Review Article

Journal of Environmental Nanotechnology

Advancements and Challenges in Solar Radiation Prediction: A Review of Machine Learning Approaches

K. C. Jayasankar[1*], G. Anandhakumar[1] and A. Kalaimurugan[2]

[1]Department of Electrical and Electronics Engineering, Saveetha School of Engineering, SIMATS, Saveetha University, Chennai, TN, India

[2]Department of Electrical and Electronics Engineering, Agni College of Technology, Chennai, TN, India

Received: 06.05.2024 Accepted: 30.05.2024 Published: 30.06.2024

*jayasankarkc3002.sse@saveetha.com

ABSTRACT

Solar irradiance prediction holds huge importance in optimizing the efficacy and dependability of solar power systems. This evaluative work focuses on an array of machine learning methodologies applied in solar irradiance prediction, encompassing conventional methodologies. It further investigates the changes in meteorological data, geographical and temporal parameters, data preprocessing, feature curation methodologies, combined learning methodologies, and composite models. This work investigates the potential advantages of combining orbital visuals and satellite sensors to augment prediction capacities. The review highlights the requirement for broader datasets, refined feature curation methodologies, and composite models to attain improved precision and resilience.

Keywords: Solar irradiance; Renewable energy; Predictive modeling; Computational methods.

1. INTRODUCTION

The elevation of solar power as a growing renewable repository reduces the dependence on fossil fuels. The accurate prediction of solar irradiance serves as a necessary seamless operation of solar energy infrastructures. Though conventional methodologies indicate the limitations in this domain, the emergence of Machine Learning (ML) algorithms denotes a transformative shift, presenting increased precision and effectiveness. This evaluative work presents a thorough investigation of the current utilizations of ML techniques in the field of solar irradiance assessment.

2. OVERVIEW OF SOLAR RADIATION FORECASTING

Clean energy goals have become vital in global energy politics recently. These aims increase renewable energy consumption, to reduce the release of greenhouse gases and mitigate climate change. Ultraviolet (UV) forecast optimizes use of solar power resources in sustainable energy management. The reliable UV forecast models are developed using ML.

It is difficult to integrate renewable energy into the electricity infrastructure. Generation can cause energy imbalances in supply and demand due to the temporary nature of energy from renewable sources. Smart grid solutions are being investigated to enhance coordination and control renewable energy production and utilization. This issue is considerably more important when integrating excess energy from renewable sources into the system (Alam *et al.* 2020).

The transient nature of energy from environmentally friendly generation produce energy disparities in both demand and supply. Another obstacle in using green power in the electrical grid is the problem of fault currents. As they can occur in unusual circumstances, such as shorts in wiring, and can cause harm to the electrical systems equipment. To minimize this issue, electrical fault limiters are being researched as an option (Alam *et al.* 2018). Such devices can help to manage the movement of currents caused by faults and protect the electrical system foundation.

Recent years have witnessed fast growing combination of wind and solar energy infrastructures to facilitate grid stability management. The automation of frequency control emerges as a essential element in upholding the equilibrium and reliability of the electrical grid. Ren *et al.* (2015) stated that various regulation strategies and techniques have been presented to overcome the issues associated with frequency regulation in both wind and solar power combined systems. These advances show ongoing investigation and innovation in the field of green energy integrating and their influence on reliability in the grid.

Within the domain of solar power forecasting, numerous algorithms and methodologies have been observed to meticulously forecast solar irradiance. A good example is the Multi-stage Multimodal Mode Decomposition approach (MEMD) linked with the Ant

Colonies Optimization (ACO) and Random Forest (RF) model, as recommended by Dong *et al.* (2014). This approach aims to predict monthly solar radiation through the employment of MEMD. The ACO technique is employed to select the most applicable intrinsic mode functions, which then act as input features for the RF model.

