




Article

A New Data-Based Dust Estimation Unit for PV Panels

Mostafa. F. Shaaban ^{1,2}, Amal Alarif ¹, Mohamed Mokhtar ^{1,3,*}, Usman Tariq ¹,
Ahmed H. Osman ¹ and A. R. Al-Ali ¹

¹ Electrical and Computer Engineering Departments, American University of Sharjah, Sharjah 26666, UAE; mshaaban@aus.edu (M.F.S.); amal.alarif@dewa.gov.ae (A.A.); utariq@aus.edu (U.T.); aosmanahmed@aus.edu (A.H.O.); aali@aus.edu (A.R.A.-A.)

² Electrical and Computer Engineering Department, University of Waterloo, Waterloo, ON N2L 3G1, Canada

³ Electrical Power and Machines Department, Faculty of Engineering, Ain Shams University, Cairo 11517, Egypt

* Correspondence: mmahmed@aus.edu or m.mokhtar@eng.asu.edu.eg

Received: 1 June 2020; Accepted: 30 June 2020; Published: 13 July 2020



Abstract: Solar photovoltaic (PV) is playing a major role in the United Arab Emirates (UAE) smart grid infrastructure. However, one of the challenges facing PV-based energy systems is the dust accumulation on solar panels. Dust accumulation on solar panels results in a high degradation in the output power. The UAE has low intensity rainfall and wind velocity; therefore solar panels must be cleaned manually or using automated cleaning methods. Estimating dust accumulation on solar panels will increase the output power and reduce maintenance costs by initiating cleaning actions only when required. In this paper, the impact of natural dust accumulation on solar panels is investigated using field measurements and regression modeling. Experimental data were collected under various real weather conditions and controlled levels of dust. Moreover, this paper proposes a data-driven approach based on machine learning to estimate the accumulated dust level on solar panels. In this approach, a dust estimation unit based on a regression tree model has been developed to estimate the dust accumulation. This unit is trained using experimental records of solar irradiance, ambient temperature, and the output power generated from solar panels as well as the amount of dust at these conditions. The proposed unit is evaluated through different case studies with a random amount of dust applied to the solar panels to demonstrate the accurate performance of the proposed unit.

Keywords: solar panels; regression models; dust accumulation; renewable energy; dust estimation

1. Introduction

Thermal power plants have a great impact on global warming. Furthermore, the fossil fuels used in such plants have an increasing cost and are limited in nature. Therefore, the need for integrating renewable energy into power grids has become an inevitable issue worldwide. Various renewable resources can be integrated into power grids, for example solar power, wind, biomass, geothermal and fuel cells.

In the Middle East, local demand for oil and gas has increased rapidly pushing the countries to look for other sources to generate electricity. Renewable energy and solar power technologies are currently considered a feasible alternative energy source in the Middle East and other countries around the world. The Middle East region has an appropriate climate and geographical location for producing the highest solar power generation in the world [1]. Therefore, solar power is the most commonly used renewable energy source in the region. The expansion of using solar energy technology has the first priority in the energy strategic plans for many countries in the region. The United Arab Emirates

is also eager to integrate solar energy into electric power grids. Currently, different solar projects have tendered in both Dubai and Abu Dhabi with expected growth of solar energy in the coming years. The United Arab Emirates (UAE) aims to generate 30% of the power demand from renewable sources by 2030. Dubai aims to install 1000 MW of concentrating solar power (CSP) technology and 4000 MW of solar photovoltaic (PV) technology by 2030 [2]. The UAE has high solar irradiance throughout the year. It is worthwhile to note that the average global solar radiation ranges from 1900 to 2300 kWh/m² [3].

Photovoltaic panels are manufactured using semi-conductor materials. The inherent properties of these materials affect the efficiency of photovoltaic systems. Therefore, the output power of solar panels depends on the materials used to manufacture the panels and the coating on the glass. Furthermore, the design of the power plants also affects the output of solar panels, such as the orientation of the installed panels, the amount of sun exposure to the power plant and the sun tracking systems. Figure 1 illustrates the factors affecting the yield of solar photovoltaic panels. In this context, there are many challenges facing the spread of solar energy in the world and particularly in the Middle East. The dust effect is the most salient factor. The dust accumulation rate in the UAE is high due to the geographic location of the country. Dust accumulation on photovoltaic panels leads to high degradation in the efficiency of the solar panel and hence a significant reduction in the output power. Dust accumulation is influenced by two factors; environmental conditions and dust properties. Dust properties refer to the weight, shape, and size of the particle, while environmental conditions include weather conditions and geographical location [4]. Wind is another factor that affects the settlement of dust. High-speed wind can remove the dust settled on the solar photovoltaic panels, while areas with low wind speed suffer from the worst degradation in solar photovoltaic output power due to the high accumulation of dust. The ambient temperature and humidity also affect dust settlement. High temperature and humidity areas suffer from higher dust accumulation on solar panels as the dust becomes wet with high humidity and sticks to the glass of the solar panels, and hence reduces their efficiency. Consequently, cleaning the panels on a daily basis guarantees high efficiency of solar panels. However, this procedure is not economical and may lead to wasted resources if no cleaning actions are required [5]. Therefore, detecting the dust levels and initiating the necessary cleaning actions is vital to increase the yield of photovoltaic panels and to reduce maintenance cost. Note that, dust estimation can be implemented using numerical methods, analyzing satellite images or machine learning models.

The effect of dust accumulation on solar photovoltaic panels is location-dependent as it depends on weather conditions and dust particles' size. Moreover, the behavior and response of the output power of PV panels' differ with different locations, dust properties, and the environment. Therefore, studies done in one country cannot be generalized to other countries. However, the process proposed in this paper can be applied in other countries but not the data and results. Thus, detecting the amount of dust and initiating the necessary cleaning action when needed is vital to increase the yield of photovoltaic farms and to reduce maintenance cost. However, the methodology applied in this paper is distinguished by generality, which can be used to detect the dust accumulation whatever the location.

In this paper, experiments are conducted to measure the PV panels' output power at various dust accumulation levels, temperature and solar irradiance. Furthermore, the outcomes obtained from experiments will be utilized in developing the dust estimation unit based on machine learning. The main contributions of this paper are as follows:

- Study the impact of dust accumulation on the output power of solar photovoltaic panels experimentally under real environmental conditions in the UAE.
- Propose a dust estimation unit based on a regression tree that estimates the amount of dust accumulated on the solar photovoltaic panel to initiate the cleaning actions.
- The detector is developed using a field measurements dataset that includes the solar irradiance, the ambient temperature, PV panels' output power as the main predictors in addition to the amount of dust as the target or response variable.
- The proposed detector is evaluated through different case studies, including premeasured amounts of dust as well as random amounts of dust applied on solar panels. Moreover, the performance

of the proposed dust estimation unit is compared with another unit based on Artificial Neural Network (ANN) to demonstrate the potential of the proposed unit.

