TOWARDS AN ACCURATE ASSESSMENT OF PERU'S PV POTENTIAL: COMPARING PHYSICAL, EMPIRICAL, AND ML-BASED MODELS ACROSS DIVERSE CLIMATES

Rafael J. Vidal[1], José R. Angulo[2], Vlada Pleshcheva[1], Juan de la Casa[3] and Jan A. Töfflinger[2]

[1] Engineering Department, Pontificia Universidad Católica del Perú, Lima, Peru.

[2] Science Department, Pontificia Universidad Católica del Perú, Lima, Peru.

[3] Grupo IDEA. Centre for Advanced Studies in Energy and Environment (CEAEMA), University of Jaen, Jaen, Spain.

Abstract

This study compares three PV modelling approaches for predicting the direct-current (DC) power output (P_{DC}) of heterojunction (HIT) systems across five Peruvian locations with diverse climates. Using multi-year, minute-resolution datasets, we evaluated the Araujo–Green (AG) empirical model, the Single Diode (SDM) physical model, and a linear regression (LR) machine learning model. Model assessment included plotting daily DC power output (P_{DC}), which revealed systematic overestimation by deterministic models (AG, SDM) and a closer alignment of (LR) predictions with observed curves, with (LR) underestimating most of the time, except for Chachapoyas. Afterwards, model performance was assessed in two ways using RMSE, MAE, MBE, NRMSE, slope (β_1), and R^2, complemented by 95% confidence interval (95% CI) analysis of slopes on a monthly and annual basis to test significance against unity and across models. The empirical model (AG) consistently overestimated output (annual NRMSE = 5.79–8.33%), while the physical model (SDM) reached "good" accuracy in most locations (e.g., NRMSE = 5.18% in Arequipa). The machine learning model (LR) achieved the best results for all locations in terms of NRMSE, reaching the "excellent" threshold (NRMSE < 5%) with values as low as 1.42% in Lima, though underestimating the most in Arequipa (MBE = –59.84 W). Confidence interval analysis showed that (AG) and (SDM) were sometimes statistically indistinguishable at monthly scale, whereas annual results confirmed all models were statistically different, with (LR) achieving annual slopes closer to unity, except for Arequipa.

Keywords: PV modelling, multi-climate PV, physical models, empirical models, machine learning

1. Introduction

Addressing the pressing challenges of climate change requires an accelerated global transition toward renewable energies (RE) (Abdulla et al., 2024). It is projected that by 2030, renewable energy sources will account for 80 % of the growth in global electricity demand (IEA, 2020), with solar PV playing the leading role due to its rapid cost declines and scalability (Kumar et al., 2023). In Peru, despite abundant solar resource potential, the share of renewables in electricity generation remains modest—only about 5 %—and institutional mechanisms for new projects have been inconsistent (Campodonico and Carrera, 2022). This underuse is not due to lack of solar availability, but rather to structural and technical barriers that hinder deployment, to overcome them, robust PV system modeling becomes indispensable to de-risk design and investment decisions and accelerate PV adoption in Peru (Meng et al., 2022).

As Kumar et al. (2018) state, accurate models are essential to capture the influence of real meteorological data on PV behavior, ensuring that predicted and observed performance align. The dependence of solar PV output on site-specific meteorological conditions—including irradiance, temperature, aerosols, humidity, and extreme weather—is well documented (Bamisile et al., 2025), and this variability is particularly relevant in a country like Peru, which spans a variety of meteorological conditions across coastal, Andean and Amazonian regions.

PV modeling approaches can be broadly categorized as empirical (Pavan et al., 2014), physical (Li et al., 2024), or machine learning (ML) based. Empirical models rely on datasheet values at standard test conditions (STC) and corrective factors to account for the differences created by plane-of-array irradiance (G_{POA}) and module temperature distinct from STC; a notable example is the Araujo–Green (AG) model, comprising eight sequential calculations (Araujo et al., 1982; Green (1982); Almonacid et al., 2011). Physical models, in contrast, derive from semiconductor

© 2026. The Authors. Published by International Solar Energy Society
Selection and/or peer review under responsibility of Scientific Committee
10.18086/swc.2025.03.76 Available at http://proceedings.ises.org

physics and range from radiation and thermal sub-models to electrical descriptions of the I–V curve (Siddiqui et al., 2022). Single Diode Model (SDM) variants are widely used due to their balance between accuracy and computational feasibility, with parameterizations ranging from four to seven variables (De la Parra et al., 2017; Pedroza-Díaz et al., 2025). While computationally efficient, empirical approaches often struggle to capture physical or environmental influences that affect PV performance, including spectral mismatch (Daxini et al., 2025), soiling losses (Conceição et al., 2022), angle-of-incidence effects and other modeling limitations (Dolara et al., 2015). Even though both of these approaches can be applied at system scale, their accuracy remains debated, as shown by Dolara et al. (2015), where parameterized SDM models can reproduce PV output with reasonable precision, but their performance is sensitive to calibration and location-specific conditions, limiting generalizability. Similarly, Das et al. (2018) highlight that system-level complexities—such as inverter clipping, curtailment, and weather uncertainty—introduce additional sources of error absent at the cell or module level. These challenges highlight the importance of critically assessing whether simple empirical or physical models suffice, or whether machine learning approaches such as linear regression (LR) can provide superior accuracy and robustness when predicting PV system performance under diverse conditions. Regarding module technology, recent trends show a shift towards n-type crystalline-silicon (c-Si) platforms (TOPCon, HJT/HIT, among others): in 2024, c-Si accounted for around 98% of production and 70% of wafers were n-type, with HIT belonging to this class (Phillips & Warmuth, 2025). In addition, it has been found that this cell technology exhibits high open-circuit voltage (V_{oc}) and lower power temperatures coefficients compared to older technologies, an advantage in hot climates (e.g., Peru's coast/lowlands) and a cue that location-specific temperature/irradiance interactions must be modeled carefully (Zeng et al., 2022).

More recently, ML methods have gained prominence in PV performance modeling, enabling flexible data-driven predictions without explicit physical assumptions (Gaviria et al., 2022). While Linear regression (LR) is the simplest supervised learning approach, more advanced techniques such as artificial neural networks (ANNs), ensemble learners, and hybrid physics–ML frameworks are gaining traction (de Oliveira Santos et al., 2024). These methods can adapt to location-specific behaviors, though they require extensive datasets and computational resources. For countries like Peru, where multi-year experimental datasets from different climates are now available, ML models offer an attractive complement to deterministic approaches.

Within this context, this work assesses three representative electrical models—AG (empirical), SDM (physical), and LR (ML)—to predict the direct-current (DC) power output of HIT-based PV systems deployed in five Peruvian locations. By combining long-term field measurements with systematic performance evaluation, the study aims to determine which modeling strategies are best suited to capture location-specific PV behavior, determine prediction errors under diverse conditions, and ultimately support the reliable expansion of solar PV across Peru.

