Data-Driven Modeling of Wind Farm Power and Revenue Generation

Ivan Karp
Department of Statistics
University of California, Los Angeles
Los Angeles, United States
ivanakarp@g.ucla.edu

Chris Qin
School of Engineering and Computer Science
Washington State University, Vancouver
Vancouver, United States
chris.qin@wsu.edu

***Abstract*—Accurate power and revenue prediction in wind farms is essential for optimizing performance and economic planning, particularly under variable and turbulent wind conditions. This study presents a data-driven model that refines traditional power curves by incorporating statistical corrections for turbulence and wake effects, enabling robust predictions of wind power generation and revenue. Using simulations in OpenFAST and turbulent inflow generated by TurbSim, the model is validated through a case study of the Pyron Wind Farm in Texas. The analysis leverages a comprehensive four-year dataset, including hourly wind speed data from ERA atmospheric data and electricity price data from the ERCOT West zone, capturing real-time generation and market conditions. Model accuracy comparisons show that the smoothed turbulence-corrected curve reduces mean absolute error (MAE) by 24 percent compared to the steady flow curve, with total predicted power production within 4 percent and predicted revenue within 1 percent of actual observations. Additionally, the modeled correlation between hourly wind speed and electricity price reveals that revenue peaks often occur at lower wind speeds, countering traditional approaches that rely on mean prices and tend to overestimate revenue.**

***Index Terms*—Data-driven method, simulation, turbulence-corrected power curves, wind energy system modeling, wind farm**

I. Introduction

As renewable energy sources, particularly wind energy, gain traction worldwide, precise data-driven modeling of wind power generation is increasingly essential. Power curves are commonly used in this context to estimate the average power output of a turbine as a function of a given wind speed. These curves are typically based on steady wind conditions and are often provided by turbine manufacturers. However, they may not accurately predict performance in turbulent wind environments, where variations in wind speed can substantially impact turbine power output.

Extending these predictions from individual turbines to entire wind farms adds further complexity. In a wind farm, wake losses arise as downstream turbines experience reduced wind speeds due to energy extracted by upstream turbines, leading to a cumulative reduction in overall power generation [1]. Traditional approaches often employ computationally demanding simulations with tools like OpenFAST [2] and TurbSim [3] to predict these effects, yet these methods can be resource-intensive and time-consuming.

NSF #2243980, REU Site: Advancing Data-Driven Deep Coupling of Computational Simulations and Experiments.

This study introduces a novel data-driven model that addresses these challenges by integrating real-world hourly wind speed and electricity price data to predict both wind farm generation and revenue under varying turbulence conditions. By incorporating statistical turbulence corrections and wake loss estimations, the model captures essential dynamics at both turbine and farm scales. Based on the authors' knowledge, this is the first study to combine turbulence correction and wake loss adjustments with hourly market data to provide simultaneous power and revenue predictions for wind farms. The framework provides robust power and revenue estimations using only general wind and terrain data, and it can also be applied to wind energy system design without the need for extensive simulations.

II. Model Derivation

A. Overview

The present model combines computational simulations with statistical corrections to improve the accuracy of power and revenue predictions for wind farms operating under turbulent conditions. This approach utilizes OpenFAST and TurbSim to simulate power generation under both steady and turbulent flows, providing a foundation for developing turbulence-corrected power curves. These curves are further adjusted to account for wake losses observed in wind farms, caused by reduced wind speeds at downstream turbines. The statistical software R is employed to fit and analyze the power curves, refining turbulence corrections and wake loss estimations. The resulting power curve, modified for both turbulence and wake effects, is then integrated with hourly wind speed and market price data, enabling detailed revenue projections. This multi-step process creates a comprehensive model capable of accurately estimating wind farm performance without relying on historical production data.

B. OpenFAST Simulation

OpenFAST, developed by the National Renewable Energy Laboratory (NREL), is an advanced open-source simulation tool for modeling the dynamic response and performance

