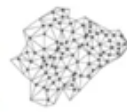




Modelagem Computacional no Sertão Mineiro



PPGMCS
Programa de Pós-Graduação em
Modelagem Computacional e Sistemas

PREVISÃO DE IRRADIÂNCIA SOLAR GLOBAL HORIZONTAL VIA REDES NEURAS PROFUNDAS: UMA ABORDAGEM COMPARATIVA EM MONTES CLAROS-MG

Matheus Rodrigues Cruz Martins¹ - matheusrcmartins@gmail.com

¹Universidade Estadual de Montes Claros, UNIMONTES - Montes Claros, MG, Brazil

Abstract. A previsão da irradiância solar é essencial para o planejamento de energia fotovoltaica e impulsiona processos atmosféricos globais. O Brasil, situado predominantemente na região intertropical, detém elevado potencial solar; contudo, a variabilidade climática impõe desafios à previsão precisa de sua disponibilidade. Este estudo propõe e compara três modelos de redes neurais profundas – Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM) e Gated Recurrent Unit (GRU), com o modelo de benchmark Smart Persistence, para a previsão da Irradiância Horizontal Global (GHI), no instante $t+1$ em Montes Claros-MG. Utilizaram-se dados históricos de satélite do National Solar Radiation Database (NSRDB), com aplicação de pré-processamento, engenharia de atributos e normalização. A avaliação baseou-se nas métricas RMSE, MAE e Coeficiente de Correlação de Pearson (r). Os resultados evidenciaram a superioridade da arquitetura GRU, que atingiu RMSE de 29,9 e r de 99,6%, demonstrando robustez em cenários de alta variabilidade e superando o modelo baseline. A abordagem comparativa elucidou vantagens e limitações de cada arquitetura, contribuindo para o desenvolvimento de ferramentas computacionais de apoio ao planejamento e à operação de sistemas fotovoltaicos.

Keywords: Previsão Solar, Redes Neurais Profundas, LSTM, GRU, Irradiância Horizontal Global

1. INTRODUÇÃO

A energia solar fotovoltaica é uma das alternativas mais promissoras na transição global para fontes de energia limpas e sustentáveis (Kumar et al., 2020). O Brasil, devido à sua localização intertropical, possui elevado potencial de incidência solar ao longo de todo o ano (Pereira et al., 2017), com destaque para Minas Gerais, cujo território apresenta ampla viabilidade de aproveitamento, conforme o Atlas Solarimétrico do estado (CEMIG, 2012). Contudo, a geração solar é intermitente e sensível às condições climáticas, o que desafia a estabilidade e a previsibilidade da produção (Pereira et al., 2017). Assim, prever com precisão

os níveis de Irradiância Solar Global Horizontal (GHI) é fundamental para o planejamento da produção, a otimização do despacho de carga e o gerenciamento de sistemas de armazenamento (Kumari and Toshniwal, 2021).

A literatura apresenta diversos métodos para a previsão da irradiância solar, classificados em modelos físicos, estatísticos e híbridos (Voyant et al., 2017). Recentemente, técnicas de aprendizado de máquina, especialmente Redes Neurais Artificiais (RNAs), têm se destacado por modelar as complexas relações não lineares dos dados meteorológicos (Mellit et al., 2022). Entre as arquiteturas de aprendizado profundo, as Redes Neurais Recorrentes (RNNs) e suas variantes *Long Short-Term Memory* (LSTM) e *Gated Recurrent Unit* (GRU) destacam-se por capturar dependências temporais em dados sequenciais, superando modelos tradicionais em previsão de séries temporais (Géron, 2017). Vale notar, contudo, que arquiteturas recorrentes mais complexas, como a LSTM, estão sujeitas a sobreajuste em bases de dados de menor porte, podendo perder desempenho até mesmo frente a modelos ingênuos (Nielsen, 2019).

Apesar dos avanços, persiste uma lacuna na avaliação comparativa sistemática dessas arquiteturas para localidades específicas frente a um algoritmo “ingênuo” de referência. Este trabalho busca preencher tal lacuna ao desenvolver, treinar e comparar três arquiteturas de redes neurais profundas (MLP, LSTM e GRU) para a previsão de GHI de curto prazo, com passo $t+1$, na cidade de Montes Claros-MG, tomando o Smart Persistence – comumente utilizado como referência para modelos de maior custo computacional – como *benchmark* na identificação da arquitetura mais eficiente e robusta para as características climáticas da região (Huertas Tato and Centeno Brito, 2019).

