frontiers in Environmental Science

ORIGINAL RESEARCH
published: 17 March 2022
doi: 10.3389/fenvs.2022.836050

Capturing Spatial Influence in Wind Prediction With a Graph Convolutional Neural Network

Zeyi Liu[1,2] *and Tony Ware*[2]*

[1]College of Systems Engineering, National University of Defense Technology, Changsha, China, [2]Department of Mathematics and Statistics, University of Calgary, Calgary, AB, Canada

Nowadays, wind power is playing a significant role in power systems; it is necessary to improve the prediction accuracy, which will help make better use of wind sources. The existing neural network methods, such as recurrent neural network (RNN), have been widely used in wind prediction; however, RNN models only consider the dynamic change of temporal conditions and ignore the spatial correlation. In this work, we combine the graph convolutional neural (GCN) with the gated recurrent unit (GRU) to do prediction on simulated and real wind speed and wind power data sets. The improvements of prediction results by GCN in all wind speed experiments show its ability to capture spatial dependence and improve prediction accuracy. Although the GCN does not perform well in short-term wind power prediction as the change of wind power data is not so smooth due to the limitation of turbine operation, the results of long-term prediction still prove the performance of GCN.

Keywords: GCN, GRU, spatial correlation, wind speed prediction, wind power prediction

OPEN ACCESS

Edited by:
Gagan Deep Sharma,
Guru Gobind Singh Indraprastha University, India

Reviewed by:
Umer Shahzad,
Anhui University of Finance and Economics, China
Sanjeet Singh,
Chandigarh University, India

***Correspondence:**
Tony Ware
aware@ucalgary.ca

Specialty section:
This article was submitted to Environmental Economics and Management, a section of the journal Frontiers in Environmental Science

Received: *15 December 2021*
Accepted: *10 January 2022*
Published: *17 March 2022*

Citation:
Liu Z and Ware T (2022) Capturing Spatial Influence in Wind Prediction With a Graph Convolutional Neural Network. Front. Environ. Sci. 10:836050. doi: 10.3389/fenvs.2022.836050

INTRODUCTION

In order to deal with the crises of fuel energy and environmental pollution, wind energy is being widely developed all over the world, and wind power is starting to play a hugely significant role in power systems (Liu and Guan, 2004). However, the large amount of wind power generation also brings great challenges to the stable operation of power systems due to the variable and stochastic nature of wind (Defferrard et al., 2016). For example, wind power fluctuation caused a challenge event reported by the Electric Reliability Council of Texas (ECROT). To mitigate the harmful impacts, accurate wind prediction is always required by system operators (Zhang et al., 2016).

Wind prediction, including wind speed and wind power prediction, uses a large amount of historical wind data to train a fitting model between inputs and outputs. Wind speed is the most relevant factor in determining wind power. In particular, the wind power curves for different types of wind turbines roughly have the same shape (Pinson, 2013); hence, one approach to modeling wind power is to model the wind speed and associate it to the wind power using wind power curves suitable for a given wind farm (Stohl et al., 1995).

Wind speed can be modeled as a spatiotemporal process as it evolves randomly in time and space. However, wind forecasting has always been a challenge due to its complex spatial and temporal dependences.

Existing wind speed forecasting models can be categorized as physical, statistical, and machine learning models. Physical models use mathematical expressions to model highly complex and nonlinear dynamics of atmospheric flow to produce numerical wind predictions (Yan and Ouyang,

