



**HAL**  
open science

# Extraction de signatures et prédiction de l'état de santé des centrales photovoltaïques

Edgar Hernando Sepúlveda Oviedo

## ► To cite this version:

Edgar Hernando Sepúlveda Oviedo. Extraction de signatures et prédiction de l'état de santé des centrales photovoltaïques. Congrès annuel de l'Ecole Doctorale GEETS 2022, Ecole Doctorale Génie Electrique, Electronique, Télécommunications et Santé, Apr 2022, Toulouse, France. hal-04290156

**HAL Id: hal-04290156**

**<https://hal.science/hal-04290156>**

Submitted on 16 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/371762057>

# Extraction de signatures et prédiction de l'état de santé des centrales photovoltaïques

Conference Paper · April 2022

---

CITATIONS

0

READS

19

1 author:



[Edgar Hernando Sepúlveda Oviedo](#)

Laboratoire d'Analyse et d'Architecture des Systèmes (LAAS)

10 PUBLICATIONS 8 CITATIONS

SEE PROFILE

# Extraction de signatures et prédiction de l'état de santé des centrales photovoltaïques

Edgar Hernando SEPULVEDA OVIEDO

LAAS-CNRS, 7 avenue du colonel Roche, 31400 Toulouse, France

Une nouvelle approche d'intelligence artificielle et une plateforme d'acquisition de données orientée vers la détection de défauts dans les systèmes PV sont proposées. L'état des panneaux est discriminé à l'aide d'un clustering hiérarchique, d'une analyse temps-fréquence et d'une extraction et réduction de dimensionnalité statistiques. La détection des défauts est effectuée à l'aide des moindres carrés partiels (PLS) et validée avec une PLS-LDA proposée. La plate-forme d'acquisition de données est capable de capturer 16 signaux à la vitesse de la milliseconde. Les résultats de cet article dépassent ceux obtenus avec les méthodes conventionnelles.

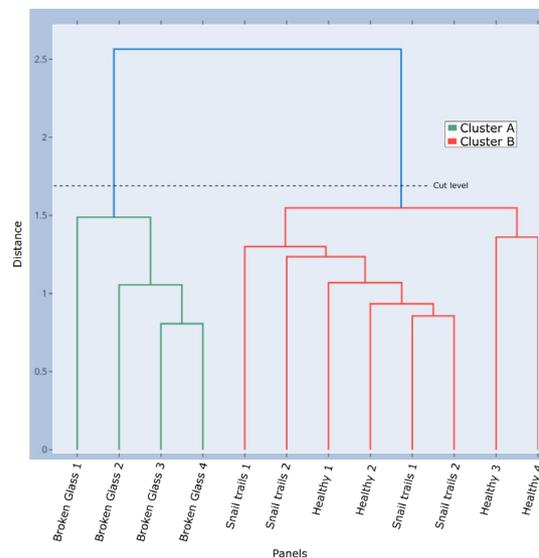
## I. INTRODUCTION

L'augmentation de l'utilisation de l'énergie photovoltaïque (PV) a mis en évidence la nécessité de détecter et de classer les défauts le plus tôt possible afin d'effectuer une maintenance préventive ou corrective du système PV. Le diagnostic des défauts nécessite l'acquisition de données et un système d'extraction de caractéristiques pour diagnostiquer les défauts. Cependant, il y a actuellement peu de discussions sur cet aspect et l'une des raisons peut être la grande complexité d'un système d'acquisition destiné à diagnostiquer des défauts qui peuvent varier en fonction des conditions météorologiques, des actions des onduleurs ou des optimiseurs, entre autres. Cet article propose une nouvelle approche pour l'extraction de signatures et la prédiction de l'état de santé des centrales PV et une plateforme d'acquisition de données avec un taux de capture à la vitesse de la milliseconde vers le diagnostic de défauts.