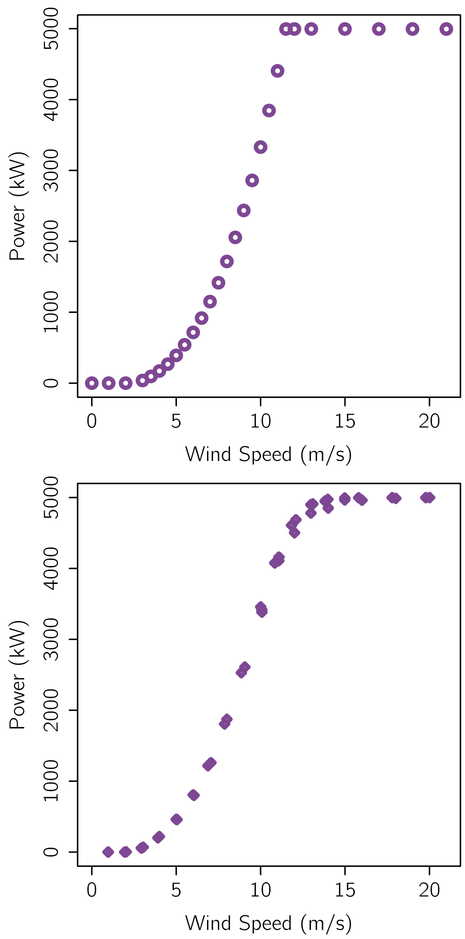

Fig. 1. Results of OpenFAST power simulation under steady flow (upper) and turbulent conditions (lower).

of wind turbines. It integrates multiple physics modules to simulate aerodynamics, structural dynamics, control systems, and electrical aspects, allowing for comprehensive time-series assessments of turbine behavior and power generation. In this study, OpenFAST is configured to simulate power generation for the NREL 5 MW reference turbine [4] under both steady and turbulent wind flows. To provide realistic turbulent wind conditions, TurbSim is used to generate time-varying flow fields based on a given hourly average wind speed. Configured according to the IEC 61400-1 standard [5], specifically turbulence class B with a reference intensity of 0.14, TurbSim employs 6–12 random seeds to produce unique flow profiles for OpenFAST, simulating one hour of turbine operation under varied turbulence. For steady flow, the wind speed while sampling was set in intervals of 1 m/s for speeds from 1 to 3 m/s, 0.5 m/s for speeds from 3 to 12 m/s, and 2 m/s for speeds from 13 to 21 m/s. For turbulent flow, the wind speed was set in intervals of 1 m/s from 1 to 15 m/s, and 2 m/s for speeds from 16-20 m/s with each wind speed being sampled at least twice with a different starting seed, and three times in the region of 10-15 m/s.

The data gathered from OpenFAST can be used to generate a parametric power curve. The choices of model are numerous [6], but only the four and five-parameter logistic functions, the double exponential function, and the Weibull CDF function were considered for fitting. Each model was fit to both the turbulent and steady flow data, upon which the Akaike Information Coefficient (AIC_C) with Correction was used as a model selection criterion. The model with the lowest AIC_C should, in theory, be the best representation of the data with a penalty to the number of parameters. In both cases, the five-parameter logistic function had the smallest AIC_C of all the tested models and was thus chosen to represent all power curves going forward.

C. Smoothing Procedure

Now that a method of representing a power curve as a continuous function has been established, it is possible to analyze the effect of turbulence as a statistical process. Beginning from the assumption that the steady flow power curve is an accurate representation of the relationship between wind speed and power output for any instant in time, it is possible to obtain a new power curve in cases where wind speed is not constant, but takes a range of values over the sampling period. From the mean and standard deviation of the wind speed measured over a given time period, it is possible to determine an approximate probability density function. In this case, the two-parameter Weibull distribution is chosen as it is commonly used and reliable for wind applications [7].

For an observed mean wind speed denoted $\bar{V}$ it is possible to determine the power output under turbulent conditions denoted P_{turb} by treating it as the expected value of the steady flow power curve P_{steady}. The equation for this expected value is

$$P_{turb}\left(\bar{V}\right) = E\left(P_{steady}(V) \mid \bar{V}, \sigma_{\bar{V}}\right) \tag{1}$$

$$= \int_0^\infty W_{\bar{V},\sigma_{\bar{V}}}(v) P_{steady}(v)\, dv \tag{2}$$

where $W_{\bar{V},\sigma_{\bar{V}}}(v)$ is the Weibull distribution with mean $\bar{V}$ and standard deviation $\sigma_{\bar{V}}$. The notation $\sigma_{\bar{V}}$ is chosen to emphasize that the standard deviation is a function of mean wind speed.

The integral in (2) does not generally have a closed-form solution. In order to obtain a parameterization for this new curve, the integral can be calculated for a sample of wind speeds upon which a new curve can be fitted. Using input wind speed values from 1 to 20 m/s in intervals of 1 m/s proves generally sufficient for an accurate fit, and will be used moving forward.

