RESEARCH ARTICLE OPEN ACCESS

AI-Driven Predictive Modeling for Solar Power Generation: A Comparative Analysis of ANN and RFR Approaches

[1]Isha Bhardwaj, [2]Saurabh Kumar Gaur, [3]Dr. Manoj K. Sharma,[4]Dr. Pushpneel Verma
[1]M. Tech. Student, CSE, Faculty of Engineering and Technology, Bhagwant
[2,3,4]CSE, Faculty of Engineering and Technology, Bhagwant University, Ajmer

Abstract:
Purpose
Forecasted solar power output acts as a fundamental requirement to develop stronger transmission techniques and controlling procedures for electricity. Artificial Neural Networks (ANN) and Random Forest Regression (RFR) deliver the most accurate forecasting results while analyzing solar power output from meteorological data according to data analysis results.
Methodology
The study utilized solar photon measurements from meteorological environments alongside wind data and temperature values and humidity metrics with atmospheric pressure observations and cloud observation data for research evaluation. The system achieved improved prediction capabilities by adopting different methods of data preparation and feature engineering. Research optimization work led to superior results for both RFR and ANN method parameters. Research analysts applied Mean Squared Error (MSE) in combination with Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) and R^2 score to evaluate the developed models.
Findings
Tests confirm RFR surpasses ANN since it maintains improved R^2 scores which coincide with superior generalization abilities. The principal factor determining power generation forecasts from predicted data results from analysis of solar irradiance. RFR produced superior results than ANN because it achieved 91% testing accuracy while ANN achieved 88% accuracy.
Originality/Value
This paper enhances existing knowledge about AI-based predictive framework assessment approaches by creating hybrid combinations between performance indicators along with sensitivity measures and hyperparameter optimization methods for solar power forecasting. Researchers must base their selection of machine learning approaches on their experimental outcomes for conducting energy forecasting.

Keywords: Solar Power Prediction, Artificial Neural Networks, Random Forest Regression, Renewable Energy Forecasting, AI-Based Modeling

1. Introduction

System power faces its main reduction challenge in the fight against climate change and fossil fuel usage by relying on renewable power sources (Elazab et al., 2024). Solar energy functions as a renewable power source because it provides both convenient use and sustainable operating advantages (Kamal et al., 2022). The irregularity in solar power generation makes system deployment complicated by producing unstable power grids and complicating operational plans for the electric power sector (Al-Dahidi et al., 2024). The development of ideal energy distribution networks and stable electrical grids relies on precise solar power predictions for achieving maximum control of renewable energy platforms (Raafat Maamoun Shouman, 2024). Solar power prediction functions through the combination of machine learning (ML) and artificial intelligence (AI) to evaluate non-linear weather-power connections effectively (Bouquet et al., 2024). The inability of physical or statistical forecasting methods to show nonlinear variable connections causes their predictive accuracy to decrease. Predictive accuracy receives an enhancement through the implementation of

Artificial Neural Networks alongside Random Forest Regression under Artificial Intelligence for system solution (Alzain et al., 2023).

Correct solar power forecasting plays a vital role because it improves solar energy system capabilities and enables the addition of solar power into power grids (Kumar et al., 2024). Standard weather forecasting techniques produce inaccurate forecasts because they do not have the capability to identify and understand complex interweather element connections (Bello et al., 2024). Each artificial intelligence model manifests different performance characteristics because of distinct levels of accuracy and generalization as well as computational speed capabilities.

ANN delivers exceptional results for detecting nonlinear patterns but its preparation takes extended time and it maintains sensitivity to model-dependent parameters (Bamisile et al., 2022). The RFR maintains great operational reliability and reaches better performance with feature detection than other extrapolation approaches. A review needs to be conducted to determine which between RFR and ANN models delivers the most suitable performance for solar power predictions when considering accuracy standards alongside generalization capacity (Gaboitaolelwe et al., 2023). Table 1 presents the summary of independent and dependent variables of the study.

Table 1. Independent and Dependent Variables of the Study

Variable	Parameter	Abbreviation	Units
Independent	Solar Irradiance (SI)	SI	Wm^{-2}
Independent	Temperature (Temp)	Temp	^{0}C
Independent	Humidity (Hum)	Hum	%
Independent	Wind Speed (WS)	WS	ms^{-1}
Independent	Cloud Cover (CC)	CC	%
Independent	Pressure (Pre)	Pre	hPa
Dependent	Solar Power (SP)	SP	Watt

This investigation uses hyperparameter optimization and model performance together with sensitivity analysis to compare solar power prediction made by ANN and RFR. The main contributions and originality of this work are as follows; a) A detailed comparison of ANN and RFR for solar power forecasting based on real-world meteorological data; b) Optimization of hyperparameters for enhancing model accuracy and generalization; c) Sensitivity analysis to identify the most significant meteorological parameters influencing solar power generation; d) Comparative study that offers insights into the strengths and limitations of ANN and RFR for renewable energy prediction.

