

Enhancing Solar Forecasting Accuracy with Sequential Deep Artificial Neural Network and Hybrid Random Forest and Gradient Boosting Models across Varied Terrains

Muhammad Farhan Hanif,* Muhammad Umar Siddique, Jicang Si,
 Muhammad Sabir Naveed, Xiangtao Liu, and Jianchun Mi

Effective solar energy utilization demands improvements in forecasting due to the unpredictable nature of solar irradiance (SI). This study introduces and rigorously tests two innovative forecasting models across different locations: the Sequential Deep Artificial Neural Network (SDANN) and the Deep Hybrid Random Forest Gradient Boosting (RFGB). SDANN, leveraging deep learning, aims to identify complex patterns in weather data, while RFGB, combining Random Forest and Gradient Boosting, proves more effective by offering a superior balance of efficiency and accuracy. The research highlights the SDANN model's deep learning capabilities along with the RFGB model's unique blend and their comparative success over existing models such as eXtreme Gradient Boosting (XGBOOST), Categorical Boosting (CatBOOST), Gated Recurrent Unit (GRU), and a K-Nearest Neighbors (KNN) and XGBOOST hybrid. With the lowest Mean Squared Error (147.22), Mean Absolute Error (8.77), and a high R^2 value (0.80) in a studied region, RFGB stands out. Additionally, detailed ablation studies on meteorological feature impacts on model performance further enhance accuracy and adaptability. By integrating cutting-edge AI in SI forecasting, this research not only advances the field but also sets the stage for future renewable energy strategies and global policy-making.

necessitates a significant shift toward renewable sources due to the detrimental environmental impact of fossil fuels.^[1,2] Solar energy, particularly effective in regions with high solar exposure such as Africa, the Southwestern US, China, Morocco, and Australia, is emerging as a crucial renewable contender.^[3] The importance of advanced solar forecasting is recognized by major energy stakeholders like the Australian Renewable Energy Agency (ARENA), the National Grid, and the California Energy Commission, offering substantial benefits.^[4] The solar Photovoltaic (PV) sector's growth, with a 22% increase in 2021, underscores its strong investment potential, dominated by China, the US, and India, with Asia expected to lead the market by 2050.^[5] Embracing this trend, Pakistan targets 30% renewable electricity by 2030, capitalizing on its ample solar resources. Advancements in solar energy utilization rely on accurate forecasting of Global Horizontal Irradiance (GHI), a complex task due to various meteorological factors.^[6,7] In response, diverse forecasting methods have

developed, including image-based models, integrations of machine learning with Numerical Weather Prediction (NWP) models,^[8,9] and deep learning (DL) techniques, especially effective for bridging short to medium-term forecasting gaps.^[10]

1. Introduction

1.1. Background

The anticipated rise in global energy demand, from 13 terawatt (TW) currently to 46 TW by the end of the 21st century,

M. F. Hanif, J. Si, M. S. Naveed, X. Liu, J. Mi
 Department of Energy & Resource Engineering
 College of Engineering
 Peking University
 Beijing 100871, China
 E-mail: farhanhanif@pku.edu.cn

M. F. Hanif, M. U. Siddique
 Department of Mechanical Engineering
 FE&T Bahauddin Zakaria University Multan
 Multan 60000, Pakistan

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adts.202301289>

DOI: [10.1002/adts.202301289](https://doi.org/10.1002/adts.202301289)

1.2. Related Work

In the realm of forecasting, the approach involves an extensive examination of various input variables, both internal and external, aiming to generate precise forecasts for a defined future period. The accuracy of these forecasts is closely linked to the precision of the input parameters and the forecast's duration, categorized into short-term ($t \leq 1$ day), medium-term ($t \leq 14$ days), and long-term ($t > 14$ days) intervals.^[11,12] A variety of methods are applied in this domain, encompassing physical, statistical, and machine learning models,^[13] each finding its utility in scenarios ranging from immediate operational planning^[14] to extensive long-term strategic planning.^[15] In the realm of long-term predictions, the introduction of DeepSI, a Bayesian deep

learning framework utilizing Global Circulation Models (GCMs) for SI forecasting, underscores the evolving sophistication in tackling climate change-related challenges.^[16]

Physical models in this area are distinguished by their dependence on fundamental attributes of the photovoltaic system, foregoing the need for historical data. Inputs such as numerical weather predictions (NWP), site-specific monitoring data, and specific characteristics of the power plant are utilized,^[17,18] leading to outputs like global horizontal irradiance (GHI) and direct normal irradiance (DNI), directly influencing the plant's production capacity.^[19] The innovative downscaling methodology Global precipitation mission (GPM)- based multitemporal weighted precipitation analysis (GMWPA), which utilizes the GPM_IMERGDF Product and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) DEM within an empirical distribution-based framework (EDBF) Algorithm, underscores the vital role of high-resolution spatial data in precipitation analysis, particularly in areas like Mainland China.^[20] The precision of meteorological data critically determines the accuracy of these models, with their predictive efficacy declining amidst imprecise weather information.^[21] This category encompasses methodologies like the ASHRAE and Hottel equations.^[22,23]

Conversely, statistical methods leverage historical data to discern patterns and relationships between input variables and power production. In the solar energy realm, these methodologies are widespread, incorporating diverse techniques like Markov Chains,^[24] fuzzy logic,^[25] and auto-regressive^[26] models such as Nonlinear Autoregressive model with eXogenous inputs (NARX)^[27] and Nonlinear Autoregressive Moving average with eXogenous inputs (NARMAX).^[28] The study on daylight utilization for energy saving in Karachi's buildings, which identified optimal dome areas for energy efficiency, serves as a practical example of how forecasting can aid in enhancing building energy conservation.^[29] Despite generally being less complex than physical models, the reliance on statistical methods on historical data facilitates more detailed modeling of specific plant characteristics.^[30,31] Nonetheless, the necessity for comprehensive plant-specific data may introduce constraints in quickly changing or growing energy landscapes.

Machine learning models in this domain deploy algorithms trained on extensive datasets of input-output pairs to tailor predictions to specific scenarios. These models encompass Decision Trees,^[32] Random Forests (RF),^[33] and their various ensembles,^[34] all adept at navigating the complex conditions of different power plant sites. Support Vector Machines (SVM)^[35] and ANN^[36] excel at identifying critical input factors and offer interpretability. Similar advancements are echoed in studies like the one employing Optimized Data-Driven Machine Learning Models, where Support Vector Regressor (SVR),^[37] Multilayer Perceptron (MLP),^[38] and RF Regressor^[39] optimized with algorithms such as Moth Flame Optimization (MFO), Grey Wolf Optimization (GWO), and Evolve Class Topper Optimization (ECTO) significantly improved SI predictions.^[40] Deep Neural Networks (DNN)^[41] are applied more selectively for their comprehensive approach to time-series data. Recurrent Networks (RNN),^[42] including specialized forms like Long Short-Term Memory (LSTM)^[43] and Gated Recurrent Unit (GRU),^[44] excel in temporal data modeling, thus improving prediction stability.

Convolutional Networks (CNN)^[45] are especially used for spatial analysis of power production data, enabling precise localized predictions.^[46]

Hybrid models that blend spatial and temporal analysis approaches have shown promise in solar energy forecasting. For instance, the ConvLSTM model combines convolutional layers within LSTM units, offering both translation invariance and an expansive analysis capability, which is crucial for accurate SI predictions.^[47] WaveNet and Temporal Convolutional Networks (TCN) utilize causal dilated convolutions to enhance temporal data interpretation.^[38,48] The Transformer model employs an attention mechanism to focus predictions on specific time steps, representing a significant advancement in forecasting methodologies.^[49] Notably, the hybrid CNN-LSTM model proposed by Ghimire et al.^[50] and Zang et al.^[51] integrates pattern recognition with time-series analysis, reducing data dependencies and improving forecasting accuracy. The integration of attention mechanisms further bolsters model performance by addressing both short-term and long-term temporal patterns effectively.^[52] Additionally, research optimizing the eXtreme Gradient Boosting (XGB) model for solar radiation forecasting has demonstrated remarkable accuracy enhancements, highlighting the efficacy of hyperparametric tuning in this field.^[53] However, it is important to acknowledge the inherent challenges associated with each model type. RNNs, for example, may face training difficulties due to vanishing gradients^[54] and numerical instability,^[55] while CNNs often require extensive training data and complex architectures to achieve comprehensive receptive fields.^[56]

1.3. Contribution

Integrating sophisticated AI and machine learning for SI forecasting faces challenges, including balancing computational efficiency with accuracy due to solar data's complexity.^[57] There is a research gap in effectively combining advanced AI with a deep understanding of meteorological impacts on SI, often leading to issues like overfitting and data inconsistency.^[58–62] Limited regional testing also restricts models' global applicability, crucial for renewable energy adoption.^[63]

Our study explores solar energy potential using innovative deep learning techniques, focusing on SDANN and Hybrid RFGB model. SDANN, with its complex multi-layered setup and ReLU activation, excels in processing sequential data. On the other hand, RFGB uniquely combines RF's^[64] robust decision-making and GB's^[65] precision. By leveraging the strengths of these models, we aim to significantly enhance SI prediction accuracy, drawing on the proven effectiveness of ensemble methods for superior performance. This approach positions our research at the forefront of advancements in SI forecasting.

Our research not only focuses on SI forecasting but also conducts detailed ablation studies to assess how different meteorological features—like temperature, humidity, wind speed, and atmospheric pressure—affect model performance. Through systematic examination, we aim to refine the accuracy and adaptability of our models, gaining deeper insights into the dynamics of SI forecasting. This approach enriches the field of renewable energy forecasting with a nuanced understanding of the relationship between environmental factors and SI predictions.

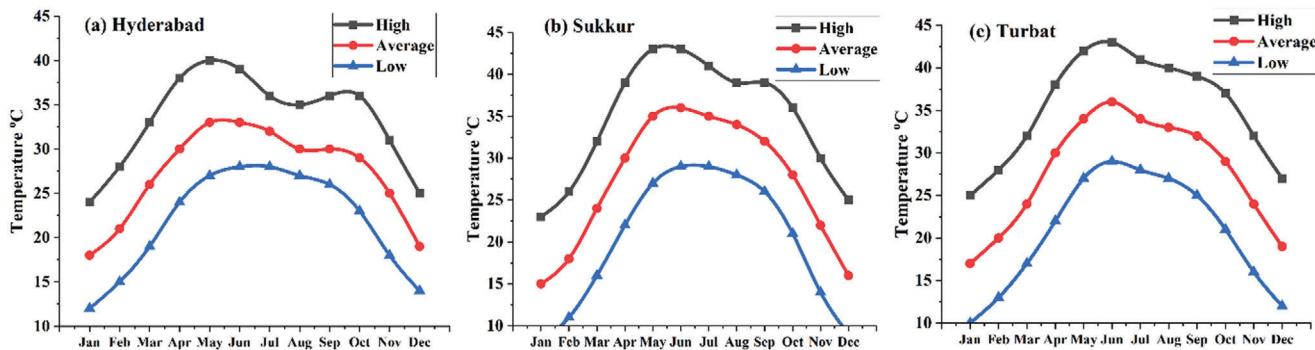


Figure 1. Annual Temperature Trends in Selected Data Sites a) Hyderabad b) Sukkur c) Turbat Reproduced with permission.^[66] weatherspark.com.

To encapsulate, the primary objective of this research is three-fold, i.e.,

- 1) To design and articulate both SDANN and RFGB modeling frameworks;
- 2) To empirically validate these models using robust statistical metrics, comparing them with existing models to ascertain their predictive prowess in the field of SI forecasting;
- 3) To evaluate the predictive accuracy of SI forecasting models across diverse geographies, understand how regional variations influence model performance and reliability; and
- 4) To investigate how various meteorological features impact the performance of forecasting models. This involved the incremental addition and analysis of various features, enhancing the understanding of their individual and collective effects on model accuracy.

The research methodology includes data gathering, model formulation, and assessment criteria, which are extensively described in Section 2. Section 3 provides obtained results and their interpretation; Section 4 provides ablation studies and Section 5 presents detailed analysis and discussion of the obtained results. Lastly, in Section 6, a summary of the research findings is presented, emphasizing the contributions made to the field of solar energy forecasting.

2. Methodology

2.1. Data Sites

The cities of Hyderabad, Sukkur, and Turbat, situated in the provinces of Sindh and Baluchistan, Pakistan, have been meticulously chosen as the testing grounds for our DL models. These locales are distinguished by their advantageous meteorological conditions, notably the substantial SI and the extensive duration of diurnal sunlight across all seasons, including the cooler winter period. The presence of such climatic features renders these sites particularly suitable for the empirical evaluation and calibration of algorithms intended for solar energy potential forecasting within our DL frameworks. As depicted in Figure 1, the temperature profiles of these cities throughout the year are instrumental in understanding the climatic dynamics relevant to solar energy potential. Figure 1 depicts that these urban centers are noted for their unique meteorological patterns, which include hot and hu-

mid summers with temperatures ranging from 23 °C to 44 °C. Specifically, June is the hottest month, recording average high temperatures of 39 °C in Hyderabad, 43 °C in Sukkur and Turbat. In contrast, the winter season is relatively mild, with temperatures fluctuating between 8 °C and 29 °C, with January being the coolest month, showing average low temperatures of 24 °C in Hyderabad, 23 °C in Sukkur, and 25 °C in Turbat.^[66]

The clear skies prevalent throughout the year offer ample opportunities for outdoor activities and contribute to the region's solar power potential. The variation in daylight hours is noteworthy; the shortest day of 2023 is expected to be December 22, with daylight hours of 10'20" in Turbat, 10'23" in Sukkur, and 10'33" in Hyderabad. Conversely, the longest day, June 21, 2023, will provide 13'57" of daylight in Turbat, 13'54" in Sukkur, and 13'43" in Hyderabad.

Complementing this graphical analysis, Table 1 furnishes detailed SI data for each city. It presents the Direct Normal Irradiations (DNI), which is the measure of solar radiation received per unit area by a surface directly facing the sun, the Diffuse Horizontal Irradiations (DIF), which is solar radiation received from the sky excluding the solar disk, and the GHI, the total amount of solar radiation received per unit area from above

Table 1. Key climatic metrics of selected data sites.

Attributes	Hyderabad	Sukkur	Turbat
Latitude	25.3924 N	27.39242 N	26.008 N
Longitude	68.3737 E	68.836899 E	63.0383 E
Elevation (m)	13	67	129
Average Temperature (°C)	28	27	27
Maximum Temperature (°C)	48.5	50.5	53.7
Minimum Temperature (°C)	3	-1.5	-0.4
Sunshine Hours in June	14	14	14
Sunshine Hours in Dec	11	11	10
Average Sunshine Hours	9.1	12	9
Direct Normal Irradiations (DNI) (kW m ⁻² day ⁻¹)	4.644	4.342	5.03
Diffuse Horizontal Irradiations (DIF) (kW m ⁻² day ⁻¹)	2.54	2.577	2.44
Global Horizontal Irradiations (GHI) (kW m ⁻² day ⁻¹)	5.692	5.515	5.781

Table 2. Characterization of satellite-derived input features utilized in proposed models.

Attribute Name	Unit	Acronym	Source	Model	Spatial Resolution
Solar Irradiance	$\text{kWh m}^{-2} \text{ day}^{-1}$	SI	CERES	SYN 1-deg	$1^\circ \times 1^\circ$
Earth's Skin Temperature	$^\circ\text{C}$	Ts	MERRA-2	M2T1NXRAD v5.12.4	$0.5^\circ \times 0.625^\circ$
Minimum Temperature	$^\circ\text{C}$	T_{\min}			
Maximum Temperature	$^\circ\text{C}$	T_{\max}			
Average Temperature	$^\circ\text{C}$	T			
Surface Pressure	kPa	PS			
Relative Humidity	%	RH			
Wind Speed	m^{-1}s	WS			

by a horizontal surface. Notably, the GHI values for all three cities surpass 5 kWh m^{-2} per day, indicating a high potential for solar energy harnessing.^[67]

Given this combination of high solar radiation, prolonged sunshine hours, and favorable temperature ranges, Hyderabad, Sukkur, and Turbat emerge as ideal test sites for the development of solar power infrastructure. This potential aligns with the increasing global demand for renewable energy sources and positions Pakistan to make a substantial contribution to the growth of solar energy within the South Asian region.

2.2. Data Collection

For the purposes of this research, the dataset was meticulously sourced from the National Aeronautics and Space Administration (NASA) Langley Research Center's (LaRC) Prediction of Worldwide Energy Resource (POWER) Project.^[68] This initiative, which receives funding through the NASA Earth Science/Applied Science Program, is renowned for its extensive and reliable long-term solar energy data. The scope of the data acquired for this study encompasses a period from 1994 to 2021, comprising a total of 10002 daily entries each, which provides a comprehensive foundation for our analysis. A variety of climatic and environmental factors can be utilized to develop models for predicting solar radiation. Based on our literature review and data analysis,^[69–73] we have selected eight input parameters for our SI prediction models, detailed description of these parameters can be seen in Table 2.

2.3. Correlation Analysis

The correlation for Sukkur dataset in the form of a pair plot, is depicted in Figure 2. The data statistics and correlation plots for the other datasets are provided as supplementary data. The pair plot provides a comprehensive visualization of the relationships between eight environmental parameters: T, T_{\max} , T_{\min} , PS, RH, WS, Ts, and SI. Each panel in the pair plot depicts a scatterplot of two parameters, with the color of the points indicating the value of a third parameter.^[74,75] Upon inspection of the pair plot, several patterns and relationships emerge. For example, the temperature parameters (T, T_{\max} , and T_{\min}) all exhibit strong positive correlations with each other, indicating that they tend to vary in unison. Conversely, the PS parameter exhibits a weak negative correlation

with the Ts, suggesting that these two parameters have an inverse relationship. Other notable relationships include the moderate negative correlation between T and RH, and the weak positive correlation between PS and T.

It is important to note that the pair plot only reveals correlations between the different parameters, not causation. It is possible that there is a third factor that is driving the observed relationships. For example, the positive correlation between T and T_{\max} may be due to the fact that both parameters are influenced by solar radiation. Overall, the pair plot provides a valuable tool for exploring and visualizing the relationships between multiple parameters. By understanding these relationships, we can better understand the available datasets to incorporate them into models. Pair plots are essential for deep learning models because they reveal correlations between different variables. This information can be used to select the best features, fine-tune model parameters, and identify outliers. SI being the target parameter in this study is the main focus of this plot. The hue determines the value of SI in a particular correlation, i.e., if the correlation is negative as shown at the bottom of plot, SI value decreases with an increase in the values of the input parameters. It is obvious from the top left of the plot that in case of a positive correlation, the SI values increase with the corresponding increase of input parameters.

