

Article

# An ANFIS-Based Modeling Comparison Study for Photovoltaic Power at Different Geographical Places in Mexico

Nun Pitalúa-Díaz <sup>1,\*</sup>, Fernando Arellano-Valmaña <sup>2</sup>, Jose A. Ruz-Hernandez <sup>2</sup>,  
Yasuhiro Matsumoto <sup>3</sup>, Hussain Alazki <sup>2</sup>, Enrique J. Herrera-López <sup>4</sup>,  
Jesús Fernando Hinojosa-Palafox <sup>1</sup>, A. García-Juárez <sup>1</sup>, Ricardo Arturo Pérez-Enciso <sup>1</sup>  
and Enrique Fernando Velázquez-Contreras <sup>1</sup>

- <sup>1</sup> Departamento de Ingeniería Industrial, Departamento de Ingeniería Química y Metalurgia, Departamento de Investigación en Física, Departamento de Investigación en Polímeros y Materiales, Universidad de Sonora, Blvd. Luis Encinas y Rosales S/N, Col. Centro. Hermosillo 83000, Sonora C.P., Mexico
  - <sup>2</sup> Facultad de Ingeniería, Universidad Autónoma del Carmen, Calle 56 No. 4 Esq. Avenida Concordia Col. Benito Juárez C.P. 24180 Cd. Del Carmen, Campeche, Mexico
  - <sup>3</sup> Departamento de Ingeniería Eléctrica, Centro de Investigación y de Estudios Avanzados del IPN, Av. Instituto Politécnico Nacional 2508, La Laguna Ticoman, C.P. 07360 Ciudad de México, CDMX, Mexico
  - <sup>4</sup> Biotecnología Industrial, Sublínea Bioelectrónica, Centro de Investigación y Asistencia en Tecnología y Diseño del Estado de Jalisco A.C., Camino Arenero 1227, Col. El Bajío del Arenal, C.P. 45019 Zapopan Jalisco, Mexico
- \* Correspondence: nun.pitalua@unison.mx; Tel.: +52-1-662-259-2159

Received: 5 June 2019; Accepted: 9 July 2019; Published: 11 July 2019



**Abstract:** In this manuscript, distinct approaches were used in order to obtain the best electrical power estimation from photovoltaic systems located at different selected places in Mexico. Multiple Linear Regression (MLR) and Gradient Descent Optimization (GDO) were applied as statistical methods and they were compared against an Adaptive Neuro-Fuzzy Inference System (ANFIS) as an intelligent technique. The data gathered involved solar radiation, outside temperature, wind speed, daylight hour and photovoltaic power; collected from on-site real-time measurements at Mexico City and Hermosillo City, Sonora State. According to our results, all three methods achieved satisfactory performances, since low values were obtained for the convergence error. The GDO improved the MLR results, minimizing the overall error percentage value from 7.2% to 6.9% for Sonora and from 2.0% to 1.9% for Mexico City; nonetheless, ANFIS overcomes both statistical methods, achieving a 5.8% error percentage value for Sonora and 1.6% for Mexico City. The results demonstrated an improvement by applying intelligent systems against statistical techniques achieving a lesser mean average error.

**Keywords:** ANFIS; statistical method; gradient descent; photovoltaic system; sustainable development

## 1. Introduction

Considerable research has been developed internationally in the field of photovoltaic systems and power generation [1–3]. Mexico is a country that receives abundant solar energy, with the northwest region being the one with the highest annual incidence of solar radiation, achieving radiation indexes between 5.6 and 6.2 kWh/m<sup>2</sup> per day; nevertheless, its advances on the photovoltaic field are scarce nowadays, although there are many possibilities of research in this topic [4].

There are many scientific reports on statistical methods to estimate power generation [5–8]. Multiple Linear Regression (MLR) includes any study area, since it finds an approach to the relation among variables [9].

A satisfactory photovoltaic power estimation involving meteorological variables was carried out in [10] in which a Gradient Descent Optimization (GDO) minimizes the error value between the real and estimated variables after many iterations; even so, it is mentioned that other techniques may obtain better results despite their lower or greater complexity.

On the other hand, intelligent techniques have acquired a worldwide reputation as simple methods to represent and replicate the behavior of a process with a not-so-understood performance. These techniques have potential to model, precisely, linear and non-linear processes, the latter being their strongest application. Some studies have used intelligent techniques to analyze photovoltaic power behavior, e.g., in [11] an artificial neural network (ANN) was used to obtain a model to forecast the photovoltaic energy; however, solar radiation was the only meteorological variable analyzed. In [12], different ANN techniques were used to perform a comparative study of systems predicting photovoltaic thermal energy data; nevertheless, due to solar radiation measurement devices being expensive and requiring periodic maintenance, some results employed global solar radiation (GSR) estimations through ANNs. In [13], it was established that a photovoltaic system is affected by many meteorological conditions; a predictive model was proposed considering outside temperature, solar radiation and direction and speed of wind. In [14], an ANN-Adaptive Neuro-Fuzzy Inference System (ANFIS)-based forecast model for predicting the photovoltaic generation and wind energy generation is presented and considers the susceptibilities to which renewable energies are exposed due to nature's vagaries. In [15], it is proved that an adaptive neuro-fuzzy inference system technique provides a reliable tool to estimate temperature from photovoltaic systems. In [16], the solar still productivity prediction aims to be improved by using an ANFIS due to its simple maintenance and ready affordability.

The importance of achieving a satisfactory model design with a minimum error between estimated and real values is crucial in precision studies or management tasks in which certain differences among them may result in economic problems or loss of information.

As can be seen, diverse statistical and intelligent methods have been applied to estimate photovoltaic power generation. However, this topic still is an open field that requires better estimation methods, in particular, those related to intelligent systems.

Since it is relevant to develop new estimation strategies in photovoltaic systems, the aim of this work is to compare ANFIS methodology with MLR and GDO as statistical approaches by comparing electrical power estimation data from photovoltaic systems [10]. It was observed that all three methods achieve a satisfactory estimation performance, but ANFIS had a better estimation capacity.

## 2. Methodology

### 2.1. Statistical Methods

#### 2.1.1. Multiple Linear Regression

To create a statistical representative model of the power generated by solar energy, the concept of multiple linear regression was applied. The relation that exists among all the input variables and the output is represented by a "linear regression". Depending on the number of input variables, the regression can be simple or multiple [6,17–19]. The purpose of multiple linear regression is to find an estimation of the real output through an equation involving all the gathered data, as shown in Equation (1).

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k + \varepsilon \quad (1)$$

where  $y$  is the estimated output,  $x_k$  is the  $k$ th input variable,  $\beta_k$  is the characteristic coefficient of every variable and  $\varepsilon$  is the error between the model and the real data.

To obtain Equation (1), a matrix “ $X$ ” and a vector “ $y$ ” are generated. The number of columns for “ $X$ ” is  $k + 1$ , with the condition that the first column is filled with ones. The dimension of the columns and the vector “ $y$ ” depends on the total quantity of data “ $n$ ”. Consequently, Equation (2) is applied to find the respective coefficient values.

