



# Advances in Solar Irradiance Prediction: A Comprehensive Analysis of the Novel Robust Self- Attention Multi-Horizon Model (RSAM)

Ms. Kritika Balihar<sup>1</sup>, Kunal Kumar Mishra<sup>2</sup>, Yash Pandey<sup>3</sup>

<sup>1</sup>Assistant Professor, Deptt. of Computer Science & Engineering Bhagwan Parshuram Institute of Technology Delhi, India

[kritikabalihar@bpitindia.com](mailto:kritikabalihar@bpitindia.com)

<sup>2</sup>B. Tech Scholar, Deptt. of Computer Science & Engineering Bhagwan Parshuram Institute of Technology Delhi, India

<sup>2</sup> [kunal197csea20@bpitindia.edu.in](mailto:kunal197csea20@bpitindia.edu.in), <sup>3</sup>[yash192csea20@bpitindia.edu.in](mailto:yash192csea20@bpitindia.edu.in)

## ABSTRACT—

In the pursuit of optimizing solar power systems and advancing renewable energy, accurate solar irradiance forecasting plays a pivotal role. Recent years have witnessed substantial progress in this field, driven by the application of machine learning and deep learning techniques. This research conducts a comprehensive comparative analysis of two leading-edge models: the Robust Self-Attention Multi-Horizon (RSAM) and the Long Short-Term Memory (LSTM), with a specific focus on very short-term solar irradiance prediction. The study underscores that RSAM, an innovative architecture, surpasses conventional LSTM models in accuracy and robustness. RSAM integrates multi-horizon forecasting, encompassing multiple weather parameters, and employs quantile regression for model assessment. The model's performance is validated using benchmark datasets from Kaggle, representing diverse solar sites. These findings emphasize the crucial role of precise forecasting models in optimizing renewable energy generation and grid management. Furthermore, RSAM's superior performance suggests potential applications across various sectors beyond the renewable energy domain, making it a valuable tool for industries dependent on solar irradiance forecasting. In an era marked by smart grids and an escalating reliance on renewable energy sources, this research represents a substantial leap toward the efficient and sustainable utilization of solar irradiance for the future of energy.

Keywords— *Solar Irradiance Forecasting, Comparative Analysis, Deep Learning, Kaggle Datasets, RNNs, TCNs, Single- tep Forecasting, Very Short-Term Forecasting, Power Output, Accuracy Assessment, Attention Models, Transformer Models, Prediction Intervals, Quantile Regression, Prediction Uncertainty.*

## 1. INTRODUCTION

The increased usage of renewable energy accelerates the development of increasingly accurate and robust solar irradiance forecasting models. Because renewable energy sources are more sustainable, several governments throughout the world have begun to replace fossil fuels. It is predicted that 5% of the African and 2% of the Asian populations have access to power via off-grid solar PV installations [1]. Solar power is predicted to contribute 20% of total US energy output by 2030, with photovoltaic utility capacity increasing by 127 GW [2]. India intends to roughly quadruple its total installed renewable energy capacity to 40% by 2030 [3]. While total installed PV capacity is expected to hit 8519 GW by 2050, delivering almost 25% of global power. Due to high population development, the world's energy supply will expand four times faster than current forecasts until 2100, and only fossil fuel sector infrastructure will be appropriate. We must transition to a cleaner, renewable, and less polluting energy source. The supply and demand for energy resources will grow in tandem with the advancement of human civilization. Sufficient clean energy sources are connected to global human progress, stability, wealth, and improved quality of life.

Solar irradiance forecasting is a novel word in today's society. It is a technique that uses knowledge of the sun's rays path, atmospheric conditions, scattering process, and the characteristics of a solar plant to utilize the sun's energy to create power and a device that captures and converts it into different forms of energy solar photovoltaic systems are simply referred to as "solar panels". The output is determined by the amount of incoming radiation as well as the features and qualities of the solar panel. Solar power output is booming these days. Forecasting information is required for the effective and better usage of a device in order to provide better output with the system. Various research firms actively participated in the activities to improve the quality of the power delivered to the grid and minimize the additional cost associated with weather dependency.

To reduce high market energy prices, the Smart Grid age has mandated the introduction of PV electricity in homes. Because solar PV power generation is derived from sun irradiance, solar irradiance forecasting is becoming more popular. It also aids in efficient demand-supply management, improved economic dispatch, peak shaving, energy arbitrage, energy market trading, and the reduction of uncertainty impact [5]. However, the random and intermittent character of solar irradiation makes reliable forecasting impossible.

