem serem descartadas.

Para controlar o que entra e o que sai dessa correia, a LSTM utiliza mecanismos reguladores chamados portas (*gates*). Conforme ilustrado no diagrama da Figura 3, uma célula LSTM possui três portas principais de entrada, de saída e de esquecimento (Kagabo *et al.*, 2024):

Figura 3 - Célula LSTM



Fonte: Rau (2021).

Porta de Esquecimento (*Forget Gate*): Essa porta analisa o estado anterior (h_{t-1}) e a entrada atual (x_t), produzindo um valor entre 0 e 1 por meio de uma função sigmoide (σ), onde 0 indica descarte total e 1 indica retenção completa. Matematicamente, é definida pela Equação (6) (Rafi *et al.*, 2021):

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (6)$$

Porta de Entrada (*Input Gate*): Essa porta determina quais informações novas serão incorporadas à memória da célula. O processo se divide em duas etapas, representadas pela

Equação (7): a porta sigmoide (i_t) seleciona quais valores serão atualizados, enquanto a tangente hiperbólica gera um vetor de candidatos ($\tilde{C}_t$) a serem adicionados à memória (Rafi e*t al.*, 2021):

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (7)$$

Com base no que foi esquecido e no que será adicionado, o estado da célula (C_t) é atualizado combinando as informações antigas com os novos candidatos, conforme a Equação (8):

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (8)$$

Porta de Saída (*Output Gate*): Por fim, a porta de saída define o que será entregue como saída no instante atual. O valor $h_t$corresponde a uma versão filtrada do estado da célula, controlada por uma função sigmoide (o_t), conforme as Equações (9) e (10):

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (9)$$

$$h_t = o_t * \tanh(C_t) \quad (10)$$

Nas equações acima, W representa as matrizes de pesos ajustadas durante o treinamento e b os vieses correspondentes (Goodfellow *et al.*, 2016).

O treinamento de uma LSTM segue a mesma lógica da Retropropagação Através do Tempo (BPTT) aplicada às RNNs. No entanto, graças ao mecanismo de portas, a LSTM não sofre com o problema do desaparecimento do gradiente, o que permite que o erro se propague até o início da sequência e que a rede aprenda relações de longo prazo (Kagabo *et al.*, 2024).

Essa capacidade tem impulsionado o uso de LSTMs em trabalhos recentes voltados à previsão de extremos climáticos. Um exemplo de destaque é o estudo de Sattari *et al*. (2025), no qual os autores empregaram um modelo profundo baseado em LSTM combinado com métodos probabilísticos para prever a vazão de rios durante tempestades extremas nos Estados Unidos. A arquitetura demonstrou capacidade de capturar a relação temporal entre o acúmulo de chuvas ao longo de vários dias e o risco de inundações graves, superando modelos físicos tradicionais em confiabilidade preditiva (Sattari *et al.*, 2025).

Apesar de seu excelente desempenho, a LSTM apresenta uma limitação prática: o elevado número de parâmetros decorrente de suas três portas implica maior custo computacional e tempo de treinamento. Para contornar essa limitação sem comprometer

expressivamente a qualidade das previsões, a *Gated Recurrent Unit* (GRU), aparece como uma opção mais simples e de treinamento mais rápido (Elabd *et al.*, 2025).

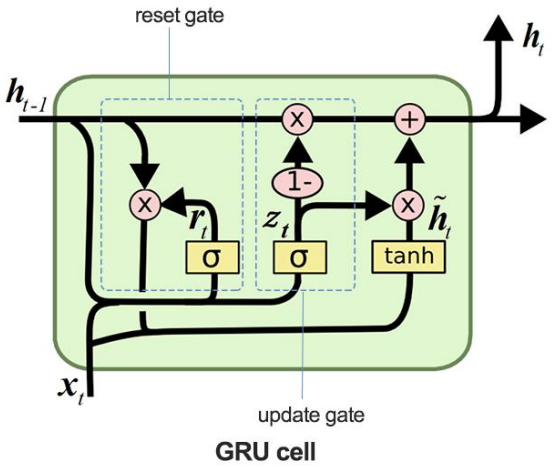
2.4.2 Gated Recurrent Unit (GRU)

Embora a LSTM tenha resolvido com eficácia o problema de esquecimento das redes recorrentes, a rede implica um custo computacional considerável. Por contar com três portas e dois estados de memória distintos, o modelo demanda maior capacidade de processamento e tempo de treinamento (Elabd *et al.*, 2025).

Para contornar essa limitação, uma arquitetura mais simples é utilizada, denominada como *Gated Recurrent Unit* (GRU) (Cho *et al.*, 2014) . A GRU alcança resultados competitivos aos da LSTM na modelagem de séries temporais, porém com uma estrutura interna simplificada.

A principal contribuição matemática da GRU reside na simplificação arquitetural em relação à LSTM, obtida pela fusão de componentes internos. Especificamente, as portas de esquecimento e de entrada da LSTM são consolidadas em uma única porta de atualização, enquanto o estado da célula e o estado oculto são unificados em um único canal de informação h_t (CHO *et al.*, 2014). Essa reestruturação, ilustrada na Figura 4, reduz substancialmente o custo computacional por passo temporal sem comprometer a capacidade de modelar dependências de longo prazo.

Figura 4 - Diagrama de uma célula GRU, destacando a ausência do estado da célula separado e as portas de Atualização e Redefinição



Fonte: Rau (2021).

O fluxo de informação na GRU é composto por duas portas:

Update Gate: o vetor z_t determina simultaneamente qual fração da memória passada será preservada e quanta informação nova será incorporada, exercendo a função combinada das portas de esquecimento e entrada da LSTM. Seu valor é calculado por uma função sigmoide sobre a concatenação do estado oculto anterior e da entrada corrente, conforme a Equação (11):

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \tag{11}$$

Reset Gate: o vetor r_t controla o quanto do contexto histórico deve ser suprimido antes da assimilação da nova entrada, permitindo que a rede ignore seletivamente informações passadas irrelevantes. É definido pela Equação (12):

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{12}$$

A partir desses vetores, a GRU computa um estado oculto candidato $\tilde{h}_t$ via função tangente hiperbólica. A porta de redefinição é aplicada diretamente sobre o estado anterior, de modo que apenas o contexto julgado relevante participa da formação da nova memória, como expresso na Equação (13):

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \tag{13}$$

O estado oculto final h_t, que acumula simultaneamente a função de memória da célula e de saída do passo temporal, é obtido por interpolação linear ponderada entre o estado anterior e o candidato, conforme a Equação (14):

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{14}$$

Essa equação evidencia o papel regulador da porta de atualização: quando $z_t \to 0$, o estado anterior é integralmente preservado; quando $z_t \to 1$, o estado candidato substitui completamente a memória corrente. Essa interpolação contínua confere à GRU a capacidade de manter dependências de longo prazo de forma seletiva e computacionalmente eficiente (Goodfellow *et al.*, 2016).

2.4.3 LSTM Bidirecional

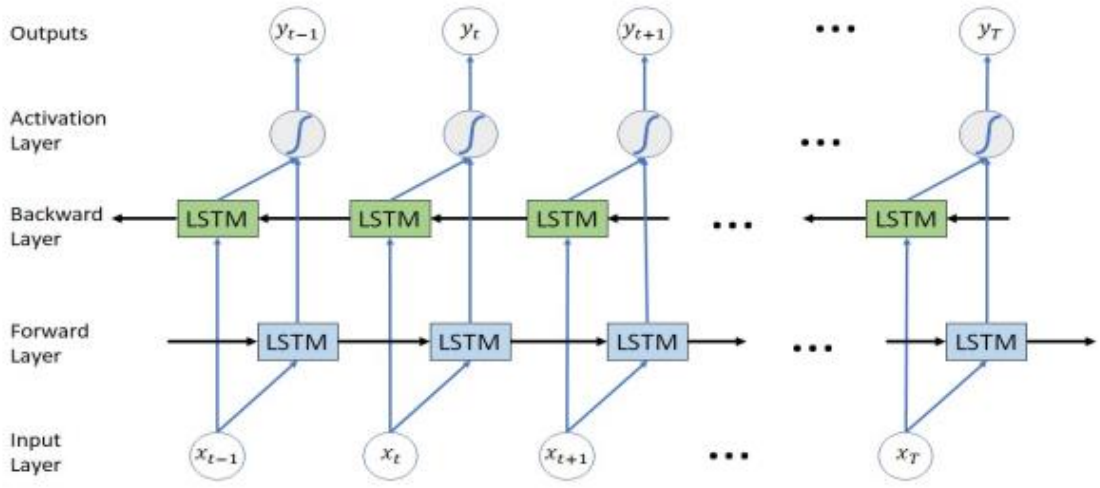
Como visto anteriormente, as arquiteturas recorrentes apresentadas fazem o processamento de uma série temporal de forma unidirecional, do primeiro ao último instante. Assim, o estado oculto codifica apenas o histórico de informações passadas.

Em problemas de modelagem atmosférica, contudo, a leitura isolada do passado costuma ser insuficiente para descrever fenômenos não-lineares. A compreensão completa da trajetória de uma anomalia climática, como a evolução de uma tempestade ou o ciclo de vida de um furacão, depende do contexto global da série, ou seja, das condições que tanto antecedem quanto sucedem o instante analisado (Schuster; Paliwal, 1997). Para suprir essa limitação, Schuster e Paliwal (1997) propuseram as Redes Neurais Recorrentes Bidirecionais (BiRNN), conceito posteriormente integrado às células de memória longa por Graves e Schmidhuber (2005), originando as arquiteturas *BiLSTM* e *BiGRU*.

2.4.3.1 Topologia e modelagem matemática

A principal característica das redes bidirecionais é a paralelização da camada oculta em duas sub-redes recorrentes independentes, que não compartilham pesos sinápticos e processam a mesma sequência de entrada em direções opostas (Schuster; Paliwal, 1997). A Figura 5 apresenta essa topologia desenrolada no tempo.

Figura 5 - Representação de uma LSTM BIDIRECIONAL



Fonte: (PRATIHER, 2018).

A camada de propagação direta processa a sequência na ordem cronológica usual ($t = 1 \rightarrow T$), produzindo um estado oculto direto $\vec{h}_t$. A Equação (15) apresenta uma formulação compacta dessa camada, na qual o operador $\mathcal{H}$ representa, de forma abstrata, todo o conjunto de operações internas da célula LSTM, sendo as portas de esquecimento, entrada e saída e a atualização do estado da célula:

$$\vec{h}_t = \mathcal{H}\left(W_{x\vec{h}}\, x_t + W_{\vec{h}\vec{h}}\, \vec{h}_{t-1} + b_{\vec{h}}\right) \tag{15}$$

Em que x_t é o vetor de entrada no instante t, $W_{x\vec{h}}$ é a matriz de pesos da entrada, $W_{\vec{h}\vec{h}}$ é a matriz de pesos recorrentes e $b_{\vec{h}}$ é o vetor de viés. Vale notar que a Equação (15) não substitui as equações das portas vistas anteriormente; ela apenas as encapsula em $\mathcal{H}$ para evidenciar o fluxo direcional da informação (Chollet, 2017).

Em paralelo, a segunda sub-rede percorre a mesma sequência $\mathcal{X}$ invertida no tempo ($t = T \rightarrow 1$), de modo que a rede tem acesso às informações "futuras" antes de calcular o estado no instante t. O estado oculto reverso $\overleftarrow{h}_t$ é dado pela Equação (16):

$$\overleftarrow{h}_t = \mathcal{H}\left(W_{x\overleftarrow{h}}\, x_t + W_{\overleftarrow{h}\overleftarrow{h}}\, \overleftarrow{h}_{t+1} + b_{\overleftarrow{h}}\right) \tag{16}$$

O ponto matematicamente relevante é a dependência recorrente do termo $\overleftarrow{h}_{t+1}$, o que quebra a causalidade estritamente unidirecional e permite incorporar o contexto posterior ao instante atual (Goodfellow *et al.*, 2016).

Concluído o mapeamento bidirecional, os estados $\vec{h}_t$ e $\overleftarrow{h}_t$são combinados para gerar a representação final do instante t. Embora a fusão possa ser feita por soma ou média, o procedimento mais comum na literatura é a concatenação dos dois vetores, produzindo um estado de dimensionalidade dupla (Graves; Schmidhuber, 2005). A saída preditiva y_t é então obtida pela Equação (17):

$$y_t = \sigma\left(W_{hy}\, [\vec{h}_t\, ;\, \overleftarrow{h}_t] + b_y\right) \tag{17}$$

onde $[\vec{h}_t\, ;\, \overleftarrow{h}_t]$denota o vetor concatenado dos estados direto e reverso, e W_{hy}é a matriz que o mapeia para a dimensão de saída.

2.4.3.2 Treinamento

O treinamento de uma BiLSTM emprega o algoritmo BPTT. Como as matrizes de pesos $W_{\overrightarrow{h}}$e $W_{\overleftarrow{h}}$são independentes, a atualização dos parâmetros de cada direção ocorre de forma isolada, com o gradiente da função de perda (por exemplo, o MSE) fluindo simultaneamente nos dois sentidos do tempo. É importante destacar que o ganho da arquitetura bidirecional está na riqueza do contexto disponível para cada instante, e não em um mecanismo adicional de combate ao desvanecimento do gradiente, esse papel continua sendo desempenhado pelos *gates* internos da célula LSTM.

No contexto da previsão de eventos meteorológicos extremos, esse acesso ampliado ao contexto pode favorecer o aprendizado de padrões em séries complexas e ruidosas. A vantagem de dispor do contexto completo da sequência foi demonstrada já em tarefas de rotulação sequencial, nas quais variantes bidirecionais superaram suas contrapartes unidirecionais (Graves; Schmidhuber, 2005). No domínio meteorológico, revisões recentes sobre aprendizado profundo aplicado a eventos extremos destacam a capacidade dessas arquiteturas de capturar as relações não-lineares características desses dados (Verma *et al.*, 2023).

Há, porém, uma restrição operacional. Como a camada reversa exige instantes futuros $(t + 1)$ para computar o instante atual (t), a *BiLSTM* pura não pode ser aplicada em sistemas de previsão em tempo real. Por isso, em tarefas de previsão meteorológica futura, essa topologia permanece restrita à etapa de pré-processamento de dados ou ao papel de componente em arquiteturas *Encoder-Decoder* (*Seq2Seq*), nas quais atua na codificação e na síntese de padrões históricos já observados (Graves; Schmidhuber, 2005).

2.5 Redes híbridas convolucional-recorrente

As seções anteriores evidenciaram uma complementaridade estrutural entre as duas grandes famílias de arquiteturas profundas aplicadas à previsão meteorológica. De um lado, as Redes Neurais Convolucionais (CNNs) destacam-se na extração automática de características espaciais, porém são "cegas" para a continuidade cronológica dos fluidos atmosféricos, tratando cada instante como um quadro estático e independente (Verma *et al.*, 2023; Materia *et al.*, 2024). De outro, as Redes Neurais Recorrentes, em especial a LSTM e a GRU, modelam com eficácia as dependências temporais de longo prazo, mas, por operarem sobre vetores unidimensionais, colapsam qualquer estrutura espacial presente no dado de entrada (Fang *et al.*, 2021; Chollet, 2017). A constatação dessas problemáticas motivou o desenvolvimento de arquiteturas híbridas

convolucional-recorrentes, que buscam reunir, em um único modelo treinável de ponta a ponta, a capacidade de percepção espacial da CNN e a memória sequencial da rede recorrente.

Na literatura, essa hibridização materializa-se em duas topologias fundamentalmente distintas, cuja diferenciação é crucial para a escolha adequada do modelo conforme a natureza do dado. A primeira, denominada *CNN-LSTM empilhada*, dispõe os módulos em série: uma sub-rede convolucional processa primeiro a entrada e extrai um vetor compacto de características; essa representação é então entregue à rede recorrente, que realiza a modelagem temporal sobre a sequência de vetores resultante (Zha *et al.*, 2022). Os trabalhos pioneiros dessa abordagem surgiram fora do domínio meteorológico, Donahue et al. (2017), com a arquitetura *Long-term Recurrent Convolutional Network* (LRCN), aplicada ao reconhecimento de ações e a descrição de vídeos, e Sainath et al. (2015), com a arquitetura CLDNN (Convolutional, LSTM, Deep Neural Network), que demonstrou no reconhecimento de fala que a combinação dos três tipos de rede explora a complementaridade entre redução de variância espectral (CNN), modelagem temporal (LSTM) e mapeamento discriminativo (camadas densas).

A segunda topologia, denominada *ConvLSTM*, promove uma fusão mais profunda entre os mecanismos: em vez de empilhar os módulos, ela substitui as multiplicações matriciais internas das portas da LSTM por operações de convolução, preservando a estrutura de grade espacial ao longo de todo o fluxo recorrente. Essa arquitetura foi proposta por Shi et al. (2015) precisamente no contexto meteorológico, o *nowcasting* de precipitação a partir de mapas de refletividade de radar e constitui o marco fundador das redes híbridas nas ciências atmosféricas. Formalmente, a célula ConvLSTM substitui o produto matricial das Equações (6) a (10) da LSTM clássica pelo operador de convolução ($*$), conforme as Equações (18) a (22), em que $\odot$ denota o produto de Hadamard:

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \odot C_{t-1} + b_i) \tag{18}$$

$$f_t = \sigma\big(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \odot C_{t-1} + b_f\big) \tag{19}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \tag{20}$$

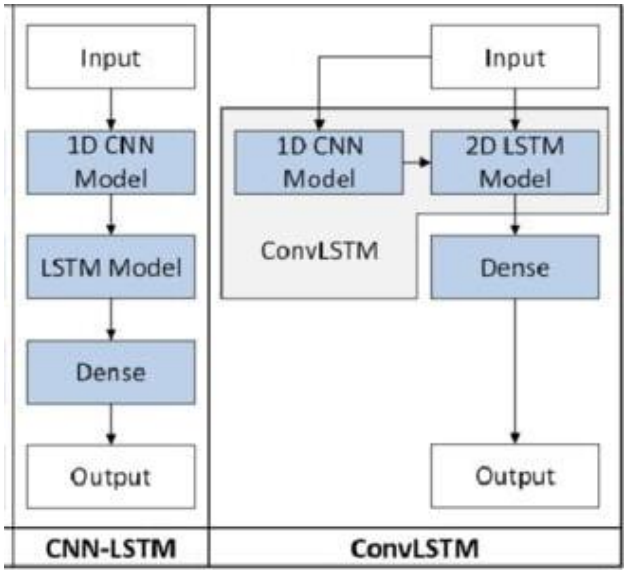
$$\sigma_t = \sigma(W_{x0} * X_t + W_{h0} * H_{t-1} + W_{c0} \odot C_t + b_o) \tag{21}$$

$$H_t = \sigma_t \odot \tanh(C_t) \tag{22}$$

Nas equações acima, X_t, H_t e C_t deixam de ser vetores e passam a ser tensores tridimensionais (altura × largura × canais), de modo que cada porta atua simultaneamente sobre

as dimensões espacial e temporal. Sob essa formulação, o treinamento das redes híbridas mantém o paradigma supervisionado já descrito, a propagação direta gera previsões que, confrontadas com as observações reais por meio de uma função de perda (como por exemplo o MSE, dada pela Equação 3), produzem gradientes propagados pela *BPTT*. Como a arquitetura é integralmente diferenciável, os pesos dos filtros convolucionais e das portas recorrentes são otimizados de forma conjunta por algoritmos como o Adam (Kigma & Ba, 2015). A inferência, por sua vez, ocorre após o congelamento dos pesos, com a rede operando de maneira autorregressiva, realimentando suas próprias previsões para horizontes superiores a um passo (Chollet, 2017). A Figura 6 mostra a diferença entre as 2 topologias de *CNN-LSTM.*

Figura 6 - Comparativo esquemático entre as duas topologias híbridas. À esquerda, a célula CNN-LSTM Empilhada (blocos convolucionais em série alimentando a célula recorrente). À direita, a célula ConvLSTM



Fonte: Widiputra e Juwono (2024).

2.5.1 Vantagens da hibridização

O ganho expressivo das redes híbridas manifesta-se principalmente em problemas espaço-temporais, isto é, quando a entrada possui estrutura espacial bidimensional que evolui no tempo. É o caso das imagens de satélite e dos campos de refletividade de radar, em que a ConvLSTM consegue rastrear simultaneamente a geometria de um sistema convectivo e o seu deslocamento vetorial ao longo dos instantes. Shi *et al.* (2015) demonstraram que a ConvLSTM captura correlações espaço-temporais de forma consistentemente superior a FC-LSTM e ao algoritmo operacional ROVER no *nowcasting* de chuva. Essa superioridade consolidou a

arquitetura como base para inúmeros sistemas posteriores de previsão imediata e, mais tarde, motivou sua combinação com topologias de segmentação como a U-Net, dando origem a modelos de previsão de precipitação a partir de dados orbitais (Fang *et al*., 2021).

De modo geral, portanto, a vantagem das redes híbridas é diretamente proporcional à quantidade de informação espacial explorável no dado de entrada, quanto mais o problema se aproxima de uma "imagem que se move no tempo", maior o benefício da convolução acoplada a recorrência. Em domínios sem essa estrutura, como a previsão de séries financeiras puramente univariadas, a literatura observa que o componente convolucional perde sua função primária de extração espacial e passa a operar apenas como um filtro de suavização e redução de dimensionalidade, reduzindo a margem de ganho sobre uma rede recorrente isolada.

2.5.2 Limitaçõe das redes híbridas para séries temporais de ponto único: ConvLSTM e CNNLSTM

A distinção topológica discutida antes tem efeito direto sobre a previsão a partir de séries temporais meteorológicas de ponto único, ou seja, dados de uma única estação de superfície, em que cada instante reúne medidas como temperatura, umidade e pressão, sem nenhuma grade espacial associada. Nesse caso, a *ConvLSTM*, que foi pensada para dados com estrutura espacial, deixa de fazer sentido. O que se usa, então, é a variante com convolução unidimensional (*1D-CNN-LSTM*): uma camada de convolução 1D percorre a janela de tempo da entrada e extrai padrões locais de cada variável (rampas, picos, repetições ao longo do dia), e a sequência resultante é entregue a LSTM, que aprende as dependências de mais longo prazo. O ponto importante é que, sem a estrutura espacial, o benefício da convolução deixa de ser garantido: ele passa a depender do horizonte de previsão e do tamanho da janela, e pode até ser nulo ou prejudicial em alguns casos.

A primeira evidência disso aparece na relação entre o ganho e o horizonte de previsão. Al-Omari, Yaghi e Alrifai (2026) compararam quatro modelos na previsão de precipitação, uma LSTM sozinha, uma CNN sozinha, o híbrido *CNN-LSTM* e um Transformer em vários horizontes. Eles observaram que a vantagem do *CNN-LSTM* sobre a LSTM pura só é significativa a partir de dois passos a frente para uma grade espacial pequena (3×6 pontos). No horizonte de um passo, todos os modelos empatam (R^2 entre 0,909 e 0,916). Mesmo assim, o resultado bate com o que se espera para a estratégia de múltiplos horizontes adotada aqui, na previsão de 1 h, o passado recente já explica bem o futuro próximo, e qualquer arquitetura dá conta, logo, a convolução pouco acrescenta. Já nos horizontes mais longos, é preciso reunir

padrões de curto prazo para projetar mais à frente, e aí a convolução pode, em certas condições, ajudar.

Além de ser apenas pequeno, o ganho da convolução às vezes desaparece de vez, há casos em que o componente convolucional piora o híbrido em relação aos modelos sozinhos. Em previsão financeira com uma única variável, Hoa, Le e Nguyen-Dinh (2025) mostraram que, na configuração com a janela mais longa, o híbrido *1D-CNN-LSTM* perde para a LSTM sozinha (R^2 0,888 contra 0,926), o contrário do que acontecia nas janelas curtas. Em outro estudo, Marinho et al. (2023) previram irradiância solar em horizontes curtos (de 5 a 30 minutos) e verificaram que, no geral, a *CNN-1D* e a LSTM isoladas foram melhores que o híbrido, com a melhor precisão sendo da *CNN-1D* pura. Em um caso específico (irradiância direta, horizonte de 20 minutos), o *CNN-LSTM* despencou: teve *forecast skill* de -73%, ou seja, ficou pior do que simplesmente repetir o último valor observado, enquanto os modelos sozinhos se mantiveram estáveis. São, portanto, dois exemplos, em áreas diferentes, em que juntar a convolução a recorrência não só não ajudou como atrapalhou.

Marinho et al. (2023) explicam que, naquele caso, os preditores carregavam pouca informação útil e a própria série de irradiância é fortemente autocorrelacionada, ou seja, o passado imediato já é o melhor previsor, e a convolução não tinha o que acrescentar. Não significa que a convolução não serve para previsão temporal, o que pode impactar negativamente é quando faltam as condições que a favorecem (como por exemplo um padrão local rico para extrair), o bloco convolucional só adiciona parâmetros e complexidade sem retorno, e, no pior caso, gera instabilidade.

Sobre o motivo de fundo, uma explicação comum na literatura aponta para as operações de redução de dimensão dentro do bloco convolucional, em especial o *pooling*. Li *et al.* (2025) afirmam de forma clara que operações como médias móveis e *pooling* descartam informações pequenas, mas importantes, ao comprimir a série, e por isso propõem a convolução dilatada como forma de evitar essa perda. É possível que esse mecanismo esteja por trás dos casos acima, afinal, os dois híbridos que falharam usam *pooling*. Portanto, a piora observada é compatível com a perda de informação por redução de dimensão descrita por Li *et al.* (2025).

Em resumo, as redes híbridas convolucional-recorrentes são eficazes para problemas com forte componente de imagem, como o *nowcasting* por satélite e radar, mas quando aplicadas a séries temporais de ponto único, elas nem sempre têm um ganho certo. A convolução pode tender a não fazer diferença no horizonte de 1 h, dominado pela autocorrelação recente, e seu benefício nos horizontes mais longos pode depender de haver padrões locais relevantes para extrair. Quando esses padrões não existem, o bloco

convolucional pode tanto não contribuir quanto piorar o desempenho em relação a uma arquitetura recorrente. Por isso, neste trabalho, a inclusão da convolução é tratada como uma hipótese a ser testada, e não como um ganho certo.

3 ABORDAGEM PROPOSTA

Conforme apresentado no Capítulo 2, embora as técnicas de ML sejam competitivas em relação aos modelos numéricos tradicionais para predição de eventos meteorológicos extremos, sua utilização ainda tem limitações recorrentes. A primeira limitação relevante é a suavização de extremos: por serem eventos de baixa probabilidade, o modelo tende a regredir à média histórica justamente nas intensidades de maior interesse operacional (Materia *et al.*, 2024). A segunda é a "caixa-preta" das redes profundas, na qual a ausência de restrições físicas explícitas e a baixa interpretabilidade dificultam a confiança operacional dos meteorologistas (Zhang *et al.*, 2025). Adicionalmente, tem-se o fato recorrente de que a predição de extremos depende de modelar adequadamente as dependências temporais da série (Jiang *et al.*, 2022). Essas limitações levaram a literatura a adotar arquiteturas cada vez mais complexas, entre elas as redes híbridas convolucional-recorrentes.

Diante disso, foi testada a adição da convolução às redes recorrentes na previsão de uma série temporal meteorológica de ponto único. Essa combinação se demonstra consolidada na literatura a partir de problemas com forte componente de imagem, como o *nowcasting* por satélite e radar, em que a convolução tem uma estrutura espacial clara para explorar. O mesmo não acontece em uma série de estação isolada, e por isso a inclusão do bloco convolucional é tratada aqui como uma hipótese a ser verificada, e não como um ganho garantido.

Para testar isso, todas as redes usaram a mesma configuração de camadas, unidades e regularização. Assim, a única diferença entre elas foi a presença do bloco convolucional. Cada modelo foi executado cinco vezes. O desvio padrão entre as execuções serviu para separar uma diferença real de um empate técnico. O ARIMA foi usado como linha de base estatística.

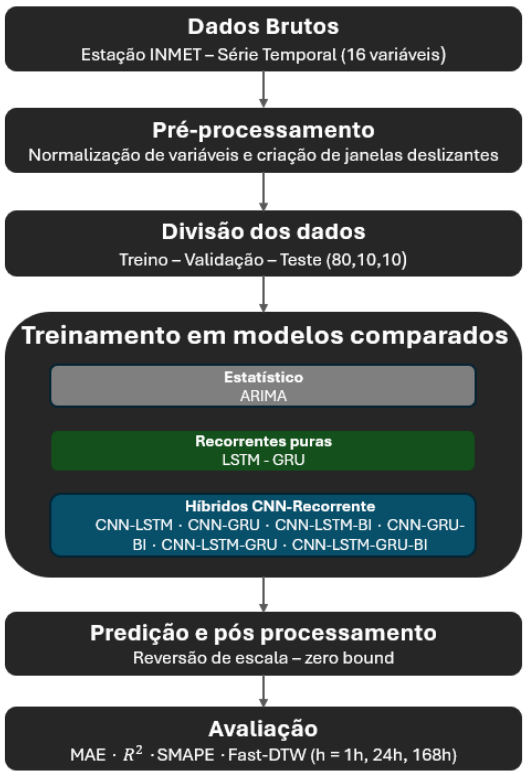
A avaliação foi feita em três horizontes: 1 hora, 24 horas e 168 horas. Isso porque o ganho da convolução pode mudar conforme a distância da previsão. Além disso, esses horizontes correspondem a janelas de decisão distintas: a janela de 1 hora é para responder imediatamente a eventos de evolução rápida, como enxurradas urbanas, e para gerir o trânsito durante o evento; a janela de 24 horas é para preparar serviços de saúde e proteger a infraestrutura urbana; a janela de 168 horas é para planejar a agricultura, onde decisões de plantio, irrigação e colheita dependem de antecipar condições adversas.

Já a suavização dos extremos não foi tratada com nenhum mecanismo específico de treinamento. Ela foi usada apenas como critério de análise, para verificar em quais variáveis os picos são atenuados.

3.1 Fluxograma da metodologia proposta

Para seguir com o treinamento, avaliação e comparação dos modelos de predição, foram seguidas as seguintes etapas mostradas na Figura 7:

Figura 7 - Pipeline de treinamento e avaliação de Modelos que foram utilizados no trabalho



Fonte: Autoria própria

A primeira etapa inicia-se com a importação e utilização do *dataset* do INMET em que há o conjunto de 16 variáveis climáticas tabuláveis da cidade de Sorocaba no período de 2000 a 2025, resultando em mais de 150 mil dados históricos. O vetor de entrada final tem 22 dimensões, sendo 16 variáveis físicas, as duas componentes seno/cosseno da direção do vento e quatro variáveis cíclicas (diária e anual). Os resultados são referentes as 16 variáveis físicas.

Após isso, foi de suma importância o pré-processamento destes dados, em que consiste na normalização dos dados e criação de variáveis cíclicas para ajudar na predição e captar padrões sazonais, além de criar as janelas deslizantes para a utilização de modelos recorrentes e híbridos.

Com os dados pré-processados, eles foram devidamente separados em treino (80%), validação (10%) e testes (10%), a fim de utilizar essa proporção na etapa de treinamento.

Após isso, foi estruturado o baseline de treinamento, que se baseia em três categorias: Estatístico, Recorrentes puras e Híbridos CNN-Recorrente. Para o modelo estático foi utilizado o ARIMA, já para as Recorrentes Puras foi utilizado LSTM e GRU, e por fim, nos modelos Híbridos CNN-Recorrentes foram treinadas as redes de CNN-LSTM, CNN-GRU, CNN-LISTM-BIDIRECIONAL, CNN-GRU-BIDIRECIONAL, CNN-LSTM-GRU e CNN-LSTM-GRU-BIDIRECIONAL.

Com o treinamento dessas arquiteturas, para cada modelo foram iniciados a predição e o pós processamento, em que consiste na rede fazer as predições para os horizontes de 1h, 24h e 168h utilizando múltiplas variáveis de entradas e resultando na predição de todas as variáveis que entraram na rede, sendo uma arquitetura de muitos para muitos.

Por fim, foram calculadas métricas de avaliação para cada rede para cada horizonte, que consiste na utilização das métricas de MAE, R^2, SMAPE, FAST-DTW, em que foram utilizados para comparação de cada modelo em cada horizonte, sendo assim possível entender qual modelo teve os melhores resultados em cada horizonte de predição de cada variável.

4 EXPERIMENTOS

Nesta seção, será abordado o Setup do experimento, *dataset* e seu pré-processamento, bem como as métricas utilizadas para avaliação de cada modelo. Além disso, serão apresentados os resultados de treinamento e predição de cada modelo, bem como seu comparativo.

4.1 Ambiente de desenvolvimento e bibliotecas utilizadas

O desenvolvimento e a execução dos experimentos foram realizados na plataforma Google Colab, que disponibiliza um ambiente de computação em nuvem com aceleração por GPU, no qual os notebooks foram executados utilizando a GPU A100 e armazenados via Google Drive. Os modelos foram desenvolvidos totalmente em Python, com utilização de bibliotecas focadas em redes neurais, normalização, visualização e manipulação de dados e cálculo de métricas.

A manipulação e a organização dos dados foram realizadas com as bibliotecas NumPy e pandas, visando facilitar as operações numéricas e o tratamento de tabelas, sendo especialmente úteis no processamento das séries temporais e no janelamento das observações. A normalização das variáveis foi feita com o MinMaxScaler, da biblioteca Sikit-learn, que reescala os dados para um intervalo comum antes da modelagem, evitando que variáveis de magnitudes distintas dominem o aprendizado devido à alta sensibilidade dos modelos de predição a dados com diferentes escalas.

As redes recorrentes LSTM e GRU e as arquiteturas híbridas Convolucional-recorrentes foram implementadas com TensorFlow e Keras, que oferecem ferramentas para construção e treinamento de modelos preditivos. Já para o ARIMA foi utilizada a biblioteca Statsmodels, sendo tradicional na análise de séries temporais.

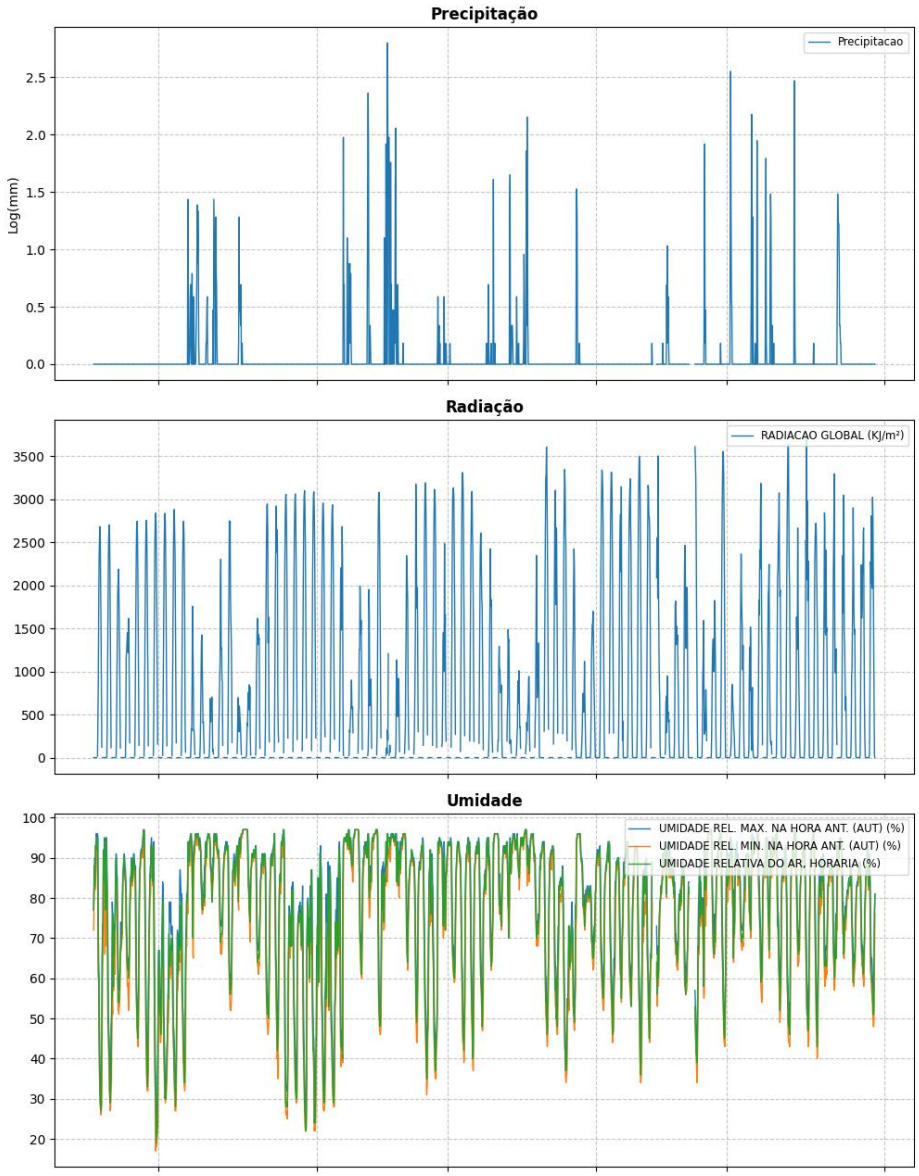
Para a avaliação dos modelos, além da biblioteca Sikit-learn, foram utilizadas as bibliotecas Tslearn e Fastdtw para o cálculo da distância baseada em *Dynamic Time Warping* (DTW), no qual avalia a similaridade de forma entre a série prevista e a observada. Por fim, a visualização dos resultados e das séries temporais foi feita com a biblioteca matplotlib, empregada na geração dos gráficos de previsão e das análises comparativas.

4.2 Dataset

Os dados utilizados neste trabalho são de origem do *dataset* do METBRA25Y, um arquivo de observações meteorológicas horárias de superfície do Brasil, derivado dos registros históricos públicos do Instituto Nacional de Meteorologia (INMET) (Lima Castro *et al.*,2026). O METBRA25Y foi realizado para dar suporte a estudos que exigem séries temporais meteorológicas em nível de estação, oferecendo variáveis padronizadas e metadados explícitos de controle de qualidade. O arquivo na sua versão 1.0.0, cobre o período de 2000 a 2025, com cobertura temporal específica por estação, em que reúne 616 códigos de estação únicos, dos quais 605 possuem coordenadas dentro de um envelope de plausibilidade do território brasileiro (Lima Castro *et al.*,2026).

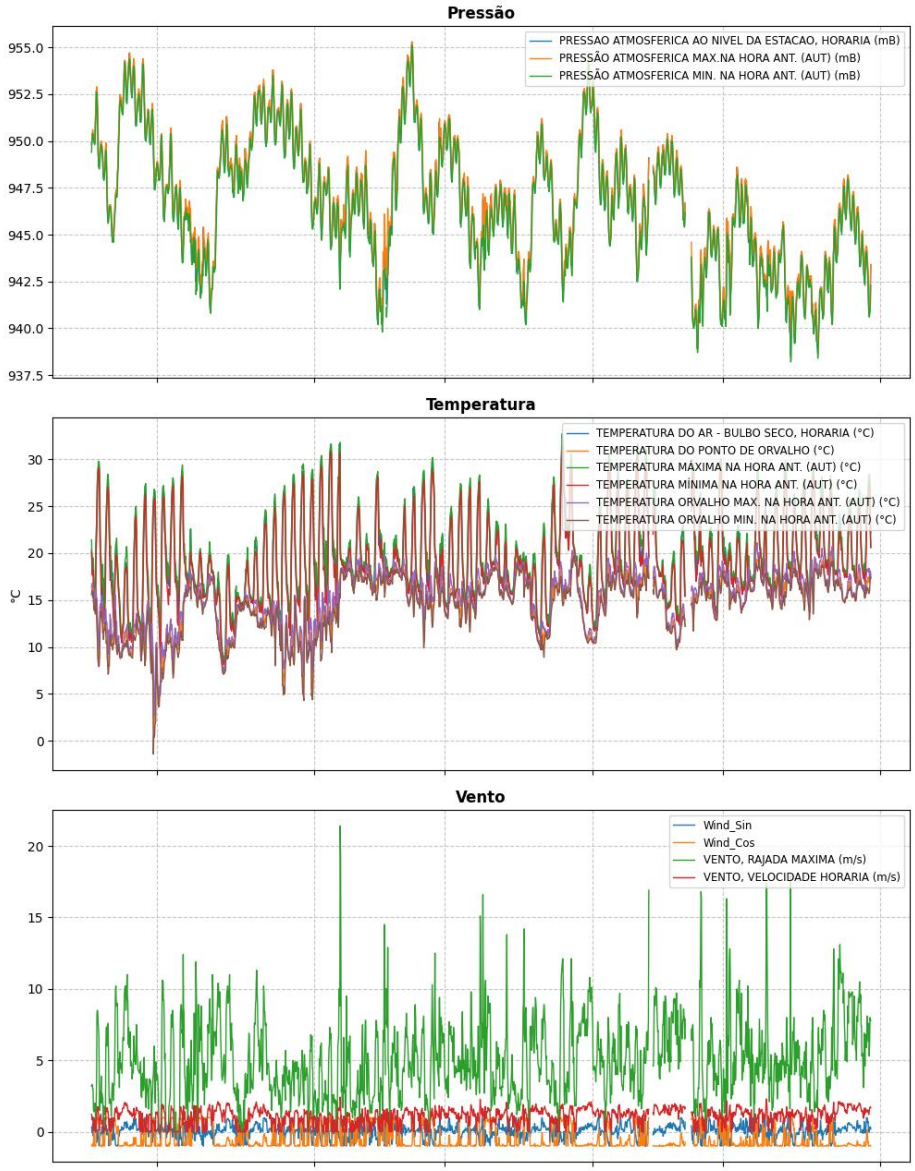
O conjunto disponibiliza as principais variáveis meteorológicas de superfície: precipitação, temperatura do ar (bulbo seco), temperatura do ponto de orvalho, umidade relativa, pressão atmosférica, velocidade do vento, rajada de vento, direção do vento e radiação solar global, além dos respectivos extremos horários de temperatura, umidade e pressão (Lima Castro *et al.*,2026), como evidenciado nas Figuras 8 e 9, as categorias das variáveis presente no *dataset* em um recorte no tempo para a cidade de Sorocaba.

