TEPES
TURKISH JOURNAL OF
ELECTRICAL POWER
AND ENERGY SYSTEMS

TEPES, Vol. 5, Issue. 3, 215-223, 2025 **DOI:** 10.5152/tepes.2025.25021

RESEARCH ARTICLE

A Hybrid Forecasting Approach for Solar Power Generation in Smart Grids Using Long Short-Term Memory and Autoregressive Integrated Moving Average

Radhakrishnan Kanthavel, Ramakrishnan Dhaya

Papua New Guinea University of Technology School of Electrical and Communications Engineering, Papua New Guinea University of Technology, Lae, Papua New Guinea

Cite this article as: R. Kanthavel and R. Dhaya, "A hybrid forecasting approach for solar power generation in smart grids using long short-term memory and autoregressive integrated moving average," *Turk J Electr Power Energy Syst.*, 2025; 5(3), 215-223.

ABSTRACT

An accurate solar energy forecast is important for the efficient operation of smart grids, especially with the increasing penetration of renewable energy sources. This paper proposes a hybrid forecasting approach that combines long short-term memory (LSTM) neural networks with autoregressive integrated moving average (ARIMA) models to improve the accuracy of short-term solar energy predictions. While ARIMA effectively captures linear temporary dependence, LSTM networks are powerful in nonlinear and long-distance pattern modeling. By integrating these two models, the proposed hybrid approach takes advantage of their complementary strengths to stop and address nonlinearity. The model is trained and tested on real-world solar power data collected from the grid-connected photovoltaic system. The evaluation metrics, such as mean absolute error, root mean squared error, and mean absolute percentage error, perform better than stand-alone ARIMA and LSTM models in the hybrid model, outpacing accuracy. Results outline the ability of hybrid intelligent models to increase the prediction of solar energy, contributing to more stable and reliable smart grid operations.

Index Terms—Autoregressive integrated moving average (ARIMA), hybrid models, long short-term memory, solar power forecasting, smart grids

I. INTRODUCTION

The increasing attention to eco-friendly energy sources has brought about an exponential increase in the installation of solar photovoltaic (PV) systems [1]. Solar energy, being a clean, inexhaustible source of power, has an intermittent nature of energy generation, which creates big challenges for power system operators, especially in smart grid environments where balancing demand and supply in real time is paramount [2]. There is a need for short-term solar power forecasting in order to ensure grid stability, better energy management, and a reduction in conventional backup generation [3]. However, these models may not be suitable for all types of data, especially when the data shows nonlinear patterns or seasonal variations. However, these methods seldom capture the random fluctuations with complex interactions typical of solar power generation data. Contrastingly, certain deep learning methods are known to do well with nonlinear time-dependent patterns; in particular, these fall into one category of recurrent networks known as the long short-term memory (LSTM) networks [4,5]. The LSTM also has limitations, so researchers have considered enhancement through hybridization with other techniques, for example, neural attention or convolutional neural networks, to improve both performance and interpretability.

The studies propose a hybrid forecasting technique that mixes the ARIMA and LSTM techniques to make use of the quality of both modeling paradigms: linear and nonlinear. The method starts with ARIMA modeling to capture linear developments of solar energy statistics, accompanied by the use of LSTM to address them. To address the limitations of stand-alone statistical or deep learning approaches, this study introduces a hybrid ARIMA-LSTM framework that leverages the strengths of both models for accurate and reliable solar power forecasting in smart grid environments.

There is a fair amount of literature that has considered solar forecasting using statistical or machine learning one-off approaches (including independent hybrid approaches). Still, there is limited

Corresponding author: Ramakrishnan Dhaya, dhayavel2005@gmail.com

Received: May 13, 2025
Revision requested: June 9, 2025
Last revision received: June 19, 2025
Accepted: July 7, 2025
Publication Date: September 3, 2025

consideration of ARIMA and LSTM in a framework of residual learning devised specifically for ease in real-time smart grid operation. While previous use of hybrid methods had a less clear separation of predictable, structured trends with linear and nonlinear components, a sequential decomposition approach of ARIMA for the structured linear trends is proposed, with the data subsequently passed to an LSTM for residual, nonlinear terms proposed, with the data. Consequently, this research is not only a step toward real, more accurate forecasts, but it is also a computationally efficient method that would enhance the implementation of hybrid methods to advance real-time operating grid forecasting, which has not been accomplished with previous hybrid approaches.

A. Research Gap

While hybrid forecasting methods do exist, the majority of approaches either integrate linear and nonlinear behaviors in a black-box model or do not scale to smart grid environments. In addition, rigorous evaluation of the residual learning framework in conjunction with strict decomposition has not been conducted on real solar PV data. This research fills these gaps by:

- Developing a structured hybrid ARIMA-E-LSTM framework whereby the LSTM is explicitly trained on the ARIMA residuals to enhance model interpretability and accuracy.
- Evaluating the hybrid ARIMA-LSTM framework and illustrating significant accuracy gains relative to other non-hybrid models, such as with mean absolute error (MAE) = 1.30 on real-world smart grid data.
- Additional practical value can be provided by going through the complexity-performance trade-offs associated with hybrid decomposition, which is important for use in smart grids in real-time.

II. METHODS

Recent research in solar energy forecasting has focused on integrating various modeling approaches to improve accuracy. Traditional statistical models such as Holt-Vinti and seasonal ARIMA (SARIMA) have been used to model seasonal solar power data, especially in areas with strong daily and seasonal radiation cycles. For example, Melit and Kalogirou [6] demonstrated the use of time-series models for solar radiation forecasting in the Mediterranean climate.

Main Points

- A hybrid model combining long short-term memory (LSTM) networks and autoregressive integrated moving average (ARIMA) is proposed for accurate solar power generation forecasting in smart grids.
- The approach captures both nonlinear temporal patterns and linear trends for improved predictive performance.
- The proposed hybrid method outperforms stand-alone LSTM and ARIMA models in terms of root mean squared error and mean absolute error metrics.
- The system provides better short-term forecasts, enabling more reliable smart grid energy management.
- The hybrid approach ensures adaptability to dynamic environmental conditions and improves forecast robustness.

