Proceedings of the ASME 2023 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2023
August 20-23, 2023, Boston, Massachusetts

IDETC2023-114762

A MULTI-FIDELITY GAUSSIAN PROCESS REGRESSION METHOD FOR PROBABILISTIC WIND FARM POWER CURVE ESTIMATION

Honglin Li*
The University of Texas at Dallas
Richardson, TX 75080
Email: honglin.li@utdallas.edu

Cong Feng†
National Renewable Energy Laboratory
Golden, CO 80401
Email: cong.feng@nrel.gov

Jie Zhang‡
The University of Texas at Dallas
Richardson, TX 75080
Email: jiezhang@utdallas.edu

ABSTRACT

Accurate estimation of the power curve for wind turbines or wind farms is crucial to ensure their efficient operation and management. However, conventional methods for power curve estimation rely either on expensive and infrequent measurements or on low-quality numerical simulations. Moreover, the majority of previous studies on power curve estimation for wind turbines or wind farms focused on deterministic estimation, which provides a point estimate of the relationship between wind speed and power generation. Nevertheless, the deterministic approach fails to consider the inherent uncertainty associated with wind energy production resulting from varying turbine characteristics. This can lead to inaccurate power generation estimation and suboptimal decisions regarding energy management. In this paper, a kernel density estimation (KDE) based Multi-Fidelity Gaussian Process Regression (MFGPR) model is proposed to fuse theoretical power curve data and the ground true measurements to create a mapping of wind speed and wind power. By conducting a case study on an actual wind farm in China, the efficacy of the proposed MFGPR model was demonstrated in characterizing the variability of wind power. The probabilistic MFGPR model was also able to generate confidence intervals that encompassed the measured power, thereby improving the accuracy and confidence in wind power estimation or wind resource assessment. Overall, the proposed MFGPR model offers a reliable approach to integrate high-fidelity ground measurements and theoretical power curve data, resulting in precise wind resource assessment and power estimation.

Keywords: Probabilistic wind power curve estimation, Multi-fidelity modeling, Gaussian process regression, Kernel density estimation.

1 Introduction

Wind energy has emerged as a promising source of renewable energy, contributing significantly to the global energy mix. However, the intermittent nature of wind poses significant challenges for grid integration and energy management. Accurately estimating the power curve of wind turbines or wind farms is essential for both wind farm operations and long-term power grid planning. This is because precise power curve estimation can decrease the uncertainty in short-term and long-term wind power forecasting [1, 2].

Most of existing research on wind turbine/farm power curve estimation has focused on deterministic estimation [3, 4, 5], which provides a point estimate of the relationship between wind speed and power generation. However, the deterministic approach does not account for the inherent uncertainty of wind energy production due to varying

*Ph.D. Student, Department of Mechanical Engineering and the Center for Wind Energy.

†Director's Postdoctoral Fellow.

‡Associate Professor, Department of Mechanical Engineering, Department of Electrical and Computer Engineering (Affiliated), and the Center for Wind Energy (Affiliated), ASME Professional Member. Address all correspondence to this author.

turbine characteristics, resulting in inaccurate power generation estimation and suboptimal energy management decisions. Probabilistic wind farm power curve estimation [6], on the other hand, characterizes the stochastic nature of the wind power generation by providing confidence intervals in wind power estimation.

Accurate wind farm power curve estimation is crucial for the efficient and reliable operations of wind farms. Probabilistic power curve estimation has the potential to help improve the economic viability and competitiveness of wind energy by reducing the uncertainty and risk associated with its operation and grid integration.

1.1 Literature Review

Accurate models of wind turbine power curves are important for understanding turbine performance and aiding operational decision-making. Lydia et al. [7] have discussed the significance of wind turbine power curve in understanding the performance of wind turbines, monitoring and controlling turbines, estimating wind energy potential, and forecasting wind power. Li et al. [8] have compared different machine learning models for estimating wind turbine power curves.

Multi-fidelity modeling (MFM) is a technique that aims to leverage information from various sources of data with different levels of fidelity to improve the accuracy of predictions [9, 10]. In the context of energy prediction, MFM has been shown to be effective in reducing the computational cost of predicting wind energy output while maintaining high accuracy. Luo et al. [11] proposed a cascaded multi-fidelity deep learning (CMF-DL) framework that utilizes high-fidelity numerical weather observation data and low-fidelity numerical weather prediction (NWP) data to improve the accuracy of photovoltaic power forecasting. An MFM based Super-Fidelity Network (SFNet) deep learning method for wind farm wake modeling was proposed in [12]. The SFNet combines advantages of analytical models and numerical models to predict high-fidelity flow fields using low-fidelity data. The model has a mean absolute error of 1.9% when trained on only 45 samples and can predict flow fields of a wind farm consisting of 100 turbines within several seconds on a standard desktop.

Gaussian Process Regression (GPR) is a popular method for modeling the output of complex systems, such as probabilistic wind power curve estimation, due to its ability to handle non-linear relationships and uncertainties. For example, Rogers et al. [13] have proposed a heteroscedastic Gaussian process model that automatically quantifies variance in predictions, to generate accurate mean predictions and higher likelihoods than a homoscedastic model on operational wind turbine data. Manobel et al. [14] proposed a wind turbine power curve estimation method using Artificial Neural Networks (ANN) with Gaussian Processes prefiltering, showing a 25% improvement in root mean square error for the predicted power compared to benchmark parametric and non-parametric methods.

