Journal of
Green Energy
Research and Innovation

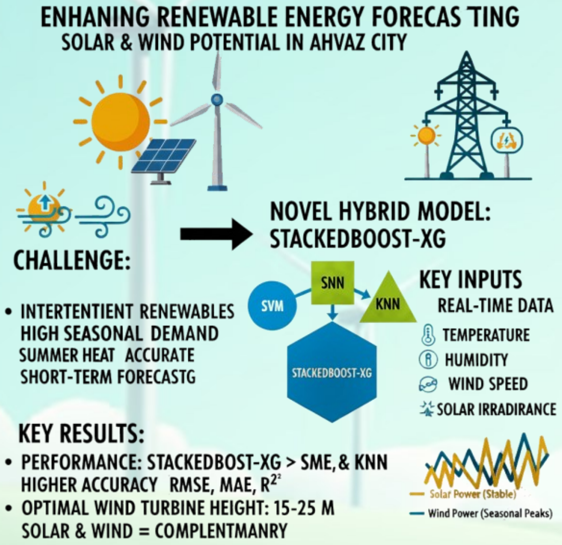
Enhancing Renewable Energy Forecasting: A Hybrid Machine Learning Approach for Solar and Wind Energy Potential in Ahvaz City

Mehdi Mohammadian Mehr, Hossein Farzin

Highlights

- StackedBoost-XG model combines SVM, KNN & XGBoost for better energy forecasting.
- Optimizes prediction accuracy, outperforming individual models in dynamic conditions.
- Application in Ahvaz City, tackling extreme climate & high summer energy demand.
- Compares SVM, KNN & StackedBoost-XG, proving superior performance in forecasting.

Graphical Abstract



Citation
M. Mohammadian Mehr, and H. Farzin, "Enhancing Renewable Energy Forecasting: A Hybrid Machine Learning Approach for Evaluating Solar and Wind Energy Potential in Ahvaz City," *Journal of Green Energy Research and Innovation*, vol. 2, no. 4, pp. 14-26, 2025.

https://doi.org/10.61882/jgeri.2.4.14

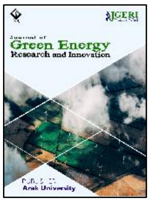
JGERI
Journal of Green Energy
Research and Innovation

Journal of Green Energy Research and Innovation 2.4 (2025) 14-26

Online ISSN: 3041-9018

Journal of Green Energy Research and Innovation

Journal Homepage: www.jgeri.araku.ac.ir



Enhancing Renewable Energy Forecasting: A Hybrid Machine Learning Approach for Solar and Wind Energy Potential in Ahvaz City

Mehdi Mohammadian Mehr, Hossein Farzin[*]

Department of Electrical Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Ahvaz 61357-43337, Iran.

ARTICLE INFO

Keywords:
Short-term forecasting,
XGBoost,
SVM,
KNN,
StackedBoost-XG,
Energy Grid Optimization,
Renewable Energy Integration.

Article History:
Received: 02 May 2025;
Revised: 15 June 2025;
Accepted: 01 July 2025.

Article type:
Research Article

* ***Corresponding author***
E-mail address
farzin@scu.ac.ir (H. Farzin)

ABSTRACT

This paper introduces a new approach for short-term forecasting of solar and wind energy potential in Ahvaz City. The method is based on the StackedBoost-XG model, a hybrid ensemble that combines Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) with XGBoost as the final estimator. The study focuses on accurately predicting energy generation using real-time meteorological data. Key inputs include temperature, humidity, wind speed, and solar irradiance factors that are crucial for reliable energy forecasting. These variables are integrated into energy production formulas to estimate outputs for both solar and wind sources. This improves prediction accuracy. The model's performance is assessed using standard evaluation metrics: RMSE, MAE, and R^2. Results indicate that StackedBoost-XG significantly outperforms the individual SVM and KNN models. It shows higher accuracy in forecasting both solar and wind energy. The research also explores the effect of wind turbine height. It finds that optimal energy output occurs at heights between 15 and 25 meters. In addition, the study highlights the importance of managing thermal losses in solar panels, especially during warmer months, to maintain system efficiency. Finally, it emphasizes the complementary nature of solar and wind energy. Solar power offers relatively stable output throughout the year, while wind energy provides higher peaks in specific seasons. By integrating both energy sources, the study proposes a promising solution to address energy demand imbalances in Ahvaz. This study introduces a hybrid forecasting method that uses advanced machine learning and weather data. Its goal is to optimize renewable energy systems and enhance the management of the energy grid.

1. Introduction

1.1. *Motivation*

In the face of rapidly growing global energy demand, the transition from traditional fossil fuel-based power generation to renewable energy sources has become a critical challenge [1,2]. Among the most promising sources of renewable energy, solar and wind power stand out due to their abundance and environmental sustainability [3-5]. However, despite their potential, the intermittent and unpredictable nature of these energy sources presents significant obstacles for integrating them into power grids [6,7]. The city of Ahvaz is located in the southwestern Iran, where summers are characterized by intense heat waves, leading to sharp increases in electricity consumption, primarily for cooling and air conditioning. This seasonal demand fluctuation further exacerbates the energy supply-demand imbalance, creating a pressing need for more efficient energy management strategies. Notably, the reliance on solar and wind energy in Ahvaz offers a sustainable alternative, but the challenge remains in reliably predicting their performance to match energy supply with peak demand [8,9]. The fluctuating nature of solar and wind energy generation, influenced by factors such as weather patterns, seasonal variations, and geographic conditions, introduces substantial uncertainty in energy forecasting [10,11]. In Ahvaz, the unpredictability of these renewable sources poses a significant challenge in planning for electricity consumption during the summer months.

Accurate forecasting of solar and wind energy output is essential for effective energy planning and improved load balancing, especially during periods of peak demand. This highlights the urgent need for reliable prediction models capable of accurately estimating the performance of solar and wind energy systems. These models help improve the integration of renewable energy sources into the power grid. They reduce the gap between energy supply and demand. In this study, advanced machine learning (ML) techniques are used to improve the accuracy of energy production forecasts from solar and wind sources in Ahvaz. Accurate forecasting of energy generation improves the sustainability of energy systems. It also enables smarter energy management. This reduces dependence on non-renewable backup power, especially during peak demand periods.

1.2. Literature Review

Global electricity demand is steadily increasing, especially during hot summer months. In Ahvaz, this leads to a critical imbalance between energy supply and demand. This issue primarily arises due to the reliance on non-renewable energy sources during peak periods. The increasing demand for sustainable energy solutions necessitates a thorough examination of the renewable energy potential of solar and wind power in Ahvaz. Accurately forecasting the energy output of these sources is crucial for effective energy management. This study aims to address these challenges by developing a mathematical model that converts meteorological parameters into energy outputs for solar panels and wind turbines. A major challenge is the intermittent and fluctuating nature of solar and wind energy. This variability makes accurate energy prediction difficult. Therefore, this study focuses on short-term forecasting using machine learning models. The goal is to improve grid management and optimize the integration of renewable energy in Ahvaz.

