International Journal of Multidisciplinary and Current Research

Research Article

ISSN: 2321-3124
Available at: http://ijmcr.com

Statistical Optimization of Ambient temperature and Irradiation conditions for Solar Photovoltaic Module performance in composite climate using data analysis and graphing workspace

R.Siddiqui[1]*, R.Kumar[1], G Perveen[1] and U.Bajpai[2]

[1]National Institute of Solar Energy, 19th Milestone, Gwalpahari, Gurgaon-Faridabad Road, Haryana-122001, India
[2]Renewable Energy Research Laboratory, Department of Physics, University of Luck now, Lucknow-226007, India

Accepted 16 Nov 2014, Available online 01 Dec 2014, **Vol.2** **(Nov/Dec 2014 issue)**

Abstract

Data analysis and graphing workspace was used to determine the optimum solar photovoltaic performance conditions in composite climate for multicrystalline technology. An effort has been made to develop a relation with the given data, where tests are operated in the site of NISE, Haryana (North Latitude 28.47 ^{o} N, East Longitude 77.03^{o}E and Elevation from sea level is 216m) as specific composite climate environment. Here data are shown considering average values of the variables (Efficiency with ambient temperature and irradiation) for a period of three years (2010-2013). The objective of this study is to determine the best set of parameters such that the regression model used in the study predicts experimental values of the dependent variable (efficiency of PV module) as accurately as possible (i.e. calculated values of efficiency of PV module should be close to experimental values). Here the regression model itself is verified to fit the observed experimental data choosing the correct mathematical form of it. The analysis indicates that both the variables (ambient temperature and irradiation) can be used to optimize efficiency of PV module for potential commercial applications.

Keywords: *Crystalline PV module, Regression Analysis, Analysis of Variance.*

1. Introduction

Solar Panels have become one of the most promising ways to handle the electrification requirements of numerous isolated consumers worldwide [1].

Technological dependency of the industrialized world on fossil fuels and the ways in which these fuels have steadily degraded the earth's environment is quite alarming.

Solar Radiation (flux) is evidently a determining factor when it comes to studying the natural potential of solar energy as a source of renewable energy. For tropical regions, the average solar radiation potential is about 16.4±1.2MJ/m^2 per day [2]. Solar Flux is described by the visible and near infrared radiation emitted from the sun. The different spectrums are described by their wavelength that range within the broad range of 0.20 to 4.0µ (microns). Terrestrial radiation is a term used to derive infrared radiation emitted from the atmosphere.
Approximately 99% of solar or short wave radiation at the earth's surface is contained in the region from 0.3 to 3.0 µm while most of terrestrial or long-wave radiation is contained in the region of 3.5to 50 µm. Outside the earth's atmosphere, solar radiation has an intensity of approximately1370watts/m^2. This is the value at mean earth-sun distance at the top of the atmosphere and is referred to as solar constant. On the surface of the earth on a clear day, at noon, the direct beam radiation will be approximately1000watts/meter^2 for many locations, at sea level. The availability of energy is affected by location (including latitude and elevation), season and time of the day. All of which can be readily determined. However, the biggest factors affecting the availability of energy are cloud and other meteorological conditions which vary with location and time. Solar energy impinging upon a transmitted medium or target is partly reflected and absorbed.

The remainder is transmitted. The relative values are dependent upon the optical properties of the transparent object and the solar spectrum. Solar radiation is partially depleted and attenuated as it traverses the atmospheric layers, preventing a substantial portion of it from reaching the earth's surface. This phenomenon is due to absorption, scattering and reflection in the upper atmosphere (stratosphere), with its thin layer of ozone and lower atmosphere troposphere within which cloud formations occur and weather conditions manifest themselves.

Radiation from the sun absorbed by the earth warms it up and its temperature rises, this situation also occurs when radiation falls on solar panels. The performance of solar panels, in terms of power output, is a function of the availability of solar energy resource in the location. Solar intensity is high modulated due to the rapidly changing cloud cover and accounts for the rapidly changing temperatures. Solar panel temperature increases more rapidly than ambient temperature. This is because low energy solar flux (photons) reaching the solar panel are absorbed as heat. Solar cells are encapsulated in black material which makes them good heat absorbers.

The gap between the solar cell and glass casing of the panel encourages greenhouse effect which equally adds to the temperature increase of solar panel. While this occurs, the ambient temperature remains relatively uniform throughout the day. The solar cell is a PN junction device which can be modeled as a diode with a photo generated current source in parallel. When this is done, the current I, flowing through the PN junction is given by equation (1).

$$I = I_o (\exp qv/\eta KT -1) \quad (1)$$

Where I_o is reverse saturation current that is dependent on temperature, K is the Boltzmann's constant; q is the electron charge, η is a diode dependent ideality factor. Equation (1) shows the inverse relationship of current and temperature [3].

1.1 Effect of temperature

The changes in ambient temperature influence the performance of the solar cell. The efficiency of the cell gets reduced with the increase of cell temperature. V_{oc} is sensitive to temperature whereas I_{sc} is not. Simple calculation may show that the cell voltage and temperature are inversely related.

For silicon cell, (dV_{oc}/dT) is approximately equal to -2 mV/°C, which means, that the efficiency of the cell drops by about 0.4 % for increase of every one degree Celsius. A silicon solar cell of 20% efficiency at 20°C will reduce to 16% at 30°C [4].

