

## Shade Detection in an Industrial Photovoltaic System with Deep Neural Networks

Claudia Navarrete-García \* Luis E. Garza-Castañón \*  
Adriana Vargas-Martínez \* Luis A. Garza-Elizondo \*\*

\* *Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias  
(e-mails: {a01032856, legarza, adriana.vargas.mtz}@itesm.mx)*

\*\* *Aixware Technologies, Monterrey, NL, México (e-mail:  
albertogarza07@gmail.com)*

---

**Abstract:** In this paper, a framework to detect shading in a Photovoltaic (PV) system is presented. This phenomenon negatively affects the production of electrical energy in a PV system and occurs when clouds or some obstacles block the sunlight hitting the panels. Early results of collected data analysis and the framework structure that will be used on this ongoing project are explained. The PV system is installed at Schneider Electric in Apodaca, N.L., México. Climatic variables, sun position and electric current and voltage from 20 solar panels were sampled every 30 seconds for five months (January to May 2017). It is shown that different data patterns are presented in sunny and partially cloudy days. After these data are given some pre-processing, they are used as inputs of a Deep Belief Network (DBN), which outputs the affected panel. This information can later be used by another system to implement corrective actions to mitigate the problem.

**Keywords:** Photovoltaic Systems, Shade Detection, Data Analysis, Deep Belief Networks

---

### 1. INTRODUCTION

The installation of photovoltaic systems has increased in recent years due to several factors: The falling prices of silicon and PV modules, technological advancements in large scale manufacturing, governmental incentives, maturation and proliferation of favourable interconnection agreements and continued technological improvement of power converter technologies (Obi et al., 2016). According to the 2015 intercensal survey made by the National Institute of Statistics and Geography (INEGI) in Mexico, there are more than 160,000 households using Solar Panels.

Besides being a clean energy, PV systems have the advantage of having low maintenance and long life, approximately 25 years. Some of the disadvantages of solar photovoltaic systems, if not properly installed, are: failures in batteries, wiring and interconnections with the network, which can potentially lead to hazardous fire. A study in the UK for a period of two years, states that energy losses caused by failures can go up to 18.9% (Silvestre et al., 2015). However, one of the most important factors, is that the efficiency of solar panels in the PV system is directly affected by the variability of weather conditions.

The response of photovoltaic systems depends on extrinsic factors such as radiation intensity, ambient temperature, cell temperature, air velocity, humidity, and the effect of shadows (Silvestre et al., 2014).

The shadow effect is a phenomenon that occurs when some of the solar cells in a photovoltaic system are partially obstructed by buildings, birds, clouds or any other object (Hosseinzadeh et al., 2016).

The loss in efficiency when shadows are present it is important because they can also cause hot spots. They can irreversibly damage the PV modules, especially when extreme shading situations are present. The use of bypass diodes could protect the modules from the hot spots, but also they create multiple power peaks, which increase the complexity of the Maximum Power Point Tracking (MPPT). Therefore, conventional detection techniques might be inefficient (Hosseinzadeh et al., 2016).

Currently, there are four main categories to solve the shadow effect. The first group includes modifying MPPT techniques. The second, refers to different configurations arrangements for connection of solar panels. The third, are the different architectures of photovoltaic systems. And the fourth refers to the use of different topologies converters (Hosseinzadeh et al., 2016).

The aim of this paper is to present a framework to detect the shading effect in an industrial PV system, early results of data analysis and experiments with a DBN classifier. The organization of the paper is as follows: First an introduction to DC faults is made; after this, a review is made about the shading effect and methods to mitigate the problem. Then a preliminary study of data collected in a real installation is presented. Also, a proposal is made of how neural networks

trained with deep belief techniques could help to predict the effects of shading on PV systems. Finally, some preliminary results obtained with a DBN classifier are shown.

## 2. DC FAULTS

Efficiency in the solar panel is critical to improve the performance of the installed PV system. To have greater efficiency is necessary to understand the effects of energy losses as well as monitoring and controlling them so they can be minimized.

