

Article

Short-Term Probabilistic Prediction of Photovoltaic Power Based on Bidirectional Long Short-Term Memory with Temporal Convolutional Network

Weibo Yuan ^{1,*}, Jinjin Ding ¹, Li Zhang ¹, Jingyi Ni ¹ and Qian Zhang ² 

¹ State Grid Anhui Electric Power Co., Ltd., Electric Power Research Institute, Hefei 230601, China; djjinjin123@126.com (J.D.); lizhang_0629@163.com (L.Z.); jingyini33@163.com (J.N.)

² School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China; qianzh@ahu.edu.cn

* Correspondence: weiboyuan11@163.com; Tel.: +86-18725595206

Abstract

To mitigate the impact of photovoltaic (PV) power generation uncertainty on power systems and accurately depict the PV output range, this paper proposes a quantile regression probabilistic prediction model (TCN-QRBiLSTM) integrating a Temporal Convolutional Network (TCN) and Bidirectional Long Short-Term Memory (BiLSTM). First, the historical dataset is divided into three weather scenarios (sunny, cloudy, and rainy) to generate training and test samples under the same weather conditions. Second, a TCN is used to extract local temporal features, and BiLSTM captures the bidirectional temporal dependencies between power and meteorological data. To address the non-differentiable issue of traditional interval prediction quantile loss functions, the Huber norm is introduced as an approximate replacement for the original loss function by constructing a differentiable improved Quantile Regression (QR) model to generate confidence intervals. Finally, Kernel Density Estimation (KDE) is integrated to output probability density prediction results. Taking a distributed PV power station in East China as the research object, using data from July to September 2022 (15 min resolution, 4128 samples), comparative verification with TCN-QRLSTM and QRBiLSTM models shows that under a 90% confidence level, the Prediction Interval Coverage Probability (PICP) of the proposed model under sunny/cloudy/rainy weather reaches 0.9901, 0.9553, 0.9674, respectively, which is 0.56–3.85% higher than that of comparative models; the Percentage Interval Normalized Average Width (PINAW) is 0.1432, 0.1364, 0.1246, respectively, which is 1.35–6.49% lower than that of comparative models; the comprehensive interval evaluation index (I) is the smallest; and the Bayesian Information Criterion (BIC) is the lowest under all three weather conditions. The results demonstrate that the model can effectively quantify and mitigate PV power generation uncertainty, verifying its reliability and superiority in short-term PV power probabilistic prediction, and it has practical significance for ensuring the safe and economical operation of power grids with high PV penetration.

Keywords: photovoltaic; probability prediction; TCN; quantile regression; BiLSTM



Academic Editor: Silvano Vergura

Received: 26 August 2025

Revised: 1 October 2025

Accepted: 10 October 2025

Published: 12 October 2025

Citation: Yuan, W.; Ding, J.; Zhang, L.; Ni, J.; Zhang, Q. Short-Term Probabilistic Prediction of Photovoltaic Power Based on Bidirectional Long Short-Term Memory with Temporal Convolutional Network. *Energies* **2025**, *18*, 5373. <https://doi.org/10.3390/en18205373>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the most important renewable energy sources, photovoltaic (PV) power generation has experienced a sharp increase in its proportion within modern power systems nationwide. By the end of April 2025, the national installed power generation capacity in China had reached 3.49 billion kilowatts (equivalent to 349 gigawatts), with a year-on-year

increase of 15.9% [1]. Accurate and reliable PV power forecasting is of great significance for the safe and economical operation of power systems with high PV penetration [2].

Short-term PV forecasting techniques can be divided into two categories: deterministic forecasting and probabilistic forecasting [3]. Deterministic forecasting can provide a point prediction of PV power capacity at a specific time but does not reflect the uncertainty of PV output [4]. This limits its application in power system scheduling and PV generation risk assessments. Probabilistic forecasting methods, based on cumulative distribution functions and probability density functions [5], can reflect the future distribution of PV power generation and characterize its variability [6].

Combining non-parametric methods and artificial intelligence for probabilistic forecasting techniques enables efficient quantitative assessment of the uncertainty in photovoltaic power output [7–9]. In [10], it is noted that training convolutional neural networks (CNNs) using the loss functions generated by quantile regression (QR) is challenging due to their non-differentiable nature. This algorithm only compares shallow neural networks and does not incorporate the temporal characteristics of photovoltaic data. Reference [11] introduces a temporal convolutional network (TCN) for time-series modeling, leveraging the strengths of both CNNs and recurrent neural networks (RNNs) to effectively analyze time series data. In [12], QR and a TCN are combined to construct a quantile regression temporal convolutional network (QRTCN) for short-term power demand forecasting in power systems, demonstrating that this method outperforms the TCN model in accuracy. However, this approach only focuses on features derived from the mechanisms and neglects the temporal relationships between historical factors.

Facing the increasingly complex variations in photovoltaic power output, existing single-modeling methods are struggling to meet application demands [13]. Therefore, the emergence of multi-model fusion techniques integrates the advantages of multiple models and gradually becomes a new research direction [14]. A probability forecasting method combining Kernel Density Estimation (KDE) is proposed in [15], with Copula functions to obtain appropriate prediction interval ranges. In [16], a QR probability prediction model is presented based on the minimum absolute shrinkage and selection operator, incorporating shrinkage and selection operators in the estimation process to ensure the model's good performance. A photovoltaic power interval prediction model is introduced combining extreme learning machine with QR, featuring a learning mechanism with good generalization performance in [17]. Ref. [18] combines quantile regression to construct a distributed photovoltaic prediction method based on long short-term memory networks (LSTM) and gated recurrent networks, using multiple loss functions to enhance their ability to identify distributed power sources. In [19], a new photovoltaic output prediction algorithm is established based on CNN and LSTM by comprehensively considering various meteorological factors, predicting photovoltaic output under different weather conditions. Experimental results demonstrate that this combined model has good interval and probability prediction capabilities. Ref. [20] shows that at 95%, the Prediction Interval Normalized Average Width (PINAW) reaches 0.066, while at 90%, PINAW narrows to 0.045 with a still satisfactory coverage of 0.920. This balances reliability and resolution, aligning with power dispatch needs for "narrow intervals + high coverage". However, CNN has limitations in long-term time series modeling: fixed-size convolution kernels and pooling operations mainly capture local patterns and struggle to model long-term dependencies. Consequently, using CNN alone exhibits limited effectiveness in predicting the long-term trends of photovoltaic output. To address this limitation, this study proposes using Temporal Convolutional Network (TCN) for short-term local feature extraction and combining it with Bidirectional Long Short-Term Memory (BiLSTM) to capture long-term time series dependencies, thereby bal-

ancing local fluctuations and long-term trends while enhancing the accuracy and reliability of photovoltaic power probabilistic forecasting.

