ORIGINAL PAPER

Enhanced renewable power and load forecasting using RF-XGBoost stacked ensemble

Rita Banik[1] · Ankur Biswas[2]

Abstract

The fast growth in renewable power generation, crucial for reducing carbon emissions in the traditional energy system, is constrained by negative environmental and economic repercussions, demanding a smarter integration with conventional energy sources. However, the seamless integration of renewable energy into grid imposes major challenges due to the inherent environment dependence of energy production, necessitating proactive forecasts to anticipate fluctuations in efficiency and dependability that might have an influence on the overall grid status and living style of the users. The objective of the work is to add a consistent dataset in the machine learning community and a new ensemble to deal with the problems pertaining to the projection and integration of renewable energy using predictive modeling. The model includes a stacked ensemble of Random Forest and XGBoost algorithms. The proposed ensemble has been successfully applied to the dataset of Agartala City. In this model, the Random Forest model is used to forecast the target variable based on input parameters, followed by the use of the XGBoost model to improve predictions through a combination of Random Forest predictions, with the goal of leveraging the strengths of multiple models while mitigating their weaknesses. The meta-model, a basic logistic regression, then learns the best combination of these predictions, allowing for the maximum potential accuracy. The model has been evaluated on R^2 and RMSE. The accuracy of 99% reveals its feasibility and superiority both in short-term and long-term predictions.

Keywords Renewable power · Forecasting · Stacked ensemble · Random Forest · XGBoost

1 Introduction

In terms of energy sustainability and safety, the world is facing many challenges. These can result in political and economic instability if they are not tackled instantly. The degradation of the deposits of fossil fuels and the effect of combustion on the environment has resulted in a growing interest in the invention of sustainable alternative renewable energy sources. The energy sector has achieved strong growth in alternate energy sources like solar photovoltaic, wave, wind, and biomass [1]. Several nations and enterprises aim to broaden their energy supply through maximizing renewable energy proportions. In the traditional system, energy generation relies on consumers energy demand, and grid

efficiency is based on the balance of energy requirements and distribution. When demand for energy exceeds energy supply, the grid is destabilized and the power reliability in some areas of the grid decreases and causes blackouts. If the demand is below supply, electricity is lost due to waste, which is extremely excessive. It is vital for both efficient grid operation and greater economic advantages that we produce the correct amount of energy just at the proper moment [2]. Moreover, due to the detrimental impact of climatic changes, it is more important than ever to push toward more sustainable technologies and practices. The International Energy Agency (IEA) estimates that the energy sector generates more than 65% of greenhouse gases (GHG), with CO_2 emissions increasing by 6%–36.3 billion tons in 2021, which is a sign of transformation in the sector [3]. Recent world conferences like COP21 established ambitious goals to mitigate the troubling effects of climate change, as tackled by strong regulations. Whereas 66% of GHG emissions are accounted by the energy sector worldwide, renewable sources of energy play an emerging role in decarbonization. But, the irregular

✉ Rita Banik
  ritabanik@iutripura.edu.in

1 Department of Electrical Engineering, ICFAI University Tripura, Agartala, Tripura, India

2 Department of Computer Science & Engineering, Tripura Institute of Technology Narsingarh, Agartala, Tripura, India

solar and wind energy generation patterns because of meteorological impacts made it difficult to implement extensively, and therefore, further studies and analyses are important. The main goal in the electricity grid is to guarantee a supply–demand equilibrium to prevent power outages and ensures that all users have access to the energy they require. This can be done with the deployment of energy storage systems and stand-by production capacities; however, such integration increases the grid's expenses. A significant amount of research has focused on the generation of energy and load predictions to ensure how much energy is needed to sustain the stability that could reduce the operating cost and encourage technological penetration with adequate accuracy [4, 5]. One approach to this issue is to improve the short-term forecastability and to integrate this information in intelligent control systems that can maximize power delivery within a smart grid.

Renewable supplies of energy such as solar and wind are extremely unpredictable, which can lead to a variation in the power grid's electricity production. This is due to the fact that the overall yield from these plants is dependent on the surrounding environment including solar irradiance, wind speed at different levels, cloud coverage, and other variables. Another significant constraint of renewable energy sources is their availability at a specific time such as solar radiation is accessible in daylight hours. Therefore, when resources are available, power must be generated and stored for future utilization while simultaneously using a proportion of the power produced. Solar and wind energy storage is costly, and therefore, the energy generation needs to be managed carefully. When renewable resource capacity is inadequate to satisfy the demand, traditional resources, including gas plants, are usually utilized to meet the energy deficit. Moreover, predicting renewable power generation is crucial in designing hybrid systems that integrate both renewable energy sources with DG systems effectively [6]. The stated problems have driven to the need for machine learning models to improve energy generation and usage planning. Several machine learning (ML) approaches are used in hybrid grid systems (integrated with renewable sources) based on the need and features of the challenge. For a renewable energy power grid with a substantial share of energy supply, short-term and medium-term demand must be devised. This might make the energy policy derivation easier, for instance by aiding to identify essential variables like the right level of spinning reserves and storage needs. The generation of renewable energy from power plants itself is also to be forecast, as the output of the power plants relies on several variables of the environment that cannot be regulated. This, in turn, requires forecasting external factors around the power plant area, such as solar irradiance, wind speed, and wind direction. The optimum location, length, height, and specification of renewable power plants are also crucial which is addressed by machine learning techniques regarding sustainable energy generation and integration. In general operations and administration of the intelligent grid, problems like fault identification, regulation, etc., is another field of implementation of ML techniques. The primary focus of this work is to analyze and synthesize the ML strategies in renewable power generation and the incorporation of renewable sources into existing power networks through an accurate forecasting. To achieve the goal, a stacked ensemble of Random Forest (RF) and XGBoost (XGB) ML algorithm is developed and applied on data of a hilly north-eastern state of India, Tripura, whose capital is Agartala. The integration of renewable energy in this state is extremely poor and hence could be a potential area for such integration. In this model, we first use the RF model to forecast the target variable based on the input parameters. Then, the XGB model is applied to further improve the prediction by combining the RF predictions. The purpose is to blend the strengths of multiple models while minimizing their weaknesses. By combining the predictions of the RF and XGB models, we can acquire additional information about the data and potentially enhance the all-around precision of our predictions. The meta-model which is a simple logistic regression (LR) then learns how to best combine these predictions to achieve the highest accuracy possible. The base models RF and XGB are trained independently on the same training set, with varied sets of hyperparameters and feature selection methods. Once trained, the models are applied to forecast the outcome for the test data. These forecasts are then employed as features to train the LR which is learned on the same training data, but instead of using the raw features, it uses the predicted outcomes of the RF and XGB as input. The goal of the LR is to determine the most effective way to aggregate the forecasts of the RF and XGB models in order to produce the most accurate predictions on the test set. Once the LR is trained, it is used to make predictions on new data by first passing the data through the RF and XGB to generate predictions, and then passing those predictions through the LR to get the final outcome. Overall, the RF and XGBoost (RF-XGB) ensemble model combines the strengths of both Random Forest and XGBoost to make more accurate predictions than either algorithm on its own. The accuracies above 99% reveal the potential of the proposed ensemble on limited datasets.

The arrangement of the paper is as follows: An introduction to the work was given in the first part of this paper. The following section summarizes the prior work in this domain of forecasting. The third section illustrates the methods used in the course of this study, beginning with the pre-processing of data needed before any prediction is carried out and following the methods to tune and build the predictive models. The fourth section describes the results and is divided into predictions for electricity consumption, wind power, solar power,

and evaluation of results. The concluding section presents a summary of the paper and suggests potential recommendations for future study.

2 Previous work

Several researches have been conducted to forecast the output of renewable power and the demand for electricity. Comprehensive analyses of electricity demand prediction [7], photovoltaic power generation, [8] and Wind energy [9] can be found. The optimization and comparative studies of renewable energy activities have been reviewed by Banos et al. [10] and Sfetsos [11]. The wide assumption here is that the respective statistical approaches are generally outperformed by artificial intelligence methods [12–14]. Such fields of operation are listed briefly below.

2.1 Predicting power generation from solar energy

Solar PV uses differ between household level and large solar photovoltaic power plants of 1 to 100 MW capacities. As solar PV has long been in use at a small domestic level, a number of research initiatives in recent years have been conducted to estimate the efficiency of PV using machine learning methods. The statistical strategies used for solar radiance are mostly Artificial Neural Networks (ANN) [15, 16], Support Vector Regression (SVR), Mixed ANN and SVR [17, 18], and Vector Autoregression (VAR) [19]. Numerical Weather Prediction (NWP) [20] is widely used techniques with good estimation capabilities in frequent solar irradiation and wind speed forecasts. Two methods are primarily used to forecast solar irradiance: model-based method and data-driven method. A model-based approach for estimation of solar irradiance incorporates environmental parameters such as altitude, temperature, humidity, wind pressure, etc. [21]. The prediction effects are reliable, but models will lead to limitations on their realistic use [22]. The other variant, i.e., data-driven method using ANN, relies only on the availability of solar data to create a connection between weather parameters and the generation of PV power. Research using ANN found that the PV generation could be successfully predicted by parameters such as solar radiation, azimuth angles, and dry-bulb temperatures. ANN models for the 24-h power forecast with projected meteorological variables of the next day [23] are built for multilayer perceptron feed-forward ANN. A robust one-day solar PV energy forecasting under various climatic conditions was tested in an enhanced back propagation (BP) ANN learning model [24]. A short-time forecast of PV grid has been based on hybridized support vector machine (SVM). To order to predict solar irradiance, a previous autoregressive integrated moving average (ARIMA) time series model was also implemented [25]. An autoregression

(AR) system is used for the estimation of solar irradiance from three French locations with a ML model on a 4-h forecast when weather conditions were unpredictable [26]. The fast implementation of deep learning methods is a common way of predicting photovoltaic power output with the least prognostic error [27]. Different architectures such as Boltzmann system, Deep Belief Network (DBN), and Recurrent Neural Network (RNN) are utilized. In order to predict solar irradiance without including previous performance information, an efficient Bayesian neural network model has been implemented [28]. A long, short-term memory network (LSTM), which demonstrated the better performance of the LSTM network in comparison to traditional networks, has been applied for the assessment of PV power [29]. Greater training time is the major downside of the LSTM network. The Gated Recurrent System (GRU) was implemented to resolve the constraint of training time while preserving the same exactness. The GRU network is currently mainly used for the problem area of classification [30]. Another common technique for prediction is the Random Forest (RF) algorithm which shows good results [31]. RF has been shown to be the most reliable model for PV generation prediction compared to other ML algorithms [32]. It is known, however, that ML models are not stable since their success depends on the learning dataset. A short-term prediction model consists of NWP model for weather reports, and based on the weather anticipated data an Artificial Intelligence (AI)-based model is applied using k-NN, ANN, ARIMA, and adaptive neuro-fuzzy inference system (ANFIS) to estimate 1–39 h PV capacity. Extreme learning machine (ELM), which is modified by various particle swarm optimization (PSO) processes, has been suggested. With the data from the BP-ANN prediction model [33], the performance of the process was evaluated. Because of ease, speed, and a strong capacity for generalization, it has become extremely popular [34]. The advantages of the individual model were also intended for hybrid prediction models. The ensemble of sub-models [35] is yet another effective method, which minimizes to a large extent the forecast error of a single model. Recently, it is increasingly popular to use ensemble data to determine solar irradiance and solar PV energy. To measure the global solar radiation to Spain in 1 h' time [36], RF is applied. Gala et al. [37] used a RF, SVR, and a Gradient Boosting Regression (GBR) to forecast solar irradiance, which considered the models to be important and optimistic.

2.2 Predicting power generation from wind energy

Wind power is a fast-expanding renewable energy source, and its popularity has grown in recent years as more governments and businesses strive to minimize their carbon impact. Predicting power generation from wind energy entails a thorough examination of many parameters in order to calculate

the potential generation of power. Wind speed, direction, temperature, and air pressure are all important meteorological elements in this forecast. Wind turbine features like as capacity, rotor diameter, and hub height are also important factors. Furthermore, geographical characteristics and the surrounding environment influence wind flow, which affects turbine efficiency. A study investigates the chaotic behavior in wind turbine system [38]. To connect wind speed to power output, mathematical models like as the power curve model are widely used. However, Wind turbine power output is directly related to wind speed. It has also been noticed that the threshold speed is only effective after that speed (to guarantee the security of the turbine); the power output is constant. Certain variables including temperature and humidity also impact air pressure, which in turn affects the generation of electricity. Such variables and the actual performance of a wind farm must, therefore, be predicted. Wind power can be estimated using,

$$P = A\rho V^3 C_p / 2 \qquad (1)$$

where 'A' is the area enclosed by the wind turbine, 'ρ,' 'V,' and 'C_p' are used for air density, wind speed, and efficiency factor, respectively.

