Energy Proceedings
Vol 49, 2024
ISSN 2004-2965

Wind speed ensemble forecasting framework based on deep neural network and dropout mechanism#

Zhendong Zhang [1*], Huichao Dai [1], Qing Zhang [1]
1 China Three Gorges Corporation, Wuhan, Hubei, China
(Corresponding Author: zhendongz@alumni.hust.edu.cn)

ABSTRACT

With the depletion of non-renewable energy, renewable energy such as wind energy has received more and more attention. Wind speed prediction plays an important role in promoting the utilization of wind energy. This paper focuses on how to realize the wind speed ensemble prediction at multiple stations. First, convolutional neural networks are introduced to wind speed prediction because of its ability to mine input features. Then, the dropout mechanism is incorporated into the model so that multiple runs can obtain multiple predictions. Next, the kernel density estimation method is used to obtain the probability density function of wind speed prediction. At the same time, in order to complete the multi-station wind speed prediction at the same time, the four-dimensional input-output structure tensor is proposed for wind speed prediction. Finally, the model proposed in this study is verified on two datasets of Tibet, China. The experimental results show that: (1) The model proposed in this study can obtain high-accuracy deterministic prediction results and appropriate ensemble prediction intervals. (2) The appropriate dropout ratio is important to neither overfit nor reduce the prediction accuracy.

Keywords: wind speed, ensemble forecasting, Convolutional neural network, dropout, Kernel density estimation

NONMENCLATURE

Abbreviations	
CNN	Convolutional neural network
ConvLSTM	Convolutional Long Short-Term Memory Network
ConvGRU	Convolutional Gated Recurrent Unit
KDE	Kernel density estimation
PDF	Probability density function
R^2	deterministic coefficient
MAE	mean of absolute error
RMSE	root mean square error
CRPS	continuous ranked probability score
Symbols	
x	input
w	weight
b	bias
y	predictions

1. INTRODUCTION

The randomness, volatility and uncertainty of wind speed increase the difficulty of wind speed prediction [1]. Wind speed prediction methods can be generally divided into three categories: physical-mechanism-driven methods, data-driven methods, and hybrid methods [2]. Physical mechanism-driven methods use mathematical physics equations to simulate and predict wind speed processes, such as numerical weather prediction [3]. Liu et al. proposed a method combining series-wise mechanism, temporal lag attention and numerical weather prediction for wind speed prediction and improved the accuracy [4]. Data-driven methods predict wind speed by mining the relationship between meteorological factors and wind speed, such as statistical models, machine learning and deep learning models [5]. Considering wind speed characteristics, Zhang et al used empirical mode decomposition and autoregressive moving average model to complete wind speed prediction, and the prediction results can lay a foundation for power grid dispatching [6]. Moreno et al. enhanced wind speed prediction through the synergistic effect of machine learning, singular spectrum analysis, and variational mode decomposition [7]. Zhang et al. obtained probabilistic forecasting results of wind speed by combining a shared weighted long-short term memory network with a Gaussian process regression model [8]. The hybrid method combines the advantages of multiple models to obtain more accurate wind speed

This is a paper for the 16th International Conference on Applied Energy (ICAE2024), Sep. 1-5, 2024, Niigata, Japan.

prediction results. Lv and Wang proposed a hybrid model combined deep learning, time series decomposition and multi-objective parameter optimization for wind speed forecasting [9]. Chen et al. employed multi-resolution feature fusion and frequency information mining for multi-step short-term wind speed predictions [10].

Most of the existing studies focus on the wind speed prediction of a single station, but few studies have completed the prediction of multiple stations at the same time. Considering that there are always errors in prediction, it is also a research hotspot to quantify the uncertainty of prediction by implementing ensemble prediction [11]. Therefore, this paper focuses on how to achieve multi-station wind speed ensemble forecast.

In this study, an ensemble forecasting framework combining convolutional neural network (CNN), dropout mechanism, and kernel density estimation (KDE) method is proposed for multi-station wind speed prediction. The remaining sections of this paper are as follows: the forecasting framework is introduced in Section 2. The forecasting metrics are described in Section 3. A case study is completed in Section 4. The conclusions are summarized in Section 5.

