

A PREDICTIVE MODEL FOR ENERGY EFFICIENCY OF SOLAR POWER PLANTS USING ARTIFICIAL INTELLIGENCE

J. R. Farmonov

Karshi State University, Department of Industrial Engineering, Lecturer
jasurfarmonov.research@gmail.com

Abstract: In this study, a prototype artificial intelligence model was developed to predict energy production based on weather parameters, using data from a 100 kW subset of the 510 kW small-scale solar power plant launched under Qarshi State University in Qarshi city. Among the five solar panel arrays of the station, the most reliable dataset from one array was selected as the basis for analysis. The study integrated the plant's recorded production data throughout 2024 with corresponding meteorological indicators — temperature, humidity, wind speed, and atmospheric pressure.

The data were cleaned and merged in a Python environment, then trained using the Random Forest Regression algorithm [1]. The model was used to forecast power generation for January–February 2025.

According to the results, the model achieved a Mean Absolute Error (MAE) of 17.5 kW, a Root Mean Square Error (RMSE) of 23.3 kW, and a coefficient of determination (R^2) of 0.043. Humidity (34%) and temperature (33%) were identified as the most influential factors. The findings confirm that meteorological variables significantly affect solar energy generation under Karshi's climatic conditions and establish a scientific foundation for the development of high-accuracy forecasting systems in future research.

Keywords: Artificial Intelligence, Solar Energy, Qarshi, Random Forest Regression, Energy Forecasting, Weather Data, Data Integration, Meteorological Factors, Data Analysis, Photovoltaic System, AI Modeling.

Introduction

In recent years, the importance of renewable energy sources—particularly solar energy—has been rapidly increasing. Climate change, the reduction of greenhouse gas emissions, and the development of sustainable energy supply systems have become among the world's top priorities. Solar energy, being an environmentally friendly, safe, and virtually inexhaustible resource, has emerged as a vital component of global energy systems.

In the Republic of Uzbekistan, solar radiation averages about 3,000 hours per year, which classifies most of the country's territory as regions with high solar potential. Specifically, Qarshi city is one of the areas with the highest solar irradiance, providing excellent conditions for the efficient operation of solar power plants. Consequently, the optimization and forecasting of solar energy generation are both practically and scientifically significant tasks.

In traditional approaches, the estimation of energy production is often based solely on technical parameters, which leads to low prediction accuracy. In recent years, Artificial Intelligence (AI) and Machine Learning (ML) technologies have introduced new methodologies to this field [2], [5], [6]. These models not only enable the tracking of production trends but also allow the creation of forecasting systems that account for weather conditions, seasonality, and environmental factors.

The primary objective of this study is to develop an AI-based model that predicts solar energy generation in Qarshi's climatic conditions using meteorological parameters and to evaluate its accuracy. The Random Forest Regression algorithm was chosen as the model, as it effectively



learns nonlinear relationships, is resistant to noisy data, and performs well even with relatively small datasets.

Numerous studies have demonstrated that AI models excel in their ability to automatically learn nonlinear dependencies [2], [3] and to uncover hidden relationships among complex meteorological variables [4], [5]. However, most existing models have been developed and optimized for the climatic conditions of countries such as the United States, India, or China. The unique environment of Uzbekistan, particularly the Kashkadarya region, with its high dust concentration, sharp temperature variations, and arid climate, prevents direct adaptation of those models. Therefore, it is essential to develop a locally tailored model based on regional meteorological and production data.

Research Methodology

The research was conducted using the production data of a 510 kW small-scale solar power plant operating under Qarshi State University, along with the corresponding meteorological data for the same period. The solar power plant consists of five panel arrays, each with a capacity of 100 kW, and each array maintains its own independent dataset.

The data used in this study cover the period from January 2024 to February 2025. These datasets include information on power generated by the solar panels (kW), power consumption, energy transmitted to the grid, and total production values.

Only one array's dataset was selected as the basis of analysis for this research. This particular array was chosen due to its data integrity and minimal data loss compared to the others. The remaining arrays were reserved for future extended studies, for which the current work serves as a foundational framework.

Meteorological data were obtained from a global meteorological observation point, including parameters such as temperature ($^{\circ}\text{C}$), relative humidity (%), wind speed (m/s), and atmospheric pressure (hPa).

