Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-14616-6

A residual ensemble learning approach for solar irradiance forecasting

Banalaxmi Brahma[1] · **Rajesh Wadhvani**[1]

Received: 14 November 2021 / Revised: 8 February 2022 / Accepted: 3 February 2023

Abstract
Solar irradiance forecasting plays an essential role in efficient solar energy systems and managing power demand sustainably. In present work, a new residual ensemble learning approach, which consists of two advanced base models, namely Deep Neural Networks (DNNs) and Recurrent Neural Networks (RNNs), is proposed for solar irradiance forecasting. A model performance depends on data utilized for modeling and the modeling approach employed on the data. This paper focuses on both these aspects of the forecast model by proposing a three module approach. Firstly, a mechanism is proposed for the collection and analysis of multiple-site data surrounding the target location. A hexagon gridding system based algorithm is proposed for selection of multiple sites neighboring the target location. Then, correlation and feature importance scores are utilized as measures for feature selection to choose the most relevant data for forecasting target solar irradiance. In the second module, a residual ensemble learning model is proposed to forecast solar irradiance. The proposed framework is inspired by the hybrid forecast mechanism that considers the linear and non-linear characteristics for modeling. Advanced DNN models of Recurrent Neural Networks are also exploited for developing an accurate and robust model. The last module performs the integration of the deep neural network information and predicts the future values of solar irradiance. For a reliable and comprehensive assessment, the proposed framework is validated with data from four different solar power sites obtained from NASA's POWER repository. The residual ensemble model is trained on past 36 years of data as input for forecasting one day ahead, four days ahead and ten days ahead values of solar irradiance. Performance evaluation is carried out by comparing the prediction results with other models, including benchmark persistence, deep neural networks, and recurrent neural network approaches on performance indexes of MSE and RMSE. The proposed model shows an improvement in forecast performance by approximately 2.5 percent in prediction error. The predictive performance and stability make the proposed residual ensemble learning approach a reliable solar irradiance prediction model.

✉ Banalaxmi Brahma
bana.brahma92@gmail.com

[1] Department of Computer Science and Engineering, Maulana Azad National Institute of Technology, Bhopal, Madhya Pradesh, India

Published online: 06 March 2023

Keywords Short-term forecasting · Sequence learning · Recurrent neural network · Mean squared error · Ensemble · Residual

1 Introduction

Drastic changes in climate and higher demand for electricity has made power generation from renewable sources a necessity. The Government of India announced a target of 100GW by 2022. For reduction of fossil fuel usage, energy from renewable resources is being integrated to the grid. India being a tropical country receives solar radiation with nearly 3000 hours of sunshine. The majority of Indian regions receive 4 – 7 kWh of solar radiation per square meters. Solar energy is a form of sustainable energy resource that is available in abundance. It causes the least harm to the environment and has turned into an important energy source recently [18]. There is a necessity of photovoltaic (PV) systems for conversion of solar energy into electricity. Grid volatality might originate due to PV power, thus, affecting grid stability [27]. Power grid operators require a production forecast for ensuring a secure and cost-effective electricity supply system. Amrouche et al. [3] have proved correlation of power output with solar irradiance and PV modules. Sharma et al. [42] stated that the forecasted solar irradiance could be helpful in different phases of production of solar energy, like stability maintenance, monitoring, regulation, management of unit commitment & scheduling.

Time series forecasting plays a significant role in the planning and operational process of numerous applications. Solar irradiance being a time series data is also characterized by its underlying patterns, linear and non-linear properties, long-term dependencies and constituent features in the data. The mechanism of feature selection plays a major role in improving the model performance. Other than this, capacity of the model in understanding the underlying data patterns affects most forecasting models' performance. This research is motivated to develop the feature selection mechanism for data selection utilizing a single feature in a multivariate manner considering data relevance and dependency. The research is also motivated towards development of a reliable model for forecasting considering the long-term dependency and non-linearity component present in time series data.

The primary objective is to predict future values of the target feature considering the variability, non-linearity and long-term dependency of the dataset and consistent model performance. To achieve this objective, the problem can be formulated in two steps:

1. For given f^{target}, find out F such that,

$$F = \{f^{L_1}, f^{L_2}, ..., f^{L_m}\} \tag{1}$$

F is a set of feature f with different values from all nearby locations. f^{L_j} represents the instance of feature for jth location. j ranges for total m nearby locations and f^{target} denotes the target location. Then identify, F' such that,

$$F' \subseteq F \tag{2}$$

where, F' are set of features strongly correlated with f^{target}.

2. Considering f^{target} and F' as final features, find out parameters of the forecast model with optimal metrics such that prediction error is minimum.

The sections ahead discuss the related work and major research contributions.

1.1 Related work

In literature, solar irradiance forecasting of a target location is done by utilizing available meteorological data of that place. The meteorological parameters consist of a wide variety of features like temperature, wind speed, precipitation, etc. In these studies, the earlier solar irradiance and other related features had been used to forecast future solar data. In research, to estimate solar radiation, Huang et al. [21], Togrul et al. [45], and Yacef et al. [52] have utilized global radiation, sunshine duration, and temperature, respectively. Some studies have used different combinations of meteorological parameters for solar irradiance forecasting, including theoretical parameters also. Mellit and Pavan [31] use average air temperature and daily solar irradiance for the day ahead forecasting, while Belaid and Mellit [5] utilized sunshine duration, solar irradiance, and ambient temperature. Rao et al. [39] analyzed different combinations of meteorological parameters for solar irradiance forecasting, including theoretical parameters. Srivastava and Lessmann [43] included precipitation, cloud cover, downward longwave radiative flux, air pressure, and several other parameters. It can be observed that there is no uniformity in the selection of the related features. Different studies utilize different combinations for development of solar irradiance forecasting models. These meteorological features are chosen based on their strong relation to the output parameter. However, often these are unnecessary and the features do not serve as much information as the solar irradiance feature itself and its past values. Instead of using different parameters, the solar irradiance information available for the surrounding locations can be utilized in the forecast of a particular location's data. Following this, Mukhoty et al. [33] proposed a sequence to sequence forecast mechanism for solar irradiance utilizing the target and neighboring location's data. The research concluded the advantage of multiple site data in case of some particular horizons. Brahma and Wadhvani [7] also proposed multi-site data for forecasting a target location's solar irradiance. The performance of multiple site data proved to be more accurate than single location's univariate data. However, the site selection was performed utilizing the POWER Regional Data Access widget. For coverage of the region, a bounding box was obtained for the locations corresponding to the surrounding of a target point location. The maximum bounding box is of (4.5×4.5) degrees with a total of 100 data points in maximum. In this work, we propose a hexagonal geospatial grid system inspired by UBER's grid system. The feature selection from multiple neighboring sites is performed utilizing not only the correlation and feature importance values but also their combinations. The final step gives the optimal combination for most relevant site selection. Then, a novel residual ensemble forecast mechanism is being proposed here instead of only forecasting and comparing on the basis of the prevalant time series forecasting mechanisms. As such, the research here focuses on data from multiple neighboring locations to predict the target data. Instead of using the conventional way of feature selection from one target location, information from multiple locations is explored for day-ahead forecasting of solar irradiance. The mechanism of collection and selection of this data and its utilization for forecast through a novel hybrid mechanism will be further discussed.

