ISBN: 978-628-96613-3-0. ISSN: 2414-6390. Digital Object Identifier: https://dx.doi.org/10.18687/LACCEI2026.1.1.1720

Daily Energy Production Variance Analysis and Statistical Distribution Modeling for Small-Scale Solar Photovoltaic Systems

Elmer Arellanos-Tafur[1]
[1]School of Engineering, Universidad Continental, Peru, earellanos@continental.edu.pe

***Abstract*—This research presents a detailed statistical analysis of daily energy production variance for the Carhuaquero Photovoltaic Solar Plant, a 565.25 kW small-scale solar installation operated by KONDU SAC in Peru. Using a complete annual dataset spanning from May 17, 2024, to May 16, 2025, this study examines the temporal variability and statistical distribution characteristics of photovoltaic energy generation. The analysis encompasses 365 consecutive daily observations, providing robust insights into production patterns and operational efficiency metrics. Statistical modeling techniques were applied to characterize the probability distributions governing daily energy output, with particular emphasis on variance analysis and distribution fitting methodologies. The results indicate that daily energy production exhibits considerable variability, with values ranging from 0.05 to 6.10 MWh/day and a coefficient of variation of 41.2%. Seasonal patterns are identified through monthly aggregation analysis, revealing distinct performance variations associated with meteorological conditions typical of the Andean region. Distribution modeling shows that energy production follows a right-skewed, leptokurtic distribution, best approximated by a Gamma model, reflecting the alternation between overcast and clear-sky conditions. These findings contribute to an improved understanding of small-scale photovoltaic system performance characteristics and provide valuable inputs for energy forecasting models and grid integration strategies. Furthermore, the results support enhanced operational planning and maintenance scheduling for similar solar installations under comparable geographical and climatic conditions.**

***Keywords*—Solar Photovoltaic Systems, Energy Production Variance, Statistical Distribution Modeling, Autocorrelation Analysis, Renewable Energy Forecasting.**

I. Introduction

Photovoltaic (PV) technology has established itself as a fundamental component of sustainable energy infrastructure worldwide, supported by sustained reductions in capital costs and progressive improvements in module conversion efficiency. Within the broader PV deployment landscape, grid-connected systems with installed capacities below 1 MW constitute a strategically significant segment of distributed generation. Their widespread adoption across diverse geographic and climatic contexts demands rigorous performance characterization to support operational planning, grid management, and long-term investment viability [1], [2].

The stochastic nature of solar resource availability introduces inherent uncertainty in daily energy output, governed by the combined influence of incident irradiance, atmospheric attenuation, seasonal dynamics, and system-level factors. Empirical studies have quantified the adverse effects of environmental degradation mechanisms on PV performance: soiling reduces module efficiency by up to 11.86% and can suppress energy yield by 8.8%, while partial shading—depending on array configuration and bypass diode architecture—may curtail instantaneous power output by as much as 92.6% [3]. Array topology also modulates system resilience; total-cross-tied (TCT) configurations have demonstrated superior performance under non-uniform illumination conditions in small-scale deployments [4]. These interacting factors underscore the need for robust statistical frameworks capable of quantifying production reliability and supporting probabilistic decision-making [5], [6].

Despite a substantial body of research addressing large-scale utility PV installations ($> 1\,$MW) in well-characterized arid and semi-arid climates, a significant knowledge gap persists regarding the statistical production behavior of small-scale systems operating under tropical mountain conditions [7]. The distinct meteorological dynamics of such environments—characterized by orographic cloud formation, pronounced seasonal precipitation regimes, and variable atmospheric humidity—generate production variability patterns that cannot be adequately represented by models calibrated on desert-region installations [8], [9]. The Andean region of Peru exemplifies this underrepresented climatic context, where the interaction between elevated altitude, tropical irradiance levels, and rainy-season cloud dynamics creates conditions with limited precedent in the indexed PV performance literature [10].

To address this deficiency, the present study delivers the first comprehensive annual statistical characterization of daily energy production variability for a grid-connected small-scale PV system integrated with hydroelectric infrastructure in the Peruvian Andes. The Carhuaquero Photovoltaic Solar Plant, operated by KONDU SAC, provides an exemplary case study in hybrid renewable generation, demonstrating the operational complementarity between solar and hydroelectric resources within a shared infrastructure perimeter [11], [12].

The remainder of this paper is organized as follows. Section II presents the theoretical framework, encompassing a description of the study site and the statistical methods employed.

.

Section III details the data acquisition protocol, preprocessing procedures, and statistical modeling approach. Section IV reports descriptive statistics, seasonal variability patterns, and distribution fitting results. Section V interprets the findings in relation to grid integration challenges, operational strategies, and existing benchmarks, and identifies the study limitations. Section VI summarizes the principal conclusions and outlines directions for future investigation.