2. ENSEMBLE WIND SPEED FORECASTING FRAMEWORK

2.1 Convolutional neural network

Convolution neural network have a strong ability to mine input features through convolutional operations. The process of convolution is shown in Fig. 1, and its output can be calculated as follows:

$$x_{t,j}^{l+1} = F\left(\sum_{q=1}^{k^l} w_{t,q}^l x_{t,r}^l + b_{t,j}^l\right); \quad r=(j-1)\cdot s^l+q; \quad j=1,2,\cdots,n^{l+1} \quad (1)$$

where $x_{t,r}^l$ and $x_{t,j}^{l+1}$ are input and output of l-th layer in t-th periods. $w_{t,q}^l$ and $b_{t,j}^l$ are weights and bias. F is activation function. n^l and n^{l+1} are input length of l-th and (l+1)-th layer. k^l and s^l are kernel size and stride length.

2.2 Dropout mechanism

Dropout is an important technique for neural network to avoid overfitting. During each batch training process, the neural network nodes are temporarily discarded from the network with a certain probability, which reduces the training time of the model on the one hand and avoids overfitting on the other hand. Since each batch of network nodes is randomly discarded, the model cannot arbitrarily fit all the random noises in the training set during the training process. The schematic diagram of the dropout mechanism is shown in Fig. 2.

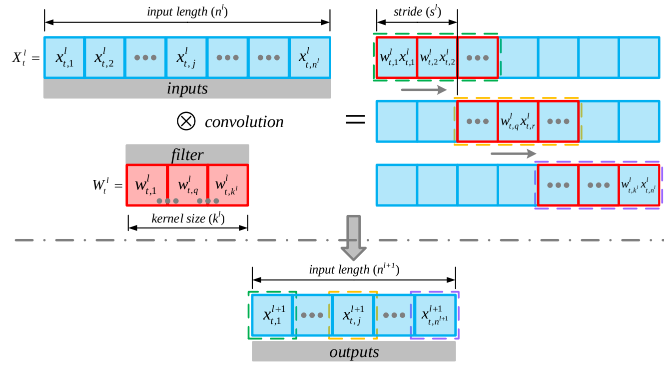
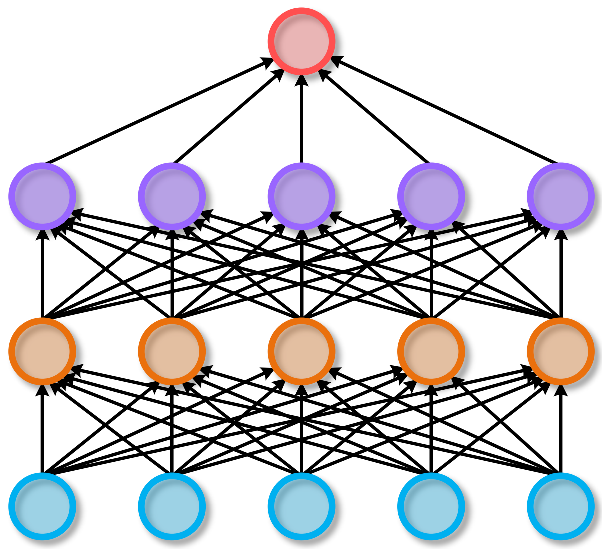
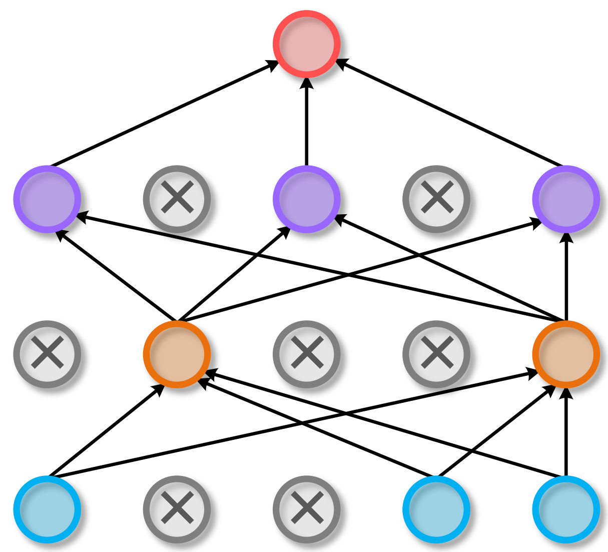


Fig. 1 Convolutional neural network



(a) Without dropout mechanism



(b) With dropout mechanism

Fig. 2 Dropout mechanism

The dropout mechanism continues to be retained during the prediction process, and part of the network nodes are still randomly dropped during each prediction, so that each prediction result has a slight difference. Multiple prediction results in ensemble forecasting can be obtained by running the prediction process several times.