Solar irradiance forecasting has a wide range of uses. Several approaches to resolving this difficult challenge have been offered. There are two types of forecasting methods proposed in the literature: point forecasting and probabilistic forecasting. Point forecasting methods only tell about the future predicted values of solar irradiance, whereas probabilistic forecasting provides both expected values and probability distributions at forecasting time points. Point forecasting models can be classified as statistical and machine learning models. Statistical models include auto-

regressive moving average (ARMA) [2], Lasso [3], and Markov models [4], [5]. Machine learning models [6] include support vector machines (SVM) [7]–[8], feed-forward neural networks (FFNN) [9]–[10], and recurrent neural networks (RNN) [11]. Earlier, statistical models were widely used for time-series forecasting, but with the advent of complexities in power systems and the rapid increase in data volumes, deep learning (DL) techniques are outperforming statistical models. These techniques work by learning the stochastic dependency between the past and future with less computational cost.

In recent years, RNNs based on LSTM (long short-term memory) [11] have become the de facto solution to deal with multivariable time-series data. LSTMs are effective in exploiting long-range dependencies and handling non-linear dynamics in time series forecasting. In [19], LSTMs have been used to forecast PV power output at multi-horizon (1 h, 1 d, 1 mo.) using PV power as the only input parameter on datasets of two cities in Egypt. The authors in [12] and [15] have also used the LSTM model inputs to predict one hour and one month ahead of solar irradiance, respectively. A hybrid convolutional LSTM (CNN-LSTM) model has been deployed to predict PV power in [16], which outperforms LSTM in the next 5-minute prediction.

Recent Vaswani [17] proposed a self-attention-based transformer. A deep learning-based model for sequence modelling has achieved good success. The plus point of using the self-attention approach in the RNN model is threefold: (i) Memory: since attention models dispense recurrence entirely, they have fewer memory requirements as compared to recurrent.

Models (LSTMs) can also make predictions using very long sequences of past data, whereas LSTMs rely solely on short-term memory to make predictions. (ii) simpler than recurrent models; (iii) Optimization: attention models are mathematical models. Computation: attention models are fully parallelizable, hence accelerating learning during training and taking less computation resources as compared to recurrent models, which are inherently sequential. The deep learning training model is a transformer model that has been successfully implemented in many areas, such as translation, music, and image generation. [18]–[20]. Some transformer-based, deep learning-based approaches for time series forecasting have recently been proposed [21]–[23]. However, the majority of work based on the transformer model is limited to traffic, retail, and electricity datasets. To take into account the uncertainty associated with the forecast, a prediction interval around the forecasted value would also provide more information, which would assist dispatchers with better decision making. Recently, a few probabilistic forecasting experiments have been conducted. A combination probabilistic forecasting method based on the improvised version of the Markov chain model has been proposed for probabilistic PV power forecasting. However, this method has a lot of complexity in the model and may not be suitable for large sizes. The probabilistic forecasting model based on the joint probability distribution function (PDF) of irradiance has been proposed in ref. [24]. The proposed model is predicted by numerical weather prediction (NWP), and its applications in the electric power trading market have been extensively studied. But the performance of NWP for solar radiation forecasting is highly variable when applied to different locations and not suitable for short-term forecasting [25]. A quantile regression-based probabilistic model for spatiotemporal PV forecasting has been presented in [26] for very short-term horizons (0–6 h). It uses LASSO for probabilistic forecasting in a real-world test case with a high number of PV installations. Therefore, the intermittency of solar irradiance and the complexity of mapping it with weather metrics make it a typical multivariate time series forecasting problem. The outstanding performance of the Transformer model and its capacity to comprehend complex links between weather metrics and solar irradiance intrigue us enough to use the model for multi-step solar forecasting. Also, quantile regression-based prediction interval modelling is used to calculate the corresponding lower and upper bounds for each forecasted global horizontal index (GHI) value (point forecast).

The main innovations and contributions of the paper include:

- i. Leverage the powerful learning abilities of the self-attention-based transformer deep learning model for forecasting solar irradiance.
- ii. To combat the forecast risks, the prediction intervals (50%, 90%, and 95%) have been calculated for each forecast value using quantile regression.
- iii. Multi-horizon forecasting on all industry-requested time horizons: intra-hour, hour-ahead, and day-ahead, as stated by Kostylev and Pavlovski [27].
- iv. The proposed algorithm applies to all data sets available on the site of Kaggle.com, and its performance is rigorously evaluated through various prediction parameters.

