eNergetics 2025
DOI: https://doi.org/10.

Original Scientific Article

Intelligent Hybrid Forecasting of Photovoltaic Power Using LSTM and XGBoost: Seasonal and Comparative Insights

Musa Terkeş[1], Alpaslan Demirci[2], Erdin Gökalp[3]

[1]Department of Electrical Engineering, Yildiz Technical University, İstanbul, Türkiye, https://orcid.org/0000-0002-4411-411X

[2]Department of Electrical Engineering, Yildiz Technical University, İstanbul, Türkiye, https://orcid.org/0000-0002-1038-7224

[3] Department of Electrical Engineering, Yildiz Technical University, İstanbul, Türkiye, https://orcid.org/0000-0002-1518-432X

[1]musa.terkes@yildiz.edu.tr, [2]ademirci@yildiz.edu.tr, [3]gokalp@yildiz.edu.tr

Abstract— **Accurate forecasting of photovoltaic (PV) power generation is essential for the reliability and cost-effectiveness of renewable-based power systems to be maintained. This study provides a comprehensive comparison of several statistical and machine learning approaches, including SARIMA, SVR, k-NN, RF, XGBoost, and LSTM networks. To enhance the consistency of predictions, a hybrid ensemble method is proposed, which integrates LSTM and XGBoost through a weighted averaging approach. The goal of this configuration is to take advantage of the temporal learning ability of LSTM and the nonlinear feature modeling strength of XGBoost. The hourly seasonal datasets were evaluated using common performance indicators, including RMSE, MAE, MAPE, and R^2. The analysis shows that each individual model performs better under certain seasonal or meteorological conditions. The LSTM–XGBoost hybrid generally yields the lowest prediction errors. It has particular effectiveness in the capture of both short-term variations and broader seasonal patterns. The results highlight the importance of hybrid intelligent systems in enhancing PV forecasting accuracy and show their potential to support more stable renewable energy operations.**

Keywords - photovoltaic power forecasting, hybrid ensemble learning, LSTM neural networks, XGBoost regression, time-series prediction.

I. INTRODUCTION

Photovoltaic (PV) energy has become a central element in the global shift away from fossil fuels, which is a major development in the energy sector. Its abundance, environmental benefits, and steadily falling costs have made it one of the most attractive renewable sources available [1], [2]. The inclusion of PV systems in modern power networks is now considered essential. This is because it helps reach sustainability goals. It also helps meet international commitments on climate change mitigation [3]. However, PV generation remains highly dependent on local weather conditions. The balance between electricity production and consumption is complicated by this variability, and grid reliability is challenged [4], [5]. Consequently, producing accurate short- and medium-term forecasts is now critical for day-to-day operations, such as reserve planning, load management, energy market participation, and system stability.

ISBN: xxx-xx-xxxxx-xx-x

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission.

Conventional statistical models often struggle to represent the nonlinear and intermittent behavior of PV output. One example is the seasonal autoregressive integrated moving-average (SARIMA) method. These limitations have encouraged the growing use of artificial-intelligence-based techniques [6]. In recent years, deep learning architectures, including convolutional neural networks (CNNs), long short-term memory (LSTMs), gated recurrent units (GRUs), and transformer models, have shown strong potential for capturing the intricate temporal and spatial structures present in PV data [7]. Spatial features are particularly well learned from irradiance or cloud-pattern information by CNNs, while sequential trends are excellently identified by recurrent models such as LSTM and GRU. More recently, research has explored ways to improve the representation of long-range temporal dependencies [8]. One approach is to use attention mechanisms in Transformer frameworks and auxiliary components like Time2Vec.