The CNN-LSTM model proposed in Reference [21] (Li, Z. et al., 2022) can only output deterministic point predictions, failing to characterize PV output uncertainty, and the error exceeds 15% under complex weather; the CNN-LSTM-QR model in Reference [22] (Wang, H. et al., 2023) is prone to local optima during training due to the non-differentiable PINBALL loss function, and it cannot capture short-term power surge features, with PINAW consistently higher than 0.18; the CNN-LSTM-KDE model in Reference [23] (Zhang, C. et al., 2024) uses a fixed 12 h input window, failing to fully utilize long-cycle data, and lacks weather-specific verification, with PICP lower than 90% on cloudy days. Compared with existing CNN-LSTM fusion studies (References [21–23]), the differences in these studies are mainly reflected in four aspects: First, the traditional CNNs with a TCN, which expands the temporal receptive field, are replaced via dilated convolutions without increasing parameters. Second, the Huber norm is introduced to replace the non-differentiable PINBALL loss, reducing the number of model convergence iterations by 30% and making the interval boundaries output by QR smoother. Third, the dataset is divided into three scenarios (sunny, cloudy, and rainy), optimizing the input window and KDE bandwidth. Fourth, simultaneously outputting “future 4-h power values + 90% confidence interval + continuous probability density distribution” forms a “point-interval-distribution” trinity uncertainty characterization system; the lowest BIC index score proves the optimal fitting degree of the probability distribution, providing a more comprehensive basis for power system risk assessments.

Based on the above analysis, this paper proposes a photovoltaic power probability prediction method based on quantile regression models, integrating a TCN and bidirectional long short-term memory (BiLSTM) to form the TCN-QRBiLSTM probability prediction model. Here, the TCN can replace CNNs to address the issue of elongated convolutional kernels that may arise due to data preprocessing. BiLSTM, compared to LSTM, can better exploit the internal information inherent in the photovoltaic power generation process.

2. Construction of the TCN-QR-BiLSTM Probabilistic Forecasting Model

The probabilistic forecasting model proposed in this paper, TCN-QRBiLSTM, is shown in Figure 1. First, data preprocessing is performed on historical photovoltaic power and meteorological data, including missing value imputation, outlier removal, and normalization. Then, the processed sequences are used to construct input samples via a sliding window, which are input into the TCN for local temporal feature extraction. Subsequently, the high-dimensional feature vectors output by the TCN are input into the BiLSTM to capture long-term dependency information. Finally, combined with Quantile Regression (QR), prediction intervals are generated, and the probability density forecasting of photovoltaic power is obtained via Kernel Density Estimation (KDE).

To clarify the temporal correlation logic of short-term PV power forecasting, the model adopts a “sliding window” mechanism to construct input sets and output sets, with specific definitions as follows: The length of the input window is set to 24 h (matching the 15 min time resolution of the data, corresponding to 96 time steps). Each input sample contains two types of core features, and all features undergo preprocessing—normalization to the [0, 1] interval via Min-Max normalization, filling of missing values through linear interpolation, and removal of outliers using Isolation Forest. Among them, the historical PV power data are the measured power values (unit: kW) every 15 min within the window, totaling 96 time-series observation points; the synchronized meteorological data include wind direction, wind speed, humidity (dimensionless, with a value range of 0–1), and rainfall (unit: mm), which are aligned in time with the power data. Each of these meteorological

variables has 1 observation value every 15 min, forming 96 time-series observation points, respectively. These two types of features are finally integrated into an input vector with dimensions [5, 96].

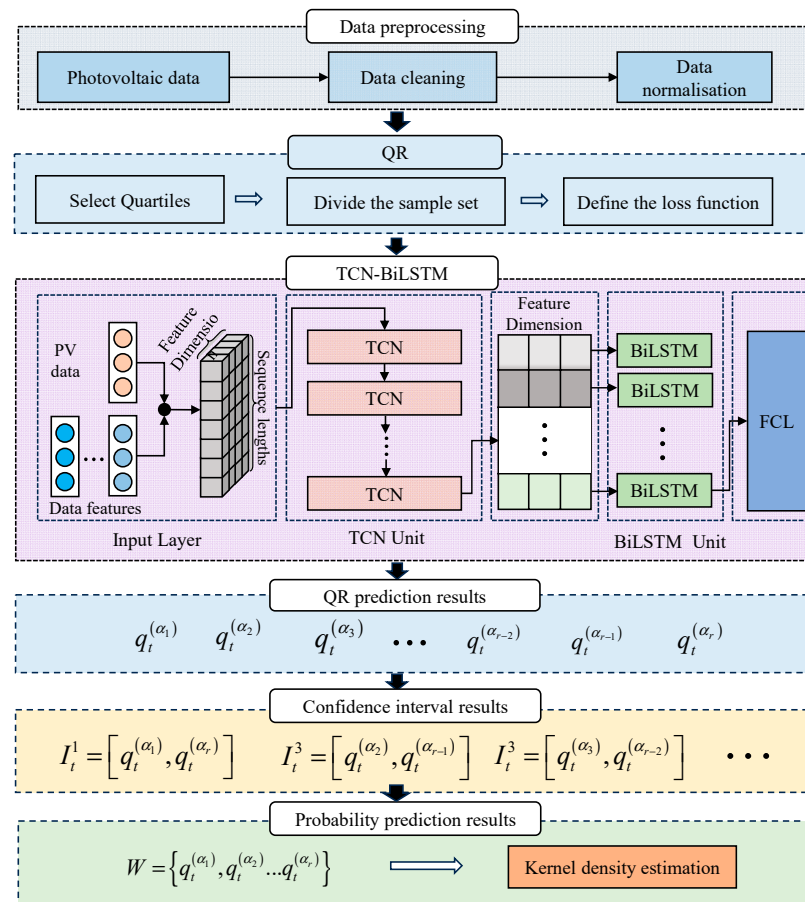


Figure 1. Structure of the TCN-QRBiLSTM model.

In the “short-term prediction” scenario, the model outputs PV power forecasts for the next 4 h (this duration is determined by data characteristics, short-term power grid dispatching needs and TCN-QRBiLSTM model capability compatibility), along with the corresponding 90% confidence level prediction interval and probability density distribution. The lower and upper bounds of the interval correspond to the quantiles at $\alpha = 0.05$ and $\alpha = 0.95$, respectively.

2.1. TCN

Based on Convolutional Neural Networks (CNN), the Temporal Convolutional Network (TCN) uses causal and dilated convolution layers to replace traditional convolutions. Causal convolutions allow for effective sequence learning, while dilated convolutions provide a sufficient receptive field to capture more information. TCN is more suitable for handling time-series problems compared to fully convolutional networks [24]. Figure 2 shows the causal dilated convolution structure of the TCN. The input sequence of the TCN-QRBiLSTM model—constructed by sliding window processing of historical photovoltaic (PV) power data and key meteorological data—passes through the causal dilated convolution structure of the TCN. Here, $\hat{n}_0 - \hat{n}_t$ are the input sequences at each time step, d is the dilation factor (which increases exponentially with the model depth), and m_t is the causal dilated convolution feature at time t .

