Evaluation of Probability Distribution Models for Renewable Energy Forecasting Under Uncertain Conditions

Thai An Nguyen
Faculty of Electrical and Electronics Engineering, HCMC University of Technology and Engineering, Ho Chi Minh City, Vietnam
ntan@hcmute.edu.vn (corresponding author)

Mi Sa Nguyen Thi
Faculty of Electrical and Electronics Engineering, HCMC University of Technology and Engineering, Ho Chi Minh City, Vietnam
misa@hcmute.edu.vn

Huy Anh Quyen
Faculty of Electrical and Electronics Engineering, HCMC University of Technology and Engineering, Ho Chi Minh City, Vietnam
anhqh@hcmute.edu.vn

Tu Duc Nguyen
Faculty of Electrical and Electronics Engineering, HCMC University of Technology and Engineering, Ho Chi Minh City, Vietnam
ducnt@hcmute.edu.vn

Received: 10 January 2026 | Revised: 12 February 2026, 9 March 2026, and 21 March 2026 | Accepted: 23 March 2026

 DOI: https://doi.org/10.48084/etasr.17452

ABSTRACT

This study explores the capability of several Probability Distribution Functions (PDFs) to effectively model and forecast power generation from renewable energy sources integrated within a Microgrid (MG). The focus is specifically on wind and solar Photovoltaic (PV) systems. The wind speed datasets are analyzed using the Weibull, Rayleigh, Lognormal, Generalized Extreme Value (GEV), and Normal distributions, while solar irradiance data are represented through the Beta, Normal, Triangular, and Lognormal distributions. Actual meteorological data are sourced from the Wind Resource Database (WRDB) and the National Solar Radiation Database (NSRDB), with a focus on the Kauai region of Hawaii. Model performance and suitability are evaluated using multiple statistical indicators, including the coefficient of determination (R^2), the Log-Likelihood, the Akaike Information Criterion (AIC), and the Root Mean Square Error (RMSE). The comparative analysis reveals that the Weibull distribution provides the most precise representation of wind speed behavior, outperforming the other models. Meanwhile, for solar irradiance, the Beta distribution demonstrates superior accuracy and adaptability, while the Normal and Triangular models are less effective due to their simplified assumptions.

Keywords-Probability Distribution Function (PDF); weibull distribution; beta distribution; renewable energy modeling; wind speed modeling; solar radiation modeling; power output estimation

I. INTRODUCTION

Solar and wind energy are the most common renewable sources in Microgrids (MG), but their intermittent, uncertain, and distributed nature complicates large-scale integration [1]. Wind power output primarily depends on wind speed, while solar power is influenced by irradiance, temperature, and weather conditions [2]. These uncertainties cause supply–demand fluctuations, impacting system efficiency and stability. To address this issue, Probability Distribution Functions (PDFs) are widely used, with Weibull and Rayleigh being the most common [3-5]. Authors in [6] found Weibull to be more accurate than Rayleigh in estimating power density, though without considering economic factors. Several studies have applied the Weibull distribution to model Wind Turbine (WT) power curves and to assess reliability, but they mainly focus on wind energy [7, 8]. Research combines Weibull (wind) and Beta (solar) distributions to capture temporal variability, using the Frank-Copula algorithm and K-means clustering [9]. In addition, authors in [10] proposed control strategies based on machine learning to enhance MG performance under uncertain conditions. However, most of these studies rely on assumptions or generalized datasets, without validating the suitability of the probability distributions against actual local data. Moreover, these studies often model wind and solar energy using a single PDF, while neglecting empirical verification based on real-world data.

This study presents a data-driven probabilistic framework for modeling uncertainty in renewable energy by systematically evaluating candidate distributions for solar irradiance and wind speed in Kauai, Hawaii. The solar irradiance data are obtained from the publicly available NSRDB database [11] and described in [12], while the wind speed data are sourced from the WRDB database [13] and detailed in [14]. Model parameters are estimated using Maximum Likelihood Estimation (MLE) and evaluated through RMSE, MAE, R^2, and AIC metrics to ensure robust model selection [15, 16]. By adopting a multi-criteria goodness-of-fit framework, the study provides a more reliable basis for probabilistic characterization of renewable resources compared with conventional single-distribution assumptions, thereby supporting improved uncertainty-aware analysis in MG applications.

II. METHOD

In renewable energy research, selecting an appropriate Probability Density Function (PDF) for modeling wind speed and solar irradiance data is crucial. An accurate distribution model enables more reliable predictions of local climatic conditions, thereby supporting more effective planning and design of renewable energy systems. The procedure for evaluating PDF models for wind speed and solar irradiance is illustrated in Figure 1.

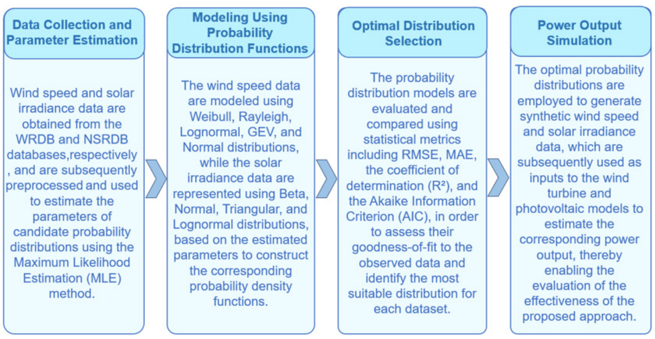

Fig. 1. Workflow for evaluating probability distribution models.