2. Methodology

2.1. Experimental setup

In order to have the required experimental data and study PV systems performance in Peru, an experimental setup was previously installed in 5 locations across the country: Lima, Tacna, Juliaca, Chachapoyas, and Arequipa. These locations present a variety of climatological conditions, which correspond to arid deserts, tropical, and Andean mountainous regions. The measurements and results from the experimental campaign were presented in Angulo et al. (2026), and the present study builds upon those measurements to carry out the electrical modeling of the monitored PV systems. These systems are based on different cell technologies; however, due to their market share and technical characteristics, this work will focus on Heterojunction Technology (HJT) also known as (HIT). We made sure that in each location, the systems are identical in module model, number of modules, model of inverter, and module temperature and plane-of-array irradiance sensors. The HIT-based PV systems, their locations across Peru, and main technical specifications are shown in Figure 1. As of now, we have recorded up to 5 years of data in Lima, and up to 4 in the remaining locations, however, as expected, each location has shown its own particularities like high levels of irradiance, low module temperatures, among others. Additionally, cleaning of the modules has been carried out as frequently as possible in order to reduce the impact of soiling. Electrical performance variables such as DC power output (P_{DC}) and environmental variables have been simultaneously recorded every minute as the IEC 61724-1 standard recommends. On the one hand, variables such as plane-of-array irradiance (G_{POA}) and module temperature (T_m) have been measured with an MS-80M pyranometer and a PT100 temperature sensors on the backside of the modules, respectively. On the other hand, a Sunny Boy inverter has been used to record electrical performance variables such as DC power output (P_{DC}) and AC power output (P_{AC}), among other variables.

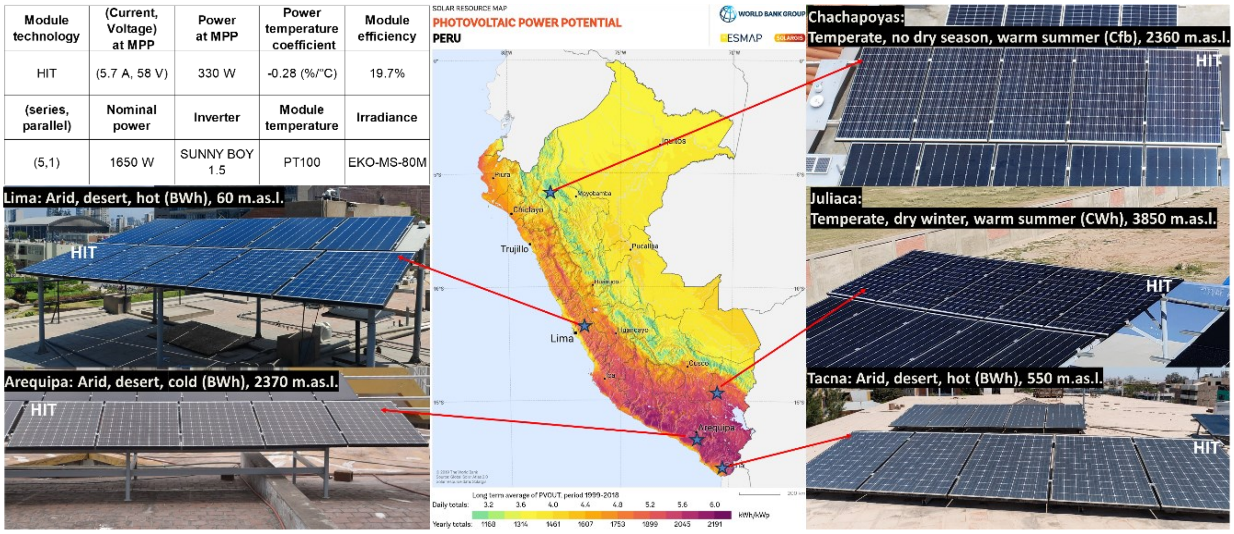

Figure 1: The installed HIT based PV systems in Peru with their technical and location related characteristics.

2.2. Data preprocessing

Pre-processing has been done to the dataset up until 2024, to ensure that the models show their actual performance by working with data that doesn't contain undesired biases. This is especially important for ML based models, since, otherwise they could learn patterns that do not relate to the PV systems behavior. The following steps describe the filters applied:

- **Inverter clipping:** Removed records with AC power ≥ 1498 W to avoid saturation, using the 1500 W nominal rating as reference.
- **Low irradiance:** In accordance with IEC61724-1:2017, removed records with $G_{POA} < 20\ W/m^2$.
- **Sensor/data errors:** Removed records when one of the required values (G_{POA}, T_m or P_{DC}) were missing due to measurement and sensor errors.

2.3 DC power output prediction

The electrical models mentioned before were implemented as Python 3.12.7 algorithms, and, specifically, the SDM was executed using pvlib-python (Holmgren et al., 2018), while, LR model was deployed using scikit-learn (Pedregosa et al., 2011). The following subsections explain the main equations used by each model.

2.3.1 Araujo-Green model (AG)

The Araujo-Green model is an empirical model based on eight sequential expressions, first at cell level, and then to PV system level. Equation 1 shows the final expression to compute DC power output for a given PV system. For further reference, (Almonacid., et al, 2011) can be consulted.

$$P_{DC_{AG}} = N_{mp}(N_{ms})N_{cp}(N_{cs})(I_{MP})(V_{MP}) \qquad \text{(eq. 1)}$$

2.3.2. Single Diode Model (SDM)

The literature presents different approaches to the SDM, depending on the number of parameters considered, such as the one presented in (De Soto et al., 2006). Equation 2, shows the implicit relation between current (I) and voltage (V), it must be noted that apart from constants, it comprises 5 parameters: (i) the photocurrent (I_L), (ii) the dark or saturation current (I_O), (iii) the series resistance (R_s), (iv) the shunt resistance (R_{sh}), and (v) the ideality factor (n). The remaining equations establish how environmental variables (G and T_m) affect the 5 parameters, and can be reviewed in (De Soto et al., 2006).

$$I = I_L - I_o\left(exp\left(\frac{V+IR_s}{N_s n V_{th}}\right) - 1\right) - \frac{V+IR_s}{R_{sh}} \qquad \text{(eq. 2)}$$

After computing the IV curve for a determined combination of (G) and (T_m), the maximum power point voltage and current (V_{MP}, I_{MP}) are extracted, and then, predicted DC power output ($P_{DC_{SDM}}$) at that point is computed following equation 3, depending on how many of the system's modules are connected in series (N_{ms}) or parallel (N_{mp}):

$$P_{DC_{SDM}} = N_{mp}(N_{ms})(I_{MP})(V_{MP}) \qquad \text{(eq. 3)}$$

2.3.3. Linear Regression (LR)

DC power output (P_{DC}) was modeled via linear regression (LR) with plane-of-array irradiance (G_{POA}) and module temperature (T_m) as the predictors; time was excluded as a covariate. System estimated DC power output ($\hat{P}_{DC}$) is computed based on the predictors and the estimated regression coefficients ($\hat{\alpha}_0$, $\hat{\alpha}_1$ and $\hat{\alpha}_2$), as shown in equation 4:

$$\hat{P}_{DC} = \hat{\alpha}_0 + \hat{\alpha}_1 G_{POA} + \hat{\alpha}_2 T_m \qquad \text{(eq. 4)}$$

2.4. Model performance assessment

After preprocessing, the three models were evaluated for their ability to predict DC power output in 2024. The two deterministic models, Araujo–Green (AG) and Single-Diode Model (SDM) were implemented according to their analytical formulations (eqs. 1 - 3). The machine-learning baseline, Linear Regression (LR), was implemented using eq. 4.

