Journal of Networking Technology

DLINE JOURNALS

Print ISSN: 0976-898X
Online ISSN: 0976-8998

JNT 2025: 16 (2)
https://doi.org/10.6025/jnt/2025/16/2/62-79

Ultra-short Time Surface Wind Prediction in Kumtag Desert Region of Xinjiang Based on Deep Learning

Xinjie Zuo[a], Chang Xu[b], Zhengbo Wang[c] and Yanfei Long*[d]
Northwest Institute of Nuclear Technology. China
[a]zuoxinjie_2023@163.com, [b]444178455@qq.com,
[c]954702331@qq.com, *[d]lyf_dyx2022@163.com

ABSTRACT

Under the influence of geographical environment and seasonal differences, wind speed changes show strong characteristics of volatility, intermittency and high variability. To make better use of wind energy in the Kumtag Desert of Xinjiang, this paper proposes a Conv-Informer model and a loss function including a trend penalty factor, combined with a data set partitioning strategy. It researches ultra-short time surface wind prediction in the Kumtag Desert based on deep learning. The results show that the data set partitioning strategy has a crucial impact on the effect of model training. In the results of Conv-Informer model training, the prediction accuracy of the wind speed above 10 m/s within 30 minutes is more than 93%, and the linear trend of wind speed change can be predicted more accurately.

Keywords: Deep Learning, Wind Speed Forecast

Received: 27 October 2024, Revised 28 December 2024, Accepted 16 January 2025

1. Introduction

As one of the most basic meteorological elements of atmospheric detection, wind plays a vital role in the Earth's atmospheric circulation, energy exchange and human production and life [1]. Influenced by geographical environment, seasonal differences and other conditions, wind speed changes show strong volatility characteristics, intermittent and highly variable, and wind speed prediction has become an essential task in weather forecasting, disaster prevention and reduction, energy strategy and other business fields [2-3].

The main methods of wind speed forecast are physical forecast based on Numerical Weather Prediction (NWP) and statistical analysis forecast based on historical data. Among them, physical prediction based on numeri cal weather prediction can solve atmospheric motion and thermodynamic equations in combination with physical conditions such as geographical location and surrounding environment, when the initial and boundary values are given, to predict future wind speed changes. However, the atmospheric physical motion process described by NWP is limited, and some parameterization schemes are provided by approximate relations or empirical formulas [4]. Moreover, many empirical assumptions are used in data assimilation to generate the initial field, making it impossible for the initial field of the numerical model to be absolutely accurate [5]. Therefore, NWP wind speed prediction inevitably has specific model errors and systematic deviations. Statistical analysis and prediction based on historical data often use the autoregressive moving average model (ARMA) [6], the autoregressive integrated moving average model (ARIMA) [7], and other statistical models. However, statistical analysis and prediction require high-quality original data, and it is difficult to effectively extract complex nonlinear features of wind speed series, so the prediction accuracy is still limited [8].

In recent years, deep learning has been proven to play an important role in the research of weather and climate prediction. Numerous studies have shown that deep learning methods can be extended from conventional machine learning tasks to atmospheric science [9-13]: Rasp et al. use 9-layer fully connected neural networks to train the parameterization process of climate models. It has positive significance for accelerating the parameterization calculation process in the model [14]. Scher et al. used convolutional neural networks (CNN) to simulate the whole physical and dynamic process of the general circulation model (GCM) [15]. Moraux's deep learning model, based on multiple tasks, realised the estimation of instantaneous precipitation probability and precipitation for satellite infrared radiometer data [16]. Research on ultra-short time prediction of surface wind speed using deep learning technology is mainly seen in the field of wind power generation: Huang et al. combined improved ICEEMDAN with adaptive noise, sample entropy (SE), optimized cyclic generalized learning system (ORBLS) and expanded sequential convolutional network (BTCN) to improve the accuracy of ultra-short time wind speed prediction [17]. Liang et al proposed a short-term wind speed prediction model combining capsule neural network (Capsnet) and Bidirectional Long Short-term Memory network (BILSTM) with Multi-object Harris Eagle Optimization (MOHHO). Historical wind speed information and multi-dimensional meteorological variables were input into the model to predict wind speed. However, the wind speed of the verification data set was mostly below 8 m/s. It is difficult to ensure the accuracy of strong wind prediction [18]; Ye et al. proposed a DynamicNet model for wind speed prediction based on an encoder-decoder architecture, and proposed corresponding multi-step prediction steps. However, their model was applied to grid data rather than single-station data, and their multi-step prediction steps were also cyclic iterations of other single-step predictions [19].

Located in the eastern part of Tarim Basin, the Kumtag Desert region of Xinjiang is a temperate desert with the Gobi Desert as its underlying surface [20]. Meteorological observation data in this region are scarce, wind energy is abundant, and great potential exists. However, there is almost no research on wind speed prediction in this region. To better use the wind energy in the Kumtag Desert area of Xinjiang, this paper focuses on the prediction accuracy of wind over 10 m/s and its changing trend. Considering that it is difficult for physical prediction based on NWP and statistical analysis and prediction based on historical data to describe the change of ultra-short time wind speed accurately, this paper uses deep learning technology and a data set partitioning strategy to research ultra-short time surface wind prediction in the Kumtag Desert.

2. Experimental Data

2.1 Automatic Weather Station Data

The movement of the atmosphere is continuous in time and space, so the wind speed of a single automatic weather station must be related to the past state in time and to the upstream and downstream stations in space. This paper selects the data of 9 upstream and downstream sites (including the sites to be predicted) as the input for model training. The experimental data are the sliding average temperature, air pressure, humidity, wind speed and wind direction data of 9 stations from December 1, 2020 to November 30, 2022, and the wind speed is decomposed into u and v components.

2.2 Data set Partitioning

To make the network model more fully learn the comprehensiveness of the input data, three different data set partitioning strategies were tried during model training:

Strategy 1

Divide the training set, verification set and test set according to the ratio of 7:2:1 by moving average data for every 5 minutes. Each data set is divided in chronological order to avoid the model learning the data features of the test set in advance during training. The model inputs 36 times steps (each time step length is 5 minutes), and outputs 6 times steps (output wind speed every 5 minutes within 30 minutes);

Strategy 2

Under the premise that the data continuous time step is 42 steps (each time step is 5 minutes), the data volume of the average wind speed of the last 6 times steps is about 462,000, 223,000 and 45,000, respectively, with 0~5 m/s, 5~10 m/s and more than 10 m/s. To enable the model to learn more wind-related features, 45,000 sets of data were selected from each of the three wind speed intervals and the training set, verification set and test set were divided according to the ratio of 7:2:1. The model input 36 times steps and output 6 times steps (that is, output wind speed every 5 minutes within 30 minutes);

Strategy 3

It is consistent with strategy 2, but to make the model learn the fluctuations on a longer time scale as much as possible, the model input is changed to 138 times steps. That is, the model input of strategies 1 and 2 is 3 hours of data, while strategy 3 is 11.5 hours.

Information about the three strategies is shown in the following table:

Data set	**Data volume**	**Period**
Training set	515,000	2020/12/01-2022/05/07
Validation set	146,000	2022/05/08-2022/09/23
Test set	73,000	2022/09/24-2022/11/30

Table 1. Data volume distribution in Strategy 1

In the model training stage, the data volume of each traversal loading training set and verification set is 50,000 and 10,000 respectively.

