



Journal of Renewable Energies

Revue des Energies Renouvelables

journal home page : <https://revue.cder.dz/index.php/rer>

Research Paper

Use an artificial intelligence method (Machine Learning) for analysis of the performance of photovoltaic systems

Hichem Hafdaoui *, El Amin Kouadri Boudjelthia, Salim Bouchakour, Nasreddine Belhaouas

Centre de Développement des Energies Renouvelables, CDER, BP 62 Route de l'Observatoire, Bouzaréah, 16340, Algiers, Algeria

ARTICLE INFO

Article history:

Received 30 May 2022

Accepted 28 September 2022

Keywords:

Photovoltaic (PV)

Monitoring

Classification

Artificial intelligence

Maximum power

ABSTRACT

The performance of a photovoltaic system depends on several parameters such as temperature, clouds, and the season, which makes the study of PV performance from monitoring databases very complex given the size of the information and the complexity of the phenomena involved. This article applies an artificial intelligence (AI) method based on machine learning (ML). For more efficient analysis, the Support Vector Machine (SVM) is used to simplify and optimize the processing of these data for the study of the performance of PV systems. More precisely, we group a multi-class data variable according to the needs of the analysis using SVMs. In this article, we present all the stages of data processing based on the application of artificial intelligence (AI). We present as an example the results obtained in the study of the performance of a 150W monocrystalline photovoltaic (PV) module after one year of monitoring.

* Corresponding author, E-mail address: hichemhafdaoui@gmail.com

Tel.: +213 675273024

ISSN: 1112-2242 / EISSN: 2716-8247



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Based on a work at <http://revue.cder.dz>.

1. Introduction

Au cours des deux dernière décennies, le marché du photovoltaïque (PV) a connu une grande croissance passant de 40.279 MW à 578.553 MW de capacité installée entre 2010 et 2019 [1]. Cette croissance est due à la réduction significative des couts des modules PV et sur l'amélioration des rendements et des performances de ces modules. Cependant, ces paramètres de fiabilités sont devenus actuellement un facteur déterminant dans la rentabilité d'un projet solaire après la baisse considérable des prix et de la marge bénéficiaire. Une étude approfondie a été menée dans ce sens par le Laboratoire national des énergies renouvelables (NREL) qui a montré qu'une augmentation du taux de dégradation annuel des modules de 0,5% à 1,5% entraînerait une augmentation de 13,6% du coût réel en électricité du site (NREL, 2019 – assumptions: 4.5% real discount rate. \$10/kW-yr average O&M expense. \$1.0/W(DC) capital cost]. Cet impact est également démontré dans un sondage réalisé par PVEL (PV Evolution Labs) auquel 70% des personnes interrogées ont répondu qu'une sous-performance de 3 à 6% est suffisante pour que leurs projets soient non viables sur le plan financier (PVEL, Solar Asset Management North America, 2016).

Dans ce contexte l'évaluation des performances pour un meilleur choix technologique et le suivie des centrales pour une détection précoce des défauts et un diagnostic en temps réel appropriés sont essentiels, non seulement pour réduire les coûts et la maintenance à temps, mais aussi pour éviter les pertes d'énergie, les dommages à l'équipement et les risques pour la sécurité.

Durant les dernières années, plusieurs travaux ont été menés pour évaluer les performances des systèmes photovoltaïques dans différentes régions du globe. D'autres ont proposé des méthodes d'évaluation et de gestion de la base de données pour mieux apprivoiser la quantité importante des données à traiter et la complexité des phénomène intervenant dans l'évaluation des performances d'un systèmes photovoltaïque.

Dans cette optique, l'IA a offert des solutions prometteuses au problème du photovoltaïque, en particulier la détection des problèmes dans les grandes centrales, permettant une plus grande flexibilité qui la rend plus robuste que les techniques traditionnelles. L'AO est un domaine d'étude de l'IA, son idée de base est l'étude d'algorithmes informatiques qui peuvent s'améliorer automatiquement grâce à l'expérience et à l'utilisation de données. L'AO consiste à créer d'énormes modèles de réseaux de neurones capables de faire des choix précis basés sur des données ; l'AO convient aux situations où les données sont complexes.

Les algorithmes AO progressent de jour en jour depuis très longtemps dans l'amélioration des algorithmes des problèmes en photovoltaïque et se sont développés dans de nombreux domaines.