II. Theoretical Framework

A. Overview of Carhuaquero Photovoltaic Solar Plant

The Carhuaquero Photovoltaic Solar Plant belongs to KONDU SAC, an emerging Peruvian energy company strategically supported by Kallpa Generación and Orazul Energy [13]. The company's flagship project, the Carhuaquero Photovoltaic Solar Plant, is located in the districts of Llama and Catache, within the provinces of Chota and Santa Cruz, in the operational area of the Carhuaquero Hydroelectric Plant [14]. The geographical context of the installation is illustrated in Fig. 1, while the main technical specifications of the system are summarized in Table I [15].

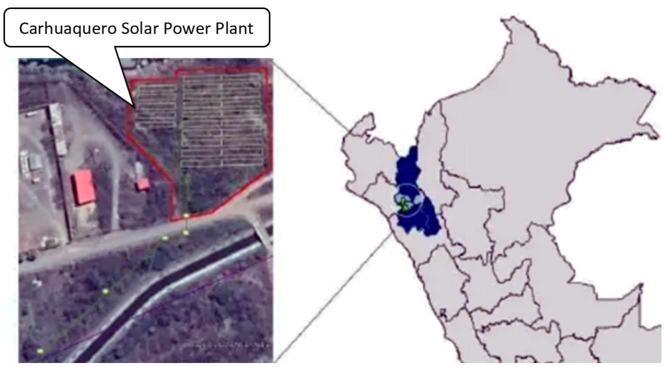

Fig. 1. Carhuaquero Photovoltaic Solar Plant Location

This project represents an innovative approach to leveraging existing infrastructure, with an installed capacity of 565.25 kW. The system consists of 950 photovoltaic modules rated at 595 W each, five 100 kVA string on-grid inverters, and a 630 kVA transformation center [16]. The physical arrangement of the photovoltaic modules is shown in Fig. 2, while a broader view of the solar installation, including its spatial distribution and surrounding environment, is presented in Fig. 3 [17].

The design highlights the complementarity between hydroelectric and solar generation within the same operational complex, reinforcing KONDU's role as a strategic actor in the diversification of renewable portfolios in Peru. Furthermore, the implementation of an underground interconnection line with existing infrastructure demonstrates a strategic approach to resource optimization and transmission cost reduction [18].

B. Gamma and Normal Distribution Modeling

Probabilistic modeling of photovoltaic energy production requires distributions capable of representing positively skewed, bounded data. The Gamma distribution, widely applied to

TABLE I
Technical Specifications Of Carhuaquero Solar Power Plant

Photovoltaic System	
System Type	Grid-Connected (On-Grid)
Inverters	5 Units, 100 kVA Each
Total Inverter Capacity	500 kVA
Inverter Model	Huawei SUN2000-100KTL
Transformation Center	10/0.48 kV, 630 kVA
PV Modules	
Number of Modules	950
Module Power	595 Wp
Technology	N-Type
Mounting Structure	
Type	Fixed Structure
Material	Aluminum
Tilt Angle	14°
DC Cabling	
Cable Section	6 mm^2
Voltage Rating	Up To 1500 Vdc
Installation Type	Outdoor (Non-Direct Burial)
Inverter System	
Standard Compliance	COES Technical Procedure No. 20 (PR-20)
Connection Type	String Inverter, Grid-Synchronized
Transformation Center	
Low Voltage Side	480 V, 60 Hz
Medium Voltage Level	10 kV
Main Transformer Power	630 kVA
Tap Changer	Off-Load Tap Changer
Transformation Ratio	10/0.48 kV

Fig. 2. Solar Panels of the Carhuaquero Photovoltaic Solar Plant

Fig. 3. 950 Photovoltaic Modules, each with a Nominal Power Rating of 595 W

energy yield characterization and renewable resource analysis, is defined as

$$f(x;\alpha,\beta)=\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}\qquad x>0 \tag{1}$$

where $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter, and $\Gamma(\cdot)$ denotes the Gamma function. This distribution accommodates right-skewed profiles characteristic of solar generation data, where low-production days associated with overcast conditions are frequent while high-production events form an extended right tail [19]. Its flexibility in capturing asymmetric structures makes it well-suited for Andean climatic regimes where cloud intermittency governs the production distribution.

The Normal distribution serves as the principal reference benchmark in distribution fitting studies [20] and is defined as

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

where μ is the mean and σ the standard deviation. Although the Normal distribution is frequently employed in PV performance studies due to its tractability and well-established probabilistic framework, its symmetric structure renders it inadequate for strongly skewed production profiles [21]. Comparison against the Gamma model through the Kolmogorov–Smirnov statistic and the Akaike Information Criterion (AIC) [22] enables objective assessment of distributional adequacy.