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & x_{31} & x_{32} & \cdots & x_{3k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = (X'X)^{-1} (X'y) \quad (2)$$

The data collected from Hermosillo, Sonora (located in the northwest region of Mexico) were obtained with the support of the University of Sonora (UNISON), while the data from Mexico City were issued by the Centro de Investigación y de Estudios Avanzados (CINVESTAV) campus Zacatenco. Solar radiation, outside temperature, wind speed and daylight hour (time) served as the input meteorological variables, while the electric power was the output for both of the PV systems. Each data sample was registered every 5 min. According to the above, the matrix “ $X$ ” and the vector “ $y$ ” from Equation (2) are shown in Equation (3).

$$X = \begin{bmatrix} 1 & \text{Solar Radiation} & \text{Temperature} & \text{Wind Speed} & \text{Time} \end{bmatrix} \quad (3)$$

$$y = [\text{Electric Power}]$$

### 2.1.2. Gradient Descent Optimization

The GDO has been used to estimate different behaviors in several studies [20,21]. As seen in [10], this method seeks the optimum  $\beta_k$  coefficients now symbolized by  $\theta$ . To achieve the minimum possible error value in Equation (4), the cost function is used.

$$J[\theta] = \frac{1}{2m} \sum_{i=1}^m \left( h_{\theta} \left( x^{(i)} \right) - y^{(i)} \right)^2 \quad (4)$$

where  $\theta = \beta$  is as mentioned in Equation (1),  $x^{(i)}$  represents the  $i$ th row of matrix “ $X$ ” and  $y^{(i)}$  is the value of the  $i$ th row of vector “ $y$ ”, both described in Equation (2). It is important to notice that, in this case, the number of columns in matrix “ $X$ ” is equal to the number of variables involved, omitting the column full of ones. The gradient descent, denoted by Equation (5), aims to converge to the cost function minimum by its partial derivative. The quickness of the convergence is given by  $\alpha$ .

$$\theta_k := \theta_k - \alpha \frac{\partial}{\partial \theta_k} J[\theta] \quad (5)$$

Equation (6) represents the substitution of Equation (4) into Equation (5) and it is known as the gradient descent implementation on linear regression method, which has to be repeated  $n$ -times until the convergence is done.

$$\theta_k := \theta_k - \alpha \frac{1}{m} \sum_{i=1}^m \left( h_{\theta} \left( x^{(i)} \right) - y^{(i)} \right) \cdot x_k^{(i)} \quad (6)$$

Once all the coefficients have been computed, they are substituted into Equation (7)

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k + \varepsilon \quad (7)$$

where  $h_{\theta}(x)$  is the estimated output,  $x_k$  is the  $k$ th input variable,  $\theta_k$  is the characteristic coefficient for every variable and  $\varepsilon$  is the error between the model and the real data.

## 2.2. Intelligent Technique

An intelligent technique consists of a dynamic learning process in order to generate an output as close as possible to the required one. Among the best intelligent systems are the artificial neural network, fuzzy systems and the adaptive neuro-fuzzy inference system [22,23]. For this contribution, an ANFIS is considered combining the strengths of a neural network and fuzzy systems, obtaining a better performance.

An artificial neural network (ANN) imitates the processing of a human brain and it has a great number of processing units (neurons or nodes) working in parallel. These nodes can be circular or squared in order to represent different adaptive capabilities. A circular node employs fixed operations that cannot be altered at any time (sum or product of inputs), unlike a squared node, which can be modified by the user (activation functions).

All the neurons are highly connected among each other through links (synapse) with weights. The network has a layer for all the inputs, a layer for one or more outputs and one or more hidden layers between the input and the output.

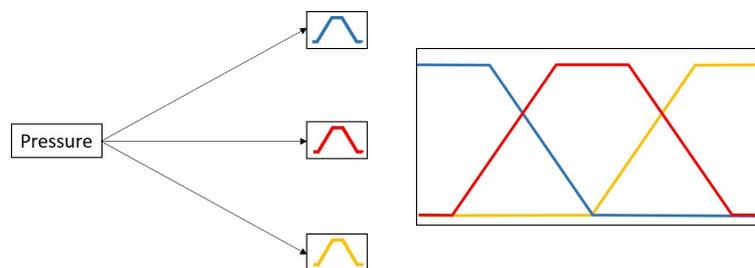
A neural network requires previous assumptions so it can learn from examples by adjusting the connection weights. The learning may be supervised if the right output is specified (being the case for this work) or non-supervised if it has to explore the relations between the patterns and learn to categorize the inputs [11,24].

A fuzzy system is able to apply conditional sentences as a human brain would. The objective of every system that uses fuzzy logic is to describe the degrees of the output sentences (given by a series of rules) according to the input ones. The “if-then” fuzzy rules are sentences with the form “if an event A happens then an event B will occur”, where both events are known as the labels of fuzzy sets from their corresponding membership functions (MF). The strength of this system is to be able to capture all the imprecise modes of the reasoning, for example:

*If the pressure is high, then the volume is small*

where, analogous to the neural network, each and every variable has a specific type of membership function according to its behavior, which can be triangular, trapezoidal, gaussian, etc., as well as to its quantity [25].

Figure 1 shows three trapezoidal membership functions for the variable called “pressure”. Each of them represents a range of values labeled as low (blue), medium (red) and high (yellow). This advantage allows to embrace all the possible data and to classify them as necessary.

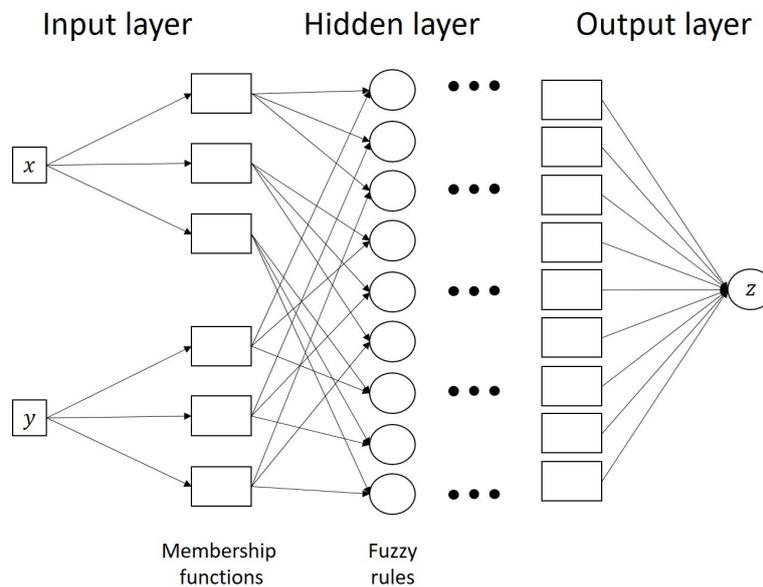


**Figure 1.** Example of three trapezoidal membership functions.

ANFIS is the union of both methods mentioned above, the neural network and fuzzy systems, in which the advantages of both cooperate in order to achieve easier and faster estimations [26].