Figura 8 - Recorte de Precipitação, Radiação e Umidade no tempo em Sorocaba presentes no METBRA25Y



Fonte: Autoria própria

Figura 9 - Recorte de Pressão, Temperatura e Vento (separada em termos de COS e SEN) no tempo em Sorocaba presentes no METBRA25Y



Fonte: Autoria própria

Um ponto central desse *dataset* é seu controle de qualidade em duas etapas. Na primeira, valores fisicamente impossíveis são transformados em valores ausentes (NaN) e marcados. Na segunda, verificações de consistência apenas sinalizam observações suspeitas, sem apagá-las, deixando a decisão de como tratá-las a cargo de quem for utilizar (Lima Castro *et al.*,2026).

Para este trabalho, selecionou-se a série da estação meteorológica automática localizada em Sorocaba - SP, cobrindo o período de 2000 a 2025 em resolução horária, no qual essa escolha caracteriza o problema como a previsão de uma série temporal multivariada de ponto único.

4.3 Pré-Processamento

Foram necessárias etapas adicionais de pré-processamento para adequar a série ao formato exigido pelos modelos de previsão. Todas as etapas buscaram preservar a sequência temporal dos dados e tratar corretamente as particularidades de cada variável.

Inicialmente, removeram-se as colunas de metadados não preditivas (identificação e coordenadas da estação) e garantiu-se que todas as variáveis fossem numéricas. Em seguida, a série foi reamostrada para uma frequência horária contínua; com isso, os instantes ausentes passam a existir explicitamente como valores nulos (NaN), em vez de serem omitidos , a fim de não criar quebras artificiais nas janelas temporais. Essa preservação dos vazios é coerente com o próprio METBRA25Y, que sinaliza, mas não descarta silenciosamente, observações ausentes.

Foram então aplicados tratamentos específicos por variável. A radiação solar global, ausente no período noturno, teve seus valores nulos noturnos preenchidos com zero, correção fisicamente coerente, pois não há radiação solar à noite. Esse tratamento é particularmente relevante porque a radiação é a variável com maior número de horas ausentes ou inválidas no METBRA25Y, conforme documentado por Lima Castro *et al.* (2026). As variáveis de precipitação e velocidade do vento, de distribuição fortemente assimétrica, foram transformadas pela função log(1+x), que comprime a cauda da distribuição e estabiliza a variância. A direção do vento, por ser uma grandeza angular na qual 0° e 360° representam o mesmo sentido, foi decomposta em suas componentes seno e cosseno, eliminando a descontinuidade artificial na fronteira dos 360°. Por fim, acrescentaram-se variáveis cíclicas derivadas do próprio instante de tempo, seno e cosseno dos ciclos diário (período de 24 h) e anual (período de 365,25 dias), que fornecem ao modelo referências explícitas de sazonalidade e, por serem calculadas diretamente da data e hora, nunca apresentam valores ausentes.

A divisão dos dados seguiu a ordem cronológica, sem embaralhamento, reservando 80% do período (mais antigo) para treino, 10% para validação e 10% para teste (mais recente). A separação temporal é essencial em previsão de séries, já que o embaralhamento permitiria ao modelo acessar informação futura durante o treino, inflando artificialmente o desempenho. A normalização foi realizada com o MinMaxScaler, no intervalo [0, 1], ajustado exclusivamente sobre o conjunto de treino e depois aplicado à validação e ao teste, de modo a impedir o vazamento de informação (data leakage) dos períodos futuros para a etapa de escalonamento.

A entrada dos modelos é construída por janelamento deslizante, cada amostra corresponde a uma janela de 336 horas consecutivas (14 dias) utilizada para prever o instante

situado no horizonte-alvo. Para lidar com as falhas reais da série sem contaminar o aprendizado, adotou-se uma estratégia de mascaramento em duas frentes. Os valores ausentes na entrada da janela recebem um preenchimento temporário: repete-se o último valor válido (*forward* e *backward fill*), apenas para permitir o cálculo. Ainda assim, duas situações levam ao descarte da janela inteira: quando o alvo a ser previsto cai sobre um valor que era originalmente ausente, ou quando mais da metade da entrada é composta por valores preenchidos artificialmente. Assim, o modelo nunca é treinado nem testado contra alvos que não existem, e janelas em sua maior parte artificiais não distorcem o resultado. A Figura 10 exemplifica o funcionamento da janela deslizante com a proteção para falhas reais sem contaminar o aprendizado.

Figura 10 - Janela deslizante com mascaramento



Fonte: Autoria própria

A definição do comprimento da janela de entrada partiu de uma necessidade prática: para prever o horizonte mais longo, de sete dias, a janela precisa conter pelo menos um ciclo semanal completo, para que o modelo tenha o padrão que deseja prever. Janelas como 24, 72 e

168 horas foram testadas e apresentaram piora progressiva nos horizontes de 24 e 168 horas, por não terem contexto suficiente para os ciclos mais longos. Janelas maiores que 336 horas, por outro lado, elevaram o uso de memória e o tempo de treinamento sem ganho proporcional nas métricas, além de diminuírem o número de amostras válidas após o mascaramento. A janela de 336 horas, que equivale a 14 dias, foi escolhida porque representa o melhor equilíbrio encontrado: ela contém dois ciclos semanais completos, abrange com folga o horizonte máximo de previsão e mantém o custo computacional dentro de limites aceitáveis.

4.4 Métricas

O desempenho dos modelos foi avaliado por quatro métricas complementares, calculadas separadamente para cada variável e cada horizonte de previsão, sempre sobre os valores já revertidos à escala física original. A escolha de métricas de naturezas distintas é deliberada: uma previsão pode ser boa em um aspecto e fraca em outro, e cada métrica ilumina uma faceta diferente da qualidade do resultado, erro de magnitude, erro relativo, ajuste de variância e erro de forma.

4.4.1 Erro Absoluto Médio (MAE)

O Erro Absoluto Médio (MAE, *Mean Absolute Error*) mede a diferença média, em valor absoluto, entre a previsão e a observação real, na unidade original de cada variável. É definido pela Equação (23):

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i | \qquad (23)$$

onde y_i é o valor observado, $\hat{y}_i$ é o valor previsto e n é o número de amostras.

4.4.2 Erro Percentual Absoluto (SMAPE)

O Erro Percentual Absoluto Médio Simétrico (*Symmetric Mean Absolute Percentage Error*) mede o erro em termos relativos (percentuais), o que permite comparar o desempenho entre variáveis de escalas diferentes. É definido pela Equação (24):

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{| \hat{y}_i - y_i |}{(| y_i | + | \hat{y}_i |)/2} \quad (24)$$

onde y_ié o valor observado, $\hat{y}_i$ é o valor previsto e n é o número de amostras.

4.4.3 Coeficiente de determinação (R^2)

O coeficiente de determinação (R^2) indica a fração da variância dos dados observados que o modelo consegue explicar. Assume valor máximo de 1 para o ajuste perfeito, aproxima-se de 0 quando o modelo não faz melhor do que prever sempre a média histórica, e pode tornar-se negativo quando a previsão é pior que essa média. É definido pela Equação (25):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (25)$$

onde $\bar{y}$ é a média dos valores observados.

4.4.4 Distância por Deformação Temporal Dinâmica (DTW)

A Distância por Deformação Temporal Dinâmica (DTW, *Dynamic Time Warping*) mede a similaridade de forma entre a série observada e a prevista, tolerando pequenos desalinhamentos no tempo, ou seja, penaliza menos uma previsão que acerta o padrão, mas está levemente deslocada. Para isso, busca, entre todos os alinhamentos possíveis entre as duas séries, aquele de menor custo total. É definida pela Equação (26):

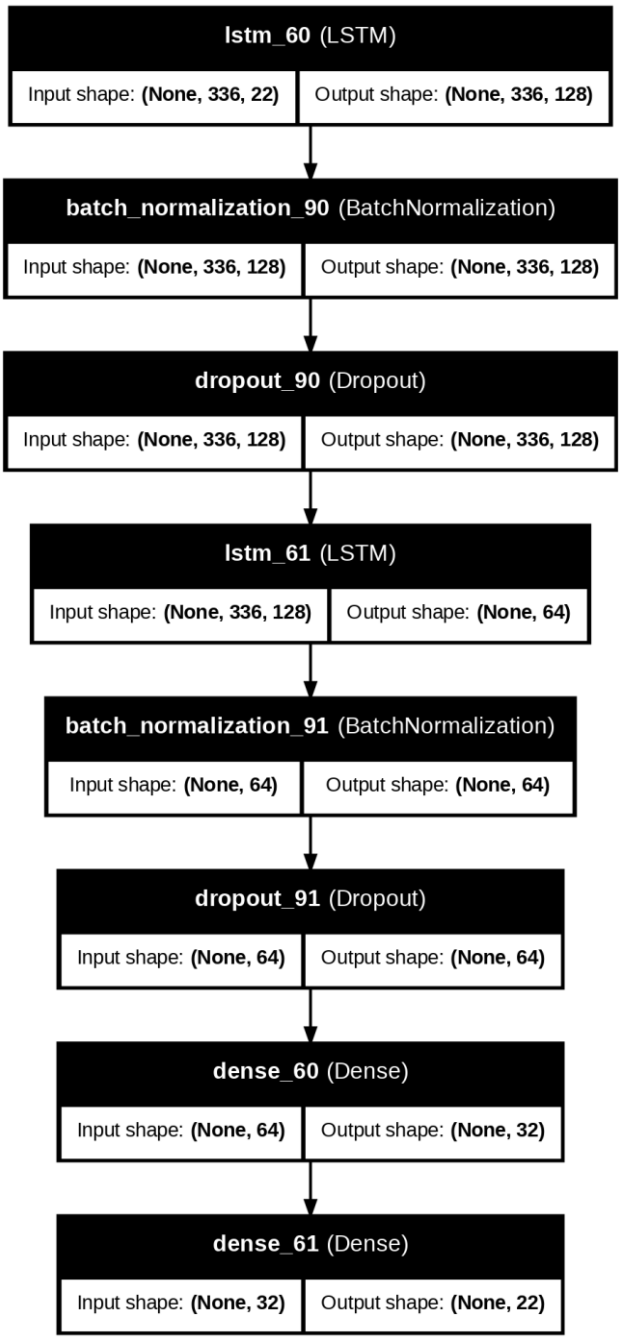
$$DTW(A, B) = \min_{w \in W} \sum_{(i,j) \in w} d(a_i, b_j) \quad (26)$$

onde Ae B são as duas séries comparadas (observada e prevista), w é um caminho de alinhamento que associa pontos de A a pontos de B, Wé o conjunto de todos os caminhos de alinhamento válidos, e $d(a_i, b_j) = | a_i - b_j |$é a distância entre o ponto $a_i$da série Ae o ponto $b_j$da série B.

4.5 Topologias

Para os experimentos com a redes recorrentes, a seguinte topologia mostrada na Figura 11 foi utilizada:

Figura 11- Topologia da LSTM utilizada nos experimentos



Fonte: Autoria própria

Os hiperparâmetros das topologias foram definidos por refinamento sucessivo em testes preliminares. Esses testes foram feitos apenas na arquitetura recorrente pura e depois os hiperparâmetros das topologias foram mantidos iguais em todas as outras configurações. Foi

iniciado com configurações maiores e diminuído o número de unidades pouco a pouco, até que as métricas de validação começaram a piorar. Configurações com 256 e 128 unidades não trouxeram ganho proporcional ao aumento de parâmetros e tempo de treinamento. Por outro lado, configurações com 64 e 32 unidades reduziram o desempenho em períodos mais longos. A melhor combinação das topologias foi usar 128 e 64 unidades, reduzindo progressivamente para a camada densa de 32 e depois para a saída de 22. Essa combinação das topologias ofereceu o melhor equilíbrio entre capacidade de representação e custo. A taxa de Dropout de 0,2 foi escolhida da mesma forma: valores maiores atrasaram a convergência sem melhorar o erro de validação, enquanto valores menores fizeram as curvas de treino e validação se aproximarem de forma indicativa de sobreajuste. Os parâmetros do bloco convolucional seguiram a mesma lógica. O kernel de tamanho 6 na primeira camada captura um padrão local de cerca de seis horas. O kernel de tamanho 3 na segunda camada, usado depois da primeira subamostragem, cobre um intervalo temporal equivalente sobre a sequência já reduzida.

A arquitetura da Figura 11 recebe como entrada um tensor de formato (336, 22), correspondente a uma janela de 336 horas (14 dias) contendo 22 variáveis por instante. A primeira camada é uma LSTM de 128 unidades configurada para retornar a sequência completa (*return_sequences=True*), de modo que sua saída preserva a dimensão temporal, resultando em um tensor (336, 128). Em seguida, aplicam-se uma camada de normalização em lote (B*atch Normalization*), que estabiliza e acelera o treinamento, e uma camada de *Dropout* com taxa de 0,2, que desativa aleatoriamente 20% das unidades para reduzir o sobreajuste; ambas mantêm o formato (336, 128).

A segunda camada recorrente é uma LSTM de 64 unidades configurada para retornar apenas o último estado (*return_sequences=False*), o que colapsa a dimensão temporal e produz um vetor de tamanho 64, sintetizando o histórico da sequência. Novamente, seguem-se *Batch Normalization* e *Dropout* de 0,2, preservando o tamanho 64. Por fim, uma camada densa de 32 neurônios com ativação ReLU realiza uma combinação não linear das características extraídas, e uma camada densa final de 22 neurônios (sem ativação) produz a saída, correspondente à previsão simultânea das 22 variáveis meteorológicas no horizonte-alvo.

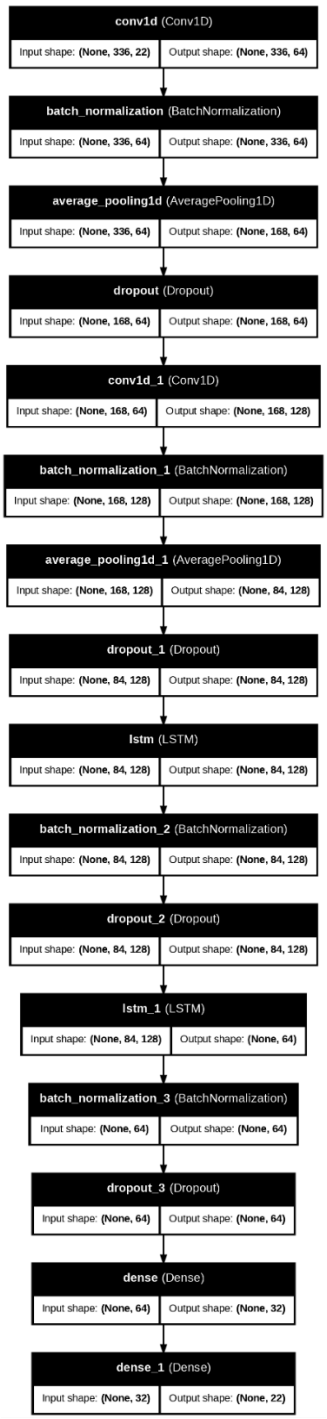
Dessa forma, as arquiteturas puramente recorrentes (LSTM e GRU) compartilham exatamente a mesma configuração de parâmetros, duas camadas recorrentes de 128 e 64 unidades, seguidas das mesmas camadas de normalização, *Dropout* e densas descritas acima, diferindo unicamente no tipo de célula recorrente empregada: a LSTM utiliza células *Long Short-Term Memory*, enquanto a GRU utiliza células *Gated Recurrent Unit*. Essa equivalência

estrutural é proposital, pois garante que qualquer diferença de desempenho observada entre os dois modelos decorra exclusivamente da célula recorrente, e não de variações de topologia.

Nas variantes bidirecionais, mantém-se a mesma quantidade de unidades-base (128 e 64), porém cada camada recorrente é encapsulada em um invólucro bidirecional, que processa a sequência simultaneamente nos dois sentidos temporais. Como consequência dessa concatenação, a dimensão de saída de cada camada é duplicada (256 e 128, respectivamente), o que também eleva o número de parâmetros treináveis.

Já as redes híbridas adotam como base a topologia apresentada na Figura 12, que contém um bloco convolucional antes da mesma recorrente que foi descrita anteriormente.

Figura 12 - Topologia da CNN - LSTM utilizada nos experimentos



Fonte: Autoria própria

Esta topologia híbrida utilizada em um dos experimentos recebe a mesma entrada de formato (336, 22) e inicia por um bloco de extração de características composto por duas camadas convolucionais unidimensionais (Conv1D), responsáveis por capturar padrões locais ao longo do eixo temporal antes da modelagem recorrente. A primeira convolução possui 64 filtros e kernel de tamanho 6, capaz de detectar padrões locais de aproximadamente seis horas, com preenchimento (*padding*) do tipo *same*, que preserva o comprimento da sequência, e

ativação ReLU. Segue-se uma *Batch Normalization* e uma camada de subamostragem por *Average Pooling* com fator 2, que reduz o comprimento da sequência pela metade, além de um Dropout de 0,2. A segunda convolução possui 128 filtros e *kernel* de tamanho 3, também com padding *same* e ReLU, seguida de nova *Batch Normalization*, de outra subamostragem por *Average Pooling* e de *Dropout* de 0,2. Após o bloco convolucional, as características extraídas são passadas para o bloco recorrente, que pode ser composto por células LSTM, GRU ou suas variantes bidirecionais, conforme a arquitetura avaliada.

A partir dessa base, derivam-se todas as demais arquiteturas híbridas que foram utilizadas nos experimentos.

5 RESULTADOS E DISCUSSÕES

Foram realizados os experimentos e seus resultados foram organizados em gráficos e tabelas com as métricas de desempenho de cada arquitetura. Para a análise, foram selecionadas quatro variáveis: Radiação Global, Temperatura do Bulbo Seco, Velocidade do Vento e Precipitação, escolhidas por representarem tanto os melhores quanto os piores resultados observados. Os demais resultados encontram-se no apêndice deste trabalho. Para o modelo estatístico ARIMA, os resultados obtidos das métricas são apresentados nas Tabelas 1 a 4, para os horizontes de 1h, 24h e 168h. Já para os valores obtidos para as redes recorrentes e híbridas, são apresentados os resultados nas Tabelas 5 a 16, para os horizontes de 1h, 24h e 168h. As curvas de predição para esses modelos são apresentadas nas Figuras 13 a 20.

Tabela 1 - Resultados das métricas para a Temperatura do Bulbo para o modelo ARIMA

Temperatura do Bulbo	**MAE**	**R^2**	**SMAPE (%)**	**DTW**
1h	0,8914	0,9284	4,77	0,8361
24h	2,4602	0,5803	12,79	2,2604
168h	3,8017	0,1727	19,40	2,8279

Fonte: Autoria própria

Tabela 2- Resultados das métricas para a Radiação Global para o modelo ARIMA

Radiação Global	**MAE**	**R²**	**SMAPE (%)**	**DTW**
1h	285,1580	0,8291	110,06	267,5054
24h	688,8496	0,2559	127,06	664,2087
168h	821,0325	−0,0449	153,45	809,2386

Fonte: Autoria própria

Tabela 3 - Resultados das métricas para a Velocidade do vento para o modelo ARIMA

Velocidade do Vento	**MAE**	**R²**	**SMAPE (%)**	**DTW**
1h	0,4816	0,6231	123,7	0,4739
24h	0,7611	0,216	136,98	0,711
168h	0,7488	0,2573	142,83	0,6851

Fonte: Autoria própria

Tabela 4 - Resultados das métricas para a Precipitação para o modelo ARIMA

Precipitação	**MAE**	**R²**	**SMAPE (%)**	**DTW**
1h	0,0868	0,4538	192,69	0,0778
24h	0,1905	-0,0032	194,36	0,175
168h	0,1822	0,003	197,03	0,1659

Fonte: Autoria própria

Tabela 5 – Resultados das métricas entre arquiteturas para a variável de Radiação global no horizonte de 1 hora

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	285,1580	0,8291	110,06	267,5054
LSTM	**135.5241 ± 4.7083**	**+0.9404 ± 0.0044**	47.69 ± 19.99	224.1966 ± 115.2630
GRU	143.9326 ± 23.2124	+0.9398 ± 0.0081	63.82 ± 27.25	232.7578 ± 283.6750
CNN-LSTM	150.4266 ± 8.2785	+0.9326 ± 0.0032	**46.84 ± 13.49**	**137.3254 ± 28.6792**
CNN-GRU	148.4706 ± 2.3638	+0.9350 ± 0.0012	55.18 ± 16.85	128.7297 ± 33.4855
CNN-LSTM-GRU	160.8187 ± 10.4006	+0.9301 ± 0.0041	78.06 ± 28.70	257.4223 ± 160.3239
CNN-BI-LSTM	144.3930 ± 8.7462	+0.9362 ± 0.0022	52.74 ± 24.37	145.6843 ± 81.6945
CNN-BI-GRU	141.1234 ± 7.1547	+0.9388 ± 0.0019	50.17 ± 24.98	149.3360 ± 62.6568
CNN-BI-LSTM-GRU	152.9123 ± 7.2917	+0.9327 ± 0.0023	69.11 ± 22.56	259.8008 ± 140.0045

Fonte: Autoria própria

Tabela 6 - Resultados das métricas entre arquiteturas para a variável de Radiação global no horizonte de 24 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	688,8496	0,2559	**127,06**	664,2087
LSTM	243.9957 ± 4.3743	+0.8322 ± 0.0019	**58.26 ± 22.14**	485.3137 ± 53.6011
GRU	**235.3160 ± 7.2983**	**+0.8413 ± 0.0023**	70.82 ± 12.52	**367.6697 ± 73.0489**
CNN-LSTM	243.1795 ± 15.6741	+0.8394 ± 0.0045	78.70 ± 25.11	495.4438 ± 123.4620
CNN-GRU	244.0959 ± 14.3629	+0.8385 ± 0.0071	74.29 ± 21.05	402.9758 ± 45.5451
CNN-LSTM-GRU	243.7092 ± 6.0490	+0.8350 ± 0.0040	63.05 ± 20.62	387.1521 ± 93.4656
CNN-BI-LSTM	254.0215 ± 12.9793	+0.8210 ± 0.0107	63.52 ± 12.12	453.5418 ± 82.3355
CNN-BI-GRU	250.0591 ± 6.8108	+0.8287 ± 0.0093	70.22 ± 15.21	431.5074 ± 38.3723
CNN-BI-LSTM-GRU	247.8321 ± 11.3678	+0.8313 ± 0.0041	70.53 ± 13.55	389.6886 ± 55.9239

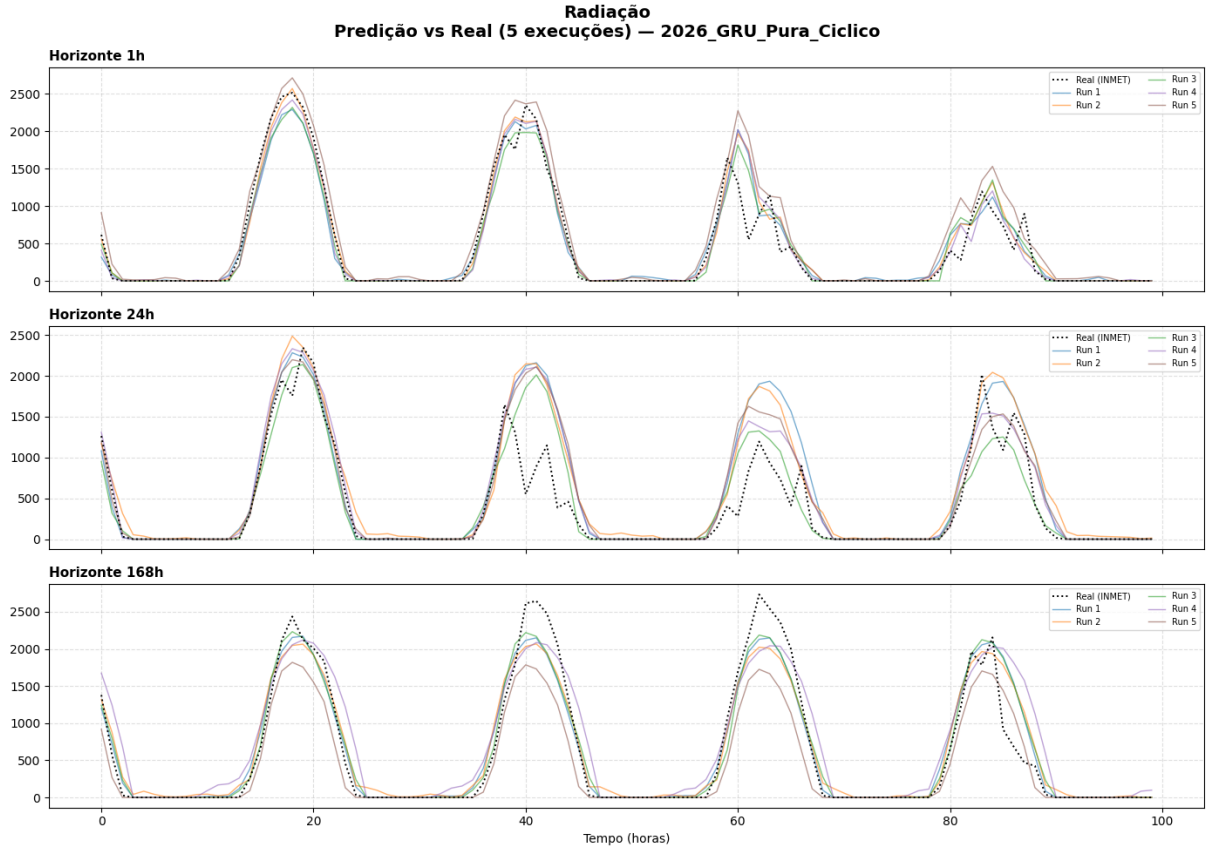
Fonte: Autoria própria

Tabela 7 - Resultados das métricas entre arquiteturas para a variável de Radiação global no horizonte de 168 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	821,0325	-0,0449	153,45	809,2386
LSTM	275.6701 ± 14.6161	+0.8081 ± 0.0100	79.14 ± 15.13	474.3191 ± 15.9818
GRU	275.9490 ± 18.8233	+0.8017 ± 0.0211	69.40 ± 18.40	566.3708 ± 157.5617
CNN-LSTM	271.8256 ± 10.9860	+0.8090 ± 0.0133	75.23 ± 21.90	513.8890 ± 163.3557
CNN-GRU	**269.2288 ± 14.2093**	**+0.8119 ± 0.0063**	73.83 ± 27.40	472.5620 ± 82.3631
CNN-LSTM-GRU	273.7385 ± 13.1633	+0.8065 ± 0.0089	70.24 ± 27.89	481.8739 ± 24.5771
CNN-BI-LSTM	284.7917 ± 10.3352	+0.8034 ± 0.0061	70.71 ± 13.33	**471.8866 ± 39.4685**
CNN-BI-GRU	278.8363 ± 9.2616	+0.8011 ± 0.0119	**66.72 ± 18.03**	542.4281 ± 74.8493
CNN-BI-LSTM-GRU	281.3100 ± 13.8784	+0.8044 ± 0.0083	74.57 ± 14.05	526.9129 ± 87.0585

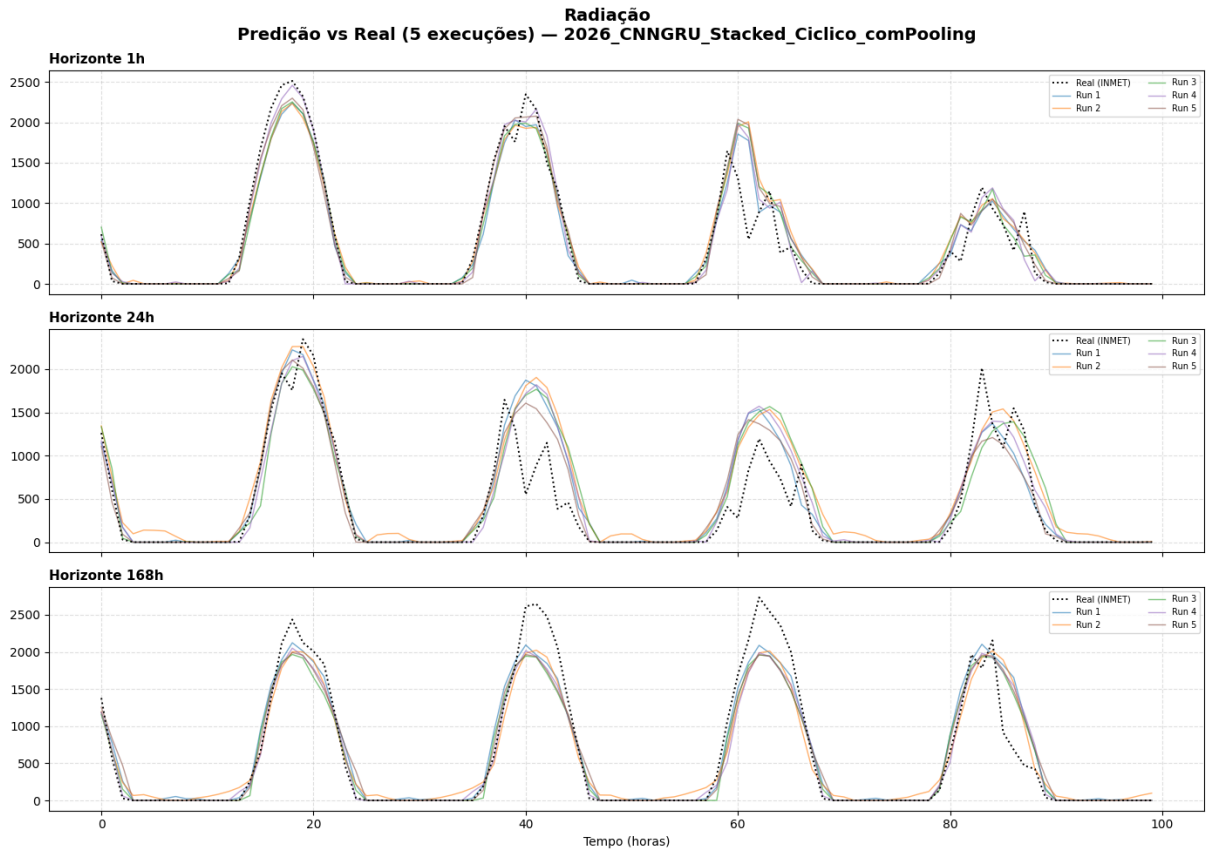
Fonte: Autoria própria

Figura 13 - Curvas de predição em comparação com o real (INMET) - GRU - Radiação Global



Fonte: Autoria própria

Figura 14 - Curvas de predição em comparação com o real (INMET) - CNN-GRU - Radiação Global



Fonte: Autoria própria

Ao analisar os resultados das métricas da radiação global entre as diferentes arquiteturas e horizontes, observou-se que ela foi a variável de melhor desempenho preditivo entre todas as analisadas, destacando-se pelos coeficientes de determinação mais altos, mesmo nos horizontes mais longos, Conforme mostram nas Tabelas 5 a 7. Seu R^2 partiu de aproximadamente 0,94 no horizonte de 1h (Tabela 5) e permaneceu em torno de 0,81 no horizonte de 168h (Tabela 7), o que evidencia uma redução bem menos acentuada do que a observada nas demais variáveis.

No horizonte de 1h, as arquiteturas recorrentes puras apresentam o melhor desempenho na maioria das métricas: a *LSTM* obteve o menor MAE (135,52) e o maior R^2 (0,9404), seguida de perto pela *GRU*. As redes híbridas que utilizam convolução ficaram, em geral, ligeiramente abaixo, com destaque, para essa variável em específico, para a *CNN-LSTM-GRU*, que obteve o menor R^2 (0,9301) e o maior SMAPE (78,06%). Ainda assim, a amplitude entre o melhor e o pior modelo foi de apenas 0,010 em R^2, o que indica poucas diferenças entre eles.

Para o horizonte de 24h, a GRU apresentou o melhor desempenho, obtendo o menor MAE (235,32), o maior R^2 (0,8413) e o menor DTW (367,67). Os modelos híbridos

mantiveram-se próximos, porém sem superar as recorrentes puras, e a *CNN-BI-LSTM* apresentou o pior R^2 (0,8210). A amplitude entre os modelos foi de 0,020 em R^2, pequena, porém maior que a observada no horizonte de 1h.

Já no horizonte de 168h, a *CNN-GRU* obteve o melhor MAE (269,23) e o melhor R^2 (0,8119), superando ligeiramente as recorrentes puras. No entanto, a amplitude entre os modelos permaneceu mínima, de apenas 0,011 em R^2.

Por fim, é importante destacar um comportamento particular da radiação global, embora o R^2 seja elevado em todos os horizontes, o SMAPE manteve-se alto e instável, variando de 47% a 79%. Isso ocorre porque, durante a noite, a radiação global é nula ou próxima de zero, e o SMAPE torna-se muito sensível a valores próximos de zero, pois qualquer diferença entre o valor previsto e o real, mesmo que irrelevante em magnitude, resulta em um erro percentual elevado. Dessa forma, o SMAPE é inflado pelos instantes noturnos, não refletindo uma falha real do modelo em prever a variável. Para variáveis com esse comportamento, métricas de erro absoluto como o MAE ou o RMSE oferecem leitura mais fiel do desempenho, uma vez que não dependem da magnitude do valor observado.

Tabela 8 - Resultados das métricas entre arquiteturas para a variável de Temperatura do bulbo no horizonte de 1 hora

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,8914	0,9284	4,77	0,8361
LSTM	0.7950 ± 0.0250	+0.9594 ± 0.0015	4.09 ± 0.19	0.5796 ± 0.0248
GRU	**0.7548 ± 0.0819**	**+0.9620 ± 0.0060**	**3.86 ± 0.48**	**0.5382 ± 0.0406**
CNN-LSTM	0.8468 ± 0.0643	+0.9530 ± 0.0064	4.26 ± 0.30	0.6405 ± 0.1268
CNN-GRU	0.8378 ± 0.0384	+0.9542 ± 0.0032	4.25 ± 0.20	0.5900 ± 0.0293
CNN-LSTM-GRU	0.8395 ± 0.0353	+0.9538 ± 0.0029	4.26 ± 0.20	0.5855 ± 0.0407
CNN-BI-LSTM	0.8285 ± 0.0489	+0.9545 ± 0.0044	4.19 ± 0.25	0.5853 ± 0.0642
CNN-BI-GRU	0.8476 ± 0.1259	+0.9534 ± 0.0105	4.32 ± 0.73	0.7103 ± 0.3203
CNN-BI-LSTM-GRU	0.8268 ± 0.0168	+0.9547 ± 0.0016	4.17 ± 0.08	0.5860 ± 0.0212

Fonte: Autoria própria

Tabela 9 - Resultados das métricas entre arquiteturas para a variável de Temperatura do bulbo no horizonte de 24 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	2,4602	0,5803	12,79	2,2604
LSTM	2.0643 ± 0.0906	+0.7428 ± 0.0123	10.47 ± 0.53	2.3185 ± 0.3169
GRU	**1.9117 ± 0.1005**	+0.7715 ± 0.0156	**9.66 ± 0.58**	**1.7287 ± 0.2915**
CNN-LSTM	1.9347 ± 0.0659	+0.7635 ± 0.0132	9.79 ± 0.34	2.0204 ± 0.2280
CNN-GRU	1.9509 ± 0.0478	+0.7574 ± 0.0097	9.87 ± 0.23	1.9640 ± 0.2006
CNN-LSTM-GRU	1.9200 ± 0.0331	**+0.7669 ± 0.0045**	9.71 ± 0.17	1.9377 ± 0.2987
CNN-BI-LSTM	2.0093 ± 0.0604	+0.7445 ± 0.0180	10.14 ± 0.24	2.0926 ± 0.2936
CNN-BI-GRU	2.5227 ± 0.0346	+0.6265 ± 0.0070	12.77 ± 0.19	2.9094 ± 0.1130
CNN-BI-LSTM-GRU	1.9922 ± 0.0306	+0.7530 ± 0.0060	10.12 ± 0.19	1.9984 ± 0.2470

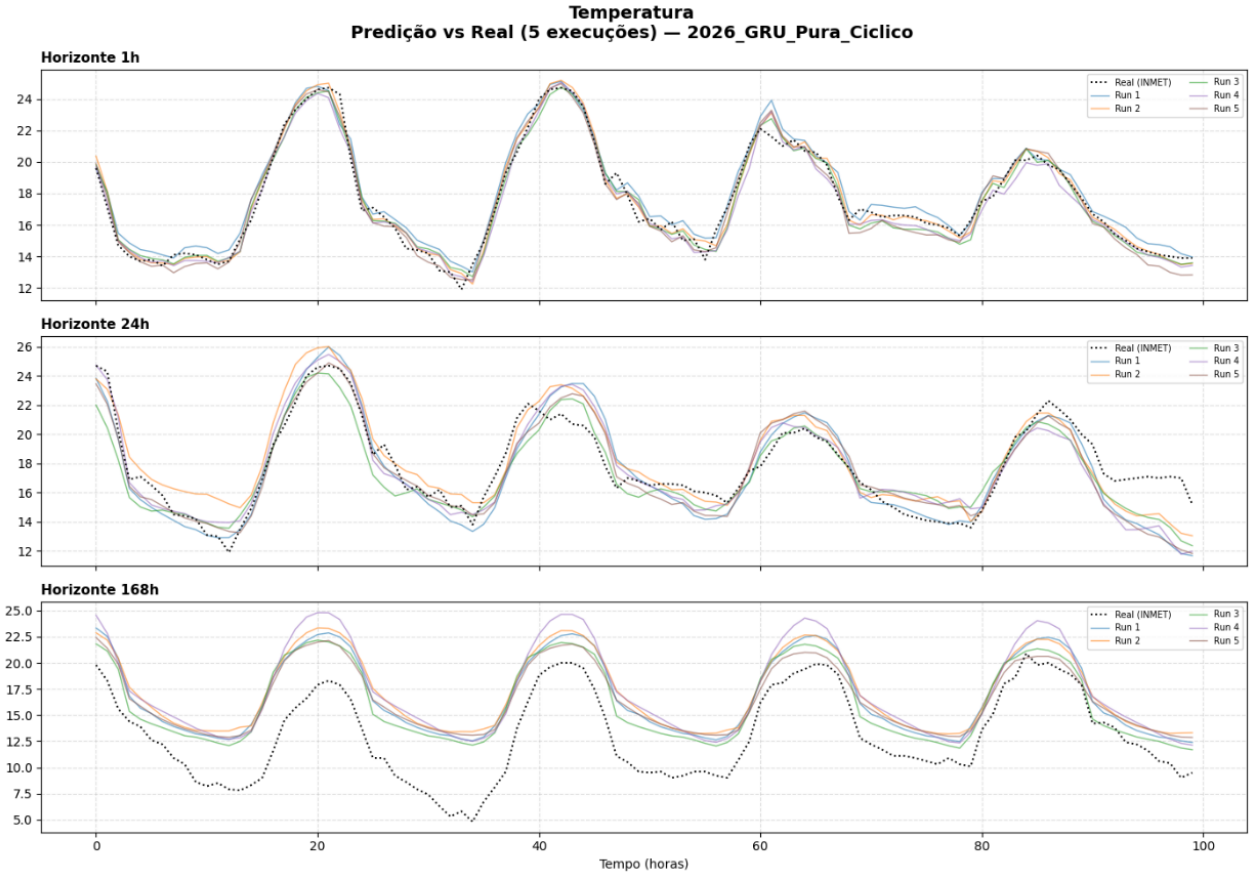
Fonte: Autoria própria

Tabela 10 - Resultados das métricas entre arquiteturas para a variável de Temperatura do bulbo no horizonte de 168 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	3,8017	0,1727	19,40	2,8279
LSTM	2.5080 ± 0.0597	+0.6306 ± 0.0181	12.70 ± 0.30	2.8756 ± 0.1479
GRU	2.4907 ± 0.0411	+0.6344 ± 0.0117	12.61 ± 0.21	2.8756 ± 0.1881
CNN-LSTM	2.4795 ± 0.0342	+0.6362 ± 0.0073	**12.59 ± 0.16**	2.7751 ± 0.1540
CNN-GRU	2.4976 ± 0.0341	+0.6341 ± 0.0056	12.66 ± 0.19	2.8237 ± 0.1456
CNN-LSTM-GRU	2.5335 ± 0.0633	+0.6217 ± 0.0159	12.84 ± 0.30	2.8236 ± 0.1714
CNN-BI-LSTM	2.5399 ± 0.0662	+0.6231 ± 0.0160	12.93 ± 0.37	**2.7140 ± 0.1286**
CNN-BI-GRU	2.5227 ± 0.0346	+0.6265 ± 0.0070	12.77 ± 0.19	2.9094 ± 0.1130
CNN-BI-LSTM-GRU	**2.4736 ± 0.0261**	**+0.6384 ± 0.0064**	12.59 ± 0.17	2.7403 ± 0.0593

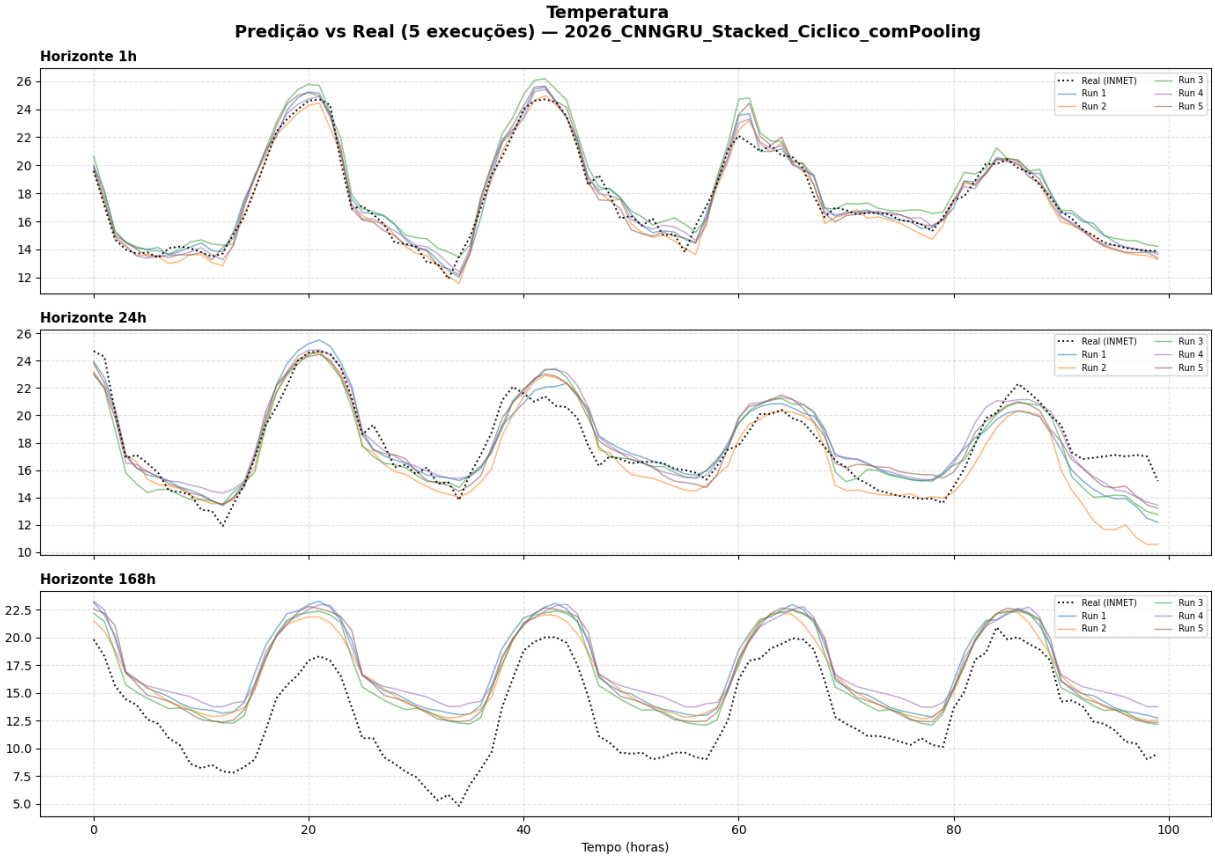
Fonte: Autoria própria

Figura 15 - Curvas de predição em comparação com o real (INMET) - GRU - Temperatura do Bulbo



Fonte: Autoria própria

Figura 16 - Curvas de predição em comparação com o real (INMET) - CNN-GRU - Temperatura do Bulbo



Fonte: Autoria própria

A temperatura do bulbo seco, variável escolhida para representar as térmicas, apresentou desempenho preditivo elevado no curto prazo, com diminuição progressiva conforme o horizonte aumenta, conforme mostrado nas Tabelas de 8 a 10. Para a GRU, arquitetura que obteve os melhores resultados nas métricas avaliadas para esta variável, o R^2 partiu de 0,96 e reduziu-se até 0,63 no horizonte de 168h.

