Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2022.29100

279 (2022) 147–151
December

Employing machine learning by classification for analysis of a monitoring database from a photovoltaic module

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

Centre de Développement des Energies Renouvelables, CDER, 16340, Algiers, Algeria, Mobile: +213 (0) 791658746; email: h.hafdaoui@cder.dz/hichemhafdaoui@gmail.com (H. Hafdaoui)

Received 3 June 2022; Accepted 20 September 2022

ABSTRACT

The use of artificial intelligence methods in data analysis facilitates and shortens the time for making decisions, especially when faults or malfunctions occur at the photovoltaic station level. The researchers face many difficulties. One of the most significant ones is in terms of collecting and analyzing the obtained results, especially for a long period of monitoring. This paper proposes a new method to analyze the results by classification using a support vector machine (SVM) classifier. In such a way, a data variable is regrouped into a multiclass for analysis using SVM. Based on the application of artificial intelligence (classification), recorded data, the power output for a given photovoltaic module (PV) technology, types, and small or large stations under any season can be analyzed and processed easily. In this paper, classification was employed to analyze the monitoring database of a photovoltaic (PV) module (260 W) over 5 months.

Keywords: Monitoring; Performance ratio; Data analysis; Classification; Support vector machine

1. Introduction

Artificial intelligence (AI) has offered promising solutions to the problem of photovoltaic, such as the detection of problems in large power plants, allowing for greater flexibility that makes AI more robust than classical methods. Furthermore, machine learning (ML) is a field of study of AI, its basic idea is the use of computer algorithms that can improve automatically through experience and the use of data. MLs are about building huge neural network models that can make precise choices based on data; ML is suitable for situations where the data is complex. ML algorithms have been improving in photovoltaic and have been developed in many fields.

Recently, several research works and machine learning and deep learning methods have been successfully applied in the field of fault diagnosis for photovoltaic generators. Lu et al. [1] proposed to use the machine learning algorithms to establish a data mining decision tree model for the operation data of photovoltaic (PV) generators, and to use the model to predict the cause of failure of a PV. Hong et al. [2] the authors proposed a new framework, consisting of image acquisition, image segmentation, defect orientation and defect warning, to address the defect limitations of PV modules. The images of the visible and infrared PV field were taken under the same conditions by a double infrared camera at low altitude. Liu et al. [3] applied Kernel Extreme Machine Learning (KEML) and support vector machine (SVM) technique to detect defects based on I-V curves under STC conditions. Bansal et al. [4] presented tool reliability models that were developed based on the amount of available failure data and were used for early detection of power generation degradation. Benedetti et al. [5] and Polo et al. [6] present a method based on artificial neural networks

* Corresponding author.

Presented at the First International Congress of Energy and Industrial Process Engineering (ICEIPE'22) 23–25 May 2022, Algiers, Algeria

1944-3994/1944-3986

(ANN) for the detection of anomalies presented is based on the comparison between the measured and predicted values of the AC energy production.

In this paper, an AI (SVM classifier) method was employed to analyze the monitoring database of a photovoltaic (PV) module polycrystalline (260 W) over 5 months. The characterization of PV module is presented in Table 1.

2. Materials and methods

The data acquisition system consisted of a set of data of diverse units and measurements under different conditions, called a database, through which an analysis system was prepared by converting it into structured data that was suitable for fast queries [7].

Table 2 shows the localization of PV modules installed in the laboratory – CDER, Ben Aknoun City, Algeria and Fig. 1 shows a photo of those PV modules in outdoor; the arrow points to the PV module on which we conducted this study.

The system collected measurements from different sensors (e.g., Isc, Voc, Pmax, Imp, Vmp, Tm, *G* and time). Using data acquisition (Keysight 34972A), the measurement system started at 06:00 until 20:00, here the following devices were employed for measurement:

Table 1
Characterization of PV module

Maximum power output of PV module (Pmax)	260 Wc
Short-circuit current of PV module (Isc)	8.6 A
Open-circuit voltage of PV module (Voc)	38 V
PV module current at maximum power point (Imp)	8.1 A
PV module voltage at maximum power point (Vmp)	31 V

Table 2
Localization of PV modules

Latitude	Longitude	Altitude
36°44′44,94″N	3°00′46,80 E	236 m

Fig. 1. PV modules in outdoor.

- Pyranometer LSI COD-DPA154
- Data acquisition Keysight technology 34972A (Section 2.1 – Data acquisition)
- Electronic charge Keysight technology N3300A (Section 2.2 – Electronic charge)
- Sensor – PT100 TC DIRECT

2.1. Data acquisition

A Keysight 34972A data acquisition system collected the measurements from the various sensors (temperature, wind speed and direction and solar irradiation), Fig. 2 shows the front face of data acquisition.