The LR model was trained using all available pre-2024 data. Specifically, the training period spanned mid-2022-2023 for Tacna, and 2021-2023 for Lima, Juliaca, Chachapoyas, and Arequipa. All models were tested on the independent 2024 dataset. Table 1 summarizes the LR training and testing datasets, including the number of days (N), the daily average number of one-minute observations, and their corresponding standard deviations (SD).

Table 1: Summary of training and testing datasets for the LR model. SD denotes the standard deviation of the daily average number of observations.

Location	Dataset							
	Train				Test			
	Years	N of days	Daily average N of observations	SD	Years	N of days	Daily average N of observations	SD
Lima	2021-2023	980	577	79	2024	317	514	83
Tacna	2022-2023	522	626	80		351	638	64
Juliaca	2021-2023	527	542	73		362	562	74
Chachapoyas	2021-2023	853	591	79		335	578	96
Arequipa	2021-2023	474	630	106		301	641	49

Model performance was assessed using two complementary approaches:

1. **Residual analysis** — the difference between observed and predicted values, $P_{DC_o} - \hat{P}_{DC_p}$; and
2. **Regression analysis** — a no-intercept linear fit between predicted and observed DC outputs.

First, residual behavior was visualized for three representative days per location (clear, partly cloudy, and overcast conditions).Then, four standard error metrics (Zhang et al., 2015) were computed using daily, monthly, and annual aggregations:: (i) Root mean square error (RMSE), (ii) mean absolute error (MAE), (iii) mean bias error (MBE), and (iv) normalized root mean square error (NRMSE), where NRMSE was calculated relative to the nominal system DC power (P_{rated} =1650 W).

To quantify systematic bias and explained variance, a no-intercept regression was fitted between the estimated and observed DC power outputs, as shown in equations (5) and (6):

$$\hat{P}_{DC_p} = \hat{\beta}_1 P_{DC_o} \qquad \text{(eq. 5)}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(P_{DC_o} - \hat{P}_{DC_p}\right)^2}{\sum_{i=1}^{n} \left(P_{DC_o}\right)^2} \qquad \text{(eq. 6)}$$

Here, $\hat{\beta}_1$ represents the slope of the regression (multiplicative bias), and R^2 quantifies the proportion of explained variance (goodness-of-fit). The intercept was constrained to zero to ensure that $\hat{\beta}_1$ directly reflects systematic bias. These metrics were computed for the same representative days and aggregated at daily, monthly, and annual scales for 2024. To determine whether $\hat{\beta}_1$ significantly differed from unity, its 95% confidence interval (95% CI) was estimated.

In summary, model assessment comprises:

Residual analysis — selected days: qualitative comparison for three representative 2024 days (clear, partly cloudy, overcast) per site; **Residual-based error metrics:** computation of RMSE, MAE, MBE, and NRMSE across daily, monthly, and annual scales; and **Regression-based evaluation:** estimation of $\hat{\beta}_1$, R^2, and 95% confidence interval (95% CI) for $\hat{\beta}_1$ to assess multiplicative bias and model fit.

The following section presents and discusses these results across locations, models (AG, SDM and LR), and temporal aggregation levels.

3. Results and discussion

3.1. Residual Analysis on Selected Days under Varying Sky Conditions

In this section, for each location we selected three days representing: (i) clear, (ii) partly cloudy, and (iii) overcast sky conditions. For each day, we plotted time series of observed and model-predicted DC power output to examine intra-day performance by sky condition and location. Figure 2 displays these series and marks an initial, qualitative comparison of model behavior across locations and conditions. For AG and SDM, it can be observed that for clear sky conditions, the largest residuals (observed-predicted) happen near solar noon (12:00 h), this could be associated to the fact that when irradiance increases, certain kinds of optical losses tend to increase as well such as the ones derived from spectral mismatch (Daxini & Wu, 2024).

In contrast, the residuals lower near sunrise (6:00 h) or sunset (18:00 h). Another interesting fact is that for clear sky conditions, deterministic models (AG and SDM) always overestimate, while the machine learning model (LR), underestimates, overestimates, or almost overlaps the observed curves, for Arequipa, Tacna and Chachapoyas, Lima and Juliaca, respectively. The results from partly cloudy sky conditions show that when large oscillations in P_{DC} occur, all the models tend to follow it, this can be related to the fact that deterministic models take into account the effects of changes in plane-of-array irradiance (G_{POA}) and module temperature (T_m), on cell or module level, and LR after the training, also partially includes those effects in α_1 and α_2. Across the analyzed locations and sky conditions, AG is the model that overestimates the most-likely reflecting its cell-level formulation-yet this bias among others from the other two models, are markedly reduced under overcast sky conditions. This improvement can be explained by the reduced effects of spectral mismatch and with extensive, stratiform cloud decks that smooth irradiance variability, thereby bringing model predictions closer to agreement to the observations (Mol et al., 2023).

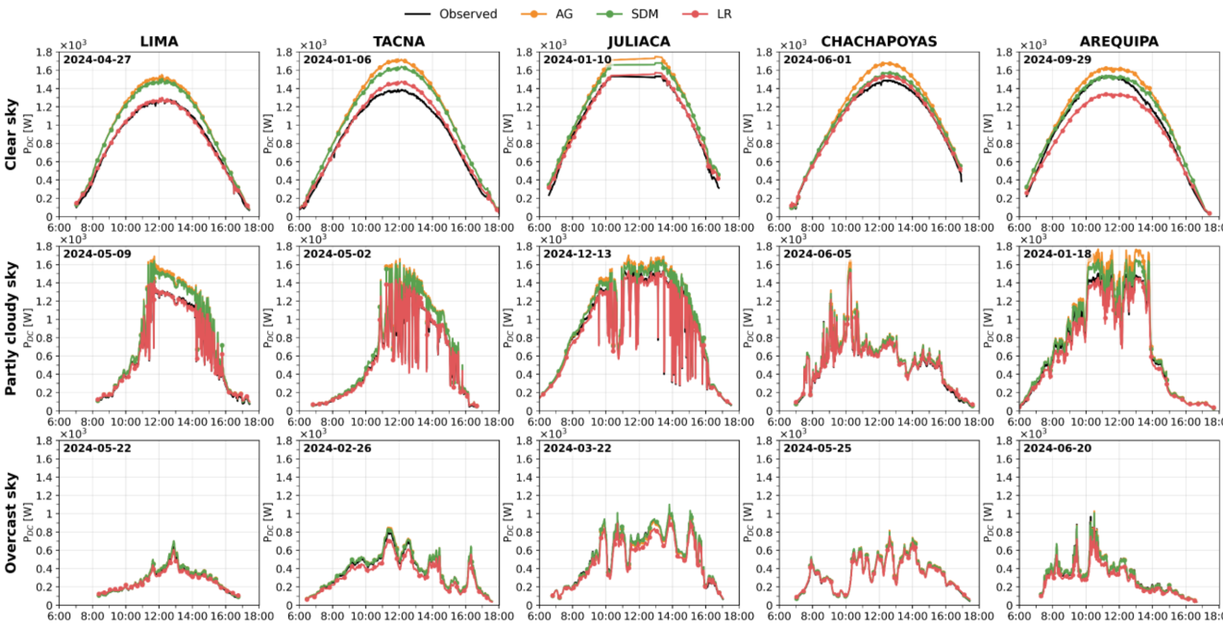

Figure 2: Observed and predicted DC power output on three selected days per location in 2024. Columns: Lima, Tacna, Juliaca, Chachapoyas, Arequipa; rows: clear, partly cloudy, overcast. Models—Araujo-Green (AG), Single Diode Model (SDM), Linear Regression (LR)—are compared at 1-min resolution over the daylight period.