Data set	Data volume	Period
Training set	93,900 (31,300 for each of the three wind speed ranges)	Average wind speed 0-5m/s: 2020/12/02-2022/02/27 Average wind speed 5-10m/s: 2020/12/01-2022/05/29 Average wind speed 10-15m/s: 2021/01/12-2022/05/29
Validation set	26,800 (0.89 million data for each of the three wind speed intervals)	Average wind speed 0-5m/s: 2022/02/28-2022/08/10 Average wind speed 5-10m/s: 2022/05/30-2022/08/09 Average wind speed 10-15m/s: 2022/05/30-2022/08/10
Test set	13,400 (0.45 million data for each of the three wind speed intervals)	Average wind speed 0-5m/s: 2022/08/11-2022/11/30 Average wind speed 5-10m/s: 2022/08/10-2022/11/29 Average wind speed 10-15m/s: 2022/08/11-2022/11/29

Table 2. Data volume distribution in Strategy 2

Data set	Data volume	Period
Training set	103,300 (34,400 for each of the three wind speed intervals)	Average wind speed 0-5m/s: 2020/12/01-2022/03/04 Average wind speed 5-10m/s: 2020/12/02-2022/05/20 Average wind speed 10-15m/s: 2021/01/15-2022/05/30
Validation set	29,500 (Data volume for each of the three wind speed ranges is 0.9800)	Average wind speed 0-5m/s: 2022/03/05-2022/08/16 Average wind speed 5-10m/s: 2022/05/212022/08/07 Average wind speed 10-15m/s: 2022/05/31-2022/08/11
Test set	14,800 (0.49 million data for each of the three wind speed intervals)	Average wind speed 0-5m/s: 2022/08/17-2022/11/30 Average wind speed 5-10m/s: 2022/08/08-2022/11/30 Average wind speed 10-15m/s: 2022/08/12-2022/11/29

Table 3. Data volume distribution in Strategy 3

3. Model and Method

3.1 Deep Learning Model

3.1.1 Informer Model

The Transformer [21] model has excellent performance in sequence modelling, but its model complexity is too high, the consumption of computing and storage resources is great, and the training speed is relatively slow. As a result, much work has been done in recent years to reduce model complexity with little impact on performance. Informer model [22] proposed a new self-attention mechanism ProbSparse and self-attention distillation module, greatly reducing the network scale. Informer is composed of Encoder and Decoder, two parts; its basic architecture is shown in Figure 1.

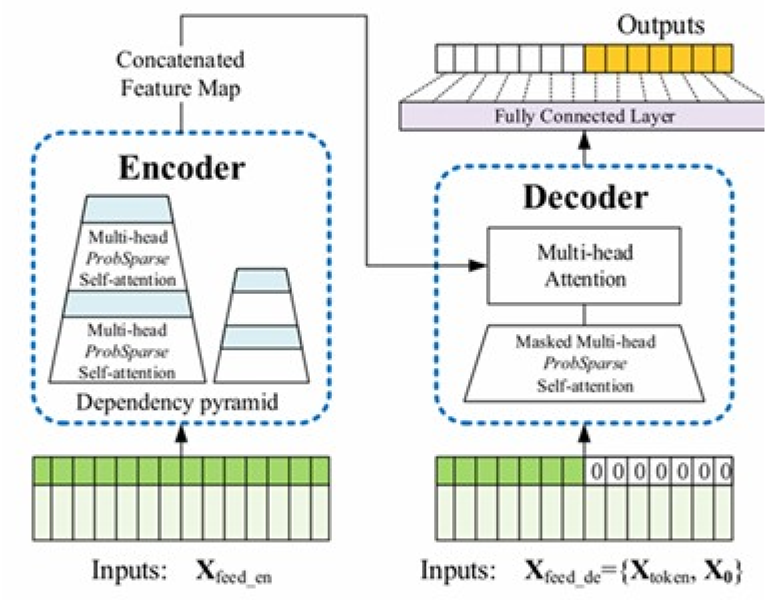


Figure 1. Basic Architecture of Informer [22]

3.1.2 LSTM-CNNs model

LSTM[23] is a classical time series prediction model. LSTM-CNN combines a double-layer LSTM and a double-layer 1-dimensional convolution layer to make the model learn the change in time and the association between different elements at different sites. Multiple dropout layers are introduced into the model to enhance its robustness. LSTM-CNNs, on the other hand, stack multiple LSTM-CNNs to deepen the network.

3.1.3 Conv-Informer Model

To better combine the features of the model, this paper proposes a Conv-Informer model which combines the Informer model and the LSTM-CNNs model. After stacking the output of LSTM-CNNs and Informer, Conv-Informer model adopts two one-dimensional convolution operations with Kernel sizes of 1 and 3 respectively and expands the channel. In order to learn the association between LSTM-CNNs and Informer output and the association between adjacent time steps, the two convolution results are stacked. Then, a 1-dimensional convolution operation with kernel size one is carried out, and the channel is compressed to 1 to converge the information of each channel as output.

3.2 Loss function with trend penalty factor

Consider this situation: The true value of wind speed at two consecutive moments is 1 m/s and 2 m/s; the predicted value of wind speed at these two moments by model 1 is 2 m/s and 1 m/s; the predicted value of wind speed at these two moments by model 2 is 2 m/s and 3 m/s; the root mean square error (RMSE) of the predicted results of the two models are the same. However, model 2 correctly predicted the change trend of wind speed, that is, the wind speed at the latter moment is 1 m/s larger than the previous moment, so model 2 is still better than model 1 despite the same RMSE of the two models.

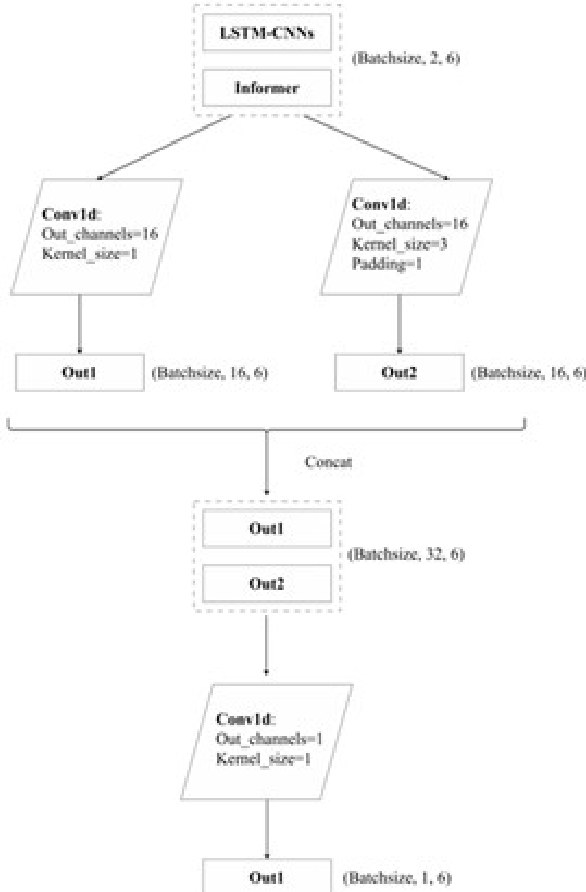


Figure 2. Basic Architecture of Conv-Informer

To make the model take into account the wind speed variation trend in the training process, this paper proposes a loss function including the trend penalty factor:

$$Loss = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log x_i - \log \hat{x}_i)^2} + \gamma\sqrt{\frac{1}{n-1}\sum_{i=2}^{n}\left[(x_i - x_{i-1}) - (\hat{x}_i - \hat{x}_{i-1})\right]^2} \quad (1)$$

In formula (1), the first term is the relative error RMSE loss function of the prediction sequence, the second term is the trend penalty term, and ³ is the hyperparameter, which controls the proportion of the trend penalty term.