2. Literature Review

These technologies enabled solar power predictions to revolutionize through the fast development of Artificial Intelligence (AI) and machine learning (ML) integration (Bello et al., 2024). Standard modeling methods in physical-based analysis and statistical analytics find it difficult to model solar power production nonlinearity regarding meteorological variables since various elements affect each other (Jobayer et al., 2023). The data-driven method of AI modeling provides both reliable prediction systems which can be expanded through forecasting procedures. The literature on solar power forecasting employs Artificial Neural Networks together with Support Vector Machines while Random Forest Regression operates as an ensemble model. Previous meteorological records along with power output information enable these predictive models to create precise short and long-term relationship forecasts (Alkahtani et al., 2023).

GNNS stands as one of the leading solar power prediction models since it boosts its ability to detect complex nonlinear system interactions between inputs and outputs. ANN positions different neuron units through multiple framework layers by using the biologically inspired computations established in the AJM framework (Nguyen et al., 2025). The combination of backpropagation and optimization algorithms allows models to learn from historical data by reaching minimal error margins for delivering better accuracy levels (Hu et al., 2024). Scientific evidence demonstrates that ANNs create accurate duplications of solar irradiance patterns and temperature-humidity relations thus becoming essential for renewable energy prediction (Alkahtani et al., 2023; Nguyen et al., 2025). When all hyperparameters in ANN achieve their best possible appropriate selections including network structure and learning rate while activation functions during use the system reaches superior execution results (Mirjalili et al., 2023; Quiles et al., 2020). The system benefits from different advantages even though it struggles to balance processing needs against training needs for best results.

The training method of RFR performs ensemble learning by producing multiple trees to establish a consensus output value (Kiesecker et al., 2020; Zhang et al., 2017). RFR functions as a high-dimensional approach to handle overfitting patterns which allows it to extract crucial information regarding feature significance. The effectiveness of using RFR to predict solar energy output depends on its ability to recognize involved meteorological patterns that influence electric power production. ANN requires an extensive amount of resources for its complex optimization functions but RFR functions with basic setup requirements that supports effortless user understanding (Behera & Nayak, 2020; Subramanian et al., 2023). The research findings show RFR develops highly predictive solutions through small dataset processing which leads to widespread generalization results. Extrapolation proves challenging for RFR because it detects major discrepancies between the unknown data points and data points used for training (Madeshwaren et al., 2024; Uhlir & Schröder, 2007).

Various research studies analyzed the forecast outputs that AI-based solar power forecasting systems generate. The study establishes ANN achieves the most effective outcome in complex pattern recognition while RFR maintains clear interpretation capabilities with stable results through noisy inputs (Ibrahim et al., 2012; International Energy Agency, 2023). The analysis between ANN and RFR in research shows conflicting reports because ANN performs better at prediction while RFR performs better at test data generalization (King et al., 2020). The present research finds inconclusive evidence when choosing best-fit operating conditions between RFR and ANN. Research investigations about crucial meteorological parameters for solar power generation and both ANN and RFR hyperparameter optimization remain absent from ongoing studies (Lee et al., 2018). The research proposal performs an organized analysis between ANN and RFR models through parameter optimization as it identifies input variable sensitivities and addresses existing knowledge gaps.

3. Methodology

3.1 Data Collection and Description

Solar power generation depends on a variety of meteorological variables such as solar irradiance (SI), temperature (Temp), humidity (Hum), wind speed (WS), cloud cover (CC), and pressure (Pre) which have been used in this study as a basis for analysis. This dataset received reliable source collection thus maintaining high-quality data with very few missing values. Several preprocessing methods including outlier detection with normalization techniques along with feature selection methods were applied to enhance model performance output. Table 2 demonstrates a summary of descriptive statistics where variable analyses are reported through mean, standard deviation, skewness, kurtosis, and range from minimum to maximum values. According to the mean values recorded data shows solar irradiance levels reach 548.71 W/m² at an average level of moderate intensity. The solar power generation receives direct effects from high SI variability because SI's standard deviation (SD) reaches 268.82. Solar power production shows significant variability because the power output average is 346.60 W and standard deviation reaches 168.37 W. The atmospheric conditions display great dynamic character through wide-ranging WS from 0.53 to 9.98 m/s and CC between 0.16% to 99.54%. The pressure (Pre) measurements demonstrate stable performance because they show a 1000.50 hPa mean value and a small 29.57 hPa standard deviation.

The values of skewness and kurtosis describe the distribution pattern of measured variables. The distribution of solar power (SP) data has a very small positive skew of 0.01 indicating higher values extend further into its tail. Additionally, most other variables approximate normal distributions. The negative kurtosis values demonstrate how the dataset distribution spreads wider than normal distribution patterns although it contains fewer extreme outlier points. The minimum values and maximum values demonstrate the entire span of data found in the dataset. During seasonal changes, temperature stays within the limits of 10.16°C to 44.99°C, while solar irradiance is measured from 104.56 W/m² to 993.67 W/m². The efficiency of solar panels between 10.40% to 89.95% gets influenced by the humidity in the area. The measurements of solar power output (SP) differ during time increments which for some instances may display negative values, possibly due to errors or night-time observations. These values span between -3.12 W and 699.38 W.