## II. Regroupement hiérarchique DTW

L'approche proposée a été testée sur 12 panneaux : 4 panneaux (healthy) , 4 panneaux (snail trails) et 4 (broken glass) instrumentés avec la plateforme industrielle TIGO avec une période d'échantillonnage d'une minute. Le clustering hiérarchique agglomératif (AHC) est utilisé pour construire deux clusters : A (panneaux gravement touchés) et B (autres panneaux). Dans cette étude, l'AHC utilise l'indice de similarité appelé Dynamic time warping (DTW), en raison de ses hautes performances dans la classification des séries temporelles [1], pour regrouper plusieurs

individus en clusters [2]. Le résultat du regroupement sous forme de dendrogramme est présenté sur la **Figure 1**.



**Figure 1. Dendrogramme du regroupement hiérarchique agglomératif des signaux de courant.**

Comme le montre la **Figure 1**, l'AHC regroupe les panneaux en ceux gravement touchés dans le groupe A (couleur verte) et les autres dans le groupe B (couleur rouge). Pour discriminer les panels healthy du cluster B en tenant compte de l'effet de l'irradiation, les signaux sont divisés en 4 fenêtres appelés : *morning*, *midday*, *afternoon*, *evening*. La génération de caractéristiques est effectuée sur chacune de ces fenêtres. Dans cet article on suppose que, si les classes sont discriminées dans au moins une des fenêtres, l'algorithme est efficace et il est possible de détecter des anomalies entre les panneaux PV du groupe B.

### III. Génération et Sélection de signatures

Ensuite, la décomposition des signaux multirésolution est effectuée, suivie de l'extraction des caractéristiques statistiques (Skewness, Kurtosis, Variance, P-P value et Energy). Certains défauts peuvent être visibles dans le domaine temporel ou dans le domaine fréquentiel. Pour les analyser simultanément, une décomposition du signal multi-résolution basée sur la transformée discrète en ondelettes est utilisée pour générer des coefficients d'approximation et de détail et ensuite l'ensemble de caractéristiques statistiques est extrait. Nous avons utilisé la corrélation et la variance pour éliminer des informations redondantes, ce qui peut réduire les performances des algorithmes de diagnostic.

La sélection basée sur la corrélation utilise la *matrice de corrélation de Pearson*. Plus une caractéristique est corrélée à une autre, moins elle fournit d'informations et peut introduire du bruit. Un seuil de corrélation  $\tau = 0,9$  est défini pour supprimer les caractéristiques corrélées. Ensuite, une sélection basée sur la variance est effectuée à l'aide de *coordonnées parallèles*. En *coordonnées parallèles* il y a autant d'axes normalisés que d'entités.

### IV. Prédiction de l'état de santé

Les caractéristiques non corrélées et avec variance significative permettent de résoudre les problèmes de prédiction et de classification de l'état de santé des panneaux PV.

### A. Régression PLS

L'algorithme PLS fournit des résultats très intéressants par rapport aux autres méthodes conventionnelles lorsque le but est la prédiction de classes [3]. De plus, PLS définit des composants latents qui peuvent ensuite être utilisés comme prédicteurs dans un problème de classification, fournissant une méthode de validation des résultats de la prédiction basée sur PLS. En ce sens, l'algorithme PLS peut être vu comme une méthode de réduction de dimension couplée à un modèle de régression. Le PLS effectue simultanément une réduction de dimensionnalité et une classification basée sur la régression. Les résultats du PLS, les métriques d'évaluation Root Mean Squared Error (RMSE) et R-Squared ( $R^2$ ) sont rapportés dans la **Figure 2**.

### B. Méthode de classification PLS-LDA

Cette méthode permet valider les résultats du PLS. Le LDA a déjà été utilisé dans la détection de défauts dans les systèmes PV [4]. L'algorithme LDA projette la matrice de données d'origine  $T$  (prédicteurs) d'un espace de grande dimension dans un nouvel espace de faible dimension qui rend la dispersion intra-classe aussi petite que possible et la dispersion inter-classe aussi grande que possible. Les résultats du PLS-LDA, la métrique d'évaluation F-Value et la matrice de confusion pour chaque fenêtre temporelle sont présentés dans la **Figure 2**.

