Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 124 (2016) 39–56

SOLAR ENERGY

www.elsevier.com/locate/solener

# Photovoltaic power pattern clustering based on conventional and swarm clustering methods

Amr A. Munshi [a,b,*], Yasser A.-R.I. Mohamed [a]

[a] Electrical and Computer Engineering Department, University of Alberta, Edmonton, Alberta, Canada
[b] Computer Engineering Department, Umm Al-Qura University, Makkah, Saudi Arabia

Received 2 July 2015; received in revised form 3 November 2015; accepted 6 November 2015
Available online 10 December 2015

Communicated by: Associate Editor Mario Medina

**Abstract**

Photovoltaic power pattern clustering is a key tool to provide information about the impacts of integrating photovoltaic systems into the electrical grid without extensive analysis and simulations. This paper investigates a set of clustering methods from different clustering categories to determine the optimum number of clusters for photovoltaic power patterns data. Furthermore, the introduction of bio-inspired optimization methods Ant Colony and Bat in clustering power patterns in presented. For the purpose of clustering and achieving efficient cluster representatives, six clustering algorithms from five different clustering categories are involved, K-means from partitional clustering, Hierarchical from agglomerative clustering, Fuzzy C-means from fuzzy clustering, Self-Organizing Maps from neural network based algorithms, and Ant Colony and Bat from bio-inspired optimization methods. In order to evaluate the clustering methods in a comprehensive manner, nine internal validity indices were employed, Davies Bouldin, Dunn, Silhouette, Bayesian Information Criterion, Xie-Beni, mean square error, clustering dispersion indicator, mean index adequacy and ratio of within cluster sum of squares to between cluster variation. The clustering results show that bio-inspired clustering methods are comparable to those conventional methods. Moreover, Bat was the most efficient and outperformed the other clustering methods.


*Keywords:* Clustering; Photovoltaic systems; Power patterns; Swarm methods

## 1. Introduction

As the demand for electrical power is increasing rapidly due to the growth of the global population and industrialization, electrical systems need to increase generation. Recent studies predict that the world's net electricity generation is expected to rise from 17.3 trillion kilowatt-hours in 2005 to 24.4 trillion kilowatt-hours (an increase of 41%) in 2015 and 33.3 trillion kilowatt-hours (an increase of 92.5%) in 2030 (International Energy Association PVPS Annual Report, 2012). Currently, most power is generated from conventional power resources, such as fossil fuels. However, there are many financial, environmental, and availability issues associated with the increasing consumption of conventional power resources. In order to overcome many of these issues, renewable energy resources can participate in producing power. The potential of using renewable energy resources as an alternative for power generation is increasing rapidly. The motivations behind this are to overcome the financial, environmental and availability issues of conventional power resources such as fossil fuels (Marwede and Reller, 2014; Goodrich et al., 2013;

* Corresponding author at: Electrical and Computer Engineering Department, University of Alberta, Edmonton, Alberta, Canada.
*E-mail address:* aamunshi@ualberta.ca (A.A. Munshi).

http://dx.doi.org/10.1016/j.solener.2015.11.010
0038-092X/

Kushiya, 2009; Sinha, 2013; Cyrs et al., 2014). Among renewable energy resources, wind energy and solar energy have recently become of extensive interest. Photovoltaic (PV) systems convert the solar energy into electrical energy. The advances in PV technology such as the ability to generate electricity for a quarter of a decade with minimal regular maintenance and the continuous reduction in capital cost has encouraged many countries in considering integrating it into their power grid systems.