Hybrid solar-wind energy systems are gaining attention as an effective solution to reduce grid pressure and support sustainable energy production. For example, one study used transformer models to forecast hybrid photovoltaic-wind systems in urban areas. The results showed a high accuracy of 90.7% for solar energy and 90.4% for wind speed. This highlights the significant role of AI-based models in improving forecasting performance for hybrid energy systems [12]. Several studies have emphasized the critical importance of forecasting for grid stability. One review highlighted the necessity of integrating forecasting methods for wind, solar, and electrical load management. It pointed out that the growing uncertainty in energy production calls for more sophisticated and integrated forecasting models to manage renewable energy sources effectively and reduce fluctuations in energy production [13]. The application of ML techniques has become a central focus in enhancing energy forecasting. A study demonstrated that combining Random Forest, Exponential Smoothing, and Long Short-Term Memory (LSTM) networks for solar power forecasting significantly improved accuracy compared to traditional methods. This underscores the effectiveness of hybrid models in optimizing the prediction of renewable energy generation [14]. Further studies also highlight the value of Earth System Models (ESMs) in long-term energy forecasting. While ESMs are valuable tools, their optimization and integration with advanced forecasting techniques are essential for accurate energy production predictions [15]. Artificial Neural Networks (ANNs) have been applied to predict the energy output of wind and solar systems based on local meteorological data. These models have been shown to be critical in determining the necessary reserve capacities in energy systems, particularly during high-demand periods, emphasizing the importance of accurate forecasting for grid planning and management [16]. The unpredictability of renewable energy sources has driven the development of advanced forecasting methods. For instance, the use of ARIMA and F-Prophet models demonstrated prediction accuracies exceeding 90%, making these models highly reliable for short-term energy forecasting and contributing significantly to renewable energy applications [17]. A comprehensive review of forecasting models has identified the need for precision in energy generation predictions. It emphasized that while physical, statistical, and hybrid models exist, ML-based approaches offer substantial improvements in accuracy. The review also stressed the importance of further research to refine forecasting models and enhance the integration of renewable energy into evolving energy markets [18].

Ensemble methods, particularly those that combine ANNs with other learning techniques, have outperformed individual models in predicting wind and solar power production. These hybrid approaches underline the importance of integrating multiple techniques to improve forecasting accuracy and ensure more reliable energy predictions [19]. The role of artificial intelligence (AI) in renewable energy forecasting is expanding. A study highlighted the potential of explainable AI (XAI) and quantum AI (QAI) to enhance the transparency and reliability of energy predictions. These technologies offer new opportunities for optimizing variable renewable energy systems in the future [20]. Moreover, a comparative study on photovoltaic power generation forecasting using various ML models, such as Support Vector Machines (SVM), Gaussian Process Regression (GPR), and Decision Trees, showed that while computational times varied, these models were effective in predicting solar power, providing valuable insights for the selection of forecasting models [21].

In conclusion, this study addresses the challenges of renewable energy forecasting in regions with high seasonal demand, like Ahvaz. The intermittent nature of solar and wind energy presents a significant challenge to grid stability, which can be mitigated by using machine learning models for short-term forecasting. By improving forecast accuracy, these models can optimize energy generation, enhance grid integration, and aid in the transition to a more sustainable energy future. The study focuses on improving solar and wind energy predictability through advanced ML techniques tailored to Ahvaz's specific needs, ensuring more efficient and sustainable energy systems

1.3. Contributions

This study makes several key contributions to renewable energy forecasting, particularly for predicting solar and wind energy in regions with significant climatic variations, such as Ahvaz:

1. The primary contribution is the introduction of the StackedBoost-XG model. This hybrid ensemble method combines SVM,

K-Nearest Neighbors (KNN), and XGBoost to optimize solar and wind energy forecasting accuracy. By leveraging the strengths of multiple machine learning techniques, this model improves prediction reliability, particularly in dynamic environments, outperforming individual models like SVM and KNN.

2. The study addresses the specific challenges of energy forecasting in Ahvaz, known for its high seasonal temperature fluctuations and significant summer energy demand. The performance of various models in this unique environment provides valuable insights into optimizing renewable energy systems in regions with complex climates and energy needs.
3. A comparative analysis of SVM, KNN, and the StackedBoost-XG model is presented. This analysis highlights the strengths and limitations of each model for forecasting solar and wind energy in Ahvaz. It demonstrates that the StackedBoost-XG model performs superiorly, offering a reliable and accurate energy prediction solution.

The paper is structured as follows: Introduction, Problem Description, Methodology, Data and Preprocessing, Results and Discussion, and Conclusion. Each section addresses key aspects of energy forecasting challenges, model development, and findings.

2. Problem Description

Accurate forecasting of solar and wind energy generation is a significant challenge, particularly in regions like Ahvaz. The city experiences extreme seasonal temperature variations and intense heat during the summer months. These harsh conditions increase energy demand, especially for cooling systems, and place considerable pressure on the energy grid. One of the main difficulties lies in the intermittent and unpredictable nature of renewable energy sources. Weather-related factors such as temperature, humidity, wind speed, and solar irradiance fluctuate continuously, making energy output hard to predict. This variability complicates energy planning and load management during peak demand periods. Due to the lack of reliable forecasting models, Ahvaz often relies on fossil fuel-based power plants. This dependency leads to higher costs and greater environmental harm, undermining efforts to reduce carbon emissions and shift toward sustainable energy systems. As global energy trends move toward renewables, the need for accurate and reliable forecasting becomes more urgent. Existing models often struggle to capture the complexity and non-linearity of energy production under dynamic weather conditions. These limitations reduce the effectiveness of integrating renewable sources into the power grid. To address these challenges, this study introduces the StackedBoost-XG model—a novel hybrid approach that combines SVM, KNN, and XGBoost algorithms. By leveraging the strengths of each method, the model significantly improves forecasting accuracy, even in complex and fluctuating environments. This approach enhances short-term energy predictions, optimizes energy planning, reduces reliance on conventional power sources, and supports the stable integration of renewables into the grid. It is especially valuable for regions like Ahvaz, where climate conditions are highly variable.

3. Methodology

This study employs a hybrid approach that combines meteorological data and ML models to predict the energy output from solar and wind sources. The methodology integrates key environmental parameters, such as temperature, humidity, and wind speed, to estimate the energy production of solar panels and wind turbines. These outputs are used to train forecasting models that can predict future energy production and optimize energy management in the region.

3.1. Solar Energy Production

To estimate the energy produced by a solar panel, Equation (1) calculates the energy output based on solar irradiance [22]:

$$E = G_t \times A \times \eta_t \tag{1}$$

E = Energy produced (Wh)
G_t = Solar irradiance on the panel at a given time (Wh/m^2)
A = Area of the solar panel (m^2)
η_t = Efficiency of the solar panel

The solar irradiance (G_t) is estimated using a time-dependent model, as Equation (2) [23]:

$$G_t = G_{max} \times \sin\left(\frac{\pi(t - t_{sunrise})}{t_{sunset} - t_{sunrise}}\right) \times (1 - (RH \times K)) \tag{2}$$

G_t = Solar irradiance at time t (Wh/m^2)
G_{max} = Daily average solar irradiance (Wh/m^2)
RH = Relative humidity
K = Humidity correction factor (typically 0.6)
$t_{sunrise}$, t_{sunset} = Times of sunrise and sunset
t = Current time of the day

The efficiency η_t of the solar panel is temperature-dependent, meaning it decreases as the temperature increases. This relationship is modeled as Equation (3) [23]:

$$\eta_t = \eta_{ref} \times (1 - \beta(T_{panel} - T_{ref})) \tag{3}$$

η_t = Efficiency of the panel at temperature
η_{ref} = Reference efficiency at 25°C
β = Temperature coefficient

T_{panel} = Actual temperature of the panel (°C)
T_{ref} = Reference temperature (25°C)

Solar panels generally operate at temperatures higher than the ambient environment due to the absorption of solar radiation, which is partly converted into heat. As a result, the panel's temperature can exceed the ambient temperature by 15 to 20°C, especially under direct sunlight. The temperature of the panel, T_{panel}, is influenced by solar radiation and ambient temperature. It is estimated using Equation (4) [23]:

$$T_{panel} = T_{ambient} + (\frac{G_t \times (1-\eta_{thermal})}{H_{thermal}}) \tag{4}$$

