Maximum Power Point Tracking with Regression Machine Learning Algorithms for Solar PV systems

P. Venkata Mahesh*‡, S. Meyyappan**, RamakoteswaraRao Alla***

*Department of Electronics and Instrumentation Engineering, Research Scholar, Annamalai University, Tamilnadu, India.
**Department of Electronics and Instrumentation Engineering, Assistant Professor, Annamalai University, Tamilnadu, India.
***Department of EEE, Associate Professor, RVR & JC College of Engineering, Andhra Pradesh, India.

(vnktmahesh@gmail.com, meys.narayan@gmail.com, ramnitkkr@gmail.com)

‡Corresponding Author; P.Venkat Mahesh, Assistant Professor, Department of EEE, RVR & JC College of Engineering, Guntur, Andhra Pradesh, India - 522019. Tel: +91 9542799302, vnktmahesh@gmail.com

Received: 26.07.2022 Accepted: 02.09.2022

Abstract- Solar panels generate energy by utilizing the sun rays on their surface, which depends on the amount of surface temperature and the strength of solar radiation. To maximize the energy conversion efficiency, the solar PV panel should operate at maximum power point (MPP). Each maximum power point tracking (MPPT) method has its unique conversion efficacy and tracking strategy of MPP. This paper describes a novel approach to operating the PV system at MPP by implementing linear and nonlinear regression-type machine learning algorithms. The data acquired from the specifications of the PV panel was used to train, validate, and test the machine learning models. These algorithms predict the maximum available power at the PV panel and its corresponding voltage for specific quantities of irradiation and temperature. These predicted values help determine the boost converter's duty cycle to work the system at the predicted MPP by regression machine learning models. The simulation results show that regression algorithms forced the PV panel to work at the predicted MPP even in the presence of changes in climatic conditions regarding solar radiation and temperature. The PV system has an efficiency higher than 95.21% with linear regression models and more than 95.32% with nonlinear regression models. The simulation results are compared with the beta MPPT and ANN approaches, which were existing MPPT methods in the literature, to show the efficacy of the proposed regression machine learning strategy.

Keywords: Boost converter, Linear and nonlinear regression, Maximum power point tracking, Photovoltaic system, Regression machine learning.

1. Introduction

Solar energy is a feasible remedy to the environmental damage caused by traditional power sources. Temperature and irradiance factors significantly influence photovoltaic module power output, although PV modules produce non-linear P-V curves due to non-uniform irradiance levels. As the PV panels are not operated at maximum power point (MPP) in general and to improve the energy conversion efficiency, it is strongly recommended to run them at MPP. This was achieved through MPPT methods. There are several methods for MPPT in the literature. Each method has its own merits and demerits.

The conventional methods are perturb & observe [1] and incremental conductance [2] methods. Both ways are commonly used due to their simplicity of implementation. However, these methods have limitations such as ambiguity owing to sensitivity to rapid changes in meteorological conditions and oscillations around the operating point. Therefore, mathematical-based techniques like the curve fitting method [3] and beta (β) MPPT [4] have been used for MPP tracking. The curve-fitting is an offline approach that requires PV module characteristics and manufacturing specifications, as well as a mathematical model and equations that describe the output characteristics. The beta MPPT is one of the fastest tracking algorithms since it uses a variable coefficient to do an initial computation of the MPP. Then, the beta value is calculated by infusing it into a conventional feedback system with constant reference using the panel voltage and current values. Along with the MPP, this approach allows faster tracking with fewer oscillations.
Measurement and comparison-based methods like the lookup table method [5] and the current sweep method [6] can be used for MPP tracking. For finding the MPP, the lookup table approach requires and stores prior knowledge of the PV panel type, technical information, and panel features under different natural conditions. The lookup table is developed based on the manufacturer's specifications or by conducting exploratory testing on the PV under different climates. This approach has the disadvantage of requiring a lot of memory for lookup table storage. The current sweep approach uses the PV arrays’ current sweep waveform to determine the desired I-V curve. The sweep is continued indefinitely until the desired MPP is reached for attaining MPP voltage. Despite its complexity, slower speed, and power losses while sweeping, this technique is effective in tracking the true MPP in some cases.