Kernel density estimation (KDE) is another technique that has been explored to help quantify the uncertainty in wind energy. For example, Juban et al. [15] proposed a KDE-based method for wind power forecasting. Zhang et al. [16, 17] proposed a KDE-based method to accurately characterize and predict the annual variation of wind conditions.

1.2 Research Objective

This paper aims to develop an enhanced probabilistic wind farm power curve estimation model using a KDE-based Multi-Fidelity Gaussian Process Regression (MFGPR) method, by combining intermittent high-fidelity measurements with continuous low-fidelity numerical simulations from models such as the theoretical power curve. By considering the varying characteristics in wind turbines and uncertain wind conditions, the KDE-based MFGPR model could predict the mean value and confidence intervals of the wind power output. The main contributions of this paper include: (1) Develop a kernel density estimation based multi-fidelity Gaussian Process Regression model for wind turbine/farm power curve estimation; and (2) Explore the impact of linear and non-linear combination methods of Gaussian kernels in multi-fidelity models on wind turbine/farm power curve estimation.

The remainder of the paper is organized as follows. The developed KDE-based MFGPR method is described in Section 2. Section 3 describes the experimental setup for the case study, and Section 4 discusses the results in probabilistic wind power curve estimation. Concluding remarks and future work are discussed in Section 5.

2 Methodology

In machine learning or deep learning modeling, it is a common practice in the data pre-processing phase to clean the data, filter out most of the noisy data, and keep the data concentrated in the theoretical range of values. This can greatly improve the model's ability to fit the theoretical values. However, adopting such an approach may result in the loss of actual data characteristics, especially when estimating models in uncertain conditions, such as power curve estimation for wind turbines, which is considerably influenced by mechanical malfunctions as well as the variability of wind speed. Simply cleaning out these discrete values would severely reduce the ability of the model to fit the real

situation. If the amount of data is simply reduced, it is challenging for the GPR model to accurately capture the characteristics within the data. On the other hand, if the full dataset is used to model the GPR model, there could exist over-fitting issues with the highly noisy data.

This paper utilizes the KDE algorithm to sample the entire dataset, benefiting from its various advantages. First, the amount of data is significantly reduced while retaining the original data distribution characteristics. Second, the sampling process only follows the density of the data distribution without assuming any underlying distribution of the data. Third, KDE estimates the probability density function (PDF) at each data point using a kernel function that assigns more weight to nearby points. This local estimation can be advantageous when the underlying distribution varies across the data space as seen in the wind energy data. The weakening or disappearance of the underlying signal due to data sampling can be compensated for by low-fidelity data.

2.1 Kernel Density Estimation (KDE)

KDE [15, 16, 18] is a non-parametric way to estimate the probability density function of a random variable. The basic idea of KDE is to estimate the density function of a random variable by placing a kernel (usually a Gaussian function) on each data point and summing the resulting kernel functions. The resulting density estimate is then normalized so that it integrates to one. KDE is a flexible method that can be used to estimate the PDF of data with any distribution, including multimodal distributions. This makes it useful for a wide range of applications where the underlying distribution is unknown or complex. Unlike parametric methods, KDE does not make any assumptions about the underlying distribution of the data. This can be an advantage when the data is highly skewed, has outliers, or comes from a distribution that is difficult to model parametrically. KDE preserves all the information in the data, unlike methods that involve binning the data or fitting a parametric distribution, which can be an advantage when analyzing data with complex distributions or small sample sizes. The density is estimated by:

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K_h\left(\frac{x-x_i}{h}\right) \tag{1}$$

where $\hat{f}_h(x)$ is the estimated density at x, $K_h(x)$ is the kernel function with bandwidth h, and x_i is the ith data point. The kernel function $K_h(x)$ is defined as:

$$K_h(x) = \frac{1}{h}\phi\left(\frac{x}{h}\right) \tag{2}$$

where $\phi(\cdot)$ is the standard normal probability density function. The bandwidth h is a smoothing parameter that controls the width of the kernel function. The choice of h has a significant impact on the smoothness of the estimated density. The standard normal probability density function $\phi(\cdot)$ is defined as:

$$\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \tag{3}$$

2.2 Gaussian Process Regression (GPR)

GPR is a flexible supervised machine learning technique used for regression analysis that models the distribution of a function using a Gaussian process (GP), which can provide accurate predictions and uncertainty estimates for complex functions [19]. The Gaussian process is defined as a collection of random variables, and any finite number of variables has a joint Gaussian distribution.

The key components of the GPR model include the mean function $m(\mathbf{x})$ and the covariance function $c(\mathbf{x},\mathbf{x}')$, which together fully specify the Gaussian process. The mean function represents the expected value of the function at each input point, while the covariance function specifies the degree of similarity between the function values at different input points. The covariance function is an appropriate kernel function parameterized by a vector of hyperparameters θ that results in a symmetric positive-definite covariance matrix $C_{ij} = c(\mathbf{x}_i,\mathbf{x}_j;\theta), C \in \mathbb{R}^{n\times n}$ [20].

The goal of GPR is to estimate the underlying function $f(\mathbf{x})$ given a set of input-output pairs $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i \in \mathbb{R}^D$ and $y_i \in \mathbb{R}$. The GPR model assumes that the output y_i is a noisy observation of the function value $f(\mathbf{x}_i)$, and that the noise follows a Gaussian distribution with a zero mean and variance σ_n^2, i.e. $f \sim \mathcal{GP}\left(f \mid \mathbf{0}, \sigma_n^2\right)$, where $\sigma_n^2 = c(\mathbf{x}_i,\mathbf{x}_j;\theta)$

The predicted function value $\hat{f}(\mathbf{x})$ at a new input point $\mathbf{x}$ is given by the posterior mean of the Gaussian process, which is a weighted sum of the observed function values:

$$\hat{f}(\mathbf{x}) = \mathbf{c}_*^\top \left(\mathbf{C} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{y} \tag{4}$$

where $\mathbf{c}_*$ is a vector of covariance between the new input point $\mathbf{x}$ and the training input points, $\mathbf{C}$ is the covariance matrix of the training input points, and $\mathbf{y}$ is a vector of the observed function values.

