Energy and AI 21 (2025) 100540

Contents lists available at ScienceDirect

Energy and AI

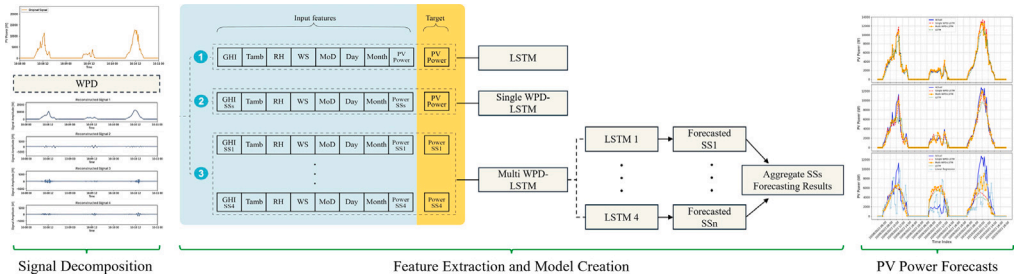
journal homepage: www.elsevier.com/locate/egyai

Multi-temporal PV power prediction using long short-term memory and wavelet packet decomposition

Amirhasan Sardarabadi [a,b], Amirhossein Heydarian Ardakani [a], Silvana Matrone [b], Emanuele Ogliari [b], Elham Shirazi [a,*]

[a] Faculty of Engineering Technology, University of Twente, Netherlands
[b] Department of Energy, Politecnico di Milano, Italy

GRAPHICAL ABSTRACT



HIGHLIGHTS

- Multi-temporal PV power forecasts for different horizons, including 15-min, 1-hour, and day ahead.
- Ensemble WPD-LSTM model enhances PV power forecasting and outperforms benchmark models.
- Wavelet Packet Decomposition reduces noise and enhances handling of non-stationarity.
- WPD offers a hierarchical decomposition, allowing forecasts to be made at different resolutions.
- Using weight optimization to reconstruct the predicted components into the final forecast.

ARTICLE INFO

Keywords:
Solar power forecasting
Wavelet packet decomposition
Long short-term memory
Deep learning models

ABSTRACT

The integration of photovoltaic (PV) systems into power grids presents operational challenges due to the inherent variability in solar power generation. Accurate PV power forecasting can help address these issues by enhancing grid reliability and energy management. This study introduces a novel hybrid deep learning approach that combines Wavelet Packet Decomposition (WPD) and Long Short-Term Memory (LSTM) networks to improve forecasting accuracy across multiple time horizons. The proposed model incorporates a dynamic weighting mechanism to optimally integrate the forecasts of decomposed subseries, effectively capturing both high- and low-frequency components of the power signal. Using real-world data from a solar parking site at the University of Twente, Netherlands, the proposed models are compared with standard LSTM, Linear Regression, and Persistence baselines across 15 min, 1-hour, and day-ahead horizons. The WPD-LSTM model with weight optimization reduces nRMSE by up to 72.5%, 52.9%, and 34.7% compared to Persistence, and by 68.6%, 36.1%, and 7.5% compared to standalone LSTM, respectively. These results highlight the effectiveness of the hybrid approach in delivering more accurate and robust PV power forecasts.

* Corresponding author.
E-mail addresses: a.heydarian@utwente.nl (A. Heydarian Ardakani), silvana.matrone@polimi.it (S. Matrone), e.shirazi@utwente.nl (E. Shirazi).

https://doi.org/10.1016/j.egyai.2025.100540
Received 27 February 2025; Received in revised form 19 May 2025; Accepted 13 June 2025
Available online 28 June 2025

1. Introduction

The transition to renewable energy sources (RES) has become essential as the world faces the consequences of climate change. The RES integration has further escalated due to the advances in RES technologies, decreasing costs, and supportive policies. By 2028, solar photovoltaic (PV) and wind energy are expected to provide 25% of the world's electricity, surpassing conventional energy sources like coal and nuclear power [1]. However, this rapid growth comes with a major challenge. Solar power generation is inherently variable, influenced by environmental factors such as solar irradiance, temperature, cloud cover, wind speed, and humidity [2]. These fluctuations in PV power output pose challenges to grid stability and reliability [3].

Unlike conventional power sources, such as fossil fuels, which can provide a steady output, solar power generation is intermittent. It peaks around midday when the sunlight is strongest and drops to zero at night. This daily fluctuation requires careful planning to manage energy flows to balance consumption during periods when solar generation is unavailable. Seasonal changes further complicate this, with winter months typically providing less solar energy as a result of shorter daylight hours and lower sun angles [4]. Moreover, weather conditions such as cloud cover can cause rapid reductions in solar irradiance, leading to sharp declines in PV output [5].

The unpredictable nature of solar power generation presents significant challenges to grid stability and reliability. As solar output fluctuates throughout the day and across seasons, it often does not align with the varying energy demands of consumers. Unlike traditional power systems, which can guarantee consistent output to meet demand, the sudden drops in solar generation can create supply shortages, risking voltage instability and grid disruptions [6]. Accurate forecasts enable more effective grid planning and operation, reduce the need for standby fossil-fueled power stations, and help minimize operational costs [7]. This not only optimizes the use of RES but also contributes to a more sustainable energy system, ultimately aiding in the reduction of greenhouse gas emissions [8].

Forecasting solar power output varies in complexity depending on the forecast horizon, with each time frame serving specific operational needs. As an example, nowcasting focuses on predictions within 15 min and plays a crucial role in real-time grid management and immediate corrective actions. Hour-ahead or intra-day forecasts, covering 15 min to 5 h, support energy trading and optimize short-term operational decisions. Meanwhile, day-ahead forecasts, which span from several hours to a full day, are essential for resource planning, economic dispatch, and grid stability. For longer-term forecasts, spanning weeks to months, applications include maintenance scheduling and financial planning, where even small inaccuracies could impact long-term financial outcomes [9].