DC faults are attributed the losses that depends on different factors like climatic conditions and PV module performance. Some of the losses that depends on climatic conditions are: the operating cell's temperature, solar irradiance level, PV shading, and the angle of irradiance. Others are related to the efficiency of the MPPT technique employed, component failure, soiling effect, etc. In the next section, the shading phenomenon is explained in detail.

## 3. SHADING EFFECT

Partial shading occurs when some cells in a module are obstructed by clouds, buildings, etc. as shown in figure 1. As the irradiation level in the shaded cells reduces, the non-shaded cells remain operating at a higher current. Since all the string current must be the same in all the series-connected cells, the result could lead to a mismatch which causes that the shaded cells operate in a reverse bias region to be able to conduct the same current as the non-shaded cells.

The bias voltage  $V_{bias}$  is the reverse voltage at which the shaded cells must operate to support the common string current. On figure 2, can be seen the behavior of a string with shaded and non-shaded zones. So, if shaded cells consume power due to the reverse voltage polarity, then the maximum extractable power from the shaded PV array decreases (Bidram et al., 2012).

A high bias voltage may also lead to thermal break down, which create a hot spot. It is also important to mention that the most shaded PV module in a chain limits the total current flowing in that chain (Firth et al., 2010).

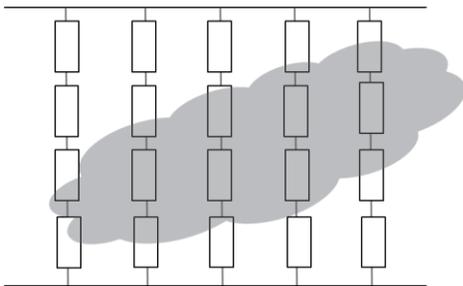


Fig 1. Representation of a photovoltaic system under shading effect.

Different approaches have been followed to deal with shading problem in PV systems.

### 3.1 Shading by stationary objects

If the object that produces the shading is stationary and did not move with time, then the position of the sun is the factor which determines the shading effect (Firth et al., 2010). In this paper, authors propose a method in which the position of the sun is calculated for in five-minutes interval, using the time, date and the latitude of the site. In this work four types of faults are detected:

- *Sustained zero efficiency fault.* Here, there is a component failure or a system isolation and it can be identified if the fault occurs for a long-time.
- *Brief zero efficiency fault.* Identifies the fault because it occurs during a short period and it could be originated by an inverter shutdown.
- *Shading fault.*
- *Non-zero efficiency, non-shading fault.* This can be the case for an inverter MPPT failure or other type of faults.

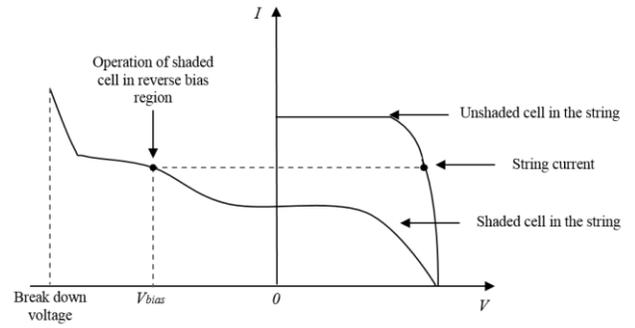


Fig 2. Current-voltage curve of a PV cell operating in a reverse bias region.

### 3.2 Partial shading

This condition has a dynamic behaviour; it depends for example, on the cloud evaluation and on the position of surrounding obstacles near the PV array. In most cases the effect of partial shading disappears quickly, so it is possible to identify the partial shading effects in terms of voltage and current.

In (Silvestre et al., 2015) the total percentage of reduction in output voltage  $\Delta V$ , due to partial shading, is computed by:

$$\Delta V = \left( \frac{V_{mo} - V_m}{V_{mo}} \right) = \left( 1 - \frac{NR_v}{NR_{vo}} \right) \quad (1)$$

Where:

$$NR_v = \frac{V_m}{V_{oc}} \quad (2)$$

$$NR_{vo} = \frac{V_{mo}}{V_{oc}} \quad (3)$$

Where  $V_m$  is the voltage of the maximum power point at the DC output of the PV generator and  $V_{oc}$  is the open circuit voltage of the PV array.  $V_{mo}$  is the maximum power point voltage of the output of the PV array in the absence of faults.