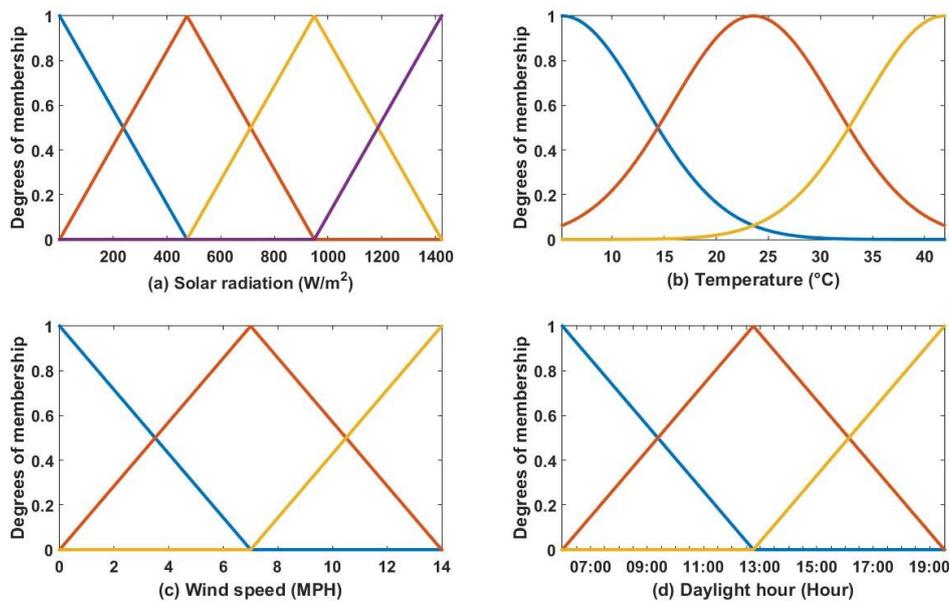
The ANFIS consists of if-then rules and input-output pairs. Moreover, for the training, the learning algorithms of neural networks are used. In order to simplify the explanations, a basic fuzzy inference system consists of two inputs ( $x$  and  $y$ ) and one output ( $z$ ) [27]. Figure 2 shows that, if each input consists of three membership functions (represented as squared nodes, because they can be modified by the user), then the input layer would have six neurons (three for each input) and,

in turn, the output layer would just have one neuron. As for the hidden layers, the first one of them will be formed by each of the if–then fuzzy rules mentioned before; by this, an ANFIS is created.



**Figure 2.** Adaptive Neuro-Fuzzy Inference System (ANFIS) structure with two inputs and three membership functions for each one.

The variables used to train the ANFIS are radiation, outside temperature, wind speed, daylight hour and electrical power, as mentioned in Section 2.1.1. Figure 3 presents the MF of input variables for the intelligent system. The behavior of each variable helps to identify the type of MF to be declared. Solar radiation, wind speed and daylight hour, due to their rapid change with respect to time, have 4, 3 and 3 triangular MF, respectively. As for temperature, which presents a slow variation with time, it has 3 gaussian MF.



**Figure 3.** Membership functions for ANFIS: (a) Triangular for solar radiation; (b) gaussian for temperature; (c) triangular for wind speed; (d) triangular for daylight hour.

According to the above, the structure of the ANFIS is presented in Figure 4. The method applied to train the ANFIS is called hybrid training. This process consists of two steps; the first one computes the result of the next linked node according to the nodes behind, i.e., based on Figure 2 each circular node depends on two rectangular nodes (MF of variables  $x$  and  $y$ ). Once all the nodes have been calculated and the output is obtained, the second step must identify the error in each node and minimize it during the training iterations in order to improve the estimation result. The first and second steps are based on least squares and gradient descent, respectively.

This ANFIS model was applied for both locations by using the MATLAB<sup>®</sup> software in order to prove its effectiveness against the statistical method, regardless of the place where the system is installed.

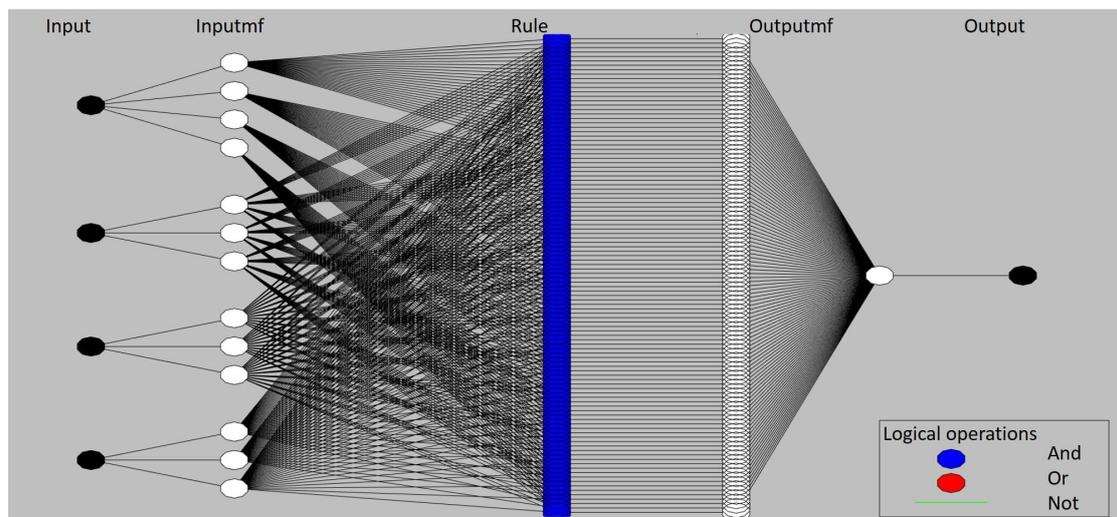


Figure 4. ANFIS structure.

### 3. Results

Seven different cases or sets of data were computed to analyze which estimation technique executes a better performance for the Hermosillo site (HS) and the Mexico City site (MCS). Every case was represented by data collected during a whole month; consequently, the first case corresponds to six months of collected data, while the seventh case stands for the accumulation of all data. According to Section 2.1.1, the time step of each estimation, regardless of the method, was five minutes. The former was considered to generate a greater amount of data allowing a better estimation result by the ANFIS. Figures 5 and 6 present the physical systems for each location, HS and MCS, respectively.



Figure 5. Physical system for the Hermosillo site (HS).



Figure 6. Physical system for the Mexico City site (MCS).

3.1. Hermosillo Site

Table 1 shows the resulting equations for the monthly and overall electrical power estimation by MLR and GDO with 1000 repetitions to minimize Equation (4), considering the data gathered from HS as mentioned in Sections 1 and 2, and by using the MATLAB<sup>®</sup> software. Figures 7 and 8 present the resulting estimation for each method considering a random period to achieve a reliable comparison and a better appreciation of the behavior (24th November to 24th December). The computational load for the overall case using MLR and GDO resulted in total durations of approximately 30 s and 10 min, respectively.

Table 1. Monthly and overall Multiple Linear Regression (MLR) and Gradient Descent Optimization (GDO) equations for HS.