C. Skewness and Kurtosis

Higher-order statistical moments such as skewness and excess kurtosis are essential for characterizing the shape of solar power distributions. Skewness is given by

$$\gamma_1=\frac{\mu_3}{\sigma^3}=\frac{E[(X-\mu)^3]}{\sigma^3}, \tag{3}$$

and excess kurtosis is expressed as

$$\gamma_2=\frac{\mu_4}{\sigma^4}-3=\frac{E[(X-\mu)^4]}{\sigma^4}-3, \tag{4}$$

where the subtraction of 3 centers the metric relative to the Normal distribution, which has a classical kurtosis of 3. A positive excess kurtosis ($\gamma_2 > 0$) indicates a leptokurtic distribution with heavier tails than the Normal, while $\gamma_2 < 0$ corresponds to a platykurtic distribution. These measures capture biases and heavy tails in solar distributions, which influence probabilistic risk assessments such as the P90 metric [23]. Their use enhances the robustness of uncertainty quantification in renewable integration [24], and they are included in evaluation metrics frameworks for solar forecasting models. Studies confirm their importance in improving photovoltaic project assessments under uncertain conditions [25].

D. Autocorrelation and Temporal Dependence

The temporal structure of photovoltaic energy production can be analyzed using the autocorrelation function (ACF), defined as:

$$\rho(k) = \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{\sigma^2}, \tag{1}$$

where k represents the time lag. The ACF reveals persistence patterns and temporal dependencies associated with meteorological conditions [26].

In many practical cases, the autocorrelation structure can be approximated by an exponential decay model:

$$\rho(k) \approx \exp(-\alpha k), \tag{2}$$

where α is the decay rate parameter. This formulation provides a compact representation of memory effects in solar generation processes. Understanding temporal dependence is crucial for improving forecasting models, energy storage sizing, and grid integration strategies [27].

E. Capacity Factor and Seasonal Variance

The capacity factor (CF) is a normalized performance metric that expresses the ratio of actual daily energy production to the maximum theoretically achievable output under continuous full-rated operation. It is defined as

$$\text{CF} = \frac{E_{\text{daily}}}{P_{\text{inst}} \times 24\,\text{h}/1000} \tag{9}$$

where E_{daily} is the measured daily energy output (MWh/day) and P_{inst} is the installed capacity (kW). For the Carhuaquero system, the nominal baseline is $565.25 \times 24/1000 = 13.566$ MWh/day. The CF provides a dimensionless index comparable across installations of different sizes and geographical contexts [28], making it the standard metric for seasonal performance benchmarking in this study.

Seasonal heteroscedasticity — the variation in production variance across meteorological seasons — is quantified through the conditional variance estimator applied to each seasonal subset S:

$$\sigma_S^2 = \frac{1}{(n_S - 1)} \sum_{i \in S} (x_i - \bar{x}_S)^2 \tag{10}$$

where n_S is the number of observations in season S, x_i is the daily production, and $\bar{x}_S$ is the seasonal mean. Comparison of σ_S^2 across the dry season (May–August), transition season (September–November), and rainy season (December–February) quantifies how production uncertainty varies throughout the annual cycle — a critical input for adaptive maintenance scheduling and energy storage dimensioning [29].

III. Methodology

The analytical framework combines descriptive statistical analysis with parametric distribution modeling and temporal correlation assessment to characterize daily energy production patterns at the Carhuaquero Photovoltaic Solar Plant. The dataset comprises 365 consecutive daily generation values (MWh/day) for the Carhuaquero plant, spanning the period from May 17, 2024, to May 16, 2025 [30], constituting a complete annual cycle. The scope of this study is confined to the statistical characterization of daily energy output.

Raw generation records were compiled in tabular form and structured as a chronological time series. Data integrity was ensured through: (i) continuity checks for missing entries; (ii) anomaly detection using the interquartile range (IQR); and (iii) validation of extreme values against operational constraints. Five near-zero generation days were retained, as they correspond to scheduled maintenance or curtailment events, preserving the temporal structure for autocorrelation analysis. Although the complete 365-day series was used for all temporal analyses (descriptive statistics, ACF, and seasonal variance), these five maintenance-related near-zero observations were excluded from probabilistic distribution fitting to avoid biasing parameter estimation, yielding a fitting subset of $N = 360$ non-maintenance observations.

Descriptive statistics—mean, median, standard deviation (σ), variance (σ^2), coefficient of variation, skewness (γ_1), and excess kurtosis (γ_2)—were computed for the annual dataset and by month to capture global and seasonal patterns.

Probability distributions were fitted via maximum likelihood estimation (MLE) for Normal, Log-Normal, Weibull, and Gamma models. Model adequacy was evaluated using the Kolmogorov Smirnov (K–S) test and Akaike Information Criterion (AIC). Graphical diagnostics, including probability density function (PDF) overlays on histograms, supported these results. Temporal dependence was assessed using the autocorrelation function (ACF) for lags 1 to 7, with significance tested against 95% confidence bounds under the white-noise null hypothesis.

