

Machine Learning–Based Prediction of Daily Solar Radiation to Support Renewable Energy Development in Coastal Regions

Mfon D. Umoh,^{1,*} Udoh F. Evans,¹ Sunday A. Akpan,² Sunday E. Otene,² and Akinola B. Olanrewaju¹

1: Directorate of Strategy, Research and Development, Maritime Academy of Nigeria, Oron, Nigeria

2: Department of General Studies, Maritime Academy of Nigeria, Oron, Nigeria

Received January 12, 2026; Accepted February 3, 2026; Published March 5, 2026

Reliable estimation of surface solar radiation is essential for climate analysis and solar energy planning, particularly in data-limited regions such as coastal Africa. This study investigates the long-term variability of surface solar radiation and evaluates the performance of machine learning models for its prediction using a comprehensive reanalysis dataset spanning 1940–2024. Five radiation components—net surface solar radiation (SSR), clear-sky net radiation (SSRC), downward surface solar radiation (SSRD), clear-sky downward radiation (SSRDC), and total surface radiation (TSR)—were analyzed to quantify the influence of atmospheric attenuation caused by clouds, aerosols, and water vapor. Five machine learning algorithms—Linear Regression (LR), Gradient Boosting (GB), Random Forest (RF), k-Nearest Neighbours (KNN), and Artificial Neural Network (ANN)—were implemented and evaluated using train–test split, k-fold cross-validation, and leave-one-out validation. The results reveal strong interannual and multi-decadal variability in solar radiation, with clear-sky radiation consistently exceeding all-sky radiation, confirming the dominant role of atmospheric modulation in the region. Among the tested models, Linear Regression achieved near-perfect predictive performance ($R^2 \approx 1.0$) with the lowest error statistics, indicating that surface solar radiation over coastal Africa is largely governed by linear radiative processes. Gradient Boosting and Random Forest also demonstrated high accuracy ($R^2 > 0.98$), while the Artificial Neural Network showed poor generalization due to overfitting. The findings demonstrate that computationally efficient and physically interpretable machine learning models can reliably estimate long-term solar radiation in coastal Africa. This provides a robust scientific basis for solar resource assessment, photovoltaic system design, and climate-resilient renewable energy planning across the region.

Keywords: Solar radiation; Machine learning; Linear regression; Ensemble methods; Coastal Africa; Renewable energy prediction

Introduction

A clean and dependable alternative to fossil fuels, solar energy is employed all around the world. Only about 5×10^4 EJ of the enormous amount of energy that the sun's beams send to Earth is readily harvestable ($1 \text{ EJ} = 10^{18} \text{ J}$) [1]. About 70% of the 342 W/m^2 of solar energy that enters the earth's atmosphere may be harvested, with the

*Corresponding author: mfonslago20@gmail.com

remaining 30% being scattered or reflected back into space. In addition to other factors like sun orientation, aerosol, temperature, wind speed, direction, and many more, the earth's rotation around the sun causes variations in solar irradiance, making solar energy an unpredictable resource [1]. The energy that the Sun emits and that travels across space to reach Earth is referred to as solar radiation. It is made up of electromagnetic waves, such as infrared (IR) radiation, ultraviolet (UV) rays, and visible light. Only a small portion of the solar energy that the Sun emits in all directions makes it to the Earth's atmosphere. The Earth's surface eventually receives some of the solar energy [2]. The critical need for energy diversification is highlighted by the depletion of traditional energy resources like fossil fuels, oil, and gas, as well as rising pollution and the consequences of climate change. A sustainable answer to these urgent problems is provided by incorporating renewable energy sources into our energy mix. We may lessen environmental damage, fight climate change, and lessen our dependency on finite fossil fuels by utilizing renewable energy sources including solar, wind, hydro, and geothermal power [3]. Accurate insolation forecasting is essential for building energy conservation. As a result, a great deal of research has been done to develop trustworthy predictive models for insolation employing machine learning techniques in addition to statistical and physical models. This is essential for reducing greenhouse gas emissions and preventing climate change, as well as for smart cities and net-zero energy buildings [4]. One area of artificial intelligence is machine learning. It is a collection of methods for examining and identifying trends in input data in order to draw conclusions or predictions. The measurement of solar radiation is being revolutionized by artificial intelligence (AI) and machine learning (ML), which are improving efficiency and accuracy. As a subset of AI, machine learning (ML) focuses on teaching algorithms to identify patterns and make predictions based on data, while artificial intelligence (AI) refers to the creation of computer systems that are capable of carrying out tasks that normally require human intelligence. In order to estimate solar radiation, machine learning models examine enormous volumes of environmental and meteorological data to provide highly accurate predictions about solar energy potential [3]. Image identification, natural language processing, robotics, spam filtering, drug development, medical diagnosis, financial analysis, bioinformatics, marketing, and even politics are among the fields in which it finds use. Bayesian networks, artificial neural networks, deep learning, clustering, and Support Vector are a few examples of traditional machine learning techniques [5]. In recent years, Solar Radiation (SR) prediction has gained special attention in academia and a wide range of ML models have been used to forecast SR [6]. In Reference [6], authors reported Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) to estimate SR using ML models. Comparative evaluation using prediction accuracy, stability, and computational time showed that XGBoost outperformed SVM for estimating SR. These ML techniques help handle large amounts of data and studying the complex pattern which makes them a promising and suitable choice for solar irradiance prediction [7]. Other prediction methods of solar radiation are based on the artificial neural network methods, linear regression, probabilistic methods, network monitoring data, fuzzy method approach, ANFIS, Multiple Linear Regressions (MLR) method, and some other methods. The conventional method commonly used for prediction is the multiple regression method. This method can be analyzed by using several independent variables so that the obtained results are more accurate [8]. The aim of this research is to predict coastal solar radiation data using Machine learning algorithms.