Method	Month	Equations
MLR	Aug	$y = -2305.4 + 1.4182x_1 + 116.4097x_2 + 8.686x_3 - 2078.9x_4$
	Sep	$y = -498.9858 + 1.8082x_1 + 36.7926x_2 + 32.4585x_3 - 1042.6x_4$
	Oct	$y = -190.6163 + 1.6737x_1 + 38.1824x_2 + 9.8075x_3 - 1075.2x_4$
	Nov	$y = -120.4414 + 2.1624x_1 + 13.5544x_2 + 18.5136x_3 - 379.8749x_4$
	Dec	$y = 26.2042 + 1.9806x_1 + 12.1240x_2 + 12.0973x_3 - 405.2137x_4$
	Jan	$y = 142.9390 + 2.1120x_1 + 18.8379x_2 + 20.2070x_3 - 792.1223x_4$
	Total	$y = 84.9221 + 2.0035x_1 + 9.7233x_2 + 15.5707x_3 - 386.3517x_4$
GDO	Aug	$y = 0.0067 + 2.2334x_1 + 0.2415x_2 + 0.0186x_3 + 0.0036x_4$
	Sep	$y = 0.0050 + 2.2366x_1 + 0.1711x_2 + 0.0182x_3 + 0.0025x_4$
	Oct	$y = 0.0064 + 2.2069x_1 + 0.1931x_2 + 0.0251x_3 + 0.0032x_4$
	Nov	$y = 0.0029 + 2.2732x_1 + 0.0764x_2 + 0.0108x_3 + 0.0014x_4$
	Dec	$y = 0.0036 + 2.1411x_1 + 0.0814x_2 + 0.0114x_3 + 0.0018x_4$
	Jan	$y = 0.0044 + 2.3776x_1 + 0.0932x_2 + 0.0138x_3 + 0.0020x_4$
	Total	$y = 0.0044 + 2.2387x_1 + 0.1156x_2 + 0.0154x_3 + 0.0022x_4$

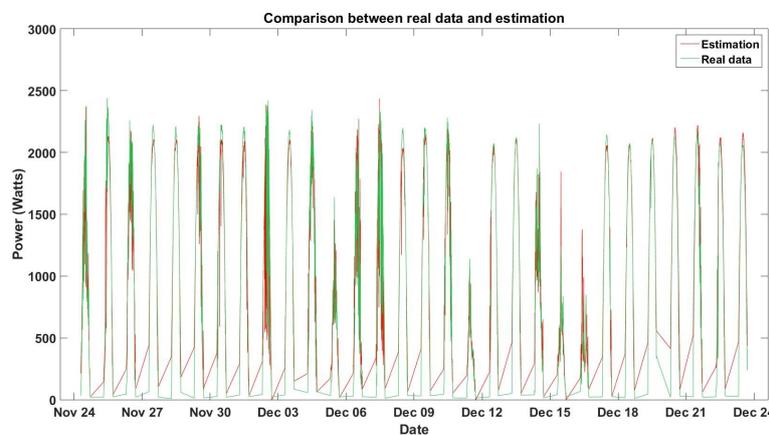


Figure 7. MLR results from HS.



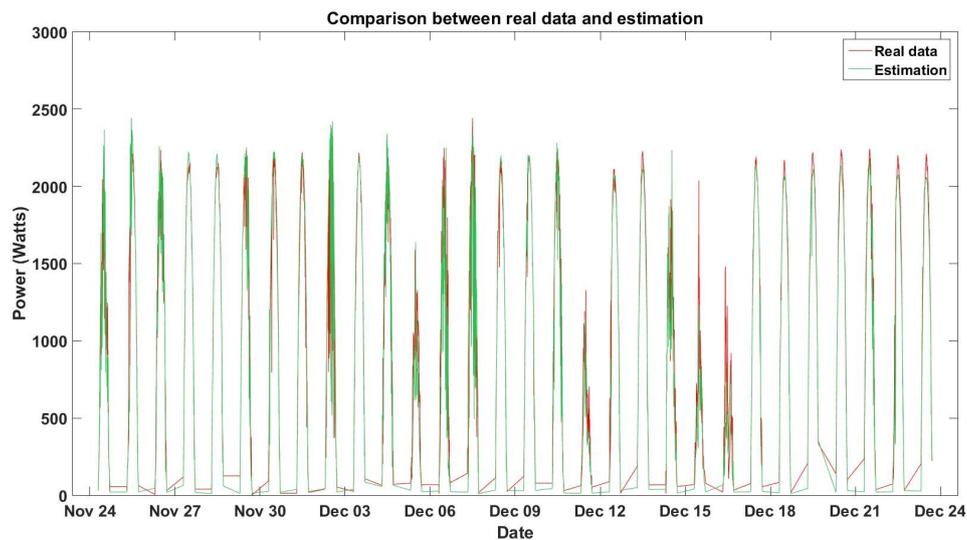


Figure 9. ANFIS results from HS.

Each case analyzed contemplates a different range of time (different month) and its respective ANFIS gets trained according to the data gathered within that period. Given that meteorological inputs vary in time, so do the ANFIS coefficients for every case; nevertheless, even if they are changing, the variation range is not too wide, as can be seen from solar radiation, outside temperature and daylight hour. As for wind speed, its variation is due to a smaller relationship with the output than with the other inputs, as well as to the different behaviors in each month.

### 3.2. Mexico City Site

Analogous to the HS case, the same procedure was applied to the data registered from the Mexico City site (MCS) mentioned in Sections 1 and 2, achieving approximately the same time duration to find the beta and theta parameters. For the MLR and GDO methods, Table 3 presents the monthly and overall estimation equations. Figures 10 and 11 depict the resulting MLR and GDO estimation in the period 24th November to 24th December (same as HS) for a better appreciation of the convergence of the method.

Table 3. Monthly and overall MLR and GDO equations for MCS.

Method	Month	Equations
MLR	Jul	$y = 727.5535 + 52.7725x_1 + 8.6209x_2 + 28.9473x_3 - 947.31789x_4$
	Ago	$y = 2527.4 + 55.7540x_1 - 110.5462x_2 + 19.4716x_3 + 849.8508x_4$
	Sep	$y = 1093.9 + 58.4495x_1 - 27.2714x_2 + 28.7514x_3 - 82.2214x_4$
	Oct	$y = -99.2276 + 57.0806x_1 - 20.0065x_2 + 52.4617x_3 + 1326.5x_4$
	Nov	$y = -1059.8 + 57.7075x_1 - 32.4315x_2 + 12.1205x_3 + 3016.3x_4$
	Dec	$y = -1430.5 + 53.7683x_1 - 15.8521x_2 + 109.1599x_3 + 2953.6x_4$
	Total	$y = 70.2756 + 55.1011x_1 + 33.8665x_2 + 90.3112x_3 - 689.5238x_4$
GDO	Jul	$y = 0.0927 + 53.4273x_1 + 2.2298x_2 + 0.1005x_3 + 0.0491x_4$
	Ago	$y = 0.1083 + 56.0178x_1 + 2.5973x_2 + 0.1441x_3 + 0.0580x_4$
	Sep	$y = 0.1159 + 59.0818x_1 + 2.7233x_2 + 0.1431x_3 + 0.0625x_4$
	Oct	$y = 0.1035 + 57.3927x_1 + 2.2950x_2 + 0.1899x_3 + 0.0560x_4$
	Nov	$y = 0.0908 + 57.2318x_1 + 1.9788x_2 + 0.1036x_3 + 0.0463x_4$
	Dec	$y = 0.0913 + 53.3033x_1 + 1.8920x_2 + 0.0441x_3 + 0.0477x_4$
	Total	$y = 0.1005 + 55.9815x_1 + 2.2910x_2 + 0.1290x_3 + 0.0533x_4$