Both sources of data were recorded in Excel format and synchronized by date. Since the weather dataset contained a time (hourly) column, the data were converted to daily average values.

Data preprocessing involved the removal of missing or erroneous values (NaN, "0", "—") and was carried out using Python with the Pandas, NumPy, and Matplotlib libraries [14]. The date columns were standardized into a unified datetime format, after which the production and weather datasets were merged using the merge() function based on the Date column.

The Random Forest Regression algorithm was employed in this study [10]. This algorithm performs prediction through the use of multiple decision trees, effectively identifying nonlinear relationships within the data.

The model utilized the following input parameters, which included both meteorological and calendar-based features: Temperature ($^{\circ}\text{C}$), Humidity (%), Wind Speed (m/s), Atmospheric Pressure (hPa), Calendar Factors: Month, Day of Year

For model evaluation, the following performance metrics were selected: MAE (Mean Absolute Error) – average prediction error, RMSE (Root Mean Square Error) – square root of mean squared error, R^2 (Coefficient of Determination) – the model's accuracy indicator

Using these metrics, the model's prediction accuracy was analyzed and compared against the actual production data.

Results and Discussion

During the research process, the developed prototype artificial intelligence model was trained on the solar power plant's production data recorded throughout 2024 and tested to forecast the



production output for January–February 2025. The model incorporated the following input features: temperature, relative humidity, wind speed, atmospheric pressure, and calendar factors (month and day of year).

After training the Random Forest Regression model, the following evaluation results were obtained:

MAE (Mean Absolute Error)	17,50 kW
RMSE (Root Mean Square Error)	23,30 kW
R ² (Coefficient of Determination)	0,043

These results indicate that the model's prediction error ranged between 17 and 23 kW on average. The relatively low R² value suggests that the model considered only primary meteorological parameters. In other words, the limited prediction accuracy at this stage can be attributed to the restricted dataset and the absence of certain meteorological variables.

Nevertheless, the model successfully identified real production trends, confirming a clear physical correlation between variations in temperature and humidity and the amount of solar energy generated [9].

Energy Production Dynamics

Figure 1 illustrates the daily variations in energy production between 2023 and 2025. As expected, energy generation was higher during the spring and summer months, while it decreased in winter due to lower solar irradiance.

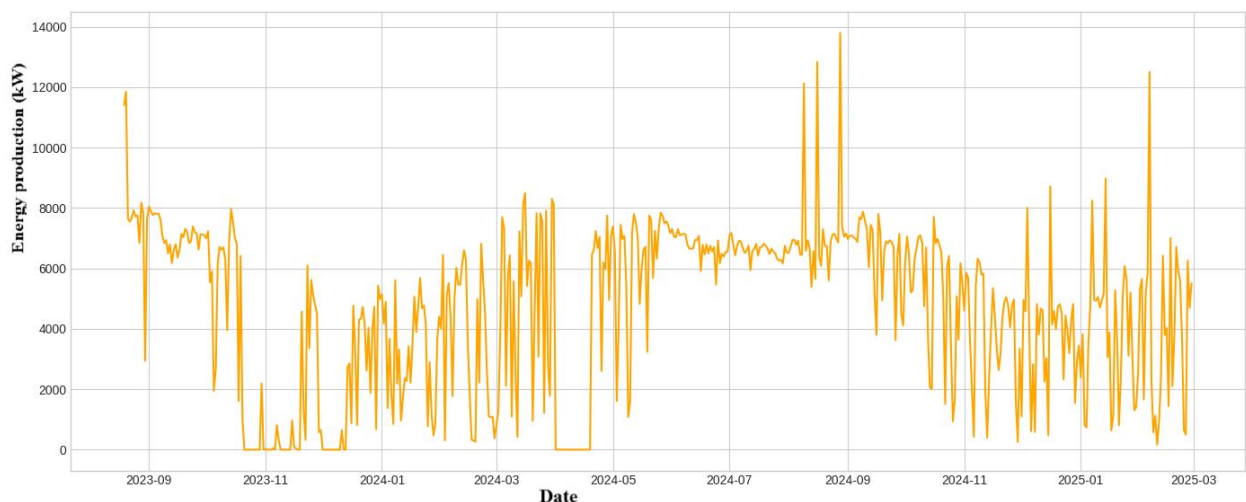


Figure 1. Daily Solar Energy Production (2023–2025)

Monthly Production Analysis

Figure 2 presents the average monthly production values. During the summer months (June–August), average energy generation ranged between 7,000–8,000 kW, while the lowest production levels were recorded in the winter months (December–January).



