ISSN: 2710-0588

Adaptive Non-Stationary PV Power Forecasting with ADWF-LSTM Networks

Yibo Wu*

School of Software, Henan Polytechnic University, Jiaozuo, China

*Corresponding Author: 212309020042@home.hpu.edu.cn

Abstract

Photovoltaic (PV) power data exhibits strong non-stationarity due to solar irradiance fluctuations, diurnal temperature variations, and operating condition changes. Traditional Discrete Wavelet Transform (DWT)-LSTM models are limited by fixed wavelet bases and static fusion strategies, failing to effectively capture dynamic multi-scale features. To address this, we propose an Adaptive Dynamic Wavelet Fusion-LSTM (ADWF-LSTM) model, which integrates an Adaptive Dynamic Wavelet Fusion (ADWF) module with LSTM. The ADWF module dynamically optimizes wavelet basis parameters and multi-scale fusion weights in a data-driven manner, while LSTM retains its strength in modeling long-term temporal dependencies. Extensive experiments on the real-world Yulara PV dataset across four seasons show that ADWF-LSTM outperforms baseline models (LSTM, DWT-LSTM, TCN-LSTM, Informer). Its average MAE, MSE, and RMSE reach 0.4669, 0.5451, and 0.7275, reducing errors by over 10% on average. With an MBE of 0.0347, the model achieves optimal bias control. Visualization results confirm its superior trend-tracking and anti-interference performance under varying seasonal data characteristics.

Keywords

Photovoltaic Power Prediction, Adaptive Wavelet Fusion, Non-stationary Prediction, LSTM.

1. Introduction

The intensifying global energy crisis and environmental pollution have positioned photovoltaic (PV) power generation as a cornerstone of the low-carbon energy transition [1]. While offering advantages like zero emissions and flexible deployment , PV output is highly volatile due to factors like irradiance and cloud cover, challenging grid stability [2]. Accurate short-term forecasting is therefore crucial for grid dispatch and renewable integration [3].

A key forecasting challenge is the non-stationary, multi-scale nature of PV power data, which exhibits intertwined long-term trends, daily cycles, and short-term fluctuations [4]. This demands models with strong multi-scale feature extraction [5]. Deep learning methods, particularly LSTM networks, are widely used for their ability to model temporal dependencies. However, LSTMs struggle to explicitly capture multi-scale features in highly non-stationary data [6].

To address this, hybrid models like DWT-LSTM combine Discrete Wavelet Transform (DWT) with LSTM. DWT decomposes the power signal into multi-scale components, improving upon standalone LSTM [7]. However, standard DWT-LSTM has two main flaws: it relies on a single, fixed wavelet basis unsuitable for all dynamic conditions, and it uses static fusion, ignoring dynamic inter-scale relationships [8].

This paper proposes an Adaptive Dynamic Wavelet Fusion-LSTM (ADWF-LSTM) model. Its core innovation is an ADWF module that dynamically selects and weights multiple complementary

ISSN: 2710-0588

wavelet bases and adaptively fuses multi-scale features. Integrated with LSTM, this enhances multi-scale feature extraction from non-stationary data while preserving LSTM's temporal modeling strengths.

The main contributions of this paper are summarized as follows:

1) The ADWF module overcomes the fixed-basis and static-fusion limitations of DWT-LSTM.

2) The ADWF-LSTM framework synergizes adaptive time-frequency analysis with temporal modeling.

3) Experiments on real PV data show the model outperforms benchmarks in accuracy and robustness.

The remainder of this paper is organized as follows: Section 2 reviews related work on PV power prediction models. Section 3 details the proposed ADWF-LSTM model, including the design of the ADWF module and the integration logic with LSTM. Section 4 presents the experimental setup, results, and analysis. Section 5 concludes the paper and discusses future work.

2. Relate Work

2.1. The Statistical Methods

The core objective of statistical approaches is to define a functional relationship that maps historical data to output power. Principal methodologies encompass time-series analysis [9], Markov Chain (MC) [10], Auto-Regressive Integrated Moving Average (ARIMA) [11], regression analysis [12], and Seasonal Auto-Regressive Integrated Moving Average (SARIMA) [13]. Dadi et al. [14] developed a piezoelectric energy harvesting system incorporated into automobile wheels, featuring sensors strategically positioned around the wheel's circumference, optimized for resonant frequency and impedance; comprehensive analysis and simulations confirmed its feasibility for power generation at typical Indian road speeds (60 km/h and 80 km/h).

2.2. The Physical Methods

Physical methodologies forecast output power by utilizing meteorological data-including solar irradiance, cloud cover, temperature, humidity, and wind speed-obtained chiefly from weather stations. These techniques construct predictive models by analyzing the relationships between meteorological parameters and power generation data. Pierro et al. [15] introduced a physics-based prediction model, though its somewhat intricate architecture led to constrained predictive accuracy. In a different approach, Lee et al. [16] presented a confidence-aware Bayesian deep learning framework designed specifically for solar irradiance forecasting. Alonso-Montesinos et al. [17] utilized an innovative sky camera system; they transformed horizontally captured digital images into irradiance data and then applied the maximum cross-correlation method for predictions. Geraldi et al. [18] improved forecast accuracy by creating a sophisticated solar radiation estimation technique that integrates additional sensors. Perez et al. [19] conducted a comprehensive assessment of multiple numerical weather prediction (NWP) models across various sites in the United States, Europe, and Canada.

