Sigma J Eng Nat Sci, Vol. 43, No. 5, pp. 1733–1744, October, 2025

Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2024.00128

Research Article

Global solar radiation prediction using machine learning approaches

M. MOHANDES[1,3], Hilal H. NUHA[2], Satria Akbar MUGITAMA[2], S. REHMAN[3,*], A. AL-SHAILKHI[1]

[1]*Department of Electrical Engineering, King Fahd University of Petroleum & Minerals, Dhahran, 31261, Saudi Arabia*
[2]*Telcom University, Bandung, 40257, Indonesia*
[3]*King Fahd University of Petroleum & Minerals (KFUPM), Interdisciplinary Research Center for Sustainable Energy Systems (IRC-SES), Dhahran, 31261, Saudi Arabia*

ARTICLE INFO

Article history
Received: 15 July 2024
Revised: 22 August 2024
Accepted: 11 September 2024

Keywords:
BiLSTM; ConvNet; GSR; GRU

ABSTRACT

Global Solar Radiation (GSR) stands as a crucial renewable energy source for electricity and heat generation without emitting greenhouse gases. Fuelled by escalating fossil fuel prices, the necessity to curb greenhouse gases (GHG) emissions, and the rapid advancement of solar technology; the role of GSR becomes pivotal in shaping the energy landscape. So, it becomes imperative to understand the variability and availability of GSR on various time scales in the temporal domain. This research conducts an in-depth comparative analysis of various machine learning models, including Long-Short Term Memory (LSTM), Bidirectional LSTM (BiLSTM), Gated Recurrent Unit (GRU), Convolutional Neural Networks (ConvNet), Multilayer Perceptron (MLP), Generalized Additive Model (GAM), Gaussian Process Regression (GPR), and Linear Regression (LR) for GSR prediction to recommend the best method/s for the purpose. Employing robust evaluation metrics such as: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Biased Error (MBE), and the coefficient of determination (R2), the study examines the predictive capabilities of these models. The numerical experimental results show that BiLSTM emerges as the standout performer, having minimal deviation from actual values and slightly positive bias. Its remarkable R^2 value (99.26%) highlights its predictive capability.

Cite this article as: Mohandes M, Nuha HH, Mugitama SK, Rehman S, Al-Shailkhi A. Global solar radiation prediction using machine learning approaches. Sigma J Eng Nat Sci 2025;43(5):1733–1744.

INTRODUCTION

Global solar radiation is important for a number of reasons. First, it is a renewable energy source that can be used to generate electricity and heat without emitting greenhouse gases (GHG). Second, it is a distributed source that is available in most parts of the world. Third, it is a cost-effective and is becoming increasingly affordable as solar technology continues to improve. In 2022, the global solar energy market was worth $160 billion. It is expected to grow to $280 billion by 2027 [1]. The growth of the solar energy

***Corresponding author.**
*E-mail address: srehman@kfupm.edu.sa; mohandes@kfupm.edu.sa

This paper was recommended for publication in revised form by Editor-in-Chief Ahmet Selim Dalkilic

market is being driven by a number of factors, including the increasing cost of fossil fuels, the need to reduce GHG emissions, and the growing availability of solar technology.

Predicting GSR ahead of time is important for several reasons. First, it can help to better understand the cycle of the Earth's energy balance and climate. Second, it can be used to optimize the design and operation of solar energy systems. Third, it can help to mitigate the effects of climate change. However, despite its periodicity, GSR modelling is challenging. As Fig. 1 shows, there are variations in hourly GSR averages over 24 hours on different days (Fig. 1a) and monthly GSR averages over 12 months in different years (Fig. 1b). The GSR data was collected at King Fahd University of Petroleum & Minerals (KFUPM) from 2012 to 2020. Research on GSR prediction is ongoing globally. Scientists are working to develop more accurate and reliable models that can be used to support the development and operation of solar energy systems [2–5]. Researchers are making efforts to develop models to estimate the effects of GSR intensities on climate change.

In recent years, numerous studies have been conducted to investigate solar radiation prediction models and their applications in various fields [6]. Sun et al. [7] studied the effect of solar radiation and other meteorological variables on the its predictability using autoregressive moving average combined with exogenous variable–generalized autoregressive conditional heteroscedasticity (ARMAX–GARCH) and multivariate GARCH (MGARCH) models. Authors reported good correlations between GSR and sunshine duration compared to ambient temperature. Alsina et al. [8] used Artificial Neural Networks (ANNs) to predict Monthly Average Daily GSR (MADGSR) over Italy and achieved better performance with MAPE ranging from 1.67% to 4.25%. The ANN best configuration included seven relevant input parameters, highlighting the efficacy of ANN for solar radiation prediction.

The authors [9] experiments in remote areas for milk storage systems using solar thermal energy and adsorption cooling. The authors [10] research by creating a pollution monitoring and forecasting system for healthy breathing and living. Sharnil The authors [11] conducted research to build a framework based on sensor fusion and deep learning for virus outbreaks. The authors [12] build a framework for the MQTT multicast messaging protocol for IoT-related problems. The authors [13] conducted research on the effect of the filling and emptying process of PCM with paraffin and AI_2O_3 additives at three locations. The authors [14] carried out experiments to analyse and theorise on exhaust manifolds with uncoated ceramics and with AI_2O_3, TiO_2 and ZrO_2 coatings.

The authors [15] conducted research to study the drying of coriander seeds in a photovoltaic thermal collector with a solar dryer. The authors [16] undertook research into the design and manufacture of a two-stage, energy-efficient, nickel-chromium-reinforced pyrolysis reactor for plastic waste applications. The authors [17] carried out research on the effects of emissions from DI-based sources that may cause adverse health effects. Kumar et al. [18] reported global challenges and emerging trends being faced by the scientific and engineering community in forecasting thermal performance degradation of WTG gearboxes. Thermal performance of WTG gearboxes degradation in relation to plant loads was carried out by experimental measurements of temperature flux measurements with condition monirtoring systems by Kumar et al. [19] Dhabale et al. [20] conducted on the multi-objective optimisation of friction stir welding process parameters for copper particle reinforcement in aluminium alloys

