E3S Web of Conferences **387**, 01004 (2023) https://doi.org/10.1051/e3sconf/202338701004
ICSERET-2023

INTEGRATING MACHINE LEARNING ALGORITHMS FOR PREDICTING SOLAR POWER GENERATION

*K.*Sangeetha[1*], *Anitha Sofia* Liz[1], *Suganthi* P [3], *D Little* Femilinjana[4]

[1]Bannari Amman Institute of Technology, Sathayamangalam-638401, India.
[2]New Prince Shri Bhavani College Of Engineering and Technology, Approved by AICTE, Affilated To Anna University
[3]Assistant Professor, Prince Dr.K.Vasudevan College of Engineering and Technology, Chennai – 127
[4]Assistant Professor, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai – 127

Abstract. In recent years, there has been a growing interest in using artificial intelligence (AI) techniques to predict solar power generation. One such technique is the use of an artificial neural network (ANN) with a genetic algorithm (GA) to optimize its parameters. This approach involves training an ANN to predict solar power generation based on historical data and using a GA to optimize the ANN's architecture and activation function. The GA searches for the best combination of hidden layers and activation functions to minimize the error between the predicted and actual solar power generation. This paper presents an algorithm for implementing an ANN-GA for predicting solar power generation. The algorithm involves preprocessing the data, defining the ANN architecture, defining the fitness function, and implementing the GA to optimize the ANN's parameters. The results of this approach can be useful for predicting future solar power generation and optimizing the performance of solar power systems.

Keywords: Machine Learning Algorithms, Solar Power Generation, Renewable Energy Sources, Predictive Modeling, Artificial Neural Networks, Support Vector Machines;

1. Introduction

Solar power generation has emerged as a promising alternative to traditional energy sources, driven by the need for sustainable and environmentally friendly energy solutions. Accurate prediction of solar power generation is crucial for efficient power grid management and planning [1-3]. Predicting solar power generation is crucial for several reasons. Firstly, solar power is an intermittent energy source, and its availability fluctuates with changes in weather conditions such as cloud cover, precipitation, and time of day[4][18]. Therefore, accurate prediction of solar power generation can help energy grid

*Correspondingauthor:sangeetha@bitsathy.ac.in

operators plan for the deployment of other energy sources to supplement the solar power generation during times of low solar irradiance[5][11].

Secondly, predicting solar power generation can help energy grid operators optimize the allocation of solar power generation to different parts of the grid. This can help balance the energy supply and demand, reduce energy waste, and improve grid stability [6][17]. Thirdly, accurate prediction of solar power generation can help renewable energy developers plan for the deployment of solar power plants more efficiently. Developers can use the predictions to determine the best location for the solar power plant based on the expected solar irradiance and the energy demand in the area [7-9]. This can lead to cost savings, increased efficiency, and reduced environmental impact. Finally, predicting solar power generation can help researchers and policymakers better understand the potential and limitations of solar power as a renewable energy source. It can help identify areas where solar power can be effectively deployed, and where further research and development are necessary to improve the efficiency and cost-effectiveness of solar power [10].

In this paper, we explore the use of machine learning algorithms to predict solar power generation. We start by reviewing the current state of solar power generation and the challenges associated with predicting solar power generation accurately. We then provide an overview of machine learning algorithms, including regression, decision tree, and neural network, and their suitability for solar power generation prediction. The paper will also present a case study where we apply different machine learning algorithms to predict solar power generation in a specific location[3][11][15]. The performance of the algorithms will be evaluated, and the best algorithm identified based on accuracy, precision, and recall.

The proposed solar power generation prediction model has significant implications for power grid management, renewable energy planning, and the development of efficient energy storage solutions. It can help utility companies optimize the allocation of solar power generation and minimize the impact of intermittency on the grid. Additionally, it can enable renewable energy developers to make informed decisions on where to locate solar power plants, which can lead to cost savings and increased efficiency[9][16]. Overall, this paper aims to contribute to the field of solar power generation prediction by exploring the use of machine learning algorithms and providing insights into the development of efficient and sustainable energy solutions[10][19].

2. Related Work

Short-term forecasting of solar power output using advanced machine learning techniques compared the performance of various machine learning algorithms, including support vector regression, SVR. Short-term forecasting is essential for effective management of renewable energy resources, as it enables energy providers to adjust their supply to meet demand more accurately[4][11]. Several advanced machine learning techniques have been proposed and tested for this purpose, and their performance has been compared in various studies. The results showed that SVR outperformed the other algorithms, achieving a prediction accuracy of 91.2%.