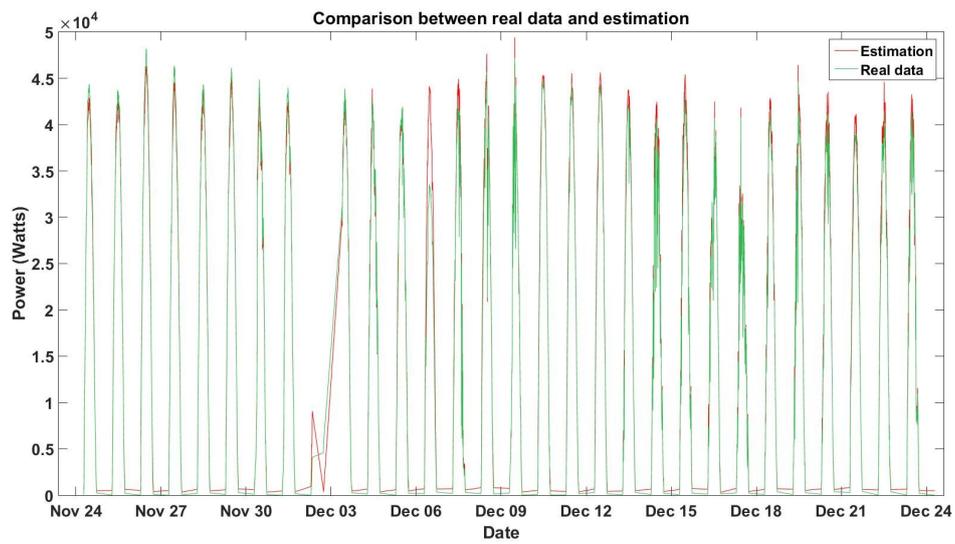


Figure 10. MLR results from MCS.

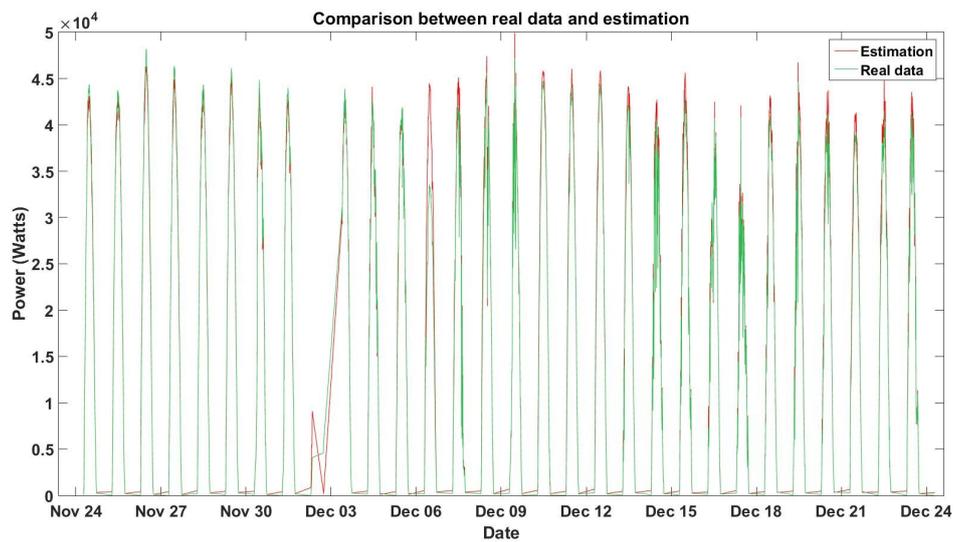
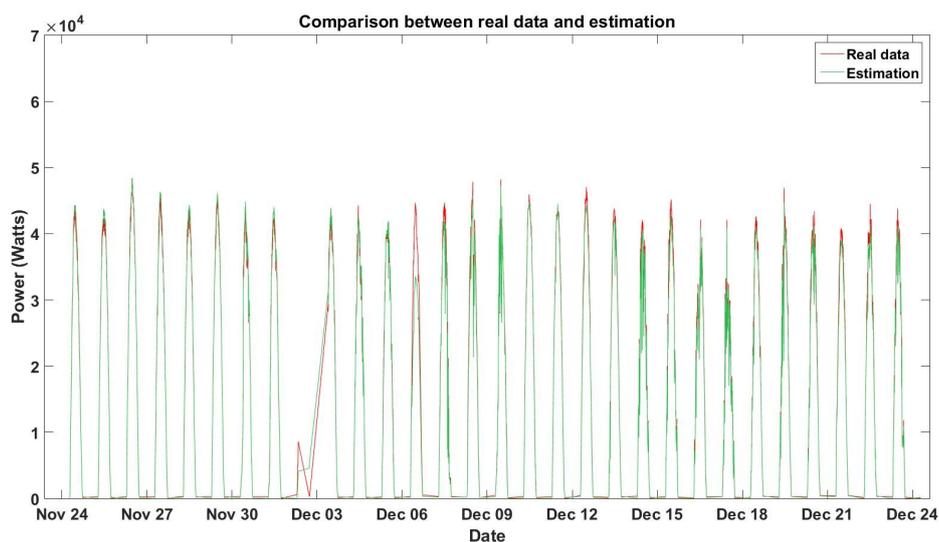


Figure 11. GDO results from MCS.

The same ANFIS structure with 1000 epoch hybrid training for the HS case was employed to obtain the estimation model of the photovoltaic system with the data collected from MCS. Table 4 shows the monthly and overall trained MF parameters. Figure 12 displays the comparison between the resulting estimation against real data with the same time range as for MLR and GDO. The time duration to train the neuro-fuzzy system was approximately the same as in the HS case.