The rest of paper is organized in six sections. Section 2 provides a background and literature review. Section 3 presents the proposed methodology, while the experimental setup is presented in Section 4. The results and the different cases used to verify the proposed model are discussed in Section 5. Finally, Section 6 includes the conclusion.

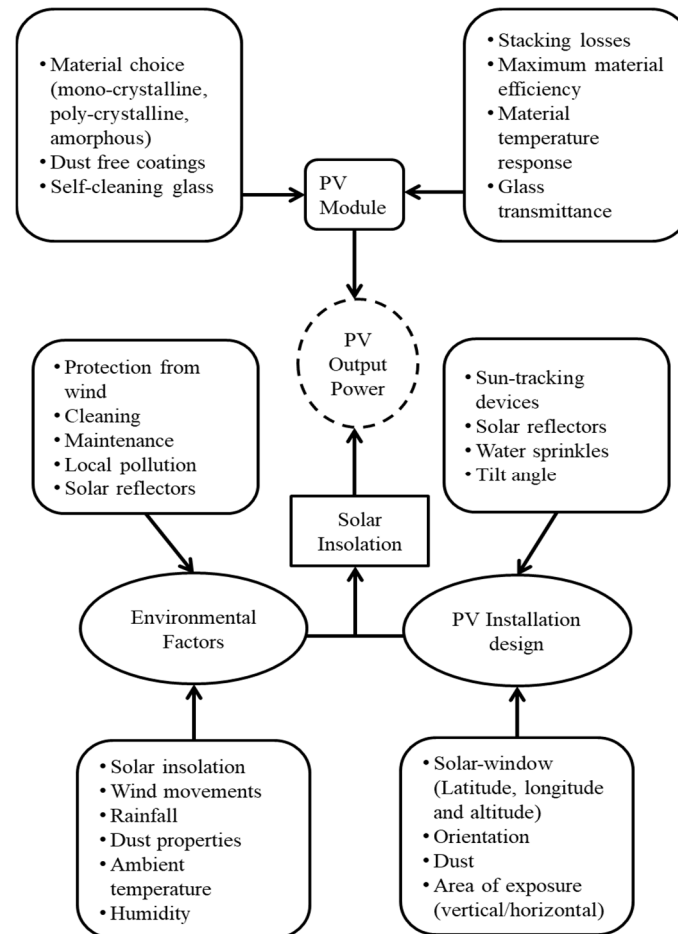


Figure 1. Factors affecting solar photovoltaic (PV) panels' output power [4].

2. Related Work

The different approaches presented in the literature to study the impact of the environment on the output of solar photovoltaic were either by conducting experiments or developing prediction models. In [6], a review study was conducted to survey the impact of dust, humidity and air velocity separately and as a group. The survey determined that the efficiency of solar panels drops significantly with fine particles compared to coarse particles. Furthermore, a larger tilt angle resulted in less dust accumulation; on the other hand, humidity resulted in more dust coagulation. In [7], experiments were implemented where the pollutant type and weight were the variables in the experiments to illustrate their impact on the PV panels' output power. In [8], three different artificial pollutants were applied on the solar photovoltaic panel with different masses. The pollutants used in the experiments were red soil, limestone and carbonaceous fly-ash particles. The results of the experiment showed that Red soil, Limestone and ash resulted in a reduction of 19%, 10% and 6% in the output power from solar panel compared to a clean panel, respectively. A similar study was conducted on a three different artificial pollutants which were red soil, sand and ash in [9]. In [3], a 23-day study was carried out in the UAE to study the effectiveness of self-cleaning coating material on solar panels. The results showed that panels

with and without this coating material have almost the same performance. In [10], another study was conducted in the UAE to study the dust effect on a solar photovoltaic panel. Dust amounts of 0.0063 g/m^2 to 0.36 g/m^2 were distributed on solar panel modules. The results showed that the relation between the dust weight and the drop-in output photovoltaic power was linear. The authors in [11] presented satellite images and a support vector machine (SVM) model to predict the solar irradiance and cloud movement. The work in [12] presented an auto-regression model to forecast the output power from solar panels for up to 36 h in Denmark. In [13], the ANN tool was used to predict the global solar irradiance from a set of inputs obtained from meteorological stations. The ANN provided a Root Mean Square Error (RMSE) of less than 20%. The authors in [14] presented a regression model based on sigmoid function to estimate the hourly diffuse solar irradiation under all weather conditions. However, the clearness index and relative optical mass were considered as the predictors. The model provided a relative RMSE in the range of 25–35%. The work in [15] presented a model based on ANN to estimate the output power for different time horizons in different seasons in Turkey. A 750 W PV panel was installed. The obtained RMSE for each time and season was recorded. In [16], four different forecasting techniques, including ANN, were developed to estimate the PV panels' output power in California. One MW solar panel field was used to forecast 1 h and 2 h ahead. The ANN offered the best performance compared to the other forecasting techniques. The presented model offered a RMSE of 15% compared to the other models, which provided an error of up to 20%. However, this research did not consider the different environmental conditions affecting the output power of the system. In [17], an experiment was conducted in Spain to study the impact of rainfall in cleaning photovoltaic modules. The study showed that the reduction in energy was 20% in solar photovoltaic power plants in periods without rain compared to only 4.4% energy reduction in rainy periods due to dust accumulation. In [18], another study was implemented for 4 months in Morocco to estimate the dust accumulation through the presence of a correlated input of output power of solar photovoltaic panel and rainfall. The rainfall data were obtained from a meteorological data center. In [19], high rate of dust, low frequency and intensity of rain were considered to study the effect of dust accumulation on the performance of PV panels in the Jazan Region. It was found that the regular dust accumulation reduced the PV efficiency by 10%. However, lower tilt angles caused higher dust accumulation than higher tilt angles. The authors in [20] presented an image processing-based system to estimate the amount of dust accumulated on the surface of PV panels, where small plasticised paper was used as indicator to make image processing. Most of the existing research focused on the effect of dust accumulation on PV panels. However, only a few studies focused on estimating the amount of dust accumulation to initiate the suitable cleaning actions.