2019). Gneiting et al. (2006) applied a second order model that uses the spatiotemporal covariance as a basis for estimation in Irish wind prediction. This model is quite suitable when the underlying process can be modeled as a Gaussian spatiotemporal field and is one basic type of spatiotemporal process used in the modeling of wind. Statistical methods rely on relevant historical data to predict future wind generation, traditionally using models such as the history average model (HA) (Liu and Guan, 2004) or the autoregressive integrated moving average model (ARIMA). A multichannel adaptive filter is used to forecast wind speed and direction using spatial correlations at multiple sites in Dowell et al. (2014). Markov chain–based statistical methods that leverage a graphic spatiotemporal learning-based model and statistical characteristics of aggregated wind generation are presented in He et al. (2014) to forecast wind generation output. As for machine learning methods, the recurrent neural network (RNN) is the most widely used neural network model in processing time series data. The long short-term memory (LSTM) model (Hochreiter and Schmidhuber, 1997) and the gated recurrent unit (GRU) model (Cho et al., 2014) are variants of the RNN that can avoid the gradient disappearance and explosion that can occur in RNN. The basic principles of the LSTM and GRU are roughly the same; they all use a gated mechanism to memorize as much long-term information as possible and are equally effective for various tasks (Chung et al., 2014). However, due to its complex structure, LSTM has a longer training time while the GRU model has a relatively simple structure, fewer parameters, and faster training ability.

However, both LSTM and GRU models only consider the dynamic change of temporal conditions and ignore the spatial dependence given that the change of wind condition is also restricted by the surrounding environment. Intuitively, the flow of wind in nearby regions affect each other. As there exists a strong correlation between wind sites located in a vicinity, the information collected from neighboring sites will help improve the prediction accuracy of target sites (Liu et al., 2020), which can be effectively handled by the convolutional neural network (CNN), which has shown its powerful ability to get spatial structural information (LeCun et al., 1998). A CNN uses filters to find relationships between neighboring inputs, which can make it easier for the network to converge on the correct solution.

The traditional CNN can obtain local spatial features in Euclidean space, but the relation network between each wind farm is in the form of a graph that is not a two-dimensional grid, which means the CNN model cannot deal with the complex topological structure of the wind relation network. However, the graph neural network (GNN) is good at handling the arbitrary graph structure data. To model the spatial correlations in graph-structured data, a graph CNN is developed to capture the spatial dependence, and a similar approach has been used in traffic prediction and shown good performance (Zhao et al., 2019).

There has been some work that applied GNN in wind prediction such as Mei Yu's (Yu et al., 2020) and Mahdi Khodayar's work (Khodayar and Wang, 2018), but all these works were based on NREL data sets, which are simulation data sets and do not provide real wind speed or power data. They showed the results in a one-time experiment only, whereas the neural network model's performance will be slightly different each time as the character of itself.

In this work, to demonstrate that the GNN model does have the ability to utilize the spatial correlations to improve accuracy in real wind prediction, we combine the graph convolutional network (GCN) with the GRU model to do prediction on five different wind data sets, including three wind speed data sets and two wind power data sets. Moreover, the GCN-GRU model is validated not only in 10 m height but also in 100 m height wind speed prediction. Furthermore, the final results are based on a collection of 10 experiments for each data set. The forecasting results show a steady state under different prediction horizons, which indicates that the GCN-GRU model can not only achieve short-term prediction, but can also be used for long-term wind prediction tasks.

REVIEW OF THE LITERATURE

Recently, given RNN's strong ability in processing time series problems, improved LSTM and GRU models have been applied in wind prediction. Zhewen Niu proposed a novel sequence-to-sequence model using the attention-based gated recurrent unit (AGRU) that improves accuracy of forecasting processes (Niu et al., 2020). A novel data-driven approach is proposed by Adam Kisvari in wind power forecasting by integrating data preprocessing, resampling, anomaly detection and treatment, feature engineering, and hyperparameter tuning based on gated recurrent deep learning models and critically compared with the algorithm of LSTM (Kisvari et al., 2021). Ruiguo Yu proposed an improved LSTM-enhanced forget-gate network model, abbreviated as LSTM-EFG, used in forecasting wind power (Yu et al., 2019). Xiaohui Yuan's simulation results showed that the PIWP obtained by the Beta-PSO-LSTM model has higher reliability and narrower interval bandwidth, which can provide decision support for the safe and stable operation of power systems (Yuan et al., 2019). In Yao Liu's work, a wind power short-term forecasting method based on discrete wavelet transform and LSTM networks (DWT_LSTM) is proposed. The discrete wavelet transform is introduced to decompose the nonstationary wind power time series into several components that have more stationarity and are easier to predict (Liu et al., 2019).