**Table 4.** Monthly and overall trained MF parameters for MCS.

Month	Trained MF Parameters										
	Solar Radiation			Outside Temperature		Wind Speed			Daylight Hour		
	a	b	c	$\sigma$	c	a	b	c	a	b	c
Jul	−356	1	358	4.274	13.02	−8.5	−0.005509	8.509	0.01522	0.3264	0.5521
	1	358	715	4.238	23	0.1882	8.494	17	0.2917	0.5493	0.8816
	355	712	1070	4.227	33.01	8.492	17	25.5	0.5612	0.7535	1.102
	715	1072	1429								
Ago	−371.7	1	373.7	4.481	14.01	−8	−0.002784	7.967	0.02276	0.4896	0.4934
	1	373.7	746.3	4.452	24.5	0.1432	7.997	16	0.343	0.5799	0.7157
	373.7	746.3	1119	4.45	34.9	7.984	16	24	0.5584	0.8143	1.099
	746.3	1119	1492								
Sep	−377.7	1	379.7	5.096	10	−10.5	−0.0001171	10.5	0.04167	0.3015	0.5584
	1	379.7	758.3	5.096	22	0.01522	10.5	21	0.2981	0.5556	0.8166
	379.7	758.3	1137	5.096	34	10.5	21	31.5	0.566	0.8042	1.069
	758.3	1137	1516								
Oct	−353.7	1	355.7	4.463	9.003	−7.5	−0.001309	7.497	0.01042	0.2663	0.4574
	1	355.7	710.3	4.453	19.5	0.0376	7.499	15	0.3247	0.4896	0.8325
	355.7	710.3	1065	4.455	30	7.504	15	22.5	0.4612	0.7494	1.052
	710.3	1065	1420								
Nov	−314.7	1	316.7	5.124	6.014	−6	0.007692	5.994	0.03837	0.3368	0.5183
	1	316.7	632.3	5.067	17.99	−0.01071	6.005	12	0.307	0.4896	0.736
	316.7	632.3	948	5.065	30.02	5.907	12	18	0.4931	0.6581	0.9685
	632.3	948	1264								
Dec	−295.7	1	297.7	5.313	3.003	−5.5	−0.0008596	5.494	0.04514	0.2633	0.5084
	1	297.7	594.3	5.31	15.5	−2.43E-10	5.499	11	0.3472	0.4868	0.7411
	297.7	594.3	891	5.314	28	5.505	11	16.5	0.4805	0.7325	0.9757
	594.3	891	1188								
Total	−377.7	1	379.7	6.813	3.01	−10.5	−0.00304	10.5	−0.008574	0.3395	0.498
	0.9999	379.7	758.3	6.791	19	0.04036	10.5	21	0.2713	0.5833	0.8143
	379.7	758.3	1137	6.783	35.01	10.5	21	31.5	0.5576	0.8047	1.115
	758.3	1137	1516								



**Figure 12.** ANFIS results from MCS.

### 3.3. Error Analysis

In order to demonstrate the impact of the intelligent technique over statistical approaches, an error analysis was applied for every result obtained in Sections 3.1 and 3.2 for HS and MCS. Considering [10],

Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were used to achieve a solid comparison between MLR, GDO and ANFIS. MAE is one of the most implemented errors in estimation studies to measure and analyze precision and is described by Equation (8), and MAPE computes the percentage error value as in Equation (9).

$$MAE = \frac{\sum_{s=1}^N |P_m - P_e|}{N} \quad (8)$$

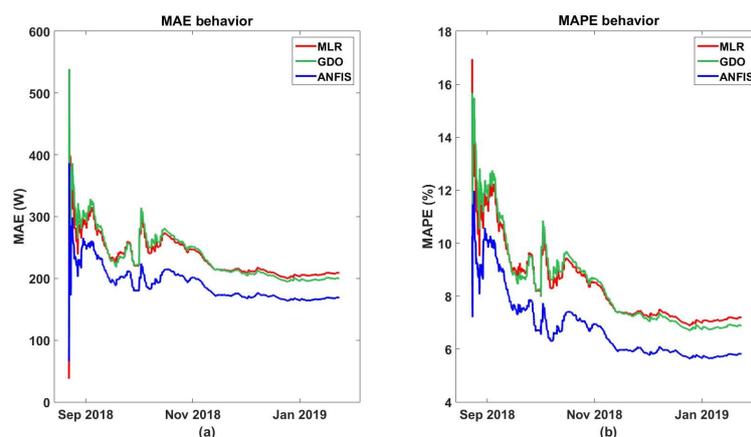
$$MAPE_{(range)\%} = \frac{\sum_{s=1}^N \left| \frac{P_m - P_e}{\max(P_m) - \min(P_m)} \right|}{N} \cdot 100 \quad (9)$$

where “ $s$ ” is the sample taken into consideration, “ $N$ ” is the total amount of data samples gathered, “ $P_m$ ” is the photovoltaic electrical power real measure and “ $P_e$ ” is the estimated value.

Table 5 presents the monthly and overall MAE and MAPE for each applied technique in HS. By the analysis of these results, an improvement of GDO over MLR can be observed by achieving a lesser error value; nonetheless, the ANFIS overcomes both MLR and GDO, successfully obtaining lower error values for every time range. By increasing “ $N$ ” in Equations (8) and (9) from the first to the last data sample, Figure 13 displays graphically the overall MAE and MAPE behavior of each implemented technique, showing and proving a clearly better performance by ANFIS against the statistical methods.

**Table 5.** Error comparison between MLR, GDO and ANFIS for HS.

Month	MLR		GDO		ANFIS	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
	(W)	(%)	(W)	(%)	(W)	(%)
Aug	296.5	11.8	299.8	11.9	23.8	0.9
Sep	222.7	8.3	203.7	7.6	96.3	3.6
Oct	276.9	9.5	268.0	9.2	168.1	5.8
Nov	142.7	5.3	138.2	5.1	96.2	3.6
Dec	190.8	7.5	178.9	7.0	129.4	5.1
Jan	192.0	7.1	184.5	6.8	126.3	4.7
Total	209.3	7.2	200.1	6.9	169.2	5.8



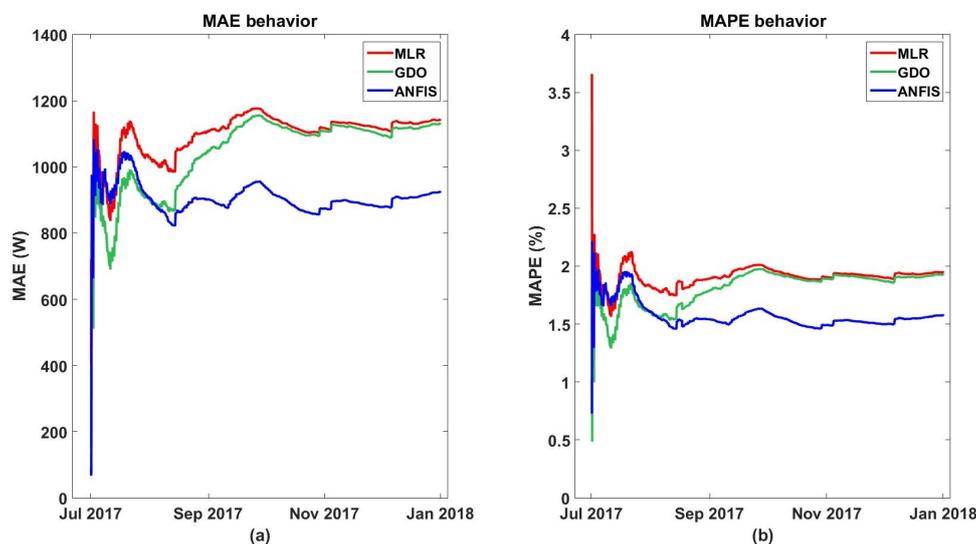
**Figure 13.** Error behavior comparison between MLR, GDO and ANFIS from HS: (a) MAE behavior; (b) MAPE behavior.