Analysis of performance of industrial solar cell depending on the temperature and humidity carried out with the help of environmental chamber, by changing the conditions inside the environmental chamber shows the variation in energy conversion efficiency of solar cell.

From the analysis, it is confirmed that as temperature increases in the range of 31^0C to 58^0C efficiency of single crystalline solar cell also varies. The light conversion efficiency of single crystalline solar cell shows 9.702 % efficiency at 31^0C as increase in temperature the conversion efficiency increases and it reaches up to 12.0459 % at 36^0C temperature however, further temperature increases from 36^0C the conversion efficiency decreases slowly & it goes up to 6.60% at 45^0C. Further increase in temperature from 45^0C there is continuously decrement in conversion efficiency & we found that at 58^0C the single crystalline silicon solar cell shows 2.37061 % conversion efficiency.

The outcome of the studies can be applied to solar cell field with respect to temperature and humidity at specific locations. [4]

The pronounced effect that the operating temperature of a photovoltaic (PV) cell/module has upon its electrical efficiency is well documented. There are many correlations expressing T_c, the PV cell temperature, as a function of weather variables such as the ambient temperature, T_a, and the local wind speed, V_w, as well as the solar radiation flux/irradiance, G_T, with material and system-dependent properties as parameters, e.g., glazing-cover transmittance,ɽ, plate absorptance, a, etc[5].

1.2 Effect of High Temperature

The temperature of PV surface rises with longer exposure period to sunlight and high ambient temperature. The elevated temperatures directly impact the PV efficiency. Atom vibrations (photons) in a p-n junction cell increase and obstruct charge carrier movement which decreases cell efficiency [6]. As a part of power system program of the International Energy Agency [EIA], a study was conducted to analyze data from 18 grid connected PV plants located on different geographic locations and it showed a direct relation between temperature and PV module efficiency. The plants were located in Austria, Germany, Italy, Japan and Switzerland. The study concluded that 17 out of 18 systems showed annual losses in efficiency due to temperature changes by 1.7% to 11.3%. The highest efficiency reduction was observed at relatively high ambient temperature of 30^oC [7]. A study was conducted on a polycrystalline PV module with solar tracker on Dhahran- east of Saudi Arabia showed similar temperature effect. The data was compared based on daily peak power output. PV module efficiency decreased from 11.6% to 10.4% when module temperature increased from 38^oC to 48^oC which corresponds to 10.3% losses in efficiency and a temperature coefficient of -0.11ΔE / $\%^o$ C [8].

An equally large number of correlations expressing the temperature dependence of the PV module's electrical efficiency, η_c, can also be retrieved, although many of them assume the familiar linear form, differing only in the numerical values of the relevant parameters which, as expected, are material and system dependent[9]. Many correlations in this category express instead the modules maximum electrical power, P_m, which is simply related to η_c through the latter's definition (η_c = P_m (under standard test conditions1)/A_{GT}, with A being the aperture area), and form the basis of various performance rating procedures.

The effect of temperature on the electrical efficiency of a PV cell/module can be traced to the former's influence upon the current, I, and the voltage, V, as the maximum power is given by

$$P_m = V_m I_m = (FF) V_{oc} I_{sc}.$$

In this fundamental expression, which also serves as a definition of the fill factor (FF), subscript m refers to the maximum power point in the module's I–V curve, while subscripts oc and sc denote open circuit and short circuit values, respectively.

1.3Temperature variation as a function of solar irradiance intensity

The PV temperature variation as a function of normal incident solar radiation, varied from 100 to 1000W/m^2 is observed resulting with the hypothesis as 1) solar radiation normal to the PV surface, 2) ambient temperature at 20^oC and 3) tilt angle fixed at 30^o.

Without incident solar radiation the module temperature should be almost equal to ambient temperature. A small radiative thermal exchange between sky and front/rear surfaces arises, due to the lower temperature of the sky with respect to the surrounding environment, as can be calculated by [10,11].

The operating temperature plays a central role in the photovoltaic conversion process [12]. Both the electrical efficiency and hence the power output of a PV module depend linearly on operating temperature, decreasing with T_c. The various correlations that have been proposed in the literature represent simplified working equations which apply to PV modules or PV arrays mounted on free standing frames, to PV/Thermal collectors, and to BIPV arrays, respectively. They involve basic environmental variables, while the numerical parameters are not only material dependent but also system dependent. Thus, care should be exercised in applying a particular expression for the electrical efficiency or the power output of PV module or array, as each equation has been developed for a specific mounting frame geometry or level of building integration. The same holds for PV module rating method, the details and limitations of which should be very clear to the prospective user. The reader therefore should consult the original sources and try to make intelligent decisions when seeking a correlation or a rating procedure to suit the needs [5].

With increase of ambient temperature there is a deficiency in electrical energy that solar cells supply than their values under ideal conditions (25 °C - 1000 W/m^2), this situation be of a high affection especially in countries of a hot climate [13].

An expression for the variation of the maximum theoretical efficiency of a solar cell with temperature is presented. The expression relates the difference in the maximum theoretical efficiencies of the cell at 0^oC and a given temperature to the difference of the fourth roots of the corresponding temperatures. Values of the maximum theoretical efficiency at various temperatures obtained from the present expression are shown to agree very well with values theoretically evaluated by [14].

The correlations expressing the PV module performance as a function of weather variables such as the ambient temperature, local wind speed, etc. have been discussed by [15] and therefore the deviation in the performance of solar module under climatic parameter as ambient temperature and wind velocity for a given location is studied.