1.1. Contributions

Despite considerable progress in solar power forecasting, challenges persist, particularly in regions with highly variable weather conditions. Traditional models often struggle to accurately predict PV power output when faced with abrupt changes caused by cloudy or rainy days, which are difficult to capture using historical PV data alone. The inherent seasonality and periodicity of solar power generation are frequently disrupted by these unpredictable weather patterns, resulting in noisy and erratic power signals.

This study aims to bridge the gaps in the literature by extending the forecasting horizon and introducing a novel application of existing methods to enhance prediction accuracy. The primary contributions of this work are as follows.

- A novel hybrid deep learning model combining WPD and LSTM is introduced to specifically address the issue of short-term LSTM predictions behaving like persistence models.
- An ensemble of multiple models is trained, and their results are combined using optimized weights, providing a more accurate forecasting approach compared to traditional methods.
- The proposed hybrid model is evaluated on a real-world case study using historical data collected from a solar parking facility located at the University of Twente in Netherlands.

The remainder of this paper is organized as follows. Section 2 reviews the literature. Section 3 details the research methodology, including data preprocessing, model description, prediction horizon, and weight optimization. Section 4 presents the results and discussion of the proposed forecasting models. Lastly, Section 5 concludes the study with key findings and implications.

2. Literature review

In short-term PV forecasting, weather conditions, including geographic location, sun position, and atmospheric conditions, are important factors in determining solar irradiance. This information could be captured by analyzing sky images. In this regard, the authors in [10] applied linear machine learning models (e.g., Bayesian Ridge (BR) regression, Stochastic Gradient Descent (SGD), and Generalized Linear Model (GLM) regression) to features extracted from sky images. Building on this approach, in [11], Convolutional Neural Networks are used to process infrared sky images as they better represent cloud features.

Using time series data has demonstrated strong potential in PV forecasting, as evidenced by models using solar irradiance, module temperature, and time-location data [12]. Studies have further identified key correlations between PV output and economic as well as geographical factors, uncovering unexploited potential in several regions [13]. Energy management frameworks for smart PV systems, supported by accurate forecasts, have greatly reduced electricity costs by enabling more efficient optimization of energy resources [14]. Additionally, prior work has demonstrated that machine learning models, including neural networks and regression approaches, can significantly enhance energy prediction accuracy and grid efficiency [15], motivating further exploration into hybrid and deep learning architectures for PV forecasting.

Sequential models such as Long Short-Term Memory (LSTM) networks, have garnered attention for their ability to capture complex patterns and dependencies in time series data [9]. Research has explored the application of LSTM networks to PV forecasting, with notable improvements in performance when hyperparameters were optimized through methods such as grid search and K-Fold cross-validation [16]. Additionally, the importance of selecting relevant input variables, including meteorological factors like solar radiation, wind speed, and cloud cover, has been emphasized as critical to improving forecast accuracy [17].

Recent studies have applied hybrid deep learning models to refine the accuracy of the forecasts. For example, a model combining Wavelet Packet Decomposition (WPD) and LSTM networks showed great performance in one-hour-ahead forecasting of PV power. However, this approach is primarily focused on short-term predictions, leaving room for exploration of longer forecasting periods [18]. Similarly, in wind power forecasting, a hybrid model merging WPD, LSTM, and CNN was employed to handle different frequency components of wind power data. The model forecasts both low and high-frequency sub-layers of wind power data and is optimized using sequential model-based optimization. The proposed model outperforms traditional approaches and improves prediction accuracy by at least 77.4% [19]. Another study introduced a new hybrid model by combining Gaussian process regression with WPD to forecast PV output. This method proved to be more accurate compared to traditional models varying seasonal conditions. While it performed well in short-term predictions, its precision decreased for forecasts [20]. The effectiveness of these hybrid methods

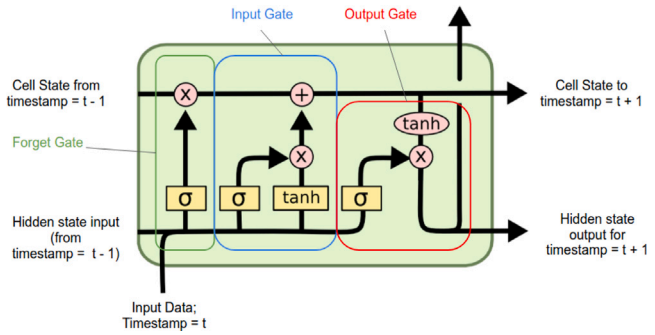


Fig. 1. LSTM cell architecture highlighting the structure of the forget, input, and output gates.

indicates that combining temporal and spatial information can lead to better forecasting performance [21].

Innovative advancements, including the integration of attention mechanisms in forecasting models, have demonstrated improved accuracy by allowing neural networks to focus on the most relevant data [22,23]. These methods have shown reliability in managing erratic and noisy time series signals, particularly during periods of seasonal variation and rapid weather changes. In addition, studies have optimized LSTM configurations for short-term forecasts by developing models tailored to specific seasons, using meteorological data from NASA to improve accuracy [24].

3. Research methodology

PV power forecasting is often influenced by complex and non-linear patterns. This limitation is more significant in regions with highly variable weather conditions, such as the Netherlands, where fluctuating weather conditions introduce considerable noise into the irradiance signal, which affects the model predictions [25].

To address these challenges, this section proposes a hybrid model for PV power forecasting. The proposed model employs LSTM architecture to capture the non-linear relationships between meteorological variables and PV power output. To further improve the model accuracy across different time horizons, WPD is applied to reduce noise and manage uncertainty in the input signal. The following sections describe the method in more detail.

3.1. Long short-term memory

LSTM networks are designed to capture long-term dependencies using a unique structure known as the cell state. The cell state acts as a memory pipeline, allowing the network to retain crucial information over long periods. As shown in Fig. 1, the flow of information within an LSTM is controlled by three primary gates: the input gate, the forget gate, and the output gate. Each gate plays a critical role in updating the cell state and hidden state as new data are processed [26].

3.1.1. Input gate

The input gate decides what new information will be added to the cell state. It consists of two parts: the sigmoid layer i_t, which determines which values to update, and the tanh layer $\tilde{C}_t$, which generates new candidate values:

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

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \tag{2}$$

where W_i, b_i, W_C, and b_C are the weight and biases for the input gate and candidate cell state, respectively.