3.3 Training Methods

For the three data set partitioning strategies designed in Section 2.2, multiple models are trained and their differences are compared. For ease of comparison, all models were trained for 200 rounds, the number of training sets for each round was 50,000, and the batchsize was set to 128.

In order to focus on the strong winds above 10 m/s, this paper adopts the prediction accuracy of strong winds above 10 m/s as the model evaluation index. It defines the prior accuracy and posterior accuracy of strong winds prediction by referring to the prior probability and posterior probability in probability theory. Among them, the prior accuracy rate represents the proportion correctly predicted by the model in all the times when strong winds appear, and the subsequent accuracy rate represents the proportion of actual strong winds in all the times when strong winds appear predicted by the model. The higher the prior/posterior accuracy rate is, the better the prediction effect of the model on strong winds.

It is worth noting that the prior/posterior accuracy should be as close as possible. If a model predicts strong winds for all cases, the prior accuracy of the model is 100%, but the posterior accuracy is relatively low. If the other model does not predict high winds for almost all cases, then the posterior accuracy of this model will be quite high, and the prior accuracy will be close to zero. Therefore, the closeness of the prior/posterior accuracy should also be used as one of the model training effect evaluation indices. The smaller the difference between the two, the better the wind prediction effect of the model and the more credible the result.

4. Results and Discussion

4.1 Data set partitioning Strategy 1 Train LSTM-CNNs and Informer

The prediction effect of LSTM-CNNs model on the test set is shown in Fig. 3. The coordinates of scatter points in the figure are (true value, predicted value), and the colors in the figure represent the degree of aggregation of scatter points; The dotted green line is the fitting line of the scattered points. The closer this fitting line is to

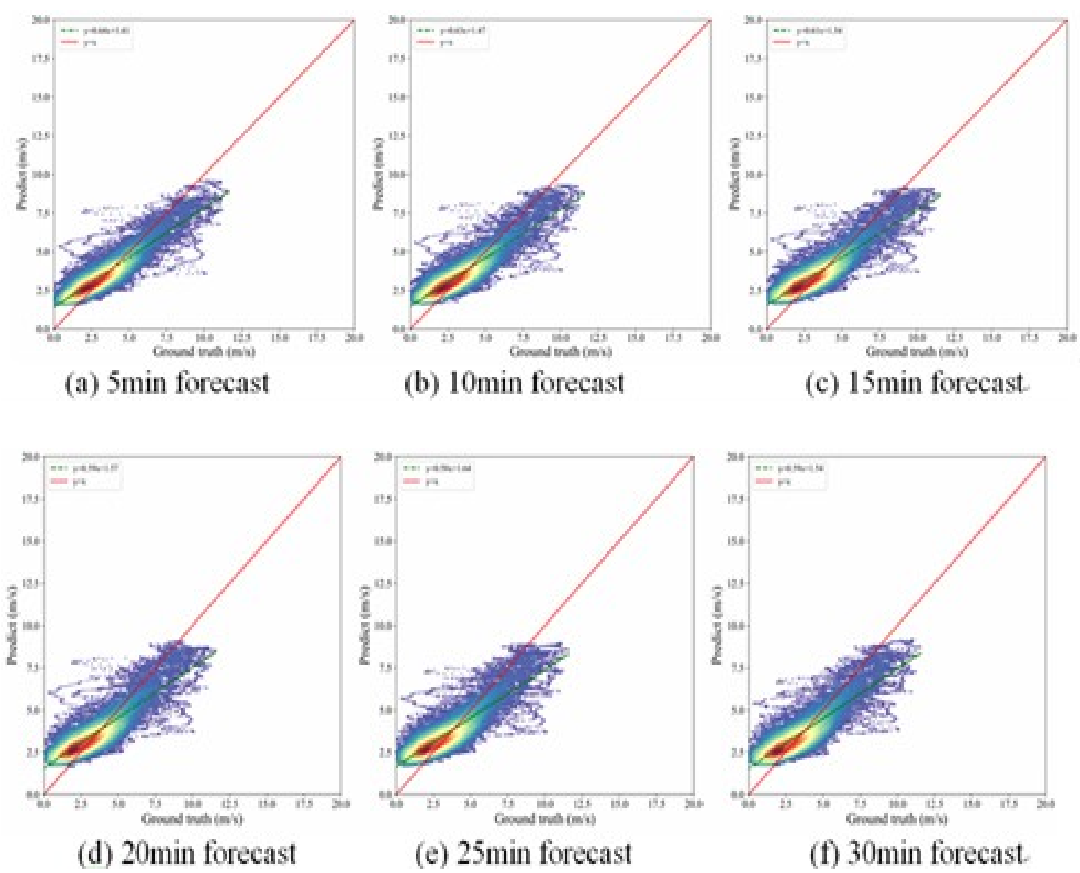


Figure 3. Training results of LSTM-CNNs under data set partitioning strategy 1

$y = x$, the more correlated the predicted value of the model is with the true value. When all the scattered points are concentrated near the fitting line, the model's prediction is more stable.

Fig. 3 shows that LSTM-CNNs have a poor fitting effect to the true value under strategy 1. Weak wind models with wind speeds below 4 m/s are larger, while those with winds above 4 m/s are significantly smaller, and there is a significant cut-off at about 2 m/s and 10 m/s. The main reason for this result is that in both the training set and the test set, the wind speed of the test site is mostly below 5 m/s, resulting in a lot of learning of wind speed characteristics below 5 m/s in the process of model training, thus ignoring the wind speed of 10 m/s. The results showed that LSTM-CNNs failed to predict high winds above 10 m/s in the test set.

Fig. 4 shows the prediction results of the Informer model, similar to the training results of the LSTM-CNN model.

