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 paramount importance in optimizing the efficacy and dependability of solar power systems. This evaluative manuscript delves into an array of machine learning methodologies applied in solar irradiance prediction, encompassing conventional methodologies. It further delves into the ramifications of meteorological data, geographical and temporal parameters, data preprocessing, feature curation methodologies, amalgamated learning methodologies, and composite models. Moreover, the manuscript investigates the latent advantages of amalgamating orbital visuals and terrestrial sensors to augment prognostic capacities. Ultimately, the review underscores the exigency for broader datasets, refined feature curation methodologies, and composite models to attain heightened precision and resilience.

Keywords: Renewable energy; Predictive modeling; Computational methods; Environmental science; Climate change adaptation.

1. INTRODUCTION

The elevation of solar power as a burgeoning renewable repository portends favourably for abating reliance on fossil fuels. Prudent anticipation of solar irradiance emerges as an imperative exigency for the seamless operation of solar energy infrastructures. Conventional methodologies bespeak limitations in this realm; nevertheless, the advent of machine learning algorithms denotes a metamorphic shift, bestowing heightened precision and efficacy. This evaluative manuscript proffers an exhaustive scrutiny of the contemporary utilizations of machine learning techniques in the field of solar irradiance assessment.

2. OVERVIEW OF SOLAR RADIATION FORECASTING

Recently, clean energy goals have become vital to global energy politics. These aims increase renewable energy consumption to reduce the release of greenhouse gases and mitigate climate change. UV forecast optimises solar power resource use in sustainable energy management. The use of machine learning has been used to create reliable ultraviolet (UV) forecast models.

But connecting renewable energy into the electricity infrastructure is difficult. The temporary nature of energy from renewable sources generation can cause energy imbalances in supply and demand. Smart grid solutions are being investigated to better coordinate and control renewable energy production and utilisation. This issue is considerably more important when integrating a lot of energy from renewable sources into the system (Alam *et al.* 2020). But getting energy from renewable sources into the electricity system is complicated.

The transient nature of energy from environmentally friendly generation might produce energy disparities in both demand and supply. Smart grid solutions are being researched in order to coordinate and control renewable energy generation and consumption. This issue is substantially more critical when adding a lot of energy from sources that are renewable into the grid (Alam *et al.* 2020). Another hurdle in using green power into the electrical grid is a problem of fault currents. Fault currents can develop at unusual circumstances, such as shorts in wiring, and can cause harm to the electrical systems equipment. To minimise this issue, electrical fault limiters are being researched as an option (Alam *et al.* 2018). Such devices can help manage the movement of currents caused by faults and protect the electrical system foundation.

Recent years have witnessed a burgeoning fascination with the amalgamation of wind and solar energy infrastructures to facilitate grid stability management. The orchestration of frequency control emerges as an indispensable element in upholding the equilibrium and reliability of the electrical grid. Various regulation strategies and techniques have been presented to overcome the issues associated with frequency regulation in both wind and solar power combined systems (Ren *et al.* 2015). These advances show ongoing investigation particularly innovation in the field of green energy integrating and its influence on reliability in the grid.

Within the domain of solar power prognostication, a plethora of algorithms and methodologies have been devised to meticulously forecast solar irradiance. A good example is the multi-stage multimodal mode decomposition approach (MEMD) linked with the optimization of ants colonies (ACO) and a random woodland (also known as model, as recommended by (Dong *et al.* 2014). This approach aims to predict monthly solar radiation through the employment of MEMD. Subsequently, the technique known as ACO is employed to select the most pertinent IMFs, which then act as input features for the RF model.

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

3. LEARNING MACHINE TECHNOLOGIES FOR UV ESTIMATION

Statistical algorithms (ML) were computational models capable of automatically discover patterns and generate predictions given data without being specifically programmed. ML algorithms can manage big datasets, capture complicated connections, and adjust to changing situations, making them excellent for ultraviolet forecasting. Some ML algorithms used for UV forecasting included artificially neural networks (ANNs), supported vector machines (SVMs), random woods (RFs), and gradient boosting computers (GBMs).