No horizonte de 1h, conforme mencionado anteriormente, a GRU obteve o melhor desempenho em todas as métricas, com o menor MAE (0,7548), o maior R^2 (0,9620), o menor SMAPE (3,86%) e o menor DTW (0,5382), seguida pela LSTM. As redes híbridas apresentaram desempenho ligeiramente inferior em comparação as recorrentes puras, e a amplitude entre o melhor e o pior modelo foi de apenas 0,009 em R^2, indicando uma diferença pequena entre as arquiteturas.

Para o horizonte de 24h, a GRU manteve o melhor desempenho nas métricas (MAE de 1,9117, R^2 de 0,7715 e DTW de 1,7287), seguida pela *CNN-LSTM-GRU*. A *CNN-BI-GRU* apresentou um desempenho anômalo em relação às demais, com R^2 de 0,6265; desconsiderando

essa arquitetura, a amplitude entre os modelos foi de aproximadamente 0,029 em R^2, permanecendo pequena, ainda que maior que a observada no horizonte de 1h.

Já para o horizonte de 168h, os modelos apresentaram desempenho praticamente equivalente. A *CNN-BI-LSTM-GRU* obteve o melhor R^2 (0,6384) e o melhor MAE (2,4736), mas a amplitude entre os modelos foi de apenas 0,017 em R^2, de modo que as diferenças não são significativas.

Em resumo, para a temperatura do bulbo seco, a GRU teve o maior destaque nos horizontes de 1h e 24h, enquanto no horizonte de 168h todos os modelos convergiram para desempenho semelhante. Logo, neste caso, a presença de convolução ou de bidirecionalidade não resultou em ganho relevante de desempenho nas predições.

Tabela 11 - Resultados das métricas entre arquiteturas para a variável de Velocidade do vento no horizonte de 1 hora

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,1924	0,6963	117,92	0,1752
LSTM	0.5304 ± 0.0534	**+0.7004 ± 0.0132**	**116.50 ± 11.50**	**0.4439 ± 0.0869**
GRU	0.5381 ± 0.0546	+0.6991 ± 0.0197	120.09 ± 8.69	0.4375 ± 0.0839
CNN-LSTM	0.5420 ± 0.0331	+0.6909 ± 0.0037	121.18 ± 7.67	0.4486 ± 0.0383
CNN-GRU	0.5615 ± 0.0382	+0.6893 ± 0.0150	122.20 ± 5.44	0.4914 ± 0.0349
CNN-LSTM-GRU	**0.5248 ± 0.0312**	+0.6976 ± 0.0062	112.65 ± 9.45	0.4517 ± 0.0265
CNN-BI-LSTM	0.5440 ± 0.0363	+0.6888 ± 0.0151	120.62 ± 6.41	0.4550 ± 0.0578
CNN-BI-GRU	0.5397 ± 0.0429	+0.6951 ± 0.0113	117.37 ± 12.62	0.4657 ± 0.0622
CNN-BI-LSTM-GRU	0.5355 ± 0.0322	+0.6940 ± 0.0102	118.83 ± 11.67	0.4419 ± 0.0742

Fonte: Autoria própria

Tabela 12 - Resultados das métricas entre arquiteturas para a variável de Velocidade do vento no horizonte de 24 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,3201	0,3950	128,17	0,2943
LSTM	1.0248 ± 0.0450	+0.2513 ± 0.0264	138.15 ± 2.02	0.8873 ± 0.0540
GRU	**1.0166 ± 0.1405**	**+0.2614 ± 0.0863**	**137.84 ± 2.47**	**0.8782 ± 0.1897**
CNN-LSTM	1.2472 ± 0.1373	+0.0668 ± 0.1231	139.94 ± 0.52	1.1216 ± 0.1406
CNN-GRU	1.0845 ± 0.0670	+0.2125 ± 0.0350	138.48 ± 2.26	0.9602 ± 0.1131
CNN-LSTM-GRU	1.1893 ± 0.1128	+0.1251 ± 0.0975	139.92 ± 0.44	1.0436 ± 0.1461
CNN-BI-LSTM	1.0908 ± 0.0820	+0.1717 ± 0.0786	139.22 ± 1.25	0.9188 ± 0.1134
CNN-BI-GRU	1.2550 ± 0.1632	+0.0503 ± 0.1434	139.62 ± 0.63	1.1001 ± 0.1844
CNN-BI-LSTM-GRU	1.1975 ± 0.0280	+0.1281 ± 0.0302	140.06 ± 0.31	1.0933 ± 0.0743

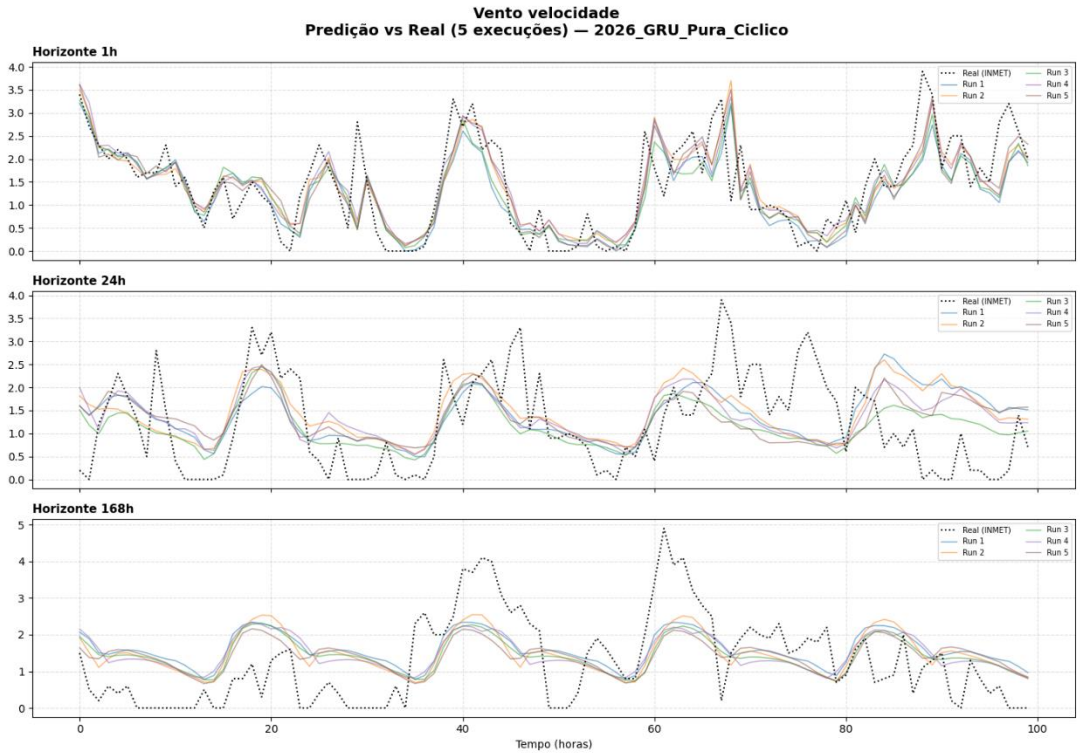
Fonte: Autoria própria

Tabela 13 - Resultados das métricas entre arquiteturas para a variável de Velocidade do vento no horizonte de 168 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,3302	0,3792	135,22	0,2942
LSTM	1.4138 ± 0.1103	-0.1231 ± 0.1226	142.39 ± 0.49	1.1213 ± 0.1853
GRU	1.4113 ± 0.0436	**-0.0997 ± 0.0445**	142.57 ± 0.16	**1.0417 ± 0.2310**
CNN-LSTM	**1.3952 ± 0.1391**	-0.1023 ± 0.1365	142.49 ± 0.50	1.1488 ± 0.2412
CNN-GRU	1.5073 ± 0.0492	-0.2224 ± 0.0631	142.22 ± 0.22	1.1424 ± 0.1594
CNN-LSTM-GRU	1.4125 ± 0.0459	-0.1060 ± 0.0526	142.37 ± 0.21	1.1196 ± 0.2094
CNN-BI-LSTM	1.4940 ± 0.1018	-0.2092 ± 0.1225	**142.16 ± 0.31**	1.2528 ± 0.2256
CNN-BI-GRU	1.4779 ± 0.0852	-0.1786 ± 0.1030	142.20 ± 0.24	1.2683 ± 0.2001
CNN-BI-LSTM-GRU	1.5004 ± 0.0618	-0.2072 ± 0.0753	142.25 ± 0.28	1.2937 ± 0.0849

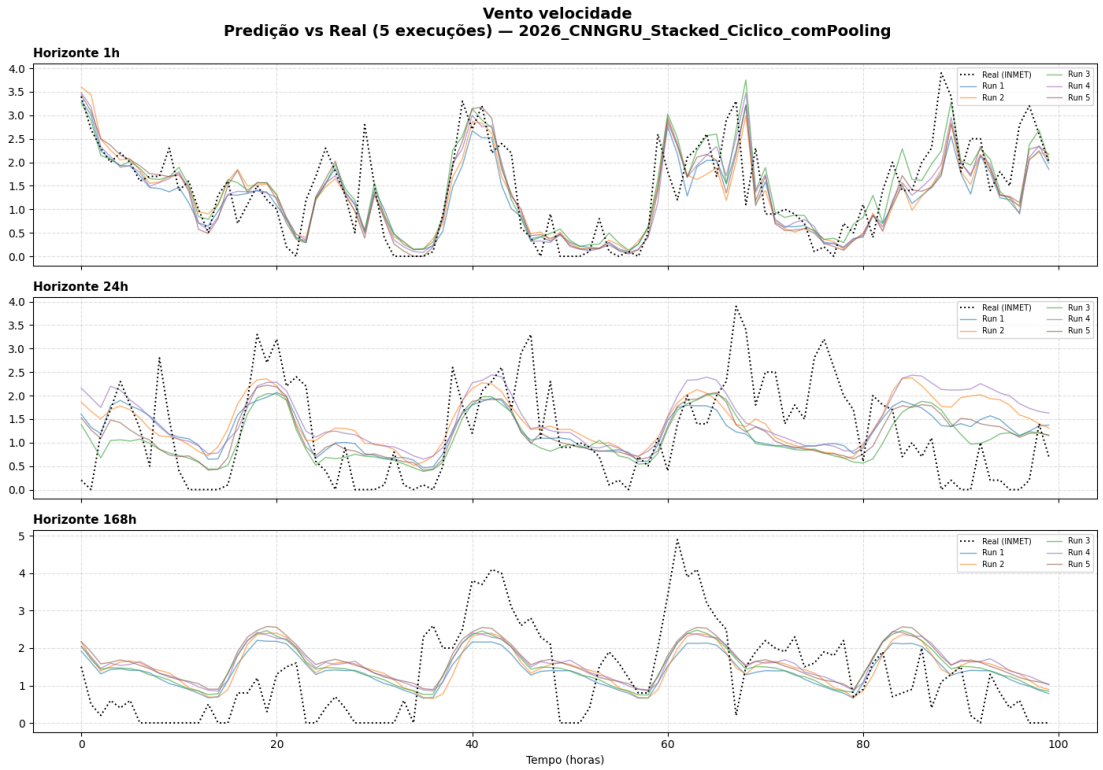
Fonte: Autoria própria

Figura 17 - Curvas de predição em comparação com o real (INMET) - GRU – Velocidade do Vento



Fonte: Autoria própria

Figura 18 - Curvas de predição em comparação com o real (INMET) - CNN-GRU – Velocidade do Vento



Fonte: Autoria própria

A velocidade do vento, escolhida por representar as variáveis de comportamento mais caótico, apresentou o pior desempenho preditivo entre as variáveis analisadas, com forte degradação conforme o horizonte aumenta, conforme mostrado nas Tabelas de 11 a 13. O R^2 partiu de aproximadamente 0,70 no horizonte de 1h, caiu para cerca de 0,25 no horizonte de 24h e tornou-se negativo no horizonte de 168h, atingindo valores em torno de -0,10 a -0,22, evidenciando a dificuldade de fazer a predição para essa variável em horizontes longos.

No horizonte de 1h, as arquiteturas recorrentes puras apresentaram o melhor desempenho, com a LSTM obtendo o maior R^2 (0,7004) e a GRU logo em seguida (0,6991). As variantes com convolução ficaram ligeiramente abaixo, e a amplitude entre o melhor e o pior modelo foi de apenas 0,012 em R^2, indicando diferenças reduzidas entre as arquiteturas neste horizonte.

No horizonte de 24h, a GRU obteve o melhor desempenho em todas as métricas (MAE de 1,0166, R^2 de 0,2614 e DTW de 0,8782), seguida pela LSTM. Neste horizonte, observa-se a maior diferença entre as arquiteturas: a CNN-LSTM registrou um R^2 de apenas 0,0668, muito abaixo da LSTM pura (0,2513), e a CNN-BiGRU obteve o pior resultado (0,0503). A amplitude

entre os modelos alcançou 0,211 em R^2, a maior observada entre todas as variáveis e horizontes analisados, evidenciando que, para o vento no médio prazo, a escolha da arquitetura passa a ter um impacto maior.

No horizonte de 168h, todos os modelos apresentaram R^2 negativo, entre -0,0997 (GRU) e -0,2224 (CNN-GRU), confirmando a perda quase total de capacidade preditiva. A GRU e a CNN-LSTM obtiveram os valores menos negativos, mas o desempenho é insatisfatório para todas as arquiteturas.

Em resumo, para a velocidade do vento, as recorrentes puras apresentaram desempenho superior nos horizontes de 1h e 24h, enquanto no horizonte de 168h nenhuma arquitetura obteve resultado satisfatório. Destaca-se ainda que, no horizonte de 24h, a introdução da convolução resultou em queda expressiva de desempenho em algumas arquiteturas, comportamento que será discutido em detalhe posteriormente.

Tabela 14 - Resultados das métricas entre arquiteturas para a variável de Precipitação no horizonte de 1 hora

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,0500	0,3785	192,22	0,0445
LSTM	**0.1384 ± 0.0038**	+0.1075 ± 0.0187	**95.88 ± 42.52**	**0.1270 ± 0.0079**
GRU	0.1474 ± 0.0092	**+0.1124 ± 0.0195**	136.55 ± 23.41	0.1334 ± 0.0059
CNN-LSTM	0.1757 ± 0.0318	+0.0630 ± 0.0210	142.52 ± 28.97	0.1681 ± 0.0281
CNN-GRU	0.1555 ± 0.0155	+0.0689 ± 0.0274	131.99 ± 50.17	0.1518 ± 0.0188
CNN-LSTM-GRU	0.1527 ± 0.0160	+0.0607 ± 0.0221	98.65 ± 50.66	0.1485 ± 0.0166
CNN-BI-LSTM	0.1617 ± 0.0220	+0.0763 ± 0.0238	124.80 ± 42.08	0.1493 ± 0.0235
CNN-BI-GRU	0.1599 ± 0.0168	+0.0744 ± 0.0270	129.42 ± 34.19	0.1509 ± 0.0147
CNN-BI-LSTM-GRU	0.1538 ± 0.0099	+0.0678 ± 0.0237	99.37 ± 46.35	0.1478 ± 0.0093

Fonte: Autoria própria

Tabela 15 - Resultados das métricas entre arquiteturas para a variável de Precipitação no horizonte de 24 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,0801	-0,0403	194,06	0,0667
LSTM	0.1865 ± 0.0172	+0.0016 ± 0.0040	166.67 ± 20.22	0.1870 ± 0.0175
GRU	**0.1787 ± 0.0084**	+0.0015 ± 0.0026	162.82 ± 18.65	**0.1784 ± 0.0096**
CNN-LSTM	0.1839 ± 0.0210	+0.0013 ± 0.0040	**155.11 ± 36.02**	0.1825 ± 0.0209
CNN-GRU	0.1905 ± 0.0155	+0.0033 ± 0.0044	167.19 ± 22.53	0.1861 ± 0.0149
CNN-LSTM-GRU	0.1863 ± 0.0179	+0.0023 ± 0.0029	162.09 ± 28.06	0.1862 ± 0.0174
CNN-BI-LSTM	0.1918 ± 0.0139	+0.0018 ± 0.0028	171.60 ± 13.61	0.1891 ± 0.0127
CNN-BI-GRU	0.2022 ± 0.0233	**+0.0054 ± 0.0037**	177.32 ± 22.51	0.1995 ± 0.0232
CNN-BI-LSTM-GRU	0.1928 ± 0.0173	+0.0028 ± 0.0031	169.06 ± 14.55	0.1908 ± 0.0157

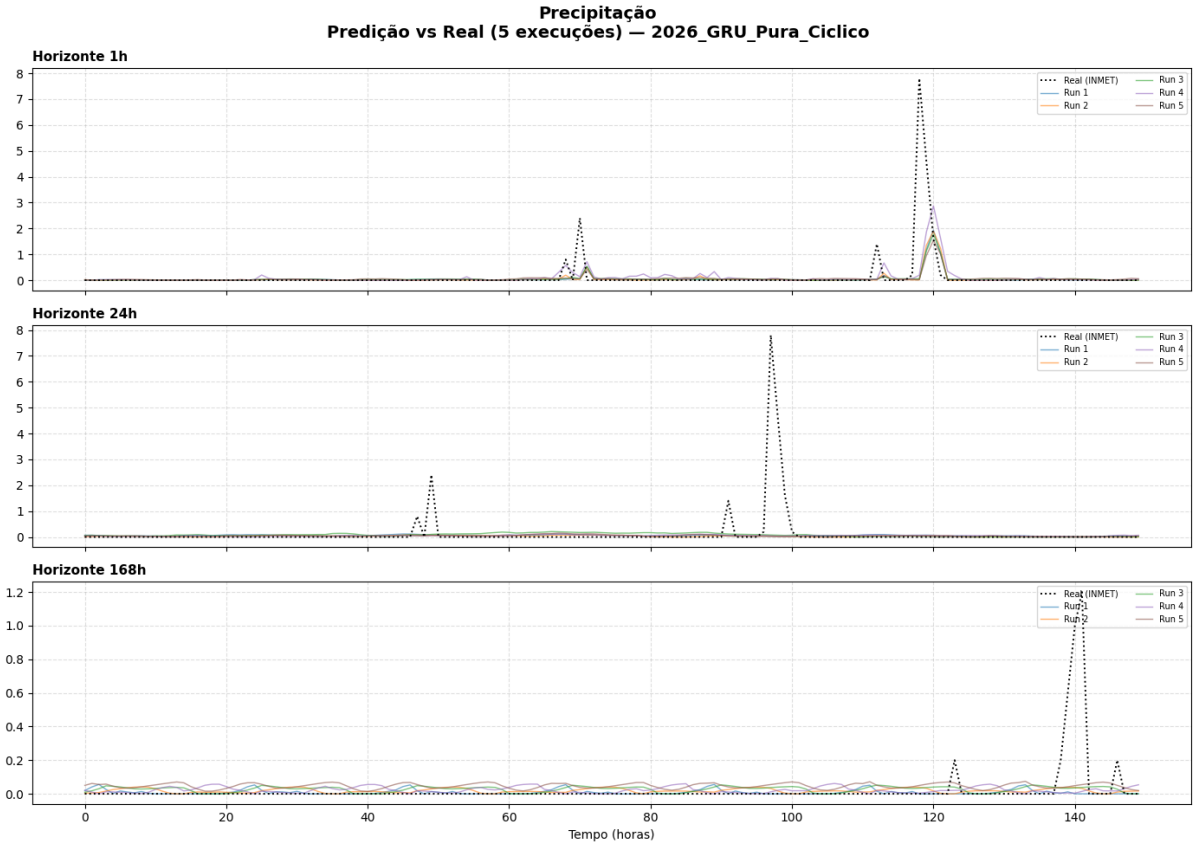
Fonte: Autoria própria

Tabela 16 - Resultados das métricas entre arquiteturas para a variável de Precipitação no horizonte de 168 horas

Modelo	MAE	R^2	SMAPE (%)	DTW
ARIMA	0,0771	-0,0275	196,82	0,0671
LSTM	0.1803 ± 0.0214	**+0.0004 ± 0.0041**	**163.55 ± 32.19**	0.1823 ± 0.0215
GRU	0.1769 ± 0.0210	-0.0024 ± 0.0024	178.34 ± 13.91	**0.1773 ± 0.0200**
CNN-LSTM	0.1884 ± 0.0176	-0.0003 ± 0.0033	185.48 ± 20.50	0.1905 ± 0.0154
CNN-GRU	0.1839 ± 0.0090	+0.0003 ± 0.0009	187.63 ± 12.09	0.1873 ± 0.0105
CNN-LSTM-GRU	**0.1768 ± 0.0058**	-0.0005 ± 0.0018	170.92 ± 21.89	0.1780 ± 0.0078
CNN-BI-LSTM	0.1796 ± 0.0170	-0.0024 ± 0.0030	168.16 ± 24.91	0.1796 ± 0.0194
CNN-BI-GRU	0.1827 ± 0.0240	-0.0014 ± 0.0046	169.19 ± 35.88	0.1831 ± 0.0269
CNN-BI-LSTM-GRU	0.1828 ± 0.0204	-0.0003 ± 0.0047	171.69 ± 15.29	0.1851 ± 0.0158

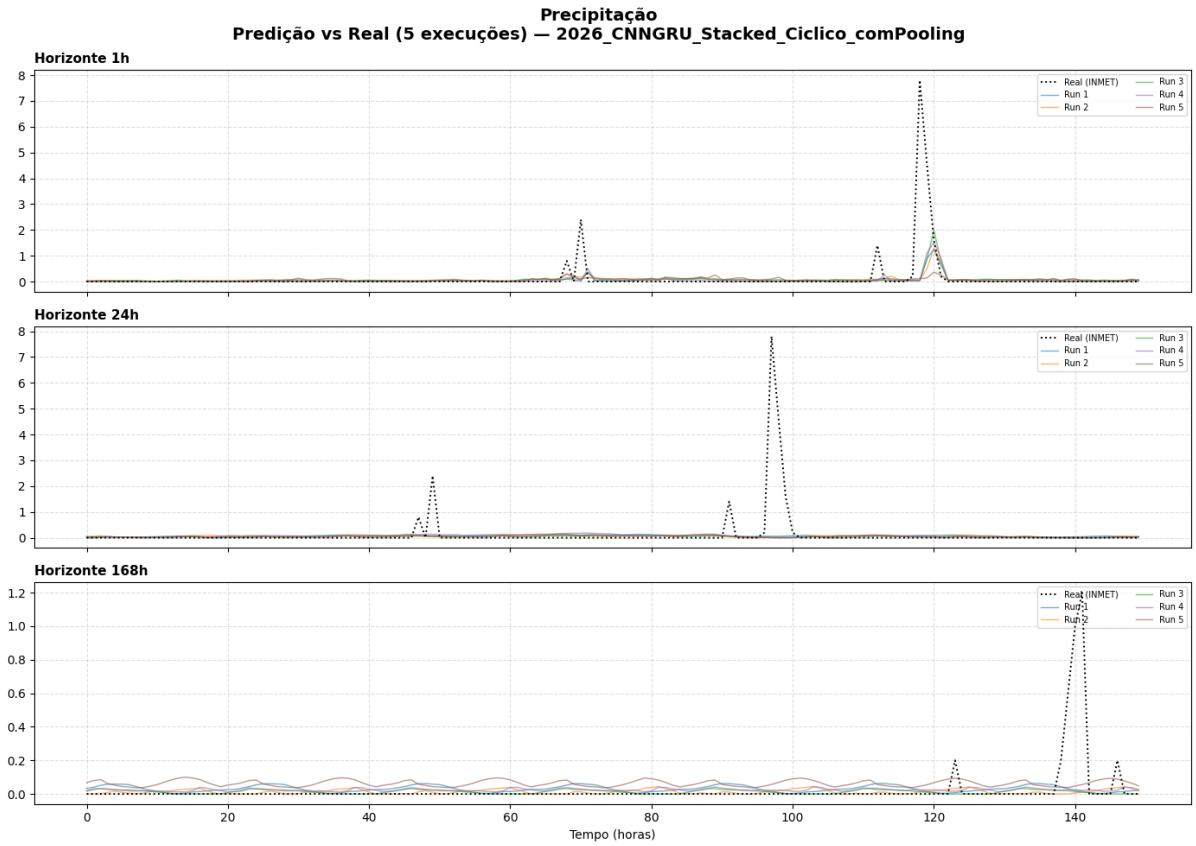
Fonte: Autoria própria

Figura 19 - Curvas de predição em comparação com o real (INMET) - GRU - Precipitação



Fonte: Autoria própria

Figura 20 - Curvas de predição em comparação com o real (INMET) - CNN-GRU - Precipitação



Fonte: Autoria própria

A precipitação, escolhida por representar a variável pluviométrica mais esparsa e de difícil previsão, apresentou o pior desempenho preditivo entre todas as variáveis analisadas, conforme mostrado nas Tabelas 14 a 16. Diferentemente das demais, seu R^2 já parte de valores baixos no horizonte de 1h, em torno de 0,11, e reduz-se a quase zero nos horizontes de 24h e 168h.

No horizonte de 1h, a GRU obteve o maior R^2 (0,1124) e a LSTM o menor MAE (0,1384), o menor SMAPE (95,88%) e o menor DTW (0,1270), com as recorrentes puras novamente à frente das variantes com convolução. A amplitude entre os modelos foi de 0,052 em R^2, a maior observada para esta variável, ainda que os valores absolutos sejam baixos.

Nos horizontes de 24h e 168h, todos os modelos apresentaram R^2 praticamente nulo, variando entre 0,0054 e valores levemente negativos. Nesses horizontes, a diferença entre as arquiteturas é irrelevante (amplitude inferior a 0,005 em R^2), e nenhum modelo consegue prever a precipitação de forma útil.

Cabe destacar um comportamento particular desta variável, embora o MAE seja baixo em todos os horizontes (em torno de 0,18), esse valor não reflete desempenho adequado. Como a precipitação é uma variável esparsa, composta em sua maior parte por valores nulos ou próximos de zero, o modelo tende a prever valores baixos de forma quase constante, o que resulta em erros absolutos pequenos na maioria dos instantes. Entretanto, o R^2 próximo de zero indica que o modelo não captura a variabilidade real da precipitação. O SMAPE elevado, superior a 150% na maioria dos casos, reforça essa limitação.

Em resumo, para a precipitação, nenhuma arquitetura obteve desempenho satisfatório, e a presença de convolução ou de bidirecionalidade não alterou esse quadro. O baixo desempenho está associado à natureza esparsa e pouco previsível da variável, e não necessariamente a escolha da arquitetura.

5.1 Análise Redes Recorrentes vs Redes híbridas

Com base nos experimentos realizados anteriormente, dois critérios foram adotados: todas as quatro métricas são consideradas em conjunto para comparação das redes híbridas e recorrentes. Além disso, adotou-se o desvio padrão entre as execuções apenas para as redes recorrentes e híbridas para distinguir uma diferença considerável de um empate técnico.

Ao analisar o desempenho do ARIMA, de maneira geral, seu desempenho foi competitivo apenas no horizonte de 1h, caindo de desempenho de forma acentuada nos horizontes mais longos. Na temperatura do bulbo seco, o R^2 partiu de 0,928 no horizonte de 1h, próximo ao das redes neurais, mas caiu para 0,173 no horizonte de 168h, evidenciando a dificuldade do modelo em capturar dependências de longo prazo. Comportamento semelhante ocorreu na radiação global, em que o R^2 partiu de 0,829 no horizonte de 1h e tornou-se negativo (-0,045) em 168h, indicando que o ARIMA não reproduz o ciclo diário em horizontes longos, ao contrário das redes neurais, que mantiveram R^2 em torno de 0,81. Já para as variáveis mais caóticas, o ARIMA apresentou desempenho comparável ou até superior ao das redes neurais. Na velocidade do vento, manteve R^2 positivo em todos os horizontes (0,623 a 0,257), enquanto as redes chegaram a valores negativos em 168h; e na precipitação, obteve R^2 de 0,454 no horizonte de 1h, bem acima do observado para as redes (0,112). Ressalta-se que o ARIMA é um modelo univariado, utilizando apenas a série da própria variável, enquanto as redes neurais dispõem das demais variáveis como entrada. Esses resultados reforçam que, para variáveis de baixa previsibilidade, a maior complexidade das redes neurais não se traduz em ganho de desempenho, podendo inclusive ser desvantajosa frente a um modelo estatístico mais simples.

Ao realizar a comparação entre as melhores arquiteturas recorrentes puras e as melhores arquiteturas híbridas, é possível notar um padrão consistente ao longo dos horizontes. Na maioria dos casos, as diferenças entre os dois tipos de redes são menores que a variação entre suas execuções, o que caracteriza empate técnico. Quando há uma diferença significativa, raramente são favorecidas as arquiteturas híbridas.

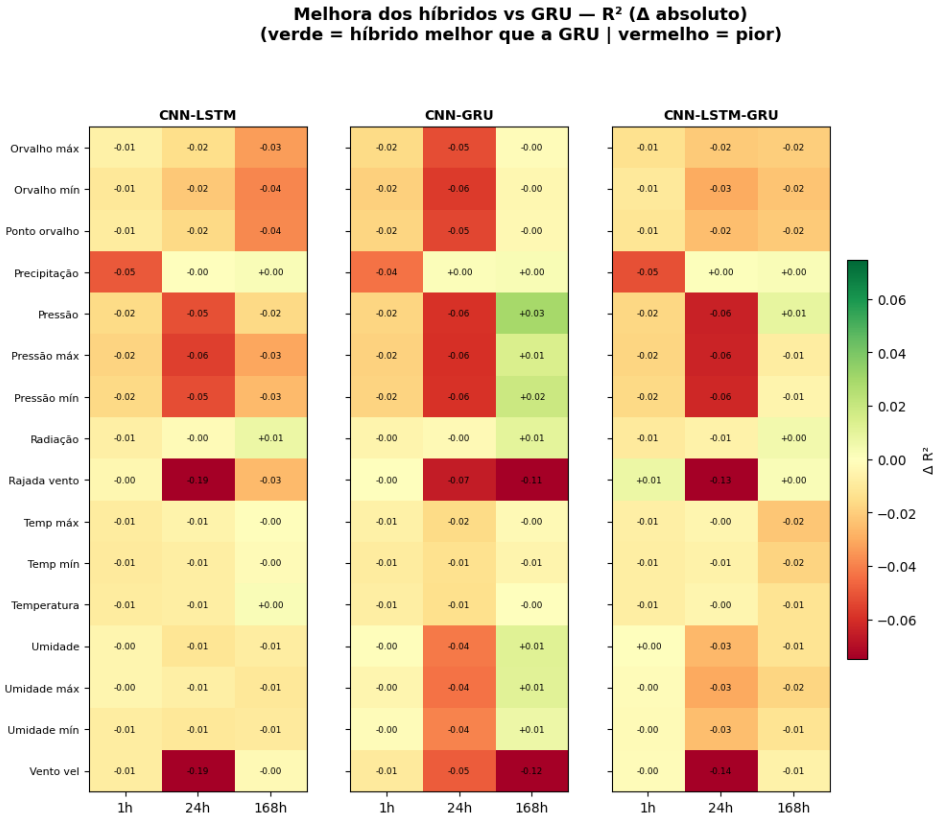
No horizonte de 1h, das 16 variáveis analisadas, 14 apresentaram empate técnico entre as melhores arquiteturas recorrentes e híbridas. Nas duas variáveis restantes (Precipitação e Rajada máxima do vento), a vantagem foi das recorrentes.

No horizonte de 24h, as recorrentes ganham em 10 das 16 variáveis, apresentando um desempenho significativamente superior, enquanto os híbridos não superam as puras em nenhuma variável, sendo assim o horizonte em que a convolução não fez diferença positiva.

Para o Horizonte de 168h, predomina novamente o empate técnico, presente em 12 das 16 variáveis. Nas quatro restantes (variáveis de pressão atmosférica e umidade relativa máxima), houve leve vantagem das arquiteturas híbridas, sendo próximas ao limite da variação entre execuções, o que não permite destacar uma superioridade real.

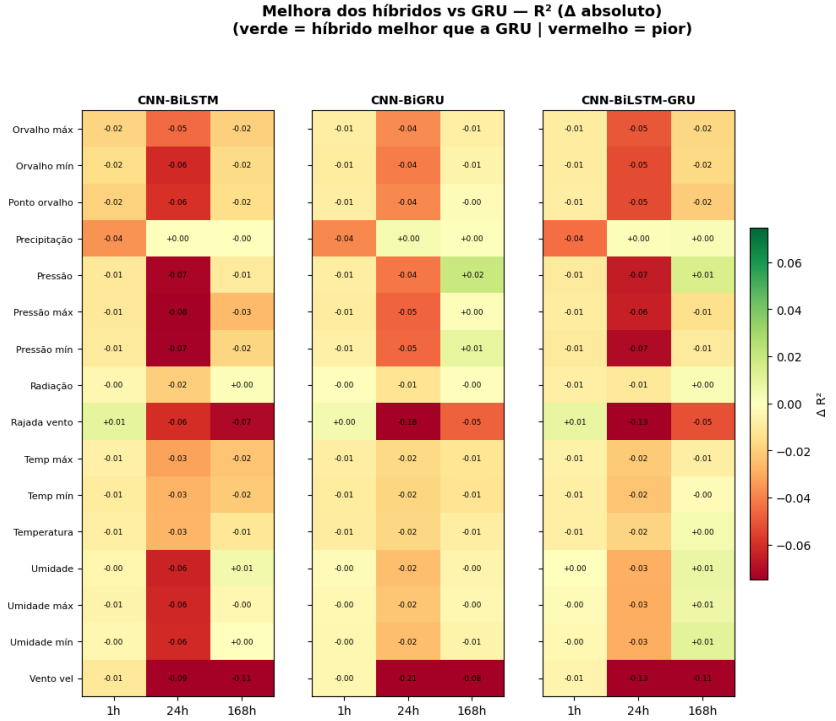
Esse comportamento descrito nos parágrafos anteriores é ilustrado na Figura 21 e 22, que apresenta a diferença de desempenho entre cada arquitetura híbrida e a GRU (arquitetura com melhor desempenho entre as recorrentes) para o coeficiente de determinação, por variável e horizonte. Predominam os valores próximos de zero (empate técnico), com a piora dos híbridos concentrada no horizonte de 24h.

Figura 21 - Heatmap da comparação de R² da GRU com as redes híbridas



Fonte: Autoria própria

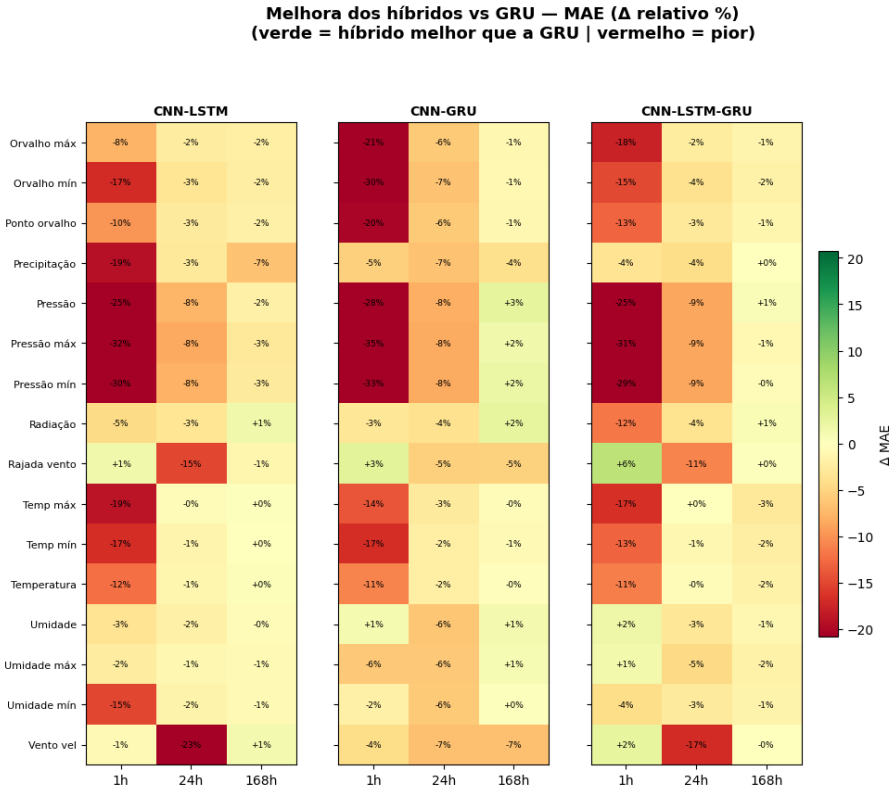
Figura 22 - Heatmap da comparação de R² da GRU com as redes híbridas bidirecionais



Fonte: Autoria própria

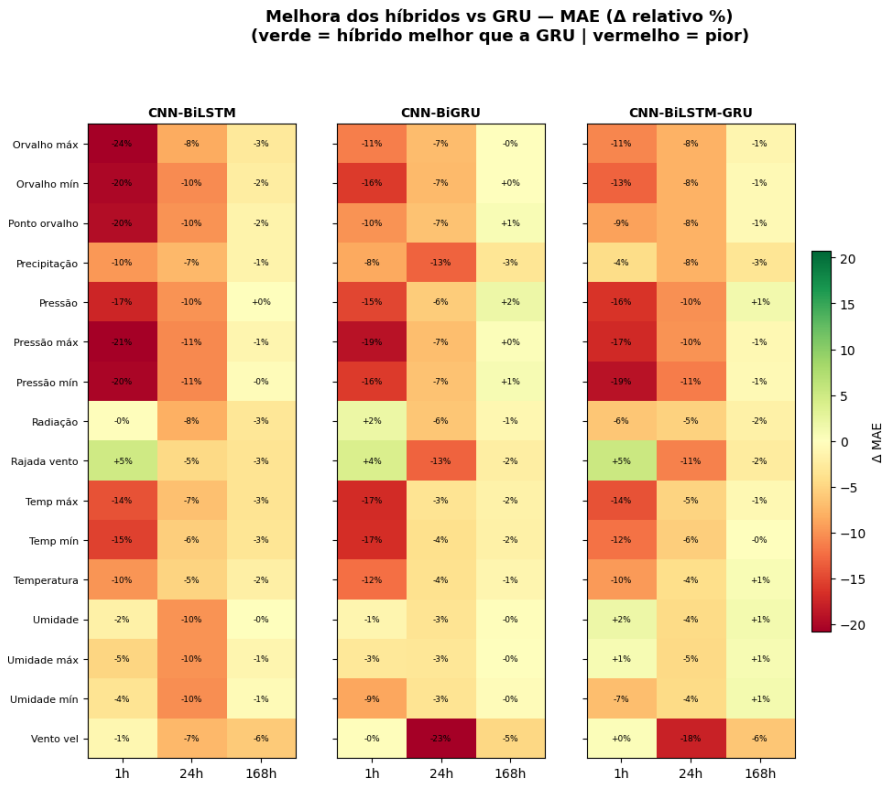
Além disso, as Figuras 23 e 24, apresentam a mesma comparação para o MAE. No horizonte de 1h, a maioria das variáveis apresenta desempenho pior nas arquiteturas híbridas, o que evidencia que a convolução prejudicou a previsão a curto prazo. Conforme o horizonte progride, essa diferença se reduz, mas permanece desfavorável aos híbridos no horizonte de 24h, em que a convolução ligada a rede recorrente não apresenta um ganho. Já para o horizonte de 168h, é predominante o empate técnico. A única exceção é a rajada máxima do vento, em que as arquiteturas híbridas apresentaram leve vantagem no MAE, especialmente no horizonte de 1h. No entanto, essa diferença é menor que a própria variação entre as execuções da arquitetura GRU, o que caracteriza também um empate técnico.

Figura 23 - Heatmap da comparação de MAE da GRU com as redes híbridas



Fonte: Autoria própria

Figura 24 - Heatmap da comparação de MAE da GRU com as redes híbridas bidirecionais



Fonte: Autoria própria

Além do padrão por horizonte, é possível notar um comportamento entre os tipos de variável. Elas podem ser divididas em dois grupos: suaves e fortemente autocorrelacionadas, como temperatura, radiação, pressão e umidade, e as caóticas e de alta variabilidade, como a velocidade do vento e a precipitação.