A. *Maximum Likelihood Estimation (MLE) Method for Distribution Parameter Estimation*

To determine the parameters of a probability distribution that best fits experimental data, the MLE method is employed.

Assume a dataset consisting of n independent observations $X = \{x_1, x_2, ..., x_n\}$, which follow a probability distribution with a PDF $f(x;\theta)$, where θ is the set of parameters to be estimated. The likelihood function is then defined as [17]:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) \tag{1}$$

By taking the derivative of the log-likelihood function with respect to the parameter θ and setting the derivative equal to zero, the parameter values that maximize the likelihood function can be obtained. The goal is to find the parameter values that make the observed data most probable [18]:

$$\frac{\partial \log L(\theta)}{\partial \theta} = 0 \tag{2}$$

B. *Probability Distribution Model for Solar Energy*

The Beta distribution is well-suited for normalized irradiance because it is defined on a finite interval, while the Lognormal distribution effectively models positively skewed behavior. The Normal distribution serves as a reference for symmetric variability, and the Triangular distribution provides a simple representation when limited information is available.

The PDF of the Beta distribution, used to model the uncertainty of solar radiation, is given by [19]:

$$f_b(s) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times s^{(\alpha-1)} \times (1-s)^{(\beta-1)} \\ \quad \text{for } 0 \le s \le 1, \alpha \ge 0, \beta \ge 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{3}$$

The Normal distribution is used to model the symmetric variation and can approximate solar radiation under certain conditions. The PDF of a normal random variable x with mean μ and standard deviation σ is given by [20]:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \sigma > 0 \tag{4}$$

The Lognormal distribution is suitable for modeling non-negative solar irradiance with asymmetric variability. The PDF of the Lognormal distribution is given by [21]:

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right], x > 0,\ \sigma > 0 \tag{5}$$

When limited statistical information is available, the Triangular distribution offers a simple and effective model for solar irradiance uncertainty. The PDF of the Triangular distribution on the interval [0,1] is defined as [22]:

$$f(x) = \begin{cases} \dfrac{2x}{c}, & 0 \le x \le c \\ \dfrac{2(1-x)}{1-c}, & c \le x \le 1 \end{cases} \tag{6}$$

where c is the mode of distribution, taking a value between 0 and 1. The function increases linearly from zero at $x = 0$ to its maximum at $x = c$, and then decreases linearly to zero at $x = 1$.

Based on (3)-(6), the MLE method is employed to estimate the parameters of all considered probability distributions. The model for determining the output power of a PV system is given as [23]:

$$P_{PV} = \begin{cases} P_{sr}\left(\dfrac{G_s^2}{G_{std} \times X_c}\right) & \text{for } 0 < G_s < X_c \\ P_{sr}\left(\dfrac{G_s}{G_{std}}\right) & \text{for } G_s \geq X_c \end{cases} \quad (7)$$

where P_{PV} is the output power of the solar energy source, P_{sr} is the rated capacity of the PV system (in kW or MW), G_s is the solar irradiance at the considered location, G_{std} is the standard solar irradiance under known environmental conditions (approximately 1000 W/m^2), and X_c represents the specific irradiance level.

Temperature is not considered in this study because solar irradiance is the primary source of uncertainty, whereas temperature mainly affects PV conversion efficiency and has a limited impact on the model's power output compared to the variability of solar irradiance.

C. Probability Distribution Model for Wind Energy

The Weibull distribution is widely used to represent wind speed statistics due to its flexibility across different wind regimes, while the Rayleigh distribution serves as a simplified special case of the Weibull. The Lognormal and Normal distributions are included as alternative models capable of describing wind variability, and the Generalized Extreme Value (GEV) distribution is considered for its ability to capture extreme wind behavior.

The Weibull probability distribution is commonly employed [24]:

$$PDF_v\left(V\right) = \left(\frac{k}{c}\right)\left(\frac{V}{c}\right)^{(k-1)} \exp\left[-\left(\frac{V}{c}\right)^k\right] \quad 0 \leq V \leq \infty \quad (8)$$

where $PDF_v\left(V\right)$ is the PDF of wind speed V (m/s), c denotes the scale parameter, and k represents the shape parameter of the Weibull.

