Machine Learning for Solar Power Forecasting

Dr. C. V. Pavithra[1], Ms. R. Divya[2], Ms. K. Bavithra[3], Ms. A. Jeyashree[4]
[1,2,3,4]Assistant Professor, Department of EEE, PSG Institute of Technology and Applied Research, India
pavithra@psgitech.ac.in, divya@psgitech.ac.in, bavithra@psgitech.ac.in, jeyashreea@psgitech.ac.in

Article Info
Page Number: 1574 - 1591
Publication Issue:
Vol 71 No. 3 (2022)

Abstract
The motivation for large scale renewable energy plants in the recent years has come along with the challenges and question of stability. Preferably, the outcome of solar panels and wind are highly intermittent, that it is difficult to rely too much on this system as a stand-alone one. When it comes to solar energy, the future of solar energy is not merely dependant on the instantaneous output. It needs to be well ascertained so that the system stability and optimum power output can be obtained. The availability of voluminous weather data and recent advancements in the computational power has enabled tremendous growth in machine learning algorithms to predict future of solar power. In the proposed paper, collection of datasets giving the information on average temperature, surface pressure, wind speed and humidity data are done. The fetched data is then used to train the model using machine learning to develop an Artificial Neural Network (ANN) Model. The data from the neural network is fed into the PV array and the output is fetched to the MPPT system so that the system sustains at maximum power, even if the load changes. Forecasting of the input parameters is then done using time series forecasting with the help of Long-Short Term Memory (LSTM). Thus, the proposed work involves data pre-processing, feature selection and training the dataset and optimization of the output data so that optimal forecasting of solar power can be achieved.

Keywords: Artificial Neural Network, Solar Power, Forecasting, Machine Learning

Article History
***Article Received**: 12 January 2022*
***Revised**: 25 February 2022*
***Accepted**: 20 April 2022*
***Publication**: 09 June 2022*

I Introduction

Renewable energy also referred as clean energy comes from natural sources. As renewable resources are abundantly available in nature, it is important to maximize the resource utilization. It is a sustainable energy that acts as an alternative to non-sustainable sources like coal, fossil fuels which is verge of extinction. The use of renewable energy has increased about 1.5% in 2020 as compared to 2019 and electricity generation increased up to 3% due to solar and wind energy compared to past years. Most popular renewable sources of energy are solar energy, hydro energy, wind energy, geothermal energy etc. Solar energy acts as one of the most valuable sources of renewable energy. It is a clean source of energy which is free from emissions [2]. As Solar Energy is abundantly available, its installation has significantly increased in the recent years. India has achieved 5th global position in Solar Power deployment. India added about 11.1 GW of solar capacity from January 2021 till November 2021. Tata Power Renewable Energy Ltd (TPREL) has commissioned a 300 MW Solar plant in Dholera, Gujarat with country's largest single-axis solar tracker system. Solar PV pace will increase drastically as compared to other renewable energy resources as it is a simple conversion of sunlight to electrical energy using solar photovoltaic cells [11].

With the growing demand of electrical energy, solar photovoltaic power plays important role in power system. Although solar energy is clearly the most abundant power resource available to modern societies, the implementation of widespread solar power utilization is so far impeded by its sensitivity to local weather conditions, intra-hour variability, and dawn and dusk ramping rates [3]. Due to the external factors like humidity, thermal cycling, ultra-violet radiation, moisture ingress, and deposited dust, the absorption of illumination by PV panel gets diminished over a period of time, which thereby reduces the electrical energy output of PV panel and hence its performance [1]. Hence, forecasting of solar power is necessary to predict solar power generation ahead of time for efficient management of electric power system. Hence, PV systems have to be designed in such a way that we maximize the use of resources available.

2 Solar Power Forecasting

The rapid growth in the grid penetration of PV is the urgent call for the power forecasting. Also, the huge investment for PV needs a better and a reliable method to forecast power. PV power forecasting can range across different temporal horizons. As there is no standard classification criterion of the temporal horizon so far, general classifications are made as very short term (Intra-hour: 15 minutes to 2 hours ahead), short-term (hour ahead: 1 to 6 hours ahead, day ahead: 1–3 days ahead), Medium-term (week to months ahead), long-term (one to several years) and lifetime forecast (until PV expected lifetime). In general, short-term forecasting includes ground based sky observations, satellite based methods and numerical weather prediction through the application of image processing and forecasting algorithms [8]. These may give predictions up to seven days ahead. Long-term forecasting usually referred to as forecasting technique predicts power in the order of weeks to years. However, this employs energetic and climatological models [4].

Solar Power forecasting is process where large data-set is collected and trained in-order to predict solar power ahead of time, which helps us provide optimal amount of energy for proper planning and scheduling of electric system. This paper aims to develop a PV system using Machine Learning to anticipate solar power with weather conditions as model inputs. Different methods have been proposed for PV power forecasting [15]. These methods can be classified as physical, heuristic, statistical and machine learning methods. Each method might have different conceptual design, implementation, application and accuracy [6, 7]. In general, forecasting methods for solar power are broadly divided into two categories: (i) physics-based models—these models predict solar power from numerical weather predictions and solar irradiation data, and (ii) statistical models—these models forecast solar power directly from historical data. In this chapter, the application and accuracy of the different methods are assessed using measured PV module power and weather data [5]. This is because, the solar panel works differently for different weather conditions.

Hence, keeping in mind the huge investment for power production using solar PV, the research problem addresses the solar power forecasting in a given area in order to predict solar power generation ahead of time. In recent days, machine learning techniques have been widely used in the fields that involve data driven problems. Hence, through the effective analysis of influencing factors of PV power output, an ANN-based approach for forecasting the power output of photovoltaic system is proposed in this paper.

3 Data Collection & Machine Learning

Machine learning is where computers have the ability to learn things without being internally programmed. Machine learning (ML) uses a computer algorithm that improves automatically through experience and learning by the use of collection data. It is a process where it gets trained over large dataset and predicts the future outputs. For example, in case of solar forecasting, average temperature, surface pressure, wind speed and humidity data is collected and fetched. The fetched data is used to train the model. We analyse the data with the required output and optimize the model. Machine Learning is classified into different categories like supervised, unsupervised and reinforcement learning. Supervised learning is where machine knows both the inputs and outputs. It is used to estimate the unknown function using labelled data. Unsupervised learning is where the labelled input data is known but we may not have a labelled output data. It is used to train machines having unlabelled data. Reinforcement learning is used where suitable actions are taken so that we can maximize the reward.

Artificial Neural network generally works based on supervised learning where we have to estimate the unknown function based on the inputs and outputs. Fig. 1 shows the neural network layer mode which consist of node layers, an input layer, one or more hidden layers, and an output layer. These nodes communicate with each other through weighted values. If the output of any individual node attains specified threshold value, that node is activated, sending data to the next layer of the network. Activation function is used for the system or model to decide if a neuron should be activated or not to attain the desired result.