Nas variáveis suaves, nota-se que, a adição da convolução não trouxe ganhos, já que essas variáveis apresentam forte dependência do passado recente, a rede recorrente, ao processar a sequência temporal, já modela adequadamente seu comportamento, tornando a extração de padrões locais pela convolução redundante. Esse resultado é coerente com o observado por Marinho *et al.* (2023), que constataram que em séries autocorrelacionadas, o valor imediatamente anterior já constitui o melhor preditor, reduzindo a contribuição de camadas adicionais de extração de características.

Para as variáveis caóticas, a convolução também não trouxe ganho, porém por conta que nesses casos, dificilmente há um padrão local consistente a ser extraído. Como a velocidade do vento e a precipitação apresentam alta variabilidade e baixa previsibilidade, os filtros convolucionais não encontram padrões locais para capturar.

Os resultados apresentados podem ser compreendidos a partir das características fundamentais de cada tipo de arquitetura. Conforme discutido na seção 2.2, a camada convolucional foi originalmente concebida para extrair padrões locais e espaciais, aplicando filtros que percorrem os dados em busca de estruturas que se repetem em diferentes posições. Já as camadas recorrentes foram projetadas especificamente para modelar dependências temporais, processando a sequência de forma ordenada e mantendo um estado interno que carrega informação ao longo do tempo.

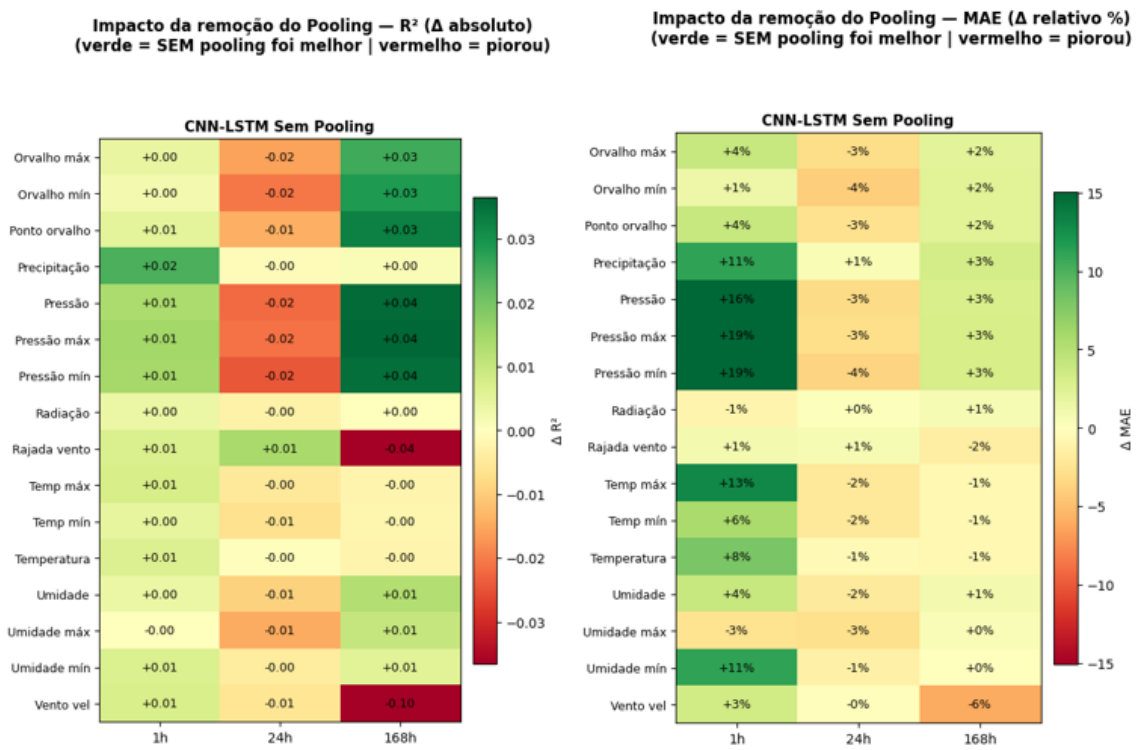
Em uma série temporal de ponto único, como a utilizada neste trabalho, dois fatores podem limitar a contribuição da convolução. A não utilização de uma estrutura espacial a ser explorada, já que os dados provêm de uma única estação, formando uma sequência unidimensional, e não uma grade espacial bidimensional para a qual a convolução oferece maior vantagem. Segundo, os padrões locais que a convolução extrairia ao longo do eixo temporal já são capturados pela camada recorrente, que modela nativamente a relação entre instantes vizinhos. Dessa forma, a convolução tende a adicionar parâmetros ao modelo sem fornecer informação complementar relevante, o que pode ajudar a explicar a ausência de ganho observada. Em uma série sem dependência espacial, o filtro unidimensional opera como um extrator de padrão local, com comportamento próximo ao de um filtro passa-baixa. Para variáveis de autocorrelação imediata alta, como temperatura e radiação, ele acrescenta parâmetros livres sem injetar sinal novo, podendo inclusive atenuar variações de alta frequência que seriam relevantes para a previsão. Esse comportamento oferece uma explicação possível para a ausência de ganho da convolução 1D observada nos experimentos.

A esse comportamento, já discutido anteriormente, soma-se a ausência de estrutura espacial, as redes convolucionais tendem a apresentar maior vantagem quando aplicadas a dados com variação no espaço, como mapas ou imagens de radar. Como este trabalho analisa os dados históricos de uma única estação meteorológica isolada, o ambiente favorável à extração de padrões pela convolução não está presente.

Logo, neste trabalho, nota-se que a CNN aliada ao *Average Pooling* provocou uma compressão dos dados ao tentar captar padrões locais. Essa compressão gerou um efeito de suavização (smoothing) que tende a atenuar variações finas e transições abruptas de curto prazo. Por outro lado, ao não utilizar a convolução seguida de *pooling* como etapa inicial, as redes recorrentes processam os dados instante a instante, preservando a resolução temporal completa da sequência. Dessa forma, evita-se que o modelo receba um resumo temporalmente impreciso, o que garante um acerto mais acurado na previsão do momento exato em que o evento meteorológico ocorre.

Para efeito de comparação na verificação da utilização de *pooling*, a Figura 25 mostra a comparação dos resultados das métricas de R^2 e MAE para uma mesma arquitetura CNN-LSTM, porém com e sem a adição de *Average Pooling*. Os resultados demonstram que o uso do *pooling* depende do horizonte, no curto prazo (1h), sua utilização prejudica a previsão, enquanto no horizonte de 168h os efeitos são mistos, chegando a ser levemente benéfico para os horizontes de 24h. Esse impacto negativo valida a afirmação de Li et al. (2025), discutida na Seção 2.5.2, de que o *pooling* comprime a série e descarta detalhes finos que são fundamentais para previsões imediatas. Já para os horizontes mais longos, em que esses detalhes finos são menos relevantes, a compressão deixa de ser prejudicial. No entanto, mesmo quando a camada de *pooling* é totalmente removida, o modelo híbrido ainda não consegue superar as redes recorrentes puras, ele apenas se aproxima do desempenho delas. Isso indica que a falta de ganho não é culpa exclusiva do descarte de informação causado pelo *pooling*, mas também da utilização da convolução ligada à rede recorrente.

Figura 25 - Heatmap da comparação de R^2 e MAE da CNN-LSTM sem *Average pooling* com a CNN-LSTM com *Average Pooling*



Fonte: Autoria própria

O teste com e sem *pooling* mostra que a perda de informação explica parte do problema, porém existem outros possíveis fatores. Em vez de o modelo receber informação de menos, ele tem capacidade de sobra. O bloco convolucional acrescenta parâmetros à rede, e cada parâmetro a mais precisa ser estimado a partir dos dados. Quando a variável tem sinal claro, isso não é problema. Quando o sinal é fraco e ruidoso, esses parâmetros tendem a se ajustar a variações do conjunto de treino que não se repetem no teste.

Petropoulos *et al.* (2022) explicam esse comportamento pelo compromisso entre viés e variância. O erro de um modelo tem três fontes, sendo ruído, variância e viés. O ruído é irredutível, existe mesmo quando o processo que gera os dados é conhecido. A variância vem da necessidade de estimar os parâmetros e cresce com a complexidade do modelo. O viés vem da má especificação e é maior em modelos simples. Por isso, até determinado limite, mais complexidade pode beneficiar o modelo, e a partir dele, passa a prejudicá-lo. O limite depende de quanta informação os dados carregam sobre o sinal. Quando o ruído domina, mesmo uma série longa oferece pouca informação útil, e modelos mais simples tendem a se sair melhor.

Isso ajuda a entender os resultados por tipo de variável. Nas variáveis suaves, a LSTM pura parece estar perto do ponto ótimo, e o bloco convolucional a empurra para além dele. Nas variáveis caóticas, o ponto ótimo desloca-se ainda mais à esquerda, favorecendo modelos ainda mais simples. Dois resultados deste trabalho são coerentes com isso: em 24h, a CNN-LSTM obteve R^2 de 0,0668 na velocidade do vento, contra 0,2513 da LSTM pura; e o ARIMA, um modelo estatístico univariado, superou todas as redes na precipitação em 1h e na velocidade do vento em 168h. Nos dois casos, o modelo mais simples venceu justamente nas variáveis menos previsíveis. No entanto, para as variáveis de vento, a comparação da distribuição entre treino e teste indicou uma queda anômala nessa variável e na rajada máxima, com a mediana da velocidade caindo de 2,1 m/s para 0,0 m/s. A inspeção dos dados brutos mostrou que a partir de fevereiro de 2023, as leituras nulas passam de 9% para 78%, com a rajada caindo na mesma proporção, o que é compatível com falha do anemômetro e não necessariamente variabilidade climática.

Para o horizonte de 24h, que foi onde as recorrentes puras mais superaram as híbridas, esse é o único horizonte em que a escolha da arquitetura influencia diretamente nos resultados de forma abrangente. Em 1h, a autocorrelação recente domina e o instante imediatamente anterior já constitui um preditor adequado, de modo que todas as arquiteturas alcançam desempenho semelhante. Em 168h, nenhuma delas recupera sinal útil e todas convergem para o padrão médio da época do ano, informação que as variáveis cíclicas já fornecem diretamente. O horizonte de 24h fica entre os dois extremos, pois exige reproduzir a forma completa do ciclo

diurno, definida pelos picos de máxima e mínima ao longo do dia. Ainda há sinal a ser extraído, mas ele não é fácil. Nesse horizonte, os dois fatores discutidos anteriormente se somam. O bloco convolucional com *pooling* reduz a resolução temporal da janela de entrada, de 336 para 84 instantes, atenuando justamente esses picos, e o que é descartado nessa etapa não é recuperável adiante. Além disso, o bloco convolucional acrescenta parâmetros que a rede precisa aprender, sem que ela tenha mais dados ou mais tempo de treino para isso. Quando o sinal é difícil de extrair, como em 24h, esses parâmetros a mais tendem a se ajustar ao ruído em vez do padrão, o que deixa muito mais visível neste horizonte do que em 1h e 168h.

Comportamento em que modelos mais simples desempenharam melhor foi relatado por Lütjens *et al.* (2025), que observaram o mesmo padrão em emulação climática, em que um modelo linear simples superou redes profundas na previsão espacial de precipitação. A causa apontada foi o ruído nos dados de treino, quando reduziram, aumentando de 3 para 50 simulações por cenário, a rede voltou a ser mais precisa. Ou seja, o que penalizava o modelo flexível era o ruído, e não a ausência de não-linearidade no fenômeno.

Um último ponto diz respeito à não-estacionariedade da série. Como mencionado no Capítulo 1, o clima apresenta concept drift, ou seja, suas propriedades estatísticas mudam ao longo dos anos. Como a divisão dos dados foi cronológica, o conjunto de teste corresponde ao período mais recente e pode ter distribuição diferente da vista no treino. Modelos mais simples tendem a ser menos sensíveis a esse tipo de mudança, o que pode ter impactado negativamente para os modelos mais complexos testados no trabalho.

Vale registrar que as duas dimensões de custo computacional apontam em direções distintas. Em tempo de treinamento, as arquiteturas híbridas tiveram vantagem nos horizontes de 24 e 168 horas, chegando a quase metade do tempo exigido pela recorrente pura. A explicação está na natureza sequencial da recorrência. A subamostragem por *Average Pooling* reduz a sequência de 336 para 84 instantes. O número de passos temporais processados em série domina o tempo total, mesmo que cada passo envolva mais parâmetros. Em número de parâmetros treináveis, a vantagem vai para as recorrentes puras. A GRU destaca-se como a arquitetura mais econômica do conjunto, com cerca de um quarto dos parâmetros das variantes híbridas bidirecionais. Como a GRU não foi superada em desempenho em nenhum dos horizontes avaliados, o resultado traz implicação prática direta. A arquitetura de melhor desempenho é também a de menor custo de armazenamento e de inferência. A GRU torna-se preferível em cenários com restrição de memória, como a execução embarcada em dispositivos de *Internet das Coisas*. O menor tempo de treinamento das híbridas só se torna vantagem prática quando o retreinamento frequente é um requisito do sistema.

6 CONCLUSÃO

Este trabalho comparou redes recorrentes puras e redes híbridas convolucional-recorrentes na previsão de uma série temporal meteorológica multivariada de ponto único. Usou-se a série horária da estação automática de Sorocaba - SP, entre 2000 e 2025, obtida do conjunto METBRA25Y. A partir de janelas de 336 horas, os modelos previram ao mesmo tempo todas as variáveis da entrada em três horizontes: 1 hora, 24 horas e 168 horas. Foram treinadas as arquiteturas LSTM e GRU e as híbridas CNN-LSTM, CNN-GRU, CNN-LSTM-GRU e suas variantes bidirecionais. O ARIMA foi usado como linha de base estatística, teve desempenho inferior às redes nas variáveis suaves, especialmente nos horizontes mais longos, mas superou todas elas na precipitação em 1 hora. Todas as redes usaram a mesma configuração de camadas, unidades e regularização. Assim, a única diferença entre os grupos foi a presença do bloco convolucional. A avaliação usou quatro métricas: MAE, R^2, SMAPE e DTW, com cinco execuções por modelo. O desvio padrão entre elas serviu para separar diferença real de empate técnico. O principal resultado mostra que a convolução não trouxe ganho. No horizonte de 1 hora, 14 das 16 variáveis terminaram em empate técnico e as duas restantes foram vencidas pelas recorrentes puras. No horizonte de 24 horas, as recorrentes puras venceram em 10 das 16 variáveis e os híbridos não superaram em nenhuma. No horizonte de 168 horas, o empate técnico voltou a predominar, em 12 das 16 variáveis.

A principal contribuição do trabalho está em testar uma ideia que a literatura costuma tratar como certa. Como mostrado na Seção 2.1, estudos recentes mostram as redes híbridas como uma evolução natural das recorrentes. Porém, na maioria das vezes, essa conclusão surge de problemas com estrutura espacial, como imagens de satélite e campos de radar, em que a convolução realmente tem o que extrair. Ao aplicar a mesma comparação a uma série de estação isolada, com topologias equivalentes e repetições suficientes para medir a variação entre execuções, foi possível mostrar que não ocorre um ganho expressivo nesse cenário. Nas variáveis suaves e muito autocorrelacionadas, como temperatura, pressão e umidade, a camada recorrente já modela sozinha a relação entre instantes vizinhos, e a convolução só repete a informação. Nas variáveis caóticas, como vento e precipitação, não existe um padrão local consistente para os filtros captarem. O teste com e sem a camada de *pooling* reforçou esse ponto. O *pooling* atrapalha a previsão de curto prazo porque comprime a série e descarta detalhes finos, mas removê-lo não foi suficiente para o híbrido superar a rede recorrente pura. Ou seja, o problema não é só a perda de informação, mas possivelmente a falta de estrutura espacial que justifique a convolução.

O trabalho também apresenta limitações que precisam ser consideradas. Os dados vêm de uma única estação meteorológica da cidade de Sorocaba. Por isso, os resultados valem para esse cenário e não podem ser aplicados automaticamente a outras regiões, outros climas ou a problemas com grade espacial. A suavização dos extremos foi usada só como critério de análise. Não houve nenhum mecanismo específico de treinamento para tratar disso, o que faz com que os modelos continuem subestimando os picos mais importantes. Os hiperparâmetros foram mantidos fixos para garantir a comparação justa entre as arquiteturas. Eles não foram otimizados para cada uma delas, o que pode ter deixado desempenho de lado. Por fim, duas variáveis se mostraram pouco previsíveis com essa abordagem. A precipitação teve R^2 próximo de zero já em 24 horas. A velocidade do vento teve R^2 negativo em 168 horas. Nesses casos, o baixo desempenho está ligado a natureza esparsa e instável das próprias variáveis, e não a escolha da arquitetura.

Como continuidade, identificam-se dois caminhos promissores. O primeiro é o *self-supervised learning*. O METBRA25Y reúne mais de 600 estações. Seria possível pré-treinar um modelo com todo esse histórico, sem precisar de rótulos, e depois ajustá-lo para uma estação específica. Isso ajudaria a contornar a escassez de exemplos de eventos extremos, que é uma das causas da suavização observada. O segundo é o uso de *Transformers*. O mecanismo de atenção liga diretamente instantes distantes da série, sem depender de um estado interno que se degrada ao longo do tempo. Como a maior perda de desempenho ocorreu justamente nos horizontes mais longos, essa é a limitação que poderia ser superada com esse tipo de arquitetura. Também seria útil repetir a comparação com dados de várias estações organizados em grade, para verificar se o ganho da convolução reaparece quando existe estrutura espacial. Além disso, seria importante testar funções de perda que deem mais peso aos valores extremos, atacando de forma direta o problema da suavização.

REFERÊNCIAS

AL-OMARI, H.; YAGHI, M. A.; ALRIFAI, L. An enhanced hybrid CNN-LSTM model for improved precipitation forecasting. **Algorithms**, [Basileia], v. 19, n. 5, art. 394, maio 2026. Disponível em: https://doi.org/10.3390/a19050394. Acesso em: 12 jan. 2026.

ALHADDAD, U.; BASUHAIL, A.; KHEMAKHEM, M.; EASSA, F. E.; JAMBI, K. Towards sustainable energy grids: a machine learning-based ensemble methods approach for outages estimation in extreme weather events. **Sustainability**, [Basileia], v. 15, n. 16, art. 12622, ago. 2023. Disponível em: https://doi.org/10.3390/su151612622. Acesso em: 12 jan. 2026.

BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning long-term dependencies with gradient descent is difficult. **IEEE Transactions on Neural Networks**, [Nova York], v. 5, n. 2, p. 157-166, mar. 1994. Disponível em: https://doi.org/10.1109/72.279181. Acesso em: 12 jan. 2026.

CENTRO NACIONAL DE MONITORAMENTO E ALERTAS DE DESASTRES NATURAIS. **Estado do clima, extremos de clima e desastres no Brasil 02|2026**. São José dos Campos: Cemaden, 2026. Disponível em: https://www.gov.br/cemaden/pt-br/assuntos/monitoramento/estado-do-clima-extremos-de-clima-e-desastres-no-brasil/estado-do-clima-extremos-de-clima-e-desastres-no-brasil-02-2026/relatorioclimaextremosdesastresbrasil2025.pdf. Acesso em: 20 ago. 2026.

CHO, K.; MERRIËNBOER, B. van; GULCEHRE, C.; BAHDANAU, D.; BOUGARES, F.; SCHWENK, H.; BENGIO, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, 2014, Doha. **Anais [...]**. Doha: Association for Computational Linguistics, 2014. p. 1724-1734. Disponível em: https://aclanthology.org/D14-1179/. Acesso em: 14 jan. 2026.

CHOLLET, F. **Deep learning with Python**. Shelter Island: Manning Publications, 2017.

DONAHUE, J.; HENDRICKS, L. A.; ROHRBACH, M.; VENUGOPALAN, S.; GUADARRAMA, S.; SAENKO, K.; DARRELL, T. Long-term recurrent convolutional networks for visual recognition and description. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [Nova York], v. 39, n. 4, p. 677-691, abr. 2017. Disponível em: https://doi.org/10.1109/TPAMI.2016.2599174. Acesso em: 22 fev. 2026.

ELABD, E.; HAMOUDA, H. M.; ALI, M. A. M.; FOUAD, Y. Climate change prediction in Saudi Arabia using a CNN GRU LSTM hybrid deep learning model in al Qassim region. **Scientific Reports**, [Londres], v. 15, n. 1, art. 16275, maio 2025. Disponível em: https://doi.org/10.1038/s41598-025-00607-0. Acesso em: 22 fev. 2026.

ELMAN, J. L. Finding structure in time. **Cognitive Science**, [S. l.], v. 14, n. 2, p. 179-211, mar. 1990. Disponível em: https://doi.org/10.1207/s15516709cog1402_1. Acesso em: 23 fev. 2026.

FANG, W.; XUE, Q.; SHEN, L.; SHENG, V. S. Survey on the application of deep learning in extreme weather prediction. **Atmosphere**, [Basileia], v. 12, n. 6, art. 661, jun. 2021. Disponível em: https://doi.org/10.3390/atmos12060661. Acesso em: 12 abr. 2026.

FUKUSHIMA, K. Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological Cybernetics**, [Berlim], v. 36, n. 4, p. 193-202, abr. 1980. Disponível em: https://doi.org/10.1007/BF00344251. Acesso em: 12 abr. 2026.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Cambridge: MIT Press, 2016.

GRAVES, A.; SCHMIDHUBER, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. **Neural Networks**, [Amsterdã], v. 18, n. 5-6, p. 602-610, jul. 2005. Disponível em: https://doi.org/10.1016/j.neunet.2005.06.042. Acesso em: 12 abr. 2026.

HE, K.; ZHANG, X.; REN, S.; SUN, J. **Deep residual learning for image recognition**. [S. l.]: arXiv, 2015. Disponível em: https://arxiv.org/abs/1512.03385. Acesso em: 13 abr. 2026.

HOA, T. T.; LE, T. M.; NGUYEN-DINH, C. H. Hybrid model of 1D-CNN and LSTM for forecasting Ethereum closing prices: a case study of temporal analysis. **International Journal of Information Technology**, [Singapura], v. 17, n. 7, p. 3999-4011, set. 2025. Disponível em: https://doi.org/10.1007/s41870-025-02472-6. Acesso em: 13 abr. 2026.

HOCHREITER, S. **Untersuchungen zu dynamischen neuronalen Netzen**. 1991. Diplomarbeit (Diploma em Informática) - Institut für Informatik, Technische Universität München, Munique, 1991. Disponível em: http://www.bioinf.jku.at/publications/older/3804.pdf. Acesso em: 07 mai. 2026.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, [Cambridge], v. 9, n. 8, p. 1735-1780, nov. 1997. Disponível em: https://doi.org/10.1162/neco.1997.9.8.1735. Acesso em: 08 mai. 2026.

HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. **Proceedings of the National Academy of Sciences**, [Washington], v. 79, n. 8, p. 2554-2558, abr. 1982. Disponível em: https://doi.org/10.1073/pnas.79.8.2554. Acesso em: 10 mai. 2026.

HUBEL, D. H.; WIESEL, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. **The Journal of Physiology**, [Londres], v. 160, n. 1, p. 106-154, jan. 1962. Disponível em: https://doi.org/10.1113/jphysiol.1962.sp006837. Acesso em: 14 mai. 2026.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. **Climate change 2023**: synthesis report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Genebra: IPCC, 2023. Disponível em: https://doi.org/10.59327/IPCC/AR6-9789291691647. Acesso em: 19 mai. 2026.

ISMAIL, W. N.; ALSALAMAH, H. A.; HASSAN, M. M.; MOHAMED, E. AUTO-HAR: an adaptive human activity recognition framework using an automated CNN architecture design. **Heliyon**, [Amsterdã], v. 9, n. 2, art. e13636, fev. 2023. Disponível em: https://doi.org/10.1016/j.heliyon.2023.e13636. Acesso em: 22 mai. 2026.

JIANG, J.; HUANG, Z.-G.; GREBOGI, C.; LAI, Y.-C. Predicting extreme events from data using deep machine learning: when and where. **Physical Review Research**, [College Park], v.

4, n. 2, art. 023028, abr. 2022. Disponível em: https://doi.org/10.1103/PhysRevResearch.4.023028. Acesso em: 23 mai. 2026.

JORDAN, M. I. **Serial order**: a parallel distributed processing approach. San Diego: University of California, Institute for Cognitive Science, 1986. (ICS Report, 8604). Disponível em: https://eric.ed.gov/?id=ED276754. Acesso em: 26 mai. 2026.

KAGABO, J.; KATTEL, G. R.; KAZORA, J.; SHANGWE, C. N.; HABIYAKARE, F. Application of machine learning algorithms in predicting extreme rainfall events in Rwanda. **Atmosphere**, [Basileia], v. 15, n. 6, art. 691, jun. 2024. Disponível em: https://doi.org/10.3390/atmos15060691. Acesso em: 01 jun. 2026.

KINGMA, D. P.; BA, J. Adam: a method for stochastic optimization. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 3., 2015, San Diego. **Anais [...]**. [S. l.: s. n.], 2015. Disponível em: https://arxiv.org/abs/1412.6980. Acesso em: 03 jun. 2026.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. In: ANNUAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 26., 2012, Lake Tahoe. **Anais [...]**. Red Hook: Curran Associates, 2012. p. 1097-1105. Disponível em: https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html. Acesso em: 03 jun. 2026.

KUMARI, S.; MUTHULAKSHMI, P. SARIMA model: an efficient machine learning technique for weather forecasting. **Procedia Computer Science**, [Amsterdã], v. 235, p. 656-670, 2024. Disponível em: https://doi.org/10.1016/j.procs.2024.04.064. Acesso em: 04 jun. 2026.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, [Nova York], v. 86, n. 11, p. 2278-2324, 1998. Disponível em: https://doi.org/10.1109/5.726791. Acesso em: 06 jun. 2026.

LI, X.; WU, J.; LIU, Y.; CHEN, Z.; WANG, L.; ZHANG, H. **A comprehensive study of CNN-LSTM architectures for weather prediction**. [S. l.]: ResearchGate, 2025. Preprint. Disponível em: https://doi.org/10.13140/RG.2.2.35641.89442. Acesso em: 06 jun. 2026.

LIMA CASTRO, M.; LUSQUINO FILHO, L.; VICHETE, W. D. **METBRA25Y**: Brazil surface meteorology archive with harmonized variables and quality control. Versão 1.0.0. [S. l.]: Zenodo, 2026. Conjunto de dados. Disponível em: https://doi.org/10.5281/zenodo.19964979. Acesso em: 07 jun. 2026.

LÜTJENS, B.; FERRARI, R.; WATSON-PARRIS, D.; SELIN, N. E. The impact of internal variability on benchmarking deep learning climate emulators. **Journal of Advances in Modeling Earth Systems**, [Washington], v. 17, n. 8, art. e2024MS004619, ago. 2025. Disponível em: https://doi.org/10.1029/2024MS004619. Acesso em: [data].
PETROPOULOS, F. *et al.* Forecasting: theory and practice. **International Journal of Forecasting**, [Amsterdã], v. 38, n. 3, p. 705-871, jul./set. 2022. Disponível em: https://doi.org/10.1016/j.ijforecast.2021.11.001. Acesso em: [data].

MARINHO, F. P.; ROCHA, P. A. C.; NETO, A. R. R.; BEZERRA, F. D. V. Short-term solar irradiance forecasting using CNN-1D, LSTM, and CNN-LSTM deep neural networks: a case

study with the Folsom (USA) dataset. **Journal of Solar Energy Engineering**, [Nova York], v. 145, n. 4, art. 041002, ago. 2023. Disponível em: https://doi.org/10.1115/1.4056122. Acesso em: 10 jun. 2026.

MATERIA, S.; PALMA GARCÍA, L.; STRAATEN, C. van; O, S.; MAMALAKIS, A.; CAVICCHIA, L.; COUMOU, D.; LUCA, P. de; KRETSCHMER, M.; DONAT, M. Artificial intelligence for climate prediction of extremes: state of the art, challenges, and future perspectives. **WIREs Climate Change**, [Hoboken], v. 15, n. 6, art. e914, nov./dez. 2024. Disponível em: https://doi.org/10.1002/wcc.914. Acesso em: 12 jun. 2026.

OLIVETTI, L.; MESSORI, G. Advances and prospects of deep learning for medium-range extreme weather forecasting. **Geoscientific Model Development**, [Göttingen], v. 17, n. 6, p. 2347-2358, mar. 2024. Disponível em: https://doi.org/10.5194/gmd-17-2347-2024. Acesso em: 15 jun. 2026.

PRATIHER, S.; CHATTORAJ, S.; AGARWAL, S.; BHATTACHARYA, S. Grading tumor malignancy via deep bidirectional LSTM on graph manifold encoded histopathological image. In: IEEE INTERNATIONAL CONFERENCE ON DATA MINING WORKSHOPS, 2018, Singapura. **Anais [...]**. [S. l.]: IEEE, 2018. p. 674-681. Disponível em: https://doi.org/10.1109/ICDMW.2018.00104. Acesso em: 20 jun. 2026.

RAFI, S. H.; NAHID-AL-MASOOD; DEEBA, S. R.; HOSSAIN, E. A short-term load forecasting method using integrated CNN and LSTM network. **IEEE Access**, [Piscataway], v. 9, p. 32436-32448, 2021. Disponível em: https://doi.org/10.1109/ACCESS.2021.3060654. Acesso em: 20 jun. 2026.

RAU, F.; SOTO, I.; ZABALA-BLANCO, D. Forecasting mobile network traffic based on deep learning networks. In: IEEE LATIN-AMERICAN CONFERENCE ON COMMUNICATIONS, 2021, Santo Domingo. **Anais [...]**. [S. l.]: IEEE, 2021. p. 1-6. Disponível em: https://doi.org/10.1109/LATINCOM53176.2021.9647788. Acesso em: 23 jun. 2026.

RAWSON, A.; BRITO, M.; SABEUR, Z.; TRAN-THANH, L. A machine learning approach for monitoring ship safety in extreme weather events. **Safety Science**, [Amsterdã], v. 141, art. 105336, set. 2021. Disponível em: https://doi.org/10.1016/j.ssci.2021.105336. Acesso em: 27 jun. 2026.

RODRIGUES, A. Desastres climáticos afetaram mais de 336 mil pessoas no país, em 2025. **Agência Brasil**, Brasília, 27 fev. 2026. Disponível em: https://agenciabrasil.ebc.com.br/meio-ambiente/noticia/2026-02/desastres-climaticos-afetaram-mais-de-336-mil-pessoas-no-pais-em-2025. Acesso em: 01 jul. 2026.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-Net: convolutional networks for biomedical image segmentation. In: INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION, 18., 2015, Munique. **Anais [...]**. Cham: Springer, 2015. p. 234-241. Disponível em: https://doi.org/10.1007/978-3-319-24574-4_28. Acesso em: 01 jul. 2026.

SAINATH, T. N.; VINYALS, O.; SENIOR, A.; SAK, H. Convolutional, long short-term memory, fully connected deep neural networks. In: IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 2015, South

Brisbane. **Anais [...]**. [S. l.]: IEEE, 2015. p. 4580-4584. Disponível em: https://doi.org/10.1109/ICASSP.2015.7178838. Acesso em: 05 jul. 2026.

SATTARI, A.; FOROUMANDI, E.; GAVAHI, K.; MORADKHANI, H. A probabilistic machine learning framework for daily extreme events forecasting. **Expert Systems with Applications**, [Amsterdã], v. 265, art. 126004, mar. 2025. Disponível em: https://doi.org/10.1016/j.eswa.2024.126004. Acesso em: 07 jul. 2026.

SCHUSTER, M.; PALIWAL, K. K. Bidirectional recurrent neural networks. **IEEE Transactions on Signal Processing**, [Nova York], v. 45, n. 11, p. 2673-2681, nov. 1997. Disponível em: https://doi.org/10.1109/78.650093. Acesso em: 12 jul. 2026.

SHI, X.; CHEN, Z.; WANG, H.; YEUNG, D.-Y.; WONG, W.-K.; WOO, W.-C. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In: ANNUAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 29., 2015, Montreal. **Anais [...]**. Red Hook: Curran Associates, 2015. p. 802-810. Disponível em: https://proceedings.neurips.cc/paper/2015/hash/07563a3fe3bbe7e3ba84431ad9d055af-Abstract.html. Acesso em: 15 jul. 2026.

SSEBYALA, S. N.; KINTU, T. M.; MUGANZI, D. J.; DRESSER, C.; DEMETRES, M. R.; LAI, Y.; MERCY, K.; LI, C.; WANG, F.; SETOGUCHI, S.; CELI, L. A.; GHOSH, A. K. Use of machine learning tools to predict health risks from climate-sensitive extreme weather events: a scoping review. **PLOS Climate**, [São Francisco], v. 3, n. 1, art. e0000338, jan. 2024. Disponível em: https://doi.org/10.1371/journal.pclm.0000338. Acesso em: 17 jul. 2026.

VERMA, S.; SRIVASTAVA, K.; TIWARI, A.; VERMA, S. **Deep learning techniques in extreme weather events**: a review. [S. l.]: arXiv, 2023. Disponível em: https://doi.org/10.48550/arXiv.2308.10995. Acesso em: 20 jul. 2026.

WIDIPUTRA, H.; JUWONO, E. Parallel multivariate deep learning models for time-series prediction: a comparative analysis in Asian stock markets. **IAES International Journal of Artificial Intelligence**, [Yogyakarta], v. 13, n. 1, p. 475-486, mar. 2024. Disponível em: https://doi.org/10.11591/ijai.v13.i1.pp475-486. Acesso em: 20 jul. 2026.

YADAV, K.; MALVIYA, S.; TIWARI, A. K. Improving weather forecasting in remote regions through machine learning. **Atmosphere**, [Basileia], v. 16, n. 5, art. 587, maio 2025. Disponível em: https://doi.org/10.3390/atmos16050587. Acesso em: 22 jul. 2026.

ZHA, W.; LIU, Y.; WAN, Y.; LUO, R.; LI, D.; YANG, S.; XU, Y. Forecasting monthly gas field production based on the CNN-LSTM model. **Energy**, [Amsterdã], v. 260, art. 124889, dez. 2022. Disponível em: https://doi.org/10.1016/j.energy.2022.124889. Acesso em: 22 jul. 2026.

ZHANG, H.; LIU, Y.; ZHANG, C.; LI, N. Machine learning methods for weather forecasting: a survey. **Atmosphere**, [Basileia], v. 16, n. 1, art. 82, jan. 2025. Disponível em: https://doi.org/10.3390/atmos16010082. Acesso em: 22 jul. 2026.

APÊNDICE A – RESULTADOS COMPLETOS

Neste apêndice, é encontrado todos os resultados para todas as métricas calculadas para todos os horizontes de predição em todas as arquiteturas que foram realizados os experimentos, assim como suas curvas de predição.

O código completo para o pipeline encontra-se disponível em: https://colab.research.google.com/drive/1e2fir1QcvvEjsfqsFQZpSsBk6q4epdFL?authuser=1#scrollTo=-rr1BxGF5Asa

LSTM

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.5411 ± 0.1233	0.9665 ± 0.0119	0.06 ± 0.01	0.4056 ± 0.1394
Pressão atmosférica máxima na hora anterior (mB)	0.4975 ± 0.1244	0.9712 ± 0.0106	0.05 ± 0.01	0.3791 ± 0.1130
Pressão atmosférica mínima na hora anterior (mB)	0.5146 ± 0.1371	0.9706 ± 0.0123	0.05 ± 0.02	0.3897 ± 0.1390
Precipitação (mm)	0.1384 ± 0.0038	0.1075 ± 0.0187	95.88 ± 42.52	0.1270 ± 0.0079
Radiação global (KJ/m²)	135.5241 ± 4.7083	0.9404 ± 0.0044	47.69 ± 19.99	224.1966 ± 115.2630
Temperatura do ar - bulbo seco, horária (°C)	0.7950 ± 0.0250	0.9594 ± 0.0015	4.09 ± 0.19	0.5796 ± 0.0248
Temperatura do ponto de orvalho (°C)	0.6684 ± 0.0586	0.9438 ± 0.0068	5.14 ± 0.48	0.4350 ± 0.0649
Temperatura máxima na hora anterior (°C)	0.6512 ± 0.0489	0.9758 ± 0.0026	3.27 ± 0.30	0.4978 ± 0.0353
Temperatura mínima na hora anterior (°C)	0.6904 ± 0.0571	0.9686 ± 0.0027	3.73 ± 0.36	0.5098 ± 0.0497
Temperatura de orvalho máxima na hora anterior (°C)	0.5837 ± 0.0783	0.9592 ± 0.0082	4.26 ± 0.62	0.4397 ± 0.0730
Temperatura de orvalho mínima na hora anterior (°C)	0.5847 ± 0.1220	0.9595 ± 0.0121	4.85 ± 1.01	0.4290 ± 0.1312

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.5686 ± 0.5454	0.9564 ± 0.0114	3.77 ± 0.80	2.2854 ± 1.0006
Umidade relativa mínima na hora anterior (%)	2.3327 ± 0.1310	0.9730 ± 0.0032	3.86 ± 0.22	1.9491 ± 0.1413
Umidade relativa do ar, horária (%)	3.1002 ± 0.2579	0.9432 ± 0.0063	4.84 ± 0.42	2.3269 ± 0.2883
Vento, rajada máxima (m/s)	1.1061 ± 0.0706	0.7425 ± 0.0171	73.21 ± 2.91	1.0095 ± 0.1354
Vento, velocidade horária (m/s)	0.5304 ± 0.0534	0.7004 ± 0.0132	116.50 ± 11.50	0.4439 ± 0.0869
Direção do vento (cosseno)	0.3161 ± 0.0103	0.5501 ± 0.0198	67.80 ± 3.13	0.2321 ± 0.0121
Direção do vento (seno)	0.5045 ± 0.0084	0.3185 ± 0.0191	122.55 ± 2.15	0.4543 ± 0.0415

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.6051 ± 0.0425	0.6903 ± 0.0151	0.17 ± 0.00	0.7657 ± 0.0977
Pressão atmosférica máxima na hora anterior (mB)	1.5738 ± 0.0421	0.6977 ± 0.0145	0.17 ± 0.00	0.7447 ± 0.0745
Pressão atmosférica mínima na hora anterior (mB)	1.5846 ± 0.0395	0.7012 ± 0.0140	0.17 ± 0.00	0.7387 ± 0.0634
Precipitação (mm)	0.1865 ± 0.0172	0.0016 ± 0.0040	166.67 ± 20.22	0.1870 ± 0.0175
Radiação global (KJ/m²)	243.9957 ± 4.3743	0.8322 ± 0.0019	58.26 ± 22.14	485.3137 ± 53.6011
Temperatura do ar - bulbo seco, horária (°C)	2.0643 ± 0.0906	0.7428 ± 0.0123	10.47 ± 0.53	2.3185 ± 0.3169
Temperatura do ponto de orvalho (°C)	1.5923 ± 0.0628	0.6892 ± 0.0151	11.99 ± 0.52	0.8996 ± 0.0618
Temperatura máxima na hora anterior (°C)	2.0874 ± 0.0926	0.7481 ± 0.0143	10.20 ± 0.51	2.4302 ± 0.3503
Temperatura mínima na hora anterior (°C)	2.0070 ± 0.1214	0.7474 ± 0.0206	10.63 ± 0.74	2.1067 ± 0.4060
Temperatura de orvalho máxima na hora anterior (°C)	1.5125 ± 0.0614	0.7124 ± 0.0157	10.81 ± 0.47	0.8535 ± 0.0586
Temperatura de orvalho mínima na hora anterior (°C)	1.6310 ± 0.0744	0.6823 ± 0.0176	12.96 ± 0.66	0.9020 ± 0.0850
Umidade relativa máxima na hora anterior (%)	7.3632 ± 0.4607	0.6404 ± 0.0299	11.03 ± 0.61	8.5397 ± 1.2995

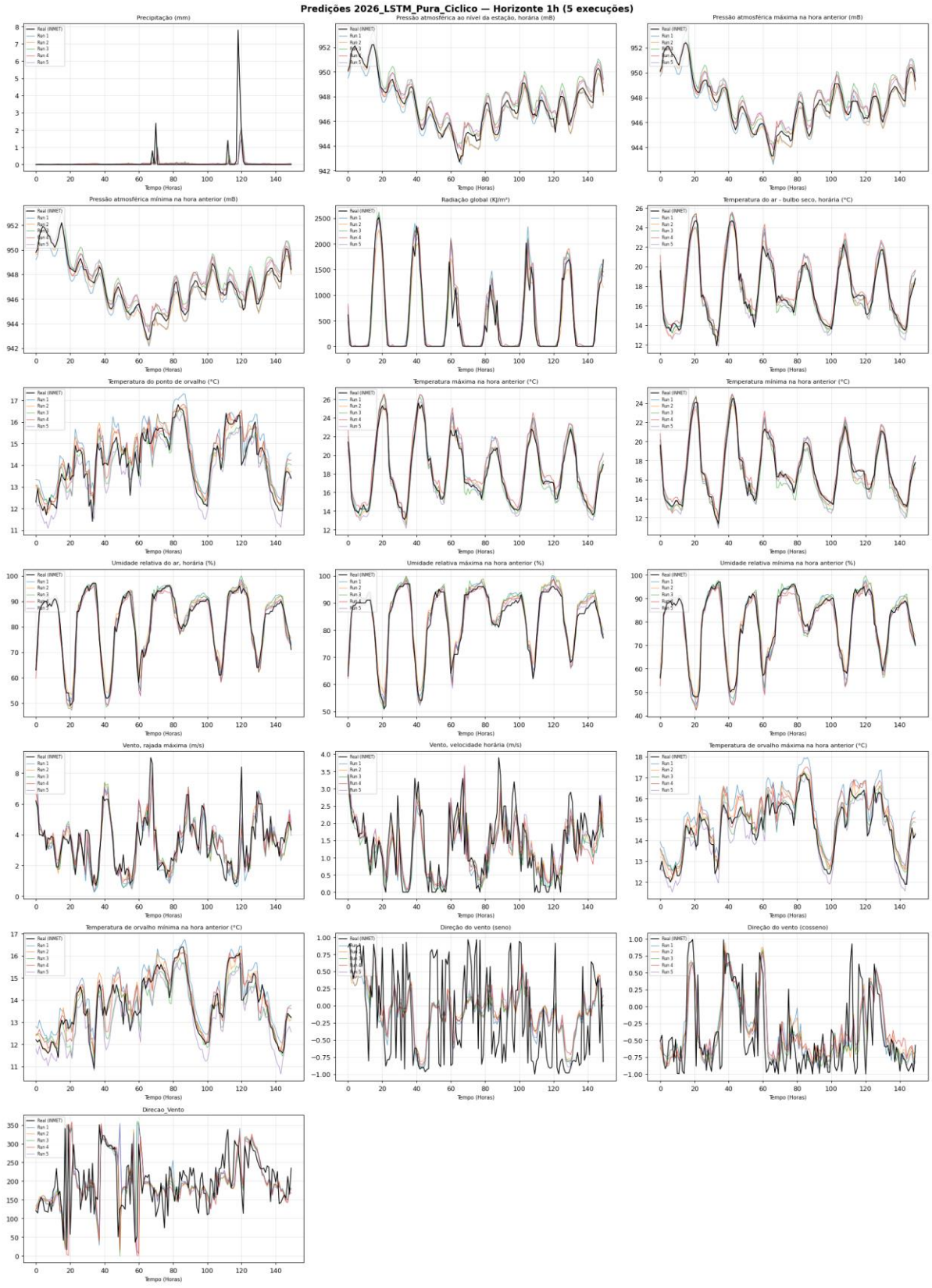
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	8.2305 ± 0.3247	0.6593 ± 0.0197	13.47 ± 0.50	9.0329 ± 1.1587
Umidade relativa do ar, horária (%)	7.8840 ± 0.3137	0.6480 ± 0.0214	12.39 ± 0.47	8.8675 ± 1.1853
Vento, rajada máxima (m/s)	2.3190 ± 0.1532	0.2048 ± 0.0781	97.42 ± 2.09	2.1220 ± 0.1127
Vento, velocidade horária (m/s)	1.0248 ± 0.0450	0.2513 ± 0.0264	138.15 ± 2.02	0.8873 ± 0.0540
Direção do vento (cosseno)	0.4386 ± 0.0043	0.2259 ± 0.0232	93.57 ± 1.44	0.3935 ± 0.0192
Direção do vento (seno)	0.6029 ± 0.0056	0.1244 ± 0.0138	148.07 ± 5.40	0.6166 ± 0.0493

