International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

A Review on Machine Learning Models for Solar Irradiation Prediction

Kapil Raykol[1] Prof. C.K.Tiwari[2] Prof.Anil Malviya[3]
M.Tech Scholor, Patel College of Science and Technology, Indore[1]
Associate Professor, Patel College of Science and Technology, Indore[2]
Assistant Professor, Patel College of Science and Technology, Indore[3]

Abstract: In recent years, one of the most developed renewable energy sources, is the solar energy, along with wind power. In the last years more investment has been made in the development of these two technologies, due to the amount of places with high irradiation and wind possibilities. However, the intermittent nature of solar irradiation increases the need for more flexible and reliable energy generation. Hence, in order to fulfill the requirements of the grid, energy system operators use conventional technologies. Therefore, if a higher penetration of renewable generators is desired, a decrease of the unreliability factors must be achieved by the design of accurate forecasters. Forecasting renewable energies' output power is not a new assignment. Several methods can be found throughout literature. Forecasting is based on predicting future figures within historical databases One application that has been grabbing attention is solar irradiation prediction using data driven models. Machine learning models can find it difficult to follow the pattern of solar irradiation owing to the fact that solar irradiation varies significantly and some may even become zero during nights. This discontinuity causes even more problems. Hence a two-fold approach has been used for solar irradiation prediction using the wavelet transform as a data processing tool for all the relevant parameters and subsequently utilizing the processed data to train a neural network. This paper presents a comprehensive review of the existing techniques in the domain.

Keywords:- Regression Learning, Solar Irradiation Forecasting, Data Pre-Processing, Mean Absolute Percentage Error (MAPE), Regression.

I.Introduction

In recent years, the requirement for energy has been continuously increasing. Utilizing renewable energy resources is a high priority within energy production and management policies in many countries.[1] The growing rate of demand for energy and the global warming phenomenon, which has raised a lot of concern about carbon dioxide emissions along with the high price of fossil fuel, has led governments to consider the utilization of new sources of energy [2]. Several nuclear power-plant disasters and their long-term effects on the next generations' health and environment have also initiated a series of debates on eliminating nuclear power from the future energy policies for some countries. Solar energy is a free and easily available source of energy and appears to be the fastest growing of renewable energy resources [3]. Solar power system penetration to the existing power system possess problems as running problems (frequency, power balance, voltage support, and quality of power), planning and economic problems (including uncertainty in solar power in to unit commitment, economic load scheduling, and spinning reserve calculations), etc.[4]

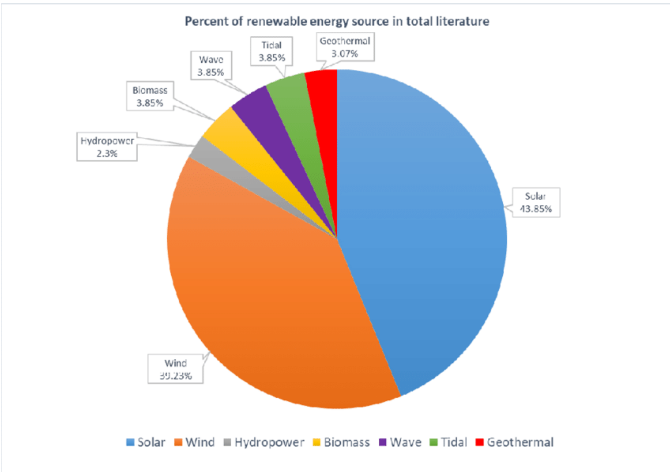

Fig.1 Share of Solar Energy
(Source: https://www.researchgate.net/figure/The-pie-chart-in-terms-of-seven-renewable-energy-sources_fig1_343953072)

Previously statistical models were used for the prediction of solar energy and thereby solar power, but it lacked accuracy due to its inability to follow complex solar energy patterns accurately. Thus the focus started shifting on Artificial Intelligence tools to forecast solar energy. The subsequent sections introduce the basics of artificial neural network, its functioning and various architectures of artificial neural networks [5].

II. Machine Learning Models.

Solar irradiation forecasting plays a crucial role in optimizing the performance of solar energy systems and ensuring efficient energy production. Several machine learning models have been developed to enhance the accuracy of solar irradiation predictions.

One commonly employed model is the Support Vector Machine (SVM), which is effective in capturing complex relationships between various meteorological parameters and solar irradiance. SVMs excel in both classification and regression tasks, making them suitable for predicting solar irradiation levels based on historical weather data [6].

