Materials Research Proceedings 64 (2026) 535-543 https://doi.org/10.21741/9781644904091-67

A Hybrid Gradient Boosting - XGBoost MPPT Strategy for Photovoltaic Systems Under Partial Shading

Rida AMINE[1,a *], Noureddine EL BARBRI[1,b] and Hatim AMEZIANE[1,c]

[1]Science and Technology for the Engineer Laboratory (LaSTI), National School of Applied Sciences Khouribga (ENSA), Sultan Moulay Slimane University,B.P. Morocco

[a]rida.amine@usms.ac.ma, [b]n.elbarbri@usms.ma, [c]h.ameziane@usms.ma

Keywords: Photovoltaic Systems, Maximum Power Point Tracking, Partial Shading, Gradient Boosting, Xgboost, Machine Learning

Abstract. Maximum Power Point Tracking is a key feature in the power conversion of Photovoltaic PV-based power systems, which enables the system to deliver maximum energy based on fluctuating environmental conditions. Conventional MPPT approaches, such as the Perturb and Observe method (P&O) and Incremental Conductance algorithm (INC), have been observed to degrade as a result of multiple local maxima in the power-voltage (P-V) curve during partial shading conditions (PSC). A novel hybrid MPPT technique based on Gradient Boosting Regression and Extreme Gradient Boosting is presented in this paper. The key idea is to exploit the strong robustness of GBR in mid-high (sunny) irradiance and the generalization capability of XGBoost at low (cloudy) irradiance, by leveraging a hysteresis-based switching scheme that takes into account the current-irradiance level and switches between models. Irradiance level from 200 to 1000 W/m² data simulation results demonstrate that the proposed hybrid controller maintains the load power very near to the theoretical maximum, obtaining instantaneous MPPT efficiencies from 95.73% up to 97.37% and global efficiency of 96.83% across the entire profile, which are largely improved over single-model GBR (83.8%) and XGBoost (78.7%) baselines referenced works.

INTRODUCTION

The growing number of photovoltaic (PV) systems integrated into the modern power grid, owing largely to environmental concerns and declining module costs, necessitates better control strategies for the purpose of energy extraction [1]. Because the current–voltage (I-V) characteristic of a PV is non-linear and also depends on environmental factors, an MPPT controller has to be specifically employed to control the PV operating point continuously.

There are many classical MPPT methods, which use Perturb and Observe (P&O) and Incremental Conductance (INC) algorithms due to their simplicity and low implementation cost [2], [3]. However, they do not work well in the rapid changing irradiance conditions and in the partially shaded conditions (PSC), especially. The partial shading array fail is generated as a result of the shadow on part of the system by cloud, vegetation or building, which results in MPP with several local maxima and a single global maximum, but this does not occur when no shading effects are considered. Due to the nature of gradient-based and hill-climbing methods, they are inherently prone to converge towards local maxima rather than the global maximum, leading to a serious energy penalty.

In order to overcome these drawbacks, several advanced MPPT methods have been developed. These include metaheuristic optimization (e.g., Particle Swarm Optimization and Genetic Algorithms) [4], fuzzy logic / neural network-based controllers, as well as more recently, data-driven methodologies based on machine learning (ML) regression/classification models. At the same time, ensemble learning with a focus on gradient boosting algorithms has proven to be

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

Materials Research Proceedings 64 (2026) 535-543 https://doi.org/10.21741/9781644904091-67

capable of modelling the challenging non-linear relationships between environmental and electrical variables and the optimum operating point [5], [6].

However, even with these developments, a single ML model may not be equally good across the whole operating point. It is also empirically observed that certain models perform better at certain ranges of irradiance or temperature due to bias-variance trade-offs and the regularization mechanism. Inspired by this, hybrid MPPT methods that combine multiple ML models and make them play with their respective advantages are potentially promising.