T_{panel} = Solar panel temperature
$T_{ambient}$ = Ambient temperature
G_t = Solar irradiance on the panel surface (Wh/m^2)
$\eta_{thermal}$ = Thermal coefficient of the panel (typically ranging from 0.8)
$H_{thermal}$: The thermal conductivity coefficient (we will assume it to be fixed at 15 W/m^2·K)

3.2. Wind Energy Production

For wind energy estimation, the mechanical power output of the wind turbine is calculated using the standard formula derived from Equation (5) [24]:

$$P = \frac{1}{2} \times \rho_h \times C_p \times v^3 \times A \tag{5}$$

P = Mechanical power output of the wind turbine (W)
ρ_h = Air density at height h
A = Swept area of the turbine blades ($A=\pi R^2$, where R is the radius of the turbine blades)
v = Wind speed (m/s)
C_p = Power coefficient of the wind turbine

The air density ρ is calculated based on Equation (6), which accounts for the presence of water vapor in humid air [25]:

$$\rho_h = \frac{P_{air}}{P_{dry} \times T} \times (1 - \frac{P_{vapor}}{P_{air}}) \tag{6}$$

P_{air}: Total air pressure (Pa)
P_{vapor}: Vapor pressure of water in humid air (Pa). This is calculated using temperature and relative humidity.
P_{dry}: Gas constant for dry air (287.05 J/(kg·K))
T: Absolute temperature (K)

To compute P_{vapor}, the saturation vapor pressure of water (P_{set}) is estimated using the Antoine equation for temperatures below 100°C (as given in Equation (7)) [26]:

$$log_{10} P_{set} = A' - \frac{B'}{T + C'} \tag{7}$$

P_{set}: Saturation vapor pressure of water (in mmHg)
T: Temperature (in °C)
A', B', and C': Antoine constants for water.

The specific values of the constants for the Antoine equation for water in the temperature range of 0°C to 100°C are as follows: $A'=8.07131$A, $B'=1730.63$B, $C'=233.426$C. The vapor pressure of water P_{Pvapo} can then be determined using the relative humidity (RH) as Equation (8) [25, 26]:

$$P_{\mathrm{Pvapo}} = P_{set} \times 133.322 \times \frac{RH}{100} \tag{8}$$

RH is the relative humidity (%), and 133.322 is the conversion factor from mmHg to Pa. Air density also varies with altitude, and this relationship is modeled as Equation (9) [27]:

$$\rho_h = \rho_0 \times e^{\frac{h}{H}} \tag{9}$$

ρ_h is the air density at height h
ρ_0 is the air density at sea level
h is the height above sea level in meters.
H is the scale height (typically around 8500 meters)

Wind speed at various heights is calculated using the Power Law, which converts the wind speed from a reference height h_0 to a desired height h, given in Equation (10) [24]:

$$V_h = V_0 \times (\frac{h}{h_0})^{\alpha} \tag{10}$$

V_h is the wind speed at height $h = 10,15,25,50$ m
V_0 is the wind speed at the reference height h_0 = 10 m
h is the desired height.
h_0 is the reference height
α is the power coefficient

The energy produced by the wind turbine is influenced by the wind speed (v) and the air density (ρ), and it can be calculated over time by integrating the power output.

4. Data and Preprocessing

Meteorological data, such as temperature, humidity, and wind speed, are collected as the primary inputs for the energy production models. These data are used to model the output of both solar panels and wind turbines, converting the raw meteorological data into energy predictions.

The following steps outline the process:

1. Data Collection: Meteorological data, including hourly temperature, humidity, wind speed, and solar irradiance, are collected from reliable sources.
2. Modeling Energy Production: The collected data are used to calculate solar energy output using the solar irradiance formula and wind energy output using the wind turbine formula.
3. Preprocessing: Data are cleaned and normalized to ensure accuracy and consistency.
4. Model Training: The energy production data, generated from the meteorological inputs, serve as the ground truth for training ML models. These models learn the relationships between weather parameters and energy outputs.
5. Forecasting: The trained models are then used to predict energy generation based on future meteorological forecasts.

These steps result in accurate, data-driven models capable of predicting energy production from solar and wind sources under varying climatic conditions.

4.1. Selection of Solar Panel and Wind Turbine

For this study, a solar panel and a wind turbine have been selected to represent the typical renewable energy systems in the region of Ahvaz. These systems will be used to evaluate the potential energy generation from solar and wind resources in the city.

The selected solar panel is the MEP600-T144-GG model from Mana Energy, an Iranian manufacturer, with a rated capacity of 575-600 watts. The panel has an area of 2278×1134 cm^2 and operates at an efficiency of 23.2% at a standard temperature of 25°C [28]. Additionally, the temperature coefficient of the panel is 0.046 **%**per degree Celsius, which implies that the efficiency will decrease slightly with increases in temperature, a crucial factor in regions with hot climates like Ahvaz. A sample of the Mana Energy panel in a solar power plant is shown in Figure 1.

The wind turbine chosen for this study is the Max600 model from NewSkyPower, with a rated capacity of 600 watts. Manufactured in China, it has a starting wind speed of 1.5 m/s and a blade diameter of 1.7 meters [29]. The power coefficient (C_p) for this model is typically between 0.25 and 0.35, making it suitable for areas with lower wind speeds. This turbine will offer useful insights into the wind energy potential in Ahvaz, which experiences lower wind speeds, especially during certain seasons of the year. The structure of the selected turbine is shown in Figure 2.

Figure 1. 140 MW Solar Power Plant in Mahallat City - Markazi Province – Iran [28].

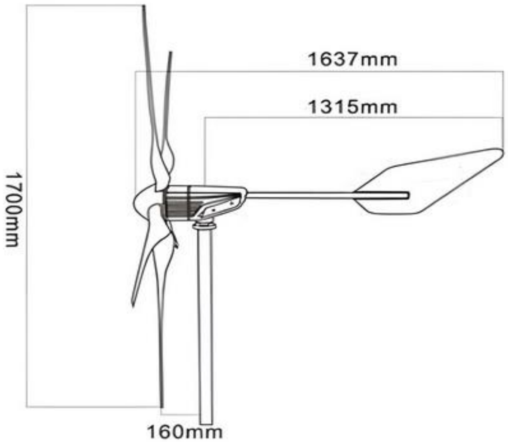


Figure 2. The wind turbine, the Max600 model from NewSkyPower [29].

These components have been selected to evaluate the potential of solar and wind energy generation under comparable capacities, allowing for an analysis of their performance in meeting the city's energy demand. By considering both solar and wind resources, this study aims to explore the optimal energy mix for Ahvaz to address its energy demand, especially during peak consumption periods.

4.2. ML and DL Models for Energy Prediction

In the context of predicting solar and wind energy production, the problem at hand is a regression task, where the goal is to predict continuous energy output values based on input features such as weather conditions and the physical attributes of energy systems. To address this issue, various ML models can be employed, each offering unique advantages. The ML techniques, such as SVM, KNN, and StackedBoost-XG, are capable of capturing the relationships between the input variables and the energy production. The models and their applications for energy prediction will be introduced in detail in the following.

4.2.1. SVM

The SVMs are powerful supervised learning algorithms used for classification and regression problems. The model works by identifying an optimal hyperplane that separates data into different classes. The key feature of SVM is the maximization of the margin between classes, ensuring the best separation between data points. The method for selecting the hyperplane is discussed in Figure 3. This hyperplane can be linear or non-linear, depending on the problem's complexity. In high-dimensional spaces, SVM can handle both linear and non-linear problems effectively. SVM's ability to work with large datasets and find complex decision boundaries makes it a suitable choice for various real-world applications.