The Rayleigh distribution is a special case of the Weibull distribution when the shape parameter is a constant value $k = 2$. The mathematical expression is defined as [25]:

$$f\left(x\right) = \frac{2x}{c^2} \exp\left[-\left(\frac{x}{c}\right)^2\right]; x > 0, c > 0 \quad (9)$$

To capture extreme variations in wind speed, the GEV distribution is adopted. This distribution is suitable for modeling tail behavior and extreme wind events, and its mathematical expression is [26]:

$$f(x) = \frac{1}{\alpha} \exp\left\{-\left[1 - \frac{\beta\left(x - x_0\right)}{\alpha}\right]^{\frac{1}{\beta}}\right\}\left[1 - \frac{\beta\left(x - x_0\right)}{\alpha}\right]^{\frac{1}{\beta}-1} \quad (10)$$

The PDFs of the Normal and Lognormal distributions for wind energy uncertainty analysis are defined analogously to those presented for solar irradiance in Section II.B, as given in (4) and (5). The parameters of the probability distributions considered for modeling the uncertainty of wind speed are estimated using the MLE method, as presented in Section A.

The relationship between the WT output power and the wind speed is expressed as [27]:

$$P_w\left(V\right) = \begin{cases} 0 & \text{for } V < V_i \text{ and } V > V_0 \\ P_r\left(\dfrac{V - V_i}{V_r - V_i}\right) & \text{for } (V_i \leq V \leq V_r) \\ P_r & \text{for } (V_r \leq V \leq V_0) \end{cases} \quad (11)$$

where P_r is the rated power of the WT (in kW or MW), and V , V_i , V_r , V_o are the current wind speed, cut-in speed, rated speed, and cut-out speed of the turbine, respectively.

III. RESULTS AND DISCUSSION

This study uses one-day wind data from the Wind Resource Database (WRDB) in Kauai, Hawaii, USA, to determine input parameters for Weibull, Lognormal, Normal, Rayleigh, and GEV distributions. The aim is to evaluate and compare the modeling capabilities of each distribution for the region's wind characteristics, thereby identifying the optimal model for accurately assessing wind energy potential. Solar irradiance data for the same location are obtained from the National Solar Radiation Database (NSRDB), with Global Horizontal Irradiance (GHI) selected as the input parameter because it represents total solar radiation, including both direct and diffuse components, and is the most widely available irradiance variable in meteorological databases. Using these data, Beta, Normal, Triangular, and Lognormal distributions are constructed to model the probabilistic characteristics of solar energy within the simulation framework. Statistical indicators, including RMSE, MAE, R^2, and AIC, are used to evaluate the goodness of fit of the probability distribution models.

The Siemens SWT-3.0-108 WT [28], with a rated capacity of 3,000 kW and a rotor diameter of 108 m, is selected for its suitability in offshore wind conditions. The PV system is designed to deliver a peak capacity of 1 MW under standard test conditions, which are defined by an irradiance of 1,000 W/m². The solar panels selected for the system are from the TOPBiHiKu7 series by Canadian Solar, each with a maximum power output of 590 W [29]. The technical specifications of the turbine and solar panel are presented in Table I. These two renewable sources help reduce reliance on diesel generators and enhance the operational efficiency of the MG.

TABLE I. TECHNICAL SPECIFICATIONS OF THE WT AND SOLAR PANEL

WT	Solar panel
Model: SWT-3.0-108	Model: CS7L-590TB-AG
Hub height: 79.5 (m)	Vmp: 34.5 (V)
Cut-in wind speed: 3 (m/s)	Imp: 17.12 (A)
Rated wind speed: 12 (m/s)	Voc: 41.7 (V)
Cut-off wind speed: 25 (m/s)	Isc: 18.01 (A)

A. *Solar Energy*

1) *Descriptive Statistical Analysis of Solar Energy Data*

Figure 2 presents the average daily solar irradiance during the first month in the form of a bar chart, illustrating significant fluctuations due to weather conditions. The GHI ranges from below 100 W/m² to over 220 W/m², with particularly high values observed between days 20 and 25.

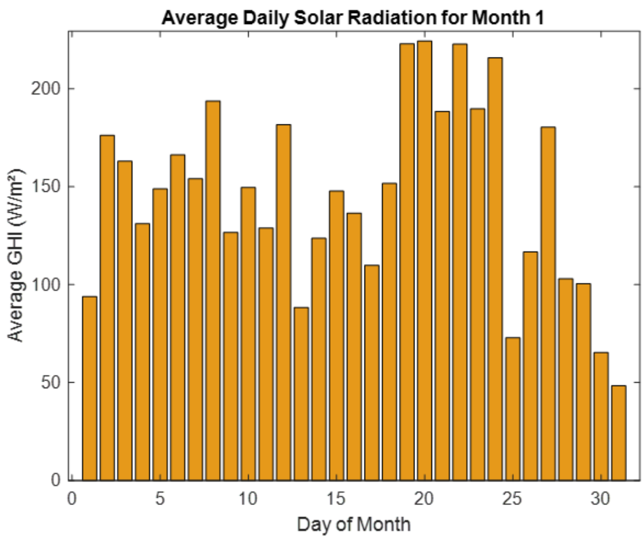

Fig. 2. Daily average solar irradiance bar chart for January.