There are basically three types of forecasting techniques, Image based Prediction, Numerical Weather Prediction (NWP), and Statistical & Machine Learning (ML) that range for short, medium and long term predictions. Here, a hybrid machine learning method

utilizing neural network architecture is intended to be used for short-term forecasting of day-ahead solar irradiance. Solar irradiance data is time series data that ranges sequentially for a time period. Traditionally, linear forecasting methods were used to be easy to compute, well understood, and provide stable forecasts. Autoregression(AR) [46], Moving Average Model(MA), ARX [30], ARMA [6], ARMAX [28], ARIMA [54], SARIMA [10], ARIMAX [24], SARIMAX [1] and Generalized Autoregressive Score (GAS) [12], [36] are models that are all linear over the previous inputs or states. Because of its linear nature these models alone are not suitable for many real applications. Here, the major limitation is the pre-assumption that data is linear due to which it is not capable of capturing nonlinear patterns. However, neural networks (NNs) utilize non linear transforming layers in which data stationarity is not necessary. NN has the interesting features of being a universal approximation for non-linear functions, strong prediction performance, and the ability to tackle unknown system modelling. Because of the universal approximation capability and distributed computing characteristic, neural networks have become one of the most influential prediction tools [11, 22, 37, 50, 53]. Neural networks have strong capability of extracting complex structures and providing an efficient tool for reconstructing a data driven noisy system, which is why it is being used for forecasting more often in recent times. These are well suited for modelling problems which need to capture dependencies and preserve knowledge through subsequent time steps. Extreme Learning Machine (ELM) [20] is also a modified version of the traditional Artificial Neural Network (ANN). Unlike the ANN, the hidden layer parameters are not required to be tuned in this framework [48]. It has the added advantage of learning speed and high efficiency through the kernel functions and activation functions for non-linearity. The ELM also shows better generalization along with a fast learning rate [55]. However, ELMs are not basically built for sequential data. Rumelhart et al. [40] work formed the basis of Recurrent Neural Network (RNN). The network assigns weights to the past input data to predict future values. The RNN is considered for time series forecasting widely due to the presence of feedback loops in the recurrent cells that addresss the temporal dependencies and temporal order of the sequences inherently.

Conceptually, statistical forecast models and artificial neural networks are non-parametric techniques which attempt to make appropriate internal representations of sequential data. In literature, hybrid models were developed by combination of different models, improving upon the forecast accuracy [17]. The basic idea of model combination is to integrate the unique features of each model into an organic whole. Both empirical and theoretical findings suggest that combination of different methods is effective in improving forecast performance [49]. The hybrid residual ensemble learning approach has been used widely in literature for solar irradiance forecasting. Ji and Chee [23] proposed an ARMA-TDNN approach for hourly forecasting of solar irradiance which achieved higher performance as compared to single ARMA and TDNN model. Voyant et al. [47] introduced an ANN-ARIMA model which outperformed the persistence, ARMA and clear sky models. [44] proposed an ARMA-GARCH model for prediction of global solar radiation on a daily scale in two regions of China. The hybrid model showed the best performance in both the locations. Gairaa et al. [13] presented an ARIMA-ANN model which also outperformed other standalone models. Further, ANN-ARX [41], NARNN-ARMAX [2], SARIMA-ANN [32] were also found to be performing better than the compared standalone models. However, the models are built only with the traditional statistical models and multilayer perceptron architecture of ANN. To the best of our knowledge, the sequential models of RNN were not utilized for the construction of hybrid models. Moreover, data from the target location is only utilized for predicting future solar radiation values. This inspired us

to build a novel residual based ensemble learning approach for forecasting solar irradiance data.

1.2 Research contribution

Here, we present a novel residual ensemble mechanism to overcome the limitations as mentioned above. A significant portion of the literature on solar forecasting mainly investigates which one is the most accurate ML model to forecast solar irradiance without analyzing the conditions under which the method might be appropriate. In this study, we propose a hybrid model, analyzing datasets from four regions taken from NASA's POWER database [35]. For this task, solar irradiance data of a region is selected for forecasting data in a single point location. That is, multi-site data is utilized for performing forecast of a single site. The question intended to being answered is how we can develop an optimized and generalized model for a particular location based on data from multiple locations such that it can be used in future solar installation operations. The main emphasis is on the correlations between data from different locations such that the appropriate data can be used for modelling. After selection of data, a novel hybrid architecture is developed for improving the forecast performance. Hence, mechanisms for feature selection are proposed in which the univariate feature of a target location along with other locations is used for construction of multivariate features. Forecast models are developed in order to model the non-linear characteristic and temporal long-term dependencies in data. To the best of our knowledge, no comprehensive study using multiple site data and hybrid residual ensemble of neural networks to forecast solar irradiance has been reported yet. Since the proposed model is newer than previously defined models, this study also opens a new window for future studies in several domains such as forecasting of renewable power by adoption of the proposed modern algorithm in present work. The main contributions of this paper include:

- Instead of utilizing given multivariate features of time series data, the proposed approach intelligently performs feature selection by utilizing only the target feature to construct multiple features. The feature selection mechanism selects data based on an approach of geographical selection of locations. The selection of locations is performed by utilizing a proposed hexagon gridding system algorithm. Our approach utilizes five standard correlation measures for data relevance in which the measures are utilized as a single measure as well as in combination. The technique used in analyzing dependencies consist of the Pearson, Spearman and Kendall correlations, the dynamic time warping distance, and the XGBoost feature importance.
- Secondly, different RNN variants, Attention-Long Short Term Memory (LSTM), CNN-Attention-LSTM and Hybrid Ensemble are developed. Also, the proposed Hybrid Ensemble model is compared with Attention-LSTM based time-series forecast models, CNN-Attention based models in which Convolutional Neural Network and Attention LSTM are applied to model the time series data.

The aim is to conduct a systematic investigation for reliable solar forecast using a hybrid architecture. The results of our study can help to enhance time-series forecasting of solar irradiance data in a more generalized and cost effective manner. The rest of the paper is organized as follows. Section 2 describes methodology while Section 3 introduces the database. Section 4 discusses the results of the experiments conducted and finally the paper is concluded in Section 5.

2 Methodology

In general, the purpose of this work is to utilize a neural network based approach for solar irradiance forecasting. For achieving this goal, a three module approach is proposed. The first module is the data selection part in which data from multiple locations are selected and accessed according to the target location. Then, the multi-site data is ranked for importance. This gives us the features based on which the residual ensemble forecast model will be developed. The output module finally performs the final forecasting.

Figure 1 represents the framework of the proposed multiple site hybrid residual ensemble based solar irradiance forecast mechanism. The data is collected for four target locations' and the neighbouring sites' solar irradiance from NASA's POWER repository. The multiple sites neighbouring the target location are selected utilizing a hexagonal region enclosing many points. This is described in details in the next section. Then, data selection is performed by analyzing the correlation and feature importance scores. Based on these scores, several combinations of sites as features are created. These feature combinations are then fed to the forecast model. The proposed forecast model utilizes the concept of hybrid residual ensemble but exploits the neural network architecture instead of the statistical autoregressive models used in literature. After modeling, the model is evaluated according to performance evaluation metrics and different feature combinations. The final model gives us the optimal combination of the multiple sites and model configuration for practical solar irradiance forecasting.

Algorithm 1 describes the methodology in a formal pseudocode template. In the algorithm, F is a set of feature f with different values from all nearby locations. f^{L_j} represents the instance of feature for jth location. j ranges for total m nearby locations and f^{target} denotes the target location. $Correlation$ denotes the different metrics for measurement of relation between target and neighbouring locations. F_i' are the set of features strongly correlated with target feature where i denotes different feature combinations. The total number of feature combinations is denoted as $featurecomb$ and the modeling steps are repeated for all these instances. The final model would denote the optimal model along with the optimal feature combination for the solar irradiance forecasting of a particular location.