Another popular approach involves the use of artificial neural networks (ANNs). These models, inspired by the structure of the human brain, can learn intricate patterns from large datasets. In solar irradiation forecasting, ANNs are adept at handling nonlinear relationships and can adapt to changing environmental conditions, making them valuable for accurate predictions.

Ensemble methods, such as Random Forests and Gradient Boosting, have also shown promise in solar irradiation forecasting. These models combine the predictions of multiple weaker models to improve overall accuracy. By leveraging the diversity of individual models, ensembles can better handle uncertainties in meteorological data and provide robust predictions [7].

Lately, deep learning techniques, particularly Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have gained traction for solar irradiation forecasting. CNNs excel in extracting spatial patterns from meteorological images, while LSTMs are adept at capturing temporal dependencies in time-series data. The combination of these architectures can enhance the model's ability to capture complex spatiotemporal relationships in solar irradiance patterns [8].

Hybrid models that combine traditional physics-based models with machine learning techniques also offer a promising avenue. By integrating domain knowledge with data-driven approaches, these models can capitalize on the strengths of both, providing accurate and reliable solar irradiance predictions.

Work on artificial neural network has been motivated right from its inception by the recognition that the human brain computes in an entirely different way from the conventional digital computer. [9] The brain is a highly complex, nonlinear and parallel information processing system. It has the capability to organize its structural constituents, known as neurons, so as to perform certain computations many times faster than the fastest digital computer in existence today. The brain routinely accomplishes perceptual recognition tasks, e.g. recognizing a familiar face embedded in an unfamiliar scene, in approximately 100-200 ms, whereas tasks of much lesser complexity may take days on a conventional computer. A neural network is a machine that is designed to model the way in which the brain performs a particular task. The network is implemented by using electronic components or is simulated in software on a digital computer. A neural network is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making it available for use. It resembles the brain in two respects [10]:

1. Knowledge is acquired by the network from its environment through a learning process.

2. Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. [11]

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Other advantages include:

1. **Adaptive learning:** An ability to learn how to do tasks based on the data given for training or initial experience.
2. **Self-Organization:** An ANN can create its own organization or representation of the information it receives during learning time.
3. **Real Time Operation:** ANN computations may be carried out in parallel, and special hardware

devices are being designed and manufactured which take advantage of this capability.

The biological model of the neuron is shown in the figure. It consists of the cell body, axon hillock, action potential, synaptic terminal, axon of pre synaptic neuron and dendrites. Signals from different parts of the body travel through different parts and reach the neuron where the neuron processes it and produces an output. It should be noted though that the output of a neuron may also be fed to another neuron. A collection of such neurons is called a neural network. The neural network can perform simple to complex tasks depending on the structure of the neural network. After studying the basic biological model of the neural network, a mathematical model is envisaged to be designated. The mathematical model for such a neural network is given by:

$$Y = f(\sum_{i=1}^{n} XiWi \; + \; \Theta) \qquad (1)$$

Where
Xi represents the signals arriving through various paths,
Wi represents the weight corresponding to the various paths
ϴ is the bias.
f is the activation function.

The diagram below exhibits the derived mathematical model of the neural network. It can be seen that various signals traverssing different paths have been assigned names X and each path has been assigned a weight W.[12] The signal traverssing a particular path gets multiplied by a corresponding weight W and finally the overall summation of the signals multiplied by the corresponding path weights reaches the neuron which reacts to it according to the bias ϴ. Finally its the bias that decides the activation function that is responsiblefor the decision taken upon by the neuralnetwork.

III. Data Filtration

Due to the randomness and flutcations in raw solar irradiation data, it is needed to filter it out. The wavelet transform can be crudely seen as a tool for the smoothening of local disturebances in the data set and is widely used as a data pre-processing tool. The mathematical formulation for the wavelet transform is given by the scaling and shifting approach of the wavelet function.
The scaling, shifting dependence can be defined as:

$$\mathbf{W\Psi\ (sc, sh) = \ ₩(\{x,t\})} \qquad \mathbf{(2)}$$

Here,
x is the space variable
t is the time variable
₩ is the transform
sc is the scaling factor
sh is the shifting factor
Thus the wavelet transform can be considered to be a shifted-scaled version of wavelet family functions.
There are several wavelet families such as haar, Mexican hat, morlet etc. The base functions differ from conventional sine/cosine functions exhibiting smoothness in the period of definition. The proposed approach uses the wavelet transform on the independent variables and then trains the neural network with the values.

