BAT inspired regression model for prediction of power loss in solar panel

Ritu Maity¹,*

¹ KIIT University, Bhubaneswar, Odisha, India
Email: maity.ritu07@gmail.com
*Corresponding Author: Ritu Maity, Email: maity.ritu07@gmail.com

Abstract

Solar energy is an increasingly popular and environmentally friendly source of renewable energy. The performance of solar panels can be significantly affected by various factors, including shading or shadowing. The shadowing effect on solar panels has been a topic of significant interest and research in the field of solar energy. Shadows cast on solar panels can have a detrimental impact on their performance, affecting their efficiency and power output. Shadows can cause hotspots, voltage drops, and current imbalances, negatively impacting the overall efficiency and output of the solar panel system. In this paper, we have used a bat inspired model to find optimum parameters which are further used in a regression model to predict the amount of power loss that can happen with the area of cells exposed to shadow in a solar panel. This model can help further researchers to know the exact amount of power loss from the shadow effect and accordingly, they can plan to mitigate the issue.

Keywords

Shadowing, Machine learning, Solar panel, Regression, Cell area, BAT algorithm

1. Introduction

In the realm of renewable energy, solar power has come up as a promising solution to meet the increasing global energy needs while lowering carbon emissions. Solar panels, the heart of solar energy systems, harness the sun's abundant energy and convert it into electricity. Solar cells are semiconductor devices that convert sunlight directly into electricity. They have a wide range of applications across various sectors due to their renewable and environmentally friendly nature solar panels are installed on rooftops of houses to generate electricity for power appliances at home [1], also
used in off-grid applications [2], solar-powered water pumps, street lighting, remote sensing and monitoring, power irrigation in agriculture, grid stabilization, energy storage [3,4], etc. However, despite their efficiency and eco-friendly nature, solar panels are not immune to certain challenges that can curtail their performance. One such formidable obstacle is the shadowing effect, a phenomenon that casts a shadow on solar panels, reducing their exposure to direct sunlight. Shadows can be caused by various factors, ranging from natural obstructions like tall buildings, trees, or mountains to human-made barriers like neighboring structures and debris. This shadowing effect induces a chain reaction of adverse consequences that lead to power loss in solar panel systems. As shadows obscure portions of the panel's surface, they significantly hinder the incoming sunlight from reaching the solar cells. This disruption affects both the photovoltaic (PV) cells and the system's overall output, leading to a decline in energy production and operational efficiency. It is important to have an intelligent system which can predict the power loss that happens from area occupied by shadow. Here we have tried to use bio inspired bat algorithm which is based on swarm intelligence technique to find the best and optimized parameters which can be further used in regression model to predict the power loss.