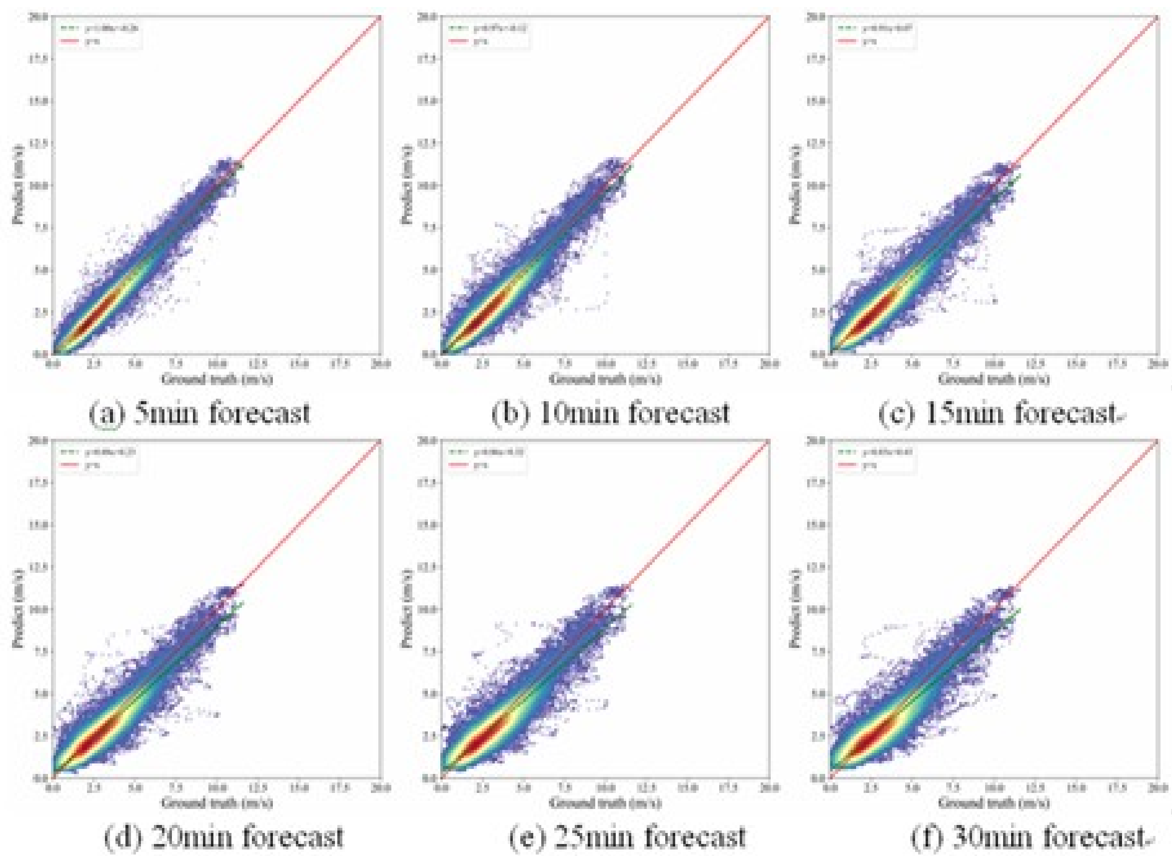


Figure 4. Informer training results under data set partitioning strategy 1

In contrast to the LSTM-CNNs model, although the wind data in the data set is relatively small, the Informer model can still learn some wind characteristics for the same data set, indicating that the Informer model is superior to the LSTM-CNNs model in wind prediction. The prediction accuracy of Informer model for winds above 10 m/s is shown in Table 4. The accuracy of Informer model for wind prediction is still low, and the posterior accuracy after the prediction step size of 15 min is significantly higher than the prior accuracy, indicating that the model's predicted value is small.

Predicted step size	**5 min**	**10 min**	**15 min**	**20 min**	**25 min**	**30 min**
Prior accuracy (%)	70.37	64.29	54.43	53.87	52.84	48.63
Posterior accuracy (%)	68.35	60.14	74.66	72.19	70.86	73.86
Difference (%)	2.02	4.43	20.23	18.32	18.02	25.23

Table 4. Prediction accuracy of high wind above 10 m/s by Informer data set partitioning strategy 1

4.2 Data set partitioning Strategy 2 Train LSTM-CNNs and Informer

Data set partitioning strategy two effectively balances the number of different wind speed levels in the data set. The prediction effect of LSTM-CNNs model on the test set is shown in Fig. 5, and the prediction accuracy rate of high winds above 10 m/s is shown in Table 5. Combining Fig. 5 and Fig. 3, it can be seen that the training results of LSTM-CNNs model have been greatly improved, the prediction effect is more stable with more concentrated scatter points. The accuracy rate of wind prediction is above 84%, and the difference of prior/ posterior accuracy is below 9%.

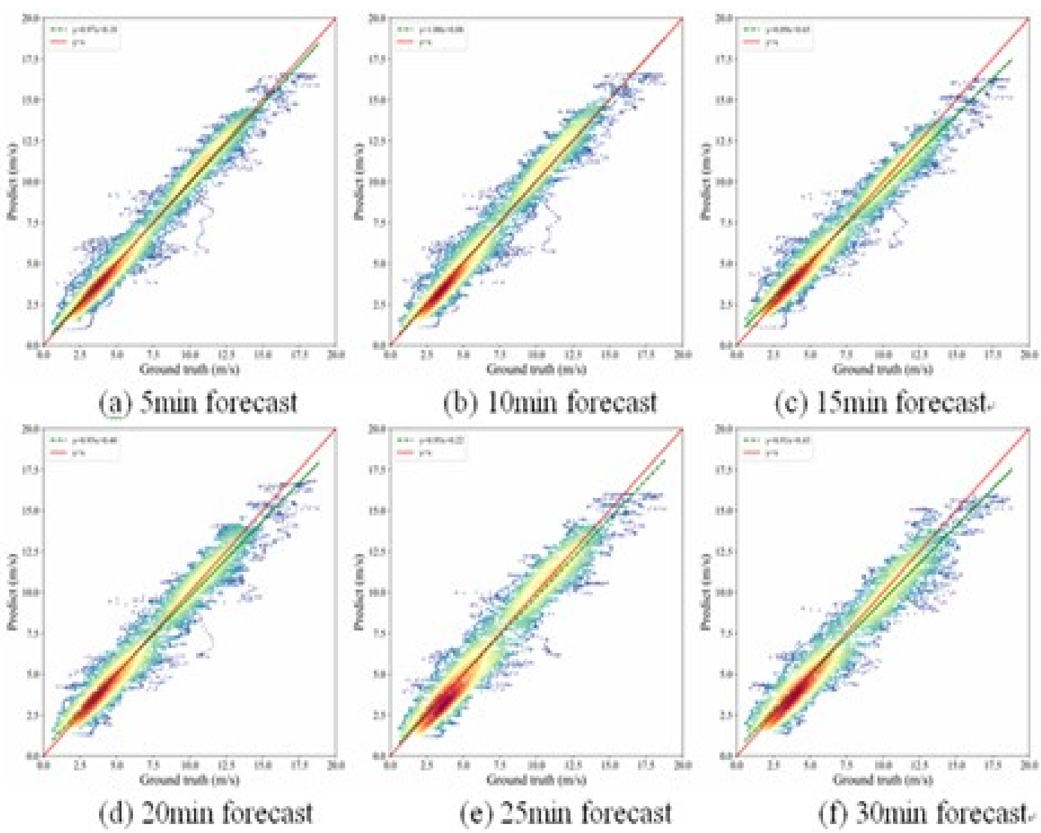


Figure 5. Training results of LSTM-CNNs under data set partitioning strategy 2

Predicted step size	5 min	10 min	15 min	20 min	25 min	30 min
Prior accuracy (%)	94.49	95.68	88.76	90.67	92.15	84.52
Posterior accuracy (%)	91.6	89.13	94.31	92.33	89.51	93.1
Difference (%)	2.89	6.55	5.55	1.66	2.64	8.58

Table 5. LSTM-CNNs prediction accuracy of high wind above 10 m/s under data set partitioning strategy 2

The prediction effect of Informer model on the test set is shown in Fig. 6, and the prediction accuracy rate of wind above 10 m/s is shown in Table 6. Compared with Table 4, the Informer model effectively improved training results from a more balanced wind speed data set, reducing the difference in the prior/posterior accuracy of wind prediction to around 10%.