The uncertainty in the predicted function value is characterized by the posterior variance of the Gaussian process, which measures the degree of uncertainty in the predicted function value at a new input point:

$$\mathrm{Var}[\hat{f}(\mathbf{x})] = c(\mathbf{x},\mathbf{x}) - \mathbf{c}_*^\top \left(\mathbf{C} + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{c}_* \tag{5}$$

where $c(\mathbf{x}, \mathbf{x})$ is the covariance between the new input point $\mathbf{x}$ and itself.

A systematic extension of the GPR framework can be employed to create probabilistic models by recursion, allowing for the amalgamation of data from sources with variable fidelity [20, 21, 22].

$$g_t(\mathbf{x}) = \rho g_{t-1}(\mathbf{x}) + \delta_t(\mathbf{x}) \tag{6}$$

Equation 6 defines a framework in which Gaussian process g_{t-1} and g_t model data at different levels of fidelity. The scaling constant ρ quantifies the correlation between the model outputs y_t and y_{t-1}, while $\delta_t(x_t)$ is a GP with a mean of $\mu\delta_t$ and a covariance function of c_t. In other words, μ_t follows a Gaussian process with the mean $\mu\delta_t$ and covariance function $c_t(\mathbf{x}_t, \mathbf{x}_t'; \theta_t)$, denoted as $\mu_t \sim \mathcal{GP}(\delta_t | \mu\delta_t, c_t(\mathbf{x}_t, \mathbf{x}_t'; \theta_t))$.

2.3 Multi-fidelity GPR Modeling

Multi-fidelity modeling is a technique used to improve the efficiency and accuracy of computer simulations by incorporating information from simulations of varying levels of fidelity. In many engineering applications, simulations of high-fidelity (i.e., detailed and accurate) models can be computationally expensive and time-consuming, while simulations of low-fidelity (i.e., simplified and approximate) models are much faster but may be less accurate. Multi-fidelity modeling seeks to combine the advantages of both high- and low-fidelity models to produce accurate predictions in a computationally efficient manner.

One common approach to multi-fidelity modeling is to use a set of low-fidelity models, denoted by $f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x})$, along with a high-fidelity model, denoted by $f_{M+1}(\mathbf{x})$. The low-fidelity models are typically less expensive to simulate, while the high-fidelity model provides the most accurate predictions but at a higher computational cost.

The goal of multi-fidelity modeling is to construct a surrogate model, denoted by $\hat{f}(\mathbf{x})$, which accurately approximates the high-fidelity model using information from the low-fidelity models. One typical method is to combine the low-fidelity models in a weighted manner with an adjustment term to account for the difference between the low-fidelity and high-fidelity models [9, 23]:

$$\hat{f}(\mathbf{x}) = \sum_{m=1}^{M} w_m(\mathbf{x}) f_m(\mathbf{x}) + \epsilon(\mathbf{x}) \tag{7}$$

where $w_m(\mathbf{x})$ are the weights assigned to each low-fidelity model, and $\epsilon(\mathbf{x})$ is the correction term that accounts for the discrepancy between the low- and high-fidelity models.

The weights $w_m(\mathbf{x})$ are typically chosen to be a function of the low-fidelity model predictions and some measure of their accuracy or fidelity, such as the discrepancy between the low- and high-fidelity models:

$$w_m(\mathbf{x}) = \frac{\omega_m(\mathbf{x})}{\sum_{i=1}^{M} \omega_i(\mathbf{x})} \tag{8}$$

where $\omega_m(\mathbf{x})$ is a fidelity measure that depends on the discrepancy between the low-fidelity model $f_m(\mathbf{x})$ and the high-fidelity model $f_{M+1}(\mathbf{x})$. One common fidelity measure is the normalized square error:

$$\omega_m(\mathbf{x}) = \frac{\|f_m(\mathbf{x}) - f_{M+1}(\mathbf{x})\|^2}{\sigma^2_{f_{M+1}}} \tag{9}$$

where $\sigma^2_{f_{M+1}}$ is the variance of the high-fidelity model predictions.

The weighted sum method is a technique that involves a linear combination of different fidelity-level models, and it is solely suitable when the models are linearly correlated. For non-linear models of different levels of fidelity, we can use Eq. 6 to fuse them together recursively by replacing the GP prior f_{t-1} with the GP posterior $f_{*_{t-1}}(\mathbf{x})$ from the previous inference level [24].