Output=Activation (Summation (Inputs*weights))

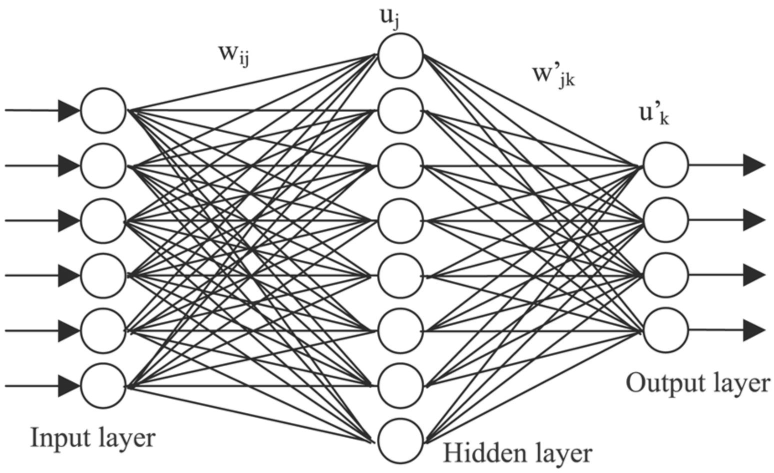


Fig. 1 Artificial Neural Networks Model

There are various algorithms to train a neural network such as gradient descent method, newton method, conjugate algorithm, quasi newton method, Levenberg-Marquardt algorithm etc. Levenberg-Marquardt algorithm considers loss function as sum of squared errors which works using gradient vector and Jacobian matrix [16]. This algorithm is mainly preferred due to its high speed of operation compared to other algorithms. In the proposed paper, Levenberg-Marquardt algorithm is implemented in MATLAB for solar power forecasting.

Initially, dataset is taken and segregated which is then initialised to the appropriate inputs and outputs. The Fig. 2 shows the histogram plot that indicates the number of times a particular output (irradiance) has occurred. Here, we can see that the irradiance value has highly occurred in the range 6.5-7.5 kW-hr/m2 day and has occurred for around 140-200 times in the dataset.

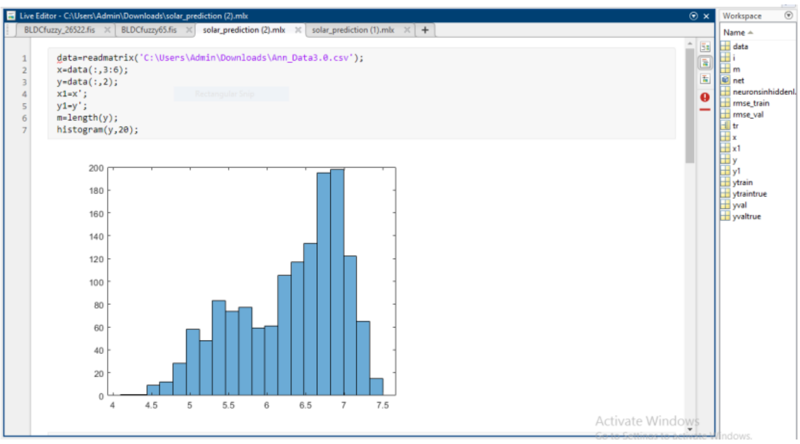


Fig. 2 Histogram plot

The MATLAB implementation includes the initialisation of number of hidden layers and then the dataset is separated to training, testing and validation set. The separated set is then trained for different neuron sizes and for a suitable number of epoch cycles which makes loss reach minimum for the system. The Root mean square error (RMSE) is calculated at every epoch cycle to analyse the performance of the system. Fig. 3 shows the Root mean square error (RMSE) of the testing set and validation set for varying number of neuron sizes. This graph is helpful for us to choose the suitable number of neuron layers for our neural network model. Here, neuron size of 10 seems to be suitable for our model.

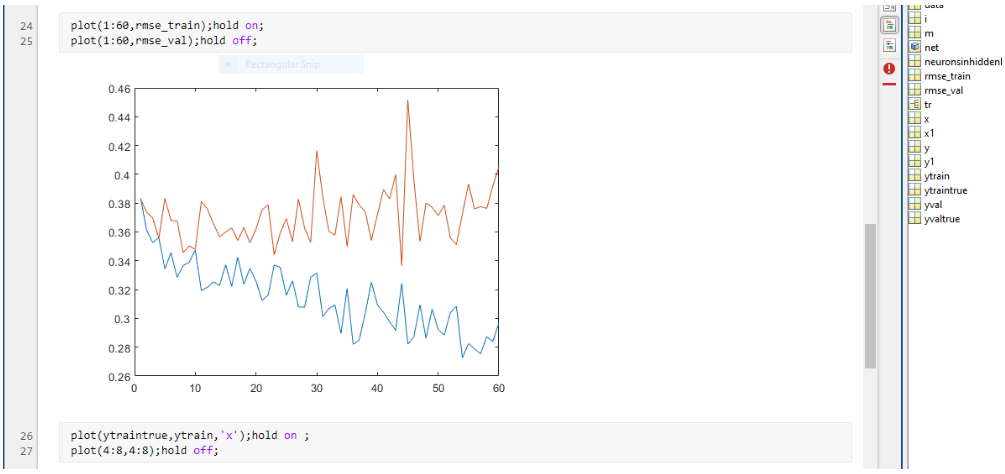


Fig. 3 RMSE Error Vs Neuron size

The Fig. 4 shows us the neural network that was framed for our model which has been executed using Levenberg-Marquardt algorithm. It is clear that each layer has certain weights, bias and sigmoid function associated with it.

Fig. 4 MATLAB model of ANN

The model has been executed for different neuron sizes and performance is analysed. Here, the system runs for certain number of epoch cycles where epoch is the iteration over the entire dataset and there are two hidden layers that where iterated for neuron size of sixty. Every model reacts differently to different neuron sizes, i.e., with change in neuron sizes different model produces different output. Levenberg-Marquardt is preferred here due to its high speed. The plot between trained value obtained by training the model and true value that needs to be obtained is observed through the regression graph as shown in the Fig. 5.

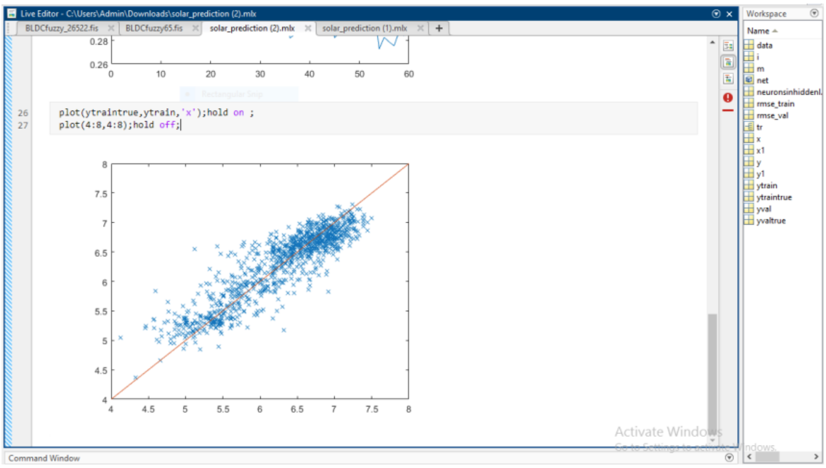


Fig. 5 Regression Graph of True value Vs Trained value plot

4 Maximum Power Point Extraction

Solar PV systems are semiconductor devices that convert sunlight into electricity energy. Solar panel or array consists of large units of PV cells which are connected to provide large units of power. These cells can be connected in series or parallel according to the requirements. Fig. 6 shows the model of a solar PV system that consists of solar Panel, DC-DC converter to track the maximum power point, inverter and loads (DC and AC loads). The Solar PV module converts sunlight energy to electricity which is the fed to a DC-DC converter to attain the voltage at maximum power point of the system to derive maximum power.