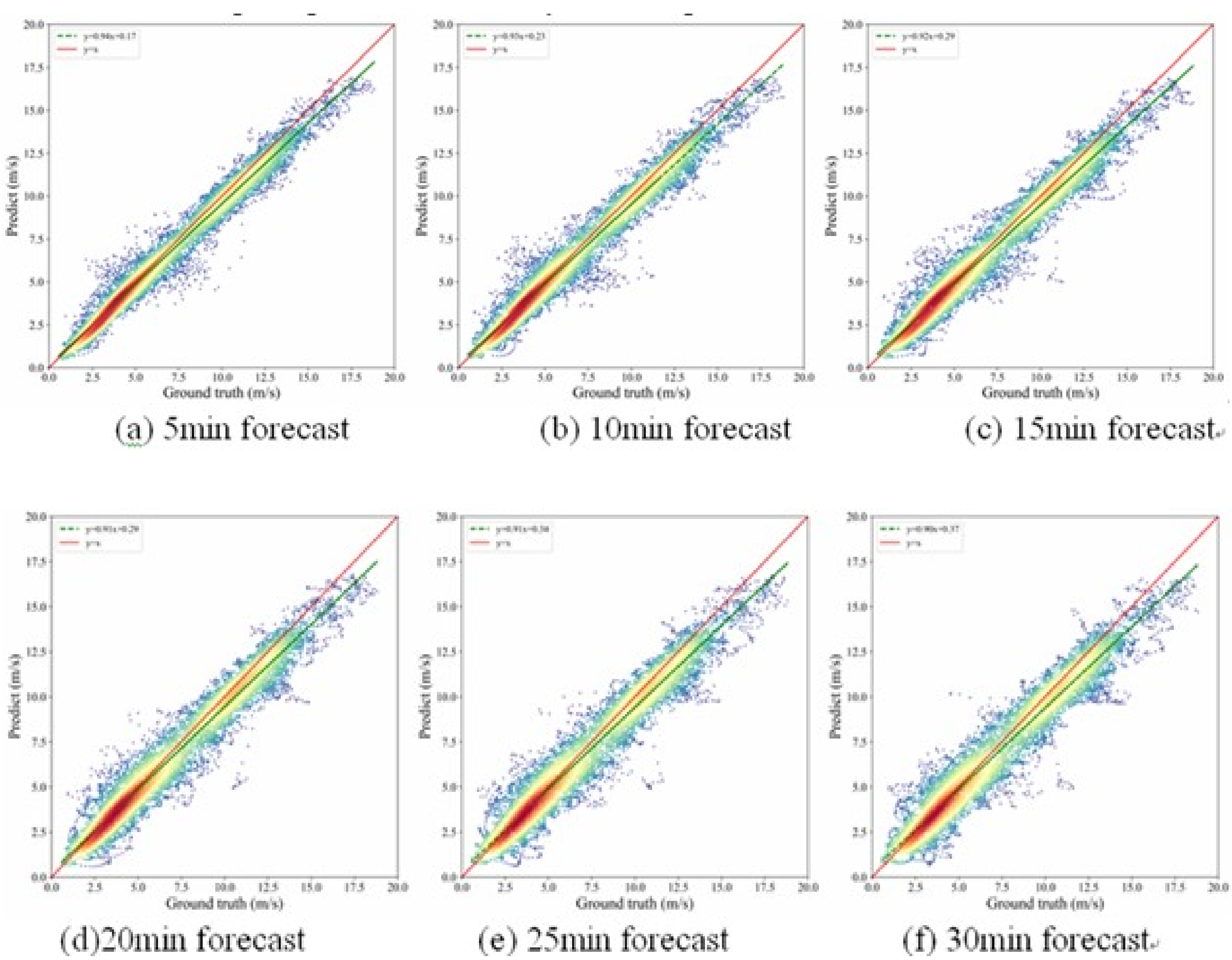


Figure 6. Informer training results under data set partitioning strategy 2

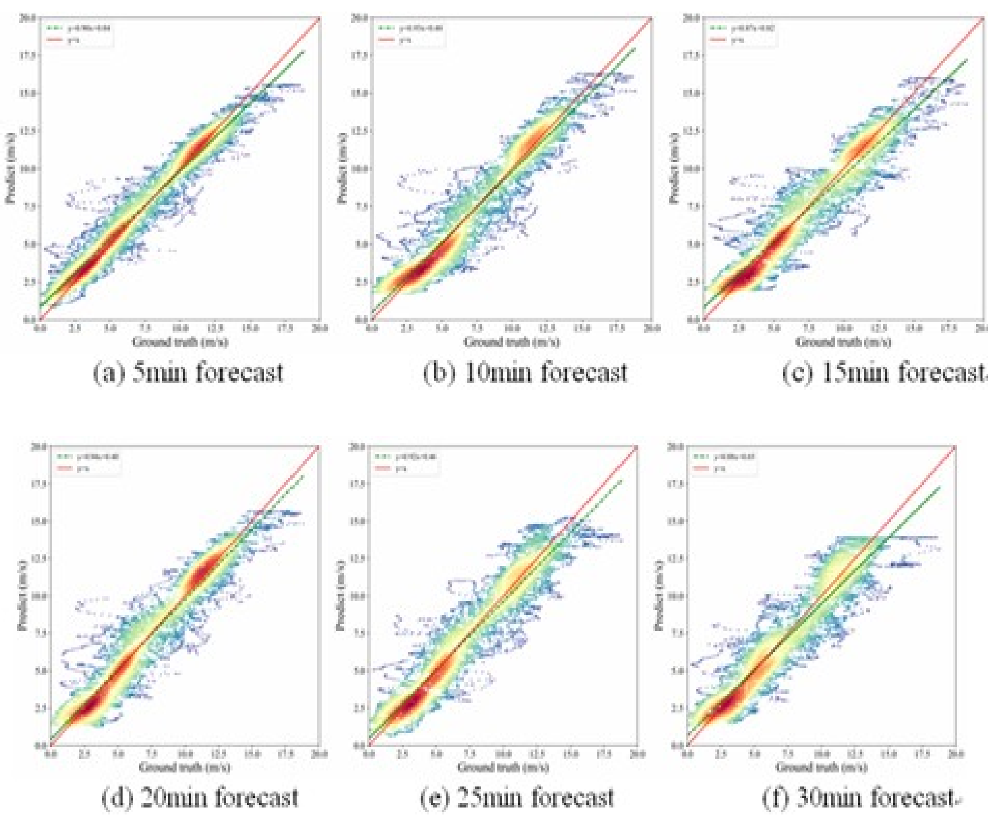
Predicted step size	5 min	10 min	15 min	20 min	25 min	30 min
Prior accuracy (%)	87.64	86.73	85.74	84.75	83.88	82.99
Posterior accuracy(%)	96.8	95.92	95.69	95.81	95.34	94.77
Difference (%)	9.16	9.19	9.95	11.06	11.46	11.78

Table 6. Informer's prediction accuracy of high wind above 10 m/s under data set partitioning strategy 2

4.3 Data set partitioning Strategy 3 Train LSTM-CNNs and Informer

To verify whether the prior/posterior accuracy difference of the model's wind prediction can be further reduced, dataset strategy 3 expands the model input step size from 36 steps (3 hours) to 138 steps (11.5 hours).

The prediction effect of the LSTM-CNNs model on the test set is shown in Fig. 7, and the prediction accuracy of the wind above 10 m/s is shown in Table 7. Fig. 7 shows that the prediction performance of the LSTM-CNNs model decreases to a certain extent after the data input length is extended. Compared with Fig. 5, the fit line of the scatter in Fig. 7 is more deviated from $y = x$, and the scatter is more divergent.