2. Literature Review

Solar panels are designed to convert light received from the sun into a constant current. This method has gained popularity due to its adaptability in different environments like deserts, mountains, and coastal areas [5]. These systems started to be produced and implemented in a less than straightforward manner. Solar radiation [6], atmospheric temperature [7], relative humidity [8], wind [9], dust [10,11], and shadow [12] all have a significant impact on photovoltaic cells since they operate in the open air exposed to a variety of air conditions. One of the major challenge faced by solar cells is on cloudy days due to shadowing the power generation get affected. The amount of electric energy produced by a collection of solar cells connected depends significantly on the temperature of the cells and the amount of solar radiation present at a specific location and time. The solar radiation intensity falling on PV cells and the cell temperature are both directly impacted by the momentary shade [13]. Due to the high starting costs of high solar cell systems, the sun energy must be as close to the panels as feasible to create the best electrical efficiency possible from the solar cell [14]. The shade creates possible safety risks in addition to reducing the amount of electricity generated [15]. An ideal comprehension of how shade affects solar cell system performance would lead to better system design.
and higher electrical efficiency. The efficiency of the system typically varies and depends on the weather because solar cells function under different conditions. PV array shading limits power output and creates a safety hazard [16]. Multiple maxima in the PV and IV curves under partially shadowed conditions confound the properties of the PV array. Multiple maxima cause the PV efficacy at maximum power point tracking (MPPT) systems to decrease, which is an issue [17]. This scenario has the disadvantage of making it impossible to distinguish between local and global peaks, which reduces output power. It is essential to comprehend how shading affects a PV array's performance because doing so can help to improve its design and efficiency [18]. Numerous studies have examined the features of PV arrays over time, as well as the impact of various operational and design variables [19,20]. Wei He studied the safety factors concerning the effect of partial shading on solar modules [21]. B. A. Alsayid [22] discussed the partial shading effect on solar cells using simulation and experimental results. M. T. Chaichan [23] has discussed the effect of environmental conditions on concentrated solar systems in deserted weather conditions. The shading effect can lead to hot spot conditions, few of the research papers concentrate on the investigation of hot spots in solar modules [24]. Solar panel monitoring is also an area of concern in the current era. Fahad Saleh M. Abdallah [25] proposed an intelligent solar panel monitoring system which is IoT based module and shading detection has been performed using an artificial neural network model. Shoaib Kamal [26] proposed optimization for solar panels using machine learning. S. Rao [27] discussed machine learning methods for monitoring, optimization, and control of solar panels. Suresh Kumar Sudabattula[28] proposed optimal allocation of solar based distributed generators using bat algorithm. Xin-She Yang[29] had proposed bat algorithm its application in various domain and its advantages over other swarm intelligence techniques. Various cutting-edge technology methods are being used to monitor the parameters of solar panels so that maximum efficiency can be achieved from the performance of solar cells. A lot of research work has been conducted in this domain but different research works have taken into consideration some of the parameters for finding the effect of shadowing on solar panels. Here we have tried to propose a machine learning model which can predict power loss in solar panels depending on the number of cells exposed to shadow which can help in advance to plan methods.
3. Solar Panels

Photovoltaic (PV) panels, sometimes referred to as solar panels, are gadgets made to convert sunlight directly into electricity. They are a key technology in the field of renewable energy and play a crucial role in generating clean and sustainable power. Solar panels are typically flat, rectangular devices consisting of multiple solar cells. These cells are usually made from semiconductor materials, such as silicon, that can generate an electric current when exposed to sunlight. They are in charge of using the photovoltaic effect to transform sunshine into power. In a solar cell, exposure to sunlight stimulates the semiconductor material, causing the electrons to flow and create an electric current. There are typically three types of solar panels [30].

i) Monocrystalline: These panels are made from a single crystal structure, resulting in high efficiency and a sleek black appearance. They tend to perform well in low-light conditions and have a longer lifespan.

ii) Polycrystalline: These panels are made from multiple crystal structures, making them slightly less efficient than monocrystalline panels but also more affordable to produce.

iii) Thin-Film: These panels are made by depositing thin layers of photovoltaic material onto a substrate, resulting in flexible and lightweight panels. While they are generally less efficient than crystalline panels, they are suitable for certain applications like curved surfaces or portable solar devices.

An off-grid solar system, also known as a standalone solar system, is a self-contained power generation and storage system that operates independently of the electrical grid. It's designed to provide electricity to locations that are not connected to the utility grid, such as remote cabins, rural areas, or even in emergencies. These systems use solar panels to capture sunlight and convert it into electricity. The generated electricity is in direct current (DC) form. Batteries are a crucial component of off-grid systems, as they store excess energy generated by the solar panels during sunny periods for use when the sun is not shining. They provide a continuous power supply, especially during night time or cloudy days. Our predictive model can help in predicting power output when solar cell is subjected to shadow effect.
4. Algorithm Description

4.1. Bat Algorithm

Bat algorithm is a bio inspired model used for solving continuous constrained problems and it can provide excellent convergence to global optimal solution. It was proposed by Yang, 2010 which takes the benefit of ability of bat to find its prey using echolocation signal. Bats can determine location of its prey by their type of movement in the echolocation region of the bat. Bats use their echolocation skills to determine distance, and they can also distinguish between insects and background obstructions. Bats always fly randomly with a particular velocity (Ui) at a place (Di) with a fixed frequency (Fmin) but variable wavelength (λ) and loudness (L) to hunt prey. The loudness of the bat's pulse might vary. The following equation 1 to 3 give information about velocity and frequency of pulse taking loudness into consideration (31)

\[
F_t = F_{\text{min}} + (F_{\text{max}} - F_{\text{min}}) \times r \quad (1)
\]

\[
U_t^t = U_t^{t-1} + (D_t^t - D_t) \times F_t \quad (2)
\]

\[
D_t^t = D_t^{t-1} + U_t^t \quad (3)
\]