This paper proposes a hybrid Gradient Boosting Regression–XGBoost MPPT algorithm in which:

- XGBoost is employed at low irradiance, where the signal-to-noise ratio is lower and regularized tree ensembles provide improved generalization;
- GBR is used at medium and high irradiance, where smoother response and high accuracy are obtained with reduced variance.

An irradiance-based switching logic with hysteresis ensures stable transitions between the two models.

Main contributions of this work are:

- Hybrid ML-based MPPT strategy using GBR and XGBoost with irradiance-aware switching technique.
- Construction of an integrated dataset based on simulation and experimental PV data under normal and shaded environments with a thoughtful feature selection.

PV Module Under Study

All simulations in this work are carried out for a commercial crystalline-silicon module Bosch Solar Energy c-Si M 60-225-16, which represents a typical 60-cell module widely used in small and medium PV systems. The module is rated at approximately 225 W maximum power under Standard Test Conditions (G=1000 W/m², T=25∘C). The main electrical and model parameters are summarized in Table I.

The module exhibits an open-circuit voltage Voc≈36v and a short-circuit current Isc≈8.7 A. The maximum power point occurs at Vmp≈28.3 V and Imp≈7.95 A. Temperature coefficients equal to -0.374 %/°C for Voc and 0.071 %/°C for Isc are used to account for the impact of cell temperature on the I-V curve. For the single-diode model, the light-generated current IL, diode saturation current I0, diode ideality factor n, series resistance Rs and shunt resistance Rsh are extracted using the parameter estimation tools described in [7].

Irradiance Dependence (Figure. 1)

Figure 1 shows the effect of irradiance on the module parameters. Showing the I-V and P-V curves at 25 °C when irradiance levels are 1000, 800, 600, 400 and 200 W/m². The short circuit current and the maximum output power increase almost linearly with irradiance, as expected, while the open circuit voltage experiences little variation. These figures illustrate the large maximum power point dependences on irradiation and confirm the necessity for an effective MPPT method which can respond quickly to sudden changes in irradiance.

Materials Research Proceedings 64 (2026) 535-543 https://doi.org/10.21741/9781644904091-67

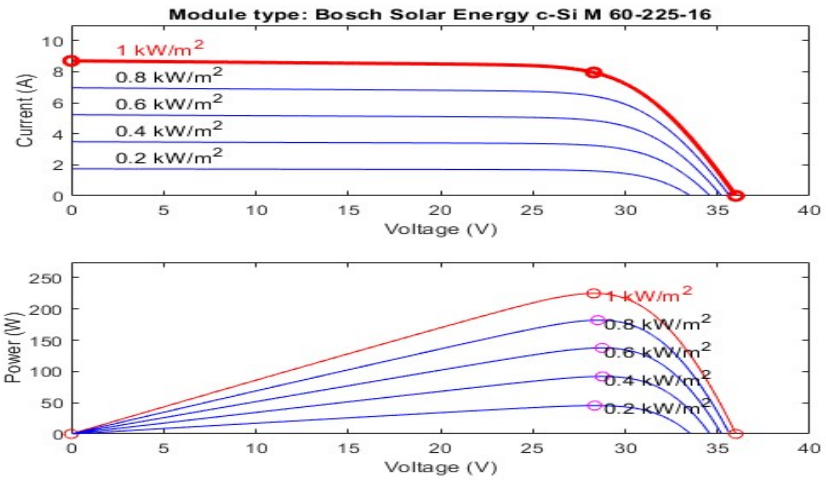


Figure 1. I–V and P–V characteristics of the Bosch c-Si M 60-225-16 module at 25 °C for different irradiance levels (1000–200 W/m²).

Temperature Dependence (Figure. 2)

The influence of temperature is presented in Figure 2, which shows I-V and P-V curves at G=1000 W/m² for two temperature values (25 °C and 45 °C). As the temperature rises, Voc, Vmp and in turn the maximum accessible power decrease significantly. This effect should be taken into account for the MPPT controller to prevent systematic efficiency losses of the module at higher temperatures.