D. Standard Deviation Models

Before applying the turbulence correction, an equation for the standard deviation $\sigma_{\bar{V}}$ as a function of mean wind speed $\bar{V}$ must be determined. A good initial guess for any site is the IEC 61400-1 Normal Turbulence Model (NTM) defined as

$$\sigma_{\bar{V}} = I_{ref}\left(0.75\bar{V} + 3.8\right) \tag{3}$$

where I_{ref} is the reference turbulence intensity at 15 m/s categorized from very high turbulence (0.18) to low turbulence (0.12) [5]. The OpenFAST trials use an alternative model

$$\sigma_{\bar{V}} = I_{ref}\left(\frac{z\bar{V}}{100} + 3.8\right). \quad (4)$$

which incorporates the hub-height z of the wind turbine [3]. Using the data collected from the TurbSim simulation, it is possible to fit a linear model to the sample standard deviation as a function of wind speed. The results of this procedure are shown in Fig. 3 where the linear regression fit is notably similar to (4) with $z = 90$. Using the result of the linear fit to determine the standard deviation, the OpenFAST steady curve can be smoothed via (2) to predict its shape under turbulent conditions. Fig. 3 shows the result of this process compared with the OpenFAST power curve.

From the graph, it appears that the smoothed curve is more accurate than if the steady flow curve were used for power estimation, but still deviates from the turbulent curve derived directly from the OpenFAST simulation. However, it remains to be seen what impact this has on power estimation. Using a generic Weibull distribution with a mean wind speed of 8 m/s and a scale parameter of 11.3 to represent the distribution of wind speed at a hypothetical site, it is possible to estimate the power output of the OpenFAST wind turbine over the course of a year. The Average Expected Power will be defined by

$$AEP = 8766\int_0^\infty W(v)P(v)\,dv \quad (5)$$

where $W(v)$ is the probability density function for the wind speed, and $P(v)$ is the power curve under investigation. The results of calculating the AEP for each of the power curves are recorded in Table I. Additionally, the mean absolute error (MAE) is calculated between the power curve predictions and the OpenFAST turbulent data, as well as the MAE exclusively for wind speed below 11 m/s. From the results, it appears that the estimated turbulent curve offers a substantial improvement over the steady flow curve when it comes to MAE. However,

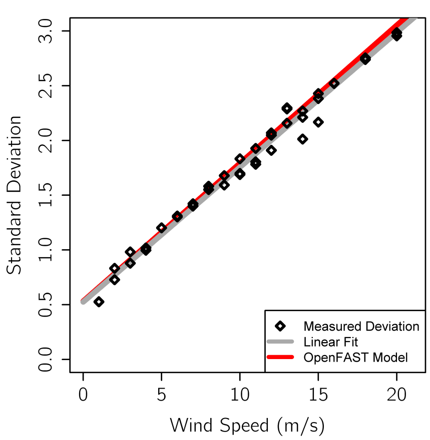

Fig. 2. OpenFAST wind speed standard deviation equation as derived from 4 and by linear fit of the OpenFAST data.

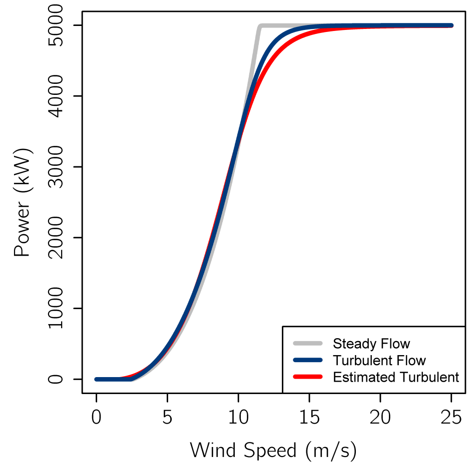

Fig. 3. Results of smoothing procedure to estimate power generation under turbulence, plotted alongside the steady flow power curve and true OpenFAST turbulent curve.

TABLE I
PREDICTED POWER PRODUCTION USING EACH CURVE FROM FIG. 3.

Model	**MAE**	**MAE** < 11	**Power (MW)**
Steady	101	75	$18,235\ (-0.7\%)$
Turb. Estimate	70	27	$18,356\ (-1.6\%)$
Turbulent	28	21	$18,070$

the steady flow curve comes closer in estimated power output because it underestimates for low wind speeds and underestimates for higher wind speeds, canceling out the error. Finally, we note that while the estimated turbulent curve has a rather high overall MAE, it becomes nearly identical for wind speeds under 11 m/s.

It appears that the observed turbulence in wind speed may be insufficient for fully accurate power curve estimation. This likely comes from the fact that while the steady flow power curve used to derive the smoothed model has a hard cap of 5 MW, the OpenFAST model does not. Instead, OpenFAST models generator torque control, which modifies blade-pitch to stabilize the power output towards the rated power of 5 MW. This leads to slightly higher power generation than our model is capable of predicting, especially in the region just below the rated wind speed, where the turbine switches from optimal to suboptimal blade-pitches most frequently. Developing a model that accurately simulates this interaction poses some major challenges, but should prove fruitful for future research.