Ces dernières années, plusieurs nouveaux travaux de recherche et de nombreux des méthodes d'apprentissage automatique et d'apprentissage en profondeur ont été appliqués avec succès dans le domaine du diagnostic de pannes pour le générateur photovoltaïque. YUN, Liu et al. [2] ont proposé l'utilisation des algorithmes d'apprentissage automatique pour établir un modèle d'arbre de décision d'exploration de données pour les données de fonctionnement des modules photovoltaïque, et d'utiliser le modèle pour prédire la cause de la panne d'un module PV photovoltaïque. HONG, Feng et al. [3] ont exploité les images satellitaires pour la détection et la prédiction des défauts dans les grandes centrales photovoltaïques. Les images du champ PV visible et infrarouge sont prises dans les mêmes conditions par une double caméra infrarouge à basse altitude. Liu, Yongjie, et al. [4] applique la technique de kernel d'apprentissage automatique (KEAO) et SVM pour détecter les défauts basé sur la courbe I-V sous les conditions STC. BANSAL, Neha et al. [5] ont développé des modèles de fiabilité en fonction de la quantité de données de défaillance disponibles et les ont utilisés pour la détection précoce de la dégradation des centrales photovoltaïques. Les résultats de [6] et [7] présentent une méthode basée sur les réseaux de neurones artificiels (ANN) pour la détection des anomalies, qui est basée sur la comparaison entre les valeurs mesurées et prédites de la production de l'énergie.

Dans ce travail, nous appliquons une technique d'AO basée sur la classification par SVM pour l'analyse optimale de la base de données d'un système photovoltaïque. Elle est basée sur la régression linéaire pour la création des intervalles (classes) homogènes ou non-homogènes.

Les SVM ont été développées dans les années 1990, elles visent à regrouper les données en classes à l'aide d'une frontière, pour que la distance entre les différents groupes de données et la frontière qui les sépare soit maximale. Cette distance est également appelée "marge" et les SVM sont ainsi appelés "séparateurs à marge large", les vecteurs supports étant les données les plus proches de la frontière [10-13].

Dans cet article, la classification est utilisée pour réduire le temps de traitement des grandes bases de données, en remplacement des méthodes manuelles qui peuvent durer très longtemps. Elle est donc une alternative qui vise l'optimisation du temps de traitement. Dans ce travail de recherche, nous appliquons cette méthode (SVM) pour analyser les résultats d'un module PV photovoltaïque de 150 W pendant une année de monitoring.

2. Machine à vecteurs de support

SVM est un algorithme d'apprentissage automatique supervisé qui peut être utilisé à la fois pour les défis de classification ou de régression. Il est utilisé dans les problèmes de classification des grandes bases de données [13,14]. Dans l'algorithme SVM, nous traçons chaque élément de données sous la forme d'un point dans un espace à n dimensions (où n est le nombre de caractéristiques), la valeur de chaque caractéristique étant la valeur d'une coordonnée particulière. Ensuite, nous effectuons une classification en trouvant l'hyper-plan qui différencie les deux classes (Fig.1) ; nous supposons que nos données ont deux classes linéairement séparables avec un hyperplan qui peut les séparer [12-15]. Dans le cas particulier de l'espace bidimensionnel, un hyperplan devient une ligne droite. Fig. 1.

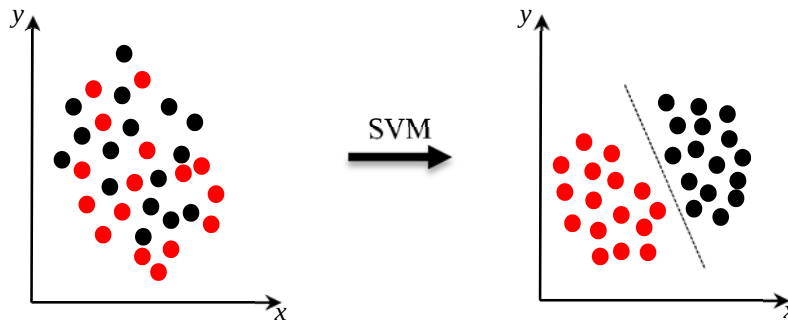


Fig. 1. Il existe de nombreux hyperplans qui peuvent diviser les données en deux classes.

La fonction SVM utilise une méthode d'optimisation pour identifier les vecteurs de support x_i poids u_i et biais b qui sont utilisés pour classer les vecteurs x selon l'équation suivante [8,9] :

$$c = \sum_i \alpha_i K(x_i, x) + b$$

La théorie sur la classification par SVM est bien expliquée dans les réf. [11-15].

3. Méthode de mesure et traitement de la base de données

Le système d'acquisition de données offre une large gamme de conditions de mesure. Les données mesurées contiennent des paramètres fournis par différentes unités de mesure qui sont enregistré à différents intervalles de temps. Ensuite, ils sont importés dans une base de données commune et préparés pour une analyse plus approfondie par transformation en structures de données appropriées pour des requêtes rapides [16].

Nos mesures ont été effectuées par une acquisition de données (Keysight 34972A). Le système collecte les mesures des différents capteurs à savoir (I_{sc} , V_{oc} , P_{max} , I_{mp} , V_{mp} , T_m , T_{amb} ,

G, WD, WS, Date , Hour, FF). Le système enregistre les données entre 05h00 et 21h00.

Fig. 2 montre le diagramme de système de mesure. Cet article s’articule sur les étapes de traitement et d’analyse de ces données.