Constant parameter-based methods such as fractional open circuit voltage (FOCV) [7, 8] and fractional short circuit current (FSSC) [8] have been used for MPP tracking. The FOCV-based MPPT takes advantage of the fact that the PV array voltage corresponding to maximum power is linearly associated with the array open circuit voltage at various irradiance and temperature levels. But, the primary drawback of this approach is that the load is detached from the panel in periodic intervals to get the PV voltage. As a result, there is a power loss. Another problem is a significant loss if the sample interval is too long. Furthermore, the FSSC technique, like the FOCV method, requires a short circuit current value to be supplied regularly, which is a disadvantage of the FSSC method since it interrupts power to the loads and reduces the efficiency of MPPT.

Trial and error methods like gradient descent [9] and variable inductor [10] approaches can also track MPP. The gradient descent approach is based on the operating voltage being between zero and the open circuit voltage, and the power change with reference to voltage is dropping functions. This approach can enhance MPPT tracking time under non-linear weather conditions. With the variable inductance technique, the MPP is acquired by impedance matching with a dc-dc converter in continuous conduction mode during steady-state operation. The required variable inductor has a specific inductance compared to the current characteristic. In addition, the inductance decreases when the current increases due to a shift in solar irradiation.

To track the MPP various methods based on optimization are genetic algorithm [11], particle swarm optimization [12, 13], ant colony optimization [14], grey wolf optimization [15], and cuckoo search [16, 17]. A function or parameter will be maximized or minimized using optimization methods. Both steady-state oscillations and tracking speed will be affected by these MPPT methods. The artificial neural network (ANN) [16, 18], fuzzy logic control [1, 16], and adaptive neuro-fuzzy interface system [1, 9] come under intellectual methods. The ANN predicts unknown data from current data by modifying multiple layers’ weight during training. The fuzzy logic control is a rule-based control. A form of ANN is an adaptive neuro-fuzzy interface system [19].

Machine learning algorithms (MLA) will predict the anonymous data if they undergo the process of training, testing, and validating the model with existing data. Typically, training 60% of the data, testing 20% of the data, and the rest are used to validate the machine learning model. The performance of the model can be measured in terms of root mean square error (RMSE), sum squared error (SSE), and R² as follows,

\[ RMSE = \left[ \frac{1}{n} \sum_{k=1}^{n} (Y_{A,k} - Y_{F,k})^2 \right]^{1/2}, \]  
\[ R^2 = 1 - \frac{\sum_{k=1}^{n} (Y_{A,k} - Y_{F,k})^2}{\sum_{k=1}^{n} (Y_{A,k} - Y_{avg})^2}, \]  
\[ SSE = \sum_{k=1}^{n} (Y_{A,k} - Y_{F,k})^2, \]

where, the actual value is \( Y_a \), the predicted value is \( Y_p \) and the samples count is \( n \).

The material used to build a solar panel affects its energy conversion and performance [20]. The PV system has several peaks during partial shading [21, 22]. The power generated by the PV panel majorly depends on irradiance and temperature. As these parameters are not constant and continuously vary with time, the maximum power generated will also vary accordingly. So, it is required to use an algorithm that will force the PV panel to operate at the MPP in less time under these variable constraints. It is possible with the use of MLA for Solar PV MPP tracking. This method will also eliminate the controller requirement for the system. Generally, the MLAs are used for classification and regression. The output of the classification MLA is a label of the data, whereas, on the other hand, the outcome of the regression MLA is a real value. PV systems for MPPT use image-based machine learning [23], a random forest-based approach [24], and reinforcement learning methods [25]. A converter is necessary to operate the PV system at MPP. The literature presents the converters like dc-de boost [2-5], buck [10], buck-boost [7], SEPIC [15], and an inverter [9]. This paper proposed a novel regression-based machine learning for the MPPT of the PV system. The trained regression MLAs were employed for MPP tracking and to operate the PV system at MPP. The data required to train the MLA was prepared from the specifications of the PV panel. The trained MLAs predicted the MPP and the required duty cycle for the converter to work the PV system at this predicted MPP. The mean efficiency was calculated by employing linear and nonlinear MLAs to show the MPP tracking accuracy of the proposed method. A comparative analysis was also carried out for all MLAs.