3.1.2. Forget gate

The forget gate determines which information from the previous cell state C_{t-1} should be discarded. This is achieved by applying a sigmoid function to the previous hidden state h_{t-1} and the current input x_t, generating a forget vector f_t with values between 0 and 1:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \tag{3}$$

where W_f and b_f are the weight and bias for the forget gate and σ denotes the sigmoid function.

3.1.3. Updating the cell state

The cell state C_t is updated by combining the previous cell state C_{t-1} and the new candidate values $\tilde{C}_t$. The forget gate f_t scales the previous cell state, and the input gate i_t scales the candidate values:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t. \tag{4}$$

This equation ensures that the LSTM retains important long-term information while incorporating relevant new information.

3.1.4. Output gate

The output gate determines the next hidden state h_t, which will be passed to the next time step and also used as part of the final output. The output gate uses the updated cell state C_t to produce the hidden state h_t:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \tag{5}$$

$$h_t = o_t \cdot \tanh(C_t), \tag{6}$$

where W_o and b_o are the weight and bias for the output gate. Through these mechanisms, LSTMs can effectively manage long-term dependencies in sequential data, making them particularly suitable for complex tasks such as time series forecasting.

3.2. Wavelet packet decomposition

The PV system under study is located in an area characterized by highly variable weather conditions, resulting in frequent fluctuations in PV power generation. These fluctuations introduce considerable noise into the power output signal, which can adversely affect the performance of forecasting methods, including LSTM.

Traditional forecasting approaches often struggle with rapidly changing data that alternates between increasing and decreasing values within short intervals. To address this challenge, WPD is used. It is a signal processing technique that extends the traditional Discrete Wavelet Transform (DWT) by providing a more comprehensive time-frequency analysis. Unlike DWT, which recursively decomposes only the low-frequency (approximation) components of a signal, WPD decomposes both the approximation and high-frequency (detail) components at each level. This process creates a complete binary tree structure, where each node represents a specific subband corresponding to a defined frequency range [27].

At each level of decomposition, the signal is convolved with a pair of quadrature mirror filters: a low-pass filter to extract lower-frequency components and a high-pass filter to extract higher-frequency components. The outputs of these filters are then downsampled to reduce redundancy, resulting in two sets of coefficients at each step. This recursive process continues to a specified depth, with the number of subbands at level n being 2^n. The ability to decompose both approximation and detail components allows WPD to capture fine-grained frequency information that is often missed in DWT, making it particularly suitable for signals with complex or non-stationary characteristics.

The selection of the wavelet function, such as Daubechies, Symlets, or Coiflets, determines the shape and resolution of the analysis, while the number of decomposition levels controls the granularity of the frequency subbands. In this work we chose to use Daubechies wavelets based on their proven ability to capture both high and low-frequency

components effectively while maintaining computational efficiency, which is critical for real-time applications. Daubechies wavelets, especially with higher orders, provide a good balance between smoothness and the ability to localize signal characteristics, making them suitable for our multi-temporal forecasting model. WPD coefficients represent the signal's energy distribution across time and frequency, making them valuable features for applications such as fault diagnosis, biomedical signal processing, and audio analysis. This theoretical foundation of WPD ensures its effectiveness in extracting detailed time-frequency characteristics for downstream analysis.

WPD is particularly effective for non-stationary signals like those found in PV power generation, as it can decompose a signal into components at multiple frequency levels. This allows the isolation of high-frequency noise from the underlying trends, thereby clarifying the signal and reducing noise. The effectiveness of WPD in managing non-stationary signals makes it an ideal choice for improving the performance of LSTM models in this study [18,20].

The original PV power output and Global Horizontal Irradiance (GHI) time series are represented as:

$$X = [P_{t-N}, P_{t-N+1}, \dots, P_t]^T, \tag{7}$$

$$Y = [G_{t-N}, G_{t-N+1}, \dots, G_t]^T. \tag{8}$$

These signals are decomposed into a set of sub-series:

$$x_1, x_2, \dots, x_i, \dots, x_m, \tag{9}$$

$$y_1, y_2, \dots, y_i, \dots, y_m, \tag{10}$$

where $m = 4$ to balance computational efficiency and decomposition effectiveness.

Each decomposed series is fed to a single-branch reconstruction method to maintain the original signal length, which is crucial for preserving temporal coherence. The reconstructed sub-series are expressed as:

$$S_{xi} = [p_{it-N}, p_{it-N+1}, \dots, p_{it}]^T, \tag{11}$$

$$S_{yi} = [g_{it-N}, g_{it-N+1}, \dots, g_{it}]^T. \tag{12}$$

This decomposition process allows for the effective isolation of different frequency components, thereby clarifying the signals and reducing noise. Fig. 3 illustrates the decomposition and reconstruction flowchart, along with a three-day analysis of the reconstructed signals, demonstrating the improvement in data quality crucial for developing precise and reliable forecasting models [19].

The original PV power signal (top panel) is decomposed using WPD, splitting it into multiple frequency components. This decomposition process is shown in the middle part of the figure, where the signal passes through low-pass and high-pass filters to isolate approximate and detailed components at different levels. Finally, the signals are reconstructed through a single-branch method, as displayed in the lower panels. Each reconstructed sub-series captures a specific frequency range, demonstrating the reduction of noise and clarification of underlying trends in the data over three days.

Wavelet decomposition allows the LSTM to focus on different timescale phenomena by isolating low, medium, and high-frequency components. LSTMs are well-suited for learning temporal dependencies, but when these dependencies occur at different time scales, the model may struggle to capture them effectively. By separating the frequency components, we enable the LSTM to specialize in learning specific patterns from each component, improving performance.

The low-frequency components capture long-term trends, such as daily solar cycles, which help the LSTM model learn broader patterns in PV power. In contrast, high-frequency components isolate short-term fluctuations, like sudden changes in irradiance caused by passing clouds or atmospheric disturbances, allowing the model to focus on these transient dynamics. This separation enables the LSTM to analyze both steady, long-term behavior and rapid, short-term variations more effectively, leading to improved forecasting accuracy. A complete flow chart summarizing the entire research methodology is provided in Fig. 2.