3. Problem Statement and Proposed Methodology

Accurate prediction of the accumulated dust on the solar panels is a vital issue for investors and grid operators. The objective of this research is to develop a Dust Estimation Unit to estimate the amount of accumulated dust on solar panels in the UAE from measured photovoltaic output power, solar irradiance, and temperature, as shown in Figure 2. Hence, cleaning procedures can be initiated if the dust accumulation is high. Machine learning is utilized to develop this detector. The first and vital step is to collect all the required data, so that a machine learning model can learn the behavior of PV panels. Then, the collected data is utilized to train different regression models and study the accuracy of each model. Finally, the proposed unit is assessed for its accuracy throughout other cases studies with various dust levels. The details for each previous step will be explained in the following subsections.

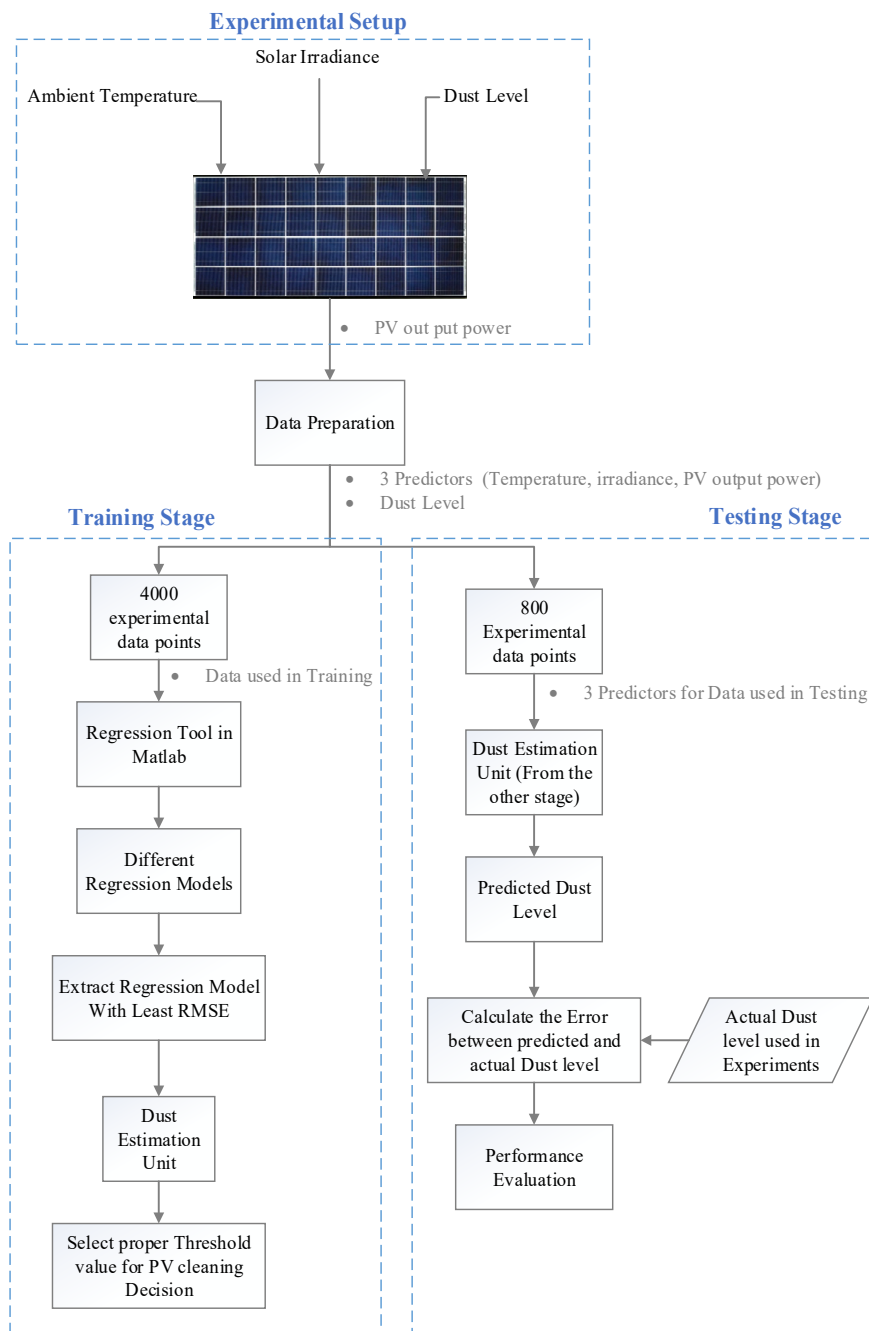


Figure 2. The proposed approach flowchart.

3.1. Data Preparation

Data preparation is the first and vital step in the model, as shown in Figure 2, where these data are fed to the regression models to learn how to predict the dust level. Regression is a type of machine learning; it is commonly used in forecasting and estimation problems [21]. Further, it is a statistical approach that is utilized to predict the relationship between one or multiple input variables called predictors and a single output variable called the response variable. Data fed to this model should be arranged in a matrix that begins with the predictors and ends with the response variable. In this research, three variables will be used as the predictors to predict a single output or response variable. The three predictors are solar irradiance, ambient temperature and the output power from PV at these conditions, while the dust level is the response variable. In order to create the required data

set that includes these readings, experiments that will be explained later were conducted at different times of the day and for multiple days to collect a wide range of data. The output power of the solar photovoltaic panel was measured with different levels of natural dust while recording the ambient temperature and solar irradiance. A total of 4800 data points were collected and these are divided into two sets. One set will be utilized in the training stage while the other set will be used in the testing stage.

3.2. Model Training

The second stage is to train the model using the experimental collected data to make the model learn how to predict the dust level with knowledge of the PV panels' output power, the ambient temperature and the solar irradiance. The collected data (predictors, response) are divided into two sets. One data set is used to develop the model in the training stage and another set is used to validate the accuracy of the model in the testing stage. The temperature, solar irradiance and the output power from PV panels' as three predictors as well as the dust level as the response variable are fed to the regression models. The regression models try to predict the dust level and compare the predicted response with the actual response. The regression models utilize 5-fold Cross-Validation in the model training. It is widely utilized in the regression models to evaluate and estimate the misclassification error. In this technique, the data used in the training stage is randomly divided into five groups. Then, one group will be used as testing data and the remaining groups will be used as training data. Therefore, the model will be trained using the four groups and evaluated using the one test group. This procedure is repeated various times with different groups to ensure that each group has been used once to test the model. After the training stage, the error between the predicted and actual response (dust level) is determined. Various regression models are utilized in the training stage. Each model can estimate the dust accumulation on solar panels with a different RMSE value. The best regression prediction model, which provides the most accurate dust estimation with minimum RMSE, is selected. After selecting the best model that predicts the dust accumulation with the least error, the model can be tested on different case studies.

3.3. Regression Models

Different models are used in the training stage to decide the best model for the proposed dust estimation unit. The regression models presented in this work are linear regression, SVM, and Decision Tree (DT).