2.3. The Artificial Intelligence Methods

Artificial intelligence (AI) learning approaches utilize the rapidly growing computational capabilities of AI algorithms to learn complex mappings between inputs and outputs, with a particular emphasis on capturing nonlinear relationships. Bamisile et al. [20] developed ANN models (trained via Levenberg-Marquardt backpropagation) for hourly solar irradiance and PV parameter forecasting in 6 Nigerian locations, achieving R values of 0.9046–0.9777 (irradiance) and 0.7768–0.8739 (PV multi-parameters). Zhang et al. [21] proposed a hybrid AP-CEEMDAN-BiLSTM method to solve the low ultra-short-term PV power forecasting accuracy under

ISSN: 2710-0588

changeable weather (due to strong randomness, uncertainty, and non-stationarity), achieving a MAPE of 2.771% on the Alice Springs dataset, which outperformed BP, single BiLSTM, and CEEMDAN-BiLSTM models. Lyu et al. [22] proposed a data-driven GRU-KAN framework for short-term PV forecasting (addressing reduced generation from extreme weather), which uses Grey Relational Analysis for feature extraction, K-Means++ for normal weather clustering, similar-day selection for extreme weather training sets, and simplified feature layer neurons; it achieved average RMSE 3.86/R^2 96.90% under normal conditions and low RMSEs (0.900–1.438) on extreme days. El Aouni et al. [23] proposed hybrid methods combining frequency-domain techniques (DWT/STFT) with LSTM, Transformers, or ANN to address the need for accurate PV power forecasting (especially 1-hour to 3-day horizons) in microgrids under varying weather conditions, with the DWT-LSTM (1-hour, HH≈0.16) and LSTM-STFT-ANN (3-day, HH=0.405) outperforming CNN-LSTM and standalone Transformers. Guo et al. [24] proposed a hybrid Attention-TCN-LSTM model optimized by the PSO algorithm for short-term PV power forecasting, integrating LSTM's sequence capture ability and Attention-TCN's meteorological feature extraction to enhance accuracy and reliability, which outperformed comparative models and facilitated renewable energy integration. Ospina et al.[25] proposed a novel model combining stationary wavelet transform (SWT) with four LSTMs and a final DNN for PV power forecasting (both large and small-scale systems), using SWT-decomposed coefficients, temperature data, and statistical features, which enhanced accuracy and reduced required features compared to existing models. Deng et al. [26] proposed a DMGformer (decoupled Informer with multi-moment guidance) for day-ahead hourly distributed PV power forecasting (addressing challenges in obtaining high-fidelity weather data), which uses a decoupled history-forecast (DHF) structure and an MMG module, achieved 24.11% MAE reduction and 1.46% accuracy improvement over the suboptimal Informer on a 500-site real dataset, with DHF and MMG enhancing LSTM/Transformer/Informer performance.

3. Method

3.1. ADWF-LSTM

The ADWF-LSTM model integrates the Adaptive Dynamic Wavelet Fusion (ADWF) module with the Long Short-Term Memory (LSTM) network, synergistically leveraging multi-scale feature extraction and long-term temporal dependency modeling to address the strong non-stationarity of PV power data. The overall architecture of the model is illustrated in **Fig. 1**.

Given the input PV power time series $x \epsilon R^{B \times T \times F}$ (where B = batch size, T = sequence length, F = feature dimension), the core target feature $x_{target} \epsilon R^{B \times T}$ is first fed into the ADWF module. Through fixed wavelet decomposition, data-driven weight generation, adaptive multi-wavelet fusion, and convolution-based enhancement, the ADWF module outputs the enhanced wavelet feature tensor $\mathrm{F}_{enhanced}$ to T_{seq} (LSTM input sequence length):

$$x_{pool} = AdaptiveAvgPool1d\left(x_{proj}^{perm}, T_{seq}\right)$$
$$\mathrm{F}_{pool} = daptiveAvgPool1d\left(\mathrm{F}_{enhanced}, T_{seq}\right)$$

The length-unified raw and wavelet features are concatenated to form the LSTM input:

$$X_{LSTM} = Concat(x_{pool}, \mathrm{F}_{pool}) \epsilon R^{B \times T_{seq} \times 2D}$$

ISSN: 2710-0588

Finally, X_{LSTM} is fed into the LSTM network to capture long-term temporal dependencies. The hidden state of the last time step is input to a linear output layer, generating the final prediction $y \epsilon R^{B \times T}$ (where P = prediction sequence length).

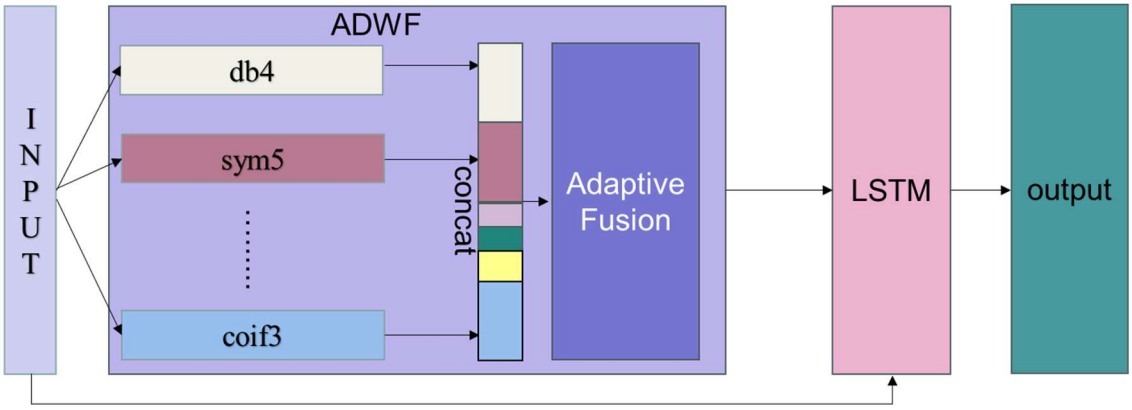

Fig. 1 Architecture diagram of the ADWF-LSTM model

3.2. Adaptive Dynamic Wavelet Fusion

3.2.1. Fixed Wavelet Basis Decomposition

Input time series are decomposed into low-frequency approximate components and high-frequency detail components using fixed wavelet filter banks, formulated as:

$$\begin{cases} cA_t = \sum_{k=0}^{L-1} h_k \cdot x_{2t-k+p} \\ cD_t = \sum_{k=0}^{L-1} g_k \cdot x_{2t-k+p} \end{cases}$$

Where $x \epsilon R^{B \times T}$ is the input time series batch with B as batch size and T as sequence length; cA and cD denote approximate (low-frequency) and detail (high-frequency) coefficients, respectively; h and g are fixed low-pass and high-pass wavelet filter coefficients; L is the filter length; $p = [(L-1)/2]$ is reflection padding to alleviate boundary effects; and T' is the coefficient length after downsampling with stride 2.