This research article is structured as: Section-2 describes the system, including the PV panel specifications.
characteristics, boost converter, and regression machine learning algorithms; methodology includes data collection, model preparation, and working of PV panel with machine learning (ML) control strategy has been described in Section-3; simulation results and discussions are provided in Section-4; Section-5 provides the comparative analysis of the proposed method with beta and ANN methods; Section-6 concludes the paper.

2. Description of System

2.1. Solar PV panel & DC-DC boost converter

Sunlight is instantly converted into electrical energy by solar cells. A PV panel is made up of several solar cells that are linked in parallel or series. Figure 1 represents a PV cell’s single diode electrical equivalent circuit [26], and the Eq.(4) gives a mathematical representation.

\[ I = I_{ph} - I_0 \left( e^{\frac{V + I R_s}{n V_T}} - 1 \right) \frac{V + I R_s}{R_{sh}}, \]  

where, the PV panel current \( I \) and the photo or light current \( I_{ph} \) are functions of irradiance \( I \) and temperature \( T \). \( I_0 \) is the diode saturation, \( V \) denotes the PV panel voltage, \( R_s \) symbolises the series resistance, the diode ideal factor is \( n \) (1 \( \leq n \leq 2 \)), thermal equivalent voltage is \( V_T \), and the shunt resistance is \( R_{sh} \).

For a photovoltaic module or array comprising \( N_s \) cells in series, and assuming all cells are identical and under uniform irradiance and temperature (i.e., generate equal current and voltage), module current and voltages are \( I_m=I_{cell} \) and \( V_m=N_s \times V_{cell} \), respectively. The single diode equation for a module or array becomes as Eq.(5).

\[ I_m = I_{ph} - I_0 \left( e^{\frac{V_m + I_m N_s R_s}{n N_s V_T}} - 1 \right) \frac{V_m + I_m N_s R_s}{N_s R_{sh}} \]  

The photocurrent \( (I_{ph}) \) in Eq.(6) is linearly dependent on solar radiation \( (G) \), and impacted by temperature . The reference solar radiation \( (G_{ref}) \) is 1000 W/m².

\[ I_{ph} = [I_{sc} + K_T (T - T_r)] \times \frac{G}{G_{ref}} \]  

where, \( K_T = \frac{I_{sc} (T_0) - I_{sc} (T_r)}{T_0 - T_r} \).

The short circuit current is \( I_{sc} \), \( T \) is real time temperature, and \( T_r (25 \, ^oC) \) is the reference temperature.

The PV cell number in the panel will decide the voltage, current, and power specifications. The PV panel with 10W maximum power, 0.62A short circuit current, 21.50V open circuit voltage, 17.50V as the voltage at MPP, and 0.57A current at MPP with 36 solar cells was simulated in this work. The PV system current and power curves to voltage are shown in Fig.2 and Fig.3, respectively.

A PV panel served pulse width modulation (PWM) controlled dc-dc boost converter is shown in Fig.4. The switch MOSFET’s duty ratio \( (D) \) regulates the supplied power from the panel to the load. An inductor \( (L) \) in the circuit increases the voltage to the appropriate output level. As a result, the voltage ripple content is reduced by the input capacitor \( (C_i) \) and output capacitor \( (C_o) \).

2.2. Regression Machine Learning Algorithms

2.2.1. Linear Regression

The linear Regression algorithm comes under supervised machine learning. It is frequently used to determine the association between variables and for prediction. The regression models vary depending on the correlation between independent and dependent variables and the number of independent variables included. Linear regression provides a solution for predicting the value of an unknown or dependent variable based on a set of independent variables known as features. So, this regression method determines a linear association between input x and output y. If the data contains just one feature, the univariate linear regression method predicts the data in 2-dimensional space with a straight line, as seen in Fig. 5(a). On the other hand, a multivariate linear regression technique generates a plane with multidimensional feature space if the data includes several features. For example, the multivariate linear regression method produces a plane when there are two features, as seen in Fig. 5(b). The linear regression planes [27] take the general form as given in Eq. (7).