The integrated approach of MEMD, ACO, and RF allows for better precision in solar radiation forecast. Another distinct strategy is the incorporation of Convolutional Neural Networks (CNNs) for irradiance from the sun prediction, as demonstrated by Dong *et al.* (2020). The CNN framework utilizes the geographical and temporal properties of sun irradiance facts to make reliable predictions. Additionally, a spatio-temporal forecasting approach for worldwide horizontal radiation using the method of Bayesian inference is also utilized.

3. MACHINE LEARNING (ML) TECHNOLOGIES FOR UV ESTIMATION

Statistical algorithms are computational models capable of automatically discovering patterns and generating predictions from given data without being specifically programmed. The ML algorithms can manage huge datasets, capture complicated connections, and adapt to changing situations, making them excellent for UV forecasting. Some ML algorithms used for UV forecasting include Artificial Neural Networks (ANNs), Supported Vector Machines (SVMs), Random forests (RFs), and Gradient Boosting Machine (GBMs).

Solar irradiance prediction plays a key role in different programs, such as producing electricity from the sun and weather prediction. Researchers have developed various models that precisely calculate the radiation from the sun according to empirical data. Their work examined historical data and built models that may be utilized for forecasting solar energy in the region. Similarly, Sharifi *et al.* (2016) evaluated and compared various temperature centric methodologies for predicting monthly global solar irradiance. They contrasted wavelet regression, ANN, Gene Expression Programming (GEP), and empirical models, illustrating the merits and demerits of each approach.

4. DATA AND PRECONDITIONING SOURCES

Accurate sunlight forecasting relies on superior input information including historical sunlight radiation metrics, weather information, and satellite imagery. The data preparation procedures, such as data cleaning, the normalization process, and feature selection, are critical to assure the accuracy and dependability of ML models. By examining data from the grid, the researchers were able to design a fault diagnosis system that can detect and identify issues in real-time, enhancing the reliability and efficiency of the grid. Ren *et al.* (2015) did a cutting-edge examination of collective methodologies.

The authors underlined the advantages of ensemble approaches in enhancing the accuracy of electricity forecasting. These research investigations demonstrate the wide-ranging uses of learning algorithms in the energy field and emphasize its potential for strengthening forecasting and diagnosis processes. Solar light predictions is an important part of energy from renewable sources systems, as it aids in optimizing the effectiveness of solar thermal power plants. In the past few decades, Artificial Intelligence (AI) models have attracted considerable interest for sunlight forecast due to their capacity to handle complicated and nonlinear interactions. An example of this AI algorithm is the physics-informed deep clustering approach, which has demonstrated excellent outcomes in various fields related to porous media and possible difficulties in heterogeneous medium (Dong *et al.* 2014; Guo *et al.* 2022a).

The technique incorporates network design searching, learning through transfer, or sensitivity testing to boost the accuracy and dependability of sun radiation prediction (Guo *et al.* 2022b). Diverse strategies such as integrated frameworks integrating machine cognition algorithms and optimization methodologies have been suggested. These frameworks endeavor to heighten the precision of solar radiation predictions by considering variables such as atmospheric circumstances, input parameters, and optimization patterns. The combined characteristic curation methodologies are recommended to improve prediction precision. In the realm of AI, gradient boost decision trees have attracted substantial attention due to their great efficiency and accuracy.

The aforementioned characteristics make LightGBM a popular choice for many different purposes, including radiation from the sun forecast (Chaibi *et al.* 2021) and the feasibility assessment of hybrid power sources. CatBoost is an algorithm specifically built to handle categorical characteristics, making it ideal for datasets with multiple data types. It implements a novel strategy called Ordered Boost, which includes the natural ranking of data by categories into the increasing process. That enables CatBoost to successfully capture the correlations among categorical characteristics and the subject variable. The adaptability of the CatBoost makes it a powerful tool in different sectors, particularly green energy studying and modelling (Prasad *et al.* 2019). Overall, the improvements in boost gradient strategies that include LightGBM and CatBoost made a significant contribution to the field of neural networks as they have opened new opportunities for handling complicated challenges across numerous domains. Several investigations have been undertaken to build models of prediction for sun exposure in Bengal. For instance, Rabbi *et al.* (2016) applied an ANN and GIS technologies to forecast solar irradiation. Bhatnagar *et al.* (2022) focused on strengthening the robustness of transmission lines employing severe gradient strengthening against

faults. Their work aims to increase the dependability as well as efficacy of the transmission system in case of faults. These studies illustrate the necessity of reliable prediction models in the realm of renewable energy and provide helpful tips for the development and optimization of wind and solar powered systems.