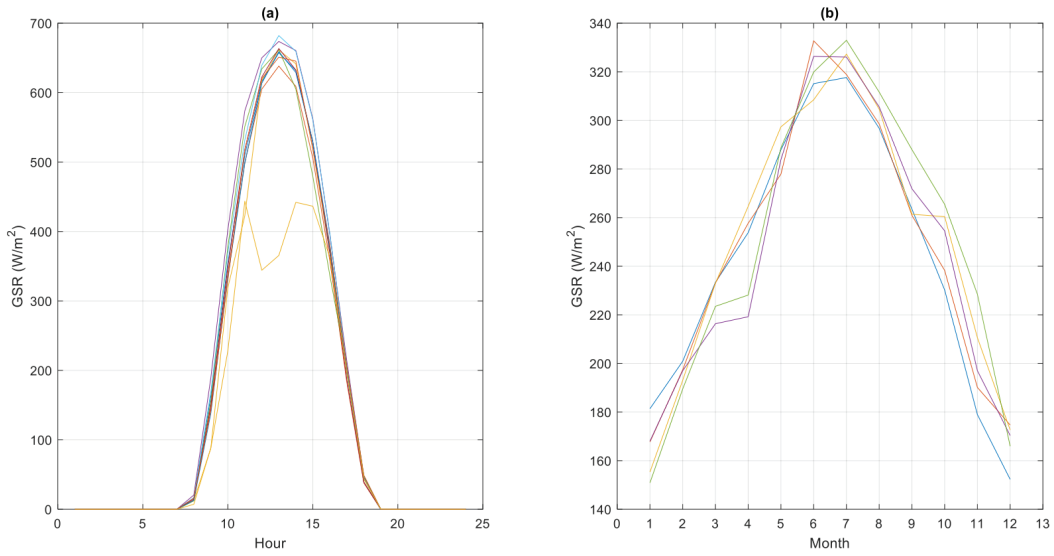

Figure 1. (a) Hourly and (b) Monthly GSR samples [6].

using Taguchi-based Grey Relational Analysis (GRA) and ANOVA. Raghavendrakumar et al. [21] investigated the analysis of central atmospheric temperature variability data from the SABER satellites.

A comprehensive perspective on machine-learning models for GSR was presented by reviewing 232 papers, exploring input parameters, feature selection methods, and model development techniques [22]. The authors reviewed machine-learning models for GSR forecasting and sources of input data. Additionally, they reviewed. Rabehi and Lalmi [23] compared the performance of different models (multi-layer perceptron (MLP), boosted decision tree, and hybrid models combining MLP with linear regression) for predicting solar radiation. The results showed that the MLP model performed the best with a normalized RMSE of 0.033 and an R^2 of 97.7%.

Ramedani et al. [24] developed four models (Support Vector Regression (SVR) with polynomial and radial basis function (RBF) kernels) to predict GSR. Results showed that SVR with RBF had better accuracy compared SVR with polynomial, ANFIS, and ANN models.

The present research compares the predictive performance of LSTM, BiLSTM, GRU, ConvNet, MLP, GAM, GPR, and LR for GSR prediction. Present study selected recurrent models (LSTM, BiLSTM, and GRU) based on their better capability of capturing long-term temporal dependencies, which is crucial for time dependedent predictions. For benchmarking purpose, the study includes the GPR and GAM regression models. For model performance evaluation, RMSE, MAE, MBE, and R^2 error matrics is used.

METHODOLOGY

This section provides the description of methods used in this paper. The methods include LSTM, BiLSTM, GRU, ConvNet, MLP, GPR, GAM, and LR. The performance of these methods is compared using the error measures, mentioned earlier.

LONG-SHORT-TERM MEMORY (LSTM) AND BIDIRECTIONAL LSTM (BILSTM)

The Long-Short-Term Memory (LSTM) neural networks [25] technique is employed in the study for GSR forecasting. This model is an advanced form of RNN and incorporates a recurrent network delay which enables the association of the network's output for a given t-th sample with prior inputs, as illustrated in Figure 2. The LSTM architecture addresses the issue of diminishing gradients through the implementation of specialized logic gates, specifically the input (i), forget (f), and output (o) gates [26]. Moreover, the model is equipped with weights that link the gates, which are the input weights U, consisting of U^f, U^i, U^o and U^c, the recurrent weights W, comprising of W^f, W^i, W^o and W^c, and the bias terms b, including b_f, b_i, b_o and b_c [26]. The input at the t-th sample, denoted as X_t along with the hidden state from the preceding sample h_{t-1}, are processed by the LSTM cell to produce the current hidden state h_t using the steps outlined below:

(1) The forget gate (f_t) determines the information to be discarded from the unit state, which is executed according to the subsequent equation:

$$f_t = \sigma\left(U^f X_t + W^f h_{t-1} + b_f\right) \tag{1}$$

where σ(z) is the sigmoid function given by:

$$\sigma(z) = \frac{1}{1+e^{-z}}. \tag{2}$$

(2) The input gate (i_t) identifies the information that should be retained in the unit and selects the values using the following equation:

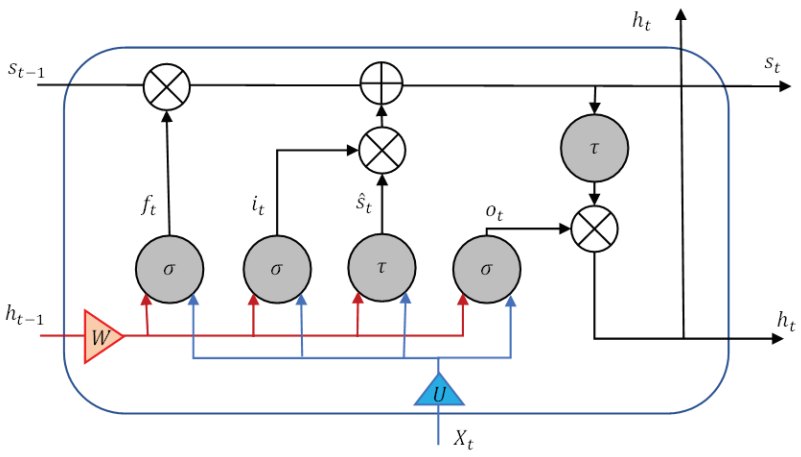

Figure 2. LSTM unit structure.

$$i_n = \sigma\left(U^i X_t + W^i h_{t-1} + b_i\right) \quad (3)$$

(3) The unit state is updated with the candidate value $\hat{s}_t$ based on the following equation 4:

$$\hat{s}_t = \tau(U^c X_t + W^s h_t + b_s) \quad (4)$$

where $\tau(z)$ is a tangent-hyperbolic function given by:

$$\tau(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (5)$$

The candidate unit state is further used to determine the unit state values as follows:

$$s_t = s_{t-1} \otimes f_t \oplus i_t \otimes \hat{s}_t \quad (6)$$

where are the operators $\oplus$ and $\otimes$ represent element-wise summation and product, respectively.

