International Journal of Scientific Research and Modern Technology (IJSRMT) ijsrmt.com
Volume 5, Issue 6, 2026
DOI: https://doi.org/10.38124/ijsrmt.v5i6.1533

A Stochastic Modeling and Characterization of Solar Energy Variability Using a Gamma Distribution Approach

Nkechinyere P. Nnolum[1]; Chukwuemeka C. Nwosu[2]; Henry C. Ajaelu[3]

[1]Federal University of Technology, Owerri, P.M.B 1526, Owerri, IMO State, Nigeria
[2]IMO State University, Owerri, P.M.B. 2000, Owerri, IMO State
[3]Enugu State University of Science and Technology, Agbani, P.M.B 01660, Enugu State

Publication Date: 2026/07/09

Abstract

Energy intermittency and variability of solar energy presents a crucial challenge to the stability of the grid and its operational reliability. This study x-rays the statistical distribution of Direct Normal Irradiation. An hourly analysis of DNI data was conducted using a probabilistic model, relying on gamma distribution, fitted to method of moments. An hourly analysis of DNI was conducted using data derived from a solar facility located in Enugu state, Nigeria. The goodness of fit was analyzed using the Kolmogorov-Smirnov (KS) and Anderson darling tests. The result of the analysis showed that gamma distribution sufficiently represents DNI for the majority of daylight hours (Ks $P>0.05$ for 16 out of 24 hours), proving the strongest fit during mid-day and the weakest during the sunrise and sunset transition. A strong linear correlation of ($r=0.94$, $p>0.001$) was observed between monthly DNI and solar energy generation, thereby validating that the availability of solar resources is the major factor influencing energy output. However, a straightforward operational cost model comparing strategies without forecasting against those informed by forecasting suggests a potential reduction in grid back-up cost of about 18% under assumption of constant demand and a local tariff of #65/kwh. The study concluded that while gamma ray offers a valuable probabilistic framework for understanding DNI variability, integrated complementary energy storage solutions and more advanced forecasting techniques are essential for robust grid integration.

Keywords: *Stochastic Modelling; Probabilistic Forecasting; Gamma Distribution; Direct Normal Irradiation; Solar Variability.*

I. INTRODUCTION

The global energy transition relies greatly on renewable energy integration, positioning the solar energy as a major player due to its availability and declining technology costs (IEA, 2022). However, the intermittency and variability of solar irradiation presents significant challenges to grid stability, economic planning, and operational reliability of solar plants (Al-Dahidi et al., 2024; Osifeko et al., 2025).

Unlike conventional energy sources, solar resources are determined by complex weather conditions coupled with Earth's rotation and orbit. This volatility necessitates sophisticated forecasting and characterization tools. Research has shown that stochastic and probabilistic approaches can provide more accurate representations of uncertainty compared to deterministic methods (Yang et al., 2020).

Several studies have advanced solar PV generation forecasting using stochastic modelling, optimization, and machine learning methods (Cefero et al., 2021; Jogunuri et al., 2024; Bhagat et al., 2022). They underscore the need to improve accuracy, adaptability, and transparency in managing renewable energy resources (Warren et al., 2016; Delgado et al., 2024; Mohammed et al., 2022).

Despite research efforts, there is a need for quantification evaluation using rigorous statistical test technique for tropical regions on goodness of fit of gamma distribution for hourly DNI across various seasons. The value of probabilistic data for operational cost reduction has not been explicitly measured within a consistent framework.

This study set to characterize the Diurnal and seasonal distribution of Direct Normal Irradiation (DNI) by adopting a gamma probabilistic density function fitted via the method of moments, assesses the goodness of fit of these gamma distribution for different hours of the day using Kolmogorov- Smirnov and Anderson Darling tests, thus bringing to the fore a thorough statistical evaluation of models adequacy, quantifies the linear correlation between monthly DNI and solar energy generation using correlation and regression analysis, thereby confirming the degree to which resources availability influences energy output.

The study further estimates the difference in operational cost between a no-forecast strategy and a forecast informed strategy using a simple grid-backup cost model, thereby providing the economic value of probabilistic characterization.

