International Journal of Electrical, Energy and Power System Engineering
Vol. 09, No. 03, October 2026
e-ISSN: 2654-4644

Explainable Machine Learning for Battery Side Power Estimation from Archived Photovoltaic Data

Giovanni Dimas Prenata[1*], Ahmad Ridho'i[2]
[1,2] Electrical Engineering Study Program, Universitas 17 Agustus 1945 Surabaya, Indonesia
[1]gprenata@untag-sby.ac.id, [2]ridhoi@untag-sby.ac.id

[*]Corresponding author, gprenata@untag-sby.ac.id

Abstract— This retrospective exploratory study applies explainable machine learning to an archived photovoltaic dataset containing 81 observations from three nominally 100 Wp modules over five measurement days. Battery side power was calculated as $P_{bat} = V_{bat} \times I_{bat}$, and V_{bat} and I_{bat} were excluded from the predictors. The original measurement sequence, instrument specifications, and calibration records were unavailable; therefore, the variables labelled V_{oc} and I_{sc} were treated as archived electrical descriptors rather than independently verified simultaneous open circuit and short circuit measurements. A five-feature base model used irradiance, temperature, time, V_{oc}, and I_{sc}. An augmented model added $V_{mp,est}$ and $I_{mp,est}$, which were deterministic transformations using module-specific datasheet ratios. Under nested leave one day out validation, the augmented Random Forest achieved RMSE = 6.9779 W, MAE = 5.4078 W, and R^2 = 0.8823, compared with RMSE = 7.2847 W for the base model. This 0.3068 W difference is reported descriptively because measurement uncertainty is unavailable. Leaving one module out, RMSE ranged from 6.4670 to 10.1599 W. Attribution results indicate reliance on irradiance and current related feature families. The findings quantify predictive patterns within this archived dataset and do not validate the original measurement system or dynamic MPPT performance.

Keywords: Battery side power, Explainable artificial intelligence, Incremental Conductance MPPT, Photovoltaic, Random Forest.

This work is licensed under a CC BY-NC. Copyright ©2026 by Author. Published by Universitas Riau.

INTRODUCTION

Photovoltaic (PV) power output is strongly influenced by environmental and electrical operating conditions. Solar irradiance determines the amount of available solar energy and has a direct relationship with the generated current, whereas module temperature affects voltage characteristics and conversion performance. Variations in these parameters can cause substantial changes in PV power under real operating conditions [1]–[3]. Consequently, accurate analysis of field PV performance requires consideration of both environmental variables and the electrical response of the module.

Machine-learning methods provide an effective approach for modeling the nonlinear

Received: Aug 8, 2026, Revised: Sept 20, 2026, Accepted: Oct 25, 2026
https://doi.org/10.31258/ijeepse.9.3.165-183

relationships among irradiance, temperature, electrical parameters, and PV power. Random Forest can capture complex nonlinear interactions among predictors without requiring a predefined analytical relationship [4], while Ridge Regression provides a regularized linear baseline that is useful when several predictors are correlated [5]. However, predictive accuracy alone does not explain how individual variables influence model output. This limitation is important in PV applications because variables such as irradiance, short-circuit current, and maximum power point current are physically related and may contain overlapping information.

Explainable artificial intelligence (XAI) provides tools for interpreting machine learning predictions. Partial dependence analysis can describe how the predicted response changes across the range of an individual feature [6], whereas Shapley based approaches quantify the relative contribution of each predictor to model output [7], [8]. Because limited experimental datasets can be sensitive to model-selection procedures, nested cross validation is useful for separating hyperparameter optimization from final performance evaluation and reducing optimistic estimation of model generalization [9].

Recent applications of XAI to solar energy modeling have demonstrated that feature attribution techniques can improve the transparency of photovoltaic prediction models [10]. Explainable analysis has also been applied to photovoltaic mapping and feature selection, showing that model explanations can help identify influential and redundant predictors [11]. Other recent studies have integrated XAI with optimized solar power forecasting [12], smart-city photovoltaic prediction [13], and meteorological feature analysis [14]. These studies generally identify irradiance as an important predictor, but the relative importance of temperature and other variables varies according to the dataset, model structure, operating conditions, and correlation among predictors. Moreover, correlated variables can redistribute feature-importance values, requiring careful interpretation of individual rankings [15].

Table 1 positions the present study against recent related works in PV-MPPT and explainable machine learning.

Table 1. Positioning against recent PV-MPPT and explainable-ML studies

Study	Data/Task	Model/MPPT	Validation/XAI	Main Distinction
Kuzlu et al. [10]	PV forecasting	ML forecasting	XAI tools	Solar forecasting interpretability
Rizk-Allah et al. [12]	Solar generation forecasting	X-LSTM-EO	Forecast metrics + LIME	Optimized XAI forecasting
Sun et al. [16]	PV MPPT simulation	Improved INC	Dynamic simulation	Tracking algorithm improvement
André et al. [19]	PV MPPT / HIL	INC feedback	Hardware-in-the-loop	Controller tracking behavior
Present study	81 field points; 3 modules; 5 days	Ridge + RF under INC	Nested LODO; PFI; grouped PFI; PDP; local sensitivity; exact Shapley; ablation; LOMO	Exploratory explanation of battery-side power with explicit leakage control

Recent PV research has advanced both Incremental Conductance MPPT and explainable solar power modeling. Improved Incremental Conductance methods primarily target tracking speed, oscillation reduction, and robustness to changing irradiance [16]–[19], whereas recent XAI studies use feature attribution to interpret solar power predictions [10]–[14]. The present exploratory study addresses a different intersection: it explains a battery side power estimator using archived field records reported as having been collected under Incremental Conductance operation, combines strict day wise external validation with multiple XAI

mechanisms, and tests the incremental value of estimated MPP predictors through ablation and leave one module out robustness analysis.

Despite these advances, a remaining gap is the joint interpretation of environmental and electrical predictors under Incremental Conductance MPPT using day wise external validation while checking whether the conclusions remain consistent across multiple XAI methods. This issue is particularly important for limited field datasets, where correlated predictors can make a single importance ranking appear more definitive than the data support.

This study investigates whether environmental variables and archived electrical descriptors can support battery side power estimation under the reported Incremental Conductance MPPT context. The target is explicitly $P_{bat} = V_{bat} \times I_{bat}$ rather than independently measured full-curve maximum power. The five-feature base model uses irradiance, temperature, measurement time, and the variables labelled V_{oc} and I_{sc} in the source archive. Because their original acquisition procedure could not be recovered, V_{oc} and I_{sc} are treated as recorded electrical descriptors rather than independently verified simultaneous open circuit and short circuit measurements. The augmented seven-feature model adds estimated maximum power point voltage and current. Battery voltage and battery current are excluded because they directly form P_{bat} and would create target leakage. Module identity is not entered as an explicit predictor; however, the augmented model indirectly uses module information through module specific datasheet ratios. Its results therefore represent a datasheet available archived data scenario.

The proposed framework compares Ridge Regression and Random Forest Regression using nested leave one day out validation and combines several complementary explanation techniques, including normalized Ridge coefficients, Pearson correlation, permutation feature importance, grouped permutation importance, partial dependence analysis, local sensitivity, and exact median reference Shapley attribution. The main objective is not only to obtain an accurate battery side power model but also to determine which environmental and electrical variables consistently contribute to its predictions and how these contributions vary across different explanation methods.

