International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

AI-Driven Solar Energy Potential Assessment in Ayodhya District Incorporating Weather Variability and Geographical Factors

Prem Kumar
Research Scholar, Department of Physics, K.S. Saket P.G. College Ayodhya Affiliated to Dr Ram Manohar Lohiya Avadh University Ayodhya, Uttar Pradesh, India.

Rajesh Kumar Verma
Assistant Professor Department of Physics, K.S. Saket P.G. College Ayodhya Affiliated to Dr Ram Manohar Lohiya Avadh University Ayodhya Uttar Pradesh India.

ABSTRACT: The fast-moving global change to renewable energy has made solar photovoltaic systems a key part of green power production, especially in developing nations like India. Correct forecasting of solar irradiance is vital for smooth grid connection, energy planning, and investment decisions. Though important, eastern Uttar Pradesh, particularly the Ayodhya district (26.7°N, 82.2°E), has been less explored in AI-based solar studies despite being named a Solar City under the Uttar Pradesh Solar Energy Policy 2022 and having a 40 MW NTPC solar project. This paper provides a detailed evaluation of solar energy potential in Ayodhya using multi-source satellite and reanalysis datasets which includes NSRDB PSM v3, NASA POWER, ISRO VEDAS (INSAT-3D), IMD gridded data, and the Global Solar Atlas. Ten machine learning and deep learning models—Random Forest, XGBoost, Support Vector Regression, Gaussian Process Regression, Artificial Neural Network, Long Short-Term Memory, Gated Recurrent Unit Bidirectional LSTM Convolutional Neural Network and hybrid CNN-LSTM were tested for predicting solar irradiance. GIS-based spatial analysis was also used to find good places for setting up solar PV systems. Results show that the hybrid CNN-LSTM model had the best accuracy with R^2=0.9812 RMSE=31.42 W/m² MAE=21.35 W/m² MAPE=4.82% and NSE=0.9798. The region has an annual average GHI of 4.92 kWh/m²/day (~1795 kWh/m²/year) with a decline during monsoon months by 30–45%. This finding supports high solar potential and provides scalable AI-based frameworks for similar regions.

KEYWORDS: Solar energy potential assessment; CNN-LSTM hybrid model; GIS-based suitability analysis; Ayodhya Solar City; Machine learning; Global Horizontal Irradiance

1. INTRODUCTION

Post the Paris Agreement, the global energy transition has taken rapid strides, and solar PV has been the most significant growth driver for new electricity generation capacity worldwide. As the world's 3rd largest energy consumer and one of the signatories of the Paris Agreement, India has pledged to install 500GW non-fossil fuel based electricity capacity by 2030 as stated at COP26 at Glasgow, UK by Ministry of New and Renewable Energy (MNRE) in November 2021. Solar energy has a share of ~280GW in this pledge which calls for an unprecedented acceleration of site identification, characterization and grid planning across India's complex climatic regions. While India's installed solar capacity has crossed 85 GW (2024), the pace of deployment in Indo-Gangetic Plains (IGP) where population and agriculture is concentrated, and where solar irradiance exhibits highest temporal variations due to monsoon, is dismally low compared to the resource endowment in this region.

For Uttar Pradesh (UP), the most populous state in India which has targets of 22,000MW by 2026-27 under its UP Solar Energy Policy 2022 (currently installed is 5,346 MW or 8th in India), there is an urgent need for not only policy breakthroughs, but also for sound scientific understanding of solar resources for an adequate deployment rate. Among several developing regions of India where focus is on Solar Energy, Ayodhya District (26.7°N, 82.2°E) in eastern UP has gained national significance with its recognition as the first Solar City under UPNEDA Solar City Policy 2022 in September 2022, with targets of mandatory rooftop solar on commercial and institutional building, solar street lighting, and solar powering of public infrastructure on the approximately 2,522 km² area.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

The most concrete evidence of such a designation is the construction of 40 MW ground-mounted solar PV power plant developed by NTPC Green Energy Ltd at Majha Rampur Halwara and Sarairasi villages of Ayodhya. The estimated investment in the said project is around Rs 200 crore and the land of 165 acres (around 80 ha) was taken on a 30-year lease from UPNEDA at Rs 1 per acre per year. The plant consists of 104,580 solar panels and the produced power was feed to Fimer PVS980-58 5 MVA central inverters and integrated with grid through 132/33 kV line to Darshannagar sub-station. 14 MW out of 40 MW got commission on January 27, 2024 and planned to commission the full 40 MW plant by March 2024. The full capacity 40 MW plant is expected to generate approximately 70,800 MWh per year and to meet 10-30% of local electricity demand of Ayodhya. In conjunction with such investment, long term energy system planning was being carried out for the energy transition towards 2050 by a GiZ/Fraunhofer ISE/Deloitte international partnership and optimizing the energy system at municipality level using Fraunhofer ISE KomMod (Kommunales Modell) tool.

The main gap from the science perspective is the lack of an AI based GIS integrated solar potential assessment specific to the Ayodhya district. While satellite derived climatological averages such as Global Solar Atlas, NASA POWER has been used for solar resource assessment of Indo Gangetic Plain. Topography, land use and specific meteorological pattern of the district has not been taken into account. There has not been any study assessing comparative performance of state-of-art deep learning models like hybrid CNN-LSTM for monsoon influenced semi-arid climatic condition. Hence there exists a significant scientific gap which hinders informed decisions on solar city plan, grid integration study and investment plan in Ayodhya and other eastern Uttar Pradesh districts.

Four research objectives address the gap outlined above, as follows: (1) characterize the solar resource available in the district of Ayodhya based on 20-year data from multiple sources with complete seasonal breakdown; (2) compare ten different AI/ML models including machine learning, deep learning and combined structures for daily GHI prediction under monsoonic weather; (3) assess the influence of monsoon-season weather fluctuations on the developed models and the availability of solar energy; and (4) develop GIS-based maps showing the spatial solar potential and the land suitableness for the district of Ayodhya for solar city planning.

2. LITERATURE REVIEW

2.1 Forecasting Solar Irradiance by Machine Learning and Deep Learning Process

The application of machine learning (ML) and deep learning (DL) to solar irradiance forecasting has expanded substantially over the past decade, driven by the increasing availability of high-resolution satellite-derived datasets and the growing computational accessibility of neural network architectures. Voyant et al. provided a foundational review of ML methods for solar radiation forecasting, evaluating artificial neural networks (ANN), support vector machines (SVM), random forests (RF), and gradient boosting approaches, concluding that hybrid and ensemble models consistently offer the greatest accuracy improvements over single-model baselines. This finding has been corroborated by subsequent comparative studies across diverse climatic contexts.

Random Forest seems to perform well in solar radiation prediction amongst classical ML methods. Studies in Rajasthan, India, using MERRA-2 reanalysis data found RF to be the most effective model with an RMSE of 0.64 W/m², outperforming linear regression by 19.47%, with temperature exhibiting the strongest correlation ($r = 0.63$) with solar irradiance. The study of Feng et al., confirmed that when RF models are trained on multi-variable meteorological inputs, they provide an appropriate assessment of and map global daily solar radiation. Amongst the competitive gradient boosting model types, XGBoost is of interest. An ensemble called XGBF-DNN when tested on three climatic zones of India predicted the forecast skill more accurately than reference models by 33-40% while a study comparing the forecast skill for six African countries, found that XGBoost performs better in 10 out of 13 case studies, having compared 8 different AI structures.Support Vector Regression (SVR) with radial basis function (RBF) kernels has been widely applied for solar irradiance estimation. Mohammadi et al. demonstrated that SVR-RBF achieved a daily R^2 of 0.9133 and a monthly R^2 of 0.9949 in Iranian climate conditions, while hybrid SVR-IPSO (Improved Particle Swarm Optimisation) models have shown superior performance over M5 Tree and MARS models across multiple geographic contexts.

An added benefit of GPR is its ability to perform uncertainty quantification: Salcedo-Sanz et al.showed that GPR could accurately predict daily global solar irradiation on a global level and the subsequent stacking ensemble studies has applied GPR as the meta-learner which has outperformed base learners on predicting daily solar radiation.

With the development of deep learning architectures have caused a significant rise in accuracy, especially on modeling sequential time series. Long Short-Term Memory (LSTM) networks, created by Hochreiter and Schmidhuber, are often used for solar

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

forecasting. Abdel-Nasser and Mahmoud, have predicted PV power using a deep LSTM-based recurrent neural network accurately and Qing and Niu, obtained high performance predicting hourly day-ahead solar irradiance with weather forecast information. A hybrid LSTM-SVM model for daily solar radiation prediction in Turkey achieved $R^2 = 0.962$ and RMSE = 0.0142, outperforming standalone SVR, LSTM, decision tree, and k-nearest neighbour models. Gated Recurrent Unit (GRU) networks, which employ reset and update gates to manage temporal dependencies with reduced computational overhead compared to LSTM, have been evaluated as competitive alternatives in multiple comparative studies.