To ensure dimensional consistency for subsequent fusion, adaptive average pooling is applied to unify the length of decomposed coefficients to a fixed value T_{pool}:

$$cA_{pool} = AdaptiveAvgPool1d(cA, T_{pool})$$
$$cD_{pool} = daptiveAvgPool1d(cD, T_{pool})$$

Where $\{cA_{pool}^k, cD_{pool}^k\}_{k=1}^K$ represents the pooled coefficients of K predefined wavelet bases.

3.2.2. Data-Driven Adaptive Weight Generation

Normalized fusion weights for the K wavelet bases are generated by a lightweight convolutional neural network, formulated as:

$$w = Softmax\left(Conv1d_2\left(AdaptiveAvgPool1d\left(GELU\left(Conv1d_1(x_{unsq})\right), 1\right)\right)\right)$$

ISSN: 2710-0588

Where $x_{unsq} \epsilon R^{B\times1\times T}$ is the input sequence with an extra channel dimension; $Conv1d_1$ and $Conv1d_2$ are 1D convolution layers for feature extraction and weight mapping; $GELU$ introduces non-linear transformations; $AdaptiveAvgPool1d(\cdot,1)$ captures global sequence features; and $w = [w_1, w_2, \dots, w_k] \epsilon R^{B\times K\times1}$ are normalized weights satisfying $\sum_{k=1}^{K} w_k = 1$.

3.2.3. Adaptive Multi-Wavelet Fusion

Pooled approximate and detail coefficients of each wavelet basis are concatenated to integrate low- and high-frequency information:

$$C^k = Concat(cA^k_{pool} \backslash unsqueeze(1), cD^k_{pool} \backslash unsqueeze(1))$$

Where $C^k \epsilon R^{B\times2\times T_{pool}}$ is the 2-channel feature tensor of the k-th wavelet basis.

Adaptive fusion of multi-wavelet features is realized via weighted summation:

$$C_{fused} = \sum_{k=1}^{K} w_k \cdot C^k$$

Where $C_{fused} \epsilon R^{B\times2\times T_{pool}}$ is the final fused wavelet feature tensor aggregating multi-scale temporal patterns from diverse wavelet bases.

3.2.4. Feature Enhancement

To improve the expressive power of fused features, a convolution-based enhancement block is applied:

$$\mathrm{F}_{enhanced} = Dropout(GELU(BatchNorm1d(Conv1d(C_{fused}, D))))$$

Where $Conv1d$ projects the 2-channel fused features to the target dimension D; $BatchNorm1d$ stabilizes the training process; and $\mathrm{F}_{enhanced} \epsilon R^{B\times D\times T_{pool}}$ is the enhanced feature tensor.

3.3. LSTM

Long Short-Term Memory (LSTM) networks constitute a specialized variant of Recurrent Neural Networks (RNNs) developed to overcome the long-term dependency limitation inherent in standard RNNs. In contrast to simple RNN architectures, which are often constrained by vanishing or exploding gradients-leading to the degradation of information over lengthy sequences-LSTMs incorporate a refined gating mechanism. This structure selectively retains, updates, and outputs information, allowing gradients to propagate more effectively across many time steps. As a result, LSTMs excel at capturing temporal relationships across extended sequences, establishing their utility in domains such as time-series forecasting, language modeling, and sequential data analysis.

Central to the LSTM network's architecture is the cell state ($\tilde{C}_t$), frequently analogized to an information "highway" that propagates through the entire sequence. In contrast to the hidden state (h_t) of standard RNNs-whose primary function is to encode short-term contextual information-the cell state serves as a long-term memory repository, preserving knowledge accrued across extended time steps. This cell state is governed by three mutually interacting gates: the forget gate, input gate, and output gate. Each gate is specifically engineered to modulate the flow of information-controlling its entry into, exit from, and internal circulation within the cell state. The structural configuration of the LSTM is illustrated in **Fig. 2**.

ISSN: 2710-0588

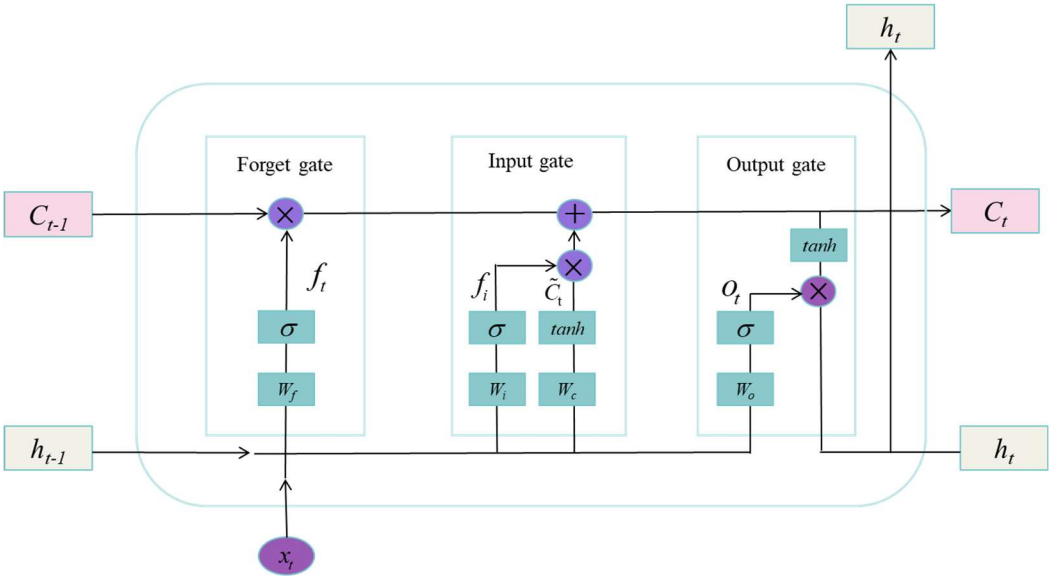

Fig. 2 LSTM Network Architecture Diagram

The cell state C_t serves as an information conduit for transporting data across time steps. The gates modulate this flow in the following way: the forget gate selects which information from the prior cell state to discard, while the input gate identifies new information for incorporation and generates a candidate cell state to enable potential updates.