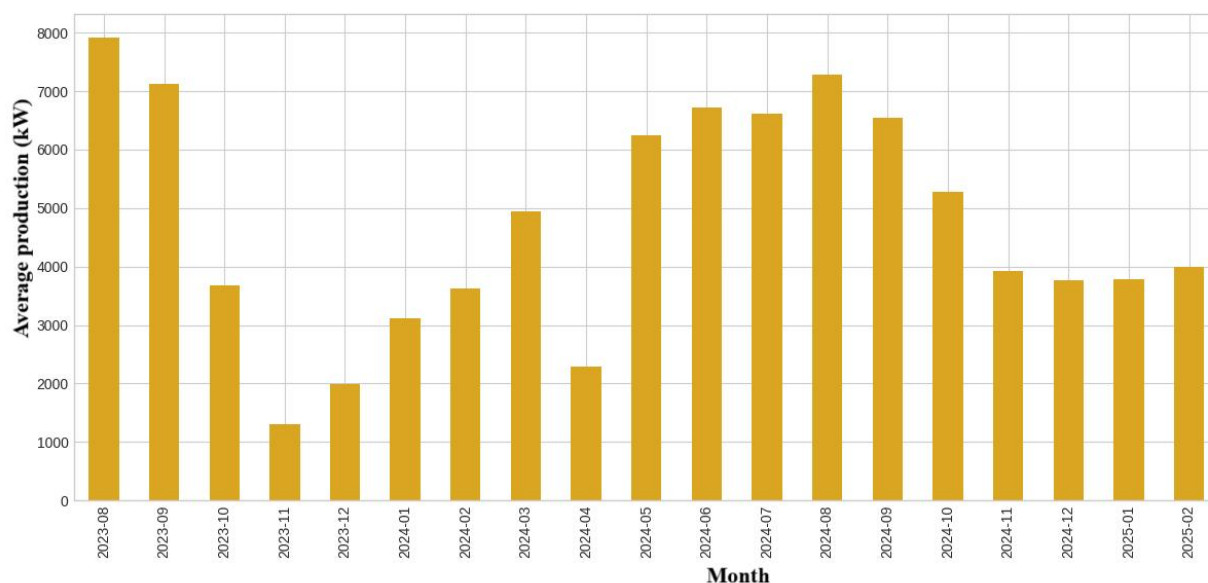


Figure 2. Average Monthly Energy Production (kW)

AI Model Results

The developed model predicted solar energy production for January–February 2025. Figure 3 compares the actual and predicted values. The model successfully captured the overall trend in production but was unable to fully reproduce some sharp fluctuations.