Machine learning methods, such as support vector machines (SVMs) and random forests, have shown promise in improving the accuracy of solar energy forecasts. Khosravi et al. [7] PV output is detected with the use of dress learning to address uncertainty in prediction.

These models are particularly effective in identifying complex relations in historical data, but comprehensive feature engineering and tuning may be required. In addition, they are often limited by their inability to model sequential dependence effectively, which is important for time-dependent energy systems. For example, Zeng and Kiao [8] introduced a wavelet transform-based hybrid model, which combines SVMs with statistical methods to handle noise and non-stationarity in the solar dataset.

Hybrid models that combine multiple forecasting techniques have been proposed as a solution to these boundaries. However, these methods often struggle with sudden changes caused by weather variability, making them insufficient for reliable short-term predictions in the dynamic smart grid environment. Ahmed et al. [9] extended SARIMA approaches by incorporating meteorological variables, but results showed limited scalability to different geographical conditions. In this context, the deep belief network (DBNs) [10] and models promoting the shield [11] have also been investigated for their strength and generalization capabilities, which are accompanied by an increase in computational complexity. Hybrid models that combine multiple forecasting techniques have been proposed as a solution to these boundaries. For example, Zeng and Kiao [8] introduced a wavelet transform-based hybrid model, which combines SVMs with statistical methods to handle noise and non-stationarity in the solar dataset. Other researchers have integrated fuzzy logic, optimization algorithms, and neural networks to increase performance [12]. Recently, Das et al. [13] despite these efforts, in this study, proposed stronger hybrid architecture, inspiring the development of the same stronger hybrid architecture, is basically an interval in the integrated model, which efficiently handles both linear and nonlinear patterns. Recent improvements in renewable electricity forecasting have more and more followed hybrid and deep mastering processes to cope with limitations in accuracy and statistical complexity. Khan et al. [14] proved the effectiveness of deep LSTM networks mixed with data preprocessing strategies for solar strength prediction, achieving improved accuracy on real international datasets. Zhang et al. [15] proposed a hybrid ARIMA-LSTM model particularly for wind speed forecasting, showing that combining statistical and neural network-based models improves generalization.

Ahmad et al. [16] delivered a CNN–BiLSTM (Convolutional Neural Networks-Bidirectional Long Short-Term Memory) framework for PV power forecasting, where convolutional layers had been used to extract temporal-spatial capabilities before sequential prediction, resulting in huge overall performance gains. In any other method, Shao et al. [17] applied a wavelet-transformed input shape mixed with gated recurrent gadgets (GRUs) for day-in-advance solar forecasting, achieving higher adaptability to non-stationary time-series information. Mahapatra et al. [18] furnished a complete evaluation of hybrid deep neural networks, highlighting their applicability in dynamic power demand environments. Additionally, Fang et al. [19] integrated interest mechanisms with LSTM models and meteorological inputs for solar energy prediction, accomplishing superior interpretability and accuracy over traditional deep learning techniques. This current research collectively helps use hybridized deep learning models for more accurate and adaptive power forecasting in smart grids.

Over the last few years, advanced hybrid approaches have also explored new combinations of deep learning, solutions using statistics, and attention-based approaches that benefit solar forecasting. Sahoo et al. [20] developed an LSTM-CNN hybrid, which was optimized for microgrid optimization in solar energy cost predictions, that takes advantage of local spatial-temporal characteristics associated with seasonal influences within the data for short-term forecasting applications. Wang et al. [21] also proposed a transformer prediction framework that included all available meteorological factors and showed success in predicting solar activity data with varying degrees of changes in weather conditions, with high levels of error variance.

Additionally, Chen et al. [22] presented and tested a GRU model that was enhanced by an attention mechanism that responsively accounted for the abundance and loss of incoming solar irradiance and demonstrated generalized high performance across many regions. Ali et al. [23] wanted to contrast black-box neural-based hybrids as compared to residual decomposition-based models. When using structured residual learning approaches—similar to the one explored here—they indicated that the residual components offered advantages for interpretability and error localization. Kaur et al. [24] presented the deployment of deep hybrid solar forecasting models for utility-grade smart grids, with recommendations regarding low latency and modular development. Just the year before, Hossain et al. [25] introduced the adaptive LSTM with attention and transfer learning for high-resolution forecasting to degrees of high complexity and a loss of modularity.

Nevertheless, while these advancements are evident, the vast majority of current models employ either black-box types or over-focus on accuracy, potentially limiting real-time use applications. This study proposes the use of structured residual modeling, where ARIMA provides a deterministic trend model and the LSTM accounts for the nonlinear distilling of the residuals. This new decomposition better handles accuracy, interpretability, and computational feasibility for real-world grid forecasting applications.

For wind speed prediction, Zhang et al. [15] proposed a hybrid model of both ARIMA and LSTM using raw data with a joint training approach. A key difference with the method is the explicit use of residual decomposition, which aids in interpretability and likelihood of performance (particularly in the non-stationary conditions of solar power). While the external features used by Zhang et al. were limited, the proposed model utilized both meteorological variables and time features, which would demonstrate both better generalization and real-time adaptability of the model for solar forecasting.

While previous studies have referenced hybrid models, there are few examples of an architecture that is residual-driven, where ARIMA is first used to capture deterministic trends as the first model, and where the LSTM is explicitly trained on the residuals from the prediction of the ARIMA model. This type of layering is advantageous because it provides interpretability as ARIMA and the LSTM are alternating, but also restricts overfitting and obtains robustness through volatile solar conditions. It is unlike the general black-box hybrids reported in the literature.

III. RESULTS

The proposed hybrid forecasting model combines the complementary strengths of ARIMA and LSTM neural networks to achieve more accurate solar power predictions. The ARIMA excels at capturing linear trends and seasonal patterns in time-series data, while LSTM is designed to model complex nonlinear relationships and long-term dependencies. This hybrid approach offers enhanced prediction performance by addressing both linear and nonlinear components of solar power generation.