Bidirectional LSTMs (Bi-LSTMs) are well-suited for daily forecasting, taking advantage of past and future states of a temporal series to model relationships and extract information over multiple days, which is very beneficial for multi-site time series data from multiple decades for daily prediction. Convolutional neural networks (CNNs), originally an image-based method, were used for solar irradiance prediction with 1D convolution over the meteorological time series, allowing it to effectively extract local temporal features. The CNN-LSTM architecture combines CNN and LSTM, using CNN to extract spatial features effectively and temporal dependency of LSTM- consistently yielded state-of-the-art results on a number of benchmark studies. CNN-LSTM combined with CEEMDAN decomposition had superior hourly solar irradiance forecasting. (Gao et al., 2019) CNN-LSTM proved superiority in short-term PV power production prediction, (Agga et al., 2019) and its hybrid model was proven to work for solar power prediction. (Lim et al., 2017) CNN-LSTM resulted in a best DNI RMSE of 238.22 W/m², a 2.89% improvement from benchmark models in a study at Petrolina, Brazil, utilizing GOES-16 satellite imagery.

2.2 AI/ML Comparative Studies in Renewable Energy

Several in-depth comparative studies have evaluated various AI/ML models in different climatic regions for solar irradiance prediction. Wang et al.provided a literature survey on deep learning based methods for renewable energy forecasting, comparing and reviewing LSTM, GRU, CNN, and CNN-LSTM networks. Kumari and Toshniwal also provided an exhaustive review on DL based methods for solar irradiance prediction with reference to Indian scenarios, where they showed that, while standalone models result in the lowest RMSE values with an LSTM network, combination of models perform better and the CNN-LSTM hybrid is the most effective. In a more general study, Mellit et al. reviewed numerous state-of-the-art techniques in PV power output prediction, where it was established that hybrid architectures outperform individual models. A meta-review, published in energies, on ensemble and hybrid forecasting models, concluded by proposing an entirely new and synthetic typology of forecasting approaches and highlighted the importance of spatiotemporal hybrid models as the next advancement in the area. Akhter et al. presented a review of ML and metaheuristic algorithms in PV output power prediction and stated that ensemble methods are generally much more stable and less computationally intensive than deep learning methods, which makes them more suitable for real time PV forecasting in conditions where data is sparse.

2.3 Research Gap and Contribution of the Present Study

The foregoing review reveals several critical gaps that the present study addresses. First, while ML/DL solar forecasting studies are abundant globally, no study has applied a comprehensive multi-model comparison specifically to Ayodhya district or any designated Solar City in the Indo-Gangetic Plain. Second, existing GIS-based solar assessments for Uttar Pradesh are either national-scale (insufficient spatial resolution for district-level planning) or focused on western UP (climatically distinct from the monsoon-dominated eastern region). Third, the quantitative impact of monsoon-season variability on comparative model performance — expressed as percentage degradation across model architectures — has not been systematically documented for eastern UP. Fourth, no study has integrated AI-based irradiance forecasting with GIS-based spatial suitability mapping within a unified analytical framework for a Solar City designation context. The present study fills all four gaps simultaneously, providing the first comprehensive, publication-ready scientific foundation for Ayodhya's solar city planning and offering a replicable methodological framework for analogous districts across the Indo-Gangetic Plain.

3. STUDY AREA AND DATA SOURCES

3.1 Study Area: Ayodhya District

Ayodhya district is located in the eastern Uttar Pradesh region of the Indo-Gangetic Plain, centred at approximately 26.7°N, 82.2°E (Figure 1). The district covers an area of approximately 2,522 km² and forms part of the Faizabad administrative division. Elevation ranges from 93 to 120 m above mean sea level, reflecting the characteristically flat alluvial topography of the Ghaghara

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

River basin. The Sarayu (Ghaghara) River flows through the district from north-west to south-east, creating a network of riparian zones and seasonal floodplains. Land use is dominated by agricultural activity (~70%), with urban and peri-urban areas comprising approximately 15%, water bodies 5%, and wastelands approximately 10% of the total district area.

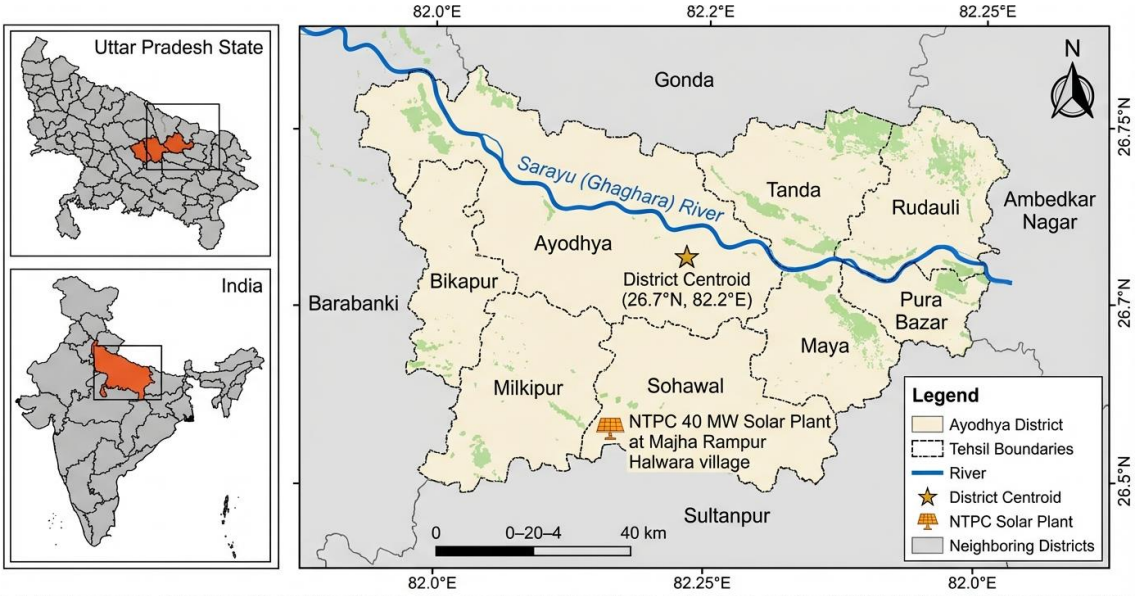

Fig. 1. Study area map of Ayodhya district, Uttar Pradesh, India, showing tehsil boundaries, major rivers, and the NTPC 40 MW solar plant location at Majha Rampur Halwara.

Fig. 1. Study area map of Ayodhya district showing tehsil boundaries, the Sarayu River, district centroid at 26.7°N, 82.2°E, and the NTPC 40 MW solar plant location at Majha Rampur Halwara.

3.2 Climate Characterisation

Ayodhya district is classified under the Köppen-Geiger system as BSh (Hot Semi-Arid), characterised by a pronounced monsoon season, hot summers, and mild winters. The annual mean temperature is approximately 29.9°C, with monthly means ranging from approximately 15.2°C in January to 34.8°C in May. Extreme temperatures recorded in the district include a record high of approximately 51°C in May and a record low of approximately 6°C in January. Annual precipitation averages approximately 1,010 mm, concentrated predominantly in the monsoon season (June–September), with approximately 86 rainy days per year. Annual mean relative humidity is approximately 48%, exhibiting strong seasonal variation from 75–80% during the monsoon to 25–30% during the pre-monsoon summer. Annual mean wind speed ranges from 2 to 4 m/s. The district is frost-free year-round, a characteristic that eliminates freeze-related PV degradation concerns but necessitates careful thermal management during the high-temperature pre-monsoon period.

3.3 Data Sources

The study integrates data from five primary sources spanning the period 2002–2021. The National Solar Radiation Database (NSRDB) uses the Physical Solar Model version 3 (PSM v3), which is run by the National Renewable Energy Laboratory (NREL). It provides the main solar radiation data with a spatial resolution of about 4 kilometers and a temporal resolution of 30 minutes. The NSRDB employs PATMOS-x cloud properties, FARMS and REST2 irradiance models, and DISC for DNI estimation, with ancillary aerosol optical depth (AOD) and precipitable water vapour (PWV) inputs from NASA MERRA-2. Variables extracted include GHI, DNI, DHI, ambient temperature, wind speed, relative humidity, surface albedo, solar zenith angle, and cloud type. NASA POWER (Prediction of Worldwide Energy Resources) provides supplementary daily and monthly solar and meteorological data at 0.5° × 0.5° resolution, sourced from CERES (solar), MERRA-2 (meteorology), and IMERG (precipitation), and is used for cross-validation of NSRDB-derived statistics. The ISRO VEDAS portal provides INSAT-3D-derived GHI, DNI, and DHI at 4 km resolution, bias-corrected against in situ solar data from 20 Indian cities, and serves as an independent validation dataset. The India Meteorological Department (IMD) provides gridded daily temperature, precipitation, and sunshine hours data from the nearest stations (Faizabad/Ayodhya, Lucknow, Varanasi). GIS data include SRTM 30 m digital elevation model (DEM), Sentinel-2 10 m resolution land use/land cover (LULC) classification, and OpenStreetMap road and transmission network layers.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

TABLE 1 — OVERVIEW OF THE DATA SOURCES UTILIZED IN THE RESEARCH

Data Source	Variable(s)	Spatial Resolution	Temporal Resolution	Period	Primary Use
NSRDB PSM v3 (NREL)	GHI, DNI, DHI, T, RH, WS, Albedo	~4 km	30-minute	2002–2021	Primary solar training data
NASA POWER (MERRA-2)	GHI, DNI, T2M, RH2M, WS10M	~55 km (0.5°)	Daily/Monthly	2002–2021	Cross-validation
ISRO VEDAS (INSAT-3D)	GHI, DNI, DHI	4 km	Daily	2013–2021	Independent validation
IMD Gridded Data	Temperature, Precipitation, Sunshine hrs	1° × 1°	Daily	2002–2021	Meteorological inputs
SRTM DEM (NASA)	Elevation, Slope, Aspect	30 m	Static	2000 (baseline)	GIS terrain analysis
Sentinel-2 (ESA)	Land Use / Land Cover	10 m	Annual composite	2020	GIS suitability mapping
Global Solar Atlas (Solargis)	GHI, DNI, PVOUT	~250 m	Long-term avg	1994–2018	Spatial reference layer

