PHOTOVOLTAIC POWER FORECASTING ALGORITHM BASED ON LSTM-KAN HYBRID ARCHITECTURE

Z. Lingfei [1,2,], C. Yangyang[1], Z. Zhanwen[1], C. Bohang[1,2], Z. Mingyu[1]*

[1] *School of Intelligence Science and Engineering, Qinghai Minzu University, Xining, China*
[2]*National Experimental Teaching Demonstration Center for Communication Engineering, Qinghai Minzu University, Xining, China*
**2004010@qhmu.edu.cn*

Keywords: POWER FORECASTING, LSTM, KAN, GATING MECHANISM, PEARSON CORRELATION

Abstract

Accurate photovoltaic (PV) power generation forecasting plays a pivotal role in enabling optimal grid dispatch operations by providing essential technical references for power system scheduling. To address the challenges of complex meteorological patterns and sudden environmental variations, a photovoltaic power forecasting model named the LSTM-KAN (Long Short-Term Memory - Kolmogorov-Arnold Network) Hybrid Architecture is proposed, with an attention mechanism introduced to enhance model robustness. To investigate the correlation between Numerical Weather Prediction (NWP) features and power output, a multi-dimensional feature selection mechanism was integrated into the data preprocessing phase. To validate the superior performance of the proposed model, photovoltaic operational data from the Alice Springs region in Australia was employed for experimental verification, with comparative analysis conducted against conventional LSTM models in terms of both prediction error and operational stability. Experimental results demonstrated that the LSTM-KAN model achieves 12.04% and 13.26% reductions in MAE and RMSE respectively compared to the baseline LSTM model when evaluated on the Dual dataset.

1 Introduction

Currently, the pressing issues stemming from excessive consumption of conventional energy resources - including resource depletion, carbon emissions, and ecological degradation - demand urgent resolution. The development of green energy has emerged as a viable solution to mitigate overreliance on conventional energy. Although photovoltaic power plants require substantial capital investment during initial construction, their subsequent operational expenditures remain relatively modest, necessitating only routine maintenance to ensure stable functioning post-commissioning. However, photovoltaic generation exhibits high susceptibility to environmental fluctuations, posing challenges for grid management and practical implementation. As demonstrated by Rettger et al., cloud cover can reduce power output by approximately 10% compared with clear-sky conditions, while elevated panel surface temperatures induce additional power degradation[1]. Therefore, investigating the correlation between meteorological factors and photovoltaic output constitutes a critical technological prerequisite for commercial viability, underscoring the paramount importance of accurate photovoltaic power forecasting.

Photovoltaic power forecasting methodologies primarily encompass statistical methods, physical approaches, and machine learning techniques. Statistical methods rely on historical data to construct predictive models, typically operating under the assumption that future photovoltaic output can be extrapolated from historical patterns. However, these methods exhibit limitations in handling extreme weather events and data outliers. Physical approaches leverage meteorological data and photovoltaic component parameters to formulate predictive models, yet require exhaustive input parameters that contribute to computational complexity. Machine learning techniques learn complex patterns of photovoltaic power output through training datasets, effectively capturing nonlinear relationships and intricate mappings between power generation and meteorological parameters. Representative approaches include Support Vector Regression (SVR) and Deep Neural Networks (DNNs) [2-3]. J Shi et al. investigated the application potential of deep learning in photovoltaic power forecasting, highlighting the synergistic integration between algorithmic frameworks and power generation system characteristics[4]. Chen et al. proposed a methodology employing Pearson correlation-based feature extraction for dimensionality reduction, thus alleviating the computational burden on Long Short-Term Memory (LSTM) models[5]. Liu et al. used the PCA method to process the input data for wind power generation prediction, reducing the dimension of the input variables and effectively screening the input variables that affect wind power generation[6].Wang et al. developed a hybrid architecture combining LSTM and Convolutional Neural Network (CNN) models, employing LSTM networks for temporal feature extraction and subsequently applying CNN architectures for spatial feature extraction, thereby improving the accuracy of time-series forecasting[7]. A. Agga et al. proposed a CNN-LSTM hybrid architecture, utilizing CNN modules to extract localized features as input for LSTM networks, with empirical validation demonstrating superior performance in both localized and global forecasting compared to standalone