The authors demonstrated the usefulness of the Industry Foundation Classes and Green Building Rating (IFC-GBR) approach through contrasting its performance with other regularly used regression algorithms. Studies demonstrated that the IFC-GBR technique outstands the remaining strategies in terms of inaccuracy and resilience. This study illustrates the prospective of the IFC-GBR method as a trustworthy tool for forecasting the tendencies of complex systems. Conclusions resulting from this study contribute to the progress of predictive modelling approaches and have changes for numerous sectors which include the chemical engineering field and material science. The IFC-GBR serves as a promising methodology for properly estimating the indeterminate diffusion activity index of chloroform in ionic liquids.

Arora *et al.* (2021) analyzed the utilization of AI techniques for assessing the construction expenditures of both reinforced and prestressed concrete bridge projects. By evaluating a dataset of previous bridge projects, the scientists built an estimation framework that could accurately forecast the costs depending on numerous factors, such as bridge length of sentence, span, and design aspects. The results indicated that the ML algorithm outperformed traditional estimation approaches, offering more accurate cost estimates and lowering the time necessary for calculations. This research emphasizes the potential of ML in transforming cost estimation procedures in bridging the field of engineering, leading to improved effectiveness and affordability in infrastructure building.

A separate investigation by the same authors investigated temperature-based ML with empirical algorithms to forecast average global sun radiation. The results suggested that these mathematical frameworks are highly effective at forecasting solar radiation. Besides computer techniques for learning, batch feature selection has also been applied for sunlight forecasting.

In the study of Bounoua *et al.* (2021), the application of the following parameters and Bayesian optimization strategies was also suggested in train arrival delays predicting. This model can be employed to increase the quality of railway operations by predicting arrival delays. Overall, these investigations show the crucial role of machine training and empirical models in effectively predicted sunlight and train timetable delays. The prediction of solar irradiance stands as a pivotal factor in optimizing the efficient utilization of sustainable energy sources. Various approaches were proposed for calculating radiation from the sun, notably empirical methods and ANN techniques. Their findings indicated that ML technologies, including ANNs, are capable of accurately estimating the sunlight and surpass classic empirical models. These findings highlight the need to adapt modern methodologies, such as ANN toward accurate solar irradiance predictions. Through the employment of algorithms leveraging ML and normalization techniques to process the data, researchers can enhance the accuracy and reliability of solar irradiance prediction, thereby facilitating the more effective utilization of sustainable energy reservoirs.

5. RESULTS

The concise presentation depicts the comparative effectiveness of various ML algorithms in forecasting solar radiation. The metrics including error standard deviation, R^2 performance, and root mean squared error for each algorithm are given in Table 1. A visual representation of comparative analysis plot portraying the above-mentioned parameters for each algorithm, enabling a clear comparison of their predictive capabilities is shown in Fig. 1.