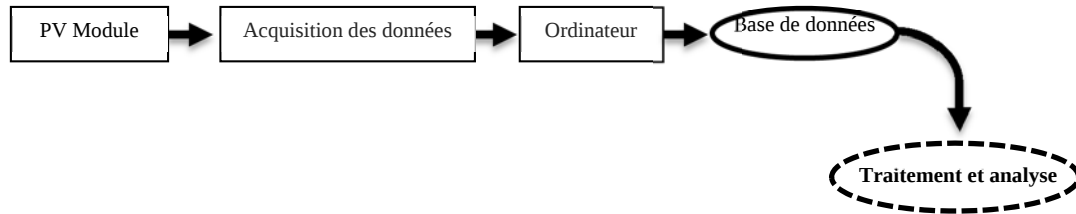


Fig 2. Diagramme de système de la mesure

Les données obtenues pendant une année ont été analysées par la méthode SVM pour faciliter leur étude. Sur la base de cette méthode (SVM Classifier), nous avons regroupé les variables dans des groupes (classes) (Fig. 3). Le diagramme de la méthode utilisé dans ce travail est présenté dans la réf [1]. L’irradiation solaire est groupée par 13 classes et l’heure en 16 classes. Les résultats obtenus sont présentés sous forme de cartes heatmap afin de simplifier et faciliter leur analyse.

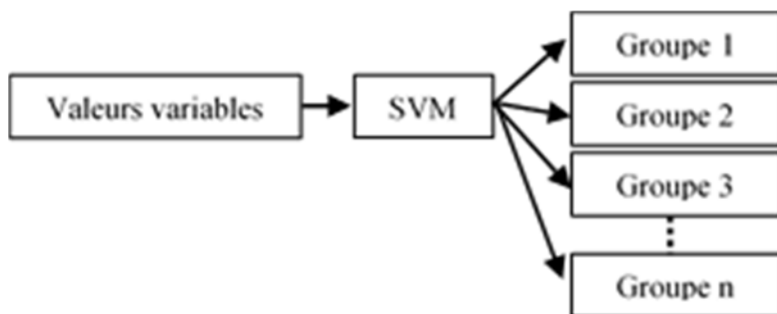


Fig. 3. SVM multi-classes unique

4. Résultats et discussion

Dans cet article nous présentons l’analyse et la classification par SVM de la base de données du monitoring d’un module photovoltaïque de 150Wc pendant une période d’une année avec un pas de mesure de 1 minute. Les paramètres mesurés dans notre système sont la courbes IV du module PV avec les paramètres électriques (P_{max} , I_{sc} , V_{oc} , I_{mp} , V_{pm} , FF), la température du module PV mesurée sur la face arrière et les paramètres du weather (l’irradiation solaire sur le plan incliné, la température ambiante, la vitesse et direction du vent)

Fig. 4 montre la puissance maximale fournie par le module en fonction de l’heure et du mois.

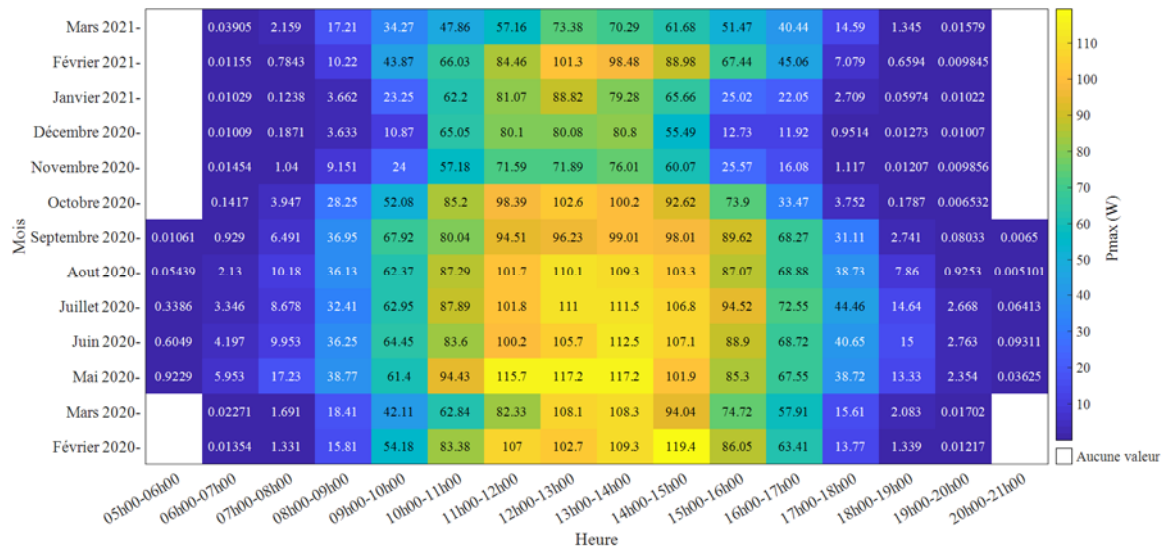


Fig 4. Évolution de P_{max} (W) fournie sur une journée en fonction du mois.