TABLE II. ESTIMATED PARAMETERS USING MLE FOR VARIOUS PROBABILITY DISTRIBUTION MODELS

Distribution	MLE estimates of the parameters
Beta	Alpha (a): 0.7622 Beta (b): 1.0978
Lognormal	Log mean (μ): 5.3271 Log std (σ): 1.1642
Triangular	Lower limit (a): 2 Upper limit (b): 821 Peak (c): 266.5
Normal	Scale (σ): 311.9655 Location (μ): 221.0559

To select the most suitable distribution for modeling solar radiation, this study compares several common distributions, including the Beta, Normal, Lognormal, and Triangular distributions (3)–(6). The distribution parameters are estimated using the MLE method as described in (1), with parameter values obtained by maximizing the likelihood through solving the partial derivatives in (2). The results of the parameter estimation via MLE are presented in Table II.

2) *Model Goodness-of-Fit Evaluation*

After parameter estimation, each distribution's PDF/CDF is used to generate theoretical values, which are compared with empirical data. Evaluation metrics (RMSE, MAE, R^2, AIC) are summarized in Table III.

The results in Table III demonstrate that the Beta distribution is the most suitable model for solar irradiance, exhibiting the lowest RMSE (0.0004) and MAE (0.0003), the highest R^2 (0.9997), and the lowest AIC (−6802.79). In contrast, the Normal ($R^2 = 0.7590$), Lognormal ($R^2 = 0.6785$), and Triangular ($R^2 = 0.4236$) distributions show weaker performance, as they do not adequately capture the bounded and asymmetric nature of solar irradiance data. These findings confirm that the Beta distribution is the most appropriate probabilistic model for solar irradiance and PV power simulation.

The close RMSE and MAE values among the considered distributions indicate that all models approximate the empirical solar irradiance data with similar average accuracy within the main data range. This is expected, as both RMSE and MAE are error-based metrics measuring mean deviations between observed and modeled values. However, subtle differences in distribution shape and tail behavior are better captured by R^2 and AIC, which explains why the Beta distribution is identified as the most suitable model.

TABLE III. GOODNESS-OF-FIT EVALUATION CRITERIA FOR PROBABILITY DISTRIBUTIONS

Distribution	RMSE	MAE	R^2	AIC
Beta	0.0004	0.0003	0.9997	-6802.7883
Lognormal	0.0006	0.0004	0.6785	-6599.6243
Triangular	0.0007	0.0006	0.4236	-6372.8394
Normal	0.0005	0.0004	0.7590	-6382.4612

3) *Comparison of Distribution PDFs*

Figure 3 presents a visual comparison between the GHI dataset and theoretical distributions, illustrated through a frequency histogram overlaid with their corresponding PDF. Each color represents a distinct probability distribution fitted to the solar radiation data. The blue curve corresponds to the Beta distribution, which demonstrates a strong fit within the concentrated density region (100–400 W/m²), effectively capturing the skewness and natural bounds of the GHI data, though it shows limitations in modeling extreme values. The purple PDF represents the Lognormal distribution, suitable for right-skewed data, peaking around 150–200 W/m² and exhibiting a long tail extending beyond 700 W/m²; however, it underestimates the density below 100 W/m². The orange curve represents the Triangular distribution, which has a simple shape, offering a reasonable model for mid-range values but lacking accuracy at both distribution tails, particularly below 100 W/m² and above 600 W/m². Finally, the red curve represents the Normal distribution, which is symmetric around 250–300 W/m², and only reflects the central tendency, failing to capture the skewness present in the data. These theoretical curves are overlaid on the frequency histogram of the measured solar irradiance, enabling a direct comparison between the modeled distributions and the observed empirical data pattern.

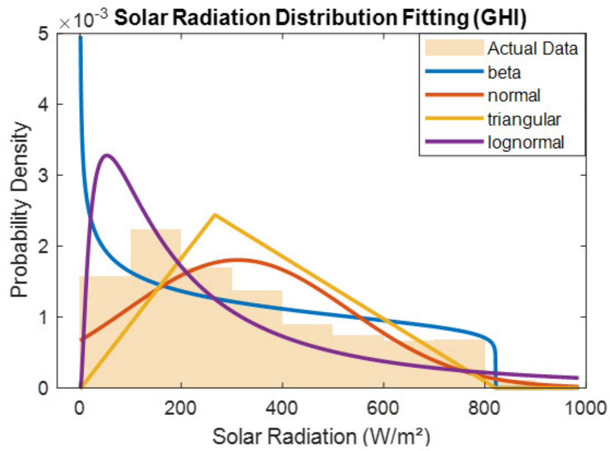

Fig. 3. Probability density functions of the distribution models.

4) Simulation of Solar Irradiance and Corresponding Output Power for the Next Day

To simulate solar irradiance at a specific location for the following day, it is assumed that the irradiance follows the Beta distribution as defined in (1). Based on this assumption, irradiance values are generated using the Beta PDF. These values are then input into (4) to compute the PV system's output power at each time step. Figure 4 illustrates the hourly solar irradiance profile simulated using the Beta distribution for the time interval between 6:00 and 19:00, which corresponds to the period of daily solar availability. The irradiance follows a bell-shaped curve, peaking around midday (13:00–14:00) and decreasing rapidly in the late afternoon, consistent with natural solar patterns. This confirms the feasibility of employing probability distributions to simulate solar irradiance data for predicting PV power generation. Additionally, the right panel of Figure 4 depicts the corresponding PV output power converted from the simulated irradiance. The power output exhibits a similar distribution, reflecting the direct relationship between irradiance and power generation, and further validating the rationale and applicability of using probability-based simulation methods. The consolidated results indicate that the Beta distribution is the most suitable model for simulating solar irradiance, owing to its high accuracy and the lowest error metrics compared to actual data. Consequently, the Beta distribution is employed to simulate hourly irradiance throughout the day.