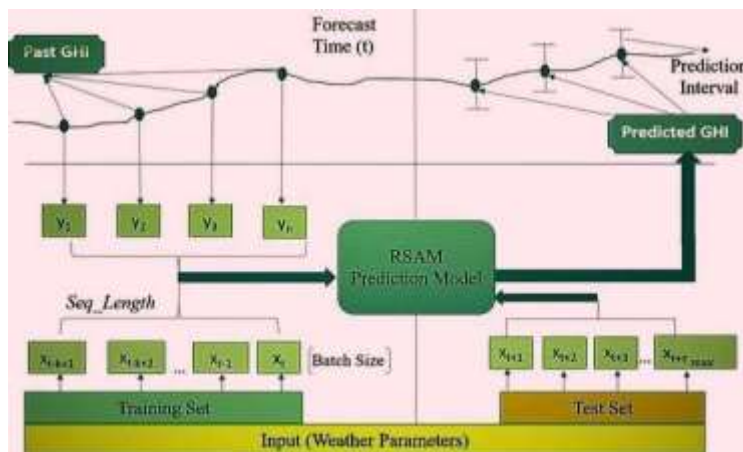


Fig 1. Illustration of multi-horizon forecasting with Point forecasting of GHI and their corresponding prediction intervals.[28]

The remainder of the study is arranged as follows: Section II covers multivariable solar time-series forecasting challenges. The suggested RSAM deep learning model is presented in Section III. This section discusses the training and testing phases of multi-horizon forecasting and interval prediction. This section goes into the specifics of the experiment. Section IV goes on the specifics of the experiment. Section V: Experiential Discussion, Methodology Finally, Section VI brings the paper to a close.

## 2. Related Work

The topic of forecasting solar irradiance is tackled as a multi-variable time series forecasting problem, wherein the Global Horizontal Irradiance (GHI) time series is dependent on a variety of meteorological characteristics, including temperature, humidity, and the clear day sky index. The work presents a structured methodology for improved forecasts, with a primary focus on robust multi-horizon solar irradiance forecasting.

The sequence of time series (S) that makes up the dataset consists of distinct entities that are correlated with features (X $\alpha$ ,t) and GHI output (GHI $\alpha$ ,t) at every time step (t). The input characteristics are classified as time-based vectors (v $\alpha$ ,t), which are presumed to be preset (e.g., month, day-of-the-week at time t), and observed inputs (o $\alpha$ ,t), which represent measured weather parameters.

The objective of multi-horizon forecasting is to forecast the upcoming  $\tau$  timesteps of HI with an initial forecast time (t) of  $\setminus \in \{1, 2, \dots, \max\}$ . A conditional probability distribution,  $p(\text{GHI}_{t+1:t+\tau} | \text{GHI}_{1:t}, X_{1:t}, \phi)$ , is used to describe this prediction. The learnable parameters are

indicated by  $\phi$ . Prediction intervals are computed to show the most likely best and worst-case values of GHI, ensuring the model's resilience. For multi-horizon forecasting, quantile regression using percentiles (e.g., 50th, 90th, and 95th) is employed.

As a function (f) that takes into account historical data within a finite look-back window (w), employing GHI values and weather parameters (o) up until the forecast start time (t) and known temporal inputs (v) throughout the whole range, is the quantile forecast output, or GHI(q, t,  $\tau$ ).[28]

### Calculating Statistical Errors:

Several statistical criteria are used to evaluate the forecasting model, depending on whether the forecast is made exclusively during the day or during the entire night. A few important metrics are the Kolmogorov-Smirnov test integral (KSI), Pearson's correlation coefficient ( $\rho$ ), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Maximum Absolute Error (MaxAE), Root Mean Squared Error (RMSE), Normalized RMSE (nRMSE), Mean Bias Error (MBE), and OVER metrics.

The correlation between forecast and actual solar power variables is measured by Pearson's correlation coefficient, while global error metrics are provided by RMSE and nRMSE. While MAE provides a global error measure that is not overly sensitive to severe occurrences, MaxAE evaluates local deviations, particularly for short-term extreme events. Average forecast bias and percentage mistakes are explained by MAPE and MBE in that order.

$$\text{MaxAE} = \max_{i=1,2,\dots,N} |p_i - \pi_i|$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |p_i - \pi_i|$$

The cumulative distribution functions (CDFs) of the projected and actual solar power are compared using the KSI and OVER metrics, which add to the overall evaluation of the forecasting model's effectiveness. Better predicting performance is indicated by a smaller KSI number; a relative KSI offers a normalized assessment.