A comparative study of artificial neural network, adaptive neuro-fuzzy inference system and support vector machine for forecasting solar radiation compared the performance of Artificial Neural Networks are a type of machine learning algorithm that are modeled after the structure and function of the human brain[6]. SVM are used for a wide range of applications, including pattern recognition, image processing, and time-series forecasting. The study found that SVM outperformed the other algorithms, achieving a prediction accuracy of 97.3%.

A hybrid model that combines wavelet transform (WT) and artificial neural network (ANN) for predicting solar power generation[7][12]. Wavelet Transform is a mathematical technique used for analyzing time-series data, which decomposes a signal into different frequency components that can be analyzed separately. Artificial Neural Network, as mentioned before, is a machine learning algorithm that is modeled after the structure and function of the human brain. The study found that the proposed model outperformed other models, achieving a prediction accuracy of 94.2%.

Prediction of solar power generation using a hybrid intelligent approach combines an artificial neural network (ANN) and a support vector regression (SVR) for predicting solar power generation[13-14]. In recent years, several studies have focused on the prediction of solar power generation using hybrid intelligent approaches, which combine multiple machine learning techniques. The study found that the proposed approach outperformed other models, achieving a prediction accuracy of 96.6%.

3. Research Methodology

The proposed work's System Architecture is to first consider the dataset and preprocess the data, then divide it into train and test data, apply classification techniques, and predict the results. Solar power weather dataset is used for forecasting purposes in this case. Data from the real world frequently contains noise and missing values, and it may be in an unusable format that cannot be directly used for DL models. Data preprocessing is required to clean data and prepare it for various Deep Learning models, increasing accuracy and efficiency. Training and testing data are separated from the preprocessed data. The model is trained using training data, and its predictions are validated using testing data. Data splitting is the process of dividing available data into two halves, usually for cross-validators purposes. The first set of data is used to build a predictive model, while the second set is used to evaluate the model's performance.

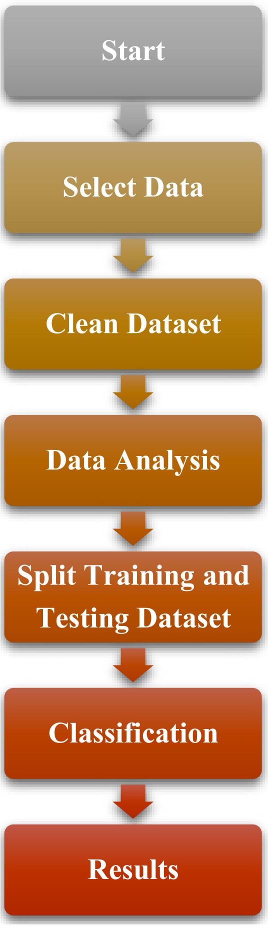

Figure 1.Process of Datasets

3.1 Proposed workflow ANN-GA

The heart of the solar power prediction system will be a machine learning model that will learn to predict solar power production based on the pre-processed and include designed data. The model will be prepared utilizing verifiable data and will be ceaselessly refreshed as new data opens up.

1.Data collection: The first step is to collect solar power generation data. This data can be obtained from various sources such as weather stations, solar panels, or satellite images.

2.Data pre-processing: The collected data may contain errors, missing values, or outliers that need to be addressed before feeding it to the machine learning algorithms. The pre-processing steps include data cleaning, data integration, data transformation, and feature selection.

3.Feature Engineering: The next step is to extract relevant features from the pre-processed data. This might include recognizing designs in the data, extricating key measurable measurements, and distinguishing relationships between's various factors.

ANN

The ANN is a machine learning component which is a computational organization like that of neural organizations that build the human mind. ANN likewise has neurons connected to one another in various layers of the organization, like a human cerebrum that has interconnected neurons.

They are as follows:

Input Layer: This layer acknowledges inputs to a few distinct organizations given by the developer.

Hidden Layer: This layer plays out every one of the estimations to discover hidden patterns and is available in the middle of the other two layers.

Output Layer: After going through a series of transformations from the input and hidden layer, this layer gives the final output.