Then the output current,  $\Delta I$  could be expressed as:

$$\Delta I = \left( \frac{I_{mo} - I_m}{I_{mo}} \right) = \left( 1 - \frac{NR_c}{NR_{co}} \right) \quad (4)$$

Where:

$$NR_c = \frac{I_m}{I_{sc}} \quad (5)$$

$$NR_{co} = \frac{I_{mo}}{I_{sc}} \quad (6)$$

$I_m$  is the current of the maximum power point at the DC output of the PV generator and  $I_{sc}$  is the short circuit current of the PV array.  $I_{mo}$  is the maximum power point current of the output of the PV array in the absence of faults.

Finally, the proportion of DC power losses due to the shading effect could be calculated as:

$$P_{loss} = \left( 1 - \frac{NR_c}{NR_{co}} \frac{NR_v}{NR_{vo}} \right) \quad (7)$$

In (Hosseinzadeh et al., 2016) authors present a method that can estimate or detect if there is a uniform or partial shading based on the comparison between the difference of the actual output power and the estimated output power:

$$|P_{PV}^{act} - P_{PV}^{est}| \leq \varepsilon_1 \quad (8)$$

Where  $\varepsilon_1$  is, the standard deviation calculated from simulated output voltage of the PV array for clear sky. Therefore, if equation (8) is satisfied, no uniform shading fault is detected, otherwise, a uniform shading fault is present in the PV system. However, to understand if there is a uniform shading fault, a calculation of a voltage ratio is requested:

$$R_v = \frac{V_{PV}^{est}}{V_{PV}^{act}} \quad (9)$$

Finally, if  $R_v > 1 + \varepsilon_2$  is accomplished, then a uniform shading fault has occurred. Where,  $\varepsilon_2$  is the output power of the PV system in clear sky.

In (Mekki, et al., 2016), is presented a method based on neural networks, to detect faults originated by partial shading. The methodology consisted in acquire data for 30 days, such as: solar irradiance, cell temperature (used as input variables), photovoltaic current and voltage (used as output neurons). The 70% of the data were used to compute the gradient and to update the network weights and bias. The rest were used to validate the model by comparing the estimated data with the actual output.

Partial shading problem leads to the creation of local maximum power points in the power-voltage profile of the PV array. To mitigate this issue, some of the techniques used are presented in the next section.

### 3.3 Techniques to mitigate partial shading

In (Chine, et al., 2014) are defined three types of solutions that can help to increase the power output of the photovoltaic system when there is a power mismatch, caused by the partial shading effects. These solutions are: Maximum Power Point Tracker, Array configurations and PV system architectures.

#### 3.3.1. Array configurations

A. *Simple Series-parallel (SP) array*. All solar cells are connected in series which are then called strings. Then, these strings are connected in parallel. The main advantages of this configuration is that it has a minimum wire size, so this helps to reduce the power loss due to wiring (Mekki, et al., 2016).

B. *Total-cross-tied (TCT) array*. All solar cells are connected in parallel creating modules. Then, in series. The advantage is that the interconnections between the PV modules enable different current flowing, so this means that it can improve the maximum power point of the PV system under partial shading (Mekki, et al., 2016).

C. *Reconfigurable PV array*. The system has a fixed part and a small adaptive bank of PV modules used for energy compensation. When the shading is detected, a switching matrix reconfigures the PV modules.

#### 3.3.2. PV system architectures

A. *Centralized architecture*. Is the most conventional architecture. Individual MPP cannot be tracked so this architecture is the one that it is more prone to shading and mismatching loss.

B. *Series-connected micro inverters*. The *dc-dc* converters are used to track individual maximum power point; the output of these converters is then fed to a central inverter.