IV. Results

A. Descriptive Statistical Analysis

As reported in Table II, the comprehensive analysis of 365 daily energy production observations provides a detailed characterization of the operational behavior of the Carhuaquero Photovoltaic Solar Plant. Daily energy production exhibits a mean of $\mu = 2.58$ MWh/day and a median of 2.60 MWh/day, with a standard deviation of $\sigma = 1.06$ MWh/day and a variance of 1.13 MWh^2/day^2. The coefficient of variation reaches 41.1%, reflecting substantial variability and inherent uncertainty in short-term production forecasting. Observed values span from 0.05 to 6.10 MWh/day (range: 6.05 MWh/day). The distribution is positively skewed ($\gamma_1 = 1.09$) and leptokurtic ($\gamma_2 = 2.84$), indicating a right-tailed, non-Gaussian behavior associated with the intermittent alternation between overcast and clear-sky conditions typical of this Andean environment.

B. Monthly Production Patterns and Seasonal Variability

Monthly production exhibits a clear seasonal pattern, as illustrated in Fig. 4. The peak capacity factor reaches 0.34 in May 2024 (4.56 MWh/day), declining to a minimum of 0.13 in February 2025 (1.73 MWh/day) representing a 162% increase relative to the minimum monthly capacity factor, driven by the Andean hydrological cycle. The dry season (May–August)

TABLE II
ANNUAL DESCRIPTIVE STATISTICS FOR DAILY ENERGY PRODUCTION

Statistical Measure	Value	Unit
Mean	2.58	MWh/day
Median	2.60	MWh/day
Standard Deviation	1.06	MWh/day
Variance	1.13	MWh^2/day^2
Minimum	0.05	MWh/day
Maximum	6.10	MWh/day
Range	6.05	MWh/day
Coefficient of Variation	41.2	%
Skewness	1.09	–
Excess Kurtosis	2.84	–

averages 3.56 MWh/day, while the rainy season (December–February) decreases to 1.95 MWh/day, representing a 45% reduction. The annual average capacity factor is 0.19, consistent with high-altitude tropical installations. The seasonal contrast depicted in Figure 4 highlights an inverse relationship between solar production and precipitation, reinforcing the operational complementarity with the co-located Carhuaquero hydroelectric facility, whose output peaks precisely when PV generation reaches its minimum.

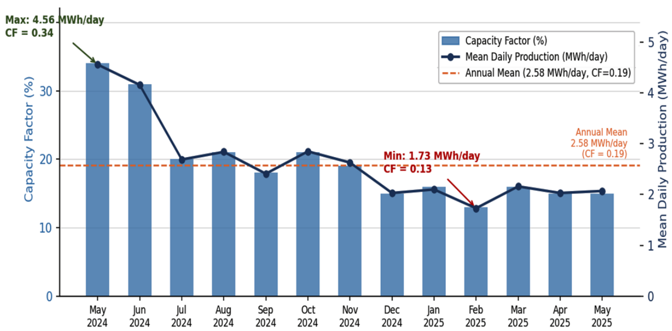

Fig. 4. Monthly capacity factors and average daily energy production

The 365-day time series reveals a strongly seasonal production profile, as shown in Fig. 5. The annual mean of 2.58 MWh/day serves as the baseline, with 81.1% of observations falling within ± 1 standard deviation (σ = 1.06 MWh/day). Peak production clusters in May–June 2024, with an abrupt 35% decline from June (4.16 MWh/day) to July (2.69 MWh/day), marking the onset of the transition season. Rainy-season months (December–February) register the lowest production levels, consistent with the pronounced troughs observed in the figure. The consecutive clustering of similar-production days provides visual evidence of the multi-day autocorrelation structure.

Monthly production analysis reveals pronounced seasonal variations throughout the annual cycle, as summarized in Table III. Peak production occurs in May 2024, with a mean

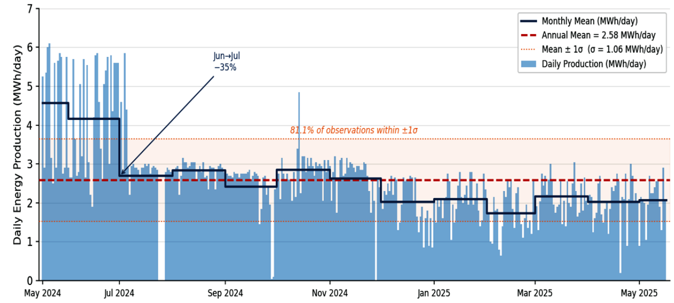

Fig. 5. Daily energy production time series (May 17, 2024 – May 16, 2025)

of 4.56 MWh/day and a capacity factor of 0.34, while the minimum is recorded in February 2025 (1.73 MWh/day; CF $=$ 0.13). The dry season (May–August) averages 3.56 MWh/day, compared to 1.95 MWh/day during the rainy season (December–February), representing a reduction of 45%. The annual average capacity factor is 0.19, calculated against the nominal baseline of 13.566 MWh/day. These patterns directly reflect the dominant influence of the Andean hydrological cycle on solar resource availability.