2.4. Data Pre-Processing

In the process of tailoring robust and effective predictive models, we have employed several data pre-processing techniques. The significance and choice of these techniques stem from their well-established efficacy in enhancing model performance. Below, we elaborate on these pre-processing steps and rationalize their adoption:

2.4.1. Data Cleaning and Handling Missing Values

A pivotal step in our pre-processing is data cleaning, which primarily deals with the challenge of missing values. Incomplete datasets, if not addressed, can substantially undermine the predictive accuracy of models, leading to outcomes that may be biased or even nonsensical. Various techniques can be utilized for the purification of data, depending on the type of data, its format, and the desired responses.^[76–78] To address this, we adopted a

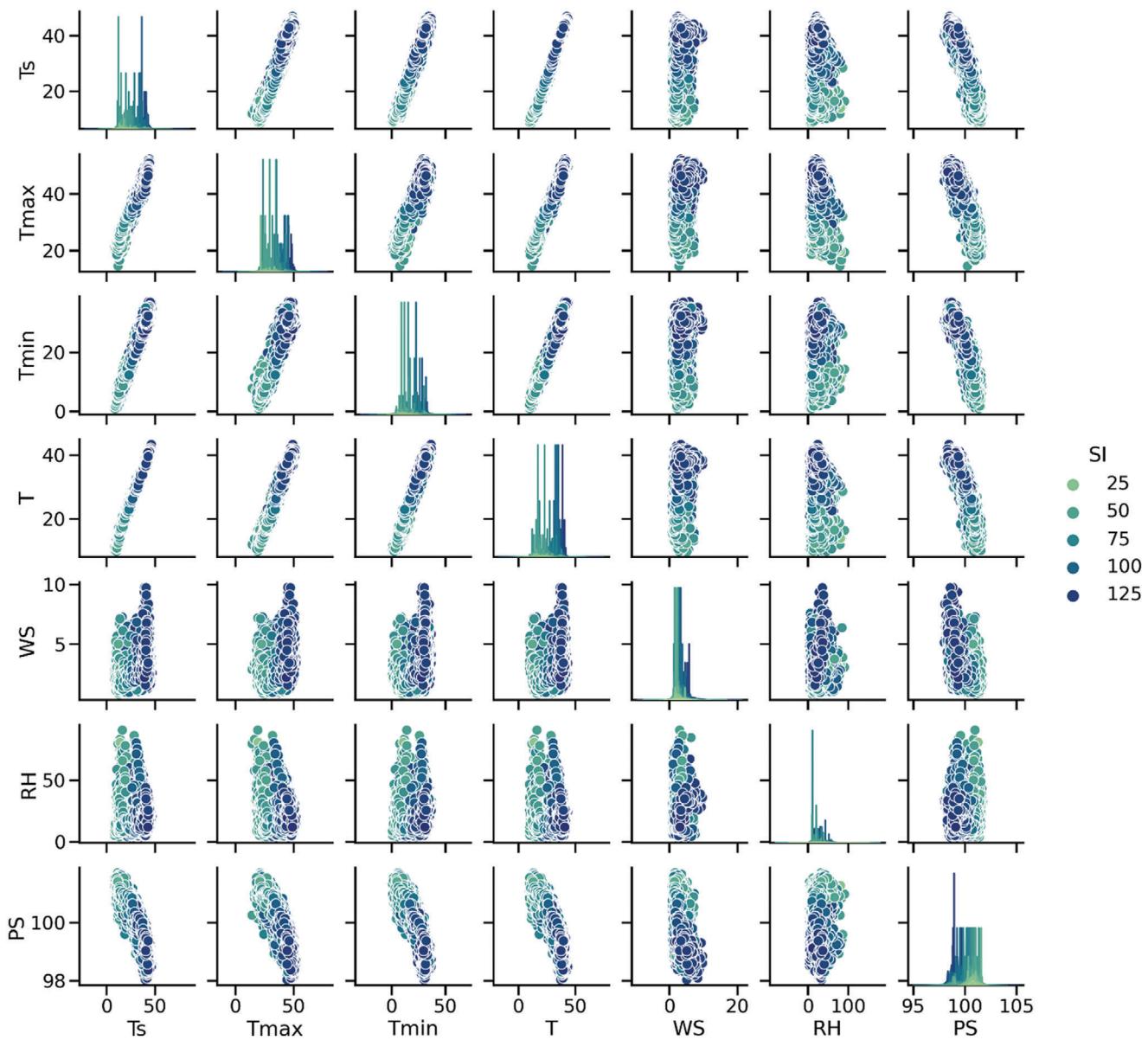


Figure 2. Pair Plot of data distribution for Sukkur.

strategy of eliminating rows with missing values. This approach, often referred to as listwise deletion, ensures data integrity and maintains consistency across the dataset. It is especially beneficial when the proportion of missing data is relatively minimal, thereby minimizing the potential loss of valuable information.

2.4.2. Feature Selection

Feature selection emerges as a pivotal pre-processing step in refining our forecasting model. This stage is critical for mitigating the risks of overfitting—particularly prevalent in complex models like SDANN—and for managing computational complexity in ensemble methods such as RF and GB. **Extraneous features can compromise model efficiency and accuracy, making judicious feature selection essential.**

Building on the insights from the correlation analysis, where we explored the relationships between environmental parameters through pair plots, we adopted a data-driven approach to feature selection. Our objective was to distill the dataset to variables that exhibit significant correlations with SI, the dependent variable of our study. This process was informed by both statistical analyses and domain expertise, ensuring that retained features are relevant for SI prediction.

Initial Correlation Analysis: Leveraging the preliminary insights from the pair plots, we conducted a correlation analysis to quantitatively assess the relationship between each potential feature and SI. **This helped us identify variables with statistically significant correlations as primary candidates for inclusion.**

Exclusion of Non-informative Features: In alignment with our focus on relevance, features that lacked substantial insights into

SI prediction, such as the “Date” column, were excluded. This decision was supported by the lack of a direct or meaningful correlation with SI, underscoring our commitment to model simplicity and performance.

Iterative Refinement: Our selection process is iterative, employing techniques such as dropping features with limited predictive value, exemplified by the exclusion of the “Date” column. This decision-making is facilitated by our code snippet, which systematically refines the dataset to focus on variables directly influencing SI.

Empirical Evaluation: The ablation studies section later elaborates on this process, detailing how features were evaluated and selected based on their predictive power for SI. This includes temperature metrics (T_s , T_{\max} , T_{\min}), PS, RH, and WS, each substantiated by their observed influence on SI variability.

Integration of Findings: By the end of our analysis, the chosen features are those demonstrated through both visual explorations and statistical validation to have significant predictive power for SI, ensuring our model is both efficient and effective.

This approach to feature selection ensures our models are built on a foundation of relevant, impactful variables. The insights gained from initial data exploration, combined with the detailed evaluations in our ablation studies, guide our final feature set, aligning with our goal of enhancing SI forecasting accuracy.

2.4.3. Data Splitting

To ensure a robust evaluation of the model’s performance, it is imperative to divide the dataset into distinct training and testing subsets. While the training set plays a crucial role in building the model, the test set offers an objective and unbiased assessment of the model’s generalization capability. For this purpose, we partitioned the data into an **80-20 distribution**, allocating the majority for training and the remainder for validation.

2.4.4. Feature Scaling

The significance of feature scaling in predictive modeling, particularly for models like SDANN, is paramount. Features measured on different scales can significantly impede the network’s convergence during the training phase, as gradient descent algorithms—commonly used in training these networks—achieve optimal performance when features are standardized. While ensemble models such as RF and GB do not inherently require feature scaling,^[79] standardizing features can contribute to model uniformity, simplifying the interpretation of hybrid models and potentially enhancing training efficiency.

To standardize our features, we employed the following standardization technique,^[80] which normalizes the features to have a mean of zero ($\mu = 0$) and a standard deviation of one ($\sigma = 1$). The mathematical formula for standardizing a feature X is given by:

$$X_{\text{scaled}} = \frac{X - \mu}{\sigma} \quad (1)$$

here: X is the original feature vector, μ is the mean of the feature vector, σ is the standard deviation of the feature vector.

This formula was applied to each feature in our dataset to ensure consistency and uniformity across the data used for training and testing our models.

Implementation: Our code implementation of feature scaling utilizes the StandardScaler from Scikit-learn, which automates the scaling process according to the formula above. In this process, the function `fit_transform`^[81] calculates the mean (μ) and standard deviation (σ) of each feature in the training set and scales the features accordingly. The same scaling parameters (μ) and (σ) are then applied to the test set using `transform`, ensuring that the model sees the test data in the same scale as the training data.

These pre-processing steps, rooted in best practices, are foundational in our machine learning endeavors, enabling our models—whether the SDANN or the hybrid approach—to effectively learn from the data and produce reliable, robust predictions.

2.5. Model Architecture

Throughout the model formulation of SDANN and the hybridized RFGB and validation stages, we leverage a suite of Python’s preeminent libraries. These encompass Matplotlib, Scikit-learn, KERAS, Seaborn, Pipeline, and Pandas.^[82,83] The system setup is anchored by the robust NVIDIA Tesla T4 GPU, equipped with NVIDIA’s Turing architecture, and is configured without display activation, indicating its dedicated use in high-performance computer operations. The system is running with a driver version of 535.104.05 and a CUDA version of 12.2, ensuring compatibility with contemporary machine-learning frameworks and libraries. The Tesla T4’s memory and processing capabilities are leveraged to support intensive computational tasks, underlined by its integration into the Google Compute Engine backend,^[84] highlighting the system’s readiness for scalable and efficient data processing tasks.

2.5.1. Sequential Deep Artificial Neural Network (SDANN)

The development of SDANN is a comprehensive process that involves several crucial steps. The six steps involved in the SDANN development process are data splitting, model architecture selection, input selection, model structure determination, model calibration/training (parameter optimization), and model validation (Figure 3b).^[85,86] The selection of inputs is a vital aspect of SDANN development, as the inclusion of too many correlated or unnecessary inputs can lead to an increase in data training time and overfitting. On the other hand, omitting crucial inputs can result in a model that fails to fully explain the input-output relationship.^[69-71,73]

The SDANN model structure determined the complete flow and structure of the model as shown in Figure 3a. The inclusion of eight pivotal parameters, such as T , T_{\max} , T_{\min} , T_s , RH, SH, WS, and PS reinforces the model’s capability to integrate a broad spectrum of variables essential for accurate forecasting. The number of input nodes in the model is equal to the number of parameters used. One of the distinguishing features of our SDANN model, setting it apart from conventional models, is its

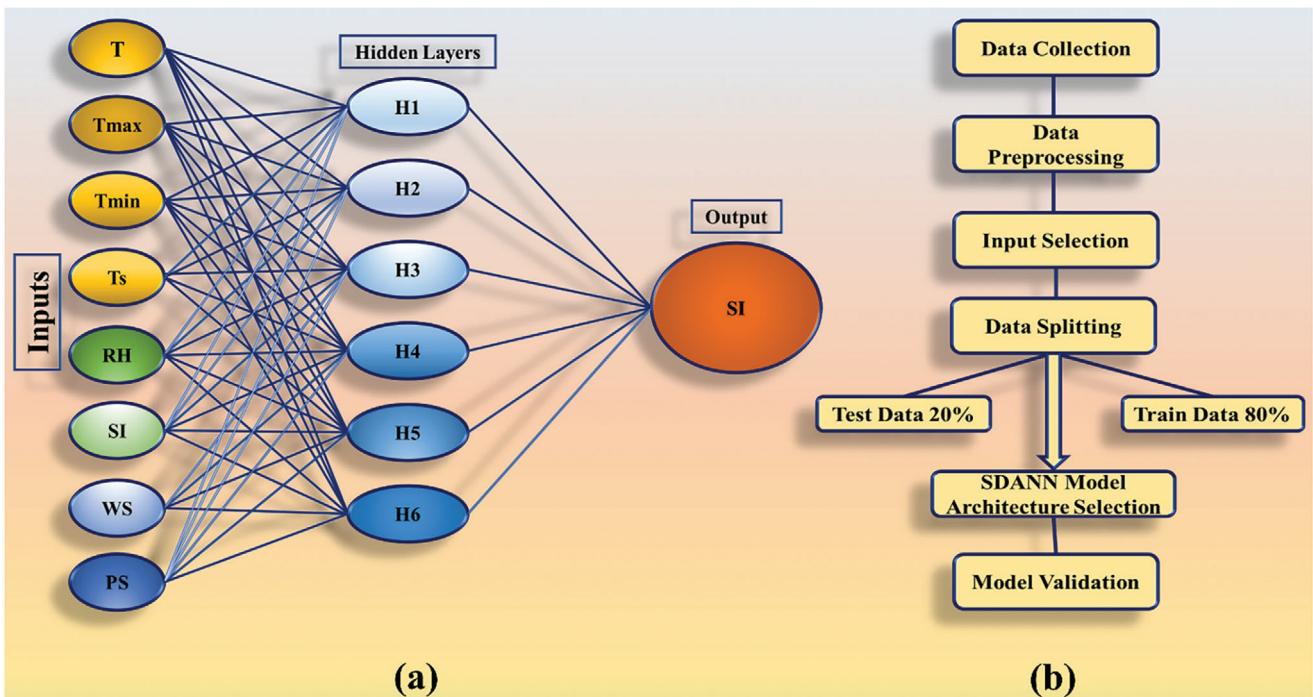


Figure 3. a) Structure of SDANN Model b) Overall SDANN Model Process.

multi-layered architecture comprising six hidden layers which is a deliberate enhancement from the traditional single-layer networks. This multi-layered approach is not arbitrary; it is grounded in the need to capture the inherent complexity of the variables that influence SI, such as temperature, solar heating, wind speed, and pressure^[87,88] ensuring the model's focus on variables essential for accurate forecasting.

In developing the SDANN model's architecture, not only did we focus on incorporating a broad spectrum of variables, but we also carefully selected the Adam optimizer for its robustness and efficiency in handling deep neural networks. The choice of Adam as the optimizer is supported by its demonstrated superiority in various deep learning applications, as highlighted in related studies, including one that assessed its performance against other renowned optimizers like SGD Nesterov, AdaGrad, and RMSProp, among others.^[89] This optimizer, selected for its proven effectiveness and compatibility with complex architectures like ours, further refines our model's capability to integrate and analyze the essential forecasting variables within its multi-layered structure.

Model Architecture and Mathematical Representation: The model initiates with an input layer corresponding to the number of selected features, followed by multiple hidden layers, and concludes with an output layer, structured as follows:

Input Layer: The input layer receives the features, setting the stage for the model's deep learning process.

Hidden Layers: The architecture progresses through six hidden layers, with the transformation at each layer (l) represented by:

$$h_l = f(W_l \cdot h_{l-1} + b_l) \quad (2)$$

where (h_l) is the output from layer (l) , (f) denotes the ReLU activation function, (W_l) is the weight matrix, (b_l) is the bias vector, and (h_{l-1}) is the input from the previous layer (with (h_0) being the initial input features).

Output Layer: The final prediction of SI is obtained from the output layer, formulated as:

$$\hat{y} = f(W_{output} \cdot h_L + b_{output}) \quad (3)$$

Model Training: Training involves optimizing the weight matrices and bias vectors across all layers to minimize the loss function, specifically the Mean Squared Error (MSE) between the predicted and actual values of SI:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (4)$$

The division of the dataset into training and testing sets, with 80% allocated for training and the remainder for validation, coupled with the employment of regularization techniques, ensures the model's robustness and generalizability.

This detailed approach, from the structured architecture to the strategic layering and meticulous training, underscores the SDANN's design sophistication. It is this architecture, grounded in both theoretical underpinnings and empirical validation, that equips the SDANN to adeptly model SI forecasting, setting it apart from conventional models.

Our journey involved navigating multiple libraries, with the Standard Scalar library being pivotal in normalizing our diverse dataset.^[82,83,90,91] The model's structure, facilitated by the Sequential feature, reflects a perfect blend of complexity and clarity. Notably, the uniform distribution employed for neuron weight initialization ensures balanced weight assignment, crucial for

Table 3. Optimal Architecture of the Purposed Models.

	Hyperparameters	Range of Values/Description
SDANN	Hidden Layers	6
	Hidden Neuron	[500, 400, 300, 200, 99, 25]
	Learning rate	0.001
	Epochs	100
	Batch Size	10
	Activation function	[ReLU]
	Solver	[Adam]
RF	Loss Function	Mean Squared Error
	Random State	42
	N Estimators	[70, 140, 210, 280]
	Max Depth	[1–10]
GB	Min Samples Leaf	[0.1, 0.2, 0.3, 0.4, 0.5]
	Random State	42
	N Estimators	[70, 90, 110, 130]
	Max Depth	[1 to 10]
Hybrid RFGB	Min Samples Leaf	[0.1, 0.2, 0.3, 0.4, 0.5]
	Initial Estimator	RF with max. depth = 2, min. samples leaf = 0.1, N estimators = 280
	N Estimators	110
	Max Depth	4
	Min Samples Leaf	0.1

model stability. With the optimizer algorithm at its helm, the model's weight values were meticulously fine-tuned to optimize performance. The complete details of the SDANN model structure can be found in **Table 3**. A uniform distribution was used to represent the weights in the neurons for every hidden layer, and an optimizer algorithm was employed to adjust the values of the weights.

2.5.2. Hybrid Random Forest Gradient Boosting (RFGB)

SI forecasting, a domain with its unique set of challenges, mandates the employment of advanced techniques to optimize predictive accuracy. In this milieu, we present the hybrid RFGB model – an innovative approach aiming to synergize the robustness of RF with the refined precision of GB.

Random Forest (RF): The working principle of the RF algorithm is based on the concept of bagging (bootstrap aggregating). In this technique, a large number of decision trees are trained on random subsets of the training data, and the outputs of all the trees are combined to produce a final output.^[92,93] The algorithm works by dividing a node into a subset of features rather than using the most significant feature, as is the case in traditional decision trees (**Figure 4**). Such diversity ensures that the model isn't overly skewed by any particular data trend, making it especially apt for handling the complex and varied data associated with SI forecasting.

The RF model operates on the principle of ensemble learning, aggregating the outputs of multiple decision trees trained on random data subsets. Mathematically, the RF model's predic-

tion (\hat{y}_{RF}) for a given input (X) can be expressed as the average output of (N) decision trees:

$$\hat{y}_{RF} = \frac{1}{N} \sum_{i=1}^N T_i \cdot (X) \quad (5)$$

where (T_i) represents the (i^{th}) decision tree in the ensemble.

Gradient Boosting: GB, on the other hand, is a three-component algorithm that consists of an optimized loss function, predictions made by a weak learner, and an additive model. The loss function is determined based on the problem being solved, and common loss functions such as the square error for regression or logarithmic loss for classification are used. The weak learner used in GB is typically a decision tree, which is trained to make predictions and minimize the loss function. The decision tree is built in an additive manner, meaning that as new trees are added to the model, existing trees are not altered. A gradient descent approach is used to reduce the loss by parameterizing the trees and moving them in a direction that minimizes the error.^[94–96]

Its prediction (\hat{y}_{GB}) is derived from iteratively adding weighted predictions (h_i) of (M) weak learners:

$$\hat{y}_{GB} = \sum_{i=1}^M \alpha_i \cdot h_i \cdot (X) \quad (6)$$

with (h_i) being the (i^{th}) weak learner's prediction, and (α_i) its corresponding weight.

Meta Model (RFGB): Our innovative RFGB model integrates the RF and GB models' strengths by initially applying RF for a robust baseline prediction, which is then refined through GB's sequential improvement technique. The hybrid model's final prediction (\hat{y}_{RFGB}) combines the RF base model's output with GB's enhancements:

$$\hat{y}_{RFGB} = \sum_{j=1}^M \alpha_j \cdot h_j \cdot (X) \quad (7)$$

Parameter Optimization: Hyperparameter tuning is critical for optimizing the RFGB model's performance. For RF, parameters such as the number of trees (N), maximum depth of trees (d), and minimum samples per leaf node (m_{sl}) are adjusted. Similarly, for GB, we fine-tune the number of stages (M), depth of trees, and minimum samples per leaf, among others, to find the optimal model configuration. This process is facilitated by grid search optimization, systematically exploring a range of parameter combinations to identify those yielding the highest predictive accuracy.

By leveraging the RF and GB models' complementary strengths, our RFGB model offers a sophisticated approach to SI forecasting. The outlined mathematical framework underpins the model's structure and operational logic, ensuring a deep understanding of its functionality without delving into specific coding implementations. The hybrid RFGB model detailed hyperparameters can be visualized in **Table 3**. Also the variation of best hyperparameters for different datasets is also presented in Appendix 1 in Supporting Information.