A. Data Preprocessing

The forecasting framework begins with the collection and preprocessing of historical solar power generation data, which includes normalization, handling missing values, and temporal feature engineering (e.g., extracting hour-of-day, day-of-week, and solar elevation angle). The dataset is then split into training and testing sets in a time-consistent manner to preserve sequence integrity.

B. Autoregressive Integrated Moving Average Component

In the first stage, an ARIMA model is trained on the normalized solar power data. The model parameters (p, d, q) are selected based on the autocorrelation function (ACF), Partial ACF, and Akaike information criterion (AIC). The ARIMA model captures the linear component of the solar power output. The proposed hybrid forecast models are designed to leverage the complementary strengths of ARIMA and LSTM neural networks for more accurate solar energy prediction. The ARIMA is an expert in capturing linear trends and seasonal patterns in time-series data, while LSTM is capable of modeling complex nonlinear relationships and long-term dependence. The model follows a two-step architecture, as illustrated in Fig. 1.

C. Residual Extraction

Once the ARIMA forecast arises, the residual chain is calculated by reducing the ARIMA output from the real values. These remain have nonlinear components that ARIMA cannot capture.

D. Long Short-Term Memory Modeling

The residues are used to train the LSTM network, which learns the underlying nonlinear dependence. The LSTM architecture consists of an input layer and one or more hidden LSTM layers, with a dropout regularization to avoid overfitting and a dense output layer. The model is trained using a sliding window approach with a suitable sequence length to catch cosmic dependence. Once the ARIMA model is applied and its forecast $\hat{y}_t^{\text{ARIMA}}$ is obtained, the residual series e_t is computed as in (1):

$$e_t = y_t - \hat{y}_t^{\text{ARIMA}} \tag{1}$$

Here, y_t is the actual solar power at time t, and $\hat{y}_t^{\text{ARIMA}}$ is the ARIMA prediction at time t. The LSTM model is trained on the residual sequence $\{e_{t-1}, e_{t-2}, \ldots, e_{t-n}\}$, using a sliding window approach, to predict the nonlinear component of the forecast $\hat{e}_t^{\text{LSTM}}$ (2):

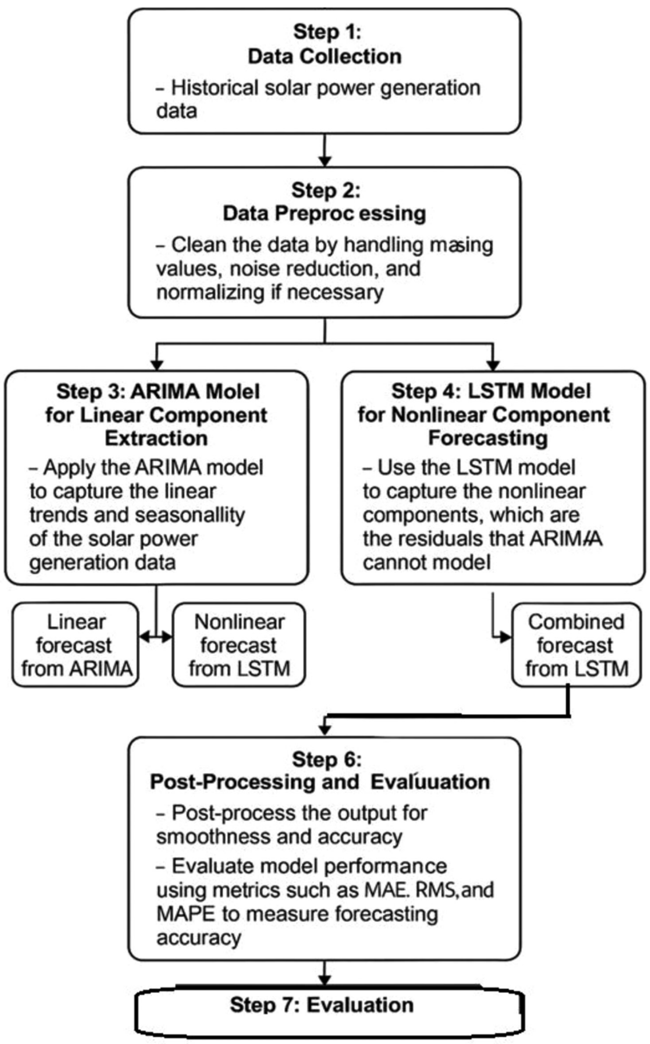

Fig. 1. Hybrid ARIMA-LSTM model architecture.

$$\hat{e}_t^{\text{LSTM}} = \text{LSTM}\left(e_{t-1}, e_{t-2}, \ldots, e_{t-n}\right) \tag{2}$$

The final hybrid forecast $\hat{y}_t$ is obtained by combining the ARIMA forecast and the LSTM predicted residual in (3):

$$\hat{y}_t = \hat{y}_t^{\text{ARIMA}} + \hat{e}_t^{\text{LSTM}} \tag{3}$$

This architecture allows the hybrid model to capture both the linear trend (via ARIMA) and nonlinear variations (via LSTM) in the solar power time series.

E. Hybrid Forecast Output

The final forecast ARIMA prediction and LSTM-pre-residue are briefly obtained. This hybrid approach allows the model to reorganize both linear and nonlinear components present in the original solar power data.

F. Evaluation Matrix

The model's performance is evaluated using the standard error matrix, which means MAE, root mean squared error (RMSE), and absolute mean absolute percentage error (MAPE). These matrices are used to compare hybrid models against Stand-alone ARIMA and LSTM models, which demonstrate the effectiveness of the hybrid approach in improving forecast accuracy.