2. METODOLOGIA

A metodologia abrange a aquisição e o pré-processamento dos dados, a engenharia de atributos, a definição das arquiteturas, o treinamento e a avaliação dos modelos. O desenvolvimento foi realizado em Python (v. 3.13), com as bibliotecas TensorFlow, Scikit-learn e Pandas, no ambiente Google Colab.

2.1 Base de Dados e Pré-processamento

Os dados meteorológicos e de irradiância solar de Montes Claros-MG, referentes a 2023, foram obtidos do *National Solar Radiation Database* (NSRDB), em resolução temporal de 30 minutos, que fornece medições de alta resolução espacial para GHI, DHI e DNI, com incerteza de aproximadamente 5% (Sengupta et al., 2018). No pré-processamento, verificou-se a existência de duplicatas e valores ausentes, descartaram-se linhas de metadados e realizou-se a conversão dos tipos de dados. Em seguida, aplicou-se engenharia de atributos com transformações cíclicas via funções seno e cosseno, para capturar padrões de sazonalidade (diários, mensais e anuais) e permitir ao modelo compreender a natureza periódica do tempo (Tancik et al., 2020). Por fim, todos os atributos de entrada foram normalizados no intervalo $[0, 1]$ pela função *MinMaxScaler* da Scikit-learn, o que melhora a convergência e o desempenho dos algoritmos (Pedregosa et al., 2011), conforme:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

2.2 Arquitetura *Smart Persistence*

O modelo *Smart Persistence* desloca continuamente as medições de GHI para instantes futuros, aplicando a média do índice de céu limpo (*GHI Clear Sky*) ao longo de uma janela temporal, sendo amplamente utilizado como *benchmark* (Kumler et al., 2019). Apesar de simples, decompõe a previsão do GHI por meio do cálculo de céu limpo, o que o diferencia de modelos ingênuos e o torna mais robusto para análises comparativas (Kumler et al., 2019). O modelo é descrito pela Equação 2, em que k_t é o índice de clareza no tempo t , GHI_{medido} é a irradiância horizontal global medida e GHI_{cs} , a irradiância de céu limpo.

$$k_t = \frac{GHI_{medido}}{GHI_{cs}} \quad (2)$$

2.3 Arquiteturas de Redes Neurais

Três arquiteturas de aprendizado profundo foram implementadas e comparadas:

1. **MLP (*Multilayer Perceptron*)**: rede *feedforward* totalmente conectada, capaz de modelar relações não lineares complexas, utilizada como *baseline*. Configuração: entrada de 144 neurônios, camada oculta de 46 e saída de 1.
2. **LSTM (*Long Short-Term Memory*)**: arquitetura recorrente com células de memória e portas de esquecimento, entrada e saída, capaz de capturar dependências de longo prazo e adequada à modelagem de dados sazonais (Olah, 2015). Configuração: três camadas, com 64, 32 e 1 neurônio de saída.
3. **GRU (*Gated Recurrent Unit*)**: variante simplificada da LSTM que combina as portas de esquecimento e entrada em uma única “porta de atualização”, resultando em maior eficiência computacional com capacidade similar à LSTM (Olah, 2015). Configuração: três camadas, com 64, 32 e 1 neurônio de saída.

A topologia das redes e seus hiperparâmetros foram definidos buscando o melhor equilíbrio entre erro de validação e custo computacional.

2.4 Métricas de Avaliação

O desempenho foi avaliado por quatro métricas padrão para problemas de regressão. O MAE mede a magnitude média dos erros absolutos, sendo menos sensível a *outliers*; o MBE avalia o viés sistemático do modelo, indicando ausência de viés quando próximo de zero; o RMSE penaliza mais fortemente erros maiores, sendo indicado quando desvios grandes são indesejáveis; e o **coeficiente de correlação de Pearson (r)** indica a intensidade e a direção da relação linear entre valores observados e preditos, variando de -1 a 1 .

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad MBE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i) \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (4)$$

Onde y_i representa os valores observados, $\hat{y}_i$, os valores preditos, e n , o número de observações. Neste artigo, o coeficiente r será avaliado de forma gráfica.

3. RESULTADOS E DISCUSSÃO

Comparam-se as três arquiteturas de aprendizado profundo (MLP, LSTM e GRU) ao *baseline* Smart Persistence. A Tabela 1 apresenta as métricas obtidas após treinamento e validação das arquiteturas desenvolvidas.