Different types of forecasts are available like physical, statistical, soft computational methods, etc. Physical simulations rely on averages and long-term forecasts because different factors including temperature, the length of the sunshine precipitation, etc. influence the wind patterns [39]. Nonetheless, these methodologies require greater processing time because of substantial volume of inputs, and their precision relies on numerical weather forecast performance. Second, the objective of statistical models is to compare past (as input) data with potential (as output) forecasts of wind characteristics. Autoregressive moving average (ARMA), ARIMA, and their hybrid models are commonly employed methodologies [40]. Machine learning technologies, such as regression models or neural networks, can improve prediction accuracy even more by capturing complex correlations in data. These models may be trained using historical meteorological data and turbine performance records. To optimize the efficiency and dependability of wind power systems, forecasting power output from wind energy requires a multidimensional strategy that integrates meteorological research, turbine technology, and modern data analytics. These methods are often used to forecast renewable energy sources [41], and the most popular choice among them is the ANN. Different variations in ANN are implemented for wind and solar power prediction, and good performance has been verified with large-scale data sets. Nevertheless, in the case of data sets where data are limited for model preparation, the SVM has obtained higher prediction results [42]. Currently, extreme learning models are successful for extremely short

online predictions [43]. Deep learning strategies have been established to resolve the poor learning skills of traditional soft computing models. Because it is possible to train large data sets, it is easier for large-scale data analysis and thus for long time-series data, which is frequently observed in the renewable energy prediction field. The convolutional neural network (CNN) and the RNN possess the capacity to learn sequence similarity and therefore are used to forecast wind speed [44].

2.3 Predicting the consumption of electricity

The load signal is a time series; its future production should be estimated based on past data for load and other forecasting parameters that potentially impact the future load. Initially, statistical methodologies like regression, multiple regression, smoothing, fuzzy logic, and ML were used for forecasting. While Hyde et al. [45] and Broadwater et al. [46] have devised a methodology based on nonlinear load regression, the initial load prediction experiments demonstrated a regression using linear regression for load prediction [47]. El-Keib et al. [48] developed short-term predictive frameworks employing exponential smoothing, while Huang [49] offered an autoregressive model for forecasting short-term consumption, among other autoregressive modeling techniques. For the very short-term to long-term predictions (horizons of less than hour, hour, week, one or more months, and one to several years), several predictive designs were created. There are three categories of modeling techniques: statistical, artificial intelligence, and hybrid. In statistical models, the output and input are connected by mathematical equations directly. These models are easy to apply and work better in short-term forecasts, but they cannot account for the non-linearity of the load series; therefore, a more intelligent method is required. Statistical techniques include ARMA, ARIMA, multiple regression, linear regression, and multiple regression using ARMA [50–53]. Techniques for artificial intelligence (AI) are closed systems with unknowable internal dynamics. This group includes three important techniques: fuzzy inference system (FIS) [54], ANN [55, 56], and SVM [57]. In FIS, the link between input and output will be determined by a set of linguistic guidelines for fuzzy structures. In contrast, training determines this relationship in SVM and ANN. The empirical risk minimization concept of SVM addresses the issue of the ANN model's inability to solve local optimization issues and its tendency to be both under- and over-fitted [58, 59]. The RF [60, 61] is another well-liked strategy that relies on training. The enhancement in RF comes from its nonlinear estimating suitability and reduced sensitivity to parameter values [62]. Hybridization is an efficient way to manage the optimal architecture design and parameter adjustment needed for all AI-based methodologies. Research by Khayatian et al. [63] applied an ANN

to estimate operational energy ratings for Italian residential structures. In order to anticipate energy performance, Ascione et al. [64] looked into how energy consumption and occupant thermal comfort are related. Papadopoulos et al. [65] also assess tree-based forecasting models to forecast building energy efficiency. Wang et al. [66] investigated a recent short-term projection of electricity utilizing RF for office building factors of envelope, temperature, and time. For estimating the hourly power usage in buildings, the research demonstrated RF outperformed regression trees and SVM. Rathore et al. [67] focused on the prediction of energy consumption in electric vehicle using different ML models like RF, XGB, etc. In this study, RF and XGB outperformed all other models. In a study [68], hourly predictions of short-term load consumption were made based on meteorological parameters and public holidays of the country, and the performances of different algorithms were compared. Another study involved ensemble of RF and XGB for electrical energy prediction for the next 24 h of load with and estimation of load for one week to a month [69].

The study conducted so far have predominantly employed individual RF and XGB models. In contrast, our approach involves the utilization of a stacked ensemble comprising both RF and XGBoost models. This novel technique improves the resilience and prediction effectiveness of our model by using the complimentary characteristics of these two popular algorithms. As a result, our work adds a new viewpoint to the current body of research by highlighting the effectiveness of ensemble methodologies when compared to individual models.

3 Methodology

There are two aspects of the research methodology. The first section explains the compilation and preprocessing of data. The second part discusses the algorithms used in machine learning and their methods of implementation.

3.1 Data pre-processing

This research aims to establish a data-driven model to generate completely accurate forecasts for wind, solar, and electricity load; hence, utmost attention is paid to collect and process data. The data utilized in this work will be described in the following parts, along with the strategies used to process it. Various data sources included in this paper like temperature, humidity, pressure, wind speed above 10 m from ground level, wind direction, and solar PV power production for each form of model training and validation have been obtained from the Ninja renewables project [70]. Here, the wind power data are limited to only 1 year in normalized form at the given site which is why we used wind power data

obtained Kaggle; the information was accumulated for a time frame of 10 years. We acknowledge that there may be potential limitations associated with merging data from different locations. For example, the characteristics of the areas where the data were collected may differ in terms of population density or other environmental factors, which could impact the comparability of the data. Finally, the only parameter used for electricity consumption predictions is the past electricity load which is the hourly load of Tripura which made available by State Load Despatch Centre (SLDC), Tripura State Electricity Corporation Limited, Agartala. The data did not have any details on the position of households, and no weather data were connected with it. The data set covered the period from January 2016 to November 2019, and the electricity consumption was calculated in MW. The proposed location corresponds to Agartala, Tripura, India, latitude 23.83 degrees north and longitude 91.28 degrees east. We combined the data from the three sources into a single dataset using a common identifier for each data point. The entire dataset is thoroughly examined to remove abnormalities, outliers, and missing values. Depending on the kind of missing data, we used imputation techniques like mean imputation or average to solve this problem. The heatmap representing the correlations within the parameters is shown in Fig. 1.

3.2 Prediction architecture

For the estimation of a time series using data-driven method, three forms of models are available, viz. the input–output approach (IO), the nonlinear autoregressive approach (NAR), and the nonlinear autoregressive with exogenous inputs (NARX). The fundamental distinction is the type of information that each system accepts as input. The IO accepts all kinds of inputs except the previous target series value. The NAR method accepts the previous values only while NARX accepts both the past values and exogenous inputs in order to build the target. The NARX method can be derived to outperform the other two methods if the target variable is correlated with the exogenous inputs since it provides additional information about the system. The three forms of model representing an association between a dependent variable '$y(t + p)$' and different lagged values of either itself or other independent variables are as follows:

$$y(t + p) = f(x(t), x(t - 1), \ldots, x(t - \mathrm{d}x)) \tag{2}$$

$$y(t + p) = f(y(t), y(t - 1), \ldots, y(t - \mathrm{d}y)) \tag{3}$$

$$\begin{aligned} y(t + p) = f(x(t), x(t - 1), \ldots, x(t - \mathrm{d}x), \\ y(t), y(t - 1), \ldots, y(t - \mathrm{d}y)) \end{aligned} \tag{4}$$

Fig. 1 Pearson correlation coefficient of the target variable with atmospheric parameters
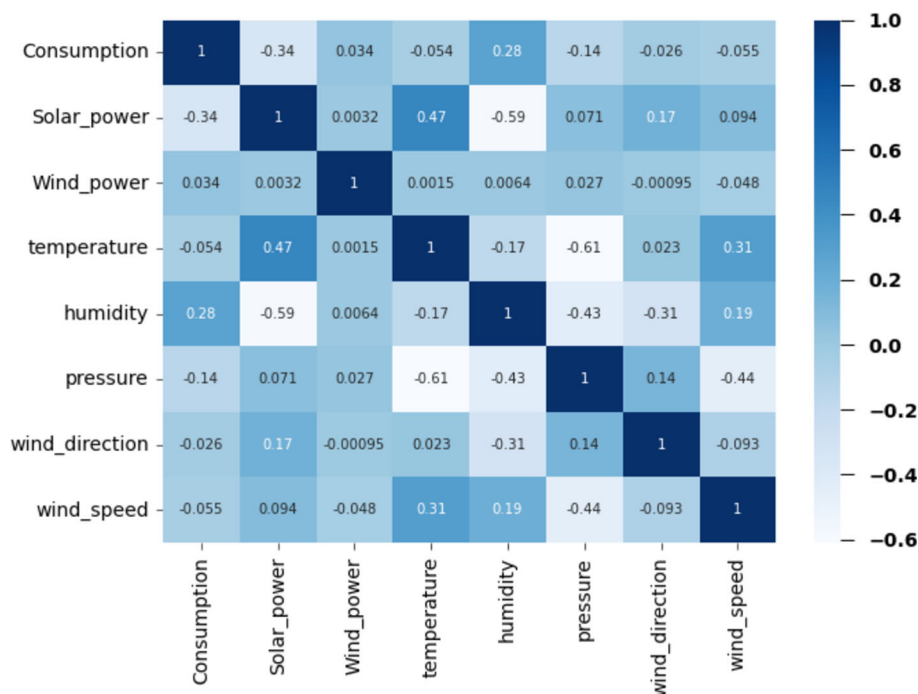
where $y(t)$ is the forecasted value, $x(t)$ is the input variables, d_y and d_x are the previous and input values of the target used for forecasting, respectively. p is the time horizon for which prediction is made. It represents the future value of $y(t + p)$ is a function of past values of a set of independent variables $x(t), x(t-1)\ldots$ or the past values of the dependent variable itself $y(t), y(t-1)\ldots$ or both.

3.3 Random Forest (RF)

The RF models act as variant of bagging techniques incorporating slight modifications. They represent an enhanced version of the bagging estimator algorithm utilizing decision trees as the base estimators. This approach involves selection of randomly chosen samples from the training data. Nevertheless, unlike bagging, where each tree has a whole collection of attributes, RF extracts a subset of features to train multiple trees for choosing the best split. This makes the trees more independent of each other, which further makes the performance of prediction better than bagging. Since each tree learns from a subset of feature attributes, it is quicker as well. On the contrary, Bagged decision trees select variables to split in a greedy manner that lessens the error. As such, the decision trees can retain some structural resemblances even with Bagging and in effect have a strong correlation in their predictions. Therefore, a combination of predictions from multiple models in the ensemble functions is better when the sub-model forecasts are uncorrelated or only very weakly correlated. To reduce the correlation among forecasts from individual sub-trees, RF modifies the sub-tree learning

method. The learning algorithm is permitted to evaluate each variable when selecting a split point for the optimal outcome. RF alters this process by restricting the learning algorithm to be tested against only a random subset of features. The ensemble learning technique known as RF involves creating random subsets of the dataset through bootstrapping, where each subset serves as the training set for an individual decision tree. At each node in the decision tree, random feature sets are chosen for the optimal split. A decision tree model is trained for each subset, and the predictions from all decision trees are averaged to generate the final prediction. This approach leverages the power of multiple models and randomization to improve the overall accuracy that helps mitigate overfitting and enhances the robustness of the prediction process.

3.4 Extreme gradient boosting (XGB)

XGB uses the principle of the original model of Gradient Boosting by Friedman [71]. It is a supervised learning problem, where the training data x_i are used to train an ensemble of 'K' trees that forecast the target attribute y' {T1 $(x_i, y_i)\ldots$ TN (x_i, y_i)} where x_i is the input vector of descriptor that is provided. The implementation of XGB through gradient boosting decision tree algorithm has proved to be a highly effective ML algorithm [72]. XGB offers superior prediction accuracy and significantly faster execution compared to traditional gradient boosting techniques. It includes L1 and L2 regularization options to mitigate overfitting and enhance overall performance. Additionally, the algorithm

incorporates sparsity-aware splitting and a weighted quantile sketch algorithm, enabling it to accommodate different types of data sparsity and handle weighted data effectively. The XGB generates a new model that analyzes the error of the earlier model and combines itself using a gradient descent method to reduce the loss in order to make the final prediction. The process of appending models continues until no additional enhancements can be made. The dataset d(x,y) with 'n' samples with 'm' features, while 'y'' is defined as the predicted value of model given by

$$y' = \sum_{j=1}^{J} f_j(x), \; f_j \in N \tag{5}$$

where 'f_j' and '$f_j(x)$' denote the tree and score given by j-th tree during forecast, respectively. The ideal leaf weight 'w_j^*'that refers to the optimal weight assigned to a leaf node in a decision tree is provided by

$$w_j^* = -\frac{G_j}{H_j + \lambda_2} \tag{6}$$

where 'G_j' represents the gradient of the loss function w.r.t. the forecasted values for the j-th leaf, 'H_j'is the second-order partial derivatives of the loss function w, and 'λ_2'is a regularization term. It can be calculated using regularization term,

$$\Phi^* = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{H_j + \lambda_2} + \lambda_2 T \tag{7}$$

where 'T' represents the total number of leaves in the decision tree or ensemble of trees.

