



OPEN ACCESS

EDITED BY

Chenghai Wang,
Lanzhou University, China

REVIEWED BY

Shuanglong Jin,
China Electric Power Research Institute
(CEPRI), China
Feng Ye,
Rutgers, The State University of New Jersey -
Busch Campus, United States

*CORRESPONDENCE

Jinman Zhang,
✉ zzzjm2046@126.com

RECEIVED 21 February 2025

ACCEPTED 26 May 2025

PUBLISHED 04 June 2025

CITATION

Chen Z, Zhang J, Zhou S, Zhao Z and Liu Y
(2025) Ultra-short-term prediction of
spatio-temporal wind speed based on a
hybrid deep learning model.
Front. Earth Sci. 13:1580945.
doi: 10.3389/feart.2025.1580945

COPYRIGHT

© 2025 Chen, Zhang, Zhou, Zhao and Liu.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Ultra-short-term prediction of spatio-temporal wind speed based on a hybrid deep learning model

Zixuan Chen^{1,2,3}, Jinman Zhang^{3,4*}, Shuang Zhou³,
Zengbao Zhao^{3,4} and Yushan Liu³

¹China Meteorological Administration Xiong'an Atmospheric Boundary Layer Key Laboratory, Xiong'an New Area, China, ²Key Laboratory of Meteorology and Ecological Environment of Hebei Province, Shijiazhuang, China, ³Hebei Provincial Meteorology Service Center, Shijiazhuang, China, ⁴Energy Meteorology Key Laboratory, China Meteorological Administration, Beijing, China

This study develops a spatio-temporal forecasting model for predicting wind speeds across the Beijing-Tianjin-Hebei region over a 4-h horizon. The model, built using advanced deep learning techniques, operates with a temporal resolution of 1 hour and a spatial resolution of 9 km. The experiments were first trained based on ConvLSTM and UNet, and improved by introducing the Self-Attention (SA) mechanism module to construct two hybrid deep learning models, Conv-SA as well as UNet-SA, respectively. The results show that the spatio-temporal predictions of the UNet model are significantly better than ConvLSTM, and the TS scores show that for the prediction of high wind, the enhancement is more than 50% for the next 4 hours. The addition of the SA module significantly improves the model prediction accuracy, and Conv-SA improves significantly, compared to ConvLSTM by more than 60%. The models were more accurate in predicting wind speeds in the region of the terrestrial than the oceanic subsurface. In addition, the model produces more accurate wind speed predictions for coastal as well as plateau regions. This study provides a new research idea for the proximity prediction of wind speed.

KEYWORDS

deep learning, wind speed, prediction, ultra-short-term, spatio-temporal

1 Introduction

Wind is an extremely important meteorological element, which can reflect different atmospheric circulation characteristics and is a key factor in climate analysis, a variety of extreme natural disasters are also affected by winds (Nan et al., 2019; Aiqing et al., 2016). Changes in wind speed can have a significant impact on the ecology of a region, while wind energy is growing in demand worldwide as a major source of electricity (Veers et al., 2019; Guoying and Xinyi, 2024). Therefore, the accurate prediction of wind speed is of great importance in the fields of energy utilization and disaster warning.

The wind speed prediction methods include physics-based models, statistical models, artificial intelligence methods and hybrid models (Chen J. et al., 2024; Yang M. et al., 2024). Numerical Weather Prediction (NWP) provides a good account of the effects of physical properties of meteorological elements on wind speed, yet its characteristics give it an advantage over localized ultrashort-term forecasts for large-scale weather forecasting (Landberg, 2001). The more commonly used statistical methods are Autoregressive

model (AR), Moving Average Model (MA), Kalman filter (KF) etc (Dai et al., 2024; Li et al., 2022). Wang et al. (2025) evaluated the accuracy of Gaussian, Lognormal, Weibull, and Gamma distribution functions in simulating the probability density function (PDF) of urban wind speed, providing a more appropriate statistical basis for wind speed prediction.

The rapid advancement of computer technology and continuous upgrades in hardware have facilitated the widespread application of artificial intelligence techniques in meteorological element forecasting (Ham et al., 2019; Chen et al., 2023; Price et al., 2025). Artificial intelligence methods exhibit robust data fitting capabilities and demonstrate proficiency in modeling complex nonlinear relationships inherent in wind speed data (Wang J. et al., 2023; Liu and Zhang, 2024). Deep learning methods based on big data can effectively predict wind speed under the premise of comprehensively considering a variety of influencing factors (Zhao et al., 2022; Zhang et al., 2019). Commonly used methods include Support Vector Machine (SVM), Artificial Neural Networks (ANN), Feed Forward Neural Networks (FFNN) (Akhtar et al., 2021; Liu et al., 2021; Masrur et al., 2016). Such methods improve can effectively learn the nonlinear characteristic changes among wind speed data (Neshat et al., 2021). For multi-step prediction of wind speed, the classic methods are Long Short Term Memory (LSTM), Convolutional Long Short Term Memory network (ConvLSTM) (Al-qaness et al., 2024; Sun et al., 2023). Yaghoubirad et al. (2023) compared the performance of several deep learning models including Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), and LSTM in a study on wind speed prediction and concluded that GRU has higher accuracy. The robust spatial data processing capability of Convolutional Neural Networks (CNN) enables accurate forecasting of meteorological disasters by comprehensively considering the influence of the surrounding geographical environment (Chen Z. et al., 2024). Frequency Filter Enhanced Dual LSTM (FDNET) is an improvement of the traditional LSTM model, and studies have shown that it effectively solves the problems of inefficiency of LSTM in dealing with long time series and sensitivity to noisy data (Mo et al., 2024). Wang et al. (2024) proposed a Conditional Local Convolution Recurrent Network (CLCRN) model, which combines temperature, pressure, and other meteorological factors for spatial and temporal prediction of wind speed. The model has good prediction performance at different prediction length. The Self-Attention (SA) mechanism can assign different weights to different features to learn the importance level among the influencing factors and improve the generalizability of the model, which has also been applied to wind speed prediction in related studies (Tian et al., 2022; Dai et al., 2023).

Integrating traditional methods with machine learning can achieve better prediction results. Liu et al. (2012) proposed a hybrid wind speed prediction model EMD-ANN based on Empirical Mode Decomposition (EMD) and ANN. The combination of LSTM and Grey Wolf Optimizer (GWO) indicates that it has better predictive performance than a single prediction model (Altan et al., 2021). Srihari and Kiran (2024) integrated the advantages of the Variational Mode Decomposition (VMD) and Sample Convolutional Interactive and Network (SCINet) architectures to predict wind speeds within different time ranges. Experiments showed that the hybrid model significantly improved the accuracy of wind speed prediction.

Hybrid deep learning methods have relatively limited research in related fields, and a reasonable combination of the advantages of different models can further improve prediction accuracy (Zheng and Wang, 2024). The existing research integrates the spatial correlation learning capabilities of CNN with the temporal feature processing advantages of LSTM, thereby enhancing the model's spatio-temporal feature learning performance (Chen et al., 2019; Wu et al., 2021). Zhu et al. (2018) proposed a Predictive Deep Convolutional Neural Network (PDCNN) method for spatio-temporal prediction of wind speed at multiple sites simultaneously. This method, which combines CNN and Multi-layer Perceptron (MLP), performs better than traditional models and can effectively learn the spatio-temporal correlation of wind speed. Jiang et al. (2023) combined Variational Mode Decomposition (VMD), Graph Neural Network (Graph Neural Network) and Temporal Convolutional Network (TCN) have achieved remarkable results in the prediction of wind speed in the next hour.

In summary, artificial intelligence methods have significant advantages in current wind speed forecasting, and combining the strengths of different models to improve the prediction accuracy is of great research significance at present. The purpose of this study is to use different deep learning methods to forecast the wind speed in the Beijing-Tianjin-Hebei (BTH) region of China in time and space under the premise of considering various meteorological factors. The main research content is as follows: (1) Based on the unique geomorphologic features of the BTH region, the hybrid deep learning model ConvLSTM-Self-Attention (Conv-SA) and UNet-Self-Attention (UNet-SA) with multiple meteorological elements is constructed to predict the spatio-temporal wind speed. (2) Explore the predictive performance of each model and quantitatively analyze the improvement of the wind speed prediction accuracy of the study area by SA mechanism under different time steps.