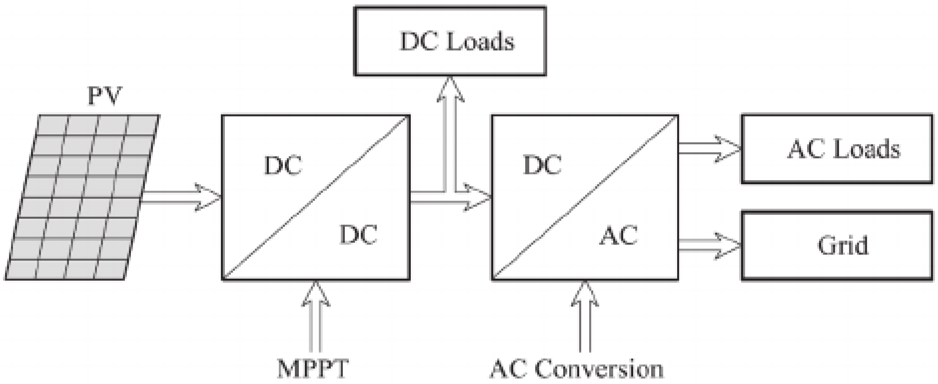


Fig. 6 Solar PV system

An illuminated solar cell can produce a certain photo voltage at a given photo current. The combination of photo current and photo voltage at which a solar cell can be operated is called as working point. The values of photo current and photo voltage at short circuit and open circuit conditions are called short-circuit current (Isc) and open-circuit voltage (Voc) respectively [10].

Solar photovoltaic panels produces varying V-I output for different solar irradiances and produce maximum power at a particular voltage Vmpp. Fig. 7 shows the Vmpp which is the voltage where maximum power point occurs. MPPT is a process of attaining maximum power from the Solar PV array by adjusting the duty cycle of the switch which in turn varies voltage to attain maximum power. Most commonly used Maximum Power Point Tracking algorithms are Perturb and Observe Method and Constant Voltage Method [9]. The concept of MPPT is to continuously observe voltage and current and update the duty cycle to attain the maximum power. It is implemented so that the system attains maximum efficiency and becomes a cost-effective tool for the system to be more economical. Fig. 8(a) and 8(b) shows the variation in solar power as the irradiance varies from 0.1kW/m2 to 1kW/m2 and temperature varies from 15^{o}C to 75^{o}C .

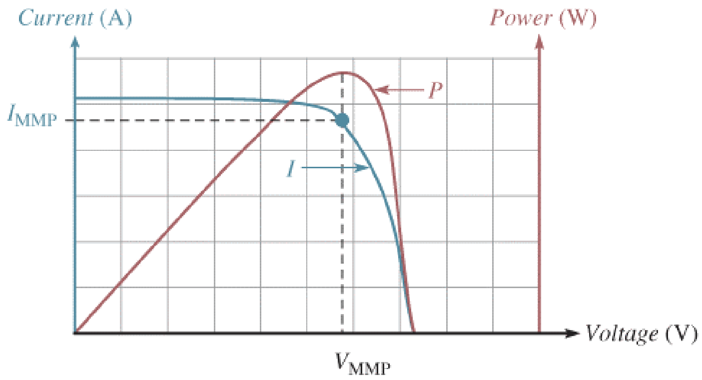


Fig. 7 V-I and P-V Plot

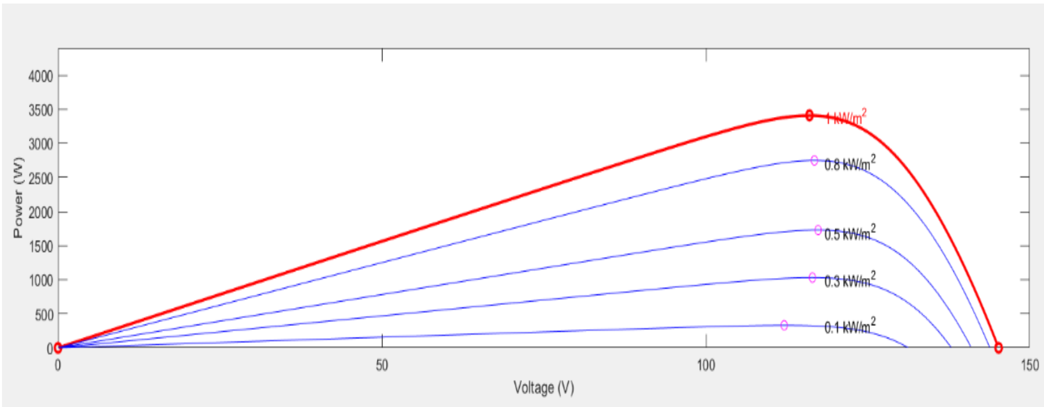


Fig. 8(a) PV Characteristics at different Irradiance value

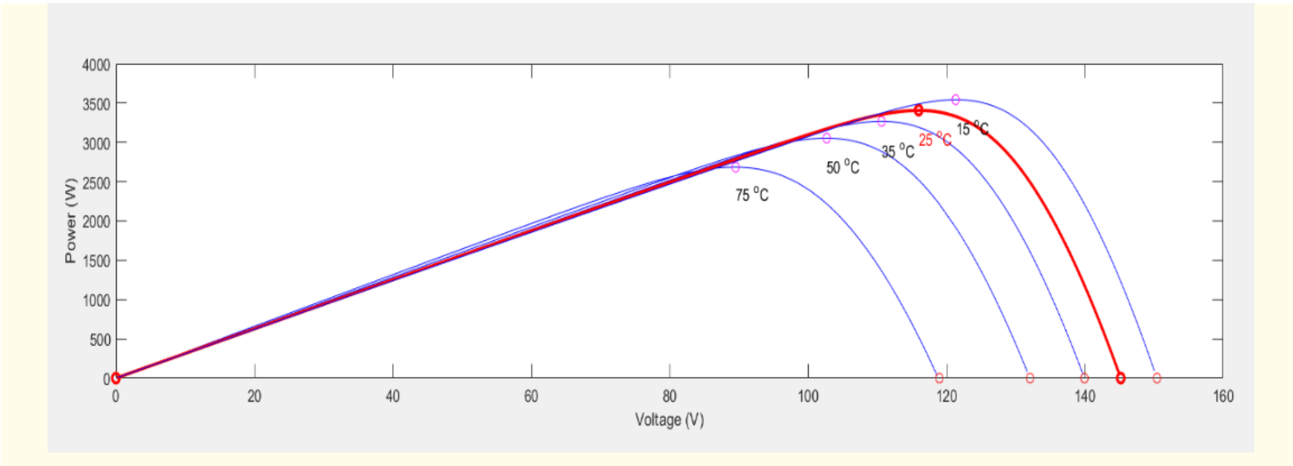


Fig. 8(b) PV Characteristics at different different Temperature

Perturb and Observe Method

Perturbing is a process of incrementing or decrementing the array voltage periodically and continuously observing the output power with its previous perturbation state. Perturb and Observe algorithm is simple in nature and does not require any formation of complex system. It varies the voltage of the system continuously till the operating point of the system reaches the Maximum