3.5 Stacking

Stacking is a machine learning strategy whereby multiple models are combined to enhance prediction accuracy. The fundamental concept is to train numerous base models on the same dataset and then utilize their predictions as inputs to a higher-level model, known as a meta-model, which produces the final prediction. When the base models have various strengths and weaknesses, stacking can be very useful since the meta-model can learn to "average out" their flaws and provide more accurate forecasts overall. Stacking may be implemented in a variety of methods, but the most popular is to utilize k-fold cross-validation to create out-of-sample predictions for each base model. These predictions are then utilized as inputs to the meta-model. This guarantees that the meta-model is trained on new data and lowers the possibility of overfitting to the training data.

For instance, a collection of base models, M = (m_1, m_2,……, m_k) is trained on a given dataset 'D,' and the resulting predictions are gathered in a matrix P of dimension ($n \times k$), where n represents number of observations in D. A meta-model m is then trained on the matrix P and the true response variable y of D, using a weighted combination of the base model predictions as inputs:

$$y' = m(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k)_2 \tag{8}$$

where y' is the predicted response variable, p_i is the vector of predictions made by base model m_i, and α_i is the weight assigned to base model m_i by the meta-model m.

3.6 The proposed RF-XGB ensemble

The present research explores how wind and solar power generation and the stochastic behavior of electricity demand are estimated by an ensemble of RF and XGB. For many factors, these approaches were selected, either because of the good results they have shown or because of the plausibility to deliver higher precision and black-box models in different contexts [73]. The RF-XGB ensemble model combines two popular machine learning algorithms: RF and XG. The RF algorithm is an ensemble algorithm that builds several decision trees, integrates their predictions, and produces a final forecast. RF is well known for initial forecasting because of its scalability, resilience, and capability to handle diverse types of data. The XGB algorithm, on the other hand, is a gradient boosting algorithm that iteratively improves a model by adding new decision trees to the ensemble. XGB is very good at using gradient boosting to optimize predictive performance, which helps to better capture complex patterns and increase accuracy. To create the RF-XGB model, we first train a RF model on the training data. This involves constructing multiple decision trees (DT) based on random subsets of the features and training them on the training data. Each decision tree in the forest independently predicts the target variable, and the final outcome is made by averaging the forecasts of all the trees in the forest. The mathematical equation for a single decision tree is written as

$$y = f(x) \tag{9}$$

where y and x represent the target variable and feature vector, respectively. The 'f' represents the decision function that translates the feature vector to the target variable which is often written as a sequence of if–else statements. The RF model uses feature values to partition the data into subsets. Multiple DT are learned using random samples of the data, and the final prediction is attained by averaging the predictions of all the trees. The RF model's mathematical equation

is written as follows:

$$y = \frac{1}{N} \sum y_i \qquad (10)$$

where 'y_i' is the i-th decision tree's forecast and 'N' is the total number of trees in the ensemble.

The XGB is then utilized to enhance the predictions of the RF model. The predictions of the RF model are utilized as features for the XGB model learned on the training set. The XGB model learns to combine the RF model's predictions in order to provide a more accurate final result. The model trains the decision trees sequentially, with each consecutive tree aiming to fix the error of the preceding tree. The following is the mathematical equation stands for the result of XGB model,

$$y = \sum_i \cdot f_i(x), \; y_i. \qquad (11)$$

where $f_i(x)$ is the forecast of the i-th DT and y is the total of the predictions of all the trees.

Finally, the stacked ensemble model employs a LR model to combine the predictions of the two models. The process starts by using the RF model to forecast the target variable based on the input information followed by improving the forecast by XGB by merging the RF predictions,

$$y = w_0 + w_1 x + w_2 \times \text{RF}_{\text{pred}} + w_3 \times \text{XGB}_{\text{pred}} \qquad (12)$$

where y is the final prediction, x is the feature vector, RF_{pred} is the prediction of the RF model on x, XGB_{pred} is the forecast of the XGB model on the combination of x and RF_{pred}, and w_0, w_1, w_2, and w_3 are the weights learned by the LR model. The stacked ensemble model's purpose is to learn the appropriate weights for merging the RF and XGB models' predictions such that the final forecast has the lowest possible error. The stacked ensemble of RF and XGB modeled on LR is shown in Fig. 2.

The precision of the stacked ensemble is improved by integrating the capabilities of the RF and XGB models. The RF model is well-known for handling noisy data and high-dimensional feature spaces, whereas the XGB model is well known for capturing complicated relationships between features and making accurate predictions. The stacked ensemble is capable to exploit the benefits of both models and increase overall accuracy by integrating the predictions of these two models using a linear regression model. The RF model may be able to preserve a few fundamental patterns in the data, but the XGB model may be able to catch more complicated patterns. The LR model then learns how to integrate these predictions efficiently to increase overall accuracy. Additionally, the stacked ensemble model can also help to reduce overfitting by combining the predictions of multiple models. Since each model is trained on a different subset of the data, they
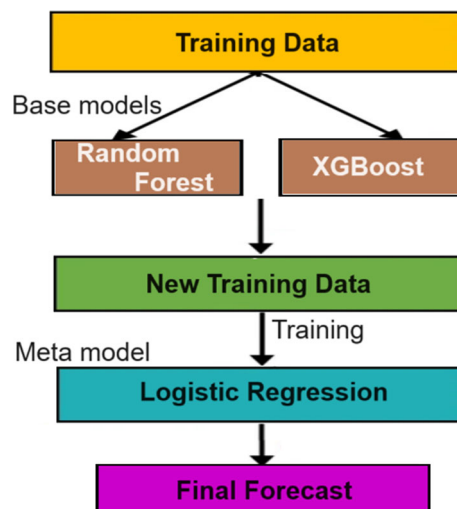

Fig. 2 Stacked ensemble of RF and XGB

may be less prone to overfitting than a single model trained on the entire dataset. The linear regression model then learns how to weight the predictions of each model to minimize the overall error on the validation set, which can further enhance the generalization ability of the model. The mathematical equation for the stacked ensemble of RF and XGB on linear regression can be represented as:

$$y' = w_1 \times y_{\text{rf}} + w_2 \times y_{\text{xgb}} + b \qquad (13)$$

where y', y_{rf}, y_{xgb} represents the predicted target value for a given input sample, RF and XGB. w_1 and w_2 are the weights assigned to the RF and XGB, respectively, learned by the linear regression model during training, and b is the bias term learned by the LR model during training.

The LR model learns the optimal weights w_1 and w_2 and bias term b during training by reducing the mean squared error (MSE) between the predicted target values and the actual target values on a validation set. The overall precision of the stacked ensemble is improved by leveraging the strengths of both the RF and XGB models, as well as by reducing overfitting through the use of multiple models and the LR model's learned weights. The general advantages of the proposed RF-XGB can be stated as:

1. Being nonparametric, it does not presume or allow data to obey any specific distribution. This reduces the time complexity in transforming data into normal distribution.
2. It can process mixed data types.
3. It includes a bias correction approach to address potential biases in renewable power and load predictions. This guarantees that the forecasting models provide impartial and reliable predictions.

4. The feature multi-collinearity does not influence model accuracy and predictive performance.

A five-fold CV technique, a common scheme for measuring model performance, was employed to confirm the model's accuracy. The investigations are run in a Python environment on a Windows GPU platform using a 16 GB RAM and 3.6 GHz Intel Core i7-4790 CPU.

4 Result and discussions

The discussed results are divided into four sections, consisting of three types of forecasts, namely electricity consumption, wind power, and solar power, followed by the last section, in which the model's effectiveness is validated. The layout of the results starts with the demonstration of different data characteristics for each prediction and follows up with the review of the results of different methods of prediction.

4.1 Electricity consumption predictions

For many factors, the forecast of electricity consumption is not similar to that of wind and solar power. Firstly, the data are collected from two independent sources. The first dataset includes data from Ninja Projects while the second dataset for electricity load is obtained from TSECL, SLDC which had no other inputs. Eventually, because this data collection has been obtained from the actual load of the state, additional data errors and ambiguity, such as instrument failures, are correlated with it. Correlation heatmap already shown earlier did not reflected major dependencies of electricity load with weather parameters. Hence, a few extra variables, such as hour, hour1 (representing peak hour of the day), week, day, month, year, etc., are appended to check the associations. To filter out unnecessary data and categorize the utmost important characteristics in the dataset, the fundamental correlation matrix with the relevance of feature is also constructed, as shown in Fig. 3.

Lagging features are historical values of a variable that are included in the feature set to forecast the present or future values of the same variable in ML and time series analysis. In time series prediction, where past measurements are used to predict future observations, lagging characteristics are frequently employed. We establish "lagged" copies of the time series in order to explore any potential serial dependencies in the data. To lag a time series is to move its values one or more-time steps ahead, or alternatively, to move the times in its index one or more steps backward. In either scenario, the delayed series' observations will appear to have occurred later in time.

By taking into consideration the impact of any intermediate lags, partial autocorrelation is a statistical concept used in time series analysis to quantify the link between an observation in a time series and its delayed values. It analyzes the correlation between two values in a time series that are separated by a certain number of time steps while accounting for any other time delays. In contrast, autocorrelation is an analytical concept employed to quantify the degree of similarity between an observation in a time series and its lagged values. In other words, it accounts the correlation among two values in a time series that are divided by a specific number of time steps. The partial autocorrelation and autocorrelation are shown in Fig. 4.

The partial autocorrelation and autocorrelation reveal the strong correlation at Lag1. Hence, lagging power of 1, 12, 24, 48, and 72 h is plotted along with rolling_4_power_mean and rolling_24_power_mean as shown in Fig. 5.

From the plotted heatmap, it has been observed that the lagged_power_1, lagged_power_24, lagged_power_48, lagged_power_72, rolling_4_power_mean, rolling_24_power_mean are the features highly correlated with the consumption. Hence, these features are also added for training the model. The model prediction ahead of 24 h is revealed in Fig. 6.

Using a straightforward procedure, energy consumption may be predicted. In addition to separating the target variable and highly correlated features, hyperparameters such as the number of estimators, learning rate, and tree depth are set before training the model which involves selecting the best values for these hyperparameters to optimize the performance of the model. The hyperparameter tuning significantly affected the overall performance. With systematic hyperparameter optimization, the model attained enhancements in its ability to extract intricate relationships and patterns in the data. Therefore, the RF and XGBoost models individually as well as in ensemble had enhanced accuracy, and minimum overfitting. Subsequently, the model was validated through a fivefold CV procedure. Lastly, a different test set may be used to assess the model's performance. This provided an additional evaluation of the model's efficiency on hypothetical data and point out any over- or underfitting problems. Figure 7 displays the comparison of consumption values at various time horizons with true and expected values of consumption using RF-XGB ensemble.