Let:

- y_i: Actual solar power at the time t
- $\hat{y}_i^{\text{ARIMA}}$ Forecast from the ARIMA model at time t
- $\in_t = \psi_t - \hat{y}_t^{\text{ARIMA}}$ Forecast of residual from LSTM
- Residual (nonlinear component)
- e_t^{LSTM} Forecast of residual from LSTM
- $\hat{y}_t = \hat{y}_t^{\text{ARIMA}} + \hat{e}_t^{\text{LSTM}}$ Final forecasted value at time t

Final hybrid forecast

Step 1: ARIMA forecasting

The ARIMA (p, d, q) model predicts the linear component using the following equation (4):

$$\psi'_t = c + \sum_{i=1}^{p} \phi_i \xi'_{t-i} + \sum_{j-1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{4}$$

Where:

- y: Differenced time series (after d differencing steps)
- ϕ_i: Autoregressive coefficients
- ϑ_j: Moving average coefficients
- ε_t: White noise error
- c. Constant term

Step 2: Residual calculation

The residual (nonlinear part) is obtained as in (5):

$$e_t = y_t - \hat{b}_t^{\text{ARDIA}} \tag{5}$$

Step 3: LSTM-based residual forecasting

The LSTM model learns from previous residuals and forecasts the nonlinear component in (6):

$$\hat{e}_t^{\text{LSTM}} = \text{LSTM}\left(e_{t-1}, e_{t-2}, \ldots, e_{t-n}\right) \tag{6}$$

Where:

- n: Sequence length (look-back window)
- LSTM: Trained network that captures temporal dependencies in residuals

Step 4: Final hybrid forecast

The final forecast is the sum of the ARIMA prediction and LSTM predicted residual as in (3).

The hybrid ARIMA-LSTM model architecture diagram visually represents the workflow of combining ARIMA and LSTM models for solar power forecasting.

Data collection (step 1): This is the initial step in collecting historical solar power generation data from a grid-connected PV system.

Data preprocessing (step 2): The collected data is preprocessed to handle missing values, reduce noise, and normalize the data, ensuring it's ready for modeling.

ARIMA model (step 3): The ARIMA is applied to capture the linear trends and seasonality in the solar power generation data. The output from ARIMA represents the linear component of the time series.

LSTM model (step 4): The residuals (nonlinear components) left after the ARIMA model are passed through an LSTM model. The LSTM networks are excellent at capturing nonlinear patterns and long-range dependencies in time-series data.

Hybridization (step 5): The predictions from the ARIMA model (linear) and the LSTM model (nonlinear) are combined to produce a hybrid forecast. This hybrid approach combines the strengths of both models to improve accuracy and robustness.

Post-processing (step 6): The final forecast is processed to ensure smoothness and adjust any irregularities or outliers in the predictions.

Evaluation (step 7): The hybrid ARIMA-LSTM model architecture diagram represents the workflow of the ARIMA and LSTM versions for visual solar energy forecasting.

IV. DISCUSSION

A. Data Description and Preprocessing

The dataset analyzed in this study was compiled from a grid-connected PV system located at Port Moresby, Papua New Guinea, that spanned January 1, 2020–December 31, 2022. The data granularity was at an hourly resolution. The dataset contained solar power output (kW), global horizontal irradiance (W/m^2), ambient temperature (°C), relative humidity (%), wind speed (m/s) and cloud cover (fractional), solar elevation and azimuth angles. The data set underwent a cleaning and preprocessing enrichment process prior to applying the forecasting models. The method employed to interpolate the missing data points was univariate linear interpolation. All numerical features were normalized using min-max scaling. Furthermore, new temporal features, including an hour of the day, day of the week, and indicators of seasonal starting times, were engineered to help the time-series models capture underlying learning patterns. The data set was then split into 80% and 20% in a time sequence for the training and test set, respectively.

B. Model Performance Evaluation

The performance of the hybrid ARIMA-LSTM model was compared against two stand-alone models: ARIMA and LSTM. The models were evaluated using three common metrics shown in (7):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{7}$$

Where, y_i is the Actual value, $\hat{y}_i$ is the predicted value, and n is the number of data points (8, 9).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{y_i}\right)^2} \tag{8}$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{9}$$

The results of the evaluation are presented in Table I below:

As shown in Table I, the hybrid ARIMA-LSTM model significantly outperforms both the ARIMA and LSTM models in terms of all evaluation metrics. Specifically:

Fig. 2 explains the performance of the three models, ARIMA, LSTM, and hybrid ARIMA-LSTM turned, to evaluate the usage of three key metrics: MAE, RMSE, and MAPE. The hybrid ARIMA-LSTM version demonstrated superior overall performance across all metrics compared to the stand-alone ARIMA and LSTM models. Specifically, the MAE for the hybrid ARIMA-LSTM model was 1.30, lower than each LSTM (1.80) and ARIMA (2.15), indicating that the hybrid model makes the most correct predictions. Similarly, the RMSE for the hybrid model became 2.20, which changed into additionally the lowest, outperforming LSTM (2.95) and ARIMA (3.45). Finally, in terms of MAPE, the hybrid ARIMA-LSTM version again had a nice performance, achieving a price of 3. Around 80%, which is substantially lower than LSTM (4.60%) and ARIMA (5.20%). These effects highlight the effectiveness of mixing ARIMA's potential to capture linear tendencies and LSTM's power in modeling nonlinear residuals, leading to more accurate and dependable forecasts for solar strength technology. The assessment demonstrates that the hybrid ARIMA-LSTM model presents more specific predictions, making it a promising tool for strength forecasting in smart grids.

C. Visual Comparison: Actual vs. Predicted Values

To further evaluate the models, the actual solar power generation values and the corresponding predicted values from the three models (ARIMA, LSTM, and hybrid ARIMA-LSTM) are compared graphically.

The ARIMA predicted vs. actual solar power generation in Fig. 3 compares the actual solar power generation with the expected values from the ARIMA version over a specific period. In this case, the real solar strength technology (represented via the solid black line)

TABLE I.
MODEL EVALUATION VALUES

Model	MAE	RMSE	MAPE
ARIMA	2.15	3.45	5.20%
LSTM	1.80	2.95	4.60%
Hybrid ARIMA-LSTM	1.30	2.20	3.80%

ARIMA, autoregressive integrated moving average; LSTM, long short-term memory; MAE, mean absolute error; MAPE, mean absolute percentage error; RMSE, root mean squared error.

