Chapter 62
Classification of Failures in Photovoltaic Systems using Data Mining Techniques

Lucía Serrano-Luján
Technical University of Cartagena, Spain

Jose Manuel Cadenas
University of Murcia, Spain

Antonio Urbina
Technical University of Cartagena, Spain

ABSTRACT

Data mining techniques have been used on data collected from a photovoltaic system to predict its generation and performance. Nevertheless, up to date, this computing approach has needed the simultaneous measurement of environmental parameters that are collected by an array of sensors. This chapter presents the application of several computing learning techniques to electrical data in order to detect and classify the occurrence of failures (i.e. shadows, bad weather conditions, etc.) without using environmental data. The results of a 222kWp (CdTe) case study show how the application of computing learning algorithms can be used to improve the management and performance of photovoltaic generators without relying on environmental parameters.

INTRODUCTION

During recent years the number of large-scale PV (photovoltaic) systems has grown worldwide. In 2010, the photovoltaic industry production more than doubled and reached a worldwide production volume of 23.5 GWp of photovoltaic modules. Business analysts predict that investments in PV technology could double from € 35-40 billion in 2010 to over € 70 billion in 2015, while prices for consumers are continuously decreasing at the same time. (European Commission, DG Joint Research Centre, Institute for Energy, Renewable Energy Unit, 2011).

DOI: 10.4018/978-1-4666-9840-6.ch062

The complexity in PV system configurations represents an additional problem in maintenance and control operations in large systems. For example, a failure in one PV module placed at a big façade is very difficult to detect. A quick detection of failures would avoid energy losses due to malfunctions of PV system and therefore improve its performance and end-user satisfaction (Roman, Alonso, Ibanez, Elorduizapatarietxe, & Goitia, 2006).

When data-mining techniques are applied to a PV database, a wide variety of relations between parameters can be found. This study relays on expert knowledge to study the possible behaviours of PV generation performance that can be affected by changes in the environment conditions; furthermore, computing learning algorithms allow us to detect and classify failures without measuring the environmental parameters.

This study focuses on a methodology to control the correct performance of each group of modules which compose a large-scale PV generator, identifying failure occurrences and its most likely causes, and it does so by a procedure that analyses the performance of a group of modules by ignoring environmental information.

DATA MINING AND RENEWABLE ENERGIES

Data Mining

Knowledge Extraction Process tries to find useful, valid, important and new knowledge about a phenomenon or activity by computational efficient procedures (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Additionally it is of paramount importance to present the results in a clear and easily interpretable way, therefore this process comprises data-mining techniques in order to extract or identify patterns from data that needs to be complemented with both pre-processing and post-processing stages. The process implies several steps:

- To determine the sources of information and the creation of a data set.
- To select, clean and transform the data that will be analysed.
- To extract knowledge in order to select and apply the most appropriate Data Mining technique.
- To evaluate, interpret, transform and represent the extracted patterns.

The concept "Data Mining" is academically considered as a step inside Knowledge extraction Process, but nevertheless, from an applied point of view, both terms are used in an equivalent way (Fayyad et al., 1996)

Basically, Data Mining process can be supervised or not supervised, depending if the entries are assigned to a finite number of discrete classes or not, and it includes the selection of the tasks to be done, for instance, classification, grouping or clustering, regression, etc. Data Mining processes find patterns that can be expressed as a model or it can easily show dependence between data in a graphical manner. The introduced model depends on its function (for instance, classification) and on the way it is represented (decision trees, rules, etc.). Besides, it should specify preference criteria to select a model inside a group of possible models and it has to specify the finding strategy to be used (that is normally determined in a particular Data Mining technique) (Fayyad et al., 1996; Hernández Orallo, Ramírez Quintana, & Ferri Ramírez, 2005; Witten & Frank, 2011)