LSTM configurations, across temporal horizons ranging from 24-hour to 7-day ahead forecasting. Quantitative results revealed that the proposed methodology achieved enhanced accuracy metrics relative to conventional LSTM-based benchmarks [8-9]. Lim et al. developed a CNN-LSTM architecture that initially employs CNN modules to classify meteorological conditions into clear-sky and cloudy categories, followed by training two distinct LSTM networks for specialized learning. This parallel computational framework demonstrated enhanced accuracy in power generation forecasting. Most machine learning methodologies are fundamentally extensions of the Multilayer Perceptron (MLP) paradigm[10]. MLP are fundamentally constructed by enveloping linear models with nonlinear activation functions to achieve transformations in nonlinear spaces, which necessitate increasing the number of hidden layers or enlarging parameter quantities to enhance prediction accuracy when handling intricate nonlinear relationships, thereby incurring elevated computational burdens.

The KAN (Kolmogorov-Arnold Network) model was proposed by researchers Liu Z et al., originating from the Kolmogorov-Arnold Representation Theorem, which asserts that any multivariate continuous function can be decomposed into finite compositions and additive superposition of univariate continuous functions[11]. Within neural network architectures, this principle is manifested through dynamically optimized activation functions and data-driven learning mechanisms, in contrast to the static functional forms of conventional models. Empirical validation demonstrates that KAN achieves dual advantages of high predictive accuracy and enhanced interpretability, with its exceptional nonlinear learning capability being underscored by Liu Z et al. in their seminal work. This research direction has garnered increasing attention within the academic community. As demonstrated by C. J. Vaca-Rubio et al. in applying the KAN to satellite traffic prediction, the framework was found to require fewer parameters than MLP models while demonstrating superior predictive performance[12]. The LSTM-KAN (Long Short-Term Memory – Kolmogorov-Arnold Network) model proposed by R. Xu et al. integrates the memory retention capability of LSTM with the nonlinear representational capacity of KAN, effectively mitigating the limitations of individual models in processing complex datasets associated with dam deformation prediction, while enhancing model interpretability[13]. Jiang et al. applied the KAN to power grid load forecasting, highlighting its superior interpretability as a white-box model[14], thereby demonstrating its applicability to photovoltaic power output prediction domains requiring stringent reliability and safety requirements.

To address the aforementioned challenges, we propose a hybrid LSTM-KAN model specifically designed for photovoltaic time-series data analysis and forecasting. The proposed network adopts a parallel structure of branched LSTM-KAN model. To mitigate computational complexity, eliminate redundant computational expenditure, and suppress spurious noise, a multi-feature selection mechanism is employed for dimensionality reduction. The LSTM module is selected for its demonstrated efficacy in capturing temporal power variation patterns, while the KAN network leverages data-driven spline functions and basis functions to autonomously learn intricate nonlinear patterns, exhibiting enhanced representational capacity with mathematical guarantees. Independent feature extraction from both pathways is concatenated to form composite features, with a dynamic fusion gate adaptively allocating feature weights. By synergistically integrating LSTM's temporal dependency modeling and KAN's nonlinear mapping capabilities, the model achieves superior comprehensive performance.

The principal contributions of this study are summarized as follows:

The model named LSTM-KAN Hybrid Architecture was proposed. To solve the problem of insufficient nonlinear fitting of the prediction model, the KAN network is added on the basis of the prediction of the LSTM network, and the feature weights are dynamically allocated in combination with the attention mechanism. A model named LSTM-KAN is proposed. In the feature fusion part of this model, a dynamic fusion gate is introduced to enable the model to reasonably allocate the feature weights of the two networks. During the implementation process, compared with the benchmark LSTM model, the MAE and RMSE of this model decreased by 27.6% and 1.4% respectively.

A preprocessing framework integrating multi-dimensional feature selection mechanisms is proposed. To solve the problem of large computational volume and multiple feature dimensions of the model, multiple feature selection mechanisms are proposed to jointly screen NWP features. In the data preprocessing part, the feature space dimension is effectively reduced while retaining highly predictive features, thereby improving the computational efficiency.