2.3 Kernel density estimation

Kernel density estimation (KDE) method does not need the prior knowledge of the data distribution to be estimated, and does not attach any assumptions to the data distribution. It is a non-parametric estimation method to study the data distribution characteristics from the data sample itself.

Assuming that the M prediction results ($[y_{t,1},y_{t,2},\cdots,y_{t,M}]$) are obtained by CNN and dropout, its probability density function (PDF) is estimated as follows:

$$\hat{f}_t(y)=\frac{1}{MB}\sum_{m=1}^{M}K\left(\frac{y_{t,m}-y}{B}\right) \tag{2}$$

where $\hat{f}_t(y)$ is the PDF of the prediction for the t-th period. B is the band width. K is the kernel function and the Epanechnikov function is usually used.

2.4 Ensemble forecasting framework

A framework for multi-station wind speed ensemble forecasting is proposed in this study, as shown in Fig. 3. In the training part, the input layer is arranged into a 4-d tensor ($[T,M,N,Q]$), where M and N are rows and columns of the spatial station, T is the number of historical periods, and Q is the number of input factors. Three convolutional neural network layers are considered in the hidden layer, and dropout layers are added between the convolutional neural network layers. The output layer is the wind speed prediction results for multiple stations. In the prediction part, multiple prediction results are obtained by running the prediction process several times, and then the probability density function is obtained by using the KDE method.

3. FORECASTING METRICS

The deterministic coefficient (R^2), the mean of absolute error (MAE), the root mean square error (RMSE) [12] are used to evaluate deterministic prediction accuracy. The continuous ranked probability score (CRPS) [13] is used to evaluate the ensemble forecasting comprehensive performance.

4. CASE STUDY

4.1 Research object and data

The study area is located in Tibet, China, with longitude varying from 81.45°E to 83.15°E and latitude from 31.55°N to 32.85°N, which covers 16 wind stations. Two datasets for 2014 are used, and one period is 1 hour.

4.2 Results and discussion

4.2.1 Comparison of forecasting metrics

Convolutional Long Short-Term Memory Network (ConvLSTM) and Convolutional Gated Recurrent Unit Network (ConvGRU) are used to compared with CNN, and these models (CNN-D, ConvLSTM-D, ConvGRU-D) combine dropout mechanisms (dropout ratio = 0.05) to complete ensemble forecasting.

The forecasting metrics of dataset 1 and dataset 2 are shown in Table. 1 and Table. 2, respectively. In dataset 1, R^2 of CNN-D is 0.972, the largest among the three models, and MAE and RMSE are 0.177m/s and 0.232m/s, respectively, the smallest among the three models, indicating that CNN-D model has the highest prediction accuracy. The CRPS of CNN-D is 0.112, which is the smallest among the three models, indicating that CNN-D model has the best ensemble prediction performance. The analysis for the metrics of dataset 2 are similar to dataset 1.

Table. 1 Forecasting metrics of dataset 1

model	R^2	MAE	RMSE	CRPS
CNN-D	**0.972**	**0.177**	**0.232**	**0.112**
ConvLSTM-D	0.968	0.192	0.250	0.116
ConvGRU-D	0.959	0.230	0.282	0.124