(a) 5min forecast (b) 10min forecast (c) 15min forecast

(d) 20min forecast (e) 25min forecast (f) 30min forecast

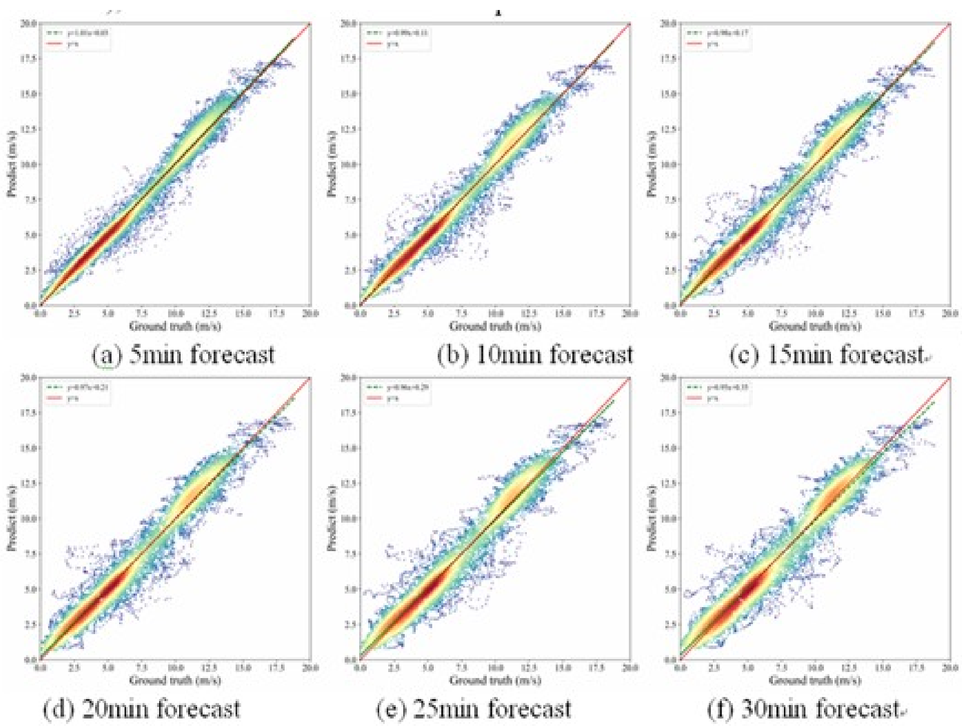
Figure 7. Training results of LSTM-CNNs under data set partitioning strategy 3

Predicted step size	**5 min**	**10 min**	**15 min**	**20 min**	**25 min**	**30 min**
Prior accuracy (%)	94.3	91.42	78.1	91.37	86.98	85.11
Posterior accuracy (%)	93.83	94.39	95.44	94.2	92.01	93.02
Difference (%)	0.47	2.97	17.34	2.83	5.03	7.91

Table 7. LSTM-CNNs prediction accuracy of high wind above 10 m/s under data set partitioning strategy 3

The prediction effect of the Informer model on the test set is shown in Fig. 8, and the prediction accuracy of the wind above 10 m/s is shown in Table 8. Fig. 8 shows that the prediction results of the Informer model are further convergent compared with strategy 2, and the slope of the fitting line (Fig. 8(f)) even for the 30 min prediction reaches 0.95. Table 8 shows that the pre-posteriori accuracy of the wind prediction is more than 93%, and the difference of the pre-posteriori accuracy is almost less than 0.5% (only the difference of 1.58% for the prediction of 5 minutes), which indicates that the model is quite reliable for the results of the wind forecast.

However, Informer's model still has a flaw: it cannot predict wind speed trends. Fig. 9(a) shows a group of wind speeds predicted by Informer, and the predicted values of the model are almost horizontal lines, which indicates that the predicted results of the Informer model are more of an average state of wind speed in the next 30 minutes, rather than reflecting its changing trend.



(a) 5min forecast (b) 10min forecast (c) 15min forecast (d) 20min forecast (e) 25min forecast (f) 30min forecast

Figure 8. Training results of Informer under data set partitioning strategy 3

Predicted step size	5 min	10 min	15 min	20 min	25 min	30 min
Prior accuracy (%)	96.42	95.26	94.83	94.58	94.1	93.63
Posterior accuracy (%)	94.84	94.90	94.91	94.62	94.14	93.33
Difference (%)	1.58	0.36	0.08	0.04	0.04	0.3

Table 8. Informer's prediction accuracy of high wind above 10 m/s under the data set partitioning strategy 3

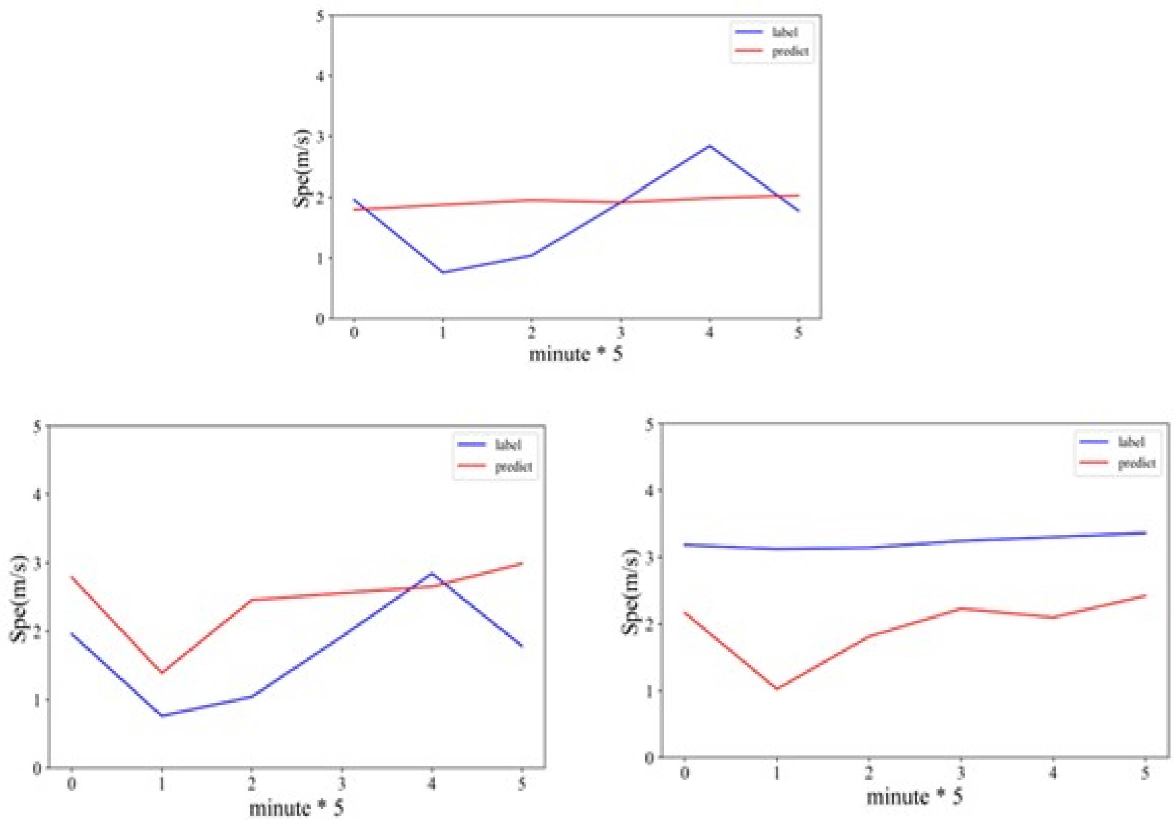


Figure 9. Example of predicted values of Informer Model (a) and LSTM-CNNs Model (b)(c) under data set partitioning strategy 3