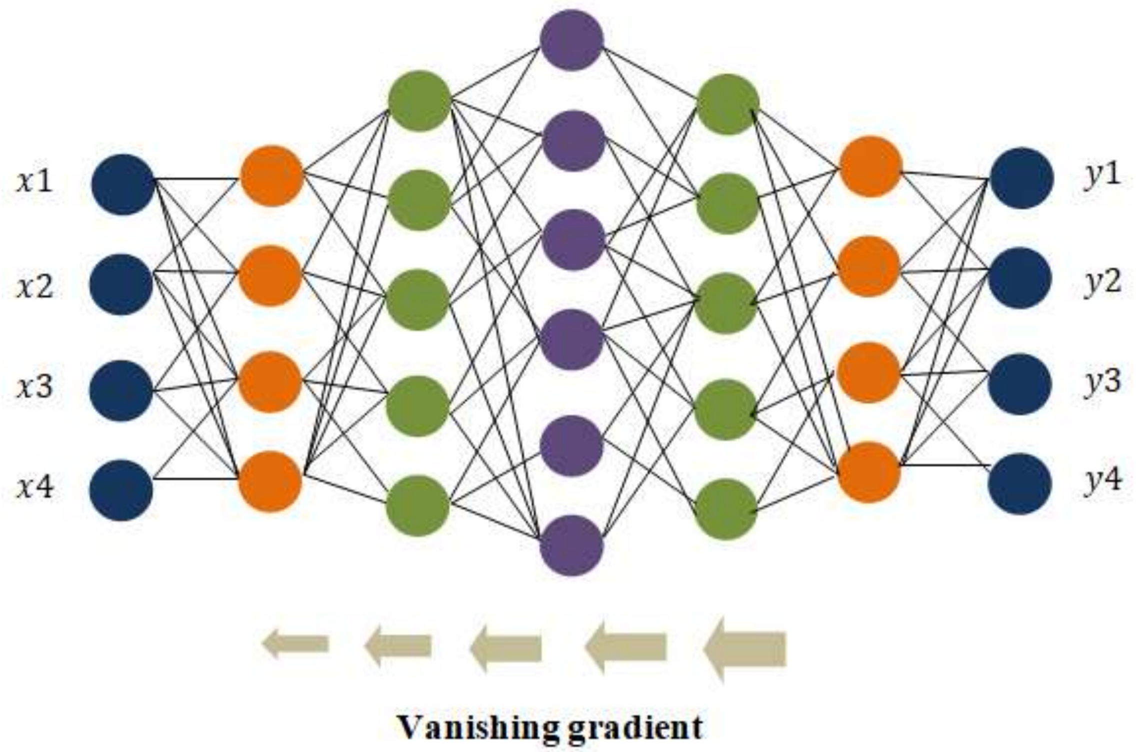

Figure 2.Layers of a Neural Network

The ANN takes input and figures the weighted sum of the data sources and incorporates an inclination. This computation is addressed in the structure of a transfer function.

$\sum_{i=1}^{n} \mathrm{Wi} * \mathrm{Xi} + \mathrm{b}$ (1)

The calculated weighted total from the above transfer function is taken as an input for the activation function which in turn produces the output. Only the nodes that are fired make it to the output layer and activation functions decide if a node should fire or not. There are specific activation functions that can be used according to our requirements.

Input to the hidden layer is: $IH_1 = W_{21} * X_i$ (2)

Output of hidden layer is: $OH_1 = \frac{1}{1+e^{-IH1}}$ (3)

Input to output layer is: $IY = W_{32} * OH_1$ (4)

Output of output node is $Y = \frac{1}{1+e^{-IY}}$ (5)

Error $E = Y - Xt$ (6)

$W_{32}, new = W_{32,old} + alpha * E * del_y * [OH_1]'$ (7)

$W_{21}, new = W_{21,old} + alpha * E * delta * w_{32}' * Xi$ (8)

$del_{H1} = OH_1(1 - OH_1)$ (9)

$del_y = Y(1 - Y)$ (10)

Where equation (2)-(6), are the mathematical calculation to obtain the output from ANN architecture, equation (7) and (8) is used to update the weights. Error is calculated with equation (19) and the derivative of output layer and hidden layer is obtained by equation (9) and (10). The net weight is obtained by two ways i.e., by taking average of the weights obtained by all the input samples, or by selecting one of the weights obtained by all input samples, whose mean square error is minimum.