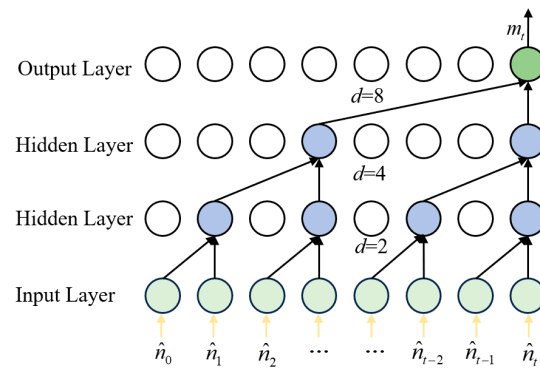


Figure 2. Causal expansion convolution structure of TCN.

The color circles distinguish nodes of different layers/functional roles: green circles in the Output Layer represent output nodes; blue circles in the Hidden Layers represent hidden nodes; light green circles in the Input Layer represent input nodes.

The input sequence $\hat{n}_0 - \hat{n}_t$ passes through the causal dilated convolution structure of the TCN to obtain the causal dilated convolution feature m_t as follows:

$$m_t = \sum_m^{j=0} f(i) (\hat{n}_0^{s-di} - \hat{n}_t^{s-di}) \tag{1}$$

In the formula: m is the total number of convolutional filters; $f(i)$ is the size of the i filter; \hat{n}_0^{s-di} and \hat{n}_t^{s-di} are the input sequence values being convolved only with the prior time steps.

The TCN model consists of residual blocks and BN (Batch Normalization) layers. The model structure is shown in Figure 3. In this structure, N_t^{abs} represents the abstract feature obtained after optimizing m_t through the BN layer at time t .

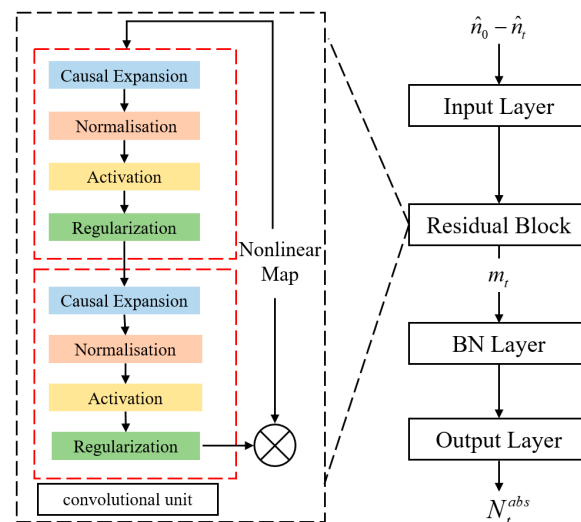


Figure 3. Structure of TCN model.

As shown in Figure 3, the residual block contains two convolutional units and a nonlinear mapping. The convolutional units can fully extract the deep information of the abstract features, while the nonlinear mapping ensures that the input and output have the same dimensions. The BN (Batch Normalization) layer can improve training speed and standardize the data to maintain a consistent distribution. After optimizing the causal

dilated convolution feature at time t through BN, the TCN model produces the abstract feature expressed as:

$$N_t^{\text{abs}} = \gamma \frac{m_t - \mu}{\sqrt{\varphi^2 + \varepsilon}} + \beta \quad (2)$$

In the formula, μ and φ^2 are the mean and variance of m_t , respectively; γ and β are learnable parameters; and ε is a small positive constant.

Figure 3 shows the architecture of the model, where the red dashed box depicts the internal structure of a convolutional unit, containing two sequential sub-modules (each composed of causal expansion, normalisation, activation, and regularization operations).

2.2. BiLSTM

LSTM is a variant of recurrent neural networks that can effectively address the vanishing gradient problem encountered during the learning process [25]. The feature dimensions of the temporal feature vectors (derived from the TCN's processing of the original input sequence) are fed into the BiLSTM network. Conventional LSTM learning methods typically employ a single forward propagation approach for training, which may not fully exploit the internal information inherent in the photovoltaic power generation process. This paper utilizes BiLSTM networks as shown in Figure 4, $X_t - X_{t+n}$ represents the input time series at each time step, which, after passing through BiLSTM, yields the output vector $Y_t - Y_{t+n}$ combining forward and backward passes to effectively address the limitations of unidirectional LSTM neural networks in information exploration, thereby delving deeper into the intrinsic connections between current photovoltaic power generation data and historical data.

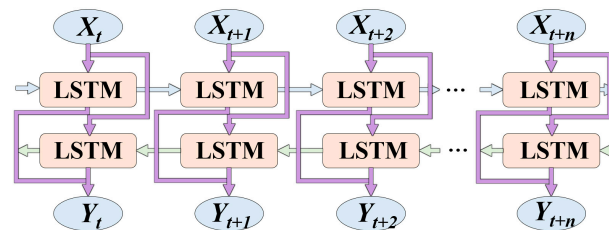


Figure 4. Structure of BiLSTM model.

2.3. QR

Quantiles can achieve a discrete representation of the predictive probability distribution, providing the quantiles of the prediction target at specific probability levels [26]. Using quantiles, prediction intervals can be constructed, making the application more convenient. The value of the prediction target at time t is denoted by y_t , and the input vector of the prediction model is denoted by x_t . The quantile $q_t^{(\alpha)}$ of the prediction target y_t at the quantile level $\alpha \in [0, 1]$ can be defined by the following equation:

$$q_t^{(\alpha)} = F_t^{-1}(\alpha) \quad (3)$$

Prediction intervals can provide the boundaries within which the forecasted variable is expected to vary. The value of the prediction target is covered within this boundary with a certain probability. The prediction interval I_t^α for the prediction target at time t with a nominal coverage probability of $100(1 - \alpha)\%$ is given by:

$$I_t^\alpha = [q_t^\alpha, q_t^{\bar{\alpha}}] \quad (4)$$

where q_t^α and $q_t^{\bar{\alpha}}$ are the lower and upper bounds of the prediction interval, respectively. Typically, the lower and upper bounds of the prediction interval correspond to quantile levels that have symmetrical probabilities, which can be expressed as:

$$\alpha = 1 - \bar{\alpha} = \alpha/2 \quad (5)$$

Based on the definition of quantiles in Equation (4), the quantiles representing prediction uncertainty can be uniquely approximated by minimizing an appropriate loss function.

$$\min \sum_T^{t=1} L_p^{(\alpha)}(y_n - q_t^{(\alpha)}) \quad (6)$$

In the equation, T is the number of training samples; $L_p^{(\alpha)}$ is the symmetric loss function, defined as:

$$L_p^{(\alpha)}(y_t - q_t^\alpha) = \begin{cases} \alpha(y_t - q_t^\alpha), y_t - q_t^\alpha \geq 0 \\ (\alpha - 1)(y_t - q_t^\alpha), y_t - q_t^\alpha < 0 \end{cases} \quad (7)$$