1329
where on a $n$-dimensional space, $y$ is the dependent data to be predicted with the features $x_1, x_2, ..., x_{n-1}, x_n$ and the regression coefficients as $\beta_0, \beta_1, ..., \beta_n$, $\beta_n \epsilon (-\infty, +\infty)$.

The model's goal in creating the best-fit regression plane or line is to predict the $y$ value so that the difference between the true and predicted value is minimum. So, it is required to update the $\beta_0, \beta_1, ..., \beta_n$ values to find the optimal value that minimizes the error ($e$), which is the difference between the true $y$ value ($y$) and predicted $y$ value ($y_{pred}$). For $N$ number of samples, the error is calculated using Eq. (8).

$$ e = \frac{1}{N} \sum_{i=1}^{N} (y_{i,\text{pred}} - y_i) $$  

(8)

The cost or objective function ($J$) of linear regression is the RMSE between $y_{pred}$ and true $y$ value as in Eq.(9).

$$ J = \frac{1}{N} \sum_{i=1}^{N} (y_{i,\text{pred}} - y_i)^2 $$  

(9)

To update $\beta_0, \beta_1, ..., \beta_n$ values to reduce $J$, the gradient descent approach is used to achieve the best fit line. The plan is to build with random $\beta_0, \beta_1, ..., \beta_n$ values and then iteratively update the values to get the lowest cost. The cost function’s gradients are used to update $\beta_0, \beta_1, ..., \beta_n$ values.

Fig. 3. Solar panel at 1000 W/m² & mentioned temperatures

a) $V$ vs $I$

$$ y = \beta_0 + \beta_1 x_1 + ... + \beta_{n-1} x_{n-1} + \beta_n x_n $$  

(7)

b) $V$ vs $P$

Fig. 4. The PV connected boost converter

The symbol $\alpha$
represents the learning rate which is a scaling factor for step sizes.

\[
\beta_0(\text{new}) = \beta_0(\text{old}) - \alpha \frac{2}{N} \sum_{i=1}^{N} (y_{i,\text{pred}} - y_i) \\
\beta_i(\text{new}) = \beta_i(\text{old}) - \alpha \frac{2}{N} \sum_{i=1}^{N} (y_{i,\text{pred}} - y_i) (x_{i,i}) \\
\ldots \\
\beta_n(\text{new}) = \beta_n(\text{old}) - \alpha \frac{2}{N} \sum_{i=1}^{N} (y_{i,\text{pred}} - y_i) (x_{n,i})
\]

(11)

Figure (6) represents the schematic procedure of the gradient descent algorithm as a flowchart. The stopping criteria of the algorithm is either the maximum number of iterations or the step size is smaller than the tolerance.

![Flowchart gradient descent algorithm](image)

**Fig. 6.** Flowchart gradient descent algorithm

2.2.2. Non-Linear Regression

The complexity of nonlinear regression is more than that of linear regression. Perhaps the most inconvenient part of the approach is the obligation to provide boundaries and initial values. The convergence of the algorithms is highly dependent on giving enough initial values. Using inappropriate initial values might lead parameter estimates to converge to a local minimum or maximum instead of a global minimum or maximum in some cases. The choice of initial values has a lesser impact on specific models, while it has a more significant impact on others. In certain circumstances, one may be able, to begin with, zeros or ones, but in most cases, better values will be required. One well-established method for obtaining a good set of initial values is to estimate them from data. The coefficients are estimated using iterative least squares estimation with specified initial values. Equation (12) shows the nonlinear regression model used in this work.

\[
y = \beta_0 + \beta_1 x_1^n + \ldots + \beta_{n-1} x_{n-1}^{n-1} + \beta_n x_n^n,
\]

(12)

with y is the data to be predicted; \(x_1, x_2, \ldots, x_{n-1}, x_n\) are the features with \(\beta_0, \beta_1, \ldots, \beta_{n-1}, \beta_n\) as regression coefficients.