4.1.1 Wind power predictions

A renewable energy source, wind power has gained prominence recently due to its capacity to reduce greenhouse gas emissions while also contributing to a more sustainable energy future. Although wind speed and direction can vary significantly across time and space, they are a major factor in wind power generation. Accurately predicting the wind power generation is crucial for the efficient operation of wind farms and integration into the power grid. There
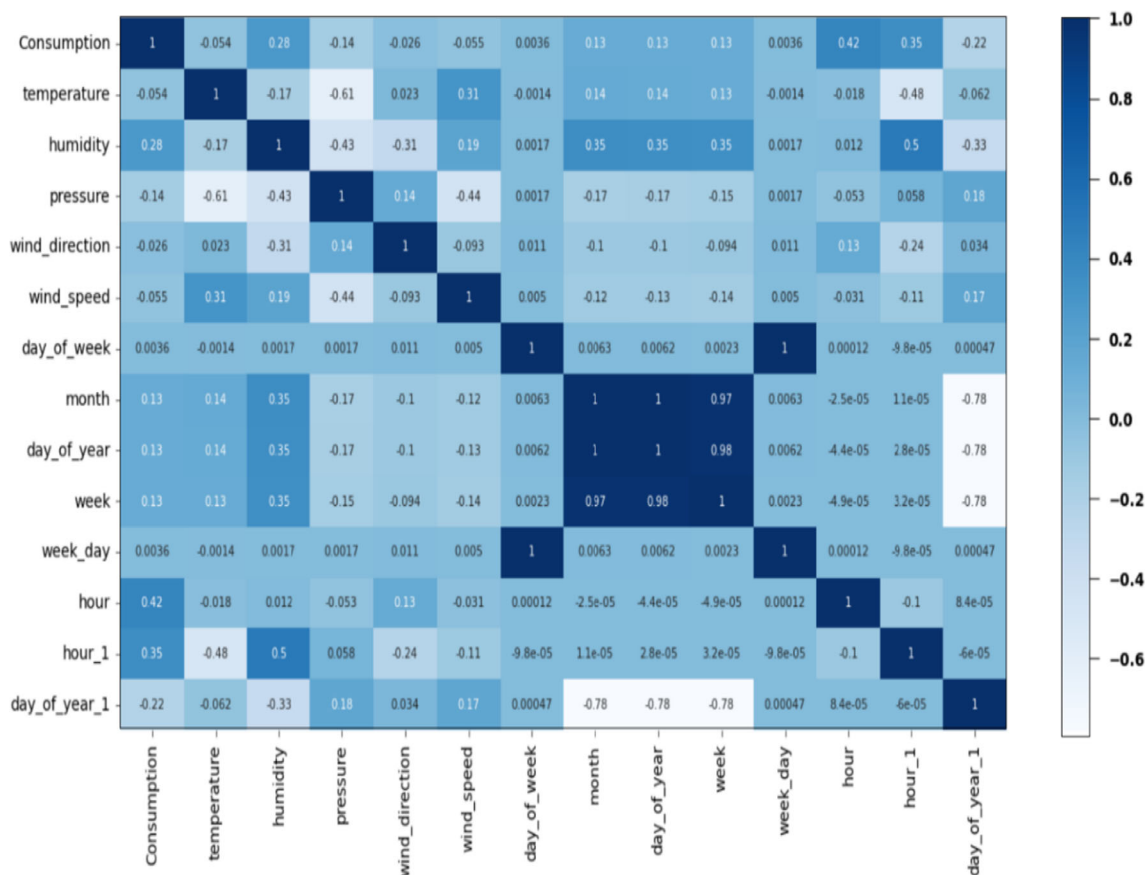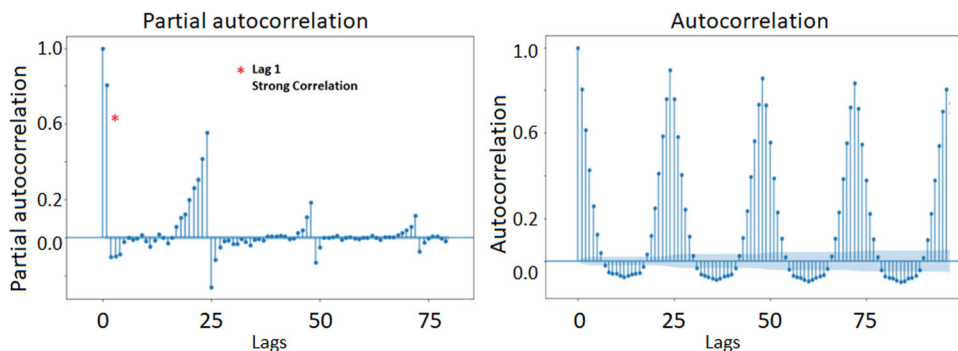
Fig. 3 Correlation coefficient of dataset parameters for prediction of consumption

Fig. 4 Partial autocorrelation and autocorrelation for prediction of consumption

are several methods for wind power prediction, including NWP, statistical models, and ML algorithms. Each method has its advantages and limitations and may be more suitable for specific applications. In the following sections, we engaged in feature engineering to extract pertinent data and provide additional features that may improve the models' predictive capacity. Finally, we plot the Pearson correlation heatmap to visualize the correlations among parameters in a dataset in the context of wind power production as shown in Fig. 8.

The heatmap does not reveal any major correlations due to mismatch of location of wind power plant and obtained

weather parameters. To address this issue, "lagged" copies of the time series are established in order to explore any potential serial dependencies in the data. Partial autocorrelation and autocorrelation statistical concept are used to quantify the link between an observation in a time series and its delayed values. It analyzes the correlation between two values in a time series that are separated by a certain number of time steps while accounting for any other time delays. The partial autocorrelation and autocorrelation on wind power are depicted in Fig. 9.

The concept reveals the strong correlation at Lag1. Hence, lagging power of 1, 12, 24, and 48 h are plotted along

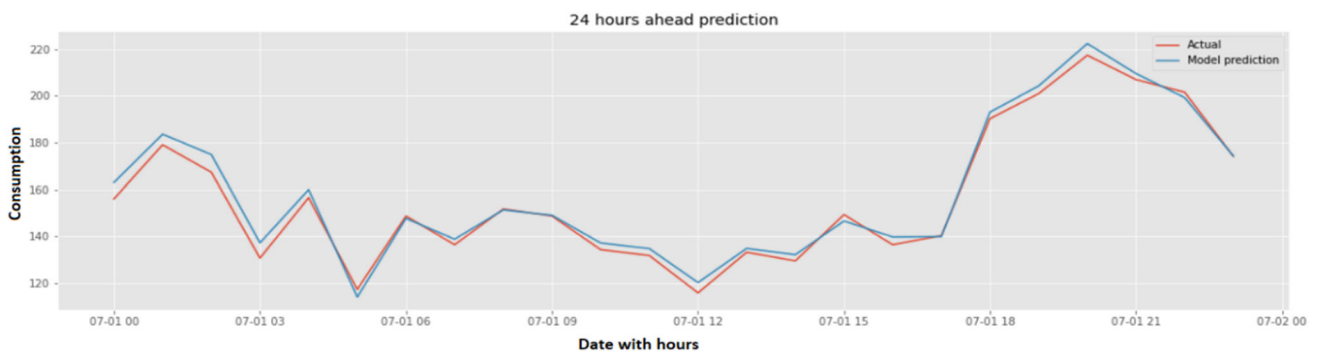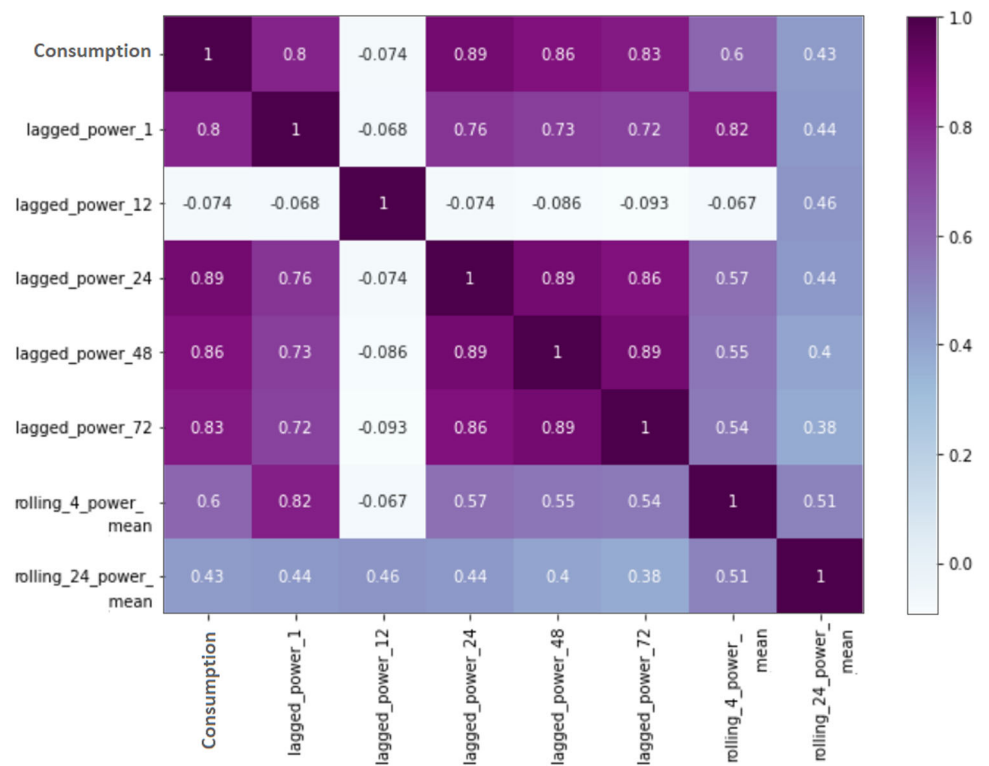**Fig. 5** Heatmap with lagged features for prediction of consumption

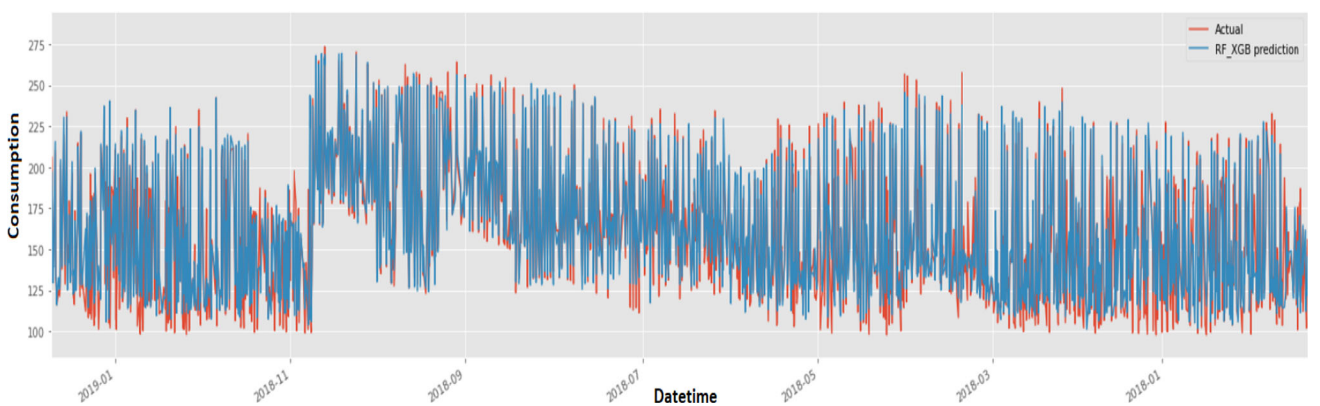**Fig. 6** Prediction of electricity consumption (24 h ahead) using RF and XGB ensemble