Power Point (MPP). Typically, P&O method is used for tracking the MPP. In this technique the PV output power is periodically measured and compared with the previous power. If the output power increases, the same process is continued otherwise perturbation is reversed. In this algorithm, perturbation is provided to the PV module or the array voltage. The PV module voltage is increased or decreased to check whether the power is increased or decreased. When an increase in voltage leads to an increase in power, this means the operating point of the PV module is on the left of the MPP [12, 13]. Hence further perturbation is required towards the right to reach MPP. Conversely, if an increase in voltage leads to a decrease in power, this means the operating point of the PV module is on the right of the MPP and hence further perturbation towards the left is required to reach MPP.

The Fig. 9 shows the MPPT P & O using MATLAB simulation model. In the model, we can see that irradiance and temperature is given as inputs and the voltage and current values is taken as output from the solar panel. Here we have considered 4 series 4 parallel string solar panel modules. The V-I and P-V characteristics of the solar panel module can be seen in the Fig. 10(a) and 10(b). The Fig. 10(b) shows the peak power value of 3410W at 1000W/m2 irradiance value. The voltage value at maximum power point is seen to be 116V. In the MATLAB simulation, the output is then passed to a bus selector where voltage (V_{pv}) and current values (I_{pv}) from the solar pv is measured using goto blocks. Then the V-I output is passed through a DC-DC converters like boost (step-up), buck (step-down), buck-boost converter etc. The system then adjusts the voltage and current values so that the system produces or extracts maximum power from the load or battery. The input and output of the voltage and current values are measured using goto blocks and are plotted.

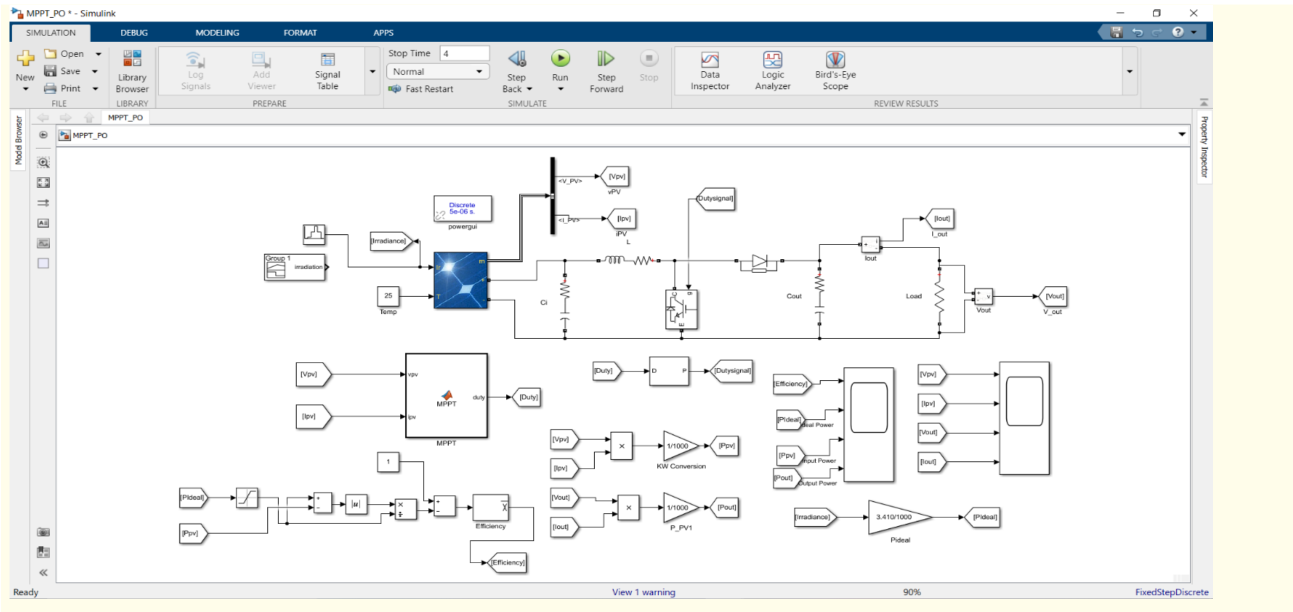
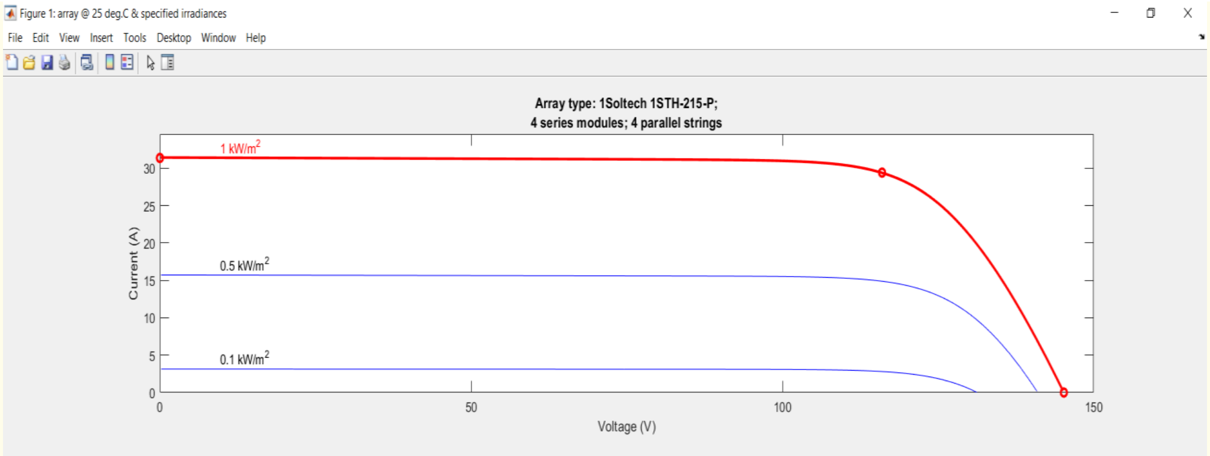
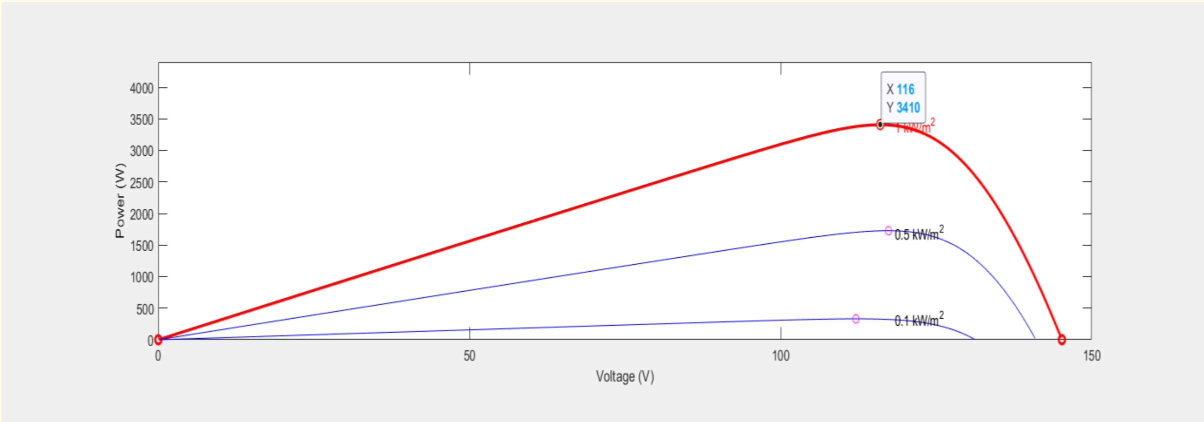


