# Previsão do Risco de Nevoeiro com Aprendizado Profundo: Estudo de Caso no Cais do Porto de Porto Alegre

**Lucas dos Santos Azevedo**[1], **Rafael Manica**[2], **Eduardo Puhl**[2]

[1]Instituto de Pesquisas Hidráulicas – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.029 – 91501-970 – Porto Alegre – RS – Brazil Porto Alegre – RS – Brazil

`lucas.azevedo@inf.ufrgs.br, manica@iph.ufrgs.br, eduardo.puhl@ufrgs.br`

***Abstract.*** *This study proposes and validates a two-step pipeline for fog alert forecasting at the Porto Alegre harbor: (i) A multivariate LSTM (Network A) forecasts meteorological variables up to 100 h ahead, based on 318,000 records (REDEMET [SBPA/SBPK, 2003–2024] and local data); (ii) A binary LSTM (Network B) estimates, for each lead time, the probability of fog occurrence using the predictors generated by Network A. In 9,000 simulations, Network B achieved AUC = 0.936, recall = 90%, and accuracy = 82.52% (7,131 true negatives, 477 true positives, 1,471 false positives, 41 false negatives), with false positives concentrated between 2–5 h under high relative humidity and weak winds in a class-imbalanced setting.*

***Resumo.*** *Este trabalho propõe e valida um pipeline em duas etapas para previsão de alerta de nevoeiro no cais de Porto Alegre. (i) Uma LSTM multivariada (Rede A) prevê, até 100 h à frente, variáveis meteorológicas com base em 318.000 dados (REDEMET [SBPA/SBPK, 2003–2024] e locais); (ii) Uma LSTM binária (Rede B) estima, para cada lead, a probabilidade de nevoeiro usando os previsores gerados pela Rede A. Em 9.000 simulações, a Rede B alcançou AUC = 0,936, recall = 90% e acurácia = 82,52% (7.131 verdadeiros negativos, 477 verdadeiros positivos, 1.471 falsos positivos, 41 falsos negativos), com falsos positivos concentrados entre 2–5 h sob umidade relativa elevada e vento fraco em cenário desbalanceado.*

## 1. Introdução

O impacto do nevoeiro sobre as operações hidroviárias e portuárias em Porto Alegre é significativo. A presença desse fenômeno constitui um dos desafios meteorológicos mais relevantes para o transporte hidroviário e portuário, modal de grande importância econômica no sul do Brasil. O nevoeiro, caracterizado pela presença de gotículas de água em suspensão que reduzem a visibilidade para menos de 1 km, apresenta comportamento fortemente dependente de processos microfísicos, termodinâmicos e da interação entre clima, relevo e dinâmica atmosférica [Gultepe et al. 2007, Silva 2012].

Porto Alegre e sua hidrovia localizam-se em uma região de transição climática, fortemente influenciada por sistemas subtropicais, pela presença de grandes corpos d'água (como o Rio Guaíba e a Lagoa dos Patos) e pelo regime de frentes frias e anticiclones subtropicais [Lima 1982, Fedorova et al. 2019, Michelon 2023]. Estudos anteriores [Willett 1928, Piva 1999, Almeida 2008] mostraram que o nevoeiro em Porto Alegre é predominantemente de dois tipos: (i) *radiação*, originado por resfriamento noturno sob céu claro e ventos fracos; e (ii) *advecção*, relacionado ao transporte de ar úmido sobre superfícies frias, frequentemente o Guaíba.

Em particular, o nevoeiro (visibilidade $<$ 4.000 m; ventos $<$ 5 m/s) tem sido associado a encalhes na hidrovia local [Michelon 2023, Scotta et al. 2025]. Motivado por esse contexto aplicado, este estudo investiga o uso de aprendizado profundo para previsão do risco de nevoeiro no cais do porto de Porto Alegre, com vistas a apoiar a gestão de riscos e a eficiência operacional.

## 2. Metodologia

Este estudo propõe um *pipeline* em duas etapas para previsão intradiária e alerta de nevoeiro no cais de Porto Alegre utilizando Redes Neurais Recorrentes do tipo *Long Short-Term Memory* (LSTM). Foram realizadas 9.000 simulações entre maio/2023 e abril/2024. Para avaliar os resultados, foram utilizadas métricas como acurácia média, *recall* médio, área sob a curva ROC (AUC), acurácia global e matriz de confusão (verdadeiros positivos e negativos, falsos positivos e negativos).