Table 6 presents the monthly and overall MAE and MAPE for each applied technique in MCS. As well as for the HS case; an improvement of GDO over MLR can be seen; nonetheless, the ANFIS overcomes again both MLR and GDO, successfully obtaining lower error values for every time range.

Analogous to Figure 13, Figure 14 displays graphically the overall MAE and MAPE behavior of each implemented technique for MCS, proving as well an improved performance by ANFIS compared to the statistical methods.

**Table 6.** Error comparison between MLR, GDO and ANFIS for MCS.

Month	MLR		GDO		ANFIS	
	MAE	MAPE	MAE	MAPE	MAE	MAPE
	(W)	(%)	(W)	(%)	(W)	(%)
Jul	588.8	1.0	590.6	1.0	334.1	0.6
Aug	1234.4	2.1	1192.9	2.0	732.9	1.3
Sep	727.6	1.3	726.3	1.3	345.8	0.6
Oct	733.6	1.3	741.8	1.3	441.4	0.8
Nov	715.5	1.3	741.4	1.4	477.6	0.9
Dec	794.1	1.7	812.8	1.7	638.3	1.4
Total	1142.2	2.0	1130.4	1.9	924.7	1.6



**Figure 14.** Error behavior comparison between MLR, GDO and ANFIS from MCS: (a) MAE behavior; (b) MAPE behavior.

#### 4. Conclusions

An ANFIS was used to estimate the photovoltaic electrical power generated by solar energy at two different locations, Hermosillo and Mexico City. The intelligent technique demonstrated a better performance than the MLR and GDO as statistical methods. The reported results for MCS showed that all three methods achieved a satisfactory estimation performance by their comparison against real measured values; nonetheless, the ANFIS system clearly displayed an improvement among the methods employed. For the case of HS, the MAE overall values with respect to GDO and MLR were 200.1 W and 209.3 W, and the MAPE overall values with respect to GDO and MLR were 6.9% and 7.2%. For the MCS, the MAE overall values with respect to GDO and MLR were 1130.4 W and 1142.2 W, and the MAPE overall values with respect to GDO and MLR were 1.9% and 2.0%. Consequently, GDO produced better results than MLR overall and in almost every monthly case; however, ANFIS outperformed both MLR and GDO in every case, achieving overall results with respect to MAE and MAPE of 169.2 W and 5.8% for HS and 924.7 W and 1.6% for MCS.

As outlined in Section 3, even when the ANFIS computational time is considerably greater than the one involved in statistical methods, the neuro-fuzzy result displays a better performance. It should be mentioned that, although the intelligent method took approximately 90 min to be trained, it is quite

an acceptable amount of time and can even be considered fast in terms of neuro-fuzzy models with months of data.

During a period of a year, the times of dusk and dawn are modified by minutes, resulting in different times to end and begin each day; as outlined in Section 3.1, for every transition day, dusk and dawn times may vary due to external variables, e.g., cloudiness. These issues generate discrepancies between the estimated and the real data, as seen in the results section, regardless the applied method; nonetheless, these may be minimized by gathering a greater amount of data to fulfil the range of changing time values.

The ANFIS proved to have a better performance than the conventional statistical methods, demonstrating that this kind of intelligent system is a potential tool to be considered for power estimation in Mexico.

**Author Contributions:** Research ideas and global design: N.P.-D., J.A.R.-H. and F.A.-V.; analysis of the results: Y.M., E.F.V.-C. and E.J.H.-L.; revision of the document: H.A., A.G.-J. and J.F.H.-P.; data gathering: R.A.P.-E.

**Funding:** This research received no external funding.

**Acknowledgments:** The authors would like to thank the University of Carmen (UNACAR), University of Sonora (UNISON), Centro de Investigación y de Estudios Avanzados (CINVESTAV) campus Zacatenco and Consejo Nacional de Ciencia y Tecnología (CONACYT), with the master scholarship program and the mobility scholarship 291249.

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

## Abbreviations

The following abbreviations are used in this manuscript:

ANFIS	Adaptive Neuto-Fuzzy Inference System
ANN	Artificial Neural Network
°C	Celsius degree. It is a temperature scale used by the International Systems of Units
GDO	Gradient Descent Optimization
GSR	Global Solar Radiation
HS	Hermosillo Site
kWh/m <sup>2</sup>	Quantity of kilowatts during an hour striking a squared meter area
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCS	Mexico City Site
MF	Membership Function
MLR	Multiple Linear Regression
MPH	Miles Per Hour
PV	Photovoltaic
W	Watt. It is a unit of power defined as a derived unit of 1 joule per second (J/s)
W/m <sup>2</sup>	Quantity of watts striking a squared meter area

## References

- Zahedi, A. Solar photovoltaic (PV) energy; latest developments in the building integrated and hybrid PV systems. *Renew. Energy* **2006**, *31*, 711–718. [[CrossRef](#)]
- Kazem, H.A.; Yousif, J.H.; Chaichan, M.T. Modeling of daily solar energy system prediction using support vector machine for Oman. *Int. J. Appl. Eng. Res.* **2016**, *11*, 10166–10172. [[CrossRef](#)]
- Awan, A.B.; Zubair, M.; P., P.R.; Abokhalil, A.G. Solar Energy Resource Analysis and Evaluation of Photovoltaic System Performance in Various Regions of Saudi Arabia. *Sustainability* **2018**, *10*, 1129. [[CrossRef](#)]
- Perea-Moreno, A.J.; Hernandez-Escobedo, Q.; Garrido, J.; Verdugo-Diaz, J.; Perea-Moreno, A.J.; Hernandez-Escobedo, Q.; Garrido, J.; Verdugo-Diaz, J.D. Stand-Alone Photovoltaic System Assessment in Warmer Urban Areas in Mexico. *Energies* **2018**, *11*, 284. [[CrossRef](#)]
- Gardiner, S.K.; Crabb, D.P. Examination of different pointwise linear regression methods for determining visual field progression. *Investig. Ophthalmol. Vis. Sci.* **2002**, *43*, 1400–1407. Available online: <https://iovs.arvojournals.org/article.aspx?articleid=2123577> (accessed on 10 July 2019).