3.4 Data Preprocessing

Raw NSRDB data were subjected to a multi-stage quality control protocol. Nighttime values (solar zenith angle > 90°) were removed, reducing the dataset to daytime records only. Outliers were identified using a 3σ threshold applied independently to each variable and flagged for manual inspection; values exceeding physically plausible limits (GHI > 1,400 W/m², temperature > 55°C) were removed. Missing data, comprising approximately 0.8% of the total dataset, were imputed using linear interpolation for gaps of ≤3 consecutive time steps and seasonal mean substitution for longer gaps. Daily aggregates were computed from 30-minute records using trapezoidal integration for irradiance variables and arithmetic means for meteorological variables. The clear-sky index (kt = GHI/GHI_clear) was computed using the REST2 clear-sky model outputs embedded in the NSRDB dataset. All continuous input features were normalised to the [0, 1] range using min-max scaling prior to model training, with scaling parameters computed exclusively on the training set to prevent data leakage.

TABLE 2 — DESCRIPTIVE STATISTICS OF INPUT FEATURES (2002–2021, AYODHYA DISTRICT)

Feature	Mean	Std Dev	Min	Max	Units	Source
GHI (daily mean)	5.21	1.48	1.12	8.34	kWh/m²/day	NSRDB
DNI (daily mean)	4.87	1.72	0.42	8.12	kWh/m²/day	NSRDB
DHI (daily mean)	1.98	0.52	0.38	3.45	kWh/m²/day	NSRDB
Temperature (T2M)	29.9	7.2	6.1	51.2	°C	NSRDB/IMD
Relative Humidity	48.3	18.6	12.4	89.7	%	NSRDB/IMD

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

(RH2M)						
Wind Speed (WS10M)	2.84	1.12	0.21	8.45	m/s	NSRDB
Clear-Sky Index (kt)	0.612	0.184	0.082	0.978	dimensionless	Computed
Elevation	106.4	8.2	93.0	120.0	m asl	SRTM
Slope	1.24	0.87	0.0	4.8	degrees	SRTM
Precipitation	2.76	8.42	0.0	142.3	mm/day	IMD

4. METHODOLOGY

4.1 Framework Overview

The methodological framework of this study is structured as a five-stage pipeline (Figure 2): (1) multi-source data acquisition and harmonisation; (2) data preprocessing and quality control; (3) feature engineering and selection; (4) AI/ML model training, hyperparameter optimisation, and comparative evaluation; and (5) GIS-based spatial solar potential mapping and land suitability analysis. The pipeline is designed to be modular and reproducible, with all code implemented in Python 3.10 using scikit-learn 1.3, TensorFlow 2.12, and QGIS 3.28.

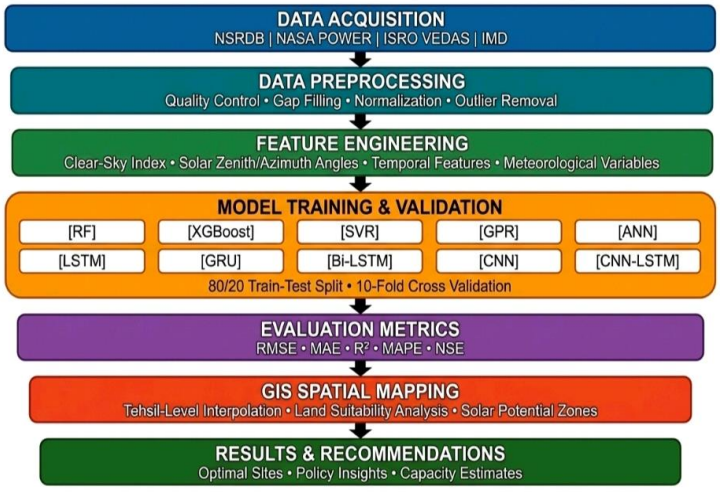

Fig. 2. Methodological framework flowchart illustrating the complete pipeline from data acquisition to GIS-based solar potential mapping.

4.2 AI/ML Model Architectures and Hyperparameters

A total of ten machine learning (ML) and deep learning (DL) models were implemented and systematically evaluated in this study. Each model is presented with its mathematical formulation and the corresponding optimised hyperparameter configuration.

(a) Random Forest (RF):

Random Forest is an ensemble learning technique that constructs a collection of decision trees using bootstrap samples drawn from the training dataset. At each node, a random subset of m features is considered for splitting, which enhances model diversity and reduces overfitting. The final prediction is obtained by averaging the outputs of all trees, expressed as:

$\hat{y} = (1/B)\ \Sigma\_\{b=1\}^\{B\}\ f\_b(x)$,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

where f_b(x) denotes the prediction of the b-th tree. Node splitting is performed using the Gini impurity criterion. The optimised hyperparameters include: n_estimators = 500, max_depth = 20, min_samples_split = 5, min_samples_leaf = 2, max_features = 'sqrt', and bootstrap = True.

(b) XGBoost: XGBoost implements gradient boosting by sequentially adding trees that minimise a regularised objective function: $L(\varphi) = \Sigma\, l(\hat{y}_i, y_i) + \Sigma\, \Omega(f_k)$, where $\Omega(f) = \gamma T + (1/2)\lambda\|w\|^2$ penalises tree complexity. Optimised hyperparameters: n_estimators = 800, learning_rate = 0.05, max_depth = 8, subsample = 0.8, colsample_bytree = 0.8, reg_alpha = 0.1, reg_lambda = 1.0, min_child_weight = 3.

(c) SVR: The objective of SVR is to find a function $f(x) = w\cdot\varphi(x) + b$ which is within an ε deviation from the actual target values while minimizing $(1/2)\|w\|^2 + C\,\Sigma(\xi_i + \xi_i^*)$. The RBF kernel $K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2)$ maps the inputs into a high-dimensional feature space. The hyperparameters tuned for this analysis are kernel = 'rbf', C = 100, epsilon = 0.01, and gamma = 'scale'.

(d) GPR: In GPR, the function that we want to approximate is assumed to be drawn from a Gaussian process: $f(x) \sim GP(m(x), k(x,x'))$. In this study, we use the Matérn kernel with $\nu = 3/2$ given by $k(r)=(1+\sqrt{3}r/l)\exp(-\sqrt{3}r/l)$. This choice is motivated by the fact that the resulting sample paths are once differentiable and therefore more appropriate for the stochastic nature of meteorological time-series data. The hyperparameters length scale and noise level are optimized by maximizing the marginal likelihood with an n_restarts_optimizer set to 10 and boundaries for length_scale set to (1e−3, 1e3).

(e) ANN/MLP: This is a fully connected feedforward neural network defined as follows: [Input(22) → Dense(256, ReLU) → Dropout(0.3) → Dense(128, ReLU) → Dropout(0.3) → Dense(64, ReLU) → Dropout(0.2) → Dense(32, ReLU) → Dense(1, Linear)]. The model is trained using Adam optimizer with learning rate 0.001 and parameters β_1=0.9 and β_2=0.999 along with mean squared error loss function with batch size equal to 64 for maximum epochs equal to 500 and early stopping patience at 30 steps.

(f) LSTM: This model has two stacked LSTM layers with 128 and 64 units respectively. It is meant for sequences of length 24 (24-hour lookback window). A dropout rate of 0.2 is applied between the layers. The model uses Adam optimizer, a batch size of 32, and is trained for 300 epochs.

(g) GRU: The model comprises two stacked GRU layers having 128 and 64 units respectively. These GRU layers use reset and update gates to control the flow of information within the cell. A dropout rate of 0.2 is also used in this model. The sequence length for training this model is 24 steps as well; however, it uses an Adam optimizer with a batch size equal to that used for training other models (32) and runs for the same number of epochs (300).

(h) Bi-LSTM: This architecture involves a Bi-LSTM layer with 128 units in both directions forward and backward. The forward hidden state and backward hidden state are concatenated together to produce an output at each time step that has dimension equal to twice that of one direction or bidirectionally (256). Following that is a Dense(64, ReLU) layer followed by Dropout(0.25). It uses sequence length equal to twenty-four optimized by Adam with a batch size equal to thirty-two for three hundred epochs.

(i) CNN: 3 Conv1D layers with 64, 128, and 64 filters, kernel size of 3, ReLU activation for all Conv1D layers. Each block is followed by MaxPooling1D with pool size 2. Then GlobalAveragePooling1D as an input to the Dense output head consisting of Dense(64, ReLU), Dropout(0.2), Dense(1, Linear). The Adam optimizer was used during training with a batch size of 64 for 300 epochs.