Machine Learning Algorithms

Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) are a class of computational models inspired by the biological neural networks of the human brain. ANNs consist of interconnected layers of nodes (neurons), where each connection has an associated weight that is adjusted during training to minimize prediction error [9].

- Structure: ANNs typically have three main layers (Figure 1):
 1. Input layer – receives the input features.
 2. Hidden layers – perform non-linear transformations through activation functions (*e.g.*, sigmoid, ReLU, tanh).
 3. Output layer – produces the final prediction (classification or regression).

Artificial Neural Network (ANN) Architecture

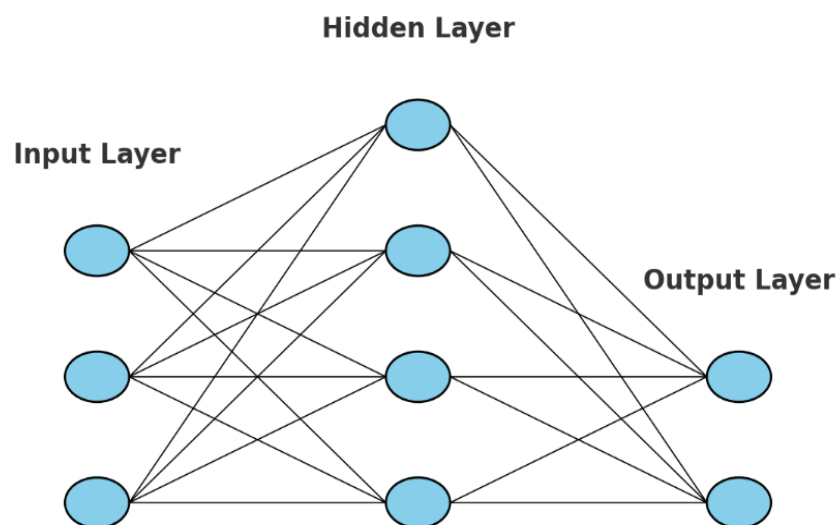


Fig 1. Artificial Neural Network (ANN) Architecture

Linear Regression

Linear Regression is one of the simplest and most widely used supervised learning algorithms for predicting a continuous output based on one or more input features. It models the linear relationship between the independent variable(s) X_n and the dependent variable Y_n .

1. Simple Linear Regression: Uses one independent variable

$$Y = a + bX$$
2. Multiple Linear Regression: Uses two or more independent variables

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$
 where:

Y = predicted value

X = input feature(s)

a = intercept

b = coefficient(s)

Modern applications of linear regression extend to healthcare analytics, finance, and environmental modeling, often enhanced with regularization techniques like Ridge and Lasso regression to handle overfitting and multicollinearity [10].

Gradient Boosting Algorithm

Gradient Boosting is a supervised machine learning technique that builds predictive models by combining an ensemble of weak learners, usually decision trees. Gradient boosting is a method standing out for its prediction of speed and accuracy, particularly with large and complex datasets. Gradient boosting is excellent for various machine learning tasks, including regression and classification. It provides high predictive accuracy by combining weak learners in an ensemble, such as decision trees, to create a strong predictor. Gradient boosting handles complex relationships in data, works well with both numerical and categorical features, and is particularly useful for dealing with largescale structured and unstructured data in real-world applications [11].

K-Nearest Neighbors (KNN)

In machine learning, the k-nearest neighbor method (k-NN) is also known as lazy learning because of its training is held up to run time. This classifier is also one of the most straightforward and simplest since classification of the datasets is based on their nearest neighbor class. The data sets are consequently allocated to the class that's more similar and k must be a positive integer. The value of k is usually small. When $k=1$, the datasets are basically allocated to the class of its nearest neighbor [12].

Random Forest

The Random Forest (RF) algorithm predicts a continuous outcome using an ensemble technique. It aggregates the predictions of multiple decision trees, improving predictive accuracy and preventing overfitting. This method, also known as bagging, generates numerous subsets of the original training data, allowing each tree in the forest to learn from a different portion of the data. Each tree selects different features to introduce randomness and ensure diversity in the model [13].

