

Modelos de Deep Learning em auxílio à previsão de variáveis meteorológicas em aeroportos na Antártica

Alana de Lima Pontes Gadelha¹, Lúcia M. A. Drummond², Leandro Santiago de Araújo²

¹Centro de Hidrografia da Marinha (CHM) – Niterói – RJ – Brazil

²Instituto de Computação, Universidade Federal Fluminense (UFF) – Niterói-RJ – Brasil

alana@id.uff.br, {lucia, leandro}@ic.uff.br

Abstract. This study investigates the use of Deep Learning models for forecasting hourly atmospheric pressure and the temperature at the SCRM Aerodrome in Antarctica, based on nearly 15 years of METAR data. The evaluated models were MLP, CNN, LSTM, and BiLSTM, with the BiLSTM achieving the lowest errors and showing strong ability to capture long-term dependencies. Forecasts were also compared with the state-of-the-art ICON-LAM numerical model: while ICON-LAM demonstrated high accuracy, the BiLSTM achieved competitive performance with much lower computational cost. These results highlight the potential of memory-based Deep Learning models as complementary tools to numerical weather prediction, enhancing forecasting and operational safety in polar regions.

1. Introdução

Muitos estudos têm explorado a aplicação de técnicas de *Deep Learning* à previsão do tempo, entretanto, permanece uma lacuna na compreensão de sua eficácia em previsões localizadas ou pontuais [Xu 2024]. Previsões meteorológicas precisas permitem uma melhor tomada de decisão e um melhor planejamento, visando a segurança de aterrissagens e decolagens de aeronaves em locais inóspitos, como o Continente Antártico [Zhang et al. 2025]. Neste contexto, a motivação deste estudo vem do desafio de prever, variáveis meteorológicas locais, as quais são cruciais para uma correta previsão do tempo em aeroportos. A previsão do tempo tradicional baseia-se principalmente em modelos de Previsão Numérica do Tempo (*Numerical Weather Prediction - NWP*), que necessitam de recursos computacionais substanciais para a realização de uma simulação [Zhang et al. 2025].

O trabalho apresenta a previsão de duas variáveis meteorológicas, temperatura do ar à superfície e pressão atmosférica reduzida ao nível médio do mar (PRNMM), por meio de séries temporais. Para isto, foram utilizados os seguintes algoritmos de *Deep Learning*: (i) Perceptron Multi-Camada (*Multi-Layer Perceptron - MLP*), (ii) Redes Neurais Convolucionais (*Convolutional Neural Network - CNN*), (iii) *Long short term Memory - LSTM*, e (iv) *Bidirectional Long short term Memory (BiLSTM)*, sendo estas duas últimas, tipos de Redes Neurais Recorrentes. Esses algoritmos foram treinados e avaliados sobre uma série temporal, com mais de 15 anos de dados meteorológicos observados, na Antártica. Estes dados foram coletados na estação meteorológica automática do aeroporto Tenente Rodolfo Marsh Martin, na Base Chilena Presidente Eduardo Frei Montalva, na Antártica ¹, cuja localização é fundamental para a segurança e o planejamento

¹código ICAO (*International Civil Aviation Organization*: SCRM)

de operações civis e militares naquela região, inclusive para o Brasil, em cumprimento ao PROgrama ANTÁRtico Brasileiro (PROANTAR)². Além da comparação com os dados observados, também foi utilizada uma série temporal de 1 ano e 6 meses proveniente do modelo numérico ICON-LAM (*ICOsahedral Nonhydrostatic - Limited Area Mode*) para comparação dos resultados, uma vez que este modelo é amplamente empregado em previsões do tempo e clima para a região da Antártica.

As principais contribuições deste trabalho são: (i) Desenvolvimento de técnicas de pré-processamento de dados meteorológicos e extração adequada de *outputs* do modelo ICON-LAM; (ii) Proposta de uso de algoritmos de *Deep Learning* para previsão de variáveis meteorológicas locais em apoio à segurança de operações aeronáuticas. Embora neste trabalho, o foco tenha sido a previsão de duas variáveis, os algoritmos podem ser adotados para quaisquer outras variáveis, tais como: Temperatura do ponto de orvalho (°Celsius), Umidade Relativa (%), dentre outras.

2. Estratégias de Pré-processamento

A seguir são descritas as estratégias implementadas para realização do pré-processamento.

2.1. Coleta de Dados e Tratamento da série temporal da estação automática

A fase de pré-processamento envolve normalização, transformação e tratamento de valores omissos. Neste estudo, os dados foram baixados, por meio da biblioteca do Python meteostat³. Foram decodificados e convertidos para o formato de séries temporais para uma melhor visualização e entendimento dos padrões atmosféricos observados.

