Short-Term Wind Speed Forecasting Using ERA5 Reanalysis: A Comparison of SARIMA and LSTM Models at the Santo Agostinho Wind Farm

Marina Paiva Browning[1], Louise da Fonseca Aguiar[1], Lívia Sancho[1,3], Priscila Esposte Coutinho[12], Marcio Cataldi[1].

1- Federal Fluminense University (UFF), Laboratory of Monitoring and Modeling of Climate Systems (LAMMOC), Brazil.

2 - Postgraduate Program in Biosystems Engineering, Federal Fluminense University, 156 Passo da Pátria St., 24210-240 - Niterói, Brazil

3 - COPPE, Universidade Federal do Rio de Janeiro, 2030 Horácio Macedo Av., 21941-594 - Rio de Janeiro, Brazil

marinabrowning@id.uff.br, louisedaguiar@gmail.com, liviasancho@gmail.com, priscila.esposte@gmail.com, mcataldi@id.uff.br.

Abstract: Renewable energy is a central component of Brazil's electricity matrix, with wind power standing out as one of the fastest-growing technologies and reaching more than 33 GW of installed capacity in 2025. Despite this rapid expansion, the temporal variability of renewable energy generation can pose operational challenges for the power system, including grid instability and potential stress on transmission infrastructure. As wind capacity continues to expand, accurately understanding both the availability and short-term predictability of wind resources becomes essential for reliable grid operation and effective energy planning. This study investigates the short-term predictability of wind speed of the Santo Agostinho Wind Complex located in Lajes, Rio Grande do Norte. ERA5 Reanalysis data from the 1990–2025 baseline period were used to characterize the regional wind resource, using wind speed at 100 m height. The spatial selection of the data was based on a local anemometric tower, improving the representativeness of the modeled wind speed. To reduce the influence of extreme values and better represent typical wind conditions, the daily median wind speed was used for the historical series. The dataset was split into training and evaluation subsets, with 80% used for training. For SARIMA, 20% was used for testing, while for LSTM, 10% was used for validation and 10% for testing. The forecasts were then used over a 30-day horizon. For comparison, metrics such as MAE, RMSE and R2 were calculated. The results enable a comparison between two different approaches, identifying which model provides more reliable forecasts for wind speed variability, with implications for power system reliability, energy planning and commercialization.

Keywords: Reanalysis data, Long Short-Term Memory (LSTM), Short-term forecasting, Deep Learning, SARIMA