Materials and Methods

Materials

The data used for this research paper consist of Net Surface Solar Radiation data (ssr), Net Surface Solar Radiation considering clear sky (ssrc), Net Surface Solar Radiation downward (ssrd), Net Surface Solar Radiation downward considering clear sky (ssrdc) and total surface Solar radiation (tsr) data. The data covers a period of eighty-four (84) years (1940 – 2024). The data which was downloaded from the CAMS-Copernicus Atmospheric data store (<https://ads.atmosphere.copernicus.eu/>), represents the coastal regions of Africa. The satellite data was downloaded as a daily recoded data. For the purpose of this research, the data is now presented and used as a monthly average data.

Method

The data was downloaded as a daily recorded data. If used this way, the dataset would be too large; therefore, the authors decided to take the monthly average of the data from 1940 to 2024. The data was then cleaned and presented in a format that will be easy to process. Python programming language was then used for the analysis of the data. Google Colaboratory (<https://colab.research.google.com/>) was then used as the Python platform for the analysis. All the graphs were also plotted using Python programming Language. The Machine Learning Algorithms used for the prediction are; Linear Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN) and Gradient Boosting (GB).

Results and Discussion

Discussion of Graphical Results

Figure 2 presents the time series trends of net surface solar radiation (SSR), net surface solar radiation under clear-sky conditions (SSRC), and net surface solar radiation downward (SSRD) over the study period (1940–2024). The graph shows noticeable temporal variability, reflecting both seasonal cycles and interannual fluctuations in solar radiation in coastal regions of Africa. SSR and SSRC generally follow similar patterns, with clear-sky values consistently higher, as expected due to the absence of atmospheric interference such as aerosols, clouds, and moisture. The SSRD series further highlights how atmospheric conditions attenuate solar flux before reaching the Earth's surface. The long-term variations observed may also be linked to broader climatic shifts and regional atmospheric changes.

Figure 3 illustrates the trends of net surface solar radiation downward under clear-sky conditions (SSRDC) and total surface radiation (TSR). The SSRDC curve provides a baseline of the maximum possible solar input in the absence of atmospheric obstructions, while TSR reflects the actual usable energy at the surface. The consistent gap between these two curves highlights the influence of atmospheric constituents, particularly cloud cover and aerosols, on solar availability. Importantly, both graphs confirm the significant role of atmospheric variability in solar radiation dynamics, underscoring the need for robust predictive models to accurately estimate insolation for renewable energy applications.

Together, the two figures provide evidence that solar radiation in coastal Africa is strongly influenced by atmospheric conditions, with clear distinctions between modeled clear-sky and actual surface radiation. These patterns justify the use of machine learning approaches in this study, as they can capture nonlinear relationships between meteorological inputs and solar radiation outcomes. Moreover, the long-term dataset (spanning more than eight decades) ensures that the models are trained on diverse conditions, enhancing their generalizability.