A série temporal totaliza mais de 15 anos de dados, começando em 01 de janeiro de 2010 até 30 de junho de 2025 e a Figura 1 apresenta o comportamento anual da temperatura e da pressão. Uma vez em posse desses dados, se iniciou a fase exploratória com o intuito de identificar a distribuição das variáveis meteorológicas, por meio do gráfico de correlação, e identificar quais delas seriam correlatas. A série temporal, foi treinada com os seguintes atributos: Temperatura do ar observada à 2 metros (°Celsius), Temperatura do ponto de orvalho (°Celsius), Umidade Relativa (%), Total de precipitação (milímetros), Direção do Vento (graus) e Velocidade do vento em (km/h), pressão atmosférica reduzida ao nível médio do mar (hPa) e “Tempo Presente”. Contudo, essa última foi removida por não apresentar correlação significativa com a temperatura do ar à 2 metros. O pré-processamento, por meio de uma matriz de correlação mostrou que as variáveis de Temperatura do ponto de orvalho e Umidade Relativa apresentaram as maiores correlações (repositório GitHub do projeto). A direção do vento apresentou correlação positiva, e por isso, várias transformações foram necessárias para adequá-los como entrada nos algoritmos de *Deep Learning*. Da mesma forma, o tempo (*timestamp*) foi adaptado para capturar padrões diários e sazonais. Por fim, todos os dados foram normalizados com a técnica de padronização (z-score), calculada por: atributo normalizado = (atributo – média) / (desvio padrão), sendo esta uma etapa essencial no pré-processamento para *Deep Learning*.

²Informações do PROANTAR: <https://www.marinha.mil.br/secirm/pt-br/proantar/sobre>

³Meteostat: <https://dev.meteostat.net/>

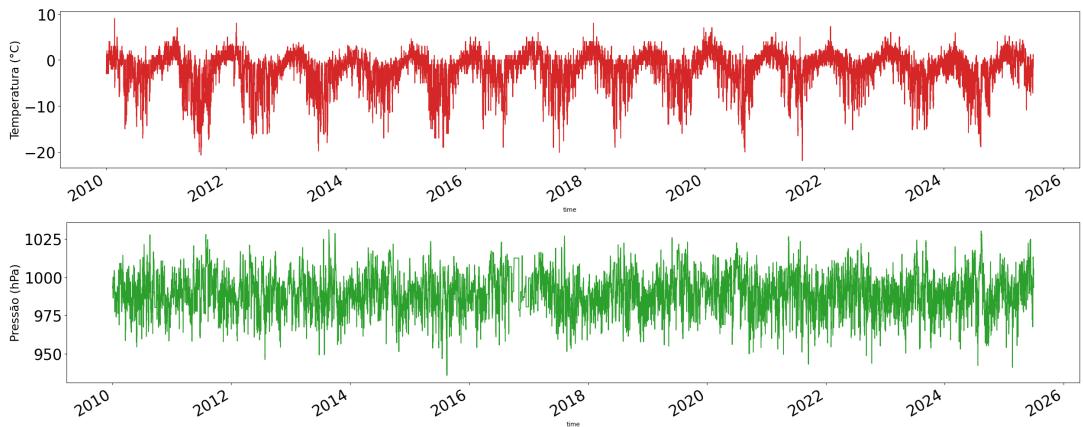


Figura 1. Série temporal da Temperatura do ar à superfície ($^{\circ}\text{Celsius}$) e da Pressão Reduzida ao Nível Médio do Mar (hPa), no aeroporto de SRCM, na Antártica, desde 01 de janeiro de 2010 até 30 de junho de 2025

2.2. Extração da série temporal do Modelo ICON-LAM

Utilizaram-se os dados do modelo numérico ICON-LAM [Iriza-Burcă et al. 2024], com resolução de 6,5 km, aplicado à Antártica. A grade utilizada foi de 1201 x 1121 pontos que cobre as latitudes de -56° a 14° e as longitudes de 270° a 345° . A série temporal foi extraída no ponto $(-62,1917^{\circ}, -58,9867^{\circ})$, correspondente as coordenadas do aeroporto SCRM, permitindo comparação direta com dados da estação meteorológica local.

3. Resultados Experimentais

A simulação do Modelo ICON-LAM foi executada de forma paralela utilizando MPI (*Message Passing Interface*), em um supercomputador HPE Apollo K6000 com 336 CPUs distribuídos em 7 nós de processamento Intel® Xeon® Gold 6248R @ 3,00 GHz, cada um com 48 cores e 197 GB de memória RAM. Enquanto que os experimentos com os algoritmos de *Deep Learning* foram realizados no ambiente do Google Colab, utilizando a biblioteca TensorFlow [Géron 2019].