(4) Next, the output o_t and the hidden state h_t is determined using the following calculation:

$$o_n = \sigma(U^o X_t + W^o h_{t-1} + b_o). \quad (7)$$

while the hidden state of the unit is given by:

$$h_t = o_t \otimes \tau(s_t). \quad (8)$$

The Bidirectional LSTM (BiLSTM) (Fig. 3) is an extension of the LSTM unit where the unit has two-way access. BiLSTM utilizes two LSTM models instead of just one, training in opposite directions on the input sequence. This means that while the first LSTM is designed to understand the sequence in a forward manner, the second LSTM is tasked with learning the sequence in reverse order. In the present approach, the LSTM is trained using LM method [27] for fast convergence. The LM algorithm is deemed particularly appropriate for this application, which necessitates a moderate quantity of units and layers. The update of weights (Δu) within the LM method is characterized by the following:

$$\Delta u = -[J^T J + \lambda \boldsymbol{I}]^{-1} J^T (y - \hat{y}) \quad (9)$$

where J represents the Jacobian matrix that maps the partial derivatives of the error function relative to the weight vector of the model [28]. The error function is evaluated by the discrepancy between the actual values (y) and the predicted outputs ($\hat{y}$). The scalar (λ) controls the magnitude of the step size, which is augmented if the updates lead to a reduction in the error function and is reduced otherwise. The methodology employs a maximum number of iterations set at $T = 200$.

GATED RECURRENT UNIT (GRU)

The GRU is a type of RNN architecture that is designed to capture long-term dependencies in sequential data while mitigating the vanishing gradient problem [29] Mathematically, the GRU (Figure. 4) can be described as follows:

(1) Update Gate (z_t): The update gate determines how much of the previous hidden state (h_{t-1}) should be retained and combined with the new candidate hidden state ($\tilde{h}_t$) at the current time step (t). It takes input features x_t and h_{t-1} and produces a value between 0 and 1 to represent the proportion of information to keep.

$$z_t = \sigma(W_z \times x_t + U_z \times h_{t-1} + b_z) \quad (10)$$

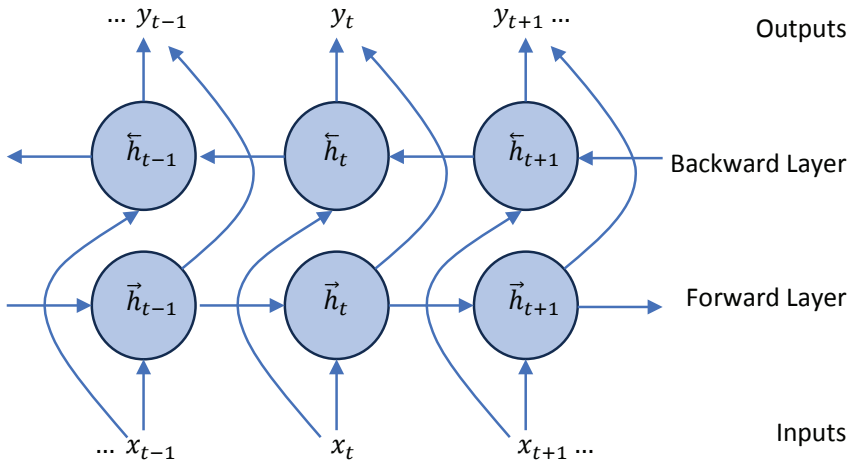

Figure 3. Bidirectional Unit.

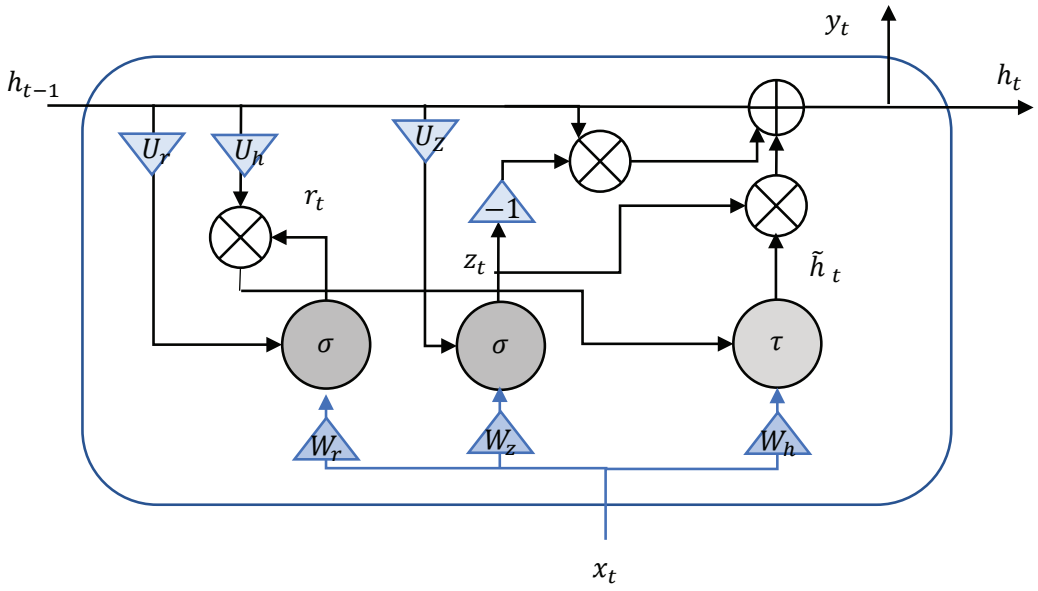

Figure 4. GRU architecture.

(2) Reset Gate (r_t) determines how much of the previous hidden state (h_{t-1}) should be ignored while calculating the new state ($\tilde{h}_t$). It produces values between 0 and 1.

$$r_t = \sigma(W_r \times x_t + U_r \times h_{t-1} + b_r) \quad (11)$$

(3) Candidate Hidden State ($\tilde{h}_t$) is the new hidden state at time step t, which combining the current input features (x_t) and previous hidden state (h_{t-1}) based r_t.

$$\tilde{h}_t = \tau(W_h \times x_t + U_h \times (r_t \times h_{t-1}) + b_h) \quad (12)$$

(4) Final Hidden State (h_t) at time step t is a combination of h_{t-1} and $\tilde{h}_t$ weighted by z_t. It determines how much information from the previous hidden the candidate hidden state should be incorporated.