Fig. 7 Assessment of forecasted and true values of electricity consumption with predicted values using RF and XGB ensemble
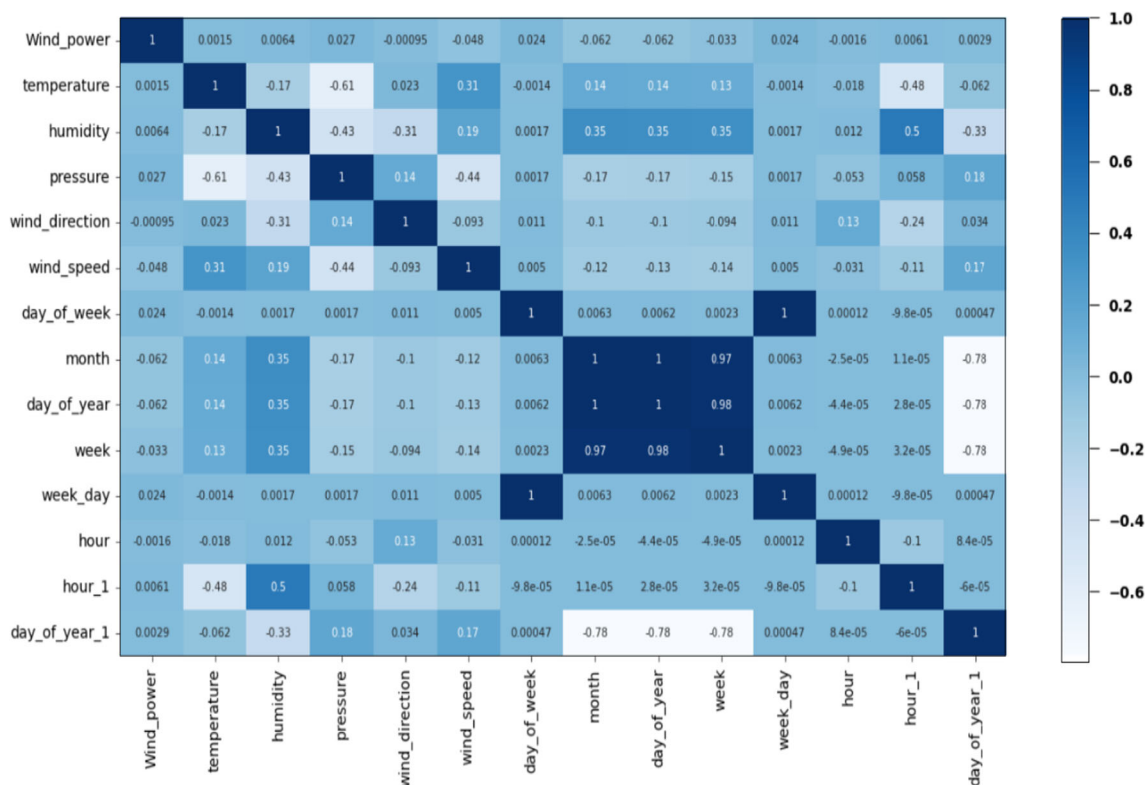
Fig. 8 Correlation coefficient of dataset parameters for prediction of wind power
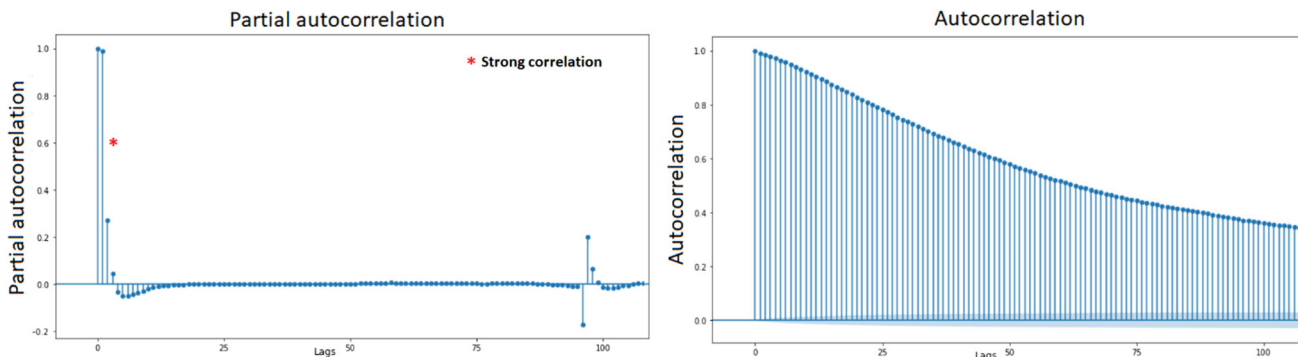

Fig. 9 Partial autocorrelation and autocorrelation for prediction of wind power

with rolling_4_power_mean and rolling_24_power_mean as shown in Fig. 10.

As the plotted heatmap reveals, there exist correlations of lagged_power_1, lagged_power_12, lagged_power_24, rolling_4_power_mean, rolling_24_power_mean with wind power. Hence, these features are also added for training the model. The model prediction ahead of 24 h is shown in Fig. 11.

A straightforward procedure is used to anticipate the wind power. The result of the correlation matrix is kept in consideration while training the model and parameters with weak correlations is dropped from the training set. The feature matrix containing highly correlated features is segregated

from the target variable, which is wind power. Subsequently, the RF-XGB model was subjected to a five-fold CV technique. The prediction of wind speed at various time intervals ahead is shown in Fig. 3. The comparison shows a very minimal variation in the target and forecasted values. Actual and projected wind power levels are compared at different time horizon as shown in Fig. 12.

4.2 Solar power predictions

For the prediction of solar power, both the approach and the steps required are almost same as those for the prediction of consumption. This is because the data (both weather

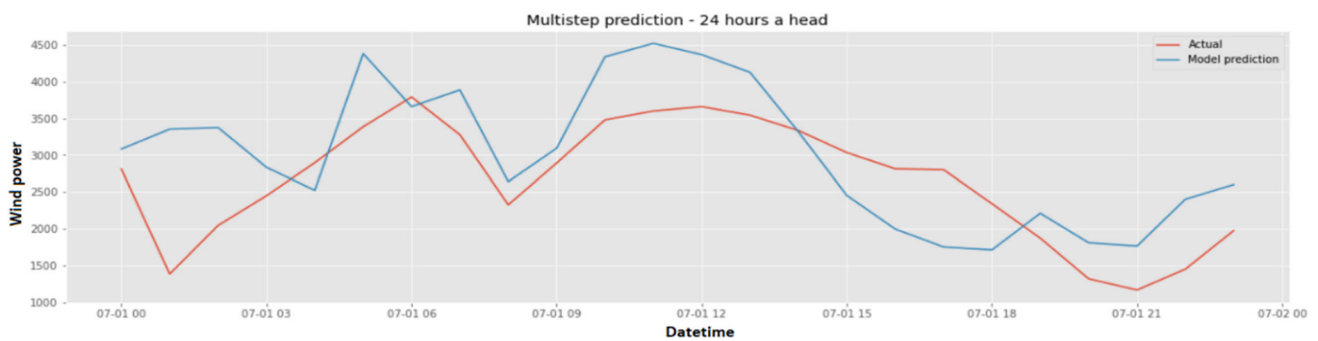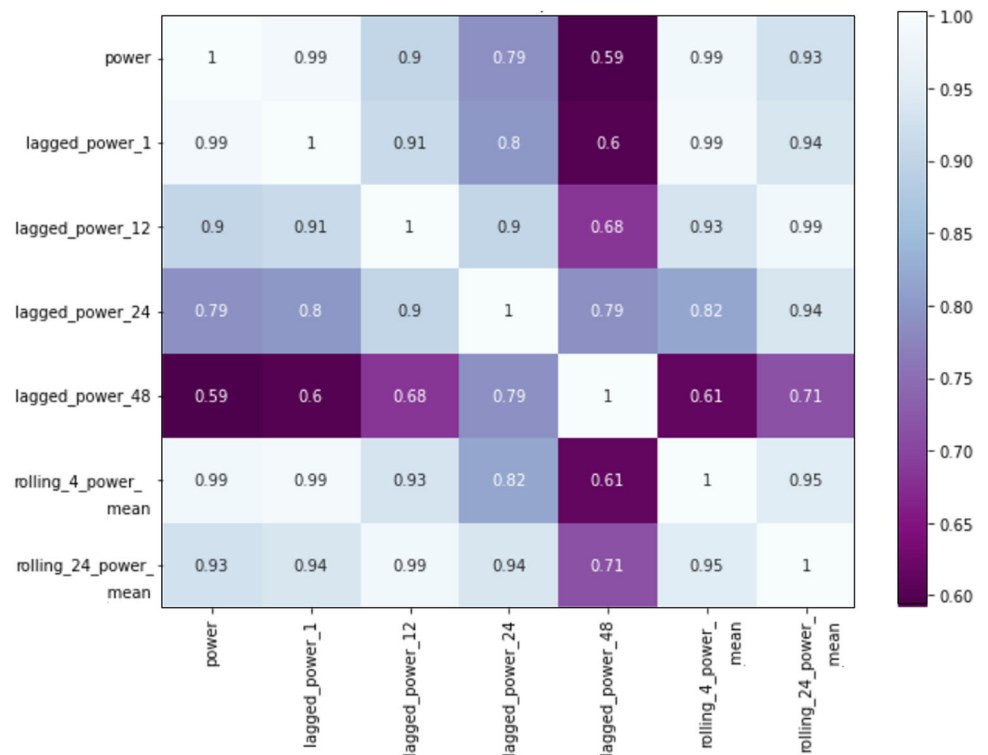**Fig. 10** Heatmap with lagged features for prediction of wind power

Fig. 11 Prediction of wind power (24 h ahead) using RF and XGB ensemble

parameters and solar power) were collected from the same source. Nevertheless, the results obtained from this data collection are very different due to the discontinuous existence of solar power. In Fig. 13, Pearson correlation heatmap for solar power correlations is indicated with each parameter.

For solar power, the two factors that appear to be marginally connected to power production are temperature and humidity. The heatmap only excludes a few parameters, while features like 'day_of_week,' 'month,' 'day_of_year,' 'week,' etc. is computed additionally. From the plotted heatmap, the attributes showed significantly weaker influence on solar power; therefore, "lagged" copies of the time series are established like wind power prediction as shown in Fig. 14.

Here, the concept reveals the strong correlation at Lag1. Hence, lagging power of 1, 12, 24, 48 and 72 h are plotted along with rolling_4_power_mean and rolling_24_power_mean as shown in Fig. 15.

As the plotted heatmap reveals, there exist correlations of lagged_power_1, lagged_power_24, lagged_power_48, lagged_power_72 and rolling_4_power_mean with solar power. Hence, these features are also added for training the model. The model prediction ahead of 24 h is shown in Fig. 16.

To accurately estimate the solar power, a simple procedure is pursued like wind power prediction. We constructed the target matrix with the features matrix with parameters that have strong correlations. Then, we implemented a five-fold CV procedure. In Fig. 16, the actual and forecasted solar power values are compared. An assessment of forecasted and true values of solar power generation using RF and XGB ensemble is shown in Fig. 17.
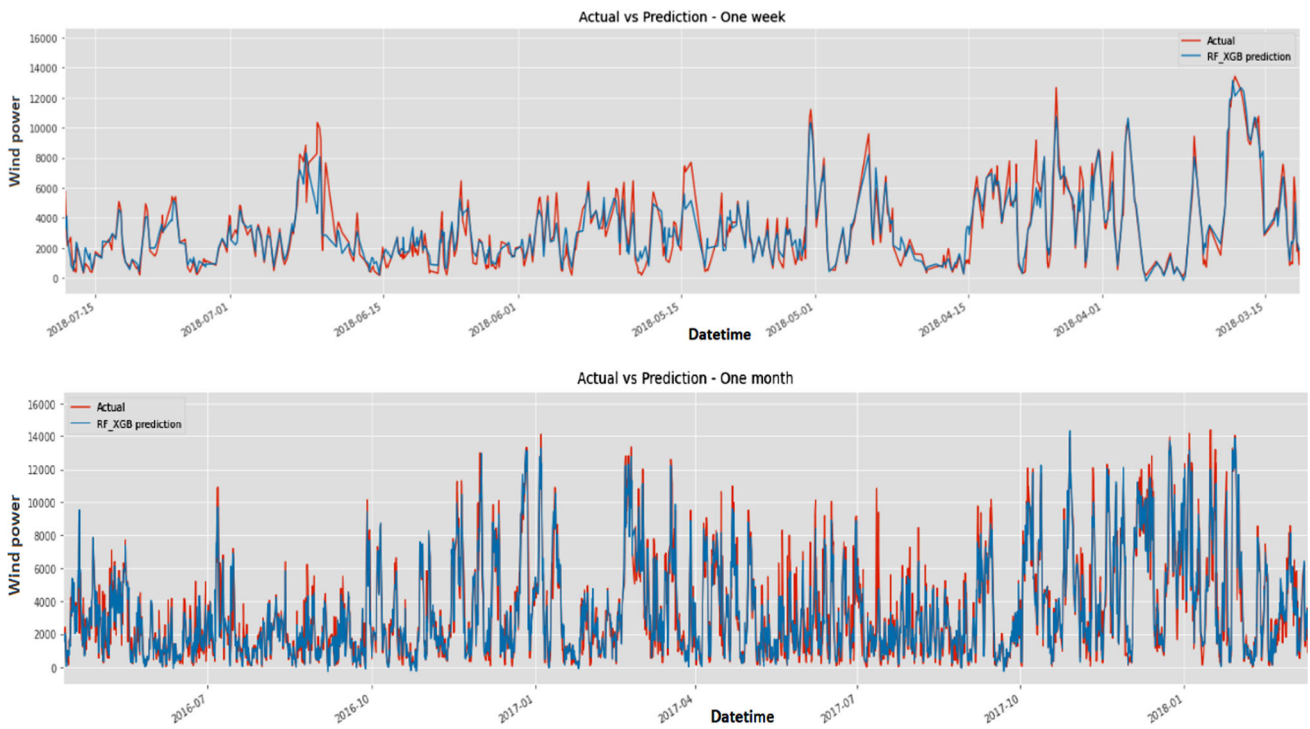
Fig. 12 Assessment of the forecasted and true values of wind power generation using RF and XGB ensemble
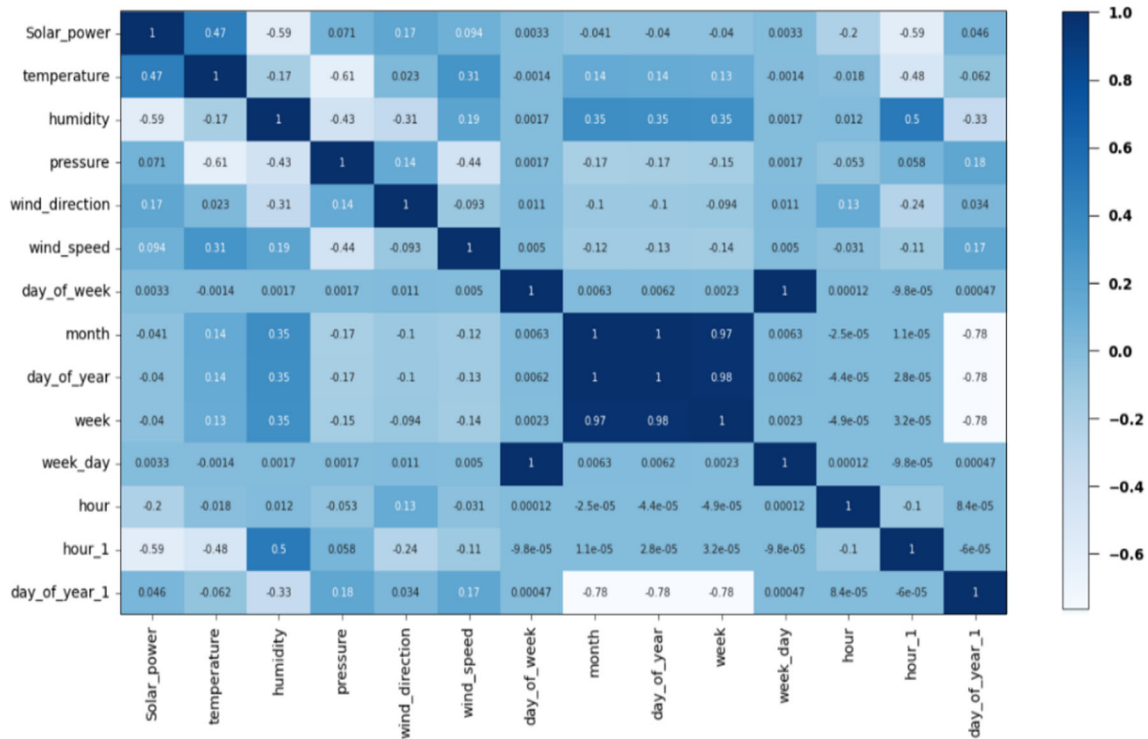