2.2. Electronic charge

The electrical characteristics of a photovoltaic module (Pmax, Isc, Voc, Imp, Vmp).

The electronic charger used in our laboratory was the Keysight N3300A. Channels 3 and 4 (N3304A, 300 W, 0–60 A and 0–60 V) were intended for the performance test at STC. For PV modules with a power lower than 300 W channel 3 was used alone. For PV modules with a power greater than 300 W the two channels were used in parallel. Fig. 3 shows the front face of electronic charge.

The measurement system diagram is shown in Fig. 4.

Machine learning was applied using the SVM classifier to analyze the data obtained during 5 months of outdoor monitoring. The ***variables*** data were converted into classes, and then grouped according to the objectives and analytical requirements.

Heatmap with a gradual color display was employed to read the data [8–11].

Fig. 2. Data acquisition (Keysight 34972A).

Fig. 3. Electronic charge (Keysight N3300A).

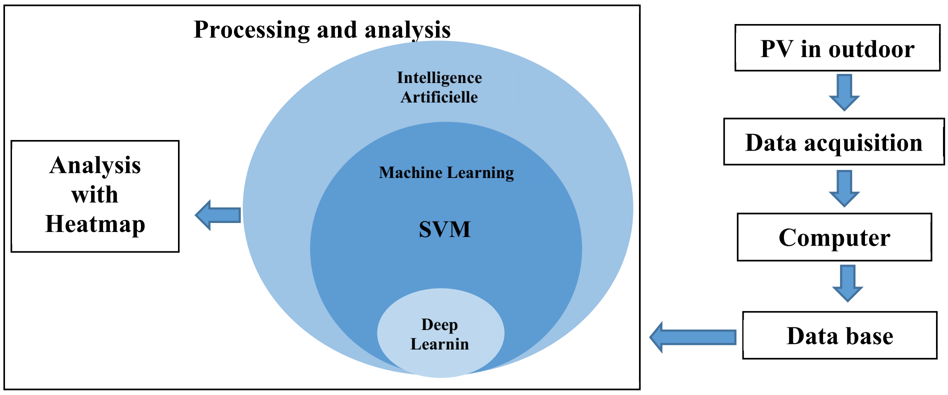

Fig. 4. Process flowsheet diagram of measurement system.

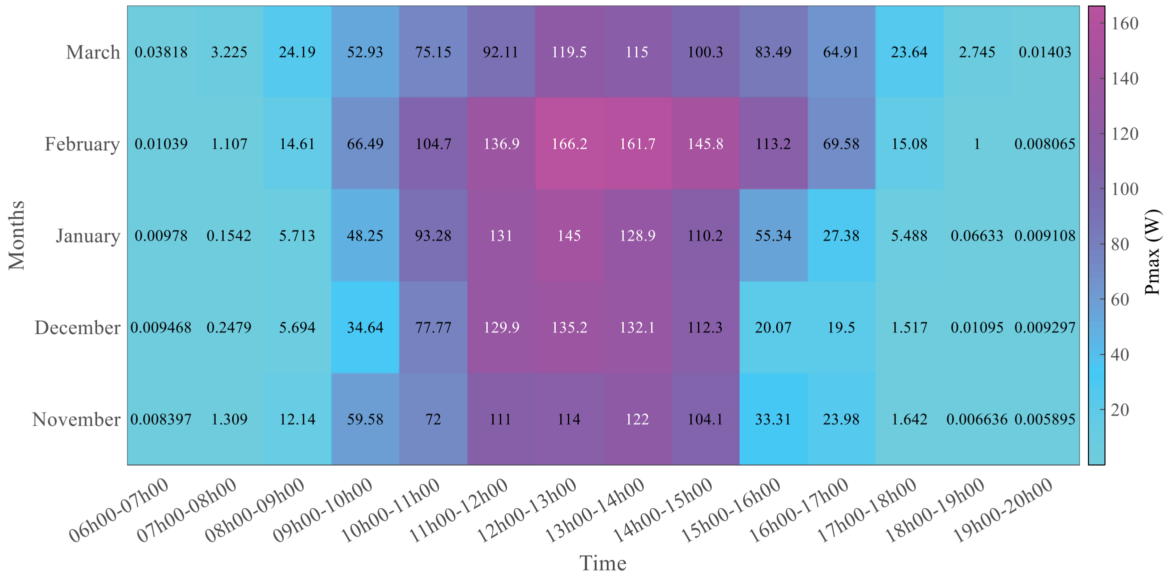

Fig. 5. Evolution of the maximum power supplied over a day in seven (5) months.