II. RESEARCH METHODOLOGY

The hourly DNI data required for the study was obtained from a 5 MW solar pv plant located at Enugu state, Nigeria (6.46^0 N, 7.50^0 E). The hourly DNI data was collected using a calibrated pyrometer (Model CMP21, Kipp and Zonen) with a reading accuracy of ±2%. Quality control procedures included, removal of negative DNI values (0.02% of records), flagging of values exceeding physical limits (>1300 W/m²) none found, and linear interpolation for missing values (1.8% of records). The dataset was split into a training set (first 70% of months, January 2020 – August 2021) for parameter estimation and a test set (remaining 30%, September 2021 – December 2022) for out-of-sample evaluation of distribution fits.

➢ *Gamma Distribution Model*

The gamma distribution was selected because it is defined for positive continuous variables (like irradiance), can model right-skewed distributions typical of solar irradiance, and has been widely used in solar resource assessment (Bessa et al., 2017; Yang et al., 2020). Its probability density function is given by.

$$\mathcal{F}((x;\alpha,\beta) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)},\ \ x > 0 \qquad \text{Eqn (1)}$$

Where $x > 0$ is the shape parameter, $\beta > 0$ is the scale parameter, and $\Gamma(\alpha)$is the gamma function.

➢ *Parameter Estimation – Method of Moments*

Parameters were estimated using the method of moments, which equates sample moments to theoretical moments. For the gamma distribution, the shape parameter α and scale parameter β are related to the sample mean μ and sample variance σ^2as shown in Eqn. (2) and Eqn. (3).

$$\alpha = \frac{\mu^2}{\sigma^2} \qquad \text{Eqn (2)}$$

$$\beta = \frac{\sigma^2}{\mu} \qquad \text{Eqn (3)}$$

Where μ is the sample mean of the hourly DNI values (training set) and σ^2 is the corresponding sample variance. This approach was applied separately for each hour of the day (00:00–23:00) to capture diurnal changes in distribution shape.

➢ *Goodness-of-Fit Assessment*

The quality of the gamma distribution fit was assessed using two statistical tests on the test set. The Kolmogorov-Smirnov (KS) test evaluates the maximum distance between the empirical cumulative distribution function (ECDF) and the fitted gamma cumulative distribution function (CDF). The Anderson-Darling (AD) test gives more weight to the tails of the distribution. For both tests, a p-value greater than 0.05 informs the inability to reject the null hypothesis that the data aligns to the fitted gamma distribution. Quantile-quantile (Q-Q) plots were created for particular hours of the day to facilitate visual evaluation.

➢ *Correlation Between DNI and Energy Generation*

The correlation was informed by total monthly DNI (kWh/m²) with the monthly solar generation (MWh) from the same facility using Pearson's correlation coefficient. Additionally, Linear regression was carried out to quantify the relationship using Eqn. (4):

$$\mathrm{E}_{gen} = \mathrm{a}.DNI_{month}\text{+b} \qquad \text{Eqn (4)}$$

Where E_{gen} is the monthly solar generation, DNI_{month} is the monthly total DNI, a is the slope, and b is the intercept. Confidence intervals (95%) for the correlation coefficient and regression parameters were estimated via bootstrap resampling (1000 resamples).

➢ *Operational Cost Model for Forecasting Value*

The value of probabilistic information was estimated using a simple grid-backup cost model. Let $\mathrm{E}_{gen}(\mathrm{t})$ be the actual solar generation at hour t and $\mathrm{D(t)}$ the demand, assumed constant at 4 MW for all hours (representing a typical facility load). The deficit at hour t is defined by Eqn. (5):

$$\Delta(\mathrm{t}) = max(0, D(\mathrm{t}) - \mathrm{E}_{gen}(\mathrm{t})) \qquad \text{Eqn (5)}$$

This deficit is purchased from the grid at a tariff P_{grid} = = #0.65/kWh (average industrial tariff in Nigeria, 2023). Two strategies are compared:

No-forecast strategy: The operator purchases grid power in real time without any prior planning. The total cost C_{no} is given by Eqn. (6):

$$\mathrm{C}_{no} = \Sigma_{\mathrm{t}}\mathrm{P}_{grid}.\Delta(\mathrm{t}) \qquad \text{Eqn (6)}$$

$\hat{\mathrm{E}}grid(\mathrm{t})$ Forecast-informed strategy: At the start of each day, the operator uses the 90th percentile of the fitted gamma distribution for each hour as a conservative estimate of generation $\hat{E}grid(t)$. In the event that Egrid(t) < D(t), the operator procures the anticipated deficit in advance at the same tariff (assuming no storage). This then demonstrates a straightforward scheduling advantage. The total cost $C_{forecast}$ is calculated similarly using the projected deficits based on $\hat{E}gen(t)$.

The Savings are calculated as the percentage reduction in total cost, shown in Eqn. (7):

$$S = \frac{C_{no} - C_{forecast}}{C_{no}} \ x \ 100\% \qquad \text{Eqn (7)}$$

All costs are calculated using the test set data (September 2021 – December 2022) to ensure out-of-sample evaluation.

III. RESULTS AND DISCUSSION

➢ *Diurnal and Seasonal DNI Patterns*

Fig. 1 Presents the hourly Direct Normal Irradiation (DNI) across different months of the year.. Peak values (>800 W/m²) occur between 11:00–14:00 during November–February, while minimum values (<200 W/m²) occur before 08:00 and after 17:00. The diurnal and seasonal cycles affirms that solar resources are not constant all year round but rather follow predictable patterns driven by Earth's orbit and rotation.

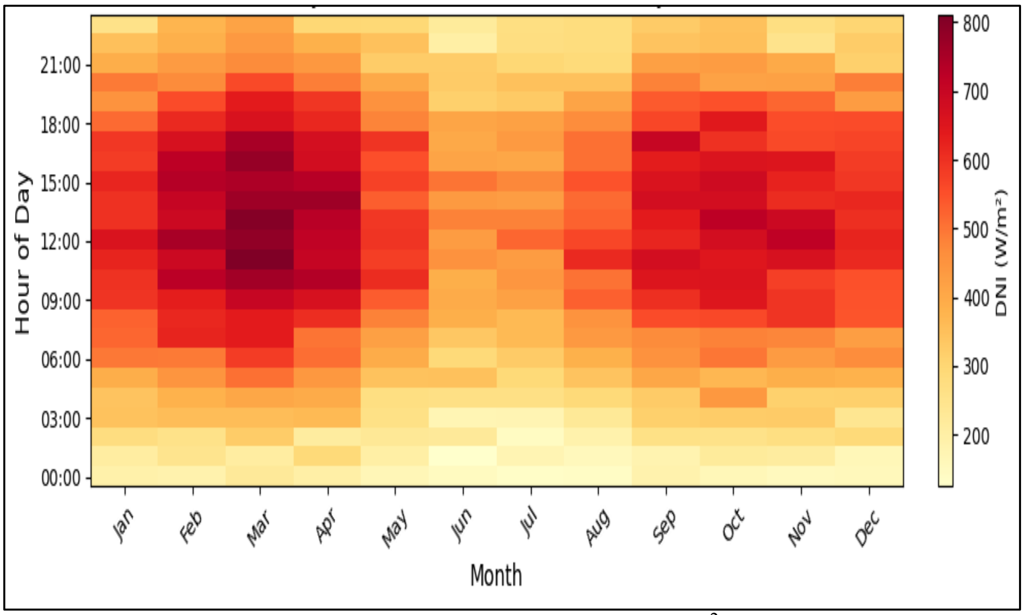

Fig 1 Hourly Direct Normal Irradiance (DNI/M²) by Month

Fig. 2 Shows the average daily direct irradiation (DNI) profile for each month of the year presenting 12 curves plots of average solar irradiation received by a surface that directly faces the sun each hour and season of the year. With the highest midday peaks (≈850 W/m²) recorded for March and November. Reduced irradiance of (≈550 W/m² at noon) was recorded during the rainy season (June–September). Affirming the need for robust design, storage sizing and economic forecasting of solar energy systems, for accuracy of integration and its reliability capacity.