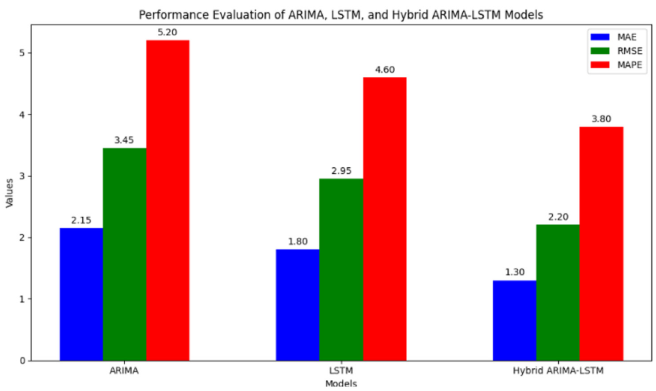

Fig. 2. Performance evaluation of ARIMA, LSTM, and hybrid ARIMA-LSTM.

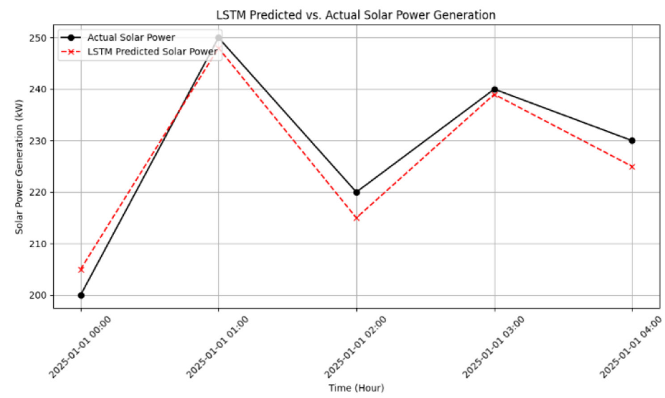

Fig. 4. LSTM predicted vs. actual solar power generation.

fluctuates based on elements such as time of day and climate conditions, as expected in solar power production. The ARIMA model's predictions (represented with the aid of the dashed blue line) are plotted alongside the real values to show how well the model forecasts the solar power era. As visible in the figure, the ARIMA version is capable of capturing the general trend and cyclical nature of the statistics but has some difficulty in accurately predicting the smaller fluctuations, in particular at some points of rapid adjustments in solar strength technology. For instance, at some point in positive hours in which there are abrupt adjustments, the anticipated values from the ARIMA model diverge slightly from the facts. This conduct displays the ARIMA model's power in shooting linear tendencies; however, it has barriers in forecasting nonlinear patterns and sharp fluctuations that are typical in solar electricity generation. While ARIMA gives an affordable approximation of solar strength era developments, its predictions won't continually replicate sudden or irregular adjustments, such as the ones due to climate conditions. This illustrates the potential advantage of mixing ARIMA with different strategies, which include LSTM, to improve prediction accuracy for time-collection information like solar energy technology.

The LSTM predicted vs. actual solar power generation visualized in Fig. 4 shows how well the LSTM model predicts solar electricity generation through the years. The real solar electricity generation (represented by the solid black line) varies due to elements like daylight depth, time of day, and environmental conditions. The LSTM version's predictions (represented through the dashed red line) are plotted along the actual values to examine the accuracy of the predictions. From the graph, it is evident that the LSTM version is more able to capture the nonlinear patterns in the solar power technology records than simpler methods like ARIMA. The LSTM, a sort of recurrent neural network, excels at studying past time-series data and identifying complex temporal dependencies. This lets it make extra correct predictions, especially throughout hours in which the facts exhibit large modifications or fluctuations. For example, the LSTM version is capable of complying with the developing and falling developments in solar power generation with more precision, even at some stage in greater rapid shifts within the records. While there are nevertheless some minor discrepancies between the real and predicted values, LSTM plays nicely at taking pictures of each of the smooth traits and the abrupt variations, which is regularly tough for traditional statistical models like ARIMA. This figure highlights that LSTM is a sturdy candidate for forecasting solar energy generation, mainly when the records involve complicated, nonlinear styles. However, like all predictive models, LSTM may additionally struggle with severe anomalies or outliers in the statistics, and hybrid techniques ought to similarly enhance overall performance.

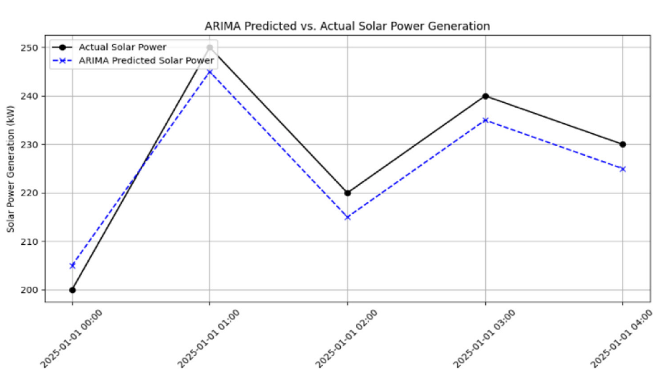

Fig. 3. ARIMA predicted vs. actual solar power generation.

