Short-Term Photovoltaic Power Forecasting: A Comparative Analysis of Machine Learning Models Across Multiple Time Intervals

Your Name

September 7, 2025

Abstract

This study presents a comprehensive evaluation of six machine learning models for short-term photovoltaic (PV) power forecasting at 5-minute, 10-minute, and 15-minute intervals. The models—XGBoost, Decision Tree, Linear Regression, Ridge Regression, Lasso Regression, and Elastic Net—are compared against a persistence baseline using a dataset from a PV site. We employ a robust preprocessing pipeline, including outlier handling, feature engineering, and hyperparameter tuning, to optimize model performance. Performance metrics (MAE, RMSE, nMAE, nRMSE, R^2, WAPE, and SMAPE) are analyzed across training, validation, testing, and seasonal datasets, with stress tests conducted for night-time and high-variability conditions. Results indicate that XGBoost and Decision Tree outperform other models, achieving high accuracy ($R^2 > 0.999$) and significant skill improvements over persistence (up to 97.28% for 15-minute forecasts). The study highlights the impact of temporal resolution on forecasting accuracy and computational efficiency, providing insights for practical PV forecasting applications.

Keywords: Photovoltaic forecasting, machine learning, short-term forecasting, time series, model comparison

1 Introduction

Photovoltaic (PV) power generation is increasingly vital in the global energy mix, driven by the need for sustainable energy sources. Accurate short-term forecasting of PV power output is critical for grid stability, energy management, and economic optimization. This study investigates the performance of six machine learning models—XGBoost, Decision Tree, Linear Regression, Ridge Regression, Lasso Regression, and Elastic Net—for forecasting PV power at 5-minute, 10-minute, and 15-minute intervals. These intervals are relevant for real-time grid operations and energy trading. The models are evaluated against a persistence baseline, a common benchmark in time-series forecasting.

The objectives are to:

- Develop a robust preprocessing and modeling pipeline for PV power forecasting.
- Compare model performance across multiple time intervals using comprehensive metrics.
- Assess model robustness under edge cases (night-time and high-variability periods).
- Analyze the impact of temporal resolution on forecasting accuracy and computational efficiency.

This paper contributes to the literature by providing a detailed comparison of model performance across different temporal resolutions, supported by extensive empirical analysis and visualizations.

2 Related Work

Short-term PV forecasting has been widely studied, with approaches ranging from statistical models to advanced machine learning techniques. **?** highlight the importance of data-driven models for capturing complex patterns in PV generation. Tree-based models like XGBoost have shown superior performance in time-series forecasting due to their ability to handle non-linear relationships [**?**]. Linear models, including Ridge and Lasso, remain popular for their interpretability and computational efficiency [**?**]. Recent studies, such as **?**, emphasize the need for robust preprocessing to handle outliers and missing data in PV datasets. However, few studies compare model performance across multiple short-term intervals (e.g., 5, 10, and 15 minutes) with a comprehensive set of metrics and edge-case analyses, which this study addresses.

3 Data and Methodology

3.1 Dataset Description

The dataset, sourced from a PV site, includes 14 features such as Active Power, Wind Speed, Global Horizontal Radiation, and Weather Temperature, recorded at 1-minute intervals. The data is resampled to 5-minute, 10-minute, and 15-minute frequencies, resulting in the dataset sizes shown in Table 1.

Table 1: Dataset Shapes Before and After Resampling

Dataset	Original Shape	5-Min Resampled	10-Min Resampled
Training	(675094, 14)	(675959, 14)	(337980, 14)
Validation	(144663, 14)	(150558, 14)	(75280, 14)
Testing	(144664, 14)	(150482, 14)	(75242, 14)

3.2 Preprocessing Pipeline

The preprocessing pipeline includes:

- **Outlier Handling:** Interquartile range (IQR) method to detect and impute outliers in Active Power.

- **Clipping and Imputation:** Clipping negative values for physical variables (e.g., Wind Speed) and forward-filling missing values.
- **Feature Engineering:** Addition of time-based features (hour, day of year, sine/cosine transformations) and lag/rolling features for Active Power (lags: 1, 12, 288; rolling windows: 12, 288).
- **Scaling:** MinMaxScaler applied to normalize features.

3.3 Models and Hyperparameter Tuning

Six models are evaluated: XGBoost, Decision Tree, Linear Regression, Ridge Regression, Lasso Regression, and Elastic Net. Hyperparameter tuning is performed using RandomizedSearchCV with TimeSeriesSplit (5 folds). The best hyperparameters are summarized in Table 2. A persistence model (lag=288) serves as the baseline.