In parallel with the development of forecasting models, researchers have introduced several complementary strategies to enhance both accuracy and robustness. Signal decomposition methods separate the time series into multiple frequency components. These methods include ensemble empirical mode decomposition (EEMD), its adaptive variant complete ensemble empirical mode decomposition with Adaptive Noise (CEEMDAN), and variational mode decomposition (VMD). These methods stabilize the predictions [9], [10]. Likewise, dimensionality-reduction and feature-selection techniques have been applied to simplify input variables and make the models more interpretable. These techniques include principal component analysis (PCA), least absolute shrinkage and selection operator (LASSO), and SHapley Additive exPlanations (SHAP) [11]. Hyperparameter-optimization approaches based on metaheuristics such as Bayesian optimization, grey-wolf optimizer (GWO), genetic algorithm (GA), multi-verse optimizer (MVO), and improved salp-swarm algorithm (ISSA) have also improved forecasting performance. In addition, clustering methods for identifying similar days have helped to capture temporal variability. These methods include K-means++ and K-medoids. They also include synthetic-data generation using generative adversarial networks (GANs) or their time-series extension (TimeGAN). This has helped to mitigate the lack of training data. Recent studies have shown that image-based CNN models trained on satellite images can sometimes do a better job than traditional meteorological inputs in very short-term forecasting tasks.

Although prediction quality has been substantially improved by these advances, several issues remain unresolved. Many studies use data from only one place and only for less than a year. This makes the models not useful for other places and times of the year [1], [2], [4], [12]. Some approaches fail to consider crucial atmospheric factors, like aerosol content, or fixate exclusively on point forecasts without accounting for uncertainty [13]. In other cases, models lack explainability mechanisms, resulting in "black-box" models that impede transparency and adoption [14], [15]. Finally, these systems can't be used in real-time or large-scale grid operations because they need too many computers [6], [7], [9], [12], [16].

Recent studies have made an attempt to address the above challenges through the combination of different model types within hybrid or ensemble frameworks. For instance, CNN-LSTM-Transformer configurations have been developed to integrate spatial, temporal, and attention-based learning to enhance overall accuracy. In a similar way, LSTM networks that have been tuned with Bayesian optimization adapt their performance to different input conditions [9]. Some researchers have also paired GAN-generated synthetic data with CNN–GRU architectures to compensate for limited training samples [8]. Concurrently, tree-based algorithms such as extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and categorical boosting (CatBoost) have been integrated into hybrid designs to capture complex nonlinear feature relations [15], [17]. Other techniques have also shown measurable improvements in prediction quality. These include attention-guided feature selection [18] and ensembles driven by sky-image CNN models.

In general, the existing work shows that one type of model usually doesn't work well in all situations. Instead, integrated and multi-layered approaches seem to offer a more balanced solution. These approaches combine deep learning with boosting methods, attention modules, uncertainty modeling, and explainable-AI components. Based on these findings, the

current study presents a hybrid framework that links an attention-based LSTM with tree regressors like XGBoost. The interpretability of the model was addressed by applying SHAP analysis to the XGBoost component of the ensemble. The relative importance of meteorological and solar features was quantified in this analysis, and insights into the model's decision process were provided. The model's transparency is enhanced and its deployment in operational energy management systems is supported by such interpretability. In this way, the goal is to produce forecasting models that are not only accurate and robust. The models should also be easier to interpret and applicable to real-world PV integration and grid operation.

II. Dataset and Modeling Framework

This research was carried out for a hydrogen refueling station in Antalya, Türkiye (36.90° N, 30.70° E). This location is known for its strong solar resource. The site receives an annual average direct normal irradiance (DNI) of about 5.23 kWh/m² per day. This varies seasonally between 3.30 and 7.48 kWh/m² per day. The mean clarity index is estimated to be around 54.6%. The PVGIS database was used to obtain estimates of PV generation. This database relies on the SARAH and ERA5 reanalysis datasets to model solar irradiance based on the geographic coordinates of the station.

The creation of hourly load-demand profiles in HOMER used a combination of stochastic and pattern-based approaches to complement the generation data. The methodological framework is unified and combines statistical, machine-learning, and deep-learning methods. This framework is designed for seasonal PV-power forecasting. Forecasting PV output is inherently difficult because deterministic solar cycles interact with random meteorological changes such as cloud movement and temperature variation. To represent both periodic and irregular behaviors, the problem is expressed as learning a mapping function $F = \mathbf{X}_t \rightarrow y_{t+h}$, $\mathbf{X}_t = \left\{ y_{t-1}, y_{t-2}, \cdots, y_{t-p} \right\}$ represents the lagged historical production values, and y_{t+h} denotes the forecasted PV output at horizon h.