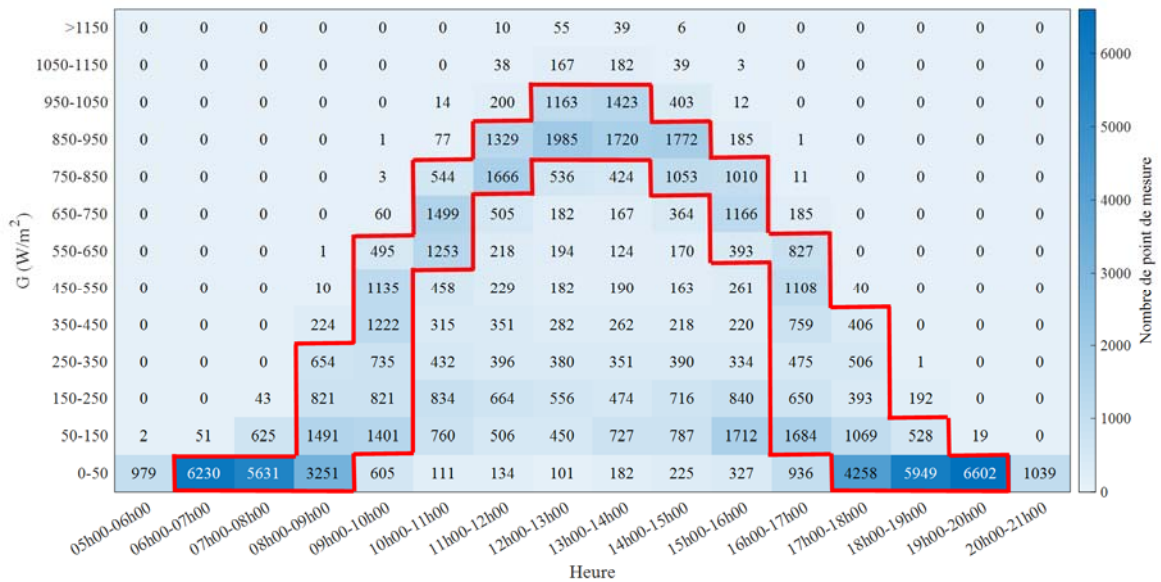


Fig. 5. Nombre de point de mesure en fonction de G (W/m^2) et l'heure pour une année.

Fig. 4. montre l'évolution de la puissance mensuelle moyenne sur une journée. La production photovoltaïque est importante entre 14h00 et 15h00 heure locale. Elle est avoisine les 79% de puissance crête durant le mois de février 2020 et entre 12h00 et 14h00 heure et environ 78% de puissance crête durant le mois de mai 2020. Cette valeur baisse en automne à cause de la présente des nuages. Cette figure montre aussi l'évolution mensuelle de la durée d'ensellement. Fig. 5 montre les nombre de point de mesure obtenu en fonction du temps et de l'irradiation.

Cette technique nous permet d'identifier les zones d'intérêt pour une analyse des performances et les zones qui représentent des valeurs non essentielles ou des phénomènes isolés. En effet, nous remarquons les carrés en bleu foncé (lignes rouges) correspondent à la courbe de

l'irradiation solaire d'un ciel claire, tandis que les point de mesure à l'intérieur de cette cloche correspondent principalement aux mesures sous conditions de nuage. Cependant, les valeurs très élevé supérieur à 1150W/m^2 peuvent correspondre à des phénomènes de réflexions ou à un passage sur le bord d'un nuage blanc. Cette présentation nous permet de délimiter les zones d'étude et réduire ainsi les valeurs qui peuvent induire en erreurs dans le calculs des différents paramètres.

Figs. 6 et 7 représentent l'évolution, en fonction de l'irradiation solaire reçue et du temps, de la puissance maximale fournie et de la température du module PV respectivement.