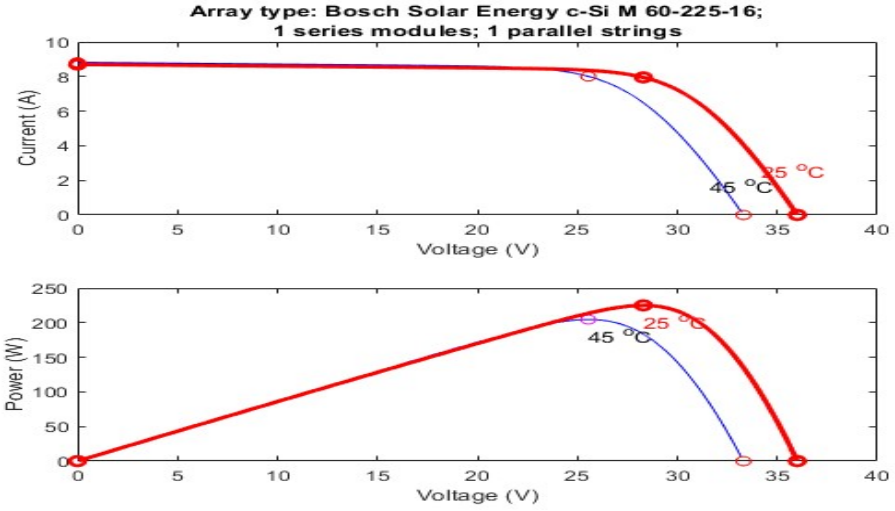


Figure 2. I–V and P–V characteristics of the Bosch c-Si M 60-225-16 module at 1000 W/m² for two temperatures (25 °C and 45 °C).

PV SYSTEM MODELING AND DATASET CONSTRUCTION

In order to assess the hybrid GBR-XGboost MPPT technique, we create a comprehensive model of the PV generator, the power converter and the operating conditions. This part briefly introduces the PV array, as well as scenarios simulated, dataset generation using NASA meteorological data, and equations for the PV equivalent circuit.

Operating Scenarios and Partial Shading

Both under uniform operating conditions and partial shading conditions (PSC), the MPPT method is examined in this paper. PSC occurs while various parts of the array are subject to different irradiance levels induced by clouds, buildings or trees, and the generated I-V and P-V

Materials Research Proceedings 64 (2026) 535-543 https://doi.org/10.21741/9781644904091-67

characteristics have several local maxima with one global maximum [7]. In this case, traditional hill-climbing methods such as Perturb and Observe (P&O) or Incremental Conductance can be stuck at a local peak and are unable to track the global MPP, resulting in considerable power waste [7].

The array is divided into several strings or blocks in order to replicate these effects, and each is exposed to a predetermined level or profile of irradiance. Common patterns consist of uniform and bi-level shading, with the more sophisticated tri-level or random shading distributions. These case studies are realised in MATLAB/Simulink with the single-diode model of Subsection II-A and utilised as benchmarks for evaluating the robustness of the hybrid GBR–XGBoost controller. The training set of the machine learning models is, however, made with realistic irradiance and temperature profiles taken from NASA data, and it is described in the next section.

Dataset Construction Using NASA Irradiance and Temperature

An important aspect of this work is that the training and test datasets for the ML-based MPPT controller are generated from satellite-measured meteorological data instead of short-term local measurements. Long-term series of solar resource and temperature are obtained from the NASA POWER (Prediction of Worldwide Energy Resources) database, which provides gridded data products for renewable energy applications worldwide.

For the geographic coordinates corresponding to the intended PV installation, the following variables are downloaded over a multi-year period:

- **ALLSKY_SFC_SW_DWN**: all-sky surface shortwave downward irradiance, used as plane-of-array irradiance G (W/m²);
- **T2M**: air temperature at 2 m height, used as ambient temperature T (°C).