3.3. Multi-temporal forecasting models

In this study, several forecasting models are developed to predict single-step, 1-h ahead, and 15-min ahead and multi-step, day-ahead PV power output. Two different approaches are used when dealing with WPD-LSTM models. These approaches differ primarily in how they handle the decomposed subseries of PV power and GHI signals and how they integrate these subseries with other input variables to generate the final predictions.

3.3.1. Single WPD-LSTM

The first model, referred to as the Single WPD-LSTM, employs a single LSTM network. In this model, the original GHI and PV power signals are first decomposed using WPD into multiple subseries. These subseries consist of one low-frequency component representing the overall trend and several high-frequency components capturing various levels of noise. After decomposition, the subseries are reconstructed, and these reconstructed signals are fed into the LSTM model alongside the original input variables, including meteorological data and temporal features. The LSTM model predicts the original PV power output by learning from the integrated information in the reconstructed subseries and other inputs. The key characteristic of this model is that it uses a single unified framework to process all the input data and predict the PV power output as a whole, leveraging both the broad trends and fine-grained fluctuations present in the reconstructed signals.

3.3.2. Ensemble multi WPD-LSTMs

The second model, known as the Ensemble Multi WPD-LSTMs, adopts a more granular approach by further specializing in the prediction task. Here, the PV power and GHI signals are again decomposed into subseries using WPD. However, instead of integrating these subseries into a single model, each LSTM in the ensemble is dedicated to a specific reconstructed subseries. For each LSTM model in the ensemble, the input variables include its corresponding reconstructed subseries along with the associated meteorological and temporal features. Each LSTM is trained to predict its dedicated subseries of the PV power output. Once all the subseries have been predicted, they are aggregated to form the final PV power forecast. The ensemble method optimizes the weights assigned to each model's prediction to minimize the overall RMSE. This approach allows each LSTM model to specialize in capturing the distinct patterns or noise characteristics of its specific subseries, potentially leading to a more accurate and robust overall forecast by combining the strengths of multiple specialized models.

By comparing these two methods — the Single WPD-LSTM, which integrates all input variables into a single model, and the Ensemble Multi WPD-LSTMs, which distributes the prediction task among multiple models dedicated to specific reconstructed subseries — the study aims to evaluate the effectiveness of signal decomposition, reconstruction, and ensemble modeling in improving forecast accuracy. This comparison is expected to provide valuable insights into how best to manage the inherent noise and variability in PV power and GHI data, ultimately leading to more reliable and accurate predictions.

3.3.3. Weight optimization for multi WPD-LSTM model

The proposed Multi WPD-LSTM model employs a weight optimization approach to enhance forecasting accuracy. This method combines the outputs of multiple LSTM networks, each trained on different subseries obtained from WPD of the original PV power and GHI signals. The final prediction is a weighted sum of individual LSTM forecasts, with weights optimized to minimize the overall RMSE [18]. Mathematically, the final forecast at time t is expressed as:

$$\hat{y}_{\text{final}}(t) = w_1\hat{y}_1(t) + w_2\hat{y}_2(t) + w_3\hat{y}_3(t) + w_4\hat{y}_4(t) \tag{13}$$

where $\hat{y}_i(t)$ represents the forecast from the ith LSTM model and w_i is its corresponding weight. The optimization problem is formulated as:

$$\min_{\mathbf{w}} \ \text{RMSE}(\hat{y}_{\text{final}}(t), y(t)) \tag{14}$$

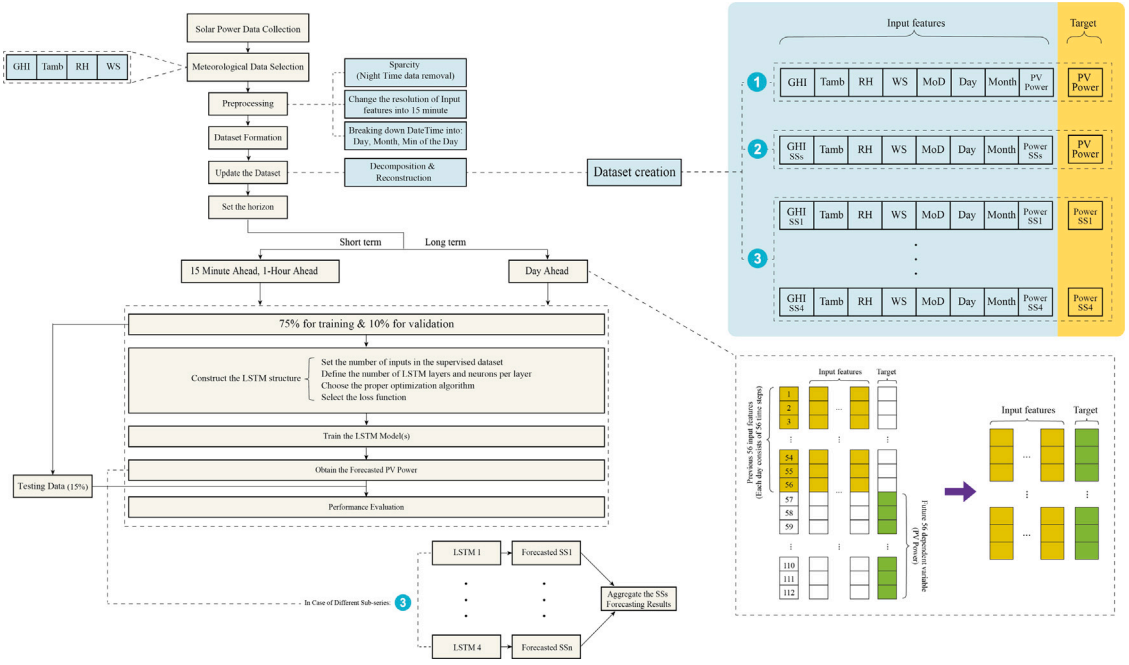


Fig. 2. Framework of the proposed LSTM-WPD PV Power Forecasting. Comprising data acquisition and preprocessing, time-series decomposition, and feature engineering. LSTM networks are trained for short-term and day-ahead forecasting, with individual sub-series predictions aggregated to form the final output in multi-model configurations.

where $\mathbf{w} = [w_1, w_2, w_3, w_4]$ is the weight vector. This optimization is solved using the Nelder–Mead method, a derivative-free technique [28]. The rationale behind this approach is that different decomposed sub-series and their corresponding LSTM models capture various aspects of the PV power signal. By optimizing the weights, the model leverages the strengths of individual LSTMs while mitigating their weaknesses, resulting in a more accurate and robust final prediction. This method addresses the challenge of combining multiple forecasts effectively, ensuring that each model's contribution is appropriately weighted to minimize the overall forecasting error.