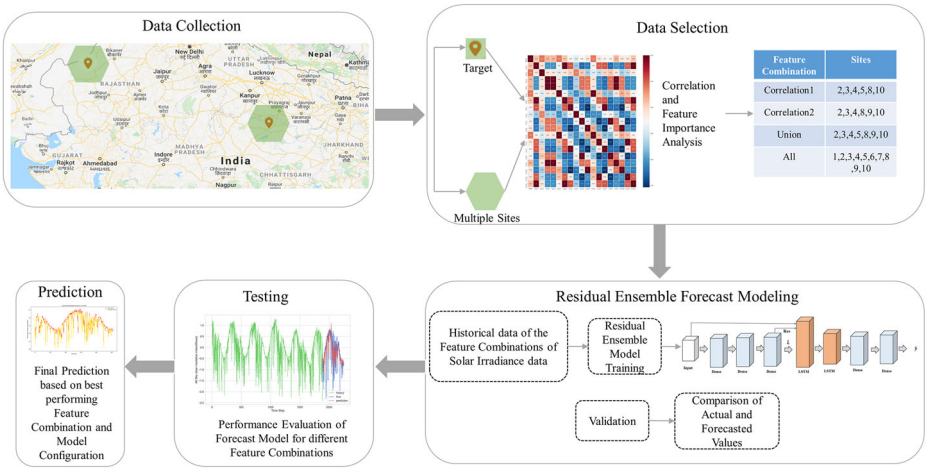

Fig. 1 Proposed framework for solar irradiance forecasting

Input: Coordinates for target location in Latitude and Longitude
Output: Solar irradiance forecast model

1: Initialize a map with the input coordinates;
2: Find neighbouring sites for target location utilizing the Hexagon Gridding mechanism described in Algorithm 2;
3: For target location with feature f^{target}, $F = \{f^{L_1}, f^{L_2}, ..., f^{L_m}\}$ is achieved from previous step;
4: Calculate the correlation and feature importance metrics ($Correlation$) between target location and neighbouring sites solar irradiance data;
5: Arrange the data from sites as feature combinations in order of higher precedence according to the $Correlation$ metrics to attain $F_i' \subseteq F$;
6: **for** $i = 1$ *to* $featurecomb$ **do**
7: Split the dataset into training, validation and test data with input F_i' and target output f^{target};
8: Initiate a forecast model and train it on training data with initial set of parameters and hyperparameters.
9: Find the optimal parameters including number of nodes, hidden layers, epochs, batch size on the validation set.
10: Compute the prediction error on the test data.
11: **end for**
12: The prediction model with optimal feature combination based on the performance metrics is then chosen as the final model.

Algorithm 1 Pseudocode for proposed solar irradiance forecasting framework.

The sections ahead discuss the components of the proposed framework in a much more detailed manner. The technologies utilized and the reasons for the choices are also discussed. Finally, the performance evaluation metrics are also introduced.

2.1 Data collection and selection module

In analyzing large spatial datasets, grid systems partition areas into identifiable grid cells, thus, playing a significant role. UBER's grid system's proposed site selection algorithm is inspired for efficient visualization and exploration of spatial data. Procuring insights and information from data require analysis of data, and due to the geographic diversity, this needs to occur at fine granularity. Hexagons were used because hexagons minimize the quantization error and also allows approximation of radiuses easily. Other choices, such as polygonal zones around areas have sizes and shapes which are unusual and not helpful for analysis. A global grid system is formed by overlaying a grid on the map. Different grids and map projections can be combined for this. Methods like the Mercator projection has distortion of size while square grids require multiple coefficients sets. Hexagons have only one distance between the center point and neighbors', in comparison to three distances for triangles or squares' two distances. This simplifies analysis and smoothing over gradients. As such, the hexagon gridding mechanism was opted for our task. Forecasting is performed using past data of the target location prevalently. This work introduces a multi-site mechanism where instead of using data only from the target location, the data from neighboring sites is also utilized.

A function is designed for the selection of neighboring sites corresponding to a target location. The steps are described in Algorithm 2. Firstly, the target location selection is performed. A distance variable d = 75 km is set for the selection of the neighboring points along with the entry of the number of layers. The map is organized in such a way that the

points are measured in an equidistant manner. Loop runs from the first layer to the maximum layer, and for the total number of points ranging, a loop is run covering equidistant locations from the target location. The calculated locations are marked, and a hexagon is drawn connecting them. Figure 2 represents this with more clarity.

```
Initialize a map utilizing the developed google maps API;
Initialize the target location;
The distance from the target location is initialized;
The maximum layers is initialized as maxlayers;
for i = 1 to maxlayer do
    NumberofPoints = 6 * i ;
    for j = 1 to NumberofPoints do
        Find coordinates of neighbouring sites from target location on given layer;
    end for
end for
```

Algorithm 2 Hexagon Gridding Mechanism Algorithm.

Figure 2 shows the target location and surrounding region selection. The center point is the target location while points in the enclosed hexagonal region represent multiple sites that surround the target location. The hexagon is used for the selection of locations for maximum coverage of surrounding sites without overlap. The multiple sites can be selected for all the neighboring locations without leaving out points. Also, the distances from the center to all the points on the hexagon are the same, which also fall on the circumference of a circle. As such, for every enclosed circle, the distance between the target location and the surrounding points will be the same.

The collected data provides us with neighboring data sites corresponding to one target location. The variables used as input for the prediction of future values are meteorological

Fig. 2 Hexagon gridding mechanism for site-selection

variables that are not always equally relevant. Often, few have a substantial influence on the target response. It is useful to learn the relative importance or correlation of the input variables in predicting the response. This selection of important features from the data collected is described in the next section.

2.1.1 Correlation

Correlation measures the strength of association between two variables along with the direction of relationship. Correlation and Causation are two different concepts that can be utilized for selecting the input features. Correlations are useful for forecasting, even when there is no causal relationship between the two variables, or when the causality runs in the opposite direction to the model, or when there is confounding. In the next step of the forecast framework, we intend to analyze this data from multiple sites to select data which are strongly correlated to the target location and would be utilized as input features in accurate forecasting of solar irradiance data. Instead of utilizing data only from the target location, the proposed framework intends to exploit the data available from the neighboring locations for predicting solar irradiance parameter's future values. Since, the same parameter is chosen from all the locations, they are strongly correlated and of substantial importance in forecasting future data. In order to perform this task, various correlation measures are utilized here. This section presents these measures which would represent the correlation between the various neighboring locations and target location, thus, denoting the feature importance in forecasting.

Pearson Correlation It measures relationship between variables. -1 denotes negative correlation, 0 implies no correlation and 1 indicates perfect correlation. It provides a measurement of global synchrony.

$$Corr_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}} \tag{3}$$

Spearman Correlation Spearman correlation is a correlation test which is non-parametric in nature. There are no assumptions regarding the data distribution and is a suitable correlation when variables are measured on a scale. Equation 4 represents the spearman correlation, where n is the observation number and d_i represents difference between corresponding variable ranks.

$$Corr = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{4}$$

Kendall Correlation Kendall correlation is a non-parametric hypothesis test for statistical dependence. The Kendall correlation is high when observations have similar rank between two variables and low when observations have dissimilar rank. It measures correspondence between two rankings. Values close to 1 indicate strong agreement while values near -1 indicate strong disagreement.

Dynamic Time Warping Dynamic Time Warping (DTW) performs computation of the path between two time series minimizing the distance between the two. DTW finds the euclidean distance between data at each frame across every other frames for calculating the minimum path between the two signals.

$$Distance = \sqrt{\sum (x_i - y_i)^2} \tag{5}$$

2.1.2 Feature importance

The feature importance of the solar irradiance values in forecasting the target location's solar irradiance value is also considered for ranking. Extreme Gradient Boosting (XGBoost) uses information gain for this task. Firstly, the construction of boosted trees is performed which then gives importance scores for all the attributes. A single decision tree makes the decision by the amount the performance measure improves by the split point, weighted by number of observations the node is responsible for. An error function indicates the performance and the feature importance is calculated by averaging across all of the model's decision trees.

All these values provide an indication of the relationship between target location and neighboring sites, such that only relevant sites are chosen.

2.2 Residual ensemble forecast module

This section introduces the proposed residual ensemble forecast methodology. The deep learning methodology of dense and recurrent neural networks is utilized which is presented below. The past solar irradiance data of the target location and the multiple-sites are the features. The final output $y_{(t+\Delta)}$ is the future value of the target location.