After a basic cleaning procedure removing missing and non-physical values, every record i is described by a pair (Gi, Ti). These climatic factors are next transferred into an approximate electrical model of the Bosch module, through analytical relationships based on STC data and temperature coefficients, following good engineering practice for c-Si PV modelling.

Let Isc,s, Imp,s, Voc,s and Vmp,s denote the short-circuit current, MPP current, open-circuit voltage and MPP voltage at STC (Gs = 1000 W/m², Ts = 25 °C). Let αIsc and βVoc be the temperature coefficients of Isc and Voc in %/°C provided by the datasheet. They are first converted to a relative current slope and an absolute voltage slope:

- **αrel = αIsc / 100 (per °C),**
- **dVoc/dT = (βVoc / 100) · Voc,s (V/°C).**

The current at the maximum power point Imp(i) for a given (Gi, Ti) is then approximated by

$$\mathbf{Imp(i) = Imp,s \cdot (Gi / Gs) \cdot [\, 1 + \alpha rel \cdot (Ti - Ts) \,],}$$

expressing the near-linear dependence of current on irradiance and its mild increase with temperature. The voltage at the maximum power point is modelled as a linear function of temperature,

$$\mathbf{Vmp(i) \approx Vmp,s + dVmp/dT \cdot (Ti - Ts),}$$

where dVmp/dT is taken close to dVoc/dT, which is a widely used approximation for crystalline-silicon modules. The corresponding maximum power is then

$$\mathbf{Pmp(i) = Vmp(i) \cdot Imp(i).}$$

Materials Research Proceedings 64 (2026) 535-543 https://doi.org/10.21741/9781644904091-67

In this work, the machine-learning input vector and target output at sample i are defined as

$$\mathbf{xi = [\ Gi, Ti\ ]^T,}$$
$$\mathbf{yi = Vmp(i).}$$

The ML models are thus trained to learn the mapping.

$$\mathbf{f : [\ G, T\ ]^T \rightarrow Vmp,}$$

This gives the reference MPP voltage as a simple function of irradiance and temperature. Features are pre-processed before the training stage, removing outliers based on physical bounds for G, T, Vmp and Pmp, and scaling the inputs G and T to zero mean and unit variance in order to enhance numerical conditioning as well as gradient boosting algorithm convergence. The final data-set, which usually includes a few thousand total samples sampling all available seasonal and diurnal variability of the model results, is randomly divided into 70% for training, 15% for validation and 15% testing according to machine learning standards [8]. This analytics-driven NASA-based dataset offers a realistic and physically consistent setting to train and test the proposed hybrid GBR-XGBoost MPPT controller.

III. PROPOSED HYBRID GBR–XGBOOST MPPT ALGORITHM

The purpose of the indicated MPPT approach described in the next section is to harvest complementary capabilities of Gradient Boosting Regression (GBR) and XGBoost to control optimal MPP voltage Vmp out of measured irradiance and temperature. Both models are trained off-line on the NASA-based dataset presented in Section II and combined on-line through a simple irradiance-aware switching block with hysteresis. The hybrid controller produces a reference voltage Vmp,ref , which is followed by the inner DC-DC converter voltage loop.

A. Architecture Overview

The proposed MPPT controller has the following structure:

- **Inputs:**
  - Irradiance G (W/m²) obtained from a sensor or estimator.
  - Temperature T (°C) corresponding to ambient or estimated cell temperature.
- **Core machine-learning models:**
  - A Gradient Boosting Regression (GBR) model, denoted fGBR(G, T).
  - An Extreme Gradient Boosting style (XGBoost) model, denoted fXGB(G, T), both implemented as an ensemble of regression trees [9].
- **Decision block:** An irradiance-based switching logic with hysteresis, which decides whether GBR or XGBoost should be used for the current operating point.
- **Output:** A reference MPP voltage Vmp, ref that is sent to the PI voltage controller of the boost converter, which adjusts the duty cycle so that the PV terminal voltage tracks this reference.