The fundamental device in a PV system is the photovoltaic cell. Photovoltaic cells are made of several types of semiconductors using different manufacturing processes. The monocrystalline and polycrystalline silicon cells are the most widely found at the present time in manufacturing photovoltaic cells. A PV cell is a semiconductor diode whose p–n junction is exposed to light. The light incident on the PV cell generates charge carriers that produce electric current. The PV cells are arranged in series to form modules. Modules are connected in series or parallel to form panels and group of panels constitute an array (Villalva et al., 2009). Generally, the output power of PV systems is influenced by the level of irradiation and the ambient temperature (Farivar et al., 2011). This leads to fluctuations in the output power of PV systems. Accordingly, the installation of these systems requires extensive studies and simulations of long historical data with sub-hourly time steps. However, dealing with such data is time consuming and computationally expensive. For this purpose developing solutions that can reduce the burden of extensive studies and simulations related to integrating PV systems into the electrical grid are of interest. For that, clustering techniques are employed to group PV power patterns that have similar features. Hence, a representative power pattern from each group can be utilized in the simulations. Moreover, accurate clustering of PV power patterns provide statistical information about the occurrence of particular patterns, and accordingly, assist in operating and planning for such systems.

An intensive research effort has been devoted to cluster power patterns and obtain representatives for these clusters during the last few years. Models based on clustering techniques were used to group electrical load patterns of customers in order to assist tariff formation (Chicco et al., 2003, 2004, 2006, 2013; Figueiredo et al., 2005; Gerbec et al., 2004; Verdu et al., 2006; Tsekouras et al., 2007), short-term forecasting purposes (Chicco et al., 2001) and in demand response programs to support management decisions (Gabaldon et al., 2010; Tsekouras et al., 2008, 2008). Also, power load clustering has been used for the classification of load profiles for ship electric consumers (Tsekouras et al., 2009) and for estimating the power load of war-ships (Tsamopoulos et al., 2014). In Ali et al. (2013), aggregate modeling of wind farms has been proposed based on the wind farm's layout and the clustering of wind speed patterns. A method to improve the management decisions of wind farms was also proposed in Duarte et al. (2012) by applying a clustering method on wind power loads.

Research interest in PV power pattern has recently increased, for analyzing the power output fluctuation effects on integrating PV systems into the electrical grid (Omran et al., 2010) and for determining the optimal location and size of PV plants (Haghdadi et al., 2012). The impact of shading on the photovoltaic power pattern was analyzed in Malathy and Ramaprabha (2015), and a recent review on modeling approaches of photovoltaic arrays based on the solar irradiance patterns is presented in Jena and Ramana (2015). In Wang et al. (2015) solar irradiance power patterns were studied in order to develop a photovoltaic power pattern recognition model. However, clustering methods to reduce the burden of those studies have not been further discussed. In the clustering process, Omran et al. (2010) adapted three clustering methods: K-means, Hierarchical and a hybrid of K-means and Hierarchical whereas, Haghdadi et al. (2012) used K-medoids and Fuzzy C-means. In Mori and Takahashi (2012) a Radial Basis Function Network model to predict short-time PV generation was proposed by clustering historical time-series PV power data, then constructing a prediction model at each cluster so that the prediction is based on data similarity. Also, Hosoda and Namerikawa (2012) proposed a predicting model by clustering historical data and using the weather forecast. The results of Omran et al. (2010), Haghdadi et al. (2012), Mori and Takahashi (2012), Hosoda and Namerikawa (2012) showed that the prediction results depend significantly on the accuracy of the clustered data. Thus, developing clustering algorithms that produce efficient partitioning of PV power data is in interest. In addition, the potential of PV in becoming a major power resource world-wide (European Photovoltaic Industry Association, 2013) motivates the investigation of applying various clustering techniques to investigate the most appropriate technique for clustering PV power patterns.

Various clustering algorithms are tested in order to establish which ones are suitable for a specific task. The comparison among those clustering algorithms has been accomplished by the utilization of clustering validity indices based on properly defined metrics and indicators. However, it is a user task to decide the final number of desired clusters. In this paper, a combination of conventional clustering algorithms and bio-inspired optimization clustering algorithms are presented. Six clustering techniques—K-means clustering, Hierarchical clustering, Fuzzy C-means (FCM) clustering, Self-Organizing Maps (SOM) clustering, Ant Colony clustering and Bat clustering—are tested, in order to investigate the most appropriate clustering method for establishing the PV power grouping process. The conventional clustering algorithms are chosen because of their extensive utilization in literature. In addition, each one represents a clustering category, except for Ant Colony and Bat which both fall in the same