3.3.1. Linear Regression

The linear regression tries to find the best possible fit between the predictors and the target. The function that represents the estimated dust level as a function of three predictors can be expressed as follows [22]:

$$\hat{y} = \beta x, \quad (1)$$

where \hat{y} is the predicted dust level; $x = [x_1 \ x_2 \ x_3]^T$ is a vector containing the temperature, irradiance, and output power from PV as the three predictors at these conditions; β represents the linear model coefficients to be determined.

The linear model coefficients are determined such that the least square error between the actual dust levels and the predicted dust levels is minimized, as written in the following equation

$$\min(y - \hat{y})^2 = (y - (\beta x))^2, \quad (2)$$

where y is the actual dust level.

3.3.2. SVM Regression

In SVM regression, a proper line or hyperplane to fit all data will be obtained by SVM. The difference between the objective function in case of Linear regression and SVM regression is that the coefficients are determined to minimize the squared coefficients, not the squared error. The predicted dust levels, the objective function, and the error constraint can be written using Equations (1), (3), and (4), respectively [23].

$$\min_{\beta} \left(\frac{1}{2} |\beta \beta^T| \right) \quad (3)$$

subject to:

$$|y - \hat{y}| \leq \varepsilon, \quad (4)$$

A slack variable, ε_n , will be defined in case no function exists to satisfy the previous constraints. Then, the equations can be written as follows:

$$\min_{\beta} \left(\frac{1}{2} |\beta \beta^T| + C \sum_{n=1}^N \varepsilon_n \right) \quad (5)$$

subject to:

$$|y - \hat{y}| \leq \varepsilon + \varepsilon_n, \quad (6)$$

where N is the number of experimental data points and C is a regularization parameter.

3.3.3. DT

It can solve problems related to both the classification and regression. The prediction space can be split into non-overlapping regions, as shown in Figure 3. Various techniques are utilized to construct the DT and calculate the number of regions like classification and regression tree (CART), and Iterative Dichotomiser 3 (ID3). These regions are called leaves, if they are not further subdivided into other regions, otherwise they are called nodes. For each region, the predicted dust level can be considered as the average of all values of dust level points utilized in the training in this region, as illustrated in Figure 3. We are only showing two predictors for the sake of simplicity of visualization, but we applied the DT model using all three predictors. The DT tries to find which predictor among all predictors is the most predictive of the response variables. The DT will use this predictor to start the tree as tree root. However, DT will ask a question related to one predictor at each node to determine the motion's direction. This procedure will continue until reaching the end of the tree [24].

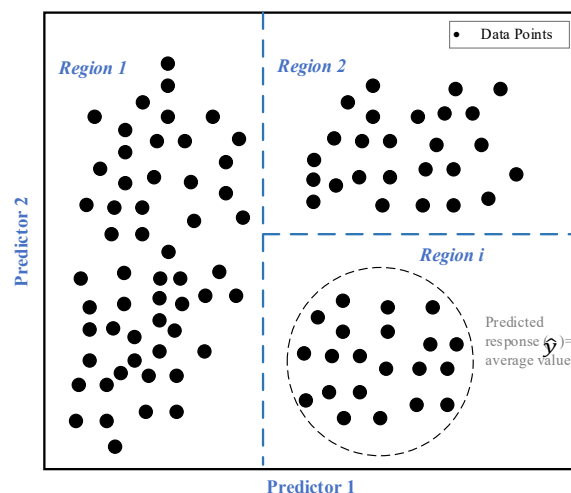


Figure 3. Decision Tree (DT) algorithm.

3.4. Dust Estimation Unit

After selecting the regression model that investigates the least RMSE, this model is extracted. Then, it can be fed by the three predictors, which are the ambient temperature, solar irradiance, and the PV panels' output power, and it would predict the accumulated dust level on the solar panel. Accurate prediction of dust is essential for the system operators to estimate the most optimal time to clean solar photovoltaic modules. A threshold value can be selected for the accumulated dust level on the solar panel, which results in a significant degradation of the energy produced by solar PV systems. Hence, cleaning procedures can be initiated once the dust accumulation level exceeds this threshold value. Utilizing the most optimal time for cleaning solar photovoltaic modules reduces unnecessary costs of continuous cleaning of solar photovoltaic modules. The final stage is testing the dust estimation unit using a data set that the model has not seen before to avoid any bias and accurately evaluate the performance of the model. The experimental data points curtailed from the original experimental data set and not used in the training stage will be used in the testing of the proposed unit.

4. Experimental Setup

A wide range of data is essential to be utilized in training and testing the model, as mentioned before. Experiments were conducted outdoors with real environmental conditions on a solar photovoltaic panel. The effect of dust on the performance of the solar photovoltaic panels has been studied under these environmental conditions. The experimental setup, consisting of a 400 Watt solar photovoltaic panel, was implemented in Dubai (at latitude 25°27'42.52" N and longitude 55°40'44.06" E). The specifications of the solar panel used in the experiment are shown in Table 1. Dust used in the experiment was collected from the same location as the experiment. The location of Dubai city was selected for the experiment as Dubai city has the largest single-site solar park in the world. The output power of the solar PV panel was measured at different times of the day under various environmental conditions of solar irradiance, temperature, and dust levels. The irradiance was measured by using the RS PRO Solar Power Meter ISM400. The best accuracy the meter can provide is $\pm 5 \text{ W/m}^2$, with a resolution of 0.1 W/m^2 . The maximum irradiance that can be detected by this meter is 2000 W/m^2 [25].

Table 1. Solar panel manufacturer specifications.

Specifications	Sun Power Manufacturer
Material	c-Si
Model	E20/435
Panel efficiency	20.1%
Max Power (W)	435
V _{mp} (V)	72.9
I _{mp} (A)	5.97
Open-circuit voltage (V)	85.6
Short-circuit Current (A)	6.43
NOCT (°C)	45
Temperature coefficient of P _{max} (%)	-0.38
Temperature coefficient of V _{oc} (%)	-0.27
Temperature coefficient of I _{sc} (%)	0.05
Panel Dimension (mm)	2067 × 1046 × 46

A high precision electronic scale with a precision of 50/0.001 g and an error range of $\pm 0.003 \text{ g}$ was used to weight the amounts of dust used in the experiments. Then, the dust was evenly distributed on the solar PV panel using a thin brush. A dust amount of 0.1 g/m^2 was evenly distributed on the panel. Then, the ambient temperature, solar irradiance, and the output power from the solar panel were measured. The same experiment was repeated for dust amounts of 0.2 g/m^2 up to 0.9 g/m^2 with incremental steps of 0.1 g/m^2 . Hence, for each instant of time, 10 data points were collected; 1 point with no dust and 9 points with various dust levels. The experiment was conducted at different

times of the day for a period of 2 months (September and October). After conducting the experiment, 4800 data points were collected to be fed to the regression model for training and testing the proposed dust estimation unit. The data set was divided randomly into 4000 data points and 800 data points. 4000 data points of the experimental data will be utilized in the training stage and divided using 5-Fold Cross Validation into 4 data groups for training and 1 data group for testing, as mentioned before; whereas the remaining 800 data points will be used in the testing stage to evaluate the potential of the proposed dust estimation unit. However, the training set has more points than the testing set to ensure that the model is more general and trained against different conditions.