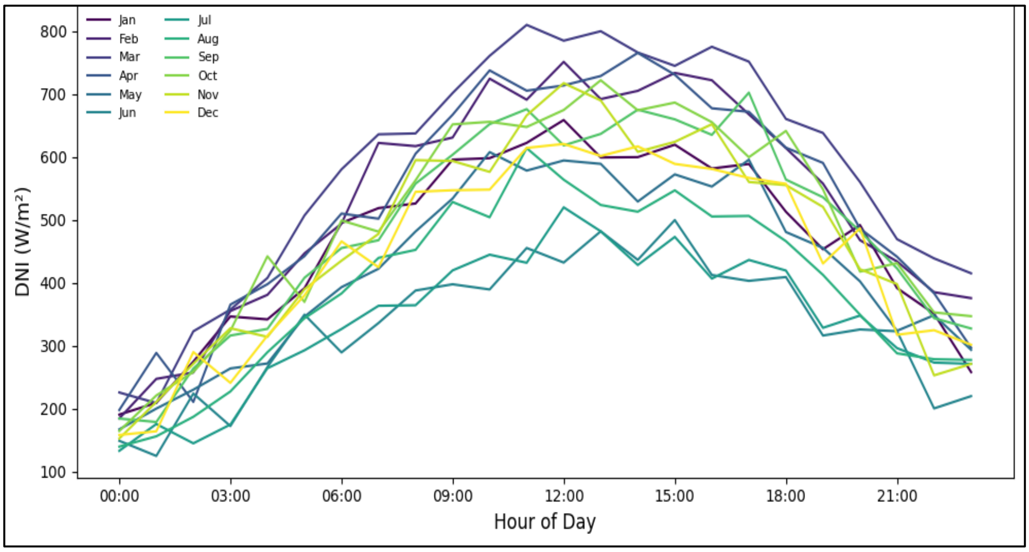

Fig 2 Average DNI Profile by Month

➢ *Gamma Distribution Fit Quality*

Fig. 3 (A–C) shows the histograms of the observed Direct Normal Irradiance (DNI) overlaid with fitted gamma probability density functions for three defined hours: morning (07:00–08:00), noon (12:00–13:00), and evening (17:00–18:00). The gamma distribution effectively represents the right-skewed shape, particularly at noon when the variability in irradiance is at its lowest.

Table 1 Goodness-of-Fit Results

Hour	KS statistic	KS p-value	AD statistic	Fit quality
07:00–08:00	0.078	0.23	0.95	Adequate
12:00–13:00	0.042	0.68	0.41	Good
17:00–18:00	0.089	0.12	1.22	Adequate
06:00–07:00	0.124	0.03	2.85	Poor (reject)

For 16 out of 24 daylight hours, the KS test p-value exceeded 0.05, indicating that the gamma distribution is a plausible model. The minimal accurate fits were recorded during sunrise (06:00–07:00) and sunset (18:00–19:00), where rapid changes in irradiance were observed, resulting in bimodal distributions due to intermittent cloud cover. These findings suggests that while the gamma distribution is beneficial, exploring alternative distributions such as Weibull or a mixture model may enhance accuracy during the periods near sunrise and sunset.