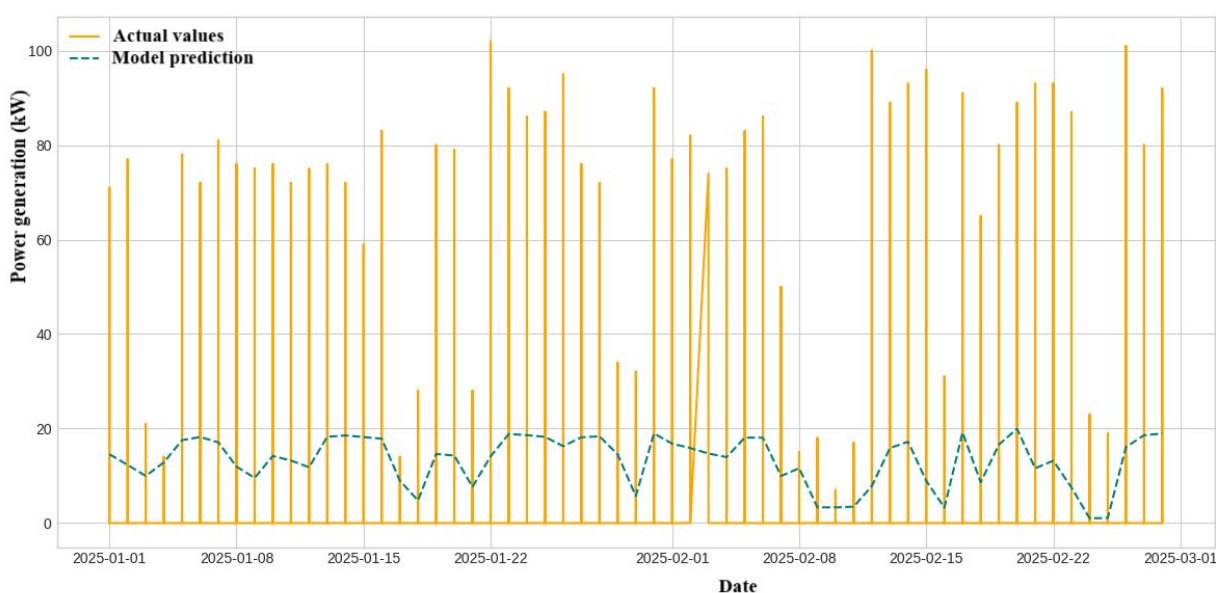


Figure 3. AI Model – Solar Energy Production Forecast (with Weather Data)

Impact of Meteorological Factors

The influence of meteorological parameters determined by the Random Forest Regression model is shown in Figure 4. The most significant factors were humidity (34%) and temperature (33%), while wind speed (10%) and atmospheric pressure (12%) had comparatively smaller effects.



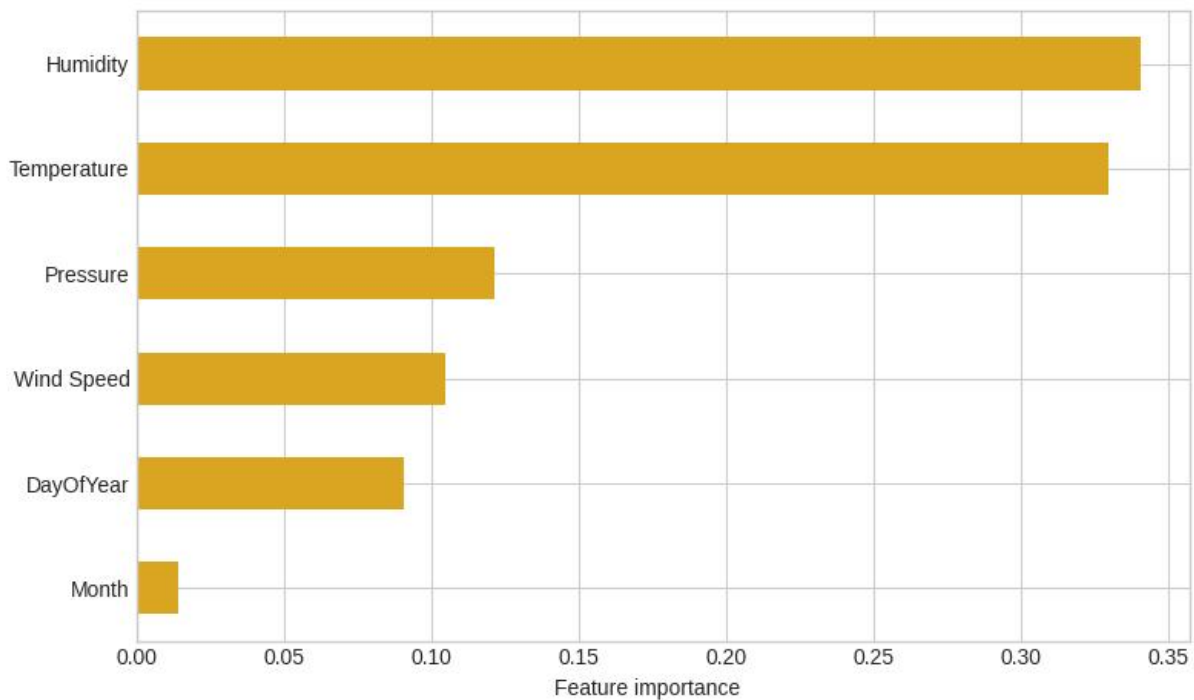


Figure 4. Impact of Meteorological Factors on Solar Energy Generation

According to the Feature Importance Analysis, the following parameters had the strongest influence on the model's predictive performance, as also depicted in Figure 4:

Factor Name	Influence Level (%)
Humidit	34,0
Temperature	32,9
Pressure	12,1
Wind Speed	10,4
Day of Year	9,0
Month	1,3

The results demonstrated that humidity and temperature exert the greatest influence on the efficiency of solar panels. When humidity increases, the moisture in the air enhances scattering of solar radiation, thereby reducing the amount of direct sunlight reaching the panels [11]. Additionally, high temperatures lower the efficiency of photovoltaic modules, as the resistance of the semiconductor layer increases with temperature.