IV. Previous Work

This section presents a summary of noteworthy contribution in the domain:
Ghimire et al. proposed a new hybrid deep learning (DL) model, the called CSVR, for Global Solar Radiation (GSR) predictions by integrating Convolutional Neural Network (CNN) with Support Vector Regression (SVR) approach. First, the CNN algorithm is used to extract local patterns as well as common features that occur recurrently in time series data at different intervals.
Estragi et al. showed that Renewable energies are the alternative that leads to a cleaner generation and a reduction in CO2 emissions. However, their dependency on weather makes them unreliable. Traditional energy operators need a highly accurate estimation of energy to ensure the appropriate control of the network, since energy generation and demand must be balanced.
Santos et al. proposed that describes the application of models to estimate the transmitted fraction of direct solar irradiation into normal incidence as a function of the atmospheric transmissivity (Kt) and the insolation ratio In the first model, the values of Kt in the hourly (h) and daily (d) partitions were correlated using polynomial regression. In the second model, and in the daily partition were correlated through linear regression.
Li et al. proposed a novel scheme for forecasting irradiance. The method considers the hourly irradiance prediction model to be the superposition of two parts: a daily average irradiance prediction model and the

irradiance amplitude prediction model. Two submodels were constructed by using deep bidirectional long short-term memory (BiLSTM) network. for 80% of the climates included in the experiment.

Boubaker et al. proposed that forecasted global horizontal irradiation (GHI) can help for designing, sizing and performances analysis of photovoltaic (PV) systems including water PV pumping systems used for irrigation applications. In this paper, various deep neural networks (DNN) models for one day-ahead prediction of GHI at Hail city (Saudi Arabia) are developed and investigated. The considered DNN models include long-shortterm memory (LSTM), bidirectional-LSTM (BiLSTM), gated recurrent unit (GRU), bidirectional-GRU (Bi-GRU), one-dimensional convolutional neural network (CNN 1D) and other hybrid configurations such as CNN-LSTM and CNN-BiLSTM.

Liu et al. proposed that residential energy scheduling of solar energy is an important research area of smart grid. The highlights of this paper are listed below. First, the weather-type classification is adopted to establish three types of programming models based on the features of the solar energy. In addition, the priorities of different energy resources are set to reduce the loss of electrical energy transmissions. Second, three ADHDP-based neural networks, which can update themselves during applications, are designed to manage the flows of electricity. Third, simulation results show that the proposed scheduling method has effectively reduced the total electricity cost and improved load balancing process. The comparison with the particle swarm optimization algorithm further proves that the present method has a promising effect on energy management to save cost.

Chettibi et al. proposed that Artificial Intelligence and machine learning concept has been used widely in the current times. Neural networks are a part of the artificial intelligence concept which is very useful for training with datasets. They possess much flexibility and accuracy in terms of performance metrics. So in this study the authors tried to use the approach of adaptive neural networks to monitor a micro grid. Grid based systems contain many intricate portions which have to be designed very carefully.

Queja et al. showed that the solar power forecasting in pleasant weather condition is a relatively common method. But when the environment has humidity and the weather is warm, then it is difficult get the accurate measure of solar power prediction. So, in order to estimate the daily global solar radiation in such a scenario, many methods and mechanisms have to be used together. In this work, SVM, ANN and ANFIS techniques have been used for the purpose. This approach is quite a strong and robust procedure to achieve the desired performance level as the merits of all methods get coupled together. However, the drawback is that designing such a concept is time consuming and little complex at the same time.

Evaluation Parameters

Since errors can be both negative and positive in polarity, therefore its immaterial to consider errors with signs which may lead to cancellation and hence inaccurate evaluation of errors. Therefore we consider mean square error and mean absolute percentage errors for evaluation. Mean Square Error is defined as:

$$\boldsymbol{MSE} = [\textstyle\sum_{i=1}^{n}(\mathbf{X}-\mathbf{X'})\text{^}\mathbf{2}]/\mathbf{n} \qquad \mathbf{(3)}$$

Mean Absolute Percentage Error is defined as:

$$\boldsymbol{MAPE} = [\textstyle\sum_{i=1}^{n}(\boldsymbol{X}-\boldsymbol{X'})/\boldsymbol{X'}]/\mathbf{n} \text{ x } \mathbf{100\%} \qquad \mathbf{(4)}$$

Here,
X is the predicted value,
X' is the actual value and
n is the number of samples.