3.4. Prediction horizon selection

Three different prediction horizons are selected for this study to address various operational needs in microgrids and energy markets: 15-min ahead, hour-ahead, and day-ahead forecasts. These horizons are chosen to comprehensively evaluate the proposed forecasting models across different timescales relevant to grid operations and energy trading [9,29].

Fig. 4 illustrates the concept of prediction horizon and its relationship with the input window length for a single-step 15-min ahead prediction. The input window, representing historical data, advances incrementally as new predictions are generated. For the 15-min ahead forecast, the model predicts the PV output power for the next immediate 15-min interval. In the case of hour-ahead predictions, a gap of four time steps (equivalent to one hour, given the 15-min data resolution) is introduced between the last time step of the input window and the forecasted value. This approach ensures an accurate representation of the hour-ahead horizon. Finally, the day-ahead forecast involves predicting 56 time steps, corresponding to the daylight hours of the subsequent day. Night-time hours are excluded from the analysis to focus on periods of active PV power generation.

4. Results and discussion

This section presents the performance outcomes of the proposed models, providing insights into their effectiveness and predictive capabilities. The performance of models is evaluated for the 15-min ahead, hour-ahead, and day-ahead forecasts.

4.1. Case study (moved to results)

This study uses historical data from a solar parking facility located on the University of Twente campus in Enschede, Netherlands for model evaluation. The dataset spans a period of 26 months, from September 1, 2021, to October 31, 2023, with measurements recorded at 15-min intervals.

To complement the PV power output data, meteorological variables such as GHI, Wind Speed (WS), Ambient Temperature (Tamb), and Relative Humidity (RH) are included. These data are sourced from the Royal Netherlands Meteorological Institute (KNMI) and initially recorded at 10-min intervals [30]. To maintain consistency with the PV power output data, the meteorological variables are adjusted to 15-min resolution using interpolation techniques. This represents modification is a form of downsampling, not upsampling. Specifically, we reduce six 10-min records per hour to four 15-min points, rather than generating artificial high-frequency data from lower-resolution measurements. This ensures that no fictitious high-resolution trends are introduced into the dataset. Moreover, meteorological variables such as temperature, irradiance, and humidity have low variability over short timescales, which justifies this interpolation approach.

During the preprocessing stage, the datetime information is decomposed into three key components: Minute of the Day, Day of the Month, and Month of the Year. This transformation is essential for capturing diurnal and seasonal variations in PV power output. Additionally, nighttime data (from 20:00 to 06:00), where PV generation is negligible, are excluded to minimize noise and enhance model performance.

4.2. Model training (moved to results)

The WPD forecasting models (i.e., Single WPD-LSTM and Ensemble Multi WPD-LSTM) utilize data preprocessed using WPD to extract time-frequency features from raw signals as described Section 3.2.

The Single WPD-LSTM model receives four input features decomposed using the WPD. These features were structured into sequences to match the temporal nature of the LSTM. The length of the input sequence is 10 for the 15-min prediction, 10 for the 1-h prediction,

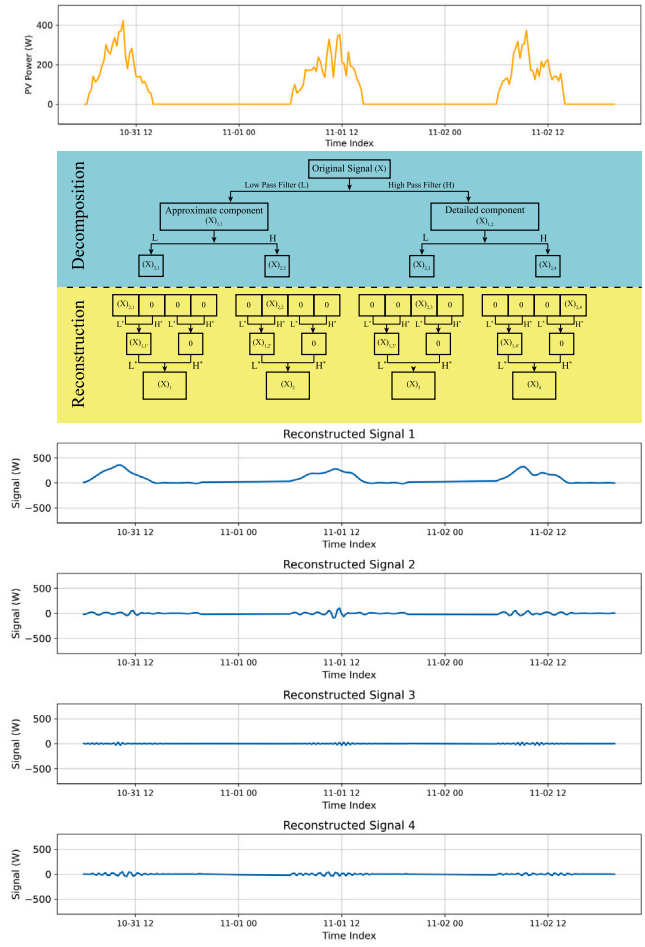


Fig. 3. Wavelet Decomposition and Reconstruction of PV Power Signal. The original PV power signal (top) is decomposed into multiple frequency sub-series using discrete wavelet transform (DWT), as illustrated in the decomposition-reconstruction diagram (middle). The lower plots show the reconstructed signals for each sub-series, capturing different frequency bands.