2 Materials and methodology

2.1 Data sources

The study data are based on a multi-source fusion observational 1-km resolution product for Hebei Province. This dataset was obtained by the Hebei Climate Center through quality-controlled ground hourly observational data, utilizing multi-grid variational analysis, fusion analysis parameter optimization, and terrain height deviation correction. The data are specifically tailored for application in the unique geomorphic environmental characteristics of Hebei. The meteorological elements used in the research include air temperature, humidity, pressure and wind speed, with a time resolution of hourly. The spatial scope of the dataset covers the BTH region, and the temporal distribution is from October 2023 to October 2024. The research area is shown in Figure 1.

In addition, the study classified the wind power levels with reference to the Chinese standard "Wind Power Levels (GB/T 28,591—2012)". For high wind samples of magnitude nine and above with less data volume, they are grouped into one category in the study, and the specific division is shown in Table 1.

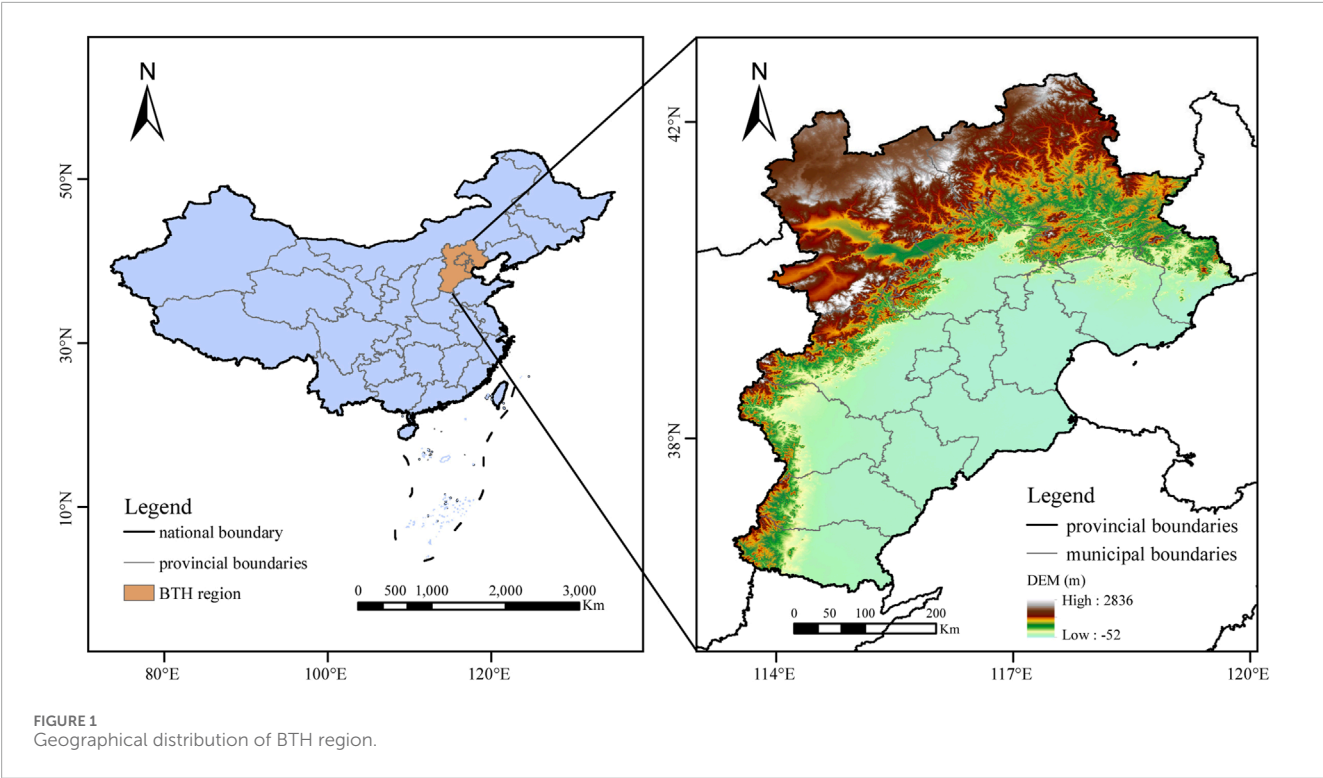


TABLE 1 Classification of wind power.

Wind power (level)	Wind speed (m/s)
0	0.0~0.2
1	0.3~1.5
2	1.6~3.3
3	3.4~5.4
4	5.5~7.9
5	8.0~10.7
6	10.8~13.8
7	13.9~17.1
8	17.2~20.7
≥9	≥20.8

2.2 Methods

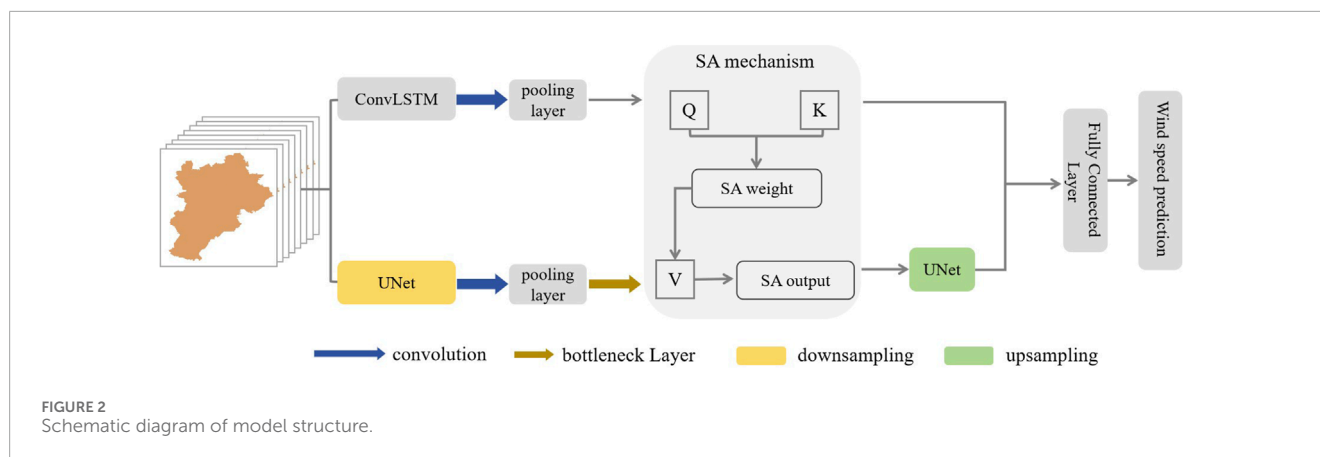
2.2.1 Deep learning models

In this study, a deep learning model is used for spatio-temporal prediction by the ConvLSTM (Shi et al., 2015) and UNet (Ronneberger et al., 2015) models. We introduce a SA mechanism (Vaswani, 2017) to further improve the model performance.

LSTM (Hochreiter, 1997) is good at handling tasks related to time series prediction. ConvLSTM improves on this basis by introducing CNN into the model structure to perform convolution operations on time steps. CNN can effectively learn the spatial distribution features of a dataset, thus ConvLSTM is commonly used for spatio-temporal prediction tasks. The UNet model principle relies on CNN, and its infrastructure consists of an encoder (downsampling) part and a decoder (upsampling) part, and effectively combines the shallow and deep features of spatio-temporal datasets by hopping layers of connectivity to enhance the accuracy of the model prediction. The experiments are based on ConvLSTM and UNet combined with the SA mechanism to propose Conv-SA and UNet-SA respectively to evaluate the model prediction performance. Two baseline models (CNN and ConvGRU) were added to the experiment for comparison to explore the predict results of wind speed under different prediction lengths.

2.2.2 SA mechanism

Attention mechanism is a method that mimics human cognitive attention, and in deep learning attention mechanism enables neural networks to assign different weights to different parts of the data, thus improving the accuracy and generalization of the model (Wang H. et al., 2023). The SA mechanism operates by establishing connections between all grid points in sequence data, computing attention weights through pairwise similarity measurements. This approach enables effective extraction of spatio-temporal features across different regions, thereby enhancing prediction accuracy. In our research, the ConvLSTM and UNet models are improved by introducing a SA mechanism module to check whether the accuracy of spatio-temporal prediction can be improved, and the structure of the model is shown in Figure 2.