Instead of using each algorithm separately, the models are used together. Each model has different strengths that help cover the weaknesses of the other models. The statistical layer of the framework is employed by the SARIMA model, which decomposes the time series into autoregressive, differencing, and moving-average elements combined with seasonal terms. Its general form can be written as follows:

$$\Phi_p(B)\Phi_P\left(B^s\right)(1-B)^d\left(1-B^s\right)^D y_t = \Theta_q(B)\Theta_Q\left(B^s\right)\varepsilon_t \quad (1)$$

Here, B represents the backshift operator, where $By_t = y_{t-1}$. The parameter denotes the seasonal period, and ε_t stands for the white-noise error term. In this formulation, y_t is the observed PV generation at time step t; Φ_p and Φ_P are the autoregressive (AR) polynomials for the non-seasonal and seasonal parts, respectively; $\Theta_q(B)$ and $\Theta_Q\left(B^s\right)$ are the corresponding moving-average (MA) polynomials. The differencing orders d and D describe the non-seasonal and seasonal components. The non-seasonal AR polynomial of order p can be expressed as:

$$\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (2)$$

where ϕ_i are AR coefficients. Seasonal AR polynomial of order P with seasonal lag s:

$$\Phi_p\left(B^s\right) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps} \quad (3)$$

The SARIMA model is especially helpful for recognizing consistent and recurring seasonal patterns. However, its strict structure limits its flexibility when sudden or irregular fluctuations occur. These fluctuations often happen in PV generation affected by short-term weather variations.

To overcome these limitations, the analysis incorporates machine learning methods that can approximate nonlinear relationships. Among them, the k-nearest neighbors (k-NN) regressor estimates future values. It does so by comparing each observation with similar historical patterns in the dataset:

$$y_{t+h} = \frac{1}{K} \sum_{i \in \mathrm{N}_K(t)} y_i \quad (4)$$

Here, $N_K(t)$ denotes the neighborhood containing the K closest samples to the target time step t. The k-NN approach is appreciated for its conceptual simplicity and transparency, which make it a useful tool for data analysis. However, its performance often declines when the system faces unusual or rapidly changing conditions because it relies heavily on historical analogues. The random forest (RF) algorithm builds an ensemble of decision trees. These trees are generated through bootstrap resampling. RF can model nonlinear relationships and reduce variance by averaging the outcomes of multiple trees. A known drawback is that the averaging process smooths the results. Sometimes it underestimates extreme values. Support vector regression (SVR) adopts a more formal approach by conceptualizing the regression task within the framework of structural risk minimization. The goal is to find a function that matches the data while making the prediction and the model less complicated. This is expressed as:

$$\min_{w,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\left(\xi_i + \xi_i^*\right) \tag{5}$$

subject to the constraint $\left|y_i - \left(w^T\phi(x_i)+b\right)\right| \le \varepsilon + \xi_i$, where $\phi(\bullet)$ represents the kernel mapping function. By introducing an ε-insensitive margin, SVR ignores small deviations while penalizing larger errors, which helps the model remain robust in the presence of noise. Here, w denotes the weight vector, and ξ_i and ξ_i^* are positive slack variables that capture prediction errors lying outside the ε-insensitive region.

Among ensemble algorithms, XGBoost applies the principle of gradient boosting. It does so to iteratively reduce residual errors. It builds an additive model that refines its predictions at each step, as expressed in the following equation:

$$y_t = \sum_{k=1}^{K} f_k(x_t),\ f_k \in F \tag{6}$$

where F represents the space of regression trees, f_k denotes the k-th regression tree within the XGBoost framework, and x_t refers to the input vector fed into the LSTM model at time step t. The tree ensemble can capture overall patterns. This ability enables it to identify broad production trends effectively. However, because of its static and rule-based nature, it may sometimes stretch production intervals. It may also lose detail in rapid temporal changes.