E. Estimating Steady Flow Curves

Using the same assumptions required to go from a steady flow curve to a turbulent flow curve, it is possible to generate a steady flow curve from a turbulent flow curve. Using least squares fitting, it is possible to find which steady flow curve most closely matches the OpenFAST turbulent curve upon applying the smoothing procedure with standard deviation

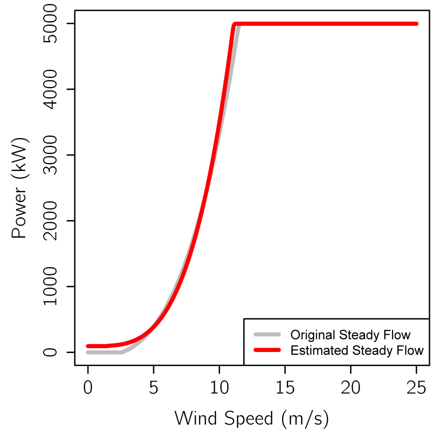

Fig. 4. Result of reverse smoothing procedure on turbulent curve.

equal to that given by the linear fit model in Fig. 2. The result of this fitting procedure is given in Fig. 4. The similarity of the two curves indicates that the process is equally applicable in both directions.

F. A Toy OpenFAST Model

Now that a sufficiently accurate model for predicting simulated power has been developed, it becomes possible to represent the simulation procedure as a random process. Using (4) with $z = 0.9$ it is possible to simulate the results of potential OpenFAST generated power estimates. A standard 10-minute sample in OpenFAST contains 4 samples per second and 2400 samples total. Thus, by randomly sampling from a Weibull distribution and taking the mean power of 2400 samples across a series of sample mean wind speeds, a collection of hypothetical power predictions can be obtained. This process is repeated 100 times for each wind speed from 3-20 m/s in intervals of 0.01.

Using the smoothed curve as the true value of the power curve, it is possible to obtain a distribution of the absolute error for each simulation as shown in the top of Fig. 5. There appears to be a distinct peak in absolute error in the region around 10 m/s such that any sampling procedure for power curves should include multiple repetitions for wind speeds in this region. A theoretical sampling procedure based on this observation is provided in Table II. Carrying out this procedure and taking the mean of the extra trials over 100 repetitions gives the bottom graph in Fig. 5 where the absolute error is considerably reduced for the middle regions. This same

TABLE II
PROPOSED SAMPLING PROCEDURE FOR POWER CURVE MODELING.

Repetitions	Wind Speed
1	3, 5-7, 15, 16, 18, 20
2	12-14
3	8-11

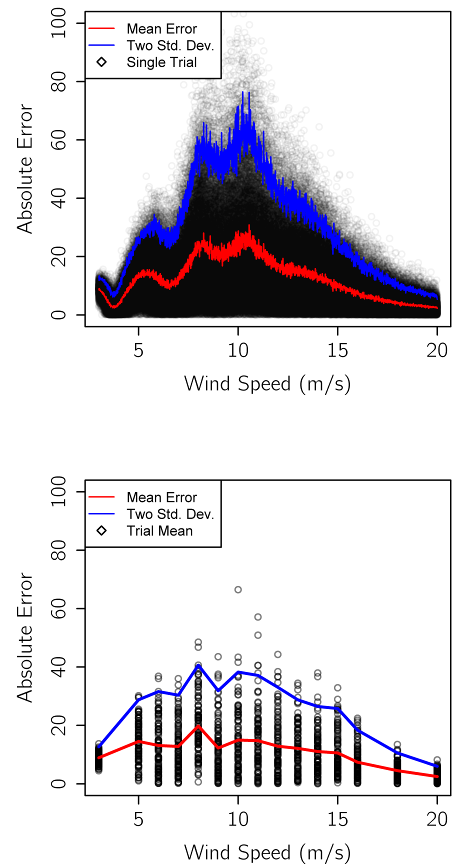

Fig. 5. Absolute errors of single trial procedure (upper) and multiple trial procedure (lower).

process can be easily modified to produce different results for different turbulence conditions and different steady flow curves if desired.

III. CASE STUDY: PYRON WIND FARM DATA

A. Data Set

The Pyron Wind Farm in Roscoe, Texas, serves as an ideal case study to validate the effectiveness of the turbulence-corrected power prediction model. By examining a dataset of hourly wind speed and power generation over four years, this analysis tests the model's ability to account for turbulence effects and wake losses at the wind farm scale. Texas, with one of the highest wind energy penetration rates in the United States, has an electricity market where prices directly reflect demand and supply dynamics. This means that fluctuations in wind power generation often coincide with real-time changes in electricity prices. Consequently, the inclusion of hourly electricity price data enables a realistic estimation of revenue, highlighting the model's capacity to address correlated fluctuations in wind power generation and electricity prices.