O GRU apresentou o melhor desempenho geral, com RMSE de 29,98 W/m² e *Skill* de 52,37%. O LSTM, apesar de sua maior complexidade computacional, obteve resultados inferiores (RMSE de 101,05 W/m²), pior até que o Smart Persistence, o que sugere sobreajuste (*overfitting*) durante o treinamento e consequente perda de generalização – comportamento associado ao maior número de parâmetros da LSTM frente ao volume de dados disponível, em contraste com a estrutura mais enxuta da GRU. O MLP, mais simples e insensível às dependências temporais, teve desempenho intermediário (RMSE de 59,77 W/m²), superando apenas o LSTM e reforçando a importância de arquiteturas especializadas em séries temporais. Assim, para a previsão de GHI com horizonte de 30 minutos à frente nesta localidade, o GRU mostrou-se a alternativa mais adequada em termos de precisão e eficiência computacional.

Table 1: Métricas médias de desempenho (RMSE e MAE em W/m²; *Skill*) em relação ao Smart Persistence.

Métrica	Smart Persist.	MLP	GRU	LSTM
RMSE (W/m ²)	63,95	59,77	29,98	101,05
MAE (W/m ²)	54,38	48,61	22,76	74,26
MBE (W/m ²)	-0,72	5,19	2,61	15,85
<i>Skill</i> (%)	—	4,91	52,37	-68,95

A Figura 1 apresenta os diagramas de dispersão entre valores observados (eixo vertical) e preditos (eixo horizontal), com o coeficiente r indicado em cada subgráfico e a linha tracejada representando a predição perfeita; a maior concentração dos pontos do GRU em torno dessa linha evidencia predições mais precisas e consistentes. Já a Figura 2 exibe a comparação temporal entre os valores previstos e os observados, permitindo avaliar qualitativamente a aderência dos modelos à dinâmica real da GHI e identificar padrões de erro sistemático e a capacidade de capturar variações ao longo do dia.

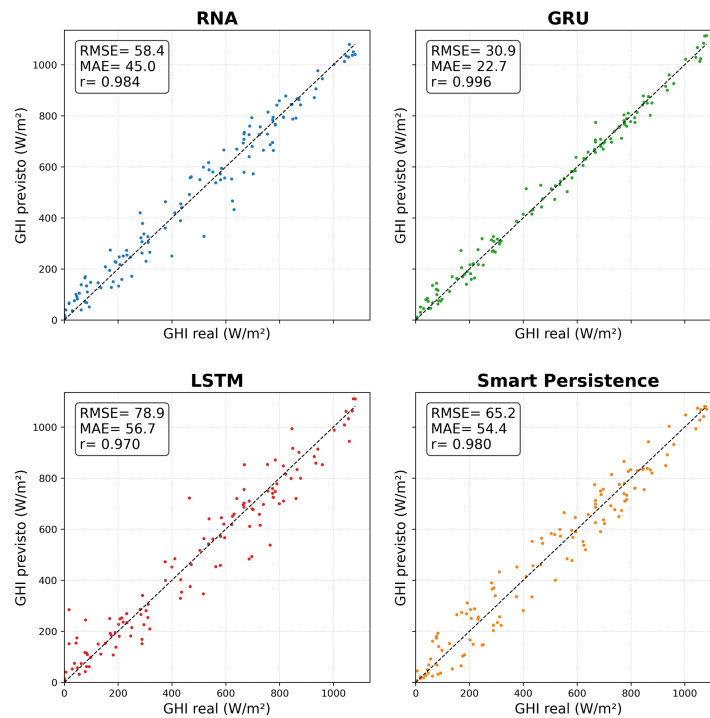


Figure 1: Comparação dos modelos de previsão: diagramas de dispersão MLP, GRU e LSTM versus Smart Persistence.