The calculation of the SA mechanism is shown in formulas 1, 2:

$$Q = W_q X, K = W_k X, V = W_v X \quad (1)$$

$$SA(Q, K, V) = \text{softmax} \frac{QK^T}{\sqrt{d_k}} \quad (2)$$

The Q , K and V , are obtained by multiplying each data X in the sequence with the three weight matrices W_q , W_k , and W_v respectively after a linear transformation. The attention weights are derived from the matrix multiplication of Q and K^T , then divide by $\sqrt{d_k}$ to stabilize the gradient. This process captures the degree of mutual influence between different data points, effectively controlling the range of attention weights to ensure smooth results and prevent the generation of extreme values. The softmax function is then applied to normalize these calculated values, after which they are multiplied with the V matrix to produce the final SA output.

2.2.3 Persistence forecasting

The study used persistence forecasting (PF) as the benchmark to evaluate the predictive performance of the deep learning model. This method assumes that the predicted value of the adjacent wind speed is equal to the current observed value, as shown in formula 3:

$$y'(t_c + h) = y(t_c) \quad (3)$$

t_c is the current time, $y'(t_c + h)$ is the wind speed forecasts at lead time h . This method can effectively verify the improvement of the model in the prediction accuracy of ultra-short-term wind speed.

2.2.4 Samples generation

The research is to predict the wind speed in the BTH region. Historical 4-h air temperature, humidity, pressure and wind speed are utilized to make spatial and temporal predictions of the next 4 hours of wind speed in BTH region. This research intends to train the deep learning methods and evaluate the prediction performance of the models for different regions in BTH. The original 1 km spatial resolution multi-source fusion observational dataset product is upscaled to 9 km to increase the model running rate and reduce the consumption of computational resources while meeting the evaluation requirements. The model predicts wind speed based on historical meteorological elements and continuously outputs the wind speed values for the next 4 h during the training process. We

conducted quality control of the research data and divided the data into samples by the sliding window method, as shown in Figure 3. We suppose that the total sample length is s , the sample contains historical temporal sequences M and predicted temporal sequences N . When the history input duration is m and the prediction duration is n , $M = \{x_{t-m}, x_{t-m+1}, \dots, x_t\}$, $N = \{x_{t+1}, x_{t+2}, \dots, x_{t+n}\}$, where t represents the current moment and x_t denotes the value of the meteorological element at the current moment. Each data sample contains 4 h of historical meteorological element data and 4 h of wind speed prediction labels. The experiment also verified all samples in the data preprocessing stage, and removed the samples with time missing to ensure the continuity of the 8-h time series in each sample. This is crucial for the model to capture the features of the time dimension. After completing the data integration process, we obtained a total of 7,264 samples that satisfied the training requirements.

The experiment divides the sample data into training set, validation set and test set with the ratio of 8:1:1, the order of the data set was randomly shuffled to ensure the robustness of the model. All sample data are normalized before training to speed up the model convergence. The formula 4 is as follows:

$$X_t^* = \frac{X_t - X^{min}}{X^{max} - X^{min}} \quad (4)$$

where t denotes the time in the dataset, X_t is the value of the meteorological element at hour t , X^{min} is the minimum value of the corresponding latitude and longitude time series, and X^{max} is the maximum value of the corresponding latitude and longitude time series.

2.2.5 Experimental setup

The composition of ConvLSTM mainly includes the input layer, the layer combining convolution operation and LSTM, the max pooling layer and the fully connected layer. Conv-SA connects SA mechanism after the Max pooling layer to further learn the important characteristics of the wind speed retained by the max pooling layer, better capture the global characteristics, and reduce the consumption of computing resources to a certain extent at the same time. For the UNet model, we used 3D convolution based on its infrastructure, enabling the model to learn the spatio-temporal characteristics of wind speed and combine SA mechanism with bottleneck. The SA mechanism is implemented with eight

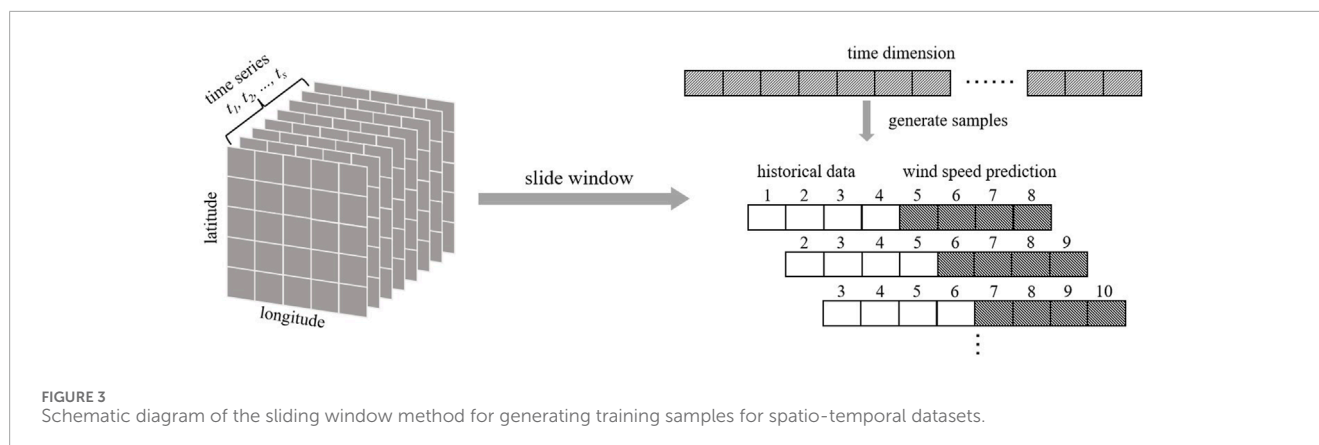


TABLE 2 Comparison of model parameters and iteration time.

Model	ConvLSTM	Conv-SA	UNet	UNet-SA
Parameters	7,219,506	192,107,716	4,800,769	874,321,241
AIS	0.029	0.033	0.038	0.078

independent heads, where each head independently computes the attention weights. During the experiment, the deep learning model used relu as the activation function.

The hyperparameters of the model were set uniformly in order to ensure the fairness of the experimental results. We use the adam optimizer. The initial learning rate is set to 5e-5, the batch size is 8, and the training epoch is 100. We employed the early stopping technique with a patience of 10 epochs. The minimum delta of the experiment is set to 0.0001. This means that the training process will only be considered improved when the change in the verification metric exceeds this threshold. We record the number of parameters of each model and the average number of seconds consumed by one iteration (AIS) in Table 2. It is worth noting that the Conv-SA and UNet-SA models have a significantly higher number of parameters than traditional methods.

2.2.6 Evaluation metrics

In this study, Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Threat Score (TS) are chosen to evaluate the model prediction performance, in which MSE is the loss function for model training. The calculation are shown in formulas 5–8

$$MSE(y, y') = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{n} \quad (5)$$

$$RMSE = \sqrt{MSE} \quad (6)$$

$$MAE(y, y') = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (7)$$

$$TS = \frac{NA}{NA + NB + NC} \quad (8)$$

where y_i is the true value, y'_i is the predicted value of the model at the corresponding moment, NA is the number of correct forecasts, NB is the number of null reports, and NC is the number of missed reports.