Fig. 9 MPPT P&O using MATLAB Simulation Model



(a)



(b)

Fig. 10(a) and 10(b) V-I and P-V characteristics of the Solar panel module

The Fig. 11 shows the continuous input irradiance condition and the characteristics output curves in the Fig. 12(a-d) and 13(a-d) shows that for continuous Irradiance input the voltage of the solar panel has boosted and current value has declined so that the system attains the same input and output power from dc-dc converter. The system attains same input and output power even when a load is connected.

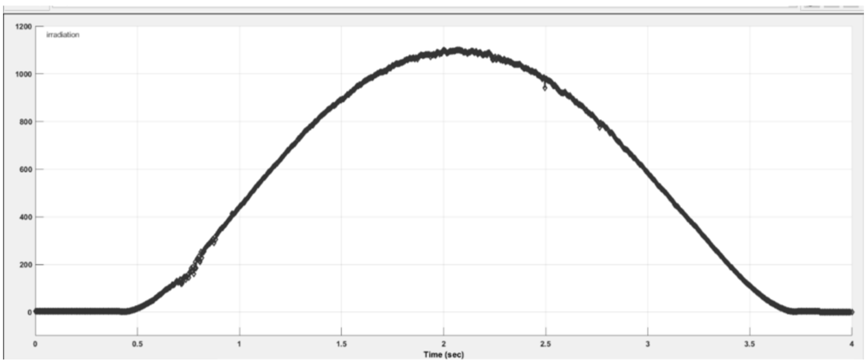


Fig. 11 Continuous Irradiance Input

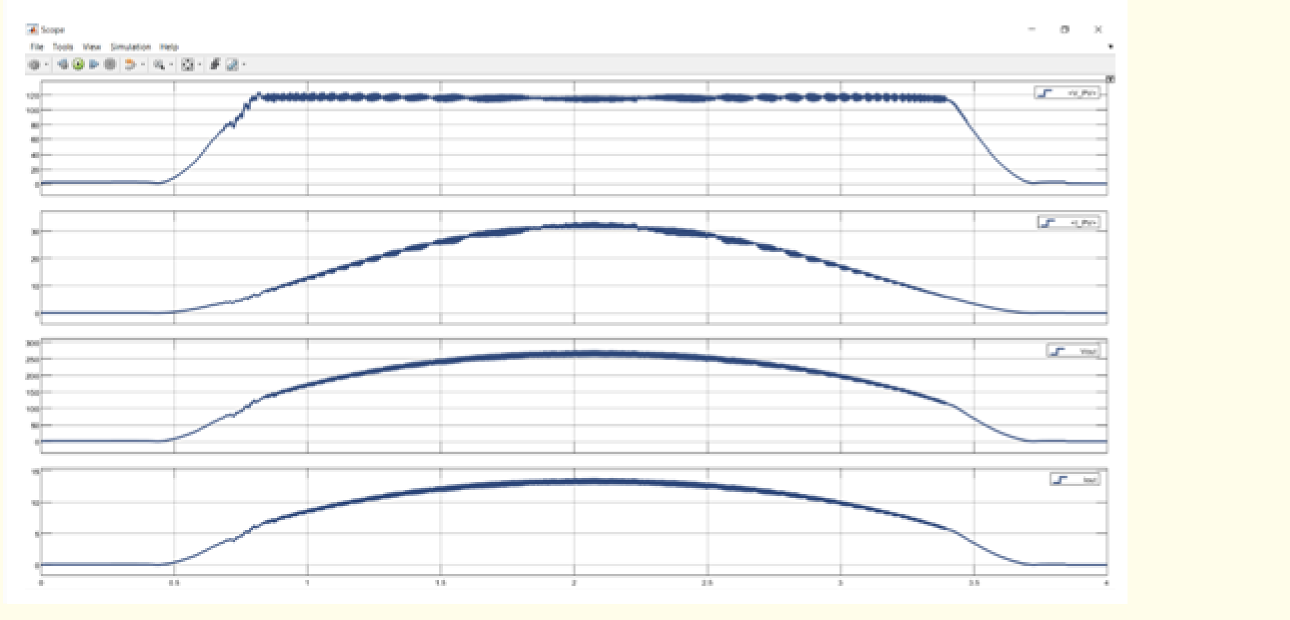


Figure 12 (a) Solar PV-voltage (b) Solar PV-Current (c) Output voltage of MPPT system (d) Output Current of MPPT system