3.2. Residual-Based Error Metrics across Models and Locations

In this section, we complete the residual-based evaluation by computing the four-error metrics defined earlier in section 2.4, for each model and location using one-minute data from 2024**.** Figure 3 reports the resulting metrics across locations and models, and Table 2 summarizes the corresponding annual averages for each model–location pair. Together, these results provide a quantitative basis for comparing accuracy and bias across locations and timescales. Regarding MAE and RMSE on a monthly basis, deterministic models achieve the lowest errors around July for Lima and Tacna, while for the other locations they show an increasing trend, reaching a maximum for Arequipa. The foregoing shows that there might be trends associated with site specific variables that deterministic models cannot describe. As for LR, there is a variety of trends: (i) for Lima and Juliaca the error curve is almost flat, (ii) for Tacna, an almost opposite trend to that of the deterministic models is observed, specifically, the lowest errors occur at the beginning and end of the year, whereas the highest errors are observed between June and July, (iii) interestingly, for Chachapoyas the trend is the same as for the other models, but with an offset most likely given by systematic bias, and (iv) for Arequipa the error curve shows the highest errors for this approach. Finally, RMSE exceeds MAE across all cases, indicating heavier-tailed residuals and occasional large residuals on specific days.

As shown previously in Figure 2, AG and SDM models tend to overestimate most of the time, this behavior, is corroborated by the monthly MBE curves staying positive across all locations. As for LR, the most stable behavior is observed for Lima and Juliaca, whereas for Tacna most of the first two months and the last month of the year the model overestimates, while for the remaining months it underestimates, these two trends can be related to the clear and overcast sky conditions shown previously for this location. For Chachapoya's case, it is observed that on a monthly basis, the overestimation persists throughout all the months, while Arequipa is the only location where a permanent underestimation, even on a daily basis, is maintained throughout the year.

NRMSE provides the most consistent basis for comparing results across locations for systems of different sizes, however, for this particular work where the systems are identical, its value lies in expressing deviations as a fraction of the system's nominal rating, and, shows the same trends as RMSE but scaled. This makes the results directly interpretable in terms of system performance prediction, since it quantifies how much each model misrepresents the DC power output relative to the system's nominal rating. NRMSE also facilitates benchmarking against literature thresholds: 10% is considered acceptable (Kyriakides et al., 2024), and Ahmed et al. (2025) propose the following classification—Excellent, NRMSE < 5%; Good, 5% ≤ NRMSE < 10%; Fair, 10% ≤ NRMSE < 15%; Poor, NRMSE ≥ 15%.

AG and SDM in Lima and Tacna are generally "good," often <10% mid-year, but systematic overestimation keeps them above the 5% "excellent" band. In Chachapoyas these two models track seasonality yet stay slightly high, approaching 5% only under favorable conditions. LR is more stable in Lima and Juliaca, occasionally near 5%; in Tacna it worsens mid-year, and in Arequipa it often exceeds 5%.

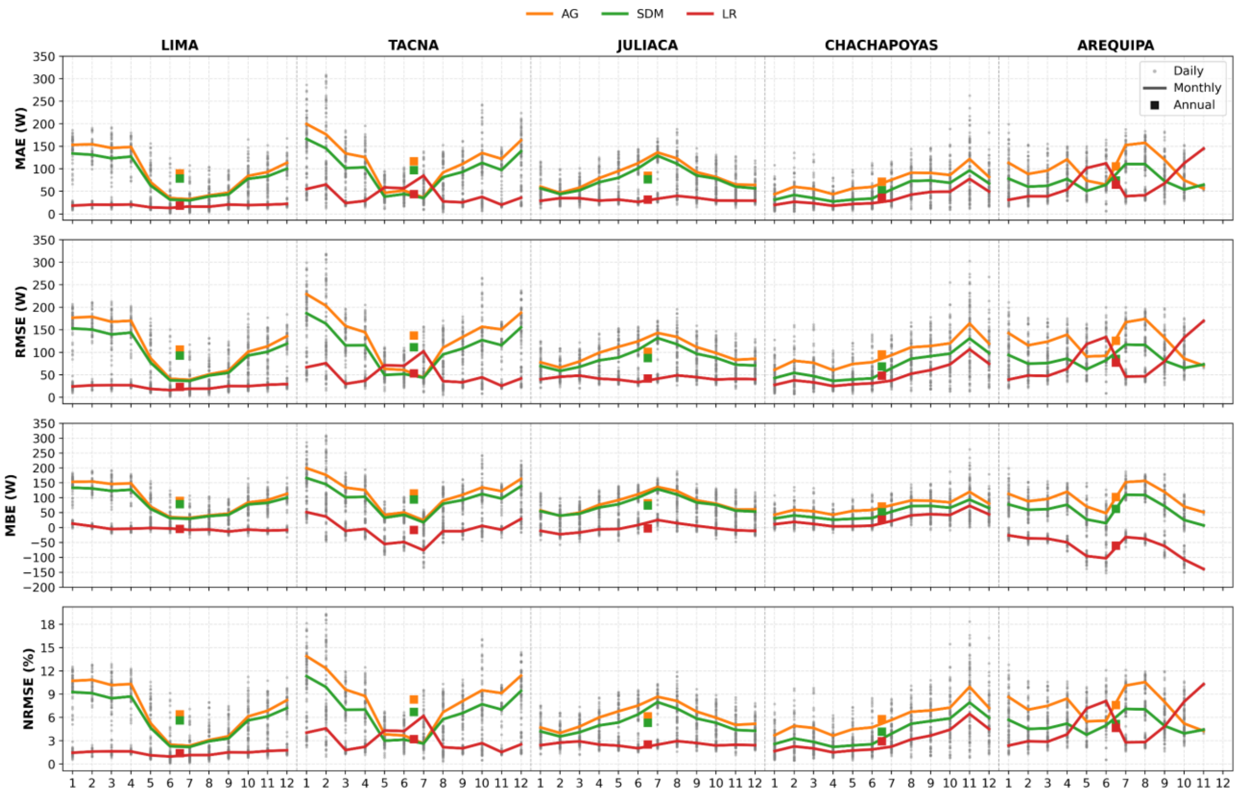

Figure 3: Residual-based error metrics (MAE-RMSE-MBE-NRMSE) for 2024 at different temporal aggregations (daily, monthly, and annual) by model and location. $NRMSE = RMSE/P_{\text{rated}}$, $P_{\text{rated}} = 1650\ W$. The x-axis lists months (1–12).

The annual basis error metrics results shown in Table 2, state that LR consistently achieves the lowest error in Lima, Tacna, Juliaca, Chachapoyas, and Arequipa with NRMSE values of 1.42%, 3.24%, 2.55% and 3.15%, and 4.68% respectively, all below the “excellent” threshold. In these locations, LR also minimizes MAE and RMSE, while AG achieves the highest annual biases, reaching 89.76 W in Lima, 115.08 W in Tacna, and 81.82 W in Juliaca. In Tacna, AG shows the highest annual RMSE, reaching 137.37 W, while LR maintains both lower RMSE and MBE, with 53.39 W and -8.07 W, respectively.