Where Di is bat position in search space, Ui indicates velocity of bat, Ft indicates frequency, r represents vector of random numbers between 0 and 1. Each bat is assigned with random frequency which varies from Fmin to Fmax. As a result, the bat method may be thought of as a frequency-tuning algorithm that provides a balanced combination of exploration and exploitation. The loudness and pulse emission rates serve as a technique for automatic control and auto-zooming into the region with promising solutions. In order to control the exploration and exploitation rate we need to vary loudness and pulse emission rate by using following equations

\[
L_{t+1}^t = \alpha \times L_t^t \quad (4)
\]

\[
P_{t+1}^t = P_0^t \times [1 - \exp(\beta t)] \quad (5)
\]

Where \( L_t^t \) represents loudness at time t, \( P_t^{t+1} \) represents rate of emission at time t+1, \( \alpha \) and \( \beta \) are the constant values.

4.1.1 Steps involved in BAT algorithm

i) Initialize a population of bats. Each bat represents a potential solution to the optimization problem.

ii) Assign random positions and velocities to the bats. Define the frequency and
loudness of each bat, which are used in the echolocation process.

iii) Evaluate the fitness of each bat in the population. The fitness function should be
defined based on the specific optimization problem you are trying to solve.

iv) Echolocation: For each bat update the frequency (F) and loudness (L) of the bat.
These parameters control the search behaviour. Generate a random solution in the
vicinity of the current bat's position. Evaluate the fitness of the new solution. If the
new solution has better fitness than the current bat's solution, update the bat's position
to the new solution.

If not, with a certain probability (controlled by loudness), adjust the frequency and
move towards the best solution found by any bat in the population.

v) Update Global Best: Keep track of the bat with the best fitness value found so far.
This bat represents the global best solution.

vi) Repeat the echolocation and update steps for a predefined number of iterations or
until a convergence criterion is met. Return the best solution found, which is the bat
with the highest fitness value in the population.

4.2 Regression Model

When establishing a relationship between one or more independent variables (also
known as features or predictors) and a dependent variable (also known as the target
or result), machine learning practitioners employ regression models. Regression
analysis seeks to develop a mathematical formula that, given the values of the
independent variables, can precisely predict the value of the dependent variable.
[32,33].

4.2.1 Linear Regression: This is one of the simplest forms of regression. It
presume that the independent variables and the dependent variable have a linear
relationship. The model seeks to minimize the difference between the anticipated
values and the actual values by fitting a linear equation to the data points.

4.2.2 Polynomial Regression: It uses polynomial equation to perform
relationship between dependent and independent variables. It's useful when the data
doesn't fit a straight line.
4.2.3 Ridge Regression and Lasso Regression: These are variations of linear regression that incorporate regularization techniques to prevent overfitting. They add penalty terms to the regression equation, which helps in controlling the complexity of the model.

4.2.4 Decision Tree Regressor: Decision Tree model that is designed to predict continuous numeric values, as opposed to categorical classes in classification tasks. Decision Trees are versatile and intuitive models that recursively partition the feature space into regions and make predictions based on the average (or some other measure) of the target values within those regions. It can handle non-linearity in the dataset.

4.2.5 Random Forest Regressor: An ensemble learning approach called a Random Forest Regressor develops and integrates various decision trees to provide a more reliable and precise regression model. It is an extension of the Decision Tree Regressor and aims to address some of its limitations, such as overfitting and instability.

5. Method

A data set of 500 samples which consists of the angle of inclination in solar panels, cell area, voltage, and current. As illustrated in Figure 1, which is a setup of a halogen lamp, power supply, and solar panel of 20W, we have gathered data from our setup in the solar lab from the solar technology trainer kit. A voltmeter and an ammeter that are connected to the setup were used to measure the voltage and current produced by the solar panel as it is kept horizontal to the halogen lamp. Data was collected by covering the portions of the solar cell in a parallel and series manner to record the change in voltage and current as the shading of the cell area changes. The dataset was gathered over the course of about a month under shade conditions that varied significantly from the experimental setup.
Once the data was collected it was compiled in the form of a data frame as shown in Figure 4 and data cleaning and data analysis steps were performed to prepare the
dataset further for applying the predictive model. We aimed to predict the power output of the solar panel as the angle of inclination of the solar panel varies and the shadowing effect on the same. To get the power output we tried to predict voltage using a machine learning regression model where in input we had taken the angle of inclination for solar panel and cell area having shadow effect and output will be the voltage. But in order to select best suited coefficients for our selected parameters we have used bio inspired meta heuristic optimization algorithm to find the global best solution for the coefficients to be used in regression model which can give us the best prediction values. After getting the voltage we can calculate the power loss in each combination of angle of the solar cell and cell area under the shadow effect.