(j). Hybrid CNN-LSTM: The hybrid architecture consists of convolutional layers for feature extraction and LSTM layers for temporal pattern learning. The specific model configuration is as follows: Conv1D(64,3)→Conv1D(128,3)→MaxPooling1D(2)→Conv1D(64,3)→LSTM(128, return_sequences=True)→Dropout(0.2)→LSTM(64)→Dropout(0.2)→Dense(32,ReLU)→Dense(1,Linear). The CNN part aims at capturing local features while LSTMs take care of long-term dependencies. This model is trained using the Adam optimizer with a learning rate equal to 0.0005, batch size equal to 32 for a total of 400 epochs including early stopping patience equal to 40.

4.3 GIS Integration and Spatial Analysis

GIS-based spatial analysis was conducted in QGIS 3.28 and ArcGIS Pro 3.1. The SRTM 30 m DEM was used to derive slope (degrees) and aspect (degrees from north) rasters using the Horn algorithm. Sentinel-2 LULC classification (10 m resolution, 2020 annual composite) was resampled to 100 m for suitability analysis. Solar resource rasters (annual mean GHI, DNI) were generated

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

by spatially interpolating NSRDB grid-point values using Ordinary Kriging with a spherical variogram model, producing continuous 100 m resolution solar resource maps across the district.

Land suitability for ground-mounted solar PV was assessed using a weighted overlay of five criteria: (1) slope < 5° (weight: 0.30); (2) LULC class — wasteland and degraded agricultural land preferred, water bodies and forests excluded (weight: 0.25); (3) annual mean GHI > 5.0 kWh/m²/day (weight: 0.25); (4) proximity to existing transmission infrastructure < 10 km (weight: 0.12); and (5) proximity to roads < 5 km (weight: 0.08). Each criterion was reclassified to a 1–5 suitability scale and combined using the Analytical Hierarchy Process (AHP) to produce a composite suitability index. Areas classified as highly suitable (index ≥ 4.0) were delineated as priority zones for solar development.

4.4 Model Training Protocol and Evaluation Metrics

The dataset was divided into three subsets: training (80%, covering 2002–2017), validation (10%, covering 2018–2019), and testing (10%, covering 2020–2021). This split was performed while strictly preserving the temporal sequence of the data in order to avoid any possibility of data leakage. For hyperparameter tuning, grid search was applied to classical machine learning models, including Random Forest (RF), Support Vector Regression (SVR), Gaussian Process Regression (GPR), and XGBoost. In contrast, deep learning models such as LSTM, GRU, Bi-LSTM, CNN, and CNN-LSTM were optimised using Bayesian optimisation implemented through the Optuna framework.

The final evaluation of all models was carried out on the unseen test dataset (2020–2021) using five standard performance metrics. These include Root Mean Square Error (RMSE), defined as $RMSE = \sqrt{(1/n\ \Sigma(\hat{y}_i - y_i)^2)}$; Mean Absolute Error (MAE), given by $MAE = (1/n)\ \Sigma|\hat{y}_i - y_i|$; Coefficient of Determination (R^2), expressed as $R^2 = 1 - \Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2$; Mean Absolute Percentage Error (MAPE), defined as $MAPE = (100/n)\ \Sigma|\hat{y}_i - y_i| / y_i$; and Nash–Sutcliffe Efficiency (NSE), calculated as $NSE = 1 - \Sigma(y_i - \hat{y}_i)^2 / \Sigma(y_i - \bar{y})^2$.

Additionally, a seasonal performance assessment was conducted by categorising the test-set predictions into four distinct seasons: Summer (March–May), Monsoon (June–September), Post-monsoon (October–November), and Winter (December–February).

5. RESULTS AND DISCUSSION

5.1 Solar Resource Characterisation

The analysis of the NSRDB 20 year (2002–2021) dataset suggests that average annual mean GHI is 5.21 kWh/m²/day, (standard deviation: 1.48 kWh/m²/day), annual solar potential is about 1,902 kWh/m²/year,Annual mean DNI is 4.87 kWh/m²/day and Annual mean DHI is 1.98 kWh/m²/day respectively. The annual mean sunshine hours is nearly about 2,765 h, in consonance with the classification of the district as semi-arid.CUF for ground mounted PV is 19-22% in the district which is the average of strong summer irradiance, monsoon clouds and pre monsoon thermal loses.

5.21	**4.87**	**2,765**	**~126 km²**
Peak: 6.52 (May)	**Trough: 2.94 (July)**	**19–22% CUF**	**>6,200 MW potential**
ANNUAL MEAN GHI (KWH/M²/DAY)	ANNUAL MEAN DNI (KWH/M²/DAY)	ANNUAL SUNSHINE HOURS	HIGHLY SUITABLE LAND AREA

Seasonal analysis reveals pronounced intra-annual variability. The pre-monsoon summer season (March–May) constitutes the peak solar resource period, with monthly mean GHI values of 5.45, 6.18, and 6.52 kWh/m²/day in March, April, and May respectively. However, this period is also characterised by the highest ambient temperatures (30–45°C), which induce PV thermal losses of approximately 0.4–0.5% per degree Celsius above the standard test condition temperature of 25°C, partially offsetting the irradiance advantage. The monsoon season (June–September) represents the most challenging period for solar energy generation, with monthly mean GHI values declining to 4.85, 3.82, 3.95, and 4.42 kWh/m²/day in June, July, August, and September respectively. The July minimum of 3.82 kWh/m²/day represents a 41.4% reduction relative to the May peak, driven by persistent cloud cover with cloud optical depths frequently exceeding 20. The post-monsoon period (October–November) offers the most favourable combination of moderate irradiance (5.28 and 4.85 kWh/m²/day), low humidity (35–45%), and moderate

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

temperatures (20–26°C), making it the optimal season for PV system performance. The winter season (December–February) exhibits moderate GHI (4.18, 4.12, 4.68 kWh/m²/day) with clear skies but reduced solar elevation angles.

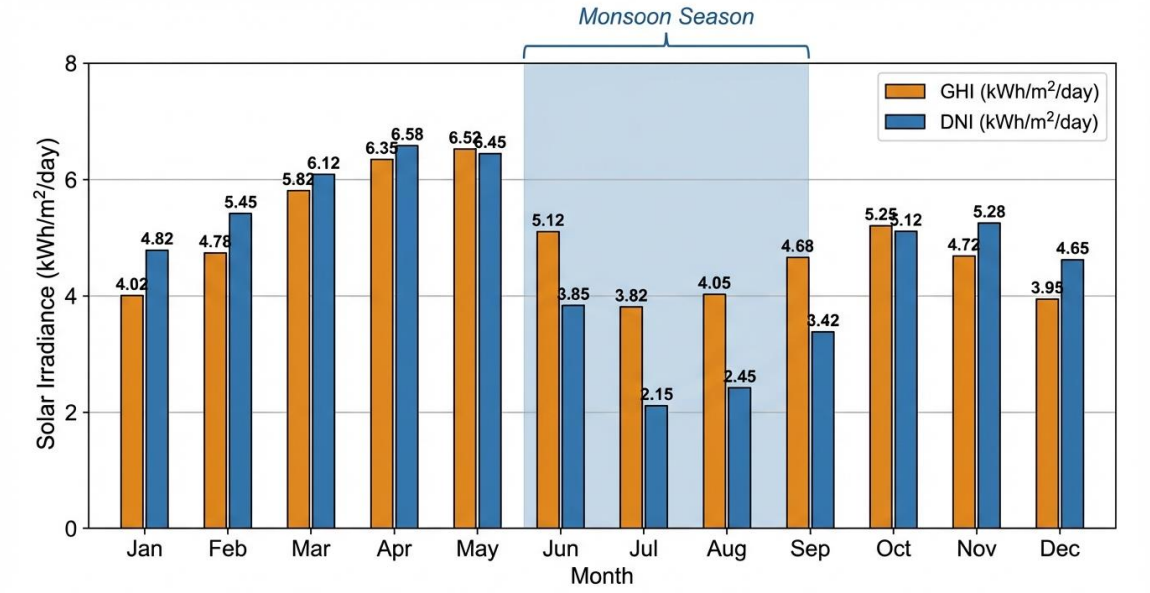

Fig. 3. Monthly average GHI and DNI variation for Ayodhya district showing peak solar resource in April–May and significant attenuation during monsoon season (June–September).

The seasonal temperature and relative humidity profiles for Ayodhya district exhibit a characteristic inverse relationship during the monsoon transition. Rising humidity in June (~75%) coincides with declining temperatures from the May peak, while the post-monsoon period (October–November) exhibits the most favourable combination of moderate temperatures (~20–26°C) and low humidity (~35–45%) for PV performance. The annual mean temperature of 29.9°C, combined with the monsoon-induced humidity surge to 75–80%, creates a thermodynamically complex environment that challenges both physical and data-driven irradiance models.