The objective of this study is to determine the best set of parameters such that the regression model used in the study predicts experimental values of the dependent variable (efficiency of PV module) as accurately as possible (i.e. calculated values of efficiency of PV module should be close to experimental values). Here the regression model itself is verified to fit the observed experimental data choosing the correct mathematical form of it.

2. PV Modules Description

The tested PV modules are based on multicrystalline solar cells **JAIN IRRIGATION** manufacturers.

The area of each solar cell is 96.72cm^2. Solar cells are arranged in 9×4series-parallel connected cells configuration.

Table 1 Specifications given by the manufacturers of **Jain Irrigation**

M/S - Jain Irrigation
Module No – WS 50
Serial No-WSAZL061002395
Module Area- 63×67.5
Cell Area- 15.6×6.2= 96.72
No of Cells –36/ Multi C-Si
P_{max} - 50W, V_{oc}-21.0V,I_{sc}-3.17A

3. PV Modules Indoor Testing

Table 2 summarizes PV modules performances prior to any outdoor exposure and established through indoor measurement with sun simulator under standard test condition(STC) as controlled indoor conditions(1000W/m^2, AM 1.5., Global Spectrum, 25^oC) using a calibrated Solar simulator.

Table 2 PV modules performances at STC Conditions

PV Module	Pmax (W)	Isc (A)	Voc (V)	Rs (Ω)	Rsh (Ω)	η (%)
Jain I.	49.9	3.249	21.84	748	89	11.7

4. PV Module Outdoor Testing In Composite Climate

The outdoor measurements were performed in the site of National Institute of Solar Energy, 19th Milestone Gwalpahari, Gurgaon–Faridabad Road, Haryana, as specific composite climate environment, characterized by high irradiation and temperature levels. The geographic characteristics of NISE site are North latitude 28.4700^{o}N, East Longitude 77.0300^{o}E, Elevation from sea level is 216 m.

Fig.1 National Institute of Solar Energy located in Gurgaon Region of Haryana (India)

An open rack is used to mount the module outside in the sun with a pyranometer installed in a specified manner. The rack is designed to minimize heat conduction from the module and to interfere as little as possible with the free radiation of heat from the front and rear surface of the module. Both the modules are positioned in a way so that it is normal to the solar beam (within $\pm 5^{o}$) at local solar noon. The bottom edge of the module is 0.6m above the horizontal plane i.e., ground level as illustrated in figure 2.

Fig.2 PV modules of **Jain Irrigation** in outdoor exposure

The reference solar cell used in our experimental investigation is based on multicrystalline silicon with integrated Pt 1000 temperature sensor. The PV modules under test receive an electrical performance (I-V), under environmental conditions for different values of solar irradiance and an ambient temperature on clear sunny day.

5. Methodology

5.1Selection of the variables

The efficiency of solar photovoltaic module obtained during the day has different values. This fluctuation results from different factors affecting the performance of solar photovoltaic module. For example, these factors may be temperature, humidity, wind velocity, cloud cover, dust, etc [16].

Previous studies have been done, where equations have been developed with the given data for different seasons of particular location called Lucknow, India consisting of composite climate, which is helpful in developing a relation of efficiency of photovoltaic modules with the major climatic parameters like temperature, wind velocity, humidity, dust etc, further this equation developed mathematically is in good correlation with the measured data

The performance of module gives a broad view of impact of climatic variables and helps to find out the efficiency of modules while knowing the climatic parameters of a particular area [17, 18].

In the present study, three year data is studied for multicrystalline PV module of **Jain Irrigation** in the site of National Institute of Solar Energy, 19th Milestone Gwalpahari, Gurgaon–Faridabad Road, Haryana, as specific composite climate environment, characterized by high irradiation and temperature levels.

Table3 Average value of Efficiency of multicrystalline Solar PV Module of (**JAIN IRRIGATION**) with average value of ambient Temperature, radiation data for the average year (2010-2013)

Month	Efficiency (average) (η) (%) Y	Ambient Temperature (average) (^{o}C) X_1	Irradiation (average) (W/m^2) X_2
Jan	9.53	22.00	535
Feb	11.79	25.00	801
March	12.18	34.00	881
April	11.36	40.25	892
May	11.00	49.44	978
June	11.12	40.36	895
July	11.16	40.47	910
Aug	11.66	43.76	954
Sept	11.83	35.65	978
Oct	10.36	39.00	697
Nov	9.83	26.80	600
Dec	8.90	21.00	473

The objective of this study is to give a more detailed description of the regression tool using regression and correlation analysis as statistical techniques to examine relationships among variables.

In a regression analysis we study the relationship, called **the regression function**, between one variable **Y**, called the **dependent variable,** and several others $\mathbf{X_i}$, called the **independent variables.**

Here, dependent variable is efficiency of PV module(Y) and independent variable is ambient temperature(X_1) and radiation (X_2) according to table 3.