For Arequipa, the results indicate that SDM performs better than AG, attaining NRMSE (5.18%), while AG records an NRMSE of 7.62%. LR performs slightly better than SDM in terms of NRMSE (4.68%), but shows the strongest underestimation, reaching -59.84 W, this reflects what was previously discussed. To sum up, LR is the overall best performer, SDM reaches its highest performance in Chachapoyas, and AG consistently delivers the poorest results, characterized by the larges RMSE and MBE values.

Table 2: Annual (2024) error metrics for all models and locations.

Location	Model	MAE (W)	RMSE (W)	MBE (W)	NRMSE (%)
LIMA	AG	89.86	106.25	89.76	6.44
	SDM	79.12	93.09	78.96	5.64
	LR	18.46	23.51	-4.31	1.42
TACNA	AG	116.80	137.37	115.08	8.33
	SDM	97.14	111.26	94.80	6.74
	LR	43.88	53.39	-8.07	3.24
JULIACA	AG	84.75	101.38	81.82	6.14
	SDM	77.23	87.52	75.09	5.30
	LR	32.04	42.04	-2.16	2.55
CHACHAPOYAS	AG	71.83	95.61	70.98	5.79
	SDM	52.67	68.74	51.21	4.17
	LR	35.68	48.56	26.86	2.94
AREQUIPA	AG	105.88	125.73	103.11	7.62
	SDM	74.26	85.53	62.91	5.18
	LR	65.14	77.17	-59.84	4.68

3.3. Regression-Based Assessment of Bias and Explained Variance

In this section, we use a no-intercept regression of $\hat{P}_{DC,p}$on $P_{DC,o}$, following equations 5 and 6, to quantify bias and explained variance, reporting $\hat{\beta}_1$, R^2, and 95% CIs for the same previously selected days and for 2024. Accordingly, Figure 4 shows how the slope ($\hat{\beta}_1$) and coefficient of determination (R^2) are affected by sky conditions, location and models. Across the five locations, and under clear and partly cloudy sky condition, all models achieved very high (R^2), close to 1, which relates to the fact that in this step an additional regression was performed between the ouputs of each model and the observations. The main distinction arises from the results regarding the slope ($\hat{\beta}_1$), the (AG) model consistently produced ($\hat{\beta}_1$values above unity, shown by its high values and relative location to the ideal dotted black line. This, correlates to previous overestimation discussion, which is a trend also shown by the (SDM), though, in a more moderately scale. Linear Regression (LR) was generally the closest to unity under clear and partly cloudy conditions, particularly in Lima, Tacna and Juliaca. An exception can be identified for Arequipa, where LR’s slope shows a markedly underestimation, in contrast to other models.

Under overcast sky, the slope trend shifted, while (R^2) values remained high across all models. (LR) model systematically underestimated output, with slopes dropping to 0.8768 in Tacna and 0.8479 in Arequipa, however, in this condition, (SDM) consistently provides slopes closer to unity, suggesting its robustness under low-irradiance conditions. (AG) still showed a positive bias, but closer to one than under other sky conditions, suggesting reduced overestimation when cloud cover was heavier. Overall, (LR) performed best in clear and partly cloudy sky conditions, (SDM) excelled under overcast conditions, and (AG) maintained a persistent tendency toward overestimation across all locations. It must be noted that, Figure 4, only provides the point estimates of the slopes, it does not establish whether the differences between models are statistically significant, nor whether each slope differs significantly from unity.

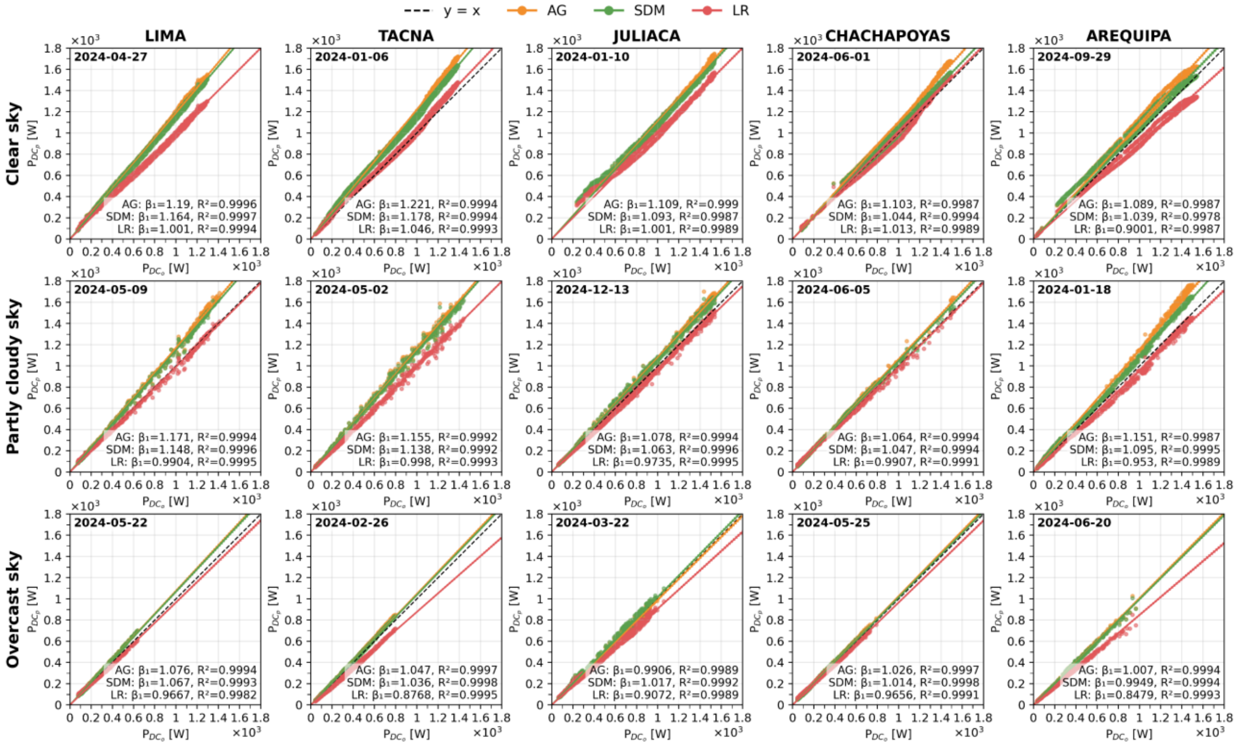

Figure 4: Predicted vs observed DC power output on selected 2024 days by location (columns: Lima, Tacna, Juliaca, Chachapoyas, Arequipa) and sky condition (rows: clear, partly cloudy, overcast). Minute-resolution pairs ($P_{DC,o}$, $\hat{P}_{DC,p}$) are shown for Araujo–Green (AG), Single-Diode Model (SDM), and Linear Regression (LR), together with the 1:1 line (black dashed). Insets report the no-intercept estimated slope $\hat{\beta}_1$and $R^2$for each model.