METHOD

The overall workflow, from archived data preprocessing and predictor construction through robustness validation and explainability analysis, is illustrated in Figure 1.

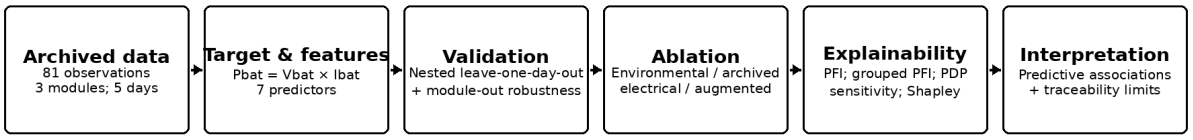


Figure 1. Archived data workflow from predictor construction to robustness validation and explainability

A. Experimental Context, Dataset Provenance, and Available Metadata

The archived Incremental Conductance dataset contains 81 observations from modules P1, P3, and P5 collected on 9, 20, 21, 23, and 24 April 2026, with 27 observations per module. According to the surviving records, the campaign was conducted at the Renewable Energy Laboratory, Electrical Engineering Study Program, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia, using discrete time slots spanning approximately 09:00–14:30 with nominal intervals of 30–60 min. Irradiance ranged from 89.61 to 1052 W/m² and module

temperature from 28.9 to 58.9 °C. The three modules were nominally rated at 100 Wp: P1, Sunasia SP100-18P polycrystalline ($V_{oc,STC}$ = 21.8 V, $I_{sc,STC}$ = 6.05 A, $V_{mp,STC}$ = 17.8 V, $I_{mp,STC}$ = 5.62 A); P3, Solartech SRM0100P polycrystalline (21.6 V, 6.33 A, 17.4 V, and 5.75 A, respectively); and P5, Kenika NMS100W monocrystalline (22.5 V, 6.00 A, 18.0 V, and 5.56 A, respectively). The records describe an MPPT solar charge controller using Incremental Conductance control with a GS MF NS60 battery on the battery side. Exact coordinates, module azimuth and tilt, converter model and topology, external load details, instrument models, sensor accuracy and resolution, calibration records, and the detailed electrical measurement sequence could not be recovered. These details are reported as unavailable rather than inferred.

This study was conducted retrospectively, and the researchers did not independently repeat or validate the original measurement campaign. The archive contains variables labelled as open-circuit voltage (V_{oc}), short-circuit current (I_{sc}), battery voltage (V_{bat}), and battery current (I_{bat}), but it does not document whether the V_{oc}- and I_{sc}-labelled values were obtained through direct circuit switching, internal controller estimation, or another acquisition procedure. It also does not establish that V_{oc}, I_{sc}, V_{bat}, and I_{bat} were measured simultaneously. For methodological transparency, V_{oc} and I_{sc} are therefore interpreted as archived V_{oc}-labelled and I_{sc}-labelled electrical descriptors; their conventional symbols are retained only for consistency with the source table. The analysis evaluates computational relationships among the archived variables and does not verify their original metrological accuracy or physical acquisition conditions. The experimental metadata recoverable from the archived field records are summarized in Table 2.

Table 2. Experimental metadata recoverable from the archived field records

Item	Recovered Information
Measurement dates	9, 20, 21, 23, and 24 April 2026
Samples	81 total; 27 per module
PV modules	P1: Sunasia SP100-18P, polycrystalline, 100 Wp, Voc=21.8 V, Isc=6.05 A, Vmp=17.8 V, Imp=5.62 A; P3: Solartech SRM0100P, polycrystalline, 100 Wp, Voc=21.6 V, Isc=6.33 A, Vmp=17.4 V, Imp=5.75 A; P5: Kenika NMS100W, monocrystalline, 100 Wp, Voc=22.5 V, Isc=6.00 A, Vmp=18.0 V, Imp=5.56 A
MPPT context	Incremental Conductance; no independent full I-V sweep reference
Archived variables	Variables labelled Voc, Isc, Vbat, Ibat, irradiance G, module temperature T, date, and time; original electrical acquisition sequence unavailable
Sampling pattern	Discrete field slots, approximately 30–60 min apart, depending on day
Observed environment	G=89.61–1052 W/m²; T=28.9°C-58.9°C
Measurement location	Renewable Energy Laboratory, Electrical Engineering Study Program, Universitas 17 Agustus 1945 Surabaya, Surabaya, Indonesia
MPPT / acquisition	Archived records identify an MPPT solar charge controller under Incremental Conductance control; instrument models, circuit-switching sequence, synchronization, and calibration records unavailable
Battery	GS MF NS60 battery
Weather context	Dry season; generally clear and rain-free during the field campaign
Unrecovered metadata	Exact site coordinates, module azimuth/tilt, detailed converter topology/model, external-load details, sensor model/accuracy/resolution, calibration records

The day wise environmental coverage of the five measurement days is presented in Table 3.

Table 3. Day-wise environmental coverage of the field campaign

Date	n	G min-max (W/m^2)	Mean G	T min-max (°C)	Mean T
9 Apr 2026	18	133.2-714.2	439.79	31.4-51.8	40.22
20 Apr 2026	18	581.1-1009.0	844.29	42.3-52.9	47.78
21 Apr 2026	15	89.61-965.0	342.70	28.9-56.9	42.33
23 Apr 2026	15	93.2-1049.0	544.77	30.4-58.9	42.17
24 Apr 2026	15	206.5-1052.0	826.25	33.2-57.9	48.14

B. Target, Estimated MPP Quantities, and Predictor Construction

Data preparation preserved the archived electrical information while preventing direct leakage from the battery side target. The target throughout this manuscript is battery side power, $P_{bat} = V_{bat} \times I_{bat}$, calculated from the recorded battery side variables rather than independently measured PV maximum power. Five complete days were successively used as outer test folds. Depending on the held out day, 63 or 66 observations were used for model development and 18 or 15 were reserved for outer testing. Internal quality control was limited to completeness and range screening, verification of the algebraic target calculation, and confirmation of the stated engineered feature transformations. These computational checks do not constitute metrological validation of the original archived observations.

The $V_{mp,est}$ and $I_{mp,est}$ columns were reconstructed as deterministic transformations of the archived V_{oc}- and I_{sc}-labelled variables using module specific datasheet ratios: $V_{mp,est} = r_{V,p}.V_{oc}$ and $I_{mp,est} = r_{I,p}.I_{sc}$. The ratios were P1: $r_V = 17.8/21.8$ and $r_I = 5.62/6.05$; P3: $r_V = 17.4/21.6$ and $r_I = 5.75/6.33$; and P5: $r_V = 18.0/22.5$ and $r_I = 5.56/6.00$. These variables do not use V_{bat}, I_{bat}, or P_{bat} and therefore do not create direct algebraic target leakage. Nevertheless, they encode module specific datasheet information, duplicate information carried by their source variables, and are not independently measured MPP quantities. The five feature base model, feature family ablation, and dependence aware interpretation are therefore essential safeguards. Battery-side photovoltaic power was defined as:

$$P_{bat,i} = V_{bat,i} \cdot I_{bat,i} \quad (1)$$

Because battery voltage and battery current directly form the target in (1), they were not included as predictors. This exclusion prevents direct target leakage. The input vector was defined as:

$$x_i = \left[G_i, T_i, t_i, V_{oc,i}, I_{sc,i}, V_{mp,est,i}, I_{mp,est,i}\right]^T \quad (2)$$

Table 4 defines the seven predictor variables used in the augmented model.

