26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Data-driven rolling forecasting for wind and solar generation: Evaluating statistical models across a range of training data lengths

Mahdi Alikhani

STEM Academic Unit, University of South Australia, Adelaide, Australia
email: alimy021@mymail.unisa.edu.au

Abstract: Accurate short-term forecasting of wind and solar farm power generation is essential for reliable grid integration, operational planning, and market participation. Classical univariate statistical models remain critical baselines due to their transparency, interpretability, and computational efficiency. By leveraging inherent temporal patterns in wind and solar output—diurnal, daily, and weekly cycles—these models provide forecasts that are both operationally reliable and scientifically interpretable.

This study systematically evaluates linear regression, ETS (including Holt-Winters), ARIMA, ARIMA with Fourier terms (using sine/cosine regressors), and Trigonometric Box-Cox ARMA Trend Seasonality (TBATS) under a rolling forecasting strategy, with particular emphasis on how forecast accuracy depends on the length of historical data available for training. Linear regression, ETS, and non-seasonal ARIMA are effective when only a few days of data (1–7 days) are available, capturing short-term persistence and local fluctuations that are common in wind speed and solar irradiance due to weather variability. Holt-Winters and SARIMA extend forecasting capability to multi-week horizons by modelling level, trend, and intra-week seasonal effects, which align with operational cycles such as weekday–weekend patterns or weekly weather regimes. For longer datasets spanning months to a year, ARIMA with Fourier terms and TBATS provide robust handling of multiple or non-integer seasonalities. Fourier-augmented ARIMA introduces deterministic seasonal regressors, reducing the need for high-order seasonal AR or MA terms, while TBATS integrates Box–Cox transformation, trigonometric seasonality, ARMA error structure, and trend components, making it especially suited to long or complex seasonal structures.

Forecast performance is assessed through one step-ahead prediction error calculation across both expanding and sliding rolling windows, residual diagnostics including Ljung–Box tests, and horizon-wise evaluation. Comparisons with common machine learning methods (e.g., random forests and LSTMs) indicate that for univariate short-term horizons, statistical models often provide competitive accuracy while maintaining interpretability and reproducibility. The results show when simple approaches suffice and when advanced seasonal models are required, offering a practical framework for selecting statistical models for short-term wind and solar forecasting across different data lengths and seasonal complexities, with guidance for both research and operational use.

Keywords: *Short-term forecasting, statistical models, historical data length, wind power generation, solar power generation, seasonality modelling, expanding rolling forecasting, sliding rolling forecasting, forecast accuracy, machine learning methods*