There is also some work that applies GNN in wind prediction, such as Mei Yu proposing the superposition graph neural network (SGNN) for feature extraction, which can maximize the use of spatial and temporal features for prediction. In the four offshore wind farms used in experiments, the mean square error of the method is reduced by 9.80%–22.53% compared with current advanced methods, and the prediction stability of the method has also been greatly improved (Yu et al., 2020). Mahdi Khodayar proposed GCDLA to capture spatial wind features as well as deep temporal features of the wind data at each wind site. Simulation results show the advantages of capturing deep spatial and temporal interval features in the proposed framework compared with the state-of-the-art deep learning models as well as shallow architectures in the recent literature (Khodayar

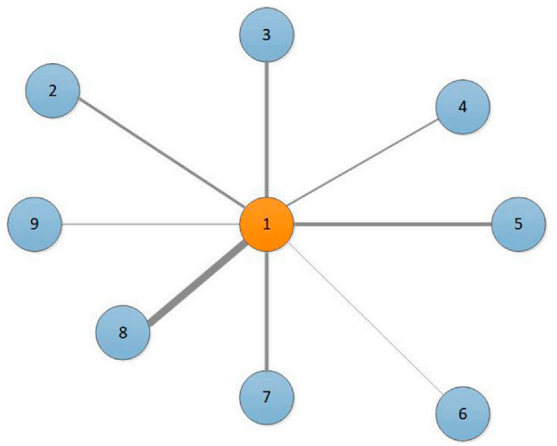

FIGURE 1 | The wind speed of the central site is affected by the wind from its surrounding site; the thickness of edges indicates the strength of the correlation.

and Wang, 2018). However, Mei Yu's work is about offshore wind, which is more stable than the wind on land as the surrounding terrain has a great influence on it. Mahdi Khodayar's work used the mutual information matrix to process the graph model. Furthermore, these two studies both worked on simulation data sets, and their results were based on a one-time experiment only, so it is hard to say that the improvement of prediction results by GNN is not occasional.

PROBLEM DEFINITION

In this paper, the goal of wind forecasting is to predict the wind speed or power in a certain period of time based on the historical wind information of several sites in a specific area. In our experiments, the wind information can be wind speed or wind power data.

Wind Network *G*

As shown in **Figure 1**, the wind of the target site may be influenced by the wind flows from its surrounding sites, we describe these wind sites in a specific area as a wind network although this network is unseen. We use an undirected graph $G = (V, E)$ to describe the topological structure of the wind network. We treat each wind site as a node, where V is a set of nodes, $V = \{v_1, v_2, \ldots, v_N\}$, N is the number of the nodes, and E is a set of edges between each wind site. The adjacency matrix A is used to represent the connection among wind sites, $A \in R^{N\times N}$. In this work, the adjacency matrix A is processed by normalization of a distance matrix, which is calculated by the distance between each wind site in the wind network. Note that distance is calculated in an "as the crow flies" sense and does not take the intervening topography into account.

Feature Matrix $X^{N\times P}$

We regard the wind speed or power as the attribute feature of the nodes in the network, expressed as $X^{N\times P}$, where P represents the number of node attribute features (the length of the historical time series) and $X_t \in R^{N\times i}$ is used to represent the speed or power on each site at time i. Again, the node attribute feature can be any