The hybrid forecast model developed here is a type of Residual Ensemble Learning Approach which is denoted in Fig. 3. $\hat{L}_t$ denote the linear component while Res denote the non-linear component. $func^1$ model is capable of modeling the linear components while $func^2$ model is meant for non-linear modeling. The process of forecasting is based on the idea that solar irradiance data is composed of both linear and nonlinear components. Forecasting of linear components require a simple linear model while the residual component uses a non-linear model. The final forecast data is obtained by the accumulation of the results of the linear and nonlinear models either additively or functionally [16, 54]. The hybrid structures of time series forecasting can be further divided into parallel, series and

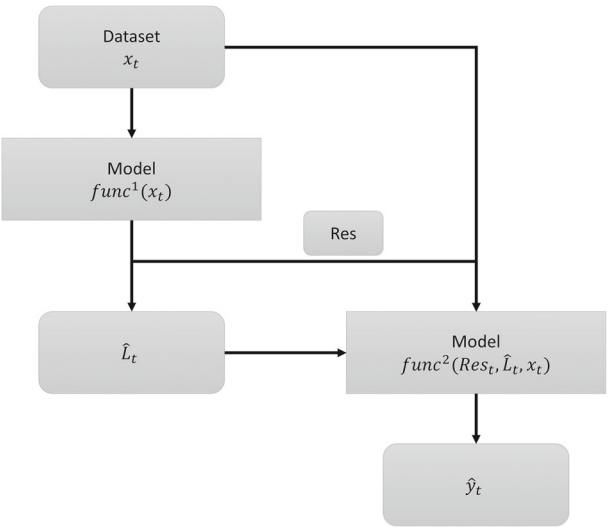

Fig. 3 Proposed ensemble mechanism

parallel-series structure [17]. Here, we intend to use a parallel-series structure of hybrid forecast mechanism.

Hybrid models are introduced since the true underlying data generating process cannot easily be determined by a single model. Identification of all the time series characteristics might require more than one model. The models in literature were derived one after the other, following the similar concept of strictly using the ARIMA and ANN models. There also exists the assumption of the relationship between the linear and non-linear components as additive which may underestimate the component relationship and degrade performance. As such, the proposed model aims to provide flexibility in the selection of the models and use the latest developments in artificial neural networks to implement an efficient and robust hybrid solar irradiance forecasting model.

In this paper, a novel residual ensemble model is proposed without assumptions of the traditional hybrid models. On the basis of previous works a time series can be said to be a combination of a linear and a non-linear component. Thus,

$$y_t = L_t + NL_t \tag{6}$$

$$y_t = func(L_t, N_t) \tag{7}$$

Here, L_t is the linear component and NL_t is the non-linear component. Previous works assume the linear and non-linear relationship to be either additive [54] or functional [25]. First, step in the hybrid model is the assumption of modelling the linear component. The residuals from the actual and predicted output of this component, from the first stage, will contain the nonlinear relationship.

$$\hat{L}_t = func^1(x_{t-1}, ..., x_{t-m}) \tag{8}$$

$$L_t = \hat{L}_t + Res \tag{9}$$

$\hat{L}_t$ is the predicted value that is obtained after the data is modelled using $func^1$ as shown in (8). Conceptually, the linear component of the data is a composition of the predicted linear component and residual as depicted in (9). These two values are results of the first stage of ensemble modeling that are used for second stage. The second stage is aimed for the proper modeling of the non-linear residuals in order to understand the underlying non-linearity. For this purpose, a functional modeling is performed utilizing the output from the first stage and the input data.

$$\hat{y}_t = func^2(Res_{t-1}, ..., Res_{t-p}, \hat{L}_t, x_{t-1}, ..., x_{t-m1}) \tag{10}$$

Here, $func^2$ is the nonlinear neural network function. Res_i and $\hat{L}_t$ components are obtained from $func^1$ model and x_i is the data being provided. p and $m1$ being integers determined in design of final neural network. These variables may also be removed in the design process of the deep neural network as a result of the data generating process and the structures in data. These intricacies of the components of data are taken care of in the implementation as well and the results are shown for the same developed model.

Unlike conventional architectures, our proposed mechanism also takes in the original input and the residual component instead of just the hidden layer output of the previous layer. The input is firstly passed through a dense network, after which the residuals are then modeled utilizing the LSTM [15, 19] architecture. Equation 11 represents $func^1$ which is the dense network operation with $\hat{x}$ representing the output of any dense layer, w_h weight of the h^{th} hidden layer and b bias.

$$\hat{x} = w_h^T x + b \tag{11}$$

The residual modeling is performed utilizing the $func^2$ function, LSTM, which is an advanced variant of Recurrent Neural Networks that improves the efficiency. Recurrent Neural Networks have the drawback of getting influenced by local instances which makes it difficult for the output to associate with the earlier inputs. The problems of exploding as well as vanishing gradients persist in RNN which is why LSTM was developed for overcoming the demerits of RNN. LSTM introduces additional components to RNN, the input, forget, and the output gate. Forward pass equations are stated in (12).

$$\begin{aligned} A_t &= \tanh(W_{cur} X_t + R_{cur} H_{t-1}) \\ I_t &= \sigma(W_{inp} X_t + R_{inp} H_{t-1}) \\ F_t &= \sigma(W_{for} X_t + R_{for} H_{t-1}) \\ O_t &= \sigma(W_{out} X_t + R_{out} H_{t-1}) \\ C_t &= I_t \odot A_t + F_t \odot C_{t-1} \\ H_t &= O_t \odot tanh(C_t) \end{aligned} \tag{12}$$

A_t processes the previous state and the current input after which I_t, the input gate, decides upon parts of A_t to be added in the C_t which is the long term state. The forget gate F_t has the responsibility to decide which parts of C_{t-1} are to be erased for removing unnecessary parts. The output gate O_t finds parts of C_t to be read and shown as output. H_t is the hidden state or the short term state shared between cells which, in principle, can contain information from arbitrary points earlier in the sequence. C_t is the long term state in which memories are dropped and added by respective gates. The $\odot$ in the equation above represent the element-wise multiplication of matrices or more popularly known as the Hadamard product.

The Dense Neural Networks and Long Short Term Memory architectures are chosen instead of the ARIMA and ANN models in conventional hybrid mechanisms, since the machine learning models tend to perform better than the statistical forecast models [34]. The DNN performs with good prediction accuracy as it has been also used for residual modeling in previous works. LSTM is ideal for forecasting tasks in time series datasets since it is able to capture long term dependencies and predict the future values in an efficient manner. Inspite of being strongly capable, efficiency of neural network depends on its ability to learn as well as generalize. Learning refers to the capability of a model for finding the underlying patterns and fit the data at hand. Generalization is more often related to the stability of a model, its performance of running on unseen data from the same generating process. In line with it, underfitting is the model being not capable of fitting present available data and overfitting is model's tuning being more than necessary. Therefore, underfitting and overfitting implies learning and generalization capability respectively. For the purpose of learning, various optimization techniques exist in literature. Neural network is trained using ADAM [26] optimization technique, finalizing the model when minimum loss is achieved. Regularization is able to control overfitting with different techniques having features of sparsity, shrinkage and capacity control. As such, extensive training is performed for generalizing the model to the task of solar irradiance forecasting.

Table 1 represents the optimized model hyperparameters of the proposed residual ensemble model. The model consists of three fully connected layers in the first part of modeling the linear component and two LSTM layers and two fully connected layers for modeling the residual component. The number of hidden neurons are shown in Table 1. The parameters are initialized as random numbers when the training is begun. The ADAM optimizer is utilized for tuning the model parameters. The initial learning rate is set at 0.001 and the batch

Table 1 Model Hyperparameters

Hyperparameter	Proposed Work
Cells per LSTM layer	50
Neurons per FC layer	50, 25
Optimizer	Adam
Initial learning rate	0.001
Batch size	32
Dropout ratio	0.2
Epochs	50

size at 32. Dropout is utilized with a ratio of 0.2 for regularization. The linear activation function is used in the output layer to get the final forecast results.

The number of these hyperparameters are only chosen after extensive search through grid search on training set. The decisions are also influenced by previous works on solar irradiance forecasting [7–9]. The model can always be made more complex with different architectures in order to accommodate non-linearity of data. But this would in turn lead to overfitting. The sequential LSTM model is utilized here since solar irradiance data is a time series data consisting of temporal dependencies. Also, the dropout regularization is included for handling the problem with variance. The proposed study is conducted using a 64 bit windows computer equipped with intel core i5-9300H@2.90 GHz, 8 GB of memory. All experiments are performed using tensorflow backend on single NVIDIA Corporation GPU GeForce GTX 1650.