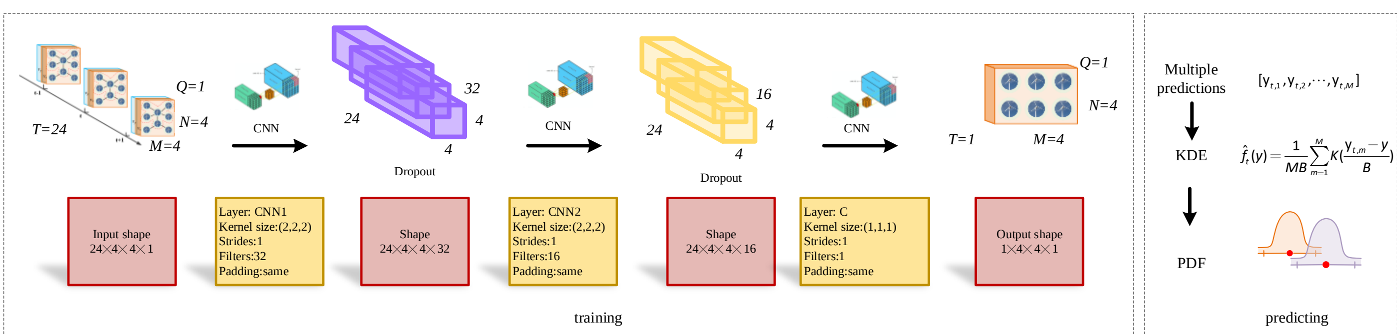


Fig. 3 Ensemble forecasting framework

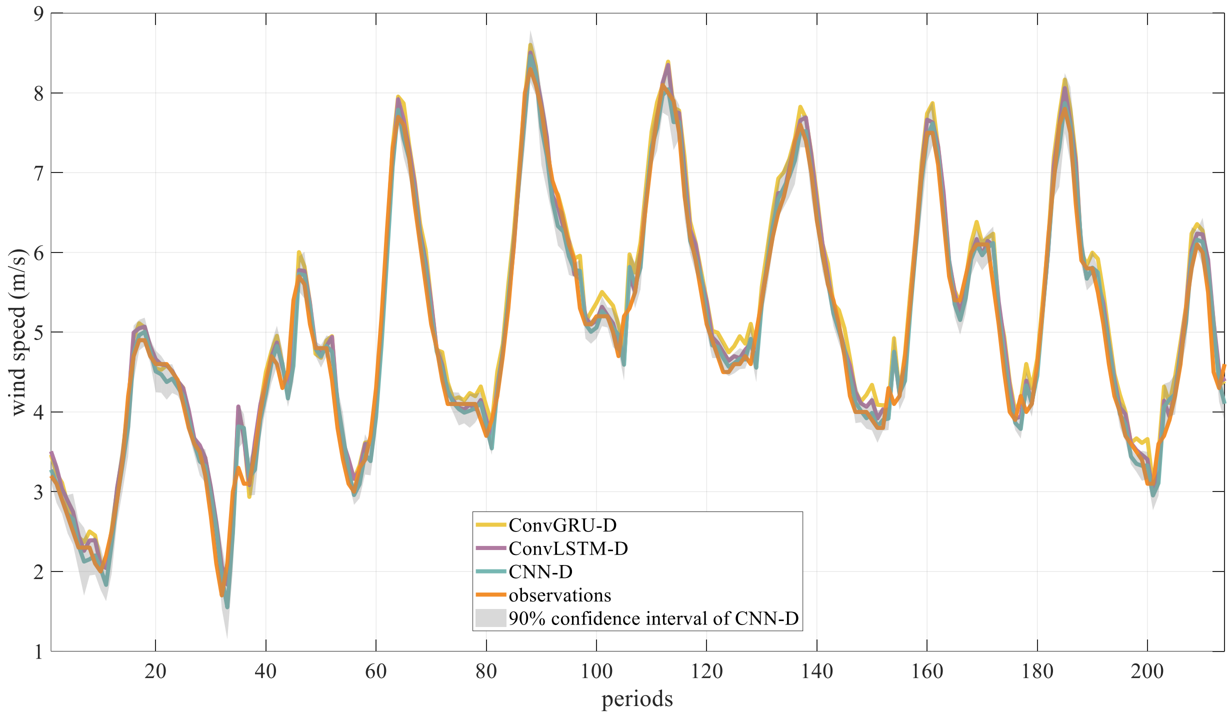
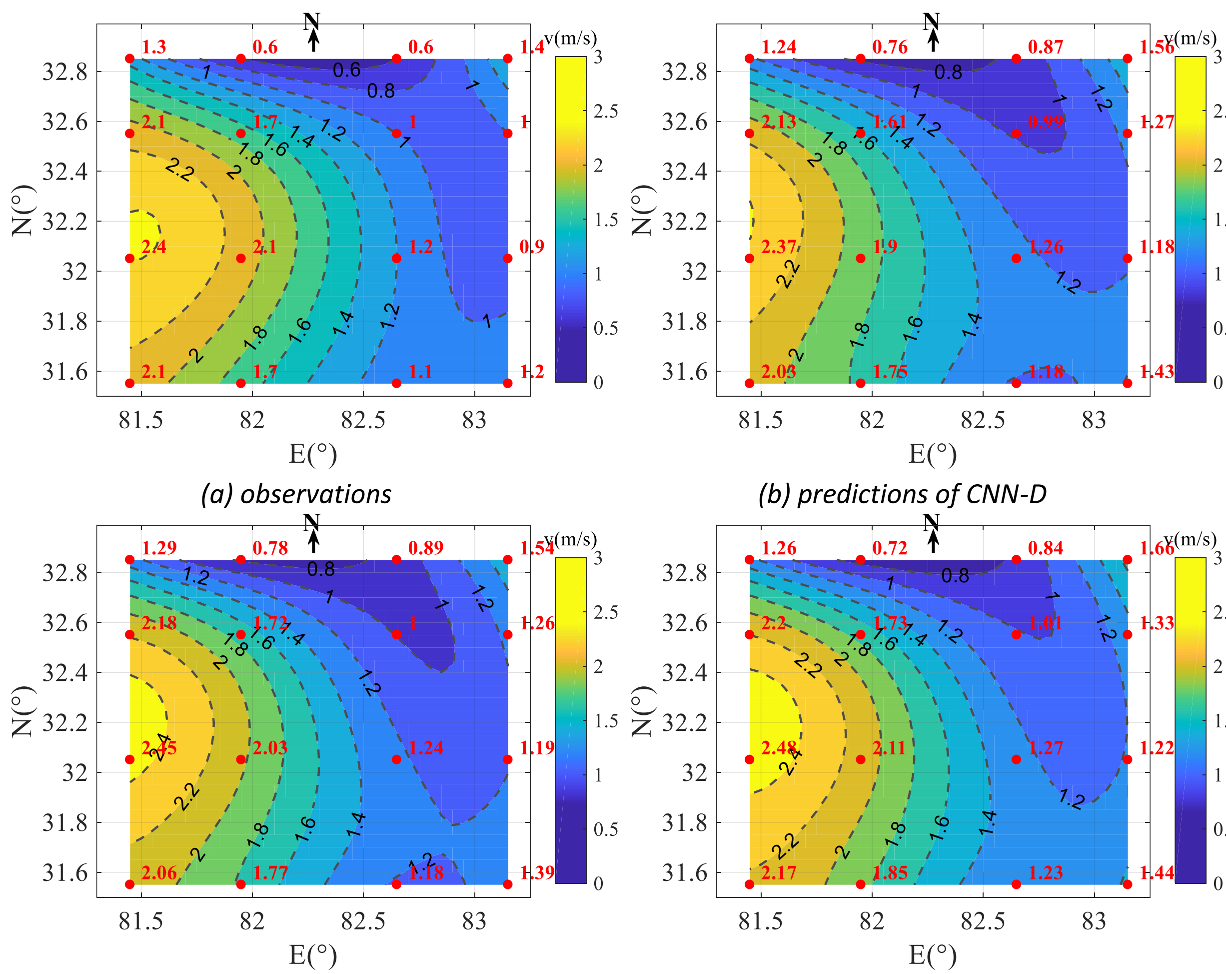