For the experiment on the NASA dataset, from preliminary studies, we observe that while the XGBoost model shows a smaller prediction error in low irradiance areas (where signal to noise ratio is low and regularization helps), it produces a better and coarse fit compared to GBR, which also achieves a slighted smoother fit in medium to high levels of irradiance. This inspires a hybrid approach in which XGBoost is employed at low irradiance and GBR at high irradiance, with the transition determined by an irradiance threshold and a hysteresis bandwidth.

B. Online Hybrid MPPT Algorithm

During real-time operation, the hybrid GBR-XGBoost MPPT runs at each sampling instant to compute the reference MPP voltage Vmp,ref . The decision is based solely on the current irradiance level and on the previous mode (GBR or XGBoost), which enables a simple and low-cost implementation.

The proposed online procedure is summarized in Algorithm 1

Algorithm 1: Hybrid GBR–XGBoost MPPT Algorithm

Input: Measured irradiance $G(k)$ [W/m^2]
Measured (or estimated) temperature $T(k)$ [°C]
Previous mode $M(k-1) \in \{\text{GBR}, \text{XGB}\}$
Trained models $f_{\text{GBR}}, f_{\text{XGB}}$
Normalization parameters $\mu_G, \sigma_G, \mu_T, \sigma_T$
Irradiance threshold G_{th} and hysteresis band ΔG
Output: Reference MPP voltage $V_{mp,\text{ref}}(k)$ [V]
Updated mode $M(k) \in \{\text{GBR}, \text{XGB}\}$

// Step 1: Acquire measurements
Read current irradiance and temperature: $G \leftarrow G(k)$, $T \leftarrow T(k)$

// Step 2: Normalize input features
$G_n \leftarrow \dfrac{G - \mu_G}{\sigma_G}$
$T_n \leftarrow \dfrac{T - \mu_T}{\sigma_T}$

// Step 3: Select model using irradiance and hysteresis
if $G < G_{\text{th}} - \Delta G$ **then**
  $M(k) \leftarrow$ XGB // Low irradiance ⇒ use XGBoost
else
  if $G > G_{\text{th}} + \Delta G$ **then**
    $M(k) \leftarrow$ GBR // High irradiance ⇒ use GBR
  else
    $M(k) \leftarrow M(k-1)$ // Keep previous mode (avoid chattering)

// Step 4: Predict reference MPP voltage
if $M(k) = GBR$ **then**
  $V_{mp,\text{ref}}(k) \leftarrow f_{\text{GBR}}(G_n, T_n)$
else
  $V_{mp,\text{ref}}(k) \leftarrow f_{\text{XGB}}(G_n, T_n)$

// Step 5: Send reference to DC–DC converter controller
Use $V_{mp,\text{ref}}(k)$ as the reference voltage for the inner PI voltage control loop of the boost converter so that $V_{pv}(k) \rightarrow V_{mp,\text{ref}}(k)$

// Step 6: Next sampling instant
Set $k \leftarrow k+1$ and repeat from Step 1

IV. RESULTS AND DISCUSSION

This section provides a study of the performance under the uniform irradiance case and compares it with the previously reported GBR-only and XGBoost-only methods [9]. Simulations are performed in MATLAB/Simulink with the Bosch c-Si M 60-225-16 module model and NASA-

based irradiance and temperature profiles as shown in Section II. Five irradiance levels (1000, 800, 600, 400 and 200 W/m²) are used, and the corresponding PV power and load power P load for each one are recorded. The results are analyzed in terms of power tracking accuracy to the theoretical maximum power, dynamic response and the overall MPPT efficiency. In the discussion, it also presents the findings in connection with the most recent data-driven MPPT development, referring to both ensemble learning and adaptive regression [9].

A. Performance of the Proposed Hybrid GBR–XGBoost MPPT

Fig. 3 shows the dynamic response of the proposed hybrid module to the same irradiance step sequence. Where the PV power output and load power demanded are represented as a time function. The five irradiations of 1000, 800, 600, 400 and 200 W/m² can be seen very clearly, and both track lines (PV power and Load power) almost overlap at the steady state, concluding a good precision MPPT as well as high conversion efficiency.