3 Results

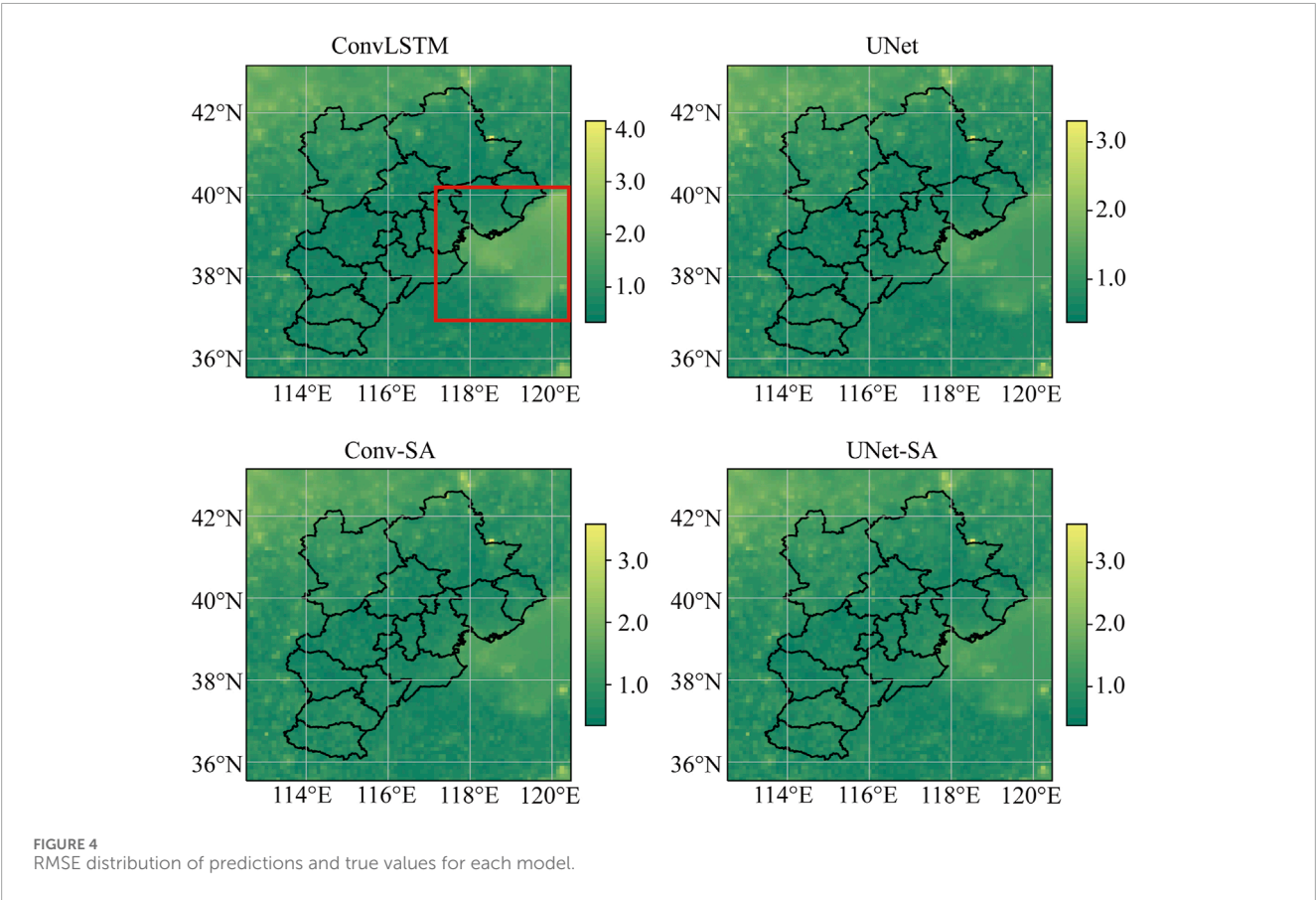
Table 3 takes RMSE as the evaluation index and lists the prediction results of wind speed for different deep learning models and PF. The prediction accuracy of ConvLSTM is slightly lower than that of the other models. However, the performance of the models is significantly better than that of the benchmark methods.

The study systematically evaluated the spatio-temporal prediction performance of various deep learning models up to a 4-h forecast horizon on a point-by-point basis. Figure 4 shows the spatial distribution of RMSE in BTH region, calculated from the prediction results and true values of ConvLSTM, UNet, Conv-SA, and UNet-SA. While the spatial distribution of the predictions of each model has a similar trend, the RMSE for the prediction of wind speed in the offshore region is obviously high (red box in the figure). This is due to the smoother subsurface of the sea which makes it easy to increase the wind speed (Chunyan et al., 2014), and fewer offshore observatories, which makes the quality of the observational fusion product relatively low, and further enhances the difficulty of model prediction. The SA mechanism substantially enhances offshore wind speed prediction accuracy in the ConvLSTM model. For the test set as a whole calculating the RMSE weakens the effect of extreme values, so the study visualizes the prediction results for different classes of wind speed to evaluate the performance of the model.

Figure 5 shows the MAE calculated separately according to different wind classes, with reference to Table 1 for the specific division criteria. As illustrated in the figure, the deviation of each model progressively increases as the wind level increases, and the prediction accuracy of the model decreases with the increase of the prediction length. When the wind is weak, the gap between the model predictions is not obvious, when the wind reaches level 6 and above, the MAE of the UNet model is lower than that of the ConvLSTM. For high wind conditions, the SA mechanism significantly reduces the MAE of the model prediction, particularly when integrated with ConvLSTM architectures. When the prediction length is long, the enhancement effect of the SA

TABLE 3 Comparison of model predictions with PF.

Methods	PF	ConvLSTM	UNet	Conv-SA	UNet-SA
RMSE	1.306	1.063	0.936	0.988	0.967



mechanism is relatively obvious, and for the prediction of the next hour, the UNet model is better than UNet-SA. This is also a shortcoming of the model, the principle of parallel operation of the SA mechanism can be more effective in learning the features of data with long time series, but the training for the samples and the consumption of computational resources is relatively large (Zhou et al., 2021; Wei et al., 2023). It is also reflected from the training process of the model that Conv-SA and UNet-SA take much more time to train than ConvLSTM and UNet. Expanding the training samples can further enhance the benefits of the SA mechanism, but synthesizing the prediction results for the next 4 h, the addition of the SA mechanism effectively improves the wind speed prediction accuracy of the model, which is particularly effective for large wind speed.

The wind speed TS score can comprehensively assess the reliability and accuracy of the forecast, reflecting the forecast by analyzing the false alarm rate, the empty alarm rate and the hit rate; the higher the TS score, the more accurate the prediction. Table 4 shows the TS scores of different deep learning models for high wind (wind of level 6 and above) at different predict length. It can be seen

that the accuracy of the model prediction decreases as the prediction length increases, which is consistent with previous results. For the prediction of high wind, the UNet model is significantly better than ConvLSTM, with 75%, 63%, 61%, and 53% improvement for the next 4 h, respectively. The prediction of ConvLSTM improved significantly after the addition of the SA mechanism, with the improvement in the prediction of the next 4 hours at 75%, 68%, 61%, and 64%, respectively. Comparing the UNet and UNet-SA models, the enhancement effect after the addition of the SA mechanism is relatively limited, mainly in the long time-series prediction (3 or 4 h).

A sample of high wind was also selected for evaluation. Figure 6 shows the wind speed forecast results for the next 4 h for each model at 8:00 a.m. on 18 October 2024, at which the wind speeds in the study area are generally high. The average wind speed in the BTH region reaches 4.97 m/s, with the proportion of gale force wind as high as 8.9%. The figure shows that there is a general overestimation of offshore wind speed and an underestimation of terrestrial wind speed by ConvLSTM. Conv-SA significantly improves the forecast results after the inclusion of the SA mechanism, but the forecast