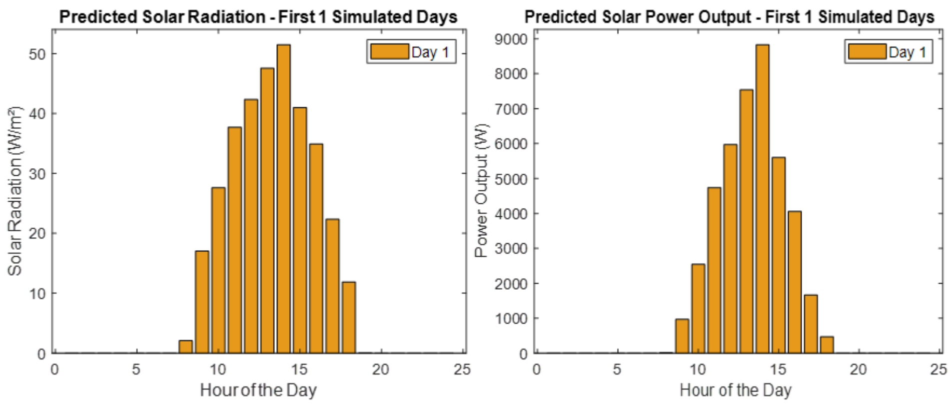

Fig. 4. Simulated solar irradiance and PV output power.

B. Wind Energy

1) Descriptive Statistical Analysis of Wind Data

Figure 5 illustrates wind speed at an 80 m height over a one-month period, revealing substantial daily variability. During the first half of the month, wind speeds range between 3–9 m/s, showing a gradual decline from day 5 to day 15. From day 20 onward, wind speed increases significantly, peaking near 17 m/s on day 31. This irregular fluctuation reflects the inherent randomness and unpredictability of meteorological data and highlights the high potential for wind energy harvesting toward the end of the month. The collected data serve as input for probabilistic modeling and are used to support MG operation optimization. To assess goodness-of-fit, the Weibull distribution is compared with Rayleigh, Lognormal, GEV, and Normal distributions. Parameters are estimated using the MLE method based on each distribution's PDF (1), and optimized by solving the system of partial derivatives (2). The other distributions are treated similarly. Estimated parameters at 80 m height are presented in Table IV.

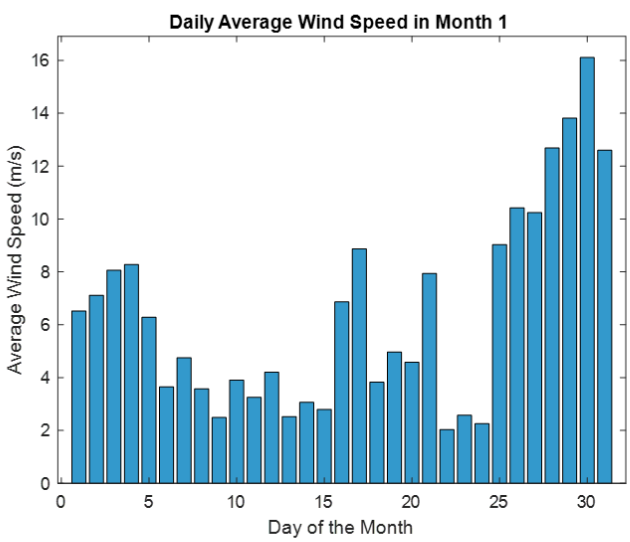

Fig. 5. Bar chart of daily average wind speed at 80 m height in January.

TABLE IV. MLE-ESTIMATED PARAMETERS FOR DISTRIBUTION MODELS AT 80 M HEIGHT

Distribution	MLE estimates of the parameters
Weibull	Scale (c): 7.1954 Shape (k):1.6042
Rayleigh	Scale (σ): 5.4201
Lognormal	Log mean (μ): 1.6226 Log std (σ): 0.7405
GEV	Shape (ξ): 0.1542 Scale (σ): 2.8992 Location (μ): 4.2820
Normal	Scale (σ): 6.4273 Location (μ): 4.1795

2) Model Goodness-of-Fit Evaluation

Statistical indicators, including RMSE, MAE, R^2, and AIC, were computed to evaluate the goodness of fit for the probability distributions. The results are presented in Table V.