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (13)$$

To use the GRU to predict GSR at time step x_{t+1}, given the previous GSR values x_1, x_2,..., x_t, the following steps are taken:

- Initialize state h_0 to a vector of zeros.
- For each time step t from 1 to T, compute the update gate (z_t), reset gate (r_t), candidate hidden state $\tilde{h}_t$, and the final hidden state (h_t) using the formulas mentioned above.
- At time step $t + 1$, use the final hidden state h_t as the context or representation of the sequence x_1, x_2,..., x_t. Then, compute the predicted GSR value x_{t+1} using this context h_t as input to a fully connected layer or regression layer.

$$\tilde{h}_t = \tau(W_h \times x_t + U_h \times (r_t \times h_{t-1}) + b_h) \quad (14)$$

BPTT is used to train the model along with weights and biases to minimize the error. These steps arerepeated over multiple epochs till the model achieves satisfactory performance.

CONVOLUTIONAL NEURAL NETWORKS (ConvNet)

The ConvNet [30], originally designed for image classification, has shown exceptional results surpassing other leading methods in that domain. Given the 2D ConvNet's impressive outcomes in image-centric tasks, a 1D variant is created to analyse 1D signals and sequences [30]. This is achieved by isolating significant features at both high and low levels within the input data sequences. The ConvNet structure consists of a sequence input layer, a convolutional layer, a fully connected layer, and a regression layer. In this context, the width of the time series, denoting the feature count, is represented by K, and its length by N [31]. The initial phase involves a convolution operation that progresses along a set direction from the start to the end of the time series, utilizing convolutional filters that match the series width but may vary in length. The primary function of the convolutional layer is to extract features from the time series, followed by the subsequent pooling stage where vectors from the convolutional layer are aggregated into a new vector in what is known as the pooling layer.

This model was enhanced by incorporating an adaptive moment estimation (ADAM) optimizer [32]. The update of each weight w_i is obtained as follows:

$$w_i = w_{i-1} - \alpha \frac{\sqrt{1-\beta_2^i}}{1-\beta_1^i} m_i \oslash \sqrt{v_i \oplus \epsilon}. \quad (15)$$

where α and β are the learning and decay rates. The operation of element-wise vector division is indicated by the symbol ⊘, and the calculation of the first moment m is executed in the subsequent manner.

$$m_i = \beta_1 m_{i-1} + (1-\beta_1)\nabla_w e(w_{i-1}). \quad (16)$$

where $\nabla_w e(w)$ symbolizes the vector differential operation of the error function e with respect to the weight w. The computation of the second moment vector v is carried out in the following way.

$$v_i = \beta_2 v_{i-1} + (1-\beta_2)\nabla_w e(w_{i-1}) \otimes \nabla_w e(w_{i-1}). \quad (17)$$

The learning rate α can assume any positive real number value, whereas the decay rates β_1 and β_2 are real numbers that are greater than 0 but less than 1.

MULTI-LAYER PERCEPTRON (MLP)

The MLP [33] is utilized as a standard for comparing the performance of LSTM. MLP possesses a notably simpler architecture compared to LSTM. Unlike LSTM, MLP employs a feed-forward structure that processes data in a sequential manner from the input layer to the output layer. The output of the k-th hidden unit, denoted as $h_k(t)$, for the t-th input vector, is mathematically expressed as follows.

$$h_k(t) = \tau(V_k x(t) + b_k) \quad (18)$$

where V_k represents the input weight matrix. In this paper, the number of hidden units used is $K = 20$. With these values for the hidden units established, the output of the MLP is determined by the following equation.

$$y(t) = Uh(t) + b \quad (19)$$

This study employs gradient descent with adaptive learning rate (lr) backpropagation for training the MLP [27]. This algorithm is applied to compute the derivatives of the error function, denoted as dE, relative to the weight and bias variables w. The performance of this approach is expressed as follows.

$$E = (y - \hat{y})^2 \quad (20)$$

Each weight is updated based on the following gradient descent equation.

$$w_i = w_{i-1} + lr \frac{\partial E}{\partial w} \quad (21)$$

During each iteration, if the output diverges from the target, the learning rate is augmented by a factor of Δlr = 1.05. Conversely, if the error escalates beyond the threshold of ΔE_{max} = 1.04, the adjustment that caused the increase in error is reversed, and the learning rate is reduced by a factor of Δlr = 0.7. The training process concludes once the maximum iteration count, set at T = 1200, is attained.

GAUSSIAN PROCESS REGRESSION (GPR)

Gaussian Process Regression (GPR) [34] is a non-parametric Bayesian regression method that builds the relationship between the target variable and the independent variables as a Gaussian process. GPR is useful with small datasets. Mathematically, GPR can be expressed as follows:

$$y_{t+1} \sim N(m_{t+1}, K_{t+1}) \quad (22)$$

where m_{t+1} is the mean function at time step $t + 1$ given by the historical GSR values $x_1, x_2, ..., x_t$. K_{t+1} is the covariance matrix at time step $t + 1$. To use GPR for predicting GSR at time step $t + 1$ (x_{t+1}) given the previous GSR values x_1, x_2,..., x_t, you need to estimate the mean function m_{t+1} and the covariance matrix K_{t+1}. This involves finding the hyperparameters of the mean and covariance functions that best fits the training data. Once the mean function m_{t+1} and the covariance matrix K_{t+1} are estimated, one can make predictions for the GSR value at time step $t + 1$ using the conditional distribution of y_{t+1} given the observed historical GSR values $x_1, x_2, ..., x_t$. The predicted GSR value at time step $t + 1$ is the mean of this conditional distribution, which is given by m_{t+1}.

GENERALIZED ADDITIVE MODEL (GAM)

Generalized Additive Model (GAM) [35] is a type of generalized linear model that extends linear regression to accommodate non-linear relationships between predictors and the response variable. Mathematically, a GAM can be represented as follows. For a regression task, let's consider a single response variable y and p predictor variables (X_1, X_2,…, X_p). The general form of a GAM can be written as.

$$y = \beta_0 + f_1(X_1) + f_2(X_2) + \cdots + f_p(X_p) + \epsilon \quad (23)$$

Where y is the response variable, β_0 is the intercept term, ($f_1(X_1)$, $f_2(X_2)$,…, $f_p(Xp)$) are smooth functions of the predictor variables. These functions are represented by splines, polynomials, or others. ϵ is the error term and follows a normal distribution for regression tasks. The GAM fitting process involves estimating the parameters β_0 and the smooth functions $f_i(X_i)$ from the training data. This is estimated by minimizing a loss function that penalizes the complexity of the model to prevent overfitting. The penalization term discourages overly complex functions, encouraging smoothness in the fitted functions.