Métricas — Horizonte 168h

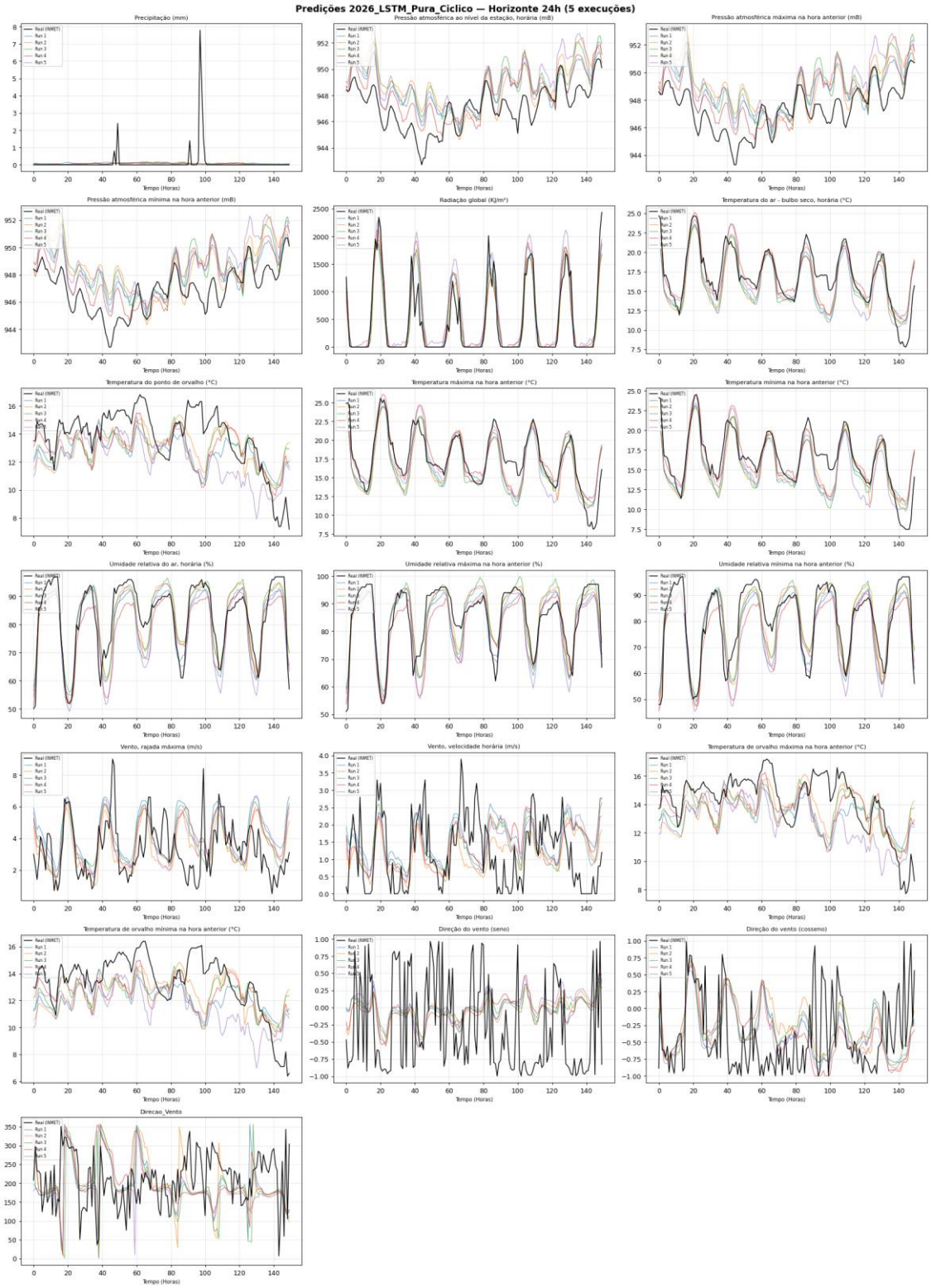
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.3342 ± 0.0423	0.3424 ± 0.0247	0.25 ± 0.01	1.7122 ± 0.0594
Pressão atmosférica máxima na hora anterior (mB)	2.3088 ± 0.0604	0.3446 ± 0.0312	0.25 ± 0.01	1.6676 ± 0.0293
Pressão atmosférica mínima na hora anterior (mB)	2.3240 ± 0.0558	0.3518 ± 0.0315	0.25 ± 0.01	1.6842 ± 0.0475
Precipitação (mm)	0.1803 ± 0.0214	0.0004 ± 0.0041	163.55 ± 32.19	0.1823 ± 0.0215
Radiação global (KJ/m²)	275.6701 ± 14.6161	0.8081 ± 0.0100	79.14 ± 15.13	474.3191 ± 15.9818
Temperatura do ar - bulbo seco, horária (°C)	2.5080 ± 0.0597	0.6306 ± 0.0181	12.70 ± 0.30	2.8756 ± 0.1479
Temperatura do ponto de orvalho (°C)	2.0398 ± 0.0515	0.5221 ± 0.0131	15.14 ± 0.37	1.5755 ± 0.0514
Temperatura máxima na hora anterior (°C)	2.5591 ± 0.0544	0.6320 ± 0.0166	12.48 ± 0.27	2.9327 ± 0.0713
Temperatura mínima na hora anterior (°C)	2.4434 ± 0.0565	0.6355 ± 0.0184	12.93 ± 0.30	2.7423 ± 0.1111
Temperatura de orvalho máxima na hora anterior (°C)	1.9713 ± 0.0551	0.5416 ± 0.0141	13.93 ± 0.37	1.4697 ± 0.0266
Temperatura de orvalho mínima na hora anterior (°C)	2.1026 ± 0.0523	0.5088 ± 0.0153	16.44 ± 0.40	1.6205 ± 0.0324
Umidade relativa máxima na hora anterior (%)	8.3086 ± 0.2873	0.5635 ± 0.0348	12.57 ± 0.40	10.3034 ± 1.8610

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.4710 ± 0.2402	0.5760 ± 0.0245	15.67 ± 0.37	11.5786 ± 1.4810
Umidade relativa do ar, horária (%)	9.0378 ± 0.2570	0.5676 ± 0.0275	14.30 ± 0.39	12.0077 ± 1.5419
Vento, rajada máxima (m/s)	2.9660 ± 0.1697	-0.1686 ± 0.1261	105.21 ± 1.20	2.6145 ± 0.2733
Vento, velocidade horária (m/s)	1.4138 ± 0.1103	-0.1231 ± 0.1226	142.39 ± 0.49	1.1213 ± 0.1853
Direção do vento (cosseno)	0.4617 ± 0.0012	0.1587 ± 0.0120	99.60 ± 1.61	0.4635 ± 0.0379
Direção do vento (seno)	0.6388 ± 0.0032	0.0470 ± 0.0047	157.51 ± 3.61	0.6491 ± 0.0145

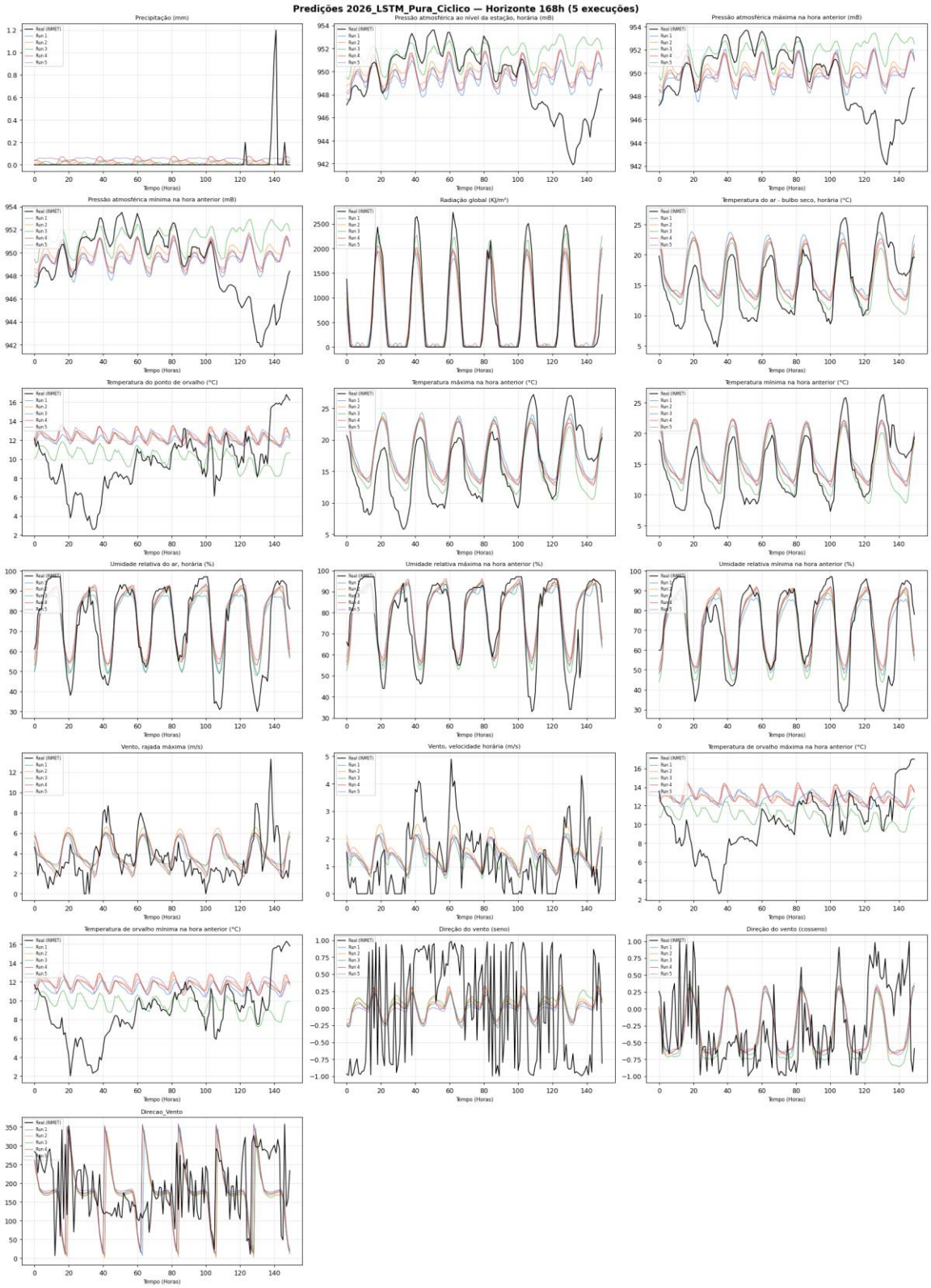
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



GRU

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.5071 ± 0.1066	0.9683 ± 0.0133	0.05 ± 0.01	0.3583 ± 0.0834
Pressão atmosférica máxima na hora anterior (mB)	0.4597 ± 0.1145	0.9728 ± 0.0132	0.05 ± 0.01	0.3351 ± 0.0890
Pressão atmosférica mínima na hora anterior (mB)	0.4654 ± 0.1220	0.9731 ± 0.0141	0.05 ± 0.01	0.3444 ± 0.0902
Precipitação (mm)	0.1474 ± 0.0092	0.1124 ± 0.0195	136.55 ± 23.41	0.1334 ± 0.0059
Radiação global (KJ/m²)	143.9326 ± 23.2124	0.9398 ± 0.0081	63.82 ± 27.25	232.7578 ± 283.6750
Temperatura do ar - bulbo seco, horária (°C)	0.7548 ± 0.0819	0.9620 ± 0.0060	3.86 ± 0.48	0.5382 ± 0.0406
Temperatura do ponto de orvalho (°C)	0.6140 ± 0.0456	0.9494 ± 0.0058	4.78 ± 0.40	0.3870 ± 0.0442
Temperatura máxima na hora anterior (°C)	0.6330 ± 0.1341	0.9758 ± 0.0074	3.15 ± 0.74	0.4897 ± 0.0833
Temperatura mínima na hora anterior (°C)	0.6401 ± 0.0629	0.9711 ± 0.0044	3.42 ± 0.34	0.4784 ± 0.0378
Temperatura de orvalho máxima na hora anterior (°C)	0.5413 ± 0.0355	0.9630 ± 0.0045	4.01 ± 0.32	0.3984 ± 0.0268
Temperatura de orvalho mínima na hora anterior (°C)	0.5210 ± 0.0643	0.9656 ± 0.0073	4.35 ± 0.56	0.3672 ± 0.0517

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.6826 ± 0.1536	0.9545 ± 0.0031	3.94 ± 0.22	2.1267 ± 0.2407
Umidade relativa mínima na hora anterior (%)	2.5617 ± 0.4969	0.9687 ± 0.0072	4.13 ± 0.73	2.3004 ± 0.8879
Umidade relativa do ar, horária (%)	3.2439 ± 0.1991	0.9408 ± 0.0036	5.02 ± 0.23	2.4180 ± 0.2351
Vento, rajada máxima (m/s)	1.2272 ± 0.1350	0.7107 ± 0.0289	77.04 ± 4.49	1.2426 ± 0.2851
Vento, velocidade horária (m/s)	0.5381 ± 0.0546	0.6991 ± 0.0197	120.09 ± 8.69	0.4375 ± 0.0839
Direção do vento (cosseno)	0.3156 ± 0.0062	0.5513 ± 0.0157	67.89 ± 1.38	0.2328 ± 0.0080
Direção do vento (seno)	0.5055 ± 0.0053	0.3194 ± 0.0101	122.67 ± 1.65	0.4331 ± 0.0074

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.6120 ± 0.0280	0.6900 ± 0.0106	0.17 ± 0.00	0.8094 ± 0.0810
Pressão atmosférica máxima na hora anterior (mB)	1.5869 ± 0.0277	0.6953 ± 0.0097	0.17 ± 0.00	0.8016 ± 0.0788
Pressão atmosférica mínima na hora anterior (mB)	1.5967 ± 0.0258	0.6997 ± 0.0091	0.17 ± 0.00	0.7939 ± 0.0724
Precipitação (mm)	0.1787 ± 0.0084	0.0015 ± 0.0026	162.82 ± 18.65	0.1784 ± 0.0096
Radiação global (KJ/m²)	235.3160 ± 7.2983	0.8413 ± 0.0023	70.82 ± 12.52	367.6697 ± 73.0489
Temperatura do ar - bulbo seco, horária (°C)	1.9117 ± 0.1005	0.7715 ± 0.0156	9.66 ± 0.58	1.7287 ± 0.2915
Temperatura do ponto de orvalho (°C)	1.5850 ± 0.0786	0.6840 ± 0.0256	11.86 ± 0.55	0.9147 ± 0.0953
Temperatura máxima na hora anterior (°C)	1.9332 ± 0.0991	0.7760 ± 0.0142	9.41 ± 0.54	1.7822 ± 0.2381
Temperatura mínima na hora anterior (°C)	1.8409 ± 0.0848	0.7797 ± 0.0141	9.71 ± 0.50	1.7169 ± 0.4277
Temperatura de orvalho máxima na hora anterior (°C)	1.5202 ± 0.0927	0.7019 ± 0.0315	10.80 ± 0.61	0.8859 ± 0.1100
Temperatura de orvalho mínima na hora anterior (°C)	1.6128 ± 0.0722	0.6818 ± 0.0221	12.75 ± 0.53	0.9408 ± 0.0962
Umidade relativa máxima na hora anterior (%)	6.8372 ± 0.1192	0.6837 ± 0.0126	10.23 ± 0.18	6.0021 ± 1.1998

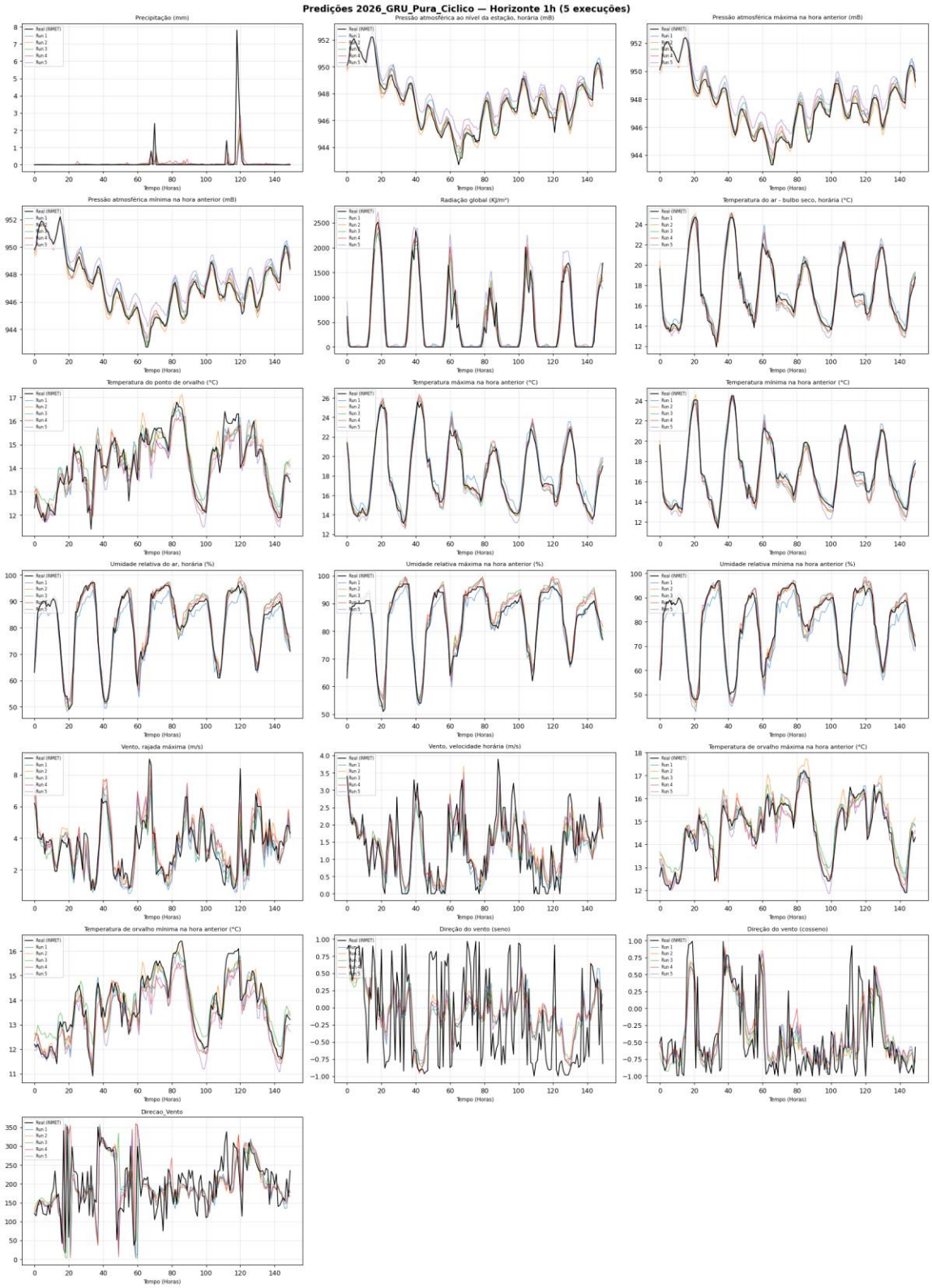
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	7.7685 ± 0.0847	0.6901 ± 0.0089	12.65 ± 0.15	7.2891 ± 1.2052
Umidade relativa do ar, horária (%)	7.3931 ± 0.0932	0.6855 ± 0.0103	11.57 ± 0.15	6.9540 ± 0.7146
Vento, rajada máxima (m/s)	2.2887 ± 0.1188	0.2271 ± 0.0666	97.27 ± 1.51	2.0585 ± 0.1146
Vento, velocidade horária (m/s)	1.0166 ± 0.1405	0.2614 ± 0.0863	137.84 ± 2.47	0.8782 ± 0.1897
Direção do vento (cosseno)	0.4395 ± 0.0053	0.2382 ± 0.0408	95.17 ± 1.68	0.4010 ± 0.0284
Direção do vento (seno)	0.6037 ± 0.0027	0.1278 ± 0.0090	147.97 ± 1.93	0.6054 ± 0.0137

Métricas — Horizonte 168h

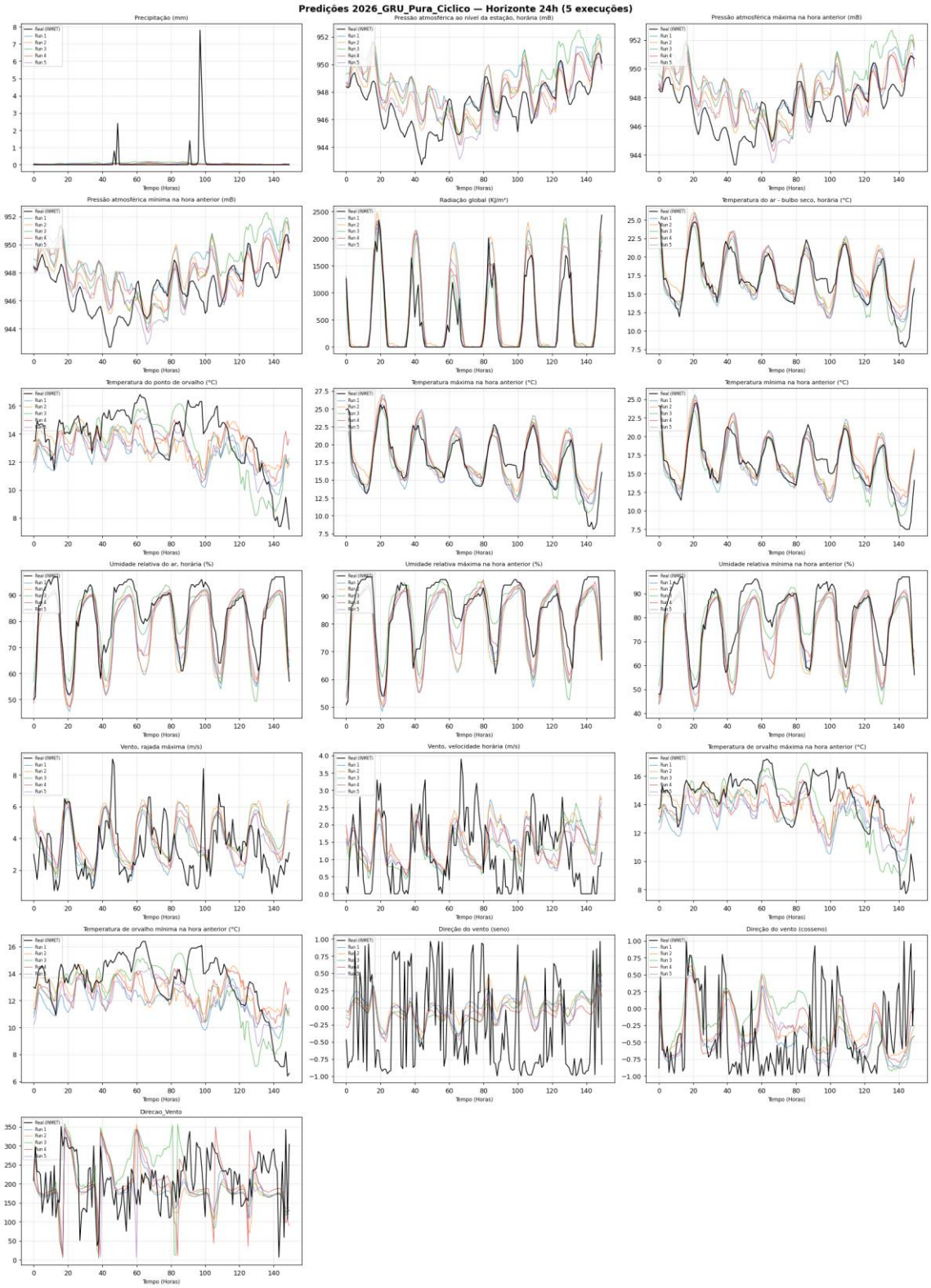
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.3217 ± 0.0880	0.3468 ± 0.0474	0.25 ± 0.01	1.6914 ± 0.0629
Pressão atmosférica máxima na hora anterior (mB)	2.2831 ± 0.0465	0.3604 ± 0.0263	0.24 ± 0.00	1.6906 ± 0.0706
Pressão atmosférica mínima na hora anterior (mB)	2.3117 ± 0.0464	0.3605 ± 0.0258	0.24 ± 0.01	1.7007 ± 0.0515
Precipitação (mm)	0.1769 ± 0.0210	-0.0024 ± 0.0024	178.34 ± 13.91	0.1773 ± 0.0200
Radiação global (KJ/m²)	275.9490 ± 18.8233	0.8017 ± 0.0211	69.40 ± 18.40	566.3708 ± 157.5617
Temperatura do ar - bulbo seco, horária (°C)	2.4907 ± 0.0411	0.6344 ± 0.0117	12.61 ± 0.21	2.8756 ± 0.1881
Temperatura do ponto de orvalho (°C)	2.0522 ± 0.0432	0.5257 ± 0.0153	15.18 ± 0.31	1.6298 ± 0.0282
Temperatura máxima na hora anterior (°C)	2.5325 ± 0.0522	0.6396 ± 0.0111	12.36 ± 0.27	2.8893 ± 0.1487
Temperatura mínima na hora anterior (°C)	2.4160 ± 0.0377	0.6429 ± 0.0102	12.79 ± 0.21	2.6461 ± 0.1173
Temperatura de orvalho máxima na hora anterior (°C)	1.9693 ± 0.0367	0.5481 ± 0.0155	13.87 ± 0.27	1.5216 ± 0.0130
Temperatura de orvalho mínima na hora anterior (°C)	2.1117 ± 0.0342	0.5151 ± 0.0131	16.45 ± 0.27	1.6557 ± 0.0515
Umidade relativa máxima na hora anterior (%)	8.3134 ± 0.2069	0.5660 ± 0.0191	12.57 ± 0.31	10.2372 ± 1.0997

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.5224 ± 0.2449	0.5728 ± 0.0164	15.74 ± 0.40	11.7964 ± 0.7334
Umidade relativa do ar, horária (%)	9.0777 ± 0.2663	0.5650 ± 0.0206	14.37 ± 0.41	11.1230 ± 1.2764
Vento, rajada máxima (m/s)	2.8569 ± 0.0584	-0.0854 ± 0.0432	104.39 ± 0.61	2.5276 ± 0.1703
Vento, velocidade horária (m/s)	1.4113 ± 0.0436	-0.0997 ± 0.0445	142.57 ± 0.16	1.0417 ± 0.2310
Direção do vento (cosseno)	0.4622 ± 0.0005	0.1634 ± 0.0038	100.68 ± 0.88	0.4437 ± 0.0423
Direção do vento (seno)	0.6413 ± 0.0012	0.0480 ± 0.0044	160.70 ± 1.62	0.6576 ± 0.0057

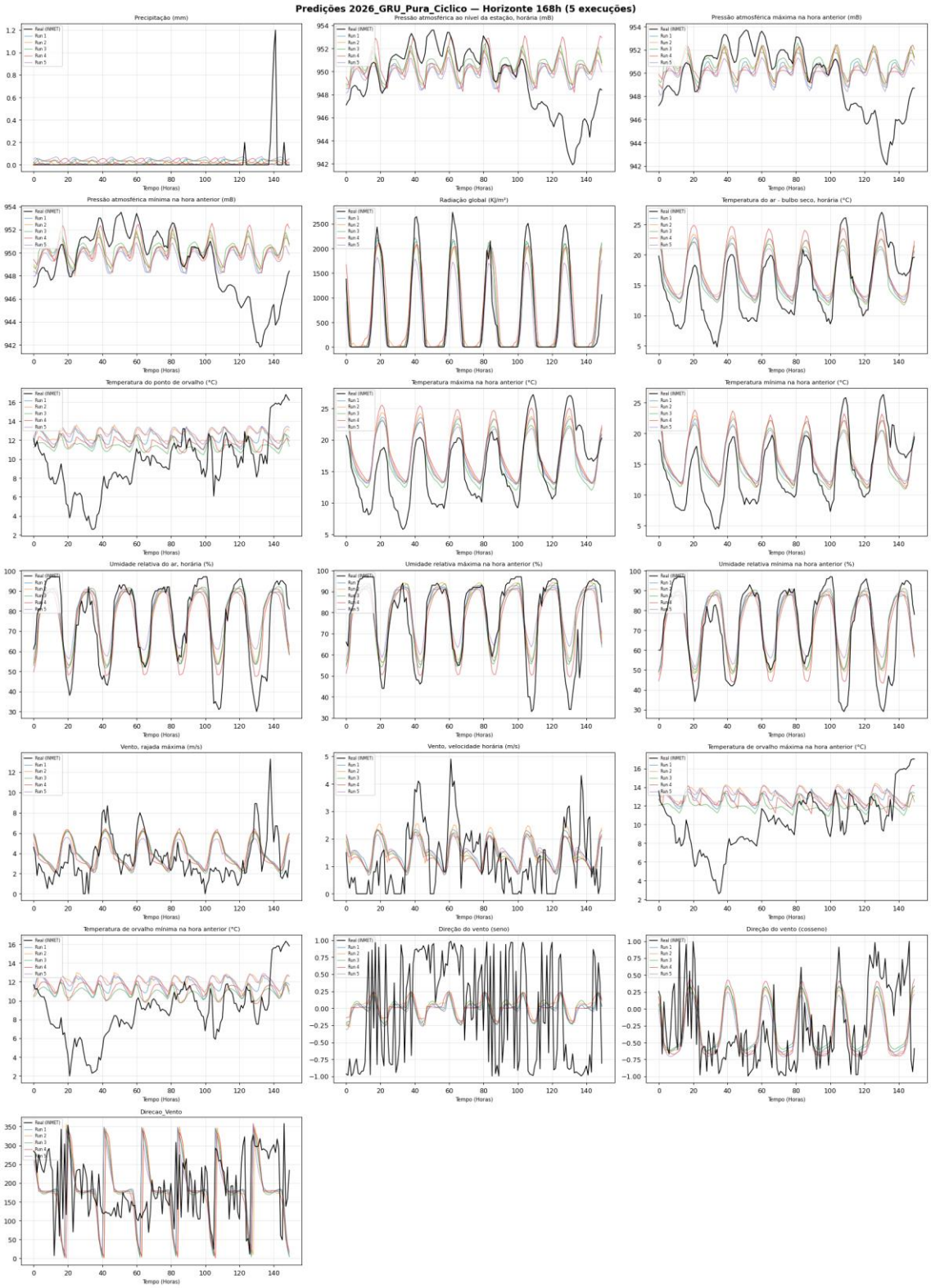
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-LSTM (com *pooling*)

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.6316 ± 0.0726	0.9521 ± 0.0094	0.07 ± 0.01	0.4570 ± 0.0755
Pressão atmosférica máxima na hora anterior (mB)	0.6068 ± 0.0793	0.9546 ± 0.0102	0.06 ± 0.01	0.4462 ± 0.0765
Pressão atmosférica mínima na hora anterior (mB)	0.6067 ± 0.0801	0.9563 ± 0.0097	0.06 ± 0.01	0.4490 ± 0.0925
Precipitação (mm)	0.1757 ± 0.0318	0.0630 ± 0.0210	142.52 ± 28.97	0.1681 ± 0.0281
Radiação global (KJ/m²)	150.4266 ± 8.2785	0.9326 ± 0.0032	46.84 ± 13.49	137.3254 ± 28.6792
Temperatura do ar - bulbo seco, horária (°C)	0.8468 ± 0.0643	0.9530 ± 0.0064	4.26 ± 0.30	0.6405 ± 0.1268
Temperatura do ponto de orvalho (°C)	0.6738 ± 0.0437	0.9405 ± 0.0055	5.15 ± 0.28	0.4404 ± 0.0349
Temperatura máxima na hora anterior (°C)	0.7521 ± 0.0968	0.9666 ± 0.0075	3.65 ± 0.49	0.6047 ± 0.1518
Temperatura mínima na hora anterior (°C)	0.7478 ± 0.0729	0.9613 ± 0.0064	3.94 ± 0.33	0.5806 ± 0.1108
Temperatura de orvalho máxima na hora anterior (°C)	0.5833 ± 0.0430	0.9563 ± 0.0053	4.19 ± 0.25	0.4164 ± 0.0298
Temperatura de orvalho mínima na hora anterior (°C)	0.6092 ± 0.0609	0.9546 ± 0.0067	5.03 ± 0.48	0.4287 ± 0.0493

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.7452 ± 0.2343	0.9503 ± 0.0063	4.12 ± 0.38	2.3159 ± 0.4117
Umidade relativa mínima na hora anterior (%)	2.9461 ± 0.3610	0.9604 ± 0.0069	4.77 ± 0.47	2.2251 ± 0.2720
Umidade relativa do ar, horária (%)	3.3544 ± 0.1659	0.9365 ± 0.0047	5.26 ± 0.27	2.3725 ± 0.1807
Vento, rajada máxima (m/s)	1.2092 ± 0.1326	0.7068 ± 0.0248	75.66 ± 5.02	1.2471 ± 0.2865
Vento, velocidade horária (m/s)	0.5420 ± 0.0331	0.6909 ± 0.0037	121.18 ± 7.67	0.4486 ± 0.0383
Direção do vento (cosseno)	0.3185 ± 0.0062	0.5502 ± 0.0112	68.73 ± 2.05	0.2438 ± 0.0122
Direção do vento (seno)	0.5052 ± 0.0075	0.3167 ± 0.0101	122.18 ± 2.91	0.4443 ± 0.0300

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7347 ± 0.0596	0.6381 ± 0.0227	0.18 ± 0.01	0.9046 ± 0.0475
Pressão atmosférica máxima na hora anterior (mB)	1.7211 ± 0.0752	0.6394 ± 0.0260	0.18 ± 0.01	0.8879 ± 0.0636
Pressão atmosférica mínima na hora anterior (mB)	1.7229 ± 0.0658	0.6474 ± 0.0244	0.18 ± 0.01	0.8912 ± 0.0572
Precipitação (mm)	0.1839 ± 0.0210	0.0013 ± 0.0040	155.11 ± 36.02	0.1825 ± 0.0209
Radiação global (KJ/m²)	243.1795 ± 15.6741	0.8394 ± 0.0045	78.70 ± 25.11	495.4438 ± 123.4620
Temperatura do ar - bulbo seco, horária (°C)	1.9347 ± 0.0659	0.7635 ± 0.0132	9.79 ± 0.34	2.0204 ± 0.2280
Temperatura do ponto de orvalho (°C)	1.6261 ± 0.0175	0.6673 ± 0.0077	12.13 ± 0.13	0.9829 ± 0.0655
Temperatura máxima na hora anterior (°C)	1.9406 ± 0.0583	0.7702 ± 0.0117	9.46 ± 0.29	2.0355 ± 0.0573
Temperatura mínima na hora anterior (°C)	1.8650 ± 0.0625	0.7719 ± 0.0131	9.86 ± 0.33	1.9202 ± 0.1886
Temperatura de orvalho máxima na hora anterior (°C)	1.5566 ± 0.0146	0.6868 ± 0.0091	11.04 ± 0.12	0.9894 ± 0.0654
Temperatura de orvalho mínima na hora anterior (°C)	1.6667 ± 0.0143	0.6596 ± 0.0079	13.12 ± 0.10	1.0042 ± 0.0557
Umidade relativa máxima na hora anterior (%)	6.9117 ± 0.0567	0.6762 ± 0.0100	10.29 ± 0.07	6.4129 ± 0.9193

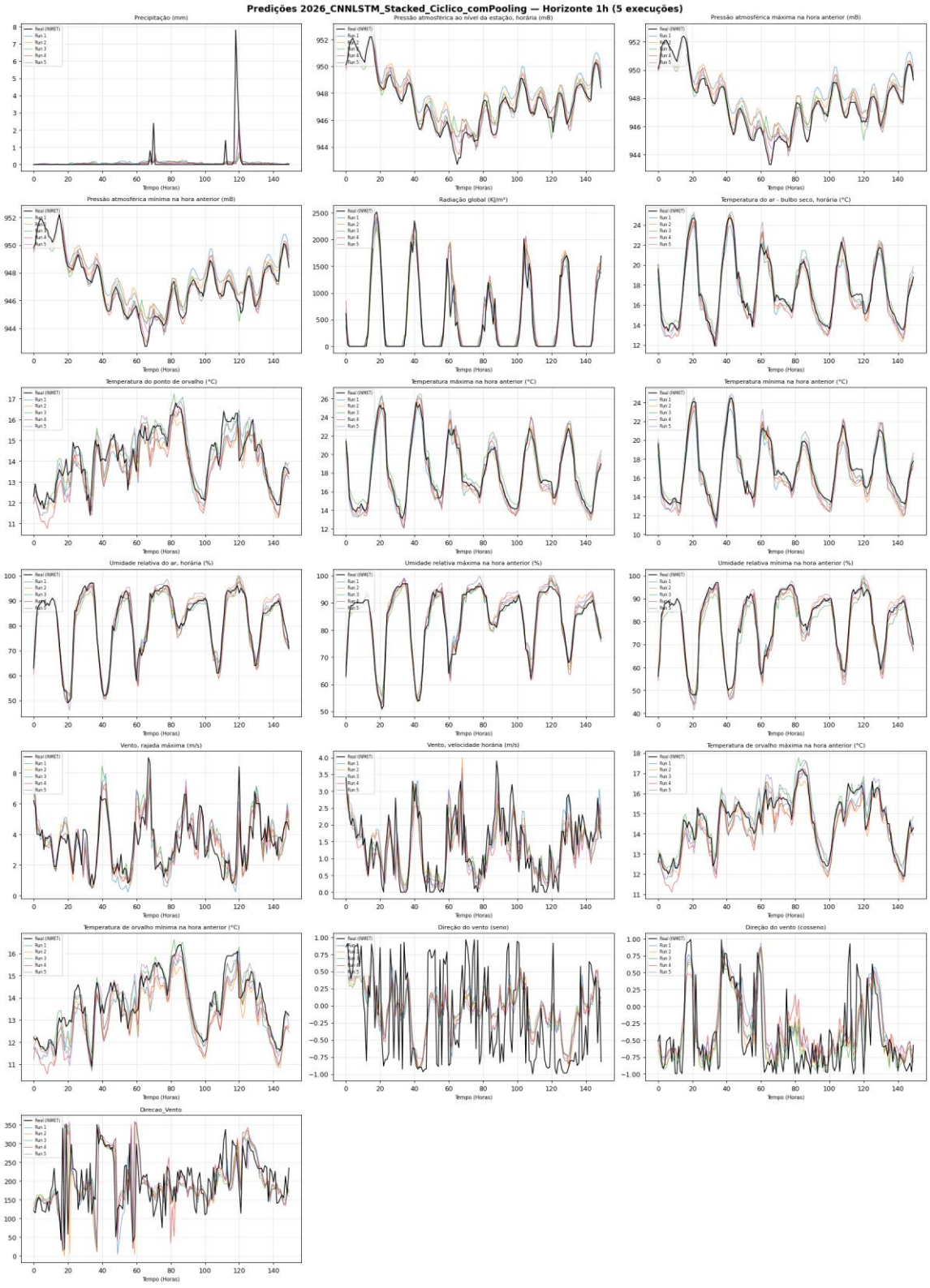
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	7.8904 ± 0.0388	0.6794 ± 0.0084	12.77 ± 0.07	7.3655 ± 0.6543
Umidade relativa do ar, horária (%)	7.5323 ± 0.0627	0.6735 ± 0.0107	11.70 ± 0.10	6.9940 ± 0.7514
Vento, rajada máxima (m/s)	2.6319 ± 0.2632	0.0329 ± 0.1590	100.91 ± 2.55	2.2929 ± 0.2778
Vento, velocidade horária (m/s)	1.2472 ± 0.1373	0.0668 ± 0.1231	139.94 ± 0.52	1.1216 ± 0.1406
Direção do vento (cosseno)	0.4413 ± 0.0034	0.2164 ± 0.0287	94.21 ± 2.15	0.3926 ± 0.0288
Direção do vento (seno)	0.6092 ± 0.0019	0.1197 ± 0.0036	150.47 ± 0.87	0.6240 ± 0.0273