C. *Parallel-connected micro inverters*. The modules are connected to the central inverter.

D. *Micro inverters*: Each of the modules has its own MPPT.

#### 3.3.3. Maximum Power Point Tracker

A. *Power Curve Slope*. This technique uses a rate of change between the power and the voltage at different points.

B. *Load-Line Maximum Power Point Tracking*. There are two types of load-line MPPT which are identified as type I and type II. In type I, the load line is defined by the ratio of  $V_{mpp}$  to  $I_{mpp}$ . In type II, a linear function brings the operating point into a vicinity of the global maximum, here it measures the last two operating points of the PV array and it has also included the number of strings that are connected in parallel. For both load line techniques, a periodic tuning is necessary.

C. *Dividing Rectangles*. Uses a Lipschitz condition to track the global maximum of the PV array power as a function of the array voltage.

D. *Power Increment*. Uses the augmented power converter as a tuneable constant power load. At each step, the amount of power is increased, resulting in progressing toward the global MPP. It is highly accurate and it is a fast tracker.

E. *Instantaneous Operating Power Optimization*. This procedure takes into consideration the irradiance and the *p-n* junction temperature; is highly accurate and is a fast tracker.

F. *Fibonacci Search*. Can find the MPP under uniform and non-uniform radiation. It consists in making iteration of two check points in the curve within an interval.

G. *Particle swarm Optimization*. It is a metaheuristic method. It is accurate and does not require a periodic tuning.

H. *Artificial Neural Networks*. Is a prediction method which uses previous data to be trained. This method will be discussed in the section 6.

#### 4. DATA MEASUREMENT AND COLLECTION

##### 4.1 Photovoltaic System at Schneider Electric

The photovoltaic system used to collect data is the solar field of Schneider Electric located in Apodaca, Nuevo León (Figure 3). The system consists of a main distribution panel connected to the grid with 9 inverters *Conext CL CL25000E* from Schneider Electric. The inverter has 5 strings of 20 modules each; a module has a maximum power of 250 KW.



Fig. 3. a) Front view of the solar field (left). b) Back view from the solar field (right). The white box on the left is the Schneider Electric Inverter.

For this work, only one string with twenty solar panels was monitored. Ten of the panels are each connected to an optimizer device (oriented to maximize power production by tracking online the MPP) and the other ten only have a monitoring system. Odd numbered panels are connected to the optimizer while pairs are to the monitoring equipment.

##### 4.2 Data acquisition

There are three data acquisition systems. The first comes directly from the inverter every 13 minutes, the variables monitored are: Temperature ( $^{\circ}\text{C}$ ), Irradiance ( $\text{W}/\text{m}^2$ ) and Wind speed ( $\text{m}/\text{s}$ ). These data are downloaded from a web page called *Conext Smart Box*.

The second data acquisition system can monitor the Voltage ( $\text{V}$ ), Current ( $\text{A}$ ), and DC output power ( $\text{W}$ ). The variables come directly from the monitoring and optimizers system installed in each of the 20 panels. The information is retrieved every 30 seconds.

In the third system, sun position is obtained every 5 minutes from a web page that takes in consideration the coordinates of the physical location of the solar installation.

As the sampling time of variables are different, it is required to create an interpolation of some data to have the same sampling time of 30 seconds.

#### 5. PRELIMINAR DATA ANALYSIS

To determine the relationship between inputs and outputs, a correlation table was created. In Table 1, the values of the correlation are shown. The variable that most affects the power is the current (0.994 correlation).