LINEAR REGRESSION (LR)

Linear regression [36] is a simple and widely used statistical method for modelling the relationship between a dependent variable and one or more independent variables. In the context of predicting global solar radiation (GSR) at time step$t + 1$ (x_{t+1}) given previous GSR values x_1, x_2,..., x_t, linear regression can be used to establish a linear relationship between the historical GSR values and the target GSR value at time $t + 1$.

$$y = \beta_0 + f_1(X_1) + f_2(X_2) + \cdots + f_p(X_p) + \epsilon \quad (24)$$

Here, β_0, β_1,..., β_t are the coefficients of the model that represent the intercept and the slopes associated with each of the historical GSR values. ε represents the error term, which captures the discrepancy between the predicted GSR value and the actual GSR value at time t+1. The resulting y_{t+1} represents the predicted GSR value at time t+1. The coefficient vector β is calculated using the following formula.

$$\beta = (X^T X)^{-1} X^T y \quad (25)$$

where X^T is the transpose of the design matrix X. The inverse of the matrix $(X^T X)^{-1}$ is generated from the multiplication of X^T and X. The matrix $X^T y$ is resulting from the multiplication of X^T and the target vector y.

DATA AND ERROR MEASURES

The measured meteorological data was obtained from the station located at 26°18'26"N latitude and 50°08'29"E longitude at King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia. The meteorological parameters (global solar radiation intensity-GSR, ambient temperature, relative humidity, wind speed, wind direction, and surface pressure) are scanned every five seconds and saved as averages over one hour. The data, used in the present study, spans over a length of nine years from 01/01/2012 to 12/31/2020. The experiment in this paper relies on a single dataset from one location, which may limit the generalizability of our findings. However, the dataset covers a substantial period of 8 years, offering robust temporal coverage that effectively captures various patterns and trends in solar radiation.

A sample of data set of GSR values (W/m^2), between 01/01/2015 and 31/12/2017 is shown in Figure 5. It is evident from the sample data set that the GSR intensity values vary from a minimum at the start of the year (winter time) and after reaching a maximum during summer time touches back the minima at the end of the year. This process is repeated from year to year but with varying magnitudes. Similarly, the daily variation of GSR intensity values, shown in Figure 6, start increasing from the early hours of the day and then keep on increasing with the passage of time and after reaching a maximum at around noon time tend to touch the minima back again at sun set. This process, as indicated in Figure 6, keeps on repeating day after day but with different values of GSR intensities.

Since the dataset is a time series, so is not shuffled when loaded. The data is divided into training, validation, and test portions of 60%, 20%, and 20%, respectively. The 24 hours lag GSR-values are considered input to give a desired input structure into a model, and the model predicts the next hour. Error measures, such as RMSE, MAE, and R^2, are commonly used to assess the accuracy and performance of predictive models, including those used for GSR prediction.

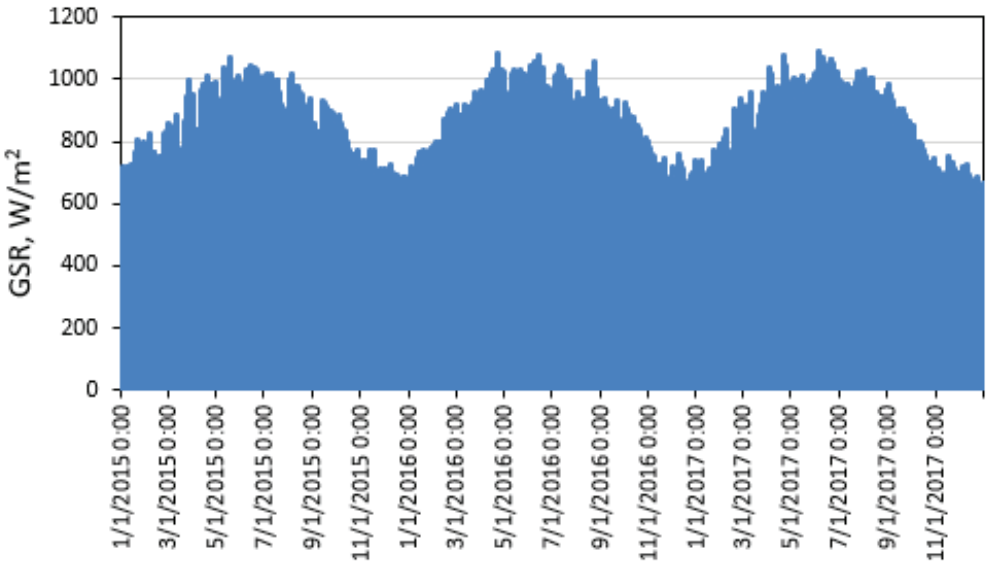

Figure 5. GSR intensity values over a period of three years (01/01/2015 to 31/12/2017).

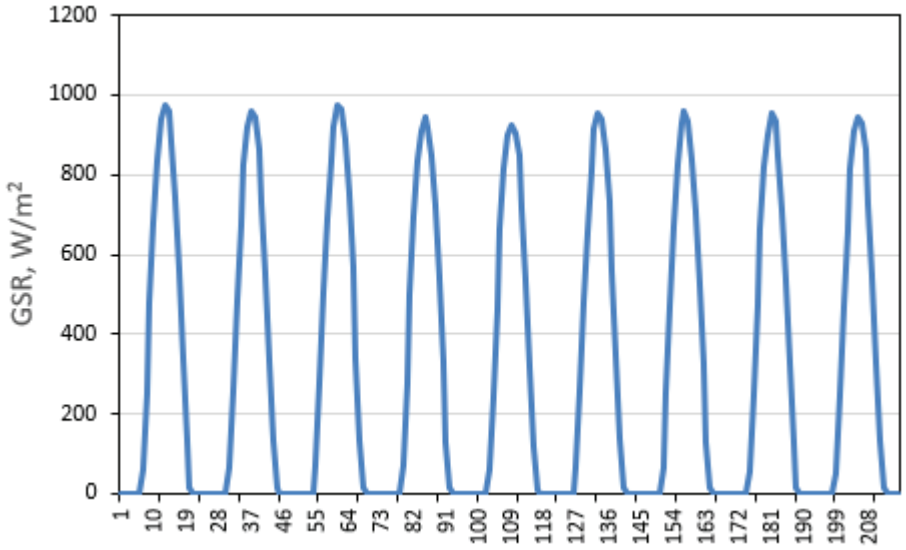

Figure 6. GSR intensity values over a period of nine days (22/06/2016 to 30/06/2016).