To represent sequential relationships and short-term variability, a LSTM-recurrent neural network (RNN) is integrated into the framework. The LSTM design is different from traditional RNNs because it addresses the vanishing-gradient limitation. It does this by using gated memory units that can retain or discard information. The following equations show this:

$$f_t = \sigma\left(W_f\left[h_{t-1}\ ,\ x_t\right] + b_f\right) \tag{7}$$

$$i_t = \sigma\left(W_i\left[h_{t-1}\ ,\ x_t\right] + b_i\right) \tag{8}$$

$$\tilde{c}_t = \tanh\left(W_c\left[h_{t-1}\ ,\ x_t\right] + b_c\right) \tag{9}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{10}$$

$$o_t = \sigma\left(W_o\left[h_{t-1}\ ,\ x_t\right] + b_o\right) \tag{11}$$

$$h_t = o_t \odot \tanh\left(c_t\right) \tag{12}$$

where x_t denotes the input vector, h_t the hidden state, and c_t the cell state of the LSTM unit. The terms f_t, i_t, and o_t represent the forget, input, and output gates, respectively. The candidate cell state is noted as $\tilde{c}_t$, while W_f, W_i, W_c, and W_o denote the weight matrices corresponding to each gate. Similarly, b_f, b_i, b_c, and b_o are the bias terms, and $\sigma(\bullet)$ indicates the sigmoid activation function. LSTM uses a gating structure to retain useful temporal information and filter out redundant signals. This enables the model's reproduction of both the smooth daily production profile and transient irregularities, such as rapid fluctuations caused by shading or load variation.

The primary contribution of this study is the creation of a hybrid structure that merges the sequence learning power of LSTM with the trend generalization ability of XGBoost. The combined prediction is expressed as a convex integration of the two model outputs:

$$y_t^{Hybrid} = \alpha y_t^{LSTM} + (1-\alpha) y_t^{XGBoost} \quad (13)$$

where the weighting factor $\alpha \in [0,1]$ is tuned to minimize the overall loss across seasonal datasets. The α coefficient controls the balance between the LSTM and XGBoost forecasts within the hybrid model. The strengths of both methods are drawn on by this structure: short-term nonlinear variations within a day are captured by the LSTM, while broader, long-term patterns are efficiently modeled by XGBoost. In combination, the slow response of the LSTM is reduced and the tendency of XGBoost to oversimplify temporal details is mitigated.

This framework is based on several assumptions. PV production is treated as a function of solar irradiance and related weather conditions, and it is assumed that enough information to estimate future outputs can be carried by the historical production record. Before training, all data were standardized to a unit scale. This ensured consistency between models and avoided bias caused by variable magnitudes. Hyperparameters for the model were tuned using a combination of grid search and validation-based tuning. The XGBoost component was identified using a grid search with five-fold time-series cross-validation (TimeSeriesSplit). This method was used to determine the optimal tree depth, learning rate, and subsampling ratios. For the LSTM network, hyperparameters were tuned empirically. These included the number of hidden units, dropout rates, and learning rate. Tuning was based on validation loss under early stopping criteria. To make sure that the optimization was consistent across models, all of the other regressors were calibrated through grid-based or heuristic parameter searches. The model can be viewed from another perspective, which is input-specific. Representative days were selected from each season to reflect differences in sunlight duration, production symmetry, and random atmospheric effects.

To measure model accuracy, common error metrics were used, including root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination (R^2). These metrics evaluate both absolute prediction accuracy and the ability to reproduce time-varying dynamics. Despite the absence of formal statistical significance tests, the performance disparities were uniformly discerned across numerous validation windows and metrics (RMSE, MAE, and R^2). This consistency indicates that the hybrid ensemble's enhancement over the baseline models is not just a coincidence but rather a result of its robust performance.

This study makes a comparison between classical and modern forecasting techniques by combining different modeling perspectives in a single mathematical form. The proposed LSTM–XGBoost hybrid is a flexible and reliable approach for seasonal PV forecasting, addressing both predictable solar pattern and random weather-driven variability.