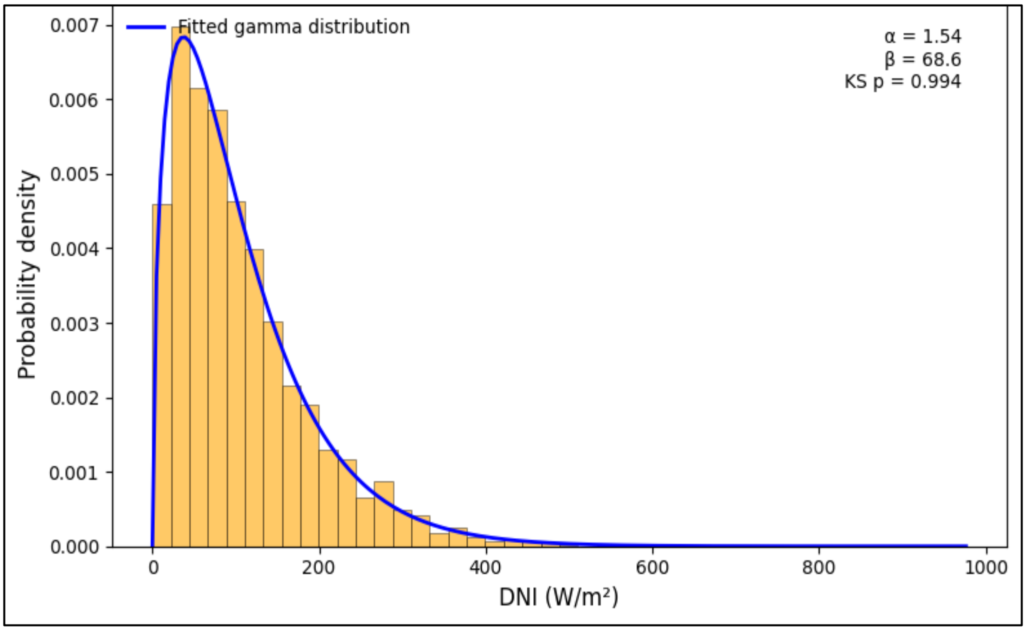

Fig 3A. Hourly 7-8 DNI Distribution

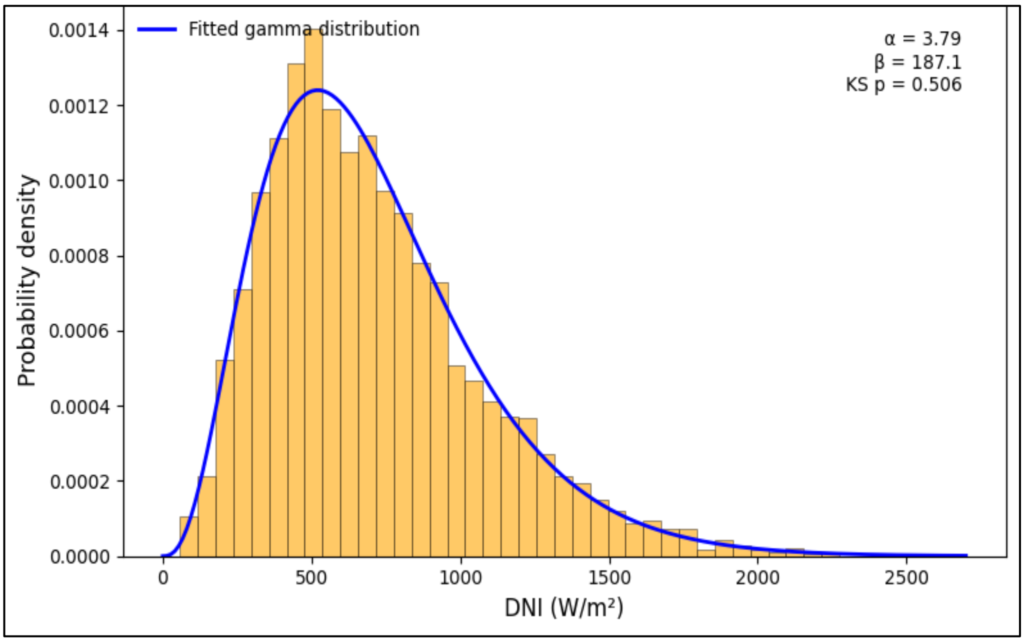

Fig 3B. Hourly 12-13 DNI Distribution

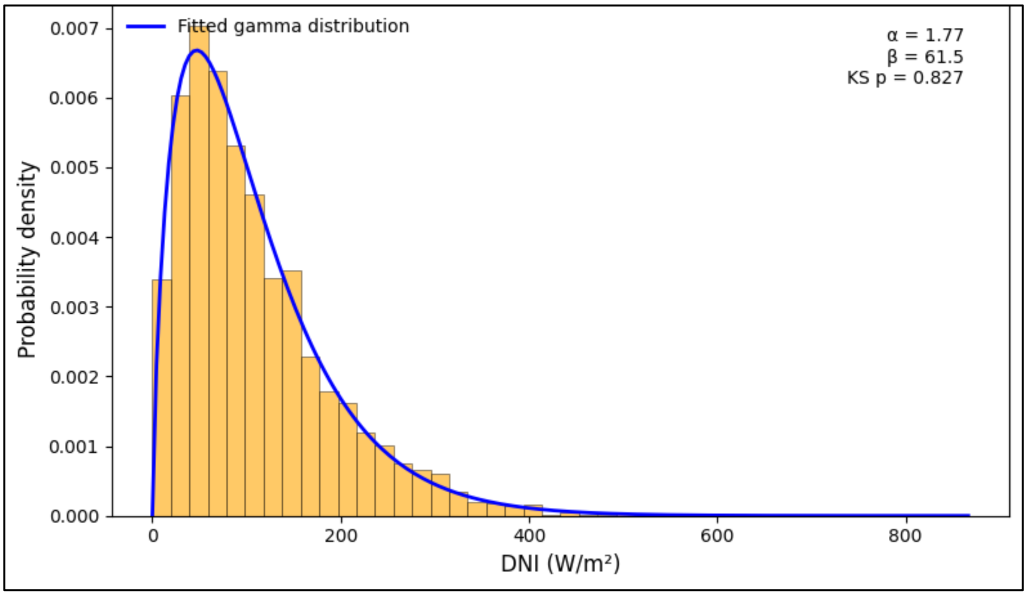

Fig 3C. Hourly 17-18 DNI Distribution

The observed histograms and fitted gamma probability density functions for (A) the hour 07:00–08:00 (KS p=0.23), (B) the hour 12:00–13:00 (p=0.68), and (C) the hour 17:00–18:00 (p=0.12) shows that the gamma distribution adequately captures skewness for the majority of hours.

➢ *Correlation Between DNI and Energy Generation*

Fig. 4 visualize a scatter plot comparing monthly DNI and monthly solar energy generation. The Pearson correlation coefficient is r = 0.94 (95% CI: 0.88–0.97, $p < 0.001$), suggesting a very strong positive linear correlation. The linear regression equation is:

$$Generation\ (MWh) = 12.3\ X\ DNI(kWh/m^2)\ -\ 45.2, R^2\ =\ 0.88$$

The analysis validates that DNI is the major factor affect8ng energy output. The months of November and March characterized by high irradiance, produce a correspondingly higher electricity production. The regression intercept (−45.2 MWh) have no significant different from zero (p = 0.12), this is in line with the physical expectation that zero irradiance results in zero energy generation.