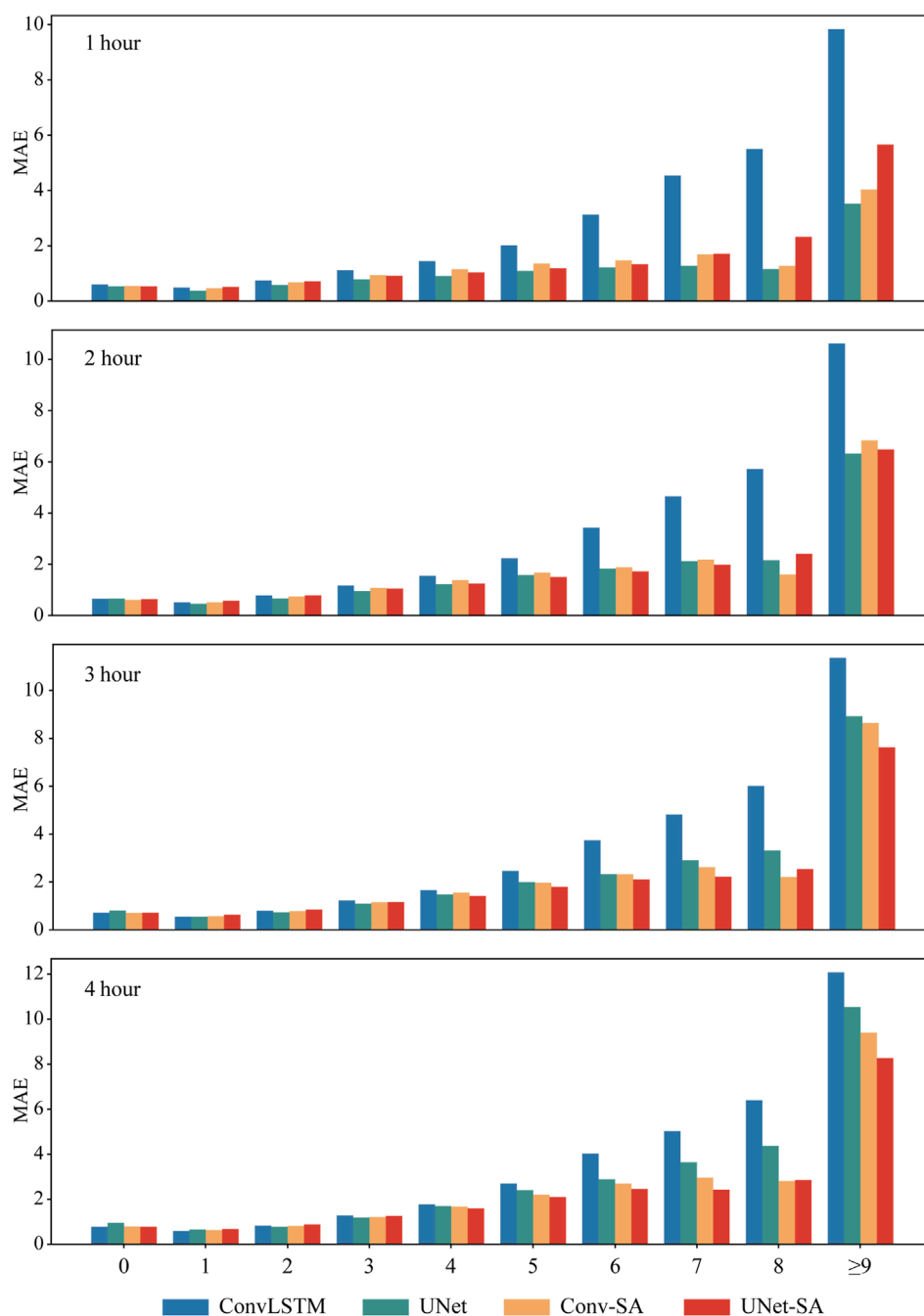


FIGURE 5
Predicted results of wind speed grading for each model, the horizontal coordinate is the wind level and the vertical coordinate is the MAE.

values of offshore wind speeds are gradually low as the predict length increases. While the UNet model demonstrates superior performance compared to ConvLSTM, the predicted values of land areas are high. The UNet-SA is relatively accurate for the offshore wind speed prediction, whereas there is an underestimation of the terrestrial wind speed prediction. For the north Hebei plateau region, where wind speed values are generally high due to topographic factors, the predictions of the UNet and Conv-SA

models are relatively accurate, and there is some overestimation in the ConvLSTM and UNet-SA models.

Figure 7 shows the comparison of the prediction results of hybrid deep learning models Conv-SA and UNet-SA with baseline models (CNN and ConvGRU). Figure 7a evaluated the RMSE of the predicted values and the true values at all wind speed levels. The performance of the model decreased with the increase of the prediction length. Conv-SA and UNet-SA achieve lower RMSE

TABLE 4 TS scores for each model at different prediction length.

Model	ConvLSTM	UNet	Conv_SA	UNet_SA
1 h TS score	0.20	0.35	0.35	0.35
2 h TS score	0.19	0.31	0.32	0.33
3 h TS score	0.18	0.29	0.29	0.31
4 h TS score	0.17	0.26	0.28	0.30

than the two baseline models. [Figure 7b](#) only retains samples with a wind speed level above 6, thereby evaluating the performance of the model under this specific condition. It is evident that the RMSE of each model has increased substantially. Meanwhile, the application of SA mechanism enhances the prediction accuracy of the hybrid deep learning model, making it significantly superior to the baseline model.

In summary, the SA mechanism can effectively improve the model prediction accuracy. Three regions with representative geomorphological characteristics, Zhangjiakou, Chengde and Qinhuangdao, were selected for further evaluation of the model, where Zhangjiakou and Chengde were chosen to be distributed in the Bashang Plateau area in the northern part of the city, and Qinhuangdao was chosen to be distributed in the eastern coast. [Figures 8, 9](#) respectively show the comparison between the predicted value and the true value of Conv-SA and UNet-SA models in each region when the predict length is 1 h and 4 h. The solid black lines in the figure are the true values, and the red and blue dashed lines are the predicted values of the model output.

The wind speed is generally high in the coastal and plateau areas. As can be seen from the figure, the selected areas in the experiment all produced different degrees of high wind in the test set. For the fluctuating trend of wind speed, Conv-SA and UNet-SA can make effective predictions. Compared with the prediction for the next 1 hour, the output stability of the model with a prediction duration of 4 hours is relatively poor, mainly manifested in the lag in predicting extreme values. Comparing the two different times, the prediction effect of UNet-SA is better than that of Conv-SA model, which is mainly reflected in the prediction result of high wind.

The deep learning models are relatively stable for wind speed prediction in the region of BTH, and the wind speed prediction in the land area is better than that in the offshore area. The advantage of the UNet model is more significant, in addition, SA mechanism can effectively improve the prediction accuracy of the model.

4 Conclusion and discussion

Based on ConvLSTM and UNet models, this study constructed a spatio-temporal prediction model of wind speed in the BTH region in the next 4 h. This study experimentally validated the potential improvement in model prediction accuracy through the incorporation of the SA mechanism. The main research conclusions are as follows.

- (1) The spatio-temporal prediction trends of the deep learning methods are similar, and the prediction accuracies of the models decrease with the prediction duration as well as with the increase of the wind level. The RMSE of the terrestrial wind speed prediction is significantly lower compared to the oceanic subsurface.
- (2) The spatio-temporal prediction accuracy of the UNet model is significantly better than that of the ConvLSTM. Notably, the addition of the SA mechanism module substantially enhances model prediction accuracy for high wind, with particularly pronounced improvements observed in the ConvLSTM.

The improvement of prediction performance by SA mechanism is significant. SA mechanism can effectively improve the prediction result of the complex landform features in the BTH region. Judging from the prediction of high winds in [Figure 6](#), Conv-SA has significantly improved the forecast accuracy of wind speed in the Bohai region and the Bashang Plateau. Furthermore, a comparison with the traditional model reveals that the improvement of SA mechanism in terms of high winds prediction accuracy is particularly noteworthy. According to the experimental results, the improvement of model performance depends on larger parameters and training time.

Wind as a meteorological factor is usually difficult to predict, the neural network can effectively learn the dynamic change law between the data, which greatly improves the previous problems in wind speed prediction ([Tian et al., 2024](#); [Li et al., 2023](#)). However, it should be noted that although the model can capture the characteristics of wind speed changes to a certain extent, the process of wind speed changes has strong nonlinear characteristics and will accumulate over time. Long-term predictions will increase the difficulty of model learning, and the prediction accuracy will also decrease as the duration increases. This is also consistent with previous studies ([Zhang et al., 2025](#); [Wu et al., 2022](#)).

The experimental findings reveal that the model's prediction error exhibits a progressive increase with rising wind levels. As the wind level escalates, the corresponding wind speed increase while the sample size becomes sparser. The high wind values that the model can learn during the training process are relatively limited, further increasing the difficulty of the model's prediction. In some terrains such as mountains and plateaus, wind speed is significantly affected by terrain factors, and it is difficult for the model to capture the variation characteristics of wind speed in such situations. In addition, the ocean is limited by the geographical environment, the observation stations are sparse, and the data quality is relatively low, which brings challenges to the model prediction of offshore wind speed ([Dong et al., 2024](#)). The wind speed at sea is generally high, which also makes the prediction error of the model significantly higher than that on land. However, the addition of SA mechanism weakens this phenomenon.