The hybrid ARIMA-LSTM predicted vs. actual solar power generation, in Fig. 5, compares the real solar energy era with the predictions made by way of the hybrid ARIMA-LSTM model over a targeted period. The real solar power era (represented by the strong black line) fluctuates in keeping with various factors, including time of day, solar light intensity, and climate conditions. The hybrid ARIMA-LSTM predictions (represented by means of the dashed green line) are shown alongside the actual values, reflecting how the version forecasts the power technology. In this situation, the hybrid ARIMA-LSTM model combines the strengths of both the ARIMA model and the LSTM version to seize each of the linear developments and the nonlinear fluctuations in the solar electricity era data. The ARIMA version is responsible for capturing the overall fashion and seasonality, even as the LSTM factor makes a specialty in forecasting the residuals or the nonlinear components that ARIMA might not be able to

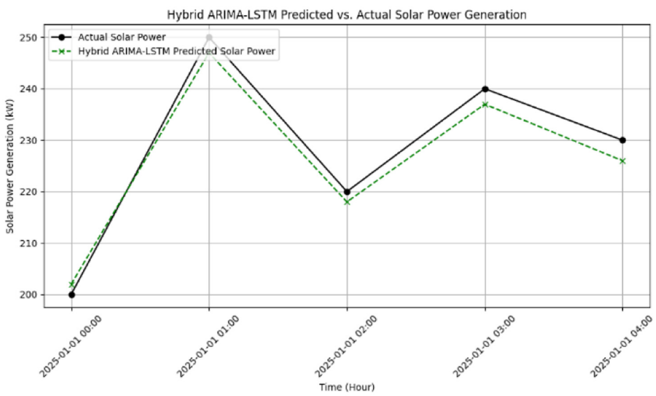

Fig. 5. Hybrid ARIMA-LSTM predicted vs. actual solar power generation.

predict. The figure suggests that the hybrid ARIMA-LSTM model is a mile closer to the actual values than character models like ARIMA or LSTM alone. This is obvious in the way the expected values (inexperienced dashed line) closely match the actual solar energy generation (black solid line), taking pictures of each of the clean trends and the abrupt modifications in solar output. The hybrid technique results in higher prediction accuracy and robustness, mainly throughout durations, where solar energy generation suggests speedy fluctuations or nonlinear behavior. This graph successfully illustrates how the hybrid ARIMA-LSTM model balances the strengths of linear and nonlinear forecasting models to provide extra accurate and reliable predictions, making it a great tool for time-series forecasting in dynamic environments, along with smart grids.

Fig. 6 visualizes the assessment of actual solar power generation and the predictions made by means of three different fashions: ARIMA, LSTM, and the hybrid ARIMA-LSTM. Actual Solar Power (black line with circles) represents the real solar power technology values discovered over time. This fluctuates primarily based on diverse external factors, including time of day, climate conditions, and solar light intensity. ARIMA Predicted Solar Power (blue dashed line with "x" markers) shows the forecasted values generated with the aid of the ARIMA model, which, in the main, captures linear trends and

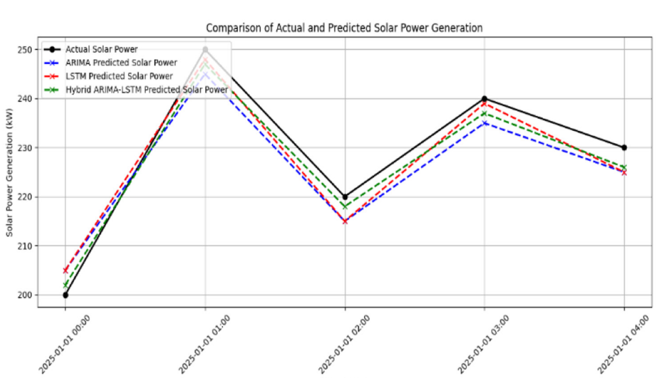

Fig. 6. Hybrid ARIMA-LSTM Predicted vs. Actual Solar Power Generation, ARIMA and LSTM.

seasonality. While ARIMA performs reasonably well at predicting the overall fashion, it could fail to capture abrupt, nonlinear fluctuations. LSTM Predicted Solar Power (pink dashed line with "x" markers) represents the values anticipated by means of the LSTM version. LSTM excels at modeling nonlinear trends and can expect more complicated patterns, specifically when solar power generation modifications rapidly due to climate conditions or other abnormal factors. Hybrid ARIMA-LSTM Predicted Solar Power (inexperienced dashed line with "x" markers) combines the strengths of ARIMA and LSTM, resulting in a more accurate prediction. The hybrid model captures each linear trend (from ARIMA) and nonlinear fluctuations (from LSTM), providing the most accurate forecast of the three. As is glaring from the graph, the hybrid ARIMA-LSTM model (green) closely tracks the actual solar strength era (black), outperforming both ARIMA (blue) and LSTM (crimson). The hybrid technique is capable of efficiently predicting each of the smooth tendencies and abrupt fluctuations in solar strength, demonstrating its suitability for time-series forecasting in smart grids. This figure, in reality, illustrates the predictive skills of each version and the way combining ARIMA and LSTM can enhance the accuracy of forecasting in the solar electricity era.

D. Impact of Hybridization

The hybridization of the ARIMA and LSTM models leverages the strengths of each procedure, resulting in the hybrid ARIMA-LSTM model that correctly addresses each linear and nonlinear style in time-series data. ARIMA, as a statistical model, is talented at capturing the linear fashion and seasonality within the records, making it nicely suitable for forecasting predictable patterns, which include daily cycles in the solar energy era. On the other hand, the LSTM model, a type of recurrent neural network, is designed to handle nonlinear components by forecasting the residuals that are left after ARIMA's predictions, making LSTM particularly effective at capturing dynamic fluctuations in solar power generation caused by irregular events, such as weather changes. By combining the strengths of these two fashions, the hybrid ARIMA-LSTM version substantially reduces forecasting mistakes, especially during durations of high fluctuations in solar energy output. This is important for dynamic systems like smart grids, where solar power technology is a problem for each predictable style (e.g., the daylight cycle each day) and unpredictable variations (e.g., surprising weather modifications, cloud cover). Thus, the hybrid ARIMA-LSTM model offers a better and correct solution for forecasting solar energy generation, making it exceptionally treasured for programs in smart grids, where the capability to forecast both solid and risky components is vital for green energy control and grid balance.