TABLE III
MONTHLY ENERGY PRODUCTION STATISTICS

Month	Mean Production (MWh/day)	Standard Deviation (MWh/day)	Capacity Factor
May 2024	4.56	1.47	0.34
June 2024	4.16	1.38	0.31
July 2024	2.69	1.25	0.20
August 2024	2.84	0.22	0.21
September 2024	2.41	0.71	0.18
October 2024	2.85	0.52	0.21
November 2024	2.63	0.63	0.19
December 2024	2.03	0.59	0.15
January 2025	2.10	0.45	0.16
February 2025	1.73	0.56	0.13
March 2025	2.16	0.47	0.16
April 2025	2.03	0.60	0.15
May 2025	2.07	0.57	0.15

C. Distribution Modeling and Statistical Characterization

Daily production follows a right-skewed, leptokurtic distribution ($\gamma_1 = 1.09$, excess kurtosis $\gamma_2 = 2.84 > 0$), inconsistent with normality, as shown in Figure 6. The Gamma distribution provides the best fit ($\alpha = 6.10$, scale $= 0.43$ MWh/day; KS $= 0.173$; AIC $= 1025.9$), although all candidate models are formally rejected ($p < 0.001$). The Normal distribution, overlaid for reference, overestimates intermediate production values (3.0–5.0 MWh/day) and underrepresents the modal peak. The dominant class $[2.5, 3.0)$ MWh/day concentrates 142 observations (39.4%), while the right tail

(> 5.0 MWh/day) reflects infrequent clear-sky events associated with the dry season, consistent with the extended tail observed in Fig. 6.

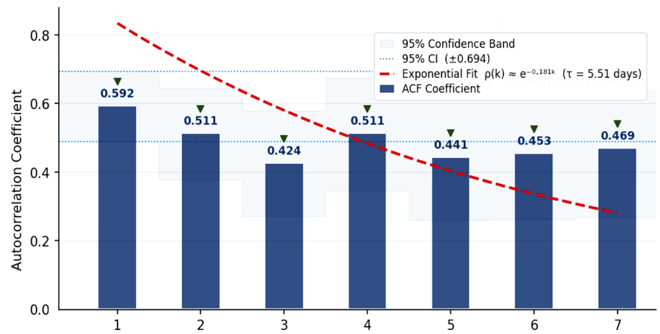

Fig. 6. Probability density distribution of daily energy production

Distribution fitting was performed via maximum likelihood estimation on $N = 360$ non-maintenance observations. Among the four candidate models, the Gamma distribution provides the best relative approximation ($\alpha = 6.10$, scale $= 0.43$ MWh/day; KS $= 0.173$; AIC $= 1025.9$), followed by the Weibull (AIC $= 1039.6$), Normal (AIC $= 1042.2$), and Log-Normal (AIC $= 1110.3$) distributions, as summarized in Table IV. All models are formally rejected by the Kolmogorov–Smirnov test ($p < 0.001$), highlighting the complex and non-standard structure of the production data. The Gamma parameters estimated on the 360-observation subset imply a mean of $\alpha \times$ scale $= 6.10 \times 0.43 = 2.62$ MWh/day, which differs marginally from the 365-observation descriptive mean of 2.58 MWh/day due to the exclusion of maintenance days. The Normal distribution with KS $= 0.208$ and AIC $= 1042.2$ exhibits the poorest performance among the parametric candidates and is therefore unsuitable as a modeling assumption for this installation.

TABLE IV
DISTRIBUTION FITTING RESULTS

Distribution Type	Parameters	K-S Statistic	P-value	AIC Value
Normal	$\mu = 2.62,\ \sigma = 1.02$	0.208	< 0.001	1042.2
Log-Normal	$\mu_{\ln} = 0.88,\ \sigma_{\ln} = 0.47$	0.182	< 0.001	1110.3
Weibull	$k = 2.59,\ \lambda = 2.93$	0.202	< 0.001	1039.6
Gamma	$\alpha = 6.10,$ scale $= 0.43$	0.173	< 0.001	1025.9

D. Variance Analysis and Temporal Correlation Structure

The autocorrelation function reveals statistically significant temporal dependence at all seven lags ($\rho_1 = 0.592$, $\rho_2 = 0.511$, $\ldots$, $\rho_7 = 0.469$), all exceeding the 95% confidence bounds, as illustrated in Fig. 7. The fitted exponential decay yields $\tau = 5.51$ days ($\rho(k) \approx e^{-0.181k}$), indicating that weather-induced anomalies persist for approximately 5–6 days. Seasonal heteroscedasticity is also evident: dry-season variance ($\sigma^2 = 1.86$ MWh^2/day^2) exceeds transition-season variance ($\sigma^2 = 0.42$ MWh^2/day^2) by a factor of 4.4, driven by the alternation of clear-sky and overcast days. These findings have direct implications for energy storage dimensioning and short-term dispatch scheduling.