Finally, a fair assessment of model performance requires multiple complementary metrics, each emphasizing a distinct aspect of predictive behavior. To ensure consistent comparisons across models and seasons, this work utilizes several standard indicators. The measure calculates the average size of prediction deviations, no matter what direction is. It is defined as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|y_t - y_t\right| \quad (14)$$

where N represents the total number of test samples, and y_t denotes the PV generation value predicted by the model at time step t. To emphasize larger prediction errors, the RMSE was adopted as a complementary accuracy measure, defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_t - y_t\right)^2} \quad (15)$$

The mean absolute percentage error (MAPE) is a metric that allows for the comparison of forecasting accuracy across datasets of different scales. The calculation is as follows:

$$MAPE = \frac{100}{N}\sum_{t=1}^{N}\left|\frac{y_t - y_t}{y_t + \varepsilon}\right|, \ \varepsilon \ll 1 \quad (16)$$

A small constant ε is added to prevent division by zero during periods of negligible PV generation. The coefficient of determination, or R^2, is a statistical measure that quantifies the proportion of variation in the observed data that is attributed to the model. It is calculated as follows:

$$R^2 = 1 - \frac{\sum_{t=1}^{N}\left(y_t - y_t\right)^2}{\sum_{t=1}^{N}\left(y_t - \bar{y}\right)^2} \qquad (17)$$

where $\bar{y}$ denotes the mean of the observed values.

To avoid bias from periods with nearly zero production, all performance metrics were calculated on test samples filtered by solar irradiance. Representative days from winter, spring, summer, and autumn were used to ensure balanced seasonal coverage during the evaluation. The final step was to compare the hybrid model's results with those of the individual models. This was done to assess the effect of combining methods and to demonstrate the practical benefit of the hybrid approach in PV forecasting.

III. Seasonal Forecasting Performance and Comparative Performance of AI Models

Fig. 1 compares the PV power output predicted by the AI models across several days selected from each season.

One of these examples, from the middle of January, shows the winter period when the sun does not make much energy. During this season, the daylight window is short and the sun doesn't provide much energy, which limits the amount of electricity produced by solar panels to a narrow time period during the day. The actual production starts around 9:00 a.m. and decreases until around 3:00 p.m. Among the models, the LSTM forecast matched the real production pattern quite well. This was particularly true around the midday peak. The LSTM forecast captured both the rise and fall of the output curve with reasonable accuracy.

In contrast to the LSTM outcomes, the XGBoost model generally initiated the production phase slightly earlier in the morning and prolonged it later into the afternoon, resulting in a noticeably broader curve. This tendency likely originates from the way tree-based algorithms generalize across similar data points. The midday peak was reduced by the gentler slopes of the smoother curves produced by the SARIMA and k-NN approaches. SARIMA's tendency to predict production too soon and underestimate the actual midday level stems from its assumption of a repeating seasonal pattern. In contrast, the hybrid LSTM + XGBoost configuration leveraged the strengths of both methods, accurately capturing both the timing and the magnitude of the peak.

Compared to the example from early March, which represented spring conditions, the current example displays much more irregular behavior. During that day, the weather shifted, leading to sudden drops in solar power generation, making it challenging to predict future energy output.