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big),\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i),$$
$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

Here, σ denotes the sigmoid activation function, h_{t-1} r refers to the hidden state from the preceding time step, and x_t represents the input at the current time step. W_f, W_i, W_c are the weight matrices corresponding to the forget gate, input gate, and candidate cell state, respectively, while b_f, b_i, b_c denote the corresponding bias vectors. i_t signifies the input gate output, and C_{t-1} is the cell state from the previous time step. The cell state undergoes an update process that fuses the output signals of the forget gate and input gate, as follows:

$$C_t = f_t \odot C_{t-1} + \ i_t \odot \tilde{C}_t$$

Here, $\odot$ indicates element-wise multiplication. Finally, the output gate produces the final hidden state based on the updated cell state:

$$o_t = \ \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), h_t = \ o_t \odot \tanh(C_t)$$

In these formulations, o_t denotes the output vector of the LSTM's output gate at the current time step t, and h_t represents the hidden state of the LSTM at time step t. W_o and b_o correspond to the weight matrix and bias vector for the output gate, respectively. This gating framework endows LSTMs with the capability to capture long-term temporal dependencies, while simultaneously alleviating the vanishing gradient problem that is intrinsic to traditional recurrent neural networks (RNNs). This unique advantage renders LSTMs highly effective for time series forecasting tasks—including photovoltaic (PV) power prediction, a key application scenario in our study.

ISSN: 2710-0588

4. Experiments

4.1. Experimental Setup

All experiments involved in this study were conducted on a personal computer using Python 3.11.7 and a 64-bit operating system with an Intel(R) Xeon(R) Gold 5218 CPU @ 2.30GHz, 64.0 GB RAM. The Huber loss function was adopted for model training because it combines the advantages of mean squared error (MSE) and mean absolute error (MAE) - it is quadratic for small errors (ensuring smooth optimization and fast convergence) and linear for large errors (being robust to outliers, which are common in PV power data due to sudden changes in weather conditions such as cloud cover or irradiance fluctuations).

Table 1. Hyperparameter Settings of the ADWF-LSTM Model

Hyperparameter	**Value**
Number of LSTM layers	2
Input size	9
Convolution kernel sizes	7
Convolution output channels	64
Sequence length	60
LSTM hidden size	128
Loss function	HuberLoss
Batch size	32
Optimizer	Adam

4.2. Data Description

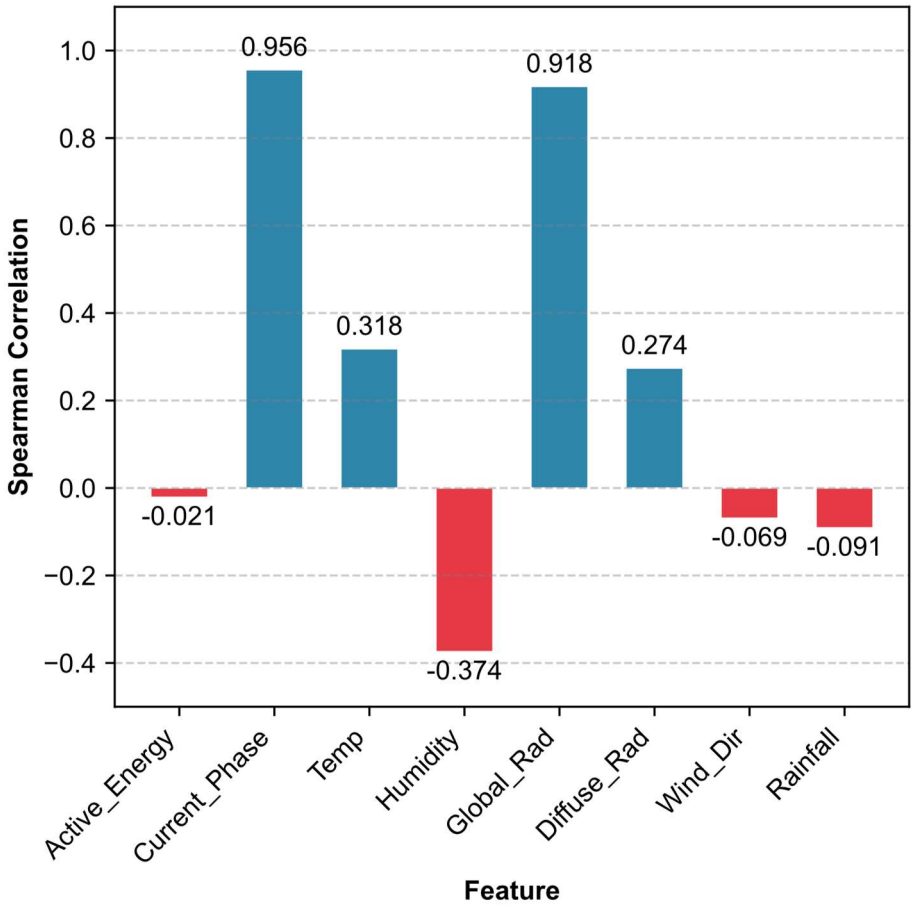

Fig. 3 Spearman Correlation Heatmap of Meteorological Features and Active Power

ISSN: 2710-0588

The dataset spans from 2013 to 2019, collected at 5-minute intervals through instrumentation specifications detailed in **Table 1**.For model development, data from 2017 to 2018 was allocated for training, while 2019 data served as the test set. Predictive variables include active power (kW), weather temperature (°C), relative humidity (%), global horizontal irradiance (W/m^2), diffuse horizontal irradiance (W/m^2), and wind direction (degrees). **Fig. 3** illustrates the Spearman rank correlation coefficients between Active_Power and meteorological variables using a bar plot derived from the 2017 subset. The analysis reveals that global horizontal irradiance exhibits the strongest positive correlation with Active_Power (ρ=0.86), aligning with photovoltaic energy conversion fundamentals. Temperatureanddiffuse horizontal irradiance show moderate positive correlations (ρ=0.29 and 0.22, respectively), while wind direction (ρ=-0.10) and daily rainfall (ρ=-0.09) display negligible associations. These results directly informed the GL-LSTM architecture: strongly correlated variables (global/diffuse irradiance, temperature) were prioritized as model inputs, whereas weakly correlated parameters (wind direction, rainfall) were excluded to minimize noise.