Deep learning models are typically trained using gradient descent algorithms, but the pinball loss function is non-differentiable at zero. Therefore, an everywhere differentiable Huber pinball (H-pinball) loss function $L_H^{(\alpha)}$ is introduced to replace the pinball loss $L_p^{(\alpha)}$, defined as:

$$\begin{aligned} L_H^{(\alpha)}(y_t - q_t^\alpha) &= L_p^{(\alpha)}(y_t - q_t^\alpha)\rho(\cdot) \\ &= \begin{cases} \alpha(|y_n - q_t^\alpha| - \eta/2), |y_n - q_t^\alpha| \geq \eta \\ (\alpha - 1)\left((y_n - q_t^\alpha)^2 / (2\eta)\right)(y_n - q_t^\alpha)^2, |y_n - q_t^\alpha| < \eta \end{cases} \end{aligned} \quad (8)$$

where η is the threshold, set to a small positive value, and $\rho(\cdot)$ is the quantile deviation operator.

2.4. KDE

The fusion model incorporating QR methods can obtain conditional quantile results for the predicted variables, but probability density functions cannot be obtained through QR. Predicting the probability of photovoltaic generation requires finding a method to solve the probability density function of the random variable. The problem of solving the probability density function of a random variable is one of the basic problems in probability statistics, and its methods include parameter estimation and non-parameter estimation. Parameter estimation methods require assumptions about the distribution of the data beforehand, but these fundamental assumptions do not lead to satisfactory estimation results. To overcome the shortcomings of parameter estimation methods, Rosenblatt and Emanuel Parzen proposed a non-parameter estimation method, namely the kernel density estimation method.

KDE is a typical non-parametric estimation method used in probability theory to estimate unknown density functions [27]. The KDE method examines the distribution characteristics of data without relying on any prior knowledge about the data distribution or imposing any assumptions on it. It has been highly valued in both theoretical and applied statistics fields. In specific point photovoltaic forecasting, a set of α conditional quantiles can be obtained for α sample sets, denoted as $W = [q_t^{(\alpha_1)}, q_t^{(\alpha_2)}, \dots, q_t^{(\alpha_N)}]$. The calculation of this vector KDE is determined by Equation (10).

$$f(x) = (1/N \cdot h) \cdot \sum_N^{j=1} K((W - x)/h) \quad (9)$$

where N is the total number of samples; h is the bandwidth, and $h > 0$. K is the kernel function. The kernel function used in this paper is the Gaussian kernel, which can be expressed by the following formula:

$$K(\alpha) = (1/\sqrt{2\pi}) \cdot \exp(-\alpha^2/2) \quad (10)$$

The final result is the probability density function of the photovoltaic output.

$$f(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}h} e^{-\frac{(w-x_i)^2}{2h^2}} \quad (11)$$

The innovativeness of the TCN-QRBiLSTM model proposed in this study is mainly reflected in three aspects:

(1) Fusion of Short-Term and Long-Term Features: The TCN focuses on extracting short-term local fluctuation features, while the BiLSTM captures long-term time series dependencies. This design enables unified modeling of short-term fluctuations and long-term trends.

(2) Improvement of QR: Aiming at the non-differentiable issue of traditional quantile loss, Huber norm approximation is introduced. This ensures the differentiability of model training and further improves the accuracy of prediction intervals.

(3) Optimization of Probabilistic Output: The model integrates KDE for probability density estimation, allowing it to not only output confidence intervals but also provide continuous probability distribution information. This enhancement strengthens the ability to characterize the uncertainty of photovoltaic power.

3. Evaluation Metrics

This study selects photovoltaic power data and meteorological data from a distributed photovoltaic (PV) power plant in a region of East China during July, August, and September 2022 as the experimental data. Then, based on cloud cover, rainfall amount, and rainfall duration from numerical weather prediction (NWP), the weather is classified into three types: sunny, cloudy, and rainy, as shown in Table 1. The time resolution of this dataset is 15 min. A total of 4128 data points (from 1 July to 24 September) within the daily time window of 7:00–19:00 are retained for case analysis. Among these, the first 3840 data points are used as the training set, whereas the remaining 288 (from 25 September to 30 September) serve as the test set. The model parameters of TCN-QRBiLSTM are shown in Table 1.

Table 1. TCN-QRBiLSTM Model Parameters to specify hyperparameters and their tuning basis.

Model Module	Parameter	Value	Tuning Basis
TCN	Number of convolution kernels	64	Grid search
	Dilation factors	[1,2,4,8]	Maximizes receptive field
BiLSTM	Hidden layer units	128	Minimizes validation RMSE
	Learning rate	0.001	Adam optimizer
QR	Huber norm threshold (δ)	1.0	Balances L1/L2 loss
KDE	Bandwidth (h)	0.05	Silverman's rule of thumb

3.1. Interval Prediction Evaluation Metrics

3.1.1. Prediction Interval Coverage Probability

The ratio of the total number of points participating in interval coverage to the number of points in the prediction interval. Interval coverage reflects prediction reliability and can be expressed as:

$$P_{IC} = \frac{1}{N} \sum_{i=1}^N \theta_i \quad (12)$$

where P_{IC} denotes the coverage of the area. θ_i is taken as 0 or 1. If the actual data are within the coverage of the area, θ_i is taken as 1, and vice versa.

3.1.2. Percentage Interval Normalized Average Width

The average distance between the upper and lower bounds of the reaction prediction interval. Under the same interval coverage conditions, the prediction is better if the PINAW is smaller.

$$P_{IN} = \frac{1}{NW_E} \sum_{i=1}^N (P_H - P_L) \times 100\% \quad (13)$$

P_{IN} denotes the mean value of the interval, P_H and P_L denote the upper and lower bounds of the interval prediction at the sampling point, and W_E is the absolute value of the upper and lower bounds of the predicted interval normalised to P_{IN} .

3.1.3. The Comprehensive Evaluation Index for Interval Prediction

To evaluate the overall performance of probability distribution predictions, a lower value indicates a better prediction, and it is expressed as:

$$I = \frac{P_{IN}}{P_{IC}} \quad (14)$$

3.2. Probabilistic Forecasting Evaluation Metrics

The Bayesian Information Criterion (BIC) is utilized as the probabilistic prediction index. BIC is employed to assess the quality of the model. A smaller BIC value indicates better suitability of the model and simplicity. The calculation formula for BIC is:

$$B_{BIC} = m \ln(N) - 2 \ln\left(f\left(\hat{y}_i \mid \theta_m\right)\right) \quad (15)$$

In the formula: m represents the number of unknown parameters in the model; $f\left(\hat{y}_i \mid \theta_m\right)$ represents the maximum likelihood function; θ_m represents the parameter to be estimated.