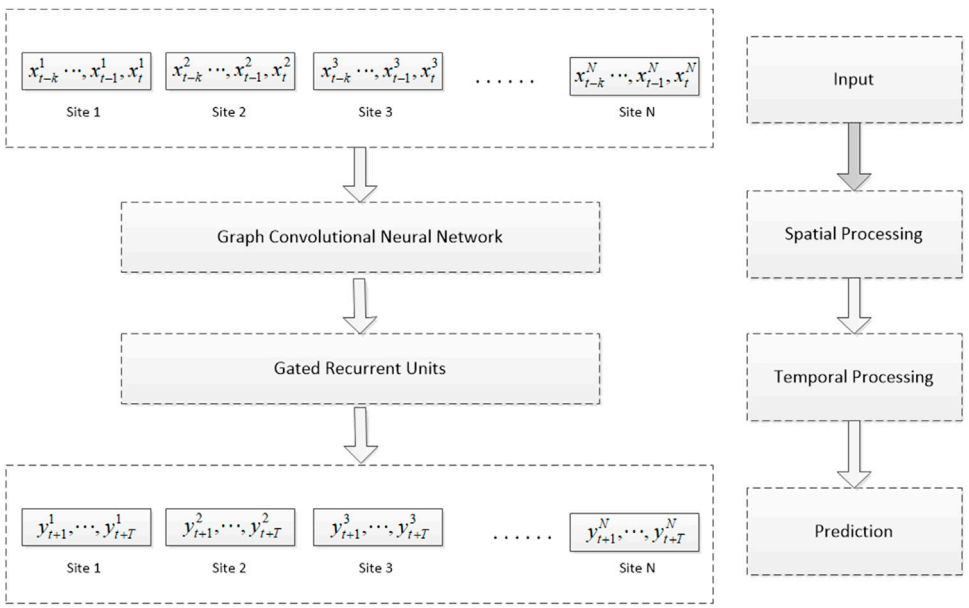

FIGURE 2 | Flow chart of the GCN-GRU model.

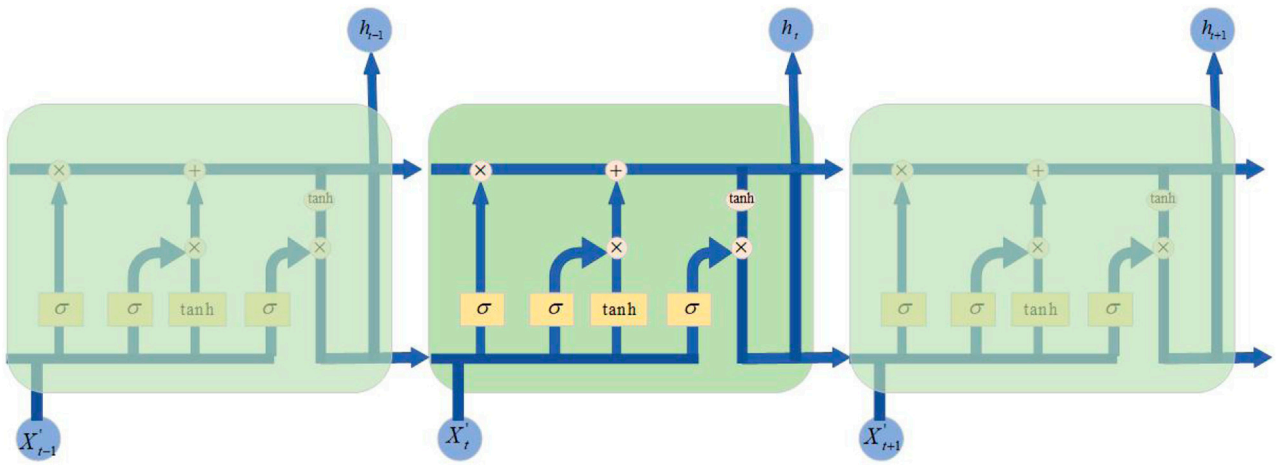

FIGURE 3 | The architecture of GRU model.

wind information, such as wind speed, wind direction, wind, or wind power.

MODEL

To evaluate the GCN's ability to capture the spatial features in wind prediction, we combine the GCN model with the GRU model and then compare the performance of the GCN-GRU model with the GRU model through several evaluation metrics. The flow chart of the GCN-GRU model is shown in **Figure 2**.