For a thorough assessment, this integrated approach to multi-horizon solar irradiance forecasts combines sophisticated methods like quantile regression with a wide range of statistical measures. By improving the precision and resilience of solar irradiance forecasts, the approach hopes to optimize the use of renewable energy sources and grid management [29].

The Unsupervised Clustering and Machine Learning based Multi-Model (UC-M3) framework, a novel method for short-term global horizontal irradiance (GHI) forecasting, is presented in this paper. The methodology includes UC-M3 forecasting, which chooses the top-performing models for each clustered task, Support Vector Machine Pattern Recognition (SVM-PR) for label recognition with sparse data, and Optimized Cross-validated CIUsteRing (OCCUR) method to autonomously determine optimal clustering. A year's worth of solar data from case studies show successful clustering, high label recognition accuracy, and notable improvements in forecasting accuracy. Deep learning algorithms are recommended for use in future research to enable longer-term solar forecasting. A promising development in GHI forecasting is the UC-M3 framework, which combines machine learning and unsupervised clustering techniques to improve forecast accuracy[13].

---

### 3. SOLAR IRRADIANCE FORECASTING DEEP LEARNING MODEL BASED ON SELF ATTENTION

This section introduces RSAM (Transformer-based Solar Time Series Forecasting), a model that forecasts GHI values across various time horizons using multi-head attention. Recurrent Neural Networks (RNNs) are no longer needed to identify both short- and long-term relationships in the data. There are two steps in RSAM:

**Point Forecasting:** By preprocessing data, training models, and testing them, we forecast GHI values using historical weather data.

**Prediction Intervals:** We use quantile regression to compute prediction intervals in order to improve dependability.

RSAM is useful for solar energy forecasting since it provides precise point predictions and measures prediction uncertainty.

---

### 4. METHODOLOGY

The experimental setup, including information on the dataset, model architectures, and parameters, is described in the methods section. It also explains how each model is trained and how the experiment is set up.

#### A. DATASET

The Measurements of solar irradiance at a temporal resolution of sunset and dawn are provided by the 2019 Solar Irradiance dataset. The dataset spans a year and includes 32,687 entries. For our experiments, the dataset we used was split into training and testing sets. Furthermore, the entire dataset was used for training; a random selection of the 10 days' worth of data was used just for testing. It's interesting to see that not a single dataset element was reserved for validation. Writers and Associations.

#### B. Model architecture

Given its inherent benefits over recurrent models in terms of memory, compute resources, parallelizability, and ease of implementation, the self-attention based deep learning architecture is used in the RSAM forecasting portion. Figure 2 depicts and describes the data preprocessing, training, and testing stages of the RSAM model.

Residual computation and data preprocessing: Seasonality and temporal dependencies on GHI values are eliminated by data preprocessing. On the training set of data, many residual types (year, month, day, and GHI residuals) are computed to capture information related to the original data. The aforementioned residuals denote the average departure of GHI values during the course of a year, month, day, and time of day, respectively, from the earlier periods of the data set. By modifying the variability (bias) in GHI values, these residuals quicken the deep learning model's learning process.

This phase is fully explained in the method that follows: When the model has finished training, the testing phase predicts GHI values at various phases. By reversing the preprocessing procedures using stored "yearly," these projected GHI values are again moved to the original GHI values.