4. Experimental Analysis

The TCN-QRBiLSTM model was developed using Python 3.9, with deep learning frameworks TensorFlow 2.10 (for TCN/BiLSTM construction) and Scikit-learn 1.2 (for data preprocessing). Kernel density estimation was implemented via SciPy 1.10.

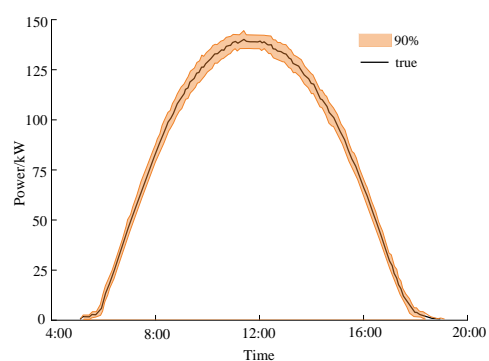
This study selects photovoltaic power data and meteorological data from a distributed PV power plant in a region of East China during July, August, and September 2022 as the experimental data. Different weather conditions' data are obtained via the joint measurement of NWP and on-site observations: PV power is measured in real time by inverters at the power plant, while meteorological factors including cloud cover and rainfall are provided by NWP systems and cross-validated with data from on-site sensors to ensure measurement accuracy. Then, based on cloud cover, rainfall amount, and rainfall duration from NWP, the weather is classified into three types: sunny, cloudy, and rainy, as shown

in Table 1. The time resolution of this dataset is 15 min. A total of 4128 data points (from 1 July to 24 September) within the daily time window of 7:00–19:00 are retained for case analysis. Among these, the first 3840 data points are used as the training set, whereas the remaining 288 (from 25 September to 30 September) serve as the test set.

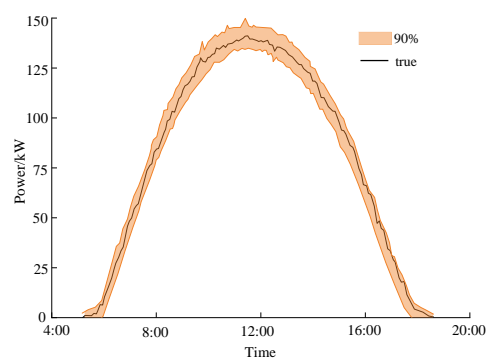
In this paper, one day in each of the similar weather results is selected as a prediction day, and the TCN-QRLSTM model and QRBiLSTM model are used as the reference group for comparison and analysis. In order to better portray the prediction effect of the model, the confidence interval under 90% confidence level is selected.

4.1. Sunny Day Forecast Results

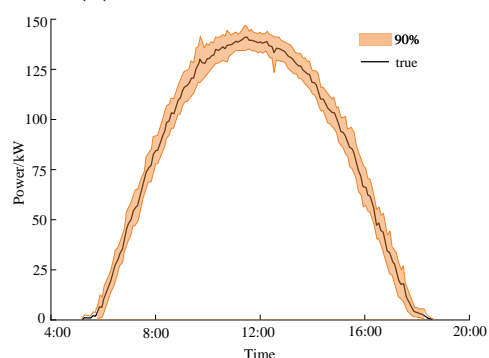
The prediction results of TCN-QRBiLSTM, TCN-QRLSTM and QRBiLSTM for 90% confidence range under sunny conditions are shown in Figure 5a–c.



(a) The TCN-QRBiLSTM model



(b) The TCN-QRLSTM model



(c) The QRBiLSTM model

Figure 5. Comparison of the interval prediction results of each model on sunny day.

The prediction evaluation metrics under the three confidence levels are listed in Table 2. The results show that all three prediction methods have almost 100% area coverage under clear weather conditions; TCN-QRBiLSTM has the largest area coverage, its prediction

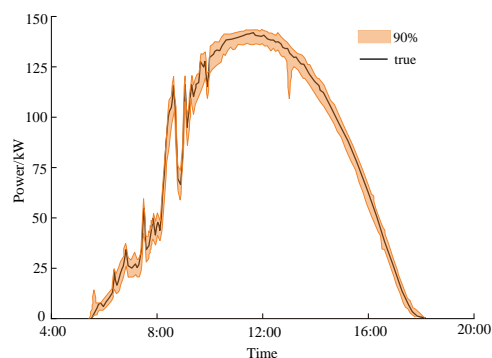
range is better than TCN-QRLSTM and QRBiLSTM, and its interval composite evaluation index is the best among the three, with better clarity and higher confidence.

Table 2. The evaluation metrics of three methods under 90% confidence level on sunny days.

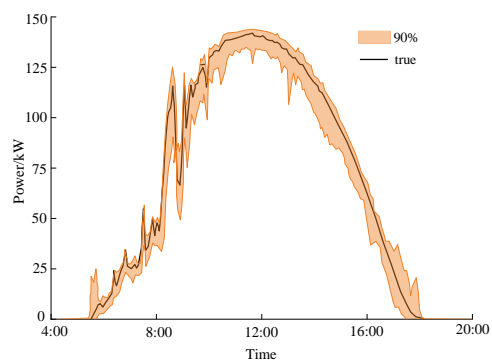
Model	90% PICP	90% PINAW	90% I
TCN-QRBiLSTM	0.9901	0.1432	0.0147
TCN-QRLSTM	0.9845	0.1567	0.0341
QRBiLSTM	0.9723	0.1631	0.0589

4.2. Cloudy Day Forecast Results

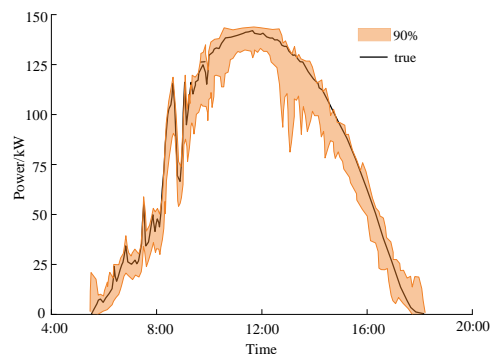
Under cloudy conditions, cloud formation and subsequent movement lead to fluctuations in solar radiation intensity, which introduce challenges for accurate forecasting. These variations can significantly affect the reliability of prediction models. Figure 6 presents the prediction intervals at the 90% confidence level for three models: TCN-QRBiLSTM, TCN-QRLSTM, and QRBiLSTM.



(a) The TCN-QRBiLSTM model



(b) The TCN-QRLSTM model



(c) The QRBiLSTM model

Figure 6. Comparison of the interval prediction results of each model on cloudy day.

As shown in the figure, the prediction intervals of TCN-QRBiLSTM are relatively narrow during stable periods and become wider during sudden changes, demonstrating the model's ability to adjust to varying uncertainty levels. Compared with the other two models, TCN-QRBiLSTM produces shorter intervals overall, indicating higher sharpness and better resolution. Furthermore, it offers improved tracking performance and responds more quickly to changes in radiation. Evaluation metrics at the 90% confidence level are summarized in Table 3, further confirming the superior accuracy and reliability of TCN-QRBiLSTM.