$$f_t(\mathbf{x}) = g_t(\mathbf{x}, f_{*_{t-1}}(\mathbf{x})) \tag{10}$$

where $g_t \sim \mathcal{GP}\left(c_t \mid \mathbf{0}, c_t\left(\left(\mathbf{x}, f_{*_{t-1}}(\mathbf{x})\right), \left(\mathbf{x}', f_{*_{t-1}}(\mathbf{x}')\right); \theta_t\right)\right)$. Since the inputs $f*_{t-1}(\mathbf{x})$ and $\mathbf{x}$ belong to inherently different spaces, sometimes the chosen structure for the covariance function of g_t may not be appropriate. To address this challenge, Perdikaris et al. [20] proposed an extension to the methodology by introducing a more structured prior for g_t, which better reflects the autoregressive nature of Eq. 10. Specifically, the covariance kernel can be decomposed as:

$$\begin{aligned} c_{t_g} = c_{t_\rho}\left(\mathbf{x}, \mathbf{x}'; \theta_{t_\rho}\right) \cdot c_{t_f}\left(f_{*_{t-1}}(\mathbf{x}), f_{*_{t-1}}\left(\mathbf{x}'\right); \theta_{t_f}\right) \\ + c_{t_\delta}\left(\mathbf{x}, \mathbf{x}'; \theta_{t_\delta}\right) \end{aligned} \tag{11}$$

where θ_{t_ρ}, θ_{t_f}, and θ_{t_δ} represent the hyperparameters of valid covariance functions, namely c_{t_ρ}, c_{t_f} and c_{t_δ}, respectively. All the kernel functions mentioned above are selected

to have the form of the squared exponential, which is also known as a radial basis function (RBF) kernel [19].

$$c_t\left(\mathbf{x},\mathbf{x}';\theta_t\right)=\sigma_t^2\exp\left(-\frac{1}{2}\sum_{i=1}^{d}w_{i,t}\left(\mathbf{x}_i-\mathbf{x}_i'\right)^2\right) \tag{12}$$

where σ_t^2 denotes the variance parameter and $(w_{i,t})_{i=1}^{d}$ represents automatic relevance determination weight corresponding to the fidelity level t.

3 Case Study

3.1 Data Summary

We have applied the proposed MFGPR method to a Spatial Dynamic Wind Power Forecasting (SDWPF) dataset, which is obtained from a real wind farm owned by the Longyuan Power Group Corp. Ltd. The dataset consists of the wind power data of 134 1.5 MW wind turbines, wind farm layout and weather conditions [25]. Table 1 summarizes the statistics of the SDWPF dataset.

TABLE 1: Statistics of the SDWPF dataset

Parameter	Value
Data length (Days)	245
Resolution	10 minutes
No. of Turbines	134
No. of Features	13
No. of Records	4,727,520

Figure 1 depicts 200 high-fidelity real data sampled from the 4,727,520 data points using the KDE algorithm, and 200 low-fidelity theoretical data points obtained from the theoretical power curve, which is a 1.5 MW wind turbine theoretical power curve data obtained from the Energy Systems Integration Group (ESIG)[26]. The cut-in, rated, and cut-out wind speeds are 3 m/s, 14 m/s, and 25 m/s, respectively. It is seen from the high-fidelity data that the actual power generation does not exactly follow the theoretical power curve, due to uncertainties, turbine conditions, and other factors.

3.2 Deterministic Evaluation Metrics

Standard metrics such as root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), normalized RMSE (NRMSE), and normalized MAE (NMAE) are used to assess the accuracy of deterministic power curve estimation in this paper. These metrics provide a quantitative measure of the deviation between predicted and actual power values. RMSE and MAE are based on the difference between the predicted and actual values, while MAPE measures the percentage deviation. NRMSE and NMAE are normalized versions of RMSE and MAE, respectively, which account for the scale of the data being analyzed. These metrics are defined as:

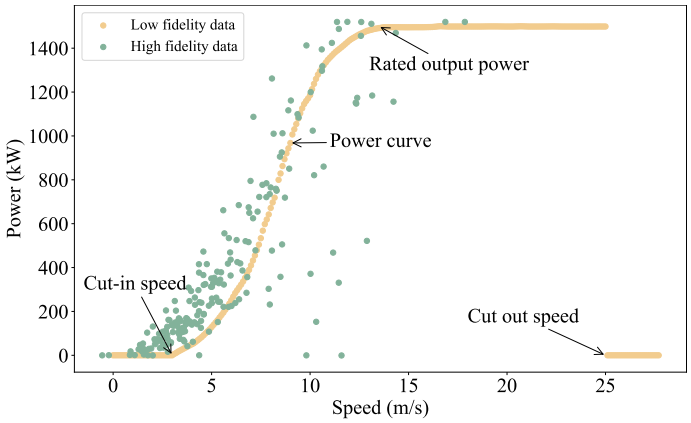

FIGURE 1: High- and low-fidelity training data used in this study

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}\left(\hat{x}_i-x_i\right)^2}{n}} \tag{13}$$

$$NRMSE=\frac{1}{x_{\max}}\sqrt{\frac{\sum_{i=1}^{n}\left(\hat{x}_i-x_i\right)^2}{n}} \tag{14}$$

$$MAE=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{x}_i-x_i\right| \tag{15}$$

$$NMAE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{x}_i-x_i}{x_{\max}}\right| \tag{16}$$

$$MAPE=\frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{x}_i-x_i}{x_i}\right| \tag{17}$$

where $\hat{x}$ represents the predicted value, x is the actual value, and n is the number of samples.

3.3 Probabilistic Evaluation Metrics

The pinball loss is a measure of the deviation between the predicted and actual values, which is commonly used in quantile regression to evaluate the accuracy of predictions [27]. The pinball loss function is defined as:

$$L_\tau(y,\hat{y}) = \begin{cases} (y-\hat{y})\tau, & \text{if } y \geq \hat{y} \\ (\hat{y}-y)(1-\tau), & \text{if } y < \hat{y} \end{cases} \tag{18}$$

where y is the actual value, $\hat{y}$ is the predicted value, and τ is the quantile of interest.