Table 4. Day wise environmental coverage of the field campaign

No	Predictor	Symbol	Interpretation
1	Irradiance	G	Available solar-energy input
2	Temperature	T	Thermal operating condition
3	Decimal time	t	Measurement time within the day
4	V_{oc}-labelled archived variable	V_{oc}	Archived voltage-related descriptor; acquisition condition unverified
5	I_{sc}-labelled archived variable	I_{sc}	Archived current-related descriptor; acquisition condition unverified
6	Estimated MPP voltage	$V_{mp,est}$	Datasheet-ratio transformation of archived V_{oc}-labelled variable
7	Estimated MPP current	$I_{mp,est}$	Datasheet-ratio transformation of archived I_{sc}-labelled variable

The archived ranges were 89.61–1052 W/m² for irradiance, 28.9–58.9 °C for temperature, 17.44–20.95 V for the V_{oc}-labelled variable, 0.12–5.84 A for the I_{sc}-labelled variable, 14.0489–17.106 V for estimated MPP voltage, and 0.1112–5.4117 A for estimated MPP current. Within each training partition, every predictor was min-max normalized:

$$x'_{ij} = \frac{x_{ij} - x_{j,min}}{x_{j,max} - x_{j,min}} \quad (3)$$

The minimum and maximum values in (3) were derived only from the relevant training partition. The same training derived transformation was subsequently applied to validation or test observations to prevent information leakage.

C. Predictive Models, Nested Day Wise Validation, and Module Out Robustness

To distinguish linear effects from nonlinear interactions under a common evaluation protocol, two complementary regressors were employed. Ridge Regression served as a regularized linear baseline, whereas Random Forest represented nonlinear relationships and predictor interactions. All hyperparameter selection was confined to the outer training days so that the held out day remained untouched until final evaluation.

Ridge Regression was used as an explainable linear baseline. Its coefficient vector was estimated by minimizing the regularized least squares objective [5]:

$$J_{Ridge} = \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x'_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \quad (4)$$

where p = 7 and λ controls coefficient shrinkage. The candidate regularization values were

$$\lambda \in \{10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\} \quad (5)$$

Random Forest Regression was used as the primary nonlinear model. Each regression tree was trained from a bootstrap sample and considered a random subset of predictors at each split, following the ensemble principle in [4]. For a forest containing B = 80 trees, the final prediction was

$$\hat{P}_{RF}(x) = \frac{1}{B} \sum_{b=1}^{B} T_b(x) \quad (6)$$

where $T_b(x)$ denotes the prediction of tree b. The hyperparameter search considered

$$d_{max} \in \{2, 3, 4, 5\} \quad (7)$$

$$n_{min} \in \{2, 3\} \quad (8)$$

$$m_{try} \in \{2, 4, 7\} \quad (9)$$

The model used nested leave one day out (LODO) validation. For each outer fold, one complete measurement day was excluded from model selection and training. Hyperparameters were selected through an inner leave one day out procedure applied only to the remaining days. The selected model was then retrained on all outer training observations and evaluated on the untouched day. This separation between model selection and final testing reduces optimistic model selection bias [9]. The selected hyperparameters varied across outer folds. Ridge selected λ = 0.1 for the first fold and λ = 1 for the remaining four folds. The selected Random Forest depths ranged from four to five, the minimum leaf size from two to three, and m_{try} from two to four. Prediction performance was assessed using RMSE, MAE, R^2, and MBE.

$$MSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2} \quad (10)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} \quad (12)$$

$$MBE = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right) \quad (13)$$

Reproducibility settings were fixed in the C++03 implementation (Microsoft Visual Studio 2010 compatible; no external machine-learning library). Random Forest used 80 bootstrap trees. The search grid was maximum depth {2,3,4,5}, minimum leaf size {2,3}, and m_{try} {2,4,7}; Ridge used λ in {10^{-8}, 10^{-7}, …, 10^{-1}, 1}. A deterministic linear congruential pseudo random generator was initialized from fixed integer seeds; outer Random Forest fits used seed 81001 + 1000 × (test-day index), and inner tuning used deterministic seeds derived from the outer day and candidate indices. Training-partition min-max scaling was applied to validation/test observations using training derived minima and maxima only. The complete fold-specific selected settings are reported in Table 5. Source code and the de-identified tabular dataset used for the analyses are available from the corresponding author upon reasonable request.

Table 5. Day wise environmental coverage of the field campaign

Test day	Ridge λ	RF depth	Min leaf	m_{try}
9 Apr	0.1	4	2	2
20 Apr	1	4	2	4
21 Apr	1	4	2	4
23 Apr	1	5	2	4
24 Apr	1	5	3	4

The reported nRMSE is range-normalized RMSE:

$$nRMSE = 100\times\frac{RMSE}{\left(y_{max} - y_{min}\right)} \quad (14)$$

where y_{min} and y_{max} are the minimum and maximum measured battery side powers in the evaluated set. WAPE is defined as:

$$WAPE = 100\times\frac{\Sigma\left|\hat{y}_i - y_i\right|}{\Sigma\left|y_i\right|} \quad (15)$$

All metrics are calculated in the original watt scale after prediction; nRMSE and WAPE are reported in percent.

D. Explainability, Ablation, and Attribution Protocol

The explanation stages were deliberately separated by validation role. PFI and grouped PFI were computed only on each untouched outer test day and then aggregated across the five folds by the mean increase in RMSE. Negative mean permutation changes were not treated as evidence of negative importance; only positive mean ΔRMSE values were normalized to sum to 100%, while non positive means were assigned 0% normalized importance and

retained in the fold level output. In contrast, one dimensional Partial Dependence Profiles (PDPs), the irradiance temperature surface, local median replacement sensitivity, and exact median reference Shapley values were computed from a final Random Forest retrained on all 81 observations after the external cross day assessment had been completed. For Shapley attribution, the reference vector is the dataset median of each of the seven predictors, absent features are replaced by those medians, all 2^7 = 128 subsets are evaluated exactly, and global importance is the normalized mean absolute Shapley magnitude. These full data explanation products are descriptive model analyses and are not used to estimate external predictive performance.

The explainability framework was designed to answer three related questions: which predictors matter globally, how predicted power changes across their operating ranges, and how individual feature contributions vary at representative operating points. Six complementary analyses were therefore combined so that interpretation did not rely on a single attribution mechanism.