4.2.2. K-NN

The K-NN algorithm is a non-parametric method used for classification and regression. This process, illustrated in Figure 4, works by identifying the "K" closest training samples to a new data point and assigning a label based on majority voting (for classification) or averaging the labels (for regression) of those neighbors. The distance between data points is calculated using various distance metrics, such as Euclidean distance, Manhattan distance, or cosine similarity. The choice of "K" is crucial as it affects the model's performance. A small "K" may lead to overfitting, while a large "K" may cause underfitting. K-NN is simple, interpretable, and often effective in problems where the relationships between data points are important.

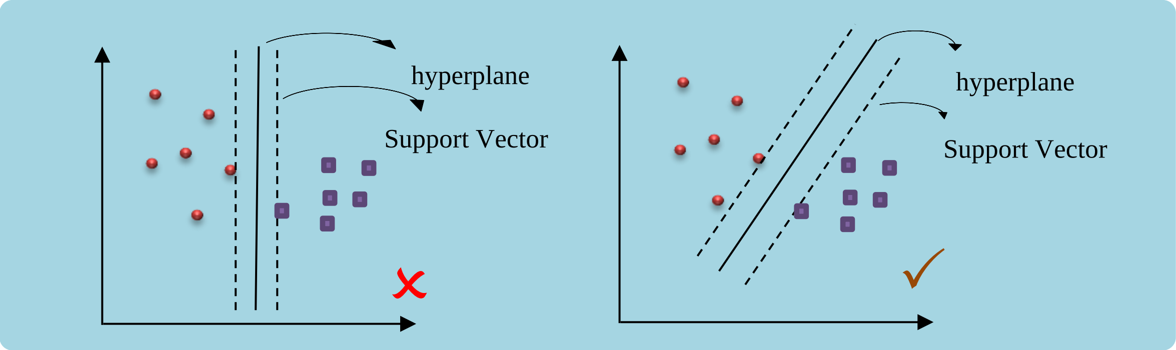


Figure 3. The method for selecting the hyperplane.

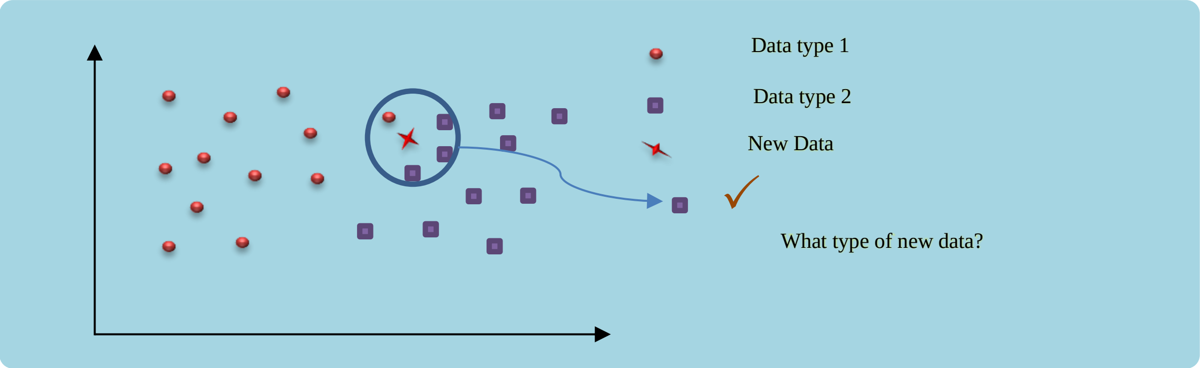


Figure 4. K-NN Algorithm Process.

4.2.3. StackedBoost-XG: Solar and Wind Energy Prediction

The StackedBoost-XG model is a hybrid and advanced approach for predicting solar and wind energy, leveraging a combination of various base models along with a more advanced final model to improve performance. Below is a more detailed and technical explanation of this model and its methods:

4.2.3.1 Base Models (SVM, KNN)

The SVM performs well on datasets that are linearly separable or clearly defined. SVM is particularly suitable for high-dimensional data (such as energy data) and when class boundaries are clearly distinct. It may not perform well with more complex or non-linear data. The KNN model excels in simulating complex, non-linear relationships between data points, relying on similar data for predictions. In noisy datasets or with additional irrelevant features, KNN may struggle. It is highly sensitive to the scaling of data and can perform poorly in noisy environments.

4.2.3.2 Final Model: XGBoost

XGBoost is an advanced algorithm designed to enhance other models through the boosting process. This algorithm excels in predictive tasks involving complex and non-linear data and incorporates features that help prevent overfitting and improve prediction accuracy [30]. Role in Model Combination: After each of the base models (SVM and KNN) generates predictions, XGBoost takes these predictions as inputs and delivers the final result based on the best possible combination of them. This final step acts as the "combiner," improving the strengths of each base model while reducing their limitations.

4.2.3.3 Model Combination (Stacking)

Stacking is a learning method where multiple base models are used simultaneously to make initial predictions, followed by another model (in this case, XGBoost) to combine these predictions and provide the final output [31]. This method typically enhances prediction accuracy by leveraging models with different characteristics. As a result, the base models compensate for each other's weaknesses, and the final model (XGBoost) intelligently combines the predictions.

4.2.3.4 Advantages of StackedBoost-XG

Leveraging the Strengths of Each Model: Each base model brings unique strengths to the table and excels at specific types of predictions. However, each also has its weaknesses. By combining these models with XGBoost, the overall prediction accuracy improves. Mitigating Model Weaknesses: The weaknesses of each base model are compensated for by other models, and the final model (XGBoost) effectively utilizes the strengths of these combinations. The combined model is shown in Figure 5, where the structure of the StackedBoost-XG framework is illustrated.

4.3. Short-Term Forecasting of Solar and Wind Energy Potential

For short-term forecasting of solar and wind energy potential, the accuracy and reliability of input data are paramount. In this study, hourly energy potential is predicted using a dataset that includes both historical and real-time meteorological data. The dataset is divided into training, validation, and test sets, as summarized in Table 1. The model uses real-time weather variables collected over time, referred to as past real-time meteorological data, to reflect actual atmospheric conditions relevant to each hourly prediction. These include temperature, humidity, solar irradiance, wind speed, and other factors detailed in Table 2. While the data is historical in terms of collection period, it retains its real-time characteristics for each recorded instance, making it suitable for developing a practical short-term forecasting model.

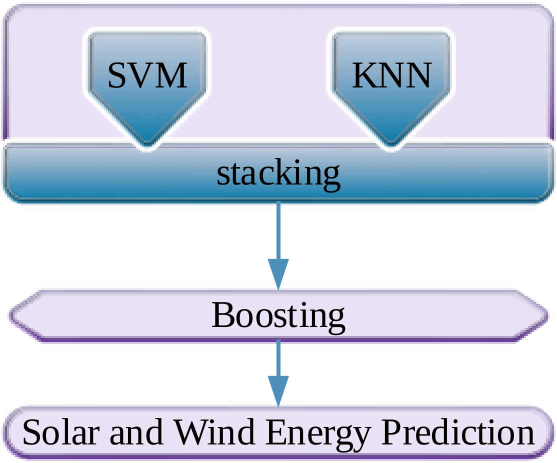


Figure 5. StackedBoost-XG: An Ensemble Model Leveraging Multiple Base.

The forecasting horizon is set to one hour ahead, classifying the model as a short-term prediction framework. Therefore, this approach integrates both past real-time meteorological observations and historical energy production values to improve the precision of hourly predictions. Ultimately, this methodology supports improved energy grid management and enhances the reliability of the renewable energy supply in Ahvaz.