2 Theoretical Foundation

2.1 Physical and statistical prediction methodologies

Physical methodologies employ solar irradiance and photovoltaic component operation parameters for forecasting, requiring detailed geolocation parameters (latitude/longitude) of PV plants and localized irradiance measurements[15], this methodology obviates the need for extensive historical datasets, rendering it particularly suitable for performance evaluation of newly constructed photovoltaic facilities.

Statistical approaches necessitate extensive historical datasets, including plant power outputs, meteorological records, and operational histories, to discern site-specific generation patterns. These methods exhibit limited portability and are constrained by static prediction frameworks that fail to capture dynamic environmental perturbations.

2.2 Long short-term memory

Machine learning methodologies exhibit superior predictive flexibility compared to conventional physical and statistical forecasting approaches, enabling dynamic capture of temporal-spatial feature dependencies. A representative example is the LSTM network, originally proposed by

Hochreiter et al, which constitutes a specialized variant of Recurrent Neural Networks (RNNs). LSTM addresses the vanishing gradient problem through gating mechanisms[16]. When photovoltaic power inputs exhibit significant intermittency, traditional RNN architectures—particularly those incorporating sigmoid-activated or tanh-activated units—may fail to effectively learn inter-data temporal dependencies under such conditions[17]. The LSTM cell architecture incorporates gating functions, which effectively address long-term dependency challenges and demonstrate superior performance in managing multivariate time-series data. Through its specialized tri-gate structure—comprising input gates, forget gates, and output gates—LSTM successfully mitigates gradient vanishing/explosion issues inherent in RNNs, thereby enhancing long-sequence dependency modeling[18-19]. The architecture diagram of the LSTM network illustrates in Fig. 1, where i_t denotes the output of the input gate, f_t represents the output of the forget gate, o_t corresponds to the output of the output gate.

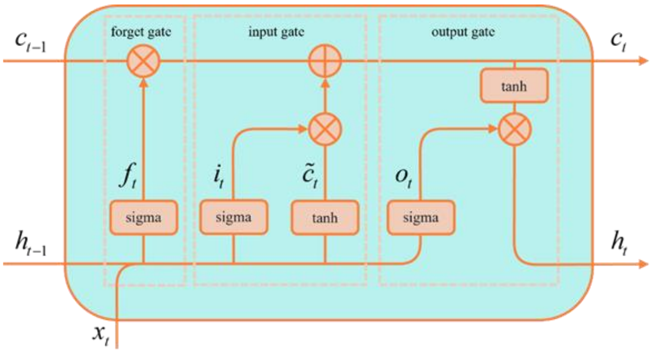

Fig. 1 LSTM network architecture diagram

The core mathematical equations and operational mechanisms are delineated as follows:

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

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{2}$$

$$\tilde{c}_t = tanh\left(W_c x_t + U_c h_t\right) \tag{3}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_{t-1} \tag{4}$$

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{5}$$

$$h_t = o_t \odot tanh\left(c_t\right) \tag{6}$$

The forget gate determines the retention proportion of historical memory, where: W_f , U_f denotes the weight matrix, b_f represents the bias term, and σ signifies the Sigmoid activation function with output constrained to [0,1]. The input gate filters newly acquired information for state updating,while the cell state update integrates outputs from both forget and input gates to refresh long-term memory storage. The output gate regulates information exposure from current cell states to downstream layers. LSTM networks demonstrate exceptional performance in multivariate time-series data analysis, particularly in domains such as natural language processing, energy forecasting, and industrial monitoring, where they have gained widespread adoption. This motivates our selection of LSTM architecture to capture temporal dependencies among feature variables.