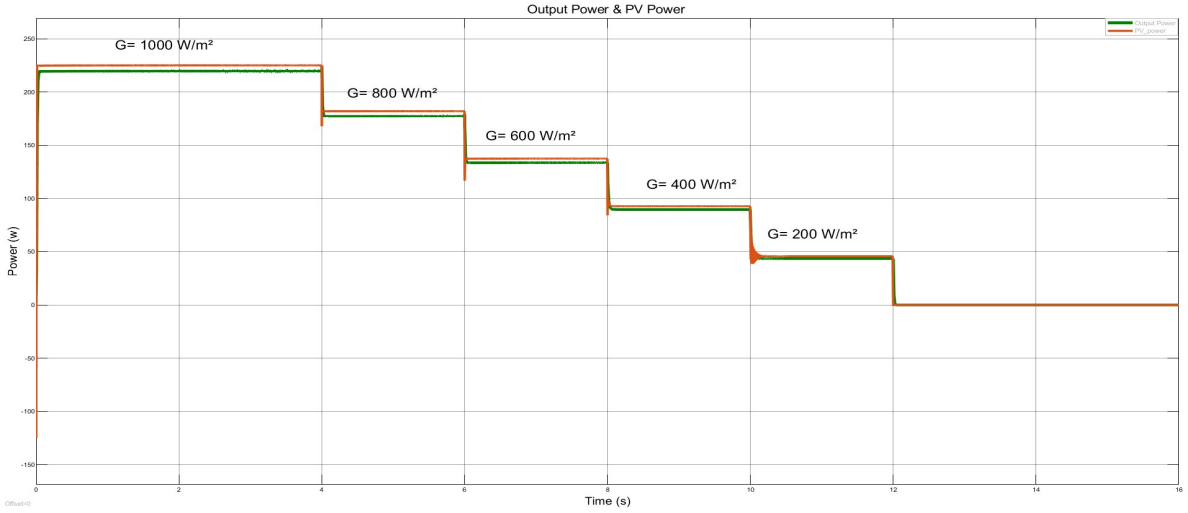


Figure 3. PV power and load power of the proposed hybrid GBR-XGBoost MPPT for a sequence of irradiance steps (1000–200 W/m²).

In Table III, the steady-state powers obtained by the hybrid method are shown for each irradiance level. We present the same theoretical powers of the fits as in Table II for direct comparison.

Table III - Steady-state PV and load power of the proposed hybrid MPPT method

Irradiance (W/m²)	Theoretical PV Power (W)	Hybrid PV Power (W)	Hybrid Load Power (W)
200	45.44	45	43.5
400	92.02	92	89.6
600	137.71	137	133.5
800	182.09	181.4	175
1000	224.98	223	219

From Table III, we can see that the hybrid controller can regulate the PV power to the theoretical MPP very closely for all irradiance levels with deviations lower than 2 W at 200 W/m² and 2-3 W at 1000 W/m². The difference between the PV power and load power is also low (usually 2-6 W), accounting mainly for converter losses. More interestingly, the hybrid scheme greatly

enhances the utilization of the PV model compared with that of Table II's single-model GBR and XGBoost approaches.

B. MPPT Efficiency Evaluation

Using the hybrid load powers of Table III and the theoretical powers of Table II, the efficiencies for each irradiance level and the global efficiency are listed in Table IV.