4.4. Evaluation Metrics for Model Comparison

To assess and compare the performance of different predictive models for solar and wind energy forecasting, key metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2) are employed. MAE measures the average magnitude of prediction errors, giving insight into how close predicted values are to actual outcomes. RMSE emphasizes larger errors, making it useful for identifying models sensitive to significant deviations. R^2 indicates the proportion of variance explained by the model, with values closer to 1 suggesting better model fit. These metrics collectively help determine the most accurate and reliable model for forecasting energy production. In addition, feature importance was evaluated using XGBoost, which provides insights into the most influential features affecting energy forecasting for both solar and wind energy [32]. To achieve this, the importance of features was computed using XGBoost's built-in feature importance method, and the results were normalized to a scale of 0 to 100 for better interpretability.

5. Results, Discussion, and Insights

Ahvaz is situated at an altitude of only 18 meters above sea level. The climate in Ahvaz is generally warm, with an average annual temperature of 26.9˚C and clear skies [33]. The area enjoys a high level of solar exposure, with an average daily solar radiation of 5012 Wh/m^2 [33]. Solar energy is most effective between 7 AM and 5 PM, allowing for an extended period of sunlight that can be harnessed for renewable energy production [33]. As shown in Table 3, wind speed at different heights was estimated using the power law relationship with the coefficient α [33]. The α value was first determined based on the mean wind speed at the reference height of 10 meters and the mean wind speed at the unknown height, according to Equation (10). This coefficient was then used to convert the wind speed from the reference height to the desired height. This relatively moderate wind speed, combined with the high levels of solar radiation, positions Ahvaz as an ideal location for harnessing both solar and wind energy. The SVR model utilizes the Radial Basis Function (RBF) kernel, effective for modeling complex, non-linear relationships in the data. The KNN model is configured with 6 neighbors, optimizing a balance between flexibility and generalization. The Stacking Model combines the predictions of SVR and KNN, using XGBoost as the final estimator, thereby enhancing predictive performance by leveraging the strengths of both base models and the boosting algorithm. The optimal parameters for these models were determined through a combination of trial and error and systematic tuning methods. These models were implemented using the scikit-learn library in Python.

The potential energy of solar panels has been calculated under two scenarios: one accounting for thermal losses and the other excluding them. The results for these scenarios are presented in Figure 6. The potential energy of the solar panels shows a slight reduction when thermal losses are considered, with the gap being more noticeable in warmer months. The impact of thermal losses is minimal during winter, with differences under 0.2 kWh. However, during summer, especially in May (0.5 kWh difference), thermal losses have a more significant effect. The percentage loss due to thermal effects can be estimated at around 0.7% to 1.5% of the total energy output, depending on the month and environmental conditions. This suggests that managing panel temperature through cooling or improved materials could reduce losses during warmer months. For the wind turbine, potential energy has been calculated across four distinct scenarios, based on different installation heights: 10 meters, 15 meters, 25 meters, and 50 meters. Each scenario reflects the energy output at varying heights, providing insight into how turbine performance changes with altitude. The results are shown in Figure 7. The data for wind turbine energy show that increasing height leads to a higher energy output.

Table 1. Dataset Splitting for Energy Forecasting.

Dataset Type	Date Range	Percentage of Data	Time Interval
Training Set	October 1, 2020 – September 30, 2021	90%	Hourly
Validation Set	October 1, 2020 – September 30, 2021	10%	Hourly
Test Set	October 2, 2021 – October 12, 2021	100%	Hourly

Table 2. Input Variables for Energy Forecasting.

Input Variables	Description
Energy Data	Hourly solar and wind energy output (Wh)
Calendar Data	Time, day, month, year, and season
Meteorological Data	Wind speed (m/s), air pressure (Pa), air density(kg/m^3), temperature (˚C), Humidity (%)

Table 3. Wind speed information.

Height (m)	Average Annual Wind Speed (m/s)	Estimated α Coefficient from 10 m
10	2.69	0
15	2.80	0.0988
25	3.02	0.1669
50	3.59	0.1793

5.1. Data-Driven Estimation of Solar and Wind Energy Potential with Thermal Losses and Optimal Heights

However, the difference between 25 meters and 50 meters is minimal, suggesting that beyond 25 meters, the increase in energy production is limited relative to the rising installation costs. The most substantial change in energy output occurs between 15 meters and 25 meters, where energy production increases significantly, making this the optimal height range for installation.

5.2. Short-Term Forecasting of Solar and Wind Energy Potential

The StackedBoost-XG model, a combination of SVM and KNN, outperforms the individual SVM and KNN models in terms of prediction accuracy for both solar energy and wind energy forecasting. This improvement can be attributed to the complementary strengths of both models within the ensemble approach. The performance of these models has been evaluated using a set of key metrics: RMSE, MAE, and R^2, across both the validation and test datasets, as presented in Table 4. All simulations were carried out on a laptop equipped with an Intel Pentium® CPU N4200 (1.1 GHz) and 8 GB of RAM. The maximum training time for the StackedBoost-XG model was approximately 19 seconds, demonstrating the model's computational efficiency and suitability for real-time or near-real-time forecasting applications.

5.2.1. Analysis of Solar Energy Forecasting

The forecasting of solar energy potential, as shown in Figure 8, demonstrates the impact of thermal losses and the accuracy of the various models in predicting solar energy output. In the case, as shown in Table 4, the StackedBoost-XG model demonstrates a significant improvement in accuracy compared to the individual models. Specifically, the RMSE of the combined model is 3.78, which is lower than both SVM (RMSE = 4.12) and KNN (RMSE = 4.08). The reduced RMSE indicates that the StackedBoost-XG model better handles both thermal losses and non-linear relationships between variables such as temperature, radiation, and historical energy production. The ability of SVM to capture global patterns and of KNN to address local variations is what allows this combined model to offer superior predictions, particularly in complex scenarios like solar energy forecasting, where both global and local factors influence the output. The SVM and KNN models show RMSE and MAE values that are higher than StackedBoost-XG, which indicates they are more sensitive to the variance in energy production and less effective at minimizing large errors during peak thermal loss months (e.g., summer).

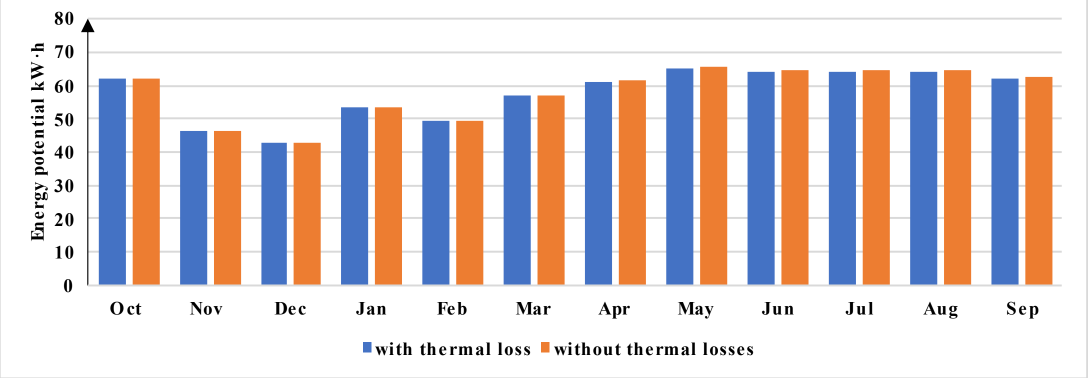


Figure 6. Potential energy of the solar panel in one year.