Table 1. Value of correlation between variables

	Sun Elevation	Azimuth	POA irradiance ( $\text{W}/\text{m}^2$ )	Ambient temperature ( $^{\circ}\text{C}$ )	Wind speed ( $\text{m}/\text{s}$ )	Voltage ( $\text{V}$ )	Current ( $\text{A}$ )	Power ( $\text{W}$ )
Sun Elevation	1.000							
Azimuth	-0.013	1.000						
POA irradiance ( $\text{W}/\text{m}^2$ )	0.705	-0.379	1.000					
Ambient temperature ( $^{\circ}\text{C}$ )	0.310	0.392	0.338	1.000				
Wind speed ( $\text{m}/\text{s}$ )	0.200	0.136	0.174	0.191	1.000			
Voltage ( $\text{V}$ )	0.107	-0.117	0.108	-0.046	0.066	1.000		
Current ( $\text{A}$ )	0.748	0.006	0.753	0.561	0.218	0.004	1.000	
Power ( $\text{W}$ )	0.741	-0.009	0.760	0.545	0.232	0.061	0.994	1.000

Also, irradiance is the value that correlates more to the output variable power, with a high positive correlation of 0.760, which is next followed by the sun elevation with a 0.741 positive correlation.

In February 2017, some days (7,8,9,11,12,17,18,21,22,24,27) were sunny and the graph Power versus Time is represented in figure 4, where the behaviour of the graph can be fitted to a second-degree polynomial. Here, it is shown that power is directly related to the irradiance. An increase in the first half of the graph is observed and then start decreasing when the hour of the day advances.

In March, there were only three days in which the power behaves “normal” for a sunny day.

In a cloudy day graph is no longer possible to interpolate as a second-degree polynomial, as it has more abrupt changes as shown in figure 5.

A detection of clear difference between the behaviour of an optimized solar panel and the monitored was observed. In February, the monitored solar panels have higher power value that the optimized, while in March the optimized solar panels have a greater power output.

The early analysis of PV system data show that the pattern behaviour of power in cloudy days, in combination with the other variables, can help us to predict the shading condition on the PV system.

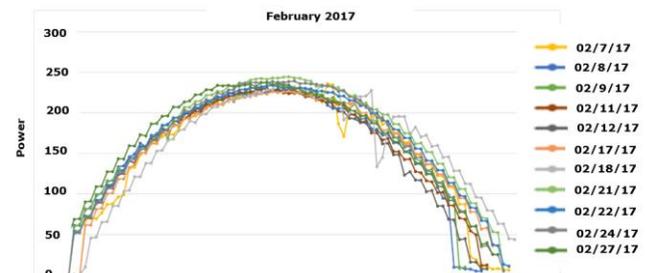


Fig. 4. Power vs. Time graph in sunny days (February)

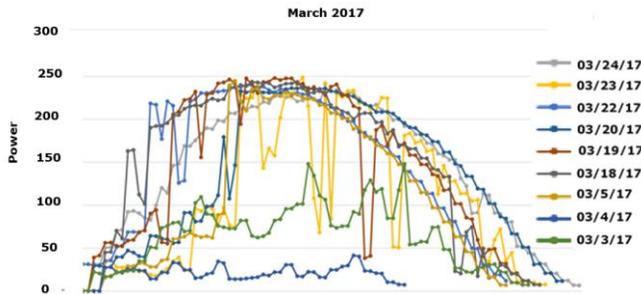


Fig. 5. Power vs. Time graph in cloudy days (March)

## 6. NEURAL NETWORKS

### 6.1 Neural networks in Photovoltaic Systems

Artificial Neural Networks (ANN) are powerful algorithms able to create accurate non-linear mapping of data. They have been used in shade detection, as was described before (Mekki, et al., 2016), but also in maximum power point tracking problems. Here, most of the applications require extra sensors to obtain data such as temperature in the panel or the solar radiation. For instance, in (Rahmann et al., 2016) they proposed a method where the ANN are used to identify a global maximum power. The method consists in measuring the current and voltage in the PV array terminals with a sampling time of few seconds and then calculating a power differential that is compared to a threshold when the irradiance distribution changes. It uses the sigmoid transfer function, and the artificial neural network was trained with the backpropagation algorithm with Levenberg-Marquardt optimization method. The inputs of the neural network are the voltage and current of all the measured points of solar irradiance and the respective temperatures.