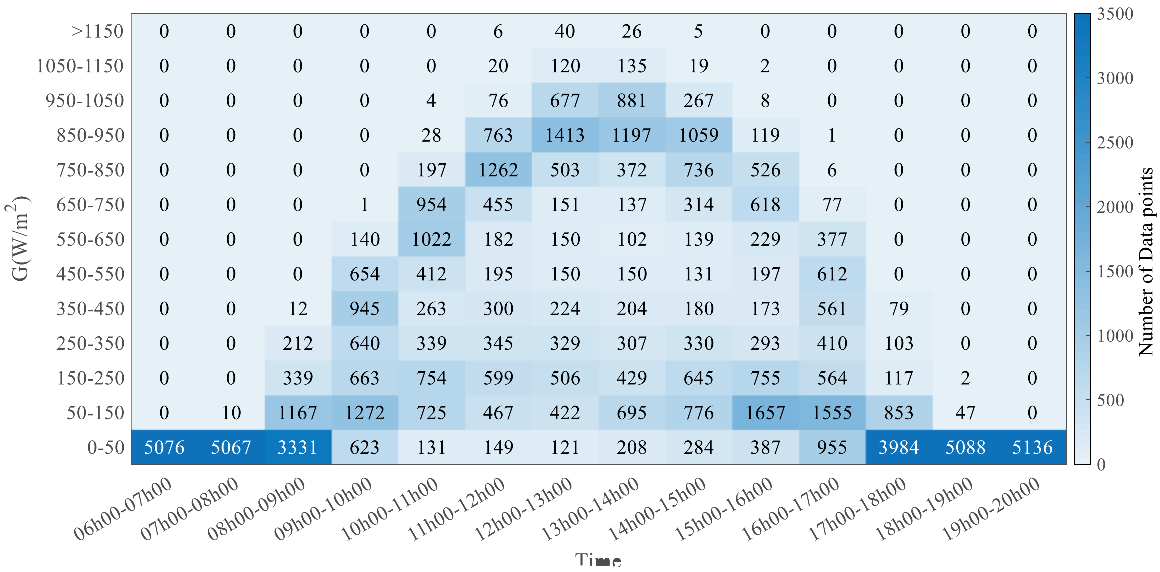

Fig. 6. Number of data points, G (W/m²) vs. time in 5 months.