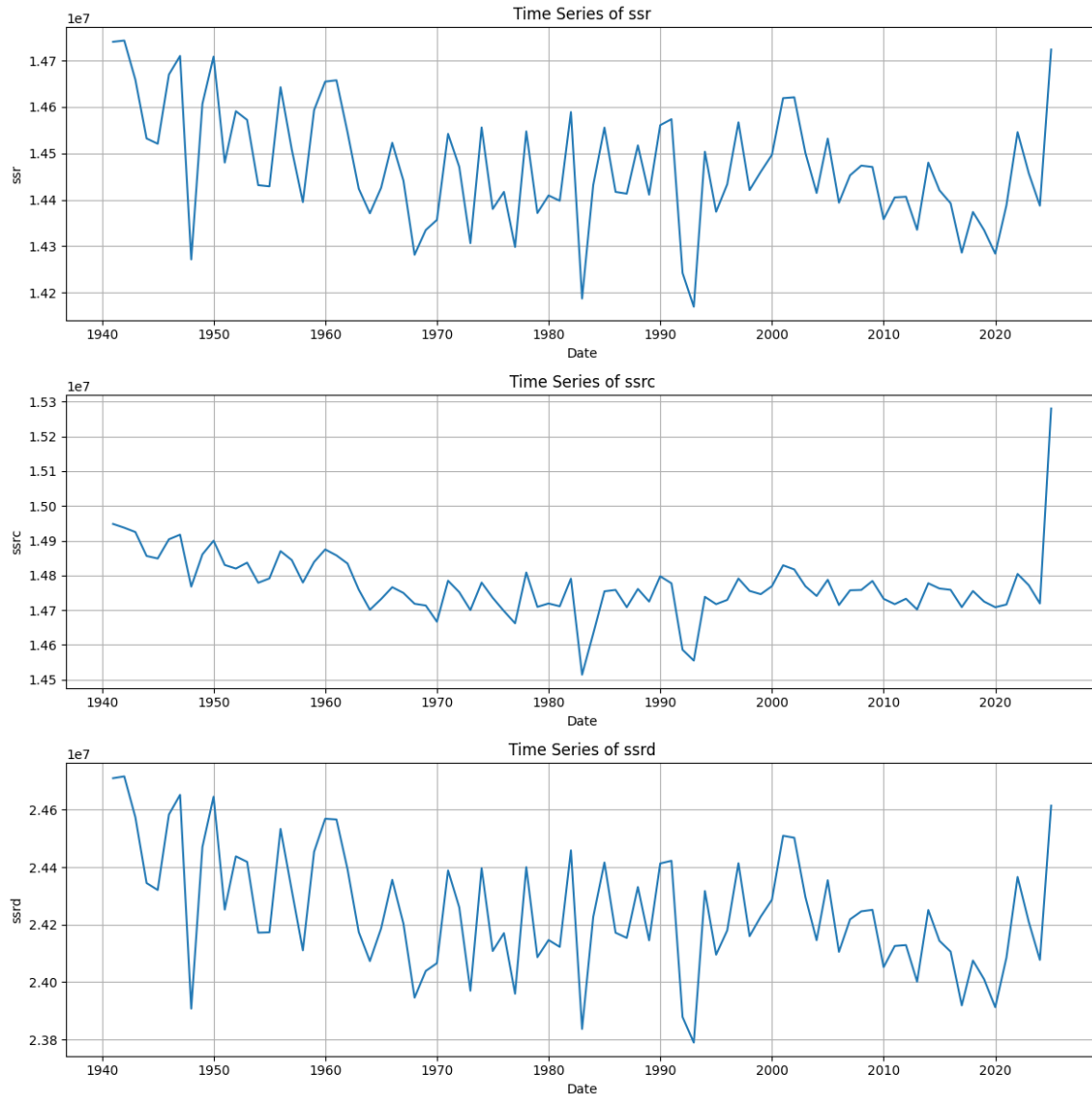


Fig 2. Time series graph of ssr, ssrc and ssrd

Model Evaluation

For models to be evaluated effectively, data has to be split into training and test sets. This is done to measure how well the model generalizes to new, previously unseen data. The main aim is not to determine how well the model fits the training set, but rather in how well it can make predictions for data that was not observed during training. The model evaluation based on train/test split of the data is shown in Table 1.

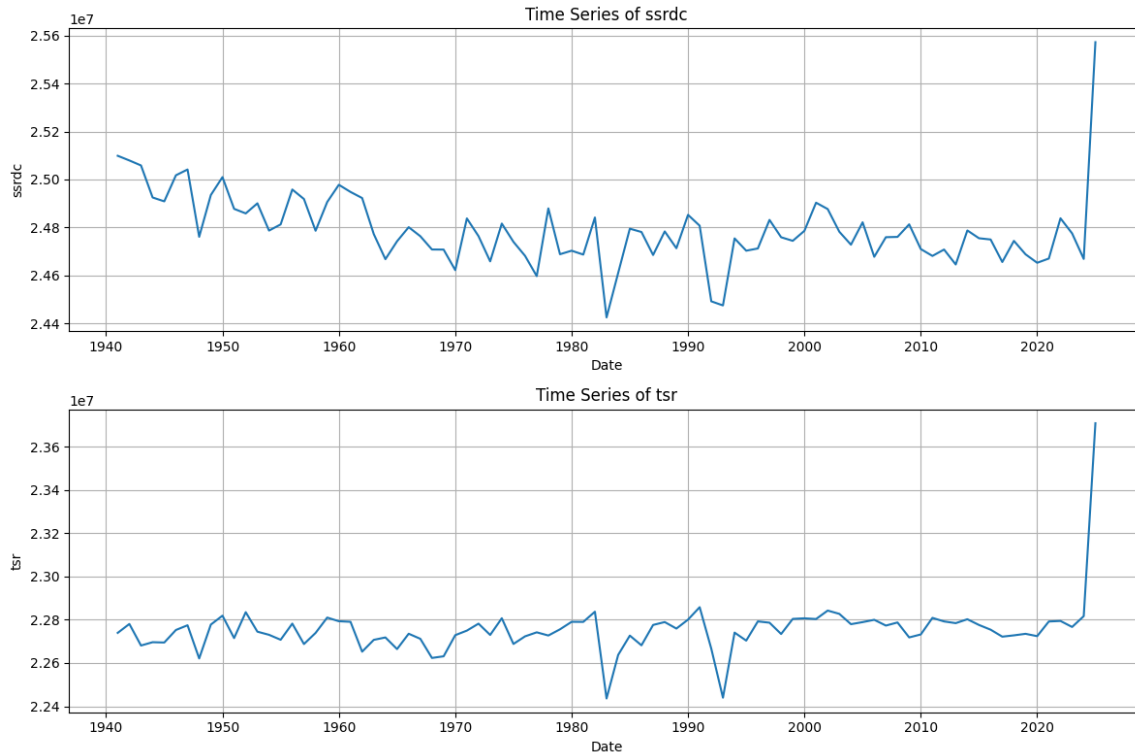


Fig 3. Time series graph of ssrdc and tsr

Table 1. Model evaluation

<i>ML Algorithms</i>	<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	<i>R²</i>
ANN	36579758540	191258.4	115195.9	0.008	-1.125
KNN	739673856	27196.9	16522.5	0.001	0.957
RF	730975651	27036.6	17844.8	0.001	0.98
GB	337703959	18376.7	11461.4	0.001	0.98
LR	3044376	1744.8	1282.3	0	1