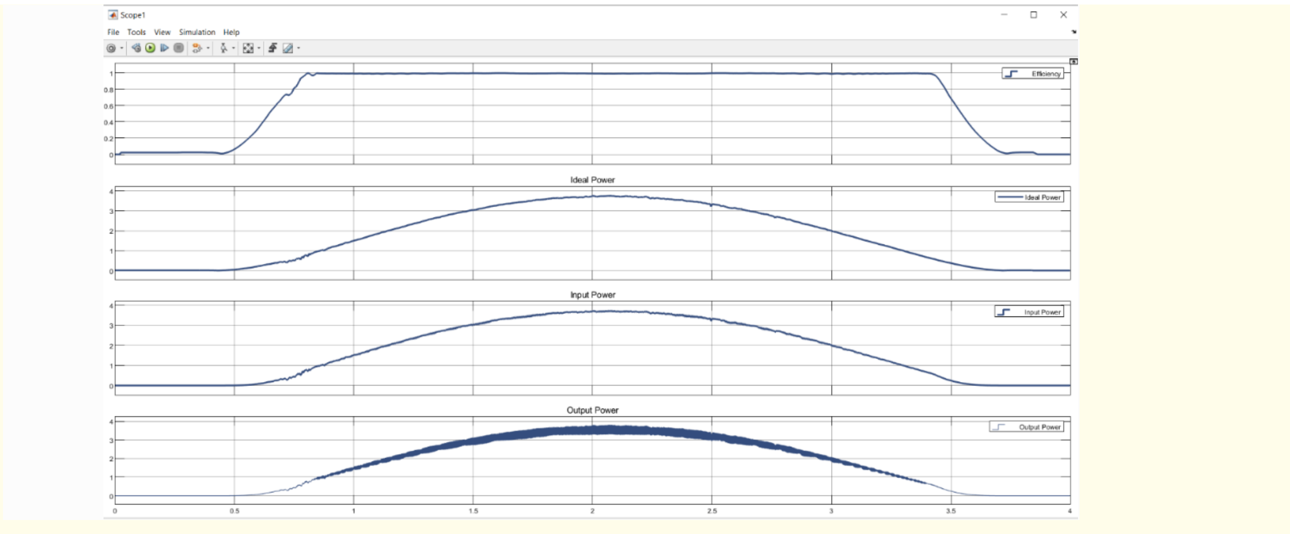
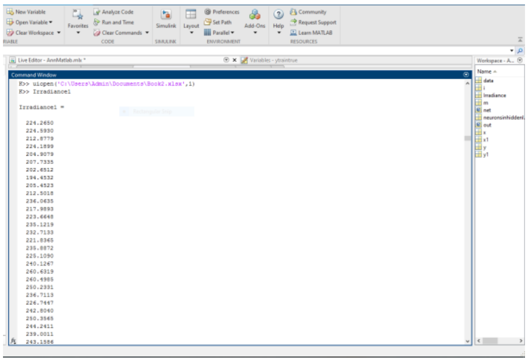
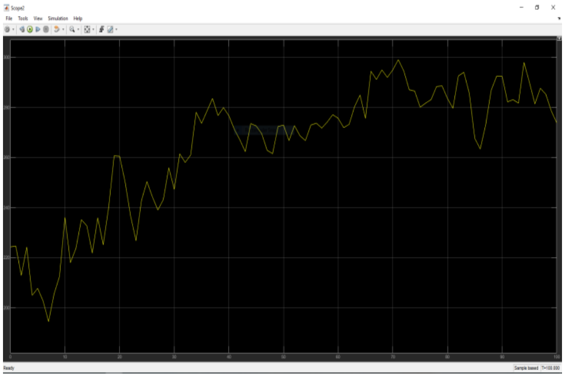


Figure 13 (a) Efficiency (b) Ideal Power plot (c) Input power plot (d) Output Power

5 Solar Power Prediction from the Irradiance data

Fig. 14(a) and 14(b) shows the dataset of predicted irradiance value that we got by training the Artificial Neural Network model and the scope of the input irradiance data.





(a) (b)

Figure 14(a) and 14(b) Dataset and Plot of Predicted Irradiance value

The irradiance data which has been predicted using Artificial Neural Network is fetched to the MPPT model using a 1D-Lookup Table in MATLAB as shown in Fig. 15 and solar power output is derived from the model.

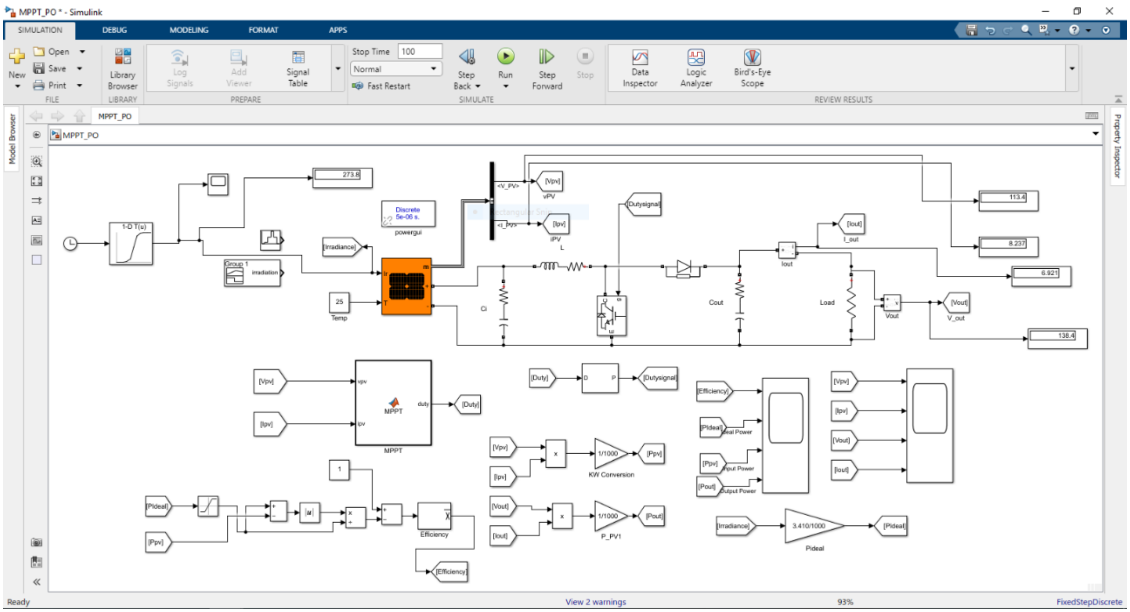


Fig. 15 MATLAB Implementation for predicting Solar Power

6 Time Series Forecasting

Time series forecasting is a process of making predictions over different time samples. It analyses time series data to make future predictions over a period of time. It is a technique for prediction of events over samples of time. The major process involved in time series forecasting is sampling the data, programming and learning and make predictions for the future data. Long Short-

Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) mainly preferred for long term dependencies i.e., for long sequence data. LSTM unlike Recurrent Neural Network which has only single neural network layer, LSTM has four neural networks later; where each neural layer has specified applications associated with it [14]. LSTM network unit is composed of a cell; forget gate, input gate and an output gate as shown in the Fig. 16

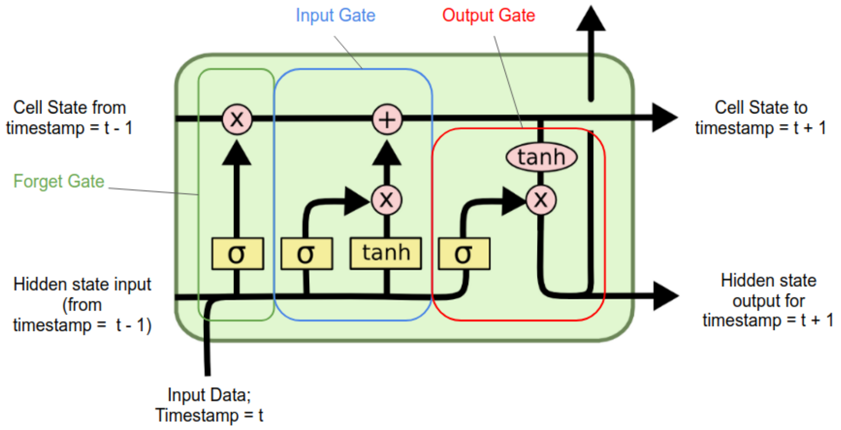


Fig. 16 Layers of LSTM

Parameters like humidity, temperature, surface pressure etc., are helpful in predicting the future solar power in time series prediction. The Fig. 17 shows the process of importing the file LSTM dataset which contains data of humidity, temperature, surface pressure and precipitation. The uploaded file is read in csv format through the command pd.read_csv() and we can have displayed the dataset.