Table 3. The evaluation metrics of three methods under 90% confidence level on cloudy days.

Model	90% PICP	90% PINAW	90% I
TCN-QRBiLSTM	0.9553	0.1364	0.1134
TCN-QRLSTM	0.9345	0.1589	0.1467
QRBiLSTM	0.9168	0.1978	0.1371

Compared with TCN-QRBiLSTM, the PICP of TCN-QRLSTM and QRBiLSTM decreases by 2.08% and 3.85%, and the PINAW improves by 2.25% and 7.75%, respectively; and the composite evaluation index of the TCN-QRBiLSTM model is the lowest at 90% confidence level. Thus, the TCN-QRBiLSTM model has the best prediction effect on the cloudy region.

4.3. Rainy Day Forecast Results

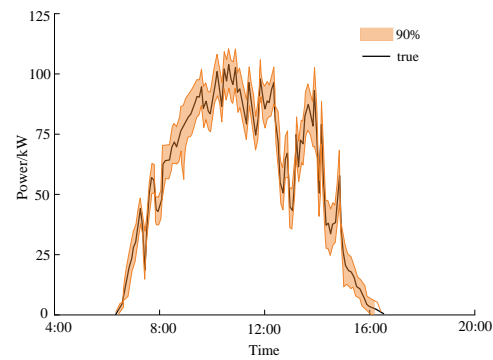
On rainy days, the sky can be obscured by clouds, especially cumulonimbus clouds, which can strongly block direct sunlight. As a result the intensity of solar radiation will fluctuate and thus the PV output power will decrease and fluctuate significantly. The predicted values of TCN-QRBiLSTM, TCN-QRLSTM and QRBiLSTM at 90% confidence range under cloudy sky are shown in Figure 7.

The predictive metrics at 90% confidence are shown in Table 4. The combined metrics of TCN-QRBiLSTM at 90% confidence are 1.42% and 2.99% lower than TCN-QRLSTM and QRLSTM, respectively. Therefore, the prediction of cloudy weather using TCN-QRBiLSTM can effectively reduce the uncertainty of PV output with strong local learning ability. It is shown that the TCN-QRBiLSTM model has a clear prediction range and high interval coverage under three kinds of weather.

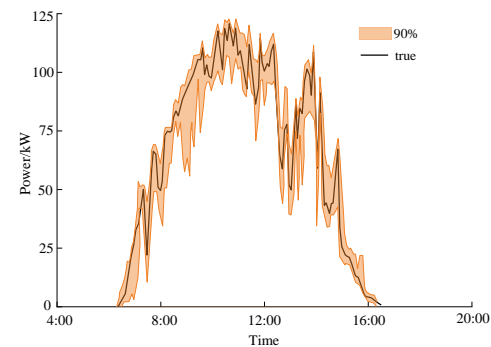
Table 4. The evaluation metrics of three methods under 90% confidence level on rainy days.

Model	90% PICP	90% PINAW	90% I
TCN-QRBiLSTM	0.9674	0.1246	0.1679
TCN-QRLSTM	0.9532	0.1678	0.2136
QRBiLSTM	0.9375	0.1895	0.1987

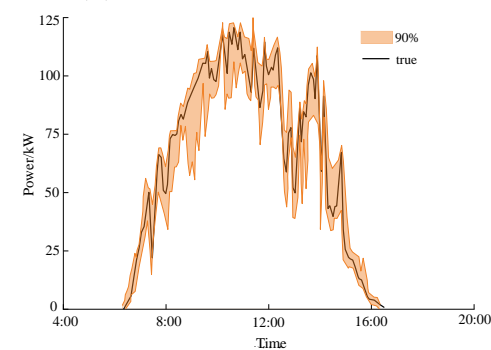
Based on Convolutional Neural Networks (CNN), the Temporal Convolutional Network (TCN) uses causal and dilated convolution layers to replace traditional convolutions. Causal convolutions allow for effective sequence learning, while dilated convolutions provide a sufficient receptive field to capture more information. TCN is more suitable for handling time-series problems compared to fully convolutional networks [20]. Figure 2 shows the causal dilated convolution structure of the TCN. Here, $\hat{n}_0 - \hat{n}_t$ are the input sequences at each time step, d is the dilation factor (which increases exponentially with the model depth), and m_t is the causal dilated convolution feature at time t .



(a) The TCN-QRBiLSTM model



(b) The TCN-QRLSTM model



(c) The QRBiLSTM model

Figure 7. Comparison of the interval prediction results of each model on rainy day.

4.4. Probabilistic Prediction Results

In this paper, data from two representative time points in the morning and afternoon were selected to describe the probability distributions under different weather conditions and times. The results obtained using kernel density estimation (KDE) are shown in Figure 8. Most of the measured data under various meteorological conditions are centrally distributed in the probability density plots, and the distribution of the points is relatively smooth. This indicates that the quantile distributions obtained using the TCN-QRBiLSTM method are reliable. After fitting the KDE for various meteorological data using three different models, the changes in their Bayesian Information Criterion (BIC) coefficients are shown in Figure 9. The BIC index of the TCN-QRBiLSTM model is the smallest, indicating that the probability density relationship obtained by this model and KDE achieves the best fit, and the model complexity is also the lowest.

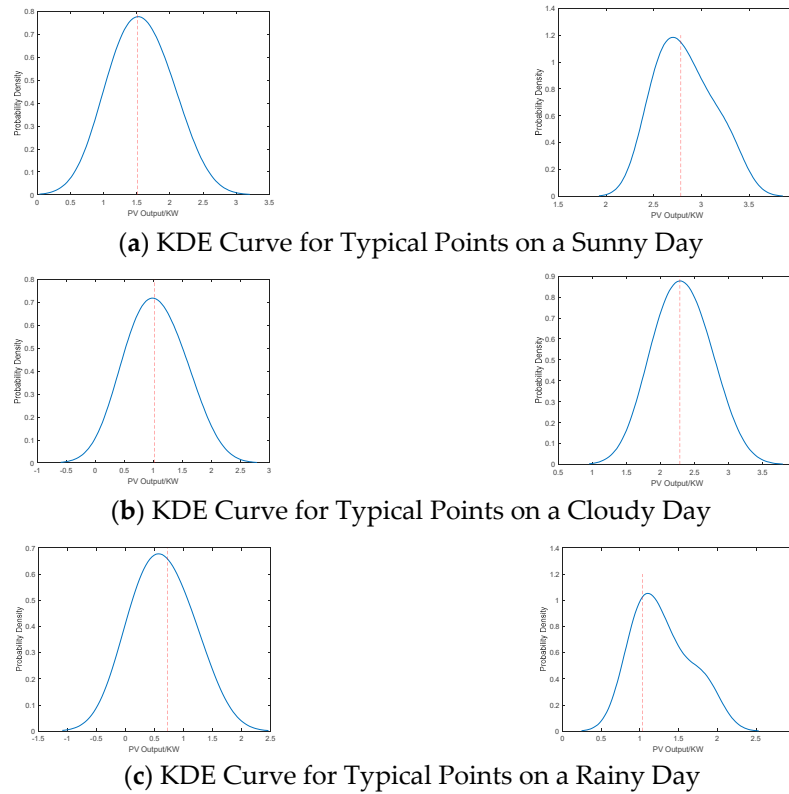


Figure 8. Probability Density Curve of the TCN-QRBILSTM Model.