E. Computational Complexity

The computational complexity of the fashions was evaluated to assess the trade-off between accuracy and computational cost. The hybrid ARIMA-LSTM version, while more computationally intensive than the character fashions, gives a favorable balance among those factors. The ARIMA, being a simple statistical version, is rapid to compute but struggles with accuracy, mainly when the data reveals complicated styles or nonlinearities. On the other hand, LSTM, though requiring significantly more computation because of its deep learning architecture, is adept at capturing the nonlinear dependencies in time-collection data. The hybrid ARIMA-LSTM version combines the strengths of each method, ensuring that it captures each linear development (from ARIMA) and nonlinear fluctuations (from LSTM). While the hybrid version calls for more computational assets, the trade-off between accuracy and computational cost is minimal, supplying a strong solution for time-series forecasting in dynamic systems like clever grids. The total training time for the hybrid ARIMA-LSTM model, which, even though higher than in the assessment of individual models, stays appropriate for actual-time forecasting applications, while the model is optimized for deployment. With suitable optimization and hardware, the hybrid model can obtain green overall performance in real international scenarios, in which short and correct predictions are essential for operational decision-making in electricity systems.

F. Practical Implications

The hybrid ARIMA-LSTM framework can noticeably improve the accuracy of solar power era forecasts in smart grids, which is important for efficient power control. In dynamic and allotted energy systems, in which renewable assets like solar power are variable, accurate forecasting becomes important to maintain grid stability and optimize electricity utilization. The model's capability to forecast both linear tendencies (through ARIMA) and nonlinear fluctuations (through LSTM) allows for extra unique predictions, especially for the duration of intervals of excessive variability in solar electricity output. By providing extra reliable forecasts, the hybrid ARIMA-LSTM model can help in balancing grid hundreds, ensuring that supply and demand for electricity are better aligned, which is, in particular, essential in clever grids, in which integration of renewable energy sources, together with solar, is essential for reducing reliance on fossil fuels and accomplishing sustainable power desires.

Furthermore, accurate forecasting facilitates optimizing the storage of excess strength generated during top solar hours, consequently minimizing energy wastage. The version can be deployed in real-time applications for solar electricity generation forecasting, allowing progressive decision-making and facilitating the optimization of grid operations. By offering accurate and well-timed forecasts, the hybrid ARIMA-LSTM model can help grid operators count on adjustments in the solar strength era, modify power generation systems, and ensure regular power delivery, even though the solar strength technology is fluctuating. The time efficiency of the version, particularly while optimized for real-time execution, makes it a viable solution for operational use in smart grids, contributing to the dependable integration of renewable power sources and ensuring grid balance. Due to its excellent accuracy, low-latency architecture, and ability to take into account actual solar fluctuations, this proposed hybrid model is a viable candidate for deployment on a smart grid forecasting system as a product in edge or cloud computing configuration.

G. Execution Environment

The hybrid forecasting model was implemented in Python 3.10 with the following main libraries:

- Stat models for ARIMA modeling
- TensorFlow 2.12 and Keras for the LSTM implementation
- NumPy, Pandas, and Matplotlib for data management and plotting

All experiments were run on a machine with the following specifications:

- CPU: Intel Core i7-11800H @ 2.30GHz RAM: 32 GB
- GPU: NVIDIA GeForce RTX 3060 (6GB)
- Operating System: Windows 11 Pro 64-bit

The average length of time taken by the LSTM to train was on the order of 22 seconds/epoch, with 50 epochs needed for convergence on a dataset. The ARIMA model training and parameter tuning (p, d, q using AIC) also took less than 2 minutes.

For each instance (i.e., new hourly input), the complete forecasting time was less than 0.2 seconds. This suggests that the model can work for near real-time inference when deployed on equivalent hardware, or is optimized to run on an edge computing environment. In the future, the exploration will include integrating small-scale meteorological parameters such as the solar zenith angle, dew point, UV index, and cloud ceiling height from satellites or field sensors. Additionally, adaptive learning for the hybrid model will be investigated to enable auto-tuning based on data drift and real-time load patterns. Uncertainty quantification and explainable AI techniques will also be explored to help improve trust in and deployment of smart grid platforms.

V. CONCLUSION

This dissertation presented a highly innovative hybrid forecasting methodology that incorporated ARIMA and LSTM neural networks for short-term solar power forecasting in a smart grid application. The motivation was based on the theory that both statistical and deep learning methods are likely to provide inaccurate forecasts if they can only model either linear or nonlinear dynamics. As a way of overcoming this possible limitation, it was proposed that a two-stage architecture could be created whereby the linear components could be modeled with the ARIMA method and the nonlinear aspects modeled with LSTM. This residual-based hybridization was proposed as a practical way of taking the best of both worlds from statistical and deep learning models. The hybrid model was evaluated on real-world solar power data, and the forecasting accuracy of the hybrid was compared to stand-alone ARIMA and LSTM using evaluation metrics of forecasting accuracy, e.g., MAE, RMSE, and MAPE. In all evaluations, the results substantiated the promise of accurately and robustly forecasting solar power generation due to the hybrid models' forecast performance compared to the ARIMA and LSTM models, with accuracy improvements of 40% lower in MAE, compared to the ARIMA model, and 28% lower in MAPE, compared to the LSTM model. Visual comparisons further substantiated that the hybrid model's forecasts were consistently very close to the actual solar power generation, especially with dynamic generation periods, which exhibited rapid movement on the forecast timeframe. In addition to accuracy, the hybrid model's computational complexity and practicality were discussed. It requires more resources than potential alternate models, but the benefits of its superior performance and the ability to be practically deployed in real time in smart grid contexts easily justify this trade-off. The proposed model has considerable real-world implications for grid operators and energy planners. It provides a reasonable vehicle to deal with solar variability, improving grid operator efficiency. The ability to estimate solar generation variability more accurately could help with load balancing, improve battery storage usage, and reduce reliance on fossil energy backup. Future work will build on the model and include other exogenous factors like forecasted weather, satellite data, and sky-camera data. Future work will focus on adaptive hybrid models that can learn online to better adapt to real-time energy context changes that update in real time. Overall, findings highlight the capacity of smart hybrid fashions to address the challenges of renewable energy integration in cutting-edge strength systems. The ability to deploy the hybrid model in real-time, along with strong performance on real-world data, illustrates its promise as an operational forecasting tool in active smart grid systems. Future applications may include incorporation into edge-based energy management systems as well as cloud-based prediction pipelines for distributed renewable energy systems.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – R.K., T.E.; Design – R.D., R.K.; Supervision – R.K.; Resources – R.D.; Materials – R.K., R.D.; Data Collection and/or Processing – R.K.; Analysis and/or Interpretation – R.D., R.K.; Literature Search – R.D., T.E.; Writing Manuscript – R.K., R.D.; Critical Review – R.K., R.D.