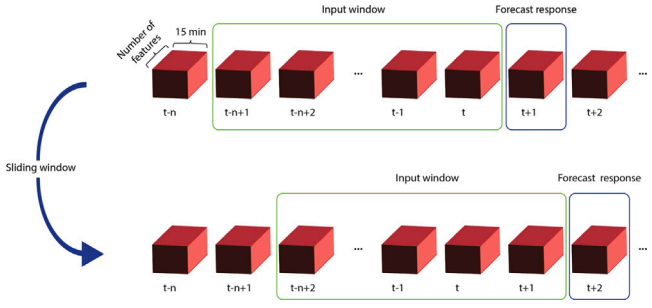


Fig. 4. Sliding window approach for supervised time series learning. The raw sequential data is transformed into input–output pairs using a sliding window mechanism. For each window, a fixed number of past observations (green boxes) are used as input features, and the subsequent data points (blue boxes) serve as prediction targets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and 112 for the day-ahead forecast. On the other hand, the Ensemble Multi WPD-LSTMs model employs a set of parallel LSTMs, each trained on different subseries of the decomposed signal. The outputs of the individual LSTMs are aggregated using a weighted sum to produce the final prediction. The single LSTM model has an inference time of 6 ms on a typical GPU, while the more complex multi WPD-LSTM model takes 35 ms. Since the forecast intervals are 15 min, both models are lightweight for training and suitable for real-time applications.

Both models were trained using a ratio of 0.75 train and 0.25 test. The architecture of each LSTM comprised one layer, with 50 units per layer, followed by two fully connected layers, one with 50 hidden neurons as the LSTM and a last one with number of neurons equal to the output sequence length. Training employed Adam optimization with an initial learning rate of 0.001, and mean squared error loss function, appropriate for regression. Regularization was achieved using a Dropout layer with a dropout rate of 0.5, applied after the LSTM layer to reduce overfitting by randomly setting 50% of the neurons to zero during training. The model was built with the Keras Python library [31] using the Sequential API. The training was conducted over 50 epochs. To determine the optimal configuration of the LSTM model, we applied a random search strategy over a predefined range of hyperparameters, including the number of LSTM units, dropout rate, learning rate, batch size, activation function, optimizer, and input sequence length. This approach was conducted independently for each forecast horizon to account for varying temporal patterns. The configuration presented in the results of this work is the one that yields the lowest RMSE on the test dataset for each horizon.

4.3. Performance evaluation

Experimental outcomes are evaluated and discussed to compare the effectiveness of conventional persistence model, LSTM, Multi WPD-LSTM without weight optimization, and Linear Regression (applied only for day-ahead predictions) against the proposed models in this research. Representative samples with different weather conditions, including partially cloudy, rainy, and sunny, are selected for evaluation. The model is tested with 3 prediction horizons: 15 min ahead, hour ahead, and day ahead. The forecasting errors are presented in the following sections.

4.3.1. Short-term predictions

Fig. 5 presents the results of short-term PV forecasting for three different weather conditions: cloudy, rainy, and sunny. In the 15-min-ahead predictions (5(a)), the proposed models effectively capture the general trends of PV power output. The predictions remain smooth under different weather conditions, including cloudy and rainy days. It can be seen that models that use decomposed input signals (single and multi WPD-LSTM) show better responsiveness to sudden weather changes. For the hour-ahead predictions (5(b)), the multi WPD-LSTM model achieves higher accuracy, particularly in capturing peak patterns. This is due to dedicating different models for processing sub-series and aggregating the information learned by all models.

4.3.2. Day-ahead predictions

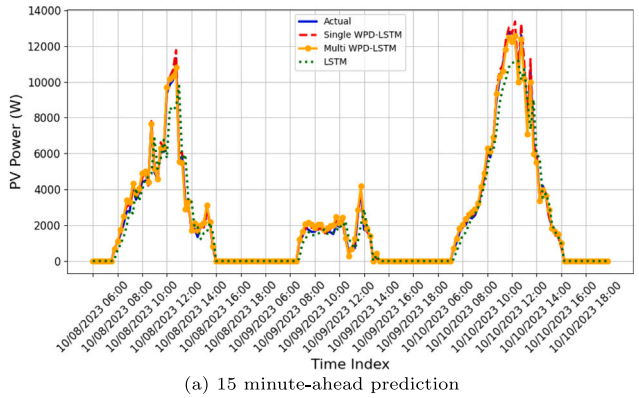
As shown in Fig. 6, day-ahead forecasting is challenging, particularly under the unpredictable weather conditions in the Netherlands. Although all predictions include some errors, WPD models show improved performance. Moreover, the multi WPD-LSTM model achieves the highest accuracy due to its higher generalization capability.

4.4. Performance comparison

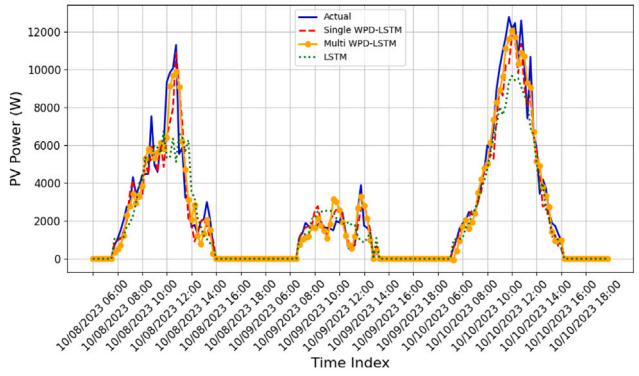
Table 1 presents the forecasting performance of different models across 15-min, 1-h, and day-ahead horizons using nRMSE, MAPE, and R^2 as evaluation metrics. The proposed Multi WPD-LSTM model with weight optimization consistently delivers the most accurate forecasts, achieving the lowest nRMSE (2.67%, 7.31%, and 12.93%), the lowest MAPE (5.37%, 13.09%, and 96.30%), and the highest R^2 scores (0.99, 0.96, and 0.63) across all three horizons. The Single WPD-LSTM model remains competitive, especially for short-term forecasting, with a 3.71% nRMSE and 4.81% MAPE at the 15-min horizon, but its performance lowers more rapidly than the multi-model version as the

Table 1
Forecasting errors for 15-min, hour, and day-ahead predictions.