Spatial Model

The target site is influenced by a complex wind network formed by its surrounding sites, and in this work, we use the GCN to capture spatial correlations. Existing GCNs can be divided into two main categories: spatial- and spectral-based GCNs (Liu et al., 2020). Spatial-based GCNs directly define the convolutional operation by operating on spatially close—the first or second order neighbors—such that the convolutional operation using different weight matrices for nodes with different degrees (Duvenaud et al., 2015). Considering the wind farms are all connected if we treat them in a graph domain, we choose to use the spectral-based GCN to capture the effect of surrounding wind farms.

In 2014, a spectral network was proposed in Bruna et al. (2013). It transforms the samples into Fourier domains to perform convolution operations through Fourier transform, and then the samples are transformed back to the graph domains through inverse Fourier transform. Specifically, the graph convolutional operation of the sample $X \in R^N$ can be defined as

$$G*X = UG_W U^T X, \tag{1}$$

where $*$ is the graph convolutional operation, U is the matrix of eigenvectors of the normalized graph Laplacian $L = I_N - D^{-\frac{1}{2}}AD^{-\frac{1}{2}} = U \wedge U^T$ with a diagonal matrix of its eigenvalues $\wedge$ and $U^T X$ being the graph Fourier transform of X, and $G_W = diag(U^T G)$ is the filter parameterized by $W \in R^n$.

In Defferrard et al. (2016), the Chebyshev spectral CNN (ChebNet) that approximate G_W by the truncated expansion of Chebyshev polynomials $T_k(X)$ up to K^{th} is proposed. To circumvent the problem of being computationally expensive, G_W can be approximated by a truncated expansion in terms of Chebyshev polynomials $T_k(X)$ up to the K^{th} order:

$$G*X \approx \sum_{k=0}^{K} W_k T_k\left(\frac{2L}{\lambda_{max}} - I_n\right)X \tag{2}$$
$$T_k(X) = 2XT_{k-1}(X) - T_{k-2}(X)T_0(X) = 1\,T_1(X) = X,$$

where $W \in R^K$ is a vector that consists of Chebyshev coefficients and λ_{max} is the largest eigenvalue. Because the operation is a K^{th} order polynomial in the Laplacian, it is K-localized.

Moreover, if we set $K = 1$ and $\lambda_{max} = 2$, we can get a linear function on the graph Laplacian spectrum:

$$G*X \approx W_0 X - W_1 D^{-\frac{1}{2}}AD^{-\frac{1}{2}}X. \tag{3}$$

Furthermore, to restrain the number of parameters and alleviate overfitting problems, we assume that $W = W_0 = W_1$ (Berg et al., 2017) and then set $\hat{A} = A + I_n$ to do renormalization to avoid vanishing gradients problems (Kipf and Welling, 2016), which lead to the following formula:

$$G*X \approx W\left(I_n + D^{-\frac{1}{2}}AD^{-\frac{1}{2}}\right)X = \hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}XW \tag{4}$$

where $\hat{D}$ denotes the diagonal matrix of node degree with $D_{ii} = \sum_{j=1}^{n}\hat{a}_{ij}$.

In summary, we use a two2-layer GCN model based on **Eq. 5** to learn spatial features from wind data and it can be expressed as

$$f(X_t, \Lambda) = \sigma(\Lambda(Relu(\Lambda X_t W_0))W_1), \tag{5}$$

TABLE 1 | The best results in Canada climate wind speed experiments.