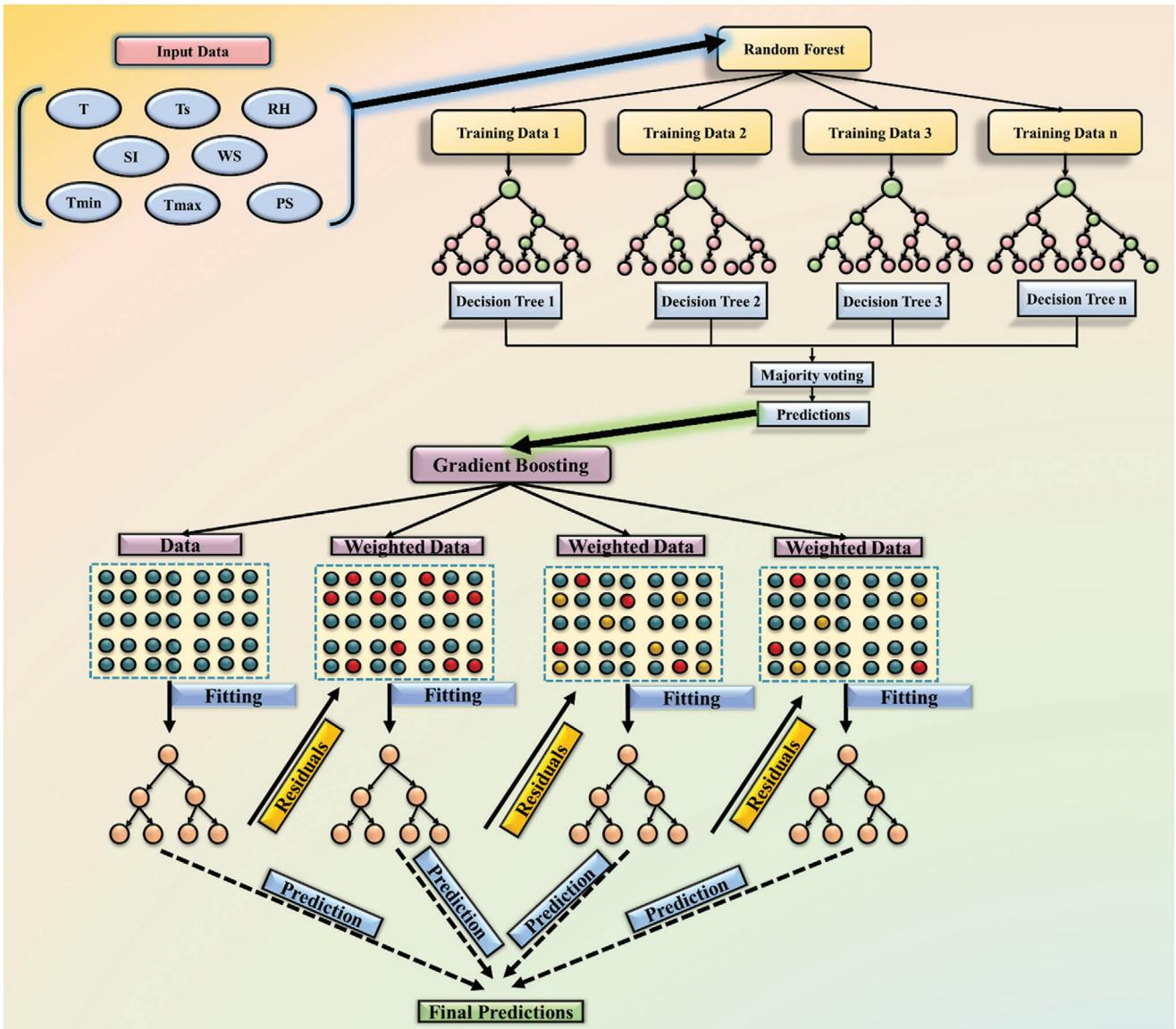


Figure 4. Visualization of Hybrid RFGB Model.

2.6. Statistical Indicators for Evaluation

The evaluation of the accuracy of a predictive model is a crucial aspect in the field of statistics and machine learning. In regression analysis, the comparison between the original and predicted values is used to determine the effectiveness of the model and to identify sources of error. Several metrics are employed to measure the error rate of prediction and the performance of the model in regression analysis. These include Mean Absolute Error (MAE), MSE, Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2).^[97]

3. Results

In this section, the outcomes of deploying advanced deep learning models for SI forecasting are meticulously detailed. Initially, the research evaluates the performance of

SDANN, a model engineered to identify complex patterns within extensive datasets, through rigorous accuracy testing. Subsequently, the effectiveness of the Hybrid RFGB model is scrutinized.

Furthermore, the comparative analysis is enriched by juxtaposing the SDANN and RFGB models against four contemporary models within the domain. This comparative study is designed to situate the contributions of the SDANN and RFGB models within the current landscape of SI forecasting methodologies. Moreover, the adaptability and efficacy of the SDANN and RFGB models are validated across diverse geographical terrains, encompassing Ahvaz (Iran),^[98] Bikaner (India),^[99] Alice Springs (Australia),^[100] Las Vegas (USA),^[101] Marrakech (Morocco),^[102] and Luxor (Egypt).^[103] This phase aims to ascertain the models' performance consistency across varied climatic and geographic conditions, thereby broadening the understanding of their global applicability in SI forecasting.

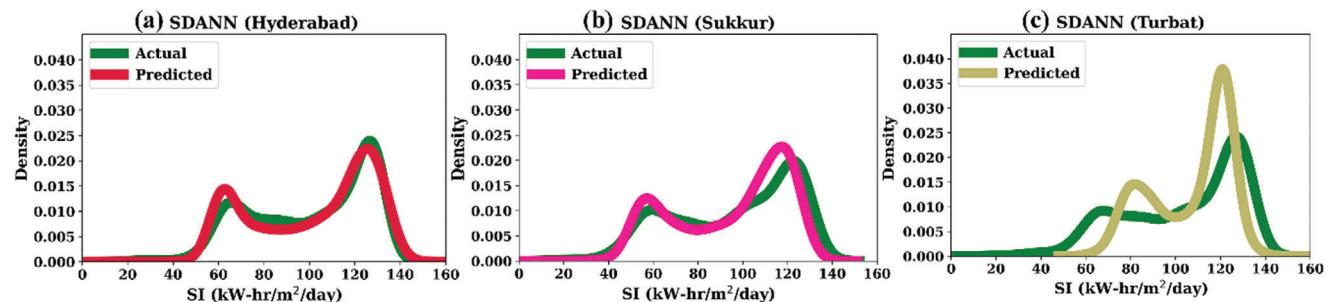


Figure 5. KDE of SDANN Predictions a) Hyderabad b) Sukkur c) Turbat.

By leveraging a comprehensive suite of statistical indicators, this research delineates a quantitative evaluation of each model's accuracy and overall performance, underpinning the significant advancements introduced through this study in the realm of solar energy forecasting.

3.1. SDANN Predictions

The SDANN model, when applied to the SI forecasting in the cities of Hyderabad, Sukkur, and Turbat, exhibits commendable predictive performance, albeit with variations reflective of regional solar profiles.

Figure 5 presents the kernel density estimations (KDE) of actual versus predicted SI values for three distinct cities—Hyderabad, Sukkur, and Turbat—using the SDANN model. The KDE is a non-parametric way to estimate the probability density function of a continuous variable, providing a smooth curve representative of the data's distribution.^[104]

In Hyderabad, the model's density plot (Figure 5a) reveals a pronounced peak at $\approx 125 \text{ kWh m}^{-2} \text{ day}^{-1}$, suggesting a high probability of occurrence around this SI value. The predicted values closely shadow the actual measurements, delineating a narrow divergence and a high kernel density overlap. Moving to Sukkur, the density plot (Figure 5b) displays a peak $\approx 120 \text{ kWh m}^{-2} \text{ day}^{-1}$, indicative of the model's precision in capturing the median irradiance values. Occasional deviations are noted, yet they remain within an acceptable range, underscoring the model's adeptness at handling the region's SI variability.

The closeness of the two curves in each subfigure is an indicator of the model's predictive accuracy. In the cases of Hyderabad and Sukkur, the actual and predicted curves are almost superimposed, suggesting that the SDANN model has a high level of accuracy in predicting SI for these cities. The peaks of the KDE curves indicate the most probable SI values, with both cities showing prominent peaks at similar SI levels, which demonstrates the model's effectiveness in these instances.

Turbat's subfigure (c), however, shows a divergence between the actual and predicted curves. The predicted curve is broader with a shifted peak, indicating variance from the actual SI values. This suggests that for Turbat, the model's predictions are less accurate compared to the other two cities. The broader spread of the predicted curve could imply overfitting, a model's sensitivity to certain variables, or variability in the city's SI not captured during the model's training.

Overall, the SDANN model demonstrates a high degree of accuracy in SI forecasting,^[105] with potential for further optimization, particularly in refining its performance in capturing the full spectrum of SI, especially in regions with more complex meteorological conditions.

3.2. RFGB Predictions

The results depicted in **Figure 6** convey the performance of the hybrid RFGB model in estimating SI for three cities: Hyderabad, Sukkur, and Turbat. The kernel density estimations (KDE) for actual and predicted SI values are presented across three subfigures.

Hyderabad (Figure 6a) demonstrates a model with high predictive accuracy, as evidenced by the substantial overlap of the actual and predicted density curves, particularly around the primary peak at $60 \text{ kWh m}^{-2} \text{ day}^{-1}$ and a secondary peak at $125 \text{ kWh m}^{-2} \text{ day}^{-1}$. This suggests that the RFGB model has a robust capacity for capturing the SI distribution in Hyderabad.

In Figure 6b, representing Sukkur, the model also displays commendable performance, with the predicted SI values closely tracking the actual data. The most prominent peak is observed around $100 \text{ kWh m}^{-2} \text{ day}^{-1}$. The minor underestimation by the model at the apex of the peak is indicative of a slight divergence but does not detract significantly from the overall accuracy.

Contrastingly, Figure 6c for Turbat shows a more pronounced discrepancy between the model's predictions and actual SI values, with the model predicting a broader and bimodal distribution of SI values. This discrepancy suggests that the model may require further tuning to accommodate the unique environmental factors or SI patterns specific to Turbat.

The RFGB model, fine-tuned through an exhaustive search of parameter space using GridSearchCV, indicates that the optimal parameters derived from the cross-validation process have led to a high degree of predictive accuracy for Hyderabad and Sukkur.^[64,65] The divergence observed in Turbat's SI predictions highlights the necessity for a more nuanced approach to model tuning, perhaps incorporating a more detailed analysis of local climatic variables or employing a region-specific model training regimen.

In summary, the RFGB model demonstrates a commendable level of accuracy in predicting SI, particularly in Hyderabad and Sukkur. The slight prediction discrepancies observed in Turbat provide a basis for future enhancements to the model,

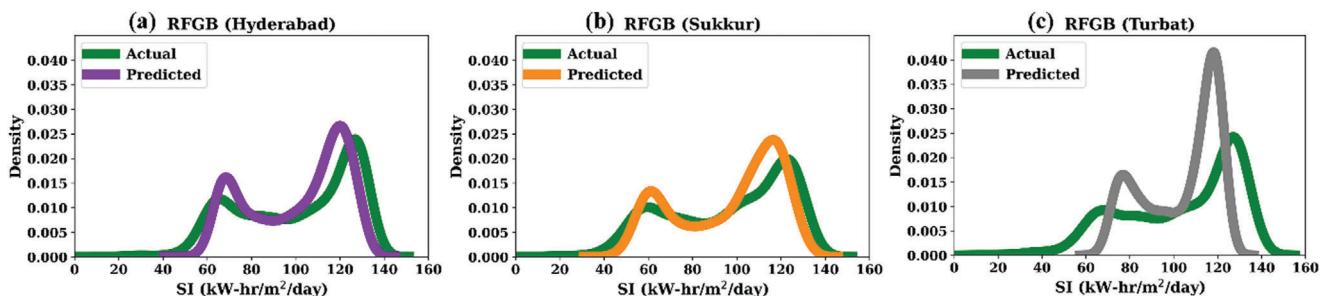


Figure 6. KDE of RFGB Predictions a) Hyderabad b) Sukkur c) Turbat.

which could involve more complex or localized approaches to data analysis and model training. These findings affirm the potential of the RFGB model as a sophisticated tool in the field of SI forecasting, with a proven capability to adapt to and capture the unique characteristics of varied regional climates.

3.3. Model Assessment

Figure 7 presents a detailed comparative analysis of the SDANN and RFGB models, employing a segment of test data for in-depth examination and the entire test dataset for frequency analysis. This approach ensures a thorough evaluation of the models'

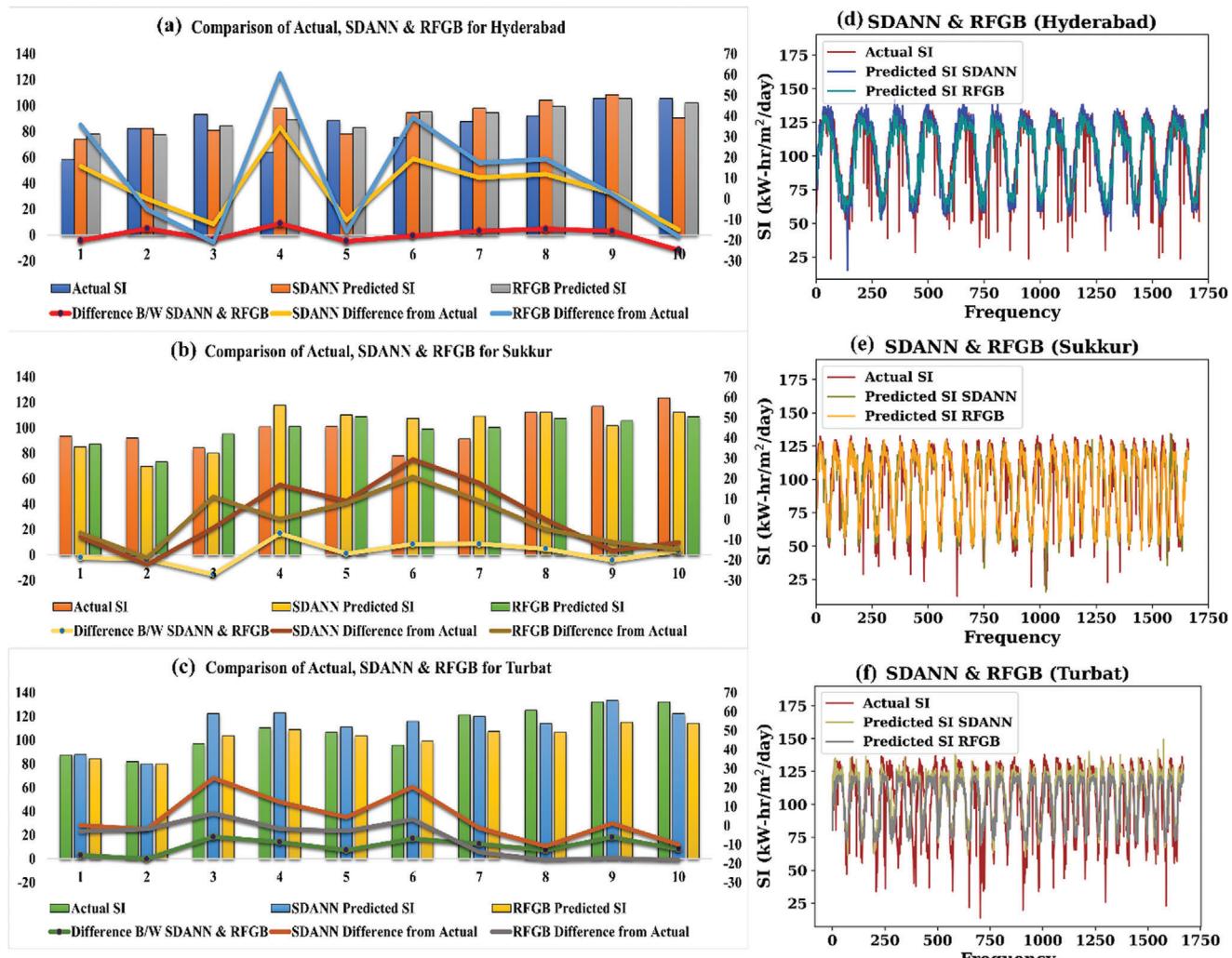


Figure 7. Bar Graph Comparison of Actual and Predicted SI: a) Hyderabad, b) Sukkur, and c) Turbat. Frequency Plot of Actual and Predicted SI: d) Hyderabad, e) Sukkur, and f) Turbat.

predictive performance for SI across Hyderabad, Sukkur, and Turbat.

In Hyderabad (Figure 7a), the SDANN model's predictions ranged from 15.1 to 145.0 kWh m⁻² day⁻¹, with an average prediction of 101.7, closely aligning with the actual SI values that spanned 14.9 to 136.0 kWh m⁻² day⁻¹ and an average of 101. The slight average difference of 0.002 suggests a high degree of accuracy. The RFGB model, meanwhile, predicted SI values within a narrower range of 53.2 to 132.2 kWh m⁻² day⁻¹, averaging at 100.8, resulting in a marginal average difference of -0.00043. The RFGB model's predictions exhibited a superior accuracy for Hyderabad, with a maximum and minimum difference of 97.93 and -41, respectively. The bar graphs reflect a close match between predicted and actual values, while the frequency plot (Figure 7d) depicts the consistency of both SDANN and RFGB models over actual data.

Sukkur's predictions (Figure 7b) indicated that both models performed similarly well, with SDANN's predicted SI values ranging from 15.8 to 134.1 kWh m⁻² day⁻¹, averaging 95.0, and RFGB's ranging from 46.7 to 129.7 kWh m⁻² day⁻¹, averaging 96.5. The actual SI values for Sukkur recorded a maximum of 133.8, a minimum of 12.4, and an average of 96.4 kWh m⁻² day⁻¹. The SDANN model had an average difference of -0.0021, while RFGB showed a slight variance of 0.0003, with maximum and minimum differences of 69.4 and -48.7, respectively. The scatter plots demonstrate RFGB's minor predictive deviations, especially at the tails of the distribution, while the frequency plot (Figure 7e) showcases both models' close prediction frequencies around the actual SI values.

In the case of Turbat (Figure 7c), the SDANN model exhibited a prediction range for SI values from 60.2 to 149.5 kWh m⁻² day⁻¹, with an average prediction at 101.7, whereas the RFGB model demonstrated predictions within a range from 68.8 to 124.8 kWh m⁻² day⁻¹, averaging at 103.5. The recorded actual SI values were 13.8, 103.9, and 137.7 kWh m⁻² day⁻¹, respectively. Here, the SDANN model showed an average difference of 0.005 from the actual values, closely followed by the RFGB model with a marginally different average difference of -0.0006. Despite the slight variance in their average differences, both models displayed commendable performance in predicting SI for Turbat, with the RFGB model illustrating a narrower gap in the maximum and minimum differences of 83.4 and -57.6, respectively. The frequency plot (Figure 7f) for Turbat indicates a significant overlap in predictions by both models with the actual SI values, highlighting their comparable accuracy in forecasting SI.

Collectively, the bar and frequency plots across all regions showcase the comparable prediction capabilities of both the SDANN and RFGB models. While the SDANN model provides valuable insights and exhibits commendable accuracy, the RFGB model also demonstrates high-quality predictions, particularly in the alignment of its forecasts with the actual SI values. These findings highlight the potential of both models as reliable and effective tools for SI forecasting, each contributing significantly to the domain of solar energy management and planning. The results emphasize the utility of integrating diverse modeling approaches to enhance the accuracy and reliability of solar forecasting systems.