### 6.2 Deep learning

Deep learning algorithms are a subset of machine learning methods. They aim to discover multiple levels of distributed representations. Recently, neural networks trained with deep learning methods have been successfully applied to different domains, such a semantic text parsing, pattern recognition in images, natural language processing, etc. Generally these algorithms can be divided into four categories according to the basic method they are derived from: Convolutional Neural Networks, Restricted Boltzmann Machines, Autoencoder and Sparse Coding (Schmidhuber, 2015). Deep neural networks have more hidden layers and weights are learned with different methods than traditional neural networks.

To evaluate the abilities of this kind of algorithms, a quantitative measure of its performance must be designed. Depending on the way to acquire knowledge from data, the machine learning algorithm can be classified into two categories: supervised learning and unsupervised learning.

### 6.2 Deep Belief Networks

Deep Belief Networks are a type of neural network. They are generative probabilistic models with many hidden layers connected to each other with many latent variables or neurons. Each neuron is connected to neurons in neighboring layers and there is no connection between neurons in the same layer. In this type of networks, the connections of the last layer are indirect and the connections between all the other layers are in direct way as shown in figure 6.

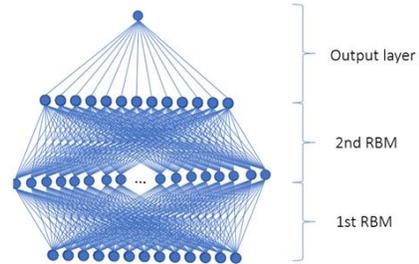


Fig. 6. Architecture of a Deep Belief Network

For the training of the network, first each pair of contiguous layers is trained as if they were a Restricted Boltzmann Machine (RBM), using the contrastive divergence method. Each layer is trained to model the distribution defined by sampling (using Gibbs sampling) the hidden neurons of the previous layer.

After modeling the joint distribution between the observed (or input) information and the intermediate hidden layers, are obtained the set of parameters which help to have multiple levels of representation for the input. After this step is performed, it is possible to fine-tune the parameters found in a supervised way as a traditional multi-layered neural network. Here, a retro propagation of the error is used, adding an extra layer of output and optimizing the weights using the descending gradient optimization method.

## 7. EXPERIMENTS AND RESULTS

Deep learning algorithms are more successful when many samples are used for training, so it is adequate for an industrial application as the one described here. Figure 6 shows the proposed framework, where inputs coming from the PV system: irradiance ( $W/m^2$ ), ambient temperature ( $^{\circ}C$ ), sun position ( $^{\circ}$ ) and electric current for each panel ( $I$ ), are pre-processed (e.g. noise elimination, sampling time adjustment) and then feed to the input layer of the DBN. The output of the DBN is the number of the affected panel. These outputs can be used by another system to mitigate the effect of shading.

For the experiments, 39,137 samples collected during May 2017 were used. A Different panel was fully covered during several days to simulate the shading. To train and test the DBN a subset with 1224 samples was used, where samples with missing values were discarded. Also, in order to have more accuracy in the classification, a specialized DBN was trained for each class, where each class corresponds to 4 shaded panels and none affected (5 classes in this dataset).

In Table 2, the parameters of the DBNs are presented. The accuracy during the training phase was very similar for each class, close to 100% (an example is shown in Figure 8). In testing mode, 100% classification efficiency was achieved for each class.