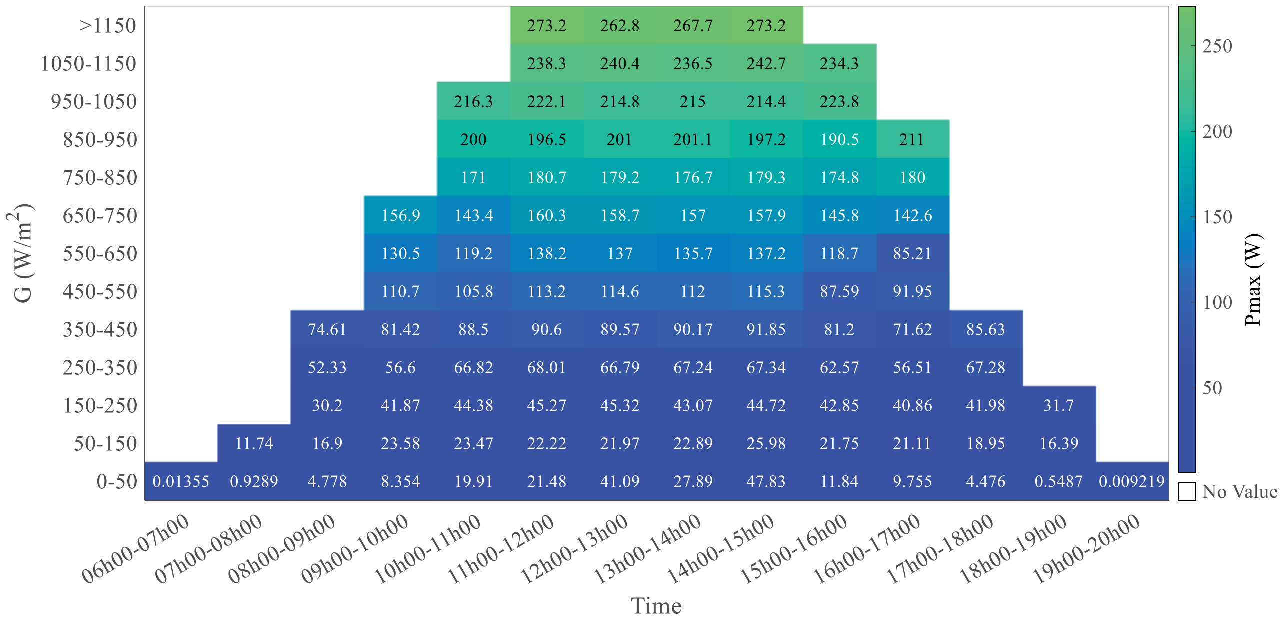

Fig. 7. PR (%) vs. G (W/m²) and time in 5 months.

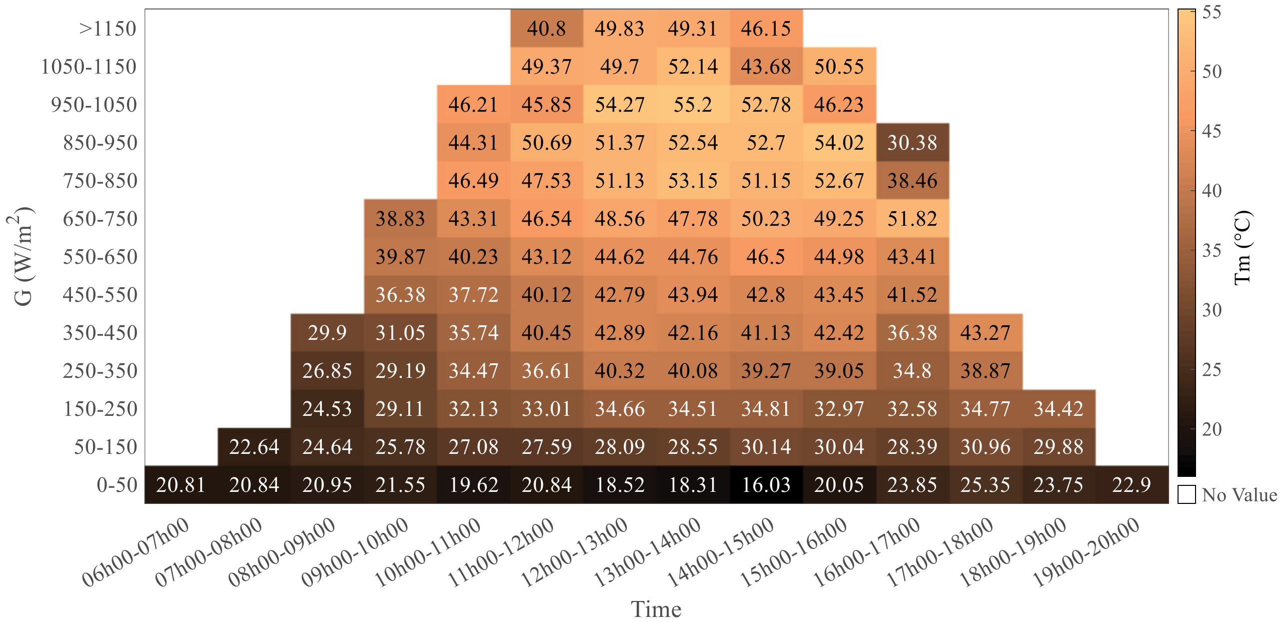

Fig. 8. Tm (°C) vs. G (W/m²) and time in 5 months.

3. Results and discussion

The system of data acquisition in the outdoors supplies a wide range of measurement conditions. This data contains parameters which are predicted by different measurement units and saved. Moreover, the parameters measures were imported into common database for preparing further analysis by transformation into data structures appropriate for fast queries [7]. Fig. 5 shows the maximal power of the PV module provided in hours and months as well as the evolution of the average monthly power over a day. Photovoltaic production was significant between 13 h00 and 14 h00 local time and was around 74% of the nominal power during the months of January and February. This value decreased in autumn because of the more frequent presence of clouds and the lower height of the sun. This figure also shows the monthly evolution of sunshine duration.