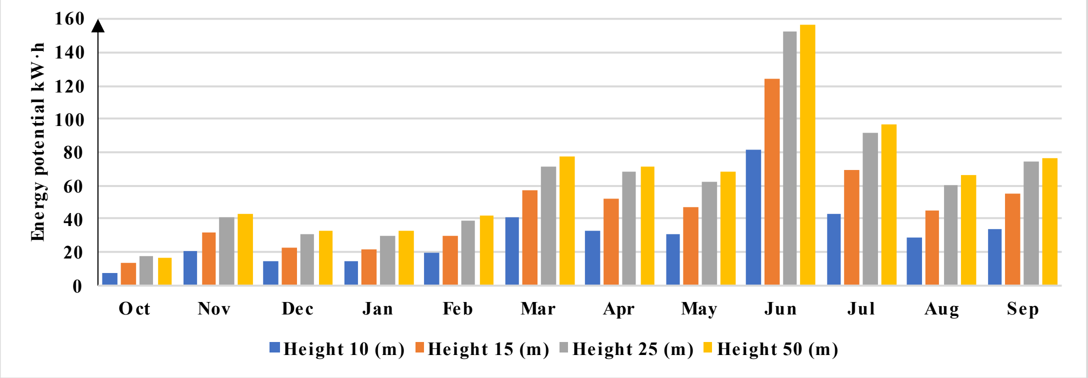


Figure 7. Potential energy of the wind turbine in one year at different heights.

5.2.2. Analysis of Wind Energy Forecasting

For wind energy forecasting, the models' performance is evaluated at different installation heights: 15 meters and 25 meters, as shown in Figures 9 and 10. The results reveal that wind energy predictions improve as the installation height increases. Similarly, in wind energy forecasting, as demonstrated in Table 4, the StackedBoost-XG model provides the lowest RMSE values and highest R^2 scores. At a 15-meter installation height, the model achieves an RMSE of 2.87, and at 25 meters, it achieves 3.09. In contrast, the individual models SVM and KNN perform less effectively, with RMSE values of 3.45 and 3.56, respectively. These results highlight the model's ability to more accurately capture the underlying dynamics of wind energy production, which can be influenced by factors such as installation height and local wind patterns. By combining SVM's global pattern recognition and KNN's sensitivity to local fluctuations, the ensemble model offers a more robust and accurate forecasting approach.

5.2.3. Performance Improvement of StackedBoost-XG in Energy Forecasting

The StackedBoost-XG model demonstrates a notable enhancement in forecasting both solar and wind energy compared to the individual models (SVM and KNN). For solar energy, the combined model shows a ~5% improvement in accuracy over the SVM model and a ~7% improvement over KNN. In wind energy forecasting, the improvements are even more pronounced. At a 15-meter installation height, the StackedBoost-XG model achieves a ~34% increase in accuracy over SVM and a ~21% improvement over KNN. At a 25-meter height, the model delivers a ~50% improvement compared to SVM and a ~48% increase compared to KNN. Overall, the StackedBoost-XG model results in a significant boost in forecasting accuracy, with an average increase of ~5-48% across both solar and wind energy predictions, making it a more reliable and efficient tool for energy forecasting.

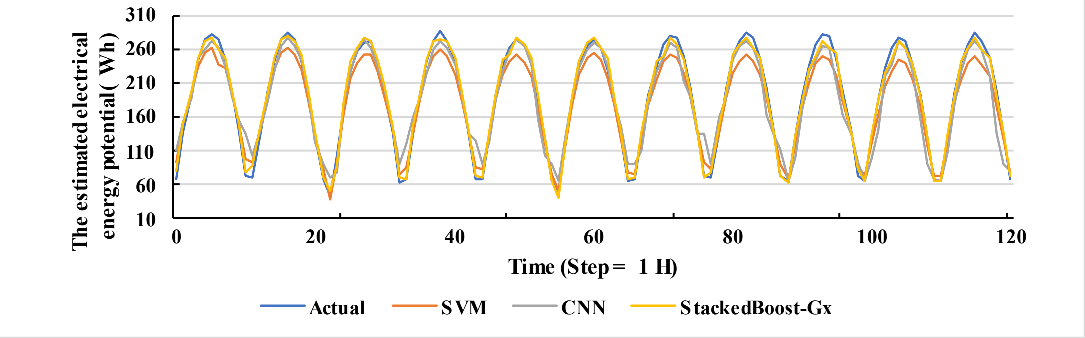


Figure 8. Forecasting the solar energy potential.

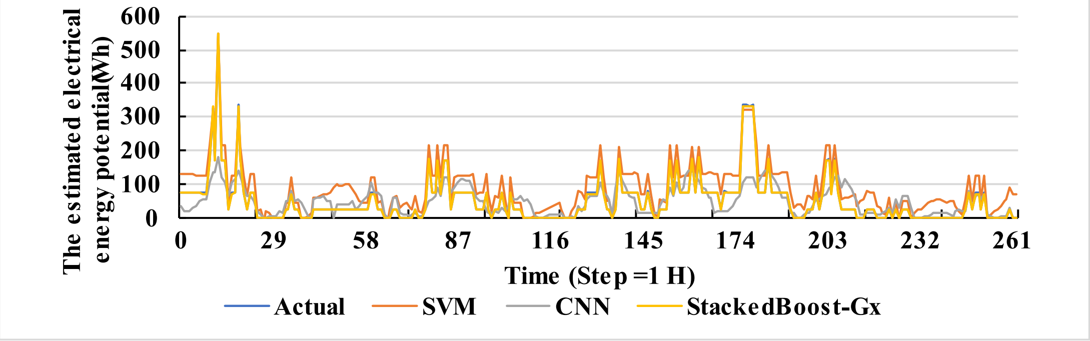


Figure 9. Forecasting the wind energy potential at a height of 15 meters.

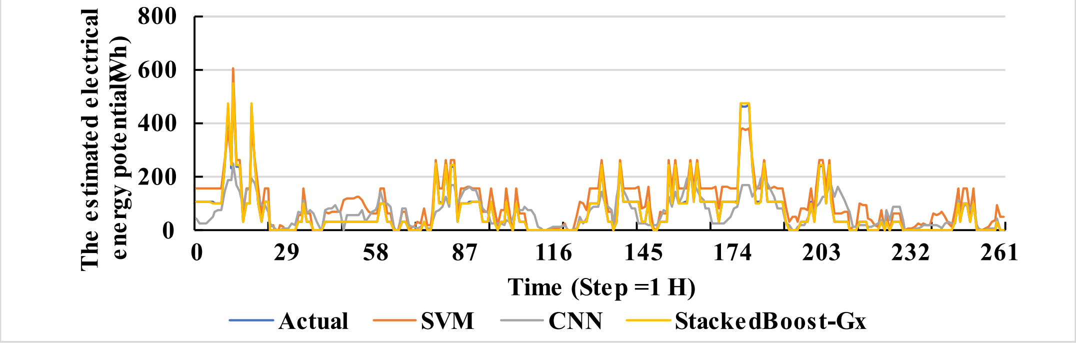


Figure 10. Forecasting the wind energy potential at a height of 25 meters.

Table 4. Comparison of Predictive Model Performance for Solar and Wind Energy Forecasting.