Table 2: Best Hyperparameters for Each Model (5-Min Interval)

Model	Best Hyperparameters
XGBoost	{subsample: 1.0, n_estimators: 100, max_depth: 5, learning_rate: 0.05, c
Decision Tree	{min_samples_split: 10, min_samples_leaf: 4, max_depth: 10}
Linear Regression	None
Ridge Regression	{alpha: 0.01}
Lasso Regression	{alpha: 0.01}
Elastic Net	{l1_ratio: 1.0, alpha: 0.01}

3.4 Evaluation Metrics

Models are evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), normalized MAE (nMAE), normalized RMSE (nRMSE), R^2, Weighted Absolute Percentage Error (WAPE), and Symmetric Mean Absolute Percentage Error (SMAPE). Skill vs. Persistence is calculated as:

$$\text{Skill}(\%) = 100 \times \frac{\text{RMSE}_{\text{persistence}} - \text{RMSE}_{\text{model}}}{\text{RMSE}_{\text{persistence}}}$$

4 Results

4.1 Model Performance Across Intervals

Table 3 summarizes the performance metrics for the 5-minute interval. XGBoost and Decision Tree consistently outperform other models, with R^2 values exceeding 0.999 across all datasets. The Decision Tree achieves the lowest MAE (0.5514) and RMSE (1.8035) on the test set, with a skill improvement of 95.9960% over persistence. Linear models (Linear, Ridge, Lasso, Elastic Net) show higher errors but are computationally efficient.

Table 3: Performance Metrics for 5-Minute Interval

Model	Dataset	MAE	RMSE	nRMSE	R^2	WAPE (%)
XGBoost	Testing	0.8577	1.7976	0.0066	0.9996	1.2777
Decision Tree	Testing	0.5514	1.8035	0.0067	0.9996	0.8213
Linear Regression	Testing	1.0910	3.8998	0.0144	0.9981	1.6251
Persistence	Testing	17.1588	45.0436	0.1662	0.7489	25.5683

For 10-minute and 15-minute intervals, similar trends are observed (see Appendix). The Decision Tree maintains superior performance, with MAE values of 0.5332 (10-min) and 0.5429 (15-min) on the test set, and skill improvements up to 97.2845% (15-min).

4.2 Seasonal Analysis

Seasonal performance (Table 4) shows that models perform best in Summer (lowest RMSE: 1.0123 for Decision Tree, 5-min) and worst in Autumn (highest RMSE: 5.4889 for Linear Regression, 5-min). This reflects higher variability in Autumn due to weather fluctuations.

4.3 Stress Test Results

Stress tests (Table 5) reveal that the Decision Tree excels in night-time conditions (MAE: 0.1204, 5-min) and high-variability periods (MAE: 0.8908, 15-min), demon-

Table 4: Seasonal Performance Metrics for 5-Minute Interval

Model	Season	MAE	RMSE	nRMSE	R^2	WAPE (%)
XGBoost	Summer	0.6858	1.2957	0.0048	0.9998	1.2209
Decision Tree	Summer	0.3348	1.0123	0.0037	0.9999	0.5960
Persistence	Summer	10.4429	34.7442	0.1281	0.8394	18.5848

strating robustness across edge cases.

Table 5: Stress Test Performance Metrics

Model	Scenario	Interval	MAE	RMSE	nRMSE	WAPE (%)
Decision Tree	Night-Time	5 min	0.1204	0.3143	0.0312	32.2315
Decision Tree	High Variability	15 min	0.8908	1.8940	0.0069	0.7788

4.4 Computational Efficiency

Table 6 shows that Linear Regression is the most computationally efficient (tuning time: 0.52s, 5-min), while XGBoost and Decision Tree require longer tuning times (231.09s and 206.78s, 5-min). However, their superior accuracy justifies the computational cost for critical applications.

Table 6: Computational Timing Summary

Model	Interval	Tuning Time (s)	Training Time (s)
XGBoost	5 min	231.09	11.95
Decision Tree	5 min	206.78	14.77
Linear Regression	5 min	0.52	0.63

5 Discussion

The results demonstrate that XGBoost and Decision Tree are the most effective models for short-term PV forecasting, with Decision Tree slightly outperforming XGBoost in MAE and RMSE across all intervals. The high R^2 values (>0.999) indicate excellent fit, particularly for 15-minute forecasts, where temporal aggregation reduces noise. Linear models, while less accurate, offer computational efficiency suitable for resource-constrained environments. Seasonal variations highlight the challenge of forecasting in Autumn, likely due to unpredictable weather patterns. Stress tests confirm the robustness of tree-based models in handling edge cases, making them suitable for real-world applications. The consistent feature selection across intervals (e.g., Wind Speed, Pyranometer$_{1})underscorestheimpor$

6 Conclusion

This study provides a comprehensive comparison of six machine learning models for short-term PV power forecasting. The Decision Tree model achieves the best performance across 5-minute, 10-minute, and 15-minute intervals, with significant improvements over the persistence baseline (up to 97.28% skill). Future work could explore hybrid models or deep learning approaches to further enhance accuracy, particularly in challenging seasons like Autumn.

Appendix

A Walk-Forward Validation Metrics

Table 7: Walk-Forward Validation Metrics

Model	Interval	RMSE	MAE	R^2	nMAE	WAPE (%)
XGBoost	5 min	1.2895	0.7252	0.9997	0.0191	1.9102
Decision Tree	5 min	1.1984	0.4417	0.9998	0.0113	1.1286