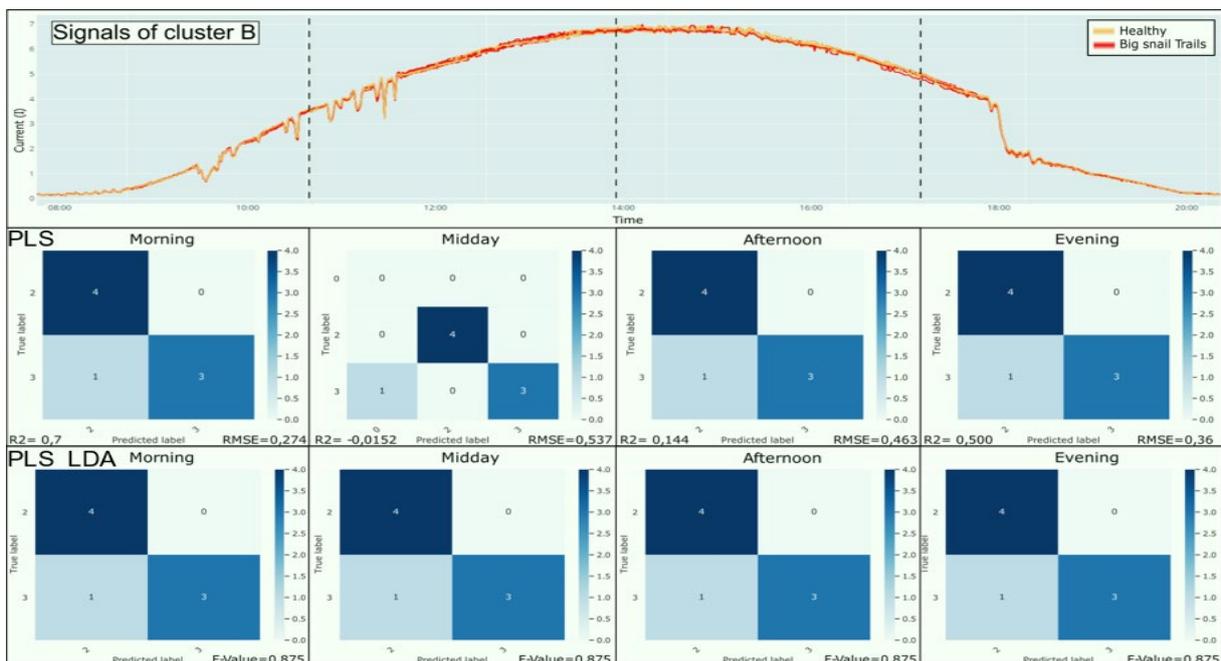
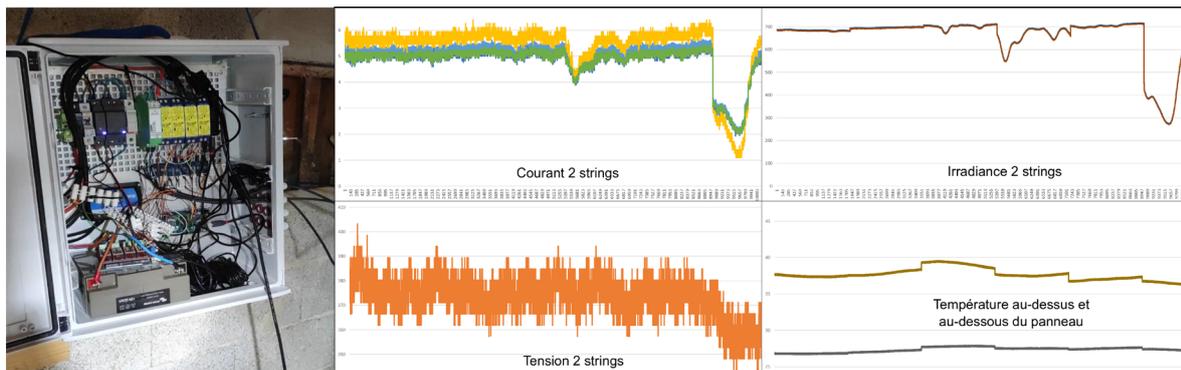