Probabilistic techniques (ML) were mathematical structures able to automatically discover patterns and create prediction based on information without being specifically designed. Machine learning (ML) systems can manage vast datasets, grasp complicated linkages, and respond to changing scenarios, making them perfect for UV forecasting. Some machine learning (ML) approaches utilized for UV forecasting were artificially constructed neural networks (ANNs), sponsored vector machine models (SVMs), unpredictable woods (RFs), and gradient-enhancing computers (GBMs).

Solar irradiance predicting plays a key role in different programs, such as producing electricity from the sun and weather prediction. Researchers have created 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. Likewise, (Sharifi *et al.* 2016) completed the evaluation and comparison of various temperature centric methodologies for prognosticating monthly global solar irradiance. They contrasted wavelet regression, artificial neural network (ANN), expression gene scripting (GEP), and empirical models, illustrating the merits and shortcomings of each approach.

4. SOURCES OF DATA AND PRECONDITIONING

Accurate sunlight forecasting relies on superior information for input, including historical sunlight radiation metrics, weather information, and satellite imagery. Data preparation procedures, such as cleaning of data, 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. Finally, (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 for algorithms for learning 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 optimising the effectiveness as well as effectiveness of solar thermal power plants. In the past few decades, AI models have attracted considerable interest for sunlight forecast due to their capacity to handle complicated and nonlinear interactions. An example of this artificial intelligence (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 amalgamated frameworks merging machine cognition algorithms and optimization methodologies have been suggested. These frameworks endeavour to heighten the precision of solar radiation prognostications by taking into account variables such as atmospheric circumstances, input parameters, and optimization paradigms. Moreover, conglomerate characteristic curation methodologies are recommended to ameliorate prognostication precision. In the realm of artificial intelligence (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 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 modeling (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 up new opportunities for handling complicated challenges across numerous domains. A number of 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.

Moreover, (Bhatnagar *et al.* 2022) concentrated 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 optimizer of wind and solar powered systems.

The authors demonstrated the usefulness of the IFC-GBR approach through contrasting how well it performed with other regularly used regression algorithms. Studies demonstrated that the IFC-GBR technique beat 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 modeling approaches and have ramifications for numerous sectors which include the chemical engineering field and material science. In summary, the IFC-GBR methodology offers a promising methodology for properly estimating the indeterminate diffusion activity index of chloroform in ionic liquids.

The researchers (Arora *et al.* 2021) delved into the utilization of artificial intelligence 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 that were entered such as bridge length of sentence, span, and design aspects. The results indicated that the machine learning algorithm outperformed traditional estimation approaches, offering more accurate cost estimates and lowering the time necessary for calculations. This research emphasizes the potential of machine learning 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 machines learning alongside empirical algorithms to foresee average global sun radiation. The results suggested that these mathematical frameworks highly effective at forecasting solar radiation. Besides to computer techniques for learning, batch feature selection has also been applied for sunlight forecasting. Furthermore, the application of the following parameters and Bayesian optimization strategies was also suggested in train arrival delays predicting (Bounoua *et al.* 2021). This model can be employed to increase the quality of railway operations by anticipating 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 neural network techniques (ANNs).

Furthermore, (Bounoua *et al.* 2021) completed a case study for Spain that calculated monthly world radiation from the sun applying observational and machine-learning methods. Their findings indicated that machine-learning technologies, including ANNs, is capable of accurately estimating the sunlight and surpass classic empirical models. These findings underscore the need of adopting modern methodologies, for instance Artificial Neural Networks toward accurate solar irradiance predictions. Through the employment of algorithms leveraging machine learning and normalization techniques to process the data, scholars can enhance the accuracy and reliability of solar irradiance prognosis, thereby facilitating the more effective utilization of sustainable energy reservoirs.

5. RESULT

The concise presentation depicts the comparative effectiveness of various machine learning algorithms in forecasting solar radiation. Metrics including error standard deviation, R^2 performance, and root mean squared error for each algorithm are encapsulated in tabular form. Moreover, a visual representation, portrayed as a comparative analysis plot, vividly showcases the R^2 performance and Root Mean Squared Error for each algorithm, enabling a clear comparison of their predictive capabilities.