The coefficients are estimated using iterative least squares estimation [28], with specified initial values. Then, the residual value \(r\), to be minimized, is calculated using Eq.(13).

\[
r = \sum_{i=1}^{N} \left(\left[\beta_0 + \beta_1 x_{i,1}^{n} + \ldots + \beta_n x_{i,n}^{n}\right] - y_i\right)^2
\]

(13)

3. Methodology

The proposed methodology involves two steps, one is data collection from the PV panel specifications, and the other is preparing the machine learning (ML) model from the collected data. The power at MPP (\(P_{mp}\)) and respective voltage (\(V_{mp}\)) are functions of \(I\) and \(T\) [29]. The \(I\) and \(T\) are features for predicting the data \(P_{mp}\) and \(V_{mp}\). The developed models are used to predict the \(P_{mp}\) and \(V_{mp}\) of the PV panel for the given \(I\) and \(T\). The predicted values are used to determine the duty cycle of the converter to work for the solar panel at MPP.

3.1. Data Collection and model preparation

The data required to train and test the model consists of \(I\), \(T\), \(P_{mp}\), and \(V_{mp}\). This data is collected from Matlab/Simulink software with the help of PV panel specifications. The schematic procedure for data collection and machine learning model preparation is shown as a flowchart in Fig.7.

3.2. MPPT Boost Converter with ML Model

For the features \(I\) and \(T\), the developed machine learning models predict the \(P_{mp}\) and \(V_{mp}\). The resistance \(R_{mp}\) equivalent to MPP was calculated using the predicted values \(P_{mp}\) and \(V_{mp}\) as in Eq.(14). The \(R_{mp}\) will be reflected between node-\(m\), and node-\(n\) in Fig.8 by adjusting the \(D\) of the converter. The \(D\) in terms of \(R_{mp}\) and load resistance \(R_0\) are given in Eq.(15).

\[
R_{mp} = \frac{V^2}{P_{mp}}
\]

(14)
where $V_{mp}$ denotes the input voltage, $V_{op}$ denotes the output voltage, $f_{sw}$ indicates the switching frequency, $\Delta I$ indicates the current ripple, and $\Delta V$ indicates the voltage ripple. The proposed regression ML MPPT control strategy with boost converter is shown in Fig.8 as a block diagram.

**Fig. 8.** Block diagram of regression ML MPPT control strategy with dc-dc boost converter

### 4. Simulation Results and Discussions

The PV panel data has been collected in the proposed manner as described in section-3 with the help of PV panel specifications. Figure 9 shows the pairwise relationship among the data. The correlation heatmap between the parameters was depicted graphically in Fig.10.

The parameters of simulation used in this work are, switching frequency, $f_{sw} = 5$ kHz, PV power, $P = 10$ W, voltage ripple, $\Delta V = 1\%$, current ripple, $\Delta I = 5\%$, boost converter inductance, $L = 34$ mH, capacitance, $C_o = 68 \mu F$, load resistance, $R_0 = 300$ $\Omega$, and input capacitance, $C_i = 1000 \mu F$.

The irradiances and temperature values are varied in the simulation at four intervals of 0.5 seconds with a total of 2 seconds to find the tracking accuracy of the system with the regression models under variable climatic conditions. The values of either $I_o$ or $T$ are varied for each interval retaining the other as constant. For example, during first 0.5 seconds, the $I_o$ is 400 W/m², and $T$ is 25°C. For the second 0.5 seconds the $I_o$ is kept constant as previous (400 W/m²), and the $T$ is raised to 35°C. For the third 0.5 seconds, the $I_o$ is increased to 800 W/m², and the $T$ is constant and is same as previous (35°C). Finally, for the last 0.5 seconds the $I_o$ is kept constant as previous (800 W/m²), and the $T$ is decreased to 25°C.