2.3 Performance evaluation metrics

The test performance is evaluated by the performance metrics of Mean Squared Error (MSE) and the Root Mean Squared Error (RMSE) which are defined in (13) and (14) below respectively.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y(i) - x(i))^2 \tag{13}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y(i) - x(i))^2} \tag{14}$$

where, N is the total number of predictions, y is the estimated variable, and x is the actual variable.

3 Application to the Indian database

As stated in Section 2, in the first module, after data collection from multiple sites the relevant sites are selected. Data from four locations are chosen as target data while the neighboring locations are chosen according to the hexagon gridding system based algorithm. The daily solar irradiance data for a period of approximately 36 years was collected. Figure 4 represents the locations on which the solar irradiance forecasting is to be performed. The map in Fig. 4 indicates the repository locations with the red triangle denoting Location 1, the green triangle pointing at Location 2, the blue triangle representing Location 3 and the

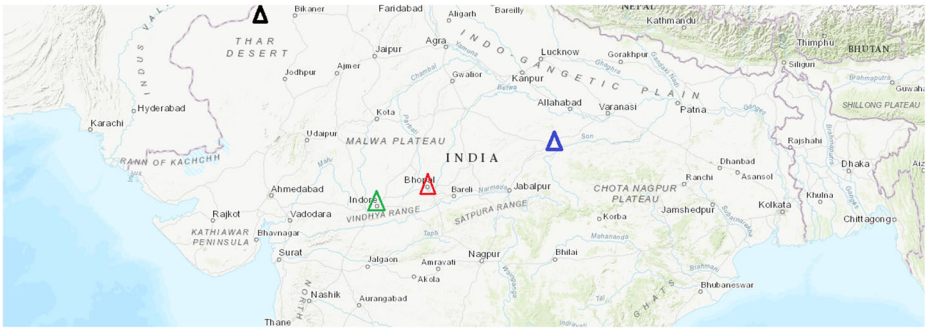

Fig. 4 Target locations selected for solar irradiance forecasting

black triangle representing Location 4. Tables 2 and 3 lists the descriptive statistics of the target solar irradiance data for the four locations.

The data utilized in the proposed work is from a publicly available website. NASA's POWER project repository [35] provides data for solar parameters on a daily, hourly or annual/monthly basis from as far as 1981 till present day. The locations can be selected by providing the latitude and longitude. Data is collected for four regions: Location 1 with 23.25991, 77.41261 (Bhopal, India) as latitude and longitude, Location 2 with 22.71961, 75.85771 (Indore, India) latitude and longitude, Location 3 with 24.470306, 81.576251 latitude and longitude (Rewa, India) and Location 4 with latitude and longitude 28.10001, 72.30181 (Bhadla, India) for a period of approximately 36 years. The dataset consists of approximately 13000 records obtained from NASA's (NASA (2020 (accessed July 20, 2020)) POWER project database, which consists of the solar radiation incident on a horizontal surface. The record comprises a time location (Year, Month, Day) and spatial location (Latitude and Longitude). Rajasthan, the largest state of India, receives maximum intensity of solar radiation, while states like Madhya Pradesh have great potential for utilizing solar energy due to their location. Therefore, data from these regions are used for experimental purposes.

4 Results and discussion

The results obtained from the experiments performed for forecasting solar irradiance utilizing multivariate data from hexagon gridding mechanism are depicted in this section. The

Table 2 Descriptive Statistics of Dataset for Location 1 and Location 2

Statistic	Value for Location 1	Value for Location 2
Total Observations	13,000	13,000
Date Range	1983–2019	1983–2019
Minimum	0.2	0.16
Maximum	8.41	8.66
Mean	5.09	5.17
Standard Deviation	1.42	1.38

Table 3 Descriptive statistics of dataset for Location 3 and Location 4

Statistic	Value for Location 3	Value for Location 4
Total Observations	13,150	13,150
Date Range	1984–2020	1984–2020
Minimum	0.170000	0.280000
Maximum	8.010000	8.580000
Mean	5.001918	5.180071
Standard Deviation	1.370618	1.392099

existing models i.e. LSTM, GRU, Bidirectional LSTM and CNN-LSTM reported in [14, 29, 38, 43, 51] are implemented on different datasets. Also, different feature selection mechanisms are implemented on the multivariate data in these works. Therefore, in order to attain unbiased performance comparison of these models with our proposed work, the existing models along with the proposed models are implemented on the same datasets with similar feature selection mechanism. The feature selection mechanism proposed in this section firstly consists of location identification utilizing the hexagon gridding system and then selecting the relevant sites with correlation scores of Pearson, Spearman, Kendall, DTW distance and XGBoost methods. The forecast results are also compared with results obtained from previous works on univariate data.

Firstly, the feature importance analysis is discussed for the multiple sites in relation to the target location. The importance scores are calculated for every site in correspondence to the target location and utilizing these scores the sites are ranked. Based on these rankings, different combinations of the sites are obtained. The model performance is then tested for selecting the best combination of features for all the target locations. The proposed hybrid models are compared to several other models and their performance metrics are shown.

Feature selection is an important task to eliminate features not important enough in forecasting target data. For modeling purpose, feature selection helps in avoiding overfitting, maintenance of memory and reducing computational cost. The proposed methodology utilizes data for solar irradiance from multiple sites in order to forecast for target location. Therefore, for efficient selection of the most influential features, importance score of each of the variables was determined.

Tables 4, 5, 6 and 7 show feature importance scores based on Pearson correlation, Spearman correlation, Kendall correlation, Dynamic Time Warping distance and XGBoost scoring for all the locations, respectively. The most relevant sites' data were selected under feature selection. Also, the table indicates the relation for 9, 8, 10 and 11 sites for Location 1, 2, 3 and 4 respectively. This was so, since, in the dataset provided by NASA's repository, out of the total selected enclosed site data, some of the sites reflected similar data. As such, to remove redundancy, those sites were not considered further. Therefore, the number of sites are different for the selected locations.

After the scores are calculated for all the sites corresponding to the target location, these are ranked. Tables 8 and 9 show the sites selected according to the rankings. The ranking is based on the correlation and feature importance values. For location 1, we have five different combinations. $Correlation_1$ denotes the combination based on the Kendall, Pearson and Spearman scores. $Correlation_2$ depicts the ranking based on the DTW and $Correlation_3$ denotes XGBoost feature importance. Then, the "union" depicts the combination of these three correlations. The combination "all" denotes all the selected

Table 4 Multi-site feature importance for Location 1

Sites	Kendall	Pearson	Spearman	DTW	XGBoost
Site 1	79.18	92.11	92.50	50.1462	35.46
Site 2	77.81	90.82	91.62	52.3369	23.15
Site 3	79.50	92.07	92.58	48.8890	25.91
Site 4	72.52	86.21	87.72	56.8826	14.15
Site 5	73.21	87.57	88.52	57.4561	11.13
Site 6	74.67	87.87	89.26	54.5266	18.22
Site 7	70.88	84.53	86.34	57.9247	20.28
Site 8	72.13	85.96	87.57	57.2281	15.15
Site 9	77.12	90.00	91.16	52.5638	21.51

Table 5 Multi-site feature importance for Location 2

Sites	Kendall	Pearson	Spearman	DTW	XGBoost
Site 1	78.61	91.80	92.01	47.6216	24.90
Site 2	74.29	88.23	88.90	53.2719	17.18
Site 3	79.13	92.25	92.34	47.7294	28.79
Site 4	79.50	92.84	92.85	45.4014	31.59
Site 5	71.54	85.35	86.85	56.9974	20.23
Site 6	72.33	86.73	87.80	53.5003	19.20
Site 7	78.81	91.55	92.09	48.1790	23.83
Site 8	74.37	87.82	89.11	53.2571	21.25