2.3 Kolmogorov-Arnold network

KAN is an innovative model derived from the Kolmogorov-Arnold Representation Theorem. The theorem's core assertion states: Any multivariate continuous function f defined on a bounded domain can be decomposed into a finite composition of univariate continuous functions combined through additive superposition. Specifically, for a smooth function, this decomposition can be explicitly formulated as shown in Equation 7.

$$f\left(x\right) = \sum_{i=1}^{2n+1} \Phi_i \left(\sum_{j=1}^{n} \phi_{i,j} x_i \right) \tag{7}$$

The innovation of KAN lies in its integration of learnable activation functions on network edges, where univariate spline functions replace the fixed linear weights of conventional architectures. These activation functions are dynamically refined during training to achieve precise approximation of complex functional relationships. As illustrated in Fig. 2, KAN adopts a two-layer architecture where edge-based activation functions (denoted by rectangular boxes) directly perform nonlinear transformations on input features. This design overcomes the limitations of traditional MLP architectures that rigidly couple linear combinations with fixed activation functions.

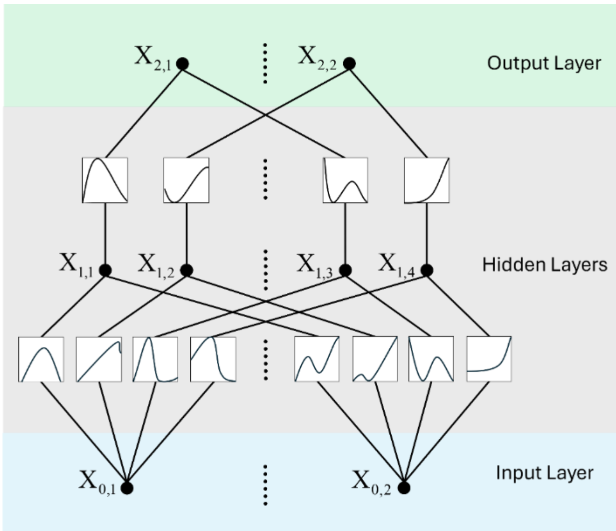

Fig. 2 KAN network architecture Diagram

KAN innovatively incorporates learnable activation functions on network edges, where these functions are dynamically adapted during training through univariate spline functions that effectively replace conventional network weight parameters[20]. Diverging from conventional methodologies, this approach not only preserves architectural flexibility but also achieves precise approximation of complex functional mappings.

3 Proposed Method

3.1 LSTM-KAN Hybrid Architecture

This paper proposes a novel LSTM-KAN predictive model that synergistically integrates LSTM's temporal dependency modeling with KAN's dynamic nonlinear representation capabilities. The LSTM module employs a two-layer architecture for temporal feature extraction from photovoltaic data, while the KAN module utilizes residual networks for enhanced feature learning. A dynamic fusion gate concatenates outputs from both streams, adaptively adjusting their contributions through gating mechanisms. Furthermore, an attention mechanism is incorporated to adaptively allocate feature weights, enabling robust predictions under complex environmental disturbances. As depicted in Fig. 3, the architecture comprises three core components:

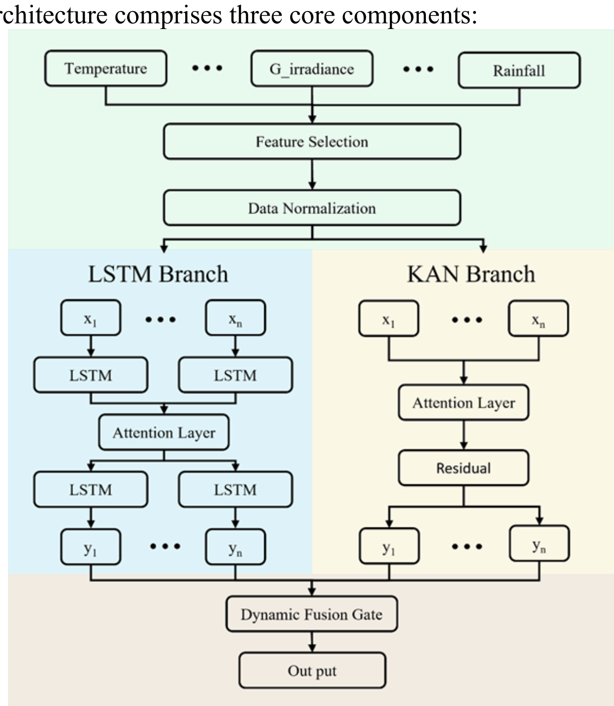

Fig. 3 LSTM-KAN Hybrid Architecture

3.2 LSTM-KAN hybrid architecture

This study employs a hybrid feature selection methodology to construct a multidimensional analytical framework based on NWP parameters, Utilizing a four-dimensional evaluation system comprising Pearson correlation coefficients, Mutual Information Criteria (MIC), Simple Linear Regression (SLR), and Recursive Feature Elimination (RFE). As shown in Table 1, the system reveals the influence of parameters such as Wind-S(Wind-Speed),T(Temperature), H(humidity), G-I(Global Irradiance), D-I(Diffuse Irradiance), Wind-D(Wind Direction), Rainfall , R-G-T(Radiation Global Tilted) and R-D-T(Radiation Diffuse Tilted) on photovoltaic power. Geometric optimization of the evaluation results is achieved through vector synthesis of feature scores. The calculation for mula for the Score is expressed as the normalized result of$\sqrt{P^2 + M^2 + S^2 + R^2}$.In predictive model construction, the precise identification of photovoltaic-relevant features coupled with the elimination of noisy and redundant variables can significantly enhance the predictive accuracy and operational reliability of the system.

Table 1. Feature Correlation Scores of Meteorological Parameters Across Three Datasets (Pearson/MIC/SLR/RFE)

Data set	Features	Pearson	MIC	SLR	RFE	Score
San Yo	T	0.420	0.208	0.177	0.286	0.214
	H	-0.46	0.184	0.208	0.571	0.300
	G-I	0.982	1.452	0.965	0.857	0.883
	D-I	0.590	0.991	0.348	0.714	0.559
	Wind_D	0.009	0.115	0.000	0.143	0.050
	Rainfall	-0.06	0.025	0.004	0.000	0.000
	R-G-T	0.997	1.744	0.994	1.000	1.000
	R-D-T	0.628	1.028	0.395	0.429	0.533
Dual	Wind-S	0.618	0.286	0.382	0.250	0.372
	T	0.500	0.220	0.250	0.625	0.395
	H	-0.48	0.220	0.232	0.375	0.307
	G-I	0.951	1.236	0.904	0.750	0.934
	D-I	0.521	0.993	0.272	0.500	0.589
	Wind-D	-0.07	0.145	0.005	0.125	0.066
	Rainfall	0.072	0.003	0.000	0.000	0.000
	R-G-T	0.964	1.000	0.930	1.000	1.000
	R-D-T	0.552	0.839	0.305	0.875	0.690
BP Solar	T	0.401	0.212	0.161	0.714	0.358
	H	-0.46	0.199	0.211	0.143	0.222
	G-I	0.973	1.377	0.946	0.857	0.924
	D-I	0.570	0.961	0.325	0.571	0.554
	Wind-D	0.007	0.121	0.000	0.429	0.170
	Rainfall	-0.06	0.035	0.004	0.000	0.000
	R-G-T	0.990	1.505	0.981	1.000	1.000
	R-D-T	0.606	1.013	0.368	0.286	0.542

As evidenced in Table 1, distinct feature selection methodologies yield divergent evaluations for identical features. hrough multi-method integrated evaluation, key features are more reliably identified. For instance, Temperature exhibits consistently high correlation coefficients with photovoltaic power output in the DKASC system[21]. Whereas Humidity demonstrates a Pearson correlation coefficient of -0.456, indicating a negative correlation with power generation--though its influence remains non-negligible. Notably, Global_irradiance and Radiation_Global_Tilted exhibit statistically significant correlations, while Wind Direction and Rainfall show weaker predictive associations. Following normalization, features with composite scores exceeding 0.1 are selected as primary inputs for both LSTM and KAN modules.