#### 4.1 Linear Regression (LR) algorithm results

Two LR models are developed with $I_o$ and $T$ as inputs, with the response to be predicted as $P_{mp}$ and $V_{mp}$. The parameters $\beta_0$, $\beta_1$, and $\beta_2$ with standard error (SE) of the developed models are tabulated in Table-1. Figure 11a depicts a minimal residual in prediction on the $P_{mp}$ plane. But Fig. 11b illustrates that for low values of $I_o$ and $T$, the
residual in prediction is significant, while for the remaining, the residual value is small on the \( V_{mp} \) plane. The linear regression between features and \( P_{mp} \) and \( V_{mp} \) is in Fig.12a and Fig.12b, respectively.

Table-2 presents the predicted values of \( P_{mp} \) and \( V_{mp} \) by LR models, including the computed duty cycle and mean efficiency in steady-state. The \( V, I, \) and \( P \) waveforms of the load and solar panel are shown in Fig.13 using LR models. These findings indicate that when the \( T \) is changed, there are mild fluctuations in the transient response and massive amplitude fluctuations when the \( I_r \) is altered. Figure 14 depicts the mean efficiency and compares predicted power and PV working power. It contends that the suggested approach effectively tracks the MPP in the steady-state.

![Fig. 9. Pairwise relationship of the variables](image)

![Fig. 10. Correlation heatmap](image)

![Fig. 12. Linear regression between a) features and \( P_{mp} \) b) features and \( V_{mp} \)](image)

### Table 1. Parameters of the developed LR models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( P_{mp} ) model (LR-1)</th>
<th>( V_{mp} ) model (LR-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>0.96585</td>
<td>0.020004</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.01002</td>
<td>1.6682×10(^{-5})</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-0.03921</td>
<td>0.0008201</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0314</td>
<td>0.202</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>1</td>
<td>0.867</td>
</tr>
</tbody>
</table>

![Fig. 11. Actual vs predicted data a) LR-1 model on \( P_{mp} \) plane b) LR-2 model on \( V_{mp} \) plane](image)

![Table 2. Predicted values by LR models, \( D \), and mean efficiency for all intervals](image)
4.2 Results with Non-Linear Regression (NLR) algorithm

Two NLR models are developed with $I_r$ and $T$ as inputs, and the response to be predicted are $P_{mp}$ and $V_{mp}$. Table 3a & 3b list the parameters of the developed models with standard error. Figure 15 illustrates the actual data and the forecasted data by the developed NLR models. Fig.15a depicts a minor residual in prediction on the $P_{mp}$ plane, whereas Fig.15b depicts that for small values of $I_r$ and $T$, the residual in prediction is significant, while for the others, the residual values are low on the $V_{mp}$ plane.

Table 3a. Parameters of the developed $P_{mp}$ (NLR-1) model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial value</th>
<th>Value</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0</td>
<td>1.5767</td>
<td>0.60167</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0</td>
<td>0.011827</td>
<td>0.00034719</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1</td>
<td>0.97717</td>
<td>0.0040856</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0</td>
<td>-2.2094</td>
<td>0.20945</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1</td>
<td>0.6315</td>
<td>0.18554</td>
</tr>
<tr>
<td>RMSE</td>
<td></td>
<td>0.0206</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table 3b. Parameters of the developed $V_{mp}$ (NLR-2) models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial value</th>
<th>Value</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>0</td>
<td>-5.8718</td>
<td>74.211</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0</td>
<td>10.796</td>
<td>69.351</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>1</td>
<td>0.029518</td>
<td>0.16441</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0</td>
<td>28.775</td>
<td>5.1844</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1</td>
<td>-0.31734</td>
<td>0.47255</td>
</tr>
<tr>
<td>RMSE</td>
<td></td>
<td>0.124</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>0.952</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 15. Actual vs predicted data a) NLR-1 model on $P_{mp}$ plane b) NLR-2 model on $V_{mp}$ plane

Table 4 shows the predicted values of $P_{mp}$ and $V_{mp}$ by NLR models, as well as the calculated duty cycle and mean efficiency in steady-state. Next, the PV panel and load $V$, $I$, and $P$ waveforms are shown in Fig.16 using NLR models. These results demonstrate that when the $T$ is changed, there are fluctuations in the transient response and substantial amplitude oscillations when the $I_r$ is altered. Finally, the mean efficiency was shown in Fig.17, comparing the predicted and PV operative power. From the results, one can be noticed that the suggested approach reliably tracks the MPP in a steady state.