VI. Conclusion:

The need for machine learning models in solar irradiation forecasting is driven by the complex and dynamic nature of meteorological conditions influencing solar energy generation. These models provide a data-driven and adaptive approach, improving accuracy, optimizing energy production, and ensuring the efficient integration of solar power into the broader energy landscape. It can be concluded form the above discussions that Artificial Neural Networks can be effectively used for solar energy prediction even though solar energy may exhibit complex time series behaviour. Various machine learning models have been discussed with their salient features. Finally the evaluation parameters used for the evaluation of any prediction model have been presented.

References

[1] S Ghimire, B Bhandari, D Casillas-Perez, “Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia”, Engineering Applications of Artificial Intelligence. Elsevier 2022, vol. 112, 104860.

[2] G Etxegarai, A López, N Aginako, “An analysis of different deep learning neural networks for intra-hour solar irradiation forecasting to compute solar photovoltaic generators' energy production”, Energy for Sustainable Development, Elsevier, 2022, vol.68, pp.1-17.

[3] C dos Santos, JF Escobedo, A de Souza, “Prediction of solar direct beam transmittance derived from global irradiation and sunshine duration using ANFIS”, International Journal of Hydrogen Energy, Elsevier, 2021, vol. 46, no. 55, 10 pp.27905-27921.

[4] C Li, Y Zhang, G Zhao, Y Ren, “Hourly solar irradiance prediction using deep BiLSTM network”, Earth Science Informatics, Springer 2021, vol. 14, pp. 299–309.

[5] S. Boubaker, M. Benghanem, A. Mellit, A. Lefza, O. Kahouli and L. Kolsi, "Deep Neural Networks for Predicting Solar Radiation at Hail Region, Saudi Arabia," in IEEE Access 2021, vol. 9, pp. 36719-36729.

[6] D. Liu, Y. Xu, Q. Wei and X. Liu, "Residential energy scheduling for variable weather solar energy based on adaptive dynamic programming," in IEEE/CAA Journal of Automatica Sinica, vol. 5, no. 1, pp. 36-46.

[7] N. Chettibi, A. Mellit, G. Sulligoi and A. Massi Pavan, "Adaptive Neural Network-Based Control of a Hybrid AC/DC Microgrid," in IEEE Transactions on Smart Grid, 2020 vol. 9, no. 3, pp. 1667-1679.

[8] RA e Silva, MC Brito, “Impact of network layout and time resolution on spatio-temporal solar forecasting”, ELSEVIER 2019, vol.163, pp. 329-337.

[9] Y. Wang, N. Zhang, Q. Chen, D. S. Kirschen, P. Li and Q. Xia, "Data-Driven Probabilistic Net Load Forecasting With High Penetration of Behind-the-Meter PV," in IEEE Transactions on Power Systems, 2018, vol. 33, no. 3, pp. 3255-3264,

[10] H. Sheng, J. Xiao, Y. Cheng, Q. Ni and S. Wang, "Short-Term Solar Power Forecasting Based on Weighted Gaussian Process Regression," in IEEE Transactions on Industrial Electronics, vol. 65, no. 1, pp. 300-308, Jan. 2018

[11] DW Van der Meer, M Shepero, A Svensson, J Widén, “Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes”, 2017 ELSEVIER., vol.213, pp. 195-207.

[12] VH Quej, J Almorox, JA Arnaldo, L Saito, “ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment”, 2017 ELSEVIER, vol.155., pp. 62-70.

[13] K Ostad-Ali-Askari, M Shayannejad, H Ghorbanizadeh-Kharazi “Artificial Neural Network for Modeling Nitrate Pollution of Groundwater in Marginal Area of Zayandeh-rood River, Isfahan, Iran”, 2017 SPRINGER., vol.21, pp.134–140,

[14] S Sepasi, E Reihani, AM Howlader, LR Roose, “Very short term load forecasting of a distribution system with high PV penetration”, 2017 ELSEVIER, vol.106, pp. 142-148.

[15] S Sepasi, E Reihani, AM Howlader, LR Roose, "Distribution Voltage Regulation Through Active Power Curtailment With PV Inverters and Solar Generation Forecasts," in IEEE Transactions on Sustainable Energy, vol. 8, no. 1, pp. 13-22, Jan. 2017.

[16] M Bou-Rabee, SA Sulaiman, MS Saleh “Using artificial neural networks to estimate solar radiation in Kuwait”, 2017 ELSEVIER, vol.72, 434-438