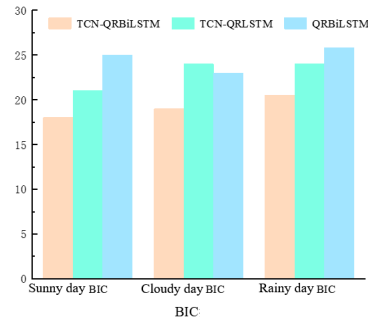


Figure 9. Three model BIC indicators.

The red dashed line represents the mean of this probability density distribution. It is the average value of the random variable characterized by this probability distribution, reflecting the central tendency of the distribution; moreover, in such a symmetric probability distribution, the position of the mean coincides with the peak point of the probability density function.

The TCN-QRBILSTM model generates confidence intervals by extracting local temporal features and capturing long-term dependencies, combined with differentiable quantile regression—resulting in quantile distributions that are not only smooth but also fully cover actual power values. Smooth probability density curves indicate that the model predictions exhibit good statistical consistency; however, adjustments are still needed by incorporating more data or external information under scenarios of extreme power fluctuations.

5. Conclusions

In order to describe and predict the uncertainties existing in interval prediction, this paper constructs a short-term PV probabilistic prediction model with TCN-QRBILSTM, and obtains the following research conclusions:

(1) Combining the TCN model with meteorological factors can effectively exploit the multidimensional stochastic characteristics such as short-term power fluctuation and long-term steady state of PV output under meteorological conditions. The reliability requirement of probabilistic prediction is satisfied both in sunny and rainy days.

(2) The combined TCN-QRBLSTM model adopts Huber's norm approximation to replace the original loss function, which makes up for the defects of the traditional interval prediction quantile loss function that cannot be minimized.

(3) The TCN-QRBLSTM combined model has high accuracy and stability. Under the conditions of 90% coverage and different meteorological conditions, the coverage of the prediction intervals of each model is very close to each other, the combined evaluation indexes of the intervals of the three weather types are the best compared with TCN-QRLSTM and QRBLSTM, and the probability density prediction indexes of BIC are the lowest while ensuring the optimal prediction intervals, so that the quality of short-term probabilistic prediction of the photovoltaic power can be improved, and therefore the proposed model in this paper has a significant superiority. The proposed model has significant superiority.

Overall, the results demonstrate that the model possesses strong adaptability, stability, and prediction accuracy under different weather conditions, reflecting its practical value in the probabilistic prediction of photovoltaic power output in complex power system environments.

6. Outlook and Future Work

Although the TCN-QRBLSTM model proposed in this paper performs excellently in short-term photovoltaic power probabilistic prediction, there is still room for further improvement. Future work can be expanded in the following aspects:

(1) Improvement of Data Accuracy: This can be achieved by incorporating meteorological data with higher temporal and spatial resolution to enhance the model's responsiveness to rapid power changes.

(2) Optimization of Model Structure: This can be achieved by integrating attention mechanisms or graph neural networks to strengthen the model's adaptive weight assignment for features across different time periods and spatial locations, thereby further improving prediction accuracy.

(3) Cross-Regional Application and Transfer Learning: This can be achieved by investigating the model's applicability in different geographical regions and exploring transfer learning methods to enhance the model's generality.

(4) Multi-Time-Scale Joint Prediction: This can be achieved by extending the model to day-ahead or multi-day prediction scenarios and realizing the integration of short-term and medium-term probabilistic prediction to increase the reference value for decision-making in photovoltaic power dispatch and system planning.

Author Contributions: Conceptualization, W.Y. and J.D.; methodology, L.Z.; software, J.N.; validation, W.Y., J.D. and L.Z.; formal analysis, W.Y.; investigation, J.D.; resources, J.N.; data curation, W.Y.; writing—review and editing, Q.Z.; visualization, L.Z.; supervision, Q.Z.; funding acquisition, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the State Grid Anhui Electric Power Co., Ltd. Science and Technology Project (B3120524003N) for its support on the research of novel distribution network optimization and scheduling technology based on distributed source-load probabilistic forecasting.

Data Availability Statement: The availability of these data is restricted. The data were obtained from the State Grid Electric Power Research Institute and, with the permission of the State Grid Electric Power Research Institute, can be obtained from the authors.

Conflicts of Interest: Authors Weibo Yuan, Jinjin Ding, Li Zhang and Jingyi Ni were employed by the company Electric Power Research Institute. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Nomenclature

Symbol	Definition	Unit/Description
PICP	Prediction Interval Coverage Probability	Dimensionless (0–1)
PINAW	Percentage Interval Normalized Average Width	Dimensionless (0–1)
I	Comprehensive evaluation index for interval prediction	Dimensionless (0– ∞); smaller = better
BIC	Bayesian Information Criterion	Dimensionless; smaller = better
α	Quantile level in quantile regression	Dimensionless (0–1)
$q_{\alpha,t}$	Quantile of PV power at time t and quantile level (α)	kW
δ	Threshold of Huber norm in improved quantile regression	Set to 1.0 (small positive value)
$\rho_{\alpha}(\cdot)$	Huber pinball loss function	Dimensionless
d	Dilation factor of the TCN	Integer (e.g., [1,2,4,8])
K	Total number of convolutional filters in the TCN	Integer (set to 64 in this study)
k_i	Size of the i -th convolutional filter in the TCN	Integer (e.g., 3)
f_t	Causal dilated convolution feature of the TCN at time t	Dimensionless (feature vector)
m_t	Abstract feature of the TCN after Batch Normalization (BN) at time t	Dimensionless (feature vector)
\hat{y}_t^{BiLSTM}	Output vector of BiLSTM at time t	Dimensionless (feature vector)
h	Bandwidth in KDE	Adaptive (determined by data distribution)
$K(\cdot)$	Kernel function in KD	Dimensionless
N	Number of training samples	Integer
y_t	Actual PV power at time t	kW
\hat{y}_t	Predicted PV power at time t	kW
x_t	Input vector of the prediction model at time t	Dimensionless (feature vector)
μ	Mean of causal dilated convolution feature f_t in BN layer	Dimensionless
σ^2	Variance of causal dilated convolution feature f_t in BN layer	Dimensionless
γ, β	Learnable parameters of the BN layer	Dimensionless
ϵ	Small positive constant in the BN layer	Set to 10^{-6} (dimensionless)