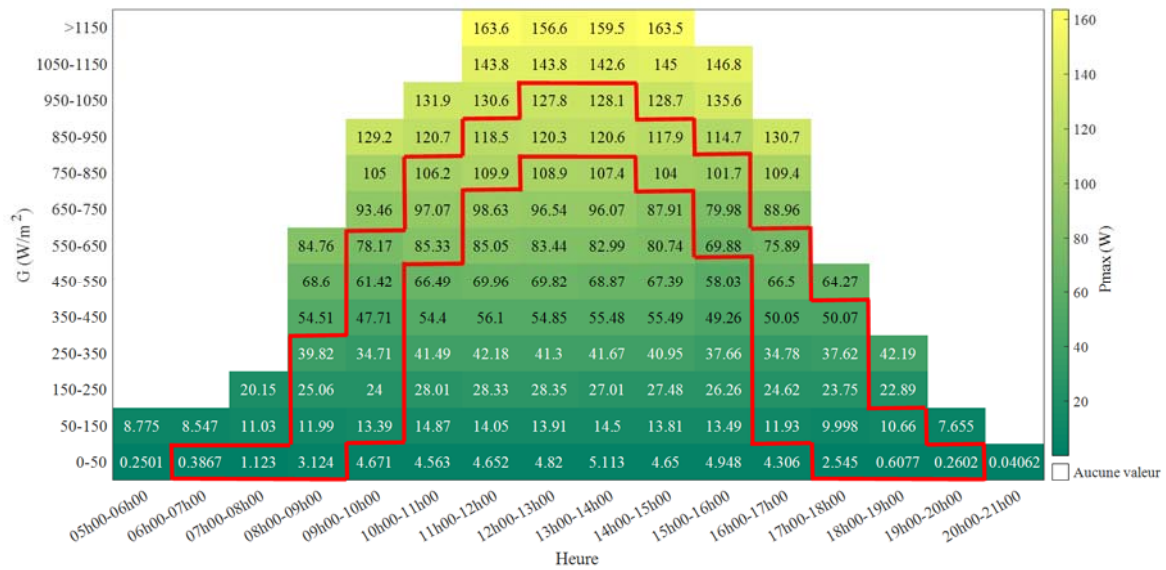


Fig. 6 . Pmax (W) en fonction de G (W/m^2) et l'heure pour une année.

Fig. 6. montre l'évolution de la puissance maximale fournie durant la journée. Sur la zone d'intérêt nous notons que la puissance maximale atteinte dans la journée est de l'ordre de 128W pour une irradiation comprise entre 950 W/m^2 et 1050 W/m^2 . Cette puissance correspond à environ 85% de puissance crête du module PV.

D'autre part, nous constatons que module PV a fonctionné environ 35% du temps sous conditions nuageuse dans la période entre 13h00 et 14h00. Ces informations peuvent servir de référence dans le dimensionnement d'un système PV, notamment pour le choix de l'onduleur. Les valeurs supérieures à 140W sont très marginales (selon la Fig. 6). Cela s'explique par la hausse de la température du module PV durant cette période (Fig. 7) la période du 13h00 à 14h00 où l'irradiation est comprise entre 950W/m^2 et 1050W/m^2 généralement.

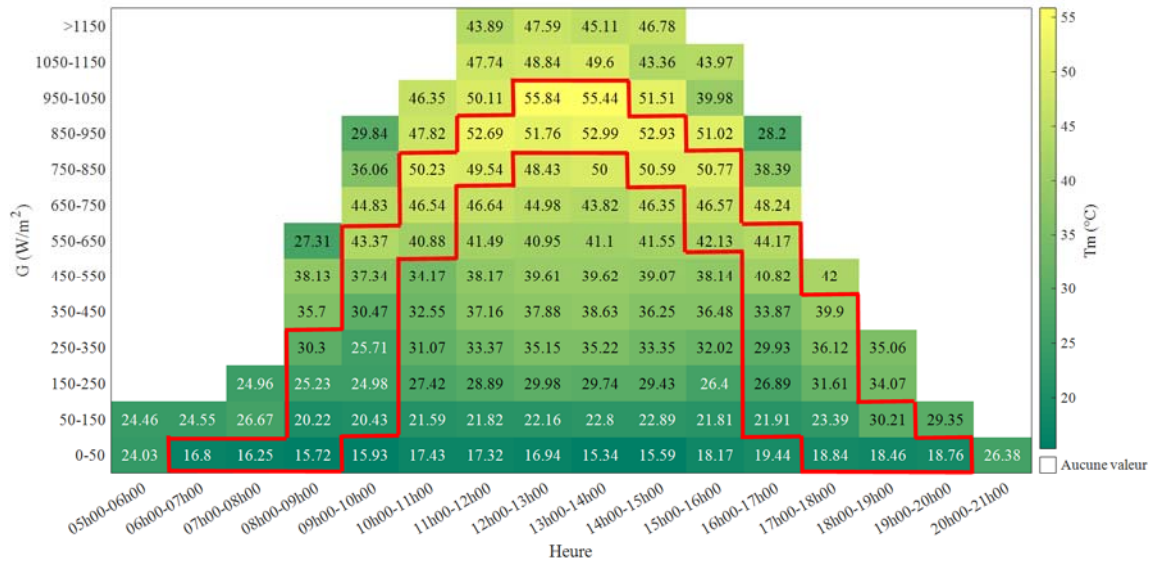


Fig. 7. T_m (°C) en fonction de G (W/m²) et l'heure pour une année.

Par ailleurs on remarque que sur la zone d'intérêt (qui correspond à un ciel clair dans Fig. 7 et qui est délimité par une ligne rouge) la température de l'après-midi est plus importante que celle de la matinée à irradiation équivalente. Cette écart de la température est dû à l'évolution journalière de la température des modules PV où il débute la journée à la température ambiante et se réchauffe au fur et à mesure de la croissance de l'irradiation solaire. Par contre durant l'après-midi le module est déjà chaud où il a atteint sa température maximale puis se refroidi longuement avec la baisse de l'irradiation solaire.

Fig. 8 montre le nombre de point de mesure en fonction de l'irradiation et de la température du module PV.