Table 2. Technical specification of the power plant at Yulara Solar System, Australia

Parameter	Value
Array Rating	10.5kW
Panel Rating	175W
Number Of Panels	2×30
Panel Type	Trina TSM-175DC01
Array Area	2×38.37 m^2
Inverter Size / Type	2×6kW,SMA SMC 6000A
Installation Completed	Thu, 8 Jan 2009
Array Tilt/Azimuth	Variable:Dual axis tracking

4.3. Data Preprocessing

Handling missing data and feature scaling are crucial preprocessing steps for time-series modeling. In this study, linear interpolation was applied to impute 26,810 missing values in the Radiation_Global_Tilted and Radiation_Diffuse_Tilted features, preserving temporal continuity and avoiding the biases associated with row deletion. Additionally, to address the significant scale disparity between features such as Active_Power (0–23 kW) and Wind_Direction (0–5530 degrees), which can destabilize gradient-based optimization and hinder convergence, Min-Max normalization was employed to rescale all features to the unit interval [0, 1].

$$x^{norm} = \frac{x - x^{min}}{x^{max} - x^{min}}$$

Here, x denotes an original data point, x^{min} and x^{max} represent the minimum and maximum values of the feature in question, and x^{norm} is the resulting normalized data point. This normalization approach scales all features to a consistent range, which in turn optimizes the model's training dynamics and enhances the accuracy of its prediction outputs.

4.4. Experimental Metrics

The performance of the proposed algorithm was assessed using a suite of standard error metrics: Mean Bias Error (MBE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).Each metric provides distinct insight into the nature of the prediction errors. MBE measures the average systematic bias, indicating whether the model consistently over- or under-predicts. MAE gives the average magnitude of errors, offering a

ISSN: 2710-0588

direct and robust interpretation of average accuracy. MSE, by squaring the errors, penalizes larger discrepancies more severely, making it sensitive to outliers. RMSE, derived as the square root of MSE, shares the same units as the target variable and is interpreted as the standard deviation of the unexplained variance; a lower RMSE signifies superior predictive precision and lower error dispersion.The formulas for calculating these metrics are provided below.

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i), MAE = \frac{1}{N}\sum_{i=1}^{N}|\tilde{y}_i - y_i|$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i)^2, RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\tilde{y}_i - y_i)^2}$$

where y_i represents the true value and $\tilde{y}_i$ represents the predicted value, with N being the total number of y_i.

4.5. Experimental Results and Analysis

Table 3. Performance metrics of different methods in different seasons

Season	Models	MBE	MAE	MSE	RMSE
Spring	*ADWF-LSTM*	*0.0347*	*0.4373*	*0.5425*	*0.7365*
	LSTM	0.1121	0.5968	0.6943	0.8332
	DWT-LSTM	-0.0978	0.6777	0.8763	0.9361
	TCN-LSTM	0.0605	0.5632	0.6349	0.7968
	Informer	-0.1074	0.5762	0.6782	0.8235
Summer	*ADWF-LSTM*	*-0.0204*	*0.3487*	*0.2741*	*0.5235*
	LSTM	0.0876	0.5794	0.5197	0.7209
	DWT-LSTM	0.0442	0.4902	0.3768	0.6138
	TCN-LSTM	-0.0968	0.3894	0.2903	0.5387
	Informer	-0.0644	0.4254	0.3164	0.5624
Autumn	*ADWF-LSTM*	*0.0832*	*0.5884*	*0.6241*	*0.79*
	LSTM	-0.0941	0.6331	0.6952	0.8337
	DWT-LSTM	-0.1087	0.6474	0.7616	0.8727
	TCN-LSTM	-0.1143	0.7768	0.9735	0.9866
	Informer	-0.1125	0.6365	0.7095	0.8423
Winter	*ADWF-LSTM*	*0.0396*	*0.4932*	*0.7397*	*0.86*
	LSTM	-0.067	0.656	0.9103	0.9541
	DWT-LSTM	-0.0643	0.6321	0.865	0.93
	TCN-LSTM	-0.0851	0.5601	0.7792	0.8827
	Informer	-0.0821	0.562	0.7831	0.8827
Average	*ADWF-LSTM*	*0.0347*	*0.4669*	*0.5451*	*0.7275*
	LSTM	-0.0127	0.6163	0.7048	0.8355
	DWT-LSTM	-0.0567	0.6118	0.7199	0.8381
	TCN-LSTM	-0.0898	0.5708	0.6695	0.8012
	Informer	-0.0892	0.5723	0.6694	0.8012

Note: The "Average" in the table represents the average values of the four seasons' metrics, not the prediction results for the entire year.

The Yulara Datase experimental results, as detailed in **Table 3**, demonstrate the superior performance of the proposed ADWF-LSTM model across all four seasons and annual averages when compared to baseline models, including TCN-LSTM, LSTM, DWT-LSTM, and Informer. The

ISSN: 2710-0588

metrics employed-MBE, MAE, MSE, and RMSE-collectively validate the model's robustness in capturing spatiotemporal dynamics and mitigating prediction biases.

The experimental results based on the Yulara dataset (**Table 3**, **Figs. 4–7**) confirm that the proposed ADWF-LSTM model outperforms baseline models (TCN-LSTM, LSTM, DWT-LSTM, Informer) across all four seasons and annual averages in terms of MBE, MAE, MSE, and RMSE, validating its superior prediction accuracy and robustness.