3.4. Comparative Analysis

In this research, we have undertaken an exhaustive comparative analysis of various machine learning models for the prediction of SI. Our study benchmarks the performance of our proposed models against a selection of four different algorithms previously suggested by other researchers. This comparison encompasses a broad spectrum of models, including XGBOOST,^[106] Cat Boost,^[107] GRU,^[32] and a hybrid model KNN-XGBOOST.^[108] Each model has been rigorously evaluated for its efficiency and precision in predicting SI. It is pertinent to highlight that during this comparative analysis, adjustments were made to the hyperparameters of these models to better align them with the specific characteristics of our dataset, as detailed in Appendix 2 (supplementary data). These customizations were essential to ensure that the comparison is not only equitable but also pertinent to the unique aspects of our dataset and the specific objectives of our study. The aim of these modifications was to provide a transparent and unbiased evaluation of how our newly proposed models, SDANN and RFGB, compare with established algorithms in the field. **This comparison using statistical indicators (Appendix 3 (supplementary data)) as illustrated in Figure 8 serves to validate the effectiveness and appropriateness of our models for SI prediction.**

Examining the Hyderabad dataset, we observe that while the XGBOOST model incurred the highest error with an MSE of 246.04, the CAT Boost model demonstrated a commendable precision with the lowest MSE of 165.33. The RFGB and SDANN models showcased their predictive robustness with very competitive MSE values of 176.21 and 177.98, respectively. This trend is mirrored in the RMSE and MAE values, where both models are in close contest, but the R² scores stand out, jointly reflecting a high correlation with actual SI values at 0.74 for both RFGB and SDANN, affirming their reliability.

The Sukkur dataset presents a similar narrative. Here, the RFGB model excels, not only improving upon its Hyderabad performance with an MSE of 147.22 but also surpassing CAT Boost's MSE of 158.45. The SDANN follows closely, ensuring tight competition with an MSE of 152.91. Both models maintain their consistency in accuracy as reflected in the RMSE and MAE scores. The RFGB model edges ahead slightly with an R² of 0.80, indicating a marginally stronger predictive accuracy than the SDANN's also impressive 0.79.

When shifting focus to the Turbat dataset, the metrics interestingly shift. Here, although the XGBOOST and CAT Boost models report lower MSE and RMSE values, their R² scores of 0.36 and 0.71, respectively, do not necessarily translate to the most reliable predictions. The RFGB and SDANN models, with closer R² scores of 0.43 and 0.44, underscore their comparable performance in this challenging environment. Despite a higher MSE and RMSE, which could be attributed to Turbat's unique climatic factors, the consistency of RFGB and SDANN in the R² scores suggests a resilience in their predictive ability.

As the scatter graphs of Figure 9 provide a visual testament to the performance of the forecasting models within the Sukkur dataset, the prowess of the RFGB and SDANN models becomes readily apparent. These graphs show that for the RFGB and SDANN models, the data points on the scatter plots closely

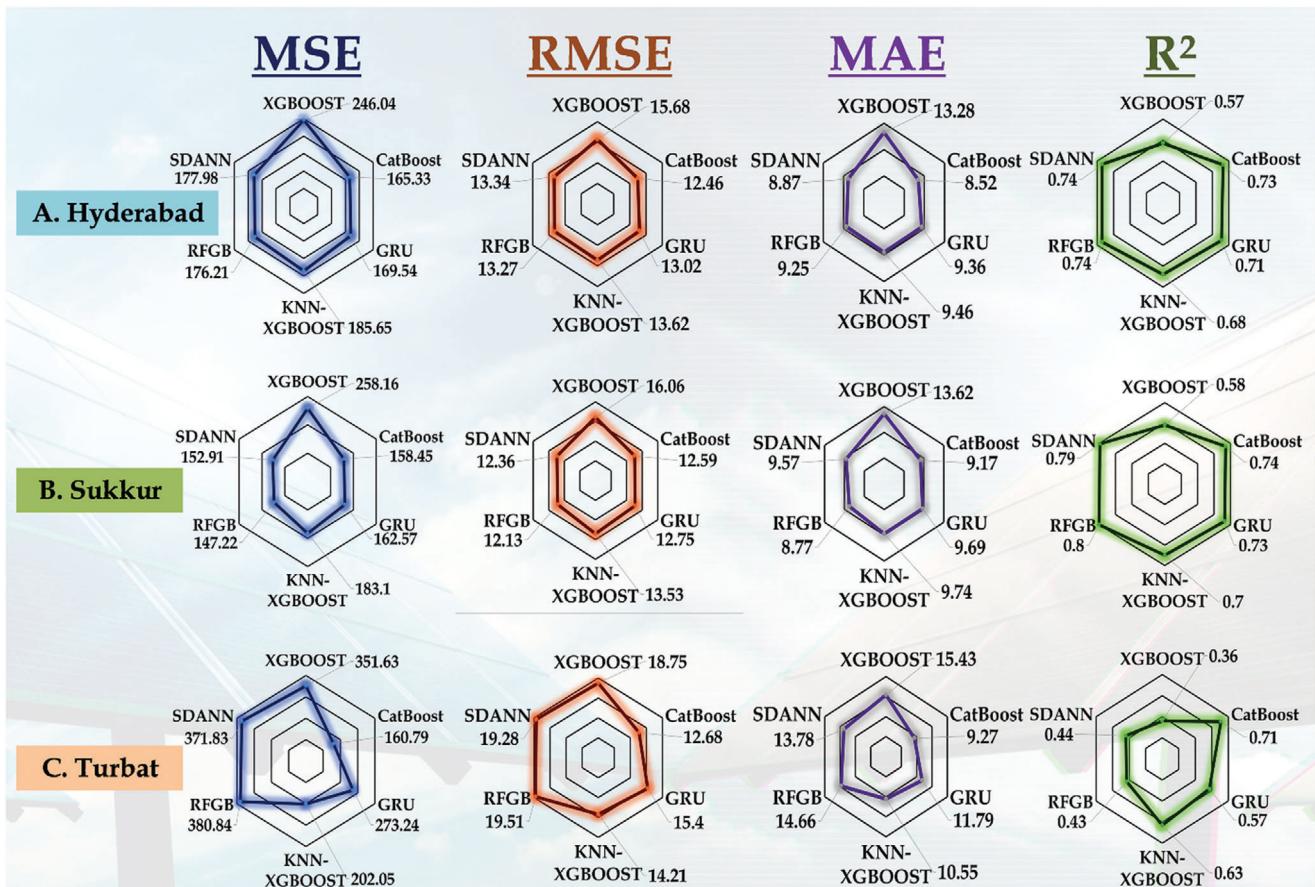


Figure 8. Statistical validation of compared Models a) Hyderabad dataset, b) Sukkur dataset, and c) Turbat dataset.

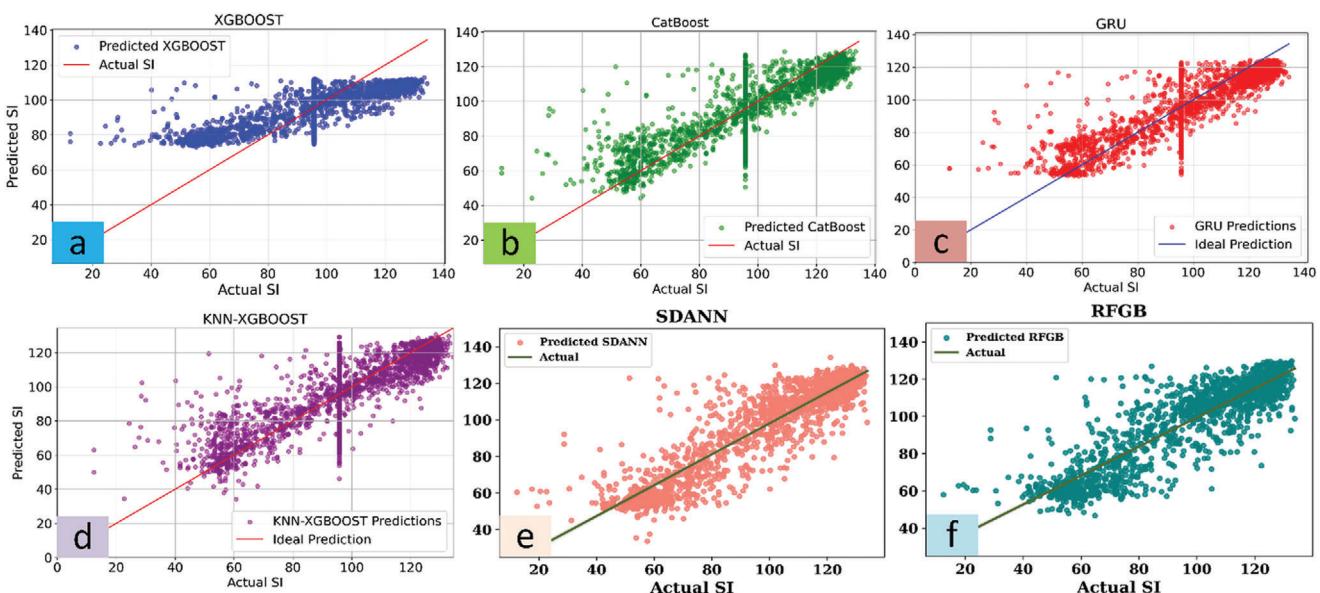


Figure 9. Scatter plot of predicted over actual SI for Sukkur dataset a) XGBOOST, b) Cat Boost, c) GRU, d) KNN-XGBOOST, e) SDANN, and f) RFGB.

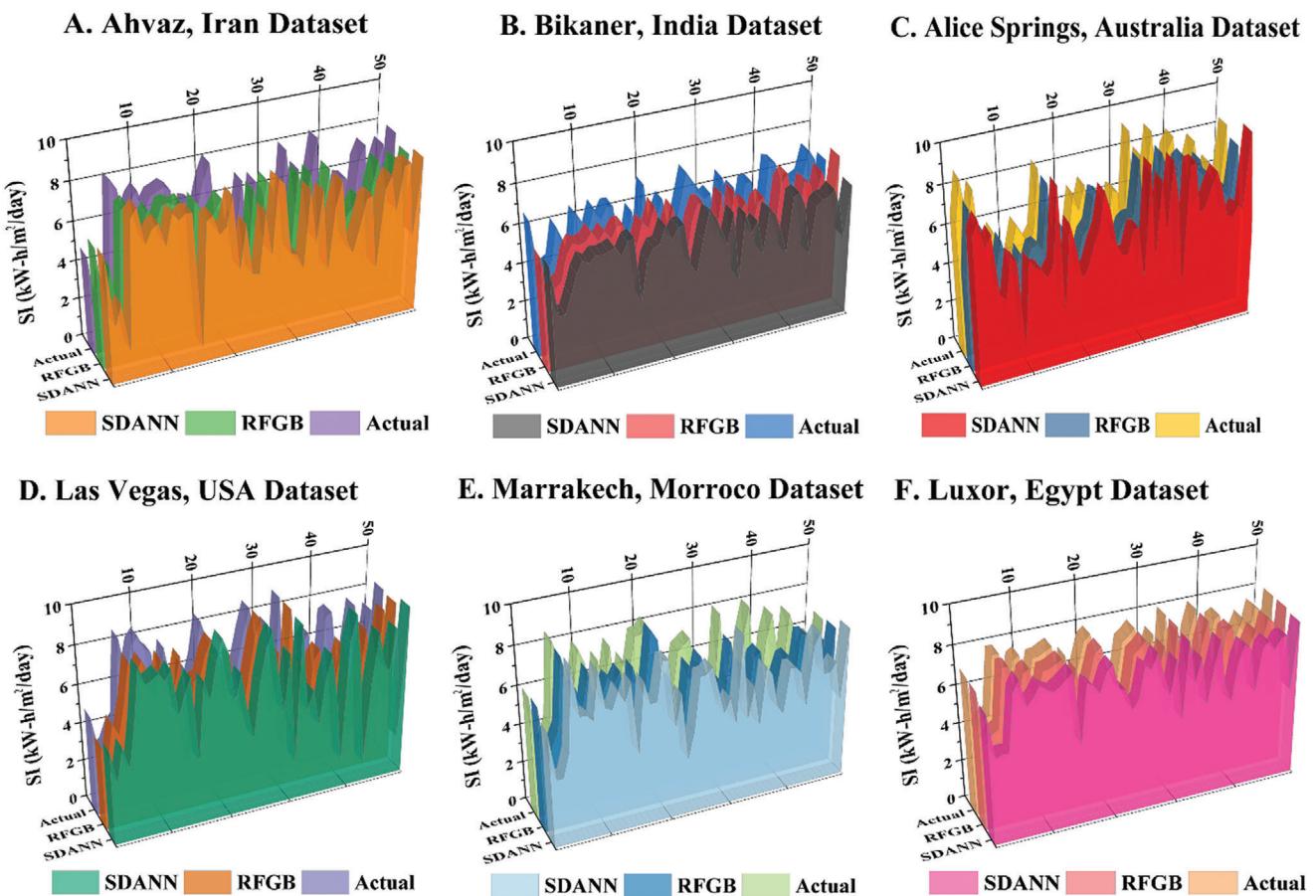


Figure 10. Comparison Plot of Actual and Predicted SI across various geographies: a) Ahvaz, b) Bikaner c) Alice Springs, d) Las Vegas, e) Marrakech, and f) Luxor.

adhere to the line of ideal prediction, underscoring their accuracy in capturing the true SI values in the Sukkur region. The proximity of the scatter points to this line in the RFGB plot is particularly indicative of its precision.

In contrast, the scatter plots for the XGBOOST, CAT Boost, GRU, and KNN-XGBOOST models reveal a more scattered distribution of points away from the ideal prediction line. This visual spread is more pronounced in the Turbat dataset plots, where prediction variance is most noticeable, reflecting these models' struggle to match the accuracy seen with the RFGB and SDANN models.

The scatter plots depicted in Figure 9 offer a clear visual comparison of the predictive success across various models within the Sukkur dataset. The RFGB and SDANN models demonstrate superior performance, as shown by the data points' close alignment to the line of perfect prediction, particularly for the RFGB model which shows slightly better precision as evidenced by lower error rates and higher R^2 values. These scatter plots provide compelling evidence of the efficacy of the RFGB and SDANN models in the field of SI forecasting. For a more comprehensive understanding, the scatter plots for the Hyderabad and Turbat datasets are made available in the supplementary data. These plots reinforce the strong performance of the RFGB and SDANN models,

not only in Sukkur but also in Hyderabad, where they both exhibit high accuracy levels.

While the RFGB and SDANN models stand out as the more reliable models for solar forecasting, the varying degrees of performance across different models and regions, especially Turbat, highlights the ongoing need for model enhancements and the possible integration of localized factors to refine their forecasting capabilities.

3.5. Evaluating the Predictive Accuracy of SI Forecasting Models Across Diverse Geographies

Upon close examination of the SI forecasting capabilities of the SDANN and RFGB models across varied geographical datasets, discernible patterns and differences in model performance emerge. The 3D bar plots, illustrated in Figure 10a-f, serve as a visual representation of these comparisons and facilitate a quantitative analysis through various performance metrics in Table 4.

Starting with the Ahvaz, Iran dataset (Figure 10a), the SDANN model registers a robust R^2 score of 0.83, signifying a strong correlation with the actual SI measurements. In contrast, the

Table 4. Statistical validation of proposed models across various geographies.

Data Set	Model	R^2	MAE	MSE	RMSE
Ahvaz, Iran	SDANN	0.83	0.594	0.666	0.816
	RFGB	0.79	0.702	0.834	0.913
Bikaner, India	SDANN	0.71	0.560	0.546	0.739
	RFGB	0.70	0.568	0.563	0.750
Alice Springs, Australia	SDANN	0.72	0.643	0.835	0.914
	RFGB	0.66	0.768	1.029	1.014
Las Vegas, USA	SDANN	0.84	0.635	0.705	0.839
	RFGB	0.80	0.735	0.883	0.939
Marrakech, Morocco	SDANN	0.75	0.676	0.804	0.897
	RFGB	0.75	0.710	0.831	0.912
Luxor, Egypt	SDANN	0.85	0.407	0.306	0.553
	RFGB	0.78	0.514	0.454	0.674

RFGB model, with an R^2 of 0.79, shows a comparable but slightly less correlation. The SDANN's precision is further evidenced by lower MAE, MSE, and RMSE values of 0.594, 0.666, and 0.816, respectively, as compared to the RFGB's 0.702, 0.834, and 0.913. The pattern of model performance carries through to the Bikaner, India dataset (Figure 10b), where both models showcase a closer R^2 value (0.71 for SDANN and 0.70 for RFGB) and near-identical MAE, MSE, and RMSE scores, suggesting a nearly equivalent predictive capability in this region.

Moving to the Alice Springs, Australia dataset (Figure 10c), the SDANN maintains its leading position with a higher R^2 score of 0.72 against the RFGB model's 0.66, indicating a more consistent fit with the actual data. The SDANN's MAE, MSE, and RMSE figures—0.643, 0.835, and 0.914—are lower than those of the RFGB model, which are 0.768, 1.029, and 1.014, respectively. However, as we shift our focus to the Las Vegas, USA dataset (Figure 10d), while the SDANN model still leads with an R^2 of 0.84 over RFGB's 0.80, the difference in MAE, MSE, and RMSE is modest, underscoring the competitive nature of the RFGB model in this American context.

In the Marrakech, Morocco dataset (Figure 10e), both models record an R^2 of 0.75, indicating their similar proficiency in capturing the variation in SI values. The SDANN model has marginally better MAE and MSE values (0.676 and 0.804) compared to the RFGB's 0.710 and 0.831, while the RMSE scores are closely matched at 0.897 and 0.912, respectively. This suggests that either model could be deployed with nearly equal effectiveness in the Moroccan setting. The trend continues with the Luxor, Egypt dataset (Figure 10f), where the SDANN outshines with the highest R^2 score observed at 0.85 and notably lower MAE, MSE, and RMSE values of 0.407, 0.306, and 0.553. The RFGB model, while trailing with an R^2 of 0.78 and error metrics of 0.514, 0.454, and 0.674, still represents a viable alternative for forecasting in the Egyptian climate.

Through this analytical narrative, it becomes evident that while the SDANN model frequently exhibits superior accuracy across the datasets, the RFGB model remains a strong contender, showcasing commendable predictive quality and of-

ferring a reliable alternative in the advanced domain of solar forecasting.

4. Ablation Studies

4.1. Feature Impact Analysis

An ablation study is conducted involving both the SDANN and the RFGB models, and a comprehensive investigation into feature^[109] set impacts on SI forecasting is conducted across three geographically distinct regions: Hyderabad, Sukkur, and Turbat detailed in Appendices 4–9 (Supporting Information). The study delves into the efficacy of different feature subsets, culminating in a comparative assessment that underscores the superior predictive capability of the full feature set configuration.

The SDANN model's feature analysis for the Hyderabad (Figure 11a) dataset began with temperature statistics as the baseline input features, yielding an R^2 of 0.71. The addition of PS initially showed a decrease in performance, with the R^2 dipping to 0.69. Yet, as the feature set expanded, the model's performance improved, particularly with the inclusion of RH and WS. The complete feature set achieved an R^2 of 0.74, MAE of 8.87, and an MSE of 177.98, indicating a robust predictive framework when utilizing all available variables.

In the Sukkur dataset (Figure 11b), the SDANN model presented a similar trend, with the full feature set reaching an R^2 of 0.79. However, a notable anomaly emerged when the subset [T_s , T_{max} , T_{min} , WS, RH, PS] was examined, where the R^2 plummeted to -12.74, suggesting an aberration in the data or model overfitting. This highlights the need for careful feature engineering and model validation to ensure the reliability of predictive insights.

The Turbat dataset (Figure 11c) revealed the challenges inherent in SI prediction for the region, with the highest R^2 , attained being 0.44 for the SDANN model using all features. The adverse R^2 values observed with certain feature subsets, such as -16.14 for the subset including PS, indicate extreme model sensitivity to the feature set composition and a potential overestimation of the model's complexity, necessitating a more refined approach to model tuning and validation.