Cross Validation

Cross validation is a statistical method of evaluation generalization performance that is more stable and thorough than using a split into training and test set. In cross validation, the data is instead split repeatedly and multiple models are trained. The most commonly used version of cross validation is k-fold cross validation, where k is a user specified number, usually 5 or 10. For this research, the k value was set to 5. The model evaluation based on 5 fold cross validation is shown in Table 2. When performing five-fold cross validation, the data is first partitioned into five parts of approximately equal size, called folds. Next, a sequence of models is trained. The first model is trained using the first fold as the test set, and the remaining folds (2-5) are used as training set. The model is built using data in folds 2-5, and then the accuracy is evaluated on fold 1. Then another model is built, this time using fold 2 as the test set and the data in folds 1, 3, 4, and 5 as the training set. This process is repeated using folds 3, 4, and 5 as the test sets.

Table 2. Evaluation metrics with cross validation

<i>ML Algorithms</i>	<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	<i>R²</i>
ANN	18505527167	136035	103520.5	0.007	-0.18
KNN	430854914.6	20757	13643.9	0.001	0.973
RF	283524785	16838.2	12218.2	0.001	0.982
GB	89874932	9480.2	7500.4	0.001	0.994
LR	3596473.5	1896.4	1410.1	0	1

Stratified K-Fold Cross Validation

In Stratified k-fold cross validation, the data is split such that the proportions between the classes are the same in each fold as they are in the whole dataset. For this research the Stratified method with number of folds equal to five (5) was used. Random sampling was used with repeat train and test set = 10. Training set size set to 66%. The evaluation is shown in Table 2.

Stratified K-Fold Cross Validation With Leave – One - Out

Another frequently used cross validation method is leave one out. Leave - one – out cross validation is a k-fold cross validation where each fold is a single sample. For each split, you pick a single data point to be the test set. The evaluation metrics using leave-one-out is shown in Table 3.

Table 3. Evaluation Metrics using Leave-One-Out

<i>ML Algorithms</i>	<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	<i>R²</i>
ANN	61138249243	247261.5	120244.6	0.008	-2.897
KNN	346772482	18621.3	12382.6	0.001	0.978
RF	280185088	16738.7	13069.2	0.001	0.982
GB	110481999	10511	7361.6	0.001	0.993
LR	3099824	1760.6	1341.1	0	1

Temporal Variability of Surface Solar Radiation

Figures 2 and 3 illustrate the long-term evolution of surface solar radiation components—net surface solar radiation (SSR), clear-sky net radiation (SSRC), downward surface solar radiation (SSRD), clear-sky downward radiation (SSRDC), and total surface radiation (TSR)—over coastal Africa from 1940 to 2024. The time series exhibits pronounced interannual and multi-decadal variability, indicating strong modulation by atmospheric processes and regional climate dynamics. The persistent separation between clear-sky radiation (SSRC and SSRDC) and all-sky radiation (SSR, SSRD, and TSR) highlights the dominant influence of clouds, aerosols, and atmospheric water vapor in controlling surface solar energy availability.

The magnitude of atmospheric attenuation is substantial, implying that a large fraction of incoming solar radiation is absorbed or scattered before reaching the surface. The relative stability of the clear-sky radiation signal compared with the higher variability of all-sky radiation further suggests that atmospheric processes, rather than changes in solar geometry, dominate regional solar variability.

The strong coherence between SSR and SSRC indicates that although atmospheric attenuation is large, it follows systematic and physically constrained patterns.

Performance of Machine Learning Models

The predictive performance of the machine learning models was assessed using three validation strategies: train–test split, k-fold cross-validation, and leave-one-out cross-validation. Across all evaluation frameworks, Linear Regression (LR) consistently outperformed the other models, achieving an R^2 value close to unity and the lowest error statistics (MSE, RMSE, MAE, and MAPE). This indicates an almost perfect agreement between observed and predicted solar radiation values.

The superior performance of LR implies that the relationship between atmospheric variables and surface solar radiation is largely linear. Consequently, the radiation field over coastal Africa appears to be governed by stable and predictable atmospheric controls.

Tree-based ensemble models, particularly Gradient Boosting (GB) and Random Forest (RF), also exhibited excellent performance with R^2 values exceeding 0.98 across all validation schemes. These models successfully captured secondary nonlinear effects arising from complex aerosol–cloud–radiation interactions. However, their slightly lower accuracy relative to LR indicates that nonlinearities play a secondary role compared with the dominant linear radiative processes.

In contrast, the Artificial Neural Network (ANN) performed poorly, with negative R^2 values across all validation approaches. This behavior reflects overfitting and weak generalization.