Table 1. The confluence of various machine learning methodologies employed in the prediction of solar irradiance

Reference	Location	Best Model	R^2	RMSE	MAE	MAPE	AER	MSLE
Dong *et al.* 2020	China	Novel CNN	-	0.014	0.063	-	0.011	-
Arora *et al.* 2021	India	ANN	0.9882	-	0.03069	6.63141	-	-
Bounoua *et al.* 2021	Morocco	RF	0.962	0.0785	0.0584	-	-	-

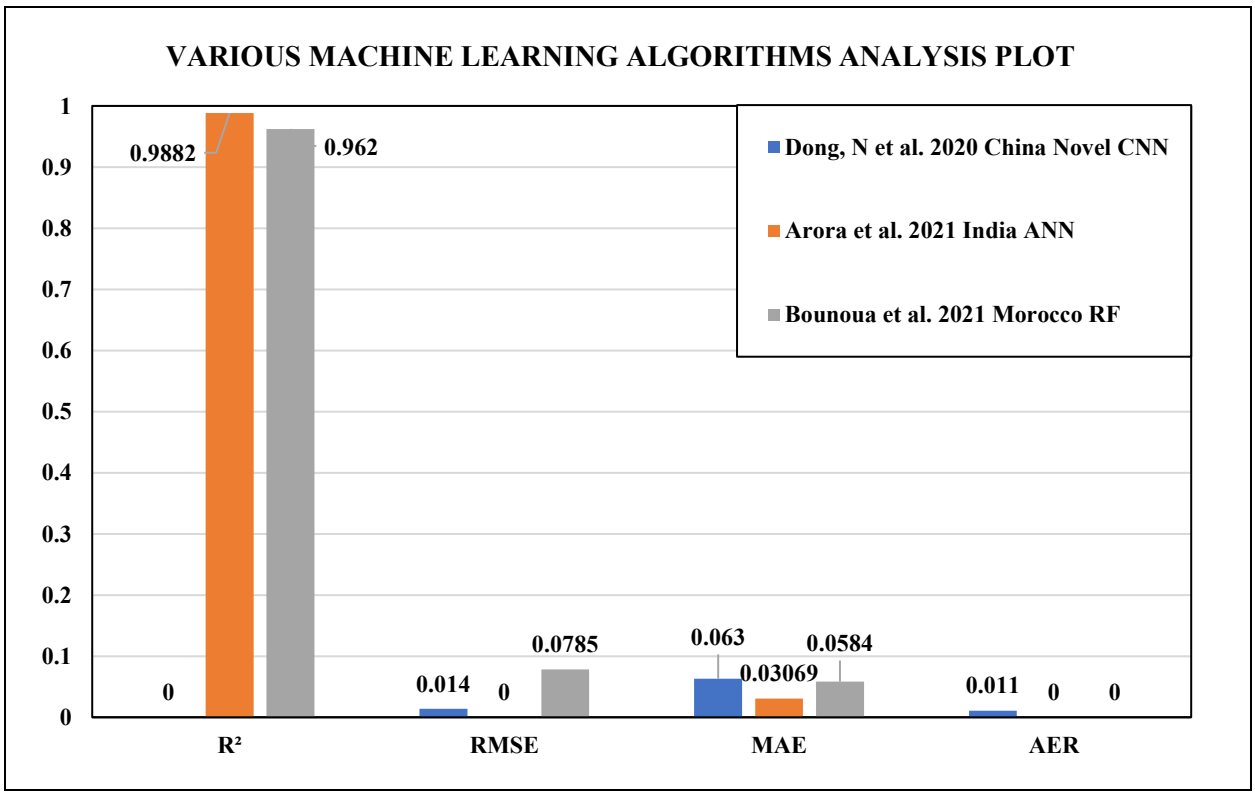


Fig. 1: Comparison analysis plot of various machine learning algorithms

6. CONCLUSION

The ML algorithms in solar radiation prediction have strong potential to improve the precision and efficacy of solar energy systems. This comprehensive observation provides an exhaustive overview of the advancements in solar radiation forecasting, highlighting the benefits, hurdles, and prospective applications of ML algorithms. Forecasting research trajectories and impediments is essential to further refine the precision, interpretability, and scalability of ML techniques for solar radiation prediction. The integration of ML algorithms into solar radiation forecasting serves as a substantial potential to optimize solar energy utilization and progress toward a sustainable future.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest.