5.2Multiple Linear Models

General Formula:

$$Y= b_o + b_1X_1 + b_2X_2 + b_pX_p \quad (1)$$

$$Y= b_o + \Sigma_i b_i X_i \quad i=1,2,\ldots p \quad (1a)$$

Polynomial (model is linear in parameters, but not in independent variables):

$Y= b_o + b_1X+ b_2X^2 + \ldots b_p X^p$, which is just a specific case of (1)

With $X_1=X$, $X_2=X^2$, $X_3=X^3$ …….. $X_p = X^p$

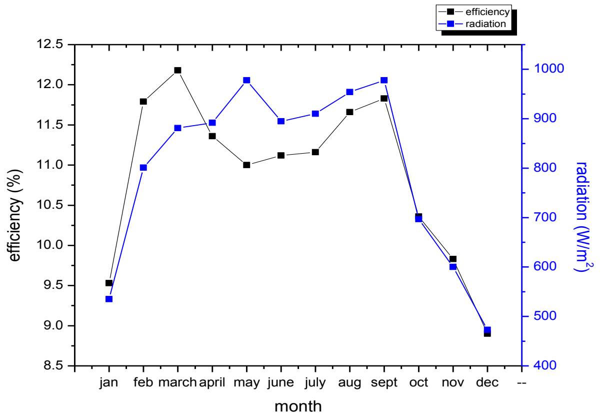

Fig. 3a. shows plotting of efficiency (%) and radiation (W/m^2) set of data for the average year (2010-2013)

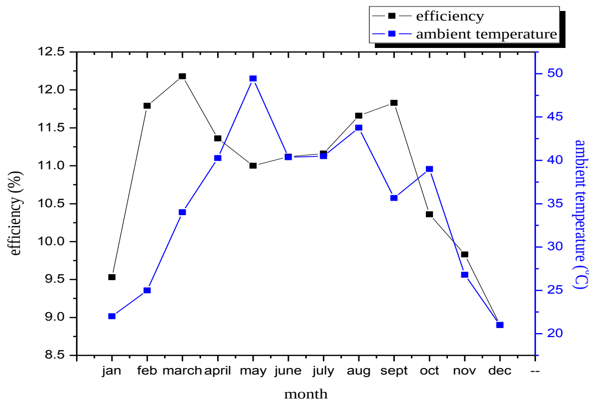

Fig. 3b. shows plotting of efficiency (%) and ambient temperature (°C) set of data for the average year (2010-2013).

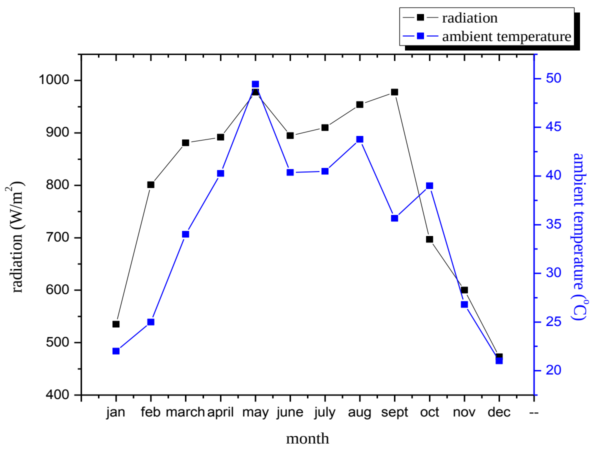

Fig. 3c. shows plotting of radiation (W/m^2) and ambient temperature (°C) set of data for the average year (2010-2013)

5.3Polynomial Regression

A common "nonlinear" model is polynomial regression. The term nonlinear is put in quotes here because the nature of this model is actually linear, this expectation can be expressed in the general regression equation:

Efficiency= a+b$_1$* radiation + b$_2$* radiation^2

and

Efficiency= a+b$_1$* ambient temperature + b$_2$* ambient temperature^2

In the above equations, a represents the intercept and b_1 and b_2 are regression coefficients. The non-linearity of this model is expressed in the term **radiation^2** and **ambient temperature^2**. However, the nature of the model is still linear, except that when estimating it, we would square the measure of radiation and ambient temperature. These types of models, where we include some transformation of the independent variables in a linear equation, are also referred to as models that are nonlinear in the variables.

6. Description of the Model used and Its Verification

Our aim is to determine the best set of parameters b_i, such that the model predicts experimental values of the dependent variable as accurately as possible. It is necessary to verify that the model itself is adequate to fit the observed experimental data.

6.1Detailed Description of Regression input and Output

The following experimental data collected for a whole year for daily monthly average values for the three variables, two independent and one dependent illustrates the discussion.

Table 4 Regression Input

Data Point j	Dependent Variable (Efficiency) Y*	Independent Variable (Ambient Temperature) X_1	Independent Variable (Radiation) X_2
1	9.53	22.00	535
2	11.79	25.00	801
3	12.18	34.00	881
4	11.36	40.25	892
5	11.00	49.44	978
6	11.12	40.36	895
7	11.16	40.47	910
8	11.66	43.76	954
9	11.83	35.65	978
10	10.36	39.00	697
11	9.83	26.80	600
12	8.90	21.00	473

Table 5 Residual Output

Observation	*Predicted Y*	*Residuals*	*Standard Residuals*
1	9.656864	-0.12686	-0.38008

2	11.53735	0.252645	0.756919
3	11.581	0.598999	1.794586
4	11.26394	0.096063	0.287801
5	11.34212	-0.34212	-1.02497
6	11.28024	-0.16024	-0.48007
7	11.39009	-0.23009	-0.68935
8	11.52109	0.138911	0.416174
9	12.23091	-0.40091	-1.20111
10	9.824201	0.535799	1.60524
11	9.854262	-0.02426	-0.07269
12	9.237938	-0.33794	-1.01245

Residual (**or error, or deviation**) is the difference between the observed value Y*of the dependent variable for the jth experimental data point (X_{1j}, X_{2j}, …X_{pj}, Y*) and the corresponding value Yj given by the regression function

$Y_j = b_o + b_1X_{1j} + b_2X_{2j}$ (in our experimental analysis) (1)

$r_j = Y_j^* - Y_j$ (2)

Parameters b ($b_o, b_1, b_2 \ldots\ldots b_p$) are part of the ANOVA output.