Cross-Validation and Model Robustness

The robustness of the machine learning models was further evaluated using k-fold and leave-one-out cross-validation. Linear Regression maintained near-perfect predictive skill under all data partitioning schemes, demonstrating exceptional stability and generalization capability. Similarly, Gradient Boosting and Random Forest models exhibited high consistency, confirming their suitability for long-term solar radiation modeling.

The strong agreement between training and cross-validated performance indicates that the long temporal span of the dataset (1940–2024) captures both short-term atmospheric variability and long-term climate trends. This minimizes sampling bias and enhances confidence in the reliability of the predictions. Figures 4 to 6 further confirm this robustness by showing close agreement between observed and model-predicted radiation values.

The observed and predicted values of ssr obtained from the different models were plotted. Figure 4 shows the observed and predicted values of ssr based on the K-NN model. Figure 5 shows the observed and predicted values of ssr based on the LR, Random forest and gradient boosting models. Figure 6 shows the observed and predicted values of ssr based on the ANN model.

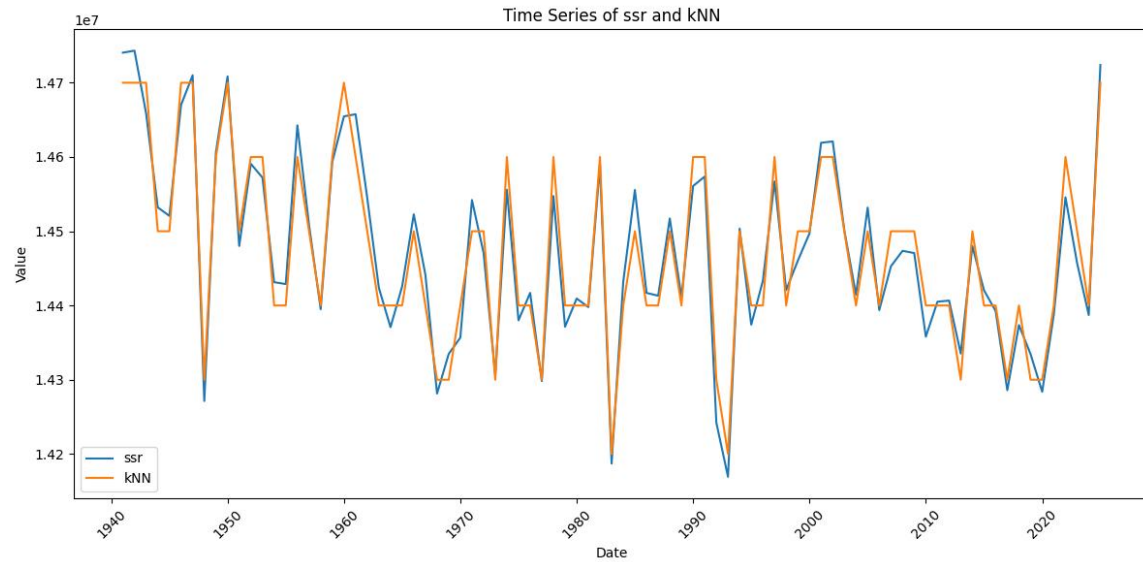


Fig 4. Observed and predicted values of ssr based on the KNN model

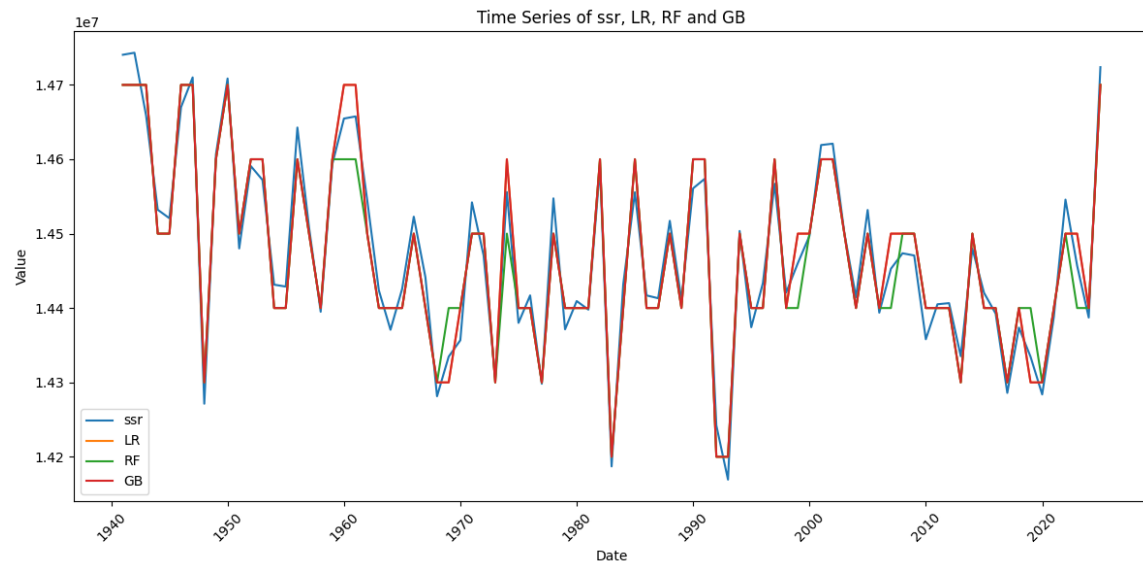


Fig 5. Observed and predicted values of ssr based on LR, random forest and gradient boosting models