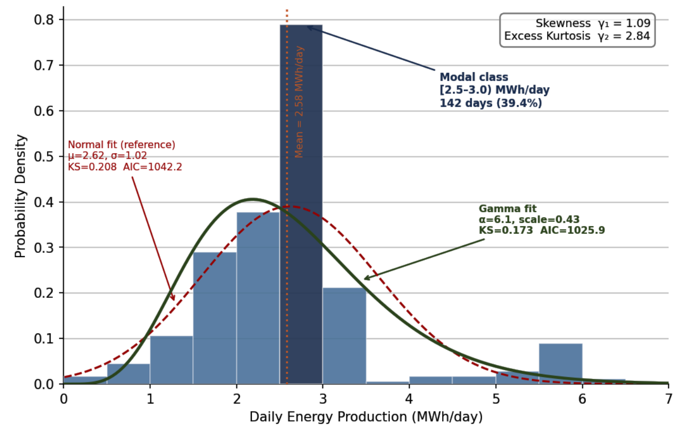

Fig. 7. Autocorrelation function of daily energy production

The sample autocorrelation function reveals statistically significant temporal dependence at all seven lags ($\rho_1 = 0.592$, $\rho_2 = 0.511$, $\ldots$, $\rho_7 = 0.469$), all exceeding the 95% confidence bounds, as summarized in Table V. The fitted exponential decay yields a characteristic time constant of $\tau = 5.51$ days, indicating that weather-induced production anomalies persist approximately 5–6 days. Seasonal heteroscedasticity is pronounced: dry-season variance ($\sigma^2 = 1.86$ MWh^2/day^2) exceeds transition-season variance ($\sigma^2 = 0.42$ MWh^2/day^2) by a factor of 4.4, while the rainy season records the minimum ($\sigma^2 = 0.31$ MWh^2/day^2). These findings have direct implications for energy storage dimensioning and short-term dispatch scheduling.

TABLE V
TEMPORAL CORRELATION ANALYSIS

Lag (Days)	Autocorrelation Coefficient	95% Confidence Interval
1	0.592	[0.489, 0.694]
2	0.511	[0.378, 0.645]
3	0.424	[0.271, 0.577]
4	0.511	[0.346, 0.675]
5	0.441	[0.260, 0.621]
6	0.453	[0.261, 0.644]
7	0.469	[0.266, 0.672]

V. DISCUSSION

A. Production Variability and Grid Integration Challenges

The coefficient of variation of 41.2% recorded at the Carhuaquero plant reflects the pronounced stochasticity inherent to small-scale PV systems operating in tropical mountain environments. This value exceeds the variability range of $25–35\%$ reported for subtropical and high-altitude installations and markedly surpasses the lower coefficients of 15–22% typical of systems in arid desert climates, where stable atmospheric conditions promote more uniform irradiance. The sustained temporal clustering of consecutive high- and low-production periods—evidenced by statistically significant autocorrelation at all seven lags—further complicates short-term energy dispatch planning, as production anomalies persist over multi-day horizons of approximately 5–6 days. These characteristics underscore the strategic value of co-locating the PV plant with the Carhuaquero hydroelectric facility, whose dispatchable generation capacity provides a natural complement to solar intermittency and supports baseload supply continuity within the SEIN.

B. Seasonal Performance and Operational Strategy

The 162% increase relative to the minimum monthly capacity factor — from 0.13 (February 2025) to 0.34 (May 2024) — reflects the dominant modulation of solar resource availability by the Andean hydrological cycle. The annual average capacity factor of 0.19, calculated on the nominal baseline of $13.566\ \mathrm{MWh/day}$, characterizes the operational yield of the system under the specific cloud-regime dynamics of this high-altitude tropical site. The inverse relationship between solar production and precipitation seasonality reinforces the operational synergy between PV and hydroelectric generation: hydroelectric output is elevated during the rainy season, when solar production reaches its minimum, while PV generation peaks during the dry season, when reservoir inflows are reduced. From an operations and maintenance perspective, the low-production rainy-season months (December–February) provide the optimal scheduling window for maintenance activities, cleaning protocols, and system inspections, as system unavailability during this period carries a comparatively lower opportunity cost.

C. Statistical Modeling and Probabilistic Forecasting

The identification of the Gamma distribution as the best relative approximation among the evaluated parametric models supported by the lowest K–S statistic (0.173) and minimum AIC score ($1{,}025.9$) reflects the strongly right-skewed, leptokurtic nature of the production data ($\gamma_1 = 1.09$, excess kurtosis $\gamma_2 = 2.84$). Although all distributions are formally rejected ($p < 0.001$) by the Kolmogorov–Smirnov test, the Gamma model provides the most adequate representation of the observed frequency structure. The pronounced positive excess kurtosis indicates that the Normal distribution substantially underestimates the frequency of high-output days in the dry season, with practical implications for yield projections and P90 revenue forecasting. The fitted Gamma parameters ($\alpha = 6.10$, scale $= 0.43\ \mathrm{MWh/day}$) provide a directly applicable basis for Monte Carlo simulation, probabilistic energy trading strategy development, and storage capacity dimensioning under production uncertainty.