5. Results and Discussions

In this section, the results from different scenarios are discussed to evaluate the performance of the proposed dust estimation unit. Experiments were conducted under real environmental conditions to collect enough data to be utilized in the training and testing of different regression models to predict the accumulated dust levels on solar photovoltaic systems. The collected data are divided into two parts, one for training that contains 4000 data points (to choose and train the appropriate model) and another part for testing that contains 800 data points. A sample of the experimental collected data under real environmental conditions is shown in Table 2. The experiments were implemented in different time periods and on different days to construct a database with various ranges of environmental conditions. The measured irradiance values through experiments were between 77.7 W/m^2 and 650.9 W/m^2 , while the temperature values were between $21.8 \text{ }^\circ\text{C}$ and $45 \text{ }^\circ\text{C}$. The output power of the solar photovoltaic panel was measured and recorded with dust levels between 0 g/m^2 to 0.9 g/m^2 with a step of 0.1 g/m^2 . The measured output power was between 2.6 W and 313 W with different dust levels, different solar irradiance and ambient temperatures. The regression models will be trained to predict the response variable based on input predictors, and hence predict dust accumulation levels based on the experimental measured values of temperature, solar irradiance, and output power from the solar panel. All 4000 collected data points from the field experiment used in the training stage are rearranged based on the dust weight, as illustrated in Figure 4, i.e., the first part of the data points from 0 to 399 consists of different combinations of output PV power, temperature, and solar irradiance, however, all these points are measured with 0 g/m^2 dust level. Various regression models are trained using these data. The first regression model used to estimate the accumulated dust level on the solar photovoltaic panel is the linear regression since the relation between dust accumulation and solar photovoltaic output power in the UAE is linear according to [10,26]. Four linear regression models are developed and utilized for dust estimation. The RMSE of linear regression models is high, as shown in Table 3, as the linear regression model is suitable for estimating dust levels with amounts less than 0.34 g/m^2 under controlled indoor environmental conditions [10]. However, for high amounts of dust accumulation under real environmental conditions, non-linear regression models must be investigated. Table 3 shows samples of the various regression models used in the dust estimation unit training and the corresponding RMSE. Among all used nonlinear regression models, Fine Tree Regression provides the best dust level estimation. The developed fine tree regression model provides an accurate dust prediction with $\text{RMSE} = 0.026737 \text{ g/m}^2$. This error is 53.13% lower compared to the lowest error obtained by the other regression models (Exponential GPR = 0.057048). Figure 5 shows the original data set and the predicted dust level using the fine tree algorithm. Figure 5 illustrates the line response of fine tree regression model. The majority of the predicted dust accumulation points by the fine tree model are located around the horizontal line, which represents a zero or very close to zero error between the actual dust levels and the predicted dust levels except for a few measured points, as shown in Figure 6. Moreover, the maximum error obtained for dust levels less than 0.6 g/m^2 is 0.35 g/m^2 . However, for dust levels more than 0.6 g/m^2 , the maximum error in prediction is 0.2 g/m^2 . Therefore, the model is reliable for estimating high dust accumulation, which is more vital in solar power plants.

Table 2. Sample of the experimental results.

Solar Irradiance (W/m ²)	Ambient Temperature (°C)	Output Power (W)	Dust (g/m ²)
650.9	33	309.28	0
650.9	33	285.43	0.1
650.9	33	182.28	0.4
650.9	33	112	0.6
650.9	33	87.37	0.8
536.9	33	252.39	0
536.9	33	50.36	0.9
158.8	27.36	63.64	0.1
516.6	35	177.11	0.3
382.6	39.42	140.65	0.2
397.1	36.42	58.52	0.6

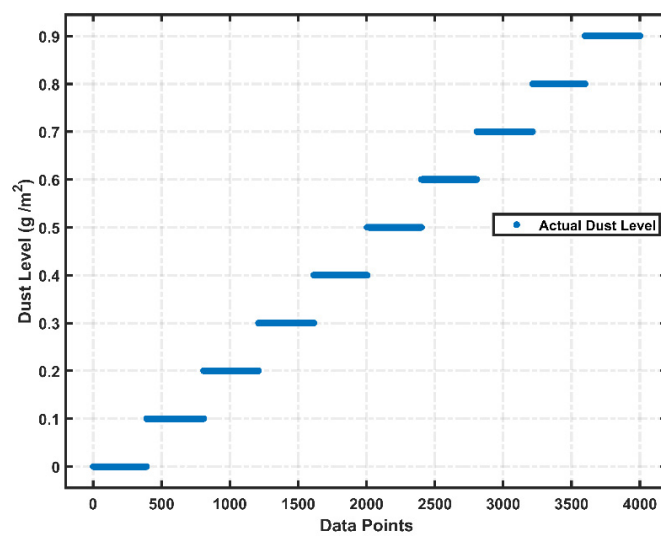


Figure 4. The original data set used for training.

Table 3. Samples of regression models used in training the dust estimation unit.

Regression Model	Model Type	RMSE (g/m ²)
Linear Regression	Linear	0.093204
Linear Regression	Stepwise Linear	0.084241
DT	Fine Tree	0.026737
DT	Coarse Tree	0.074296
SVM	Linear SVM	0.094035
SVM	Medium Gaussian SVM	0.068671
Ensemble	Boosted Trees	0.064575
Gaussian Process Regression	Squared Exponential GPR	0.063416
Gaussian Process Regression	Exponential GPR	0.057048

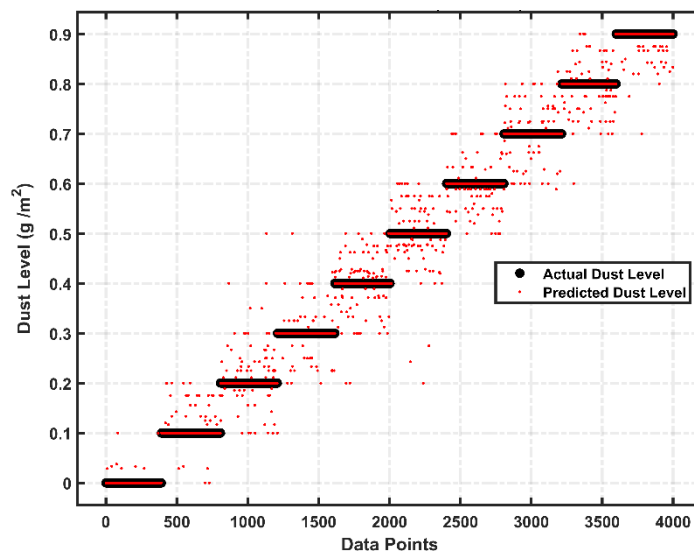


Figure 5. The original and predicted dust levels based on the Fine Tree algorithm.