On the contrary, some of the results predicted by LSTM-CNNs model can basically reflect the changing trend of wind speed correctly (see Fig. 9(b)). However, a considerable part of the results also reflect the wrong trend of change (Fig. 9(c)), which is also the reason why the prediction results of LSTM-CNNs model are more divergent than those of Informer model. The reason why the LSTM-CNNs model incorrectly predicts the change trend of wind speed may be that part of the wind speed fluctuation is caused by atmospheric turbulence, which is difficult to predict because of its high randomness. Moreover, the LSTM-CNNs model is more likely to incorrectly predict the future change trend of wind speed after input data containing random fluctuations. Such predictions are less effective than the Informer model, which lacks variation. In other words, the prediction results of the Informer model tend to be more like the average state of future wind speed, while the LSTM-CNNs model is rich in changes, but it is also easy to predict the change trend of wind speed incorrectly.

4.4 Data set partitioning Strategy 3 Train Conv-Informer

In order to make the training model more usable, this paper hopes to find a compromise that makes the model predict the wind speed change trend as much as possible, and predict the future average state for those cases that may contain random fluctuations. Therefore, this paper proposes a Conv-Informer model combining LSTM-CNNs and Informer. In the Conv-Informer model, LSTM-CNNs and Informer directly use the weights trained by their respective models as pre-trained models to speed up model convergence.

The prediction effect of Conv-Informer model on the test set is shown in Fig. 10, and the prediction accuracy of the wind above 10 m/s is shown in Table 9, with the prediction accuracy reaching more than 93%.

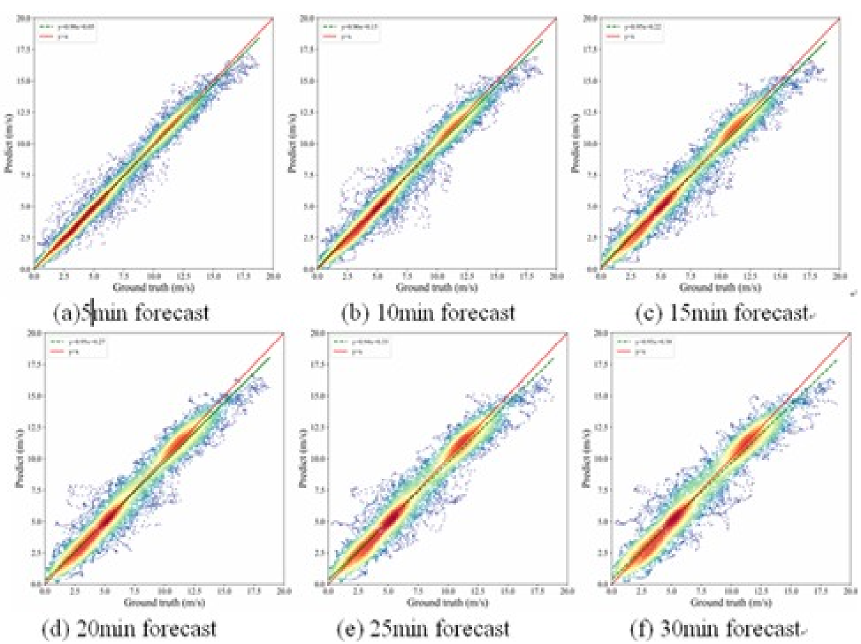


Figure 10. Training results of Conv-Informer under data set partitioning strategy 3

Predicted step size	5 min	10 min	15 min	20 min	25 min	30 min
Prior accuracy (%)	95.09	94.29	94.07	94.22	93.86	93.09
Posterior accuracy (%)	95.58	94.75	94.69	94.72	94.26	93.42
Difference (%)	0.49	0.46	0.62	0.5	0.4	0.33

Table 9. Conv-Informer's prediction accuracy of high wind above 10m/s under data set partitioning strategy 3

Compared with Informer model, the prediction effect of Conv-Informer model is slightly decreased, but in exchange for the ability to predict the linear trend of future wind speed. As shown in Fig. 11, Conv-Informer model generally predicted the future change trend of wind speed, but for the same set of data, Informer model failed to predict the trend, while LSTM-CNNs model predicted the trend incorrectly. However, the LSTM-CNNs model correctly predicts the complex nonlinear trend of wind speed, while the Conv-Informer model can only predict the linear trend.