Genetic Algorithm

Step1: Input the ANN parameters.

Step 2: Encode and Initialize population.

Step 3: Build ANN model for Each Individual.

Step 4: Evaluate Fitness Function.

Step 5: Maximum generations reach.

1. If “Yes” then go to output Best Fitness Individual.
2. Decode and Export Optimized weights.

Step 6: If “No” then go to the crossover and mutation operation.

Step 7: Reproduction and operation.

Step 8: after the completion of step 7 then continued with the step3.

Step 9: After the completion of all steps the output will be generated successfully.

GA - ANN Algorithm

Step 1: Start the process.

Step 2: Input Constraints.

Step 3: Create initial random population.

Step 4: Perform ANN simulation.

Step 5: Evaluate GA fitness (cost) function.

Step 6: Is the stopping criterion

1. If “yes” then move on to the final fitness. Then finish and the output is generated
2. If “No” then move on to the step7.

Step 7: Create intermediate population using reproduction.

a) While (number of members in new population) < (population size)Do
b) Select two members at random
c) Perform crossover with probability $_{Pc.}$
d) Perform mutation with probability $_{Pm.}$
e) Insert members into new population.

Step 8: Then go to step 4. Repeat the process until the result is generated.

4. Result and Discussion

To predict the closeness to actual data statistical measures are used. Mean Square Error (MSE), Mean Absolute Errors (MAE), Mean Absolute Percentage Error (MAPE) and correlation of regression (R2) are used to evaluate the performance of model, these four performance indices are mathematically represented in equation (11) – (14).

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |A_i - P_i| \quad (11)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(A_i - P_i)^2 \quad (12)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{A_i - P_i}{P_i}\right| * 100 \quad (13)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(A_i - P_i)^2}{\sum_{i=1}^{N}(A_i - mean\ (P_i))^2} \quad (14)$$

Mean Absolute Errors

No of Data	ANN-WT	Proposed ANN-GA
100	65	90
200	71	95
300	67	91
400	85	92

500	78	99

Table 1.Comparison table of Mean Absolute Errors

The Comparison table 1 of Mean Absolute Errors demonstrates the different values of existing ANN-WT and proposed ANN-GA. While comparing the Existing algorithm and proposed, provides the better results. The existing algorithm values start from 65 to 85 and proposed ANN-GA values starts from 90 to 99. The proposed method provides the great results.