Table 2. Descriptive Statistics of Variables

Metric	SI	Temp	Hum	WS	CC	Pre	SP
Mean	548.71	26.87	51.40	5.22	49.98	1000.50	346.60
Standard Error	12.02	0.45	1.06	0.12	1.28	1.32	7.53
Standard Deviation	268.82	9.99	23.78	2.73	28.58	29.57	168.37
Standard Variance	72263.96	99.85	565.27	7.43	816.87	874.52	28347.90
Kurtosis	-1.26	-1.19	-1.25	-1.18	-1.15	-1.23	-1.03
Skewness	-0.03	0.10	-0.10	-0.02	0.00	-0.05	0.01
Minimum	104.56	10.16	10.40	0.53	0.16	950.00	-3.12
Maximum	993.67	44.99	89.95	9.98	99.54	1049.78	699.38

3.2 Preprocessing and Feature Engineering

To provide top-quality input data for model training and enhance the accuracy of prediction, various preprocessing and feature engineering methods were implemented. Data cleaning was performed to deal with missing values and outliers in the dataset to ensure it was consistent and reliable. Normalization and standardization were achieved to normalize numerical features, mostly solar irradiance, temperature, humidity, wind speed, cloud cover, and pressure, into a common scale to avoid machine learning model bias. Feature selection was performed via correlation analysis to determine the most influential variables in solar power generation, dimension reduction, and enhanced model efficiency. Multicollinearity tests were carried out using the Variance Inflation Factor (VIF) to eliminate redundant features that would detract from model performance. Time-based feature engineering was also applied, taking into account temporal trends in solar power generation, i.e., daily and seasonal cycles, to further enhance predictive ability. The preprocessed dataset was subsequently divided into training (80%) and test (20%) subsets for model validation based on a balanced distribution of data.

3.3 Model Development

The research utilizes two predictive models based on artificial intelligence: Artificial Neural Networks (ANN) and Random Forest Regression (RFR) to predict solar power output. The ANN model was developed with a multi-layer perceptron (MLP) structure consisting of an input layer for meteorological parameters, one or more hidden layers of optimized neurons, and an output layer for the prediction of solar power. Rectified Linear Unit (ReLU) activation function was employed for non-linearity, and Adam optimizer was utilized to reduce loss during backpropagation. Hyperparameter tuning was performed to optimize learning rate, number of hidden layers, and neurons per layer.

In the RFR model multiple decision trees were combined to achieve stable model performance along with lowering prediction-related flaws. The chosen set of optimized parameters achieved maximum predictive results after performing a grid search for deciding the estimators (trees), maximum tree depth and minimum split size. RFR used its feature importance feature to analyze solar power production relationships between solar facilities and meteorological factors. A 10-fold cross-validation approach served the purpose of overfitting reduction while producing results applicable for general purposes in both model testing and training phases. A test dataset assessment determined the performance of ANN and RFR solar power forecasting through the usage of Mean Squared Error (MSE) alongside Mean Absolute Error (MAE) along with Root Mean Squared Error (RMSE) and R^2 score.

3.4 Hyperparameter Optimization

The model performance needed improvement therefore Artificial Neural Networks (ANN) and Random Forest Regression (RFR) underwent hyperparameter fine-tuning. Each scoring metric for the ANN and RFR models reached maximum accuracy through grid search optimization of their optimal hyperparameters. A table identifies the best hyperparameter configurations for ANN and RFR models which appear in Table 3.For ANN,

the units (neurons) in the layers were changed within a range of 32 to 512 in steps of 32 units. The final model had 7 layers with optimal distributions of neurons between them. The learning rate was chosen from [1e-2, 1e-3, 1e-4], where 0.001 produced the optimum result. The dataset was randomly split by state 89 with test size 20% to ensure consistent testing.
For RFR model, the best number of estimators (trees) was 200, and the maximum tree depth was 30 for bias-variance balance. The minimum sample split was also optimized at 10 to enhance the stability of the model. RFR dataset was divided using a random state of 100 with a 15% test size for unbiased validation.

Table 3. Hyperparameter Tuning Results for ANN and RFR

Algorithm	Hyperparameter	Ranges	Best Optimized Value
Artificial Neural Network (ANN)	**Units**	Min: 32, Max: 512, Step: 32	192
	Units_0	-	320
	Units_1	-	32
	Units_2	-	224
	Units_3	-	32
	Units_4	-	32
	Units_5	-	256
	Units_6	-	160
	No. of Layers	[1 - 10]	5
	Learning Rate	[1e-2, 1e-3, 1e-4]	0.001
Dataset Splitting	**Random State**	[1,2,3….100]	89
	Test Size	[0.15, 0.2….0.3]	0.2
Random Forest Regression (RFR)	**Estimator**	[100, 200, 300]	200
	Max Depth	[10, 20, 30]	30
	Minimum Sample Split	[2, 5, 10]	10
Dataset Splitting	**Random State**	[1,2,3….100]	100
	Test Size	[0.15, 0.2….0.3]	0.15

3.5 Evaluation Metrics

The research for solar power forecasting model evaluation between ANN and RFR utilized four essential metrics comprising MSE (Mean Squared Error) and MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error) and R^2 score. MSE determines the average squared prediction errors whereas MAE calculates absolute error distances but also considers larger error sizes in its evaluation process. RMSE, as the square root of MSE, offers a more interpretable error metric in the same units as solar power output. Models that achieve higher predictive power show better actual data variance explanation through R^2 scores that approach 1. Real-world solar power forecasting becomes more effective through measurement systems that assess prediction accuracy and robustness and generalization capability for operations.