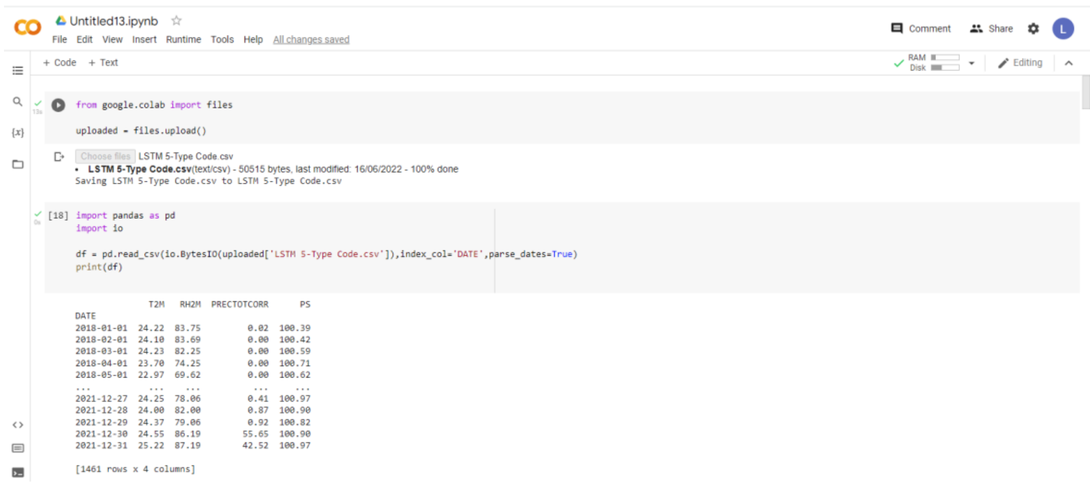


Fig. 17 File Import

Libraries like numpy, which is used to store the dataset in array format which makes it easier to work with large dataset and also consumes less memory and matplotlib, which is used for data visualization and for plotting dataset in graphs are imported. The Fig. 18 shows the graphs of four parameters i.e. humidity, temperature, surface pressure and precipitation.

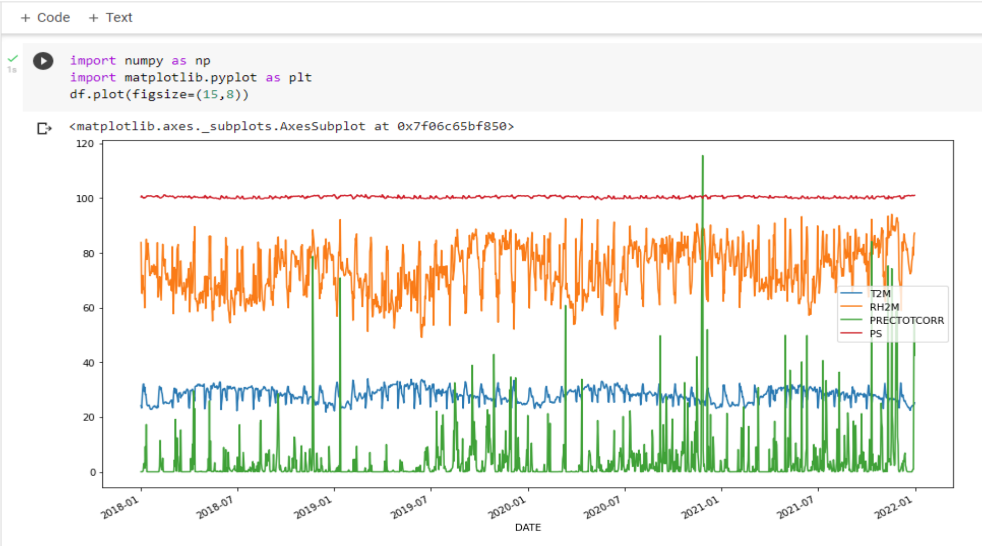


Fig. 18 Graph of input parameters

The library called as sklearn is imported to scale data between certain values, which is preferable when working with large dataset. The dataset above has been scaled to value between 0 and 1. Time series generator from keras library is imported to take upon certain number of inputs and helps the model predict certain number of outputs. Here we have taken 8 data as inputs which is used to predict one output. From the output shown in the Fig. 19, it is seen that the input array has eight values which are used to predict one output as seen in output array.

```
[24] from keras.preprocessing.sequence import TimeseriesGenerator

[25] input = 8
     output =1
     generator = TimeseriesGenerator(strain, strain, length=input, batch_size=1)
     i,o = generator[0]
     print(f'Input Array:\n{i.flatten()}')
     print(f'Output Array: \n {o}')

     Input Array:
     [0.80562007 0.80422666 0.77078495 0.58499768 0.47747329 0.39340455
      0.47909893 0.60102183]
     Output Array:
      [[0.74895495]]
```

Fig. 19 Implementation of Time Series Generator

Similarly, the LSTM model which has 100 layers is created. Here, Relu activation is used which will output the input directly if positive, else it will output zero and a dense layer is created for the output which has one layer as we predicting for a single output. The model is then executed for certain number of epoch cycles to minimize the loss. The Fig. 20 shows the plot of loss function value with respect to the epoch cycles and we see that the loss is decreasing as we increase the number of epoch cycles.

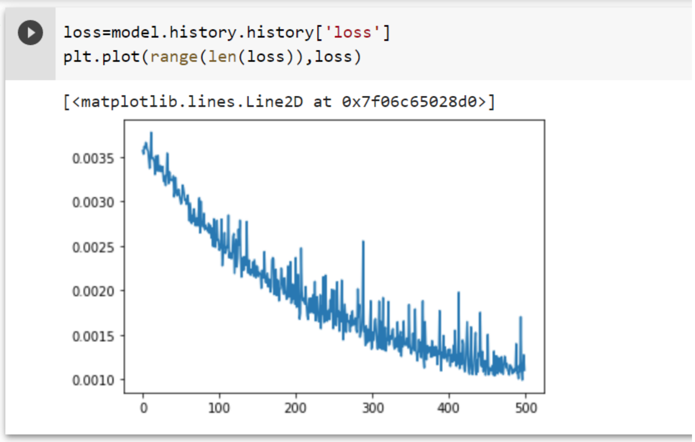
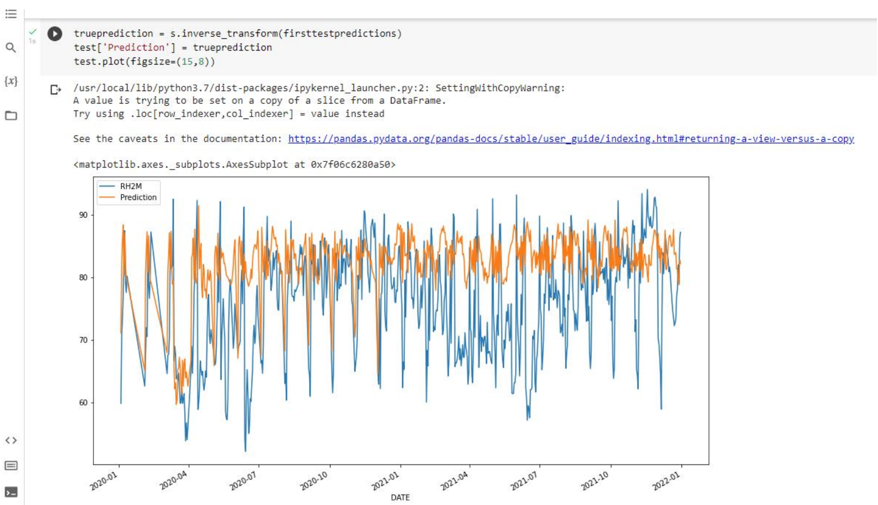
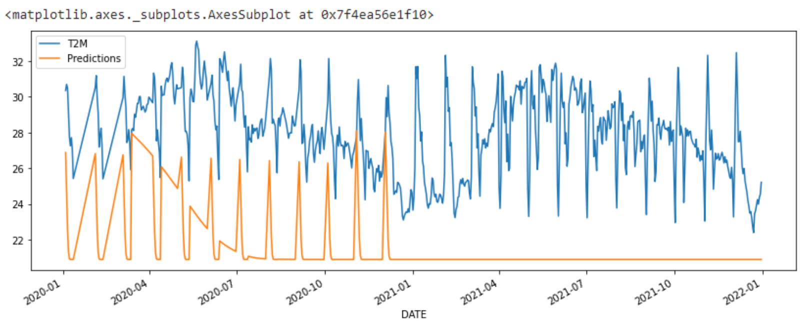