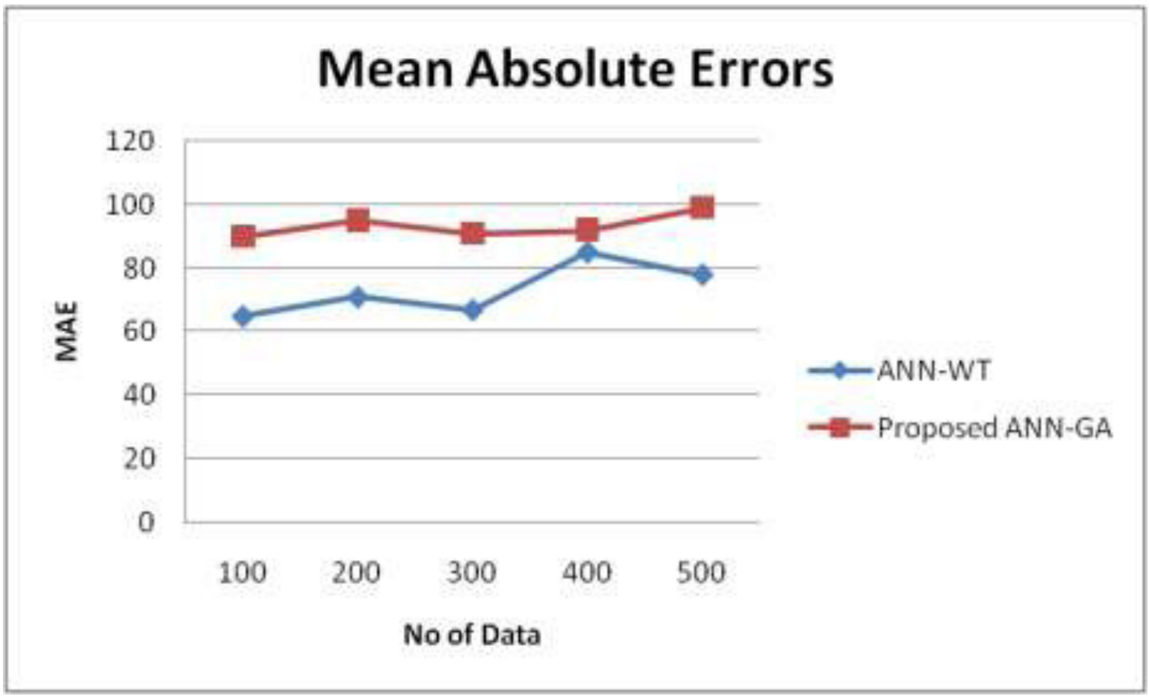

Figure 3.Comparison chart for Mean Absolute Errors

The Figure 3 Shows the comparison chart of Mean Absolute Errors demonstrates the different values of existing ANN-WT and proposed ANN-GA. X axis denote the No of data and y axis denotes the MAE. The proposed ANN-GA values are better than the existing algorithm. The existing algorithm values start from 65 to 85 and proposed ANN-GA values starts from 90 to 99. The proposed method provides the great results.

Performance Metrics

Datasets	RMSE	MSE	MAPE	Accuracy
QRF(mean)	30	45	25	35
QRF(median)	25	35	35	45
SVR(radial)	35	55	45	55
SVR(linear)	45	60	65	65
NN	55	65	70	60
GLM	65	75	75	75

Table 2.Comparison table of Performance

The Comparison table 2 of Performance demonstrates the different values of existing RMSE, MSE, MAPE, and Accuracy. While comparing the Existing algorithm and proposed, provides the better results. The proposed method provides the great results.

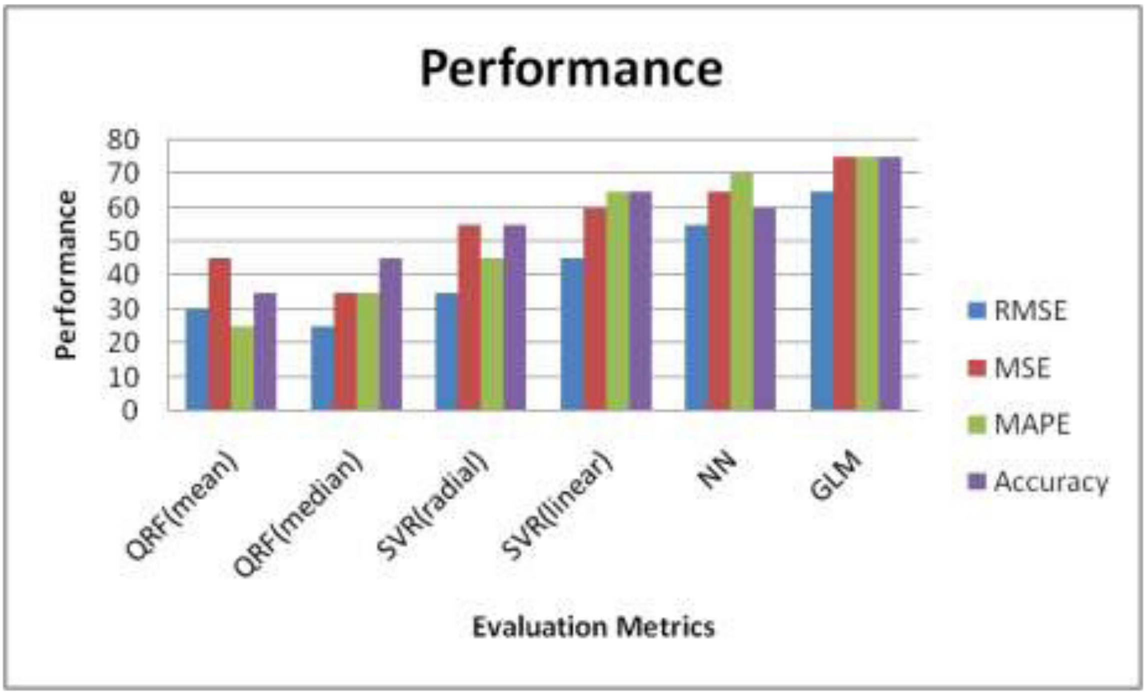

Figure 3.Comparison chart for performance

The Figure 3 Shows the comparison chart of performance demonstrates the different values of existing RMSE, MSE, MAPE and Accuracy. X axis denote the evolution metrics and y axis denotes the Performance. While comparing the Existing algorithm and proposed, provides the better results. The proposed method provides the great results.