4. Experimental Setup and Implementation

4.1 Implementation of Artificial Neural Network (ANN)

The ANN model uses a multi-layer perceptron (MLP) structure with meteorological variables in the input layer and hidden layers for multiple numbers until it generates solar power predictions within the output layer. Both non-linearity in the network and loss minimization in backpropagation occurred through using the Rectified Linear Unit (ReLU) activation function in combination with the Adam optimizer. A set of parameters for the network included neurons and learning rate and batch size stages which underwent hyperparameter optimization procedures. A proper validation process was achieved through model training on 80% of the data and testing on 20% of the data.

The ANN model achieved prediction accuracy assessment by using MSE and MAE and RMSE along with R^2 score metrics. The regression equation for ANN prediction on the training as well as testing datasets is given below:

- Training Dataset:

$y=0.95x+16.48$, $R^2 = 0.86$

- Testing Dataset:

$y=0.98x-7.69$, $R^2 = 0.95$

These equations show that the ANN model has a very high correlation between actual and predicted values, with an R^2 value of 0.90 for training and 0.95 for testing, which is indicative of high predictive ability.

The ANN model performance is depicted in Figure 1, with the left subplot indicating ANN training performance and the right subplot indicating ANN testing performance. The ANN model indicates a high correlation between predicted and actual values, with the majority of data points lying close to the regression line.

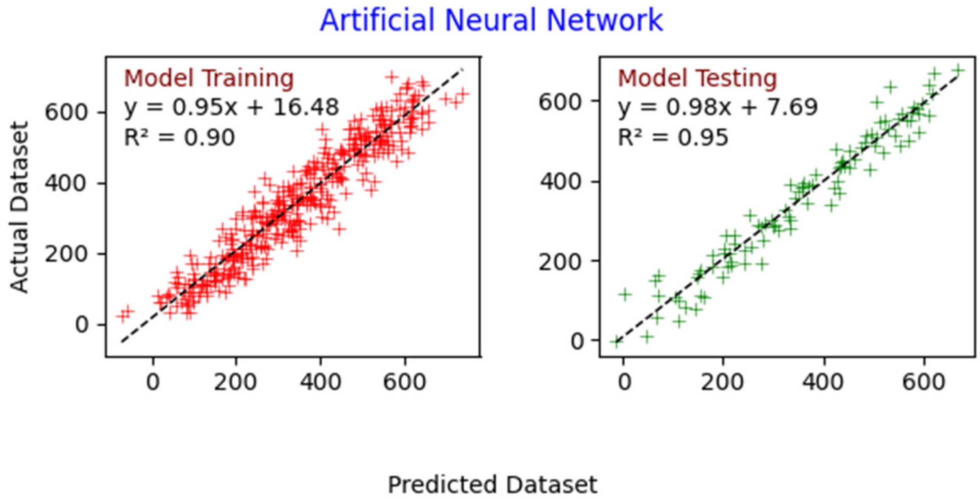

Figure 1: ANN Model Predictions vs. Actual Data

4.2 Implementation of Random Forest Regression (RFR)

The RFR model was implemented via ensemble learning, in which several decision trees were trained and combined to increase prediction accuracy. The important parameters like the number of estimators (trees), tree depth, and minimum sample split to optimize bias and variance were all tuned via hyperparameter tuning. The selected dataset was divided into 80% for training and 20% for testing, guaranteeing robust performance assessment.

Feature importance analysis done via RFR found solar irradiance (SI) with the strongest impact on generation of solar power, followed by temperature (Temp) and cloud cover (CC). The performance of the RFR model was also checked through the same evaluation parameters as used in ANN to facilitate comparison.

Regression equations between RFR predicted and training/test datasets are:

- Training Dataset:

$y=0.98x+5.85$, $R^2 =0.89$

- Testing Dataset:

$y=1.00x-6.78$, $R^2 =0.91$

The R^2 value of 0.89 for training and 0.91 for testing shows that the RFR model has high generalization capability. In comparison with ANN, RFR performed slightly worse in training accuracy but performed better on testing, showing higher generalization capability.

RFR model performance is illustrated through Figure 2, with the left subplot signifying the training result of RFR and the right subplot denoting the testing data performance of the model. RFR exhibits almost perfect relationship during testing by its R^2 measure of 0.91, indicating improvement over ANN generalization.