Fig. 4 Predictions of station 11 on dataset 1



(a) observations *(b) predictions of CNN-D*

(c) predictions of ConvLSTM-D *(d) predictions of ConvGRU-D*

Fig. 5 Predictions of 142-nd period on validation set 2

Table. 2 Forecasting metrics of dataset 2

model	R^2	MAE	RMSE	CRPS
CNN-D	**0.976**	**0.145**	**0.199**	**0.077**
ConvLSTM-D	0.974	0.153	0.209	0.081
ConvGRU-D	0.973	0.160	0.210	0.079

4.2.2 Comparison of prediction results

Taking station 11 of dataset 1 as an example, the predictions of the three models and observations are drawn in Fig. 4. It can be intuitively seen that the predictions of the CNN-D model are closest to the observations, indicating that the CNN-D model has the highest prediction accuracy. At the same time, the prediction interval of CNN-D model is wide in the period with large prediction error, and narrow in the period with small error, indicating that the ensemble prediction interval of CNN-D model is suitable.

4.2.3 Prediction results of multiple stations

Taking the 142nd period of validation set 2 as an example, the multi-station prediction results and observations are plotted in Fig. 5. It can also be intuitively seen that the multi-station wind speed prediction distribution of CNN-D model is closest to the distribution of observations, indicating that CNN-D model has the highest prediction accuracy.