RMSE [37] is calculated by taking the square root of the average of the squared differences between the predicted GSR values (y_p) and the actual GSR values (y).

$$RMSE = \sqrt{\frac{\Sigma_{n=1}^{N}(y_p(n) - y(n))^2}{N}} \quad (26)$$

MAE [38] is another error metric used to measure the average absolute difference between the predicted GSR values and the actual GSR values.

$$MAE = \frac{\Sigma_{n=1}^{N}|y_p(n) - y(n)|}{N} \quad (27)$$

MBE is used to measure the average difference between the predicted GSR values and the actual GSR values. It provides the model's tendency to under and overestimation.

$$MBE = \frac{\Sigma_{n=1}^{N}(y_p(n) - y(n))}{N} \quad (28)$$

R^2 [39–41] measures the proportion of variance in the GSR values that can be explained by the predictive model. The R^2 values range between 0 and 1.

$$R^2 = 1 - \frac{\Sigma_{n=1}^{N}(y_p(n) - y_t(n))^2}{\Sigma_{n=1}^{N}(y_t(n) - y_m)^2} \quad (29)$$

RESULTS AND DISCUSSION

The GSR values predicted using different proposed models are compared with the measured values and the resulting error parameters are summarized in Table 1. The error values obtained for different methods show that BiLSTM performs the best in terms of RMSE, indicating that the predicted values have the smallest overall deviation from the measured

Table 1. Models Accuracy

No	Methods	RMSE	MAE	MBE	R^2 (%)
1	BiLSTM [31]	**30.84**	17.21	11.65	**99.26**
2	LSTM [25]	45.87	23.97	14.36	98.17
3	GRU [29]	45.039	21.99	-9.16	98.25
4	ConvNet [30]	57.75	36.86	21.68	97.23
5	MLP [27]	40,60	15.76	1.77	98.40
6	GPR [34]	43.77	**14.69**	**-0.15**	98.14
7	GAM [35]	41.09	17.98	2.01	98.36
8	LR [36]	44.83	19.75	3.33	98.05

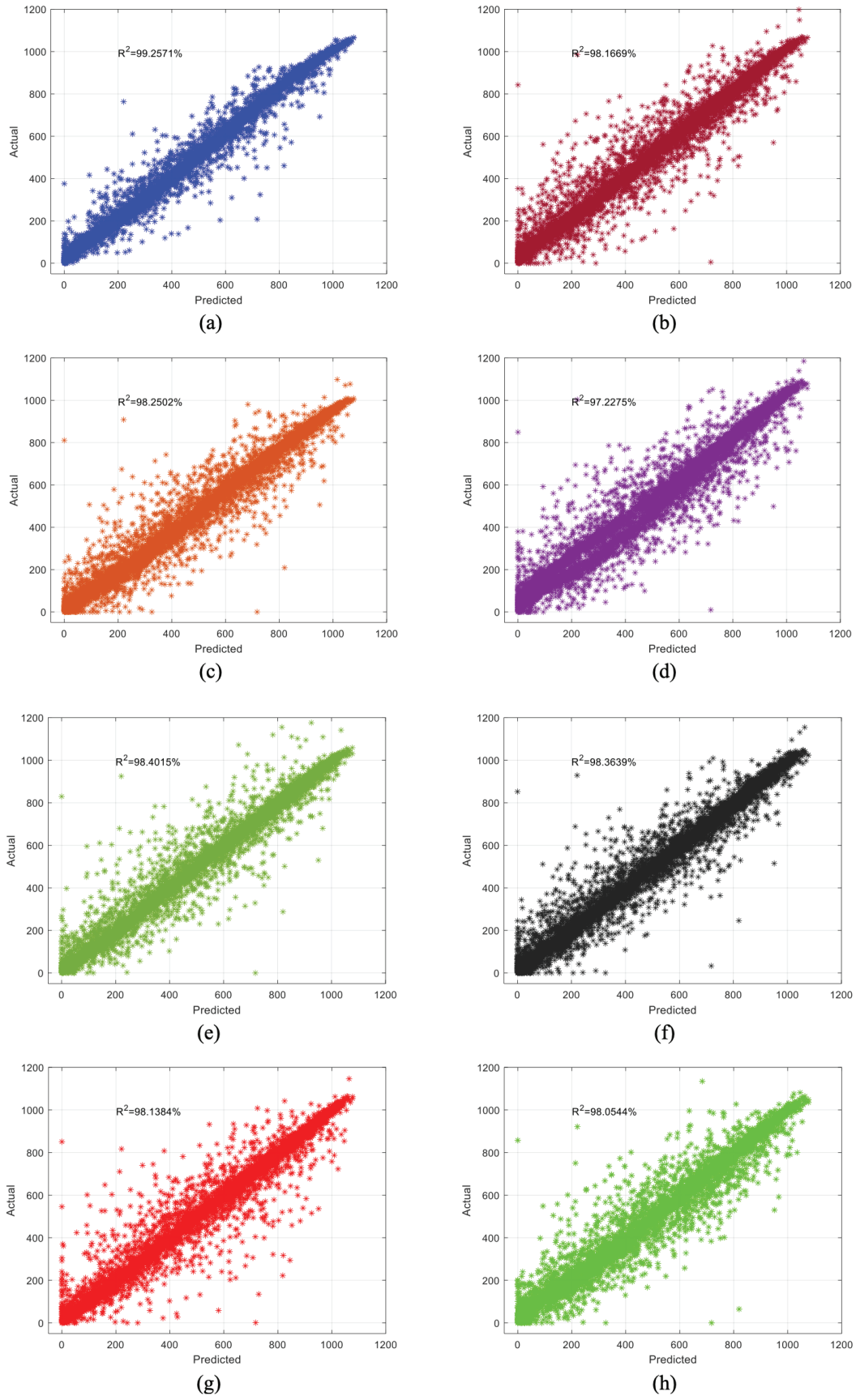

Figure 7. Scatter Plot of predicted and actual data using (a) BiLSTM, (b) LSTM, (c) GRU, (d) ConvNet, (e) MLP, (f) GPR, (g) GAM, and (h) LR.