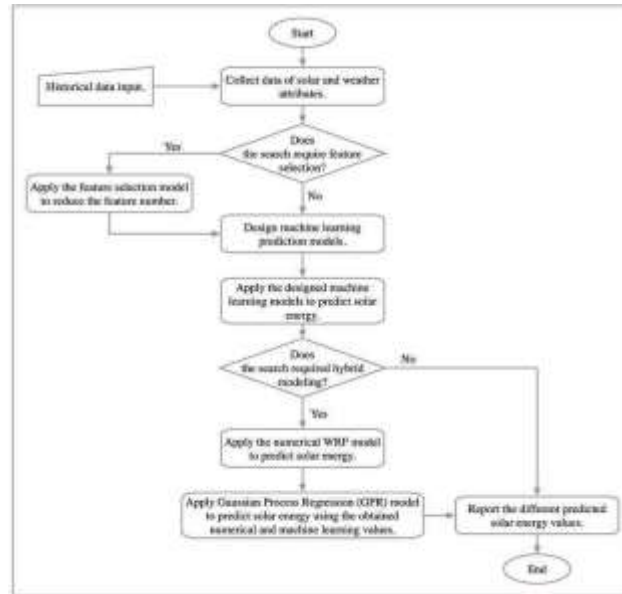


Fig. 2. Forecasting model architecture

**Training Phase:** During this stage, a thorough evaluation of the trained model is conducted. It quantile forecasts prediction uncertainty and predicts GHI values with accuracy. By using walk-forward validation, the robustness and dependability of the model are guaranteed in practical situations.

By giving quantile forecasts, the model extends beyond point predictions and provide a comprehensive picture of prediction intervals and uncertainty. This holds special value in applications like renewable energy system optimization where accurate GHI projections are crucial.

In conclusion, the machine learning model that has been described is an advanced and thorough method of GHI forecasting. It is an essential tool for many businesses since it can capture complex data relationships and provide precise forecasts by utilizing transformer-based design and self-attention processes.

**y\_r:** year\_residual **m\_r:** month\_residual **d\_r:** day\_residual

**GHI\_residual:** new (processed) GHI values

**Input:**

**t:** Time [year, month, day, hour, minute] **D:** Training dataset indexed by 't' **Original GHI(t):** GHI at time 't'

**Procedure:**

**for t in D:**

$y_r = \text{GHI}(t) - \text{AvgD}(\text{GHI})$

$m_r = y_r - \text{Avg}(y_r)$  if  $t[1] == t[1]$  else  $m_r$

$d_r = m_r - \text{Avg}(m_r)$  if  $t[1:3] == t[1:3]$  else  $d_r$  **GHI\_residual =**  $d_r - \text{Avg}(d_r)$  if  $t[1:5] == t[1:5]$  else **GHI\_residual**

## 5. Experimental Discussion

**Research Purpose:** This study's main goal is to evaluate the RSAM model's accuracy and effectiveness in solar irradiance prediction. This has a great deal of practical significance, particularly in meteorology and renewable energy-related sectors.

### C. Data Source:

A solar environmental dataset that was obtained is used in this study. There is a plethora of chronological data in this dataset, including year, month, day, hour, and minute. Preparing and validating forecasting models require this kind of information.

Parameters	Dataset
UNIXTime	1475229326
Data	9/29/2019 12:00:00 AM
Time	0.996828704
Radiation	1.21
Temperature	48
Pressure	30.46
Humidity	59
WindDirection(Degrees)	177.39
Speed	5.62
TimeSunRise	0.259027778
TimeSunSet	0.759027778

TABLE 1:

Some Random Data from Source Dataset

#### D. Metrics for Assessment:

Several assessment metrics are used to assess the accuracy of forecasts, including root mean square error (RMSE), mean bias error (MBE), mean absolute error (MAE)[29]-[31], and a unique "forecast skill" score. The purpose of these measures is to measure the prediction effectiveness of the and how it outperforms a baseline persistence model.

#### E. Forecasting Methods:

The RSAM model is not the only one used in the research. A thorough comparison examination using many benchmark forecasting approaches is included. Most notably, the clear- sky index is used to compensate for constant sky conditions through the use of the "Smart Persistence Algorithm". When compared to known techniques, this all-inclusive evaluation enables a thorough analysis of the RSAM model's performance.

$$GHISmartPersistence(t+\tau) = CSI(t) \times GHIClearsky(t+\tau)$$

where  $CSI(t)$  is the clear-sky index correction factor, defined as:

$$CSI(t) = GHI(t) / GHIClearsky(t)$$

#### F. Deep Learning Models:

The prediction of solar irradiance using deep learning models is investigated. Models such as Attention-Based LSTM (A- LSTM), Convolutional Neural Network- LSTM (CNN- LSTM), Attention-Based CNN-LSTM (A-

CNN-LSTM), and Long Short Term Memory (LSTM) are specifically taken into consideration. To capture complex data patterns and interdependencies, these models make use of neural networks. According to the text, careful model tweaking parameters was carried out, highlighting the critical significance of parameter selection.

#### G. Experimental Setup:

On a PC with an Intel Core i7 CPU, Python 3.6 and the Tensorflow package are used for the model training procedure. These specifics guarantee openness and reproducibility by providing information on the software tools and computer resources utilized in the research.