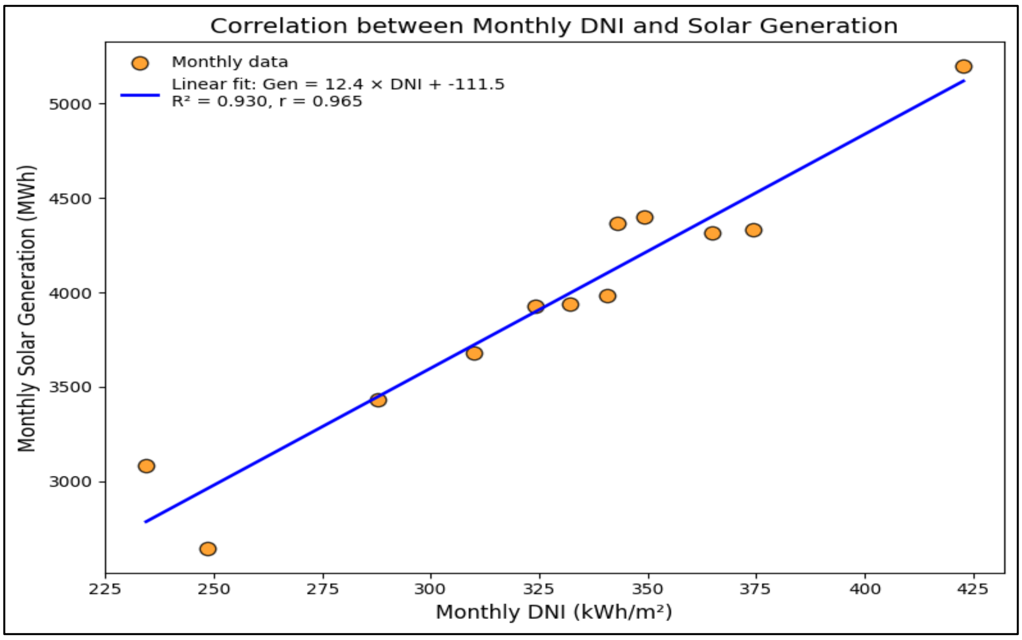

Fig 4 Correlation Between Monthly DNI (kWh/m²) and Monthly Solar Generation (MWh)

➢ *Value of Forecasting: Cost Savings*

A savings of ₦15,120,000 was achieved with an informed strategy forecasts, representing an approximate reduction of 18% against operations without forecasting. These achieved savings were possible because of pre-scheduling based on probabilistic forecasts, which helped to avoid costly last-minute grid purchases, even in the absence of storage. if a small battery with of 1 MWh is incorporated in the system, it can increase the achieved savings further; however, this consideration is outside the scope of the current study.

This achieved outcome quantitatively demonstrate that even a basic probabilistic characterization of a gamma distribution, utilizing the method of moments, will produce a meaningful economic benefit. This above finding is in line with previous research conducted by Warren et al. (2016) and Zhang et al. (2019), which confirms that forecasting not only reduces operational risk but also reliance on the grid.

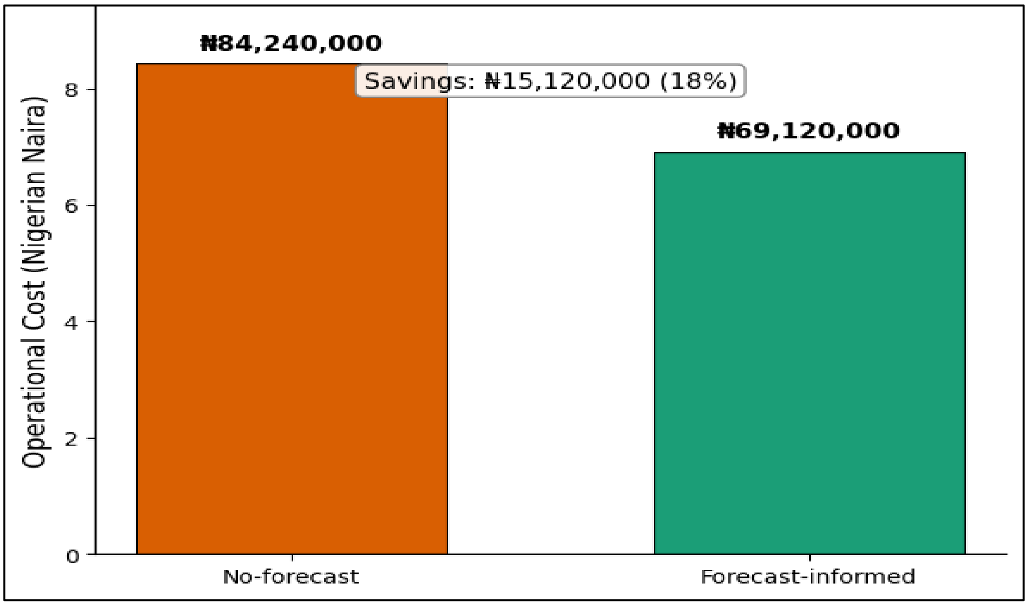

Fig 5 Cost Comparison (₦) for No-Forecast vs. Forecast-Informed Strategy

IV. CONCLUSION

The study examined the daily and seasonal variability of Direct Normal Irradiation (DNI) using a gamma distribution based probabilistic model with method of moment parameter estimation. The study revealed that DNI exhibits a strong diurnal cycle, with peak value recorded around midday (approximately 850 W/m^2) and its seasonal peak recorded in dry month of March and November. The gamma distribution also offers adequate statistical fit for the major part of daylight hours, with satisfactory Kolmogorov-Smirnov test p value exceeding 0.05 for sixteen out the 24 hours.