Figure 5a presents the results of computing ($\hat{\beta}_1$) for the 3 models for the 5 locations across all days of 2024. (AG) model shows the highest bias, defined by a higher median than other models, in addition, shows some outliers for Chachapoyas that could be related to oscillation in plane-of-array irradiance measurements. (SDM) achieves better predictions since it shows lower medians than (AG), and also narrower distribution, indicating reduced day-to-day variability, and lower overestimation in its outliers. The (LR) model achieves the medians closest to unity, and it always underestimates except for Chachapoyas. The greatest dispersion of daily slopes is observed in Tacna, whereas Lima, Juliaca and Chachapoyas, exhibit outlier presence for this model.

Figure 5b presents the distribution of daily (R^2) values for the three models across the five locations during 2024. Overall, the medians remain close to unity, confirming that the models capture most of the variability in measured DC power output, however it must be noted that this valid for the (R^2) corresponding to the additional regression, as for each model's coefficient of determination, it achieves lower values. Clear location-dependent differences can still be observed. In Lima and Tacna, the distributions are highly compact, with short boxes and whiskers indicating very stable performance, with (R^2) almost always above 0.98. Juliaca shows a wider spread than the coastal locations, with medians slightly further from one and outliers extending down to around 0.96, suggesting greater day-to-day variability, though still maintaining generally high explanatory power. Chachapoyas occupies an intermediate position, with broader interquartile ranges and numerous outliers between 0.95 and 0.90, pointing to more frequent moderate differences between observations and predictions, under certain conditions. Arequipa, while showing relatively compact boxes, except for (SDM), exhibits the lowest extreme outliers, with some values reaching 0.90-0.92. Comparing models, (SDM) yields the narrowest distributions for Lima and Tacna, while (AG) produces similar medians but wider spreads in Chachapoyas, and values further from 1, except for Arequipa. (LR) achieves competitive central values but is more sensitive to site variability, as seen in its greater dispersion in Juliaca and Arequipa. Finally, the presence of outliers across all locations might be explained by local meteorological fluctuations for certain days.

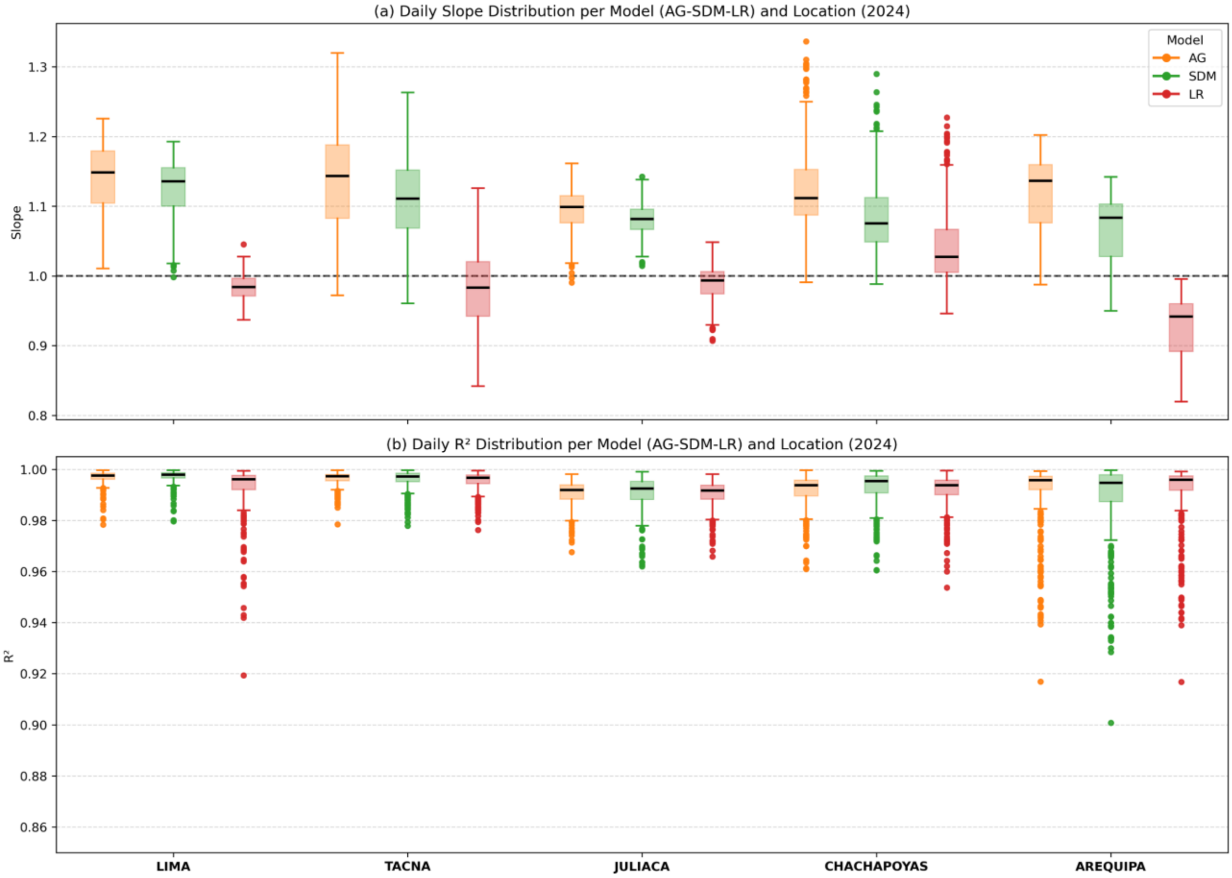

Figure 5: Daily distributions of (a) slope $\widehat{\beta}_1$and (b) $R^2$from the no-intercept regression of predicted on observed DC power in 2024. Models are: Araujo-Green (AG), Single Diode Model (SDM) and Linear Regression (LR). Boxplots summarize per-day estimates (thick black line = median; box = IQR; whiskers = 1.5×IQR); dots indicate outliers beyond the whiskers. The horizontal dashed line in (a) marks $\widehat{\beta}_1 = 1$(no bias).

Figure 6 presents, for each location, the monthly slope estimates ($\hat{\beta}_1$) alongside their annual counterparts, allowing assessment of whether each model's slope differs from unity and visual comparison of slopes across models at both monthly and annual levels. The results state that whether the 95% CI of the slope is analyzed on a monthly or annual basis, both deterministic models (AG-SDM) are statistically different from the machine learning model (LR), however, this is not always the case between (AG) and (SDM), due to their variable overlapping present in all locations, except in Arequipa. In Lima, the deterministic models overlap between June and October, which correlates to their error curves shown in Figure 3, approximating each other in this same time period, and in fact, aligns with the usual overcast days (e.g., 2024-05-22) also present in this period, as shown in Figure 4. As for Juliaca, they overlap the first and last three months of 2024, in addition, in Chachapoyas there is a slight overlap observed in December. It must be noted that aligning with the exemplary days and error metrics results, Arequipa exhibits some months with heavy underestimation for (LR), while other locations like Lima and Juliaca, have slopes closer to unity most of the months, and on an annual basis these two models are statistically closest to unit, while on this broader basis, all the models are statistically different from each other. (AG) model still exhibits the highest annual slopes for all locations, while (LR) consistently shows an annual slope closer to unity compared to other models, except for Arequipa where it is second to (SDM), this is related to the notable underestimation of the slope which might be caused to measurement errors or location-specific behaviors that the model cannot capture. Finally, while (AG) and (SDM) often exhibit similar behavior, the overlap of their confidence intervals in several locations shows that they are not statistically different from each other, reinforcing that both deterministic models share common limitations not captured by their formulations, limitations that a machine learning model, such as (LR) can adapt to, albeit at the expense of higher computational cost, requiring large amounts of data, and an operative PV system where data can be recorded. It is worth noting that, despite the differences in slope values, the deterministic models follow the same trend as the machine learning model, which demonstrates their ability to adapt to changes in plane-of-array irradiance and module temperature; however, they still exhibit a systematic bias, due to the previously discussed unconsidered losses, that causes these models to overestimate and deviate from the ideal slope of 1.