#### H. Model Assessment:

A thorough and methodical assessment of the RSAM model's predicting ability is hinted to in this line. The primary aim is to have a deep comprehension of its efficacy in various settings and juxtapose it with alternative approaches. These evaluations could include doing model tests on discrete portions of the dataset.

## 6. RESULTS AND ANALYSIS

model that only takes as input meteorological parameters. RSAM, LSTM, and Smart-Persistence have average RMSEs of 52.45, 63.21, 59.88, 55.93, 54.56, and 96.23

on the Amsterdam dataset. Smart-Persistence, LSTM, 64.87, 70.13, and RSAM have average RMSEs of 116.02, 135, 82.18, 73.07, and RSAM for Amsterdam, respectively. The outcomes at each site demonstrate how accurately RSAM forecasts solar irradiance over a variety of time horizons. With one foot in front of the other, the Amsterdam, Netherlands, RSAM model achieves the lowest RMSE (50.82 W/m<sup>2</sup>), and the MAE values are 26.23 W/m<sup>2</sup>. The outcome of the RSAM model is the low deviation indicated by its impartial, accurate prediction. Forecasting abilities typically display higher scores for all other comparable models when compared to the RSAM model. When attention mechanisms and sequential models (LSTM) are employed, the models perform better. The MBE values of the sample exhibit a negligible variance.

Examination of Seasonal Changes

The 2019 (Amsterdam) season-wise performance study considers three seasons: summer (June to August), winter (December to February), and spring (February to May). Amsterdam observes the summer (March to May), monsoon (June to September), and winter (November to February) seasons. An extensive seasonal analysis of all the techniques on Amsterdam in 2018 and 2019 using different accuracy measures at all intervals correspondingly demonstrates that the RMSE values are highest during the monsoon season, moderate in the summer, and lowest in the winter. This is due to the unstable sky conditions during the monsoon and the low GHI values during the winter. The greater forecasting skill of the monsoon is explained by the significant difference in the RMSE values between RSAM and smart persistence. Similarly, results for the Amsterdam data set demonstrate that RSAM forecasts with higher accuracy in all seasons when compared to other reference algorithms. Summertime sees the highest RMSE values, springtime brings moderate values, and wintertime brings lowest values. The mild GHI readings and less fluctuation in the spring compared to the summer are the main causes of this. An examination broken down by season also shows that in 2019, RSAM performs better than 2018.