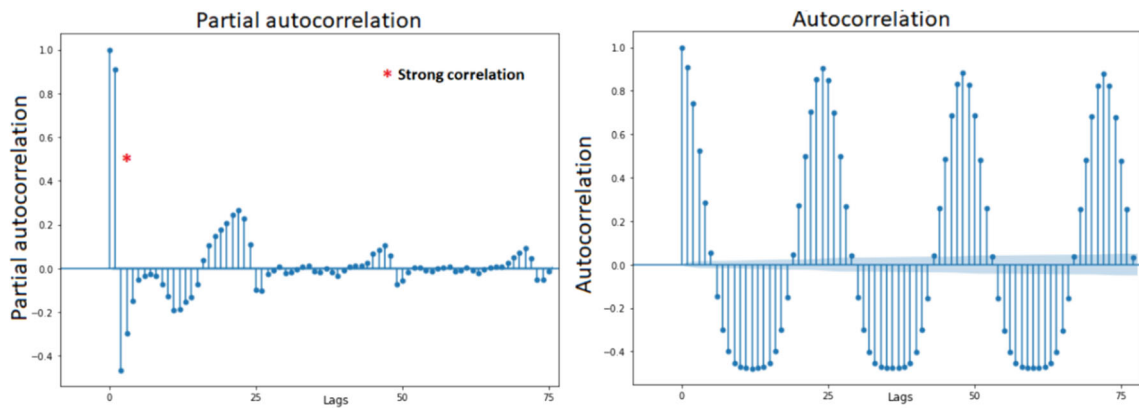**Fig. 13** Correlation coefficient of dataset parameters for prediction of solar power

Fig. 14 Partial autocorrelation and autocorrelation for prediction of solar power
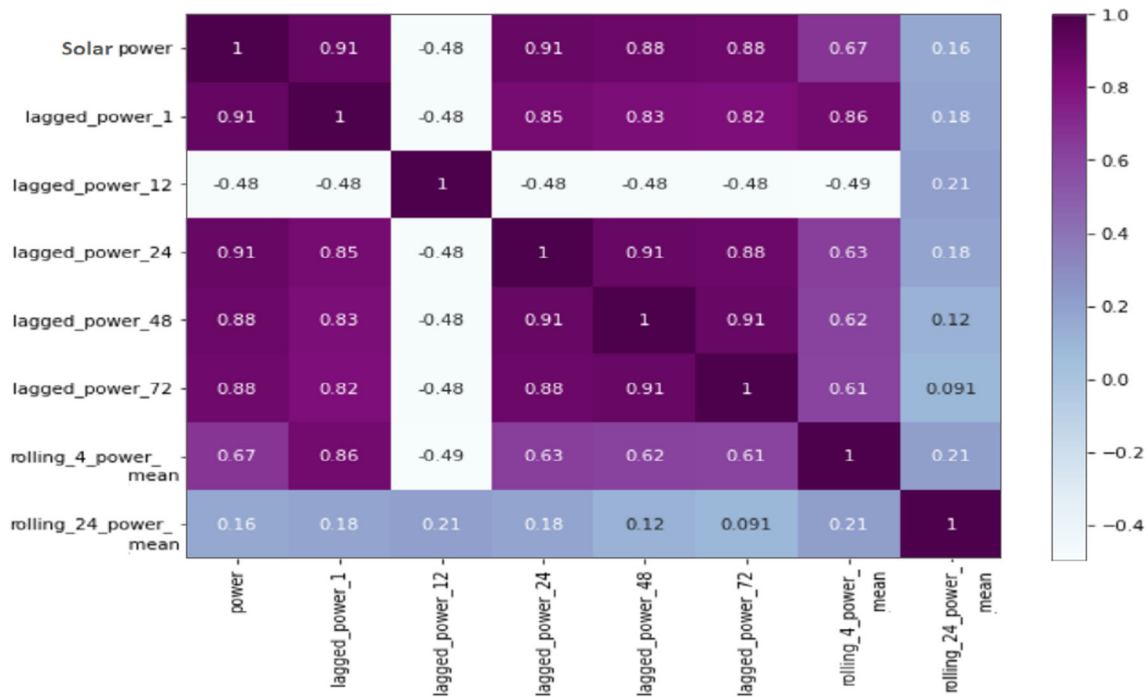

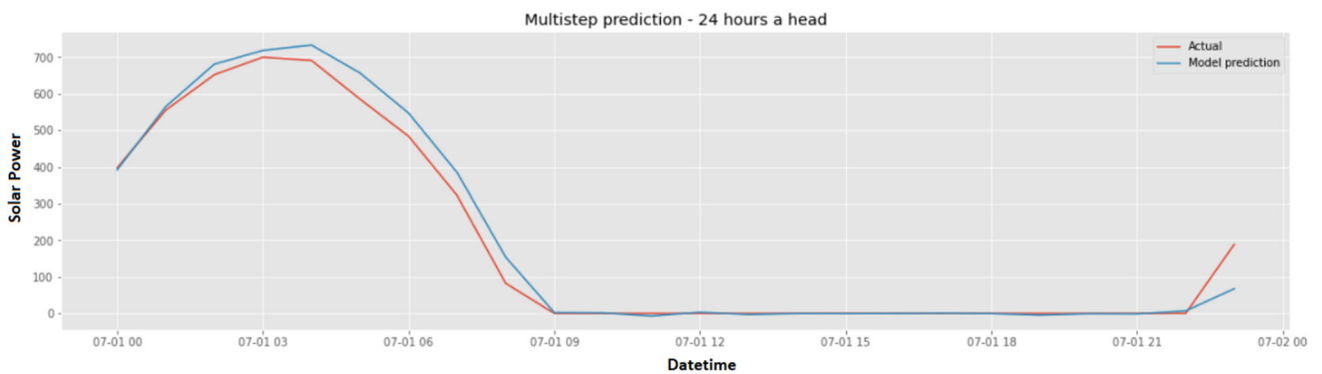**Fig. 15** Heatmap with lagged features for prediction of solar power

Fig. 16 Prediction of solar power (24 h ahead) using RF and XGB ensemble
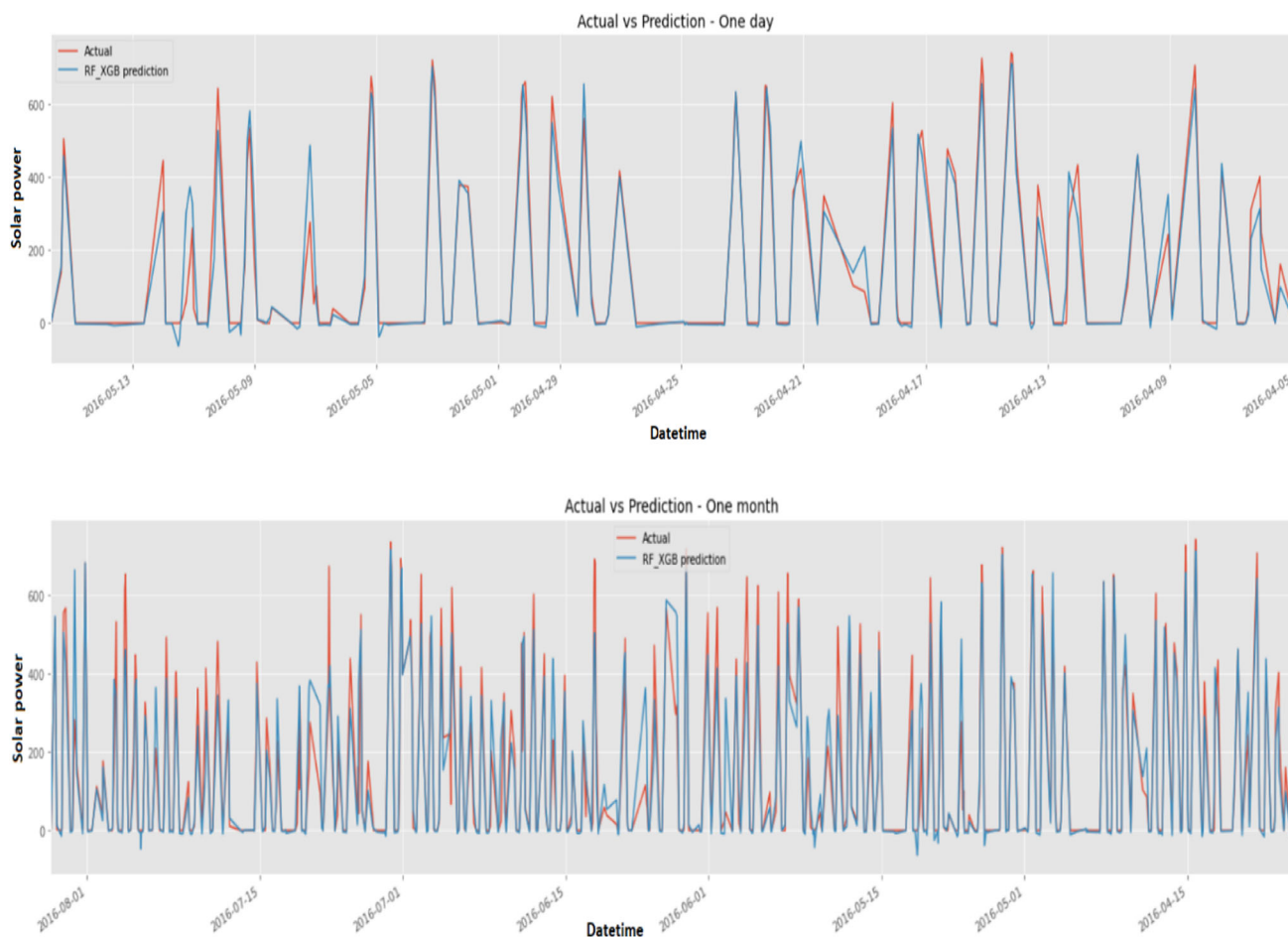
Fig. 17 Assessment of forecasted and true values of solar power generation using RF and XGB ensemble

4.3 Evaluation metrics

Several evaluation metrics like mean squared error (MSE), mean absolute error (MAE), coefficient of determination (R^2), and root mean square error (RMSE) are employed to assess the effectiveness of the suggested model [74]. MSE is used to assess how well an estimate or prediction performs. The average squared difference between actual and projected values is what is measured. By multiplying the total squared difference between the actual and anticipated values by the number of samples, the MSE is determined. The MSE equation is:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{n} (y - y')^2 \tag{14}$$

where n represents the observations, y and y' are the actual and predicted values in the i-th sample. MSE is frequently used in regression problems to assess the precision of the model. The model performs better at predicting the dependent variable when the MSE is smaller. The value of MSE can be significantly impacted by huge mistakes since it is sensitive to outliers.