Table 6 Multi-site feature importance for Location 3

Sites	Kendall	Pearson	Spearman	DTW	XGBoost
Site 1	65.79	80.33	81.24	64.9824	19.17
Site 2	73.71	88.59	89.80	55.1607	30.91
Site 3	73.45	87.42	89.09	56.0488	26.86
Site 4	71.99	86.51	87.91	58.0971	18.16
Site 5	75.15	89.82	90.43	54.2944	28.16
Site 6	71.04	86.08	87.00	58.1367	21.32
Site 7	69.25	84.57	86.18	59.6273	20.19
Site 8	79.83	92.76	93.20	46.7653	44.56
Site 9	69.94	83.97	85.69	57.9506	24.50
Site 10	79.22	92.52	92.85	47.8552	32.13

Table 7 Multi-site feature importance for Location 4

Sites	Kendall	Pearson	Spearman	DTW	XGBoost
Site 1	71.78	86.31	87.76	51.8487	20.14
Site 2	70.03	84.01	85.94	54.9354	19.12
Site 3	71.29	85.77	87.44	51.2778	22.25
Site 4	65.95	80.26	82.59	59.4847	21.17
Site 5	75.15	89.10	90.16	47.7146	24.33
Site 6	71.62	87.30	88.32	53.2410	23.30
Site 7	78.73	91.94	92.55	44.6313	27.42
Site 8	79.09	92.38	92.93	44.5774	29.21
Site 9	73.78	88.51	89.50	49.0666	25.33
Site 10	80.09	93.11	93.53	41.8541	31.93
Site 11	80.84	93.38	93.83	42.8047	38.75

neighboring sites taken in consideration. For location 2, four different combinations are selected. $Correlation_1$ denotes the combination based on the Kendall, Pearson, Spearman and DTW scores. $Correlation_2$ depicts the ranking based on the XGBoost feature importance. Then, the "union" depicts the combination of the correlations. For location 3, we have four different combinations. $Correlation_1$ denotes the combination based on the Kendall, Pearson and Spearman scores. $Correlation_2$ depicts the ranking based on the DTW and XGBoost feature importance. Then, the "union" depicts the combination of the two correlations, $Correlation_1$ and $Correlation_2$. The combination "all" denotes all the selected neighboring sites taken in consideration. In case of Location 4, only two combinations are present since all the correlation scores indicate similar ranking. This combination is proposed for selecting the best combination of the multiple sites instead of random selection. This ensures that only highly relevant features are considered for forecasting target solar irradiance data.

Tables 10, 11, 12 and 13 represent the results of the models for a horizon of 1, 4 and 10 days ahead forecasting. The developed models are compared on the basis of the MSE and RMSE metrics. The existing models of solar irradiance forecasting i.e., LSTM, GRU, Bidirectional LSTM, CNN-LSTM, Attention LSTM and CNN-Attention methods reported in [8, 9, 14, 29, 38, 43, 51] are implemented on different datasets. Therefore, the performance comparison of these models with our proposed work would have been biased if all the models were not trained on the same data. As such the existing and the proposed models are trained with data from the four specified locations. The models are then tested for performance comparison. Tables 10–13 prove the superiority of the proposed model. Standard

Table 8 Multi-location feature combinations for Location 1 and 2

Combination	Location 1 Sites	Location 2 Sites
$Correlation_1$	1, 2, 3, 5, 6, 9	1, 2, 3, 4, 7, 8
$Correlation_2$	1, 2, 3, 4, 6, 9	1, 3, 4, 5, 7, 8
$Correlation_3$	1, 2, 3, 6, 7, 9	
Union	1, 2, 3, 4, 5, 6, 7, 9	1, 2, 3, 4, 5, 7, 8
All	1, 2, 3, 4, 5, 6, 7, 8, 9	1, 2, 3, 4, 5, 6, 7, 8

Table 9 Multi-location feature combinations for Location 3 and 4

Combination	Location 3 Sites	Location 4 Sites
$Correlation_1$	2, 3, 4, 5, 8, 10	5, 7, 8, 9, 10, 11
$Correlation_2$	2, 3, 4, 8, 9, 10	
Union	2, 3, 4, 5, 8, 9, 10	
All	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

persistence model was also implemented for the locations. The persistence model is not able to forecast the target data efficiently. The additive combination of the linear and non-linear component was also considered for modeling the hybrid model. However, it was found that this form of combination might lead to some of the characteristics of data being lost. The model was not able to detect some part of the data generating process which might be either noise or some form of non-linearity. The proposed model combines the residual component along with the forecast component and input features for attaining accurate forecast performance.

In case of Location 1 it is the "union" combination of the proposed model which performs the best for a horizon of 10 days ahead while for horizon length of 1 and 4, "all" combination of features perform the best. Utilizing all the multi-site data always is also an option, but it leads to unnecessary overhead and complexity which can be reduced by utilizing only the relevant components. In case of Location 2, for 1 day ahead forecasting "all" combination performs the best while for 4 and 10 days ahead forecasting the "union" combination works better than other models. For Location 3, "all" combination performs the best for 1 day ahead forecasts. For a horizon length of 4 and 10 days, the $Correlation_1$ performs better than other combinations. Lastly, for Location 4, "all" combination performs the best in all horizons. Thus, the hexagon gridding system was capable of identifying relevant sites for forecasting target solar irradiance data. This proves the significance of the multi-site data selection and the residual ensemble based modeling for forecasting irradiance data of a particular location.

Table 14 shows the comparison of our proposed model with two of our previous works. Model 1 refers to the univariate form of forecasting discussed in [8], Model 2 refers to the multivariate form of forecasting proposed in [7] and Model 3 is our proposed work. It can be observed that the current model proposed performs optimally out of the three.

Table 10 Performance evaluation for Location 1

Model	1 day		4 days		10 days	
	MSE	RMSE	MSE	RMSE	MSE	RMSE
LSTM	8.268	8.339	11.491	10.122	13.876	11.131
GRU	7.980	8.293	11.543	10.175	14.797	11.546
Bidir-LSTM	7.681	8.142	11.459	10.061	13.521	10.872
CNN-LSTM	8.001	8.303	11.578	10.153	14.137	12.049
Attention LSTM	8.068	8.485	11.319	10.046	15.014	11.569
CNN-Attention	8.422	8.517	10.959	10.024	13.167	10.683
Hybrid Ensemble	**7.022**	**5.067**	**10.059**	**9.561**	**12.547**	**9.791**

Boldface entries represent the proposed hybrid ensemble model's result

Table 11 Performance Evaluation for Location 2

Model	1 day		4 days		10 days	
	MSE	RMSE	MSE	RMSE	MSE	RMSE
LSTM	6.433	7.450	9.212	8.995	11.188	9.918
GRU	6.861	7.691	9.370	9.055	11.558	10.078
Bidir-LSTM	6.365	7.412	9.237	8.975	11.118	9.878
CNN-LSTM	7.072	7.808	9.635	9.175	13.698	10.978
Attention LSTM	6.386	7.423	9.398	9.065	11.028	9.848
CNN-Attention	6.716	8.154	9.376	9.055	14.038	11.118
Hybrid Ensemble	**5.576**	**5.907**	**8.266**	**7.835**	**10.448**	**9.548**

Boldface entries represent the proposed hybrid ensemble model's result

Table 12 Performance evaluation for Location 3

Model	1 day		4 days		10 days	
	MSE	RMSE	MSE	RMSE	MSE	RMSE
LSTM	11.735	10.565	14.606	11.505	15.694	12.583
GRU	11.719	10.556	14.839	11.594	16.197	12.775
Bidir-LSTM	11.691	10.544	14.758	11.563	16.130	12.752
CNN-LSTM	11.517	10.464	15.342	11.785	16.304	11.813
Attention LSTM	11.824	10.593	14.744	11.558	16.215	11.782
CNN-Attention	11.785	10.587	14.242	11.747	16.406	11.849
Hybrid Ensemble	**10.722**	**9.459**	**12.624**	**10.462**	**14.518**	**11.677**