Métricas — Horizonte 168h

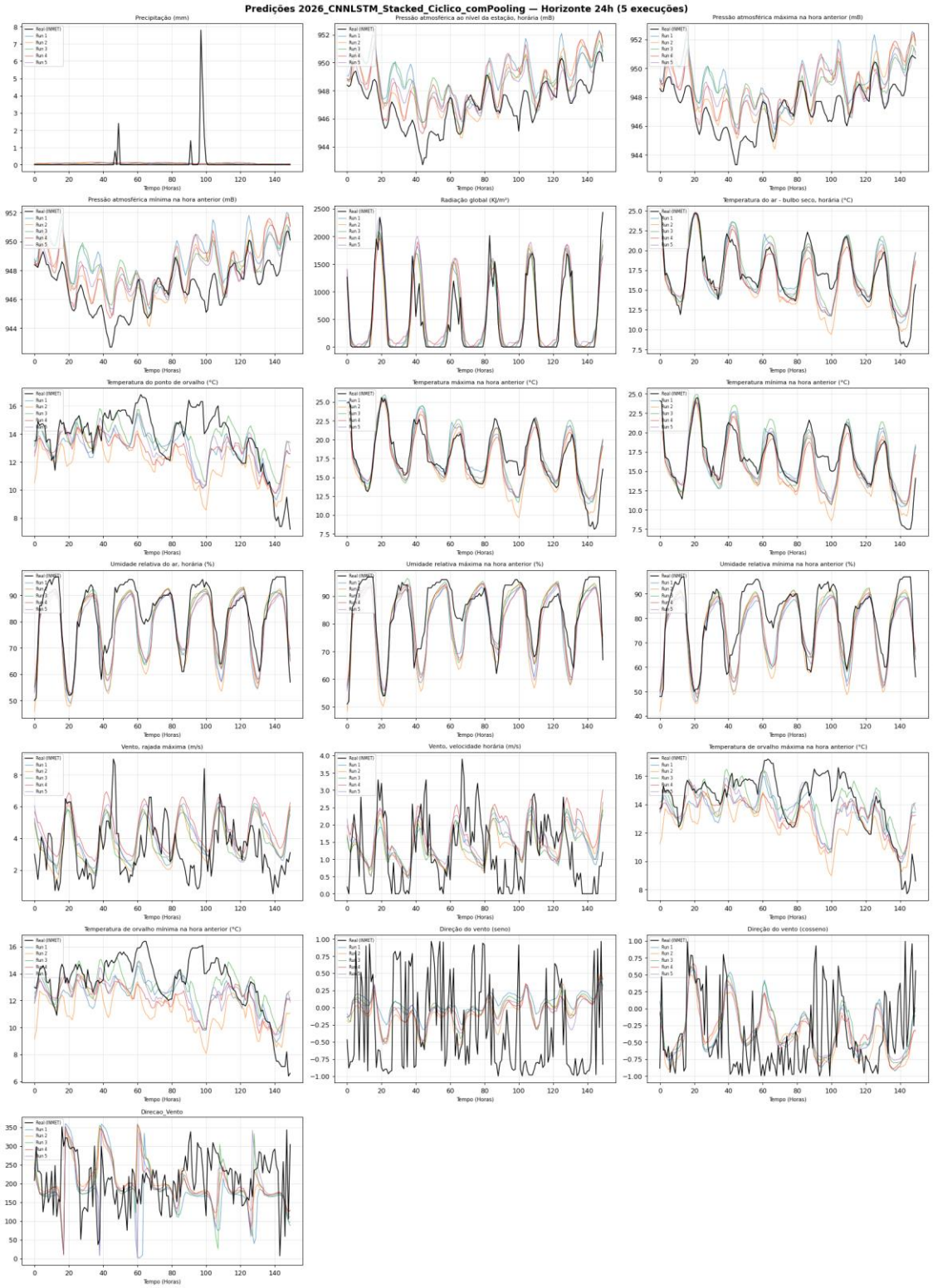
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.3651 ± 0.1424	0.3304 ± 0.0678	0.25 ± 0.01	1.6771 ± 0.0636
Pressão atmosférica máxima na hora anterior (mB)	2.3493 ± 0.1392	0.3293 ± 0.0658	0.25 ± 0.01	1.6538 ± 0.0637
Pressão atmosférica mínima na hora anterior (mB)	2.3722 ± 0.1422	0.3346 ± 0.0653	0.25 ± 0.01	1.6694 ± 0.0574
Precipitação (mm)	0.1884 ± 0.0176	-0.0003 ± 0.0033	185.48 ± 20.50	0.1905 ± 0.0154
Radiação global (KJ/m²)	271.8256 ± 10.9860	0.8090 ± 0.0133	75.23 ± 21.90	513.8890 ± 163.3557
Temperatura do ar - bulbo seco, horária (°C)	2.4795 ± 0.0342	0.6362 ± 0.0073	12.59 ± 0.16	2.7751 ± 0.1540
Temperatura do ponto de orvalho (°C)	2.0890 ± 0.0538	0.4874 ± 0.0356	15.45 ± 0.36	1.5660 ± 0.0517
Temperatura máxima na hora anterior (°C)	2.5262 ± 0.0431	0.6388 ± 0.0085	12.34 ± 0.19	2.9345 ± 0.1899
Temperatura mínima na hora anterior (°C)	2.4140 ± 0.0356	0.6412 ± 0.0085	12.80 ± 0.17	2.6351 ± 0.0709
Temperatura de orvalho máxima na hora anterior (°C)	2.0084 ± 0.0381	0.5143 ± 0.0289	14.14 ± 0.25	1.4948 ± 0.0476
Temperatura de orvalho mínima na hora anterior (°C)	2.1573 ± 0.0492	0.4762 ± 0.0369	16.80 ± 0.34	1.6429 ± 0.0775
Umidade relativa máxima na hora anterior (%)	8.3693 ± 0.3478	0.5555 ± 0.0406	12.66 ± 0.46	9.9677 ± 0.7287

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.5984 ± 0.3877	0.5630 ± 0.0389	15.88 ± 0.56	11.5467 ± 1.5677
Umidade relativa do ar, horária (%)	9.1096 ± 0.4015	0.5562 ± 0.0443	14.42 ± 0.55	11.5647 ± 1.1011
Vento, rajada máxima (m/s)	2.8910 ± 0.1899	-0.1115 ± 0.1295	104.70 ± 1.49	2.6786 ± 0.0954
Vento, velocidade horária (m/s)	1.3952 ± 0.1391	-0.1023 ± 0.1365	142.49 ± 0.50	1.1488 ± 0.2412
Direção do vento (cosseno)	0.4616 ± 0.0038	0.1653 ± 0.0090	99.90 ± 1.63	0.4393 ± 0.0351
Direção do vento (seno)	0.6408 ± 0.0021	0.0457 ± 0.0046	159.90 ± 2.51	0.6485 ± 0.0093

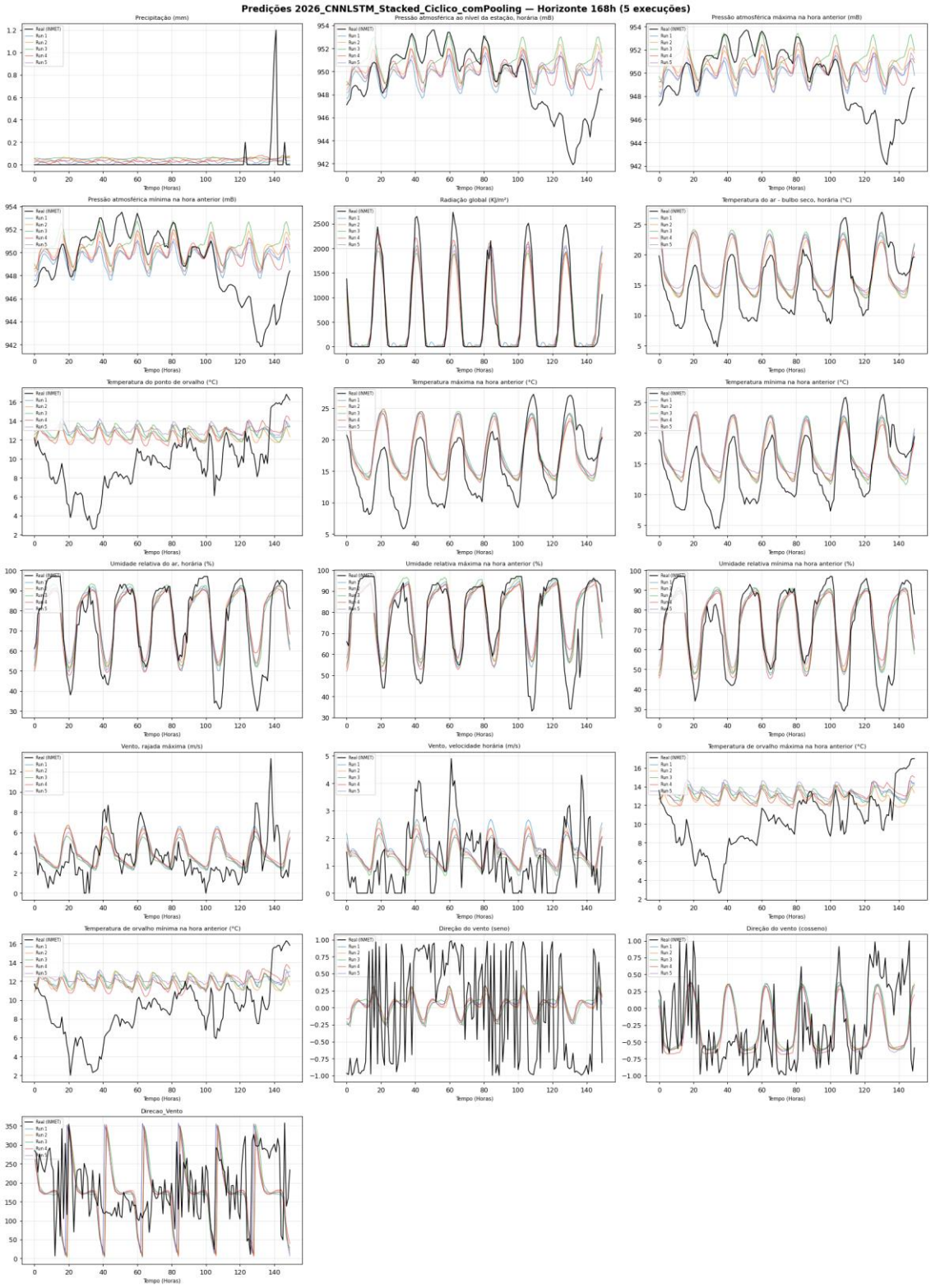
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-GRU

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.6491 ± 0.0724	0.9503 ± 0.0091	0.07 ± 0.01	0.4707 ± 0.0904
Pressão atmosférica máxima na hora anterior (mB)	0.6204 ± 0.0702	0.9537 ± 0.0089	0.07 ± 0.01	0.4552 ± 0.0786
Pressão atmosférica mínima na hora anterior (mB)	0.6208 ± 0.0796	0.9550 ± 0.0088	0.07 ± 0.01	0.4528 ± 0.1008
Precipitação (mm)	0.1555 ± 0.0155	0.0689 ± 0.0274	131.99 ± 50.17	0.1518 ± 0.0188
Radiação global (KJ/m²)	148.4706 ± 2.3638	0.9350 ± 0.0012	55.18 ± 16.85	128.7297 ± 33.4855
Temperatura do ar - bulbo seco, horária (°C)	0.8378 ± 0.0384	0.9542 ± 0.0032	4.25 ± 0.20	0.5900 ± 0.0293
Temperatura do ponto de orvalho (°C)	0.7380 ± 0.1159	0.9315 ± 0.0184	5.63 ± 0.91	0.4954 ± 0.1047
Temperatura máxima na hora anterior (°C)	0.7219 ± 0.0376	0.9688 ± 0.0031	3.54 ± 0.20	0.5303 ± 0.0169
Temperatura mínima na hora anterior (°C)	0.7478 ± 0.0839	0.9618 ± 0.0059	3.98 ± 0.45	0.5387 ± 0.0626
Temperatura de orvalho máxima na hora anterior (°C)	0.6552 ± 0.1369	0.9468 ± 0.0192	4.72 ± 1.01	0.4777 ± 0.1256
Temperatura de orvalho mínima na hora anterior (°C)	0.6750 ± 0.0851	0.9460 ± 0.0135	5.54 ± 0.71	0.4907 ± 0.0634

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.8461 ± 0.2685	0.9499 ± 0.0058	4.16 ± 0.39	2.1557 ± 0.1724
Umidade relativa mínima na hora anterior (%)	2.6092 ± 0.0528	0.9670 ± 0.0014	4.29 ± 0.15	2.0387 ± 0.0520
Umidade relativa do ar, horária (%)	3.2096 ± 0.0674	0.9402 ± 0.0020	5.01 ± 0.07	2.3059 ± 0.0900
Vento, rajada máxima (m/s)	1.1908 ± 0.0988	0.7106 ± 0.0184	75.37 ± 3.71	1.2038 ± 0.2264
Vento, velocidade horária (m/s)	0.5615 ± 0.0382	0.6893 ± 0.0150	122.20 ± 5.44	0.4914 ± 0.0349
Direção do vento (cosseno)	0.3145 ± 0.0021	0.5549 ± 0.0067	67.10 ± 0.66	0.2329 ± 0.0048
Direção do vento (seno)	0.5033 ± 0.0043	0.3230 ± 0.0076	121.51 ± 1.85	0.4451 ± 0.0128

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7425 ± 0.0439	0.6309 ± 0.0149	0.18 ± 0.01	0.9044 ± 0.0756
Pressão atmosférica máxima na hora anterior (mB)	1.7189 ± 0.0402	0.6353 ± 0.0138	0.18 ± 0.00	0.8759 ± 0.0588
Pressão atmosférica mínima na hora anterior (mB)	1.7304 ± 0.0434	0.6405 ± 0.0140	0.18 ± 0.01	0.8978 ± 0.0655
Precipitação (mm)	0.1905 ± 0.0155	0.0033 ± 0.0044	167.19 ± 22.53	0.1861 ± 0.0149
Radiação global (KJ/m²)	244.0959 ± 14.3629	0.8385 ± 0.0071	74.29 ± 21.05	402.9758 ± 45.5451
Temperatura do ar - bulbo seco, horária (°C)	1.9509 ± 0.0478	0.7574 ± 0.0097	9.87 ± 0.23	1.9640 ± 0.2006
Temperatura do ponto de orvalho (°C)	1.6790 ± 0.0829	0.6302 ± 0.0446	12.47 ± 0.59	1.0833 ± 0.1184
Temperatura máxima na hora anterior (°C)	1.9834 ± 0.0478	0.7599 ± 0.0094	9.65 ± 0.21	2.0919 ± 0.2314
Temperatura mínima na hora anterior (°C)	1.8774 ± 0.0490	0.7661 ± 0.0098	9.91 ± 0.25	1.9085 ± 0.1980
Temperatura de orvalho máxima na hora anterior (°C)	1.6099 ± 0.0918	0.6495 ± 0.0464	11.38 ± 0.63	1.0461 ± 0.1291
Temperatura de orvalho mínima na hora anterior (°C)	1.7177 ± 0.0820	0.6248 ± 0.0447	13.44 ± 0.62	1.1401 ± 0.1491
Umidade relativa máxima na hora anterior (%)	7.2554 ± 0.1975	0.6404 ± 0.0231	10.85 ± 0.26	7.3466 ± 1.0452

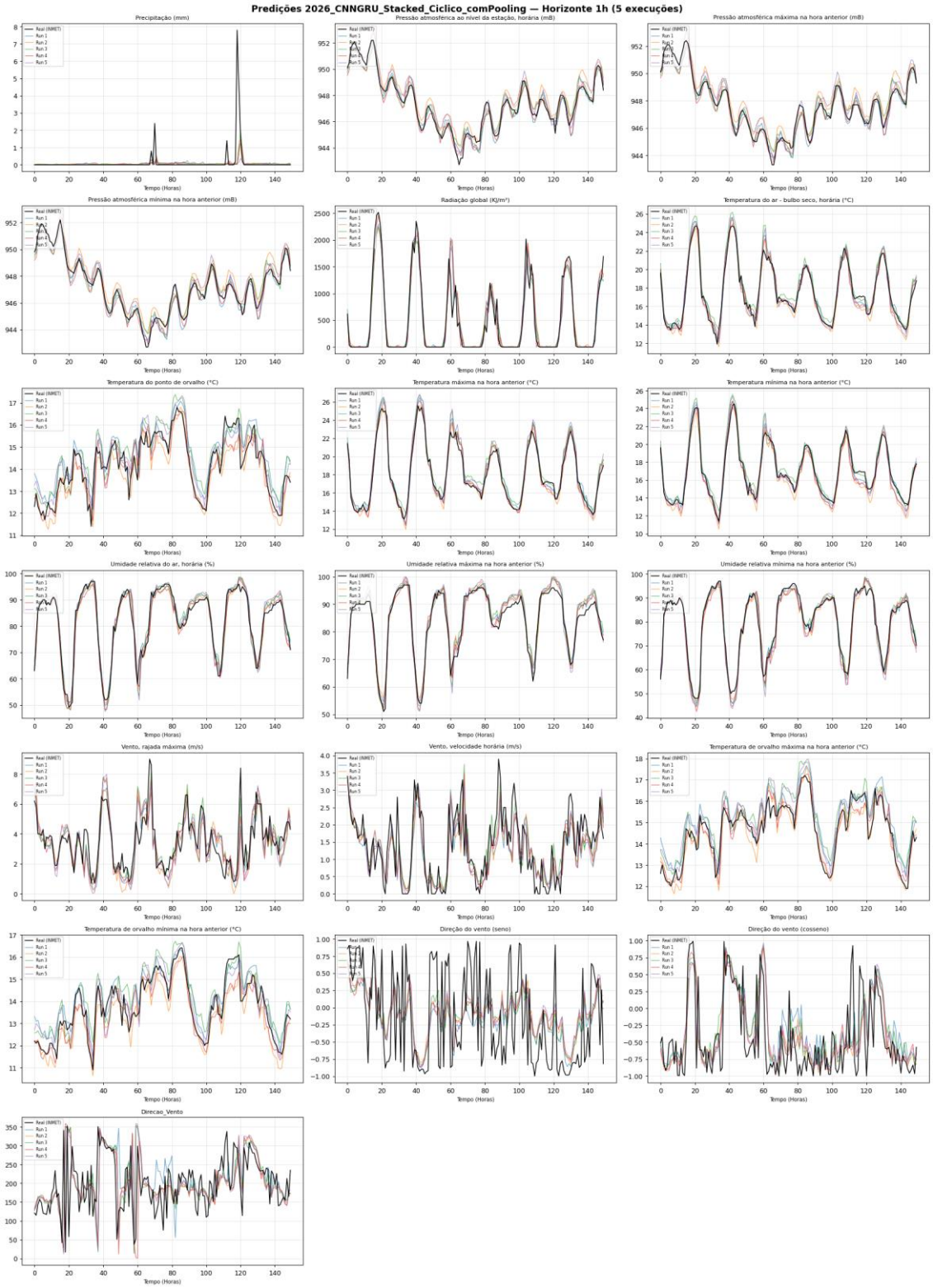
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	8.2320 ± 0.1699	0.6509 ± 0.0214	13.44 ± 0.30	8.5780 ± 1.5584
Umidade relativa do ar, horária (%)	7.8574 ± 0.1629	0.6437 ± 0.0218	12.30 ± 0.24	8.2846 ± 1.4607
Vento, rajada máxima (m/s)	2.4113 ± 0.1624	0.1613 ± 0.0733	98.74 ± 2.26	2.0894 ± 0.1199
Vento, velocidade horária (m/s)	1.0845 ± 0.0670	0.2125 ± 0.0350	138.48 ± 2.26	0.9602 ± 0.1131
Direção do vento (cosseno)	0.4500 ± 0.0023	0.2110 ± 0.0277	98.11 ± 2.06	0.4008 ± 0.0192
Direção do vento (seno)	0.6109 ± 0.0043	0.1131 ± 0.0108	151.08 ± 2.19	0.6488 ± 0.0561

Métricas — Horizonte 168h

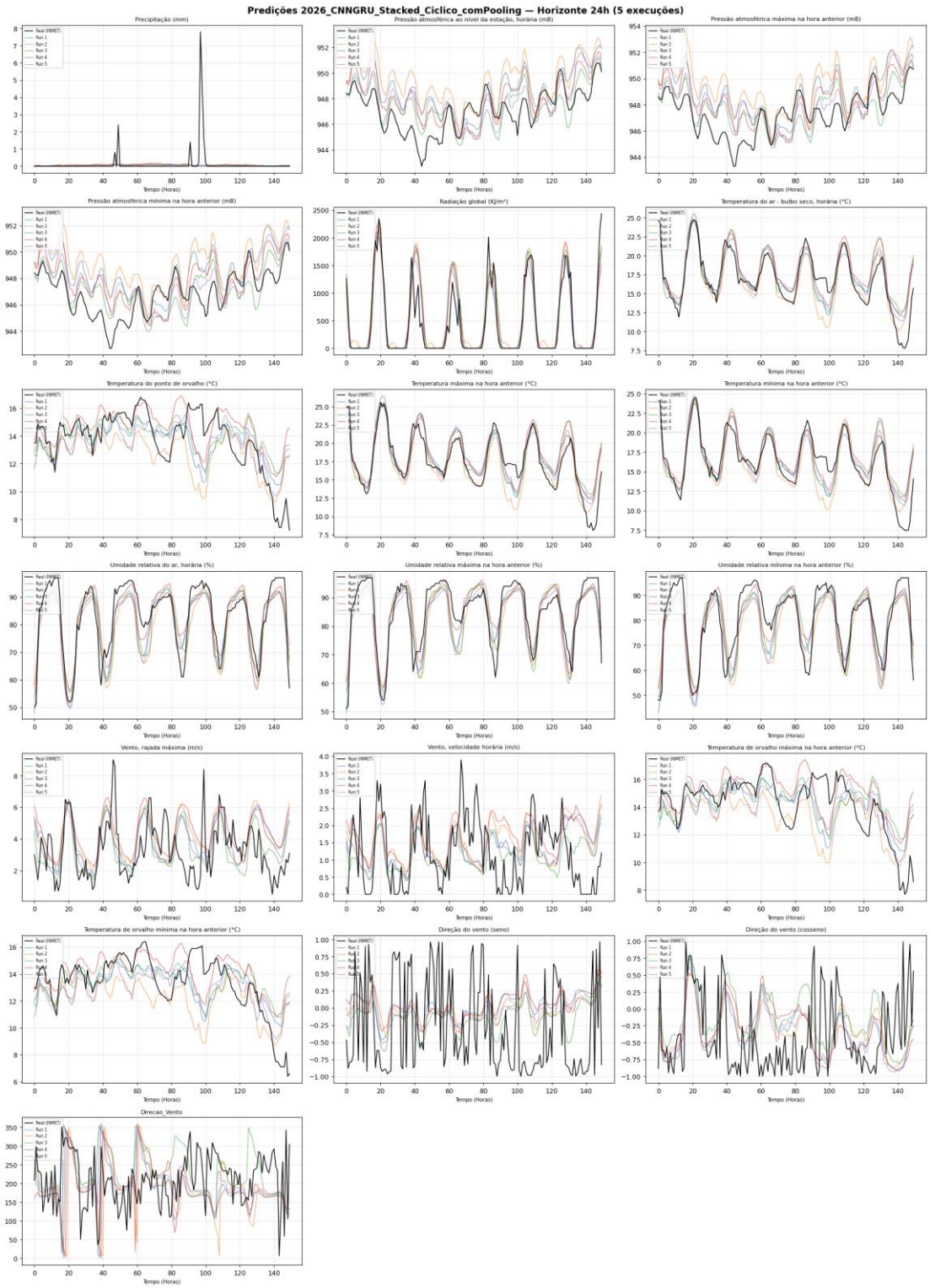
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.2606 ± 0.0157	0.3761 ± 0.0095	0.24 ± 0.00	1.7081 ± 0.0758
Pressão atmosférica máxima na hora anterior (mB)	2.2469 ± 0.0074	0.3746 ± 0.0085	0.24 ± 0.00	1.6847 ± 0.0710
Pressão atmosférica mínima na hora anterior (mB)	2.2662 ± 0.0145	0.3796 ± 0.0115	0.24 ± 0.00	1.6865 ± 0.0709
Precipitação (mm)	0.1839 ± 0.0090	0.0003 ± 0.0009	187.63 ± 12.09	0.1873 ± 0.0105
Radiação global (KJ/m²)	269.2288 ± 14.2093	0.8119 ± 0.0063	73.83 ± 27.40	472.5620 ± 82.3631
Temperatura do ar - bulbo seco, horária (°C)	2.4976 ± 0.0341	0.6341 ± 0.0056	12.66 ± 0.19	2.8237 ± 0.1456
Temperatura do ponto de orvalho (°C)	2.0739 ± 0.0616	0.5226 ± 0.0161	15.33 ± 0.45	1.6600 ± 0.0440
Temperatura máxima na hora anterior (°C)	2.5427 ± 0.0288	0.6371 ± 0.0061	12.40 ± 0.14	2.9745 ± 0.1680
Temperatura mínima na hora anterior (°C)	2.4349 ± 0.0414	0.6379 ± 0.0069	12.88 ± 0.23	2.6704 ± 0.1836
Temperatura de orvalho máxima na hora anterior (°C)	1.9890 ± 0.0612	0.5469 ± 0.0168	13.99 ± 0.41	1.5641 ± 0.0336
Temperatura de orvalho mínima na hora anterior (°C)	2.1313 ± 0.0626	0.5116 ± 0.0189	16.59 ± 0.48	1.6649 ± 0.0417
Umidade relativa máxima na hora anterior (%)	8.2336 ± 0.1455	0.5772 ± 0.0120	12.45 ± 0.20	10.0401 ± 1.1839

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.4926 ± 0.2004	0.5792 ± 0.0136	15.68 ± 0.29	11.4199 ± 0.9949
Umidade relativa do ar, horária (%)	8.9713 ± 0.2184	0.5767 ± 0.0176	14.21 ± 0.31	10.7671 ± 1.8057
Vento, rajada máxima (m/s)	3.0081 ± 0.1058	-0.1958 ± 0.0712	105.61 ± 0.88	2.8132 ± 0.1902
Vento, velocidade horária (m/s)	1.5073 ± 0.0492	-0.2224 ± 0.0631	142.22 ± 0.22	1.1424 ± 0.1594
Direção do vento (cosseno)	0.4603 ± 0.0010	0.1586 ± 0.0027	99.67 ± 0.90	0.4596 ± 0.0391
Direção do vento (seno)	0.6401 ± 0.0016	0.0468 ± 0.0019	158.42 ± 2.47	0.6464 ± 0.0102

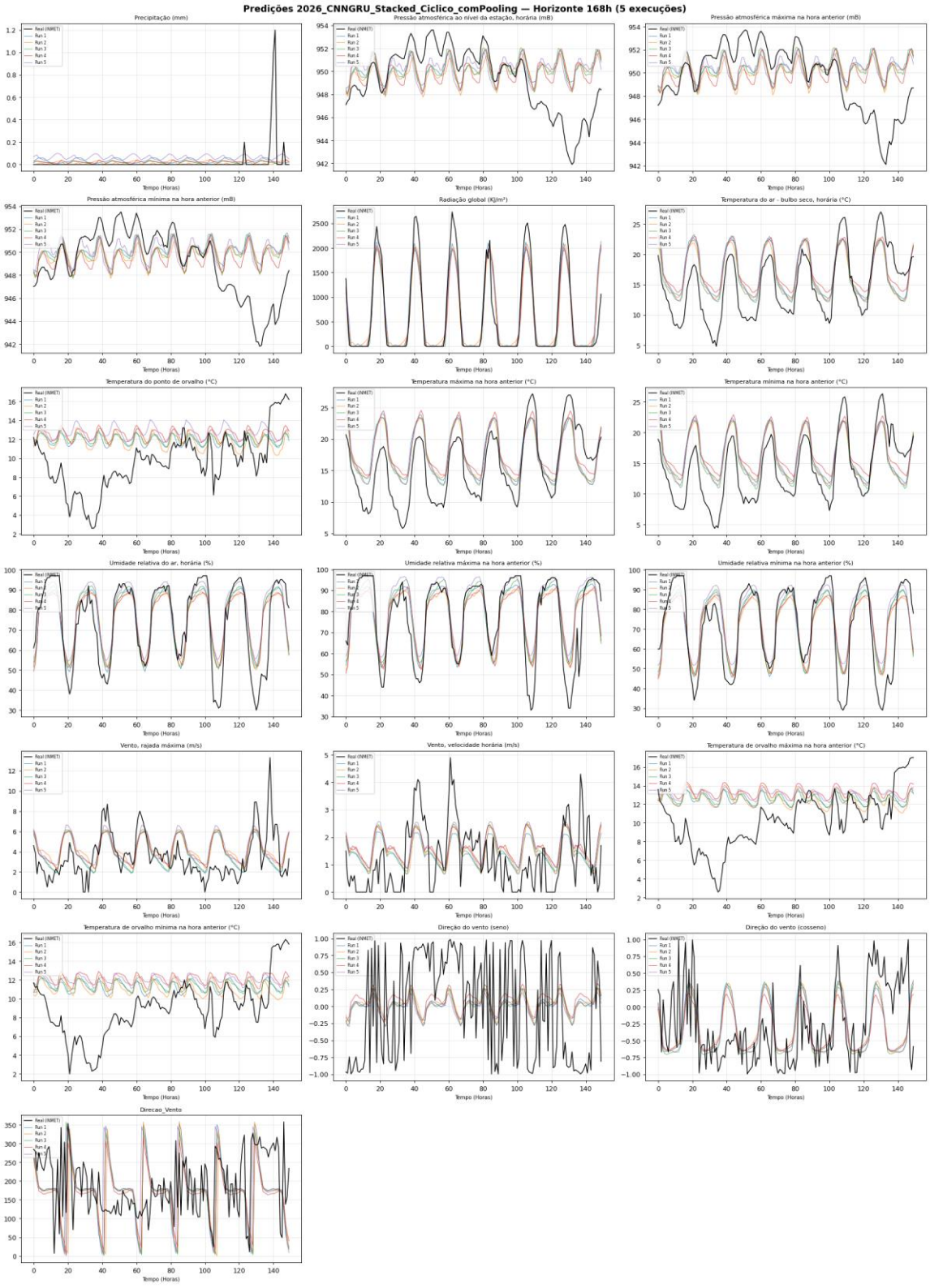
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-LSTM-GRU

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.6341 ± 0.1113	0.9513 ± 0.0160	0.06 ± 0.01	0.4344 ± 0.0694
Pressão atmosférica máxima na hora anterior (mB)	0.6018 ± 0.1167	0.9549 ± 0.0166	0.06 ± 0.01	0.4206 ± 0.0748
Pressão atmosférica mínima na hora anterior (mB)	0.6005 ± 0.1213	0.9565 ± 0.0162	0.06 ± 0.01	0.4199 ± 0.0822
Precipitação (mm)	0.1527 ± 0.0160	0.0607 ± 0.0221	98.65 ± 50.66	0.1485 ± 0.0166
Radiação global (KJ/m²)	160.8187 ± 10.4006	0.9301 ± 0.0041	78.06 ± 28.70	257.4223 ± 160.3239
Temperatura do ar - bulbo seco, horária (°C)	0.8395 ± 0.0353	0.9538 ± 0.0029	4.26 ± 0.20	0.5855 ± 0.0407
Temperatura do ponto de orvalho (°C)	0.6928 ± 0.0349	0.9375 ± 0.0053	5.31 ± 0.23	0.4527 ± 0.0325
Temperatura máxima na hora anterior (°C)	0.7383 ± 0.0590	0.9682 ± 0.0041	3.62 ± 0.34	0.5435 ± 0.0431
Temperatura mínima na hora anterior (°C)	0.7235 ± 0.0463	0.9630 ± 0.0039	3.82 ± 0.21	0.5185 ± 0.0429
Temperatura de orvalho máxima na hora anterior (°C)	0.6371 ± 0.0456	0.9496 ± 0.0060	4.60 ± 0.28	0.4632 ± 0.0392
Temperatura de orvalho mínima na hora anterior (°C)	0.5981 ± 0.0385	0.9552 ± 0.0051	4.92 ± 0.28	0.4176 ± 0.0231

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.6452 ± 0.1567	0.9531 ± 0.0038	3.95 ± 0.27	2.0028 ± 0.1041
Umidade relativa mínima na hora anterior (%)	2.6714 ± 0.0965	0.9665 ± 0.0024	4.35 ± 0.17	2.0601 ± 0.0399
Umidade relativa do ar, horária (%)	3.1844 ± 0.0548	0.9409 ± 0.0016	4.97 ± 0.12	2.2400 ± 0.0552
Vento, rajada máxima (m/s)	1.1483 ± 0.0470	0.7182 ± 0.0124	73.26 ± 1.91	1.0781 ± 0.0713
Vento, velocidade horária (m/s)	0.5248 ± 0.0312	0.6976 ± 0.0062	112.65 ± 9.45	0.4517 ± 0.0265
Direção do vento (cosseno)	0.3166 ± 0.0020	0.5519 ± 0.0062	68.03 ± 0.81	0.2423 ± 0.0174
Direção do vento (seno)	0.5002 ± 0.0088	0.3252 ± 0.0091	120.52 ± 3.39	0.4363 ± 0.0243

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7513 ± 0.0569	0.6262 ± 0.0252	0.18 ± 0.01	0.8868 ± 0.0791
Pressão atmosférica máxima na hora anterior (mB)	1.7240 ± 0.0589	0.6323 ± 0.0257	0.18 ± 0.01	0.8573 ± 0.0895
Pressão atmosférica mínima na hora anterior (mB)	1.7344 ± 0.0593	0.6375 ± 0.0246	0.18 ± 0.01	0.8649 ± 0.0794
Precipitação (mm)	0.1863 ± 0.0179	0.0023 ± 0.0029	162.09 ± 28.06	0.1862 ± 0.0174
Radiação global (KJ/m²)	243.7092 ± 6.0490	0.8350 ± 0.0040	63.05 ± 20.62	387.1521 ± 93.4656
Temperatura do ar - bulbo seco, horária (°C)	1.9200 ± 0.0331	0.7669 ± 0.0045	9.71 ± 0.17	1.9377 ± 0.2987
Temperatura do ponto de orvalho (°C)	1.6322 ± 0.0316	0.6591 ± 0.0170	12.17 ± 0.20	0.9888 ± 0.0707
Temperatura máxima na hora anterior (°C)	1.9315 ± 0.0169	0.7717 ± 0.0018	9.40 ± 0.09	1.9510 ± 0.2695
Temperatura mínima na hora anterior (°C)	1.8598 ± 0.0175	0.7738 ± 0.0014	9.83 ± 0.11	1.8961 ± 0.1368
Temperatura de orvalho máxima na hora anterior (°C)	1.5580 ± 0.0350	0.6809 ± 0.0178	11.04 ± 0.22	0.9740 ± 0.0335
Temperatura de orvalho mínima na hora anterior (°C)	1.6733 ± 0.0249	0.6519 ± 0.0157	13.17 ± 0.15	0.9966 ± 0.0232
Umidade relativa máxima na hora anterior (%)	7.1545 ± 0.2317	0.6533 ± 0.0223	10.60 ± 0.36	7.0042 ± 1.3536

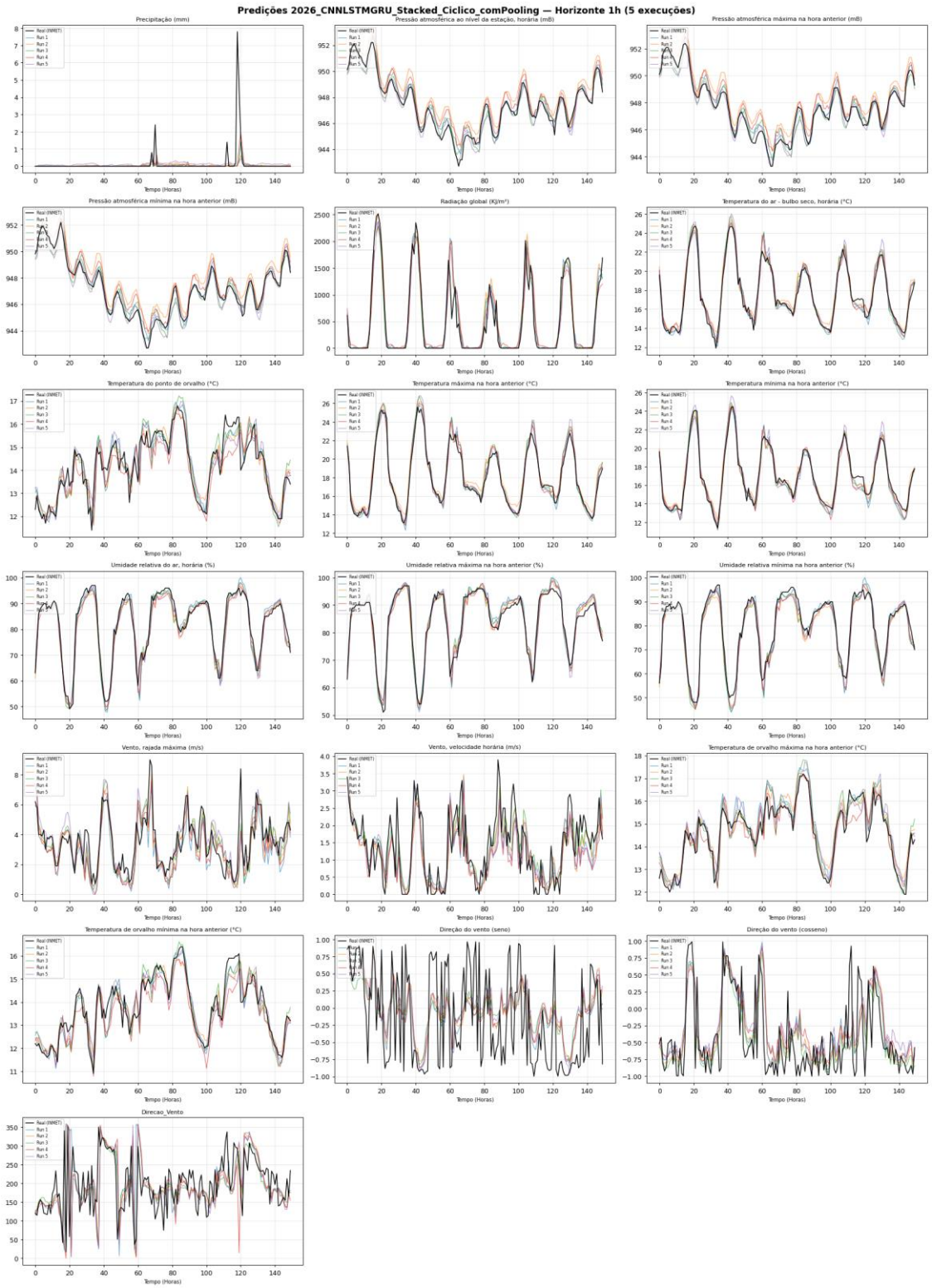
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	7.9783 ± 0.2615	0.6651 ± 0.0192	12.95 ± 0.45	7.3425 ± 1.8702
Umidade relativa do ar, horária (%)	7.6293 ± 0.2145	0.6592 ± 0.0175	11.87 ± 0.36	7.2641 ± 1.9577
Vento, rajada máxima (m/s)	2.5384 ± 0.1294	0.0950 ± 0.0793	100.14 ± 1.28	2.2995 ± 0.1076
Vento, velocidade horária (m/s)	1.1893 ± 0.1128	0.1251 ± 0.0975	139.92 ± 0.44	1.0436 ± 0.1461
Direção do vento (cosseno)	0.4415 ± 0.0051	0.2308 ± 0.0064	95.28 ± 2.31	0.3942 ± 0.0147
Direção do vento (seno)	0.6104 ± 0.0046	0.1135 ± 0.0057	149.91 ± 3.03	0.6233 ± 0.0066