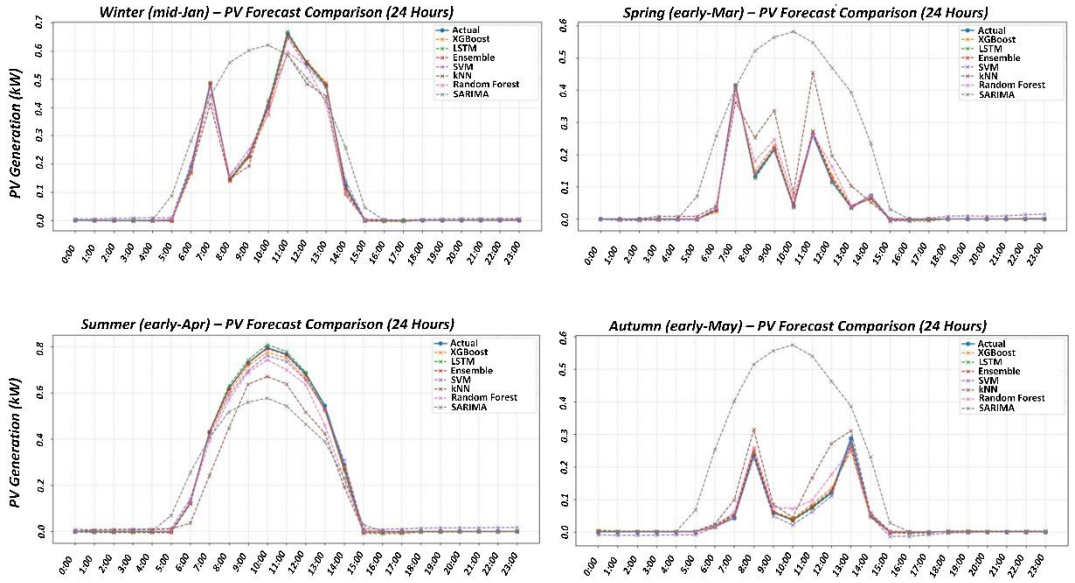


Figure 1. Predicted and actual PV power by season

TABLE I. SEASONAL PERFORMANCE COMPARISON OF FORECASTING MODELS FOR PV POWER PREDICTION

Season	Metrics	AI Model						
		SARIMA	**k-NN**	**RF**	**SVR**	**XGBoost**	**LSTM**	**Ensemble**
Winter (mid-Jan)	RMSE	0.1308	0.0294	0.0229	0.0085	0.0095	0.0089	0.0059
	MAE	0.0667	0.0152	0.0119	0.0075	0.005	0.0046	0.0033
	R^2	0.6106	0.9804	0.988	0.9983	0.9979	0.9982	0.9992
Spring (early-Mar)	RMSE	0.2027	0.0586	0.017	0.007	0.0068	0.0053	0.0054
	MAE	0.1157	0.0306	0.0081	0.0058	0.0037	0.0037	0.0036
	R^2	-	0.6778	0.9727	0.9954	0.9956	0.9974	0.9972
Summer (early-Apr)	RMSE	0.0993	0.0831	0.033	0.0184	0.0097	0.0067	0.0073
	MAE	0.0583	0.0508	0.0186	0.0165	0.0053	0.0041	0.0045
	R^2	0.8903	0.9233	0.9879	0.9962	0.999	0.9995	0.9994
Autumn (early-May)	RMSE	0.2216	0.0416	0.0172	0.0088	0.0099	0.0074	0.0068
	MAE	0.1287	0.0188	0.0088	0.0074	0.0046	0.0057	0.0044
	R^2	-	0.6858	0.9464	0.986	0.9823	0.99	0.9915

The k-NN and SVM models depend heavily on historical analogues. They could not keep pace with these rapid shifts. The models produced overly smooth estimates. The LSTM network was able to detect some of the sudden changes, but it was slow react. XGBoost, on the other hand, tended to overshoot morning production and occasionally provided non-zero predictions during periods of no sunlight. Even though there were problems, the hybrid model still made most of the short production drops with small mistakes, especially between 11:00 and 14:00.

The profile of production in the early April example, which represents summer, is more symmetrical, with higher values as daylight hours increase. During this period, production begins around 6:00–7:00. It ends around 14:00–15:00. The production during the afternoon reaches a peak with a distinct rise and fall that resembles a classic PV production curve. In this case, the LSTM model performed well in predicting the timing and magnitude of the production curve. It accurately captured the morning rise and the afternoon peak.

The overall production trend was tracked by the XGBoost model, but relatively aggressive predictions were made in the morning, and a slight delayed decline was shown in the afternoon. SARIMA was used to predict the early morning production, but the production values were suppressed, resulting in a flatter curve that did not represent the peak well. The k-NN and RF models, which rely on averages from historical data, smoothed out production increases and decreases, particularly modeling the afternoon decline in a delayed and flattened manner. This is a common result for models that are not very sensitive to sudden curve changes. Nevertheless, the LSTM+XGBoost hybrid AI model generated highly accurate predictions. These predictions were made across the entire time range. The time range included the morning start, noon peak, and the end of production. The hybrid model achieved balanced and stable prediction performance for both short-term fluctuations and overall production patterns.