TABLE V. STATISTICAL EVALUATION OF WIND SPEED AND WIND POWER BASED ON PROBABILITY DISTRIBUTIONS

Distribution	RMSE	MAE	R²	AIC
Weibull	0.0096	0.0087	0.9205	-10194.6123
Rayleigh	0.0197	0.0168	0.6661	-9128.7826
Lognormal	0.0131	0.0096	0.8522	-9732.9573
GEV	0.0098	0.009	0.9166	-10157.1698
Normal	0.0228	0.0198	0.5512	-8906.6644

The Weibull distribution demonstrates the best performance, with values of RMSE, MAE, R^2, and AIC at 0.0096, 0.0087, 0.9205, and –10194.61, respectively, indicating low error and a high degree of agreement with the empirical data. The GEV distribution ranks second, with comparable metrics but a higher AIC (–10157.17), which reduces its relative priority. The Lognormal distribution exhibits moderate accuracy (R^2 = 0.8522), while the Rayleigh (R^2 = 0.6661) and Normal (R^2 = 0.5512) distributions show the poorest fit, with significantly higher error values and AIC scores. Accordingly, the Weibull distribution is identified as the most appropriate model for simulating wind speed in this dataset.

3) Comparison of Distribution PDFs

Figure 6 illustrates the probability distribution of actual wind speed (blue histogram) along with the theoretical distribution curves for Weibull, Rayleigh, Lognormal, GEV, and Normal. It is evident that the Weibull distribution (blue curve) closely matches the empirical data across the peak and both tails of the distribution, demonstrating its strong capability in modeling wind speed. The GEV (purple curve) and Lognormal (yellow curve) distributions also show relatively good alignment, particularly at the peak and the right tail, with GEV providing a smoother and more stable fit. In contrast, the Rayleigh distribution (red curve) and the Normal distribution (green curve) exhibit noticeable deviations: the Rayleigh distribution skews leftward, underestimating high wind speeds, while the Normal distribution fails to capture the right-skewed nature of the actual data, resulting in significant error in the tail region.

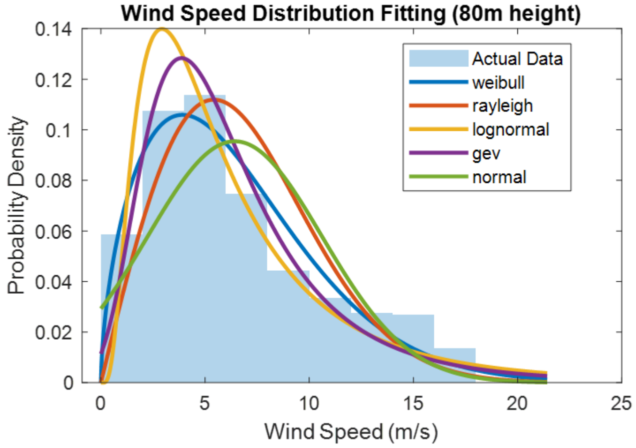

Fig. 6. Probability density functions of the wind speed distribution models.

In summary, the Weibull distribution is identified as the optimal model for wind speed at the study site, showing strong visual and statistical fit. While GEV and Lognormal also perform well—with GEV being close to Weibull—Rayleigh, and Normal yield the poorest results, with higher errors and AIC, making them unsuitable for realistic wind speed modeling.

4) Simulation of Wind Speed and Corresponding Output Power for the Following Day

Following the evaluation of various probability distributions based on the collected data, the Weibull distribution is identified as the most appropriate model for simulating wind speed in the study area, according to statistical criteria. To predict wind speed for the next day, random data samples are generated from the Weibull distribution using parameters estimated via the MLE method. This approach preserves the probabilistic characteristics of the original dataset and enables system evaluation under stochastic conditions.

Hourly wind speed data are generated, showing significant variability, which reflects the inherent randomness and temporal fluctuation of wind energy sources. Based on these wind speeds, the corresponding power output is calculated according to the wind turbine's power curve using (6). As depicted in Figure 7, a strong correlation exists between the two variables: as wind speed increases, the generated power also rises, peaking near 3 MW during hours with strong wind. Conversely, power output decreases to near zero during low-wind periods. This confirms the validity of using probability distributions for simulation purposes and demonstrates the method's applicability in forecasting and operational planning of wind energy systems.

Despite the promising results, this study has several limitations. The probabilistic analysis of wind speed and solar irradiance is based on a limited dataset from a single location. Additionally, the modeling framework relies on a predefined set of probability distributions and statistical indicators, which, although widely used, may not capture the full range of distributional behaviors of renewable energy resources.

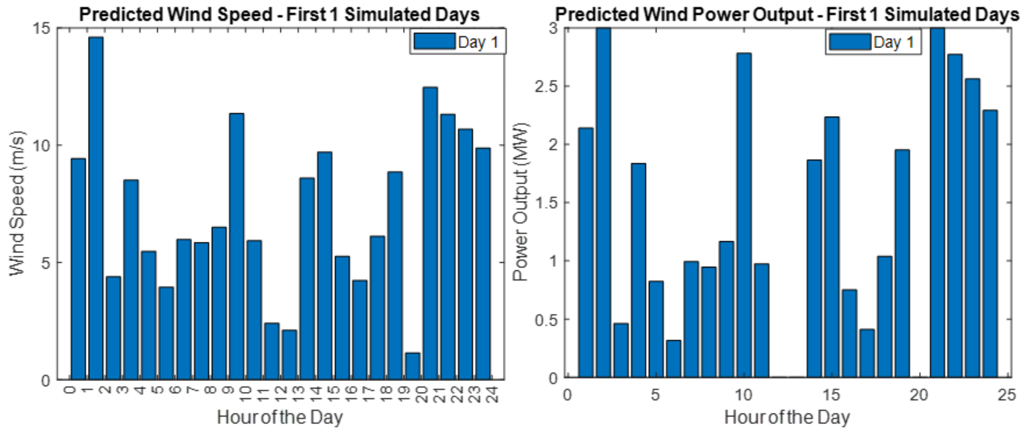

Fig. 7. Hourly forecast of wind speed and corresponding wind power output.