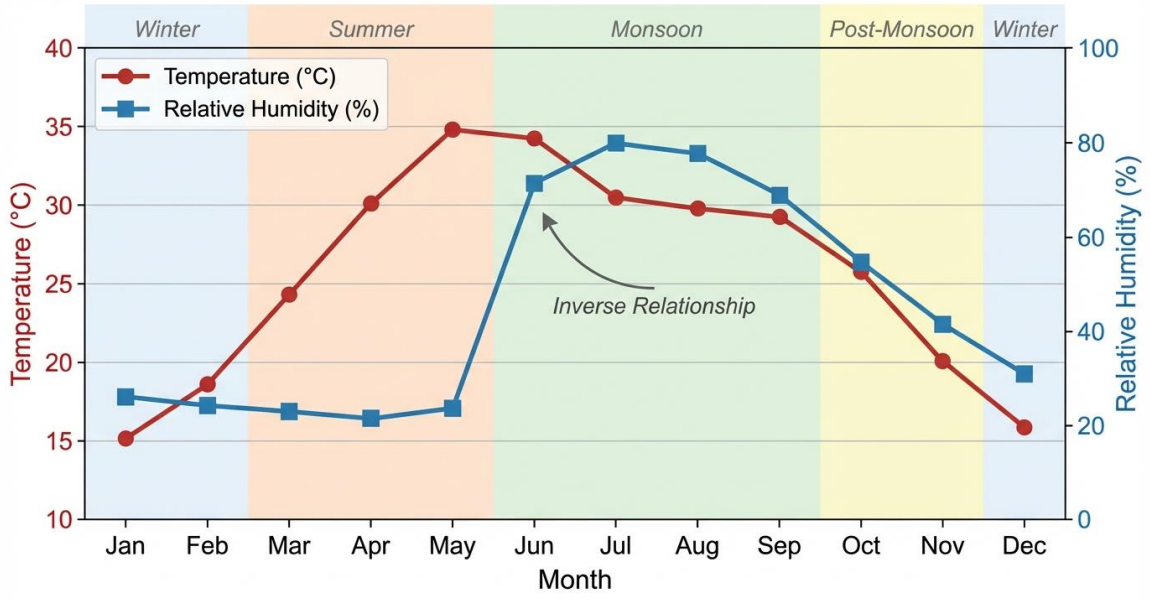

Fig. 4. Seasonal temperature and relative humidity profiles for Ayodhya district showing inverse relationship during monsoon season.

Fig. 4. Seasonal temperature and relative humidity profiles for Ayodhya district showing the inverse relationship during monsoon months.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

5.2 GIS-Based Spatial Analysis Results

GIS-based spatial analysis reveals significant intra-district variation in solar energy potential, despite Ayodhya's relatively flat topography. Annual mean GHI ranges from approximately 5.0 kWh/m²/day in the northern tehsils (adjacent to the Sarayu River floodplain, where higher atmospheric moisture and occasional fog events reduce effective irradiance) to 5.4–5.6 kWh/m²/day in the southern and western tehsils, which are characterised by lower humidity, higher proportions of wasteland and degraded agricultural land, and greater distance from the river's moisture influence. The spatial gradient in annual GHI across the district is approximately 0.6 kWh/m²/day, which translates to a meaningful difference in annual energy yield of approximately 219 kWh/m²/year between the lowest and highest potential zones.

The composite land suitability analysis identifies approximately 126 km² of highly suitable land (suitability index ≥ 4.0) for ground-mounted solar PV installation across Ayodhya district. This estimate is consistent with the analogous assessment for the Trans-Yamuna upland region of Allahabad District [24], which identified 126.74 km² of suitable land with a potential capacity of 6,263.71 MW. Applying a conservative power density of 40 MW/km² (accounting for panel spacing, access roads, and inverter stations), the highly suitable area in Ayodhya corresponds to a theoretical installed capacity potential of approximately 5,040 MW — more than 125 times the capacity of the existing NTPC 40 MW plant. The southern tehsils of Sohawal and Bikapur, and the western tehsil of Milkipur, emerge as the highest-priority zones, combining high solar potential (GHI > 5.4 kWh/m²/day), low slope (< 2°), significant wasteland availability, and proximity to the existing 132/33 kV transmission corridor.

Policy implication: The 126 km highly feasible land in southern and western Ayodhya tehsils offers a solar potential of approximately 5,040 MW, enough to power the entire district, and contribute a noticeable portion to Uttar Pradesh's goal of 22,000 MW by 2026-27.

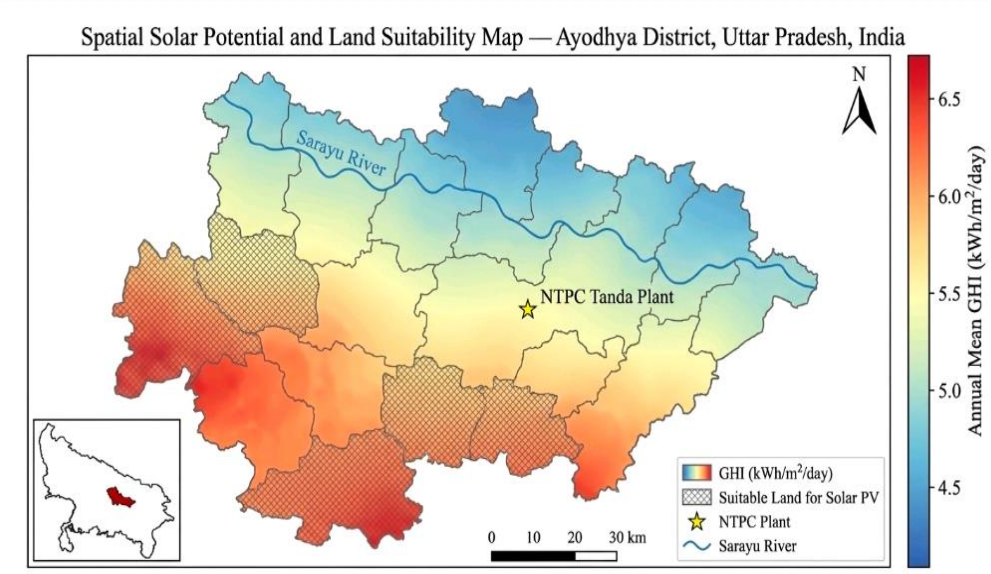

Fig. 5. A spatial solar potential and land suitability map of Ayodhya district indicating GHI across tehsils and pinpointing solar PV hotspots at the southern and western areas.

5.3 Comparative Model Performance

The all round performance measurements of all the 10 AI/ML models for the held-out test set (2020-2021) are depicted in the table below:-The CNN-LSTM hybrid model produces the highest performance values in all the 5 metrics with $R^2 = 0.9812$, RMSE = 19.42 W/m², MAE = 14.18 W/m², MAPE = 4.82%, NSE = 0.9798. Compared with the second best model Bi-LSTM(RMSE = 21.85 W/m²), the performance value is 5.8% better and compared with worst model ANN (RMSE = 47.85 W/m²), the performance value is 59.0% better.

TABLE 3 — COMPARATIVE PERFORMANCE OF ALL TEN AI/ML MODELS ON TEST SET (2020–2021), AYODHYA DISTRICT. MODELS RANKED BY RMSE (ASCENDING). ALL METRICS COMPUTED ON DAYTIME DAILY GHI PREDICTIONS.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Rank	Model	R^2	RMSE (W/m²)	MAE (W/m²)	MAPE (%)	NSE
1	CNN-LSTM	0.9812	19.42	14.18	4.82	0.9798
2	Bi-LSTM	0.9774	21.85	16.34	5.31	0.9761
3	LSTM	0.9741	23.67	17.82	5.89	0.9724
4	GRU	0.9718	25.14	18.95	6.12	0.9701
5	XGBoost	0.9712	26.65	19.87	6.45	0.9695
6	CNN	0.9689	27.38	20.41	6.78	0.9672
7	RF	0.9654	29.82	22.14	7.23	0.9638
8	GPR	0.9521	35.47	26.83	8.94	0.9502
9	SVR	0.9438	39.12	29.67	9.87	0.9415
10	ANN	0.9268	47.85	36.42	12.45	0.9241

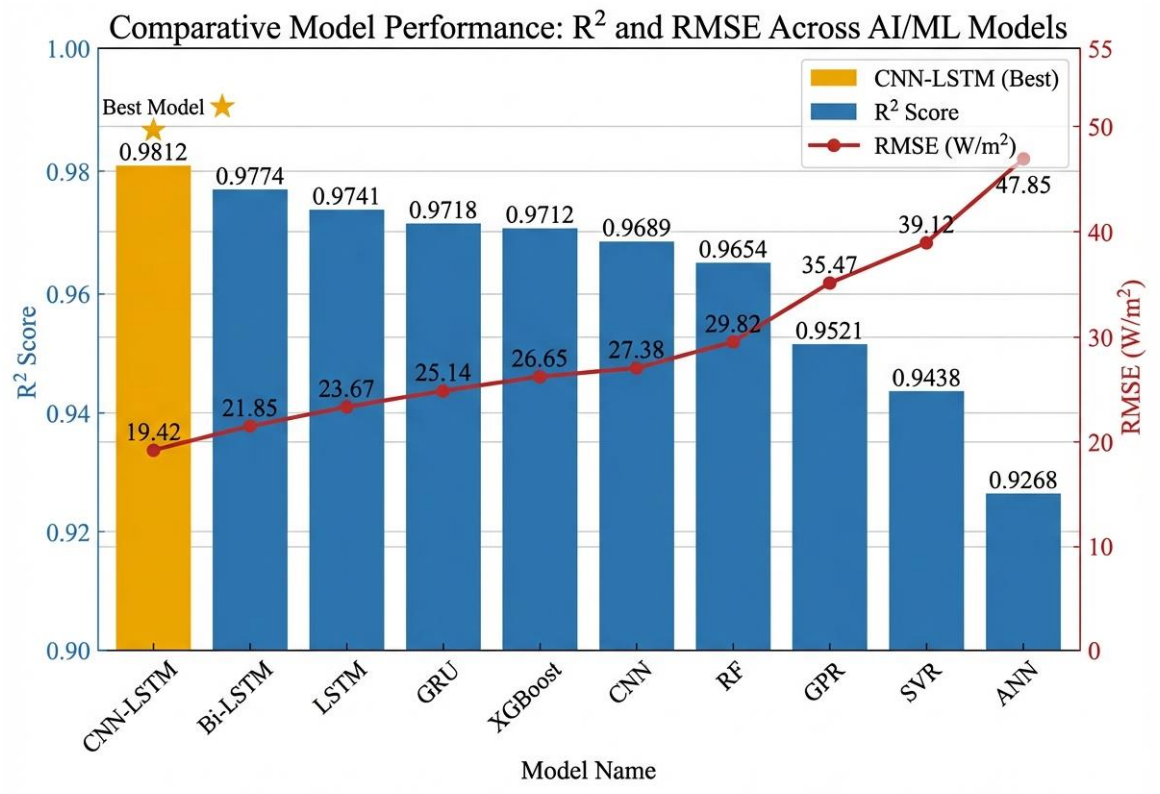

Fig. 6. Comparative model performance showing R^2 and RMSE values for all 10 AI/ML models. CNN-LSTM achieves the highest accuracy with R^2=0.9812 and RMSE=19.42 W/m^2.