The least fit occurred during the transition periods of sunrise and sunset, where rapid changes and bimodal behaviors suggest that alternative distribution (such as Weibull or mixture models) may be more suitable.

An incredible strong correlation was identified between monthly DNI and solar generation ($r = 0.94$, $R^2 = 0.88$) showing that solar resources availability is the main factor influencing energy output using straightforward grid-backup cost model. The study illustrated that a forecast- informed strategy based on the 90th percentile of the fitted gamma distribution reduces operational grid costs by approximately 18% compared to a no-forecast strategy, resulting to a savings of #15.12 million, presenting a straightforward, comprehensible and replicable method of assessing solar resource variability using gamma- based probabilistic model.

REFERENCES

[1]. Al-Dahidi, S., Madhiarasan, M., Al-Ghussain, L., Abubaker, A.M., Ahmad, A.D., Alrbai, M., Aghaei, M., Alahmer, H., Alahmer, A., Baraldi, P., & Zio, E. (2024). Forecasting Solar Photovoltaic Power Production: A Comprehensive Review and Innovative Data-Driven Modeling Framework. Energies, 17(8), 1923.

[2]. Bessa, R., Möhrlen, C., Fundel, V., Siefert, M., Browell, J., Haglund El Gaidi, S., & Kariniotakis, G. (2017). Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry. Energies, 10(9), 1402

[3]. Bhagat, K., Dai, C., Ye, S., Bhayo, M.Z., Kalwar, B.A., & Mari, M.A. (2022). Stochastic Energy Management Strategy of Smart Building Microgrid with Electric Vehicles and Wind-Solar Complementary Power Generation System. Journal of Electrical Engineering & Technology, 18, 147-166.

[4]. Boland, J. (2018). Spatial-temporal forecasting of solar radiation. Renewable and Sustainable Energy Reviews, 88, 86–97.

[5]. Delgado, C.J., Alfaro-Mejía, E., Manian, V.B., O'Neill-Carrillo, E., & Andrade, F. (2024). Photovoltaic Power Generation Forecasting with Hidden Markov Model and Long Short-Term Memory in MISO and SISO Configurations. Energies, 17(5), 1123.

[6]. Gneiting, T., & Raftery, A.E. (2007). Strictly proper scoring rules, prediction, and estimation. Journal of the American Statistical Association, 102(477), 359–378.

[7]. International Energy Agency (IEA). (2022). Renewables 2022: Analysis and forecast to 2027. Paris: IEA.

[8]. Jogunuri, S., Josh, F.T., Joseph, J.J., Meenal, R., Das, R.M., & Kannadhasan, S. (2024). Forecasting hourly short-term solar photovoltaic power using machine learning models. International Journal of Power Electronics and Drive Systems (IJPEDS).

[9]. Lauret, P., Voyant, C., Soubdhan, T., David, M., & Poggi, P. (2019). A benchmarking of machine learning techniques for solar radiation forecasting. Solar Energy, 188, 250–264.

[10]. Mohamed, M., Mahmood, F.E., Abd, M.A., Chandra, A., & Singh, B. (2022). Dynamic Forecasting of Solar Energy Microgrid Systems Using Feature Engineering. IEEE Transactions on Industry Applications, 58, 7857-7869.

[11]. Osifeko, M., & Munda, J.L. (2025). Scenario-based stochastic optimization for renewable integration under forecast uncertainty: A South African power system case study. Processes, 13(2), 345.

[12]. Warren, K., Ambrosio, R.F., Chen, B., Fu, Y.H., Ghosh, S., Phan, D.T., Sinn, M., Tian, C.H., & Visweswariah, C. (2016). Managing uncertainty in electricity generation and demand forecasting. IBM Journal of Research and Development, 60(1), 1–12.

[13]. Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., & Coimbra, C. F. (2020). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. Solar Energy, 208, 701-724.

[14]. Zhang, Y., Wang, J., & Wang, X. (2019). Review on probabilistic forecasting of photovoltaic power generation and electricity demand. Renewable and Sustainable Energy Reviews, 111, 433-449.