To achieve a more comprehensive representation of wind and solar energy variability, future studies could incorporate long-term, multi-year datasets. Additionally, exploring a broader range of probabilistic models and advanced analytical techniques, along with cross-regional validation, would significantly enhance the robustness and applicability of the proposed framework.

IV. CONCLUSIONS

This study developed a data-driven probabilistic framework for modeling solar irradiance and wind speed using real meteorological datasets and Maximum Likelihood Estimation (MLE). Model performance was rigorously assessed through multiple goodness-of-fit criteria, including Root Mean Square Error (RMSE), MAE, R^2, and Akaike Information Criterion (AIC), enabling objective and reliable selection of probability distributions.

For solar irradiance, the Beta distribution demonstrated the best overall performance, achieving the lowest RMSE (0.0004) and MAE (0.0003), the highest R^2 (0.9997), and the lowest AIC (−6802.79), effectively capturing the bounded and asymmetric nature of the data. For wind speed, the Weibull distribution provided the most accurate fit (RMSE = 0.0096, MAE = 0.0087, R^2 = 0.9205, AIC = −10194.61), outperforming the Rayleigh, Lognormal, GEV, and Normal models.

This study offers several key contributions to the probabilistic modeling of renewable energy resources. It presents a systematic comparison of multiple probability distributions using real high-resolution data, introduces a multi-criteria evaluation framework for robust model selection, and generates realistic renewable energy scenarios based on the optimal distributions. These findings provide a reliable foundation for uncertainty characterization and support enhanced forecasting, planning, and optimal operation of renewable energy systems and Microgrids (MG).

DECLARATION OF COMPETING INTERESTS

The authors declare no competing interests.

ACKNOWLEDGMENT

This study belongs to the project with grant No: T2025-204 funded by the Ho Chi Minh City University of Technology and Engineering, Vietnam.

DATA AVAILABILITY

Solar radiation and wind speed datasets are available at [11, 13] and further discussed in [12, 14], respectively.

REFERENCES

[1] M. M. R. Ahmed *et al*., "Mitigating Uncertainty Problems of Renewable Energy Resources Through Efficient Integration of Hybrid Solar PV/Wind Systems Into Power Networks," *IEEE Access*, vol. 12, pp. 30311–30328, 2024, https://doi.org/10.1109/ACCESS.2024.3370163.

[2] A. R. Krishna, A. V. Kumar, A. G. Krushna, J. Shanmugapriyan, and C. Keertana, "The Study of Solar and Wind Power Systems under Different Weather Conditions," in *Proceeding of the International Conference on Sustainable Green Energy Technologies* (*ICSGET*), Hyderabad, India, June 14–15, 2024, vol. 547, 2024, Art. no. 03009, https://doi.org/10.1051/e3sconf/202454703009.

[3] E. Gardashov and R. Gardashov, "On the derivation of an analytical expression for wind power probability distribution function and capacity factor of turbine," *International Journal of Sustainable Energy*, vol. 43, no. 1, Dec. 2024, Art. no. 2390447, https://doi.org/10.1080/14786451.2024.2390447.

[4] S.-D. Kwon, "Uncertainty analysis of wind energy potential assessment," *Applied Energy*, vol. 87, no. 3, pp. 856–865, Mar. 2010, https://doi.org/10.1016/j.apenergy.2009.08.038.

[5] R. Xezile and N. Mbuli, "Review of Application of Probability Distribution Functions in Assessment of Variable Load, PV Solar and Wind," in *5th International Conference on Electrical, Computer, Communications and Mechatronics Engineering* (*ICECCME*), Zanzibar, Tanzania, Oct. 16–-19, 2025, pp. 1–6, https://doi.org/10.1109/ICECCME64568.2025.11277431.

[6] T. B. M. J. Ouarda *et al*., "Probability distributions of wind speed in the UAE," *Energy Conversion and Management*, vol. 93, pp. 414–434, Mar. 2015, https://doi.org/10.1016/j.enconman.2015.01.036.

[7] B. El Kihel, N. E. El Kadri Elyamani, and A. Chillali, "Wind energy potential assessment using the Weibull distribution method for future energy self-sufficiency," *Scientific African*, vol. 26, Dec. 2024, Art. no. e02482, https://doi.org/10.1016/j.sciaf.2024.e02482.