Table IV - MPPT efficiencies of the proposed hybrid method (based on load power)

Irradiance (W/m²)	Theoretical PV Power (W)	Hybrid Load Power (W)	MPPT Efficiency η_i (%)
200	45.44	43.5	95.73
400	92.02	89.6	97.37
600	137.71	133.5	96.94
800	182.09	175	96.11
1000	224.98	219	97.34
Global	**Σ = 682.24**	**Σ = 660.60**	**96.83**

As a result, the achieved global MPPT efficiency provided by the proposed hybrid controller is approximately 96.8 %, which is much higher than 83.8 % and 78.7 % of GBR-only and XGBoost-only approaches in [9]. The advantage is especially pronounced at low irradiance (200-400 W/m²), where the hybrid method preserves efficiencies above 95%, while the other techniques experienced significant tracking errors. These enhancements demonstrate the appropriateness of mixing multiple learning models in an irradiance-aware way, as recently observed for stacked ensemble MPPT teens [10], and the latest adaptive regression-based controllers like ALR-CWM [11].

In summary, the results suggested that the proposed hybrid GBR-XGBoost MPPT controller offers precise tracking under a broad range of irradiance conditions and possesses fast dynamic response with well-damping character, as well as high energy capture. This underlines that it is a potential data-driven alternative to traditional algorithms and single-model machine learning systems in realistic PV systems.

References

[1] Esram, T., & Chapman, P. L. (2007). Comparison of photovoltaic array maximum power point tracking techniques. IEEE Transactions on energy conversion, 22(2), 439-449. https://doi.org/10.1109/TEC.2006.874230

[2] Femia, Nicola, et al. "Optimization of perturb and observe maximum power point tracking method." IEEE transactions on power electronics 20.4 (2005): 963-973. https://doi.org/10.1109/TPEL.2005.850975

[3] Hohm, D. P., and M. E_ Ropp. "Comparative study of maximum power point tracking algorithms." Progress in photovoltaics: Research and Applications 11.1 (2003): 47-62. https://doi.org/10.1002/pip.459

[4] Ishaque, Kashif, and Zainal Salam. "A review of maximum power point tracking techniques of PV system for uniform insolation and partial shading condition." Renewable and Sustainable Energy Reviews 19 (2013): 475-488. https://doi.org/10.1016/j.rser.2012.11.032

[5] Friedman, Jerome H. "Greedy function approximation: a gradient boosting machine." Annals of statistics (2001): 1189-1232. https://doi.org/10.1214/aos/1013203451

[6] Chen, Tianqi. "XGBoost: A Scalable Tree Boosting System." Cornell University (2016). https://doi.org/10.1145/2939672.2939785

[7] Ahmed, Jubaer, and Zainal Salam. "A critical evaluation on maximum power point tracking methods for partial shading in PV systems." Renewable and Sustainable Energy Reviews 47 (2015): 933-953. https://doi.org/10.1016/j.rser.2015.03.080

[8] Goodfellow, Ian, et al. Deep learning. Vol. 1. No. 2. Cambridge: MIT press, 2016.

[9] R. Amine, N. El Barbri and H. Ameziane, "Optimizing Photovoltaic Performance: A Comparative Analysis of XGBoost and Gradient Boosting for MPPT," 2025 International Conference on Electrical Systems & Automation (ICESA), Troyes, France, 2025, pp. 1-6, doi: 10.1109/ICESA66763.2025.11280867. https://doi.org/10.1109/ICESA66763.2025.11280867

[10] R. Amine, N. El Barbri and H. Ameziane, "Stacked Ensemble Machine Learning for MPPT Under Partial Shading: A High-Accuracy Data-Driven Approach," 2025 International Conference on Electrical Systems & Automation (ICESA), Troyes, France, 2025, pp. 1-6, doi: 10.1109/ICESA66763.2025.11280669. https://doi.org/10.1109/ICESA66763.2025.11280669

[11] Amine, R., & Ameziane, H. (2025, June). An Enhanced MPPT Algorithm Based on Adaptive Linear Regression with Contextual Weight Memory (ALR-CWM) for PV Systems. In 2025 17th International Conference on Electronics, Computers and Artificial Intelligence (ECAI) (pp. 1-4). IEEE. https://doi.org/10.1109/ECAI65401.2025.11095559