A plot of residuals is very helpful in detecting an obvious correlation between the residuals and the independent variable.

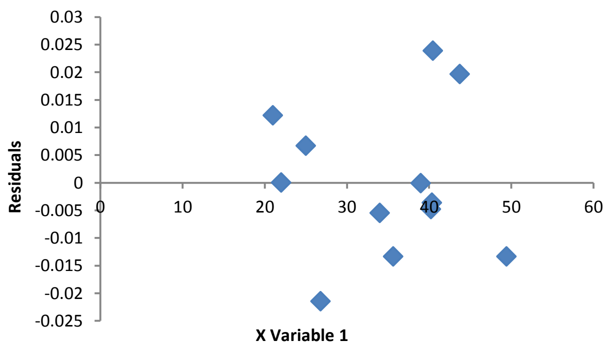

Fig. 4a. Correlation between the residual and the independent variable (ambient temperature)

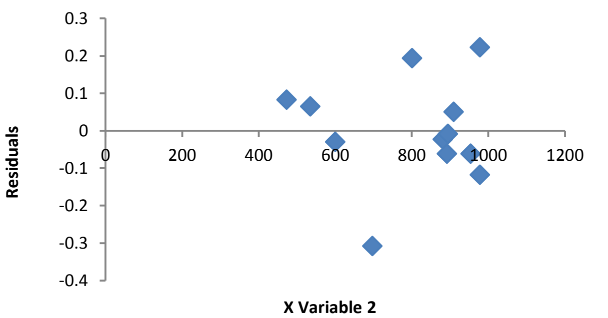

Fig. 4b Correlation between the residual and the independent variable (irradiation)

However, the fact that the residual look random and there is no obvious correlation with the variable X does not necessarily mean by itself that the model is adequate. More tests are needed.

Standard (or standardized) residual is a residual scaled with respect to the *standard error (deviation)* S_y in a dependent variable

$r_j' = r_j / S_y$ (2a)

The quantity S_y is part of the *"Regression Statistics"* output.

ANOVA Output

There are two tables in ANOVA (Analysis of variance)

Table 6 ANOVA output (part I)

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	10.29796	5.148979	37.81353	4.17E-05
Residual	9	1.225509	0.136168		
Total	11	11.52347			

Table 6a ANOVA output (part II)

	Intercept	X Variable 1(ambient temperature)	X Variable 2(radiation)
Coefficients	6.903722	0.007796	-0.06445
Standard Error	0.509808	0.001095	0.021388
t Stat	13.5418	7.117329	-3.01346
P-value	2.73E-07	5.56E-05	0.014634
Lower 95%	5.750456	0.005318	-0.11283
Upper 95%	8.056988	0.010274	-0.01607

6.2Coefficients

Coefficients are listed in the second table of ANOVA (Table 6a). These coefficients allow the program to calculate **predicted** values of the dependent variable Y (efficiency of PV module) which were used above in formula (2) and are a part of **residual output** (Table 5).

6.3Sum of Squares

In the **ANOVA** regression output one will find three types of sum of squares (Table 6)

Total Sum of Squares

SS_T (Total sum of squares)

$= SS_E + SS_R$ (3)

Where

SS_E - residual (error) Sum of Squares

SS_R - regression sum of squares

It is obvious that **SS_T** is the sum of squares of deviations of the experimental values of dependent variable Y* (efficiency of the PV module) from its average value. $\mathbf{SS_T}$ could be interpreted as the sum of deviations of Y* from the simplest possible model.

Residual (or error) sum of squares (SS_E)

$\mathbf{SS_E}$ could be viewed as the due –to – random- scattering- of- Y*- about – predicted –line contributor to the total sum of squares SS_T. This is the reason for calling the quantity "due to error (residual) sum of squares".

Regression Sum of Squares (SS_R)

$\mathbf{SS_R}$ is the sum of squares of deviations of the *predicted-by-regression-model* values of dependent variable (i.e. Efficiency of PV module) Y from its average experimental value $Y^*_{avg..}$ It accounts for addition of p(no.) variables (X_1, X_2,X_3,…..X_P) to the simplest possible model, here there is a transformation from the "non regression model" to the true regression model, so $\mathbf{SS_R}$ is also called as " due to regression sum of squares".

Mean square (variance) and degrees of freedom

The general expression for the mean square of an arbitrary quantity q is

$$MS_q= SS_q / df \quad (4)$$

SS_q is the sum of squares and df is the **number of degrees of freedom** associated with quantity SS_q. MS is also referred to as the **variance.** The number of degrees of freedom could be viewed as the difference between the number of observations n and the number of constraints (fixed parameters associated with the corresponding sum of squares).