The Pyron wind farm in Roscoe, TX is cited to contain 166 GE 1.5 MW wind turbines with a combined rated capacity of 249 MW [8]. However, observing satellite imagery and power output, it is found to contain only 163 turbines for a total of 244.5 MW. The data used included hourly power generation and energy prices (averaged from 15 minute intervals) for the years 2018 to 2021 [9], ERA5 atmospheric wind speed adjusted for air density, and steady flow power curve generation [10]. Using this data, it is possible to construct a power curve for the entire wind farm and evaluate its performance. Hours which report generation of below 5 kW at wind speeds above 8 m/s are ignored for model evaluation as they likely represent periods of mechanical failure, abnormal weather events, or are otherwise unsuitable for evaluating typical generation performance.

B. Wake Loss Estimation

The dynamics of wind farms are complicated by the presence of wake losses. As wind flows through a turbine, its speed decreases as some of its kinetic energy is extracted by the rotor blades. This reduction in wind speed results in a loss of power (wake loss) for downwind turbines when compared to atmospheric data predictions. Wake loss increases with the size and density of turbines, such that the approximate wake loss can be calculated as a function of the power density of a wind farm. Badger et al. offer an estimate of the relationship between power density and wake loss using the Kinetic Energy Budget of the Atmosphere model [11]. A degree four polynomial is fit to the results of these estimates in Fig. 6.

Unlike most offshore wind farms, onshore wind farms are often non-uniform in density, as is the case for the Pyron wind farm. This means that the effects of wake loss may vary greatly in different sections of the farm. This would result in individual turbines having different power curves depending on their location. Using satellite image data from the US Wind Turbine Database [12] it is possible to split a farm into multiple sections and estimate the power density in each. This is done for the Pyron wind farm in Fig. 7. Note that the Pyron farm is surrounded by multiple wind farms which are counted when considering wake loss while not contributing to reported power production.

Using the fit from Fig. 6 it is then possible to calculate the percent loss of power expected in each region. Assuming that this loss is directly due to a slowing of the wind, it is then possible to calculate the percent wind speed associated with the percent power production. This involves first fitting a Weibull distribution to the site wind speed data, then calculating the expected average power without wake loss, and solving to find what percent wind speed would result in the predicted wake loss. This procedure gives the percent wind speeds shown in Fig. 7.

After the percent wind speed in each region has been calculated, a new power curve for each of the 163 turbines can be determined by multiplying the input wind speed of the steady flow curve by the percent wind speed associated with its region. Summing up all of these power curves then produces a power curve for the entire wind farm adjusted for wake loss. This curve for the Pyron wind farm is shown in yellow in Fig. 8.

C. Results of Pyron Farm Power Estimation

Now that a power curve with wake loss adjustment has been obtained, it is possible to properly apply the turbulence correction and use the resulting power curve to provide production estimates. The IEC NTM estimate from (3) for the value of $\sigma_{\bar{V}}$ is used with a value of I_{ref} equal to 0.12 signifying low turbulence determined by considering the site's relatively simple geography. The results of the smoothing procedure are given in Fig. 8 alongside the steady flow curves with and without wake loss. These power curves can then be used to predict the power production by using the observed wind speed to estimate the expected power. This is done in Table III for the initial steady flow curve and the final smoothed curve with the percent error from the raw power provided as well. An additional power curve was fit to the data via least squares estimation, which will serve as a baseline for the best possible fit of the data

Overall, it appears that the smoothing and wake loss procedures provide a significant improvement over the steady

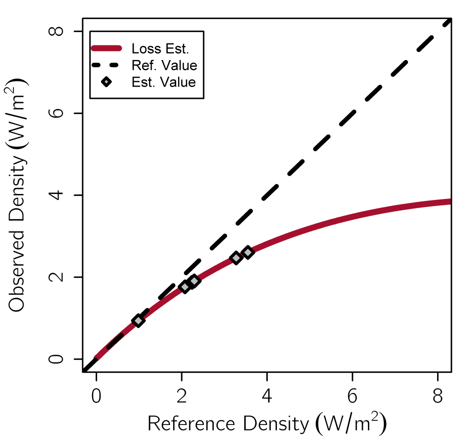

Fig. 6. Estimate of wake loss from power density.

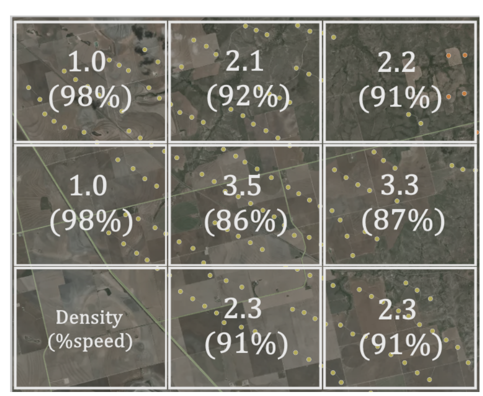

Fig. 7. Satellite image of Pyron wind farm split with observed density and estimated percent wind speed.