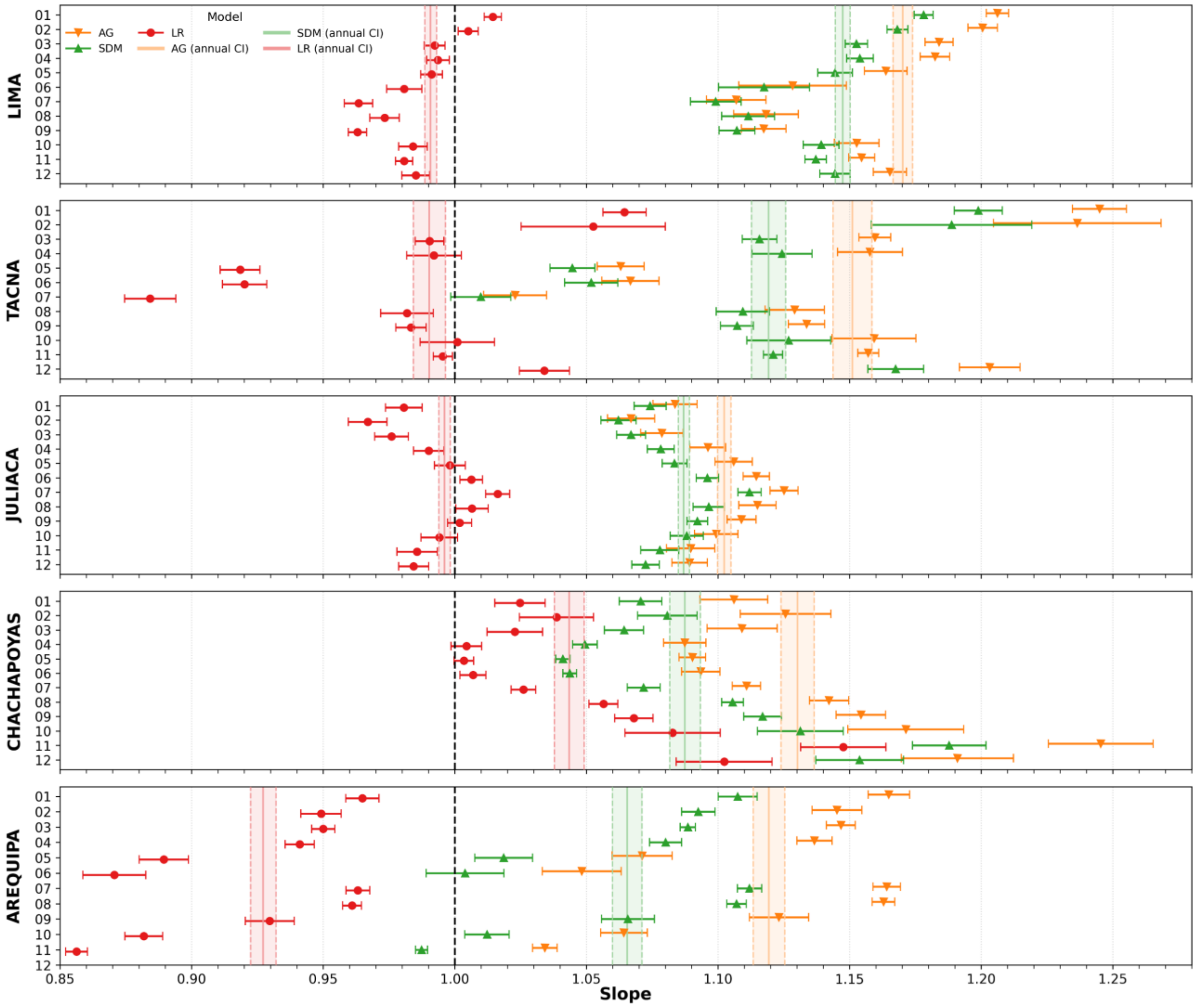

Figure 6. Monthly slope estimates $\hat{\beta}_1$ with 95% confidence intervals (95%CI's) from the no-intercept regression of predicted on observed DC power for 2024, by location (rows) and model: Araujo-Green (AG), Single Diode Model (SDM) and Linear Regression (LR). The y-axis lists months (01–12) and the vertical dashed line marks $\hat{\beta}_1 = 1$ (no bias). Shaded vertical bands with center ticks show each model's annual slope and its 95% CI for reference.

4. Conclusions

This study has provided a comprehensive comparison of three representative approaches for predicting the DC power output of HIT-based PV systems operating under Peru's diverse climatic conditions: the empirical Araujo–Green (AG) model, the physical Single Diode Model (SDM), and a machine learning–based Linear Regression (LR). By applying a consistent experimental setup across five geographically distinct locations, and evaluating the models using both error metrics and regression-based indicators slope ($\hat{\beta}_1$) and coefficient of determination (R^2), and, statistical tools, the work has demonstrated the extent to which deterministic and machine learning models can capture PV behavior, as well as their systematic limitations.

The results show that deterministic models, particularly AG, consistently exhibit a positive bias that leads to overestimation of DC output. For instance, AG recorded annual mean bias errors (MBE) of 89.76 W in Lima, 115.08 W in Tacna, and 104.96 W in Arequipa. SDM performed moderately better, with lower errors, such as RMSE of 85.53 W and NRMSE of 5.18% in Arequipa, compared to AG's RMSE of 125.73 W and NRMSE of 7.62%. LR, by contrast, achieved the lowest annual errors in all five locations, with NRMSE values of 1.42% in Lima, 3.24% in Tacna, 2.55% in Juliaca, and 2.94% in Chachapoyas, and 4.68% in Arequipa. According to the classification proposed in the literature, LR consistently reached the "excellent" level in all locations, while SDM attained "good" performance in Arequipa and AG remained largely in the "good" to "fair" range. Confidence interval tests confirm that LR and deterministic models are statistically different from each other across most conditions, with LR providing slopes closest to unit, except for Arequipa, while SDM shows greater robustness under overcast conditions, achieving slopes close to unity Arequipa. Nonetheless, both deterministic models follow similar trends to LR, indicating that they remain valid approximations when limited data or computational resources prevent the deployment of ML approaches. The persistence of systematic bias in AG and, to a lesser extent, in SDM, highlights the importance of refining these formulations to reduce their consistent overestimation. This could be done by incorporating certain PV system losses into these models.

Overall, this work underlines the importance of integrating multiple modelling strategies to assess PV performance in complex environments. For Peru, the findings imply that machine learning can substantially enhance prediction accuracy in most climates, while refined deterministic models may still serve as reliable baselines, especially in data-scarce contexts. Future work can be done adding more variables to LR or switching to more advanced ML techniques and hybrid physics–ML frameworks, incorporate degradation effects over time, among other possible improvements Such developments will be key to achieving robust, scalable, and location-adapted PV performance prediction, supporting Peru's upcoming solar energy expansion as well as the worldwide acceleration of PV adoption and the broader decarbonization agenda.