### 2.1. Rede A — LSTM multivariada para previsão dos previsores

A Rede A foi concebida para gerar, a cada ciclo, as 100 horas futuras das variáveis meteorológicas de interesse, contemplando direção e velocidade do vento (`dir_vel` e `vnt_vel`), pressão atmosférica ao nível da estação (`qnh`), umidade relativa (`ur`), ponto de orvalho (`po`) e temperatura do ar (`tt`). Para isso, utilizou-se um conjunto composto por 306.250 registros da API-REDEMET, referentes às estações SBPA e SBPK no período de 2003 a 2024, complementados por 12.589 medições *in loco* no cais do porto, coletadas entre 2021 e 2024. Os dados foram organizados em janelas de 100 *lags*, de modo que as entradas assumiram a forma de tensores $(\text{amostras}, 100, \text{features})$ com séries previamente normalizadas.

A arquitetura da rede foi estruturada com uma camada inicial LSTM de 256 unidades associada a um *dropout* de 0,2, seguida por um bloco de sete camadas LSTM empilhadas com 256 unidades cada, cada qual acompanhada de *dropout* de 0,1, resultando em uma saída densa linear que produz, para cada variável, vetores de 100 *leads*. O treinamento foi realizado com o otimizador Adam e a função de perda baseada no erro quadrático médio (MSE), com até 500 épocas. Para prevenir sobreajuste e assegurar estabilidade, foram empregados *callbacks* como *EarlyStopping*, *ReduceLROnPlateau* e *ModelCheckpoint*.

### 2.2. Rede B — Classificador de nevoeiro

A Rede B foi projetada para estimar, a cada horizonte $h \in \{1, \ldots, 100\}$, a probabilidade de ocorrência de nevoeiro, definida como

$$p_{t+h} = \Pr(R = 1 \mid x_{t+h}),$$

em que o rótulo binário $R \in \{0, 1\}$ é estabelecido segundo a seguinte regra operacional:

$$R = \begin{cases} 1, & \text{se } \big(ur \geq 95\% \wedge \Delta T \leq 2\,^{\circ}\text{C} \wedge vnt\_vel \leq 5\,\text{m/s}\big) \;\vee\; (Vis \leq 4000\,\text{m}) \;\vee\; (S = 1) \\ 0, & \text{caso contrário} \end{cases}$$

com $\Delta T = tt - po$ representando a diferença entre a temperatura do ar e o ponto de orvalho previstos pela Rede A. A entrada da Rede B foi composta por vetores de preditores previstos e derivados ($tt_h, po_h, ur_h, vnt\_vel_h$), definidos por busca em grade de forma a maximizar a AUC.

A arquitetura final foi implementada com duas camadas LSTM de 128 células, sendo a primeira configurada para retornar sequências, seguidas de um *dropout* de 0,5 entre as camadas, culminando em uma camada densa com ativação sigmoide e um único neurônio para estimar a probabilidade de ocorrência. O treinamento utilizou a perda binária (*binary cross-entropy*), o otimizador Adam e ajuste de pesos de classe para lidar com desbalanceamento, em 60 épocas por *lead*. A avaliação baseou-se na métrica principal AUC-ROC, complementada por acurácia, matriz de confusão e *recall*. Quando necessário, o limiar $\tau$ foi calibrado por critérios como Youden-J ou custo-operacional, em vez de fixado rigidamente em 0,5.

### 2.3. Baselines, interpretabilidade e pipeline

Como referência, foram treinados modelos baseline (GLM/logística, MLP e *naive* de classe majoritária). A GLM serve de comparador interpretável (sinais e intervalos de confiança dos coeficientes), mas a decisão operacional adota a Rede B por desempenho agregado superior (AUC e *recall* por lead). Em produção, o *pipeline* executa: (i) a Rede A para prever 100 *leads* de `tt`, `ur`, `qnh`, `vnt_vel`, `vnt_dir` e `po`; (ii) o cálculo de $\Delta T$ e demais derivados; (iii) a Rede B para estimar $p_{t+h}$ e classificar $R_{t+h}$ conforme o limiar $\tau$; e (iv) a publicação das probabilidades/classes por *lead* e a consolidação de janelas de alto risco para apoio à decisão.

