Hierarchical clustering and dynamic time warping for fault detection in photovoltaic systems

Edgar Hernando Sepúlveda Oviedo, Louise Travé-Massuyès, Audine Subias, Corinne Alonso, Marko Pavlov

▶ **To cite this version:**

Edgar Hernando Sepúlveda Oviedo, Louise Travé-Massuyès, Audine Subias, Corinne Alonso, Marko Pavlov. Hierarchical clustering and dynamic time warping for fault detection in photovoltaic systems. X Congreso Internacional CIMM Ingeniería Mecánica, Mecatrónica y Automatización, May 2021, Bogotá (virtual), Colombia. hal-03355362

HAL Id: hal-03355362

https://hal.science/hal-03355362

Submitted on 27 Sep 2021

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.

HIERARCHICAL CLUSTERING AND DYNAMIC TIME WARPING FOR FAULT DETECTION IN PHOTOVOLTAIC SYSTEMS

Edgar Hernando Sepúlveda Oviedo (1,2), Louise Travé-Massuyès (1), Audine Subias (1), Corinnne Alonso (1), Marko Pavlov (2)

1. LAAS_CNRS, Université deToulouse, CNRS,INSA,UPS,Toulouse, France
2. Quantom SAS, France

Núcleo Temático: Mecatrónica y Automatización

Introduction

Safety and energy efficiency of PV plants can be affected by failures in any component of the plant if the degradation is not detected and corrected quickly. This is why fault detection and diagnosis (FDD) methods have a critical role to play in this application domain. FDD methods are classified into two large groups: i) based on models; and ii) based on data. In the first group, a high level of expert knowledge is necessary. In the second, a large volume of data is required to train the machine learning algorithms. This paper proposes to experiment Dynamic Time Warping (DWT) followed by Hierarchical Clustering (HC) as a data-driven approach. The results of this method are compared with the diagnosis labels assessed by visual inspection of the panels.

Methods

Our approach is made up of three phases: i) Data acquisition (voltage and current); ii) Feature Extraction (DWT); and iii) Clustering (HC). In the first phase, we capture the voltage and current signals every minute. In the second, we calculate the DWT similarity metric that has been widely used in clustering and classification [Li, 2021] [Wang *et al*., 2019]. This metric allows to determine the similarity even between out-of-phase signals [Jeong *et al*., 2011]. The result of the DWT is then used in the HC algorithm to group the data in a hierarchical way. In this study we use agglomerative HC assuming that each observation is a cluster and the pairs of clusters merge as they move up the hierarchy. We present the results in the form of a dendrogram in Figure 1.

Results

In Figure 1, we observe the result of HC on the current (I) and voltage (V) signals for one day. The green clusters correspond to the panels identified as healthy. The red clusters correspond to the faulty panel clusters.

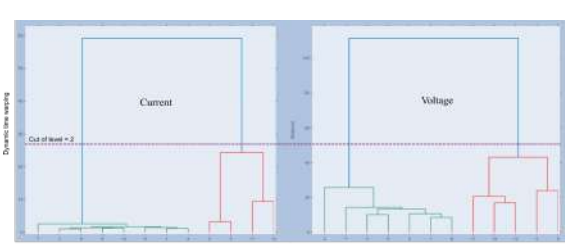

Figure 1: Cluster dendrogram for current (I) and voltage (V).

Since the panels are labelled in two groups: **Healthy** (7 panels) or **Faulty** (5 panels) by visual inspection, we define the cut-off level of the dendrogram in two to compare the results. Figure 2 illustrates how many signals HC grouped for variables (I) and (V) in each group.

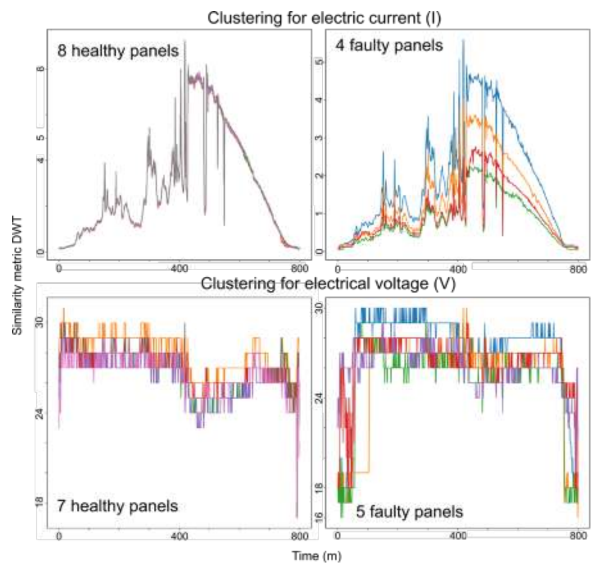

Figure 2: Signals grouped by HC algorithm.

Discussion

In Figure 2, we observe that HC was able to identify all healthy panels in both current and voltage. However, in current (I), HC identifies an extra panel as healthy. This is because the extra panel has a corrosion of the encapsulation surface sheet, this fault is called "Snail Trails" for its optical appearance. This fault does not significantly decrease the performance of solar panels (emulates healthy behaviour) and it is therefore difficult to detect at the electrical signal level. However, it can be the cause of fractures or micro-cracks in the modules that reduce the production of the panel.

It is interesting to note that this approach does not require the extraction of multiple characteristic features, which reduces the computational cost. HC groups the signals using a single variable, which could reduce the number of sensors for fault diagnosis. Obviously, the final result of the HC depends on the level at which the clusters are cut [Nielsen, 2016].

In order to improve preventive maintenance times, HC was tested with different time windows and we obtained the same results even with 3 minute windows. In addition, as we can see in Figure 1 in the faulty clusters if we increase the cut-off level we could determine the level of impact of the fault. These two aspects are vital to establish a priority in preventive maintenance.

References

Jeong, Y. S *et al*., Weighted dynamic time warping for time series classification. Pattern Recognition, 44:2231-2240, 2011

Li, H. Time works well: Dynamic time warping based on time weighting for time series data mining. Information Sciences, 547:592-608, 2021.

Nielsen, F. Hierarchical clustering. In Introduction to hpc with mpi for data science. Cham: Springer International Publishing. 195–211, 2016

Wang, X *et al*., Clustering of interval-valued time series of unequal length based on improved dynamic time warping. Expert Systems with Applications,125:293–304, 2019.