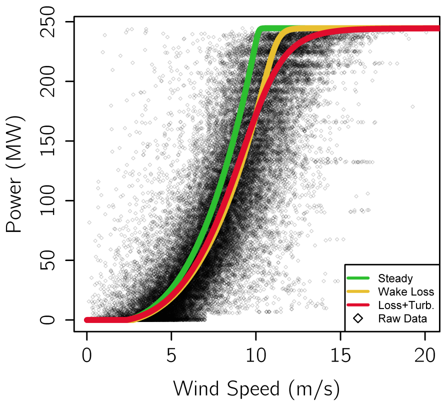

Fig. 8. Pyron farm power curves under steady flow with no wake loss adjustment, under steady flow with wake loss adjustment, and under turbulent flow with wake loss adjustment.

flow curve when it comes to long-term power prediction. However, it is important that the estimated power curve is not only accurate in the long-term, but also offers an informative estimate at any given wind speed. Thus, the MAE over the course of all four years in the sampling period is calculated for each of the four power curves. From the results in Table IV it appears that the smoothed wake loss curve outperforms all others curves, including the least squares estimate. Thus, the simulated OpenFAST results appear to translate well to the reality of wind farm dynamics.

The modeling approach thus discussed can be considered "data blind" in the sense that no power generation data is used in creating the smoothed model. Thus, certain simple improvements could be easily applied but are intentionally omitted. For example, the wind farm has a theoretical rated power of 244.5 MW, but only ever reaches this value three times in four years in spite of average wind speeds above the rated speed being present in 3,750 hours. Taking the mean of the top 0.05% of power generation values indicates that a value of 242.1 MW would be more appropriate. This small change decreases the percent error in power prediction by almost 1.2% with a marginal decrease in MAE alongside it. The largest improvements would likely come from obtaining accurate on-site turbulence estimates. In fact, if hour-by-hour turbulence estimates were available, power generation for each hour could be predicted using (2) without needing a turbulence function or even generating a power curve.

TABLE III
RESULTS OF POWER CURVE PREDICTIONS BY YEAR (1,000 MWH).

Year	Raw	Steady Flow	Smoothed	Least Sq.
2018	774	978 (+26%)	808 (+4.3%)	780 (+1.1%)
2019	784	991 (+27%)	812 (+3.6%)	785 (+0.5%)
2020	797	1,002 (+26%)	818 (+2.7%)	789 (-0.7%)
2021	748	949 (+27%)	775 (+3.7%)	752 (+1.0%)
Total	3,100	3,930 (+26%)	3,222 (+3.5%)	3,111 (+0.4%)

TABLE IV
MEAN ABSOLUTE ERROR FOR EACH POWER CURVE.

Curve	MAE
Steady Flow	36.8
Wake Loss	29.3
Smoothed	27.8
Least Sq.	28.3

D. Randomness of Power Estimation

Now that a reasonable power curve estimate has been obtained, it is important to understand the long-term variability of both the wind farm power production and potential power estimates. This is done by splitting the data into bins by wind speed and fitting a probability density function to each bin. In this case, the four-parameter Beta distribution is chosen as it works well for a variety of bounded data. The parameters determining the upper and lower bound of the distribution are set to constant values of 0 and 244.507 respectively, as these are the maximum and minimum values observed in the data. This gives the formula

$$\frac{x^{\alpha-1}(244.5-x)^{\beta-1}}{244.5^{\alpha+\beta-1}B(\alpha,\beta)} \tag{6}$$

where the parameters α and β are to be fit to the data and $B(\alpha, \beta)$ is the Beta function which acts as a normalizing constant. Once these values have been calculated in each bin, it is possible to fit a function to estimate the intermediate values. This is done using least squares spline fitting with degrees of freedom determined via cross validation. The results of this estimation are given in Fig. 9.

These predicted α and β values can then be used to take any observed wind speed and output a randomly generated power

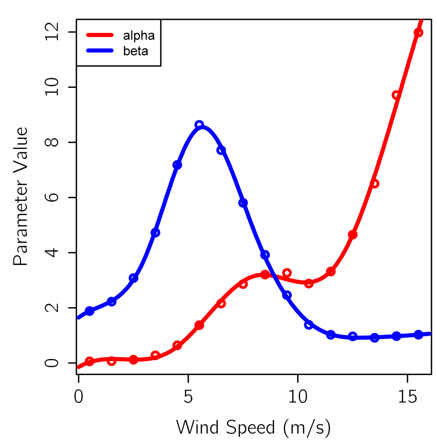

Fig. 9. Alpha and beta estimate fitted on binned values.