## 3. Resultados

Os resultados da Rede A mostram o padrão físico esperado de degradação com o horizonte: RMSE/MAE aumentam da lead 1 para a lead 100, mais claramente para a temperatura (p.ex. 2,56/1,97 °C para 4,84/3,74 °C) e o ponto de orvalho (de 2,26/1,73 °C para 4,43/3,50 °C). Já para a pressão, a degradação é mais contida (4,02/3,06 e 5,92/4,64 hPa), refletindo estabilidade barométrica; a umidade relativa quase dobra de valor (10,33/8,04% para 17,20/13,45%), bem como a velocidade do vento (de 2,75/2,21 para 4,00/2,88 nós), indicando persistência temporal. A direção do vento é a variável mais desafiadora, com erros altos em ambos os horizontes (124,81/92,52 $\rightarrow$ 135,47/109,62 graus), devido ao caráter angular e à variabilidade acentuada (mod 360°).

Em 9.000 simulações na Rede B, o classificador LSTM atingiu acurácia média de 82,52%, *recall* médio de 90% e AUC média de 0,936, com 7.131 verdadeiros negativos, 477 verdadeiros positivos, 1.471 falsos positivos e 41 falsos negativos. Na análise mensal (mai/2023–abr/2024), a acurácia global variou de 69,74% a 99,05%, o *recall* de 0,8889 a 1,0 e a AUC de 0,7988 a 0,9845, com acerto de 100% para nevoeiros densos, moderados e leves na maioria dos meses. Os falsos positivos concentraram-se na madrugada (2h–5h), sob $ur \geq 95\%$ e $vnt\_vel \leq 5$ m/s, sem queda marcada de visibilidade; em visibilidades intermediárias (2–5 km), as taxas de erro oscilaram entre 2,381% e 60%, e em valores superiores a 8 km ocorreram erros pontuais (p.ex. 50% em out/2023), enquanto o *Fog-Alert* mostrou utilidade acima de 4 km.

A baixa incidência de falsos negativos (geralmente $\leq$ 8/mês) indica alta sensibilidade a eventos críticos (Vis $\leq 1000$ m), embora meses com poucos casos (p.ex., fev/2024:

3 verdadeiros positivos; dez/2023: 1 verdadeiro positivo) tragam maior incerteza. Em síntese, o modelo é robusto e operacionalmente confiável, com prioridade de melhorias na calibração de limiares (por sazonalidade/horário) para reduzir falsos positivos sem sacrificar a sensibilidade.

## 4. Conclusões

Conclui-se que a combinação entre uma LSTM multivariada ajustada localmente e um classificador binário dedicado possibilita a emissão de alertas de nevoeiro confiáveis para horizontes de curto e médio prazo. O modelo proposto apresenta desempenho robusto, alta sensibilidade a eventos críticos e aplicabilidade prática, oferecendo suporte às decisões operacionais no ambiente portuário de Porto Alegre.

## Referências

Almeida, L. A. F. (2008). Análise de frequência da ocorrência de nevoeiro no aeroporto internacional salgado filho – porto alegre/rs. Relatório de pesquisa, Instituto de Controle do Espaço Aéreo (ICEA).

Fedorova, N., Levit, V., and Afonso, J. M. S. (2019). Estudo de baixa visibilidade no aeroporto de porto alegre: processos sinóticos e termodinâmicos. *Revista Ibero-Americana de Ciências Ambientais*, 10(6):131–145.

Gultepe, I., Tardif, R., Michaelides, S. C., Cermak, J., Bott, A., Bendix, J., Müller, M. D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., and Cober, S. G. (2007). Fog research: A review of past achievements and future perspectives. *Bulletin of the American Meteorological Society*, 88(8):1111–1126.

Lima, J. S. (1982). Método de previsão de ocorrência e de duração de nevoeiro em porto alegre: método objetivo. In *Congresso Brasileiro de Meteorologia (CBMET)*, Pelotas, Brasil.

Michelon, C. R. (2023). Estudo de casos de acidentes de navegação na hidrovia do rio guaíba. Dissertação de mestrado, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre.

Piva, E. D. (1999). Estudo de nevoeiros de radiação em porto alegre. *Revista Brasileira de Meteorologia*, 14(1):33–45.

Scotta, F., Michelon, C. R., Puhl, E., Manica, R., Jr., E. E. T., Fick, C., Lenhard, C., Silva, T., and Nunes, J. C. R. (2025). Environmental causes of navigation accidents in the guaíba river waterway. *Brazilian Journal of Water Resources*. Aceito para publicação.

Silva, A. O. (2012). Análise e previsão de visibilidade baixa no aeroporto de maceió. Dissertação de mestrado, Universidade Federal de Alagoas, Instituto de Ciências Atmosféricas, Maceió. Orient.: N. Fedorova; Coorient.: V. Levit.

Willett, H. C. (1928). Fog and haze, their causes, distribution, and forecasting. *Monthly Weather Review*, 56(11):435–468.