3.3 Dual-Branch processing layer and dynamic fusion gate

The proposed LSTM-KAN model processes preprocessed features through dual pathways: a bidirectional LSTM branch captures dynamic evolution patterns of environmental parameters via temporal gating mechanisms, whereas the KAN branch employs adaptive spline networks. Specifically, after feature selection and standardization (e.g., temperature, global irradiance), the first LSTM layer utilizes forget gates to filter irrelevant historical data and input gates to regulate update intensity, thereby modeling hysteresis effects in parameters like humidity. The second LSTM layer captures long-term dependencies through cell state propagation, while an attention layer computes time-step-specific hidden state weights via learnable parameter matrices, enhancing feature representation at critical temporal nodes (e.g., daily solar irradiance peaks). In the KAN pathway, adaptive spline networks construct nonlinear mappings using B-spline basis functions. Each neuron fits complex feature-power relationships (e.g., saturation characteristics in irradiance-power curves). Attention mechanisms dynamically weight input features to prioritize critical factors like temperature anomalies, and residual connections stabilize deep network training by fusing raw features with high-order representations. A dynamic fusion gate adaptively allocates weights between the dual pathways through learnable gating functions (Equations 8-9).

$$g = \sigma\left(W_g \times \text{ReLU}\left(W_c \times \left[h_{lstm}; h_{kan}\right] + b_c\right) + b_g\right) \tag{8}$$

$$h = g \cdot h_{lstm} + \left(1-g\right) h_{kan} \tag{9}$$

h_{lstm} is the temporal· vector extracted by the LSTM, h_{kan} is the feature vector from the KAN, W_c represents the projection matrix for the concatenated vect or, W_g is the matrix for generating gating scalars, σ denotes the Sigmoid activation function.

4 Experiments and Results Analysis

4.1 Experimental setup

This study utilizes three photovoltaic datasets from the DKASC system in Alice Springs, Australia: San Yo (ground-mounted fixed-tilt, January 1, 2018 - September 1, 2018), Dual (dual-axis tracking, March 1, 2015 - July 31, 2015), BP Solar (roof-mounted fixed-tilt, January 20, 2018 - July 31, 2018). The datasets record parameters including current, power, temperature, humidity, global horizontal irradiance (G_irradiance), and direct normal irradiance (D_irradiance) at 5-minute intervals. Experimental validation was conducted by comparing the performance of the proposed LSTM-KAN model against classical LSTM and standalone KAN architectures. All models were trained using the Mean Squared Error (MSE) loss function with an Adam optimizer (learning rate: 1e-3, batch size: 1024, epochs: 1000). Input features were standardized via Z-score normalization and split into training-test sets with an 8:2 ratio.

4.2 Evaluation Metrics

To analyze the predictive performance of different algorithms across datasets, we select MAE (Mean Absolute Error), RMSE (Root Mean Square Error), and R^2 (Coefficient of Determination) as evaluation metrics. MAE provides the average level of prediction bias, MRSE reveals the stability of the prediction, and R2 reflects the degree of model fit.

$$MAE = \frac{\sum_{i=1}^{N}\left|y_i - y_i\right|}{N} \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i - y_i\right)^2}{N}} \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - y_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \overline{y}\right)^2} \tag{12}$$

In the above formulas: N is the total number of samples, i is the counting variable, y_i is the true value, (y_i)^is the model's predicted value and y¯ is the sample mean of the dependent variable.

4.3 Evaluation metrics

Experimental results on the San Yo dataset demonstrate that LSTM-KAN model achieves a marginal improvement in MAE compared to standalone LSTM. However, with the incorporation of attention mechanisms, the LSTM-KAN-A model reduces MAE and RMSE by 27.6% and 1.4%, respectively, compared to the baseline LSTM model, and also outperforms LSTM-KAN model. This indicates that introducing attention mechanisms enhances the dynamic adjustment of weights between LSTM model and residual networks. Furthermore, on the more complex Dual dataset, LSTM-KAN model reduces MAE, RMSE, and R^2 by 12.04%, 13.26%, and 1.34%, respectively, compared to standalone LSTM model, while LSTM-KAN-A model exhibits superior performance over LSTM-KAN model after attention integration. Table 2 compares the predictive performance (MAE, RMSE, R^2) of the four models across three distinct datasets.