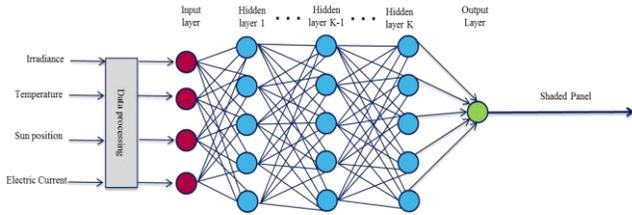


Fig. 7. Block diagram of the system with Deep Belief Networks

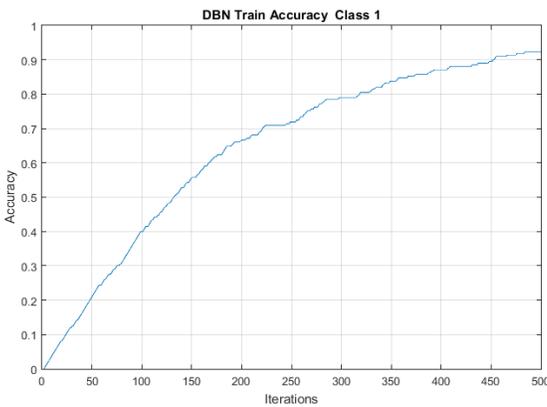


Fig. 8. Accuracy of the DBN for class 1 during training

Table 2. DBN Parameters

Network structure	22x30x30x30x1
Number of training examples	70%
Number of testing examples	30%
Learning rate (Unsupervised phase)	0.15
Iterations (Unsupervised phase)	800
Learning rate (Supervised fine-tuning)	0.075
Iterations (Supervised fine-tuning)	500

## 8. CONCLUSIONS

The advances and experiments on the implementation of shade detection in an industrial PV system have been presented in this paper. One string with 20 panels was instrumented to monitor the relevant variables. An early study on data collected during several months show the difficulties inherent to a real world application: nonlinear behaviour, noise, lack or excess of data, sampling time mismatch, etc. The data analysis study helped us to select the critical variables which will feed the DBN to detect when shading phenomenon is present in a panel. These powerful algorithms are more efficient than other traditional neural networks, mainly because of the existence of more hidden layers and also a robust training processing phase (providing there are sufficient data to input).

Experiments conducted to simulate the shading effect in the panels deliver a dataset that later was used to train and test the accuracy of Deep Belief Networks. Excellent results were obtained in the correct classification of the status of panels. Future work will address the problem of missing information in dataset and experiments with partially shaded panels.

## REFERENCES

- Bidram, A., Davoudi, A., & Balog, R. S. (2012). Control and circuit Techniques to Mitigate Partial Shading Effects in Photovoltaic Arrays. *IEEE Journal of Photovoltaics*, 2(4), 532-546.
- Chine, W., Mellit, A., Pavan, A. M., & Kalogirou, S. (2014). Fault detection method for grid-connected photovoltaic plants. *Renewable Energy*, 66, 99-110.
- Firth, S., Lomas, K., & Rees, S. (2010). A simple model of PV system performance and its use in fault detection. *Solar Energy*, (84), 624-635.
- Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. (2016) Deep Learning. *Deep Learning Book*. MIT
- Hosseinzadeh, M., & Rajei SImase, F. (2016). Determination of maximum solar power under shading and converter faults-A prerequisite for failure-tolerant power management systems. *Simulation Modelling Practice and Theory*, (62), 14-30.
- Mundo-Hernández, J., De Celis Alonso, B., Hernández-Álvarez, J., & De Celis-Carillo, B. (2014). An overview of solar photovoltaic energy in Mexico and Germany. *Renewable and Sustainable Energy Reviews*,(31), 639-649.
- Mekki, H., Mellit, A., & Salhi, H. (2016). Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. *Simulation Modelling Practice and Theory*, (67), 1-13.
- Obi, M., & Bass, R. (2016). Trends and challenges of grid-connected photovoltaic systems-A review. *Renewable and Sustainable Energy Reviews*, (58), 1082-1094.
- Rahmann, C. (2016). Mitigation Control Against Partial Shading Effects in Large-Scale PV Power Plants. *IEEE Transactions on Sustainable Energy*, 7(1), 173-180.
- Silvestre, S., Aires da Silva, M., Chouder, A., Guasch, D., & Karatepe, E. (2014). New procedure for fault detection in grid connected PV systems based on the evaluation of current and voltage indicators. *Energy Conversion and Management*, (86), 241-249.
- Silvestre, S., Kichou, S., Chouder, A., Nofuentes, G., & Karatepe, E. (2015). Analysis of current and voltage indicators in grid connected photovoltaic systems working in faulty and partial shading conditions. *Energy*, (86), 42-50.
- Schmidhuber, J. (2015). Deep Learning in Neural Networks: An Overview. *Neural Networks* (61), 85-117.