Comparatively, the RFGB model exhibited consistent improvement in performance with the inclusion of additional features across all datasets. For Hyderabad (Figure 11d), the incremental feature inclusion led to a maximum R^2 of 0.74, and for Sukkur (Figure 11e), the full feature set achieved an R^2 of 0.80. However, similar to the SDANN, the Turbat dataset (Figure 11f) posed a greater challenge, with a maximum R^2 of 0.43, indicating room for model enhancement or alternative predictive strategies.

Overall, Figure 11 of feature analysis for both SDANN and RFGB models, clearly illustrates that the use of a comprehensive suite of input features generally provides the best results in SI forecasting. This finding emphasizes the importance of utilizing a full range of available environmental data to improve model accuracy. Nevertheless, the anomalies and challenges noted, especially in the Turbat dataset, indicate that a one-size-fits-all approach may not be universally applicable. Thus, model customization and further exploration of feature interactions remain pivotal for advancing SI forecasting in varied geographical contexts.

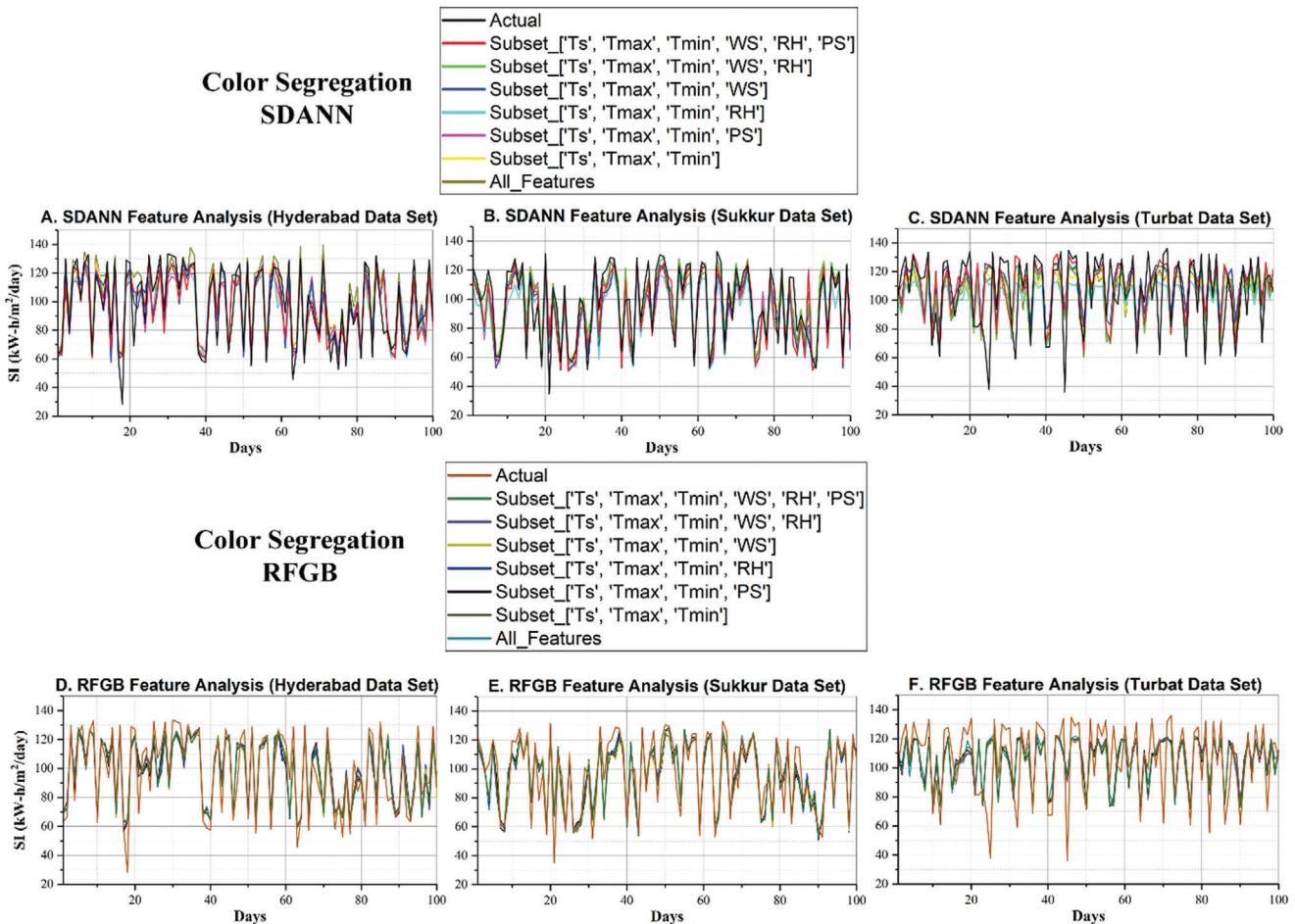


Figure 11. Line Plot for Feature Impact analysis; a) SDANN Hyderabad, b) SDANN Sukkur, c) SDANN Turbat, d) RFGB Hyderabad, e) RFGB Sukkur, and e) RFGB Turbat.

4.2. SHAP Analysis

Building on the ablation studies, the SHapley Additive exPlanations (SHAP)^[110] analysis provides a deeper interpretative layer to the RFGB model's functionality by measuring the impact of each feature on the model's predictions for SI. SHAP analysis is a sophisticated method used to interpret machine learning models.^[111] It explains the output of any model by computing the contribution of each feature to a prediction. SHAP values provide an understanding of how much each predictor variable impacts the model's output. In essence, SHAP values can explain the prediction of an instance by quantifying the contribution of each feature to the prediction.

The SHAP values, as visualized in the plot for Sukkur dataset in **Figure 12**, reflect the magnitude and direction of a feature's impact on the model output. Features that push the prediction higher are represented with positive SHAP values, while those that lower the prediction are indicated by negative values. The color gradient from blue to pink denotes the low to high range of the feature values.

In this analysis, we observe that temperature-related features— T_{\min} , T_{\max} , and T_s —have a mixed influence on the model's predictions. T_{\max} particularly stands out with its

predominantly positive SHAP values, indicating that higher maximum temperatures are usually associated with higher predicted values of SI, which is consistent with the general understanding of SI dynamics.

The PS feature predominantly displays positive SHAP values, suggesting its critical role in increasing the model's SI predictions. SH and RH display a complex interaction with the SI

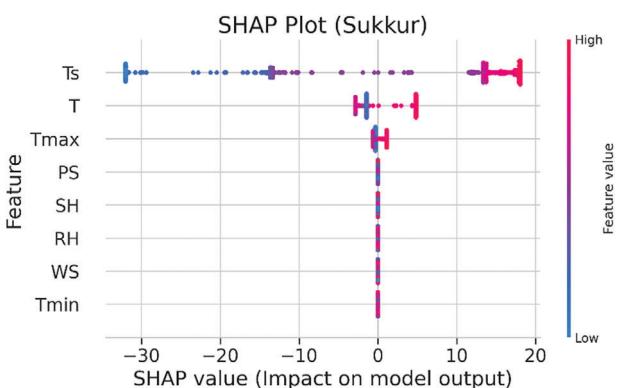


Figure 12. SHAP Analysis Plot for Sukkur Dataset.

predictions. Their SHAP values are distributed across both the positive and negative spectrum, suggesting that the relationship between humidity and SI is not straightforward and may depend on the interaction with other factors within the model. WS, with its cluster of points leaning toward the negative side of the SHAP value spectrum, suggests an inverse relationship with SI predictions. This could imply that in the context of Sukkur, higher wind speeds might correlate with factors that reduce SI, such as increased cloud cover or atmospheric disturbances.

The distribution of SHAP values across the features not only reveals the individual impact of each feature but also hints at the underlying complexity of their interactions. For example, tightly grouped SHAP values indicate a consistent impact of a feature on the model output, while a wider dispersion suggests variability in influence, which could be due to the interplay with other features in the model.

This SHAP analysis is instrumental in understanding the contribution of each feature within the RFGB model and thus can be leveraged to fine-tune the model for even greater accuracy in SI prediction. The insights drawn from this analysis are especially useful for stakeholders in the solar energy sector, providing a clearer picture of the factors that significantly influence SI.

5. Discussion and Analysis

Our investigation into SI forecasting through advanced machine learning models, specifically the RFGB and SDANN models, underscores a pivotal advancement in predictive capabilities across various geographical locations. These models, set against the backdrop of a comprehensive comparison with contemporaries such as XGBOOST, CatBOOST, GRU, and KNN-XGBOOST, alongside thorough ablation studies, have elucidated a rich landscape of insights.

The nuanced calibration of the RFGB model, balancing complexity with the imperative to avoid overfitting, has been instrumental in its standout performance. This precision in model design is mirrored in its adeptness at capturing intricate environmental patterns, which is particularly notable in datasets where these characteristics are well-represented. Similarly, the SDANN model, with its deep neural network architecture, has shown an impressive aptitude for accuracy across diverse settings, evidencing the power of advanced deep learning in SI forecasting. The scatter plots in Figure 9 and supplementary data for the Hyderabad and Turbat datasets visually reinforce the high degree of predictive accuracy achieved by both models, particularly within the Sukkur dataset where their data points closely align with the line of ideal prediction.

Ablation studies further illuminate the significant impact of deliberate feature selection on the RFGB model's performance, showcasing systematic improvements in accuracy as the feature set expanded. This meticulous approach to model refinement reveals the importance of aligning model inputs with the nuanced dynamics of climatic factors, yet also highlights areas where enhancements are needed, particularly in environments with unique complexities.

Despite the demonstrated strengths, it is imperative to acknowledge certain limitations. Tailoring these models to specific regional climates might necessitate adjustments, a process that is less a limitation and more an opportunity for optimization

and regional customization. Additionally, the complexity inherent in these sophisticated models, especially the RFGB, introduces computational demands that could challenge resource constraints.

Future directions could involve exploring more diverse datasets, particularly for regions presenting complex meteorological conditions, to further refine model accuracy. Advanced feature engineering and the potential application of techniques such as transfer learning to the SDANN model may offer new avenues for enhancing predictive versatility and effectiveness.

In essence, this research not only highlights the robust predictive performance of the RFGB and SDANN models but also marks a significant contribution to the domain of renewable energy forecasting. The insights gleaned from this study serve as a foundation for future innovations in technology and sustainable energy policy, driving forward the capabilities of SI forecasting in support of renewable energy development and management.

6. Conclusion

The comprehensive analysis presented in this research article delineates the efficacy of advanced deep learning models for SI forecasting, underscored by quantitative assessments derived from extensive data across three different datasets. Our investigation leveraged the SDANN and RFGB models, revealing significant insights into their predictive capacities and comparing them with established models like XGBOOST, CAT Boost, GRU, and KNN-XGBOOST. Additionally, the ablation study conducted further illuminated the impact of various features on the models' performance. The keynote conclusions are as follows:

- 1) RFGB model showcased exceptional predictive capabilities, distinguishing itself with remarkable efficiency across different datasets. Notably, the Sukkur dataset, achieved an impressive R^2 value of 0.80, alongside the lowest Mean Squared Error (MSE) of 147.22 and Mean Absolute Error (MAE) of 8.77.
- 2) SDANN model, renowned for its intricate architecture, demonstrated high precision in SI forecasting, particularly excelling in the Hyderabad dataset where it mirrored actual SI values with significant accuracy. With a compelling R^2 value of 0.74, akin to RFGB's, and closely matched MSE and MAE values in regions such as Hyderabad and Sukkur, SDANN proved its adeptness in capturing the nuanced dynamics of SI.
- 3) Geographical validations across Ahvaz, Bikaner, Alice Springs, Las Vegas, Marrakech, and Luxor showcased the adaptability of both SDANN and RFGB models to diverse climates, underlining their predictive accuracy's sensitivity to regional factors. The SDANN model, in particular, demonstrated its versatility, despite variations in performance indicators such as R^2 and MAE, emphasizing the importance of tailoring models to specific geographical conditions.
- 4) Comparative analysis revealed the SDANN and RFGB models' strengths against traditional forecasting methods like XGBOOST, CAT Boost, GRU, and KNN-XGBOOST. Specifically, RFGB demonstrated a slight edge in performance, with both models surpassing others in metrics such as R^2 values. For example, in certain datasets, RFGB achieved an R^2 value of 0.80, closely followed by SDANN, showcasing their superior

- handling of complex data interplays compared to XGBOOST, CAT Boost, GRU, and KNN-XGBOOST.
- 5) Ablation studies provided critical insights into the impact of feature selection on model performance. For the RFGB model, the systematic addition of features such as temperature statistics, atmospheric pressure (PS), relative humidity (RH), and wind speed (WS) incrementally improved its accuracy. In particular, the inclusion of all selected features elevated the model's R^2 to 0.762 in Hyderabad and to an optimal 0.801 in Sukkur, affirming the pivotal role of comprehensive feature analysis in bolstering predictive precision.

The development and evaluation of the RFGB and SDANN models significantly advance SI forecasting, essential for improving solar energy predictions and optimizing power generation. Both models demonstrated robust performances, with the RFGB model excelling in adaptability and accuracy across diverse conditions, and the SDANN model showcasing deep learning strengths, albeit with room for improvement in complex environments. These findings highlight the need for model customization to regional climates and suggest exploring broader datasets and advanced techniques to enhance future model efficiency and adaptability. This research contributes valuable insights into SI forecasting, with implications for global renewable energy strategies.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in POWER | Data Access Viewer at <https://power.larc.nasa.gov/data-access-viewer/>, reference number 0.

Keywords

artificial neural network (ANN), deep learning (DL), gradient Boosting (GB), machine learning (ML), random Forest (RF), solar forecasting

Received: December 30, 2023

Revised: April 7, 2024

Published online:

- [4] B. Zakeri, K. Paulavets, L. Barreto-Gomez, L. G. Echeverri, S. Pachauri, B. Boza-Kiss, C. Zimm, J. Rogelj, F. Creutzig, D. Ürge-Vorsatz, D. G. Victor, M. D. Bazilian, S. Fritz, D. Gielen, D. L. McCollum, L. Srivastava, J. D. Hunt, S. Pouya, *Energies (Basel)* **2022**, 15, 61141.
- [5] F. Sher, O. Curnick, M. T. Azizan, *Sustainability* **2021**, 13, 1.
- [6] M. Mele, A. R. Gurrieri, G. Morelli, C. Magazzino, *Environ. Sci. Pollut. Res. Int.* **2021**, 28, 41127.
- [7] M. Shahbaz, B. A. Topcu, S. S. Sarigül, X. V. Vo, *Renew Energy* **2021**, 178, 1370.
- [8] M. J. B. Kabeyi, O. A. Olanrewaju, *Front. Energy Res.* **2022**, 9.
- [9] N. Alsagr, S. van Hemmen, *Environ. Sci. Pollut. Res. Int.* **2021**, 28, 25906.
- [10] Y. Demirel, *Energy* **2016**, 441.
- [11] J. Wang, J. Yao, H. Chen, X. Li, J. Lin, *Compar. Analysis Power Load Forecast. ML Meth.* **2023**, 12604, 126042.
- [12] A. S. Elberawi, M. Belal, *Future Comput. Informat. J.* **2021**, 6, 45.
- [13] T. Ahmad, D. Zhang, C. Huang, *Energy* **2021**, 231, 120911.
- [14] F. Gerges, M. C. Boufadel, E. Bou-Zeid, H. Nassif, J. T. L. Wang, *Proced. 2023 7th Int. Conf. Mach. Learn. Soft Comput.* **2023**, 102.
- [15] P. Kumari, D. Toshniwal, *J. Clean Prod.* **2021**, 318, 128566.
- [16] F. Gerges, M. C. Boufadel, E. Bou-Zeid, H. Nassif, J. T. L. Wang, *Knowl. Inf. Syst.* **2024**, 66, 613.
- [17] D. Harada, P. Chinnavornrungrungsee, S. Kittisontirak, N. Chollacoop, S. Songtrai, K. Sriprapha, J. Yoshino, T. Kobayashi, *Jpn. J. Appl. Phys.* **2023**, 62, SK1056.
- [18] R. A. A. Ramadhan, Y. R. J. Heatubun, S. F. Tan, H. J. Lee, *Renew Energy* **2021**, 178, 1006.
- [19] P. Seshadri, B. P. Bagavat, A. Kumar B, K. H., K. G., S. S., *Electric Power Comp. Syst.* **2021**, 49, 1379.
- [20] S. Ullah, Z. Zuo, F. Zhang, J. Zheng, S. Huang, Y. Lin, I. Iqbal, Y. Sun, M. Yang, L. Yan, *Remote Sens (Basel)* **2020**, 12, 1.
- [21] S. Gbémou, J. Eynard, S. Thil, E. Guillot, S. Grieu, *Energies* **2021**, 4, 3192.
- [22] A. K. Mandal, R. Sen, S. Goswami, B. Chakraborty, *Symmetry (Basel)* **2021**, 13.
- [23] M. Golam, R. Akter, J. M. Lee, D. S. Kim, *IEEE Geosci. Remote Sens. Lett.* **2022**, 19, 1.
- [24] A. T. Dahiru, C. W. Tan, A. L. Bukar, K. Y. Lau, C. L. Toh, S. Salisu, *2021 IEEE Conf. Energy Convers. (CENCON)* **2021**, 214.
- [25] M. Viswavandy, B. Sarangi, S. Mohanty, A. Mohanty, *Adv. Intell. Syst. Comput.* **2020**, 990, 751.
- [26] M. A. Hassan, N. Bailek, K. Bouchouicha, S. C. Nwokolo, *Renew Energy* **2021**, 171, 191.
- [27] I. Sansa, Z. Boussaada, N. M. Bellaaj, *Energies (Basel)* **2021**, 14, 6920.
- [28] C. McHugh, S. Coleman, D. Kerr, *Mach. Learn. Appl.* **2022**, 9, 100383.
- [29] I. Iqbal, I. Ahmed, *ICESP* **2015**, 2014, 1.
- [30] M. Massaoudi, I. Chihi, L. Sidhom, M. Trabelsi, S. S. Refaat, H. Abu-Rub, F. S. Oueslati, *IEEE Access* **2021**, 9, 36571.
- [31] H. Gao, S. Qiu, J. Fang, N. Ma, J. Wang, K. Cheng, H. Wang, Y. Zhu, D. Hu, H. Liu, J. Wang, *Sustainability* **2023**, 15.
- [32] A. Mellit, A. M. Pavan, V. Lugh, *Renew Energy* **2021**, 172, 276.
- [33] R. Kim, K. M. Kim, J. H. Ahn, *J. Korean Soc. Environm. Engin.* **2021**, 43, 347.
- [34] N. M. Sabri, M. El Hassouni, *Energ. Sourc., Part A: Recov., Utilizat., Environ. Effects* **2022**, 44, 6303.
- [35] Y. Ledmaoui, A. El Fahli, A. Elmaghraoui, M. El Aroussi, R. Saadane, A. Chehri, *2023 Int. Conf. Digital Age Technol. Adv. Sustainable Develop. (ICDATA)* **2023**, 120.
- [36] M. F. Hanif, M. S. Naveed, M. Metwaly, J. Si, X. Liu, J. Mi, M. F. Hanif, M. S. Naveed, M. Metwaly, J. Si, X. Liu, J. Mi, *AIMS Energy* **2024**, 12, 350.

- [1] J. L. Holechek, H. M. E. Geli, M. N. Sawalhah, R. Valdez, *Sustainability* **2022**, 14.
- [2] S. A. Qadir, H. Al-Motairi, F. Tahir, L. Al-Fagih, *Energy Reports* **2021**, 7, 3590.
- [3] D. Bogdanov, M. Ram, A. Aghahosseini, A. Gulagi, A. S. Oyewo, M. Child, U. Caldera, K. Sadovskaya, J. Farfan, L. De Souza Noel Simas Barbosa, M. Fasihi, S. Khalili, T. Traber, C. Breyer, *Energy* **2021**, 227, 120467.