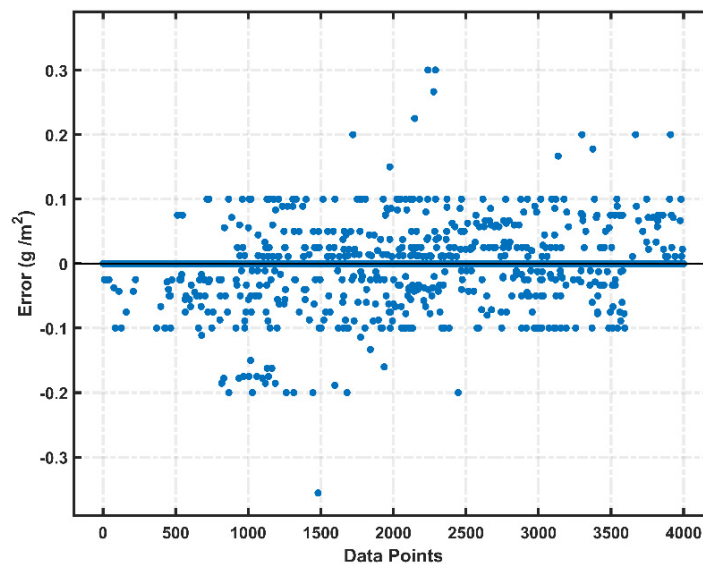


Figure 6. Fine Tree model line response of dust accumulation.

5.1. Case 1: Premeasured Dust Levels

Two case studies are conducted to assess the performance of dust estimation unit based on the fine tree regression model. The first case study is conducted using premeasured dust levels for 800 data points curtailed from the original data set. The amount of dust used in this case study is between 0 g/m^2 and 0.9 g/m^2 with an incremental step of 0.1 g/m^2 . The three predictors, which are the ambient temperature, solar irradiance and PV panels' output power, are fed to the dust estimation unit to predict the accumulated dust level on the solar panel. These values are compared to the original dust levels obtained from the experiment and hence, the error is determined, as shown in Figure 7. Table 4 illustrates samples for the experimental dataset utilized in case 1, in addition to the predicted dust level by the proposed unit. The error between the actual and predicted dust levels for small amounts of accumulated dust on solar panels is higher than the error in case of large amounts of accumulated dust on solar panels. Moreover, the highest error between the measured dust level and the estimated dust level is 0.3 g/m^2 . Hence, the model provides accurate dust estimation, particularly with high dust accumulation, which is vital for the operation of solar plants.

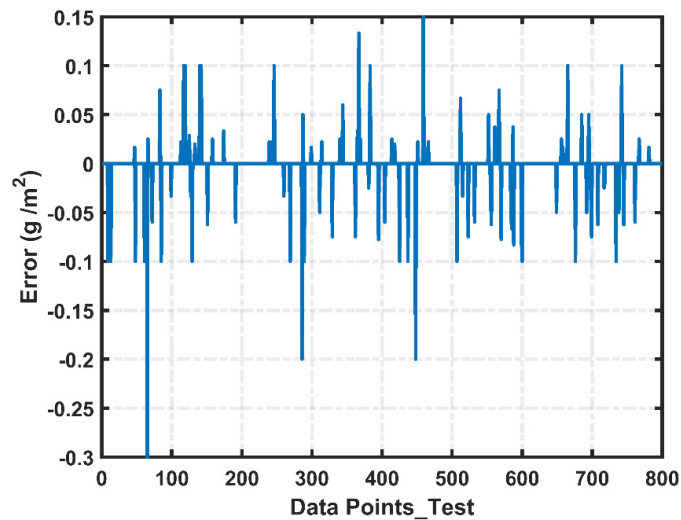


Figure 7. Error between measured and estimated dust values.

Table 4. Sample of experimental results used in case study 1 and the predicted dust level at these conditions.

Solar Irradiance (W/m ²)	Ambient Temperature (°C)	Output Power (W)	Actual Dust (g/m ²)	Predicted Dust [Error] (g/m ²)
564	25.91	81.97	0.8	0.7 [0.1]
536.9	36.49	124.06	0.5	0.517 [−0.017]
457.7	30.54	121.07	0.5	0.2 [0.3]
310	33.71	95.38	0.4	0.425 [−0.025]
650.9	32.99	104.62	0.7	0.7 [0]
520.1	38.56	248.22	0	0 [0]
338	31.82	124.64	0.1	0.175 [−0.075]
158.8	36.93	44.62	0.5	0.4 [0.1]
507.4	36.611	81.33	0.9	0.9 [0]

5.2. Case 2: Random Dust Levels

Another case study is conducted to evaluate the accuracy of the proposed dust estimation unit based on the fine tree regression model. Experiments were performed again to measure the output power of solar panels with different solar irradiance, ambient temperatures, and random dust accumulation amounts. In this case, random dust amounts were spread uniformly on the solar photovoltaic panel. The output power of the solar photovoltaic panel, the solar irradiance, and the ambient temperature were measured at each instant. A total of 35 different data points were collected. The experiments were conducted outdoors under real environmental conditions. A sample of the collected data from experiments is illustrated in Table 5. The collected data are fed into the proposed dust estimation unit, based on fine tree regression model, to estimate the dust levels on the solar photovoltaic panel. The inputs to the proposed unit are the three predictors: ambient temperature, solar irradiance, and output solar photovoltaic power. The proposed model predicts the accumulated dust levels on the solar photovoltaic panel. The predicted and measured dust levels are shown in Figure 8. For high dust levels, the error is between 0.02 g/m² and 0.01 g/m². For low amounts of dust accumulation, the range of error is between of 0.06 g/m² and 0.01 g/m². Estimating high amounts of dust is more important for its impact on solar photovoltaic output power compared to the effect of low dust accumulation. The results reveal the ability of the proposed dust estimation unit to detect the dust level with low error.

Table 5. Sample of experimental results used in case study 2.