Wind speed contributes positively by cooling the panels, which helps maintain stable performance. Atmospheric pressure, on the other hand, affects air density — an increase in pressure changes the absorption and transmission characteristics of solar radiation. Both atmospheric pressure and calendar factors serve as auxiliary variables representing seasonal variations.

Overall, the model successfully identified how solar power generation in Qarshi's climatic conditions depends on meteorological factors. The observed variations in the model correspond



well with physical and climatic principles, demonstrating the practical feasibility of forecasting energy production based on weather parameters.

Conclusion

This study presented the initial practical results of forecasting solar power generation using artificial intelligence under the climatic conditions of Qarshi. Based on data from one of the five 100 kW panel arrays of a 510 kW solar power plant, an AI model was developed that integrated meteorological parameters (temperature, humidity, wind speed, pressure) and calendar factors. The Random Forest Regression model predicted solar energy generation with an average error of 17–23 kW, while also identifying the relative influence of weather variables.

The analysis revealed that humidity and temperature were the most significant factors affecting solar panel efficiency. Increased humidity intensifies the scattering of solar radiation, thereby reducing the amount of light reaching the panels, whereas higher temperatures decrease the efficiency of photovoltaic elements due to increased semiconductor resistance. Wind speed had a moderately positive cooling effect, stabilizing panel performance.

Although the model's accuracy ($R^2 = 0.043$) remains modest, these findings demonstrate the potential of AI technologies in the energy sector. Even with a limited dataset, the model was able to capture physical correlations and forecast production behavior, confirming the practical feasibility of AI-based prediction in solar energy systems.

Future work will focus on the following directions:

- Incorporating additional parameters such as solar irradiance, cloud cover, dust accumulation, and panel surface temperature;
- Enhancing the model architecture using XGBoost, LightGBM, or LSTM neural networks [7], [8], [9];
- Integrating the model into a real-time monitoring system to enable operational forecasting and full participation of all panel arrays;
- Expanding the dataset (e.g., incorporating complete 2025 data) to achieve a comprehensive two-year analysis, enabling higher model precision.

Ultimately, this approach has significant practical importance for improving the efficiency, management, and planning of solar power plants not only in Karshi, but also throughout Uzbekistan [12], [13], [15].

References

1. L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
2. A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Progress in Energy and Combustion Science*, vol. 34, pp. 574–632, 2008.
3. C. Voyant, G. Notton et al., "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569–582, 2017.
4. G. Li and J. Shi, "Application of artificial neural networks for forecasting solar radiation data," *Renewable Energy*, vol. 35, no. 12, pp. 2801–2812, 2010.
5. M. Khandakar, M. R. Islam, and M. Alam, "Hybrid AI models for solar power prediction based on weather data," *Energy Reports*, vol. 6, pp. 1040–1051, 2020.
6. A. Diagne et al., "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.
7. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proc. 22nd ACM SIGKDD*, 2016.



8. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
9. M. Ahmed et al., "Predicting PV output using hybrid deep learning approaches," Energy Conversion and Management, vol. 252, 2022.
10. K. Kim and S. Park, "Optimization of photovoltaic output prediction using ensemble learning," Energies, vol. 14, no. 3, 2021.
11. N. Ozturk and E. Yetis, "Impact of temperature and humidity on PV panel efficiency," Solar Energy Materials and Solar Cells, vol. 174, pp. 119–127, 2018.
12. A. A. Abdullaev and B. S. Kholmatov, "Climate analysis of solar energy in the southern regions of Uzbekistan," UzEnergy Journal, vol. 4, no. 2, pp. 33–41, 2022.
13. Ministry of Energy of Uzbekistan, "Renewable Energy Development Strategy 2030," Tashkent, 2023.
14. Python Software Foundation, "Python Documentation – Pandas, Scikit-learn, Matplotlib," 2024.
15. OpenAI Research, "GPT-5 Technical Capabilities in AI-assisted scientific analysis," 2025.