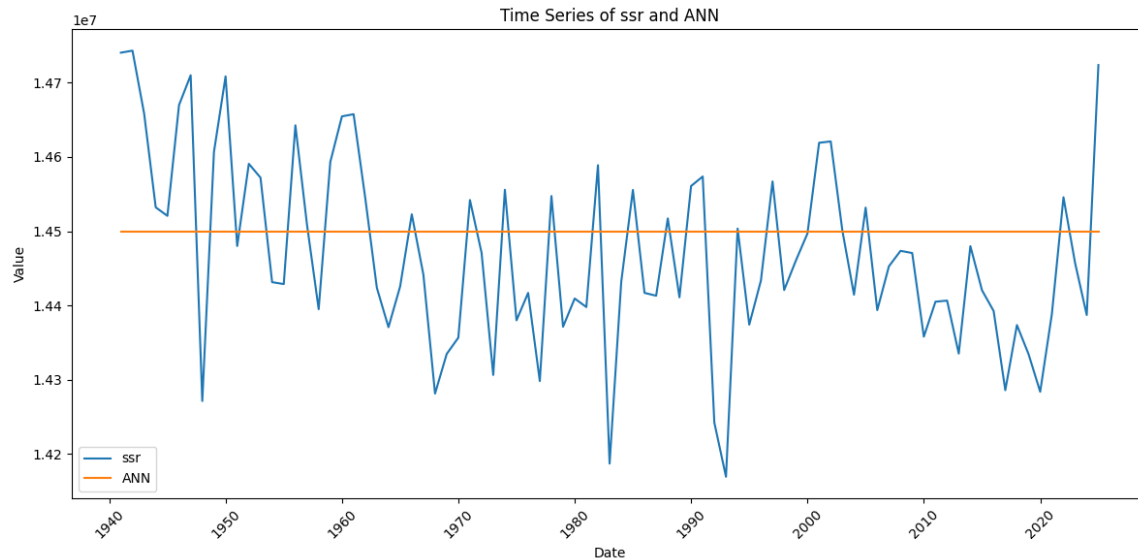


Fig 6. Observed and predicted values of SSR based on ANN model.

Implications for Solar Energy Development

The exceptionally high accuracy achieved by the machine learning models has important implications for renewable energy development in coastal Africa. Reliable estimation of surface solar radiation is fundamental for photovoltaic system design, solar farm siting, and grid integration. The dominance of linear behavior further simplifies operational forecasting, allowing policymakers and energy planners to deploy computationally efficient and interpretable models without compromising accuracy.

Moreover, the use of a long-term climate dataset ensures that the predictions remain robust under changing atmospheric and climatic conditions, making the modeling framework suitable for climate-resilient solar energy planning across the region.

CONCLUSIONS

This study investigated the prediction of daily solar radiation in coastal regions of Africa using multiple machine learning algorithms, including Linear Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), and Gradient Boosting (GB). The results revealed that Linear Regression outperformed all other models, consistently achieving an R^2 of 1.0 and the lowest error values across different evaluation techniques. This indicates that the relationship between solar radiation and the selected meteorological inputs is strongly linear.

Among the nonlinear models, Gradient Boosting and Random Forest demonstrated excellent performance, achieving R^2 values above 0.98, with Gradient Boosting slightly outperforming Random Forest in terms of error minimization. KNN also produced competitive results but with comparatively higher error values. In contrast, the Artificial Neural Network underperformed, likely due to the dataset's relatively low complexity and the model's need for larger data volumes and advanced parameter optimization.

Based on these findings, the following recommendations are made:

1. Adoption of Linear Regression for practical applications: Given its simplicity, efficiency, and superior accuracy, LR is recommended for solar radiation prediction in coastal regions where interpretability and low computational cost are important.
2. Use of ensemble methods for enhanced accuracy: Gradient Boosting and Random Forest should be considered in applications where small-scale variations and nonlinear interactions are critical, such as in smart grid integration, solar farm optimization, and real-time energy forecasting.
3. Further exploration of deep learning techniques: While ANN underperformed in this study, future work using larger datasets (*e.g.*, high-resolution hourly or minute-level data) and optimized architectures may improve its predictive accuracy and better capture nonlinear patterns.
4. Integration into energy planning and policy: Reliable solar radiation prediction models can support policymakers, energy developers, and utility operators in optimizing solar energy utilization, improving grid stability, and reducing reliance on fossil fuels in coastal African regions.