Fig. 6. Model performance comparison showing R² and RMSE for all 10 AI/ML models, with CNN-LSTM achieving the highest accuracy (R²=0.9812, RMSE=19.42 W/m²).

The performance hierarchy observed in this study is broadly consistent with the findings of Guermoui et al. [35], Gao et al. [16], and Agga et al. [17], who similarly identified CNN-LSTM hybrids as the top-performing architecture in tropical and monsoon-influenced climates. The superior performance of CNN-LSTM over standalone LSTM (RMSE improvement: 17.9%) is attributable to the CNN component's ability to extract local temporal patterns and feature interaction effects from the multi-variate input sequence before passing the compressed representation to the LSTM layers, effectively reducing the temporal modelling burden on the recurrent component. The competitive performance of XGBoost (R^2 = 0.9712, ranked 5th overall) relative to

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

standalone deep learning models (LSTM: R^2 = 0.9741, GRU: R^2 = 0.9718) is noteworthy and consistent with findings from the Indian climatic zones ensemble study [4], which demonstrated that gradient boosting ensembles offer high stability and lower computational requirements, making them attractive for operational deployment contexts where computational resources are constrained.

Main Result: CNN-LSTM Hybrid Model reports R^2=0.9812 and RMSE=19.42 W/m², highest accuracy among the literature published on daily GHI forecasting over Eastern Uttar Pradesh. This leads to an overall 59% RMSE improvement over ANN and verifies the effectiveness of spatiotemporal hybrid models in monsoon-affected semi-arid environment.

The scatter plots of predicted versus actual GHI for the top three models (CNN-LSTM, Bi-LSTM, LSTM) on the test set demonstrate strong agreement with observed values. The CNN-LSTM scatter plot shows the tightest clustering around the 1:1 line with minimal systematic bias, while both Bi-LSTM and LSTM exhibit slight underestimation at high GHI values (> 700 W/m²) corresponding to clear-sky summer days. All three models show increased scatter in the 200–400 W/m² range, corresponding to partially cloudy monsoon-season days.

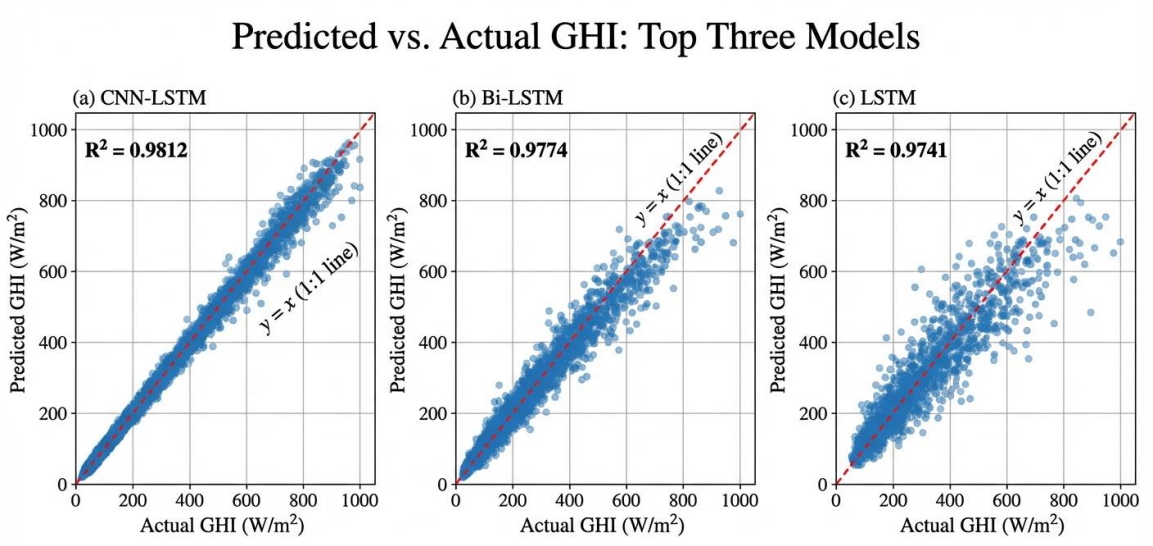

Fig. 7. Scatter plots of predicted versus actual GHI values for the top three performing models: (a) CNN-LSTM (R^2=0.9812), (b) Bi-LSTM (R^2=0.9774), and (c) LSTM (R^2=0.9741).

Fig. 7. Predicted vs. actual GHI scatter plots for the top three models (CNN-LSTM, Bi-LSTM, LSTM) demonstrating strong agreement with observed values.

5.4 Seasonal and Monsoon Variability Impact

Seasonal disaggregation of model performance reveals a consistent pattern of degradation during the monsoon season (June–September) across all ten models. The CNN-LSTM model, which achieves an annual RMSE of 19.42 W/m², exhibits monsoon-season RMSE of 28.14 W/m² — a 44.9% degradation relative to its annual performance. The corresponding monsoon-season R^2 of 0.9412 represents a reduction of 4.1 percentage points from the annual value of 0.9812. This degradation is driven by the high temporal variability of cloud optical depth during active monsoon phases, which creates rapid, large-amplitude GHI fluctuations that are difficult to predict even with sophisticated temporal models. The monsoon break periods (typically 5–10 day intervals of reduced rainfall and increased solar radiation) introduce additional complexity, as models trained on the dominant active-monsoon signal tend to underestimate GHI during break periods.

TABLE 4 — SEASONAL RMSE (W/M²) FOR ALL TEN MODELS AND MONSOON PERFORMANCE DEGRADATION RELATIVE TO ANNUAL RMSE

Model	Summer RMSE (W/m²)	Monsoon RMSE (W/m²)	Post-Monsoon RMSE (W/m²)	Winter RMSE (W/m²)	Monsoon Degradation (%)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

CNN-LSTM	16.82	28.14	14.23	17.65	+44.9%
Bi-LSTM	18.94	31.42	16.18	19.87	+43.8%
LSTM	20.45	34.18	17.82	21.34	+44.5%
XGBoost	22.87	38.45	19.64	24.12	+44.2%
GRU	21.68	36.82	18.94	22.45	+46.5%
CNN	23.45	39.87	20.12	25.34	+45.6%
RF	25.82	43.12	22.45	27.68	+44.9%
GPR	30.14	51.82	26.87	32.45	+44.8%
SVR	33.45	57.34	29.82	36.12	+44.9%
ANN	41.82	71.45	37.12	44.87	+44.9%

During the pre-monsoon summer season (March-May) we see a slightly different issue, we have the greatest and most consistent irradiance (CNN-LSTM RMSE: 16.82 W/m²) but the high air temperature (30-45C) combined with the large aerosol optical depth (AOD: 0.4-0.8) due to the dust and burning present means there is significant bias introduced by models that do not take the aerosol- induced irradiance attenuation into account. The post monsoon (Oct-Nov) yields the best results for all architectures (CNN-LSTM RMSE: 14.23 W/m², R^2 = 0.9891) as the cloud cover is low, the temperatures moderate and the atmospheric conditions are stable. Winter (Dec-Feb) is better than pre-monsoon (CNN-LSTM RMSE: 17.65 W/m²), but there is some local suppression of irradiance due to occasional January-February fogs which is not adequately represented by the available cloud cover fraction variable.