In spring, ADWF-LSTM achieves MBE=0.0347, MAE=0.4373, MSE=0.5425, and RMSE=0.7365, outperforming all comparative models. As shown in **Fig. 4**, its predicted curve closely tracks the actual power output with minimal deviations during mild fluctuations. Spring PV data features gentle volatility and distinct periodicity, which the ADWF module's adaptive wavelet fusion effectively captures, while DWT-LSTM's limited temporal dependency modeling leads to the highest RMSE (0.9361).

Summer delivers the optimal overall prediction results, with ADWF-LSTM exhibiting the strongest advantage: MAE=0.3487, MSE=0.2741, RMSE=0.5235—10.4% and 5.6% lower in MAE and MSE than the second-best TCN-LSTM. **Fig. 5** illustrates that ADWF-LSTM's predicted curve aligns quickly with true values even during short-term irradiance fluctuations, benefiting from summer's stable periodic data that matches its feature fusion mechanism.

ADWF-LSTM maintains leadership in autumn (MAE=0.5884, MSE=0.6241, RMSE=0.79), while TCN-LSTM performs abnormally poorly (MAE=0.7768, MSE=0.9735). **Fig. 6** shows TCN-LSTM's predicted curve deviates drastically from actual data due to autumn's discontinuous non-stationary features, whereas ADWF-LSTM's adaptive processing sustains stable performance.

Winter poses the greatest prediction challenge, with all models recording the highest errors. ADWF-LSTM still retains an advantage (MAE=0.4932, MSE=0.7397, RMSE=0.86), 11.9% lower in MAE than TCN-LSTM. **Fig. 7** reveals that winter's intense, irregular fluctuations lead to over-smoothed predictions in traditional LSTM, while ADWF-LSTM's dynamic fusion enhances anti-interference capability to capture key trends.

ADWF-LSTM achieves an average MAE=0.4669, MSE=0.5451, RMSE=0.7275-over 10% lower than other models on average-with MBE=0.0347 (closest to 0, indicating optimal bias control). Visualizations confirm its superior trend-tracking and anti-interference performance across seasons. The core advantage lies in the ADWF module's adaptive adjustment to seasonal data differences, synergizing time-frequency feature extraction with LSTM's temporal modeling to deliver robust full-season PV power prediction.