Table 1. The confluence of various machine learning methodologies employed in the prognostication 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	-	-	-

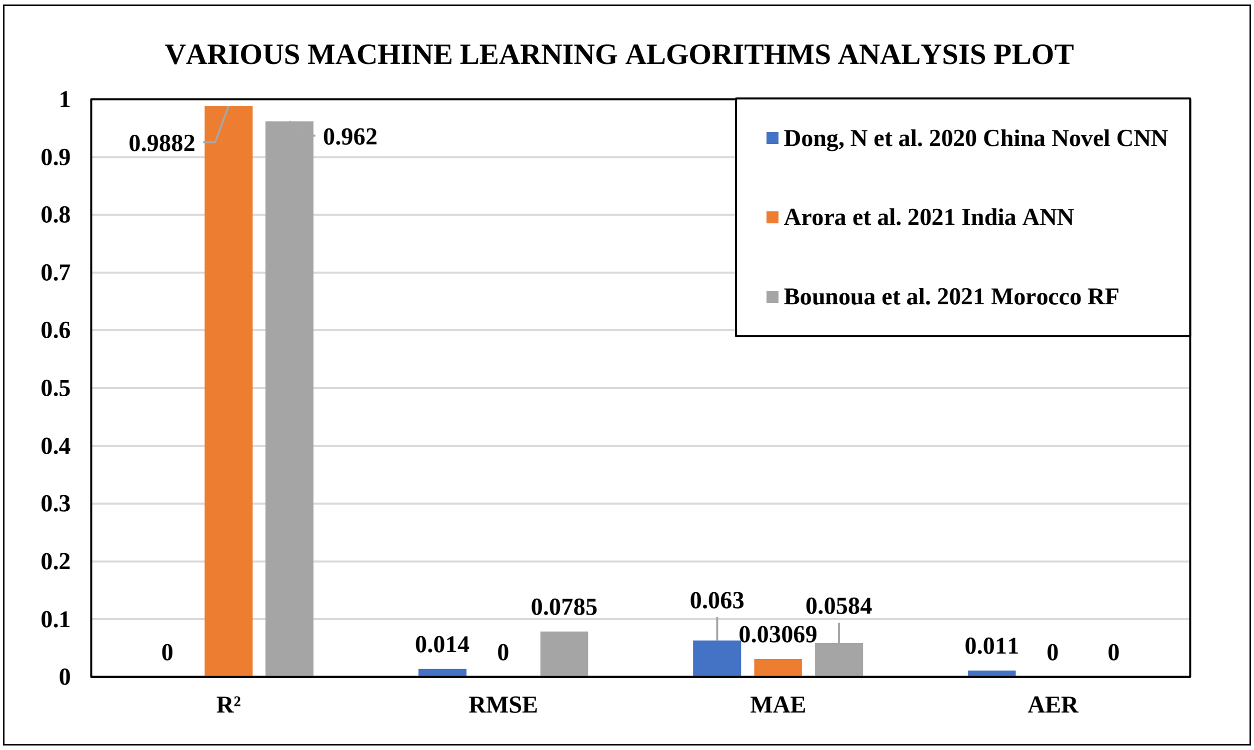

Fig. 1: The comparison analysis plot of various machine learning algorithms

6. CONCLUSION

Machine learning (ML) algorithms in solar radiation prediction signifies a promising avenue to heighten the precision and efficacy of solar energy systems. This comprehensive scrutiny delivers an exhaustive overview of the cutting-edge in solar radiation forecasting, accentuating the benefits, hurdles, and prospective applications of ML algorithms. Foreseeing future research trajectories and impediments is pivotal to further refine the precision, interpretability, and scalability of ML techniques for solar radiation forecasting. In essence, the integration of ML algorithms into solar radiation forecasting harbours substantial potential to optimize solar energy utilization and hasten the progression 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.

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