In probabilistic estimation, sharpness refers to the degree of concentration or spread of the forecast distribution around its mean. A sharp forecast means that the predicted distribution is concentrated around the mean, while a flat forecast means that the distribution is spread out. The sharpness of a probabilistic forecast can be measured using various metrics, such as the variance, standard deviation, or entropy of the forecast distribution. The smaller the value of these metrics, the sharper the forecast [28]. Mathematically, sharpness can be defined as follows:

$$S = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{M}(\hat{p}_{ij} - \bar{p}_j)^2 \tag{19}$$

where S is the sharpness, N is the number of forecast samples, M is the number of forecast probabilities in each sample, $\hat{p}_{ij}$ is the forecast probability for the j-th category in the i-th sample, and $\bar{p}_j$ is the average probability for the j-th category across all forecast samples.

Reliability of probabilistic estimation refers to the degree to which the forecast probabilities match the observed frequencies [29]. It measures how well the forecast probabilities reflect the actual outcomes. A well-calibrated probabilistic forecast is considered reliable. The most commonly used measure of reliability is the probability integral transform (PIT) histogram, which is the cumulative distribution function (CDF) evaluated at the observation. If the forecast is well-calibrated, the PIT values should be uniformly distributed between 0 and 1. Mathematically, the reliability can be expressed as:

$$Re = \frac{1}{N}\sum_{i=1}^{N}\int_{-\infty}^{\infty}[\hat{F}_i(z) - I(Y_i \leq z)]^2 dP(z) \tag{20}$$

where N is the number of observations, Y_i is the actual observation, $\hat{F}_i(z)$ is the forecast CDF at z, and I is the indicator function.

4 Results and Discussion

In the case study, we have compared the performance of three power curve estimation models, including a nonlinear kernel based MFGPR, a linear kernel based MFGPR, and a standard GPR method. All three models were tested on 9 wind turbines that were selected randomly.

The deterministic power curve estimation results are summarised in Table 2, and a smaller metric value indicates better performance. The results show that, on average, the nonlinear kernel fusion method utilizing the MFGPR approach performed better than the standard GPR method for the first eight wind turbines, demonstrating a decrease of 3.9% in RMSE and 3.5% in MAE.

Figure 2 also displays a comparison of the three models' performance in terms of deterministic power curve estimation for the 9 wind turbines across 50 consecutive time intervals. The figure shows that the non-linear kernel fusion approach consistently outperforms the other two models, namely the standard GPR and the linear kernel fusion approach. The non-linear kernel fusion approach produces the most accurate estimation for the power output of the wind turbines in the vast majority of the time intervals. In contrast, the standard GPR and the linear kernel fusion approaches show less consistent performance, with both models producing higher errors in certain time intervals. Overall the non-linear kernel fusion approach is more robust and reliable in predicting the power output of wind turbines.

Figures 3 and 4 provide further insights into the performance of the three models for the probabilistic estimation of wind turbine power output. Figure 3 shows that the non-linear kernel fusion method produces more accurate probabilistic estimation results, especially for extreme value points. This suggests that the non-linear kernel fusion approach is better able to capture the underlying patterns and uncertainties in the data, resulting in more accurate probabilistic estimation.

Figure 4 compares the performance of the three models across three probabilistic evaluation criteria: sharpness, reliability, and pinball loss. The results show that the standard GPR achieves the best results for sharpness across all nominal proportions, indicating that it produces more tightly concentrated predictive distributions. For reliability, the non-linear kernel fusion method performs worse than the linear fusion method for smaller nominal proportions, but performs better for larger nominal proportions, with both models outperforming the standard GPR model. This suggests that the non-linear approach is better able to pro-

TABLE 2: Deterministic power estimation results using nonlinear kernel, linear kernel, and standard GPR

		T1	T2	T3	T4	T5	T6	T7	T8	T9
Nonlinear kernel MFGPR	RMSE (kW)	188.75	225.88	225.97	210.56	201.77	232.98	181.60	210.72	250.11
	NRMSE (%)	0.12	0.15	0.15	0.14	0.13	0.15	0.12	0.14	0.16
	MAE (kW)	112.27	117.05	143.92	115.39	122.34	134.12	105.77	127.14	155.75
	NMAE (%)	0.07	0.08	0.09	0.08	0.08	0.09	0.07	0.08	0.10
Linear kernel MFGPR	RMSE (kW)	212.61	239.48	238.97	222.44	212.92	244.84	203.28	212.45	262.50
	NRMSE (%)	0.14	0.16	0.16	0.15	0.14	0.16	0.13	0.14	0.17
	MAE (kW)	150.48	162.02	178.65	150.01	158.50	171.81	145.24	162.13	198.50
	NMAE (%)	0.10	0.11	0.12	0.10	0.10	0.11	0.10	0.11	0.13
Standard GPR	RMSE (kW)	201.07	226.36	241.07	216.55	209.82	238.48	198.96	217.74	248.67
	NRMSE (%)	0.13	0.15	0.16	0.14	0.14	0.16	0.13	0.14	0.16
	MAE (kW)	113.00	126.19	146.09	118.69	126.31	139.77	106.74	132.85	166.71
	NMAE (%)	0.07	0.08	0.10	0.08	0.08	0.09	0.07	0.09	0.11

duce probabilistic estimations that are consistent with the true outcomes. For pinball loss that measures the accuracy of probabilistic estimation, the non-linear kernel fusion method achieves better results than the other two models in extreme cases.