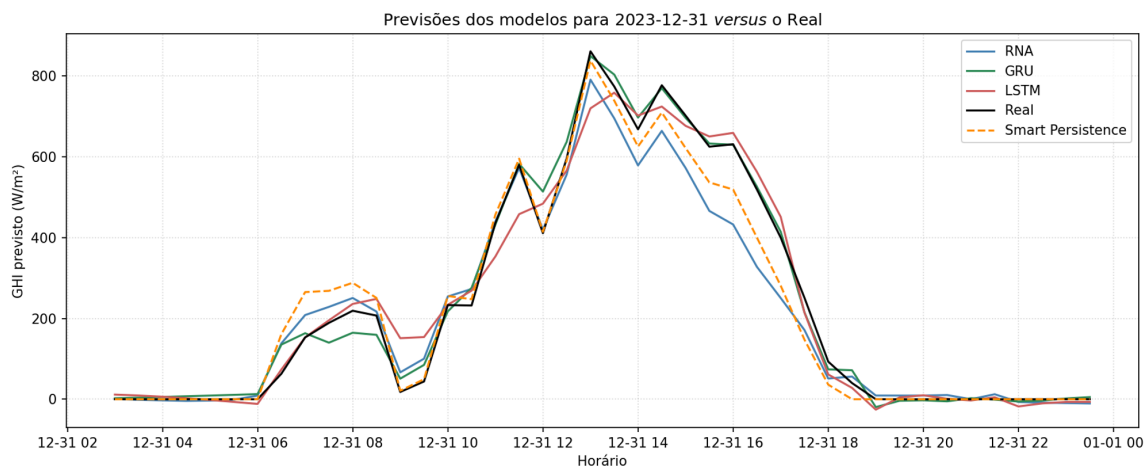


Figure 2: Comparação das previsões obtidas: valores observados versus valores preditos.

4. CONCLUSÃO

Foram implementadas e comparadas três arquiteturas de redes neurais profundas para previsão da Irradiância Solar Global Horizontal (GHI) em Montes Claros-MG: MLP, LSTM, GRU, tomando o *Smart Persistence* como *benchmark*. O GRU apresentou desempenho superior, com RMSE de aproximadamente 30,0 W/m², baixo viés (MBE < 3,0 W/m²) e ganho de 52% sobre o *Smart Persistence*, evidenciando sua capacidade de modelar dependências temporais com baixo custo computacional e de capturar a alta variabilidade típica de dados solarimétricos,

posicionando-se como a solução mais adequada entre os modelos avaliados em termos de precisão, eficiência computacional e capacidade de representar a dinâmica do conjunto de dados.

Trabalhos futuros poderão aplicar o modelo a diferentes localidades e períodos sazonais, além de incorporar variáveis meteorológicas adicionais para aprimorar a acurácia das previsões.

Agradecimentos

O autor agradece ao Programa de Pós Graduação em Modelagem Computacional e Sistemas e à FAPEMIG.

REFERENCES

- CEMIG (2012). Atlas solarimétrico de minas gerais. Technical report, Companhia Energética de Minas Gerais, Belo Horizonte.
- Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly Media, Sebastopol, CA.
- Huertas Tato, J. and Centeno Brito, M. (2019). Using smart persistence and random forests to predict photovoltaic energy production. *Energies*, 12(1).
- Kumar, N. M., Chopra, S. S., Chand, A. A., Elavarasan, R. M., and Shafiullah, G. M. (2020). Hybrid renewable energy microgrid for a residential community: A techno-economic and environmental perspective in the context of the sdg7. *Sustainability*, 12(10):3944.
- Kumari, P. and Toshniwal, D. (2021). Deep learning models for solar irradiance forecasting: A comprehensive review. *Journal of Cleaner Production*, 318:128566.
- Kumler, A., Xie, Y., and Zhang, Y. (2019). A physics-based smart persistence model for intra-hour forecasting of solar radiation (pspi) using ghi measurements and a cloud retrieval technique. *Solar Energy*, 177:494–500.
- Mellit, A., Pavan, A. M., Ogliari, E., Leva, S., and Lughi, V. (2022). Advanced methods for photovoltaic output power forecasting: A review. *Applied Sciences*, 12(2):1–27.
- Nielsen, A. (2019). *Practical Time Series Analysis: Prediction with Statistics and Machine Learning*. O'Reilly Media.
- Olah, C. (2015). Understanding LSTM networks. <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. Acesso em: 21 jun. 2026.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830.
- Pereira, E. B., Martins, F. R., Gonçalves, A. R., Costa, R. S., Lima, F. J. L., Rüther, R., Abreu, S. L. d., Tiepolo, G. M., Pereira, S. V., and Souza, J. G. d. (2017). *Atlas Brasileiro de Energia Solar*. Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2 edition.
- Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., and Shelby, J. (2018). The national solar radiation data base (nsrdb). *Renewable and Sustainable Energy Reviews*, 89:51–60.
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., and Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. In *Advances in Neural Information Processing Systems*, volume 33, pages 7537–7547.
- Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., and Foulloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569–582.