- [37] P. Pawar, N. Mithulanthan, M. Q. Raza, *GPECOM* **2020**, 2020, 226.
- [38] X. Huang, Q. Li, Y. Tai, Z. Chen, J. Zhang, J. Shi, B. Gao, W. Liu, *Renew Energy* **2021**, 171, 1041.
- [39] K. Anuradha, D. Erlapally, G. Karuna, V. Srilakshmi, K. Adilakshmi, *E3S Web Conf.* **2021**, 309, 01163.
- [40] M. Kumar, K. Namrata, N. Kumar, G. Saini, *J. Grid Comput.* **2023**, 21, 28.
- [41] S. Boubaker, M. Benghanem, A. Mellit, A. Lefza, O. Kahouli, L. Kolsi, *IEEE Access* **2021**, 9, 36719.
- [42] S. Nosouhian, F. Nosouhian, A. K. Khoshouei, A Review of Recurrent Neural Network Architecture for Sequence Learning: Comparison between LSTM and GRU, <https://www.preprints.org/manuscript/202107.0252/v1>.
- [43] C. H. Liu, J. C. Gu, M. T. Yang, *IEEE Access* **2021**, 9, 17174.
- [44] A. Rai, A. Shrivastava, K. C. Jana, *Optik (Stuttgart)* **2022**, 252, 168515.
- [45] H. Raju, S. Das, *Solar Physics* **2021**, 296, 1.
- [46] M. Jaihuni, J. K. Basak, F. Khan, F. G. Okyere, T. Sihalath, A. Bhujel, J. Park, D. H. Lee, H. T. Kim, *ISA Trans* **2022**, 121, 63.
- [47] P. Kumari, D. Toshniwal, *Appl. Energy* **2021**, 295, 117061.
- [48] S. Ziyabari, Z. Zhao, L. Du, S. K. Biswas, *IEEE Trans Ind Appl* **2023**, 59, 5293.
- [49] E. M. Al-Ali, Y. Hajji, Y. Said, M. Hleili, A. M. Alanzi, A. H. Laatar, M. Atri, *Mathematics* **2023**, 11, 676.
- [50] S. Ghimire, R. C. Deo, N. Raj, J. Mi, *Appl. Energy* **2019**, 253, 113541.
- [51] H. Zang, L. Liu, L. Sun, L. Cheng, Z. Wei, G. Sun, *Renew. Energy* **2020**, 160, 26.
- [52] B. Gu, X. Li, F. Xu, X. Yang, F. Wang, P. Wang, *Sustainability (Switzerland)* **2023**, 15.
- [53] M. Kumar, K. Namrata, N. Kumar, *Adv. Electr. Electr. Engin.* **2023**, 20, 549.
- [54] O. Jogunola, A. S. Ajagun, B. Adebisi, A. M. Aibinu, J. A. Ojo, The 5th Int. Conf. Fut. Netw. Distri. Sys. **2021**, 214.
- [55] S. Sivarajan, S. D. S. S. Jebaseelan, *Int. J. Electr. Electr. Engin.* **2023**, 10, 221.
- [56] J. Yan, Y. Sha, Y. Zhang, T. Li, J. Zhang, 2023 3rd Int. Conf. Energ., Power Electr. Engin. (EPEE) **2023**, 261.
- [57] P. Kumari, D. Toshniwal, *J. Clean. Prod.* **2021**, 318, 128566.
- [58] C. Halton, E. Rasure, *Predictive Analytics: Definition, Model Types, and Uses, Investopedia*, **2023**.
- [59] S. Manju, M. Sandeep, *J. Clean Prod.* **2019**, 230, 116.
- [60] S. Ahmad, M. Parvez, T. A. Khan, O. Khan, *Environm. Challeng.* **2022**, 9, 100634.
- [61] Ü. Ağbulut, A. E. Gürel, Y. Biçen, *Renewable Sustainable Energy Rev.* **2020**, 135, 110114.
- [62] S. Islam, N. K. Roy, *Energy Reports* **2023**, 9, 6063.
- [63] B. K. Puah, L. W. Chong, Y. W. Wong, K. M. Begam, N. Khan, M. A. Juman, R. K. Rajkumar, *Renew Energy* **2021**, 164, 908.
- [64] C. G. Villegas-Mier, J. Rodriguez-Resendiz, J. M. Álvarez-Alvarado, H. Jiménez-Hernández, Á. Odry, *Micromachines (Basel)* **2022**, 13, 1406.
- [65] O. M. Mbah, C. I. Madueke, R. Umunakwe, M. N. Agba, *J. Engin. Sci.* **2022**, 9, E1.
- [66] Weatherspark.com Compare the Climate and Weather in Sukkur, Hyderabad, and Turbat, <https://weatherspark.com/compare/y/106589~106562~106188/Comparison-of-the-Average-Weather-in-Sukkur-Hyderabad-and-Turbat>.
- [67] Solargis Global Solar Atlas 2.0 (2018) GHI Map of Pakistan Utilizing Solargis Data, Energy Sector Management Assistance Program (ESMAP), World Bank Group, **2007**, 1.
- [68] NASA, POWER Data Access Viewer v2.0.0 **2023**, <https://power.larc.nasa.gov/beta/data-access-viewer/>.
- [69] Ü. Ağbulut, A. E. Gürel, Y. Biçen, *Renew. Sustainab. Energy Rev.* **2021**, 135, 110114.
- [70] Z. Pang, F. Niu, Z. O'Neill, *Renew. Energy* **2020**, 156, 279.
- [71] Z. Wang, C. Tian, Q. Zhu, M. Huang, *Energies (Basel)* **2018**, 11, 68.
- [72] M. A. Munir, A. Khattak, K. Imran, A. Ulasyar, A. Khan, 1st Int. Conf. Electr., Commun. Comp. Engin., **2019**, 24.
- [73] R. Gallo, M. Castangia, A. Macii, E. Macii, E. Patti, A. Aliberti, *Eng Appl Artif Intell* **2022**, 116, 105493.
- [74] J. W. Emerson, W. A. Green, B. Schloerke, J. Crowley, D. Cook, H. Hofmann, H. Wickham, *J. Computat. Graph. Statis.* **2013**, 22, 79.
- [75] N. J. Cox, *Stata J* **2009**, 9, 621.
- [76] Z. Lv, W. Deng, Z. Zhang, N. Guo, G. Yan Proceed. - 2019 IEEE Intl. Conf Parallel Distrib. Process. Appl., Big Data Cloud Comput., Sustain. Comput. Commun., Soc. Comput. Netw, **2019**, 802.
- [77] C. Zhang, S. Lin, D. Zhang, *ACM Int. Conf. Proceed. Series* **2022**, 77.
- [78] P. Wang, Y. Liu, Q. Sun, Y. Bai, C. Li, *Sustainability (Switzerland)* **2022**, 14, 12286.
- [79] T. T. Le, W. Fu, J. H. Moore, *Bioinformatics* **2020**, 36, 250.
- [80] Z. Sultana, M. M. Rahman, L. Nahar, *J. Appl. Computer Sci. Mathem.* **2022**, 16, 18.
- [81] P. P. Palmes, A. Kishimoto, R. Marinescu, P. Ram, E. Daly **2021** Designing Machine Learning Pipeline Toolkit for AutoML Surrogate Modeling Optimization, arXiv preprint arXiv:2107.01253.
- [82] S. Gholizadeh, *J. Robot. Automat. Res.* **2022**, 3.
- [83] I. Stančin, A. Jović, 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE, pp. 977–982.
- [84] Welcome to Colaboratory - Colaboratory, <https://colab.research.google.com/>.
- [85] W. Wu, G. C. Dandy, H. R. Maier, *Environ. Model. Soft.* **2014**, 54, 108.
- [86] H. R. Maier, A. Jain, G. C. Dandy, K. P. Sudheer, *Environ. Model. Software* **2010**, 25, 891.
- [87] T. Vujičić, T. Matijević, J. Ljucović, A. Balota, Z. Ševarac, Central European Conference on Information and Intelligent Systems, **2016**.
- [88] P. Raut, A. Dani, *Correlat. Bet. Num. Hidden Layers Accu. Artific. Neural Net.* **2020**, 513.
- [89] I. Iqbal, G. A. Odesanmi, J. Wang, L. Liu, *Appl. Artif. Intell.* **2021**, 35, 697.
- [90] A. Holzinger, Machine Learning for Health Informatics, Springer International Publishing, **2016**, pp. 1–24.
- [91] F. Psallidas, Y. Zhu, B. Karlas, M. Interlandi, S. Krishnan, B. Kroth, V. Emani, W. Wu, C. Zhang, M. Weimer, A. Floratou, C. Curino, K. Karanasos, *ACM SIGMOD Record* **2022**, 51, 30.
- [92] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintha, S. Kundu, *IEEE Transact. Image Process.* **2018**, 27, 4012.
- [93] M. Schonlau, R. Y. Zou, *Stata Journal* **2020**, 20, 3.
- [94] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, *Artif. Intell. Rev.* **2021**, 54, 1937.
- [95] V. K. Ayyadvara in *Pro Machine Learning Algorithms*, Apress, New York **2018**, 117.
- [96] G. Biau, B. Cadre in *Advances in Contemporary Statistics and Econometrics: Festschrift in Honor of Christine Thomas-Agnan*, Springer International Publishing, Berlin **2021**, 23.
- [97] Scikit-Learn Metrics and Scoring: Quantifying the Quality of Predictions, https://scikit-learn.org/stable/modules/model_evaluation.html.
- [98] M. Guermoui, F. Melgani, K. Gairaa, M. L. Mekhalfi, *J Clean Prod* **2020**, 258, 120357.
- [99] N. Krishnan, K. R. Kumar, C. S. Inda, *J Clean Prod* **2023**, 388, 135860.
- [100] S. Sheik Mohammed, F. Titus, S. B. Thanikanti, S. M. Sulaiman, S. Deb, N. M. Kumar, *Sustainability* **2022**, 14, 3498.
- [101] Y. Wang, D. Millstein, A. D. Mills, S. Jeong, A. Ancell, *Sol. Energy* **2022**, 231, 846.
- [102] B. Belmahdi, M. Louzazni, M. Marzband, A. El Bouardi, *Forecasting* **2023**, 5, 172.

- [103] M. A. Ali, A. Elsayed, I. Elkabani, M. E. Youssef, G. E. Hassan, *Environ Dev Sustain* **2023**, 1.
- [104] Y. C. Chen, *Biostat Epidemiol* **2017**, 1, 161.
- [105] L. Benali, G. Notton, A. Fouilloy, C. Voyant, R. Dizene, *Renew Energy* **2019**, 132, 871.
- [106] X. Li, L. Ma, P. Chen, H. Xu, Q. Xing, J. Yan, S. Lu, H. Fan, L. Yang, Y. Cheng, *Energy Rep.* **2022**, 8, 1087.
- [107] J. Fan, X. Wang, F. Zhang, X. Ma, L. Wu, J. *Clean Prod.* **2020**, 248, 119264.
- [108] S. Husain, U. A. Khan, *Environ. Prog. Sustain Energy* **2022**, 41, e13782.
- [109] J. Wang, H. Zhong, X. Lai, Q. Xia, Y. Wang, C. Kang, *IEEE Trans. Smart Grid* **2019**, 10, 1417.
- [110] Y. G. Lee, J. Y. Oh, D. Kim, G. Kim, J. *Electr. Engin. Technol.* **2023**, 18, 579.
- [111] C. Cakiroglu, S. Demir, M. Hakan Ozdemir, B. Latif Aylak, G. Sariisik, L. Abualigah, *Expert. Syst. Appl.* **2024**, 237, 121464.

ADVANCED THEORY AND SIMULATIONS

Supporting Information

for *Adv. Theory Simul.*, DOI 10.1002/adts.202301289

Enhancing Solar Forecasting Accuracy with Sequential Deep Artificial Neural Network and Hybrid Random Forest and Gradient Boosting Models across Varied Terrains

Muhammad Farhan Hanif^{}, Muhammad Umar Siddique, Jicang Si, Muhammad Sabir Naveed, Xiangtao Liu and Jianchun Mi*

Supplementary Data

Enhancing Solar Forecasting Accuracy: Deep Learning and Hybrid Modeling Approaches across Varied Terrains

Muhammad Farhan Hanif^{a,b}, Muhammad Umar Siddique^b, Jicang Si^a, Muhammad Sabir Naveed^a, Xiangtao Liu^a, Jianchun Mi^a*

^aDepartment of Energy & Resource Engineering, College of Engineering, Peking University, Beijing 100871, China

^bDepartment of Mechanical Engineering, FE&T Bahauddin Zakariya University Multan- 60000, Pakistan

*Corresponding Author's e-mail address: farhanhanif@pku.edu.cn

Appendix 1. Best Hyperparameters of RF and GB across various datasets.

Data Set	Model	max_depth	min_samples_leaf	n_estimators	Best Accuracy
Hyderabad, Pakistan	RF	4	0.1	210	0.691
	GB	7	0.1	130	0.748
Sukkur, Pakistan	RF	3	0.1	140	0.714
	GB	8	0.1	130	0.773
Turbat, Pakistan	RF	2	0.1	280	0.407
	GB	4	0.1	110	0.430
Ahvaz, Iran	RF	3	0.1	280	0.648
	GB	8	0.1	130	0.796
Bikaner, India	RF	3	0.1	70	0.612
	GB	8	0.1	130	0.709
Alice Springs, Australia	RF	3	0.1	280	0.529
	GB	6	0.1	130	0.690
Las Vegas, USA	RF	3	0.1	280	0.642
	GB	8	0.1	130	0.834
Marrakech, Morocco	RF	3	0.1	140	0.494
	GB	7	0.1	130	0.753
Luxor, Egypt	RF	4	0.1	140	0.633
	GB	8	0.1	130	0.796

Appendix 2. Hyperparameters of comparative models

Hyperparameter	XGBoost	CatBoost	GRU	KNN-XGBoost
colsample_bytree	0.3	-	-	0.8
Learning Rate	0.1	0.03	0.001	0.1
Max Depth	5	6	-	3
Alpha (L1 Regularization)	10	-	-	-
n_estimators	10	1000	-	100
Iterations	-	1000	-	-
l2_leaf_reg	-	3	-	-
border_count	-	254	-	-
Loss Function	-	'RMSE'	Mean Squared Error	-
Eval Metric	-	'RMSE'	-	-
colsample_bylevel	-	1	-	-
Number of Layers	-	-	2	-
Number of Units	-	-	50 per layer	-
Activation Function	-	-	'tanh'	-
Optimizer	-	-	Adam	-
KNN n_neighbors	-	-	-	5
KNN weights	-	-	-	'uniform'
KNN metric	-	-	-	'euclidean'
Subsample	-	-	-	0.8

Appendix 3. Statistical metrics comparison of proposed and validation models.

Data Set	Statistical Indicator	XGBOOST	CAT Boost	GRU	KNN-XGBOOST	RFGB	SDANN
		MSE	165.33	169.54	185.65	176.21	177.98
Hyderabad	RMSE	15.68	12.46	13.02	13.62	13.27	13.34
	MAE	13.28	8.52	9.36	9.46	9.25	8.87
	R2	0.57	0.73	0.71	0.68	0.74	0.74
	MSE	258.16	158.45	162.57	183.10	147.22	152.91
Sukkur	RMSE	16.06	12.59	12.75	13.53	12.13	12.36
	MAE	13.62	9.17	9.69	9.74	8.77	9.57
	R2	0.58	0.74	0.73	0.70	0.80	0.79
	MSE	351.63	160.79	273.24	202.05	380.84	371.83
Turbat	RMSE	18.75	12.68	15.4	14.21	19.51	19.28
	MAE	15.43	9.27	11.79	10.55	14.66	13.78
	R2	0.36	0.71	0.57	0.63	0.43	0.44

Appendix 4. Statistical validation of RFGB feature Analysis (Hyderabad dataset)

Input Features	R2	MAE	MSE	RMSE
[Ts, Tmax, Tmin]	0.69	10.33	211.95	14.56
[Ts, Tmax, Tmin, PS]	0.71	9.89	201.25	14.19
[Ts, Tmax, Tmin, RH]	0.69	10.27	210.74	14.52
[Ts, Tmax, Tmin, WS]	0.71	9.6	195.13	13.96
[Ts, Tmax, Tmin, WS, RH]	0.72	9.58	192.69	13.88
[Ts, Tmax, Tmin, WS, RH, PS]	0.73	9.42	188.88	13.74
All Features	0.74	9.25	176.21	13.27

Appendix 5. Statistical validation of RFGB feature Analysis (Sukkur dataset)

Input Features	R2	MAE	MSE	RMSE
[Ts, Tmax, Tmin]	0.789	9.26	153.44	12.38
[Ts, Tmax, Tmin, PS]	0.794	9.10	150.18	12.25
[Ts, Tmax, Tmin, RH]	0.797	8.95	147.56	12.15
[Ts, Tmax, Tmin, WS]	0.787	9.22	155.11	12.45
[Ts, Tmax, Tmin, WS, RH]	0.796	8.87	148.35	12.18
[Ts, Tmax, Tmin, WS, RH, PS]	0.799	8.78	146.70	12.11
All Features	0.801	8.77	147.2	12.13

Appendix 6. Statistical validation of RFGB feature Analysis (Turbat dataset)

Input Features	R2	MAE	MSE	RMSE
[Ts, Tmax, Tmin]	0.401	15.42	400.25	20.01
[Ts, Tmax, Tmin, PS]	0.409	15.16	394.69	19.87
[Ts, Tmax, Tmin, RH]	0.400	15.42	400.84	20.02
[Ts, Tmax, Tmin, WS]	0.402	14.40	399.37	19.98
[Ts, Tmax, Tmin, WS, RH]	0.400	15.14	396.22	19.90
[Ts, Tmax, Tmin, WS, RH, PS]	0.407	15.14	396.22	19.09
All Features	0.434	14.66	380.84	19.51

Appendix 7. Statistical validation of SDANN feature Analysis (Hyderabad dataset)