Solar Irradiance (W/m ²)	Ambient Temperature (°C)	Output Power (W)	Dust (g/m ²)
606.36	33	225.31	0.14
506.18	31	198.97	0.14
546.1	32	207.03	0.22
546.1	32	179.52	0.27
606.36	33	198.64	0.32
609.8	33	139.15	0.48
609.8	33	95.17	0.69
516.8	33	90.15	0.71
516.8	33	45.35	0.89
506.18	31	39.65	0.92

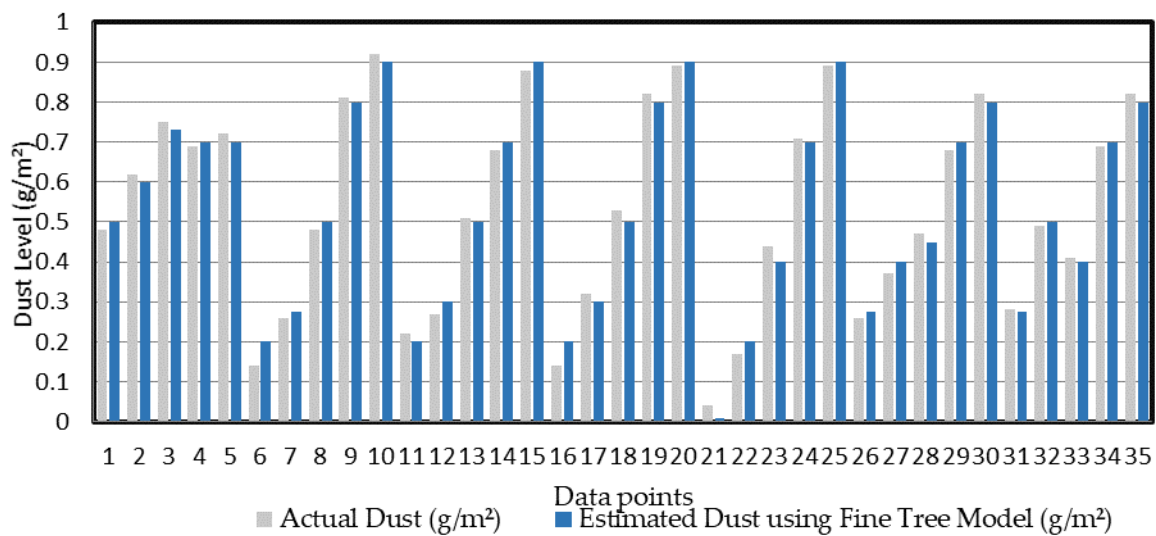


Figure 8. The actual and predicted dust levels for case study 2.

5.3. Performance Evaluation of the Proposed System

To assess the performance of the proposed dust estimation unit, a 0.6 g/m² dust level was selected as the threshold value or decisive value to initiate the cleaning actions. A level of 0.6 g/m² was selected to represent high accumulation level, where the power is degraded by more than 60% at this dust level compared to the zero-dust case at the same environmental conditions. Fifty data points with dust levels less than 0.6 g/m², and 50 data points with dust levels higher than or equal 0.6 g/m² collected from experiments, are fed to the proposed unit to evaluate its performance. The proposed unit will determine the need for cleaning actions according to the predicted dust level compared to the threshold value. The following performance parameters are determined to illustrate the potential of the proposed dust estimation unit.

$$\text{Sensitivity (Detection Rate)} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Negative Predictive Value (NPV)} = \frac{TN}{TN + FN} \quad (10)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$\text{False Alarm (FA)} = 1 - \text{Specificity} \quad (12)$$

where, TP is the number of true positives, which means there is a need for cleaning and the proposed unit decides the starting of cleaning actions; FN is the number of false negatives, which means there is a need for cleaning and the proposed unit decides no cleaning actions; TN is the number of true negatives, which means there is no need for cleaning and the proposed unit decides no cleaning actions; and FP is number of false positives, which means there is no need for cleaning and the proposed unit decides the starting of cleaning actions.

The performance of the proposed dust estimation unit based on fine tree regression model is compared with another detector based on ANN to demonstrate the potential of the proposed unit. We used an ANN as it is a popular prediction model in many approaches. We used a three layer-feed-forward network to predict the relationship between the target variable, dust level, and the three predictors; temperature, irradiance and the output power. There are many variants of the backpropagation algorithm to train neural networks; the Bayesian Regularization algorithm is one such method. Although it takes more time, it is characterized by good generalization for difficult, small or noisy datasets. We chose the other parameters of the network (number of hidden nodes, etc.), using a validation set. The detector based on ANN provides an accurate estimation with RMSE equal to 0.0592. Table 6 illustrates the value of the performance parameters for both detectors. The results reveal the potential of the proposed unit to initiate the cleaning actions only when required, with a low false alarm rate.

Table 6. Performance parameters of the proposed unit.

Performance Parameters	Fine Tree	ANN
Sensitivity	86%	80%
Specificity	99%	96%
Precision	98.85%	95.24%
Negative Predictive Value	87.61%	82.76%
Accuracy	92.5%	88%
False Alarm	1%	4%

6. Conclusions

This paper presents a dust estimation unit to estimate the dust accumulation levels on solar panels and, hence, initiate the cleaning actions. First, experiments were conducted under real environmental conditions in order to investigate the effect of dust accumulation on solar photovoltaic panels and to collect the dataset required to develop the proposed dust estimation unit. The output power of the solar PV panel was measured at different times of the day under various real environmental conditions of solar irradiance, temperature, and dust levels through the conducted experiments. Second, the collected data from the experiments are used in training different regression models. The model with the least RMSE is utilized in developing the dust estimation unit. The regression model based on the fine tree algorithm provides the least RMSE. Therefore, the dust estimation unit is developed based on the fine tree algorithm. A threshold value is selected and once the dust level exceeds this value, which leads to high degradation in the output power from solar panels, the cleaning procedure is initiated. Various case studies are performed to assess the accuracy of the proposed model to predict the amount of dust accumulation on solar panels. A comparison between the performance of the proposed unit and another unit based on ANN is presented. The results reveal that the proposed fine tree dust estimation unit is able to predict the dust level with a low false alarm rate and high performance. Analysis of economic yield derived from the obtained results in this paper will be investigated as a core of a new paper.