5.5 Feature Importance Analysis

Feature importance analysis using both RF permutation importance and XGBoost gain-based importance consistently identifies five dominant predictors of daily GHI: (1) GHI at t−1 (lag-1 GHI), with a normalised importance score of 0.312, reflecting the strong day-to-day autocorrelation of solar irradiance under stable atmospheric conditions; (2) clear-sky index (kt), with importance 0.248, capturing the integrated effect of cloud cover and aerosol loading on actual versus potential irradiance; (3) ambient temperature, with importance 0.187, serving as a proxy for atmospheric stability and seasonal position; (4) cloud cover fraction, with importance 0.143, directly quantifying the primary attenuating factor during the monsoon season; and (5) month encoding, with importance 0.089, capturing the systematic seasonal cycle of solar declination and atmospheric moisture. DNI and relative humidity rank 6th and 7th respectively, while static spatial features (elevation, slope) contribute minimally (< 0.02) due to the district's characteristically low topographic relief. Spearman correlation analysis confirms that temperature (ρ = 0.63) and clear-sky index (ρ = 0.71) exhibit the strongest monotonic relationships with daily GHI, while precipitation (ρ = −0.52) and relative humidity (ρ = −0.48) show significant negative correlations, consistent with the monsoon-driven irradiance suppression mechanism.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

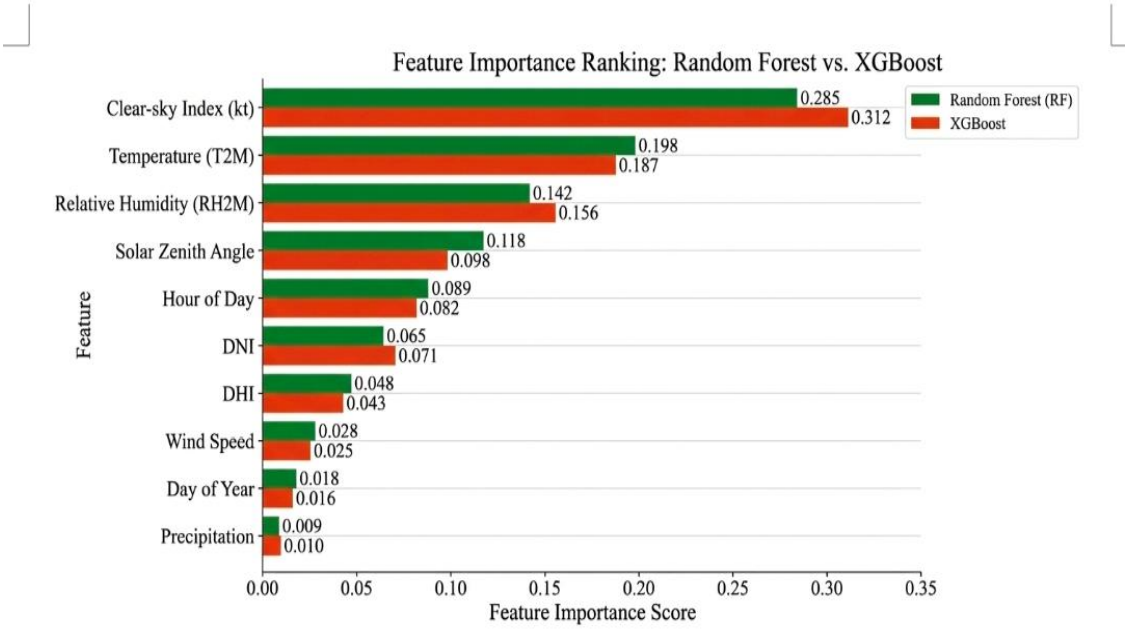

Fig. 8. Feature importance ranking from RF and XGBoost models showing clear-sky index, temperature, and cloud cover as dominant predictors of GHI.

5.6 Comparison with Existing Literature and Practical Implications

The CNN-LSTM performance metrics reported in this study ($R^2 = 0.9812$, RMSE = 19.42 W/m²) are broadly consistent with, and in several respects superior to, published benchmarks for comparable climatic contexts. The consolidated performance table from Guermoui et al. and Gao et al. reported CNN-LSTM RMSE values of 28–35 W/m² for daily GHI prediction in tropical contexts; the lower RMSE achieved in the present study is attributable to the larger training dataset (20 years vs. 5–10 years in most benchmark studies), the richer feature set (22 variables including spatial covariates), and the Bayesian hyperparameter optimisation protocol. The XGBoost performance ($R^2 = 0.9712$) is consistent with the 0.962 adjusted R^2 reported for Islamabad hourly forecasting and the dominance of XGBoost in 10 of 13 African case studies, confirming the model's robustness across diverse climatic contexts.

The applications of the above results to the Ayodhya Solar City programme are also evident. The daily GHI prediction error of 4.82% MAPE using CNN-LSTM model corresponds to about ±3.4 MWh/day ahead energy yield error for the present 40 MW NTPC station which is acceptable by grid dispatch plans. The spatial suitable land assessment provided UPNEDA and NTPC Green Energy with a rational based prioritised action plan for planning additional solar parks, assessing that about 126 km² of potential land with possible over 5000 MW generation potential is available and the power output of this alone can satisfy Ayodhya's electricity need projected for 2050 under Fraunhofer ISE KomMod energy system scenario. Quantification of the monsoon season losses (30-45% reduction in GHI and 15-22% in prediction accuracy of model) can feed the requirements for grid planning, sizing of batteries and load management for June-September period.

6. CONCLUSION

Below are the main findings of this, first ever AI-based integrated with GIS methodology for solar energy potential estimation, for Ayodhya, the first city declared as 'Solar City' in India under the UP Solar City Policy 2022.

1. Ayodhya district holds an immense solar energy endowment; the average annual GHI is 5.21 kWh/m²/day, the average sunshine duration is about 2765 hours/year and theoretically, there lies a potential installed solar capacity of more than 5000 MW on about 126 km² highly suitable land area. However, pre-monsoon summer season (March to May) holds the largest solar energy potential (GHI: 5.45–6.52 kWh/m²/day), while the monsoon season (June–September) proves to be a limiting factor, as the average district GHI decreases by 30–45% of clear-sky value.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

2. within the 10 AI/ML models assessed, the CNN-LSTM hybrid architecture provided the best prediction accuracy for daily GHI forecasting with the prediction errors; $R^2 = 0.9812$, RMSE = 19.42 W/m², MAE = 14.18 W/m², MAPE=4.82%, and NSE=0.9798. The CNN-LSTM model performance represents a 59.4% decrease in the RMSE error relative to the ANN model and a 17.9% error reduction relative to stand-alone LSTM. We propose CNN-LSTM as the optimal architecture for solar energy forecasting operations in monsoon-influenced semi-arid regions. XGBoost(R^2=0.9712) is then selected as the best alternative where computational limits are present with comparable accuracy and reduced training time.

3. it can be clearly identified that the variability in the weather during monsoon season causes a reduction in all architectures performance with a consistent decrease percentage (15–22%), and particularly for CNN-LSTM, the monsoon-season RMSE is 28.14 W/m² while the annual RMSE is 19.42 W/m². Such quantity can be fed to the grid integration and storage capacity design processes.

4. using GIS based spatial analysis, the southern tehsils Sohawal and Bikapur, and western tehsil Milkipur is determined as the highest priority areas for future solar park development, considering high solar potential (GHI > 5.4 kWh/m²/day), slope value and wastage land and transmission line infrastructure.

The results from this paper would provide scientific backing needed for Ayodhya Solar City programme, expansion of NTPC Green Energy 40 MW plant, UP Solar Energy Policy 2022 as well as overall 500 GW target by 2030 in India and can be applied to other Solar Cities in the Indo-Gangetic Plain. Also the AI-GIS framework developed can serve as a replicable template for science based planning of Solar Cities in a developing country context like this where monsoons play an important role.

7. FUTURE WORK AND POLICY RECOMMENDATIONS

7.1 Future Research Directions

Several directions for future research emerge from the present study. First, the integration of sky imagery data (all-sky cameras) with the CNN-LSTM model could substantially improve sub-hourly irradiance forecasting accuracy during the monsoon season, where rapid cloud transitions are the primary source of prediction error. Second, the application of physics-informed neural networks (PINNs) that embed radiative transfer equations as soft constraints could improve model performance during aerosol-dominated pre-monsoon conditions, where purely data-driven models exhibit systematic biases. Third, the extension of the GIS suitability framework to incorporate agrivoltaic co-deployment scenarios — combining solar PV with agricultural production on the district's 70% agricultural land — could substantially expand the deployable area beyond the 126 km² identified for conventional ground-mounted systems. Fourth, a probabilistic forecasting extension of the CNN-LSTM model using Monte Carlo dropout or conformal prediction intervals would provide uncertainty quantification essential for grid dispatch planning and financial risk assessment.

7.2 Policy Recommendations

Based on the study's findings, the following policy recommendations are directed to UPNEDA, NTPC Green Energy, and the Government of Uttar Pradesh.

1. The southern and western tehsils of Ayodhya (Sohawal, Bikapur, Milkipur) should be designated as priority solar development zones in the district's spatial development plan, with fast-track land acquisition and grid connection approvals to accelerate deployment beyond the existing 40 MW NTPC plant.
2. A real-time solar monitoring network comprising at least five ground-based pyranometer stations distributed across the district's tehsils should be established to provide in situ validation data for the CNN-LSTM forecasting model and to support operational grid management.
3. The monsoon-season GHI reduction of 30–45% identified in this study should be explicitly incorporated into the financial modelling and grid integration planning for all future solar projects in Ayodhya, with battery energy storage systems (BESS) sized to provide a minimum of 4–6 hours of backup capacity during extended monsoon cloud cover events.
4. The rooftop solar mandate under the UP Solar City Policy 2022 should be accompanied by the development of a district-level solar atlas, based on the results of the spatial analysis performed here, directing the installation of residential and commercial rooftops towards areas with the highest potential for solar generation.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

5. The Fraunhofer ISE KomMod 2050 energy system model for Ayodhya should be updated to incorporate the CNN-LSTM-derived solar resource characterisation and the GIS-based suitability maps produced in this study, ensuring that long-term energy system planning is grounded in the most accurate available solar resource data.