Figure 5 depicts the NRMSE results of the non-linear MFGPR model for all turbines in the wind farm. With the exception of a few turbines for which relatively large values are obtained, the non-linear MFGPR model has accurately estimated the power curve for the majority of wind turbines.

5 Conclusion

In this study we proposed and investigated the the multi-fidelity Gaussian process regression (GPR) approach for wind farm power curve estimation. By comparing the results of three different GPR models on nine randomly selected wind turbines, the study found that the non-linear kernel fusion method MFGPR consistently outperformed the standard GPR model, achieving lower errors across multiple metrics such as RMSE, reliability, and pinball loss. Moreover, the non-linear kernel fusion method MFGPR demonstrated superior performance for both deterministic and probabilistic power curve estimation.

Overall, these findings highlight the potential of the multi-fidelity GPR approach, and in particular, the non-linear kernel fusion method, for improving the accuracy and reliability of wind energy estimation. This has significant implications for the efficient operation of wind farms and the maximization of wind energy utilization. Therefore, this study provides a valuable contribution to the field of wind energy forecasting and emphasizes the potential of the proposed method for future applications.

Potential future research directions include (i) examining the applicability of the MFGPR model in forecasting wind power uncertainty based on existing wind farms, and (ii) exploring ways to enhance the accuracy of estimation while reducing computational costs through the integration of additional techniques.

ACKNOWLEDGMENT

This paper was developed based upon funding from the Alliance for Sustainable Energy, LLC, Managing and Operating Contractor for the National Renewable Energy Laboratory for the U.S. Department of Energy.

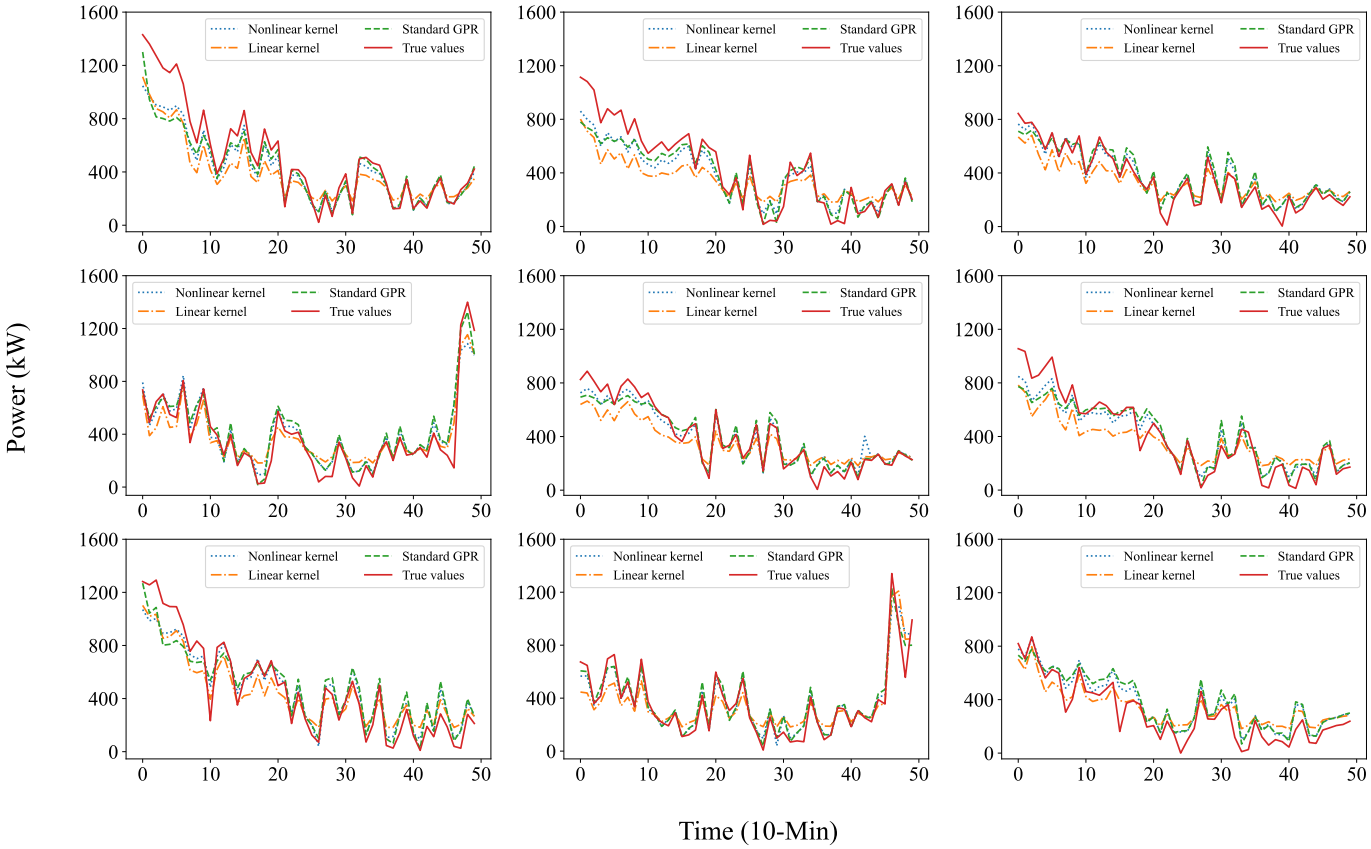

FIGURE 2: Deterministic single turbine power curve and power generation estimation for nine randomly selected wind turbines. Row 1: T1 – T3; Row 2: T4 – T6; Row 3: T7 – T9

