A Methodology to Improve the Predictability of Offshore Wind Energy Generation: Evidence from Great Britain

Kevin F Forbes[1]

1. Energy and Environmental Data Science, Malahide, CO DUBLIN, Ireland

In a power grid, a generation scheduling error measures the difference between the metered generation by an electricity producer or group of producers and the corresponding scheduled generation. In the case of wind energy, large positive scheduling errors can lead to the economic waste that occurs when a system operator curtails wind energy generation. Negative errors that are large in absolute value can reduce the power grid's resiliency to exogenous shocks. Given these consequences, it is troubling to note that wind energy scheduling errors in the British offshore, a region slated for significant expansion, are significantly larger than the errors for combined cycle gas turbines (CCGT) or nuclear. Specifically, the weighed-mean-absolute-percentage-error (WMAPE) in the scheduling errors for offshore wind in 2021 was 7.86%, while the WMAPEs for CCGT and nuclear were 1.70% and 4.13%, respectively. The difference between offshore wind energy and the CCGT WMAPEs suggests that the planned expansion of Great Britain's offshore wind resources coupled with reduced reliance on CCGT may present significant operational challenges in matching electricity supply with demand. It may be worth noting that National Grid ESO, the system operator in Great Britain, reports that the error in its day-ahead wind energy forecasts is only about 4%. Unfortunately, this statistic is open to question because it is calculated by dividing the mean absolute error by the installed capacity. According to the published literature, this can give rise to false confidence.

With the above facts in mind, the analysis of this paper aims at developing an algorithm to improve the predictability of offshore wind energy generation. The analysis proceeds by first recognizing that the generation data are significantly autoregressive and highly volatile at times. Based on these properties, a machine learning approach known as ARCH/ARMAX (Autoregressive Conditional Heteroskedasticity/ Autoregressive Moving Average with Exogenous Inputs) model is formulated. The model's exogenous inputs include simulated meteorological variables (e.g., simulated air pressure, wind speeds, and air density). Another key exogenous input is the final physical notification (FPN) that wind farm operators provide to the system operator one hour prior to real-time. The ARCH effects are modeled using seven variables, while the autoregressive nature of the data is modeled using 23 ARMA (autoregressive-moving average) terms. These ARMA terms model the relationship between the current and previous outcomes and thus play a critical role in improving predictability. The model is estimated using 30-minute data from Jan 1, 2019, through Dec 31, 2021. The preliminary out-of-sample predictions have a WMAPE of about three percent. These results indicate that the modeling approach may be useful in reducing the operational uncertainty associated with offshore wind energy generation.