Method	15-min ahead			Hour-ahead			Day-ahead		
	nRMSE	MAPE	R^2	nRMSE	MAPE	R^2	nRMSE	MAPE	R^2
Persistence	9.70%	30.05%	0.90	15.52%	73.72%	0.59	19.81%	125.99%	−0.07
LSTM	8.51%	14.56%	0.92	11.44%	20.19%	0.86	13.98%	96.31%	0.42
Single WPD-LSTM	3.71%	4.81%	0.99	8.88%	11.03%	0.95	13.80%	99.73%	0.46
Multi WPD-LSTM without weight optimization	3.07%	32.12%	0.91	8.02%	28.31%	0.82	13.11%	138.47%	0.39
Multi WPD-LSTM with weight optimization	**2.67%**	**5.37%**	**0.99**	**7.31%**	**13.09%**	**0.96**	**12.93%**	**96.30%**	**0.63**



(a) 15 minute-ahead prediction



(b) Hour-ahead prediction

Fig. 5. Short-term PV power forecasting under different weather conditions: cloudy (left), rainy (middle), and sunny (right).

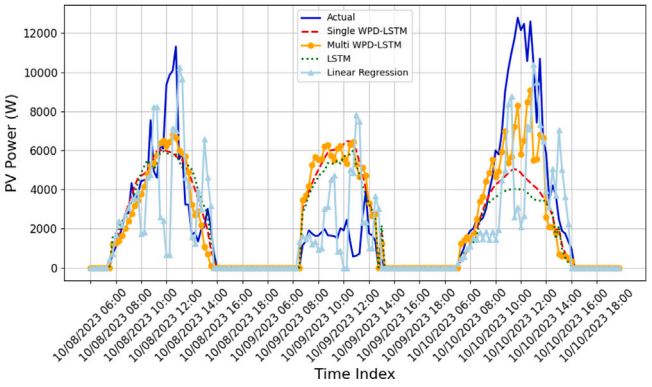


Fig. 6. Long-term PV power forecasting for different weather conditions: cloudy (left), rainy (middle), and sunny (right).

forecast horizon extends. The comparison between the optimized and non-optimized Multi WPD-LSTM models clearly shows the benefits of weight optimization, as the non-optimized variant performs worse in all metrics (most notably in MAPE, where it reaches 32.12% for 15-min

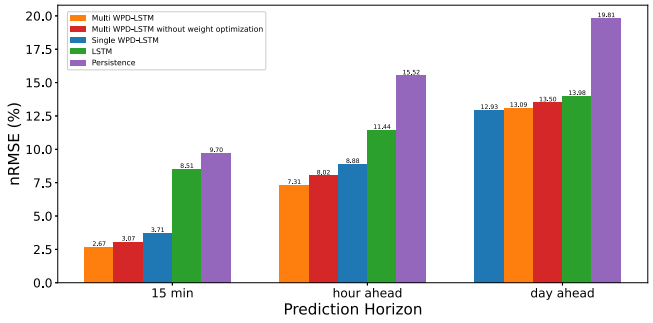


Fig. 7. Comparison of nRMSE across forecasting models and prediction horizons.

and 138.47% for day-ahead predictions). Traditional LSTM and persistence models have poor performances in comparison, particularly for longer horizons; for example, persistence yields a negative R^2 (−0.07) and extremely high MAPE (125.99%) for day-ahead forecasts, while LSTM shows better but still inferior performance. This performance trend is further illustrated in Fig. 7, highlighting the superiority of the proposed Multi WPD-LSTM model, which outperforms traditional methods across all horizons.

5. Conclusion

This study addresses the challenges of solar power intermittency, which impact the stability, safety, and economic efficiency of power systems. A hybrid approach, combining WPD and LSTM models, is proposed to improve the accuracy of PV power output forecasting across three different horizons. The results showed that the proposed WPD-LSTM models outperform the benchmark models (persistence, LR, and LSTM), particularly for short-term forecasts, i.e., 15-min and 1-h ahead predictions. The effectiveness of the model is most evident when each of the decomposed-reconstructed PV power series is forecasted independently and then combined using a weight optimization approach, resulting in lower error values. However, the inclusion of decomposed subseries resulted in only marginal improvements for day-ahead forecasting, as anticipated, given the complexities of making accurate long-term predictions in PV systems.

Notably, standalone LSTM models failed to achieve the level of accuracy provided by the proposed WPD-LSTM models. This underscores the critical role of WPD in overcoming the limitations of traditional LSTM models, particularly in short-term forecasting. For day-ahead predictions, taking in more historical data improves model robustness and its ability to capture long-term trends. However, for 15-min ahead forecasting, where immediate environmental conditions play a larger role, relying solely on recent data proved more effective than using longer historical patterns.

Future work could explore the integration of attention mechanisms into the LSTM architecture to enhance the ability of the model to focus on the most relevant time steps or features, particularly under variable weather conditions. Additionally, extending the forecasting horizon beyond day-ahead to cover multi-day or weekly predictions would provide further value for grid operations. To improve predictive accuracy and model robustness, more advanced hyperparameter

optimization techniques, such as Bayesian optimization or automated tuning frameworks, could be applied. Incorporating more meteorological inputs would also better performance. Finally, implementing online learning or adaptive retraining strategies would allow the model to adjust continuously to seasonal shifts and changing environmental conditions without requiring full retraining.

CRediT authorship contribution statement

Amirhasan Sardarabadi: Writing – original draft, Visualization, Validation, Software, Methodology, Conceptualization. **Amirhossein Heydarian Ardakani:** Writing – original draft, Validation, Methodology. **Silvana Matrone:** Writing – original draft, Software, Resources, Methodology. **Emanuele Ogliari:** Writing – review & editing, Visualization, Supervision, Data curation, Conceptualization. **Elham Shirazi:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research utilizes data from the SlimPark Living Lab, located on the University of Twente campus in Enschede.