COPYRIGHT

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

REFERENCES

Alam, M., Abido, M., El-Amin, I., Fault Current Limiters in Power Systems: A Comprehensive Review, *Energies* 11(5), 1025 (2018). https://doi.org/10.3390/en11051025

Alam, M. S., Al-Ismail, F. S., Salem, A., Abido, M. A., High-Level Penetration of Renewable Energy Sources Into Grid Utility: Challenges and Solutions, *IEEE Access* 8, 190277–190299 (2020). https://doi.org/10.1109/ACCESS.2020.3031481

Arora, I., Gambhir, J., Kaur, T., Data Normalisation-Based Solar Irradiance Forecasting Using Artificial Neural Networks, *Arab. J. Sci. Eng.* 46(2), 1333–1343 (2021). https://doi.org/10.1007/s13369-020-05140-y

Bhatnagar, M., Yadav, A., Swetapadma, A., Enhancing the resiliency of transmission lines using extreme gradient boosting against faults, *Electr. Power Syst. Res.* 207, 107850 (2022). https://doi.org/10.1016/j.epsr.2022.107850

Bounoua, Z., Ouazzani Chahidi, L., Mechaqrane, A., Estimation of daily global solar radiation using empirical and machine-learning methods: A case study of five Moroccan locations, *Sustain. Mater. Technol.* 28, e00261 (2021). https://doi.org/10.1016/j.susmat.2021.e00261

Chaibi, M., Benghoulam, E. M., Tarik, L., Berrada, M., Hmaidi, A. El, An Interpretable Machine Learning Model for Daily Global Solar Radiation Prediction, *Energies* 14(21), 7367 (2021). https://doi.org/10.3390/en14217367

Dong, H., Yang, L., Zhang, S., Li, Y., An Improved Prediction Approach on Solar Irradiance of Photovoltaic Power Station, TELKOMNIKA Indones, *J. Electr. Eng.*, https://doi.org/10.11591/telkomnika.v12i3.4017

Dong, N., Chang, J.-F., Wu, A.-G., Gao, Z.-K., A novel convolutional neural network framework based solar irradiance prediction method, *Int. J. Electr. Power Energy Syst.* 114, 105411 (2020). https://doi.org/10.1016/j.ijepes.2019.105411

Guo, H., Zhuang, X., Chen, P., Alajlan, N., Rabczuk, T., Analysis of three-dimensional potential problems in non-homogeneous media with physics-informed deep collocation method using material transfer learning and sensitivity analysis, *Eng. Comput.* 38(6), 5423–5444 (2022a). https://doi.org/10.1007/s00366-022-01633-6

Guo, H., Zhuang, X., Chen, P., Alajlan, N., Rabczuk, T., Stochastic deep collocation method based on neural architecture search and transfer learning for heterogeneous porous media, *Eng. Comput.* 38(6), 5173–5198 (2022b). https://doi.org/10.1007/s00366-021-01586-2

Prasad, R., Ali, M., Kwan, P., Khan, H., Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation, *Appl. Energy* 236, 778–792 (2019). https://doi.org/10.1016/j.apenergy.2018.12.034

Rabbi, K. M., Nandi, I., Saleh, A. S., Faisal, F., Mojumder, S., Prediction of solar irradiation in Bangladesh using artificial neural network (ANN) and data mapping using GIS technology, In: 2016 4th International Conference on the Development in the in Renewable Energy Technology (ICDRET). IEEE, pp 1–6, (2016).https://doi.org/10.1109/ICDRET.2016.7421482

Ren, Y., Suganthan, P. N., Srikanth, N., Ensemble methods for wind and solar power forecasting—A state-of-the-art review, *Renew. Sustain. Energy Rev.* 50, 82–91 (2015). https://doi.org/10.1016/j.rser.2015.04.081

Sharifi, S. S., Rezaverdinejad, V., Nourani, V., Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: A comparative study of selected temperature-based approaches, *J. Atmos. Solar-Terrestrial Phys.* 149, 131–145 (2016). https://doi.org/10.1016/j.jastp.2016.10.008