3.1. Comparação dos Modelos de Deep-Learning e o Modelo ICON-LAM

Foram utilizados 15 anos de dados METAR⁴, divididos em 70% (aproximadamente 10 anos e 6 meses de dados), 20% (cerca de 3 anos de dados) e 10% (aproximadamente 1 ano e 6 meses de dados), respectivamente. A série foi organizada com janelas deslizantes (3, 6, 12, 24 e 48 passos de tempo) aplicadas tanto no horizonte de entrada quanto de saída. Todos os modelos foram treinados por 20 épocas, sendo o erro quadrático médio (*Mean Squared Error - MSE*) como função de custo; a MLP foi configurada com 2 camadas ocultas de 32 neurônios, tendo a ReLU (*Rectified Linear Unit*) como função de ativação. A CNN com uma camada convolucional (*kernel = 3*), a LSTM com 1 camada unidirecional de 32 neurônios e a BiLSTM com 1 camada bidirecional (32 neurônios, para o futuro da série e 32 para o passado da série, totalizando 64 em cada batch). A Figura 2 mostra um exemplo (um batch) de previsão de PRNMM e de temperatura, ambas horárias, realizada pelos modelos de *Deep Learning*, para o dia 19 de abril de 2024 às 18

⁴METAR (*Meteorological Aerodrome Report*), dados coletados por estações meteorológicas automáticas nos aeroportos.

horas até o dia 20 de abril às 18 horas, com um horizonte total de previsão de 24 horas. Embora a LSTM apresente bons resultados, a BiLSTM leva em consideração o passado e o futuro da série, ou seja, no exemplo mais promissor de nossos testes, valores de 48 steps anteriores, antes de realizar a previsão futura de cada variável; Demonstrou também, desempenho superior em suas métricas *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE) e *Root Mean Squared Error* (RMSE) Figura 3. Além dos dados dos modelos de *Deep Learning*, foram plotados os dados do modelo ICON-LAM após cada simulação, como mais um dado a ser comparado. O intuito aqui foi verificar se as simulações de modelos de *Deep Learning*, se assemelhariam as plotagens extraídas em termos de latitude e longitudes das variáveis meteorológicas.

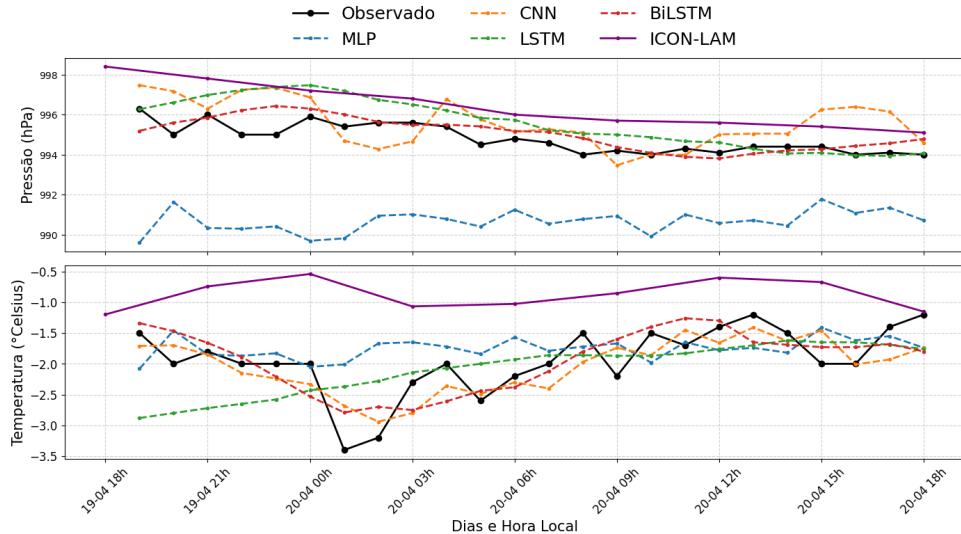


Figura 2. Comparação dos modelos de *Deep-Learning* para uma janela de previsão de 24 horas a frente, comparado com os dados reais observados da série temporal, e com os dados previstos pelo modelo numérico ICON-LAM.