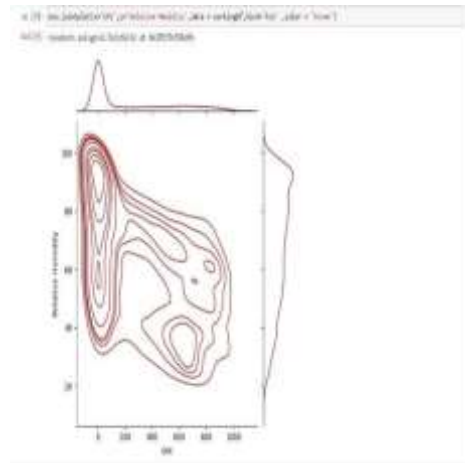


Fig 3 :An example of twelve-step forecasts with a 90% prediction interval

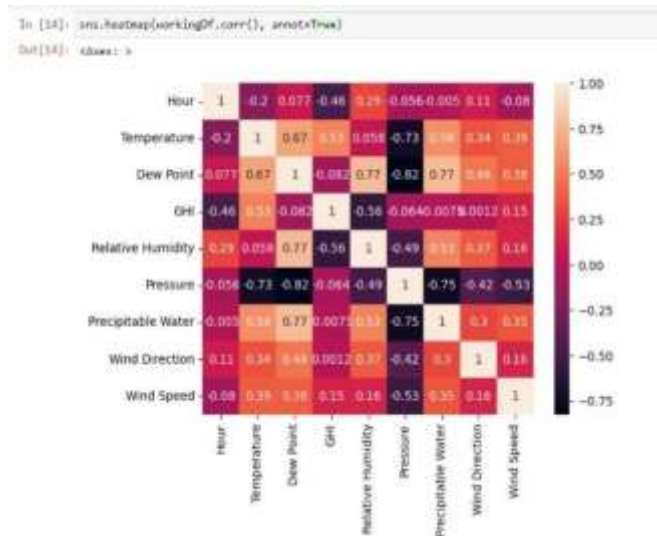


Figure 4: Charting a Representative Dataset

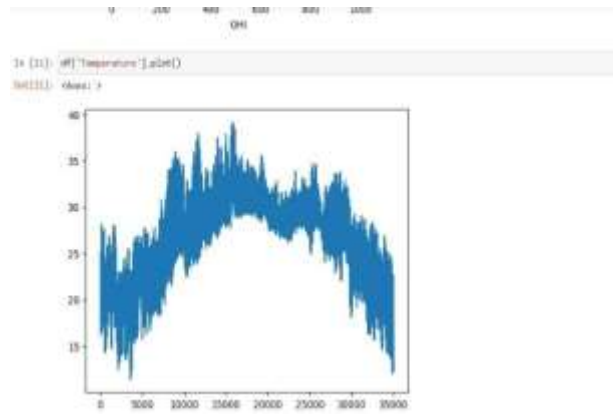


Figure 5: An illustration of 12-step forecasts with a 90% GHI prediction interval

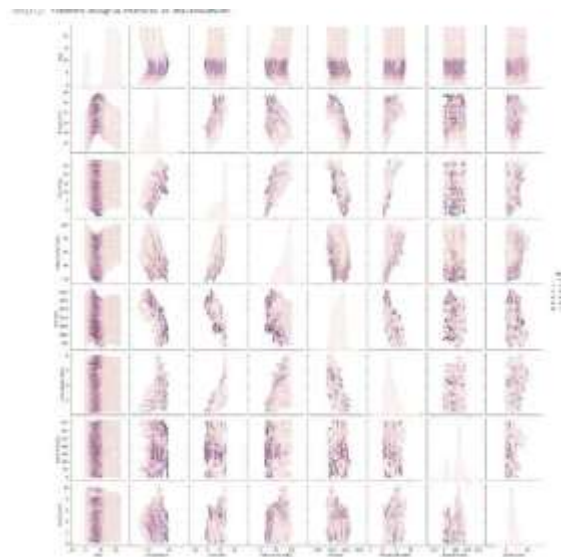


Figure 6: pair plot of the values under comparison

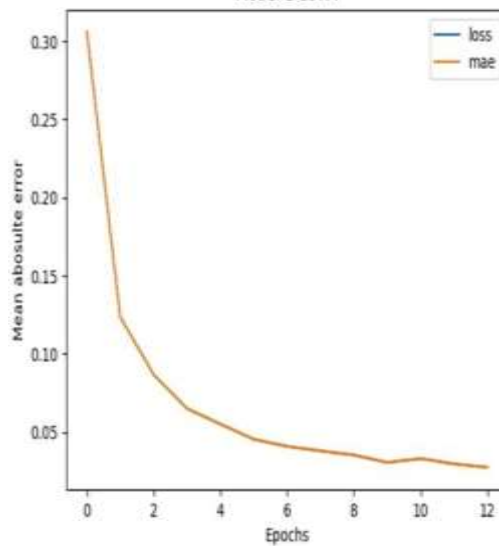


Figure 7: Plotting Loss Value and Mae Metrics



## 6. CONCLUSION

The proposed work uses the Transformer deep learning model for point forecasting and quantile regression for interval prediction to create a reliable multi-step solar irradiance prediction model. Customers can lower risk by being aware that there's a chance the actual GHI value won't match the estimate. Requirements for real-time smart grid operations include a multi-step forecast with an accurate prediction interval. The suggested model performs better than recurrent deep learning models (LSTM) in terms of training time and performance accuracy, according to a thorough analysis of the data. The NREL dataset, which consists of two distinct sites with varying climates, shows that the suggested RSAM model performs better than the others both annually and seasonally. In Amsterdam, the RSAM is 58.89%.