Data set		Canada climate wind speed–9 stations data set			Canada climate wind speed–18 stations data set		
Step	**Method**	**RMSE**	**MAE**	**Accuracy**	**RMSE**	**MAE**	**Accuracy**
1 h	GCN-GRU	5.8995 ± 0.0246	4.346 ± 0.0135	0.7981 ± 0.0008	3.3798 ± 0.0365	2.5845 ± 0.0394	0.8046 ± 0.0021
	GRU	5.9826 ± 0.0271	4.4271 ± 0.0199	0.7952 ± 0.0009	3.6912 ± 0.0218	2.7659 ± 0.0205	0.7866 ± 0.0013
	ARIMA	6.2386	4.5878	0.7864	3.8255	2.8427	0.7788
	SVR	6.1082	4.4884	0.7809	3.6949	2.724	0.7815
	HA	8.5224	6.3859	0.6556	5.2101	3.9004	0.6409
2 h	GCN-GRU	6.9127 ± 0.0825	5.153 ± 0.0609	0.7636 ± 0.0028	3.8814 ± 0.0404	2.9394 ± 0.0348	0.7757 ± 0.0023
	GRU	7.0757 ± 0.0648	5.2697 ± 0.0357	0.758 ± 0.0022	4.3106 ± 0.0544	3.3003 ± 0.0859	0.7564 ± 0.0021
	ARIMA	7.2344	5.3152	0.7526	4.3832	3.2449	0.7466
	SVR	7.0682	5.2229	0.7416	4.2192	3.1079	0.7481
	HA	9.6091	7.3293	0.6121	6.0377	4.5414	0.5839
3 h	GCN-GRU	7.8451 ± 0.0578	5.8835 ± 0.0398	0.7323 ± 0.002	4.2518 ± 0.0233	3.206 ± 0.0155	0.7548 ± 0.0013
	GRU	7.8767 ± 0.1378	5.8586 ± 0.095	0.7313 ± 0.0047	4.6189 ± 0.0286	3.574 ± 0.0267	0.7409 ± 0.0016
	ARIMA	8.0936	5.9261	0.7238	4.9996	3.8501	0.7195
	SVR	7.9023	5.8444	0.7073	4.643	3.4364	0.7215
	HA	10.8836	8.4161	0.5616	6.6606	5.0407	0.5419

TABLE 2 | The results of NREL wind experiments.

Data sets		NREL wind speed data set			NREL wind power data set		
Step	**Method**	**RMSE**	**MAE**	**Accuracy**	**RMSE**	**MAE**	**Accuracy**
1 h	GCN-GRU	0.9847 ± 0.0086	0.7151 ± 0.0085	0.8896 ± 0.001	0.1053 ± 0.0006	0.071 ± 0.0012	0.8084 ± 0.0011
	GRU	1.0195 ± 0.0084	0.7445 ± 0.0067	0.8857 ± 0.0009	0.105 ± 0.0009	0.0684 ± 0.0017	0.8076 ± 0.001
	ARIMA	1.0677	0.7707	0.8803	0.2192	0.1271	0.6417
	SVR	1.047	0.7537	0.8804	0.1166	0.0882	0.7718
	HA	1.7979	1.3933	0.7913	0.1767	0.1217	0.6625
2 h	GCN-GRU	1.2761 ± 0.0266	0.9407 ± 0.0188	0.8549 ± 0.0018	0.1317 ± 0.0004	0.0903 ± 0.0009	0.7604 ± 0.0008
	GRU	1.3521 ± 0.0257	1.0013 ± 0.0224	0.8481 ± 0.003	0.1325 ± 0.0012	0.0904 ± 0.0015	0.7573 ± 0.0012
	ARIMA	1.3448	0.9764	0.8494	0.2446	0.1487	0.6006
	SVR	1.3207	0.9602	0.8478	0.1426	0.108	0.7184
	HA	2.4512	1.949	0.7157	0.2276	0.1693	0.5654
3 h	GCN-GRU	1.5575 ± 0.0488	1.1631 ± 0.0385	0.8246 ± 0.0042	0.1494 ± 0.0009	0.1063 ± 0.0015	0.7272 ± 0.0017
	GRU	1.5938 ± 0.0347	1.2006 ± 0.0312	0.8181 ± 0.0035	0.1522 ± 0.0014	0.108 ± 0.0022	0.7208 ± 0.0015
	ARIMA	1.5725	1.1489	0.8239	0.266	0.1675	0.5648
	SVR	1.5373	1.1337	0.8213	0.15986	0.1202	0.6836
	HA	0.9847 ± 0.0086	0.7151 ± 0.0085	0.8896 ± 0.001	0.2642	0.2077	0.4937

where X_t represents the features matrix at time t, Λ is the adjacency matrix calculated by the distance between each wind site, $\Lambda = \hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}$ denotes preprocessing step in **Eq. 4**, $\hat{A} = A + I_N$ is a matrix with self-connection structure, and $\hat{D}$ is a degree matrix. W_0 and W_1 represent the weight matrix in the first and second GCN layer, and $\sigma(\cdot)$ and $Relu()$ represent the activation function.