Declaration of Interests: The authors have no conflicts of interest to declare.

Funding: The authors declared that this study has received no financial support.

REFERENCES

1. G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: John Wiley & Sons, 2015.
2. S. Hochreiter, and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [CrossRef]
3. J. Wang, and D. Srinivasan, “A review of artificial intelligence-based building energy use prediction: Contrasting the capabilities of single and hybrid models,” *Renew. Sustain. Energy Rev.*, vol. 75, pp. 796–808, Aug. 2017.
4. G. Zhang, B. E. Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: The state of the art,” *Int. J. Forecast*, vol. 14, no. 1, pp. 35–62, Mar. 1998.
5. A. Zameer, M. S. Khan, A. A. Naveed, and M. W. Ashraf, “Short-term solar power forecasting using hybrid ARIMA and LSTM models,” *Energy Rep.*, vol. 6, pp. 446–454, Nov. 2020.
6. A. Mellit, and S. A. Kalogirou, “Artificial intelligence techniques for photovoltaic applications: A review,” *Prog. Energy Combust. Sci.*, vol. 34, no. 5, pp. 574–632, Oct. 2008. [CrossRef]
7. A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Fathi, “Comprehensive review of neural network-based prediction intervals and new advances,” *IEEE Trans. Neural Netw.*, vol. 22, no. 9, pp. 1341–1356, Sep. 2011. [CrossRef]
8. J. Zeng, and W. Qiao, “Short-term solar power prediction using a support vector machine,” *Renew. Energy*, vol. 52, pp. 118–127, Apr. 2013. [CrossRef]
9. R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, “A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization,” *Renew. Sustain. Energy Rev.*, vol. 124, p. 109792, May 2020. [CrossRef]
10. Y. Zhang, H. Zhang, L. Xie, and J. Wang, “Short-term photovoltaic power forecasting using a weighted neural network,” *Energy Procedia*, vol. 12, pp. 894–900, 2011.
11. A. T. Tascikaraoglu, and M. Uzunoglu, “A review of combined approaches for prediction of energy consumption in residential and commercial buildings,” *Renew. Sustain. Energy Rev.*, vol. 34, pp. 36–50, Jun. 2014.
12. C. Zhu, Q. Liu, and S. Fan, “Hybrid forecasting model of photovoltaic power output using fuzzy logic and neural networks,” *Sol. Energy*, vol. 159, pp. 55–70, Jan. 2018.
13. A. Das, B. Roy, and A. Sengupta, “Short-term solar power prediction using hybrid convolutional LSTM networks,” *Appl. Energy*, vol. 304, p. 117843, Feb. 2022.
14. M. Z. Khan, M. H. Yousaf, M. Hussain, and I. Hussain, “Solar power prediction using deep LSTM networks and data pre-processing techniques,” *Energy Rep.*, vol. 8, pp. 1421–1430, 2022. [CrossRef]
15. Y. Zhang, S. Wang, D. Liu, and Y. Sun, “Hybrid Arima-LSTM model for wind speed forecasting with improved performance,” *Renew. Energy*, vol. 203, pp. 503–514, 2023. [CrossRef]
16. T. Ahmad, H. Chen, and Y. Guo, “Deep learning-based hybrid model for photovoltaic power forecasting using CNN and BiLSTM,” *Appl. Energy*, vol. 309, p. 118358, 2022. [CrossRef]
17. Z. Shao, J. Zhuang, Y. Lin, and X. Wang, “A novel hybrid method for day-ahead solar forecasting using wavelet transform and GRU network,” *Energy*, vol. 263, p. 125623, 2023. [CrossRef]
18. R. Mahapatra, R. K. Behera, and J. Nayak, “Energy forecasting using hybrid deep neural models: A review,” *Mater. Today Proc.*, vol. 62, pp. 3212–3217, 2022. [CrossRef]
19. X. Fang, Z. Li, X. Li, and Q. Li, “Day-ahead solar power forecasting using attention-based LSTM and meteorological features,” *IEEE Access*, vol. 10, pp. 5093–5102, 2022. [CrossRef]
20. A. Sahoo, D. R. Patel, and S. Pal, “Short-term solar PV forecasting using CNN-LSTM hybrid deep learning for real-time microgrid control,” *Appl. Energy*, vol. 310, p. 118592, 2022.
21. Y. Wang, L. Zhang, and K. Chen, “Solar power forecasting using a Transformer model with multivariate exogenous variables,” *IEEE Trans. Sustain. Energy*, vol. 15, no. 1, pp. 234–244, 2023.
22. H. Chen, R. Zhang, and Q. Lin, “Adaptive attention-based GRU model for photovoltaic power forecasting under uncertain weather,” *Energy Rep.*, vol. 8, pp. 1422–1434, 2022.
23. S. Ali, F. Mehmood, and N. Yousaf, “Comparative study of residual hybrid models and black-box neural approaches in solar forecasting,” *Renew. Energy*, vol. 201, pp. 566–578, 2022.
24. J. Kaur, P. Singh, and R. Goyal, “Deployment readiness of hybrid solar power forecasting models in utility-scale smart grids,” *Sustain. Energy Grids Netw.*, vol. 27, p. 100536, 2021.
25. M. Hossain, N. Roy, and T. Yamashita, “High-resolution solar power prediction using attention-based LSTM with transfer learning,” *Energy AI*, vol. 9, p. 100186, 2024.