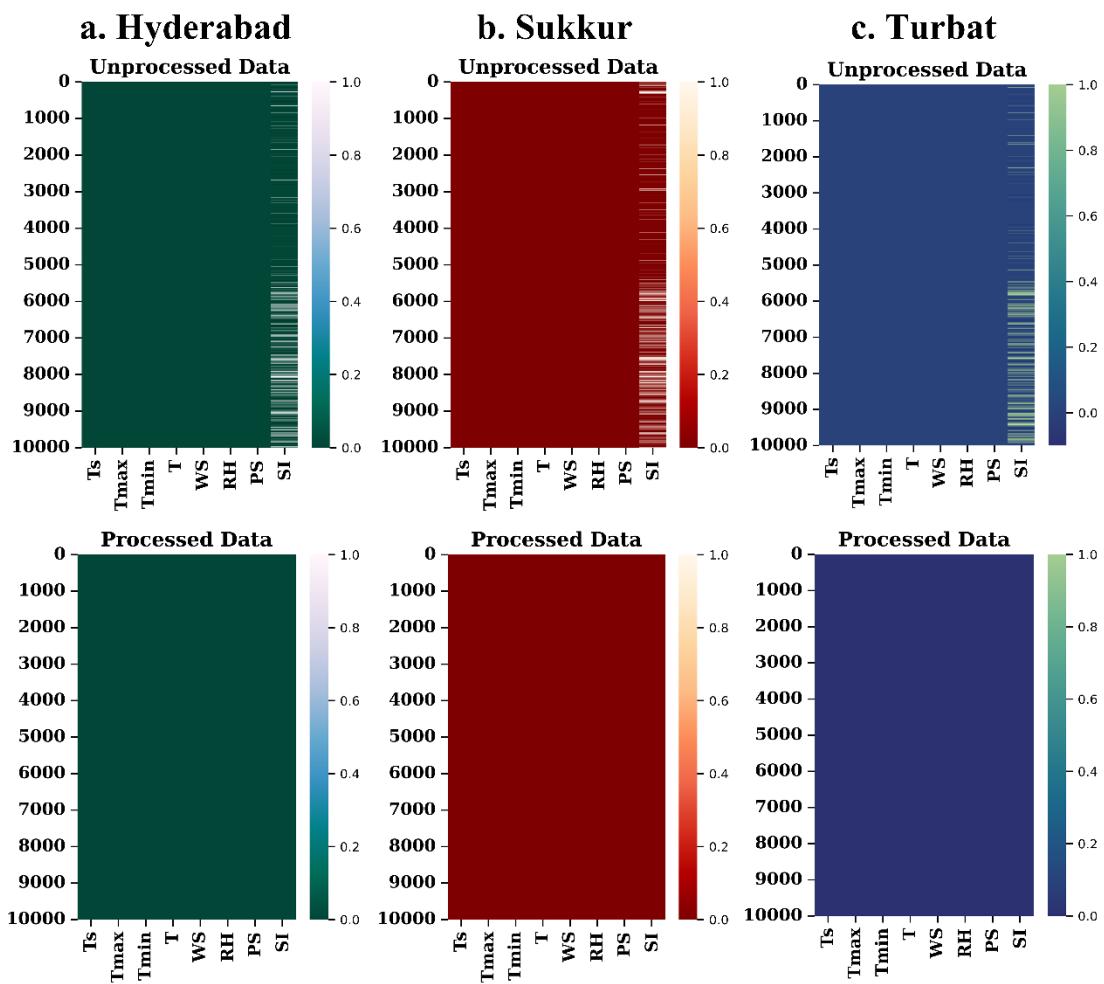
Input Features	R2	MAE	MSE	RMSE
[Ts, Tmax, Tmin]	0.71	9.80	198.85	14.10
[Ts, Tmax, Tmin, PS]	0.69	9.51	213.49	14.61
[Ts, Tmax, Tmin, RH]	0.72	9.49	190.20	13.79
[Ts, Tmax, Tmin, WS]	0.72	8.92	187.64	13.70
[Ts, Tmax, Tmin, WS, RH]	0.75	8.40	172.10	13.11
[Ts, Tmax, Tmin, WS, RH, PS]	0.74	8.95	178.33	13.35
All Features	0.74	8.87	177.98	13.34

Appendix 8. Statistical validation of SDANN feature Analysis (Sukkur dataset)

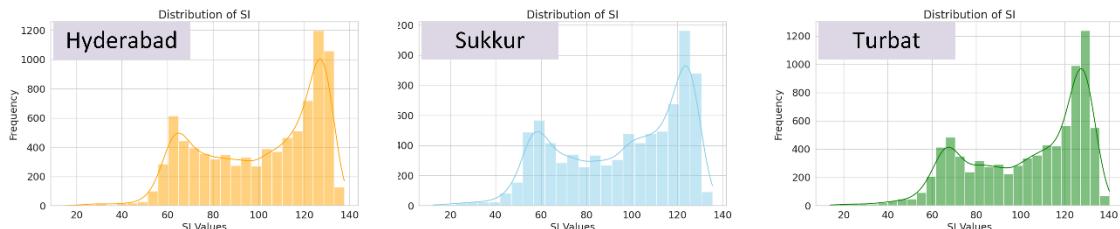
Input Features	R2	MAE	MSE	RMSE
[Ts, Tmax, Tmin]	0.80	8.86	142.98	11.95
[Ts, Tmax, Tmin, PS]	0.81	8.30	138.42	11.76
[Ts, Tmax, Tmin, RH]	0.80	8.76	142.21	11.92
[Ts, Tmax, Tmin, WS]	0.79	9.05	148.14	12.17
[Ts, Tmax, Tmin, WS, RH]	0.81	8.06	137.31	11.72
[Ts, Tmax, Tmin, WS, RH, PS]	-12.74	96.36	10013.64	100.0
All Features	0.79	9.57	152.91	12.36

Appendix 9. Statistical validation of SDANN feature Analysis (Turbat dataset)

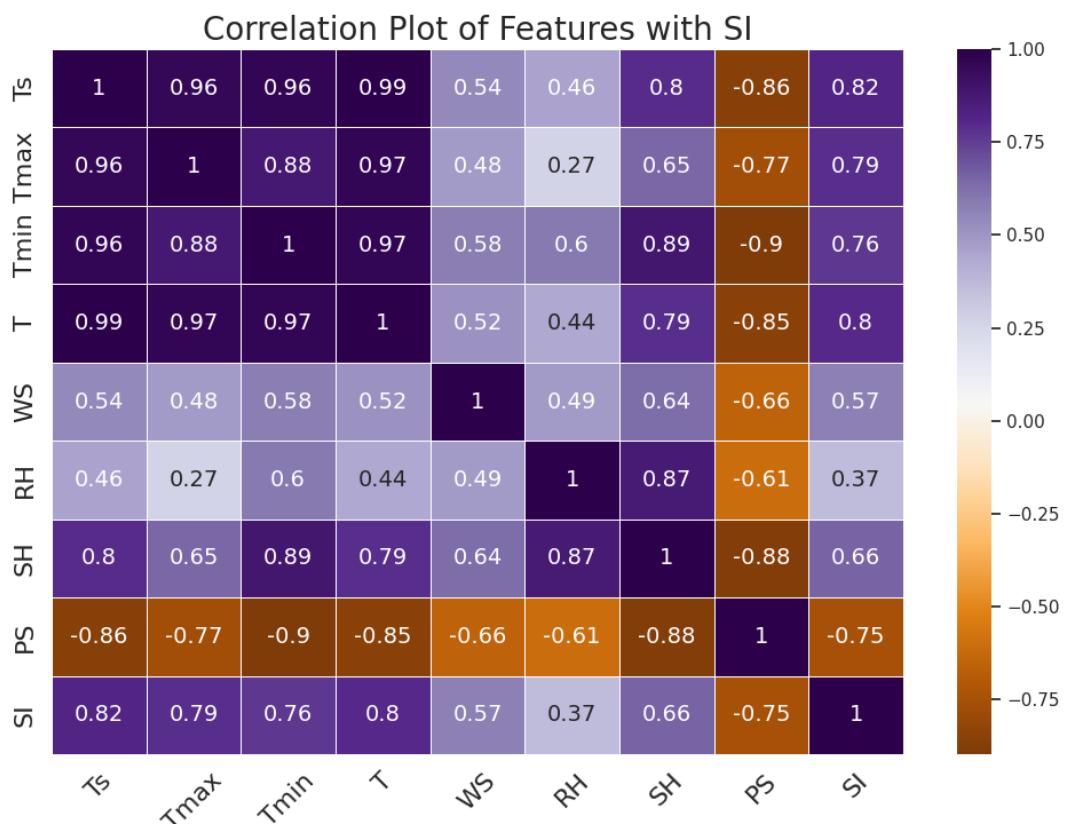
Input Features	R2	MAE	MSE	RMSE
[Ts, Tmax, Tmin]	0.41	15.51	389.69	19.74
[Ts, Tmax, Tmin, PS]	-16.14	103.87	11459.20	107.04
[Ts, Tmax, Tmin, RH]	0.37	16.74	416.76	20.41
[Ts, Tmax, Tmin, WS]	0.38	15.00	412.74	20.31
[Ts, Tmax, Tmin, WS, RH]	0.41	14.47	392.17	19.80
[Ts, Tmax, Tmin, WS, RH, PS]	0.40	13.85	399.91	19.99
All Features	0.44	13.78	371.83	19.28



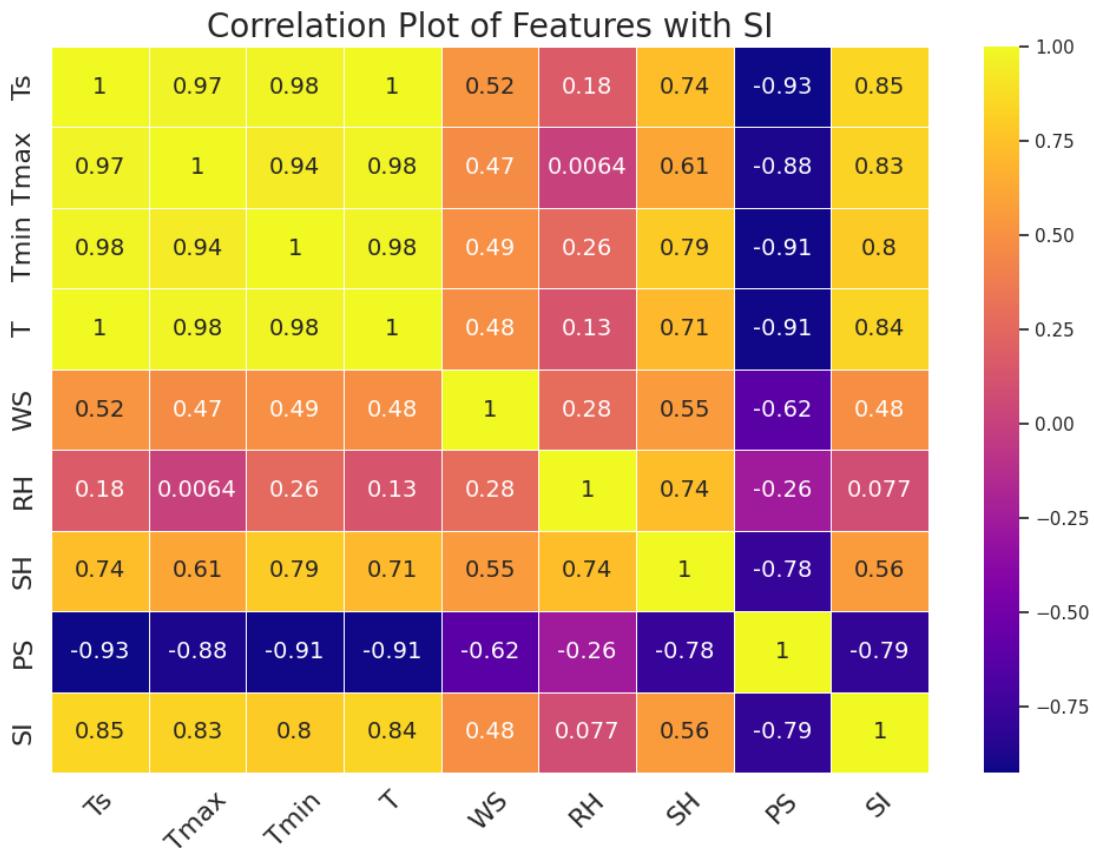
Supplementary Figure 1. Comparison of Unprocessed & Processed Data: (a) Hyderabad (b) Sukkur and (c) Turbat.



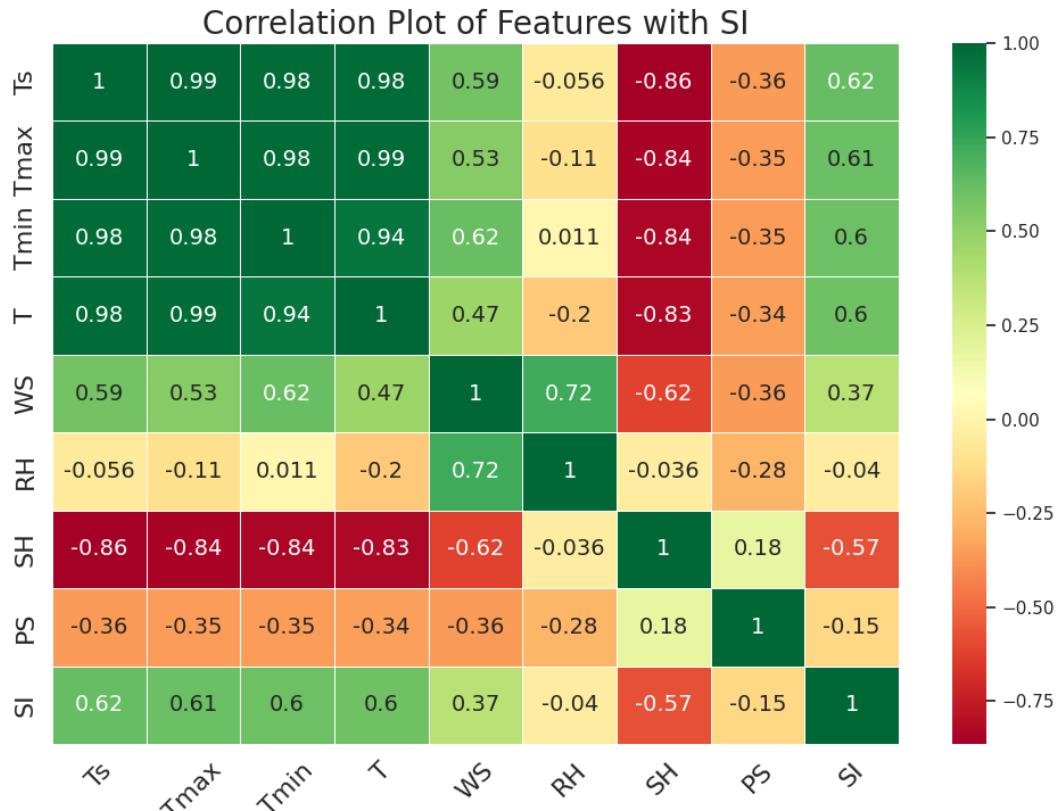
Supplementary Figure 2. SI distribution plots for understudy datasets



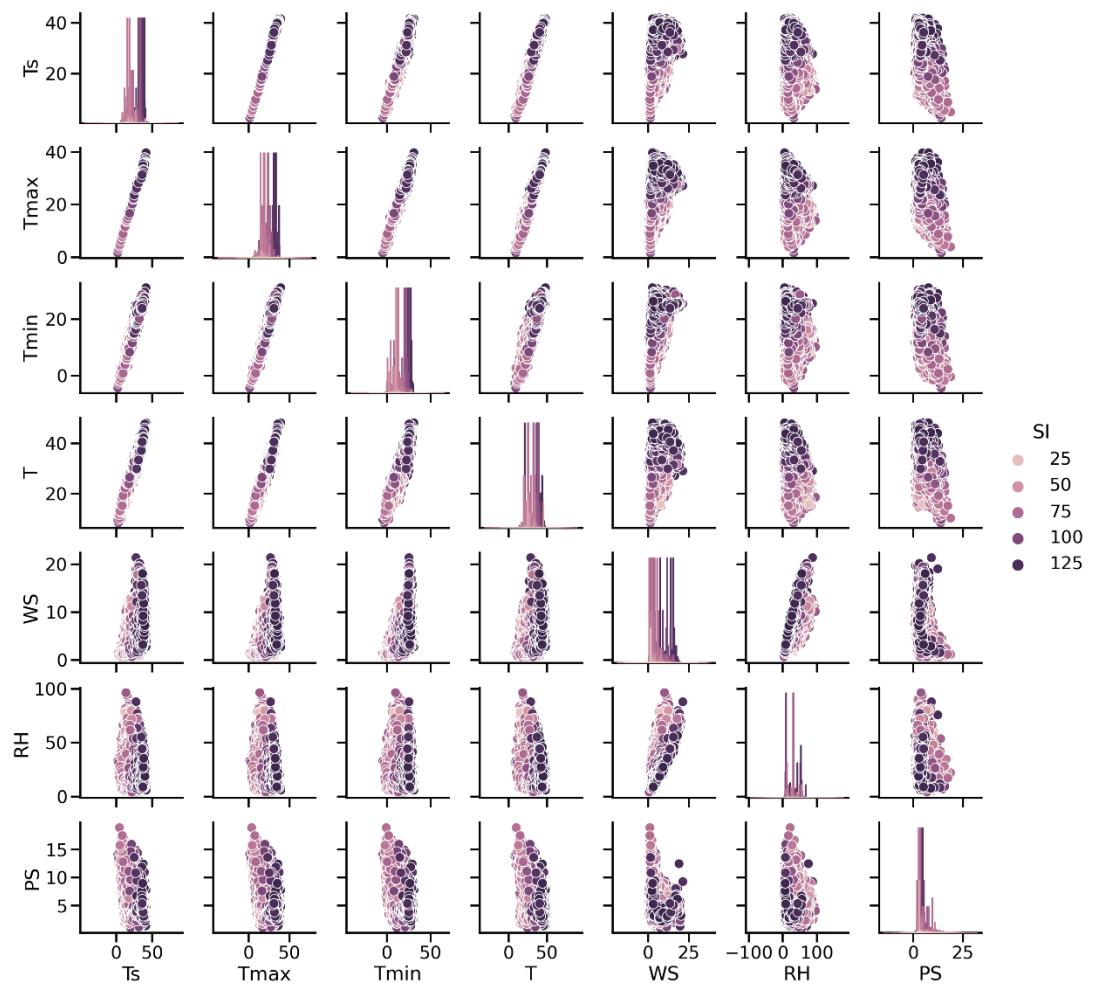
Supplementary Figure 3. Correlation Plot for Hyderabad SI



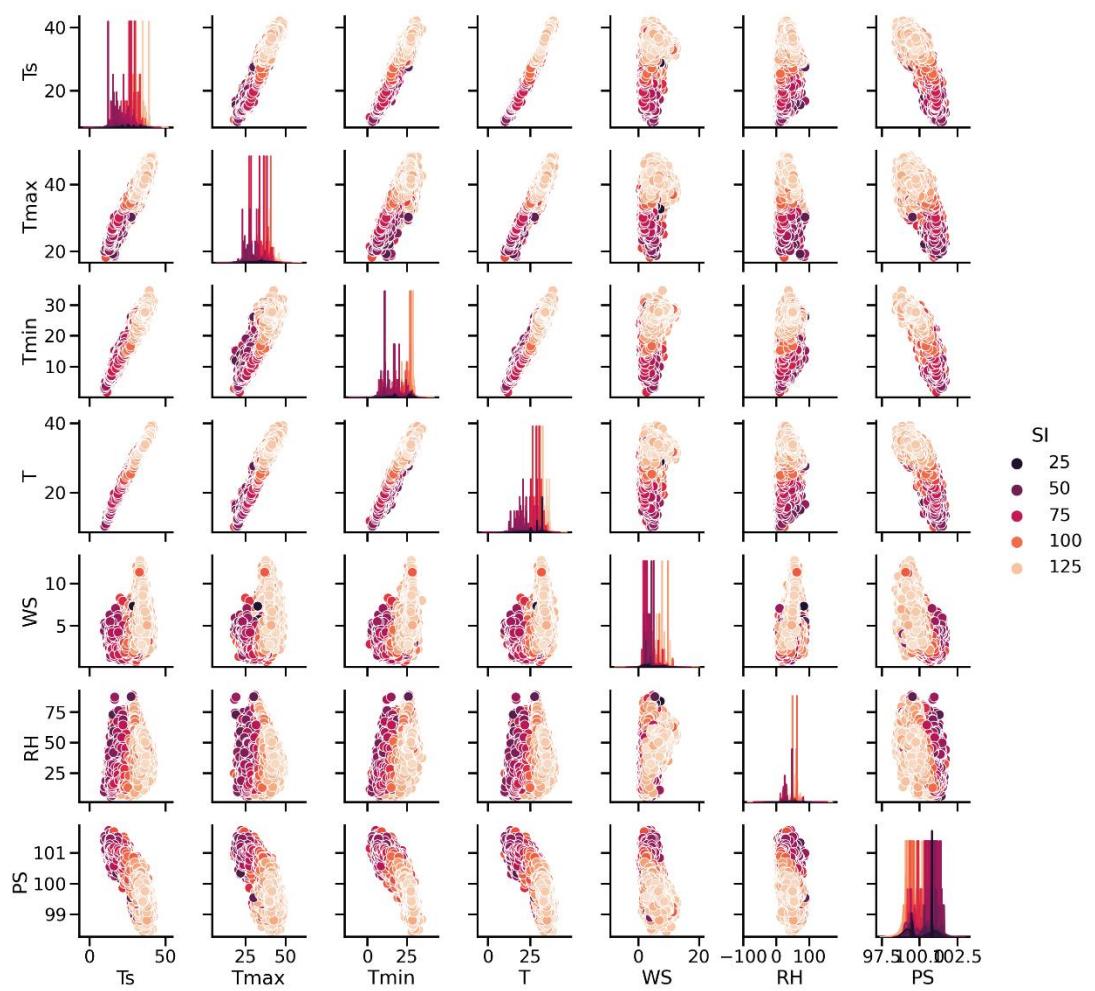
Supplementary Figure 4. Correlation Plot for Sukkur SI