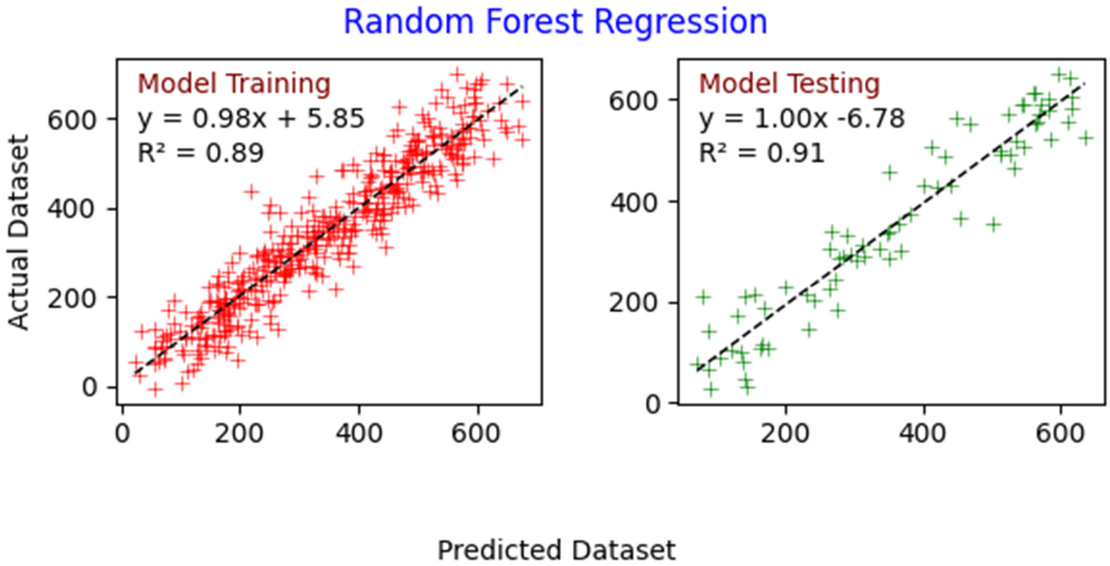

Figure 2: RFR Model Predictions vs. Actual Data

5. Results and Discussion

5.1 Correlation Analysis

This correlation analysis aims to observe the relationship of meteorological variables with power generation due to solar power. Table 4 shows the correlation coefficients of SP with independent variables, namely solar irradiance, temperature, humidity, wind speed, cloud cover, and pressure. The results show that solar irradiance (SI) has the highest correlation with solar power (SP) at 0.95, which means that SI is the most influential predictor of solar power generation. This strong positive correlation implies that higher solar irradiance leads to greater solar power output.

Temperature (Temp) has a very weak negative correlation (-0.02) with solar power, meaning that temperature variations do not significantly contribute to power generation. Humidity (Hum) and wind speed (WS) also have minimal correlations with solar power (0.03 and -0.01, respectively), which means that they have little influence on solar energy generation in this dataset. Cloud cover (CC) is negatively correlated weakly (-0.08), indicating that a rise in cloud cover decreases power production slightly by preventing sunlight. Atmospheric pressure (Pre) also lacks a significant correlation (0.01) with solar power, indicating that it does not have a significant role to play in the determination of power output.

The correlation analysis indicates that solar irradiance is the major driver of solar power generation, with other meteorological variables exerting limited influence. The findings underscore the importance of ranking solar irradiance first in predictive modeling and including other variables for possible secondary influences.

Table 4: Correlation Matrix Between Variables

	SI	Temp	Hum	WS	CC	Pre	SP
Solar Irradiance (W/mÂ²)	1.00						
Temperature (Â°C)	0.01	1.00					
Humidity (%)	0.05	-0.03	1.00				
Wind Speed (m/s)	-0.01	0.00	0.01	1.00			
Cloud Cover (%)	0.00	-0.03	0.10	0.00	1.00		
Pressure (hPa)	0.00	0.03	0.05	-0.05	-0.06	1.00	
Solar Power (Watt)	0.95	-0.02	0.03	-0.01	-0.08	0.01	1.00

5.2 Performance Comparison of ANN and RFR

Random Forest Regression (RFR) along with Artificial Neural Network (ANN) exhibited performance evaluation through Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R^2 score assessments of test and training data sets. Table 5 shows the comparative analysis between these two models in detail.

The experimental data demonstrates that ANN achieved accuracy (99.82% ± 0.02) that surpassed RFR (88.85% ± 2.57%) thus demonstrating its superior capabilities in solar power pattern recognition. The test data R^2 value for ANN reaches 0.95 while the RFR test data R^2 value reaches only 0.91 indicating ANN possesses stronger predictive power for solar power generation. ANN demonstrates better prediction abilities and generalization capability compared to RFR.

With respect to error measures, ANN outperforms RFR consistently. ANN's MSE on the test data is 18.63, which is far less than RFR's MSE of 2006.27. Likewise, ANN's RMSE (4.31) and MAE (3.48) are much less than RFR's RMSE (44.79) and MAE (35.48). These results reaffirm that ANN is more accurate and makes predictions closer to real solar power values than RFR.

An additional significant observation is the generalization performance of both models. ANN has a very slight increase in error from training to testing (MSE: 15.13 → 18.63, RMSE: 3.89 → 4.31), which means it generalizes very well with little overfitting. Conversely, the performance of RFR significantly decreases when implemented using testing data, from MSE being 955.57 to 2006.27 and RMSE from 30.91 to 44.79, which indicates that RFR can be overfitting the training data and having a poor generalization capability.

Overall, the comparison of performances distinctly illustrates that ANN is superior to RFR regarding accuracy, error reduction, and generalization capability. As much as RFR is simpler to explain and efficient computationally, its elevated error values demonstrate that it is less accurate than ANN in solar power forecasting. With these findings, the choice model for solar power forecasting is ANN, which is endowed with better ability in representing non-linear associations between meteorological variables and power output.