Total mean square MS_T (total variance)

$$\mathbf{MS_T} = \mathbf{SS_T} / (n-1) \quad (5)$$

$\mathbf{SS_T}$ is associated with the model, which has only one constraint (parameter b_o), therefore the number of degrees of freedom in this case is:

$$df_T = n-1 \quad (5a)$$

Residual (error) mean square MS_E (error variance)

$$\mathbf{MS_E} = SS_E / (n-k) \quad (6)$$

SS_E is associated with the random error around the regression model (1), which has **k=p+1 parameters** (one per each variable out of p variables total plus intercept). It means there are k constraints and the number of degrees of freedom is:

$$df_E = n-k \quad (6a)$$

Regression mean square MS_R (regression variance)

$$\mathbf{MS_R= SS_R / (k}-1) \quad (7)$$

The number of degrees of freedom in this case can be viewed as the difference between the total number of degrees of freedom (df_T = n-1) (5a) and the number of degrees of freedom for residuals df_E (6a).

$$df_R = df_T – df_E = (n-1) – (n-k) \quad (7a)$$

$$df_R = k-1= p \quad (7b)$$

Test of Significance and F- numbers

The F-number is the quantity which can be used to test for the statistical difference between two variances. For example, if we have two random variables R and E, the corresponding F-number is:

$$F_R = MS_R/MS_E \quad (8)$$

In our analysis F-number is 37.81353(F_R), the variances MS_R and MS_E are defined by an expression of type (4).In order to tell whether **two variances are statistically different**, we determine the corresponding probability from F- distribution function:

$$P= P (F_R, df_R, df_E) \quad (9)$$

The quantities df_R, df_E – degrees of freedom for numerator and denominator- are parameters of this function.

The probability **P** given by (9) is a probability that the variances MS_R and MS_E are statistically **indistinguishable.** On the other hand, **1-P** is the probability that they are **different** and is often called **confidence level.** Conventionally, a reasonable confidence level is 0.95 or higher. If it turns out that **1-P< 0.95**, we say that MS_R and MS_E are statistically the same. If 1-P **>0.95**, we say that at least with the 0.95 (or 95%) confidence MS_R and MS_E are different. The higher the confidence level, the more reliable our conclusion, calculating the procedure numerically we get

Solving eq (9)

$$P= P (F_R, df_R, df_E) \quad (9)$$

P= FDIST (37.81353, 2, 9)

= 4.17157E-05
While calculating 1-P we get

1-P > 0.95= 0.99995

Here the confidence level is much higher; signifying the conclusion to be more reliable.

There are several F-tests related to regression analysis. Here three most common ones have been discussed. They deal with the significance of parameters in the regression model.

6.4 Significance test of all coefficients in the regression model

This test is performed to check with what level of confidence we can state that AT LEAST ONE of the coefficients **b** (b_1, b_2…bp) in the regression model is significantly different from zero.

After determining F_R i.e. F- number for the whole regression (part of regression output (as shown in table 6)).

The second step is to determine the numerical value of the corresponding probability P_R (also a part of regression output)

Finally we can determine the confidence level **1-P** (as calculated from equation 9). At this level of confidence, the variance "due to regression" $\mathbf{MS_R}$ (5.148979) (from Table 6)) is statistically different from the variance "due to error" $\mathbf{MS_E}$ **(**0.136168) (from Table 6)). In its turn it means that the addition of **p** variables (X_1, X_2… X_p) to the simplest model where Y = b_o (dependent variable Y is just a constant) is a statistically significant improvement of the fit. Thus, at the confidence level not less than **1-P** we can say: "At least ONE of the coefficients in the model is significant". **The higher the F_R the more accurate the corresponding model.**

6.5 Significance test of subset of coefficients in the regression model

With what level of confidence can we be sure that at least ONE of the coefficients in a selected subset of all the coefficients is significant? So it is necessary to test a subset of the last **m** coefficients in the model with a total of p coefficients (b_1, b_2,….b_p).

Here we need to consider two models:

$Y=b_o+b_1X_1+b_2X_2+b_pX_p$
(unrestricted) (10)

$Y=b'_o + b'_1X_1 + b'_2X_2 + b'_{p-m}X_{p-m}$ (restricted) (11)

These models are called unrestricted (10) and restricted (11) respectively. We need to perform ***two separate*** least square regression analyses for each model.

The regression output for the unrestricted model is already presented in Table 6. To test whether the quadratic terms are significant, in this case restricted model is considered where the equation comes out to be as

$Y= b_o + b_1X_1$ (restricted model) (12)

The subset of parameters consists of two parameters and m=2. By analogy with the input table for the unrestricted model (Table 4) we prepare for the restricted model:

Table7 Regression input for restricted model (considering independent variable X_1= Ambient Temperature)

Data Point # j	Dependent Variable (Efficiency) Y*	Independent Variable (Ambient Temperature) X_1
1	9.53	22.00
2	11.79	25.00
3	12.18	34.00
4	11.36	40.25
5	11.00	49.44
6	11.12	40.36
7	11.16	40.47
8	11.66	43.76
9	11.83	35.65
10	10.36	39.00
11	9.83	26.80
12	8.90	21.00

We perform an additional regression using this input table and as part of ANOVA.