First, normalized Ridge coefficients provided a directionally interpretable linear baseline; their magnitudes were interpreted cautiously because several predictors are correlated. Second, Pearson correlation was used to describe the pairwise linear association between each predictor and battery-side power:

$$r_{xy} = \frac{\sum_{i=1}^{N}\left(x_i-\underline{x}\right)\left(y_i-\underline{y}\right)}{\sqrt{\sum_{i=1}^{N}\left(x_i-\underline{x}\right)^2\sum_{i=1}^{N}\left(y_i-\underline{y}\right)^2}} \quad (16)$$

Third, Permutation Feature Importance (PFI) was calculated on each outer test fold. Predictor j was randomly permuted 20 times while all other predictors remained unchanged. Importance was defined as

$$I_j^{PFI} = \frac{1}{R}\sum_{r=1}^{R}\left(RMSE_{j,r}^{perm} - RMSE_{base}\right) \quad (17)$$

with R = 20. A larger positive increase in RMSE indicates stronger predictive reliance on that feature. Grouped permutation importance used the same procedure but permuted related predictors simultaneously. Four groups were evaluated:

To examine the descriptive contribution of the engineered estimated MPP pair, an additional Random Forest ablation used the same nested day wise protocol for four feature sets: environmental only {G, T}; archived electrical descriptors {V_{oc}, I_{sc}}; a five feature base set {G, T, time, V_{oc}, I_{sc}}; and the augmented seven feature set. A complementary leave one module out (LOMO) analysis trained on two modules and tested on the third; RF hyperparameters were selected only from the two module development data using inner day wise folds. In this LOMO scenario, the held-out module's datasheet ratios were assumed to be known at inference so that $V_{mp,est}$ and $I_{mp,est}$ could be constructed. The analysis therefore evaluates transfer to an unseen module with available datasheet coefficients, not fully module agnostic deployment.

$$G_1 = \{G, T\}, \quad G_2 = \{t\}, \quad G_3 = \{V_{oc}, I_{sc}\}, \quad G_4 = \{V_{mp,est}, I_{mp,est}\} \quad (18)$$

Fourth, one dimensional PDPs were generated. For predictor x_j, the partial dependence was calculated as [6]:

$$PD_j(z) = \frac{1}{N}\sum_{i=1}^{N} f\left(z, x_{i,-j}\right) \quad (19)$$

where z is a grid value and $x_{i,-j}$ denotes the remaining predictors. A two dimensional partial dependence surface was similarly generated for irradiance and temperature:

$$PD_{G,T}(g,\tau) = \frac{1}{N}\sum_{i=1}^{N} f\left(g, \tau, x_{i,-\{G,T\}}\right) \quad (20)$$

Fifth, local median-replacement sensitivity was calculated for representative low-, median-, and high-power observations:

$$C_{ij}^{local} = f\left(x_i\right) - f\left(x_{i,j\leftarrow\tilde{x}_j}\right) \quad (21)$$

where the modified predictor is replaced by its dataset median. Finally, an exact median reference Shapley decomposition was calculated. With seven predictors, all possible subsets were explicitly evaluated for every observation:

$$2^7 = 128 \quad (22)$$

The contribution of predictor j was then obtained from its weighted marginal contribution across all feature coalitions [7], consistent with additive feature attribution principles used in modern XAI [8]:

$$\phi_j = \sum_{S\subseteq F_j} \frac{|S|!(p-|S|-1)!}{p!}[f(S\cup j) - f(S)] \quad (23)$$

where F_j denotes the predictor set excluding feature j. Absent predictors were replaced by their median reference values. The resulting decomposition satisfies

$$f(x) = f\left(\tilde{x}\right) + \sum_{j=1}^{7} \phi_j \quad (24)$$

The running output confirmed reconstruction errors only at floating point precision, demonstrating numerical completeness of the exact decomposition. The complete workflow therefore consisted of reading the 81 observations, constructing the leakage free predictor set, holding out one complete day, tuning Ridge and Random Forest on the remaining days, retraining the selected models, predicting the untouched day, repeating the procedure across all five days, and finally evaluating global and local explanations using Ridge coefficients, correlation, PFI, grouped PFI, PDP, local sensitivity, and exact median-reference Shapley attribution.

RESULT AND DISCUSSION

A. Cross-Day Predictive Performance

Cross day performance was evaluated exclusively from outer fold predictions, so every reported value corresponds to a measurement day that was excluded from model selection and fitting. This evaluation directly tests whether the learned relationships generalize across changing daily operating conditions rather than merely fitting pooled observations. Across the 81 outer test observations, Random Forest produced the lowest aggregate error and the highest coefficient of determination, as summarized in Table 6.

Table 6. Overall nested cross day prediction performance

Model	RMSE (W)	MAE (W)	R^2	nRMSE (%)	WAPE (%)	MBE (W)
Ridge	7.7308	6.0747	0.8556	11.0846	19.1321	−0.5411
Random Forest	6.9779	5.4078	0.8823	10.0051	17.0317	0.3751

Relative to Ridge, Random Forest (RF) reduced overall RMSE by 9.74%, MAE by 10.98%, and WAPE by approximately 10.97%, while increasing R^2 from 0.8556 to 0.8823. The lower errors indicate that the nonlinear ensemble captured relationships that could not be represented completely by the additive linear Ridge model. More importantly, Random Forest obtained a lower RMSE in all five held out days, demonstrating that its aggregate improvement was not generated by a single favorable test partition. The day level results are detailed in Table 7.

Table 7. Cross-day outer test performance

Test day	Ridge RMSE (W)	RF RMSE (W)	Ridge(R^2)	RF(R^2)
1	9.9732	8.1148	0.6248	0.7516
2	7.3019	6.8550	0.5947	0.6428
3	4.8443	4.5004	0.9153	0.9269
4	7.4270	6.8640	0.8797	0.8973
5	7.7952	7.7560	0.8563	0.8577

The fold by fold RMSE comparison between Ridge Regression and Random Forest across the five held out days is shown in Figure 2.

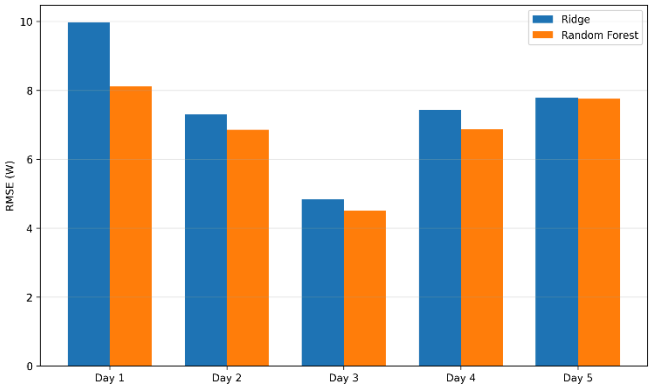


Figure 2. Cross-Day RMSE

The strongest generalization occurred on day 3, where Random Forest achieved an RMSE of 4.5004 W and R^2 = 0.9269. The most difficult condition was day 1, with an RMSE of 8.1148 W and R^2 = 0.7516. These differences indicate that model difficulty changed according to the operating conditions represented by each day. A paired Wilcoxon signed rank comparison of the five daily RMSE values yielded p = 0.0625. Therefore, although RF achieved lower RMSE in every fold, the small number of independent day level comparison units does not justify claiming conventional statistical significance at α = 0.05. The more defensible conclusion is that Random Forest consistently produced the lower observed cross day error and the best aggregate performance.