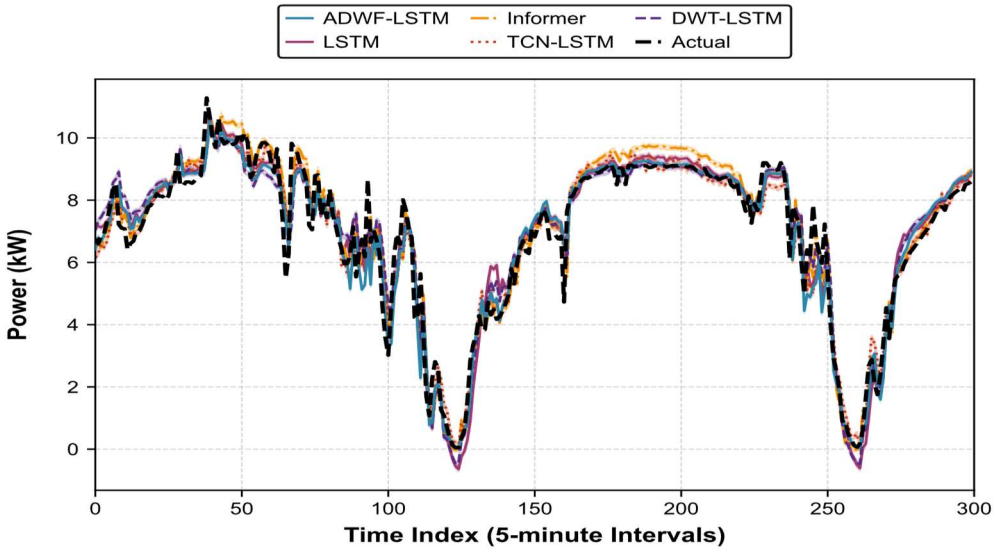

Fig.4 Spring Power Prediction Comparison

ISSN: 2710-0588

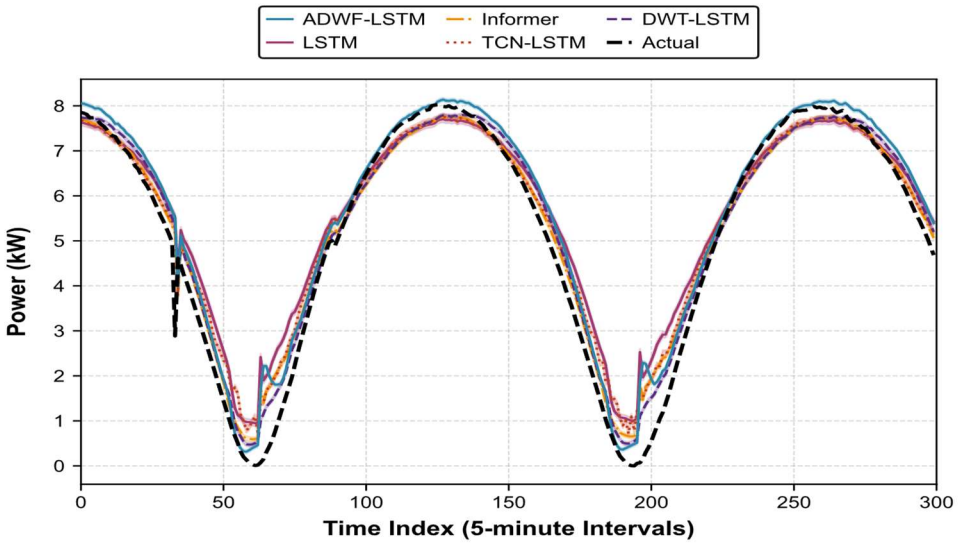

Fig.5 Summer Power Prediction Comparison

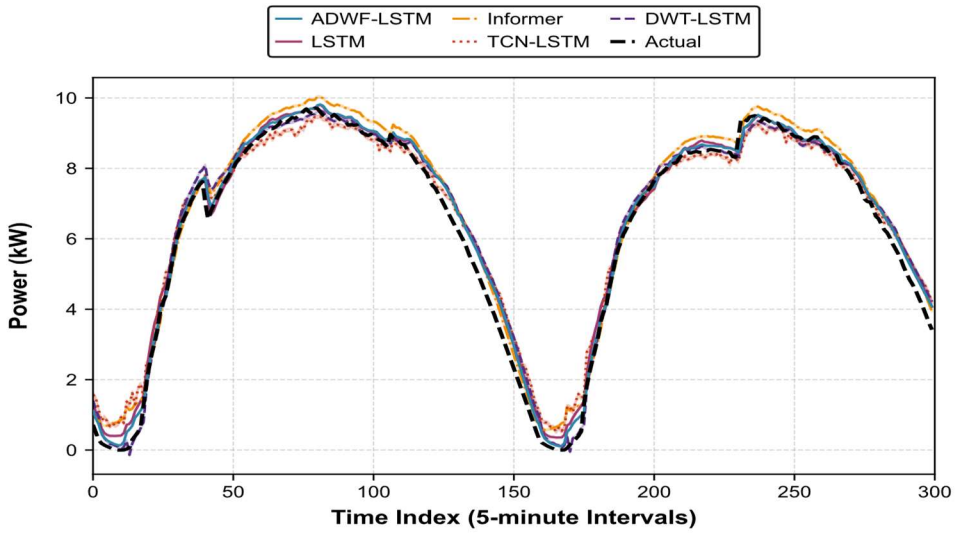

Fig.6 Autumn Power Prediction Comparison

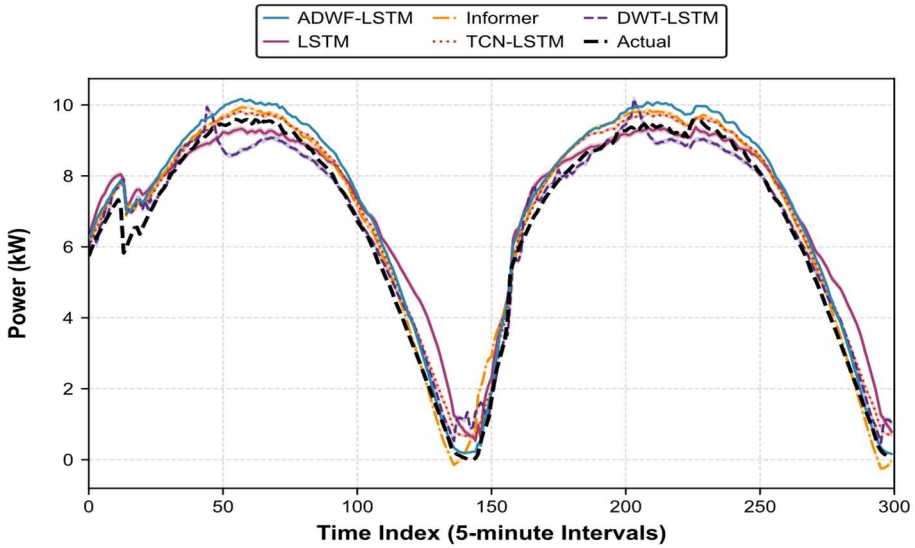

Fig.7 Winter Power Prediction Comparison

ISSN: 2710-0588

5. Conclusion

This paper addresses the challenge of strong non-stationarity in photovoltaic (PV) power data, which arises from factors such as solar irradiance fluctuations, diurnal temperature differences, and equipment operating condition changes. Traditional DWT-LSTM models are limited by fixed wavelet bases and static fusion strategies, leading to insufficient adaptability to dynamic multi-scale features. To overcome these limitations, an ADWF-LSTM model integrated with an Adaptive Dynamic Wavelet Fusion (ADWF) module is proposed.

The ADWF module achieves two key innovations: first, it abandons the single fixed wavelet basis and adopts multiple wavelet bases with complementary time-frequency characteristics, generating data-driven weights to dynamically adjust the contribution of each wavelet basis; second, it realizes adaptive optimization of multi-scale feature fusion weights based on input data characteristics, effectively integrating multi-dimensional time-frequency information. By synergizing the ADWF module with LSTM's long-term temporal dependency modeling capability, the ADWF-LSTM model comprehensively enhances the extraction of non-stationary features and the expression of temporal patterns.

Extensive experiments on the real-world Yulara PV dataset validate the effectiveness of the proposed model. Comparative results across four seasons show that ADWF-LSTM outperforms baseline models (LSTM, DWT-LSTM, TCN-LSTM, Informer) in all evaluation metrics. Its average MAE, MSE, and RMSE reach 0.4669, 0.5451, and 0.7275 respectively, reducing errors by over 10% on average compared to other models. The MBE of 0.0347 is closest to 0, indicating superior systematic bias control. Visualization results further confirm that ADWF-LSTM maintains strong trend-tracking ability and anti-interference performance under diverse seasonal data characteristics, whether for stable periodic summer data or volatile winter data.

Future research directions can be expanded in two aspects: first, optimizing the structure of the ADWF module by introducing attention mechanisms to further enhance the selection accuracy of wavelet bases and fusion weights; second, extending the model to multi-step ultra-short-term PV power prediction scenarios and exploring its adaptability to different types of PV plants and complex meteorological environments, to provide more comprehensive technical support for the efficient integration of renewable energy into power grids.

References

[1] Wen Y ,Wu J ,Lin P , et al.The role of offshore wind and solar PV resources in global low-carbon transition.[J].Science advances,2025,11(43):eadx5580.DOI:10.1126/SCIADV.ADX5580.

[2] Zhang W ,Zhang Z ,Dai Y , et al.Research on reactive power compensation control method for improving the voltage stability of photovoltaic station area[J].Frontiers in Energy Research,2024,121396074-1396074.DOI:10.3389/FENRG.2024.1396074.