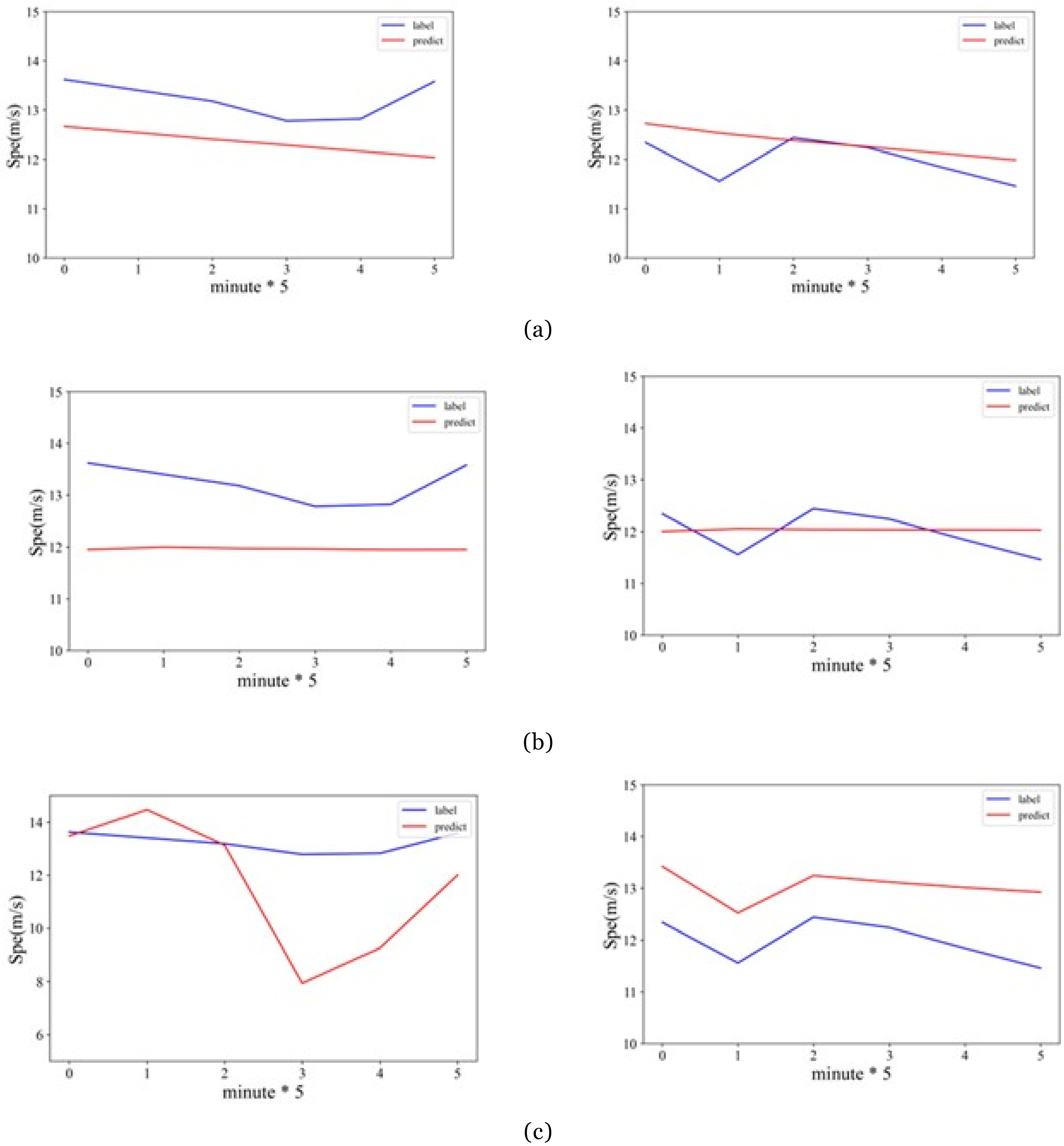


Figure 11. Training results of Conv-Informer model (a), Informer model (b) and LSTM-CNNs model (c) under data set partitioning strategy 3

5. Conclusion and Future Prospect

The variation of wind speed is characterized by strong fluctuation, intermittency and high variability. The Kumtag Desert region of Xinjiang has abundant wind resources. Considering that there is almost no research on wind speed prediction in this region at present, in order to make better use of it, this paper uses deep learning technology to carry out single-station ultra-short time surface wind prediction for this region, and the specific work is as follows:

(1) Using three different data set partitioning strategies, it is proved that whether the proportion of different wind speed levels in the data set is uniform has a crucial impact on the model training effect;

(2) A loss function containing trend penalty factor is proposed, so that the wind speed variation trend can be taken into account in the training process of the model;

(3) Experimental results show that Informer model tends to predict the average state of future wind speed; LSTM-CNNs model can reflect the nonlinear change trend of future wind speed to a certain extent, but it is easy to predict the change trend incorrectly. In this paper, an improved Conv-Informer model based on Informer and LSTM-CNNs model is proposed to forecast ultra-short time wind speed, which can predict the linear trend of future wind speed change under the premise of ensuring the prediction effect. The prediction accuracy rate of the model for winds above 10 m/s can reach more than 93% within 30 minutes.

However, the Conv-Informer model cannot predict the nonlinear trend of wind speed change, and the next step is to improve this problem further.

References

[1] Tang, Hao., Zhou, YaMan., Yang, Jieyao., et al. (2022). Analysis of mesoscale characteristics of strong winds in the Baili Wind region of Xinjiang. *Journal of Desert & Oasis Meteorology*, *16*(4), 1–8.

[2] He, Qiuping., Guo, Naxian., Chengfeng, A. et al. (2023). Research progress of short-term wind speed prediction model. *Qinghai Science and Technology*, *18*(6), 146–149.

[3] Ma, Junyan., Yuan, Yiping., Chai, Tongbing. (2021). Short-term wind speed prediction at fan hub based on combined neural networks. *China Mechanical Engineering*, *32*(17), 2082–2089.

[4] Liu, Donghai., Huang, Jing., Liu, Juan., et al. (2023). Review and prospect of parameterization schemes for typical mesoscale numerical prediction models. *Advance in Earth Science*, *38*(4), 349–362. *(in Chinese)*

[5] Feng, J., Ding, R Q., Liu, D Q. (2014). The application of nonlinear local Lyapunov vectors to ensemble predictions in Lorenz systems. *Journal of the Atmospheric Sciences*, *71*(9), 3554–3567.

[6] Liu, H., Chen, C. (2019). Multi-objective data-ensemble wind speed forecasting model with stacked sparse autoencoder and adaptive decomposition-based error correction. *Applied Energy*, *254*(9), 113686.

[7] Liu, X. L., Lin, Z., Feng, Z M. (2021). Short-term offshore wind speed forecast by seasonal ARIMA - a

comparison against GRU and LSTM. *Energy*, *227*, 120492.

[8] Liang, T., Chai, C. J., Sun, H. X., et al. (2022). Wind speed prediction based on multi-variable Capsnet-BILSTM-MOHHO for WPCCC. *Energy*, *250*, 123761.

[9] Zeng, Anjie., Hua, Wei., Yan, Zhongwei., et al. (2019). Spatio-temporal forecasting strategy of meteorological elements based on deep learning: Comparison of direct forecasting and iterative forecasting. *Climatic and Environmental Research*, *28*(5), 89–100.

[10] Zhang, Jingguo. (2023). *Research on correction method of numerical weather forecast precipitation products based on deep learning* (Master's thesis). Nanjing University of Information Science and Technology.

[11] Fang, Wei., Qi, Mei-Han. (2023). High spatiotemporal resolution precipitation prediction method based on deep learning. *Journal of Earth Sciences and Environment*, *45*(3), 706–718. (in Chinese)

[12] Yan, Junjie., Qu, Jianhua., Yuan, Mingge., et al. (2023). Green light channel construction method of Fengyun-4 satellite based on deep learning. *Remote Sensing Technology and Application*, *38*(1), 214–226. (in Chinese)

[13] Yuan, Kai., Pang, Jing., Li, Wujie., et al. (2023). Application evaluation of deep learning model for radar echo prediction in Wuhan in flood season in 2021. *Arid Meteorology*, *41*(1), 173–185. (in Chinese)

[14] Rasp, S., Pritchard, M S., Gentine, P. (2018). Deep learning to represent sub-grid processes in climate models. *Proceedings of the National Academy of Sciences*, *115*(39), 9684–9689.

[15] Scher, Sebastian. (2018). Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. *Geophysical Research Letters*, *45*(22), 12616–12622.

[16] Moraux, A., Dewitte, S., Cornelis, B., et al. (2019). Deep learning for precipitation estimation from satellite and rain gauges measurements. *Remote Sensing*, *11*(21), 2463.

[17] Huang, L. N., Wang, Q. Y., Huang, J. H., et al. (2022). A novel hybrid predictive model for ultra-short-term wind speed prediction. *Energies*, *15*, 4895.

[18] Liang, Tao., Chai, Chunjie., Sun, Hexu., et al. (2022). Wind speed prediction based on multi-variable Capsnet-BILSTM-MOHHO for WPCCC. *Energy*, *240*, 123761.

[19] Ye, Ru., Li, Xiaotong., Ye, Yuanming., et al. (2022). DynamicNet: A time-variant ODE network for multi-step wind speed prediction. *Neural Networks*, *152*, 118–139.

[20] Zhang, Yu., Ma, Lichun., Wang, Kai. (2022). Research progress of Quaternary environmental evolution in Lop Nur Dry Salt Lake. *Advance in Earth Science*, *37*(2), 149–164.

[21] Vaswani, A., Shazeer, N., Kaiser, L., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*.

[22] Zhou, H Y., Zhang, S H., Peng, J Q., et al. (2021). Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(12), 11106–11115.

[23] Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.