B. Feature Set Ablation and Leave One Module Out Robustness

The ablation results show a small descriptive difference between the base and augmented models. Environmental variables alone produced RMSE = 9.4142 W (R^2 = 0.7858), the archived electrical descriptors {V_{oc}, I_{sc}} produced 7.2326 W (R^2 = 0.8736), and the five-feature base model produced 7.2847 W (R^2 = 0.8718). Adding $V_{mp,est}$ and $I_{mp,est}$ reduced RMSE to

6.9779 W and increased R^2 to 0.8823, a difference of 0.3068 W (approximately 4.2%). Because sensor accuracy, resolution, and calibration records are unavailable and no inferential test was performed for this ablation contrast, the difference is not presented as a statistically or metrologically established improvement. The complete ablation results under nested leave one day out validation are presented in Table 8.

Table 8. Random Forest feature-set ablation under nested leave one day out validation

Feature set	Predictors	RMSE (W)	MAE (W)	R^2
Environmental	G, T	9.4142	7.9122	0.7858
Measured electrical	V_{oc},I_{sc}	7.2326	5.5071	0.8736
No estimated MPP	G, T, time, V_{oc}, I_{sc}	7.2847	5.6018	0.8718
Complete	All seven	6.9779	5.4078	0.8823

LOMO validation was more demanding because the tested module was absent from model fitting, although its datasheet ratios remained available for constructing the two engineered predictors. Holding out P1 yielded RMSE = 6.4670 W and R^2 = 0.9036; holding out P3 yielded 8.2506 W and R^2 = 0.7256; and holding out P5 yielded 10.1599 W and R^2 = 0.8096. The spread indicates material module to module heterogeneity and limits the findings to transfer across the tested modules under the stated datasheet available scenario. Table 9 presents the leave one module out validation results.

Table 9. Leave one module out Random Forest robustness results

Held-out module	n test	RF depth/min leaf/m_{try}	RMSE (W)	MAE (W)	R^2
P1	27	4/2/4	6.4670	4.7170	0.9036
P3	27	5/2/4	8.2506	6.0489	0.7256
P5	27	5/2/2	10.1599	8.3908	0.8096

The scatter plot of measured versus predicted battery side power for the Random Forest model across all outer test folds is shown in Figure 3.

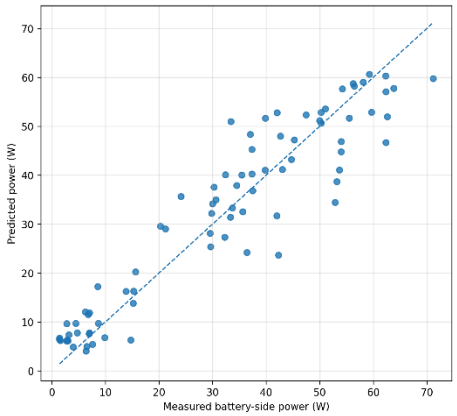


Figure 3. Random Forest measured versus predicted power

The measured versus predicted distribution confirms that the Random Forest reproduced the broad power range while retaining error at several intermediate and high power observations. The R^2 value of 0.8823 indicates that the augmented model explained a substantial portion of unseen day variation without directly using V_{bat} or I_{bat}. It did not use panel identity as an explicit predictor, but it did use module-specific datasheet ratios through $V_{mp,est}$ and $I_{mp,est}$. Because Random Forest achieved the lower observed RMSE on every held out day, the nonlinear explainability analyses in the following subsections focus on the Random Forest model, while the Ridge coefficients are retained as a complementary linear reference. This separation keeps the interpretation aligned with the model that provided the strongest cross

day predictive evidence.

C. Global Explanation and Cross Method Consistency

Having established Random Forest as the stronger cross day predictor, the next analysis examined whether independent explanation mechanisms identified the same dominant variables. Permutation feature importance measures the loss of predictive performance after disrupting a feature, whereas exact median reference Shapley values decompose predictions relative to a common reference. Comparing both views is important in this dataset because irradiance and current related electrical variables contain overlapping physical information. Table 10 summarizes the resulting global rankings.

Table 10. Leave one module out Random Forest robustness results

Rank PFI	Feature	PFI (%)	Rank Shapley	Exact Shapley (%)
1	Irradiance	34.97	3	25.50
2	$I_{mp,est}$	28.42	2	31.17
3	I_{sc}	28.11	1	31.60
4	V_{oc}	4.77	4	7.11
5	Temperature	2.64	6	1.08
6	$V_{mp,est}$	1.07	5	2.70
7	Time	0.00	7	0.85

The comparison between PFI rankings and exact median reference Shapley rankings is illustrated in Figure 4.

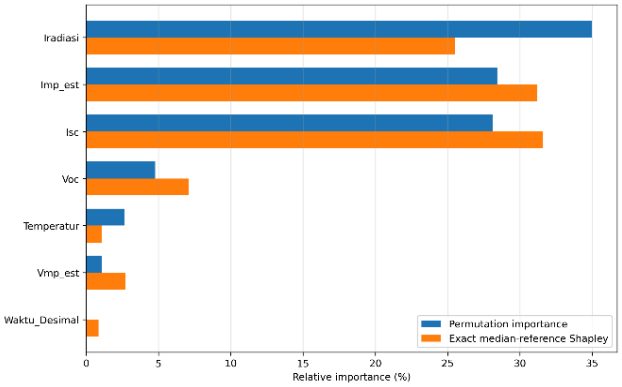


Figure 4. PFI versus exact median reference Shapley importance

Three features, namely irradiance, I_{sc}, and $I_{mp,est}$, dominated both explanation methods. They accounted for 91.51% of positive PFI importance and 88.26% of global Shapley importance. The ranking order, however, was not identical. PFI placed irradiance first, whereas Shapley placed I_{sc} first. This difference is scientifically meaningful rather than contradictory. Permutation importance quantifies the degradation in predictive performance when a feature is disrupted, while Shapley values allocate model output relative to a particular reference configuration. Consequently, features containing overlapping information can exchange ranking positions while remaining consistently dominant.

The Spearman correlation between the PFI and Shapley ranks was approximately ρ_s = 0.821, indicating strong overall agreement. The explanation of consistency output shows that $I_{mp,est}$, V_{oc}, and time received identical ranks from the two methods, while the maximum rank difference was only two positions. The grouped analysis further clarified the result. The grouped permutation importance values are listed in Table 11, and Figure 5 illustrates the relative contribution of each feature group.

Table 11. Grouped permutation importance

Rank	Feature Group	Members	Importance (%)
1	Environment	Irradiance, temperature	38.48
2	I-V endpoint electrical	V_{oc}, I_{sc}	32.32
3	Estimated MPP electrical	$V_{mp,est}$, $I_{mp,est}$	29.19
4	Temporal	Time	0.00

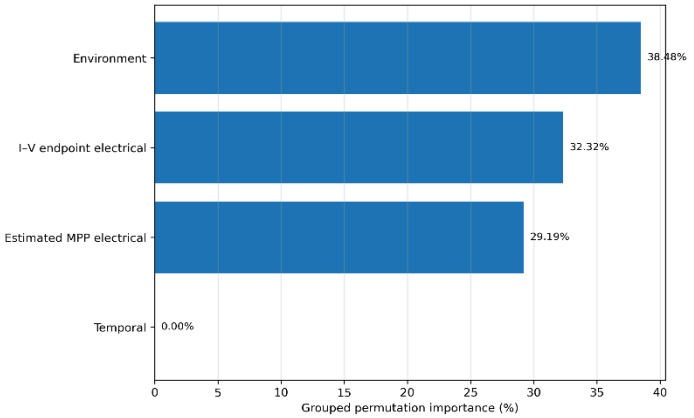


Figure 5. Grouped predictive importance

The environmental group provided the largest single grouped contribution, but its 38.48% share should not lead to the conclusion that electrical parameters are secondary. Together, the two electrical groups accounted for 61.51% of grouped model reliance, indicating that the electrical response remained at least as important as environmental forcing when interpreted collectively. The result supports a two stage physical interpretation: irradiance establishes the available solar energy forcing, whereas current and voltage related electrical parameters encode how the modules respond to that forcing within the archived operating records.