Boldface entries represent the proposed hybrid ensemble model's result

Table 13 Performance evaluation for Location 4

Model	1 day		4 days		10 days	
	MSE	RMSE	MSE	RMSE	MSE	RMSE
LSTM	3.283	5.229	5.075	6.6074	5.603	7.054
GRU	3.093	5.085	4.863	6.473	5.995	7.288
Bidir-LSTM	3.402	5.317	5.117	6.633	5.808	7.177
CNN-LSTM	3.317	5.254	4.708	6.373	5.565	7.032
Attention LSTM	3.201	5.167	4.862	6.473	5.539	7.016
CNN-Attention	3.221	5.183	5.210	6.691	5.402	6.932
Hybrid Ensemble	**2.782**	**4.211**	**3.079**	**5.152**	**3.574**	**6.031**

Boldface entries represent the proposed hybrid ensemble model's result

Table 14 Comparison to previous work

Dataset	Model	1 day		4 days		10 days	
		MSE	RMSE	MSE	RMSE	MSE	RMSE
Location 1	Model 1 [8]	8.12	8.74	12.02	10.60	15.00	12.15
	Model 2 [7]	7.23	5.36	10.53	9.78	12.99	10.08
	Model 3	**7.02**	**5.06**	**10.05**	**9.56**	**12.54**	9.79
Location 2	Model 1 [8]	6.05	7.73	9.56	9.87	12.27	11.00
	Model 2 [7]	5.82	6.21	8.56	8.11	10.87	9.86
	Model 3	**5.57**	**5.90**	**8.26**	**7.83**	**10.44**	**9.54**
Location 3	Model 1 [8]	11.65	10.84	14.25	11.69	15.79	12.54
	Model 2 [7]	10.95	9.83	12.87	10.94	14.90	11.89
	Model 3	**10.72**	**9.45**	**12.62**	**10.46**	**14.51**	**11.67**
Location 4	Model 1 [8]	3.48	5.76	4.21	6.52	4.92	6.73
	Model 2 [7]	2.99	4.67	3.39	5.53	3.80	6.24
	Model 3	**2.78**	**4.21**	**3.07**	**5.15**	**3.57**	**6.03**

Boldface entries represent the proposed hybrid ensemble model's result

Other conventional hybrid ensemble models [4, 11, 25, 54] were also implemented and tested for significance in solar irradiance forecasting. It was found that the traditional statistical autoregressive models are not able to accurately predict future values. The model is not able to capture the non-linear characteristics with these models. Also, other forms of residual ensemble such as the additive modeling or the utilization of moving average components was not able to detect the data generation process. The proposed residual ensemble model could outperform these models. The results also suggest that multiple neighboring sites data indeed act as important features in forecasting target solar irradiance data. The proposed method of site selection is capable of extracting relevant site data which is further enhanced by the correlation and feature extraction ranking mechanism. Overall, the proposed multivariate forecast mechanism with hexagon gridding system proves to be an efficient solar irradiance forecast model.

5 Conclusion

Solar irradiance forecasting is performed in existing literature through statistical, machine learning, or hybrid models by utilizing features only from the target location's data. While experimenting, it has been observed that data from neighboring sites leads to improvement upon performance of a single target location's data. The proposed site selection method can extract multiple neighboring sites covering maximum area without overlapping, and no place is left out. Instead of utilizing all the selected neighboring sites, which would lead to unnecessary overhead and an increase in model complexity, the different feature combinations made through correlation could conclude upon the best performing model. In this work, correlation scores and feature importance values help in the identification of relevant sites. It has been found that the Pearson, Spearman and Kendall correlations, the dynamic time warping distance, and the XGBoost feature importance are capable of selecting the most relevant sites for forecasting.

Further, the data generating process cannot be easily identified by a single model. The hybrid residual ensemble inspired mechanism is proposed for the identification of data characteristics. LSTM, with its robust gating mechanism, was able to model the temporal and long term dependencies in the residual data characteristic. The experiments were performed on datasets from four solar power plants in India for evaluation. The proposed model shows an improvement in forecast performance for the four target locations in different horizons of one day ahead, four days ahead and ten days ahead. The improvement is by approximately 2.5 percent in prediction error when compared to previous works. Experimentally, we have proved that the proposed model improved forecasting performance, which can be used commercially for analyzing regions for solar irradiance data for future solar power plant planning.

Acknowledgements The data used in the research were obtained from the NASA Langley Research Center (LaRC) POWER Project funded through the NASA Earth Science/Applied Science Program. The dataset is available at the website https://power.larc.nasa.gov/data-access-viewer/.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Aburto L, Weber R (2007) Improved supply chain management based on hybrid demand forecasts. Appl Soft Comput J 7(1):136–144. https://doi.org/10.1016/j.asoc.2005.06.001
2. Alanazi M, Mahoor M, Khodaei A (2017) Two-stage hybrid day-ahead solar forecasting. In: 2017 North american power symposium, NAPS 2017, https://doi.org/10.1109/NAPS.2017.8107319
3. Amrouche B, Sicot L, Guessoum A, Belhamel M (2013) Experimental analysis of the maximum power point's properties for four photovoltaic modules from different technologies: Monocrystalline and polycrystalline silicon CIS and CdTe. https://doi.org/10.1016/j.solmat.2013.08.010
4. Babu CN, Reddy BE (2014) A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data. Appl Soft Comput 23:27–38. https://doi.org/10.1016/j.asoc.2014.05.028. https://www.sciencedirect.com/science/article/pii/S1568494614002555
5. Belaid S, Mellit A (2016) Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate. Energy Convers Manag 118:105–118. https://doi.org/10.1016/j.enconman.2016.03.082
6. Box GEP, Jenkins GM, Reinsel GC, Ljung GM (2015) Time Series Analysis: forecasting & Control. https://doi.org/10.1016/j.ijforecast.2004.02.001
7. Brahma B, Wadhvani R (2020) Solar irradiance forecasting based on deep learning methodologies and multi-site data. Symmetry 12(11):1–20. https://doi.org/10.3390/sym12111830
8. Brahma B, Wadhvani R (2021) Visualizing solar irradiance data in ArcGIS and forecasting based on a novel deep neural network mechanism. Multimedia Tools Appl. https://doi.org/10.1007/s11042-021-11025-5
9. Brahma B, Wadhvani R, Shukla S (2021) Attention mechanism for developing wind speed and solar irradiance forecasting models. Wind Eng 45(6):1422–1432. https://doi.org/10.1177/0309524X20981885
10. Brockwell PJ, Davis RA (2002) Introduction to Time Series and Forecasting - Second Edition
11. Büyükşahin Ü. Ç., Ertekin Ş. (2019) Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. Neurocomputing 361:151–163. https://doi.org/10.1016/j.neucom.2019.05.099. https://www.sciencedirect.com/science/article/pii/S0925231219309178
12. Creal D, Koopman SJ, Lucas A (2013) Generalized autoregressive score models with applications. J Appl Econom 28(5):777–795. https://doi.org/10.1002/jae.1279
13. Gairaa K, Khellaf A, Messlem Y, Chellali F (2016) Estimation of the daily global solar radiation based on Box-Jenkins and ANN models: A combined approach. https://doi.org/10.1016/j.rser.2015.12.111
14. Ghimire S, Deo RC, Raj N, Mi J (2019) Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. Appl Energy 253. https://doi.org/10.1016/j.apenergy.2019.113541