D. Benchmarking and Regional Context

The performance metrics derived from the Carhuaquero dataset establish an empirical reference for the emerging small-scale PV sector in the Peruvian Andes, where comparable annual performance data remain scarce in the indexed literature. The annual capacity factor of 0.19 and the temporal autocorrelation structure—characterized by a decay time constant of $\tau = 5.51$ days—reflect the dominant influence of persistent synoptic weather systems on multi-day production variability at this site. These findings provide a quantitative foundation for calibrating generation models employed in SEIN integration studies and support the development of evidence-based technical standards for small-scale PV deployment across the Andean region.

E. Study Limitations

First, the research did not include meteorological variables such as in-plane irradiance, ambient temperature, or precipitation, limiting the analysis to statistical inference without physical decomposition of production drivers. Second, the dataset covers only one annual cycle, which may not fully represent inter-annual variability or long-term degradation and soiling effects. Third, generalizability is constrained by the specific configuration of the Carhuaquero system, as well as the local climatic conditions.

VI. Conclusion

This study of daily energy production at the Carhuaquero Photovoltaic Solar Plant provides key insights into small-scale solar performance in the Andean region of Peru. Analysis of 365 observations reveals high variability, with a coefficient of variation of 41.2%, underscoring significant forecasting and grid integration challenges. Temporal patterns confirm sustained multi-day autocorrelation ($\tau = 5.51$ days) and pronounced seasonal transitions that inform operational decision-making. Seasonal analysis indicates capacity factors ranging from 0.13 to 0.34, reflecting the dominant influence of the Andean hydrological cycle on solar resource availability. Distribution analysis identifies the Gamma model as the best-fitting candidate, consistent with the right-skewed, leptokurtic production distribution. Temporal correlation reveals persistent weather-driven anomalies affecting energy storage dimensioning and dispatch scheduling. Overall, the results enhance the understanding of photovoltaic performance in Andean environments and provide data-driven benchmarks for operational planning in Peru. These findings support adaptive strategies to manage seasonal variability, optimize maintenance scheduling during low-production periods, and inform the development of regional technical standards. Future work should extend the analysis to multiple sites and incorporate meteorological variables to quantitatively decompose the drivers of production variance.

References

[1] P. D. Rigo, J. Siluk, D. P. Lacerda, G. Rediske, and C. B. Rosa, "A model for measuring the success of distributed small-scale photovoltaic systems projects," *Solar Energy*, vol. 206, pp. 1–13, 2020.

[2] P. D. Rigo, J. Siluk, D. P. Lacerda, C. B. Rosa, and G. Rediske, “Is the success of small-scale photovoltaic solar energy generation achievable in Brazil?,” *Journal of Cleaner Production*, vol. 241, 2019.

[3] R. J. Mustafa, M. R. R. Gomaa, M. Al-Dhaifallah, and H. Rezk, “Environmental Impacts on the Performance of Solar Photovoltaic Systems,” *Sustainability*, vol. 12, no. 2, 2020.

[4] S. Sudarsono, T. Rusianto, and S. Kristiyama, “An Effect of Photovoltaic (PV) Arrangement on Performance of Small-scale Solar Energy System,” *E3S Web of Conferences*, vol. 328, 2022.

[5] W. Charfi, M. Chaabane, H. Mhiri, and P. Bournot, “Performance evaluation of a solar photovoltaic system,” *Energy Reports*, vol. 4, pp. 323–329, 2018.

[6] E. Al-Meshaiei, A. Al-Habaibeh, and B. Shakmak, “Rapid evaluation of micro-scale photovoltaic solar energy systems using empirical methods combined with deep learning neural networks to support systems' manufacturers,” *Journal of Cleaner Production*, vol. 244, 2020.

[7] P. Ramanan, K. Murugavel, and A. Karthick, “Performance analysis and energy metrics of grid-connected photovoltaic systems,” *Energy for Sustainable Development*, vol. 51, pp. 1–10, 2019.

[8] A. O. M. Maka and T. Chaudhary, “Performance investigation of solar photovoltaic systems integrated with battery energy storage,” *Journal of Energy Storage*, vol. 84, 2024.

[9] P. D. Rigo, J. C. M. Siluk, D. P. Lacerda, V. Thomasi, G. Rediske, and C. B. Rosa, “Evaluation of the Success of a Small-Scale Photovoltaic Energy System,” *IEEE Latin America Transactions*, vol. 17, no. 9, pp. 1477–1484, 2019.

[10] A. Barbón, P. Fortuny Ayuso, L. Bayon, and J. Fernández-Rubiera, “Non-uniform illumination in low concentration photovoltaic systems based on small-scale linear Fresnel reflectors,” *Energy*, vol. 239, 2022.

[11] T. W. B. Riyadi, M. Effendy, B. R. Utomo, and A. T. Wijayanta, “Performance of a photovoltaic-thermoelectric generator panel in combination with various solar tracking systems,” *Applied Thermal Engineering*, vol. 226, 2023.