Table 4. Predicted values by NLR models, $D$, and mean efficiency for all intervals

<table>
<thead>
<tr>
<th>Time (sec)</th>
<th>Predicted $P_{mp}$ (W)</th>
<th>Predicted $V_{mp}$ (V)</th>
<th>Duty cycle $D$</th>
<th>% Mean efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 0.5</td>
<td>4.1058</td>
<td>17.3732</td>
<td>0.4995</td>
<td>95.32</td>
</tr>
<tr>
<td>0.5 - 1</td>
<td>3.6165</td>
<td>16.3239</td>
<td>0.5044</td>
<td>95.79</td>
</tr>
<tr>
<td>1 - 1.5</td>
<td>7.6128</td>
<td>16.5903</td>
<td>0.6528</td>
<td>98.45</td>
</tr>
<tr>
<td>1.5 - 2</td>
<td>8.0122</td>
<td>17.6395</td>
<td>0.6402</td>
<td>97.42</td>
</tr>
</tbody>
</table>

5. Comparative analysis

This section compares the proposed LR and NLR machine learning MPPT approach results with the existing methods of $\beta$-MPPT and perceptron artificial neural network (ANN).
5.1 $\beta$-MPPT method

The fundamental idea behind the $\beta$-MPPT [4] approach is to monitor an intermediate variable called $\beta$ instead of the change in power and as follows,

$$\beta = \ln\left(\frac{i_{pv}}{v_{pv}}\right) - C \times v_{pv}$$  \hspace{1cm} (18)

$$C = \frac{q}{NnKT}$$  \hspace{1cm} (19)

where, $i_{pv}$ and $v_{pv}$ stand for the output current and voltage of the PV module, respectively. The diode constant is $C$, $q$ is the electron charge ($1.6 \times 10^{-19}$ coulomb), $n$ is the diode ideal factor, Boltzmann constant is $K$ ($1.38 \times 10^{-23}$ J/K), the pn-junction temperature is $T$ in Kelvin and $N$ is the number of PV cells in the module.

$\Delta D = F \times (\beta - \beta_g)$  \hspace{1cm} (20)

with $F$ as the scaling factor.

The PV module’s temperature and irradiance affect the $\beta$ parameter’s range. In this study, the $\beta_{min} = 15.45$ and $\beta_{max} = 19.02$ are used. The average value of these is used as $\beta_g = 17.24$ with a scaling factor $F = 0.01$.

Figure 19 compares the power response of the $\beta$-MPPT ($P_{\beta}$), the maximum power predicted by LR ($P_{mp}$), and the developed control technique using the LR ($P_{LR}$) and NLR ($P_{NLR}$) algorithms. For low irradiance levels, the $\beta$-MPPT operates at MPP, but for high irradiance values, this technique fails and exhibits a significant amount of error, preventing the PV system from working at MPP. However, the system can be operated at MPP in a steady state with more precision using the proposed regression ML approach compared with $\beta$-MPPT.

5.2 Artificial neural network method

The architecture of the perceptron ANN model is given in Fig 20. This architecture has two inputs, ten hidden layer neurons, two output layer neurons, and two outputs. For the ANN model [16, 18], the two inputs are $I$ and $T$, and the two outputs are $P_{mp}$ and $V_{mp}$. The same data was used and decomposed in 60%, 20%, and 20% to train, validate, and test the ANN model. The activation functions $tansig$ and $purelin$ were used in the hidden layer and the output layer of ANN. This nonlinear and linear collection of operations is
employed to produce practical ANN training. The function of \textit{tansig} is given as \text{Eq.}(21), and the principle to update the weights \((W_{i,j})\) is shown in \text{Eq.}(22). The function of \textit{purelin} is \text{Eq.}(23), and the principle to update the weights \((W_{j,o})\) is in \text{Eq.}(24).