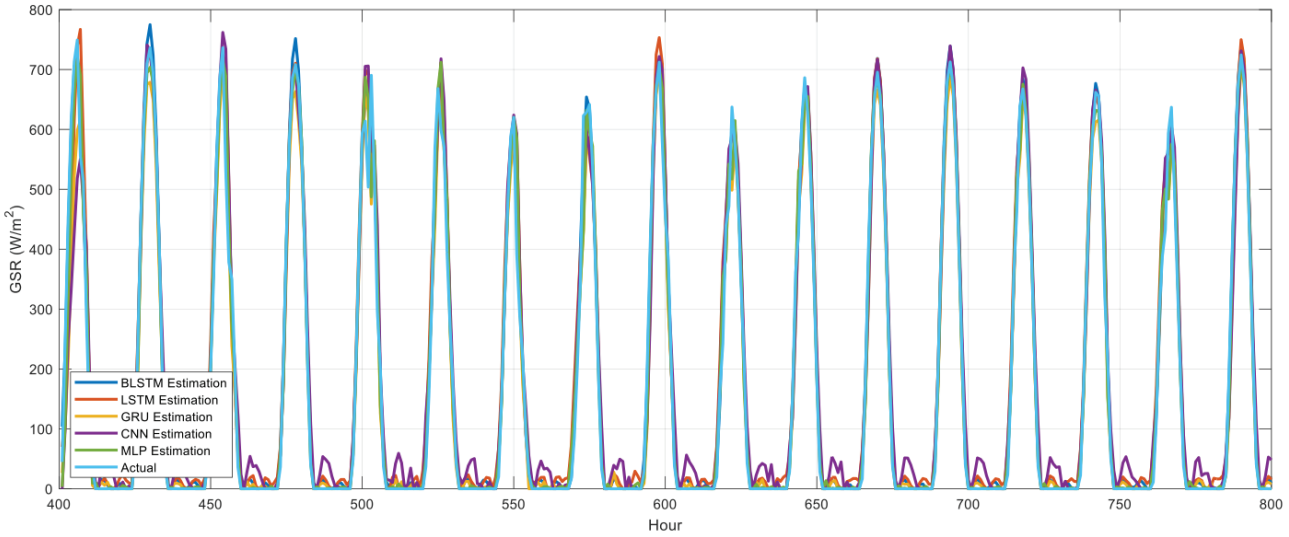

Figure 8. Comparison of predicted and actual data using BiLSTM, LSTM, GRU, ConvNet, and MLP.

values. Additionally, it has the highest R^2, suggesting that BiLSTM method resulted in the closest agreement between the predicted and measures GSR values. LSTM is outperformed by BiLSTM, with larger values of RMSE and MAE relative to other models. However, R^2 of 98.17% is slightly lower than BiLSTM. GRU model performed poorly with high RMSE and negative MBE values. However, a R^2 value of 98.25% comparable to LSTM, represents a good fit. ConvNet methos, slightly under performed than the BiLSTM, LSTM, and GRU models in terms of RMSE and MAE. The associated R^2 value of 97.23% is observed to be the lowest, showing a weaker correlation, Table 1.

Traditional machine learning approach, like MLP, also demonstrated a relatively good performance with the lowest magnitude of MAE. The coefficient of determination value is found to be moderate with R^2 of 98.40%, suggesting a good fit to the data. GPR performed reasonably well with low MAE and RMSE values. The lower MBE and high R^2 value of 98.14% show a minor bias and a good fit to the data, respectively. GAM showed a good performance with low values of RMSE and MAE. The MBE and R^2 values of 2.01 and 98.36% show a positive bias and a good fit. The model LR performed slightly below expectations with higher RMSE and MAE values compared to some other techniques.

The plots in Figure 7, show that BiLSTM (Figure 7a) is the best- model, closely followed by MLP (Figure 7e) and GAM (Figure 7f). Overall, these models show better perfromance to predict GSR values with better accuracy. On the other hand, LSTM and GRU also performed reasonably well, but with varying degrees of biases and accuracies. Figure 8 shows the comparison of the predicted and actual GSR values obtained using BiLSTM, LSTM, GRU, ConvNet, and MLP methods.

CONCLUSIONS

The present sudy carried out state of the art analysis of predicting the Global Solar Radiation using different modern machine learning methods. Accurate GSR prediction ahead of time is crucial for better and commercial implementation of solar energy applications. These applications may include power production, solar cooling and heating, water desalination, green hydrogen production and so on. The study explored deep learning architectures (BiLSTM, LSTM, and GRU) and traditional machine learning methods (ConvNet, MLP, GPR, GAM, and LR). Among these, BiLSTM is found to be the best performing model, exhibiting the minimum RMSE and maximum R^2 values. It demonstrated the ability of retaininglong-term dependencies of solar radiation patterns for accurate prediction of GSR ahead of time. MLP and GAM models followed interms of robust performance and capturing temporal patterns of GSR. The recurrent models (BiLSTM, LSTM, and GRU) showed effectiveness in predicting the time-dependent patterns of GSR. Other deep network model, like ConvNet, also provided comparable performance. The identification of BiLSTM as the most accurate model, the present study provides a reliable method for predicting GSR high accuracy. This, for sure, lead for more efficient resources allocation and utilization, energy management, and improved decision-making in solar energy applications.

ACKNOWLEDGEMENTS

This research is supported by King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia and PPM Telkom University, Indonesia.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

[1] Snapshot 2021 - IEA-PVPS. Available from: https://iea-pvps.org/snapshot-reports/snapshot-2021/ Accessed on Dec 1, 2023.

[2] Mutlu B, Özyörük B. A research on mathematical model approaches in biomass supply chain. Sigma J Eng Nat Sci 2024;42:945–955. [CrossRef]

[3] İbrahim D, Sari M, Gülen S, Balacescu A. An investigation on the estimation of the impact factors of pandemic deaths with artificial neural network and multiple regression algorithms: Covid-19 case. Sigma J Eng Nat Sci 2024;42:667–678. [CrossRef]

[4] Koçak E, Özsoy VS, Örkcü HH. Modeling of wind speed using differential evolution: Istanbul case. Sigma J Eng Nat Sci 2024;42:642–652. [CrossRef]

[5] Kartal F, Kaptan A. Artificial neural network and multiple regression analysis for predicting abrasive water jet cutting of Al 7068 aerospace alloy. Sigma J Eng Nat Sci 2024;42:516–528. [CrossRef]

[6] Olawoyin-Yussuf D, Mohandes M, Liu B, Rehman S. Solar radiation forecasting using attention-based temporal convolutional network. Mater Res Proceed 2024;43:88–95. [CrossRef]

[7] Sun H, Zhao N, Zeng X, Yan D. Study of solar radiation prediction and modeling of relationships between solar radiation and meteorological variables. Energy Convers Manag 2015;105:880–890. [CrossRef]

[8] Alsina E, Bortolini M, Gamberi M, Regattieri A. Artificial neural network optimisation for monthly average daily global solar radiation prediction. Energy Convers Manag 2016;120:320–329. [CrossRef]