15. Greff K, Srivastava RK, Koutnik J, Steunebrink BR, Schmidhuber J (2017) LSTM: A search space odyssey. IEEE Trans Neural Netw Learn Syst 28(10):2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924
16. Guermoui M, Melgani F, Gairaa K, Mekhalfi ML (2020) A comprehensive review of hybrid models for solar radiation forecasting. https://doi.org/10.1016/j.jclepro.2020.120357
17. Hajirahimi Z, Khashei M (2019) Hybrid structures in time series modeling and forecasting: a review. Eng Appl Artif Intell 86:83–106. https://doi.org/10.1016/j.engappai.2019.08.018
18. Heng J, Wang J, Xiao L, Lu H (2017) Research and application of a combined model based on frequent pattern growth algorithm and multi-objective optimization for solar radiation forecasting. Appl Energy 208:845–866. https://doi.org/10.1016/j.apenergy.2017.09.063
19. Hochreiter S, Schmidhuber J (1997) Long Short-Term memory. Neural Comput 9(8):1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
20. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: Theory and applications. Neurocomputing 70(1):489–501. https://doi.org/10.1016/j.neucom.2005.12.126. https://www.sciencedirect.com/science/article/pii/S0925231206000385. Neural Networks
21. Huang J, Korolkiewicz M, Agrawal M, Boland J (2013) Forecasting solar radiation on an hourly time scale using a Coupled AutoRegressive and Dynamical System (CARDS) model. Sol Energy 87(1):136–149. https://doi.org/10.1016/j.solener.2012.10.012
22. Jaeger H, Haas H (2004) Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. Science 304(5667):78–80. https://doi.org/10.1126/science.1091277. https://science.sciencemag.org/content/304/5667/78
23. Ji W, Chee KC (2011) Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. Sol Energy 85(5):808–817. https://doi.org/10.1016/j.solener.2011.01.013
24. Kariniotakis G (2017) Renewable energy forecasting: From models to applications. Renewable Energy Forecasting: From Models to Applications
25. Khashei M, Bijari M (2011) A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. Appl Soft Comput 11(2):2664–2675. https://doi.org/10.1016/j.asoc.2010.10.015. https://www.sciencedirect.com/science/article/pii/S1568494610002759. The Impact of Soft Computing for the Progress of Artificial Intelligence
26. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv:1412.6980
27. Kumar DS, Yagli GM, Kashyap M, Srinivasan D (2020) Solar irradiance resource and forecasting: a comprehensive review. IET Renew Power Gener 14(10):1641–1656. https://doi.org/10.1049/iet-rpg.2019.1227
28. Li Y, Su Y, Shu L (2014) An ARMAX model for forecasting the power output of a grid connected photovoltaic system. Renew Energy 66:78–89. https://doi.org/10.1016/j.renene.2013.11.067
29. Liu Y, Qin H, Zhang Z, Pei S, Wang C, Yu X, Jiang Z, Zhou J (2019) Ensemble spatiotemporal forecasting of solar irradiation using variational Bayesian convolutional gate recurrent unit network. Appl Energy 253. https://doi.org/10.1016/j.apenergy.2019.113596
30. Mateo F, Carrasco JJ, Sellami A, Millán-Giraldo M., Domínguez M., Soria-Olivas E (2013) Machine learning methods to forecast temperature in buildings. Expert Syst Appl 40(4):1061–1068. https://doi.org/10.1016/j.eswa.2012.08.030
31. Mellit A, Pavan AM (2010) A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. Sol Energy 84(5):807–821. https://doi.org/10.1016/j.solener.2010.02.006
32. Mukaram MZ, Yusof F (2017) Solar radiation forecast using hybrid SARIMA and ANN model
33. Mukhoty BP, Maurya V, Shukla SK (2019) Sequence to sequence deep learning models for solar irradiation forecasting. In: 2019 IEEE Milan PowerTech, PowerTech 2019, https://doi.org/10.1109/PTC.2019.8810645
34. Narvaez G, Giraldo LF, Bressan M, Pantoja A (2021) Machine learning for site-adaptation and solar radiation forecasting. Renew Energy 167:333–342. https://doi.org/10.1016/j.renene.2020.11.089. https://www.sciencedirect.com/science/article/pii/S0960148120318395
35. (2020) NASA POWER project dataset from Renewable Energy archive. https://power.larc.nasa.gov. Accessed 20 July 2020
36. Neves C, Fernandes C, Hoeltgebaum H (2017) Five different distributions for the Lee–Carter model of mortality forecasting: A comparison using GAS models. Insur Math Econ 75:48–57. https://doi.org/10.1016/j.insmatheco.2017.04.004
37. Niska H, Hiltunen T, Karppinen A, Ruuskanen J, Kolehmainen M (2004) Evolving the neural network model for forecasting air pollution time series. Eng Appl Artif Intell 17(2):159–167. https://doi.org/10.1016/j.engappai.2004.02.002. https://www.sciencedirect.com/science/article/pii/S0952197604000119. Intelligent Control and Signal Processing

38. Qing X, Niu Y (2018) Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 148:461–468. https://doi.org/10.1016/j.energy.2018.01.177
39. Rao KDVK, Premalatha M, Naveen C (2018) Analysis of different combinations of meteorological parameters in predicting the horizontal global solar radiation with ANN approach: a case study. Renew Sust Energ Rev 91:248–258. https://doi.org/10.1016/j.rser.2018.03.096
40. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536. https://doi.org/10.1038/323533a0
41. Shamshirband S, Mohammadi K, Piri J, Petković D., Karim A (2016) Hybrid auto-regressive neural network model for estimating global solar radiation in Bandar Abbas, Iran. Environ Earth Sci 75(2):1–12. https://doi.org/10.1007/s12665-015-4970-x
42. Sharma A, Kakkar A (2018) Forecasting daily global solar irradiance generation using machine learning. Renew Sust Energ Rev 82:2254–2269. https://doi.org/10.1016/j.rser.2017.08.066. http://www.sciencedirect.com/science/article/pii/S1364032117312121
43. Srivastava S, Lessmann S (2018) A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Sol Energy 162:232–247. https://doi.org/10.1016/j.solener.2018.01.005
44. Sun H, Yan D, Zhao N, Zhou J (2015) Empirical investigation on modeling solar radiation series with ARMA-GARCH models. Energy Convers Manag 92:385–395. https://doi.org/10.1016/j.enconman.2014.12.072
45. Togrul IT, Onat E (2000) A comparison of estimated and measured values of solar radiation in Elazig, Turkey. Renew Energy 20(2):243–252. https://doi.org/10.1016/S0960-1481(99)00099-3
46. Tsay RS (2005) Analysis of Financial Time Series Second Edition. https://doi.org/10.1002/0471264105
47. Voyant C, Muselli M, Paoli C, Nivet ML (2013) Hybrid methodology for hourly global radiation forecasting in Mediterranean area. Renew Energy 53:1–11. https://doi.org/10.1016/j.renene.2012.10.049
48. Wang L, Li X, Bai Y (2018) Short-term wind speed prediction using an extreme learning machine model with error correction. Energy Convers Manag 162:239–250. https://doi.org/10.1016/j.enconman.2018.02.015. https://www.sciencedirect.com/science/article/pii/S0196890418301110
49. Wang L, Zou H, Su J, Li L, Chaudhry S (2013) An ARIMA-ANN hybrid model for time series forecasting. Syst Res Behav Sci 30(3):244–259. https://doi.org/10.1002/sres.2179
50. Wang X, Han M (2015) Improved extreme learning machine for multivariate time series online sequential prediction. Eng Appl Artif Intell 40:28–36. https://doi.org/10.1016/j.engappai.2014.12.013. https://www.sciencedirect.com/science/article/pii/S0952197614003054
51. Wojtkiewicz J, Hosseini M, Gottumukkala R, Chambers TL (2019) Hour-ahead solar irradiance forecasting using multivariate gated recurrent units. Energies 12(21). https://doi.org/10.3390/en12214055
52. Yacef R, Mellit A, Belaid S, Şen Z. (2014) New combined models for estimating daily global solar radiation from measured air temperature in semi-arid climates: Application in Ghardaïa, Algeria. Energy Convers Manag 79:606–615. https://doi.org/10.1016/j.enconman.2013.12.057
53. Zemouri R, Racoceanu D, Zerhouni N (2003) Recurrent radial basis function network for time-series prediction. Eng Appl Artif Intell 16(5):453–463. https://doi.org/10.1016/S0952-1976(03)00063-0. https://www.sciencedirect.com/science/article/pii/S0952197603000630
54. Zhang PG (2003) Time series forecasting using a hybrid ARIMA and neural network model. https://doi.org/10.1016/S0925-2312(01)00702-0
55. Zhou Y, Zhou N, Gong L, Jiang M (2020) Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine. Energy 204:117894. https://doi.org/10.1016/j.energy.2020.117894. https://www.sciencedirect.com/science/article/pii/S036054422031001X