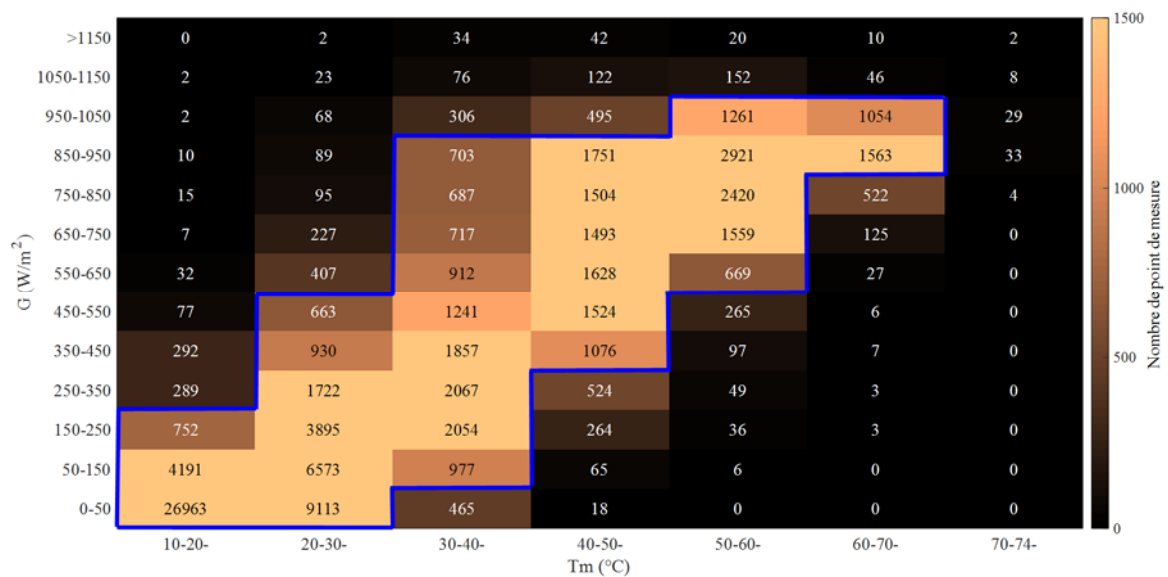


Fig. 8. Nombre de point de mesure en fonction de G (W/m²) et T_m (°C) pour une année.

Fig.8. montre l'évolution de la température du module PV en fonction de l'irradiation où nous avons une concentration des points de mesures supérieur à 650 (plus d'une journée) qui se situe sur une trajectoire linéaire. D'autre part, nous remarquons l'existence de point de mesure correspondant à des valeurs aberrantes comme une température du module PV comprise entre 60°C et 70°C correspondant à une irradiation de 250W/m² et 350W/m². Croiser ces valeurs avec des valeurs constatées sur la Fig. 7 (deux cases avec une température du module (35,15 et 35,22) supérieur à la moyenne dans la même irradiation) montre que ces températures sont probablement dues à des passages éphémères de nuages où l'irradiation diminue brusquement alors que la température diminue d'une manière beaucoup plus lente. Cet exemple montre l'intérêt de cette présentation qui réside essentiellement dans le ciblage des zones d'intérêt qui représentent les phénomènes dominant dans l'analyse des performances d'un systèmes PV et l'isolation des phénomènes marginaux mais qui peuvent induire des erreurs d'interprétation.

Fig. 9 montre l'évolution de la puissance maximale en fonction de l'irradiation et de la température du module PV. Elle montre clairement les grandes tendances classiques du comportement d'un module photovoltaïque tel que l'évolution décroissante de la puissance maximale avec l'évolution de la température du module PV. Ces valeurs sont représentatives comme le montre la Fig. 10 qui compare l'évolution de la puissance maximale fournie en fonction de la température à partir des figures heatmap et sa comparaison avec les valeurs obtenues par des mesures directes de l'évolution de Pmax en fonction de la température du module PV effectués sur deux autres modules PV de la même marque. Ces derniers ont été mesuré selon la section 4.4 Measurement of température coefficients (MQT 04) de la norme IEC 61215 (Terrestrial photovoltaic (PV) modules – Design qualification and type approval – Part 2: Test procedures).