Figure 2. Prédiction des résultats de l'état de santé des panneaux PV. Signaux de current du cluster B (a). Résultats PLS (b). Résultats PLS-LDA (c).



**Figure 3. Acquisition des données. Nouvelle plateforme d'acquisition de données (millisecondes). Signaux capturés sur une centrale photovoltaïque industrielle (35 minutes).**

## V. ACQUISITION DES DONNEES

En raison de la grande difficulté à détecter les défauts des signaux échantillonnés toutes les minutes, cette étude propose une plateforme d'acquisition de données robuste dotée d'un ensemble de capteurs de haute précision qui enregistrent à la fois la production d'énergie et les conditions météorologiques de fonctionnement de la centrale PV. La plateforme capture : tension, courant, d'irradiation et de température.

Pour la construction de la plate-forme d'acquisition de données, il est nécessaire de sélectionner les capteurs, la carte d'acquisition de données et la carte de traitement des données. Ensuite, il faut évaluer les capteurs pour vérifier la correspondance du comportement avec celui reporté dans les fiches techniques. Notre plateforme est composée de 16 capteurs analogiques, d'une carte Arduino pour l'acquisition de données et d'une carte Raspberry PI 4 qui contient le système embarqué de détection de défauts. Les mesures de courant comportent plusieurs capteurs pour des mesures de validation redondantes. La **Figure 3** présente le prototype de plate-forme d'acquisition de données orientée vers la détection de défauts et un exemple de 35 minutes capturant 16 capteurs sur centrale PV industrielle.

## VI. CONCLUSION

L'approche proposée dans cet article est capable de discriminer clairement les 3 types de panneaux en retenant les informations essentielles et en éliminant les informations redondantes ou non pertinentes. Comme le montre la figure 3, dans le fenêtre de temps *Midday*, l'apparition d'une nouvelle étiquette

(*classe 0*) peut signifier qu'en dépit d'appartenir au même type de défaut, son niveau d'affectation est différent ou ce panneau présente une combinaison de défauts. Les défauts qui surviennent sous faible irradiation (*Morning* et *Evening*) sont les plus difficiles à diagnostiquer, cependant, c'est dans ces cas que notre approche est la plus performante. L'utilisation de fenêtres temporaires permet d'analyser l'impact et l'évolution des défauts dans le temps.

La plate-forme proposée a été testée dans des centrales PV industrielles et a montré de hautes performances en termes de qualité des données. Le nombre d'enregistrements avec la nouvelle plateforme est passé de 60 à 1800 enregistrements par heure et d'une variable mesurée à 4 en augmentant la richesse des signaux pour le diagnostic des défauts des systèmes PV. Notre système est autonome en énergie puisqu'il possède son propre alimentateur PV, ce qui le rend polyvalent et facilement installable dans toutes les centrales PV.

## REFERENCES

- [1] J. Lines, A. Bagnall, *Time series classification with ensembles of elastic distance measures*, Data Min Knowl Disc 29 2015 (2011) 565–592.
- [2] S. Rani, G. Sikka, *Recent techniques of clustering of timeseries data. A survey*, International Journal of Computer Applications 52 (2012) 1–9.
- [3] Y. Liu, W. Rayens, *PLS and dimension reduction for classification*, Computational Statistics 22 (2007) 189–208.
- [4] S. Fadhel, *Efficacité énergétique et surveillance d'un microgrid à courant continu alimenté par des panneaux photovoltaïques*, Theses, Université Paris-Saclay, 2020.