Simple Pearson correlation provides additional context. $I_{mp,est}$, I_{sc}, and irradiance had correlations with P_{bat} of 0.9315, 0.9295, and 0.9099, respectively. Temperature showed a relatively strong marginal correlation of 0.7159, whereas V_{oc} and $V_{mp,est}$ had moderate correlations of approximately 0.43. Because $I_{mp,est}$ is a fixed ratio transformation of I_{sc} within each panel, these two predictors are structurally correlated rather than independent physical drivers. Accordingly, the grouped PFI and feature set ablation are treated as the primary safeguards against over-interpreting individual percentages. The reported PFI and Shapley percentages quantify model reliance or attribution within this feature set; they do not partition physical causation.

The final Ridge coefficients followed a similar pattern: irradiance had the largest normalized coefficient, 0.2489, followed by $I_{mp,est}$ at 0.2286 and I_{sc} at 0.2230. Temperature had a coefficient of only 0.0166. This apparent difference between the marginal temperature correlation and its small conditional model contribution is important. Temperature co-varies with the environmental operating regime, but once irradiance and more directly responsive archived electrical descriptors are available, comparatively little unique predictive information remains attributable to temperature alone. Therefore, a high univariate correlation should not automatically be interpreted as high conditional feature importance. The convergence of PFI, Shapley attribution, Pearson correlation, and Ridge coefficients supports a consistent high level interpretation, but ranking alone does not reveal how predicted power changes across the operating range. The next analysis therefore examines nonlinear response shape and local contributions to determine whether the dominant predictors behave similarly under low-, intermediate-, and high power conditions.

D. Nonlinear Response and Local Explanations

Global importance establishes which predictors the model relies on, but it does not reveal the shape of their nonlinear response or whether the dominant contribution changes with the operating point. Partial dependence profiles were therefore used to examine marginal response patterns, while local median replacement analysis was used to evaluate representative low-, median-, and high power observations. Irradiance provided the clearest global nonlinear response, as shown in Figure 6.

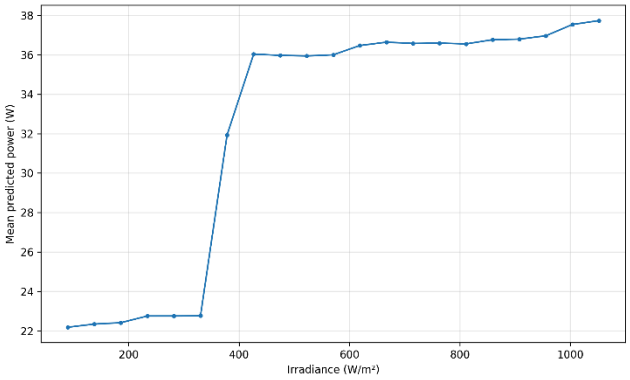


Figure 6. Irradiance partial dependence

Mean Random Forest prediction increased from approximately 22.19 W at 89.61 W/m² to 37.73 W at 1052 W/m². The response was not perfectly linear. Predicted power remained relatively flat in the lowest irradiance region, then increased sharply around the intermediate regime before gradually approaching a higher power plateau. This shape is consistent with the tree ensemble structure: Random Forest represents response changes through locally piecewise partitions rather than a single smooth linear coefficient. Current related PDPs showed a similarly strong response. The PDP for the I_{sc}-labelled variable is presented in Figure 7.

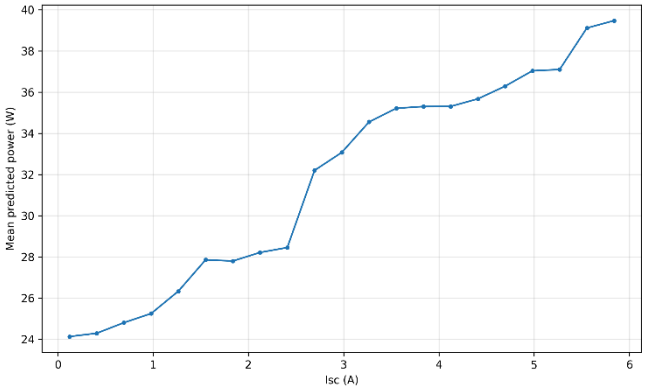


Figure 7. PDP for I_{sc}

The mean prediction increased from 24.13 W at I_{sc} = 0.12 A to 39.48 W at I_{sc} = 5.84 A. The corresponding PDP for estimated MPP current is shown in Figure 8.

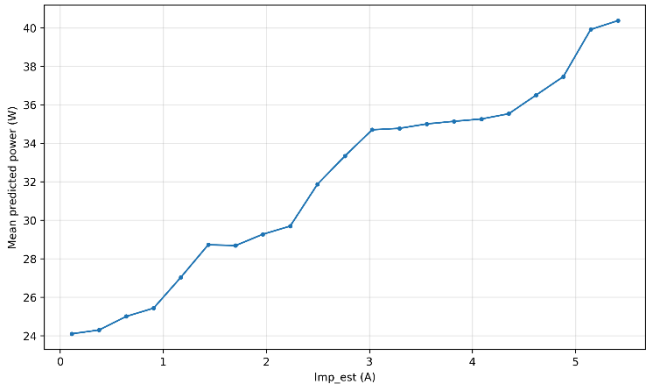


Figure 8. PDP for $I_{mp,est}$

For $I_{mp,est}$, the PDP increased from 24.10 W at 0.1112 A to 40.38 W at 5.4117 A. The magnitude and similarity of these responses explain why I_{sc} and $I_{mp,est}$ occupied the highest global Shapley ranks. Both current variables carry direct information about the instantaneous electrical response of the module. Temperature produced a very different PDP. The temperature PDP is shown in Figure 9.

Across the observed range from 28.9 to 58.9 °C, mean predicted power varied only from approximately 31.13 to 31.77 W, with a maximum near 31.84 W in the intermediate region. The limited conditional variation does not imply that temperature has no physical effect on a photovoltaic module. Rather, it indicates that within this particular seven feature model and this limited field dataset, most of the predictive information associated with temperature is already represented through irradiance and the electrical response variables. Figure 10 displays the two dimensional irradiance temperature response surface, which reinforces this interpretation.