6. Verma, S.; Bartosova, A.; Markus, M.; Cooke, R.; Um, M.J.; Park, D.; Verma, S.; Bartosova, A.; Markus, M.; Cooke, R.; et al. Quantifying the Role of Large Floods in Riverine Nutrient Loadings Using Linear Regression and Analysis of Covariance. *Sustainability* **2018**, *10*, 2876. [CrossRef]
7. Grimaccia, F.; Mussetta, M.; Zich, R. Neuro-fuzzy predictive model for PV energy production based on weather forecast. In Proceedings of the 2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011), Taipei, Taiwan, 27–30 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 2454–2457. [CrossRef]
8. Kashyap, Y.; Bansal, A.; Sao, A.K. Solar radiation forecasting with multiple parameters neural networks. *Renew. Sustain. Energy Rev.* **2015**, *49*, 825–835. [CrossRef]
9. He, J.; Zelikovsky, A. MLR-tagging: Informative SNP selection for unphased genotypes based on multiple linear regression. *Bioinformatics* **2006**, *22*, 2558–2561. [CrossRef] [PubMed]
10. Ruz-Hernandez, J.A.; Matsumoto, Y.; Arellano-Valmaña, F.; Pitalúa-Díaz, N.; Cabanillas-López, R.E.; Abril-García, J.H.; Herrera-López, E.J.; Velázquez-Contreras, E.F. Meteorological Variables' Influence on Electric Power Generation for Photovoltaic Systems Located at Different Geographical Zones in Mexico. *Appl. Sci.* **2019**, *9*, 1649. [CrossRef]
11. Ding, M.; Wang, L.; Bi, R. An ANN-based Approach for Forecasting the Power Output of Photovoltaic System. *Procedia Environ. Sci.* **2011**, *11*, 1308–1315. [CrossRef]
12. Yousif, J.H.; Kazem, H.A.; Alattar, N.N.; Elhassan, I.I. A comparison study based on artificial neural network for assessing PV/T solar energy production. *Case Stud. Therm. Eng.* **2019**, *13*, 100407. [CrossRef]
13. Wen-Tao, Z.; Shuai, W.; Xin-Hui, D. Research of power prediction about photovoltaic power system: Based on bp neural network. *Multi-Criteria Anal. Air Pollut. Urban Environ. Due Road Traffic* **2017**, *18*, 1614–1623. Available online: [https://www.researchgate.net/profile/Claudia\\_Moisa/publication/323144140-SUSTAINABLE\\_DEVELOPMENT\\_THROUGH\\_CONVERSION\\_TO\\_ORGANIC\\_AGRICULTURE\\_-\\_IMPLICATIONS\\_ON\\_THE\\_FINANCIAL\\_INDICATORS\\_OF\\_FIRMS/links/5c25f7bfa6fdccfc706d4a64/SUSTAINABLE-DEVELOPMENT-THROUGH-CONVERSION-TO-ORGANIC-AGRICULTURE-IMPLICATIONS-ON-THE-FINANCIAL-INDICATORS-OF-FIRMS.pdf#page=313](https://www.researchgate.net/profile/Claudia_Moisa/publication/323144140-SUSTAINABLE_DEVELOPMENT_THROUGH_CONVERSION_TO_ORGANIC_AGRICULTURE_-_IMPLICATIONS_ON_THE_FINANCIAL_INDICATORS_OF_FIRMS/links/5c25f7bfa6fdccfc706d4a64/SUSTAINABLE-DEVELOPMENT-THROUGH-CONVERSION-TO-ORGANIC-AGRICULTURE-IMPLICATIONS-ON-THE-FINANCIAL-INDICATORS-OF-FIRMS.pdf#page=313) (accessed on 10 July 2019).
14. Kuamr, K.R.; Kalavathi, M.S. ANN-ANFIS Based Forecast Model for Predicting PV and Wind Energy Generation. In Proceedings of the World Congress on Engineering 2016, WCE 2016, London, UK, 29 June–1 July 2016; Volume I. Available online: [http://www.iaeng.org/publication/WCE2016/WCE2016\\_pp322-327.pdf](http://www.iaeng.org/publication/WCE2016/WCE2016_pp322-327.pdf) (accessed on 10 July 2019).
15. Bassam, A.; May Tzuc, O.; Escalante Soberanis, M.; Ricalde, L.; Cruz, B.; Bassam, A.; May Tzuc, O.; Escalante Soberanis, M.; Ricalde, L.J.; Cruz, B. Temperature Estimation for Photovoltaic Array Using an Adaptive Neuro Fuzzy Inference System. *Sustainability* **2017**, *9*, 1399. [CrossRef]
16. Mashaly, A.F.; Alazba, A.A. ANFIS modeling and sensitivity analysis for estimating solar still productivity using measured operational and meteorological parameters. *Water Sci. Technol. Water Supply* **2018**, *18*, 1437–1448. [CrossRef]
17. Song, Y.; Park, M.; Song, Y.S.; Park, M.J. A Study on Estimation Equation for Damage and Recovery Costs Considering Human Losses Focused on Natural Disasters in the Republic of Korea. *Sustainability* **2018**, *10*, 3103. [CrossRef]
18. Clack, C.T.M. Modeling Solar Irradiance and Solar PV Power Output to Create a Resource Assessment Using Linear Multiple Multivariate Regression. *J. Appl. Meteorol. Climatol.* **2017**, *56*, 109–125. [CrossRef]
19. Montgomery, D.C. *Design and Analysis of Experiments*; John Wiley & Sons: Hoboken, NJ, USA, 2017; ISBN 978-1-119-11347-8. Available online: <https://www.wiley.com/en-mx/Design+and+Analysis+of+Experiments2C+9th+Edition-p-9781119113478> (accessed on 10 July 2019).
20. Bairán, J.M.; Marí, A.; Cladera, A. Analysis of shear resisting actions by means of optimization of strut and tie models taking into account crack patterns. *Hormigón Y Acero* **2018**, *69*, 197–206. [CrossRef]
21. Arjoune, Y.; Mrabet, Z.; Kaabouch, N.; Arjoune, Y.; Mrabet, Z.E.; Kaabouch, N. Multi-Attributes, Utility-Based, Channel Quality Ranking Mechanism for Cognitive Radio Networks. *Appl. Sci.* **2018**, *8*, 628. [CrossRef]
22. Chen, S.H.; Jakeman, A.J.; Norton, J.P. Artificial Intelligence techniques: An introduction to their use for modelling environmental systems. *Math. Comput. Simul.* **2008**, *78*, 379–400. [CrossRef]

23. Matich, D.J. *Redes Neuronales: Conceptos básicos y aplicaciones*. Universidad Tecnológica Nacional México. Available online: <ftp://decsai.ugr.es/pub/usuarios/castro/Material-Redes-Neuronales/Libros/matich-redesneuronales.pdf> (accessed on 10 July 2019).
24. Aggarwal, S.; Saini, L. Solar energy prediction using linear and non-linear regularization models: A study on AMS (American Meteorological Society) 2013–14 Solar Energy Prediction Contest. *Energy* **2014**, *78*, 247–256. [[CrossRef](#)]
25. Pitalúa-Díaz, N.; Penaloza, U.C.; Ruz-Hernandez, J.A.; Jimenez, R.L. *Introducción a Los Sistemas Inteligentes*; Departamento de Editorial Universitaria (UABC): Mexicali, Baja California, Mexico, 2008; ISBN 978-607-7753-47-6. Available online: <https://libreriauabc.com/products/introduccion-a-los-sistemas-inteligentes> (accessed on 10 July 2019).
26. Jang, J.S. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 665–685. [[CrossRef](#)]
27. Boyacioglu, M.A.; Avcı, D. An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange. *Expert Syst. Appl.* **2010**, *37*, 7908–7912. [[CrossRef](#)]



© 2019 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/>).