Fig. 20 Error Graph

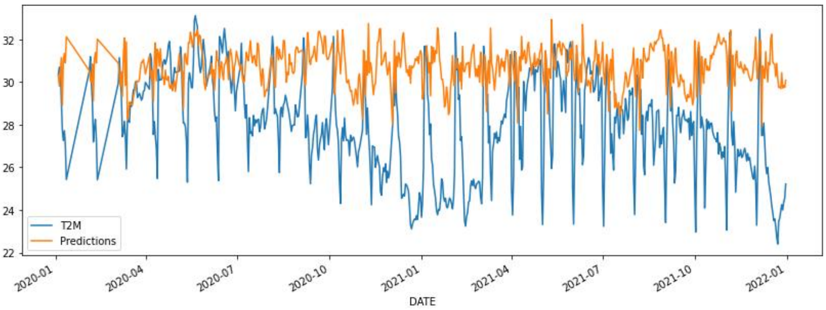
Finally, the last eight values of the training set is used to predict the first value of the testing set and then append the prediction. Then the next consecutive eight values are taken to predict the next sequence of data. This process is iterated for over the length of the test set. The Fig. 21(a), 21(b) and 21(c) shows the graph of predicted value of the test dataset for relative humidity, temperature and pressure trained using LSTM model.



(a)



(b)



(c)

Fig. 21(a), 21(b) and 21(c) Predicted Output plot Relative Humidity

7 Conclusion

The work presented in this paper focused on the performance analysis of photovoltaic systems to meet the global power requirement with sustainability and high stability. The challenges in the solar power prediction and the methods to assess the solar power based on the data collection and training the artificial neural network model is clearly discussed in this paper. Forecasting of the collected input parameters is done using time series forecasting with the help of Long-Short Term Memory. Thus prediction of solar power ahead of time which is the need of the hour is well executed in the proposed work. This work would help the consumers to establish a solar PV in their workstation to get an idea of the amount of solar power that can be achieved from a system through effective forecasting.

References

[1] J. Jurasz, F.A. Canales, A. Kies, M. Guezgouz, A. Beluco, “A review on the complementarity of renewable energy sources: Concept, metrics, application and future research directions”, Solar Energy, Volume 195, 2020, Pages 703-724, ISSN 0038-092X.

[2] K. R. Padiyar; Anil M. Kulkarni, "Solar Power Generation and Energy Storage", Dynamics and Control of Electric Transmission and Microgrids , IEEE, 2019, pp.391-414.

[3] Hassan, M.A., Bailek, N., Bouchouicha, K. et al , "Evaluation of energy extraction of PV systems affected by environmental factors under real outdoor conditions", Theor Appl Climatol (2022).

[4] Haoying Ye, Bo Yang, Yiming Han, Nuo Chen, "State-Of-The-Art Solar Energy Forecasting Approaches: Critical Potentials and Challenges", Frontiers in Energy Research, Volume 10, 15 March 2022.

[5] Yang, B., Zhu, T., and Cao, P, "Classification and Summarization of Solar Irradiance and Power Forecasting Methods: a Thorough Review", 2021.

[6] Tuohy, A., Zack, J., Haupt, S. E., Sharp, J., Ahlstrom, M., Dise, S.,"Solar Forecasting: Methods, Challenges, and Performance", IEEE Power Energ. Mag. 13 (6),2015, 50–59.

[7] Hugo T.C.PedroCarlos F.M.Coimbra," Assessment of forecasting techniques for solar power production with no exogenous inputs", Solar Energy, Volume 86, Issue 7, July 2012, Pages 2017-2028.

[8] Fernandez-Jimenez, L. A., Muñoz-Jimenez, A., Falces, A., Mendoza-Villena, M., Garcia-Garrido, E., Lara-Santillan, P. M., "Short-term Power Forecasting System for Photovoltaic Plants. Renew. Energ". 2012, 44, 311–317.

[9] Manish Kumar , Charvi Saroj, "Maximum Power Point Tracking using Perturb & Observation Technique", Research Paper, Electrical Engineering, India, Volume 10 Issue 7, July 2021, Pages: 451 – 456.

[10] Pavithra, C. V., and C. Vivekanandan. "Bidirectional three port converter for power flow management of PV/Battery-Fed elevator system." Journal of Vibroengineering 20.1 (2018): 732-747.

[11] C V Pavithra, V Chenniyappan, "Modified Standalone Single Stage Three Port Converter with Domain Distribution Control for Renewable Energy Applications", Asian Journal of Research in Social Sciences, 2017, Vo. 7, No. 3, pp. 945-963.

[12] Salman, S., AI, X. & WU, Z. Design of a P-&-O algorithm based MPPT charge controller for a stand-alone 200W PV system. Prot Control Mod Power Syst 3, 25 (2018).

[13] Abdelhakim Belkaid, Ilhami Colak, Korhan Kayisli, "Implementation of a modified P&O-MPPT algorithm adapted for varying solar radiation conditions", September 2017, Electrical Engineering 99(3).

[14] Yu Wang, "A new concept using LSTM Neural Networks for dynamic system identification," 2017 American Control Conference (ACC), 2017, pp. 5324-5329

[15] Dr. Mohammadreza Aghaei, "Solar Radiation - Measurement, Modeling and Forecasting Techniques for Photovoltaic Solar Energy Applications", IntechOpen, 2010, Pages 807-821,ISSN 0038-092X.

[16] A G Farizawani, M Puteh, Y Marina, A Rivaie, "A review of artificial neural network learning rule based on multiple variant of conjugate gradient approaches", Journal of Physics", 1529 (2020).