MAE is a statistical metric used to assess the precision of an estimate or a forecast. The average absolute difference of true and anticipated values is what is measured. By adding together, the total absolute disparities between the values that were anticipated and those that were actually obtained, MAE is determined. The MAE equation is:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{n} |y - y'| \tag{15}$$

where n stands for number of observations; y and y' are the true and predicted values in the i-th sample. Regression analysis frequently makes use of MAE, which offers a simple-to-understand measurement of the average absolute variance of the expected and the true values. Because MAE does not square the errors, it is less susceptible to outliers than MSE. Yet, because it handles all faults equally, it is less sensitive to significant errors.

R^2 score is a statistical measure that indicates the amount of the variability of the dependent variable that is represented

by the model's independent factors. It has a value between 0 and 1, with larger values suggesting that the regression model fits the data better. It comes from

$$R^2 = 1 - \frac{\sum (y - y')^2}{\sum (y - y_{\text{mean}})^2} \qquad (16)$$

RMSE is the metric that measures the average difference between the forecasted and true values of the dependent variable. It is defined by square root of MSE,

$$\text{RMSE} = \sqrt{\sum \frac{(y' - y)^2}{N}}. \qquad (17)$$

A comparative analysis of various models for electricity consumption, wind and solar power prediction in terms of MSE, MAE, R^2, and RMSE is shown in Tables 1, 2 and 3, respectively. A combined analysis is also presented in Table 4 with a graphical representation of R^2-score in Fig. 17 to confirm the efficacy of the proposed methodology.

We aim to forecast the consumption of electricity, wind power, and solar power in a particular area based on historical data. We have selected seven different models to compare: linear regression, decision tree, lasso model, LightGBM, catboost, RF and XGB. We have also proposed and developed a stacked ensemble of RF and XGB for the purpose. The evaluation metric used for comparability of the models is R^2, MSE, MAE, and RMSE, which computes the average variance of the predicted consumption and the actual consumption. We obtained an accuracy of 0.98 for prediction of consumption and solar power, while 0.99 for wind power. The reported accuracy is highest among all the models including lone RF and XGB. The error metric of MSE, MAE, and RMSE is also lowest among all the models. In prediction of electrical energy consumption, the proposed model performed better that lone RF itself by 1%. In wind power prediction, Catboost performed better than all other models which is superseded by our proposed model by accuracy of 5%. In case of solar power prediction, LightGBM performed better than all other models which is superseded by our proposed model by accuracy of 1%.

Based on the findings, it is revealed that the proposed ensemble achieved the best results in wind power, solar power, and electricity consumption predictions, with RMSEs of 188.69, 25.92, and 4.90 and R^2 score of 0.99, 0.98, and 0.98. The suggested RF-XGB ensemble had a higher grade in the comparison study than other models for predicting in low volume datasets. Finally, a comparative study of different models in terms of R^2 for forecasting wind power, solar power, and electrical load forecasting is presented in Fig. 18.

A comparative analysis of the proposed work with existing studies in the literature based on RF and XGB is shown in Table 5.

The model evaluation metrics for RF and XGB algorithms were recoded. With an R^2 value of 0.89, Random Forest was able to attain MAE values of 90% and 95%. XGBoost showed an R^2 value of 0.91 and an RMSE of 5.9%. A comparison between the Random Forest and XGBoost ensembles also revealed better results, with an RMSE of 0.955 and an R^2 of 0.91. The Random Forest and XGBoost combined stacked ensemble fared better, with an R^2 of 0.98 and an RMSE of 4.909. These findings demonstrate the effectiveness of the proposed ensemble methods in raising prediction accuracy.

5 Conclusion & future scope

With traditional energy supplies such as coal, oil, and natural gas dwindling as well as the environmental impact of burning fossil fuels, governments and businesses gradually depend on the production of renewable sources of energy. Wind and solar are exciting renewable energy sources, which have been growing rapidly. An intrinsic characteristic of these resources is that the ability to generate energy cannot be managed entirely or even predicted. To ensure that renewable energy sources are integrated effectively into the grid, it is essential to have appropriate forecasting and management methods in place. A grid integrating renewable energy sources must be tracked continuously and must be able to predict unexpected changes in energy supply and demand. In this paper, predictive analytical methods have been successfully implemented that predicts wind power, solar power, and electricity consumption for the state using a stacked ensemble of RF and XGB. By combining the predictions of these two models using a LR model, the stacked model is capable to leverage the strengths of both models and improve the overall accuracy. The RF model is capable of capturing a portion of the simpler trends, while the XGB model may be inclined of capturing the more complex trends. The LR model then learns how to optimally combine these predictions to improve the overall accuracy. The variance of these predictions is quantified by R^2 and RMSE. The expertise gained from the data-driven models is thought to contribute to a better investment and use of power storage systems, leading to an economically viable solution enabling even higher rates of renewable energy penetration in the grid.

Eventually, it is important to point out that when considering upcoming energy market several additional opportunities and problems must be dealt with. All these possibilities can be incorporated into a predictive smart grid control system, and these research findings can be applied immediately, promoting, and contributing to the shifting of power industry to new and sustainable modes. More study is needed to assess and compare the limits of the existing ensemble for forecasting with more complicated models and optimization-based training approaches. Further study is needed to emphasize

Table 1 An analysis of various models aimed at forecasting the consumption of electrical energy

Model	MSE	MAE	RMSE	R^2-score
Linear regression	811.29	22.456	28.483	0.48
Decision tree	355.64	13.232	18.858	0.77
Lasso model	811.45	22.463	28.486	0.48
LightGBM	273.52	12.085	16.538	0.82
Catboost	105.27	6.986	10.260	0.93
Random Forest	34.479	3.4861	5.8720	0.97
XGBoost	205.84	10.226	14.347	0.87
RF-XGB ensemble (proposed)	24.105	3.1235	4.9098	0.98

Table 2 An analysis of various models aimed at forecasting the wind power

Model	MSE	MAE	RMSE	R^2-score
Linear regression	983,530.97	573.14	991.73	0.89
Decision tree	768,724.51	513.65	876.76	0.91
Lasso model	983,548.09	573.19	991.73	0.89
LightGBM	918,529.56	579.55	958.39	0.89
Catboost	458,124.32	407.16	676.84	0.94
Random Forest	714,336.42	493.04	845.18	0.92
XGBoost	630,774.50	466.61	794.21	0.92
RF-XGB ensemble (proposed)	35,606.98	109.58	188.69	0.99

Table 3 An analysis of various models aimed at forecasting the solar power

Model	MSE	MAE	RMSE	R^2-score
Linear regression	29,830.92	139.49	172.71	0.47
Decision tree	8924.67	48.289	94.470	0.84
Lasso model	29,950.54	140.06	173.06	0.46
LightGBM	1356.75	18.616	36.834	0.97
Catboost	4741.02	37.404	68.855	0.91
Random Forest	8103.24	46.876	90.018	0.85
XGBoost	6041.42	42.361	77.726	0.89
RF-XGB ensemble (proposed)	671.88	13.991	25.920	0.98

Table 4 A combined analysis of wind power, solar power and load consumption prediction

| Model	Wind power		Solar power		Electrical load	
	RMSE	R^2-score	RMSE	R^2-score	RMSE	R^2-score
Linear regression	991.73	0.89	172.710	0.47	28.483	0.48
Decision tree	876.76	0.91	94.470	0.84	18.858	0.77
Lasso model	991.73	0.89	173.060	0.46	28.486	0.48
LightGBM	958.39	0.89	36.834	0.97	16.538	0.82
Catboost	676.84	0.94	68.855	0.91	10.260	0.93
Random Forest	845.18	0.92	90.018	0.85	5.872	0.97
XGBoost	794.21	0.92	77.726	0.89	14.347	0.87
RF-XGB ensemble (proposed)	188.69	0.99	25.920	0.98	4.909	0.98
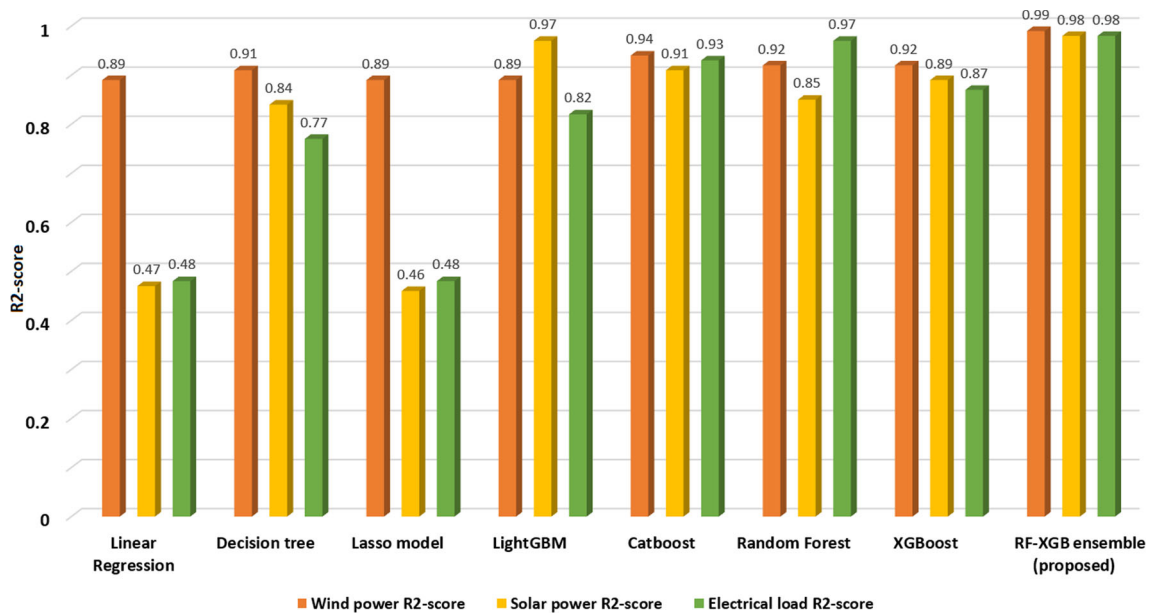
Fig. 18 Comparison of different prediction models

Table 5 A comparative analysis of proposed work with existing studies in the literature

Study	Prediction parameter	Model	Metrics
[65]	Energy performance	Random forest	MAE—90% & 95%
[66]	Building energy	Random forest	R^2—0.89
[67]	Energy prediction	Random forest	RMSE—5%
[67]	Energy prediction	XGBoost	RMSE—5.9%
[68]	Electricity demand	Random forest	RMSE—2075.54
[68]	Electricity demand	XGBoost	RMSE—2038.54
[69]	Electrical energy	Random forest & XGBoost ensemble	R^2—0.91, RMSE—0.955
Proposed	Renewable power & electricity consumption	Stacked ensemble of Random Forest & XGBoost	R^2—0.98, RMSE—4.909

mostly on assessment of enhancing the suggested technique's maximal forecasting horizons with enhanced accuracy in big multi-step forecasting. While there are still challenges to overcome in modeling accurate forecasts, improvements in ML and AI are bringing up new prospects for boosting forecast accuracy and lowering uncertainty. Further research and development in this field will be required in the future to allow significant deployment of renewable energy technology and to help the global effort to combat climate change.

Declarations

References

1. Ang T-Z, Salem M, Kamarol M, Das HS, Nazari MA, Prabaharan N (2022) A comprehensive study of renewable energy sources: Classifications, challenges, and suggestions. Energ Strat Rev 43:100939. https://doi.org/10.1016/j.esr.2022.100939
2. Dey S, Sreenivasulu A, Veerendra GTN, Rao KV, Babu PSSA (2022) Renewable energy present status and future potentials in India: an overview. Innov Green Dev 1(1):100006. https://doi.org/10.1016/j.igd.2022.100006

3. IEA (2022) Global Energy Review: CO$_2$ Emissions in 2021. IEA Paris. https://www.iea.org/reports/global-energy-review-co2-emissions-in-2021-2.