[12] A. Glick, N. Ali, J. Bossuyt, M. Calaf, and R. B. Cal, “Utility-scale solar PV performance enhancements through system-level modifications,” *Scientific Reports*, vol. 10, no. 1, 2020.

[13] Orazul Energy Perú S.A. and Subsidiaries, *Unaudited Consolidated Financial Statements Q2 2018*. Lima, Peru: Orazul Energy Perú, 2018. [Online]. Available: https://inkiaenergy.com/wp-content/uploads/2019/09/ORAZEN-FS-Q2-2018.pdf

[14] Ministerio de Energía y Minas del Perú, “Resolución Directoral Nº 012-2024-MINEM/DGE,” *El Peruano*, Lima, 2024. [Online]. Available: https://busquedas.elperuano.pe

[15] ENVIRONMENTHG S.A.C., *Resumen ejecutivo: Declaración de Impacto Ambiental (DIA) del Proyecto Central Solar Fotovoltaica Carhuaquero*, preparado para KONDU S.A.C., Perú, March 2023.

[16] C. R. Castañeda, J. E. Valverde, and M. P. Díaz, “Technical feasibility and optimization of photovoltaic solar panels in the central area of Peru,” in *Sustainable Development with Renewable Energy. ICEER 2023. Lecture Notes in Networks and Systems*, vol. 947, Springer, 2024, pp. 37–48. doi: 10.1007/978-3-031-54394-4_4

[17] Kallpa Generación, “C.S.F. Carhuaquero,” *Project Portfolio*, [Online]. Available: https://kallpageneracion.com.pe/portafolio/p-s-carhuaquero/. [Accessed: Dec. 14, 2025].

[18] Ministerio de Energía y Minas del Perú, “Memoria descriptiva – Línea de interconexión subterránea Carhuaquero,” MINEM, Lima, 2024. [Online]. Available: https://www.gob.pe/minem

[19] R. A. López-Meza, L. Hernández-Callejo, and L. Omar-Jaramillo-Rosa, “Determination of photovoltaic power by modeling solar radiation with gamma distribution in the CEDER microgrid,” *Rev. Fac. Ing. Univ. Antioquia*, no. 102, pp. 107–153, Jan. 2022, doi: 10.17533/udea.redin.20200579.

[20] A. Guwaeder and R. Ramakumar, “Statistical Analysis of PV Insolation Data,” 2018.

[21] S. Y. Kim, B. Sapotta, G. Jang, Y.-H. Kang, and H.-G. Kim, “Prefeasibility study of photovoltaic power potential based on a skew-normal distribution,” *Energies*, vol. 13, no. 3, p. 676, 2020, doi: 10.3390/en13030676.

[22] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. New York, NY, USA: Springer, 2002, doi: 10.1007/b97636.

[23] S. Pasari and V. S. S. K. Nandigama, “Statistical modeling of solar energy,” in *Enhancing Future Skills and Entrepreneurship*, K. Sangwan and C. Herrmann, Eds. Cham, Switzerland: Springer, 2020, doi: 10.1007/978-3-030-44248-4_16.

[24] C. Min, “Investigating the Effect of Uncertainty Characteristics of Renewable Energy Resources on Power System Flexibility,” *Applied Sciences*, vol. 11, no. 12, p. 5381, 2021, doi: 10.3390/app11125381.

[25] A. Abdulkarim, S. M. Abdelkader, and D. J. Morrow, “Statistical analyses of wind and solar energy resources for the development of hybrid microgrid,” in *Proc. 2nd Int. Congr. Energy Efficiency and Energy Related Materials (ENEFM2014)*, A. Oral, Z. B. Oral, and M. Ozer, Eds. Cham, Switzerland: Springer, 2015, doi: 10.1007/978-3-319-16901-9_2.

[26] A. Skartveit and J. A. Olseth, “The probability density and autocorrelation of short-term global and beam irradiance,” *Solar Energy*, vol. 49, no. 6, pp. 477–487, 1992, doi: 10.1016/0038-092X(92)90155-4.

[27] R. Aguiar and M. Collares-Pereira, “TAG: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation,” *Solar Energy*, vol. 49, no. 3, pp. 167–174, 1992, doi: 10.1016/0038-092X(92)90068-L.

[28] N. Bolson, P. Prieto, and T. Patzek, “Capacity factors for electrical power generation from renewable and nonrenewable sources,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 119, no. 52, Art. no. e2205429119, 2022, doi: 10.1073/pnas.2205429119.

[29] International Electrotechnical Commission, *IEC 61724-1:2017 Photovoltaic System Performance – Part 1: Monitoring*, Geneva, Switzerland: IEC, 2017.

[30] COES - Economic Operation Committee of the National Interconnected System, ”Information Portal - Generation,” COES, May 1, 2025. [Online]. Available: https://coes.org.pe/Portal/portalinformacion/generacion. [Accessed: Jul 8, 2025].