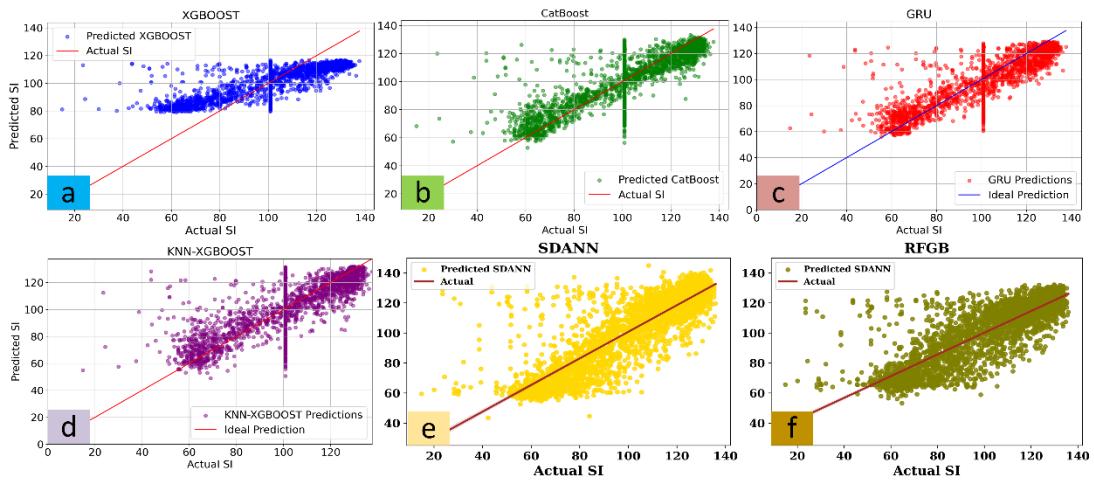
Supplementary Figure 5. Correlation Plot for Turbat SI



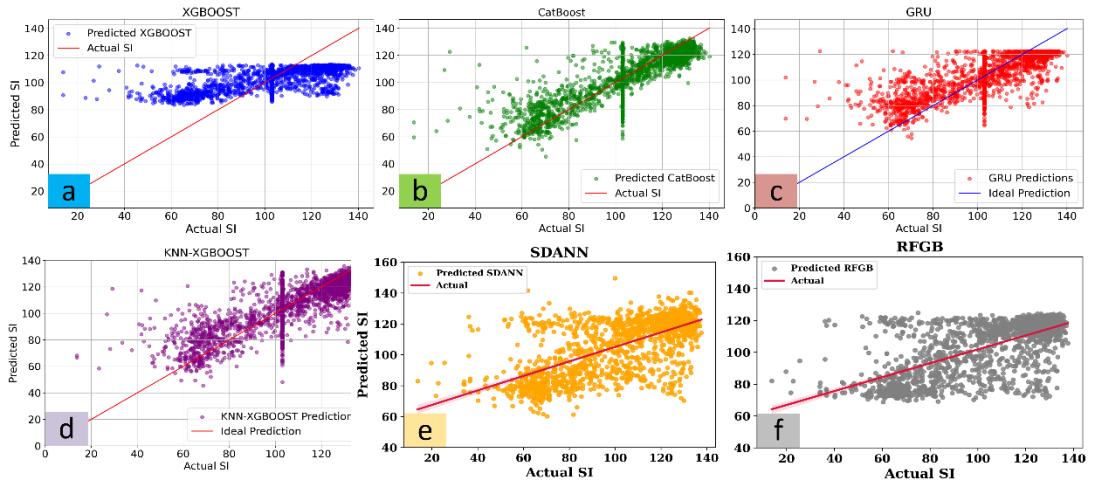
Supplementary Figure 6. Pair Plot of data distribution for Hyderabad



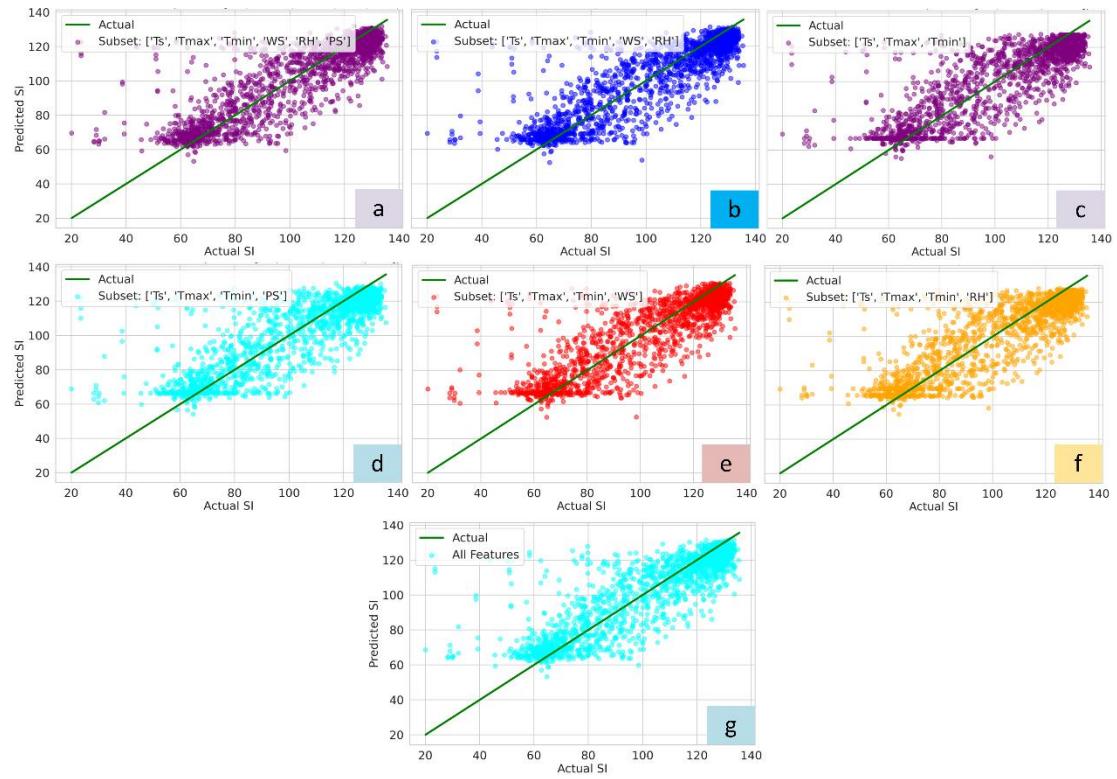
Supplementary Figure 7. Pair Plot of data distribution for Turbat



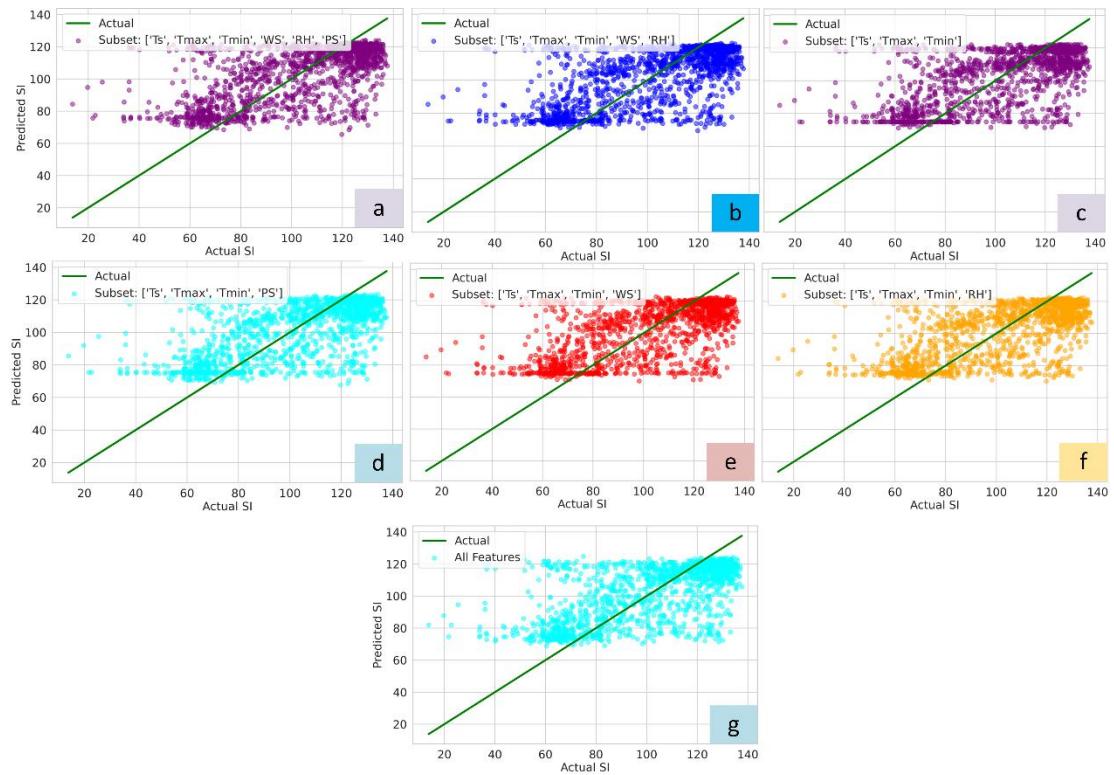
Supplementary Figure 8. Scatter plot of predicted over actual SI for Hyderabad dataset (a) XGBOOST, (b) Cat Boost, (c) GRU, (d) KNN-XGBOOST, (e) SDANN and (f) RFGB



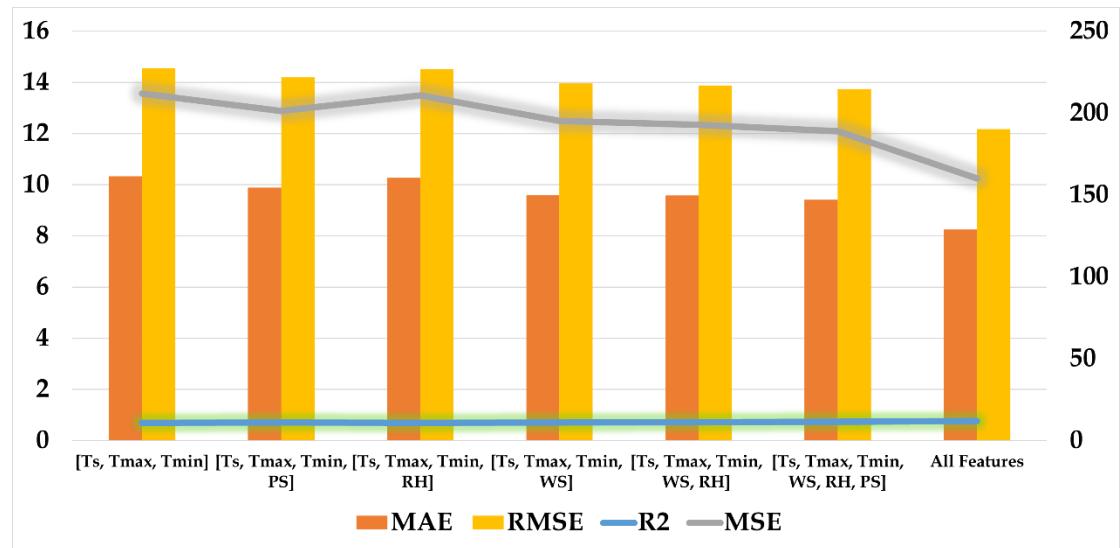
Supplementary Figure 9. Scatter plot of predicted over actual SI for Turbat dataset (a) XGBOOST, (b) Cat Boost, (c) GRU, (d) KNN-XGBOOST, (e) SDANN and (f) RFGB



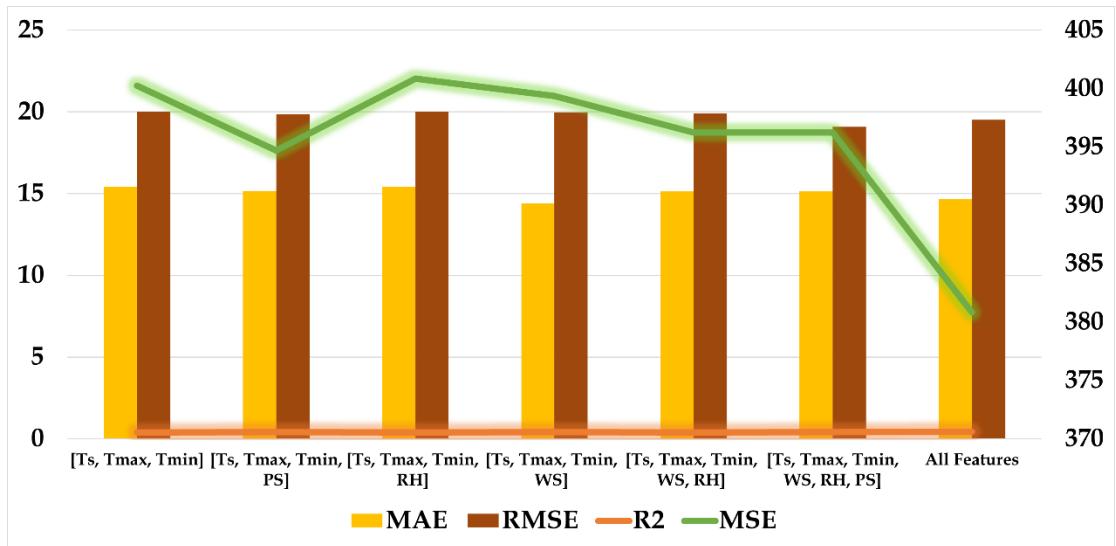
Supplementary Figure 10. Scatter plot of predicted over actual SI of various feature sub-sets (Hyderabad dataset).



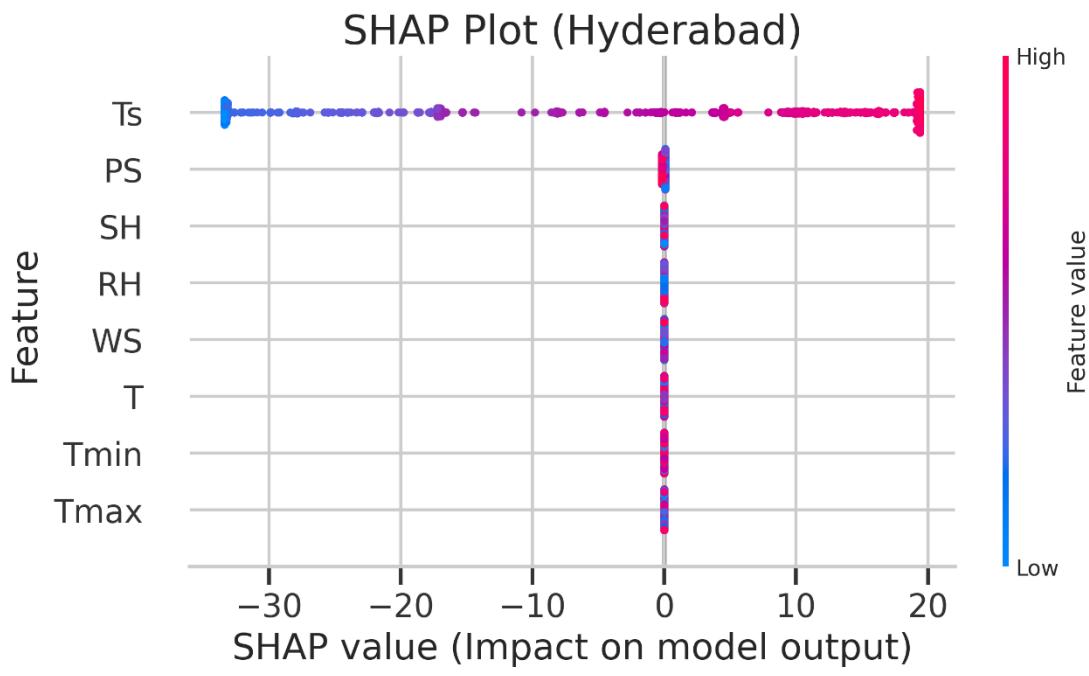
Supplementary Figure 11. Scatter plot of predicted over actual SI of various feature sub-sets (Turbat dataset).



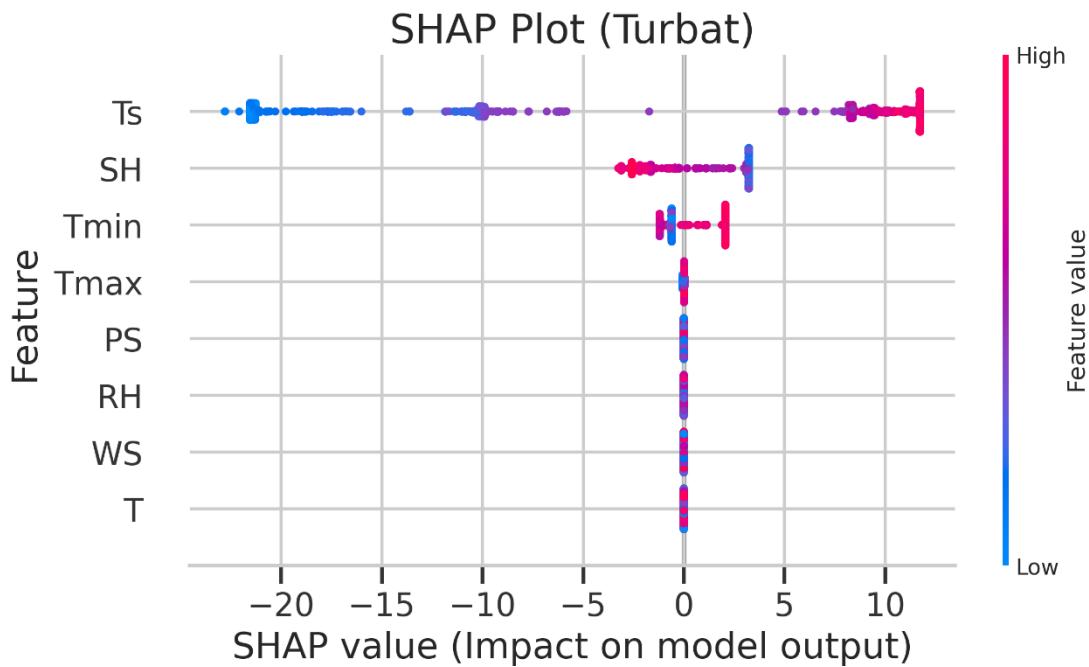
Supplementary Figure 12. Validation metrics plot of various feature sub-ssets for Hyderabad dataset.



Supplementary Figure 13. Validation metrics plot of various feature sub-sets for Turbat dataset.



Supplementary Figure 14. SHAP Analysis Plot for Hyderabad Dataset



Supplementary Figure 15. SHAP Analysis Plot for Turbat Dataset