\[
\phi(v_j) = \frac{2}{1 + e^{-v_j}} - 1 \tag{21}
\]

For adjusting the weights among the nodes of the input layer and hidden layer, the delta rule is provided as,

\[
W_{i,j} \leftarrow W_{i,j} + l_i \phi'(v_j) e_j x_i \tag{22}
\]

\[
\phi(v_o) = v_o \tag{23}
\]

with \(x_i\) to be \(i^{th}\) \((i = 1 \& 2)\) input node present value, \(W_{i,j}\) to be the weight among \(i^{th}\) input node and \(j^{th}\) hidden node, the learning rate is \(l_i\) which ranges from 0 to 1, \(j^{th}\) hidden node’s error value is \(e_j\), the weighted sum of inputs and the bias \(b_j\) is \(v_j\), and the \textit{tansig} function’s first derivative is \(\phi'(v_j)\). For adjusting the weights among the nodes of the hidden layer and output layer, the delta rule is provided as,

\[
W_{j,o} \leftarrow W_{j,o} + l_o \phi'(v_o) e_o y_i \tag{24}
\]

with \(y_i\) to be \(j^{th}\) \((j = 1, 2, ..., 10)\) hidden node present value, \(W_{j,o}\) is the weight between \(j^{th}\) output node and \(j^{th}\) hidden node, \(o^{th}\) output node’s error value is \(e_o\), the weighted sum of hidden layer nodes with their bias \(b_o\) is \(v_o\), and the \textit{purelin} function’s first derivative is \(\phi'(v_o)\).

To train, validate, and test the ANN model, the Lavenberg-Marquardt optimization algorithm was used in this study. For updating weights \((W_{i,j}^*\) and \(W_{j,o}^*)\), the rule is provided in \text{Eq.}(25). The training, testing, and validation results in Fig.21 demonstrate that the \(R^2\) value is nearly 1.

\[
W_{i,j}^* = W_{i,j} - [J^T J - l_i I]^{-1} (J^T e) \tag{25}
\]

\[
W_{j,o}^* = W_{j,o} - [J^T J - l_o I]^{-1} (J^T e) \tag{25}
\]

where the Jacobian matrix \(J\), the identity matrix is \(I\), the cumulative error vector is \(e\), the parameter of learning is \(l_p\).

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{fig19}
\caption{Power response comparison of \(P_{mp}\), \(P_{LR}\), \(P_{NLR}\) & \(P_{beta}\).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{fig20}
\caption{ANN model architecture with ten hidden layer nodes}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{fig21}
\caption{Regression plots of training, validation, and testing}
\end{figure}

Figure 22 compares the power response of the ANN \((P_{nn})\), the maximum power predicted by LR \((P_{mp})\), and the developed control technique using the LR \((P_{LR})\) and NLR \((P_{NLR})\) algorithms. If there is a sudden and massive rise in irradiance, the ANN response has an overshoot (observe zoomed portion of Fig.22), and the system operates little far away from the MPP. As shown in Fig.22, as a whole, the proposed method gives a superior response in tracking the MPP compared with the ANN method.
6. Conclusions

The regression machine learning algorithms are applied in this paper for an isolated PV panel to extract maximum power from the PV panel. The developed linear and nonlinear regression algorithms predict available maximum power at the PV panel and respective voltage for a given temperature and irradiance. These predicted values determine the duty cycle required for the boost converter to deliver maximum available power to load by matching the reflection resistance. The mean efficiency is calculated for all four intervals to show the efficacy of the proposed method. In the steady state, the MPPT efficiency is more than 95.21% with the linear and higher than 95.32% with the nonlinear regression machine learning models. The simulation results show better accuracy in tracking the maximum power point on the steady-state under variable climatic conditions with the proposed method compared to the beta MPPT and ANN MPPT methods in the literature.

As a future part of the work, various types of partial shading patterns can be considered for analysing the proposed regression machine learning approach and hardware implementation.

References


[12] Nusaif, Aman Ismael, and Anas Lateef Mahmood, "MPPT Algorithms (PSO, FA, and MFA) for PV System under Partial Shading Condition, Case Study: BTS in