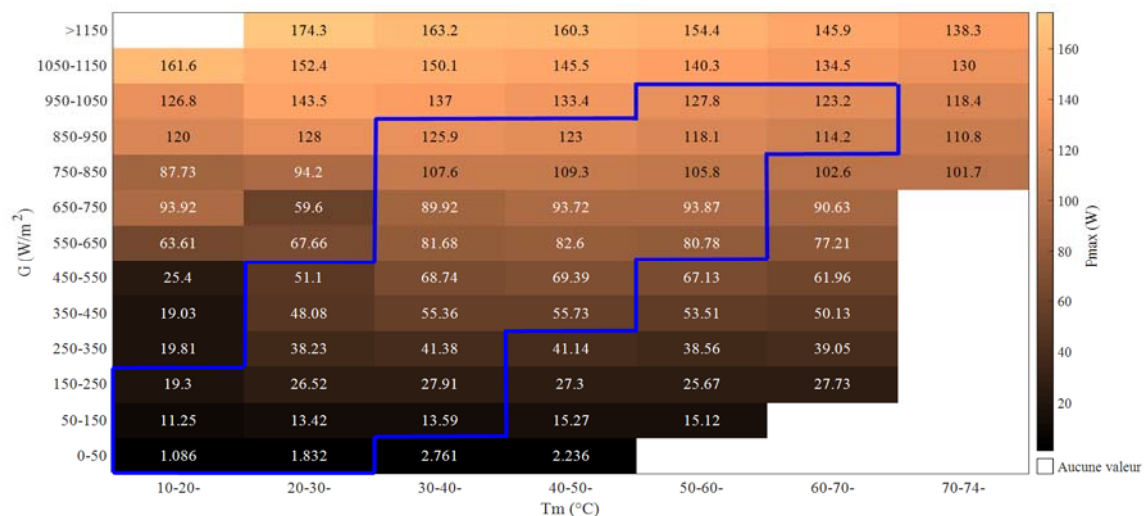


Fig. 9. Evolution de Pmax (W) fournie en fonction de G (W/m²) et Tm (°C) pour une année.

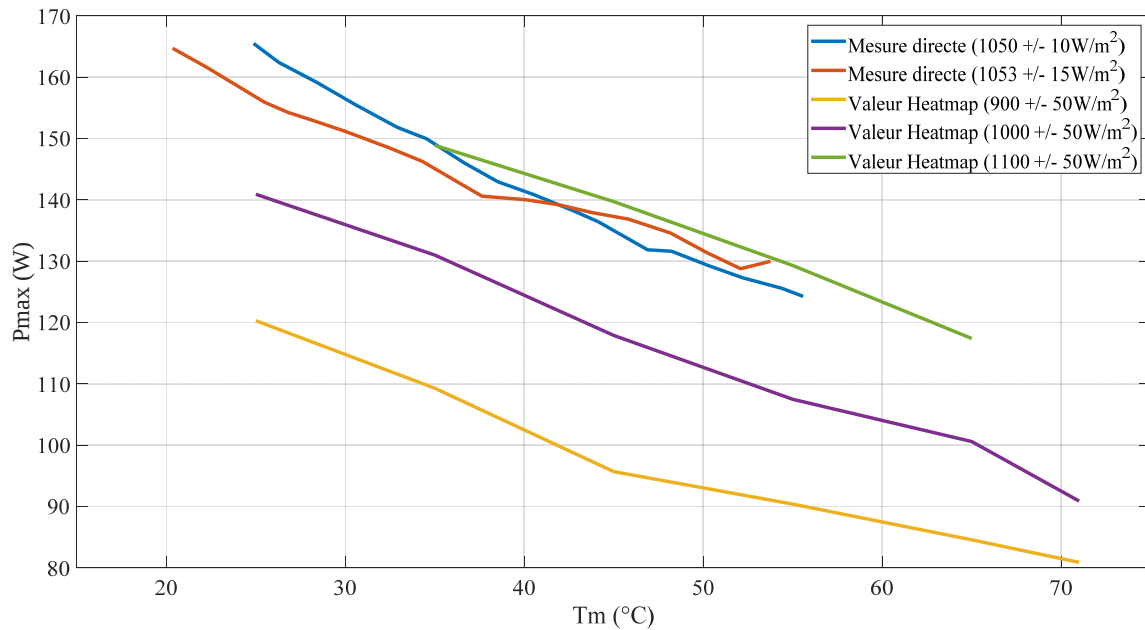


Fig. 10. Évolution de Pmax (W) fournie en fonction de Tm (°C).

Le coefficient de température obtenu par mesure direct de l'évolution de la puissance mesuré en fonction de la température selon la norme IEC 61215 donne pour le module 1 une valeur de 0,42%/°C, pour le module 2 une valeur de 0,32%/°C et pour le module 3 placé dans le système du monitoring (à partir des valeurs du heatmap) une valeur de 0,36%/°C.

Tableau 1 : coefficient de température des modules photovoltaïques.

PV module	Coefficient de température (%/°C)
PV module 1 à partir des valeurs du heatmap	0.36
PV module 2 par mesure direct selon la norme IEC 61215	0.32
PV module 3 par mesure direct selon la norme IEC 61215	0.42

La comparaison entre les valeurs obtenu selon la norme IEC 61215 et les valeurs de la classification SVM montre que ces dernières peuvent constituer une base fiable pour l'analyse des paramètres PV et que les erreurs dues aux différents phénomènes lié aux mesures outdoor sont atténuées par la quantité importantes des données limitées dans des conditions des des zones d'intérêts spécifiques.