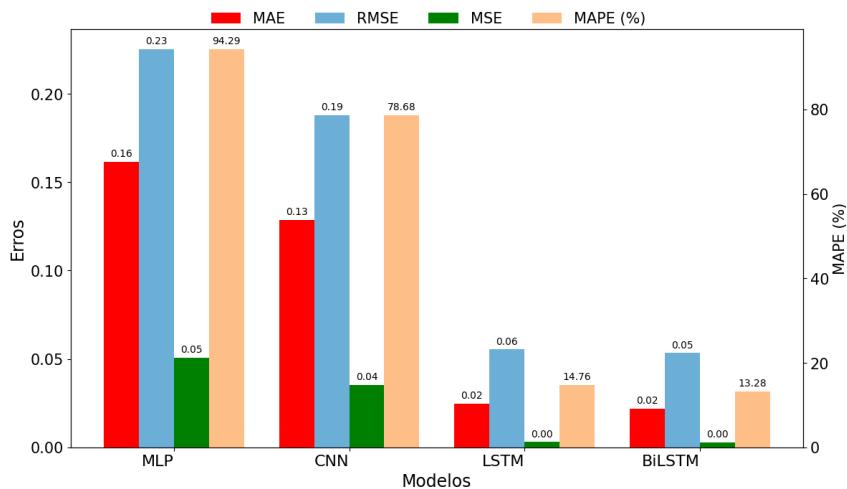


Figura 3. Métricas de erros dos modelos de *Deep Learning*

Os tempos de execução do conjunto de treinamento e teste/inferência foram coletados, conforme apresentado na Tabela 1. Observou-se que, na inferência, modelos

leves (MLP/CNN) apresentaram maior latência em GPU do que em CPU. Isso se deve ao custo fixo de transferência host-device (CPU-GPU) e de inicialização de kernels, que é amortizado apenas com lotes maiores ou computação mais intensa. Em janelas curtas e redes rasas, a CPU mantém alta eficiência. Em contrapartida, modelos recorrentes (LSTM/BiLSTM) possuem maior densidade computacional por amostra, explorando melhor o paralelismo da GPU, o que explica os menores tempos de teste na GPU para essas arquiteturas.

Modelo	Treino CPU	Treino GPU	Teste CPU	Teste GPU
MLP(24 Inputs)	73,618	62,407	1,141	2,614
CNN(3 Inputs)	119,969	95,317	1,125	1,612
LSTM(12 Inputs)	1038,652	635,180	2,968	1,742
BiLSTM(24 Inputs)	1829,633	580,637	3,867	2,074

Tabela 1. Tempos de execução em segundos, para treinamento e testes/inferências dos modelos, tanto em CPU quanto em GPU.

4. Conclusão

O foco deste estudo foi apresentar uma alternativa de previsão de variáveis meteorológicas em pontos localizados, como é o caso de um aeroporto SCRM, onde as condições atmosféricas precisam ser avaliadas em apoio à segurança de operações aeronáuticas. Atualmente, as ferramentas utilizadas são os NWP, tais como o ICON-LAM. Como estes modelos são robustos, eles necessitam de uma infraestrutura computacional substancial e um tempo mínimo de 40 minutos de tempo de execução para a área mencionada neste trabalho. O intuito, foi comparar arquiteturas de redes neurais profundas, tais como MLP, CNN, LSTM e BiLSTM, visando entender quais delas apresentariam um desempenho competitivo aos NWP. Os experimentos demonstraram que as redes recorrentes, especialmente a BiLSTM, apresentaram desempenho significativamente superior nas métricas de erro (MAE, MAPE e RMSE), destacando sua capacidade de capturar dependências temporais de longo prazo e padrões sazonais. A configuração com maior profundidade temporal (48 horas de dados de entrada e 12 horas de previsão) apresentou, dentre todos os testes, os melhores resultados, com um tempo de execução de 10 minutos.

Referências

- [Géron 2019] Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Sebastopol, CA, 2nd edition.
- [Iriza-Burcă et al. 2024] Iriza-Burcă, D., Andrei, D., Dragoş, C., Mihăilă, A., Rugină, C., Dima, M., Pripón, C., and Mitrea, M. (2024). Comparison of cosmo and icon-lam high-resolution numerical forecast for romanian territory: Case studies and evaluation. *Romanian Journal of Physics*.
- [Xu 2024] Xu, M. (2024). Comparative analysis of machine learning models for weather forecasting: A heathrow case study. *Dean & Francis*. ISSN 2959-6157.
- [Zhang et al. 2025] Zhang, H., Liu, Y., Zhang, C., and Li, N. (2025). Machine learning methods for weather forecasting: A survey. *Atmosphere*, 16(1):82. Open Access under CC BY 4.0 License.