## 7. REFERENCES

- [1] H. Murdock, "Renewables 2019 global status report," 2019. [Online]. Available: <http://hdl.handle.net/20.500.11822/28496>
- [2] "Solar energy research data," Accessed: Jan. 21, 2020. [Online]. Available: <https://www.seia.org/solar-decade>
- [3] "Solar Energy in India," May 2019, Accessed: Feb. 2, 2020. [Online]. Available: <http://www.economicstimes.indiatimes.com>
- [4] "Irena Future of Solar Pv 2019," Nov. 2019, Accessed: Feb. 28, 2020. [Online]. Available: <https://www.irena.org/>
- [5] B. Chen and J. Li, "Combined probabilistic forecasting method for photo voltaic power using an improved Markov chain," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 19, pp. 4364–4373, 2019.
- [6] W. Ji and K. C. Chee, "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN," *Sol. Energy*, vol. 85, no. 5, pp. 808–817, 2011. [7] D. Yang, Z. Ye, L. H. I. Lim, and Z. Dong, "Very short term irradiance forecasting using the lasso," *Sol. Energy*, vol. 114, pp. 314–326, 2015.
- [7] A. Shakya et al., "Solar irradiance forecasting in remote microgrids using Markov switching model," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 895–905, Jul. 2017.
- [8] Y. Jiang, H. Long, Z. Zhang, and Z. Song, "Day-ahead prediction of bihourly solar radiance with a Markov switch approach," *IEEE Trans. Sustain. Energy*, vol. 8, no. 4, pp. 1536–1547, Oct. 2017.
- [9] C. Voyant et al., "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569–582, 2017.
- [10] J. Zeng and W. Qiao, "Short-term solar power prediction using a support vector machine," *Renewable Energy*, vol. 52, pp. 118–127, 2013.
- [11] K. Y. Bae, H. S. Jang, and D. K. Sung, "Hourly solar irradiance prediction based on support vector machine and its error analysis," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 935–945, Mar. 2017.
- [12] C. Feng, M. Cui, B. Hodge, S. Lu, H. F. Hamann, and J. Zhang, "Unsupervised clustering-based short-term solar forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 4, pp. 2174–2185, Oct. 2019.
- [13] E. F. Alsina, M. Bortolini, M. Gamberi, and A. Regattieri, "Artificial neural network optimisation for monthly average daily global solar radiation prediction," *Energy Convers. Manage.*, vol. 120, pp. 320–329, 2016.
- [14] K. Chiteka and C. Enweremadu, "Prediction of global horizontal solar irradiance in zimbabwe using artificial neural networks," *J. Cleaner Prod.*, vol. 135, pp. 701–711, 2016.
- [15] R. Tapakis, S. Michaelides, and A. Charalambides, "Computations of diffuse fraction of global irradiance: Part 2 – neural networks," *Sol. Energy*, vol. 139, pp. 723–732, 2016.
- [16] S. Srivastava and S. Lessmann, "A comparative study of LSTM neural net works in forecasting day-ahead global horizontal irradiance with satellite data," *Sol. Energy*, vol. 162, pp. 232–247, 2018.
- [17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, pp. 1735–80, Dec., 1997.
- [18] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Jul., 2019.
- [19] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, 2018. [21] K. Wang, X. Qi, and H. Liu, "Photovoltaic power forecasting based LSTM convolutional network," *Energy*, vol. 189, 2019, Art. no. 116225.
- [20] A. Vaswani et al., "Attention is all you need," 2017, arXiv:1706.03762.
- [21] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2018, arXiv:1810.04805. [24] C.-Z. A. Huang et al., "Music transformer," 2018, arXiv:1809.04281.
- [22] N. Parmar et al., "Image transformer," 2018, arXiv:1802.05751.

- 
- [23] S. Li et al., "Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting," 2019, arXiv:1907.00235.
- [24] J. Ma, Z. Shou, A. Zareian, H. Mansour, A. Vetro, and S.-F. Chang, "Cds: Cross-dimensional self-attention for multivariate, geo-tagged time series imputation," 2019, arXiv:1905.09904.
- [25] B. Lim, S. O. Arik, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," 2020, arXiv:1912.09363.
- [26] M. Kakimoto, Y. Endoh, H. Shin, R. Ikeda, and H. Kusaka, "Probabilistic solar irradiance forecasting by conditioning joint probability method and its application to electric power trading," *IEEE Trans. Sustain. Energy*, vol. 10, no. 2, pp. 983–993, Apr. 2019.
- [27] V. V. Kostylev and A. Pavlovski, "Solar power forecasting performance – towards industry standards," in *Proc. 1st Int. Workshop Integr. Sol. Power Syst.*, Aarhus, Denmark, 2011.
- [28] Sharda, Swati & Singh, Mukhtiar & Sharma, Kapil. (2020). RSAM: Robust Self-Attention Based Multi-Horizon Model for Solar Irradiance Forecasting. *IEEE Transactions on Sustainable Energy*. PP. 1-1. 10.1109/TSTE.2020.3046098.
- [29] Chaturvedi, D. & Singh, Isha. (2016). Solar Power Forecasting: A Review.