ConvLSTM and UNet, as commonly used basic models, can effectively predict wind speed in time and space. However, meteorological conditions in BTH region are complex, they tend to ignore the influence of terrain changes and have relatively limited predictive capabilities for high wind ([Xue et al., 2024](#)). Attention-based models such as transformer need to be trained on a large amount of data in order to demonstrate good performance. High-resolution meteorological data for short-term forecasting are

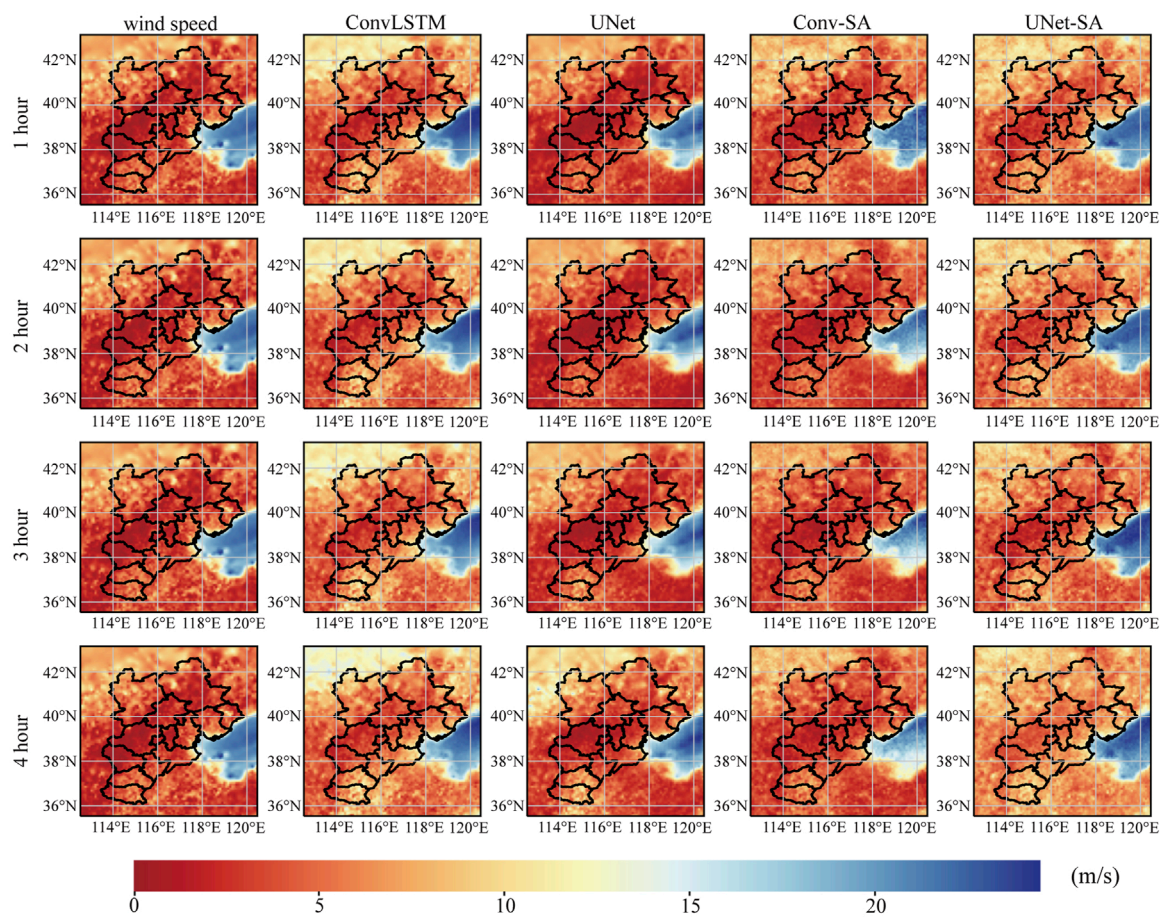


FIGURE 6
Wind speed predict for the next 4 h by model at 8:00 a.m. on 18 October 2024, the color bar indicates the magnitude of the wind speed.

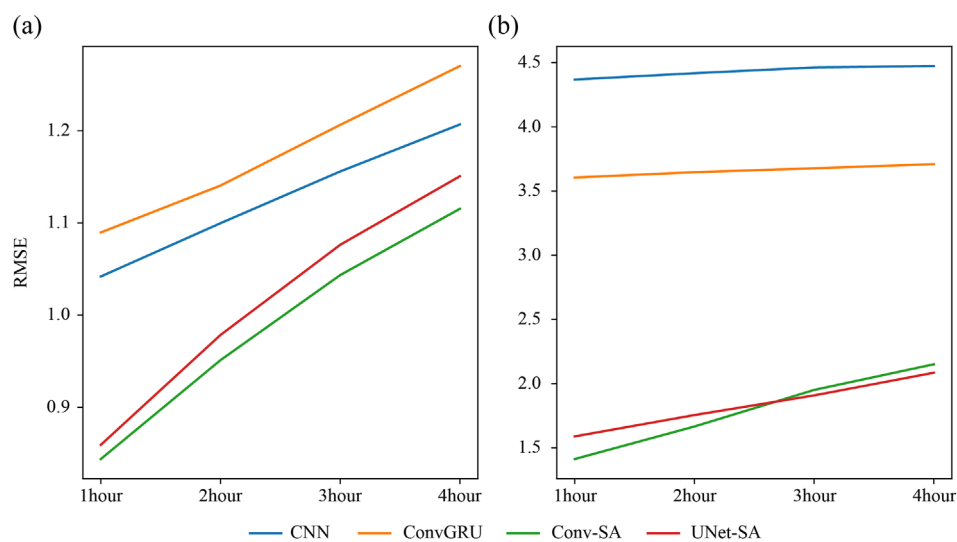


FIGURE 7
(a) The prediction results of each model under different prediction lengths. (b) The prediction results of high wind of each model under different prediction lengths.

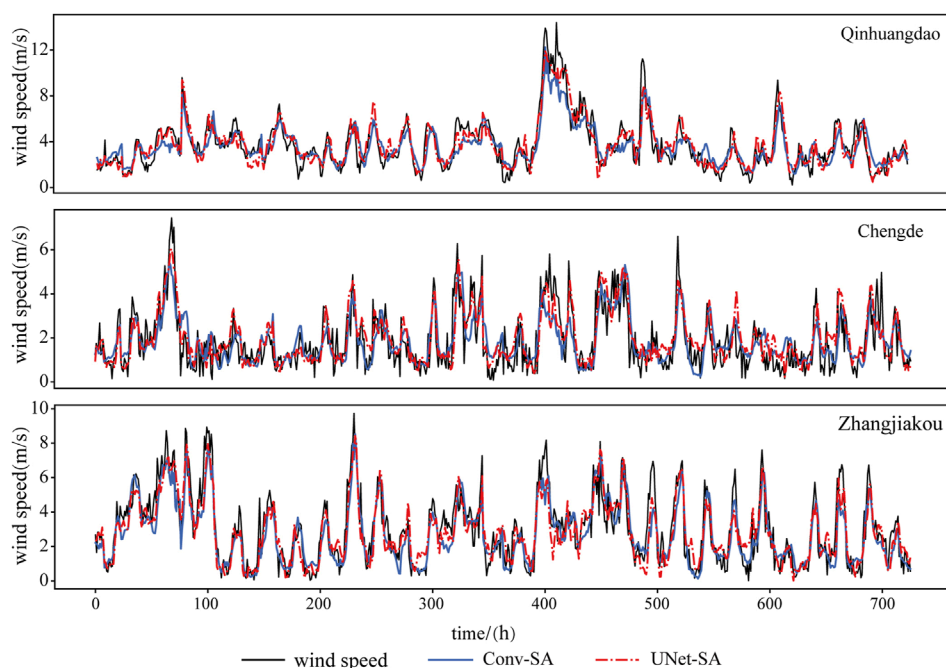


FIGURE 8
A time series comparison of predicted and true values with a predict length of 1 h.