Table 2. Predictive Performance Comparison of Four Models

Dataset	Predictive Model	MAE/KW	RMSE/KW	R2/%
San Yo	LSTM	0.0623	0.0972	0.9978
	KAN	0.1499	0.1845	0.992
	LSTM-KAN	0.0621	0.1119	0.9971
	LSTM-KAN-A	**0.0451**	**0.0958**	**0.9979**
Dual	LSTM	1.1574	2.2416	0.9487
	KAN	1.2201	2.193	0.9509
	LSTM-KAN	1.0181	1.9441	0.9614
	LSTM-KAN-A	**0.9144**	**1.7873**	**0.9674**
BP Solarl	LSTM	0.0456	0.1043	0.9941
	KAN	0.0588	0.1077	0.9937
	LSTM-KAN	0.0446	0.0963	0.9949
	LSTM-KAN-A	**0.0339**	**0.0928**	**0.9953**

Fig 4 to 6 compare the prediction results of four models against actual photovoltaic power across different datasets, with one representative day selected for each comparison. Taking the San Yo dataset as an example, the predictions of all four models on the August 28 test set are closely aligned with ground truth values. However, as evidenced by the zoomed-in views, the LSTM-KAN-A model most accurately tracks the actual power output.

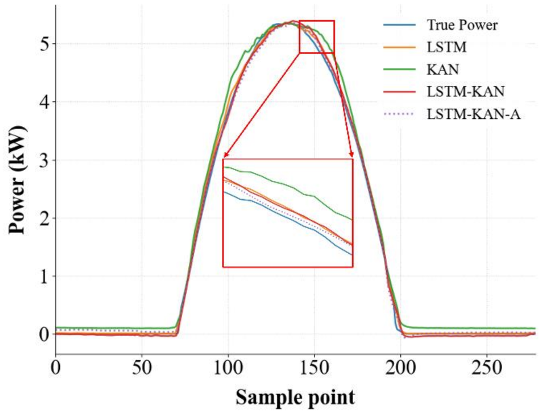

Fig. 4 Experimental Comparison Curves for the San Yo Dataset

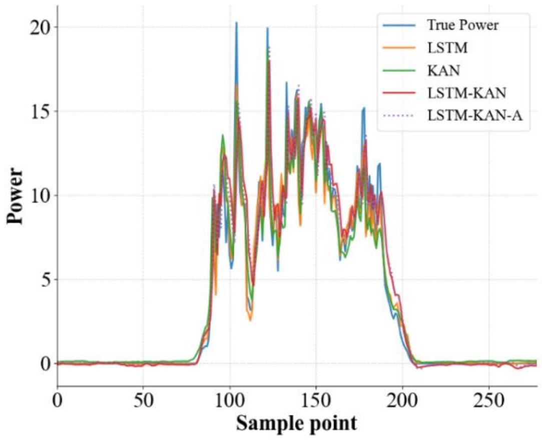

Fig. 5 Experimental Comparison Curves for the Dual Dataset

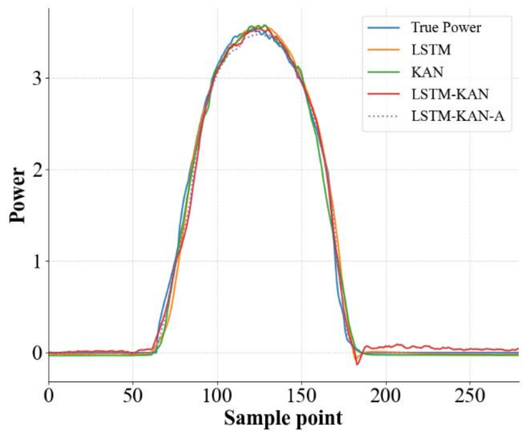

Fig. 6 Experimental Comparison Curves for the BP Solarl Dataset

5 Conclusion

To address photovoltaic power prediction challenges, this study proposes an LSTM-KAN Hybrid Architecture by integrating the KAN with superior nonlinear approximation capabilities. The proposed framework incorporates attention mechanisms to adaptively allocate feature weights, thereby enhancing robustness against transient disturbances (e.g., irradiance fluctuations and cloud occlusion events). Additionally, it employs multi-criteria feature selection mechanisms to jointly screen NWP features, effectively reducing model parameter dimensionality. Experimental validation demonstrates the LSTM-KAN model's superior performance on the DKASC photovoltaic system data from Alice Springs, Australia, showing promising potential for engineering applications in renewable energy dispatch systems, such as photovoltaic power prediction and grid integration.