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (y - yp)^2
\]  

(6)

Figure 3: Figure shows the data set collected from the solar panel experimental setup

Here the objective function was taken as mean squared error which need to be minimized
The parameters of BAT algorithm i.e., population size, frequency, loudness of bat need to be initialized and convergence criteria need to be set. Fitness function is evaluated with each bat solution by applying regression model. Using python module we have imported pybatopt package which has bat algorithm function. We have built a fitness function, parameters of which are then passed to bat algorithm function to give the best optimized parameters after 80 iterations based on minimizing the mean squared error and enhancing the R squared value. The flow chart below shows the steps in involved in bat algorithm based regression model.

```
Data collection, data cleaning, preprocessing, feature selection
  ↓
Split the dataset into training and testing
  ↓
Initialize bat algorithm parameters and define fitness function
  ↓
Identify the bat with best fitness function
  ↓
Update the position and frequency of bat and extract the optimized hyper parameters
  ↓
Train the regression model using the optimized hyper parameters
  ↓
Model evaluation using Mean squared error and R squared value
```

Figure 4: Flow diagram for steps involved in bat algorithm based regression model

Then we have tried to apply different regression models after getting optimized parameters from bat algorithm i.e., linear regression, polynomial regression, lasso regression, and ridge regression models were applied to the dataset. By applying various regression models predictions were generated and the accuracy of the model was checked by calculating the mean squared error and R squared error.
6. Results and Discussion

Once the data set was made available by applying data cleaning and data analysis process bat algorithm was applied to generate best hyperparameters taking those values the machine learning models were applied. The accuracy of the models was checked by evaluating MSE and R2 Score results which have been tabulated in the table given below.

<table>
<thead>
<tr>
<th>Model Used</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression</td>
<td>85%</td>
</tr>
<tr>
<td>Lasso Regression</td>
<td>87%</td>
</tr>
<tr>
<td>Ridge Regression</td>
<td>87.8%</td>
</tr>
<tr>
<td>Polynomial Regression</td>
<td>95%</td>
</tr>
<tr>
<td>Decision Tree Regressor</td>
<td>93%</td>
</tr>
<tr>
<td>RandomForestRegressor</td>
<td>96%</td>
</tr>
</tbody>
</table>

From the above results, we can conclude that the random forest regressor gave the best results and can be used as a predictive model for the prediction of power output in solar panels. By performing the bat algorithm based regression analysis on the dataset we can further also conclude that an angle of inclination of 30 degrees to 45 degrees gave the best result and when the exposed cell area increases we got maximum voltage and power output. The graph in Figures 5 and 6 below shows how the angle of inclination of solar panel and voltage obtained varies as well as the relationship between exposed cell area to light and voltage generation.

Figure5: Figure shows the graph between the angle of inclination of the solar panel and the voltage
7. CONCLUSION

Photovoltaic cells are subject to a range of environmental factors when they are outside or above building roofs, including the shadows cast by surrounding buildings, clouds, and dust. This research examined the impact of shade on the variations in the power output by using an experimental setup in the solar lab to propose and predictive model which can help people further to estimate the power loss if the area of the cell of the solar panel under shade is known. In order to get the best results from the predictive model a bio inspired meta heuristic algorithm ie BAT algorithm was used to generate optimal hyper parameters which can be applied in the regression models to predict the voltage output and made a comparison based on their accuracy achieved. Finally, it was found the random forest method gave the best result which can help in the effective prediction of power output. The results generated showed that as the shaded area increases power output decreases and the best inclination angle for solar cells was found to be 30 degrees to 45 degrees using a predictive machine learning model.

References