References

- National Energy Administration. China's New-Type Energy Storage Development Report (2025) Released by the National Energy Administration. *China Nonferrous Metall.* **2025**, *54*, 146.
- Sangrody, H.; Zhou, N.; Zhang, Z. Similarity-Based Models for Day-Ahead Solar PV Generation Forecasting. *IEEE Access* **2020**, *8*, 104469–104478. [[CrossRef](#)]
- Mei, F.; Gu, J.; Lu, J.; Lu, J.; Zhang, J.; Jiang, Y.; Shi, T.; Zheng, J. Day-Ahead Nonparametric Probabilistic Forecasting of Photovoltaic Power Generation Based on the LSTMQRA Ensemble Model. *IEEE Access* **2020**, *8*, 166138–166149. [[CrossRef](#)]
- Tang, D.; Li, J.; Zeng, F.; Li, Y.; Yan, C. Bayesian parameter estimation of SST model for shock wave-boundary layer interaction flows with different strengths. *Chin. J. Aeronaut.* **2023**, *36*, 217–236. [[CrossRef](#)]
- Dai, Y.; Lu, Z.; Xiong, W.; Yuan, X.; Xu, Y.; Tan, Z. Power quality disturbance classification method based on CDAE and TCN/BLSTM model. *Smart Power* **2023**, *51*, 59–66.
- Huang, Q.; Wei, S. Improved quantile convolutional neural network with two-stage training for daily-ahead probabilistic forecasting of photovoltaic power. *Energy Convers. Manag.* **2020**, *220*, 113085. [[CrossRef](#)]
- Xiang, M.J. Probabilistic Prediction of Photovoltaic Power Generation Based on Ground-Based Cloud Mapping. Master's Thesis, Zhejiang University, Hangzhou, China, 2022.
- Li, J.; Liu, Q. Short-term Photovoltaic Power Forecasting Using SOM-based Regional Modelling Methods. *Chin. J. Electr. Eng.* **2023**, *9*, 158–176. [[CrossRef](#)]
- Gao, Y.; Wu, H.; Zhang, J.; Zhang, H.; Zhang, P. A probabilistic prediction model for photovoltaic power day-ahead based on combinatorial deep learning. *China Power* **2024**, *57*, 100–110.
- Pang, H.; Gao, J.; Du, Y. A short-term load probability density forecasting method based on time-convolution network quantile regression. *Grid Technol.* **2020**, *44*, 1343–1350.

11. Wan, C.; Cui, W.; Song, Y. Probabilistic prediction of new energy power systems: Basic concepts and mathematical principles. *Chin. J. Electr. Eng.* **2021**, *41*, 6493–6509.
12. Gao, Q. Multi-temporal Scale Wind Power Forecasting Based on Lasso-CNN-LSTM-LightGBM. *EAI Endorsed Trans. Energy Web* **2024**, *11*, 5792. [[CrossRef](#)]
13. Karthiga, M.; Sountharajan, S.; Suganya, E.; Sankarananth, S. Sentence semantic similarity model using convolutional neural networks. *EAI Endorsed Trans. Energy Web* **2021**, *8*, 35.
14. Yao, C.W.; Yang, P.; Liu, Z.J. A load forecasting method based on CNN-GRU hybrid neural network. *Power Grid Technol.* **2020**, *44*, 3416–3424.
15. Wang, K.; Du, H.; Jia, R.; Liu, H.; Liang, Y.; Wang, X. Short-term interval probability prediction of photovoltaic power based on similar day clustering and QR-CNN-BiLSTM model. *High Volt. Technol.* **2022**, *48*, 4372–4388.
16. Liu, H.; Ling, N.; Luo, Z.; Sun, Z. Power grid short-term load forecasting method based on TCN- LSTM and meteorological similar day sets. *Smart Power* **2022**, *50*, 30–37.
17. Zhang, Z.; Duan, Z.; Zhang, L. Photovoltaic power generation prediction and optimization configuration model based on GPR and improved PSO algorithm. *EAI Endorsed Trans. Energy Web* **2024**, *11*, 3809. [[CrossRef](#)]
18. Usha, S.; Geetha, P.; Geetha, A.; Balamurugan, K.S.; Selciya, S. A novel concept of solar photovoltaic partial shading and thermal hybrid system for performance improvement. *EAI Endorsed Trans. Energy Web* **2024**, *11*, 4943. [[CrossRef](#)]
19. Singh, P.; Jha, M.; Sharaf, M.; El-Meligy, M.A.; Gadekallu, T.R. Harnessing a Hybrid CNN-LSTM Model for Portfolio Performance: A Case Study on Stock Selection and Optimization. *IEEE Access* **2023**, *11*, 104000–104015. [[CrossRef](#)]
20. Zhang, Y.; Hu, T. Ensemble Interval Prediction for Solar Photovoltaic Power Generation. *Energies* **2022**, *15*, 7193. [[CrossRef](#)]
21. Obiora, C.N.; Ali, A. Hourly Photovoltaic Power Forecasting Using CNN-LSTM Hybrid Model. In Proceedings of the 2021 62nd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), Riga, Latvia, 14–15 October 2021; pp. 1–6. [[CrossRef](#)]
22. Liu, Y.; Zhang, Q.; Li, R.; Hui, H.; Fan, L.; Wu, J. PV power probability density prediction based on the long short-term memory network with quantile regression. In Proceedings of the 2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA), Ningbo, China, 18–22 August 2023; pp. 800–804. [[CrossRef](#)]
23. Cao, R.; Wang, K.; Wu, F. Research on the Probability Prediction Method of Electricity Price Response Load based on CNNLSTM and Quantile Regression Fusion. In Proceedings of the 2023 IEEE 7th Conference on Energy Internet and Energy System Integration (EI2), Hangzhou, China, 15–18 December 2023; pp. 1–9. [[CrossRef](#)]
24. Nguyen, C.; Hoang, T.M.; Cheema, A.A. Channel Estimation Using CNN-LSTM in RIS-NOMA Assisted 6G Network. *IEEE Trans. Mach. Learn. Commun. Netw.* **2023**, *1*, 43–60. [[CrossRef](#)]
25. Selim, T.; Elkabani, I.; Abdou, M.A. Students Engagement Level Detection in Online e-Learning Using Hybrid EfficientNetB7 Together With TCN, LSTM, and Bi-LSTM. *IEEE Access* **2022**, *10*, 99573–99583. [[CrossRef](#)]
26. Yang, Y.; Zhang, H.; Peng, S.; Su, S.; Li, B. Wind Power Probability Density Prediction Based on Quantile Regression Model of Dilated Causal Convolutional Neural Network. *Chin. J. Electr. Eng.* **2023**, *9*, 120–128. [[CrossRef](#)]
27. Wei, S.; Liu, X.; Shi, H.; Gan, J. Remaining Useful Life Prediction of High-Dimensional Kernel Density Estimation with Adaptive Relative Density Window Width Considering Multisource Information Fusion. *IEEE Sens. J.* **2024**, *24*, 6548–6563. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.