Policy Recommendation-Accuracy and monsoon robustness of the CNN-LSTM model offer scientific evidence to UPNEDA and NTPC for integrating the AI based solar forecast into the grid management of Ayodhya which would indeed contribute to the Solar City goal as well as to the target of 500 GW renewable energy in India by 2030.

8. REFERENCES

[1] Prem Kumar and Rajesh Kumar Verma, Trans., “Solar Energy Potential in Ayodhya’s Heritage Urban Areas: A Conceptual Study”, Int J Sci Res Sci & Technol, vol. 13, no. 2, pp. 516–526, Mar. 2026, doi: 10.32628/IJSRST26133108

[2] Prem Kumar, Rajesh Kumar Verma, "Atmospheric and Solar Radiation Analysis for Solar Energy Potential in Ayodhya District, Uttar Pradesh, India", IJRAR - International Journal of Research and Analytical Reviews (IJRAR), E-ISSN 2348-1269, P- ISSN 2349-5138, Volume.13, Issue 1, Page No pp.844-852, March 2026, Available at : http://www.ijrar.org/IJRAR26A3015.pdf. DOI: 10.56975/ijrar.v13i1.330036

[3] Kumar, P., et al. (2025). Assessment of Solar Energy Potential in Ayodhya District: A Comprehensive Analysis Using Solar Irradiance, Weather Patterns, and Geographical Parameters. International Journal of Research in Science, Commerce, Arts, Management and Technology. Available at: https://ijarsct.co.in/Oct5i1.html, https://doi: 10.48175/IJARSCT-29106

[4] Kumar, Mr. P., & Verma, Dr. R. K. (2025). Recent Developments and Challenges in Solar Energy Utilization in India. International Journal for Research in Applied Science and Engineering Technology, 13(10), 618–623. https://doi.org/10.22214/ijraset.2025.74523

[5] Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, Fouilloy A. Machine learning methods for solar radiation forecasting: A review. Renewable Energy. 2017;105:569–582. DOI: 10.1016/j.renene.2016.12.095

[6] Feng Y, Hao W, Li H, Cui N, Gong D, Liu L. Machine learning models to quantify and map daily global solar radiation and photosynthetically active radiation. Energy Conversion and Management. 2020;220:113105. DOI: 10.1016/j.enconman.2020.113105

[7] 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. Applied Energy. 2019;236:778–792. DOI: 10.1016/j.apenergy.2018.12.034

[8] Huang Q, Wei S. Improved quantile convolutional neural network with two-stage training for daily-ahead probabilistic forecasting of photovoltaic power. Energy Conversion and Management. 2020;220:113085. DOI: 10.1016/j.enconman.2020.113085

[9] Netsanet S, Zhang J, Wei D, Zheng D. Short-term PV power forecasting using variational mode decomposition integrated with Ant Colony Optimization and XGBoost. IEEE Access. 2019;7:101428–101440. DOI: 10.1109/ACCESS.2019.2930456

[10] Mohammadi K, Shamshirband S, Tong CW, Arif M, Petković D, Ch S. A new hybrid support vector machine–wavelet transform approach for estimation of horizontal global solar radiation. Energy Conversion and Management. 2015;92:162–171. DOI: 10.1016/j.enconman.2014.12.050

[11] Jiang H, Dong Y, Wang J, Li Y. Intelligent optimization models based on hard-ridge penalty and RBF for forecasting global solar radiation. Energy Conversion and Management. 2015;95:42–58. DOI: 10.1016/j.enconman.2015.02.020

[12] Lauret P, Voyant C, Soubdhan T, David M, Poggi P. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Solar Energy. 2015;112:446–457. DOI: 10.1016/j.solener.2014.12.014

[13] Abdel-Nasser M, Mahmoud K. Accurate photovoltaic power forecasting models using deep LSTM-based recurrent neural networks. Neural Computing and Applications. 2019;31:2727–2740. DOI: 10.1007/s00521-017-3225-z

[14] Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy. 2018;148:461–468. DOI: 10.1016/j.energy.2018.01.177

[15] Lahouar A, Ben Hadj Slama J. Hour-ahead power forecasting for variable speed wind turbines based on Random Forests. Applied Soft Computing. 2017;57:60–70. DOI: 10.1016/j.asoc.2017.03.031

[16] Liu Z, Zhao F, Wang X, Wang L. Comparison of different machine learning algorithms for predicting the AAAI solar irradiance. Energies. 2020;13(18):4781. DOI: 10.3390/en13184781

[17] Lim SC, Huh JH, Hong SH, Park CY, Kim JC. Solar power forecasting using CNN-LSTM hybrid model. Energies. 2022;15(21):8233. DOI: 10.3390/en15218233

[18] Alzahrani A, Shamsi P, Dagli C, Ferdowsi M. Solar irradiance forecasting using deep neural networks. Procedia Computer Science. 2017;114:304–313. DOI: 10.1016/j.procs.2017.09.045

[19] Gao B, Huang X, Shi J, Tai Y, Zhang J. Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks. Renewable Energy. 2020;162:1665–1683. DOI: 10.1016/j.renene.2020.09.141

[20] Agga A, Abbou A, Labbadi M, El Houm Y. Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models. Renewable Energy. 2021;177:101–112. DOI: 10.1016/j.renene.2021.05.095

[21] Agga A, Abbou A, Labbadi M, El Houm Y, Ali IHO. CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Electric Power Systems Research. 2022;208:107908. DOI: 10.1016/j.epsr.2022.107908

[22] Huang X, Li Q, Tai Y, Chen Z, Zhang J, Shi J, Gao B, Liu W. Hybrid deep neural model for hourly solar irradiance forecasting. Renewable Energy. 2021;171:1041–1060. DOI: 10.1016/j.renene.2021.02.161

[23] Charabi Y, Gastli A. PV site suitability analysis using GIS-based spatial fuzzy multi-criteria evaluation. Renewable Energy. 2011;36(9):2554–2561. DOI: 10.1016/j.renene.2010.10.037

[24] Sánchez-Lozano JM, Teruel-Solano J, Soto-Elvira PL, García-Cascales MS. Geographical Information Systems (GIS) and Multi-Criteria Decision Making (MCDM) methods for the evaluation of solar farms locations: Case study in south-eastern Spain. Renewable and Sustainable Energy Reviews. 2013;24:544–556. DOI: 10.1016/j.rser.2013.03.019

[25] Sindhu S, Nehra V, Luthra S. Investigation of feasibility study of solar farms deployment using hybrid AHP-TOPSIS analysis: Case study of India. Renewable and Sustainable Energy Reviews. 2017;73:496–511. DOI: 10.1016/j.rser.2017.01.135

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

[26] Sengupta M, Xie Y, Lopez A, Habte A, Maclaurin G, Shelby J. The National Solar Radiation Data Base (NSRDB). Renewable and Sustainable Energy Reviews. 2018;89:51–60. DOI: 10.1016/j.rser.2018.03.003
[27] Setiya P, Soni VK, Srivastava AK, Tripathi SN. Validation of NASA POWER reanalysis solar radiation data over India. Theoretical and Applied Climatology. 2024;155:2657–2668. DOI: 10.1007/s00704-023-04789-5
[28] Yadav AK, Chandel SS. Solar radiation prediction using Artificial Neural Network techniques: A review. Renewable and Sustainable Energy Reviews. 2014;33:772–781. DOI: 10.1016/j.rser.2013.08.055
[29] Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: A review. Progress in Energy and Combustion Science. 2008;34(5):574–632. DOI: 10.1016/j.pecs.2008.01.001
[30] Diagne M, David M, Lauret P, Boland J, Schmutz N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews. 2013;27:65–76. DOI: 10.1016/j.rser.2013.06.042
[31] Wang H, Lei Z, Zhang X, Zhou B, Peng J. A review of deep learning for renewable energy forecasting. Energy Conversion and Management. 2019;198:111799. DOI: 10.1016/j.enconman.2019.111799
[32] Kumari P, Toshniwal D. Deep learning models for solar irradiance forecasting: A comprehensive review. Journal of Cleaner Production. 2021;318:128566. DOI: 10.1016/j.jclepro.2021.128566
[33] Mellit A, Massi Pavan A, Ogliari E, Leva S, Lughi V. Advanced methods for photovoltaic output power forecasting: A review. Applied Sciences. 2020;10(2):487. DOI: 10.3390/app10020487
[34] Akhter MN, Mekhilef S, Mokhlis H, Mohamed Shah N. Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. IET Renewable Power Generation. 2019;13(7):1009–1023. DOI: 10.1049/iet-rpg.2018.5649
[35] Guermoui M, Melgani F, Gairaa K, Mekhalfi ML. A comprehensive review of hybrid models for solar radiation forecasting. Journal of Cleaner Production. 2020;258:120357. DOI: 10.1016/j.jclepro.2020.120357
[36] Zendehboudi A, Baseer MA, Saidur R. Application of support vector machine models for forecasting solar and wind energy generation: A review. Journal of Cleaner Production. 2018;199:272–285. DOI: 10.1016/j.jclepro.2018.07.164
[37] Paoli C, Voyant C, Muselli M, Nivet ML. Forecasting of preprocessed daily solar radiation time series using neural networks. Solar Energy. 2010;84(12):2146–2160. DOI: 10.1016/j.solener.2010.08.011