5. Acknowledgements

This work was supported by CONCYTEC through the PROCIENCIA program within the framework of the "Proyectos de Investigación Aplicada 2025-02", under contract PE501094457-2025.

6. References

Abdulla, H., Sleptchenko, A., Nayfeh, A., 2024. Photovoltaic systems operation and maintenance: A review and future directions. Renew. Sustain. Energy Rev. 195, 114342.

Almonacid, F., Rus, C., Perez-Higueras, P., Hontoria, L., 2011. Calculation of the energy provided by a PV generator. Comparative study: Conventional methods vs. artificial neural networks. Energy 36(1), 375-384.

Angulo, J.R., Berastain, A., Conde, L.A., Carhuavilca, A., Pleshcheva, V., Montes-Romero, J., García, M.A., Campos-Falcon, V., Montoya, A., Gosgot, W., Coaquira, E., Puma, P., Alfaro, E., Vidal, R.J., Beltran, N.J., Chirinos, L., Cataño, M., Condori, R., Palo-Tejada, E., Barrena, M., Polo, C.A., Espinoza, R., Muñoz-Cerón, E., De la Casa, J., Töfflinger, J.A., 2026. Yield and performance analysis of PERC, HIT, and CIGS photovoltaic systems in five Peruvian city-climates. Sol. Energy 304, 114204.

Araujo, G.L., Sanchez, E., Marti, M., 1982. Determination of the two-exponential solar cell equation parameters from empirical data. Sol. Cells 5(2), 199-204.

Bamisile, O., Acen, C., Cai, D., Huang, Q., Staffell, I., Van Deventer, W., Horan, B., Stojcevski, A., 2025. The environmental factors affecting solar photovoltaic output. Renew. Sustain. Energy Rev. 208, 115073.

Campodonico, H., Carrera, C., 2022. Energy transition and renewable energies: Challenges for Peru. Energy Policy 168, 113097.

Conceicao, R., Pascoa, J.C., Oliveira Panao, M.R., Collares-Pereira, M., 2022. Soiling effect in solar energy conversion systems: A review. Renew. Sustain. Energy Rev. 161, 112348.

Das, U.K., Tey, K.S., Seyedmahmoudian, M., Mekhilef, S., Idris, M.Y., Van Deventer, W., Horan, B., Stojcevski, A., 2018. Forecasting of photovoltaic power generation and model optimization: A review. Renew. Sustain. Energy Rev. 81, 912-928.

Daxini, R., Anderson, K., Stein, J., Theristis, M., 2025. Photovoltaic module spectral mismatch losses due to cell-level EQE variation. IEEE J. Photovolt. 15(3), 458-464.

Daxini, R., Wu, Y., 2024. Review of methods to account for the solar spectral influence on photovoltaic device performance. Energy 286, 129461.

De la Parra, I., Munoz, M., Lorenzo, E., Garcia, M., Marcos, J., Martinez-Moreno, F., 2017. PV performance modelling: A review in the light of quality assurance for large PV plants. Renew. Sustain. Energy Rev. 78, 780-797.

De Oliveira Santos, L., AlSkaif, T., Barroso, G.C., de Carvalho, P.C.M., 2024. Photovoltaic power estimation and forecast models integrating physics and machine learning: A review on hybrid techniques. Sol. Energy 284, 113044.

De Soto, W., Klein, S.A., Beckman, W.A., 2006. Improvement and validation of a model for photovoltaic array performance. Sol. Energy 80(1), 78-88.

Dolara, A., Leva, S., Manzolini, G., 2015. Comparison of different physical models for PV power output prediction. Sol. Energy 119, 83-99.

Ferkous, K., Menakh, S., Guermoui, M., Bellaour, A., Bekkar, B., Rabehi, A., Agajie, T.F., Benghanem, M., 2025. Optimized solar power forecasting: A multi-decomposition framework using VMD and swarm techniques. AIP Adv. 15(9).

Gaviria, J.F., Narvaez, G., Guillen, C., Giraldo, L.F., Bressan, M., 2022. Machine learning in photovoltaic systems: A review. Renew. Energy 196, 298-318.

Green, M.A., 1982. Solar Cells: Operating Principles, Technology, and System Applications. Prentice-Hall, Englewood Cliffs.

Holmgren, W.F., Hansen, C.W., Mikofski, M.A., 2018. pvlib python: A python package for modeling solar energy systems. J. Open Source Softw. 3(29), 884. https://doi.org/10.21105/joss.00884

Kumar, M., Malik, P., Chandel, R., Chandel, S.S., 2023. Development of a novel solar PV module model for reliable power prediction under real outdoor conditions. Renew. Energy 217, 119224.

Kumar, R., Singh, S.K., 2018. Solar photovoltaic modeling and simulation: As a renewable energy solution. Energy Rep. 4, 701-712.

Li, B., Chen, X., Jain, A., 2024. Power modeling of degraded PV systems: Case studies using a dynamically updated physical model (PV-Pro). Renew. Energy 236, 121493.

Mayer, M.J., Grof, G., 2021. Extensive comparison of physical models for photovoltaic power forecasting. Appl. Energy 283, 116239.

Meng, B., Loonen, R.C.G.M., Hensen, J.L.M., 2022. Performance variability and implications for yield prediction of rooftop PV systems - Analysis of 246 identical systems. Appl. Energy 322, 119550.

Mol, W.B., van Stratum, B.J., Knap, W.H., van Heerwaarden, C.C., 2023. Reconciling observations of solar irradiance variability with cloud size distributions. J. Geophys. Res. Atmos. 128(5), e2022JD037894.

Pavan, A.M., Mellit, A., Lughi, V., 2014. Explicit empirical model for general photovoltaic devices: Experimental validation at maximum power point. Sol. Energy 101, 105-116.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. 12, 2825-2830.

Pedroza-Diaz, A., Rodrigo, P.M., Davalos-Orozco, O., De-la-Vega, E., Valera-Albacete, A., 2025. Review of explicit models for photovoltaic cell electrical characterization. Renew. Sustain. Energy Rev. 207, 114979.

Philipps, S., Warmuth, W., 2025. Photovoltaics Report [Internet]. Freiburg (Germany): Fraunhofer Institute for Solar Energy Systems ISE; 29 May 2025 [cited 13 Oct 2025]. Available from: https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf

Siddiqui, M.U., Siddiqui, O.K., Alquaity, A.B., Ali, H., Arif, A.F.M., Zubair, S.M., 2022. A comprehensive review on multi-physics modeling of photovoltaic modules. Energy Convers. Manag. 258, 115414.

Theocharides, S., Makrides, G., Georghiou, G.E., 2024. PV generation forecasting utilizing a classification-only approach. EPJ Photovolt. 15, 12.

Zeng, Y., Peng, C.W., Hong, W., Wang, S., Yu, C., Zou, S., Su, X., 2022. Review on metallization approaches for high-efficiency silicon heterojunction solar cells. Trans. Tianjin Univ. 28(5), 358-373.

Zhang, J., Florita, A., Hodge, B. M., Lu, S., Hamann, H. F., Banunarayanan, V., & Brockway, A. M. (2015). A suite of metrics for assessing the performance of solar power forecasting. *Solar Energy*, *111*, 157-175.