Métricas — Horizonte 168h

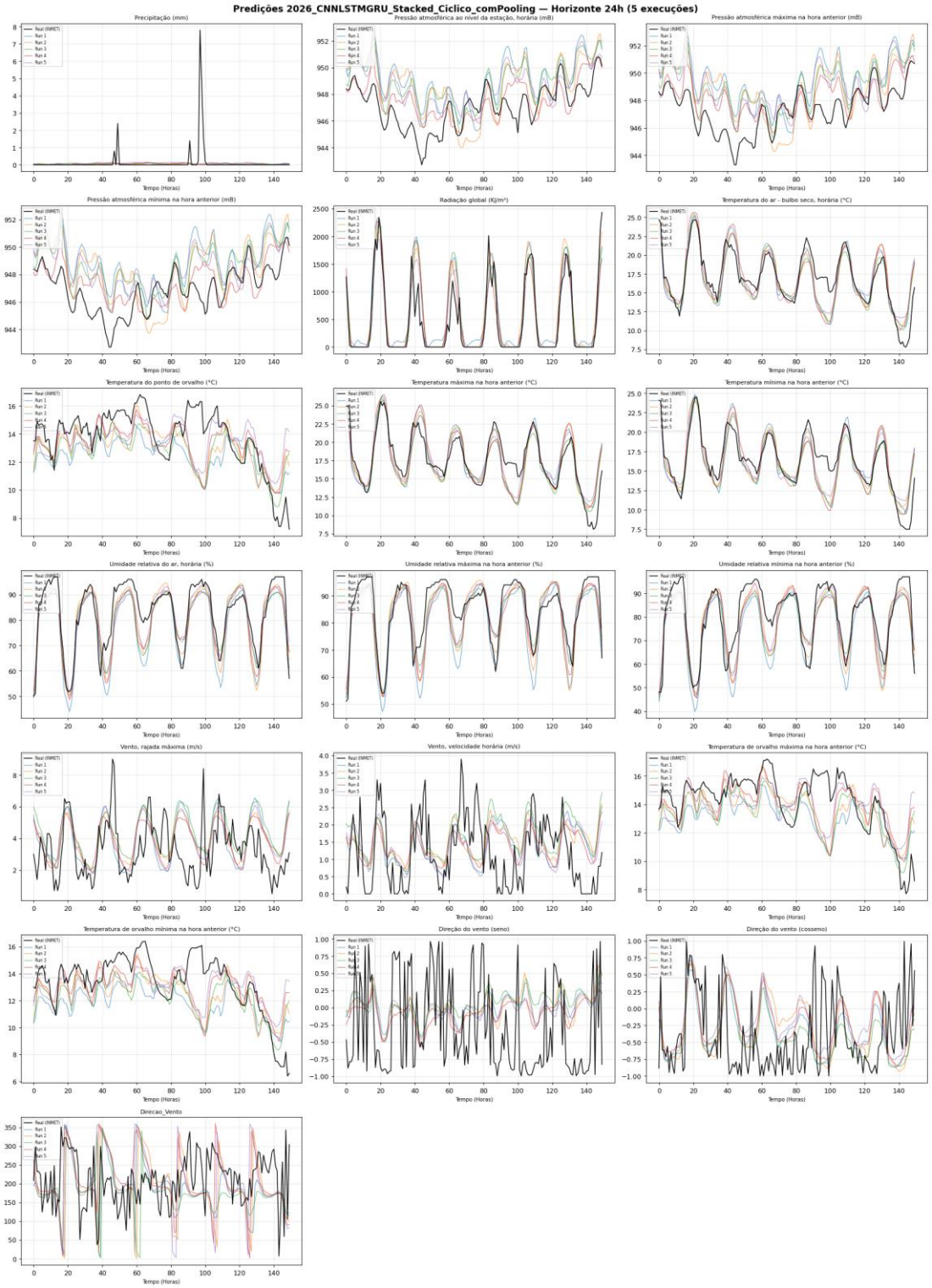
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.3083 ± 0.0383	0.3558 ± 0.0219	0.25 ± 0.01	1.7254 ± 0.0825
Pressão atmosférica máxima na hora anterior (mB)	2.3023 ± 0.0363	0.3517 ± 0.0195	0.24 ± 0.01	1.7087 ± 0.0577
Pressão atmosférica mínima na hora anterior (mB)	2.3223 ± 0.0381	0.3554 ± 0.0234	0.25 ± 0.01	1.7140 ± 0.0784
Precipitação (mm)	0.1768 ± 0.0058	-0.0005 ± 0.0018	170.92 ± 21.89	0.1780 ± 0.0078
Radiação global (KJ/m²)	273.7385 ± 13.1633	0.8065 ± 0.0089	70.24 ± 27.89	481.8739 ± 24.5771
Temperatura do ar - bulbo seco, horária (°C)	2.5335 ± 0.0633	0.6217 ± 0.0159	12.84 ± 0.30	2.8236 ± 0.1714
Temperatura do ponto de orvalho (°C)	2.0786 ± 0.0643	0.5044 ± 0.0265	15.41 ± 0.47	1.6447 ± 0.0388
Temperatura máxima na hora anterior (°C)	2.6055 ± 0.0773	0.6172 ± 0.0196	12.70 ± 0.35	2.9125 ± 0.1957
Temperatura mínima na hora anterior (°C)	2.4696 ± 0.0559	0.6248 ± 0.0155	13.06 ± 0.28	2.6643 ± 0.1133
Temperatura de orvalho máxima na hora anterior (°C)	1.9979 ± 0.0567	0.5281 ± 0.0209	14.09 ± 0.39	1.5267 ± 0.0188
Temperatura de orvalho mínima na hora anterior (°C)	2.1474 ± 0.0688	0.4920 ± 0.0281	16.75 ± 0.52	1.6758 ± 0.0652
Umidade relativa máxima na hora anterior (%)	8.4602 ± 0.3104	0.5483 ± 0.0329	12.80 ± 0.46	10.8686 ± 1.2595

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.6619 ± 0.2347	0.5605 ± 0.0236	15.98 ± 0.39	12.4646 ± 1.2044
Umidade relativa do ar, horária (%)	9.1771 ± 0.3019	0.5524 ± 0.0296	14.54 ± 0.47	11.7221 ± 0.9792
Vento, rajada máxima (m/s)	2.8480 ± 0.1437	-0.0832 ± 0.1014	104.27 ± 1.30	2.4608 ± 0.0973
Vento, velocidade horária (m/s)	1.4125 ± 0.0459	-0.1060 ± 0.0526	142.37 ± 0.21	1.1196 ± 0.2094
Direção do vento (cosseno)	0.4619 ± 0.0031	0.1698 ± 0.0114	101.03 ± 1.21	0.4501 ± 0.0303
Direção do vento (seno)	0.6408 ± 0.0014	0.0472 ± 0.0046	159.57 ± 4.58	0.6498 ± 0.0129

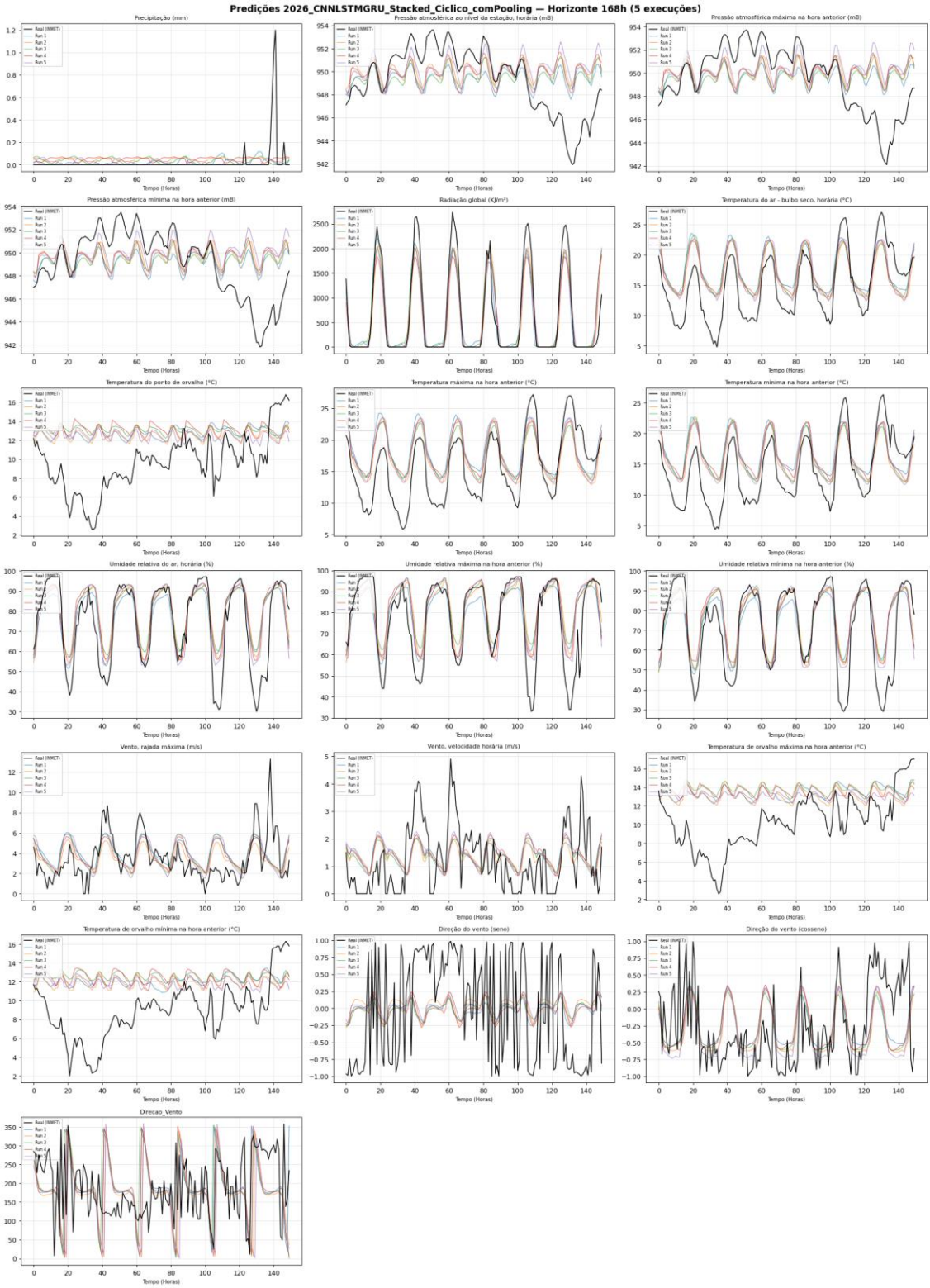
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-BiLSTM

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.5956 ± 0.0425	0.9579 ± 0.0056	0.06 ± 0.00	0.4254 ± 0.0365
Pressão atmosférica máxima na hora anterior (mB)	0.5568 ± 0.0396	0.9622 ± 0.0052	0.06 ± 0.00	0.3999 ± 0.0307
Pressão atmosférica mínima na hora anterior (mB)	0.5591 ± 0.0452	0.9636 ± 0.0053	0.06 ± 0.00	0.4000 ± 0.0358
Precipitação (mm)	0.1617 ± 0.0220	0.0763 ± 0.0238	124.80 ± 42.08	0.1493 ± 0.0235
Radiação global (KJ/m²)	144.3930 ± 8.7462	0.9362 ± 0.0022	52.74 ± 24.37	145.6843 ± 81.6945
Temperatura do ar - bulbo seco, horária (°C)	0.8285 ± 0.0489	0.9545 ± 0.0044	4.19 ± 0.25	0.5853 ± 0.0642
Temperatura do ponto de orvalho (°C)	0.7342 ± 0.0373	0.9303 ± 0.0062	5.61 ± 0.28	0.4958 ± 0.0365
Temperatura máxima na hora anterior (°C)	0.7231 ± 0.0684	0.9689 ± 0.0048	3.53 ± 0.35	0.5435 ± 0.0584
Temperatura mínima na hora anterior (°C)	0.7378 ± 0.0482	0.9621 ± 0.0047	3.92 ± 0.26	0.5307 ± 0.0338
Temperatura de orvalho máxima na hora anterior (°C)	0.6688 ± 0.0484	0.9445 ± 0.0076	4.81 ± 0.35	0.5196 ± 0.0614
Temperatura de orvalho mínima na hora anterior (°C)	0.6250 ± 0.0292	0.9503 ± 0.0045	5.13 ± 0.25	0.4642 ± 0.0684

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.8149 ± 0.3214	0.9487 ± 0.0092	4.20 ± 0.55	2.4658 ± 0.7787
Umidade relativa mínima na hora anterior (%)	2.6527 ± 0.1474	0.9650 ± 0.0043	4.38 ± 0.31	2.0187 ± 0.0552
Umidade relativa do ar, horária (%)	3.3025 ± 0.1971	0.9363 ± 0.0060	5.21 ± 0.37	2.2989 ± 0.1008
Vento, rajada máxima (m/s)	1.1664 ± 0.0672	0.7203 ± 0.0146	74.71 ± 2.04	1.1615 ± 0.1838
Vento, velocidade horária (m/s)	0.5440 ± 0.0363	0.6888 ± 0.0151	120.62 ± 6.41	0.4550 ± 0.0578
Direção do vento (cosseno)	0.3190 ± 0.0068	0.5383 ± 0.0217	67.80 ± 1.45	0.2425 ± 0.0158
Direção do vento (seno)	0.4990 ± 0.0050	0.3182 ± 0.0108	118.19 ± 2.24	0.4470 ± 0.0293

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7726 ± 0.0607	0.6176 ± 0.0297	0.19 ± 0.01	0.9046 ± 0.0568
Pressão atmosférica máxima na hora anterior (mB)	1.7551 ± 0.0613	0.6202 ± 0.0292	0.19 ± 0.00	0.8983 ± 0.0661
Pressão atmosférica mínima na hora anterior (mB)	1.7661 ± 0.0635	0.6259 ± 0.0289	0.19 ± 0.00	0.9098 ± 0.0685
Precipitação (mm)	0.1918 ± 0.0139	0.0018 ± 0.0028	171.60 ± 13.61	0.1891 ± 0.0127
Radiação global (KJ/m²)	254.0215 ± 12.9793	0.8210 ± 0.0107	63.52 ± 12.12	453.5418 ± 82.3355
Temperatura do ar - bulbo seco, horária (°C)	2.0093 ± 0.0604	0.7445 ± 0.0180	10.14 ± 0.24	2.0926 ± 0.2936
Temperatura do ponto de orvalho (°C)	1.7421 ± 0.0819	0.6251 ± 0.0406	12.89 ± 0.46	1.0747 ± 0.0683
Temperatura máxima na hora anterior (°C)	2.0625 ± 0.0521	0.7437 ± 0.0148	10.00 ± 0.19	2.2725 ± 0.3167
Temperatura mínima na hora anterior (°C)	1.9440 ± 0.0680	0.7518 ± 0.0187	10.24 ± 0.29	1.9432 ± 0.2040
Temperatura de orvalho máxima na hora anterior (°C)	1.6475 ± 0.0487	0.6566 ± 0.0267	11.61 ± 0.26	1.0562 ± 0.0790
Temperatura de orvalho mínima na hora anterior (°C)	1.7808 ± 0.0467	0.6205 ± 0.0261	13.92 ± 0.38	1.0873 ± 0.0826
Umidade relativa máxima na hora anterior (%)	7.5170 ± 0.1769	0.6219 ± 0.0151	11.22 ± 0.24	8.9071 ± 1.2681

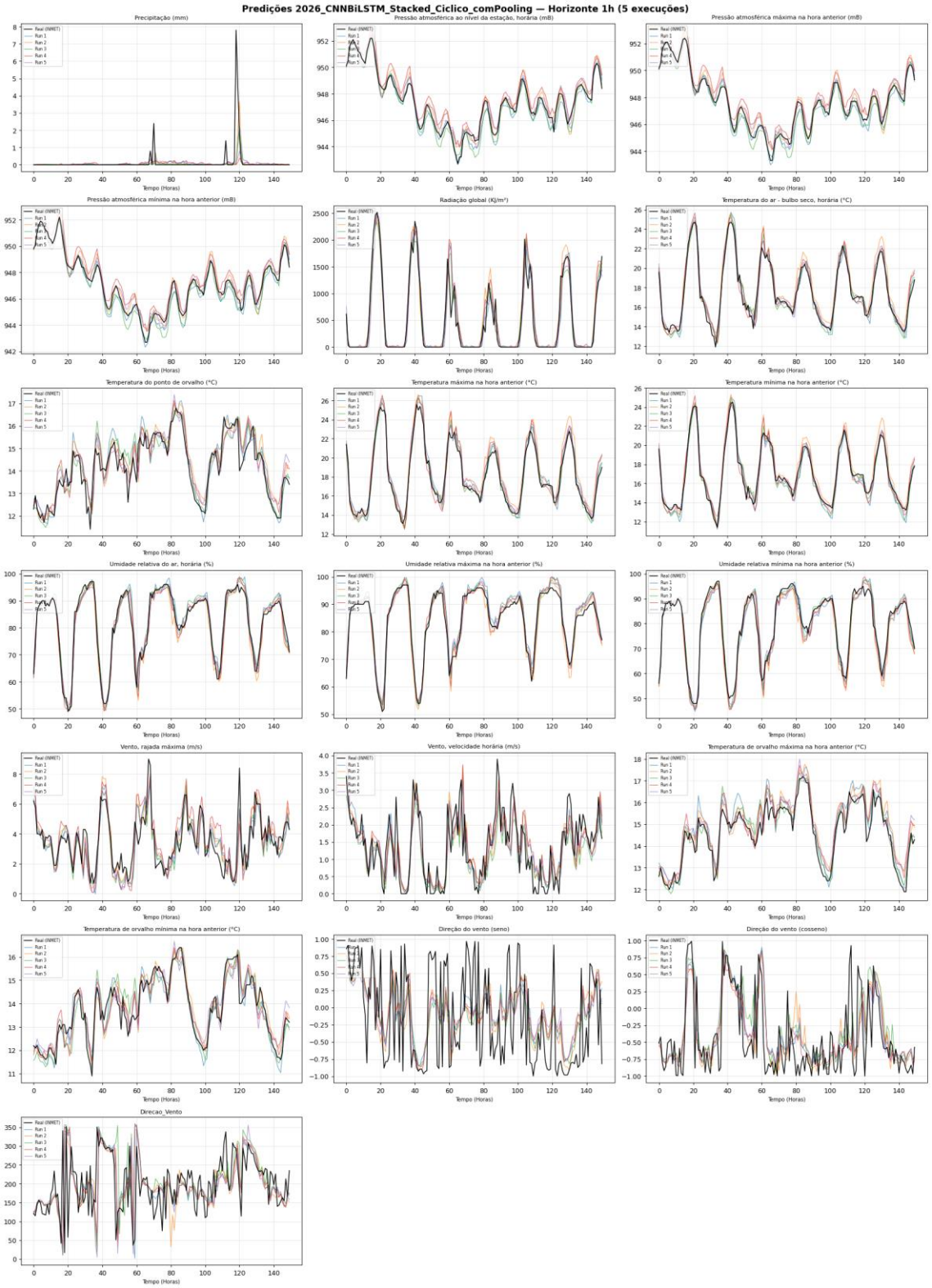
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	8.5678 ± 0.1561	0.6283 ± 0.0092	14.00 ± 0.28	9.5116 ± 1.2983
Umidade relativa do ar, horária (%)	8.1329 ± 0.1597	0.6220 ± 0.0105	12.73 ± 0.24	9.5419 ± 0.6482
Vento, rajada máxima (m/s)	2.3959 ± 0.0795	0.1666 ± 0.0419	98.36 ± 1.03	2.1403 ± 0.1440
Vento, velocidade horária (m/s)	1.0908 ± 0.0820	0.1717 ± 0.0786	139.22 ± 1.25	0.9188 ± 0.1134
Direção do vento (cosseno)	0.4524 ± 0.0072	0.1706 ± 0.0358	95.27 ± 2.75	0.4115 ± 0.0255
Direção do vento (seno)	0.6137 ± 0.0051	0.0991 ± 0.0160	149.96 ± 1.62	0.6307 ± 0.0296

Métricas — Horizonte 168h

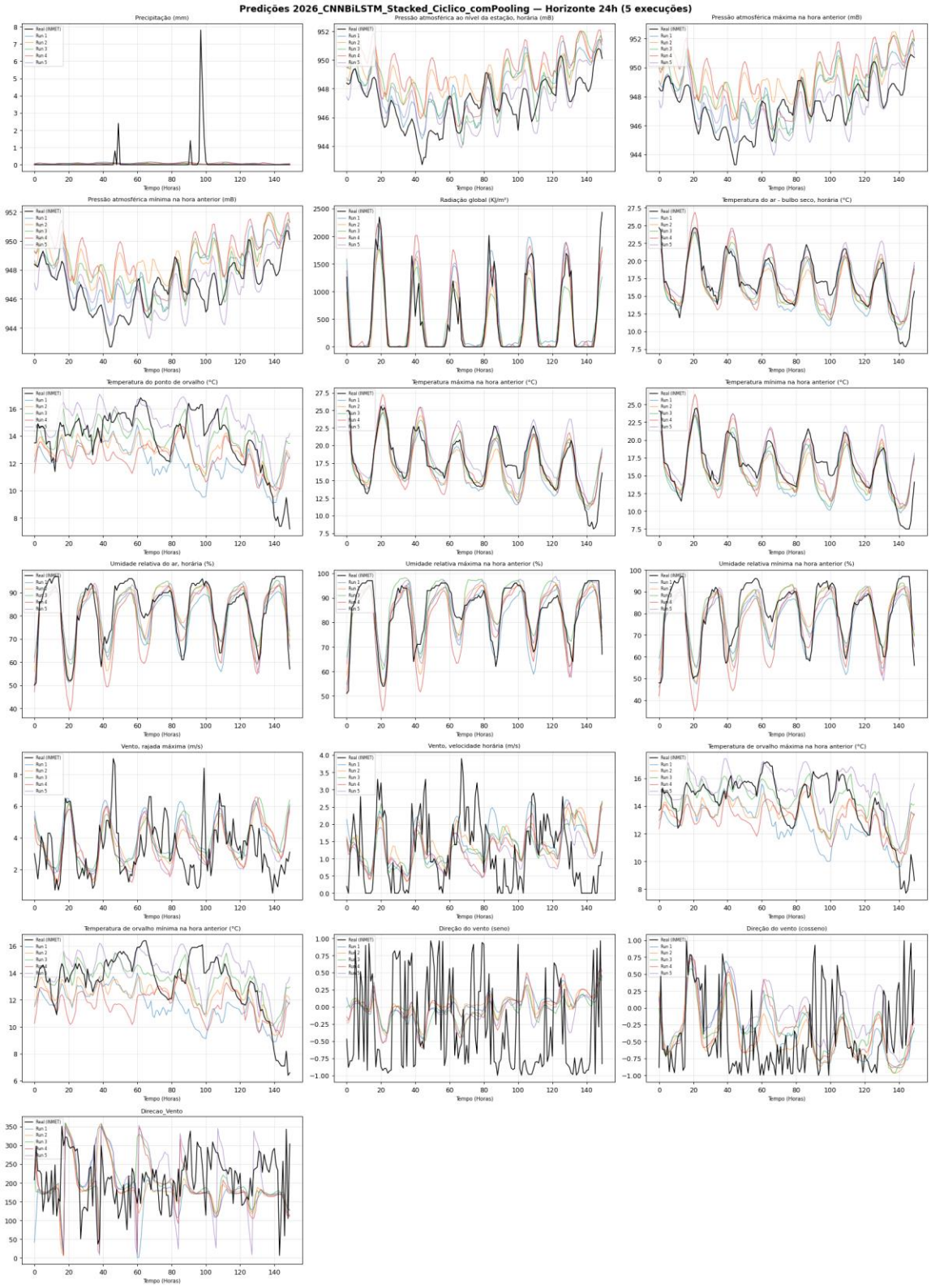
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.3188 ± 0.0460	0.3372 ± 0.0322	0.25 ± 0.01	1.7004 ± 0.0809
Pressão atmosférica máxima na hora anterior (mB)	2.3101 ± 0.0550	0.3346 ± 0.0377	0.24 ± 0.01	1.6613 ± 0.0690
Pressão atmosférica mínima na hora anterior (mB)	2.3219 ± 0.0503	0.3425 ± 0.0349	0.25 ± 0.01	1.6945 ± 0.0817
Precipitação (mm)	0.1796 ± 0.0170	-0.0024 ± 0.0030	168.16 ± 24.91	0.1796 ± 0.0194
Radiação global (KJ/m²)	284.7917 ± 10.3352	0.8034 ± 0.0061	70.71 ± 13.33	471.8866 ± 39.4685
Temperatura do ar - bulbo seco, horária (°C)	2.5399 ± 0.0662	0.6231 ± 0.0160	12.93 ± 0.37	2.7140 ± 0.1286
Temperatura do ponto de orvalho (°C)	2.0852 ± 0.0871	0.5106 ± 0.0269	15.42 ± 0.57	1.6031 ± 0.0532
Temperatura máxima na hora anterior (°C)	2.6209 ± 0.0479	0.6165 ± 0.0135	12.85 ± 0.29	3.0099 ± 0.2255
Temperatura mínima na hora anterior (°C)	2.4972 ± 0.0946	0.6215 ± 0.0247	13.26 ± 0.52	2.7284 ± 0.1419
Temperatura de orvalho máxima na hora anterior (°C)	2.0241 ± 0.0983	0.5285 ± 0.0309	14.23 ± 0.61	1.5093 ± 0.0406
Temperatura de orvalho mínima na hora anterior (°C)	2.1611 ± 0.0828	0.4994 ± 0.0259	16.81 ± 0.57	1.6741 ± 0.0568
Umidade relativa máxima na hora anterior (%)	8.4126 ± 0.2016	0.5624 ± 0.0234	12.68 ± 0.26	10.0156 ± 1.0148

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.5806 ± 0.1889	0.5728 ± 0.0162	15.81 ± 0.29	11.3965 ± 0.6943
Umidade relativa do ar, horária (%)	9.0801 ± 0.1735	0.5701 ± 0.0172	14.35 ± 0.25	11.0823 ± 1.1320
Vento, rajada máxima (m/s)	2.9551 ± 0.0912	-0.1570 ± 0.0652	105.11 ± 0.79	2.5709 ± 0.1767
Vento, velocidade horária (m/s)	1.4940 ± 0.1018	-0.2092 ± 0.1225	142.16 ± 0.31	1.2528 ± 0.2256
Direção do vento (cosseno)	0.4611 ± 0.0014	0.1518 ± 0.0133	98.90 ± 1.13	0.4576 ± 0.0477
Direção do vento (seno)	0.6430 ± 0.0046	0.0395 ± 0.0062	160.75 ± 4.86	0.6496 ± 0.0105

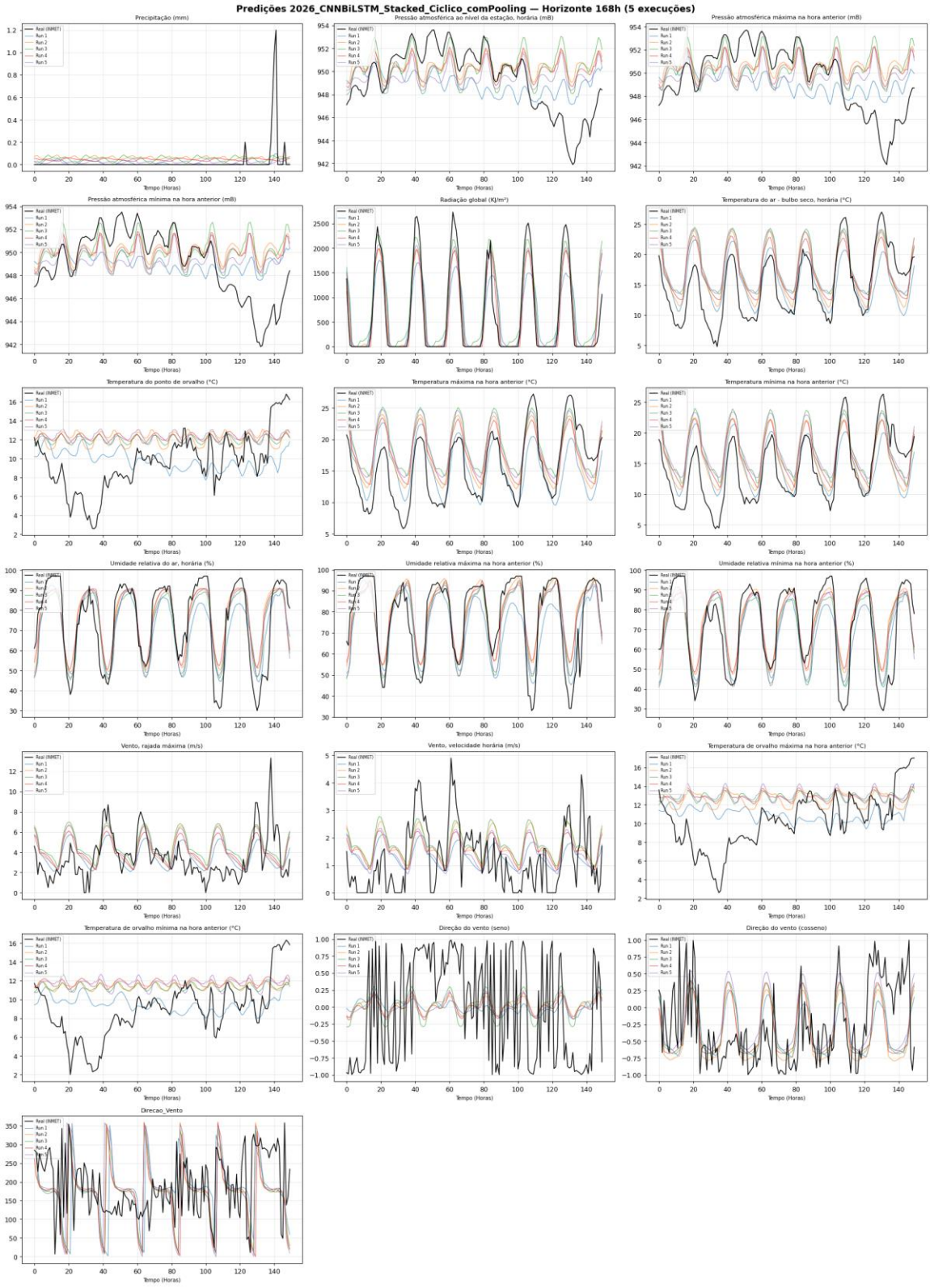
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-BiGRU

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.5830 ± 0.0626	0.9604 ± 0.0068	0.06 ± 0.00	0.4099 ± 0.0472
Pressão atmosférica máxima na hora anterior (mB)	0.5472 ± 0.0611	0.9645 ± 0.0065	0.06 ± 0.01	0.3971 ± 0.0360
Pressão atmosférica mínima na hora anterior (mB)	0.5389 ± 0.0609	0.9666 ± 0.0054	0.06 ± 0.01	0.3922 ± 0.0578
Precipitação (mm)	0.1599 ± 0.0168	0.0744 ± 0.0270	129.42 ± 34.19	0.1509 ± 0.0147
Radiação global (KJ/m²)	141.1234 ± 7.1547	0.9388 ± 0.0019	50.17 ± 24.98	149.3360 ± 62.6568
Temperatura do ar - bulbo seco, horária (°C)	0.8476 ± 0.1259	0.9534 ± 0.0105	4.32 ± 0.73	0.7103 ± 0.3203
Temperatura do ponto de orvalho (°C)	0.6748 ± 0.0451	0.9415 ± 0.0060	5.13 ± 0.35	0.4459 ± 0.0418
Temperatura máxima na hora anterior (°C)	0.7405 ± 0.1784	0.9679 ± 0.0124	3.65 ± 0.96	0.6342 ± 0.2802
Temperatura mínima na hora anterior (°C)	0.7477 ± 0.1004	0.9619 ± 0.0088	3.99 ± 0.58	0.6279 ± 0.2447
Temperatura de orvalho máxima na hora anterior (°C)	0.6025 ± 0.0776	0.9549 ± 0.0086	4.32 ± 0.55	0.4490 ± 0.0655
Temperatura de orvalho mínima na hora anterior (°C)	0.6035 ± 0.0205	0.9566 ± 0.0031	4.93 ± 0.17	0.4324 ± 0.0228

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.7644 ± 0.3125	0.9515 ± 0.0056	4.02 ± 0.39	2.4820 ± 1.0771
Umidade relativa mínima na hora anterior (%)	2.7844 ± 0.3109	0.9647 ± 0.0049	4.58 ± 0.50	2.4068 ± 0.3805
Umidade relativa do ar, horária (%)	3.2812 ± 0.1756	0.9388 ± 0.0034	5.09 ± 0.24	2.4718 ± 0.3408
Vento, rajada máxima (m/s)	1.1780 ± 0.0597	0.7151 ± 0.0136	74.74 ± 1.90	1.1322 ± 0.2046
Vento, velocidade horária (m/s)	0.5397 ± 0.0429	0.6951 ± 0.0113	117.37 ± 12.62	0.4657 ± 0.0622
Direção do vento (cosseno)	0.3212 ± 0.0045	0.5405 ± 0.0119	69.00 ± 1.18	0.2422 ± 0.0065
Direção do vento (seno)	0.5008 ± 0.0057	0.3218 ± 0.0076	119.50 ± 2.31	0.4396 ± 0.0187

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7061 ± 0.0389	0.6476 ± 0.0131	0.18 ± 0.00	0.8485 ± 0.0675
Pressão atmosférica máxima na hora anterior (mB)	1.6956 ± 0.0384	0.6482 ± 0.0136	0.18 ± 0.00	0.8627 ± 0.0773
Pressão atmosférica mínima na hora anterior (mB)	1.7021 ± 0.0454	0.6538 ± 0.0147	0.18 ± 0.00	0.8523 ± 0.0800
Precipitação (mm)	0.2022 ± 0.0233	0.0054 ± 0.0037	177.32 ± 22.51	0.1995 ± 0.0232
Radiação global (KJ/m²)	250.0591 ± 6.8108	0.8287 ± 0.0093	70.22 ± 15.21	431.5074 ± 38.3723
Temperatura do ar - bulbo seco, horária (°C)	1.9851 ± 0.0581	0.7543 ± 0.0112	10.05 ± 0.29	2.2677 ± 0.2495
Temperatura do ponto de orvalho (°C)	1.6892 ± 0.0673	0.6456 ± 0.0082	12.54 ± 0.44	1.0432 ± 0.0312
Temperatura máxima na hora anterior (°C)	1.9992 ± 0.0431	0.7595 ± 0.0096	9.74 ± 0.20	2.1691 ± 0.1906
Temperatura mínima na hora anterior (°C)	1.9128 ± 0.0348	0.7621 ± 0.0091	10.10 ± 0.18	2.1469 ± 0.2445
Temperatura de orvalho máxima na hora anterior (°C)	1.6268 ± 0.0524	0.6641 ± 0.0065	11.46 ± 0.33	1.0595 ± 0.0281
Temperatura de orvalho mínima na hora anterior (°C)	1.7288 ± 0.0917	0.6408 ± 0.0180	13.54 ± 0.66	1.0679 ± 0.0706
Umidade relativa máxima na hora anterior (%)	7.0552 ± 0.1848	0.6614 ± 0.0184	10.50 ± 0.30	6.8546 ± 1.9412

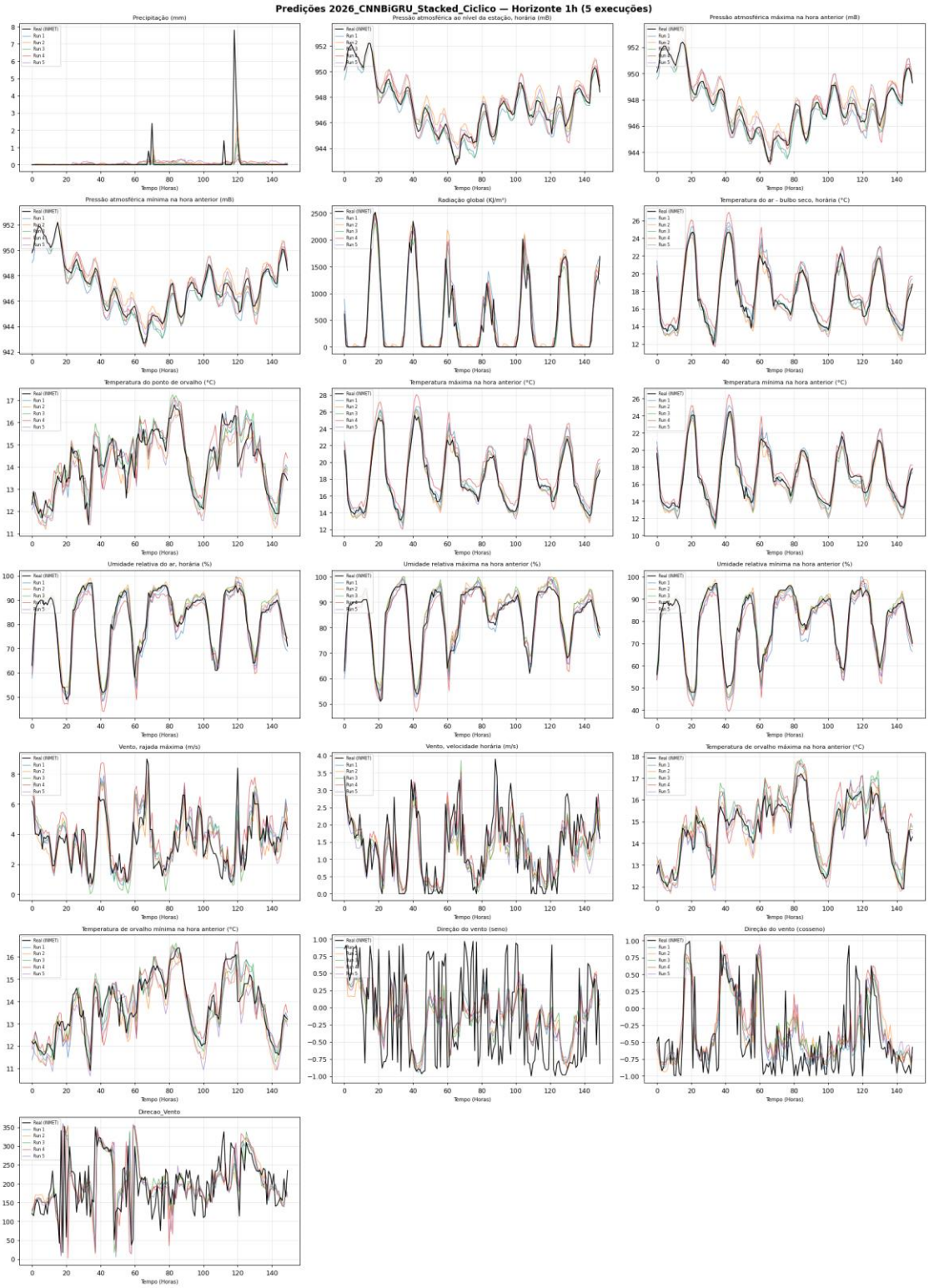
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	8.0355 ± 0.1858	0.6656 ± 0.0143	13.07 ± 0.39	8.3185 ± 2.2279
Umidade relativa do ar, horária (%)	7.6509 ± 0.2063	0.6609 ± 0.0177	11.92 ± 0.37	8.0830 ± 1.9701
Vento, rajada máxima (m/s)	2.5870 ± 0.2982	0.0463 ± 0.1610	100.27 ± 3.55	2.1811 ± 0.2507
Vento, velocidade horária (m/s)	1.2550 ± 0.1632	0.0503 ± 0.1434	139.62 ± 0.63	1.1001 ± 0.1844
Direção do vento (cosseno)	0.4397 ± 0.0040	0.2178 ± 0.0198	93.15 ± 2.26	0.3985 ± 0.0369
Direção do vento (seno)	0.6102 ± 0.0024	0.1147 ± 0.0020	150.19 ± 2.08	0.6259 ± 0.0168

Métricas — Horizonte 168h

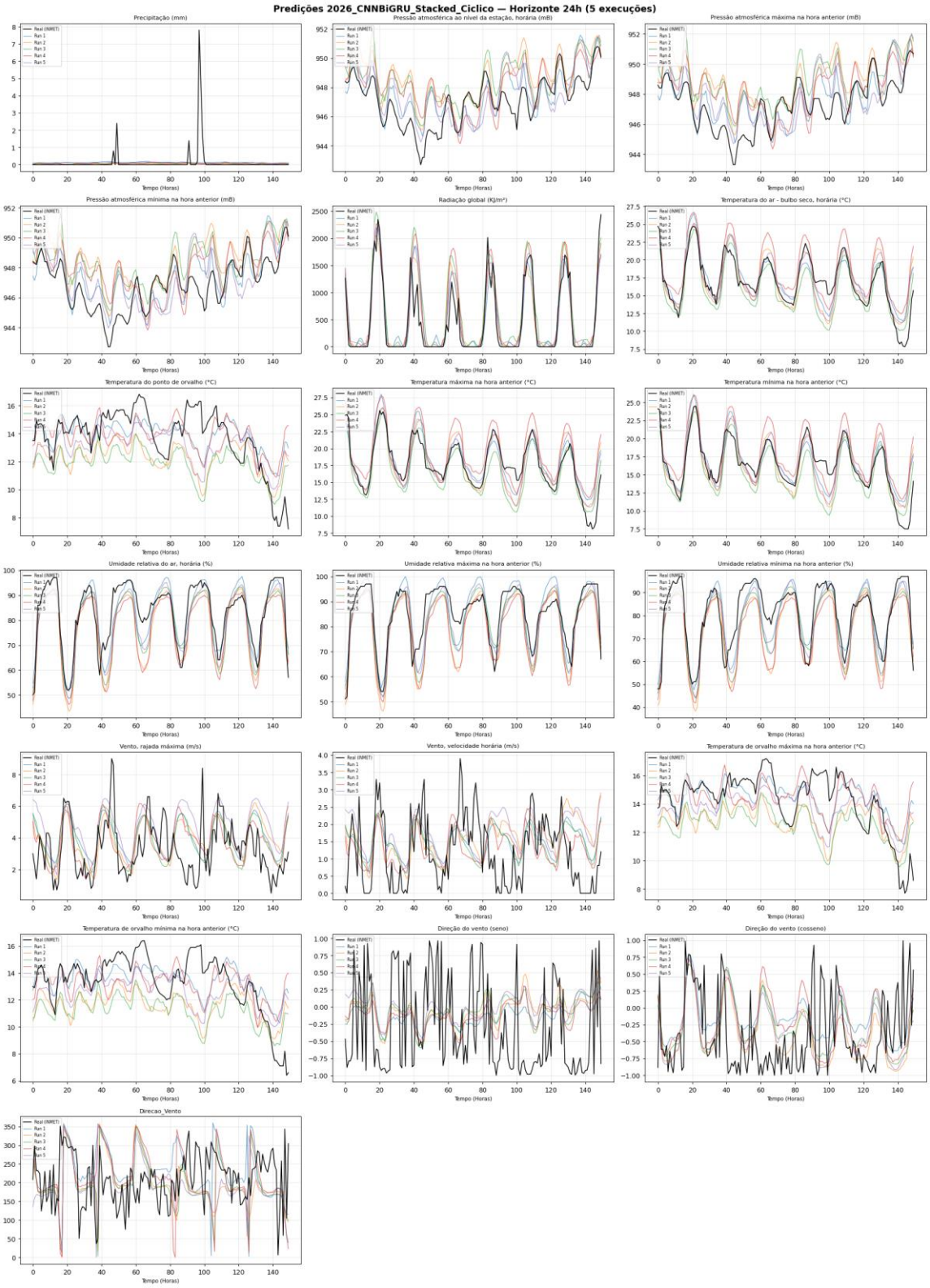
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.2769 ± 0.0364	0.3666 ± 0.0158	0.24 ± 0.00	1.6927 ± 0.0342
Pressão atmosférica máxima na hora anterior (mB)	2.2743 ± 0.0384	0.3619 ± 0.0209	0.24 ± 0.00	1.6631 ± 0.0457
Pressão atmosférica mínima na hora anterior (mB)	2.2912 ± 0.0258	0.3696 ± 0.0127	0.24 ± 0.00	1.6644 ± 0.0393
Precipitação (mm)	0.1827 ± 0.0240	-0.0014 ± 0.0046	169.19 ± 35.88	0.1831 ± 0.0269
Radiação global (KJ/m²)	278.8363 ± 9.2616	0.8011 ± 0.0119	66.72 ± 18.03	542.4281 ± 74.8493
Temperatura do ar - bulbo seco, horária (°C)	2.5227 ± 0.0346	0.6265 ± 0.0070	12.77 ± 0.19	2.9094 ± 0.1130
Temperatura do ponto de orvalho (°C)	2.0379 ± 0.0377	0.5242 ± 0.0141	15.08 ± 0.23	1.6535 ± 0.0262
Temperatura máxima na hora anterior (°C)	2.5775 ± 0.0274	0.6276 ± 0.0031	12.56 ± 0.15	2.9934 ± 0.2691
Temperatura mínima na hora anterior (°C)	2.4624 ± 0.0124	0.6302 ± 0.0049	13.00 ± 0.03	2.7245 ± 0.1191
Temperatura de orvalho máxima na hora anterior (°C)	1.9720 ± 0.0333	0.5411 ± 0.0139	13.89 ± 0.20	1.5778 ± 0.0276
Temperatura de orvalho mínima na hora anterior (°C)	2.1108 ± 0.0543	0.5098 ± 0.0218	16.44 ± 0.35	1.6598 ± 0.0633
Umidade relativa máxima na hora anterior (%)	8.3159 ± 0.2739	0.5620 ± 0.0210	12.57 ± 0.31	10.5118 ± 1.2290