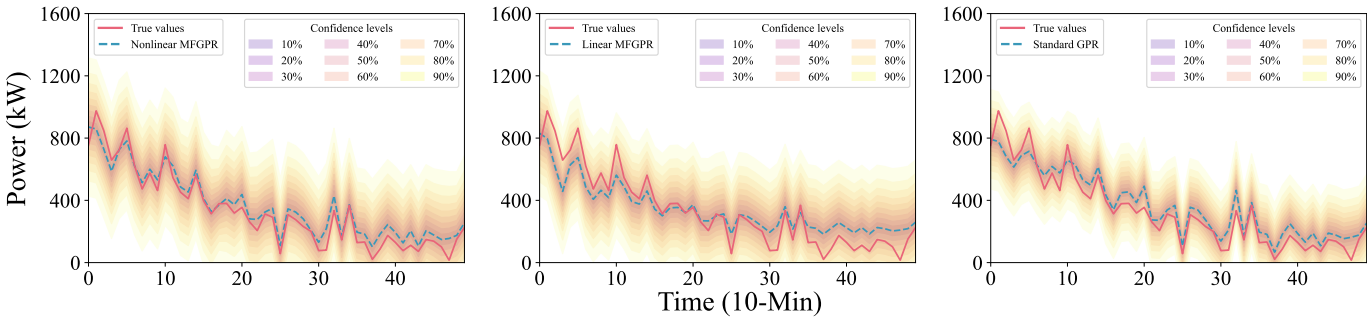

FIGURE 3: Probabilistic power curve and power generation estimation for one selected wind turbine. From left to right, the three subplots show the results of probabilistic estimates of the nonlinear kernel MFGPR, linear kernel MFGPR, and standard GPR models on the same turbine.

REFERENCES

[1] Arwade, S. R., Lackner, M. A., and Grigoriu, M. D., 2011. "Probabilistic models for wind turbine and wind farm performance". *Journal of Solar Energy Engineering,* ***133***(4).

[2] Olaofe, Z., and Folly, K., 2012. "Wind power estimation using recurrent neural network technique". In IEEE Power and energy society conference and exposition in africa: intelligent grid integration of renewable energy resources (PowerAfrica), IEEE, pp. 1–7.

[3] Chang, T.-P., Liu, F.-J., Ko, H.-H., Cheng, S.-P., Sun, L.-C., and Kuo, S.-C., 2014. "Comparative analysis

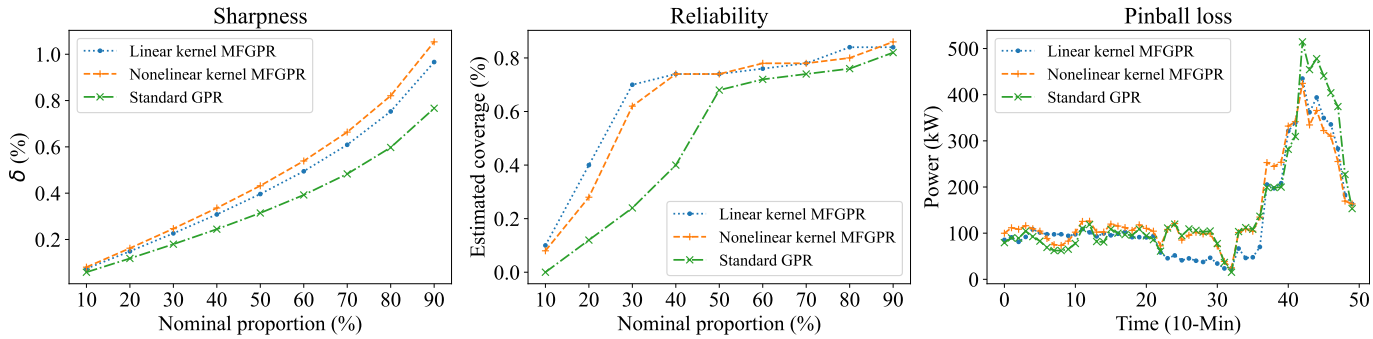

FIGURE 4: The sharpness, reliability, and pinball loss metrics for evaluating the performance of probabilistic power estimation

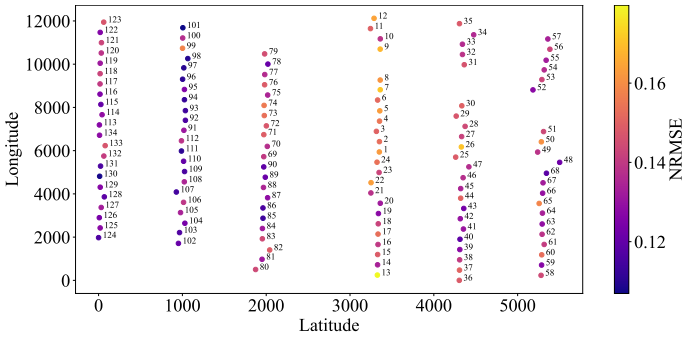

FIGURE 5: The NRMSE results of all turbines in the wind farm based on the nonlinear MFGPR power curve estimation (the units of latitude and longitude are meter)

on power curve models of wind turbine generator in estimating capacity factor". *Energy,* ***73***, pp. 88–95.

[4] Lydia, M., Selvakumar, A. I., Kumar, S. S., and Kumar, G. E. P., 2013. "Advanced algorithms for wind turbine power curve modeling". *IEEE Transactions on sustainable energy,* ***4***(3), pp. 827–835.