Author Contributions: Conceptualization, M.F.S. and U.T.; methodology, A.A., M.F.S. and U.T.; software, A.A. and M.M.; validation A.A., M.M. and M.F.S.; formal analysis, A.A. and M.M.; investigation, A.A., M.F.S., and M.M.; resources, A.A., A.H.O. and M.F.S.; data curation, A.A., M.F.S. and U.T.; writing—original draft preparation, A.A.,

M.M. and M.F.S.; writing—review and editing, A.R.A.-A., A.H.O., M.M., M.F.S. and U.T.; visualization, M.M., and M.F.S.; supervision, M.F.S. and U.T.; project administration, M.F.S.; funding acquisition, A.R.A.-A., A.H.O., M.F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This publication was supported by research grant # EFRG18-GER-CEN-10 from the American University of Sharjah.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ministry of Energy—Electrical Affairs Department. *Use of Renewable Energy in the Gulf Countries*; Environmental center for Arab towns (ECAT), Envirocities eMagazine: Dubai, UAE, 2012; Volume 1, pp. 4–11.
2. Government of UAE. “The United Arab Emirates Government Portal,” UAE Government Renewable Energy. 18 December 2018. Available online: <https://government.ae/en/about-the-uae/leaving-no-one-behind/7affordableandcleanenergy#key-achievements-towards-affordable-and-clean-energy> (accessed on 23 April 2019).
3. Mokri, A.; AlAli, M.; Emziane, M. Solar energy in the United Arab Emirates: A review. *Renew. Sustain. Energy Rev.* **2013**, *28*, 340–375. [[CrossRef](#)]
4. Mani, M.; Pillai, R. Impact of dust on solar photovoltaic (PV) performance: Research status, challenges and recommendations. *Renew. Sustain. Energy Rev.* **2010**, *14*, 3124–3131. [[CrossRef](#)]
5. Maghami, M.; Hizam, H.; Gomes, C.; Radzi, M.; Rezadad, M.; Hajighorbani, S. Power loss due to soiling on solar panel: A review. *Renew. Sustain. Energy Rev.* **2016**, *59*, 1307–1316. [[CrossRef](#)]
6. Mekhilef, S.; Saidur, R.; Kamalisarvestani, M. Effect of dust, humidity and air velocity on efficiency of photovoltaic cells. *Renew. Sustain. Energy Rev.* **2012**, *16*, 2920–2925. [[CrossRef](#)]
7. Ahmed, Z.; Kazem, H.; Sopian, K. Effect of Dust on Photovoltaic Performance: Review and Research Status. *Latest Trends Renew. Energy Environ. Inform.* **2013**, 193–199.
8. Kaldellis, J.; Fragos, P.; Kapsali, M. Systematic experimental study of the pollution deposition impact on the energy yield of photovoltaic installations. *Renew. Energy* **2011**, *36*, 2717–2724. [[CrossRef](#)]
9. Kazem, H.; Al-Bahri, S.; Al-Badi, S.; Al-Mahkladi, H.; Al-Waeli, A. Dust Effect on the Performance of Photovoltaic. In *Advanced Materials Research*; Trans Tech Publications: Stafa-Zurich, Switzerland, 2014; Volumes 875–877, pp. 1908–1911.
10. Hachiche, A.A.; Al-Sawafta, I.; Said, Z. Impact of Dust on the Performance of Solar Photovoltaic (PV) Systems under United Arab Emirates Weather Conditions. *Renew. Energy* **2019**, *141*, 287–297. [[CrossRef](#)]
11. Jang, H.S.; Bae, K.Y.; Park, H.; Sung, D.K. Solar Power Prediction Based on Satellite Images and Support Vector Machine. *IEEE Trans. Sustain. Energy* **2016**, *7*, 1255–1263. [[CrossRef](#)]
12. Bacher, P.; Madsen, H.; Nielsen, H. Online short-term solar power forecasting. *Sol. Energy* **2009**, *83*, 1722–1783. [[CrossRef](#)]
13. Gutierrez-Corea, F.; Manso-Callejo, M.; Moreno-Re, M. Forecasting short-term solar irradiance based on artificial neural networks and data from neighboring meteorological stations. *Sol. Energy* **2016**, *134*, 119–131. [[CrossRef](#)]
14. Ruiz-Arias, J.A.; Alsamamra, H.; Tovar-Pescador, J. Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions. *Energy Convers. Manag.* **2010**, *51*, 881–893. [[CrossRef](#)]
15. İzgi, E.; Öztöpal, A.; Yerli, B.; Kaymak, M. Short–mid-term solar power prediction by using artificial neural networks. *Sol. Energy* **2012**, *86*, 725–733. [[CrossRef](#)]
16. Coimbra, C.; Pedro, H. Assessment of forecasting techniques for solar power production with no exogenous inputs. *Sol. Energy* **2012**, *86*, 2017–2028.
17. Zorrilla-Casanova, J.; Piliouguine, M. Analysis of dust losses in photovoltaic modules. In Proceedings of the World Renewable Energy Congress, Linköping, Sweden, 8–13 May 2011.
18. Dahlioui, D.; Laarabi, B.; Sebbar, M.A.; Barhdadi, A.; Dambrine, G.; Menard, E.; Boardman, J. Soiling effect on photovoltaic modules performance: New experimental results. In Proceedings of the 2016 International Renewable and Sustainable Energy Conference (IRSEC), Marrakech, Morocco, 14–17 November 2016; pp. 111–114.
19. Rachid, K.; Hamid, E. Solar Cells Performance Reduction under the Effect of Dust in Jazan Region. *J. Fundam. Renew. Energy Appl.* **2017**, *7*, 1–4.

20. Tribak, H.; Zaz, Y. Dust Soiling Concentration Measurement on Solar Panels based on Image Entropy. In Proceedings of the 2019 7th International Renewable and Sustainable Energy Conference (IRSEC)—IEEE, Agadir, Morocco, 27–30 November 2019; pp. 1–4. [[CrossRef](#)]
21. Bishop, C. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006; pp. 137–173.
22. Zhang, J.; Chung, C.Y.; Han, Y. Online Damping Ratio Prediction Using Locally Weighted Linear Regression. *IEEE Trans. Power Syst.* **2016**, *31*, 1954–1962. [[CrossRef](#)]
23. Ertekin, S.; Bottou, L.; Giles, C.L. Nonconvex Online Support Vector Machines. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 368–381. [[CrossRef](#)] [[PubMed](#)]
24. Zheng, C.; Malbasa, V.; Kezunovic, M. Regression tree for stability margin prediction using synchrophasor measurements. *IEEE Trans. Power Syst.* **2013**, *28*, 1978–1987. [[CrossRef](#)]
25. RS Components Ltd. RS PRO Solar Power Meter ISM400. 2019. Available online: <https://uk.rs-online.com/web/p/solar-power-meter/1232218/> (accessed on 20 August 2019).
26. ElMouatasim, A.; Darmane, Y. Regression Analysis of a Photovoltaic (PV) System in FPO. In Proceedings of the AIP Conference Proceedings 2056, 020008, Ouarzazate, Morocco, 8–10 May 2018.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).