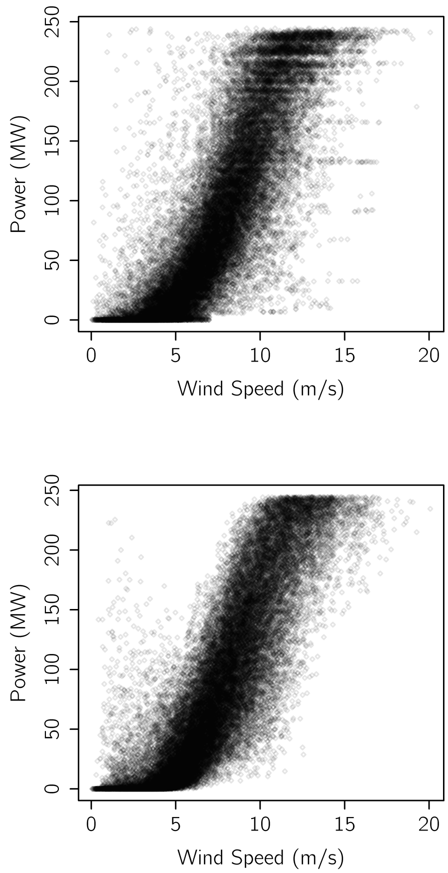

Fig. 10. Raw power production by wind speed from the Pyron farm (upper) and simulated power production via the dynamic Beta distribution (lower).

estimate from the associated Beta distribution. This process was performed on all observed wind speed values at the Pyron farm, with the results placed beneath the raw power data in Fig. 10. Summing up the simulated estimates produces a four-year power production estimate of about 3,103 MW, whereas the true value over this period was 3,104 MW. Thus, this random simulation appears to offer a reasonable approximation of the observed power production.

E. Revenue Estimation

Using the power curve obtained for the Pyron wind farm, it is possible to obtain annual revenue estimates from regional energy prices. Bechman et al. use Danish energy market data to construct an energy value density distribution to determine which wind speeds are responsible for the greatest proportion of generated revenue [13]. For the Pyron wind farm the revenue density will be given by

$$R(v) = P_{turb}(v)W(v)f(v) \tag{7}$$

where v is the wind speed, $P_{turb}(v)$ is the Pyron turbulent flow power curve, $W(v)$ is the Weibull density function of wind speed, and $f(v)$ is the price function of the regional market. The price function $f(v)$ was determined using hourly energy price values for the years 2018-2020 from the ERCOT West zone, where the Pyron wind farm is located. Prices above \$1000/MWh were discarded as outliers, accounting for 41 observations of over 26,000 total. Data from 2021 was not included, as the severe winter storm in February caused massive spikes in energy prices of a magnitude and duration not seen in previous years [14]. It should be noted that while removing these outliers greatly improves model accuracy, these periods of extreme price account for a significant proportion of wind farm profits. Thus, the following discussion is only meant to investigate model accuracy under standard conditions and more sophisticated corrections may be required to obtain fully accurate revenue estimates. A least-squares spline model with degrees of freedom determined by cross validation was then fit to the energy data, after which the revenue density was calculated using (7).

The results of this procedure can be seen in Fig. 11 where it appears that the wind farm may be unable to capitalize on periods of high wind speed due to energy price dips from competing farms. In the case of a dynamic energy market, this causes the revenue density to aggregate towards lower

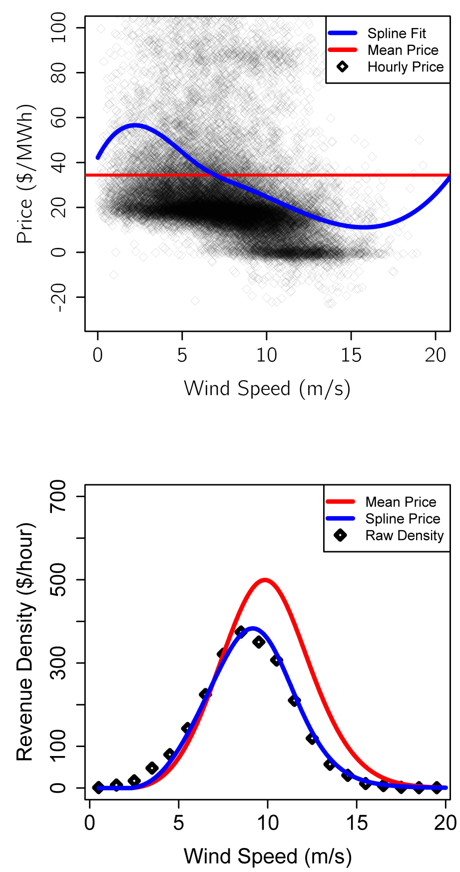

Fig. 11. Competing estimates for $f(v)$ (left), and revenue density function of the Pyron farm using different price and power estimates (right).

wind speeds where prices are higher compared to the case of a static market with energy price equal to the overall mean.