Temporal Model

After spatial processing by the GCN model, we get a new feature matrix $X_t^{'}$, and we can have new time series data to do time prediction. In this work, as our key goal is to evaluate the GCN's performance in wind prediction, a simpler and faster model to do time forecasting is needed. As the GRU model has a relatively simple structure and fewer parameters than LSTM, we chose the GRU model to obtain temporal dependence from the wind data.

As shown in **Figure 3**, h_{t-1} denotes the hidden state at time $t-1$; $X_t^{'}$ denotes the wind information at time t; r_t is the reset gate, which is used to control the degree of ignoring the status information at the previous moment; u_t is the update gate, which is used to control the degree to which the status information at the previous time is brought into the current status; c_t is the memory content stored at time t; and h_t is output state at time t. The GRU obtains the wind status at time t by taking the hidden status at time $t-1$ and the current wind information as inputs. While capturing the wind information at the current moment, the model still retains the changing trend of historical wind information and has the ability to capture temporal dependence.

$$\begin{aligned} u_t &= \sigma\left(W_u\left[X_t^{'}, h_{t-1}\right] + b_u\right) \\ r_t &= \sigma\left(W_r\left[X_t^{'}, h_{t-1}\right] + b_r\right) \\ c_t &= tanh\left(W_c\left[X_t^{'}, (r_t * h_{t-1})\right] + b_c\right) \end{aligned} \tag{6}$$

TABLE 3 | NRGStream wind power experiments.

Data sets		NRGStream wind power data set		
Step	**Method**	**RMSE**	**MAE**	**Accuracy**
1 h	GCN-GRU	32.3046 ± 0.15	22.1363 ± 0.3941	0.8233 ± 0.0008
	GRU	32.2158 ± 0.1266	21.3319 ± 0.4801	0.8238 ± 0.0007
	ARIMA	33.7329	20.4416	0.8155
	SVR	32.5486	20.0331	0.821
	HA	27.6618	13.7631	0.6493
2 h	GCN-GRU	21.223 ± 15.2436	15.4518 ± 11.0381	0.7591 ± 0.0023
	GRU	30.9518 ± 16.0936	21.9504 ± 11.6982	0.7592 ± 0.0015
	ARIMA	45.6305	27.3199	0.7499
	SVR	44.144	27.0784	0.7536
	HA	34.6944	18.4608	0.5594
3 h	GCN-GRU	21.3579 ± 16.2888	15.8231 ± 12.2517	0.7158 ± 0.0066
	GRU	29.5493 ± 19.6803	21.2965 ± 14.4141	0.7103 ± 0.005
	ARIMA	54.9295	33.1688	0.6997
	SVR	13.2146	7.4584	0.7105
	HA	39.3619	22.5405	0.5011

TABLE 4 | The average value of prediction results of all wind stations in Canada Climate wind speed experiments.