4.2.4 dropout ratio analysis

By changing the dropout ratio from 0 to 0.6 in steps of 0.05, the prediction accuracy of the CNN-D model on the two datasets is shown in Fig. 6. When the dropout mechanism is not used (dropout ratio=0), there is a certain overfitting phenomenon, and the prediction accuracy is lower than that of scenarios with dropout ratio of 0.05. However, with the increase of dropout ratio, more and more nodes are abandoned during the training process, resulting in a decline in prediction accuracy. Therefore, an appropriate dropout ratio is important to ensure neither overfitting nor degradation of prediction accuracy.

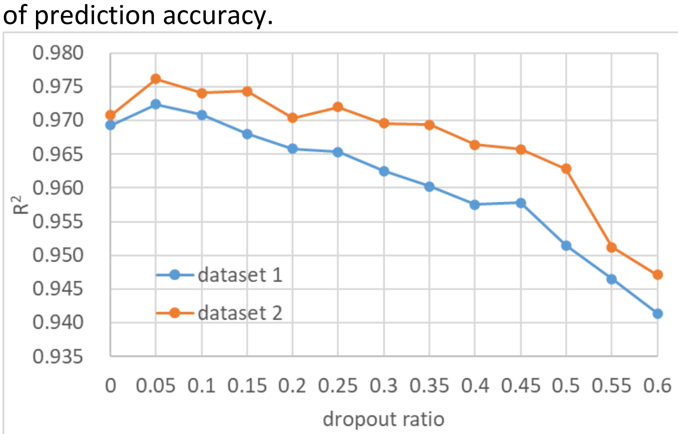


Fig. 6 Dropout ratio analysis

5. CONCLUSIONS

In this study, a multi-station ensemble wind speed forecasting framework based on the convolutional neural network, dropout mechanism and kernel density estimation method is proposed and verified on the Tibet datasets. The experimental results show that the CNN-D model has high deterministic prediction accuracy and suitable ensemble prediction interval.

ACKNOWLEDGEMENT

This study is supported by National Natural Science Foundation of China (No. 52409007). Special thanks to the anonymous reviewers and editors for their meaningful comments.

REFERENCE

[1] Heng J, Hong Y, Hu J, Wang S. Probabilistic and deterministic wind speed forecasting based on non-parametric approaches and wind characteristics information. Appl Energ 2022;306: 118029.

[2] Wang H.Z, Wang G.B, Li G.Q, Peng J.C, Liu Y.T. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energ 2016;182: 80-93.

[3] Zhang J, Draxl C, Hopson T, Monache LD, Vanvyve E, Hodge B. Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods. Appl Energ 2015;156:528-541.

[4] Liu C, Zhang X, Mei S, Zhou Q, Fan H. Series-wise attention network for wind power forecasting considering temporal lag of numerical weather prediction. Appl Energ 2023;336: 120815.

[5] Khodayar M, Wang J. Spatio-temporal graph deep neural network for short-term wind speed forecasting. IEEE Trans Sustain Energ 2019;10:670–81.

[6] Zhang Y, Zhao Y, Kong C, Chen B. A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. Energ Convers Manage 2020;203: 112254.

[7] Moreno S.R, Seman L.O, Stefenon S.F, Coelho L.S, Mariani V.C. Enhancing wind speed forecasting through synergy of machine learning, singular spectral analysis, and variational mode decomposition. Energ 2024;292: 130493.

[8] Zhang Z, Ye L, Qin H, Liu Y, Wang C, Yu X, et al. Wind speed prediction method using Shared Weight Long Short-Term Memory Network and Gaussian Process Regression. Appl Energ 2019;247:270–84.

[9] L S.X, Wang L. Deep learning combined wind speed forecasting with hybrid time series decomposition and

multi-objective parameter optimization. Appl Energ 2022;311: 118674.
[10] Chen Q, He P, Yu C, Zhang X, He J, Li Y. Multi-step short-term wind speed predictions employing multi-resolution feature fusion and frequency information mining. Renew Energ 2023;215:118942.
[11] Zhang Z, Qin H, Liu Y, Yao L, Yu X, Lu J, et al. Wind speed forecasting based on Quantile Regression Minimal Gated Memory Network and Kernel Density Estimation. Energ Convers Manage 2019;196:1395–409.
[12] Liu Y, Qin H, Zhang Z, Pei S, Jiang Z, Feng Z, et al. Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. Appl Energ 2020;260:114259.
[13] Zhang Z, Tang H, Qin H, Luo B, Zhou C, Zhou H. Multi-step ahead probabilistic forecasting of multiple hydrological variables for multiple stations. J Hydrol. 2023;617.