Taking the integral of the energy density function gives the Annual Expected Revenue by

$$AER = 8776 \int_0^\infty R(v)dv \tag{8}$$

which outputs a value of $\$20,126,376$ for the Pyron wind farm. Alternatively, multiplying the hourly power generation of the wind farm by the energy price of each hour and weighting up to account for the removed hours gives a raw revenue of $\$20,284,931$, or about 0.08% greater than that AER method. This indicates that the error associated with turbulent curve's power estimation is significantly diminished for estimating revenue, likely because high wind speeds where the error is greatest contribute least to the revenue density.

Using the mean price instead of the price function $f(v)$ yields a significantly greater revenue of $\$27,459,420$ indicating that the average price provides a significant overestimate. Additionally, the median of the revenue density curve with mean price is reached at 9.9 m/s while with flexible prices it is reduced to 9.1 m/s. Further, the flexible price curve indicates that more than 50% of all revenue is generated at wind speeds between 7.5 and 10.7 m/s. This corroborates previous findings indicating that wind power providers should seek to extract more energy from lower wind speeds in order to maximize profits in competitive markets [13].

IV. Conclusion

This study introduces a data-driven model that enhances wind power prediction accuracy by incorporating turbulence and wake effects, validated through simulations and a detailed case study of the Pyron Wind Farm. By refining traditional power curves with statistical adjustments for turbulence, the model offers a reliable framework for predicting both power output and revenue under variable wind conditions. The case study demonstrates the model's effectiveness in addressing the real-world variability of wind farm performance, providing a practical tool for assessing both operational efficiency and value delivered in wind energy systems.

This model holds significant potential as a fast, adaptable solution for data-driven wind energy assessments, particularly valuable in scenarios where historical data is limited or unavailable. As renewable energy demands continue to grow, such models will be critical for optimizing wind farm design, operations, and economic planning across diverse wind environments.

References

[1] C. Anderson, *Wind turbines theory and practice*. Cambridge: Cambridge University press, 2020.

[2] OpenFAST v3.5.3. Accessed Oct. 27, 2024. [Online]. Available: https://github.com/openfast/openfast/

[3] B. Jonkman and M. B. Jr., *TurbSim User's Guide*. National Renewable Energy Laboratory, Sep. 2006.

[4] J. Jonkman, S. Butterfield, W. Musial, and G. Scott, *Definition of a 5-MW Reference Wind Turbine for Offshore System Development*. National Renewable Energy Laboratory, Feb. 2009.

[5] I. 61400-1, *"Wind turbines-Part 1: Design requirements." 3rd edition*. International Electrotechnical Commission, Aug. 2005.

[6] V. Sohoni, S. C. Gupta, and R. K. Nema, "A critical review on wind turbine power curve modelling techniques and their applications in wind based energy system," *Journal of Energy*, 2016.

[7] D. Yu, *Data Science for Wind Energy*. Boca Raton: CRC Press, 2020.

[8] "Pyron Wind Farm LLC", Power Utility Comission of Texas. [Online]. Available: https://www.puc.texas.gov/industry/electric/directories/pgc/report_pgc.aspx?ID=PGSQL01DB1245458100178

[9] Electric reliability council of Texas (ERCOT). Accessed Oct. 22, 2023. [Online]. Available: https://www.ercot.com

[10] D. Millstein, S. Jeong, A. Ancell, and R. Wiser, "A database of hourly wind speed and modeled generation for us wind plants based on three meteorological models," *Scientific Data*, 2023.

[11] J. Badger, M. Imberger, P. Volker, A. Kleidon, S. Germer, and J. Minz, "Making the most of offshore wind: Re-evaluating the potential of offshore wind in the german north sea," *Agora Energiewende, Agora Verkehrswende, Technical University of Denmark and Max-Planck-Institute for Biogeochemistry*, 2020.

[12] B. D. Hoen, J. E. Diffendorfer, J. T. Rand, L. A. Kramer, C. P. Garrity, and H. E. Hunt, "United states wind turbine database v7.0 (may 10, 2024): U.S. geological survey," *American Clean Power Association and Lawrence Berkeley National Laboratory data release*, 2018.

[13] A. Bechmann, T. Barlas, and H. A. Madsen, "Incorporating electricity prices in wind turbine design: Introducing the AEV metric," *Journal of Physics: Conference Series*, 2024.

[14] A. Lindstrom and A. Gorski, "Average texas electricity prices were higher in february 2021 due to a severe winter storm," *U.S. Energy Information Administration*, 2024.