Data availability

The data that has been used is confidential.

References

[1] Citaristi I. International energy agency—iea. In: The Europa directory of international organizations 2022. Routledge; 2022, p. 701–2.
[2] Pavan AM, Ogliari E, Leva S, Lughi V. Advanced methods for photovoltaic output power forecasting: A review. Appl Sci 2020;10(2):487.
[3] Bošnjaković M, Santa R, Crnac Z, Bošnjaković T. Environmental impact of PV power systems. Sustain 2023;15(15).
[4] Yang D, Wang W, Gueymard CA, Hong T, Kleissl J, Huang J, Perez MJ, Perez R, Bright JM, Xia X, van der Meer D, Peters IM. A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality. Renew Sustain Energy Rev 2022;161:112348.
[5] Bazionis IK, Kousounadis-Knousen MA, Georgilakis PS, Shirazi E, Soudris D, Catthoor F. A taxonomy of short-term solar power forecasting: Classifications focused on climatic conditions and input data. IET Renew Power Gener 2023;17:2411–32.
[6] Shirazi E, Sark W, Reinders A. Deployment of solar photovoltaic systems in distribution grids in combination with batteries. In: Photovoltaic solar energy: From fundamentals to applications. vol. 2, 2024, p. 435–54.
[7] Jakoplić A, Franković D, Kirinčić V, Plavšić T. Benefits of short-term photovoltaic power production forecasting to the power system. Optim Eng 2021;22:49144–56.
[8] Iheanetu KJ. Solar photovoltaic power forecasting: A review. Sustain 2022;14(24).
[9] Alcañiz A, Grzebyk D, Ziar H, Isabella O. Trends and gaps in photovoltaic power forecasting with machine learning. Energy Rep 2023;9:447–71.
[10] Shirazi E, Gordon I, Reinders A, Catthoor F. Sky images for short-term solar irradiance forecast: A comparative study of linear machine learning models. IEEE J Photovolt 2024;14(4):691–8.
[11] Ogliari E, Sakwa M, Cusa P. Enhanced Convolutional Neural Network for solar radiation nowcasting: All-Sky camera infrared images embedded with exogeneous parameters. Renew Energy 2024;221:119735.
[12] Shirazi E, Ozkalay E, Reinders A. PV DC yield determined by deep neural networks: the case of building integrated PV. In: Proceedings of 8th world conference on photovoltaic energy conversion. 8th world conference on photovoltaic energy conversion, vol. 780-783, EU PVSEC; 2022, p. 780–4. http://dx.doi.org/10.4229/WCPEC-82022-3BV.3.61.
[13] Raghoebarsing A, Farkas I, Atsu D, Boddaert S, Moser D, Shirazi E, Reinders A. The status of implementation of photovoltaic systems and its influencing factors in European countries. Prog Photovolt, Res Appl 2023;31:113–33.
[14] Watari D, Taniguchi I, Goverde H, Manganiello P, Shirazi E, Catthoor F, Onoye T. Multi-time scale energy management framework for smart PV systems mixing fast and slow dynamics. Appl Energy 2021;289:116671.
[15] Nayyef ZT, Abdulrahman MM, Kurdi NA. Optimizing energy efficiency in smart grids using machine learning algorithms: A case study in electrical engineering. SHIFRA 2024;2024:46–54.
[16] Konstantinou M, Peratikou S, Charalambides AG. Solar photovoltaic forecasting of power output using LSTM networks. Atmosphere 2021;12(1).
[17] Son N, Jung M. Analysis of meteorological factor multivariate models for medium- and long-term photovoltaic solar power forecasting using long short-term memory. Appl Sci 2020;11:316.
[18] Li P, Zhou K, Lu X, Yang S. A hybrid deep learning model for short-term PV power forecasting. Appl Energy 2020;259:114216.
[19] Hanifi S, Zare-Behtash H, Cammarano A, Lotfian S. Offshore wind power forecasting based on WPD and optimised deep learning methods. Renew Energy 2023;218:119241.
[20] Ferkous K, Guermoui M, Bellaour A, boulmaiz T, Bailek N. Enhancing photovoltaic energy forecasting: a progressive approach using wavelet packet decomposition. Clean Energy 2024;8(3):95–108.
[21] Li G, Xie S, Wang B, Jiantao X, Li Y, Du S. Photovoltaic power forecasting with a hybrid deep learning approach. IEEE Access 2020;8:175871–80.
[22] Zhou H, Zhang Y, Yang L, Liu Q, Yan K, Du Y. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. IEEE Access 2019;7:78063–74.
[23] Matrone S, Ogliari E, Nespoli A, Leva S. Electric vehicle supply equipment day-ahead power forecast based on deep learning and the attention mechanism. IEEE Trans Intell Transp Syst 2024.
[24] Dhaked DK, Dadhich S, Birla D. Power output forecasting of solar photovoltaic plant using LSTM. Green Energy Intell Transp 2023;2(5):100113.
[25] Alshafeey M, Csáki C. Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods. Energy Rep 2021;7:7601–14.
[26] Van Houdt G, Mosquera C, Nápoles G. A review on the long short-term memory model. Artif Intell Rev 2020;53.
[27] Gokhale M, Khanduja DK, et al. Time domain signal analysis using wavelet packet decomposition approach. Int'L J Commun Netw Syst Sci 2010;3(03):321.
[28] Ozaki Y, Yano M, Onishi M. Effective hyperparameter optimization using Nelder-Mead method in deep learning. IPSJ Trans Comput Vis Appl 2017;9:20.
[29] Alshafeey M, Csáki C. The impact of input data resolution on neural network forecasting models for wind and photovoltaic energy generation using time series data. Environ Prog Sustain Energy 2022;42.
[30] Venkateswari R, Sreejith S. Factors influencing the efficiency of photovoltaic system. Renew Sustain Energy Rev 2019;101:376–94.
[31] Chollet F, et al. Keras. 2015, URL https://github.com/fchollet/keras.