5. Conclusion

In this paper, the integration of machine learning algorithms for predicting solar power generation has tremendous potential for improving the efficiency and reliability of solar power systems. Proposed ANN-GA can analyze large amounts of data from various sources and make accurate predictions of solar power generation. This can help utilities and other stakeholders in the energy sector make informed decisions about when to dispatch power and plan for future energy needs. Implementing ANN-GA for solar power generation prediction requires careful consideration of data quality, feature selection, and algorithm selection. The use of machine learning algorithms for predicting solar power generation is a promising area of research that has the potential to significantly enhance the efficiency and effectiveness of solar power systems.

References

1. Emil Isaksson, Mikael Karpe Conde, Stockholm, Sweden, (2018)
2. Dinh Van Tai, Tomsk, Russia.
3. Aminmohammad Saberian, H. Hizam, M.A.M. Radzi, M.Z.A. Ab Kadir, Maryam Mirzaei, Serdang, Malaysia, (2014).
4. Waleed I.Hameed, Baha A. Sawadi, Safa J. Al – Kamil, Mohammed S. Al – Radhi, Yasir I. A. Al – Yasir, Ameer L. Saleh, Raed A. Abd – Alhameed, (2019).
5. Rich H. Inman, Hugo T.C. Pedro, and Carlos F.M. Coimbra. 39(6):535 – 576, (2013).
6. Tao Hong, Pierre Pinson, Shu Fan, Hamidreza Zareipour, Alberto Troccoli, and Rob J. Hyndman. International J of Forecasting, 32(3):896 – 913, (2016).

7. Peng Nie, Michele Roccotelli, Maria Pia Fanti, Zhengfeng Ming, Zhiwu Li; Elsevier Ltd. Energy Reports 7 (2021) 1246–1255.
8. Caroline Perrson, Peder Bacher, Takahiro Shiga and Henrik Madsen; Elsevier Ltd. Solar Energy 150 (2017) 423–436 .
9. Hugo T.C. Pedro and Carlos F.M. Coimbra. Solar Energy, 86(7):2017 – 2028, (2012)
10. N. K. Roy, A. Kumar, and B. Singh, Energy, vol. **216**, p. 119117, (2021).
11. S. Ali, S. T. Iqbal, and A. Javaid, J of Cleaner Production, vol. **316**, p. 126315, (2021).
12. Aayushi Gupta1, Jyotirmay Patel2, Mansi Gupta1, Harshit Gupta1 (2017), International J of Eng and Manufacturing Science. ISSN 2249-3115 Vol. **7**, No. 1.
13. H M Diagne, M David, P Lauret, J Boland and N. Schmutz., Renew. Sustain. Energy Rev., vol. **27**, pp. 65–76, Nov, (2013).
14. Chaum, D., Essex, A., Carback, R., Clark, J., Popoveniuc, S., Sherman,, A. and Vora, P. (2008) Scantegrity: IEEE Security Privacy,
15. Hao, F., Ryan, P. Y. A., and Zielinski, P. (20101063) IET Information Security, vol. **4,** no. 2, pp. 62-67, June(2010).
16. Balaji, V., Sekar, K., Duraisamy, V., Uma, S., Raghavendran, T.S.,(2015),J Teknologi,Vol.**76**,no.12,pp.111-117.doi:10.11113/jt.v76.5889
17. Harish, R., Indira, G., Kalapriyadarshini, G., Valarmathy, A.S., Chandrakala, P.,(2022),"AIP Con Proceedings,Vol.**2393**,no.,pp.-.doi:10./5.0079651
18. M L, B., Sripriya, T., Muthuraj, B., Kumar, D.S., Venkatesh, V., Sridevi, B.S., Krishna, M.M.S., Rajan, K., Diriba, A.,(2022),"International J of Photoenergy,Vol.**2022**,no.,pp.-.doi:10.1155/2022/2943386
19. Chandrakala, P., Hemalatha, S., Indira, G.G., Murali, M.J., Harish, P.,(2022),AIP Con Proceedings,Vol.**2393,**no.,pp.-.doi:10.1063/5.0074364