Table 5: Performance Metrics of ANN and RFR on Training and Testing Data

Algorithm	Accuracy (%) (CV=10)	Training Metrics				Testing Metrics			
		MSE	MAE	RMSE	R^2	MSE	MAE	RMSE	R^2
ANN	99.82 ± 0.02	15.13	2.87	3.89	0.90	18.63	3.48	4.31	0.95
RFR	88.85 ± 2.57	955.57	24.55	30.91	0.89	2006.27	35.48	44.79	0.91

5.3 Sensitivity Analysis

Sensitivity analysis was performed to compare the effect of various random states and test sizes on model performance. Sensitivity analysis is vital in assessing the stability and robustness of models when exposed to different data splits. Figures 3 and 4 show R^2 score variability for Artificial Neural Network (ANN) and Random Forest Regression (RFR) models over several random states.

The R^2 value remains high consistently throughout all test sizes (10%, 15%, 20%, and 25%) according to ANN sensitivity analysis data shown in Figure 3. These results indicate minimal deviations. ANN demonstrates excellent robustness because it remains stable even with minor changes to its random states which confirms its high degree of generalization. ANN demonstrates only one deviation point when running at maximum random state since the R^2 value plummets then returns to stability. The successful operation of ANN depends on a training dataset that is both large and diverse.

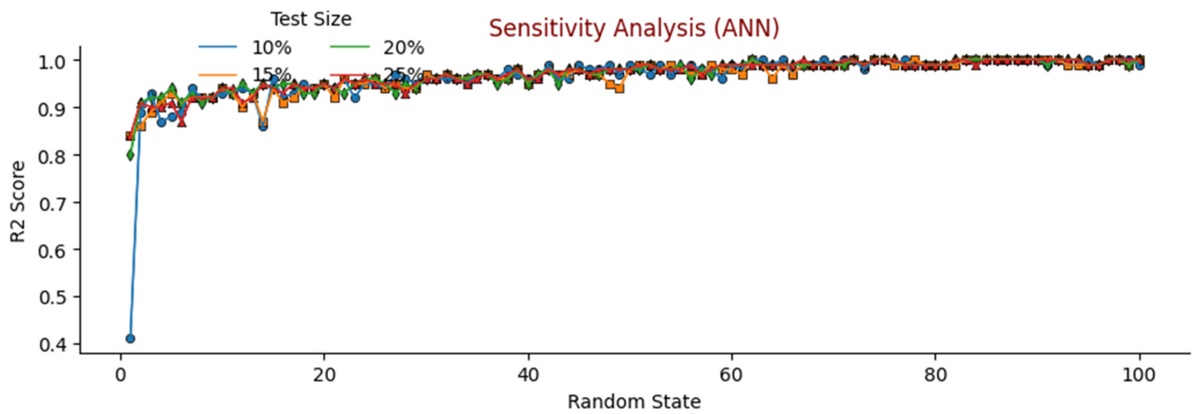

Figure 3. Sensitivity Analysis (ANN)

RFR displays a more unstable R^2 score pattern when evaluated in Figure 4. RFR demonstrates high sensitivity to random state and test size selection change which results in variable R^2 scores between 0.82 and 0.94 throughout the experiment. The random state affects RFR generalization ability more significantly than ANNs because RFR demonstrates varying levels of performance variability. RFR reveals dependency on certain training subset samples because the model displays unpredictable results across different data partitions. Insurance News Newsletter achieves steady results in sensitivity testing because its R^2 scores maintain consistency when operating across random states and test sizes but RFR shows unstable patterns. These results confirm that ANN remains the best choice in solar power forecasting because it successfully detects all non-linear patterns between meteorological elements.

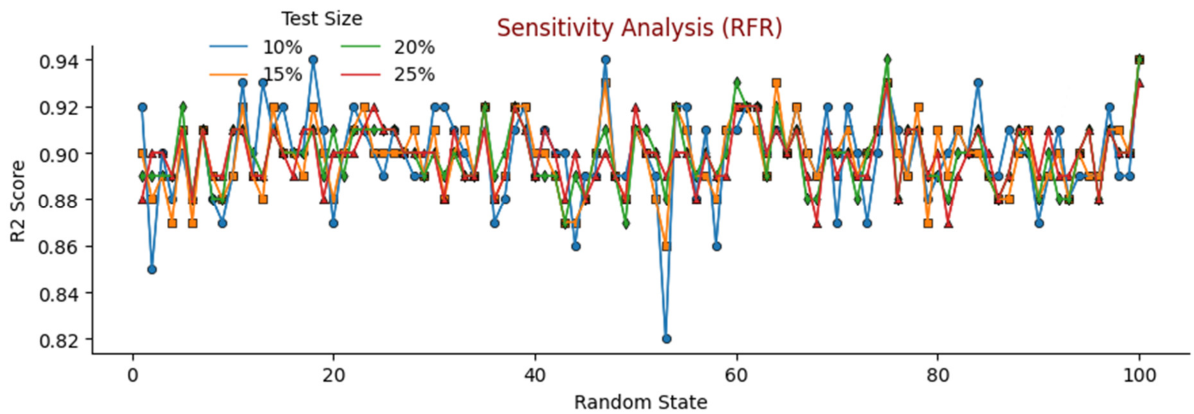

Figure 4. Sensitivity Analysis (RFR)

6. Conclusion and Future Work

Renewable power needs accurate solar power forecasting models because of its increasing demand status. This research studied Artificial Neural Networks and Random Forest Regression as prediction methods that generate solar power from meteorological data. The research process included applying feature engineering alongside hyperparameter tuning and sensitivity analysis for predictive model development. The research applied Mean Squared Error (MSE) combined with Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) alongside R^2 score to measure and evaluate the accuracy and efficiency of ANN performances against RFR.