[3] Ghali Y ,Sathyajith M ,Joao L .Direct and indirect short-term aggregated turbine- and farm-level wind power forecasts integrating several NWP sources[J].Heliyon,2023,9(11):e21479-e21479.DOI:10.1016/J.HELIYON.2023.E21479.

[4] ZouH ,YangC ,MaH , et al.Short-term power prediction of distributed PV based on multi-scale feature fusion with TPE-CBiGRU-SCA[J].IET Generation, Transmission & Distribution,2024,18(20):3200-3220.DOI:10.1049/GTD2.13266.

[5] Chaoui A K ,Fadil E H ,Choukai O , et al.A Wavelet–Attention–Convolution Hybrid Deep Learning Model for Accurate Short-Term Photovoltaic Power Forecasting[J].Forecasting,2025,7(3):45-45.DOI:10.3390/FORECAST7030045.

[6] Shen A ,Lin Y ,Peng Y , et al.DSC-CBAM-BiLSTM: A Hybrid Deep Learning Framework for Robust Short-Term Photovoltaic Power Forecasting[J].Mathematics,2025,13(16):2581-2581.DOI:10.3390/MATH13162581.

ISSN: 2710-0588

[7] Liu Y ,Guan L ,Hou C , et al.Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform[J].Applied Sciences,2019,9(6):1108.DOI:10.3390/app9061108.

[8] Yousaf M Z , Singh A R , Khalid S ,et al.Bayesian-optimized LSTM-DWT approach for reliable fault detection in MMC-based HVDC systems[J].Scientific Reports, 14(1):1-24[2025-12-13].DOI:10.1038/s41598-024-68985-5.

[9] Das K U ,Tey S K ,Seyedmahmoudian M , et al.Forecasting of photovoltaic power generation and model optimization: A review[J].Renewable and Sustainable Energy Reviews,2018,81912-928. https://doi.org/10.1016/j.rser.2017.08.017.

[10] Karol ,César A ,Víctor E , et al.Intra-hour photovoltaic forecasting through a time-varying Markov switching model[J].Energy,2023,278(PB). https://doi.org/10.1016/j.energy.2023.127952.

[11] Laurentiu F ,Alexandru D ,Dan C , et al.Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models[J].International Journal of Photoenergy,2021. https://doi.org/10.1155/2021/6777488.

[12] Dahidi A S ,Alrbai M ,Alahmer H , et al.Enhancing solar photovoltaic energy production prediction using diverse machine learning models tuned with the chimp optimization algorithm[J].Scientific Reports,2024,14(1):18583-18583. https://doi.org/10.1038/s41598-024-69544-8.

[13] Lwesya F ,Kibambila V .A Comparative Analysis of the Application of Seasonal ARIMA and Exponential Smoothing methods in short run Forecasting Tourist Arrivals in Tanzania[J].European Journal of Business and Management,2017,9(10):56-69.

[14] Dadi P S , Meera R , Rajkumar K ,et al.Enhance Energy Efficiency based Automobile Wheels by using Piezoelectric Energy Harvesting Technology and Sustainable Power Generation[C]//2024 2nd International Conference on Networking and Communications (ICNWC).[2025-10-29].DOI:10.1109/ICNWC60771.2024.10537585.

[15] Pierro M ,De Felice M ,Maggioni E , et al. Data-driven upscaling methods for regional photovoltaic power estimation and forecast using satellite and numerical weather prediction data[J].Solar Energy,2017,158:1026–1038. https://doi.org/10.1016/j.solener.2017.09.068.

[16] Lee H ,Lee BT.Confidence-aware deep learning forecasting system for daily solar irradiance[J]. IET Renew Power Gener,2019,13(10):1681–1689. https://doi.org/10.1049/iet-rpg.2018.5354.

[17] Alonso-Montesinos J ,Batlles F ,Portillo C.Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images[J]. Energy Convers Manag,2015,105:1166–1177. https://doi.org/10.1016/j.enconman.2015.09.001.

[18] Geraldi E ,Romano F ,Ricciardelli E.An advanced model for the estimation of the surface solar irradiance under all atmospheric conditions using msg/seviri data[J]. IEEE Trans Geosci Remote Sens,2012,50(8):2934–2953. https://doi.org/10.1109/TGRS.2011.2178855.

[19] Perez R ,Lorenz E ,Pelland S , et al.Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe[J]. Solar Energy,2013,94:305–326.

[20] Bamisile O , Oluwasanmi A , Obiora S ,et al.Application of deep learning for solar irradiance and solar photovoltaic multi-parameter forecast[J].Energy Sources Part A Recovery Utilization and Environmental Effects, 2020.DOI:10.1080/15567036.2020.1801903.

[21] Zhang J ,Hao Y ,Fan R , et al.An Ultra-Short-Term PV Power Forecasting Method for Changeable Weather Based on Clustering and Signal Decomposition[J].Energies,2023,16(7):DOI:10.3390/EN16073092.

[22] Lyu J ,Zhu G ,He C .Data-driven short-term photovoltaic power forecasting under extreme weather conditions using GRU-KAN model[J].Renewable Energy,2026,257124756-124756.DOI:10.1016/J.RENENE.2025.124756.

[23] Aouni E A ,Naimi E S ,Ayat Y .Machine learning-based photovoltaic power and energy prediction in time–frequency domain[J].Electrical Engineering,2025,107(11):1-12.DOI:10.1007/S00202-025-03272-8.

[24] Guo Y ,Song Y ,Lai Z , et al.Learning Coupled Meteorological Characteristics Aids Short-Term Photovoltaic Interval Prediction Methods[J].Energies,2025,18(2):308-308.DOI:10.3390/EN18020308.

ISSN: 2710-0588

[25] Ospina J ,Newaz A ,Faruque O M .Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model[J].IET Renewable Power Generation,2019,13(7):1087-1095.DOI:10.1049/iet-rpg.2018.5779.

[26] Deng R ,Wang Y ,Xu P , et al.A high-precision photovoltaic power forecasting model leveraging low-fidelity data through decoupled informer with multi-moment guidance[J].Renewable Energy,2025,250123391-123391.DOI:10.1016/J.RENENE.2025.123391.