The early May example representing autumn reveals a more complex production profile: production starts late in the morning, suddenly peaks, and then declines. Deep learning methods are at an advantage when using this structure in comparison to classical statistical models. SARIMA and RF models were unable to capture this sudden change. Instead, they provided an "average day" prediction. The LSTM model accurately predicted the peak, but it did not adequately capture the subsequent decline rate. The rapid decline after production was exaggerated by XGBoost, resulting in negative prediction biases. These two tendencies were balanced by the hybrid model, and both the peak and subsequent decline were successfully modeled in terms of timing and magnitude. Overall, behind the hybrid model keeps relevant

information from earlier moments within its cell states, the model can follow short-term variations, such as the morning-to-evening transitions or brief production drops, that often appear in solar generation. This relates to how XGBoost interprets long-term statistical behavior. The gradient-boosting design of the algorithm helps it capture broader production trends that are learned from historical data, and it can adjust to hourly fluctuations. When the outputs of both models are combined through weighted averaging, the short-term and long-term errors compensate for one another. This produces forecasts that remain steady and trustworthy across the day (see Table I).

Overall, the hybrid ensemble exhibited both strong predictive capabilities and sufficient computational efficiency for practical applications. To ensure robust generalization, the hyperparameters of the XGBoost and LSTM components were optimized via grid search and time-series cross-validation. All the information that was put in was made the same using MinMax scaling, which made it so that things would happen more smoothly, and the numbers that were used for this were saved so that the same process could be used again. The enhanced PV power forecast accuracy further reinforces the model's practical relevance, extending beyond its statistical performance. This enhanced accuracy supports operational decision-making processes related to grid stability management and renewable integration.

IV. Conclusion and Future Suggestions

The objective of this work was the testing of various models for PV power prediction, including statistical, machine learning, and deep learning approaches. It also introduced a hybrid model that combines LSTM and XGBoost. The analysis showed that simple time-series models such as SARIMA perform well when seasonal patterns are strong. Algorithms such as SVR, k-NN, and RF demonstrate optimal performance in stable conditions; however, they exhibit a decline in accuracy when production patterns change rapidly. When combined into a weighted ensemble (LSTM and XGBoost), the two models complemented each other, resulting in improved stability and accuracy of predictions across different seasons.

The findings suggest that hybrid AI systems can contribute to enhancing the reliability of power grids and improve the planning of renewable energy. The results' practical implications underscore the potential for enhanced PV power forecasts to assist grid operators in preserving voltage stability and refining strategies for integrating renewable energy. More reliable scheduling of storage and dispatchable units is achieved by the reduced prediction error of the hybrid ensemble. The computational framework of LSTM-XGBoost exhibited effective convergence when employing the Adam optimizer and grid-based hyperparameter optimization for the constituent sub-models. Sensitivity analyses confirmed the stability of performance under different normalization and scaling methods, ensuring the approach's reproducibility and robustness.

Also in the future, researchers could also use methods like Bayesian inference or quantile regression to create confidence intervals for PV power forecasts. This would help make more reliable decisions about operations. Further research could explore attention-based or transformer models, and the inclusion of physics-based layers may improve the interpretability of the predictions. The models should be evaluated on datasets from multiple regions and climate zones so that their general applicability can be assessed. In general, the study shows that using a mix of different AI methods can make renewable energy forecasts more accurate and reliable.

References

[1] Li, P., Zhou, K., Lu, X., & Yang, S. (2020). A hybrid deep learning model for short-term PV power forecasting. Applied Energy, 259, 114216. https://doi.org/10.1016/j.apenergy.2019.114216

[2] Li, G., Xie, S., Wang, B., Xin, J., Li, Y., & Du, S. (2020). Photovoltaic Power Forecasting With a Hybrid Deep Learning Approach. IEEE Access, 8, 175871–175880. https://doi.org/10.1109/ACCESS.2020.3025860

[3] Akhter, M. N., Mekhilef, S., Mokhlis, H., Ali, R., Usama, M., Muhammad, M. A., & Khairuddin, A. S. M. (2022). A hybrid deep learning method for an hour ahead power output forecasting of three different photovoltaic systems. Applied Energy, 307, 118185. https://doi.org/10.1016/j.apenergy.2021.118185

[4] Xiang, X., Li, X., Zhang, Y., & Hu, J. (2024). A short-term forecasting method for photovoltaic power generation based on the TCN-ECANet-GRU hybrid model. Scientific Reports, 14(1), 6744. https://doi.org/10.1038/s41598-024-56751-6