[9] Sur A, Sah RP, Pandya S. Milk storage system for remote areas using solar thermal energy and adsorption cooling. Mater Today Proc 2020;28:1764–1770. [CrossRef]

[10] Pandya S, Sur A, Solke N. Pollution weather prediction system: smart outdoor pollution monitoring and prediction for healthy breathing and living. Sensors 2020;20:5448. [CrossRef]

[11] Pandya S, Sur A, Solke N. COVIDSAVIOR: a novel sensor-fusion and deep learning based framework for virus outbreaks. Front Public Health 2021;9:797808. [CrossRef]

[12] Pamdya S, Mistry M, Kotecha K, Sur A, Parikh P, Shah K, et al. A novel multicast secure MQTT messaging protocol framework for IoT-related issues. In: Proc 2nd Int Conf Comput Commun Cyber-Security. 2021. p. 339–359. [CrossRef]

[13] Muthuraja S. Effect of charging and discharging process of PCM with paraffin and Al2O3 additive subjected to three point temperature locations. J Ecol Eng 2022;23:32–42. [CrossRef]

[14] Kumar KS, Kalos PS, Akhtar MN, Shaik S, Sundara V, Fayaz H, et al. Experimental and theoretical analysis of exhaust manifold by uncoated and coated ceramics (Al2O3, TiO2 and ZrO2). Case Stud Therm Eng 2023;50:103465. [CrossRef]

[15] Baskaran S, Kumar KS. Drying investigation of coriander seeds in a photovoltaic thermal collector with solar dryer. Therm Sci 2023;27:659–668. [CrossRef]

[16] Kumar KS, Muniamuthu S. Design and fabrication of nickel-chromium reinforced 2-stage energy efficient pyrolysis reactor for waste plastics applications. Int J Ambient Energy 2024;45:2288148. [CrossRef]

[17] Naushad RP, SRS K, Edayadulla. The effect of emissions from DI-based sources causing hazardous health effects. In: Recent advances in mechanical engineering. 2024. p. 1–13. [CrossRef]

[18] Kumar KS, Surakasi R, Mohan A, Tharanisrisakthi BT, Muniamuthu S, et al. Emerging trends and global challenges to predict drop in thermal performance of WTG gearbox. J Therm Eng 2024;10:657–669. [CrossRef]

[19] Kumar KS, Muniamuthu S, DR P, Mohan KR. Measurement of temperature flow analysis by condition monitoring system for WTG gear box to evaluate the thermal performance associated with plant load factor. J Therm Eng 2022;9:970–978. [CrossRef]

[20] Dhabale RB, Kurkute VK, Unde SS, Jadhav P, Dhabale VR, Sur A. Multi objective optimizations of friction stir welding process parameters for reinforcement of copper particles in aluminum alloy using Taguchi based grey relational analysis (GRA) and ANOVA. J Phys Conf Ser 2023;2604:012001. [CrossRef]

[21] Raghavendrakumar E, Kamalakar V, Kumar KS. An investigation in temperature data analysis of middle atmospheric variation from SABER satellite. Nat Environ Pollut Technol 2024;23:443-450. [CrossRef]
[22] Zhou Y, Liu Y, Wang D, Liu X, Wang Y. A review on global solar radiation prediction with machine learning models in a comprehensive perspective. Energy Convers Manag 2021;235:113960. [CrossRef]
[23] Rabehi MGA, Lalmi D. Hybrid models for global solar radiation prediction: a case study. Int J Ambient Energy 2020;41:31–40. [CrossRef]
[24] Ramedani Z, Omid M, Keyhani A, Shamshirband S, Khoshnevisan B. Potential of radial basis function based support vector regression for global solar radiation prediction. Renew Sustain Energy Rev 2014;39:1005–1011. [CrossRef]
[25] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735–1780. [CrossRef]
[26] Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 2019;323:203–213. [CrossRef]
[27] Islam MS, Mohandes M, Rehman S. Vertical extrapolation of wind speed using artificial neural network hybrid system. Neural Comput Appl 2017;28:2351–2361. [CrossRef]
[28] Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 1963;11:431–441. [CrossRef]
[29] Chung J, Gulcehre C, Cho K, Bengio Y. Gated feedback recurrent neural networks. In: Proc 32nd Int Conf Mach Learn. 2015. p. 2067–2075.
[30] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86:2278–2324. [CrossRef]
[31] Lawal A, Alhems L. Wind speed prediction using hybrid 1D CNN and BLSTM network. IEEE Access 2021;PP:1. [CrossRef]
[32] Al-Shaikhi A, Nuha H, Lawal A, Mohandes M. Vertical wind profile estimation using hybrid convolutional neural networks and bidirectional long short-term memory. Arab J Sci Eng 2023;13:1213. [CrossRef]
[33] MathWorks. traingd - gradient descent backpropagation. Available at: https://www.mathworks.com/help/deeplearning/ref/traingd.html Accessed on Oct 18, 2024.
[34] Schulz E, Speekenbrink M, Krause A. A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions. J Math Psychol 2018;85:1–16. [CrossRef]
[35] Nuha H, Mohandes M, Rehman S, Al-Shaikhi A. Vertical wind speed extrapolation using regularized extreme learning machine. FME Trans 2022;50:412–421. [CrossRef]
[36] Huang DMSJQ, Wang D. A simple method for partitioning total solar radiation into diffuse/direct components in the United States. Int J Green Energy 2018;15:497–506. [CrossRef]
[37] Ahamed S, Tanino K, Guo H. Cloud cover-based models for estimation of global solar radiation: a review and case study. Int J Green Energy 2021;19. [CrossRef]
[38] Rehman S, Nuha HH, Al-Shaikhi A, Akbar S, Mohandes M. Improving performance of recurrent neural networks using simulated annealing for vertical wind speed estimation. Energy Eng 2023;120:775–789. [CrossRef]
[39] Mohandes M, Nuha H, Islam MS, Schulze FH. Accuracy of wind speed predictability with heights using recurrent neural networks. FME Trans 2021;49:908–918. [CrossRef]
[40] Mohandes M, Rehman S, Nuha H, Schulze FH. Predictability of wind speed with heights using recurrent neural networks. In: Proc 4th Int Symp Adv Electr Commun Technol 2021. p. 1–4. [CrossRef]
[41] Nuha H, Pahlevi R, Mohandes M, Rehman S, Al-Shaikhi A, Tella H. Vertical wind speed estimation using generalized additive model (GAM) for regression. In: Proc CICN. 2022. p. 691–695. [CrossRef]