Model	Energy Type	Validation Set				Test Set			Limitations
		RMSE Wh	MAE Wh	R²	Time (s)	RMSE Wh	MAE Wh	R²	
SVM	Solar Energy	11.83	9.8	0.97	0.14	18.43	15.80	0.94	Thermal losses considered
	Wind Energy	39.84	32.68	0.87	1.06	41.69	37.94	0.65	Performance at 15 meters
	Wind Energy	40.51	34.34	0.90	1.72	43.49	38.10	0.78	Performance at 25 meters
KNN	Solar Energy	12.51	9.24	0.97	0.02	21.41	16.21	0.92	Thermal losses considered
	Wind Energy	29.99	14.26	0.92	0.05	50.86	29.81	0.49	Performance at 15 meters
	Wind Energy	34.09	17.77	0.93	0.05	65.81	40.61	0.51	Performance at 25 meters
StackedBoost-XG	Solar Energy	11.04	8.76	0.99	1.06	7.35	5.65	0.99	Thermal losses considered
	Wind Energy	2.87	1.02	0.99	14.76	0.91	0.63	0.99	Performance at 15 meters
	Wind Energy	5.18	2.30	0.99	18.36	3.09	1.79	0.99	Performance at 25 meters

5.3. *Sensitivity Analysis of Key Factors Affecting Solar and Wind Energy Forecasting*

The sensitivity analysis, presented in Figure 11, highlights the varying importance of features in predicting solar and wind energy potential across three scenarios:

For the solar panel scenario, time (100) is the most influential feature, indicating that solar energy generation is predominantly driven by the daily sunlight cycle. Temperature (1) and humidity (72) are also important, with humidity being particularly noteworthy. High humidity or fog can obstruct sunlight, reducing the amount of solar radiation that reaches the panels. Therefore, the higher importance of humidity reflects its impact on solar panel output. Season and day contribute minimally (0 and 0.25, respectively), suggesting that solar energy is less influenced by seasonal variations or specific days. In the wind turbine scenarios, wind speed and air density emerge as the dominant factors. For the wind turbine at 15 m height, wind speed (59.49) and air density (100) are highly influential, and the trend is similar for the wind turbine at 25 m height (58.63 and 100). This underscores the reliance of wind energy on real-time environmental conditions such as wind speed and air density, which directly affect turbine efficiency. Time and humidity have a lower impact, further confirming that wind energy is primarily driven by atmospheric conditions.

6. Conclusions

This study evaluated the solar and wind energy potential in Ahvaz using a combination of machine learning techniques and energy system simulations. The region's high solar irradiance and moderate wind speeds make it a promising site for hybrid renewable energy systems. Results showed that the 600W wind turbine produces approximately 340.56 kWh/year at 10 meters, increasing to 523.59 kWh at 15 meters, 681.48 kWh at 25 meters, and 719.03 kWh at 50 meters. However, the marginal gain from 25 to 50 meters is relatively small, suggesting that 25 meters represents an optimal trade-off between energy yield and installation cost. For solar energy, the 600W solar panel generates 691.04 kWh/year under optimal conditions. The panel's output showed consistent energy production across the year, with slight reductions during warmer months due to thermal losses. Wind energy, in contrast, showed greater seasonal variability but significantly higher output when installed at appropriate heights. Therefore, combining both sources could ensure balanced, year-round energy availability, particularly valuable during peak demand periods. From a forecasting perspective, the StackedBoost-XG model, which integrates SVM and KNN with XGBoost as a meta-learner, consistently outperformed individual models. It achieved accuracy improvements of 5–7% for solar energy, and 21–50% for wind energy, depending on height. These improvements demonstrate the advantage of ensemble learning for capturing both global and local data patterns in non-linear, multivariate systems. Additionally, feature sensitivity analysis indicated that time and humidity are critical for solar energy, while wind speed and air density dominate wind energy predictions. These insights are crucial for informed site selection, system design, and adaptive energy management strategies. Overall, the study underscores the value of machine learning in optimizing renewable energy systems and recommends a hybrid solar-wind configuration with wind turbines installed between 15–25 meters, coupled with predictive ML tools, as the most effective strategy for sustainable energy planning in Ahvaz. Future research could validate the model in diverse climatic regions, integrate it with smart grid systems for real-time energy management, and assess the hybrid system's cost-effectiveness and environmental impact.

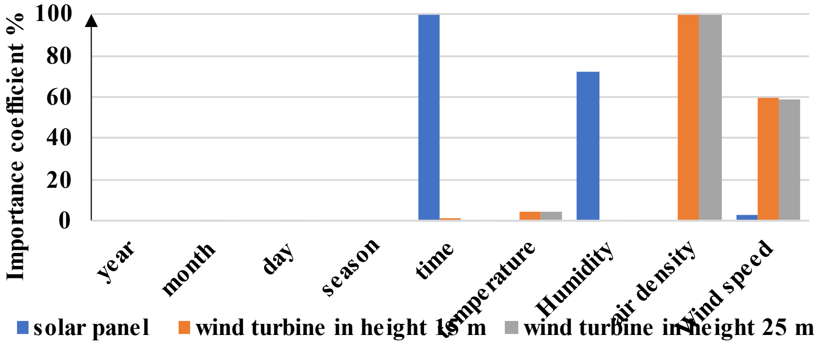


Figure 11. The importance coefficient of features in predicting solar and wind potential energy.

Nomenclature

Symbol	Description
E	Energy output from solar panel or wind turbine (Wh)
G_t	Solar irradiance on the panel at time t (Wh/m^2)
A	Area of solar panel (m^2) or swept area of wind turbine blades (m^2)
η_t	Efficiency of solar panel at time t (dimensionless)
G_{max}	Daily average solar irradiance (Wh/m^2)
RH	Relative humidity (dimensionless)
K	Humidity correction factor (dimensionless)
$t_{sunrise}$	Time of sunrise (hours)
t_{sunset}	Time of sunset (hours)
t	Current time of the day
η_{ref}	Reference efficiency at 25 ˚C (dimensionless)
β	Temperature coefficient (˚C)
T_{panel}	Temperature of solar panel (˚C)
$T_{ambient}$	Ambient temperature (˚C)
T_{ref}	Reference temperature (25 ˚C)
$H_{thermal}$	Thermal conductivity coefficient (W/m^2·K)
$\eta_{thermal}$	Thermal coefficient of panel (dimensionless)
P	Mechanical power output of wind turbine (W)
ρ	Air density (kg/m^3)
C_p	Power coefficient of wind turbine (dimensionless)
v	Wind speed (m/s)
R	Radius of wind turbine blades (m)
A',B',C'	Antoine constants for water
P_{air}	Total air pressure (Pa)
P_{vapor}	Vapor pressure of water in humid air (Pa)
P_{dry}	Gas constant for dry air (287.05 J/(kg·K))
T	Absolute temperature (K)
P_{set}	Saturation vapor pressure of water (mmHg)
h	Height above sea level (m)
H	Scale height (m)
V_h	Wind speed at height h (m/s)
V_0	Wind speed at reference height h_0 (m/s)
h_0	Reference height (m)
α	Power coefficient for wind speed conversion (dimensionless)

Symbol	Description
ML	Machine Learning
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
XGBoost	Extreme Gradient Boosting
StackedBoost-XG	An Ensemble Model Leveraging Multiple Base
ANN	Artificial Neural Network
ESM	Earth System Model
XAI	Explainable Artificial Intelligence
QAI	Quantum Artificial Intelligence

References

[1] M. J. B. Kabeyi, and O. A. Olanrewaju, "Sustainable Energy Transition for Renewable and Low Carbon Grid Electricity Generation and Supply," *Frontiers in Energy Research*, vol. 9, 2022.

[2] J. L. Holechek, H. M. E. Geli, M. N. Sawalhah, and R. Valdez, "A Global Assessment: Can Renewable Energy Replace Fossil Fuels by 2050?," *Sustainability*, vol. 14, no. 8, 4792, 2022.

[3] L. Amjith, and B. Bavanish, "A Review on Biomass and Wind as Renewable Energy for Sustainable Environment," *Chemosphere*, vol. 293, 133579, 2022.