5. Conclusion

Nous avons apporté dans cet article une méthode de classification par l'intelligence artificielle pour l'analyse des performances des systèmes photovoltaïques. Les résultats obtenus montrent que cette méthode est efficace pour le traitement des grandes bases de données du monitoring

photovoltaïque. Elle permet également d'effectuer des analyses ciblées selon l'objectif de l'analyse, d'une part, et l'importance des informations à analyser, d'autre part.

En effet, la complexité des phénomènes qui influence les performances des systèmes photovoltaïques est apprivoisée par le ciblage des zones selon l'importance des phénomènes mises en jeux et l'identification et l'isolation des phénomènes marginaux.

Cette technique de classification permet également d'avoir une vue d'ensemble sur le comportement des systèmes photovoltaïque et peut être utilisé dans la comparaison des différentes technologies photovoltaïques ou pour un dimensionnement optimisé.

Nous avons également montré la fiabilité des paramètres obtenus avec le ciblage des conditions de mesure par l'exemple du calcul du coefficient de température. En effet, les résultats obtenus par la méthode de classification sont comparables aux résultats obtenus par les mesures directs selon la norme IEC 61215.

6. Références

- [1] Hafdaoui, Hichem, et al. "Analyzing the performance of photovoltaic systems using support vector machine classifier." *Sustainable Energy, Grids and Networks* 29 (2022): 100592. doi : 10.1016/j.segan.2021.100592.
- [2] YUN, Liu, BOFENG, Yan, DAN, Qian, et al. Research on Fault Diagnosis of Photovoltaic Array Based on Random Forest Algorithm. In: 2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA). IEEE, 2021. p. 194-198. doi : 10.1109/ICPECA51329.2021.9362559.
- [3] HONG, Feng, SONG, Jie, MENG, Hang, et al. A novel framework on intelligent detection for module defects of PV plant combining the visible and infrared images. *Solar Energy*, 2022, vol. 236, p. 406-416. DOI: 10.1016/j.solener.2022.03.018.
- [4] LIU, Yongjie, DING, Kun, ZHANG, Jingwei, et al. Intelligent fault diagnosis of photovoltaic array based on variable predictive models and I–V curves. *Solar Energy*, 2022, vol. 237, p. 340-351. DOI: 10.1016/j.solener.2022.03.062.
- [5] BANSAL, Neha, JAISWAL, Shiva Pujan, et SINGH, Gajendra. Long term performance assessment and loss analysis of 9 MW grid tied PV plant in India. *Materials Today: Proceedings*, 2022. doi : 10.1016/j.matpr.2022.01.263.
- [6] DE BENEDETTI, Massimiliano, LEONARDI, Fabio, MESSINA, Fabrizio, et al. Anomaly detection and predictive maintenance for photovoltaic systems. *Neurocomputing*, 2018, vol. 310, p. 59-68. doi : 10.1016/j.neucom.2018.05.017.

- [7] POLO, Fernando A. Olivencia, BERMEJO, Jesús Ferrero, FERNÁNDEZ, Juan F. Gómez, et al. Failure mode prediction and energy forecasting of PV plants to assist dynamic maintenance tasks by ANN based models. *Renewable Energy*, 2015, vol. 81, p. 227-238. doi : 10.1016/j.renene.2015.03.023.
- [8] RUIZ-GONZALEZ, Ruben, GOMEZ-GIL, Jaime, GOMEZ-GIL, Francisco Javier, et al. An SVM-based classifier for estimating the state of various rotating components in agro-industrial machinery with a vibration signal acquired from a single point on the machine chassis. *Sensors*, 2014, vol. 14, no 11, p. 20713-20735. DOI: 10.3390/s141120713.
- [9] FARID, Nahla, ELBAGOURY, Bassant, ROUSHDY, MOHAMED, et al. A comparative analysis for support vector machines for stroke patients. *Rec Adv Inf Sci*, 2013, p. 71-76.
- [10] RAY, Sunil, BANSAL, S., GUPTA, A., et al. Understanding Support Vector Machine algorithm from examples (along with code). *Analytics Vidhya*, 2017, vol. 13, p. 19.
- [11] C.J.C. Burges, “A Tutorial On Support Vector Machines for Pattern Recognition,” *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121-167, 1998. DOI: 10.1023/A:1009715923555.
- [12] B. Schoˆlkopf, I. Guyon, and J. Weston, “Statistical Learning and Kernel Methods in Bioinformatics,” *Artificial Intelligence and Heuristic Methods in Bioinformatics* 183, P. Frasconi and R. Shamir, eds., pp. 1-21, IOS Press, 2003.
- [13] S. Gunn, “Support Vector Machines for Classification and Regression,” *ISIS Technical Report MP-TR-98-05*, Image Speech and Intelligent Systems Group, Univ. of Southampton, 1998.
- [14] K.K. Chin, “Support Vector Machines Applied to Speech Pattern Classification,” master’s thesis, Eng. Dept., Cambridge Univ., 1999.
- [15] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge Univ. Press, 1999.
- [16] IEC 61853, Photovoltaic (PV) module performance testing and energy rating – Part 1: Irradiance and temperature performance measurements and power rating.