[5] Pei, S., and Li, Y., 2019. "Wind turbine power curve modeling with a hybrid machine learning technique". *Applied Sciences,* ***9***(22), p. 4930.

[6] Yun, E., and Hur, J., 2021. "Probabilistic estimation model of power curve to enhance power output forecasting of wind generating resources". *Energy,* ***223***, p. 120000.

[7] Lydia, M., Kumar, S. S., Selvakumar, A. I., and Kumar, G. E. P., 2014. "A comprehensive review on wind turbine power curve modeling techniques". *Renewable and Sustainable Energy Reviews,* ***30***, pp. 452–460.

[8] Li, S., Wunsch, D. C., O'Hair, E., and Giesselmann, M. G., 2001. "Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation". *J. Sol. Energy Eng.,* ***123***(4), pp. 327–332.

[9] Song, X., Lv, L., Sun, W., and Zhang, J., 2019. "A radial basis function-based multi-fidelity surrogate model: exploring correlation between high-fidelity and low-fidelity models". *Structural and Multidisciplinary Optimization,* ***60***, pp. 965–981.

[10] Wang, X., Liu, Y., Sun, W., Song, X., and Zhang, J., 2018. "Multidisciplinary and multifidelity design optimization of electric vehicle battery thermal management system". *Journal of Mechanical Design,* ***140***(9), p. 094501.

[11] Luo, X., and Zhang, D., 2023. "A cascaded deep learning framework for photovoltaic power forecasting with multi-fidelity inputs". *Energy*, p. 126636.

[12] Li, R., Zhang, J., and Zhao, X., 2022. "Multi-fidelity modeling of wind farm wakes based on a novel superfidelity network". *Energy Conversion and Management,* ***270***, p. 116185.

[13] Rogers, T., Gardner, P., Dervilis, N., Worden, K., Maguire, A., Papatheou, E., and Cross, E., 2020. "Probabilistic modelling of wind turbine power curves with application of heteroscedastic gaussian process regression". *Renewable Energy,* ***148***, pp. 1124–1136.

[14] Manobel, B., Sehnke, F., Lazzús, J. A., Salfate, I., Felder, M., and Montecinos, S., 2018. "Wind turbine power curve modeling based on gaussian processes and artificial neural networks". *Renewable Energy,* ***125***, pp. 1015–1020.

[15] Juban, J., Siebert, N., and Kariniotakis, G. N., 2007. "Probabilistic short-term wind power forecasting for the optimal management of wind generation". In 2007 IEEE Lausanne Power Tech, IEEE, pp. 683–688.

[16] Zhang, J., Chowdhury, S., Messac, A., and Castillo, L. "Multivariate and multimodal wind distribution model based on kernel density estimation". In ASME 2011 5th International Conference on Energy Sustainability, American Society of Mechanical Engineers Digital Collection, pp. 2125–2135.

[17] Zhang, J., Chowdhury, S., Messac, A., and Castillo, L.,

2013. “A multivariate and multimodal wind distribution model”. *Renewable Energy,* ***51***, pp. 436–447.

[18] Parzen, E., 1962. “On estimation of a probability density function and mode”. *The annals of mathematical statistics,* ***33***(3), pp. 1065–1076.

[19] Seeger, M., 2004. “Gaussian processes for machine learning”. *International journal of neural systems,* ***14***(02), pp. 69–106.

[20] Perdikaris, P., Raissi, M., Damianou, A., Lawrence, N. D., and Karniadakis, G. E., 2017. “Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling”. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* ***473***(2198), p. 20160751.

[21] Kennedy, M. C., and O’Hagan, A., 2000. “Predicting the output from a complex computer code when fast approximations are available”. *Biometrika,* ***87***(1), pp. 1–13.

[22] Le Gratiet, L., and Garnier, J., 2014. “Recursive co-kriging model for design of computer experiments with multiple levels of fidelity”. *International Journal for Uncertainty Quantification,* ***4***(5).

[23] Fernández-Godino, M. G., Park, C., Kim, N.-H., and Haftka, R. T., 2016. “Review of multi-fidelity models”. *arXiv preprint arXiv:1609.07196*.

[24] Cutajar, K., Pullin, M., Damianou, A., Lawrence, N., and González, J., 2019. “Deep gaussian processes for multi-fidelity modeling”. *arXiv preprint arXiv:1903.07320*.

[25] Zhou, J., Lu, X., Xiao, Y., Su, J., Lyu, J., Ma, Y., and Dou, D., 2022. “Sdwpf: A dataset for spatial dynamic wind power forecasting challenge at kdd cup 2022”. *arXiv preprint arXiv:2208.04360*.

[26] ESIG. 1.5 mw wind turbine theoretical power curve. [Online] Available at: `https://www.esig.energy/wiki-main-page/general-electric-1-5-mw-series/`. [Accessed: 01 Mar. 2023].

[27] Sun, M., Feng, C., Chartan, E. K., Hodge, B.-M., and Zhang, J., 2019. “A two-step short-term probabilistic wind forecasting methodology based on predictive distribution optimization”. *Applied energy,* ***238***, pp. 1497–1505.

[28] Feng, C., Cui, M., Hodge, B.-M., and Zhang, J., 2017. “A data-driven multi-model methodology with deep feature selection for short-term wind forecasting”. *Applied Energy,* ***190***, pp. 1245–1257.

[29] Sun, M., Feng, C., and Zhang, J., 2019. “Conditional aggregated probabilistic wind power forecasting based on spatio-temporal correlation”. *Applied Energy,* ***256***, p. 113842.