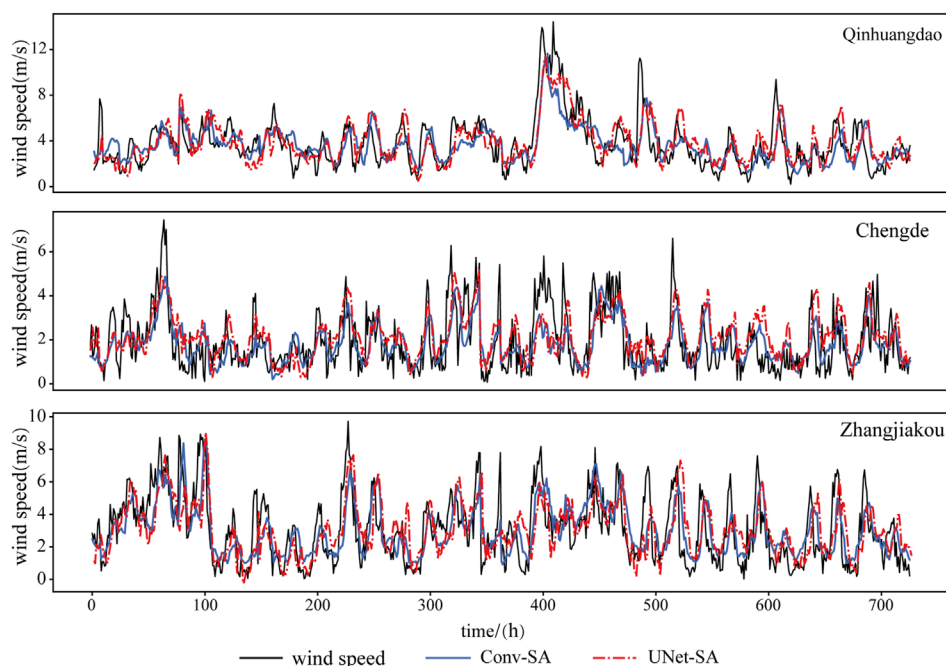


FIGURE 9
A time series comparison of predicted and true values with a predict length of 4 h.

relatively limited, and samples of high winds are particularly scarce. The main contribution of the research lies in integrating the advantages of different modules. From the perspective of model structure, the basic model extracts the spatial characteristics of

local meteorological elements through convolution operations. The addition of the SA mechanism can effectively combine local and global features. Dynamic allocation of weights is conducive to learning the details ignored by convolution operations, such as

the variation process of strong winds and the fluctuations of wind speed caused by the influence of mountainous terrain, and focuses on improving the prediction of wind speed at high levels. At the same time, the study takes into account the effects of multiple meteorological factors on the variation of wind speed to further improve the reliability of the model predictions (Wang et al., 2024).

This study shows the advantages of hybrid deep learning model in prediction, but there is still room for further improvement. In the experiment, Conv-SA and UNet-SA models require much more parameters to train than the basic model, and they also consume more time. The parallel computation principle of SA mechanism makes them extremely complex in time and space (Keles et al., 2023). In addition, the study is relatively deficient in the consideration of large time scales. Wind speeds are susceptible to seasonal as well as inter-annual variations and exhibit different fluctuation characteristics, which is weakened by the sample size of the dataset for model training (Yang X. et al., 2024; Han et al., 2018).

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

ZC: Conceptualization, Data curation, Formal Analysis, Methodology, Software, Visualization, Writing – original draft. JZ: Conceptualization, Funding acquisition, Supervision, Writing – review and editing. SZ: Formal Analysis, Writing – review and editing. ZZ: Writing – review and editing. YL: Writing – review and editing.

References

- Aiqing, F., Jiangbo, G. A. O., Shaohong, W. U., Yanhua, L. I. U., Xiaojia, H. E., 何霄嘉, et al. (2016). A review of storm surge disaster risk research and adaptation in China under climate change. *Prog. Geogr.* 35, 1411–1419. doi:10.18306/dlkxjz.2016.11.011
- Akhtar, I., Kirmani, S., Ahmad, M., and Ahmad, S. (2021). Average monthly wind power forecasting using fuzzy approach. *IEEE Access* 9, 30426–30440. doi:10.1109/ACCESS.2021.3056562
- Al-qaness, M. A., Ewees, A. A., Aseeri, A. O., and Abd Elaziz, M. (2024). Wind power forecasting using optimized LSTM by attraction–repulsion optimization algorithm. *Ain Shams Eng. J.* 15, 103150. doi:10.1016/j.asej.2024.103150
- Altan, A., Karasu, S., and Zio, E. (2021). A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. *Appl. Soft Comput.* 100, 106996. doi:10.1016/j.asoc.2020.106996
- Chen, J., Guo, Z., Zhang, L., and Zhang, S. (2024a). Short-term wind speed prediction based on improved Hilbert–Huang transform method coupled with NAR dynamic neural network model. *Sci. Rep.* 14, 617.
- Chen, K., Han, T., Gong, J., Bai, L., Ling, F., Luo, J., et al. (2023). Fengwu: pushing the skillful global medium-range weather forecast beyond 10 days lead. *arXiv Prepr. arXiv:2304.02948*. doi:10.48550/arXiv.2304.02948
- Chen, Y., Zhang, S., Zhang, W., Peng, J., and Cai, Y. (2019). Multifactor spatio-temporal correlation model based on a combination of convolutional neural network and long short-term memory neural network for wind speed forecasting. *Energy Convers. Manag.* 185, 783–799. doi:10.1016/j.enconman.2019.02.018
- Chen, Z., Wang, G., Wei, X., Liu, Y., Duan, Z., Hu, Y., et al. (2024b). Basin-scale daily drought prediction using convolutional neural networks in fenhe river Basin, China. *Atmosphere* 15, 155. doi:10.3390/atmos15020155
- Chunyan, S., Jianhua, L. I., and Sudan, F. A. N. (2014). Numerical study of terrain and underlying surface effect on Bohai gale. *Meteorol. Mon.* 40, 1338–1344. doi:10.7519/j.issn.1000-0526.2014.11.006
- Dai, X., Liu, G., and Hu, W. (2023). An online-learning-enabled self-attention-based model for ultra-short-term wind power forecasting. *Energy* 272, 127173. doi:10.1016/j.energy.2023.127173
- Dai, Y., Zhang, M., Jiang, F., Zhang, J., Liu, M., and Hu, W. (2024). Wind speed multi-step prediction based on the comparison of wind characteristics and error correction: focusing on periodic thermally-developed winds. *Eng. Appl. Artif. Intell.* 136, 108924. doi:10.1016/j.engappai.2024.108924
- Dong, D., Wang, S., Guo, Q., Ding, Y., Li, X., and You, Z. (2024). Short-term marine wind speed forecasting based on dynamic Graph embedding and spatiotemporal information. *J. Mar. Sci. and Eng.* 12, 502. doi:10.3390/jmse12030502
- Guoying, Z., and Xinyi, Z. (2024). Changes in wind speed in China from 1981 to 2020 and its relationship with atmospheric circulation. *Acta Sci. Nat. Univ. Pekin.* 60, 431–441. doi:10.13209/j.0479-8023.2024.029

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the China Meteorological Administration Innovation development project (CXFZ 2024J038).

Acknowledgments

We thank the researchers or teams who provided the basic data. We also thank the authors, reviewers, and editors who made amendments to the article.