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.5611 ± 0.3619	0.5675 ± 0.0299	15.80 ± 0.45	10.9276 ± 0.3454
Umidade relativa do ar, horária (%)	9.0975 ± 0.2764	0.5601 ± 0.0229	14.39 ± 0.34	11.6036 ± 1.0312
Vento, rajada máxima (m/s)	2.9201 ± 0.1323	-0.1328 ± 0.0824	104.76 ± 1.15	2.6810 ± 0.1057
Vento, velocidade horária (m/s)	1.4779 ± 0.0852	-0.1786 ± 0.1030	142.20 ± 0.24	1.2683 ± 0.2001
Direção do vento (cosseno)	0.4624 ± 0.0051	0.1686 ± 0.0053	100.40 ± 1.86	0.4183 ± 0.0109
Direção do vento (seno)	0.6415 ± 0.0017	0.0399 ± 0.0084	158.79 ± 6.11	0.6533 ± 0.0120

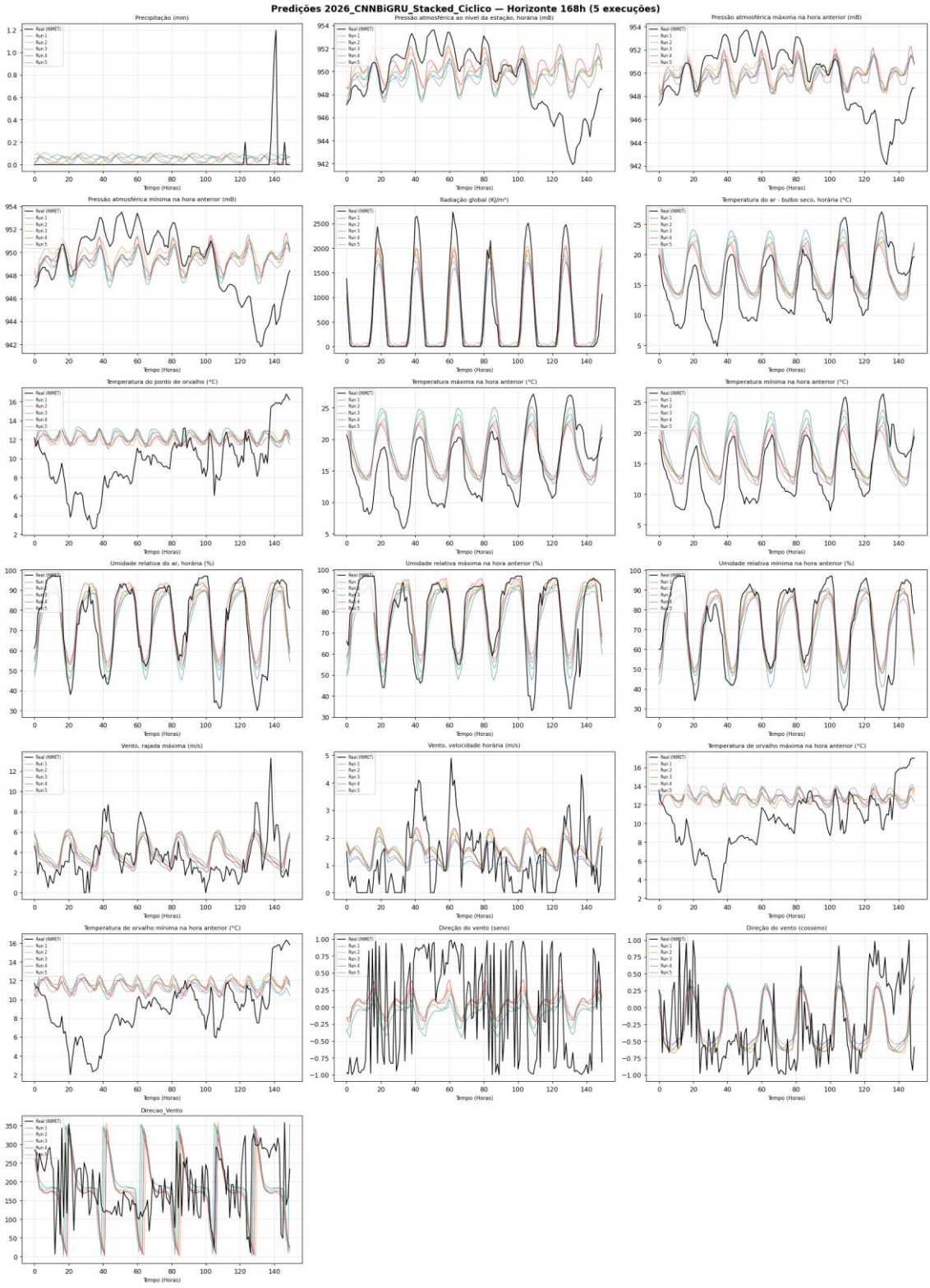
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-BiLSTM-GRU

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.5897 ± 0.0592	0.9586 ± 0.0073	0.06 ± 0.00	0.4030 ± 0.0392
Pressão atmosférica máxima na hora anterior (mB)	0.5383 ± 0.0512	0.9644 ± 0.0056	0.06 ± 0.01	0.3864 ± 0.0413
Pressão atmosférica mínima na hora anterior (mB)	0.5541 ± 0.0597	0.9635 ± 0.0071	0.06 ± 0.01	0.3769 ± 0.0333
Precipitação (mm)	0.1538 ± 0.0099	0.0678 ± 0.0237	99.37 ± 46.35	0.1478 ± 0.0093
Radiação global (KJ/m²)	152.9123 ± 7.2917	0.9327 ± 0.0023	69.11 ± 22.56	259.8008 ± 140.0045
Temperatura do ar - bulbo seco, horária (°C)	0.8268 ± 0.0168	0.9547 ± 0.0016	4.17 ± 0.08	0.5860 ± 0.0212
Temperatura do ponto de orvalho (°C)	0.6688 ± 0.0337	0.9412 ± 0.0052	5.11 ± 0.25	0.4312 ± 0.0275
Temperatura máxima na hora anterior (°C)	0.7231 ± 0.0275	0.9695 ± 0.0023	3.53 ± 0.15	0.5374 ± 0.0117
Temperatura mínima na hora anterior (°C)	0.7179 ± 0.0245	0.9636 ± 0.0021	3.79 ± 0.14	0.5128 ± 0.0176
Temperatura de orvalho máxima na hora anterior (°C)	0.5998 ± 0.0313	0.9546 ± 0.0044	4.31 ± 0.22	0.4405 ± 0.0237
Temperatura de orvalho mínima na hora anterior (°C)	0.5891 ± 0.0317	0.9570 ± 0.0046	4.87 ± 0.27	0.4082 ± 0.0241

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.6600 ± 0.1375	0.9530 ± 0.0032	3.96 ± 0.20	2.0258 ± 0.1016
Umidade relativa mínima na hora anterior (%)	2.7377 ± 0.1730	0.9653 ± 0.0032	4.52 ± 0.28	2.1541 ± 0.1082
Umidade relativa do ar, horária (%)	3.1833 ± 0.0933	0.9409 ± 0.0024	5.00 ± 0.17	2.2643 ± 0.0843
Vento, rajada máxima (m/s)	1.1628 ± 0.0857	0.7191 ± 0.0195	74.06 ± 3.14	1.1196 ± 0.2046
Vento, velocidade horária (m/s)	0.5355 ± 0.0322	0.6940 ± 0.0102	118.83 ± 11.67	0.4419 ± 0.0742
Direção do vento (cosseno)	0.3212 ± 0.0081	0.5371 ± 0.0144	68.91 ± 2.36	0.2494 ± 0.0326
Direção do vento (seno)	0.5033 ± 0.0067	0.3137 ± 0.0103	120.79 ± 2.70	0.4540 ± 0.0246

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7758 ± 0.0742	0.6243 ± 0.0250	0.19 ± 0.01	0.9130 ± 0.0963
Pressão atmosférica máxima na hora anterior (mB)	1.7461 ± 0.0507	0.6307 ± 0.0182	0.18 ± 0.01	0.8991 ± 0.0859
Pressão atmosférica mínima na hora anterior (mB)	1.7786 ± 0.0719	0.6287 ± 0.0243	0.19 ± 0.01	0.9098 ± 0.0948
Precipitação (mm)	0.1928 ± 0.0173	0.0028 ± 0.0031	169.06 ± 14.55	0.1908 ± 0.0157
Radiação global (KJ/m²)	247.8321 ± 11.3678	0.8313 ± 0.0041	70.53 ± 13.55	389.6886 ± 55.9239
Temperatura do ar - bulbo seco, horária (°C)	1.9922 ± 0.0306	0.7530 ± 0.0060	10.12 ± 0.19	1.9984 ± 0.2470
Temperatura do ponto de orvalho (°C)	1.7097 ± 0.0626	0.6312 ± 0.0300	12.73 ± 0.41	1.0335 ± 0.0778
Temperatura máxima na hora anterior (°C)	2.0315 ± 0.0322	0.7547 ± 0.0082	9.94 ± 0.19	2.0338 ± 0.2799
Temperatura mínima na hora anterior (°C)	1.9435 ± 0.0235	0.7565 ± 0.0061	10.30 ± 0.13	1.8962 ± 0.2131
Temperatura de orvalho máxima na hora anterior (°C)	1.6387 ± 0.0615	0.6519 ± 0.0305	11.62 ± 0.39	1.0412 ± 0.0725
Temperatura de orvalho mínima na hora anterior (°C)	1.7390 ± 0.0586	0.6291 ± 0.0272	13.65 ± 0.38	1.0585 ± 0.0716
Umidade relativa máxima na hora anterior (%)	7.1534 ± 0.2343	0.6543 ± 0.0268	10.63 ± 0.31	7.2675 ± 1.2106

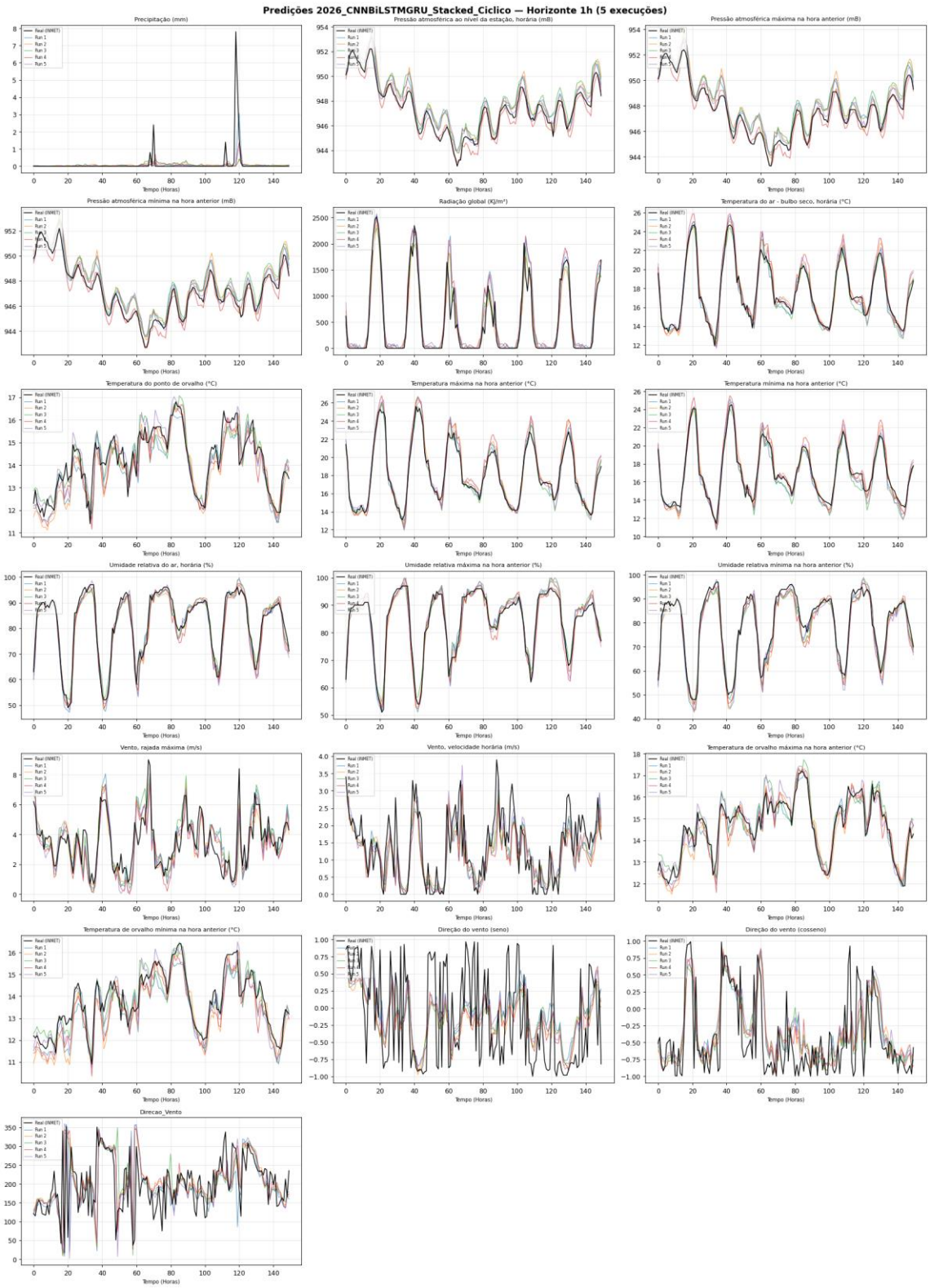
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	8.1176 ± 0.2110	0.6607 ± 0.0220	13.18 ± 0.34	7.5421 ± 1.4664
Umidade relativa do ar, horária (%)	7.7246 ± 0.1952	0.6565 ± 0.0216	12.02 ± 0.30	7.5779 ± 1.0848
Vento, rajada máxima (m/s)	2.5439 ± 0.1462	0.0974 ± 0.0785	100.37 ± 1.59	2.2900 ± 0.2441
Vento, velocidade horária (m/s)	1.1975 ± 0.0280	0.1281 ± 0.0302	140.06 ± 0.31	1.0933 ± 0.0743
Direção do vento (cosseno)	0.4405 ± 0.0022	0.2311 ± 0.0040	94.55 ± 1.46	0.3838 ± 0.0159
Direção do vento (seno)	0.6126 ± 0.0043	0.1092 ± 0.0108	149.67 ± 3.25	0.6254 ± 0.0215

Métricas — Horizonte 168h

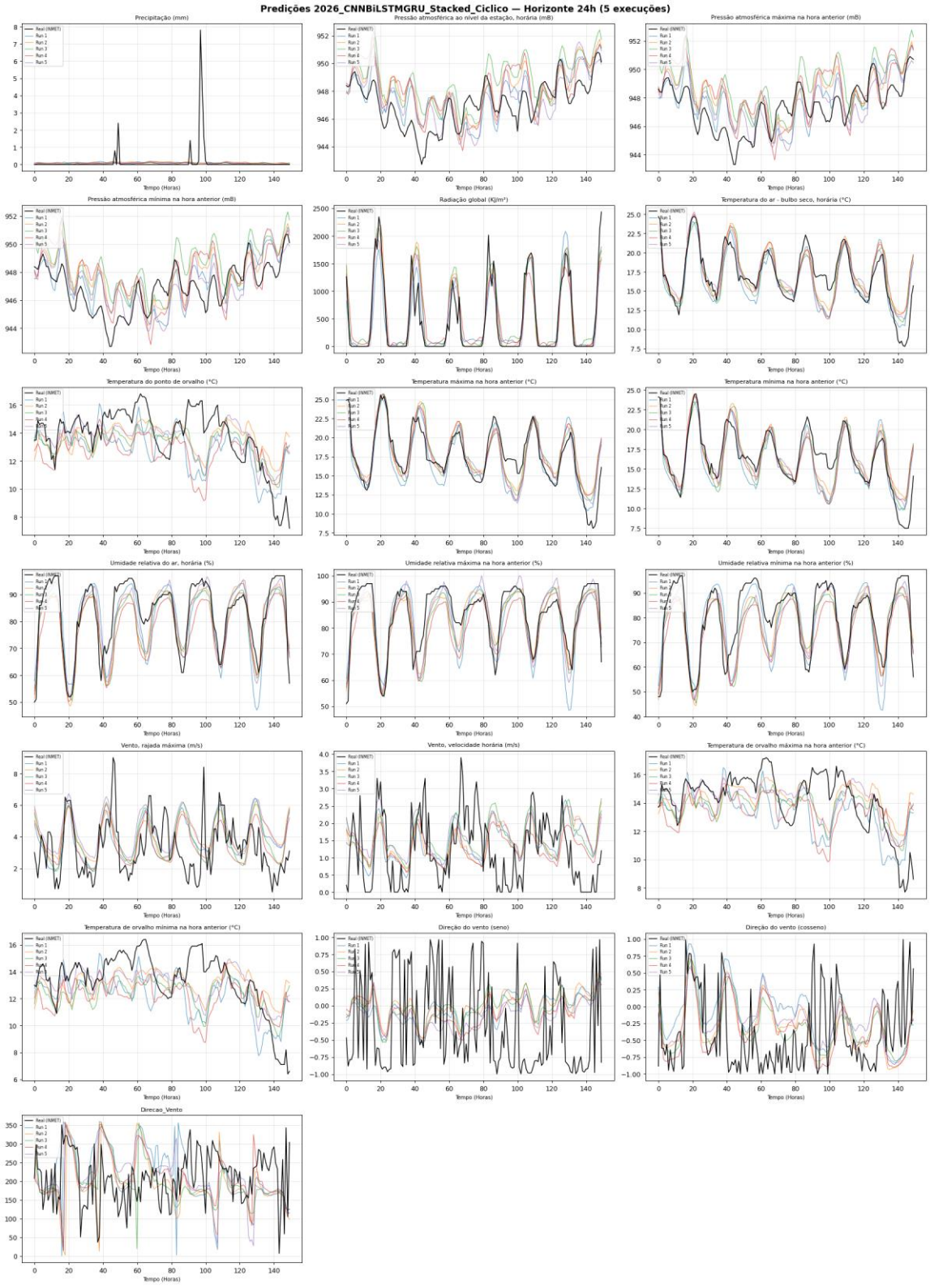
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.2909 ± 0.0340	0.3615 ± 0.0139	0.24 ± 0.00	1.6618 ± 0.0609
Pressão atmosférica máxima na hora anterior (mB)	2.3036 ± 0.0577	0.3462 ± 0.0325	0.24 ± 0.00	1.6486 ± 0.0528
Pressão atmosférica mínima na hora anterior (mB)	2.3291 ± 0.0682	0.3510 ± 0.0263	0.24 ± 0.01	1.6556 ± 0.0859
Precipitação (mm)	0.1828 ± 0.0204	-0.0003 ± 0.0047	171.69 ± 15.29	0.1851 ± 0.0158
Radiação global (KJ/m²)	281.3100 ± 13.8784	0.8044 ± 0.0083	74.57 ± 14.05	526.9129 ± 87.0585
Temperatura do ar - bulbo seco, horária (°C)	2.4736 ± 0.0261	0.6384 ± 0.0064	12.59 ± 0.17	2.7403 ± 0.0593
Temperatura do ponto de orvalho (°C)	2.0655 ± 0.0465	0.5049 ± 0.0153	15.31 ± 0.32	1.5728 ± 0.0284
Temperatura máxima na hora anterior (°C)	2.5535 ± 0.0758	0.6320 ± 0.0199	12.52 ± 0.36	2.7572 ± 0.1111
Temperatura mínima na hora anterior (°C)	2.4178 ± 0.0685	0.6409 ± 0.0172	12.85 ± 0.36	2.6105 ± 0.2112
Temperatura de orvalho máxima na hora anterior (°C)	1.9948 ± 0.0485	0.5309 ± 0.0128	14.06 ± 0.30	1.5299 ± 0.0404
Temperatura de orvalho mínima na hora anterior (°C)	2.1269 ± 0.0420	0.4983 ± 0.0096	16.59 ± 0.33	1.6667 ± 0.0517
Umidade relativa máxima na hora anterior (%)	8.2386 ± 0.1636	0.5724 ± 0.0143	12.45 ± 0.22	9.6655 ± 0.7629

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.3998 ± 0.1082	0.5835 ± 0.0084	15.55 ± 0.19	10.6796 ± 0.5064
Umidade relativa do ar, horária (%)	8.9663 ± 0.1721	0.5729 ± 0.0171	14.20 ± 0.27	9.9896 ± 0.8920
Vento, rajada máxima (m/s)	2.9266 ± 0.0741	-0.1370 ± 0.0406	104.97 ± 0.81	2.6718 ± 0.1309
Vento, velocidade horária (m/s)	1.5004 ± 0.0618	-0.2072 ± 0.0753	142.25 ± 0.28	1.2937 ± 0.0849
Direção do vento (cosseno)	0.4615 ± 0.0021	0.1664 ± 0.0099	100.14 ± 0.84	0.4203 ± 0.0066
Direção do vento (seno)	0.6416 ± 0.0018	0.0422 ± 0.0048	159.13 ± 1.61	0.6563 ± 0.0268

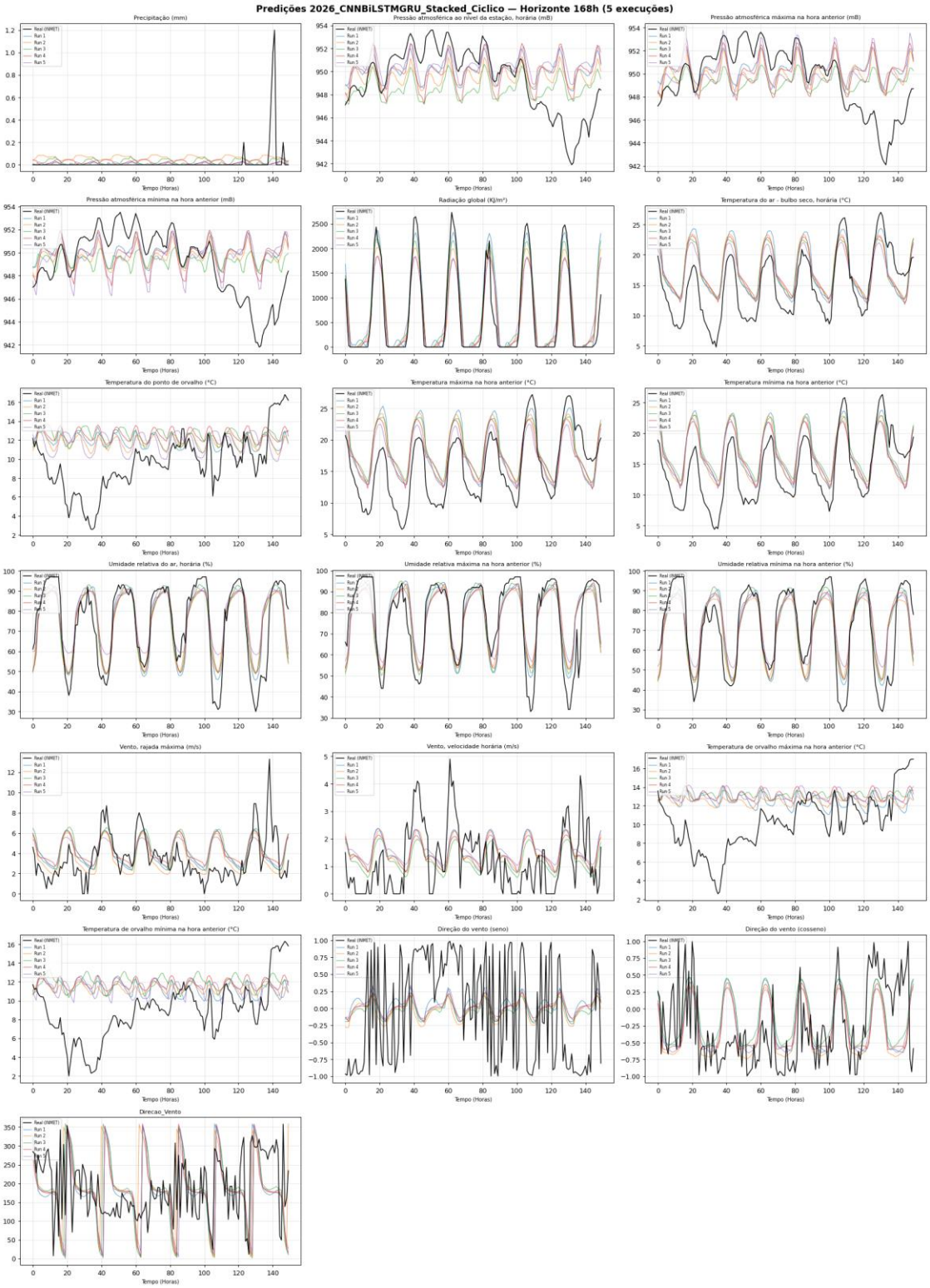
Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h



CNN-LSTM (sem *pooling*)

Métricas — Horizonte 1h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	0.5295 ± 0.0714	0.9657 ± 0.0091	0.06 ± 0.01	0.3723 ± 0.0535
Pressão atmosférica máxima na hora anterior (mB)	0.4920 ± 0.0790	0.9692 ± 0.0096	0.05 ± 0.01	0.3507 ± 0.0548
Pressão atmosférica mínima na hora anterior (mB)	0.4909 ± 0.0804	0.9708 ± 0.0096	0.05 ± 0.01	0.3497 ± 0.0536
Precipitação (mm)	0.1562 ± 0.0148	0.0875 ± 0.0097	128.53 ± 55.91	0.1460 ± 0.0121
Radiação global (KJ/m²)	151.8605 ± 7.0166	0.9361 ± 0.0023	70.49 ± 22.68	202.3097 ± 84.5145
Temperatura do ar - bulbo seco, horária (°C)	0.7787 ± 0.0522	0.9594 ± 0.0048	3.92 ± 0.25	0.5431 ± 0.0276
Temperatura do ponto de orvalho (°C)	0.6466 ± 0.0276	0.9456 ± 0.0033	4.91 ± 0.21	0.4187 ± 0.0289
Temperatura máxima na hora anterior (°C)	0.6546 ± 0.0628	0.9741 ± 0.0049	3.18 ± 0.28	0.4882 ± 0.0388
Temperatura mínima na hora anterior (°C)	0.7054 ± 0.0866	0.9657 ± 0.0066	3.71 ± 0.41	0.5205 ± 0.0686
Temperatura de orvalho máxima na hora anterior (°C)	0.5600 ± 0.0369	0.9606 ± 0.0039	4.01 ± 0.27	0.4007 ± 0.0339
Temperatura de orvalho mínima na hora anterior (°C)	0.6009 ± 0.0518	0.9568 ± 0.0054	4.92 ± 0.40	0.4324 ± 0.0374

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa máxima na hora anterior (%)	2.8175 ± 0.3968	0.9501 ± 0.0078	4.08 ± 0.50	2.7442 ± 1.2577
Umidade relativa mínima na hora anterior (%)	2.6204 ± 0.2396	0.9672 ± 0.0039	4.30 ± 0.35	2.1141 ± 0.3445
Umidade relativa do ar, horária (%)	3.2172 ± 0.1543	0.9402 ± 0.0040	4.99 ± 0.19	2.2717 ± 0.0914
Vento, rajada máxima (m/s)	1.1985 ± 0.0503	0.7137 ± 0.0078	76.34 ± 1.70	1.2555 ± 0.1349
Vento, velocidade horária (m/s)	0.5235 ± 0.0221	0.6983 ± 0.0052	120.71 ± 6.37	0.4478 ± 0.0093
Direção do vento (cosseno)	0.3208 ± 0.0093	0.5275 ± 0.0240	68.28 ± 2.49	0.2429 ± 0.0193
Direção do vento (seno)	0.4990 ± 0.0055	0.3231 ± 0.0075	119.94 ± 2.73	0.4327 ± 0.0215

Métricas — Horizonte 24h

Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	1.7899 ± 0.0787	0.6160 ± 0.0245	0.19 ± 0.01	0.9382 ± 0.1316
Pressão atmosférica máxima na hora anterior (mB)	1.7721 ± 0.0785	0.6181 ± 0.0237	0.19 ± 0.01	0.9381 ± 0.1259
Pressão atmosférica mínima na hora anterior (mB)	1.7864 ± 0.0775	0.6228 ± 0.0242	0.19 ± 0.01	0.9440 ± 0.1178
Precipitação (mm)	0.1829 ± 0.0164	0.0006 ± 0.0038	162.89 ± 18.82	0.1811 ± 0.0145
Radiação global (KJ/m²)	242.8444 ± 8.1977	0.8363 ± 0.0018	57.48 ± 22.86	332.8257 ± 52.4088
Temperatura do ar - bulbo seco, horária (°C)	1.9452 ± 0.0167	0.7630 ± 0.0052	9.84 ± 0.07	2.0042 ± 0.0957
Temperatura do ponto de orvalho (°C)	1.6716 ± 0.0513	0.6533 ± 0.0190	12.48 ± 0.37	1.0147 ± 0.0820
Temperatura máxima na hora anterior (°C)	1.9821 ± 0.0297	0.7653 ± 0.0056	9.66 ± 0.14	2.0489 ± 0.2277
Temperatura mínima na hora anterior (°C)	1.9053 ± 0.0373	0.7653 ± 0.0070	10.06 ± 0.21	1.8532 ± 0.1341
Temperatura de orvalho máxima na hora anterior (°C)	1.6052 ± 0.0585	0.6708 ± 0.0261	11.41 ± 0.42	0.9944 ± 0.0687
Temperatura de orvalho mínima na hora anterior (°C)	1.7346 ± 0.0992	0.6388 ± 0.0376	13.65 ± 0.73	1.0421 ± 0.0985
Umidade relativa máxima na hora anterior (%)	7.1092 ± 0.3405	0.6614 ± 0.0280	10.63 ± 0.54	7.0499 ± 1.6527

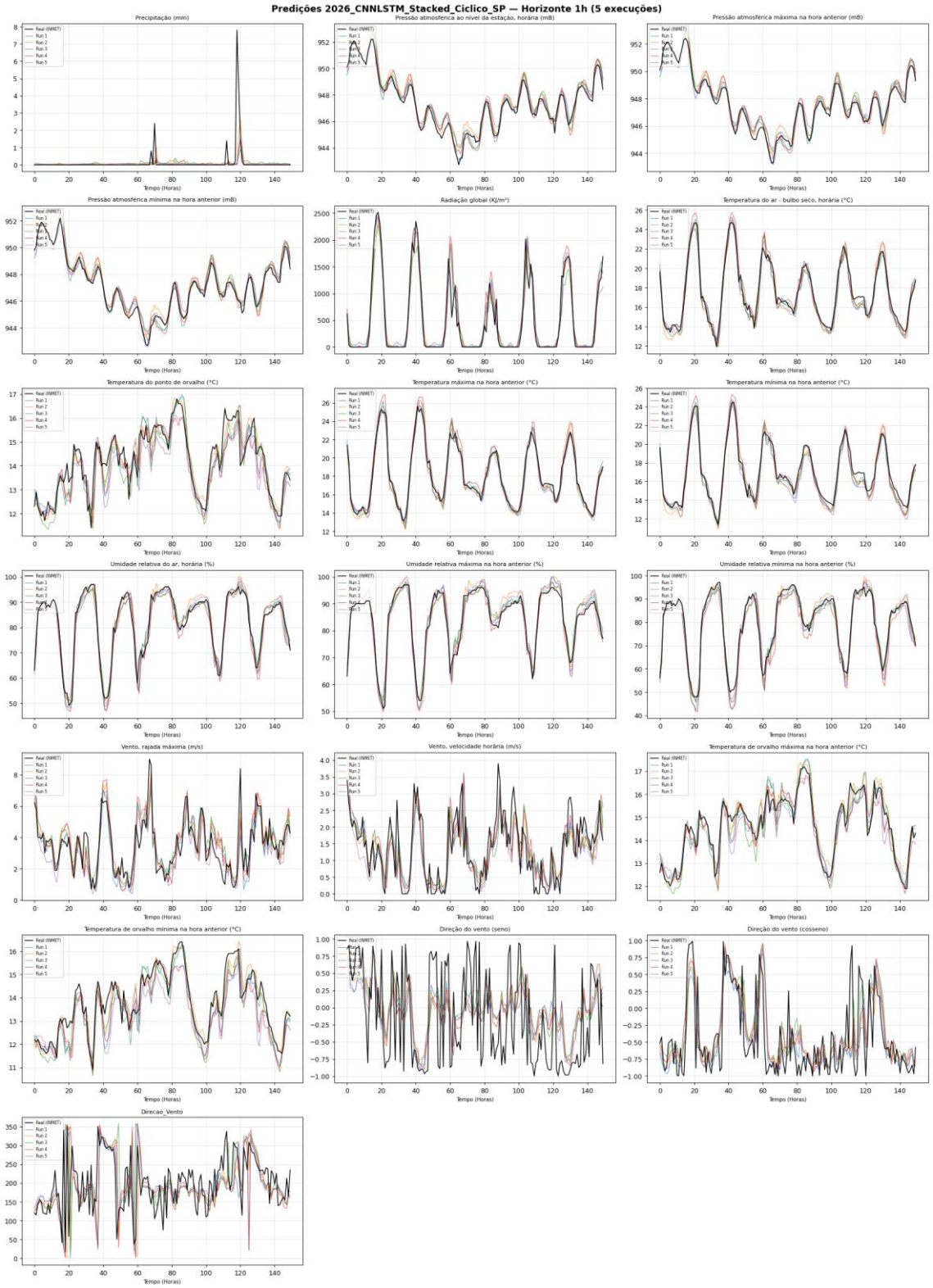
Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	8.0015 ± 0.2934	0.6749 ± 0.0209	13.07 ± 0.50	8.2956 ± 2.2801
Umidade relativa do ar, horária (%)	7.6796 ± 0.3350	0.6642 ± 0.0252	12.01 ± 0.55	7.7751 ± 1.9355
Vento, rajada máxima (m/s)	2.6153 ± 0.1508	0.0471 ± 0.0997	100.88 ± 1.29	2.2140 ± 0.0984
Vento, velocidade horária (m/s)	1.2487 ± 0.1355	0.0610 ± 0.1291	139.96 ± 0.30	1.0999 ± 0.1166
Direção do vento (cosseno)	0.4410 ± 0.0048	0.2253 ± 0.0153	94.14 ± 2.15	0.3942 ± 0.0162
Direção do vento (seno)	0.6127 ± 0.0021	0.1051 ± 0.0164	150.01 ± 2.98	0.6262 ± 0.0225

Métricas — Horizonte 168h

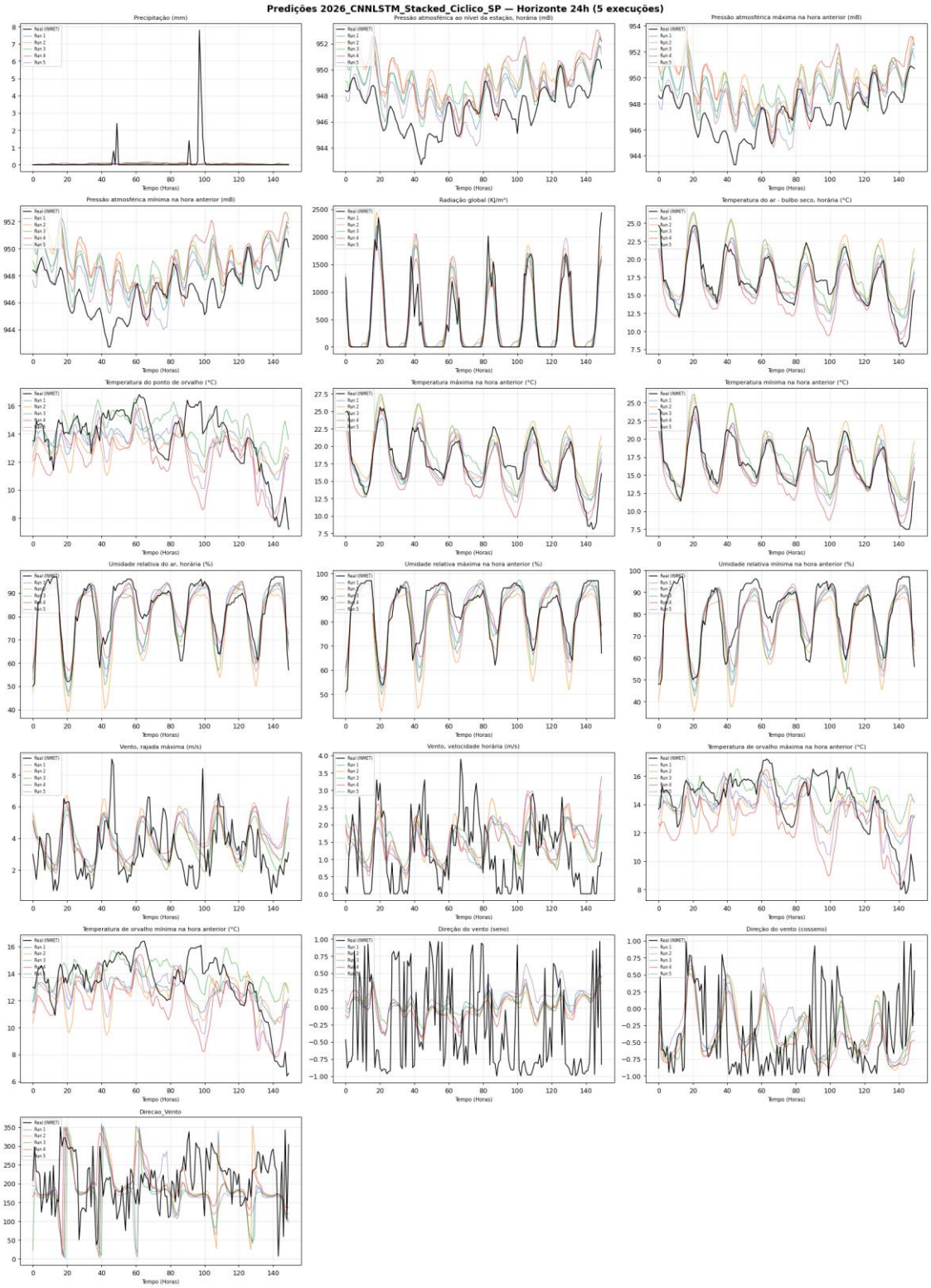
Variável	MAE	R^2	SMAPE (%)	DTW
Pressão atmosférica ao nível da estação, horária (mB)	2.2952 ± 0.0288	0.3662 ± 0.0142	0.24 ± 0.00	1.6929 ± 0.0510
Pressão atmosférica máxima na hora anterior (mB)	2.2757 ± 0.0302	0.3663 ± 0.0151	0.24 ± 0.00	1.6682 ± 0.0699
Pressão atmosférica mínima na hora anterior (mB)	2.2993 ± 0.0246	0.3699 ± 0.0150	0.24 ± 0.00	1.6870 ± 0.0641
Precipitação (mm)	0.1823 ± 0.0057	0.0007 ± 0.0014	192.68 ± 8.63	0.1842 ± 0.0089
Radiação global (KJ/m²)	270.1542 ± 18.0293	0.8092 ± 0.0136	73.21 ± 16.60	463.0251 ± 41.3023
Temperatura do ar - bulbo seco, horária (°C)	2.4969 ± 0.0691	0.6337 ± 0.0176	12.65 ± 0.33	2.8899 ± 0.1832
Temperatura do ponto de orvalho (°C)	2.0420 ± 0.0153	0.5202 ± 0.0096	15.13 ± 0.09	1.6091 ± 0.0254
Temperatura máxima na hora anterior (°C)	2.5432 ± 0.0602	0.6365 ± 0.0146	12.40 ± 0.29	3.0316 ± 0.1892
Temperatura mínima na hora anterior (°C)	2.4303 ± 0.0754	0.6391 ± 0.0194	12.85 ± 0.39	2.7723 ± 0.1114
Temperatura de orvalho máxima na hora anterior (°C)	1.9650 ± 0.0144	0.5399 ± 0.0094	13.85 ± 0.09	1.4898 ± 0.0326
Temperatura de orvalho mínima na hora anterior (°C)	2.1084 ± 0.0155	0.5048 ± 0.0107	16.45 ± 0.08	1.6385 ± 0.0378
Umidade relativa máxima na hora anterior (%)	8.3311 ± 0.1323	0.5656 ± 0.0165	12.60 ± 0.20	9.8603 ± 0.9718

Variável	MAE	R^2	SMAPE (%)	DTW
Umidade relativa mínima na hora anterior (%)	9.5941 ± 0.2193	0.5686 ± 0.0186	15.84 ± 0.34	11.8706 ± 1.6530
Umidade relativa do ar, horária (%)	9.0286 ± 0.1546	0.5688 ± 0.0153	14.30 ± 0.25	11.0121 ± 1.5688
Vento, rajada máxima (m/s)	2.9383 ± 0.1054	-0.1498 ± 0.0787	104.94 ± 0.86	2.6385 ± 0.1485
Vento, velocidade horária (m/s)	1.4803 ± 0.0574	-0.2016 ± 0.0756	142.19 ± 0.15	1.0642 ± 0.1674
Direção do vento (cosseno)	0.4621 ± 0.0031	0.1639 ± 0.0183	100.78 ± 2.36	0.4624 ± 0.0286
Direção do vento (seno)	0.6397 ± 0.0010	0.0489 ± 0.0023	158.79 ± 2.32	0.6521 ± 0.0159

Curvas de predição — Horizonte 1h



Curvas de predição — Horizonte 24h



Curvas de predição — Horizonte 168h