[8] P. Lencastre, A. Yazidi, and P. G. Lind, "Modeling Wind-Speed Statistics beyond the Weibull Distribution," *Energies*, vol. 17, no. 11, 2024, Art. no. 2621, https://doi.org/10.3390/en17112621.

[9] Z. Cai and S. Zuo, "Optimal System Analysis for Hybrid Wind-Solar-Pumped Storage Systems under Renewable Output Uncertainty,"

International Journal of Frontiers in Engineering Technology, vol. 6, no. 4, 2024, https://doi.org/10.25236/IJFET.2024.060419.

[10] S. Patty and T. Malakar, "Performance analysis of machine learning based prediction models in assessing optimal operation of microgrid under uncertainty," *Results in Control and Optimization*, vol. 15, June 2024, Art. no. 100407, https://doi.org/10.1016/j.rico.2024.100407.

[11] "NSRDB: National Solar Radiation Database." National Laboratory of the Rockies (NLR). https://nsrdb.nlr.gov/.

[12] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The National Solar Radiation Data Base (NSRDB)," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 51–60, June 2018, https://doi.org/10.1016/j.rser.2018.03.003.

[13] "WRDB: Wind Resource Database." National Laboratory of the Rockies (NLR). https://wrdb.nlr.gov/.

[14] C. Draxl, A. Clifton, B.-M. Hodge, and J. McCaa, "The Wind Integration National Dataset (WIND) Toolkit," *Applied Energy*, vol. 151, pp. 355–366, Aug. 2015, https://doi.org/10.1016/j.apenergy.2015.03.121.

[15] Y. Kassem, H. Camur, and A. A. S. Mosbah, "Wind Resource Evaluation in Libya: A Comparative Study of Ten Numerical Methods for the Estimation of Weibull Parameters using Multiple Datasets," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13388–13397, Apr. 2024, https://doi.org/10.48084/etasr.6882.

[16] Z. Wang and W. Liu, "Wind energy potential assessment based on wind speed, its direction and power data," *Scientific Reports*, vol. 11, no. 1,Aug. 2021, Art. no. 16879, https://doi.org/10.1038/s41598-021-96376-7.

[17] R. B. Millar, *Maximum likelihood estimation and inference : with examples in R, SAS and ADMB*, 1st ed. Chichester, United Kingdom: John Wiley & Sons, Ltd, 2011.

[18] R. H. Norden, "A Survey of Maximum Likelihood Estimation," *International Statistical Review / Revue Internationale de Statistique*, vol. 40, no. 3, pp. 329–354, Dec. 1972, https://doi.org/10.2307/1402471.

[19] R. Mullen, L. Marshall, and B. McGlynn, "A Beta Regression Model for Improved Solar Radiation Predictions," *Journal of Applied Meteorology and Climatology*, vol. 52, no. 8, pp. 1923–1938, 2013, https://doi.org/10.1175/JAMC-D-12-038.1.

[20] K. Krishnamoorthy, *Handbook of Statistical Distributions with Applications*, 1st ed. Boca Raton, Florida, USA: Chapman and Hall/CRC, 2006.

[21] B. E. Smith and F. J. Merceret, "The Lognormal Distribution," *The College Mathematics Journal*, vol. 31, no. 4, pp. 259–261, 2000, https://doi.org/10.1080/07468342.2000.11974155.

[22] W. E. Stein and M. F. Keblis, "A new method to simulate the triangular distribution," *Mathematical and Computer Modelling*, vol. 49, no. 5–6, pp. 1143–1147, Mar. 2009, https://doi.org/10.1016/j.mcm.2008.06.013.

[23] P. Megantoro *et al.*, "Modeling the uncertainties and active power generation of wind-solar energy with data acquisition from telemetry weather measurement," *Results in Engineering*, vol. 25, Mar. 2025, Art. no. 104392, https://doi.org/10.1016/j.rineng.2025.104392.

[24] P. Wais, "A review of Weibull functions in wind sector," *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 1099–1107, Apr. 2017, https://doi.org/10.1016/j.rser.2016.12.014.

[25] F. Homa, M. Khetan, Mohd. Arshad, and P. Mishra, *Distribution theory: Principles and Applications*, 1st ed. Boca Raton, Florida, USA: CRC Press, 2023.

[26] J. A. R. Villaseñor, *Frequency analyses of natural extreme events: A Spreadsheets Approach*, 1st ed. Cham, Switzerland: Springer Nature, 2021.

[27] M. Ebeed and Shady H. E. Abdel Aleem, "Overview of uncertainties in modern power systems: uncertainty models and methods," in *Uncertainties in Modern Power Systems*, Ahmed F. Zobaa and Shady H. E. Abdel Aleem, Eds., 1st ed. London, United Kingdom: Academic Press, 2021, ch. 1, pp. 1–34, https://doi.org/10.1016/B978-0-12-820491-7.00001-3.

[28] "Siemens SWT-3.0-108 - Manufacturers and turbines" The Wind Power, https://www.thewindpower.net/turbine_en_920_siemens_swt-3.0-108.php.

[29] Canadian Solar, "TOPBiHiKu7 BIFACIAL TOPCON," CS7L-590TB-AG datasheet, Oct. 2022.