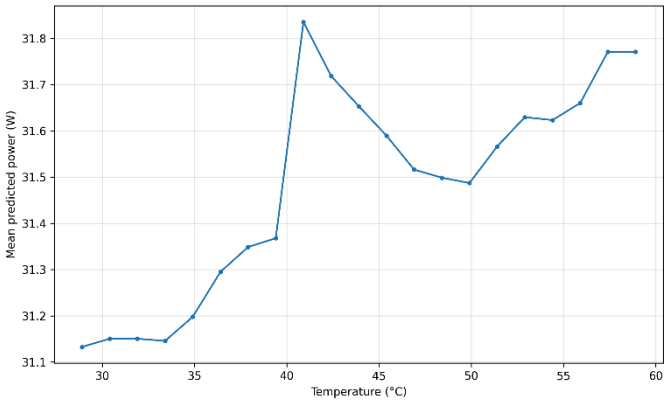


Figure 9. Temperature partial dependence

At the lowest irradiance of 89.61 W/m², predicted power ranged only from 21.63 W at 28.9 °C to 22.39 W at 58.9 °C. The surface therefore changes predominantly along the irradiance axis rather than the temperature axis. Representative local contributions for low-, median-, and high-power observations are reported in Table 12, and the local contributions across the three representative operating points are visualized in Figure 11. Local sensitivity analysis reveals that the dominant predictor can also depend on the operating point.

The low-power observation was particularly sensitive to irradiance. Replacing 133.2 W/m² with the dataset median increased the modeled output sufficiently to produce a local contribution of −14.25 W. At the high power observation, however, the strongest local contributions were associated with I_{sc} and $I_{mp,est}$, at 7.16 W and 6.44 W. This distinction

between global and local explanations is important. Irradiance can be the most important predictor according to global PFI, while an individual high power prediction can be dominated by current related electrical variables.

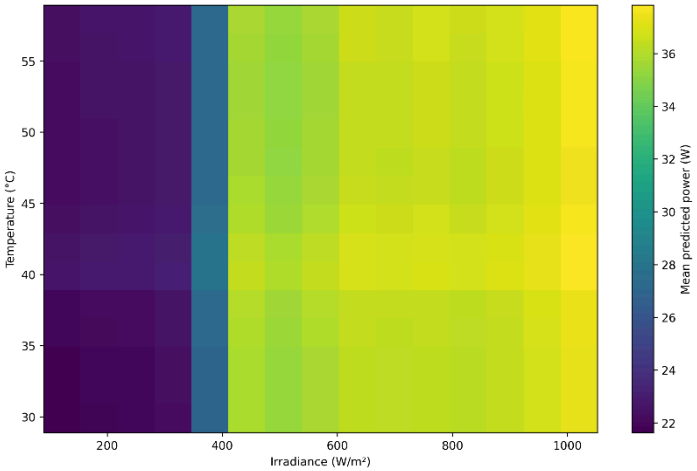


Figure 10. Irradiance temperature response surface

Table 12. Selected local median replacement contributions

Representative Observation	Feature	Original Value	Median	Local Contribution (W)
Low power, (P=8.64) W	Irradiance	133.2 W/m²	661.5 W/m²	−14.2546
	I_{sc}	0.91 A	2.85 A	-8.1995
	$I_{mp,est}$	0.8433 A	2.5889 A	-8.4401
	Temperature	31.4 °C	44.6 °C	-0.5087
Median power, (P=33.673) W	Irradiance	820.3 W/m²	661.5 W/m²	-0.0701
	V_{oc}	18.5 V	19.85 V	-0.3796
High power, (P=50.225) W	I_{sc}	5.00 A	2.85 A	7.1649
	$I_{mp,est}$	4.6446 A	2.5889 A	6.4430
	V_{oc}	20.11 V	19.85 V	1.8523
	Irradiance	965 W/m²	661.5 W/m²	0.6897

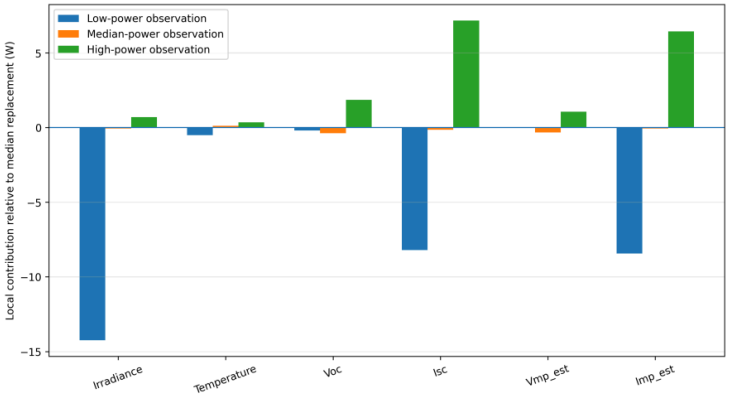


Figure 11. Irradiance temperature response surface

The sign of the local contribution is also informative: negative values at the low power point indicate that the observed low irradiance and current levels reduced the prediction relative to their median reference replacements, whereas positive current contributions at the high power point indicate the opposite behavior. Because these dominant predictors are physically correlated, the stability and causal limits of their attributions require separate consideration.

E. Interpretation Reliability, Correlated Predictors, and Limitations

Before translating attribution rankings into physical conclusions, the reliability of the explanation framework must be considered. Two issues are central: whether the exact Shapley decomposition is numerically complete and whether correlated predictors permit a causal interpretation of individual importance percentages.

The exact Shapley decomposition provides an additional consistency check because all 2^7 = 128 feature coalitions were evaluated rather than sampled. For each of the 81 observations, the reference prediction plus the seven Shapley values reproduced the complete Random Forest prediction with only floating point level numerical residual. Individual examples in the output show reconstruction errors of zero or approximately 10^{-15}.

This numerical completeness does not, however, make the attributions causal. The results demonstrate model attribution relative to the median reference baseline. In photovoltaic systems, irradiance, I_{sc}, and $I_{mp,est}$ are physically related. Strongly correlated predictors can redistribute permutation or Shapley importance because multiple features may contain similar information. Previous methodological studies have highlighted this limitation of marginal variable-importance interpretation [15].

For that reason, agreement across explanation methods is treated as descriptive evidence of model reliance rather than proof of independent effects. In particular, separately attributing importance to V_{oc} and $V_{mp,est}$ or to I_{sc} and $I_{mp,est}$ can double count closely related information because each estimated variable is a deterministic, module-scaled transformation of its archived source counterpart. Grouped PFI and feature set ablation reduce, but do not eliminate, this dependence problem. Fold level uncertainty intervals for PFI were not estimated; normalized importance percentages and rankings should therefore not be interpreted as inferentially stable quantities.

The principal limitation is incomplete measurement traceability. The original instrument models, sensor accuracy and resolution, calibration records, circuit switching procedure, and time intervals among the V_{oc}-, I_{sc}-, V_{bat}-, and I_{bat}-labelled observations were not available. Consequently, the physical acquisition conditions of the V_{oc}- and I_{sc}-labelled variables could not be independently verified, and simultaneity among the electrical variables cannot be assumed. This limitation affects the physical interpretation and external applicability of the model. The reported error metrics demonstrate predictive performance within the archived dataset, but they do not validate the original measurement system or establish that comparable accuracy would be obtained using newly acquired, calibrated, or real-time measurements.