6 Acknowledgments

This work was supported by the Key Research and Development Transformation Program of Qinghai Province (Grant No. 2025-0204-GXC-0024) and the Kunlun Talent Program of Qinghai Province.

7 References

[1] Rettger P, Keshner M, Pligavko A K, et al. Dynamic management of power production in a power system subject to weather-related factors: U.S. Patent Application 12/320,715[P]. 2010-8-5.

[2] Pan M, Li C, Gao R, et al. Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization[J]. Journal of Cleaner Production, 2020, 277: 123948.

[3] Wolff B, Kühnert J, Lorenz E, et al. Comparing support vector regression for PV power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data[J]. Solar Energy, 2016, 135: 197-208.

[4] Shi J, Zhang J. Ultra short-term photovoltaic refined forecasting model based on deep learning[J]. Electric Power Construction, 2017, 38(6): 28-35.

[5] Chen H, Chang X. Photovoltaic power prediction of LSTM model based on Pearson feature selection[J]. Energy Reports, 2021, 7: 1047-1054.

[6] Shijian L, Yongjun H, Fuchao L. Application of the combined model in short-term wind power forecasting[J]. International Journal of Smart Grid and Clean Energy, 2016, 5(3).

[7] Wang K, Qi X, Liu H. Photovoltaic power forecasting based LSTM-Convolutional Network[J]. Energy, 2019, 189: 116225.

[8] Agga A, Abbou A, Labbadi M, et al. Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models[J]. Renewable Energy, 2021, 177: 101-112.

[9] Agga A, Abbou A, Labbadi M, et al. CNN-LSTM: An efficient hybrid deep learning architecture for predicting

short-term photovoltaic power production[J]. Electric Power Systems Research, 2022, 208: 107908.
[10] Lim S C, Huh J H, Hong S H, et al. Solar power forecasting using CNN-LSTM hybrid model[J]. Energies, 2022, 15(21): 8233.
[11] Liu Z, Wang Y, Vaidya S, et al. Kan: Kolmogorov-arnold networks[J]. arXiv preprint arXiv:2404.19756, 2024.
[12] Vaca-Rubio C J, Blanco L, Pereira R, et al. Kolmogorov-arnold networks (kans) for time series analysis[J]. arXiv preprint arXiv:2405.08790, 2024.
[13] Xu R, Liu X, Wei J, et al. Predicting the Deformation of a Concrete Dam Using an Integration of Long Short-Term Memory (LSTM) Networks and Kolmogorov–Arnold Networks (KANs) with a Dual-Stage Attention Mechanism[J]. Water, 2024, 16(21): 3043.
[14] Jiang B, Wang Y, Li Y, et al. A Novel Interpretable Short-term Probabilistic Load Forecasting Method based on Kolmogorov-Arnold Networks[J]. Authorea Preprints.
[15] Almonacid F, Pérez-Higueras P J, Fernández E F, et al. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator[J]. Energy conversion and Management, 2014, 85: 389-398.
[16] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
[17] Yu Y, Si X, Hu C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural computation, 2019, 31(7): 1235-1270.
[18] Sun S, Wang S, Wei Y. A new ensemble deep learning approach for exchange rates forecasting and trading[J]. Advanced Engineering Informatics, 2020, 46: 101160.
[19] Xiang S, Qin Y, Zhu C, et al. Long short-term memory neural network with weight amplification and its application into gear remaining useful life prediction[J]. Engineering Applications of Artificial Intelligence, 2020, 91: 103587.
[20] Vaca-Rubio C J, Blanco L, Pereira R, et al. Kolmogorov-arnold networks (kans) for time series analysis[J]. arXiv preprint arXiv:2405.08790, 2024.
[21] Desert Knowledge Australia Centre , https://dkasolarcentre.com.au/download?location=alice-springs，last accessed 2022/10/15.