In conclusion, this study demonstrates that machine learning methods, particularly Linear Regression and ensemble models, provide accurate and reliable tools for solar radiation prediction. Their application holds significant potential for advancing renewable energy integration, improving efficiency in solar energy systems, and supporting sustainable development goals in Africa.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interests regarding the publication of this paper.

ACKNOWLEDGMENTS

The authors acknowledge that artificial intelligence (AI)–based tools were used during the preparation of this manuscript for language refinement and drafting support. All scientific content, interpretations, data accuracy, and references were reviewed, verified, and approved by the authors, who take full responsibility for the integrity and validity of the work. AI tools were used solely as editorial assistance and did not contribute to the study design, data analysis, or scientific conclusions.

REFERENCES

- [1] Tajjour, S., Chandel, S. S., Alotaibi, M. A., Malik, H., Márquez, F. P. G., & Afthanorhan, A. (2023). Short-Term Solar Irradiance Forecasting Using Deep Learning Techniques: A Comprehensive Case Study. *IEEE Access*, 11, 119851-119861. doi:<https://doi.org/10.1109/ACCESS.2023.3325292>
- [2] Hissou, H., Benkirane, S., Guezzaz, A., Azrour, M., & Beni-Hssane, A. (2023). A Novel Machine Learning Approach for Solar Radiation Estimation. *Sustainability*, 15(13), 10609. doi:<https://doi.org/10.3390/su151310609>

- [3] Benatallah, D., Zahouani, A., Yahiaoui, A., Djeldjli, H., Nasri, B., Benatallah, A., & Benabdelkrim, B. (2024). Estimating of Solar Radiation Based on Machine Learning Approaches under Algerian Desert Climate. *Instrumentation Mesure M étrologie*, 23(5), 363-374. doi:<https://doi.org/10.18280/i2m.230504>
- [4] Sinha, N., Kulkarni, R., Thakkar, P., & Shiragapur, B. (2023). DEEP LEARNING BASED MODEL FOR PREDICTION OF INCOMING SOLAR RADIATION. *Journal of Engineering Technologies and Innovative Research*, 10(5), 285-289.
- [5] Senekane, M., & Taele, B. M. (2016). Prediction of solar irradiation using quantum support vector machine learning algorithm. *Smart Grid and Renewable Energy*, 7(12), 293-301. doi:<http://dx.doi.org/10.4236/sgre.2016.712022>
- [6] Abdelhag, M. E., Elmarud Ali, S. E., Amin, S. T., Khan, A. A., & Metwally, A. A. (2022). Machine Learning Based Model For Solar Radiation Prediction. *Advanced Engineering Science*, 54(2), 2583–2591.
- [7] Tandon, A., Awasthi, A., Pattanayak, K. C., Tandon, A., Choudhury, T., & Kotecha, K. (2025). Machine learning-driven solar irradiance prediction: advancing renewable energy in Rajasthan. *Discover Applied Sciences*, 7(2), 107. doi:<https://doi.org/10.1007/s42452-025-06490-8>
- [8] Suyono, H., Santoso, H., Hasanah, R. N., Wibawa, U., & Musirin, I. (2018). Prediction of solar radiation intensity using extreme learning machine. *Indonesian Journal of Electrical Engineering and Computer Science*, 12(2), 691-698. doi:<http://doi.org/10.11591/ijeecs.v12.i2.pp691-698>
- [9] Rahman, S., Rahman, S., & Bahalul Haque, A. K. M. (2021). Prediction of Solar Radiation Using Artificial Neural Network. *Journal of Physics: Conference Series*, 1767(1), 012041. doi:<https://doi.org/10.1088/1742-6596/1767/1/012041>
- [10] Ayyoubzadeh, S. M., Ayyoubzadeh, S. M., Zahedi, H., Ahmadi, M., & R Niakan Kalhori, S. (2020). Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. *JMIR Public Health and Surveillance*, 6(2), e18828. doi:<https://doi.org/10.2196/18828>
- [11] Okafor, I. R., Stephen, E., Chen, G., Jones, G., Onyema, O., & Ezekiel, A. (2024). Prediction Of Daily Solar Radiation In Southeastern Nigeria Using Extreme Gradient Boosting Model. *World Journal of Engineering Research and Technology*, 10(10), 61-108.
- [12] Ramli, N. A., Hamid, M. F. A., Azhan, N. H., & Ishak, M. A. A.-S. (2019). Solar power generation prediction by using k-nearest neighbor method. *AIP Conference Proceedings*, 2129(1), 020116. doi:<https://doi.org/10.1063/1.5118124>
- [13] Rafi, M. A. I., & Sohan, M. R. (2024). *Enhanced Solar Radiation Prediction with Machine Learning: A Comprehensive Analysis of Meteorological Data Using Random Forest*. Paper presented at the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT).

Article Copyright: © 2026 Mfon D. Umoh, Udoh F. Evans, Sunday A. Akpan, Sunday E. Otene, Akinola B. Olanrewaju. This is an open access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and distribution provided the original author and source are credited.