[4] A. Rahman, O. Farrok, and M. M. Haque, "Environmental Impact of Renewable Energy Source Based Electrical Power Plants: Solar, Wind, Hydroelectric, Biomass, Geothermal, Tidal, Ocean, and Osmotic," *Renewable and Sustainable Energy Reviews*, vol. 161, 112279, 2022.

[5] Y. Ebazadeh, R. Alayi, E. Jamali, and A. Behvandi, "Modeling and Optimization of The Photovoltaic System Connected to the Grid," *Journal of Green Energy Research and Innovation*, vol. 2, no. 1, pp. 32–43, 2025.

[6] A. Q. Al-Shetwi, "Sustainable Development of Renewable Energy Integrated Power Sector: Trends, Environmental Impacts, and Recent Challenges," *Science of The Total Environment*, vol. 822, 153645, 2022.

[7] F. Masteri Farahani, A. Kazemi, and A. Hedayati Aghmashadi, "Optimal Site Selection of Solar Power Plant Stations Using GIS-ANP and Genetic Optimization Algorithm in Markazi Province, Iran," *Journal of Green Energy Research and Innovation*, vol. 1, no. 4, pp. 47–63, 2024.

[8] C. Sweeney, R. J. Bessa, J. Browell, and P. Pinson, "The Future of Forecasting for Renewable Energy," *WIREs Energy and Environment*, vol. 9, no. 2, 2019.

[9] A. H. Karamali, A. Daeichian, S. Rezaei, and A. Reihanian, "Assessing Wind Energy Potential in Markazi Province, Iran: A Data-Driven Approach with AI Algorithms,"

Journal of Green Energy Research and Innovation, vol. 2, no. 2, pp. 26–35, 2025.
[10] R. Meenal, D. Binu, et al., "Weather Forecasting for Renewable Energy System: A Review," *Archives of Computational Methods in Engineering*, vol. 29, no. 5, pp. 2875–2891, 2022.
[11] M. H. Shakoor, "A Survey on Renewable and New Sources of Energies for Electricity Power Production and Its Challenges," *Journal of Green Energy Research and Innovation*, vol. 1, no. 4, pp. 64–85, 2024.
[12] A. Javaid, M. Sajid, E. Uddin, A. Waqas, and Y. Ayaz, "Sustainable Urban Energy Solutions: Forecasting Energy Production for Hybrid Solar-Wind Systems," *Energy Conversion and Management*, vol. 302, 118120, 2024.
[13] H. Wang, N. Zhang, et al., "A Comprehensive Review for Wind, Solar, and Electrical Load Forecasting Methods," *Global Energy Interconnection*, vol. 5, no. 1, pp. 9–30, 2022.
[14] G. Chen, and Z. Ji, "A Review of Solar and Wind Energy Resource Projection Based on the Earth System Model," *Sustainability*, vol. 16, no. 8, 3339, 2024.
[15] V. Z. Manusov, S. K. Khaldarov, and B. V. Palagushkin, "Short-Term Forecasting of Wind and Solar Power Generation," *Journal of Physics: Conference Series*, vol. 2131, no. 5, 052050, 2021.
[16] A. Vennila, S. Balambigai, et al., "Predicting Solar and Wind Power Production with the Weather Map Data Approach," *2024 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT)*, pp. 1–6, 2024.
[17] V. P. Singh, A. Kumar, C. S. Meena, and N. Dutt, "Understanding Forecasting Models for Renewable Energy Generation and Market Operation," *Forecasting Methods for Renewable Power Generation*, pp. 95–130, 2025.
[18] M. Abdul Baseer, A. Almunif, I. Alsaduni, and N. Tazeen, "Electrical Power Generation Forecasting from Renewable Energy Systems Using Artificial Intelligence Techniques," *Energies*, vol. 16, no. 18, 6414, 2023.
[19] N. E. Benti, M. D. Chaka, and A. G. Semie, "Forecasting Renewable Energy Generation with Machine Learning and Deep Learning: Current Advances and Future Prospects," *Sustainability*, vol. 15, no. 9, 7087, 2023.
[20] B. N. Rao, M. Praveen, and D. R. Babu, "A Review on the Role of AI in Optimizing Renewable Energy Grid Management," *International Journal of Scientific Research in Engineering and Management*, vol. 8, no. 11, pp. 1–13, 2024.
[21] L. A. Yousef, H. Yousef, and L. Rocha-Meneses, "Artificial Intelligence for Management of Variable Renewable Energy Systems: A Review of Current Status and Future Directions," *Energies*, vol. 16, no. 24, 8057, 2023.
[22] D. Y. Goswami, *Principles of Solar Engineering*, CRC Press, 2022.
[23] J. A. Duffie, W. A. Beckman, and N. Blair, *Solar Engineering of Thermal Processes, Photovoltaics and Wind*, John Wiley & Sons, 2020.
[24] J. F. Manwell, J. G. McGowan, and A. L. Rogers, *Wind Energy Explained: Theory, Design and Application*, John Wiley & Sons, 2010.
[25] R. D. Zucker and O. Biblarz, *Fundamentals of Gas Dynamics*, John Wiley & Sons, 2019.
[26] J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, Elsevier, 2006.
[27] C. David Cooper, and F. Chris Alley, "Air Pollution," *Electrical Engineering Handbook*, 2004.
[28] manaenergypak, "Solar Power Plant Projects,".
[29] newskypower, "Products,".
[30] P. P. Deka and J. Weiner, *XGBoost for Regression Predictive Modeling and Time Series Analysis: Learn How to Build, Evaluate, and Deploy Predictive Models with Expert Guidance*, Packt Publishing Ltd, 2024.
[31] J. Brownlee, *Ensemble Learning Algorithms with Python: Make Better Predictions with Bagging, Boosting, and Stacking*, Machine Learning Mastery, 2021.
[32] B. Zhang, Y. Zhang, and X. Jiang, "Feature Selection for Global Tropospheric Ozone Prediction Based on the BO-XGBoost-RFE Algorithm," *Scientific Reports*, vol. 12, no. 1, 2022.
[33] globalsolaratlas, "Map,".

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, have been completely observed by the authors.

Bibliography

Mehdi Mohammadian Mehr received a Bachelor's degree in Electrical Engineering, specializing in Power Systems, from Shahid Chamran University of Ahvaz, from 2018 to 2022. He then pursued a Master's degree in Electrical Engineering with a concentration in Power Systems at the same university, from 2022 to 2024. His academic background emphasizes energy systems, particularly in the areas of renewable energy integration, smart grids, and advanced forecasting techniques. His research interests include the application of artificial intelligence in power system optimization, distribution network design and enhancement, electric load forecasting and control, power system reliability and resilience, and the integration of energy storage and renewable energy resources into power grids.
Email: Eng.mohammadianmehr@gmail.com
ORCID: 0009-0003-7270-7849
Contribution Statement: Investigation, Methodology, Software, Roles/Writing - original draft.

Hossein Farzin received the BSc and PhD degrees in Electrical Engineering from Sharif University of Technology, Tehran, Iran, in 2011 and 2016, respectively. He was a postdoctoral researcher at Sharif University of Technology, from 2016 to 2017. He is currently an Associate Professor in the Electrical Engineering Department, Shahid Chamran University of Ahvaz, Ahvaz, Iran. His research interests include microgrids design and optimization, integration of distributed energy resources and electric vehicles in smart grid, and power system reliability and resilience. Dr. Farzin ranked 2nd in Iran’s nationwide universities entrance exam in 2007, and is ranked among the world’s top 2% most cited researchers in 2021 and 2023. He has authored more than 60 journal and conference papers, and serves as an editor of the Scientia Iranica journal.
Email: Farzin@scu.ac.ir
ORCID: 0000-0002-4433-1982
Contribution Statement: Supervision, Validation, Writing-review & editing.