Table 8 Regression ANOVA output for the restricted model

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.400	3.400	4.1857	0.067972
Residual	10	8.123	0.812		
Total	11	11.52			

From Table 6 and Table 8 we have:

$\mathbf{SS_E = 1.225509}$
(Error sum of squares; unrestricted model)

$\mathbf{MS_E = 0.136168}$
(Error mean square; unrestricted model)

$\mathbf{df_E = n-k = 9}$
(Degrees of freedom; unrestricted model)

$\mathbf{SS'_E = 8.123266}$
(Error sum of squares; restricted model)

Now we are able to calculate $F_{m=2}$:

$F_{m=2}$= {(**8.123266-1.225509)/2}/ 0.136168**
= 25.328

Using the Microsoft Excel function for the F-distribution we determine the probability $P_{m=2}$:

$P_{m=2}$ = FDIST ($F_{m=2}$, m, n-k)

$P_{m=2}$ = FDIST (25.328, 2, 9)

= 0.0002012

Finally we calculate the level of confidence

1- $P_{m=2}$ = 1-0.0002012
= 0.999798

Here 1- P_m is big enough (greater than 0.95) we state that ***other coefficients in the subset are significant to a good extent.***

Table9 Regression input for restricted model (considering independent variable X_1= Radiation)

Data Point # j	Dependent Variable (Efficiency) Y*	Independent Variable (Radiation) X_1
1	11.70	535
2	11.68	801
3	11.67	881
4	11.39	892
5	10.83	978
6	10.45	895
7	10.50	910
8	10.54	954
9	10.69	978
10	11.31	697
11	11.57	600
12	11.64	473

We perform an additional regression using this input table and as part of ANOVA.

Table 10 Regression ANOVA output for the restricted model

	Regression	Residual (error)	Total
df'	1	10	11
SS'	9.0614	2.462	11.5234
MS'	9.0614	0.2462	
F'	36.8045		
Significance F'	0.000121		

From Table 6 and Table 10 we have:

SS_E = 1.225509
(Error sum of squares; unrestricted model)
MS_E = 0.136168
(Error mean square; unrestricted model)

df_E = n-k = 9
(Degrees of freedom; unrestricted model)

SS'_E = 2.4620
(Error sum of squares; restricted model)

Now we are able to calculate $F_{m=2:}$

$F_{m=2}$= **2.4620-1.225509)/2}/ 0.136168**

= 4.5403

Using the Microsoft Excel function for the F-distribution we determine the probability $P_{m=2}$:

$P_{m=2}$ = FDIST ($F_{m=2}$, m, n-k)

$P_{m=2}$ = FDIST (4.5403, 2, 9)

= 0.043314

Finally we calculate the level of confidence

1- $P_{m=2}$ = 1-0.043314
= 0.956687

Here 1-Pm is also big enough (greater than 0.95), we state that other coefficient in the subset is significant also to a good extent.

6.6 Significance test of an individual coefficient in the regression model

In our illustration P_0 = 2.73E-07, and P_2 = 0.014634, Table 6a corresponds to fairly high confidence levels, 1-P_0 = 0.99999 and 1-P_2=0.98536. This suggests that parameters b_o and b_2 are significant. The confidence levels for b_1 (1-P_1=1- 5.56E-05) = 0.99994 are high, which means that it is significant.

6.7 Confidence Interval

For the unrestricted model, the lower and upper 95% limits for intercept are "5.75045"and "8.05698" respectively. The fact that with the 95% probability zero does not fall in this interval is consistent with our conclusion of significance of b_0 made in the course of F-testing of individual parameters. The confidence levels at the 95% for b_1 do not include zero. This also agrees with the F-Test of individual parameters.

6.8 Regression Statistics Output

Table 11 Regression Statistics Output

Multiple R	0.945331
R Square (R^2)	0.893651
Adjusted R^2(R^2_{adj})	0.870018
Standard Error (S_y)	0.369009
Observations (n)	12

The information contained in the "Regression statistics" output characterizes the "goodness" of the model as a whole. The quantities listed in this output can be expressed in terms of the regression F-number $\mathbf{F_R}$ (Table 6).

Standard Error (Sy)

$$Sy = (MS_E)^{1/2}$$

$\mathbf{MS_E}$ is an error variance discussed before (equation 6). Quantity S_Y is an estimate of the *standard error (deviation)* of experimental values of the dependent variable Y* with respect to those predicted by the regression model.

Coefficient of Determination R^2 (or R Square):

$$R^2 = SS_R / SS_T = 1- SS_E / SS_T$$

$\mathbf{SS_R}$, $\mathbf{SS_E}$ **and** $\mathbf{SS_T}$ are regression, residual (error) and total sum of squares.

The coefficient of determination is a measure of the regression model as whole. The closer R^2 is to one, the better the model (1) describes the data. In the case of a perfect fit R^2 =1.

Adjusted coefficient of determination R^2 (or Adjusted R Square):

$$R^2_{adj} = 1-\{SS_E / (n-k)\} / \{ SS_T / (n-1)\}$$

$\mathbf{SS_E}$ and $\mathbf{SS_T}$ are the *residual (error)* and *total sum of squares*. The significance of $\mathbf{R^2_{adj}}$ is basically the same as that of $\mathbf{R^2}$ (the closer to one the better).

Multiple Correlation coefficient R

The fact that $\mathbf{R^2_{adj} = 0.870018}$ in our illustration is fairly close to 1 (Table 11) suggests that overall model is **GOOD** to fit the experimental data presented in Table 3.

Conclusion

As clear from **Significance test of subset of coefficients in the regression model,** the regression analysis done for the restricted model considering independent variable as ambient temperature and radiation respectively, the level of confidence is **0.99994 and 0.98536 respectively,** which confirms that Efficiency of multicrystalline PV module was influenced more by ambient temperature than irradiation.

In conclusion of the F-Test discussion, it should be noted that in case we remove even one significant variable from the model, we need to test the model once again, since coefficients which were significant in certain cases **might** become insignificant after removal and vice versa.