4. Suganthi L, Samuel AA (2012) Energy models for demand forecasting—a review. Renew Sustain Energy Rev 16:1223–1240

5. Khan AR, Mahmood A, Safdar KZA, Khan NA (2016) Load forecasting, dynamic pricing, and DSM in the smart grid: a review. Renew Sustain Energy Rev 54:1311–1322

6. Hiron N, Busaeri N, Sutisna S, Nurmela N, Sambas A (2021) Design of Hybrid (PV-Diesel) system for tourist island in Karimunjawa Indonesia. Energies 14(24):8311. https://doi.org/10.3390/en14248311

7. Mellit A, Kalogirou S (2008) Artificial intelligence techniques for photovoltaic applications: a review. Prog Energy Combust Sci 34:574–632

8. Inman RH, Pedro HTC, Coimbra CFM (2013) Solar forecasting methods for renewable energy integration. Prog Energy Combust Sci 39:535–576

9. Jung J, Broadwater RP (2014) Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 31:762–777

10. Baños R, Manzano-Agugliaro F, Montoya FG, Gil C, Alcayde A, Gómez J (2011) Optimization methods applied to renewable and sustainable energy: a review. Renew Sustain Energy Rev 15:1753–1766

11. Sfetsos A (2000) A comparison of various forecasting techniques applied to mean hourly wind speed time series. Renew Energy 21:23–35

12. Martín L, Zarzalejo LF, Polo J, Navarro A, Marchante R, Cony M (2010) Prediction of global solar irradiance based on time-series analysis: Application to solar thermal power plants' energy production planning. Sol Energy 84:1772–1781

13. Martín L, Zarzalejo LF, Polo J, Navarr A, Marchante R, Cony M (2010) Prediction of global solar irradiance based on time-series analysis: application to solar thermal power plants' energy production planning. Sol Energy 84:1772–1781

14. Fernandez-Jimenez LA, Muñoz-Jimenez A, Falces A, Mendoza-Villena M, Garcia-Garrido E, Lara-Santillan PM et al (2012) Short-term power forecasting system for photovoltaic plants. Renewable Energy 44:311–317

15. Ceci M, Corizzo R, Fumarola F, Malerba D, Rashkovska A (2017) Predictive modeling of PV energy production: how to set up the learning task for a better prediction? IEEE Trans Ind Inf 13:956–966

16. Abdel-Nasser M, Mahmoud K (2019) Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Comput Appl 31:2727–2740

17. Ozgoren M, Bilgili M, Sahin B (2012) Estimation of global solar radiation using ANN over Turkey. Expert Syst Appl 39:5043–5051

18. Koca A, Oztop HF, Varol Y, Koca GO (2011) Estimation of solar radiation using artificial neural networks with different input parameters for the Mediterranean region of Anatolia in Turkey. Expert Syst Appl 38:8756–8762

19. Bessa RJ, Trindade A, Miranda V (2015) Spatial-temporal solar power forecasting for smart grids. IEEE Trans Industr Inf 11:232–241

20. Chen Z, Troccoli A (2017) Urban solar irradiance and power prediction from nearby stations. Meteorol Z 26:277–290

21. Wan C, Zhao J, Song Y (2015) Photovoltaic and solar power forecasting for smart grid energy management. J Power Energy Syst 1:38–46

22. Das UK, Tey KS, Seyedmahmoudian M, Mekhilef S, Idris MYI, Van Deventer W, Horan B, Stojcevski A (2018) Forecasting of photovoltaic power generation and model optimization: a review. Renew Sustain Energy Rev 81:912–928

23. Omar M, Dolara A, Magistrati G, Mussetta M, Ogliari E, Viola F (2016) Day-Ahead Forecasting for Photovoltaic Power Using Artificial Neural Networks Ensembles. In: Proceedings of the 2016 IEEE International Conference on Renewable Energy Research and Applications, ICRERA 2016, Birmingham, UK, 20–23 November 2016, pp 1152–1157.

24. Ding M, Wang L, Bi R (2011) An ANN-based approach for forecasting the power output of photovoltaic system. Proced Environ Sci 11:1308–1315

25. Kanagasundaram A, Valluvan R, Atputharajah A (2018) A study on solar PV power generation influencing parameters using captured data from faculty of engineering, university of jaffna solar measuring station, International Conference On. Solar Energy Materials, Solar Cells & Solar Energy Applications, Jan 2018.

26. Lauret P, Voyant C, Soubdhan T, David M, Poggi P (2015) A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol Energy 112:446–457

27. Neo YQ, Teo TT, Woo WL, Logenthiran T, Sharma A (2017) Forecasting of photovoltaic power using deep belief network. In: Proceedings of the TENCON 2017—2017 IEEE Region 10 Conference, Penang, Malaysia, 5–8 November, 1189–1194.

28. Yacef R, Benghanem M, Mellit A (2012) Prediction of daily global solar irradiation data using Bayesian neural network: a comparative study. Renew Energy 48:146–154

29. Qing X, Niu Y (2018) Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 148:461–468

30. Liu B, Fu C, Bielefield A, Liu QY (2017) Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network. Energies 10:1453

31. Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, Fouilloy A (2017) Machine learning methods for solar radiation forecasting: A review. Renew Energy 105:569–582

32. Zamo M, Mestre O, Arbogast P, Pannekoucke O (2014) A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: deterministic forecast of hourly production. Sol Energy 105:792–803

33. Behera MK, Majumder I, Nayak N (2018) Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. Eng Sci Technol Int J 21:428–438

34. Huang GB, Zhu Q, Siew C (2006) Extreme learning machine: theory and applications. Neurocomputing 70:489–501

35. Baek MK, Lee D (2018) Spatial and temporal day-ahead total daily solar irradiation forecasting: ensemble forecasting based on empirical biasing. Energies 11:70

36. Urraca R, Antonanzas J, Martinez MA, Martinez-de-Pison FJ, Torres FA (2016) Smart baseline models for solar irradiation forecasting. Energy Convers Manage 108:539–548

37. Gala Y, Fernandez A, Diaz J, Dorronsoro JR (2016) Hybrid machine learning forecasting of solar radiation values. Neurocomputing 176:48–59

38. Sambas A, Mohammadzadeh A, Vaidyanathan S (2023) Ayob AFM (2023) Investigation of chaotic behavior and adaptive type-2 fuzzy controller approach for permanent magnet synchronous generator (PMSG) wind turbine system. AIMS Math 8(3):5670–5686. https://doi.org/10.3934/math.2023285

39. You Q, Fraedrich K, Min JZ, Kang SC, Zhu XH, Pepin N, Zhang L (2014) Observed surface wind speed in the Tibetan Plateau since 1980 and its physical causes. Int J Climatol 34:1873–1882. https://doi.org/10.1002/joc.3807

40. Zhao ED, Zhao J, Liu LW, Su ZY, An N (2016) Hybrid wind speed prediction based on a self-adaptive ARIMAX Model with an exogenous WRF simulation. Energies 9(1):7. https://doi.org/10.3390/en9010007

41. Leva S, Dolara A, Grimaccia F, Mussetta M, Ogliari E (2017) Analysis and validation of 24 hours ahead neural network forecasting

of photovoltaic output power. Math Comput Simul 131:88–100. https://doi.org/10.1016/j.matcom.2015.05.010

42. Yang L, He M, Zhang JS, Vittal V (2015) Support-vector-machine-enhanced Markov model for short-term wind power forecast. IEEE Trans Sustain Energy 6:791–799. https://doi.org/10.1109/TSTE.2015.2406814

43. Wan C, Xu Z, Pinson P, Dong ZY, Wong KP (2014) Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans Power Syst 29:1033–1044. https://doi.org/10.1109/TPWRS.2013.2287871

44. Liu H, Mi XW, Li YF (2018) Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manage 156:498–514

45. Hyde O, Hodnett PF (1997) An adaptable automated procedure for short-term electricity load forecasting. IEEE Trans Power Syst 12:84–94

46. Broadwater RR, Sargent A, Yarali A et al (1997) Estimating substation peaks from load research data. IEEE Trans Power Delivery 12:451–456

47. Ross G, Galiana D (1987) Short-Term load forecasting. Proc IEEE 75:1558–1573

48. El-Keib AA, Ma X, Ma H (1995) Advancement of statistical-based modeling techniques for short-term load forecasting. Electr Power Syst Res 35:51–58

49. Huang SR (1997) Short-term load forecasting using threshold autoregressive models. IEEE proc gener Transm Distrib 144:477–481

50. Goia A, May C, Fusai G (2010) Functional clustering and linear regression for peak load forecasting. Int J Forecast 26(4):700–711

51. Amral N, Ozveren C, King D (2007) Short term load forecasting using multiple linear regression. In: UPEC 2007. 42nd International universities power engineering conference, 2007: 1192–1198.

52. Pappas S, Ekonomou L, Karamousantas D, Chatzarakis G, Katsikas S, Liatsis P (2008) Electricity demand loads modeling using autoregressive moving average (ARMA) models. Energy 33(9):1353–1360

53. Lee CM, Ko CN (2011) Short-term load forecasting using lifting scheme and ARIMA models. Expert Syst Appl 38(5):5902–5911

54. Mastorocostas P, Theocharis J, Bakirtzis A (1999) Fuzzy modeling for short term load forecasting using the orthogonal least squares method. IEEE Trans Power Syst 14(1):29–36

55. Mandal P, Senjyu T, Funabashi T (2006) Neural networks approach to forecast several hours ahead electricity prices and loads in a deregulated market. Energy Convers Manage 47(15–16):2128–2142

56. Senjyu T, Takara H, Uezato K, Funabashi T (2002) One-hour-ahead load forecasting using neural network. IEEE Trans Power Syst 17(1):113–118

57. Lin CT, Chou LD (2013) A novel economy reflecting short-term load forecasting approach. Energy Convers Manage 65:331–342

58. Wu CH, Tzeng GH, Lin RH (2009) A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. Expert Syst Appl 36(3):4725–4735

59. Wang J, Li L, Niu D, Tan Z (2012) An annual load forecasting model based on support vector regression with a differential evolution algorithm. Appl Energy 94:65–70

60. Nagi J, Yap KS, Nagi F, Tiong SK, Ahmed SK (2011) A computational intelligence scheme for the prediction of the daily peak load. Appl Soft Comput 11(8):4773–4788

61. Cheng YY, Chan P, Qiu ZW (2012) Random forest-based ensemble system for short-term load forecasting. In: 2012 International conference on machine learning and cybernetics (ICMLC) 1: 52–56.

62. Krawczak M, Popchev I, Rutkowski L et al (2015) Intelligent systems'2014. Adv Intell Syst Comput 323:821–828

63. Khayatian F, Sarto L, Dall'O' G (2016) Application of neural networks for evaluating energy performance certificates of residential buildings. Energy Build 125:45–54

64. Ascione F, Bianco N, De Stasio C, Mauro GM, Vanoli GP (2017) Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: a novel approach. Energy 118:999–1017

65. Papadopoulos S, Azar E, Woon WL, Kontokosta CE (2017) Evaluation of tree-based ensemble learning algorithms for building energy performance estimation. J Build Perform Simul 1493:1–11

66. Wang Z, Wang Y, Zeng R, Srinivasa RS, Ahrentzen S (2018) Random Forest-based hourly building energy prediction. Energy Build 171:11–25

67. Rathore H, Meena HK, Jain P (2023) "Prediction of EV energy consumption using random forest and XGBoost," 2023 International Conference on Power Electronics and Energy (ICPEE), Bhubaneswar, India: 1–6. doi: https://doi.org/10.1109/ICPEE54198.2023.10060798.

68. Gökçe MM, Duman E (2022) Performance comparison of simple regression, random forest and XGBoost algorithms for forecasting electricity demand, 2022 3rd International Informatics and Software Engineering Conference (IISEC), Ankara, Turkey: 1–6, doi: https://doi.org/10.1109/IISEC56263.2022.9998213.

69. Banik R, Das P, Ray S, Biswas A (2021) Prediction of electrical energy consumption based on machine learning technique. Electr Eng 103:1–12. https://doi.org/10.1007/s00202-020-01126-z

70. Staffell I, Pfenninger S (2016) Using bias-corrected reanalysis to simulate current and future wind power output. Energy 114:1224–1239

71. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Ann Stat 29(5):1189–1232

72. Emami S, Emami H, Parsa J (2023) LXGB: a machine learning algorithm for estimating the discharge coefficient of pseudo-cosine labyrinth weir. Sci Rep 13:12304

73. Nadkarni SB, Vijay GS, Kamath RC (2023) Comparative Study of random forest and gradient boosting algorithms to predict airfoil self-noise. Eng Proc 59(1):24

74. Zafar MH, Khan NM, Mansoor M, Mirza AF, Moosavi SKR, Sanfilippo F (2022) Adaptive ML-based technique for renewable energy system power forecasting in hybrid PV-Wind farms power conversion systems. Energy Convers Manage 258:115564. https://doi.org/10.1016/j.enconman.2022.115564