Classification of monitoring data to analyze using heatmaps is important. Fig. 6 shows the number of data points which represent measurements in each class. The maximum point analyzed was 5136 at $0 > G > 50$ W/m², between 19 h00–20:00. For reading heatmaps with each other in Fig. 6, for example the data point 1,197 at $850 > G > 950$ W/m² between (13 h00–14:00) the maximal power point in

Fig. 7 was (201.1 W) and at the same point, the PV module reached a temperature of 52.54°C in the same condition and time (Fig. 8); those values (201.1 W and 52.54°C were the calculation rate of 1197 values. It was noticed that the Pmax was between 80% and 90%, so it can be argued that 20% of the energy was lost in 80 d. To classify the data and to eliminate the not acceptable points in any irradiation, the heat map based on classification facilitated the current study and measurement to choose the logical point. Another point was selected by analyzing each two points and calculating the difference between them. Any data can be investigated with the same procedure and analysis method.

4. Conclusions

Monitoring techniques allow stakeholders to track the production of solar PV systems, evaluate their performance, detect problems, and prepare reports. It provides complete solutions for single source monitoring of solar power plants from the planning stage and delivery of all components to the operational stage of remote monitoring technologies.

AI methods analyze and extract all the photoelectric effects to support research work and determine in detail problems such as the degradation of PV modules and solar stations over large periods. This can give a new insight in analyzing the results in the field of photovoltaic systems. This method can be applied to compare different photovoltaic systems in terms of the technologies used, by tracking and monitoring all phenomena. It is recommended that a comparison should be made between several different technologies. Future research work should l focus on AI methods to analyze and extract photoelectric effects to determine details about degradation of panels and solar stations over large time periods. This will give new insight in the field of photovoltaic systems.

Symbols

SVM	—	Support vector machine
ML	—	Machine learning
ANN	—	Artificial neural networks
AI	—	Artificial intelligence
Isc	—	Short-circuit current of PV module, A
Voc	—	Open-circuit voltage of PV module, V
Pmax	—	Maximum power output of PV module, W
Imp	—	PV module current at maximum power point, A
Vmp	—	PV module voltage at maximum power point, V
Tm	—	Temperature of PV module, measured at the backsheet, °C
G	—	Measured in-plane irradiance, W/m^2
Time	—	Hour of measurement

References

[1] X. Lu, P. Lin, S. Cheng, Y. Lin, Z. Chen, L. Wu, Q. Zheng, Fault diagnosis for photovoltaic array based on convolutional neural network and electrical time series graph, Energy Convers. Manage., 196 (2019) 950–965.

[2] F. Hong, J. Song, H. Meng, R. Wang, F. Fang, G. Zhang, A novel framework on intelligent detection for module defects of PV plant combining the visible and infrared images, Sol. Energy, 236 (2022) 406–416.

[3] Y. Liu, K. Ding, J. Zhang, Y. Lin, Z. Yang, X. Chen, Y. Li, X. Chen, Intelligent fault diagnosis of photovoltaic array based on variable predictive models and I–V curves, Sol. Energy, 237 (2022) 340–351.

[4] N. Bansal, S.P. Jaiswal, G. Singh, Long term performance assessment and loss analysis of 9 MW grid tied PV plant in India, Mater. Today:. Proc., 60 (2022) 1056–1067.

[5] M. De Benedetti, F. Leonardi, F. Messina, C. Santoro, A. Vasilakos, Anomaly detection and predictive maintenance for photovoltaic systems, Neurocomputing, 310 (2018) 59–68.

[6] F.A. Olivencia Polo, J.F. Bermejo, J.F. Gómez Fernández, A. Crespo Márquez, Failure mode prediction and energy forecasting of PV plants to assist dynamic maintenance tasks by ANN based models, Renewable Energy, 81 (2015) 227–238.

[7] D. Magare, O. Sastry, R. Gupta, H. Mohammed, B. Bora, Y. Singh, Estimation of operating condition frequency distribution and power prediction by extending application of IEC 61853-1 standard for different technology photovoltaic modules, J. Renewable Sustainable Energy, 10 (2018) 033507, doi: 10.1063/1.5023084.

[8] C.J.C. Burges, A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discovery, 2 (1998) 121–167.

[9] G. Camps-Valls, D. Tuia, L. Bruzzone, J.A. Benediktsson, Advances in hyperspectral image classification: earth monitoring with statistical learning methods, IEEE Signal Process Mag., 31 (2014) 45–54.

[10] R. Ruiz-Gonzalez, J. Gomez-Gil, F.J. Gomez-Gil, V. Martínez-Martínez, 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, 14 (2014) 20713–20735.

[11] H. Hafdaoui, E.A. Kouadri Boudjelthia, A. Chahtou, S. Bouchakour, N. Belhaouas, Analyzing the performance of photovoltaic systems using support vector machine classifier, Sustainable Energy Grids Networks, 29 (2022) 100592, doi: 10.1016/j.segan.2021.100592.