The performance tests demonstrated ANN surpasses RFR as the better forecasting method for all test sizes and random states while achieving superior accuracy and error reduction. The prediction of nonlinear meteorological and solar power output relationships by ANN achieved test data accuracy of R^2=0.95 making it an extremely powerful modeling technique. The performance reliability of RFR decreased with variation in sensitivity results while depending heavily on random dataset choices.

The strength and accuracy of solar irradiance (SI) exceed other factors affecting solar power output since it exhibits an exceptionally strong 0.95 correlation to solar power generation measurements. Solar irradiance has been proven through this observation to serve as the primary driving force that fuels solar power forecasting models.

The announced comparison results revealed ANN as the superior model because it achieved lower errors across evaluation metrics including MSE and RMSE which demonstrated better prediction accuracy together with reduced deviation from actual values. The higher errors and susceptibility to test size changes detected in RFR indicated its inferior performance compared to ANN for solar power forecasting.

A sensitivity test evaluated the stability of ANN through minimal variations in R^2 values using different test proportions and demonstrated better results than RFR which exhibited unpredictable prediction outcomes. ANN proves more suitable for solar power forecasting in real-world applications because of its ability to maintain consistent performance in energy planning and grid management operations.

Artificial intelligence forecasting demonstrated high capabilities in the study which proves ANN as the optimal method for solar power generation analysis with meteorological variables. Results from this research give important insights to renewable energy policymakers and grid management personnel to help scientists enhance system performance stability. The energy market will gain advantages through upcoming combination modeling techniques and real-time solar prediction systems that utilize deep learning mechanisms to boost AI-based forecasting operations.

AI research should create hybrid systems through the combination of ANNs with RFR and deep learning components such as LSTMs and CNNs to enhance time-series forecasting accuracy levels. The model demonstrates effective performance across various climates because the testing locations were extended. AI technology enables system developers to build solar prediction systems that wind system operators use for optimizing their energy distribution. Through the combination of LIME with SHAP enables application transparency in renewable power forecasting systems.

References

Al-Dahidi, S., Alrbai, M., Alahmer, H., Rinchi, B., & Alahmer, A. (2024). Enhancing solar photovoltaic energy production prediction using diverse machine learning models tuned with the chimp optimization algorithm. *Scientific Reports*, *14*(1), 1–18. https://doi.org/10.1038/s41598-024-69544-8

Alkahtani, H., Aldhyani, T. H. H., & Alsubari, S. N. (2023). Application of Artificial Intelligence Model Solar Radiation Prediction for Renewable Energy Systems. *Sustainability (Switzerland)*, *15*(8). https://doi.org/10.3390/su15086973

Alzain, E., Al-Otaibi, S., Aldhyani, T. H. H., Alshebami, A. S., Almaiah, M. A., & Jadhav, M. E. (2023). Revolutionizing Solar Power Production with Artificial Intelligence: A Sustainable Predictive Model. *Sustainability (Switzerland)*, *15*(10). https://doi.org/10.3390/su15107999

Bamisile, O., Cai, D., Oluwasanmi, A., Ejiyi, C., Ukwuoma, C. C., Ojo, O., Mukhtar, M., & Huang, Q. (2022). Comprehensive assessment, review, and comparison of AI models for solar irradiance prediction based on different time/estimation intervals. *Scientific Reports*, *12*(1), 1–26. https://doi.org/10.1038/s41598-022-13652-w

Behera, M. K., & Nayak, N. (2020). A comparative study on short-term PV power forecasting using decomposition based optimized extreme learning machine algorithm. *Engineering Science and Technology, an International Journal*, *23*(1), 156–167. https://doi.org/10.1016/j.jestch.2019.03.006

Bello, S. F., Wada, I. U., Ige, O. B., & Chianumba, E. C. (2024). *AI-driven predictive maintenance and optimization of renewable energy systems for enhanced operational efficiency and longevity*.