The dataset is also limited to 81 observations, three modules, and five measurement days, with LOMO RMSE ranging from 6.4670 to 10.1599 W. Measurement uncertainty cannot be quantified, so the 0.3068 W RMSE difference between the base and augmented models cannot be judged against the uncertainty of the measurement chain. The target is battery-side power calculated from the archived V_{bat} and I_{bat} values, not independently measured full curve maximum power. $V_{mp,est}$ and $I_{mp,est}$ are deterministic transformations using module-specific datasheet ratios, not independent MPP measurements. PFI and Shapley values describe model reliance within a dependent feature set and do not identify causal effects. No claim is made regarding dynamic MPPT efficiency, simultaneous physical relationships among the four electrical variables, universal technology superiority, or broad seasonal generalization. Confirmation requires synchronized measurements, documented switching procedures, calibrated instruments, independent I–V/MPP references, and longer multi-site campaigns. Nevertheless, the combination of nested cross-day evaluation and multiple mutually complementary XAI methods provides a stronger basis for interpretation than either an accuracy only analysis or a single feature ranking method. This is particularly

important for limited experimental datasets where apparent feature dominance may otherwise result from a specific model, test split, or explanation technique.

CONCLUSION

Within the archived dataset, Random Forest provided the lowest observed cross day error for battery side power, achieving RMSE = 6.9779 W, MAE = 5.4078 W, and R^2 = 0.8823; however, the five day Wilcoxon comparison remained non-significant at $\alpha = 0.05$. Feature set ablation yielded a descriptive RMSE reduction from 7.2847 W for the five feature base model to 6.9779 W for the augmented model. Because the engineered variables are deterministic transformations based on module specific datasheet ratios and measurement uncertainty is unavailable, this difference is not claimed as statistically or metrologically established. Leaving one module out validation produced RMSE values of 6.4670 W (P1), 8.2506 W (P3), and 10.1599 W (P5) when the held out module's datasheet ratios were available at inference. This spread demonstrates appreciable module heterogeneity and limits technology level generalization. PFI, grouped PFI, Ridge coefficients, PDP, local sensitivity, and exact median reference Shapley analysis consistently indicated model reliance on irradiance and current related feature families. Because the engineered variables are deterministic transformations of archived descriptors, individual importance percentages may redistribute or duplicate related information and are therefore interpreted only as descriptive predictive associations. The findings constitute a retrospective computational case study and do not validate the original measurement system, simultaneous relationships among V_{oc}, I_{sc}, V_{bat}, and I_{bat}, or dynamic MPPT performance. Future work should employ synchronized acquisition, documented circuit switching, calibrated sensors, complete orientation and converter/load metadata, uncertainty aware comparisons, independent I–V/MPP references, a greater number of modules, and longer seasonal campaigns.

REFERENCES

[1] W. De Soto, S. A. Klein, and W. A. Beckman, "Improvement and validation of a model for photovoltaic array performance," Solar Energy, vol. 80, no. 1, pp. 78–88, 2006, doi: 10.1016/j.solener.2005.06.010.
[2] T. Huld and A. M. Gracia Amillo, "Estimating PV module performance over large geographical regions: The role of irradiance, air temperature, wind speed and solar spectrum," Energies, vol. 8, no. 6, pp. 5159–5181, 2015, doi: 10.3390/en8065159.
[3] S. Dubey, J. N. Sarvaiya, and B. Seshadri, "Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world — A review," Energy Procedia, vol. 33, pp. 311–321, 2013, doi: 10.1016/j.egypro.2013.05.072.
[4] L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
[5] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics, vol. 12, no. 1, pp. 55–67, 1970, doi: 10.1080/00401706.1970.10488634.
[6] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," The Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.
[7] M. Li, H. Sun, Y. Huang, and H. Chen, "Shapley value: from cooperative game to explainable artificial intelligence," Autonomous Intelligent Systems, vol. 4, no. 1, Art. no. 2, 2024, doi: 10.1007/s43684-023-00060-8.
[8] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, vol. 30, 2017, pp. 4768–4777.
[9] S. Varma and R. Simon, "Bias in error estimation when using cross-validation for model selection," BMC Bioinformatics, vol. 7, Art. no. 91, 2006, doi: 10.1186/1471-2105-7-91.

[10] M. Kuzlu, U. Cali, V. Sharma, and Ö. Güler, “Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools,” IEEE Access, vol. 8, pp. 187814–187823, 2020, doi: 10.1109/ACCESS.2020.3031477.
[11] E. Gomes, A. Esteves, H. Morais, and L. Pereira, “Leveraging explainable artificial intelligence in solar photovoltaic mappings: Model explanations and feature selection,” Energies, vol. 18, no. 5, Art. no. 1282, 2025, doi: 10.3390/en18051282.
[12] R. M. Rizk-Allah, L. M. Abouelmagd, A. Darwish, V. Snasel, and A. E. Hassanien, “Explainable AI and optimized solar power generation forecasting model based on environmental conditions,” PLOS ONE, vol. 19, no. 10, Art. no. e0308002, 2024, doi: 10.1371/journal.pone.0308002.
[13] O. Petrosian and Y. Zhang, “Solar power generation forecasting in smart cities and explanation based on explainable AI,” Smart Cities, vol. 7, no. 6, pp. 3388–3411, 2024, doi: 10.3390/smartcities7060132.
[14] S. Nematzadeh and V. Esen, “Explainable machine learning and predictive statistics for sustainable photovoltaic power prediction on areal meteorological variables,” Applied Sciences, vol. 15, no. 14, Art. no. 8005, 2025, doi: 10.3390/app15148005.
[15] C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, “Conditional variable importance for random forests,” BMC Bioinformatics, vol. 9, Art. no. 307, 2008, doi: 10.1186/1471-2105-9-307.
[16] C. Sun, J. Ling, and J. Wang, “Research on a novel and improved incremental conductance method,” Scientific Reports, vol. 12, Art. no. 15700, 2022, doi: 10.1038/s41598-022-20133-7.
[17] I. Owusu-Nyarko, M. A. Elgenedy, I. Abdelsalam, and K. H. Ahmed, “Modified variable step-size incremental conductance MPPT technique for photovoltaic systems,” Electronics, vol. 10, no. 19, Art. no. 2331, 2021, doi: 10.3390/electronics10192331.
[18] M. Derbeli, C. Napole, O. Barambones, J. Sanchez, I. Calvo, and P. Fernández-Bustamante, “Maximum power point tracking techniques for photovoltaic panel: A review and experimental applications,” Energies, vol. 14, no. 22, Art. no. 7806, 2021, doi: 10.3390/en14227806.
[19] S. André, F. Silva, S. Pinto, and P. Miguens, “Novel incremental conductance feedback method with integral compensator for maximum power point tracking: A comparison using hardware in the loop,” Applied Sciences, vol. 13, no. 7, Art. no. 4082, 2023, doi: 10.3390/app13074082.

BIOGRAPHIES OF AUTHORS



Giovanni Dimas Prenata received his B.Eng. in Electrical Engineering (Power Systems) from Petra Christian University, Surabaya, in 2009 and his M.Eng. in Electrical Engineering (Telematics) from Institut Teknologi Sepuluh Nopember, Surabaya, in 2018. His research interests include artificial intelligence, biomedical signal processing, computational modeling, and intelligent algorithms for energy and healthcare applications.



Ahmad Ridho’i received his B.Eng. in Electrical Engineering (Electronics) from Universitas 17 Agustus 1945 Surabaya in 1995 and his M.Eng. in Electrical Engineering (Electrotechnics) from the Bandung Institute of Technology in 2000. His research interests include intelligent control, disaster mitigation, early-warning systems, and community-based monitoring technologies.