[5] Zhang, J., Liu, Z., & Chen, T. (2023). Interval prediction of ultra-short-term photovoltaic power based on a hybrid model. Electric Power Systems Research, 216, 109035. https://doi.org/10.1016/j.epsr.2022.109035

[6] Zheng, J., Du, J., Wang, B., Klemeš, J. J., Liao, Q., & Liang, Y. (2023). A hybrid framework for forecasting

power generation of multiple renewable energy sources. Renewable and Sustainable Energy Reviews, 172, 113046. https://doi.org/10.1016/j.rser.2022.113046

[7] Agga, A., Abbou, A., Labbadi, M., Houm, Y. El, & Ou Ali, I. H. (2022). CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Electric Power Systems Research, 208, 107908. https://doi.org/10.1016/j.epsr.2022.107908

[8] Li, Q., Zhang, X., Ma, T., Liu, D., Wang, H., & Hu, W. (2022). A Multi-step ahead photovoltaic power forecasting model based on TimeGAN, Soft DTW-based K-medoids clustering, and a CNN-GRU hybrid neural network. Energy Reports, 8, 10346–10362. https://doi.org/10.1016/j.egyr.2022.08.180

[9] Wang, L., Mao, M., Xie, J., Liao, Z., Zhang, H., & Li, H. (2023). Accurate solar PV power prediction interval method based on frequency-domain decomposition and LSTM model. Energy, 262, 125592. https://doi.org/10.1016/j.energy.2022.125592

[10] Ge, L., Xian, Y., Yan, J., Wang, B., & Wang, Z. (2020). A Hybrid Model for Short-term PV Output Forecasting Based on PCA-GWO-GRNN. Journal of Modern Power Systems and Clean Energy, 8(6), 1268–1275. https://doi.org/10.35833/MPCE.2020.000004

[11] Talaat, M., Said, T., Essa, M. A., & Hatata, A. Y. (2022). Integrated MFFNN-MVO approach for PV solar power forecasting considering thermal effects and environmental conditions. International Journal of Electrical Power & Energy Systems, 135, 107570. https://doi.org/10.1016/j.ijepes.2021.107570

[12] Qu, Y., Xu, J., Sun, Y., & Liu, D. (2021). A temporal distributed hybrid deep learning model for day-ahead distributed PV power forecasting. Applied Energy, 304, 117704. https://doi.org/10.1016/j.apenergy.2021.117704

[13] Kumar, A., Rizwan, M., & Nangia, U. (2020). A Hybrid Intelligent Approach for Solar Photovoltaic Power Forecasting: Impact of Aerosol Data. Arabian Journal for Science and Engineering, 45(3), 1715–1732. https://doi.org/10.1007/s13369-019-04183-0

[14] Abou Houran, M., Salman Bukhari, S. M., Zafar, M. H., Mansoor, M., & Chen, W. (2023). COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications. Applied Energy, 349, 121638. https://doi.org/10.1016/j.apenergy.2023.121638

[15] Kim, H., Park, S., Park, H.-J., Son, H.-G., & Kim, S. (2023). Solar Radiation Forecasting Based on the Hybrid CNN-CatBoost Model. IEEE Access, 11, 13492–13500. https://doi.org/10.1109/ACCESS.2023.3243252

[16] Wang, T., Xu, Y., Qin, Y., Wang, X., Zheng, F., & Li, W. (2025). Short-term PV forecasting of multiple scenarios based on multi-dimensional clustering and hybrid transformer-BiLSTM with ECPO. Energy, 334, 137654. https://doi.org/10.1016/j.energy.2025.137654

[17] S. Ghimire, B. Bhandari, D. Casillas-Pérez, R. C. Deo, and S. Salcedo-Sanz, "Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in Queensland, Australia," Eng. Appl. Artif. Intell., vol. 112, p. 104860, 2022, doi: https://doi.org/10.1016/j.engappai.2022.104860.

[18] Mubarak, H., Hammoudeh, A., Ahmad, S., Abdellatif, A., Mekhilef, S., Mokhlis, H., & Dupont, S. (2023). A hybrid machine learning method with explicit time encoding for improved Malaysian photovoltaic power prediction. Journal of Cleaner Production, 382, 134979. https://doi.org/10.1016/j.jclepro.2022.134979