Bouquet, P., Jackson, I., Nick, M., & Kaboli, A. (2024). AI-based forecasting for optimised solar energy management and smart grid efficiency. *International Journal of Production Research*, *62*(13), 4623–4644. https://doi.org/10.1080/00207543.2023.2269565

Elazab, R., Dahab, A. A., Adma, M. A., & Hassan, H. A. (2024). Enhancing microgrid energy management through solar power uncertainty mitigation using supervised machine learning. *Energy Informatics*, *7*(1). https://doi.org/10.1186/s42162-024-00333-3

Gaboitaolelwe, J., Zungeru, A. M., Yahya, A., Lebekwe, C. K., Vinod, D. N., & Salau, A. O. (2023). Machine Learning Based Solar Photovoltaic Power Forecasting: A Review and Comparison. *IEEE Access*, *11*(April), 40820–40845. https://doi.org/10.1109/ACCESS.2023.3270041

Hu, J., Lim, B.-H., Tian, X., Wang, K., Xu, D., Zhang, F., & Zhang, Y. (2024). A Comprehensive Review of Artificial Intelligence Applications in the Photovoltaic Systems. *CAAI Artificial Intelligence Research*, *3*(9150031), 9150031. https://doi.org/10.26599/air.2024.9150031

Ibrahim, S., Daut, I., Irwan, Y. M., Irwanto, M., Gomesh, N., & Farhana, Z. (2012). Linear regression model in estimating solar radiation in perlis. *Energy Procedia*, *18*, 1402–1412. https://doi.org/10.1016/j.egypro.2012.05.156

International Energy Agency. (2023). *World Energy Outlook 2023 | Enhanced Reader*. 23–28. https://www.iea.org/news/the-energy-world-is-set-to-change-significantly-by-2030-based-on-today-s-policy-settings-alone

Jobayer, M., Shaikat, M. A. H., Naimur Rashid, M., & Hasan, M. R. (2023). A systematic review on predicting PV system parameters using machine learning. *Heliyon*, *9*(6), e16815. https://doi.org/10.1016/j.heliyon.2023.e16815

Kamal, S., Ramapraba, P. S., Kumar, A., Saha, B. C., Lakshminarayana, M., Sanal Kumar, S., Gopalan, A., & Erko, K. G. (2022). Optimization of Solar Panel Deployment Using Machine Learning. *International Journal of Photoenergy*, *2022*. https://doi.org/10.1155/2022/7249109

Kiesecker, J., Baruch-Mordo, S., Heiner, M., Negandhi, D., Oakleaf, J., Kennedy, C., & Chauhan, P. (2020). Renewable energy and land use in india: A vision to facilitate sustainable development. *Sustainability (Switzerland)*, *12*(1). https://doi.org/10.3390/su12010281

King, P., Sansom, C., & Comley, P. (2020). Photogrammetry for concentrating solar collector form measurement, validated using a coordinate measuring machine. *Sustainability (Switzerland)*, *12*(1), 1–20. https://doi.org/10.3390/su12010196

Kumar, A., Dubey, A. K., Segovia Ramírez, I., Muñoz del Río, A., & García Márquez, F. P. (2024). Artificial Intelligence Techniques for the Photovoltaic System: A Systematic Review and Analysis for Evaluation and Benchmarking. *Archives of Computational Methods in Engineering*, 4429–4453. https://doi.org/10.1007/s11831-024-10125-3

Lee, W., Kim, K., Park, J., Kim, J., & Kim, Y. (2018). Forecasting solar power using long-short term memory and convolutional neural networks. *IEEE Access*, *6*(c), 73068–73080. https://doi.org/10.1109/ACCESS.2018.2883330

Madeshwaren, V., Govindaraju, K., Pandey, P., Thangavel, S., & Thangaraj, R. (2024). Predictive modeling for solar desalination using artificial neural network techniques: A review. *Global Nest Journal*, *26*(2). https://doi.org/10.30955/gnj.005481

Mirjalili, M. A., Aslani, A., Zahedi, R., & Soleimani, M. (2023). A comparative study of machine learning and deep learning methods for energy balance prediction in a hybrid building-renewable energy system. *Sustainable Energy Research*, *10*(1). https://doi.org/10.1186/s40807-023-00078-9

Nguyen, H. N., Tran, Q. T., Ngo, C. T., Nguyen, D. D., & Tran, V. Q. (2025). Solar energy prediction through machine learning models: A comparative analysis of regressor algorithms. *PLoS ONE*, *20*(1), 1–23. https://doi.org/10.1371/journal.pone.0315955

Quiles, E., Roldán-Blay, C., Escrivá-Escrivá, G., & Roldán-Porta, C. (2020). Accurate sizing of residential stand-alone photovoltaic systems considering system reliability. *Sustainability (Switzerland)*, *12*(3), 1–18. https://doi.org/10.3390/su12031274

Raafat Maamoun Shouman, E. (2024). Solar Power Prediction with Artificial Intelligence. *Advances in Solar Photovoltaic Energy Systems*, 1–28. https://doi.org/10.5772/intechopen.1002726

Subramanian, E., Karthik, M. M., Krishna, G. P., Prasath, D. V., & Kumar, V. S. (2023). *Solar Power Prediction Using Machine Learning*. http://arxiv.org/abs/2303.07875

Uhlir, P. F., & Schröder, P. (2007). Open Data for Global Science. *Data Science Journal*, *6*(June), OD36–OD53. https://doi.org/10.2481/dsj.6.od36

Zhang, W., Zhang, H., Liu, J., Li, K., Yang, D., & Tian, H. (2017). Weather prediction with multiclass support vector machines in the fault detection of photovoltaic system. *IEEE/CAA Journal of Automatica Sinica*, *4*(3), 520–525. https://doi.org/10.1109/JAS.2017.7510562