The proposed model equation illustrated the quantitative effect of both the variables in addition to the interactions of the variables with the efficiency of the PV module.

Based on the outdoor conditions, the experimental efficiency content was in very good agreements with the predicted value, as shown from table 5 where residual (error) and standard residual (error) are shown for each value of predicted efficiency for the PV module.

The results of this analysis indicate that both the parameters (ambient temperature and irradiation) can be used to optimize efficiency of PV module for potential commercial applications.

Nomenclature

The following notation is used in this work

Y -Dependent variable (efficiency of the PV module) predicted by a regression model.

Y*-Dependent variable (efficiency of the PV module) experimental value.

p- Number of independent variables (ambient temperature and irradiation)(number of coefficients)

X_i- ith independent variable from total set of p variables (i=1,2…p).

b_i - ith coefficient corresponding to X_i (i=1,2…p).

b_o - intercept (or constant)

k- Total no of parameters including intercept (constant)(k=p+1)

n- Number of observations (experimental data points)

i=1,2…p – independent variables index

j= 1, 2,..n – Data points index.

References

[1]. Ettah E.B, Obiefuna J. Nwabueze, Njar G.N, (2011) The relationship between solar radiation and the efficiency of solar panels in Port Harcourt, Nigeria, *International Journal of Applied Science and Technology,* 1(4), 124-126.

[2]. Akpabio I. E, Udoimuk A. B, (2002) Characteristic distribution of total, diffuse and direct solar radiation at Calabar*, GJPAS,* 9(1), 139-146.

[3]. Ettah E.B, Eno E.E, Udoimuk A.B, (2009) The effects of Solar panel temperature on the power output efficiency in Calabar, Nigeria*, Journal of Association of Radiographers of Nigeria*, 23,16-22.

[4]. Katkar A.A, Shinde N.N, Kohli G.C, Gaikwad S.P, (2009) Evaluation of Industrial Solar Cell with respect to temperature*, Journal of Association of Radiographers of Nigeria*, 23, 16-22.

[5]. Skoplaki E, Palyros J.A, (2009) On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/ power correlations, *Solar Energy,* 83, 614-624.

[6]. Baras A etal, (2012), Opportunities and challenges of Solar Energy in Saudi Arabia, In *proc. World Renewable Energy forum(WREF)*2012, ed. By C. Fellows (Curran Associates), pp.4721

[7]. Nordmann T and Clavadetscher L, (2003), Understanding temperature effects on PV system performance, *Proc. 3rd World Conf on PV Energy Conversion,* 3, 2243.

[8]. Adinoyi M.J and S.A.M. Said, (2013) Effect of Dust Accumulation on the Power Outputs of Solar Photovoltaic Modules, *Renewable Energy,* 60, 633.

[9]. Nishioka K, Hatayama T, Urakoa Y, Fuyuki T, Hagihara R, Watanabe M, (2003) Field –Test Analysis of PV System Output Characteristics focusing on module temperature, *Solar Energy Materials and Solar Cells*, 75, 665-671.

[10]. Barolhi M, Grandi G and Tina G.M, (2012), Comparison of PV cell Temperature estimation by different thermal exchange calculation methods, *International Conference on Renewable Energies and Power Quality* (ICREPQ'12), Santiago de Compostela (Spain), 28th to 30th March.

[11]. Silva J.P, Nofuentes G, Munoz J.V, (2010) Spectral Reflectance patterns of Photovoltaic modules and their thermal effects, *Journal of Solar Engineering*, 132

[12]. Alonso Garcia M.C, Balenzategui J.L, (2010) Estimation of Photovoltaic module yearly temperature and performance based on nominal operation cell temperature calculations, *Renewable Energy,* 29.

[13]. Al Sabounchi A.M, (1998) Effect of ambient temperature on the demanded energy of solar cells at different inclinations, *Renewable Energy*, 14, 149-155.

[14]. Ravindra N. M, Srivasatva V.K, (2012) Temperature dependence of the maximum theoretical efficiency in solar cells, *Solar Energy Materials and solar Cells*,1, 107-109.

[15]. Siddiqui R, Bajpai U, (2012) Deviation in the performance of Solar Module under Climatic parameters as Ambient Temperature and Wind Velocity in Composite Climate, *International Journal of Renewable Energy Research,* 2(3), 486-490.

[16]. Furushima K, Nawata Y, Sadatomi M, (2006), Prediction of photovoltaic power output considering weather effects, In *ASES Conference SOLAR - Renewable Energy key to climate Recovery,* July 7-13, Denver, Colorado.

[17]. Siddiqui R, Bajpai U, (2012) Statistical Analysis of Solar Photovoltaic Module output with Temperature, Humidity and Wind Velocity in Composite Climate, *European Journal of Scientific Research,* 80 (4), 447-456.

[18]. Siddiqui R, Bajpai U, (2012) Correlation between thicknesses of dust collected on photovoltaic module and difference in efficiencies in composite climate, *International Journal of Energy and Environment Engineering*, 3(1), 1-7. doi:10.1186/2251-6832-3-26.

[19]. Resser V.V, Fuentes M.K, (1984), Linear regression analysis of flat plate photovoltaic system performance data. IN : *Photovoltaic Solar Energy Conference, Proceedings of the fifth International Conference,* Athens, Greece, October17-21, 1983(A85-11301 02-44). Dordrecht, D. Reidel Publishing Co., p.623-627.