Conflict of interest

The 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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ham, Y.-G., Kim, J.-H., and Luo, J.-J. (2019). Deep learning for multi-year ENSO forecasts. *Nature* 573, 568–572. doi:10.1038/s41586-019-1559-7
- Han, L., Wang, J., Wang, G., Wang, Z., and Wu, M. (2018). Spatial and temporal characteristics of average wind speed in the wind erosion region of northern China. *Arid. Land Geogr.* 41, 963–971. doi:10.12118/j.issn.1000-6060.2018.05.08
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* MIT-Press 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735
- Jiang, W., Lin, P., Liang, Y., Gao, H., Zhang, D., and Hu, G. (2023). A novel hybrid deep learning model for multi-step wind speed forecasting considering pairwise dependencies among multiple atmospheric variables. *Energy* 285, 129408. doi:10.1016/j.energy.2023.129408
- Keles, F. D., Wijewardena, P. M., and Hegde, C. (2023). “On the computational complexity of self-attention,” in *Proceedings of the 34th international conference on algorithmic learning theory, 2023 proceedings of machine learning research*. Editors A. SHIPRA, and O. FRANCESCO (Cambridge, MA: PMLR), 597–619. doi:10.48550/arXiv.2209.04881
- Landberg, L. (2001). Short-term prediction of local wind conditions. *J. Wind Eng. Industrial Aerodynamics* 89, 235–245. doi:10.1016/S0167-6105(00)00079-9
- Li, D., Yu, X., Liu, S., Dong, X., Zang, H., and Xu, R. (2022). Wind power prediction based on PSO-Kalman. *Energy Rep.* 8, 958–968. doi:10.1016/j.egyr.2022.02.077
- Li, K., Shen, R., Wang, Z., Yan, B., Yang, Q., and Zhou, X. (2023). An efficient wind speed prediction method based on a deep neural network without future information leakage. *Energy* 267, 126589. doi:10.1016/j.energy.2022.126589
- Liu, H., Chen, C., Tian, H. Q., and Li, Y. F. (2012). A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renew. Energy* 48, 545–556. doi:10.1016/j.renene.2012.06.012
- Liu, H., and Zhang, Z. (2024). Development and trending of deep learning methods for wind power predictions. *Artif. Intell. Rev.* 57, 112. doi:10.1007/s10462-024-10728-z
- Liu, X., Zhou, J., and Qian, H. (2021). Short-term wind power forecasting by stacked recurrent neural networks with parametric sine activation function. *Electr. Power Syst. Res.* 192, 107011. doi:10.1016/j.epsr.2020.107011
- Masrur, H., Nimol, M., Faisal, M., and Mostafa, S. M. G. (2016). in *Short term wind speed forecasting using Artificial Neural Network: a case study* (IEEE), 1–5. doi:10.1109/ICISSET.2016.7856485
- Mo, Y., Wang, H., Yang, C., Yao, Z., Li, B., Fan, S., et al. (2024). FDNet: frequency filter enhanced dual LSTM network for wind power forecasting. *Energy* 312, 133514. doi:10.1016/j.energy.2024.133514
- Nan, W., Qing-long, Y., and Ju-ju, L. (2019). The long-term trend of surface wind speed in China from 1979 to 2014. *J. Nat. Resour.* 34, 1531. doi:10.31497/zrzyxb.20190715
- Neshat, M., Nezhad, M. M., Abbasnejad, E., Mirjalili, S., Tjernberg, L. B., Garcia, D. A., et al. (2021). A deep learning-based evolutionary model for short-term wind speed forecasting: a case study of the Lillgrund offshore wind farm. *Energy Convers. Manag.* 236, 114002. doi:10.1016/j.enconman.2021.114002
- Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T. R., El-Kadi, A., Masters, D., et al. (2025). Probabilistic weather forecasting with machine learning. *Nature* 637, 84–90. doi:10.1038/s41586-024-08252-9
- Ronneberger, O., Fischer, P., and Brox, T. (2015). “U-net: convolutional networks for biomedical image segmentation,” in *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III* (Springer), 234–241. doi:10.1007/978-3-319-24574-4_28
- Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C. (2015). Convolutional LSTM network: a machine learning approach for precipitation nowcasting. *Adv. neural Inf. Process. Syst.* 28. doi:10.48550/arXiv.1506.04214
- Srihari, P., and Kiran, T. (2024). VMD-SCINet: a hybrid model for improved wind speed forecasting. *Earth Sci. Inf.* 17, 329–350. doi:10.1007/s12145-023-01169-3
- Sun, H., Song, T., Li, Y., Yang, K., Xu, D., and Meng, F. (2023). EEMD-ConvLSTM: a model for short-term prediction of two-dimensional wind speed in the South China Sea. *Appl. Intell.* 53, 30186–30202. doi:10.1007/s10489-023-05042-0
- Tian, C., Niu, T., and Wei, W. (2022). Developing a wind power forecasting system based on deep learning with attention mechanism. *Energy* 257, 124750. doi:10.1016/j.energy.2022.124750
- Tian, G., Le Coz, C., Charantonis, A. A., Tantet, A., Goutham, N., and Plougonven, R. (2024). Improving sub-seasonal wind-speed forecasts in Europe with a non-linear model. *arXiv Prepr. arXiv:2411.19077*. doi:10.48550/arXiv.2411.19077
- Vaswani, A. (2017). Attention is all you need. *Adv. Neural Inf. Process. Syst.* doi:10.48550/arXiv.1706.03762
- Veers, P., Dykes, K., Lantz, E., Barth, S., Bottasso, C. L., Carlson, O., et al. (2019). Grand challenges in the science of wind energy. *Science* 366, eaau2027. doi:10.1126/science.aau2027
- Wang, H., Qin, H., Liu, G., Liu, S., Qu, Y., Wang, K., et al. (2023a). A novel feature attention mechanism for improving the accuracy and robustness of runoff forecasting. *J. Hydrology* 618, 129200. doi:10.1016/j.jhydrol.2023.129200
- Wang, J., Gao, D., and Zhuang, Z. (2023b). An optimized deep nonlinear integrated framework for wind speed forecasting and uncertainty analysis. *Appl. Soft Comput.* 141, 110310. doi:10.1016/j.asoc.2023.110310
- Wang, M., Wang, J., Yu, M., and Yang, F. (2024). Spatiotemporal wind speed forecasting using conditional local convolution and multidimensional meteorology features. *Sci. Rep.* 14, 26219. doi:10.1038/s41598-024-78303-8
- Wang, W., Li, Y., and Ikegaya, N. (2025). Statistical models incorporating mean and standard deviation to predict probability distributions of pedestrian-level wind speed in a realistic urban area. *Build. Environ.* 279, 113034. doi:10.1016/j.buildenv.2025.113034
- Wei, X., Wang, G., Schmalz, B., Hagan, D. F. T., and Duan, Z. (2023). Evaluation of transformer model and self-attention mechanism in the Yangtze River basin runoff prediction. *J. Hydrology Regional Stud.* 47, 101438. doi:10.1016/j.ejrh.2023.101438
- Wu, Q., Guan, F., Lv, C., and Huang, Y. (2021). Ultra-short-term multi-step wind power forecasting based on CNN-LSTM. *IET Renew. Power Gener.* 15, 1019–1029. doi:10.1049/rpg2.12085
- Wu, Z., Yin, H., He, H., and Li, Y. (2022). Dynamic-LSTM hybrid models to improve seasonal drought predictions over China. *J. Hydrology* 615, 128706. doi:10.1016/j.jhydrol.2022.128706
- Xue, W., Hui, Y. U., Tang, S., and Huang, W. (2024). Relationships between terrain features and forecasting errors of surface wind speeds in a mesoscale numerical weather prediction model. *Adv. Atmos. Sci.* 41, 1161–1170. doi:10.1007/s00376-023-3087-5
- Yaghoubirad, M., Azizi, N., Farajollahi, M., and Ahmadi, A. (2023). Deep learning-based multistep ahead wind speed and power generation forecasting using direct method. *Energy Convers. Manag.* 281, 116760. doi:10.1016/j.enconman.2023.116760
- Yang, M., Guo, Y., Fan, F., and Huang, T. (2024a). Two-stage correction prediction of wind power based on numerical weather prediction wind speed superposition correction and improved clustering. *Energy* 302, 131797. doi:10.1016/j.energy.2024.131797
- Yang, X., Delworth, T. L., Jia, L., Johnson, N. C., Lu, F., and McHugh, C. (2024b). Skillful seasonal prediction of wind energy resources in the contiguous United States. *Commun. Earth and Environ.* 5, 313. doi:10.1038/s43247-024-01457-w
- Zhang, Z., Lin, L., Gao, S., Wang, J., Zhao, H., and Yu, H. (2025). A machine learning model for hub-height short-term wind speed prediction. *Nat. Commun.* 16, 3195. doi:10.1038/s41467-025-58456-4
- Zhang, Z., Ye, L., Qin, H., Liu, Y., Wang, C., Yu, X., et al. (2019). Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression. *Appl. energy* 247, 270–284. doi:10.1016/j.apenergy.2019.04.047
- Zhao, E., Sun, S., and Wang, S. (2022). New developments in wind energy forecasting with artificial intelligence and big data: a scientometric insight. *Data Sci. Manag.* 5, 84–95. doi:10.1016/j.dsm.2022.05.002
- Zheng, J., and Wang, J. (2024). Short-term wind speed forecasting based on recurrent neural networks and Levy crystal structure algorithm. *Energy* 293, 130580. doi:10.1038/s41598-024-51252-y
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., et al. (2021). “Informers: beyond efficient transformer for long sequence time-series forecasting,” in *Proceedings of the AAAI conference on artificial intelligence*, 11106–11115. doi:10.48550/arXiv.2012.07436
- Zhu, Q., Chen, J., Zhu, L., Duan, X., and Liu, Y. (2018). Wind speed prediction with spatio-temporal correlation: a deep learning approach. *Energies* 11, 705. doi:10.3390/en11040705