Data set		Canada climate wind speed–9 stations data set			Canada climate wind speed–18 stations data set		
Step	**Method**	**RMSE**	**MAE**	**Accuracy**	**RMSE**	**MAE**	**Accuracy**
1 h	GCN-GRU	6.0233 ± 0.0167	4.3714 ± 0.0174	0.7542 ± 0.0007	3.4313 ± 0.0836	2.6277 ± 0.0915	0.7538 ± 0.007
	GRU	6.1196 ± 0.0165	4.4138 ± 0.0115	0.7501 ± 0.0007	3.4717 ± 0.0186	2.5915 ± 0.0249	0.7534 ± 0.0015
	ARIMA	6.5397	4.6785	0.7326	3.7026	2.7231	0.7381
	SVR	6.3514	4.509	0.725	3.5468	2.6136	0.7401
2 h	GCN-GRU	6.7617 ± 0.0391	4.9569 ± 0.033	0.7245 ± 0.0015	3.7748 ± 0.0356	2.8443 ± 0.0359	0.7306 ± 0.0029
	GRU	6.9655 ± 0.0431	5.0784 ± 0.0221	0.716 ± 0.0018	3.9477 ± 0.0136	2.9254 ± 0.0097	0.72 ± 0.0009
	ARIMA	7.3284	5.2737	0.7007	4.1924	3.079	0.7034
	SVR	7.1502	5.1283	0.6835	4.0198	2.9614	0.7026
3 h	GCN-GRU	7.4551 ± 0.0253	5.5149 ± 0.0224	0.6972 ± 0.001	4.0913 ± 0.0337	3.0755 ± 0.0382	0.7087 ± 0.0027
	GRU	7.6366 ± 0.0603	5.5905 ± 0.0324	0.6894 ± 0.0025	4.2831 ± 0.0166	3.1722 ± 0.0133	0.6967 ± 0.0012
	ARIMA	7.9607	5.7519	0.6756	4.5974	3.3767	0.6753
	SVR	7.7627	5.6278	0.6517	4.3796	3.2382	0.6724

TABLE 5 | The average value of prediction results of all wind stations in NREL wind experiments

Data sets		NREL wind speed data set			NREL wind power data set		
Step	**Method**	**RMSE**	**MAE**	**Accuracy**	**RMSE**	**MAE**	**Accuracy**
1 h	GCN-GRU	1.0418 ± 0.0056	0.7581 ± 0.004	0.879 ± 0.0006	0.1099 ± 0.0004	0.0731 ± 0.0008	0.7894 ± 0.0008
	GRU	1.0548 ± 0.0038	0.7639 ± 0.0031	0.8775 ± 0.0005	0.1103 ± 0.0005	0.0717 ± 0.001	0.7885 ± 0.001
	ARIMA	1.1071	0.7956	0.8713	0.2342	0.1367	0.6103
	SVR	1.0899	0.7787	0.8706	0.1196	0.0909	0.7528
2 h	GCN-GRU	1.348 ± 0.0162	0.9978 ± 0.015	0.8436 ± 0.0019	0.1347 ± 0.0004	0.0928 ± 0.0007	0.742 ± 0.0007
	GRU	1.4087 ± 0.0251	1.0516 ± 0.0251	0.8365 ± 0.0029	0.136 ± 0.0005	0.0935 ± 0.0011	0.7396 ± 0.001
	ARIMA	1.3754	0.9948	0.8403	0.2577	0.1578	0.5717
	SVR	1.3559	0.9816	0.8375	0.1432	0.1088	0.7021
3 h	GCN-GRU	1.6184 ± 0.0264	1.2156 ± 0.0231	0.8121 ± 0.0031	0.1519 ± 0.0006	0.1086 ± 0.0013	0.7081 ± 0.0011
	GRU	1.6799 ± 0.0207	1.2762 ± 0.0191	0.805 ± 0.0025	0.1545 ± 0.0005	0.1116 ± 0.0012	0.7033 ± 0.0011
	ARIMA	1.5947	1.1641	0.8148	0.2776	0.1761	0.5379
	SVR	1.5673	1.1579	0.8103	0.1596	0.1207	0.6669

$$h_t = u_t * h_{t-1} + (1 - u_t) * c_t$$

In these equations, $X_t^{'}$ is the new wind information processed by GCN, W and b represent the weights and deviations in the training process.

Loss Function

In the training step, the goal of the GCN-GRU model is to learn the well-fitted parameters by minimizing the error between the real wind data and the predicted wind value. We use Y_t and $\hat{Y}_t$ to denote the real wind and the predicted speed or power data, respectively. The loss function is shown in **Eq. 7**. The first term is used to minimize the error between the real traffic speed and the prediction. The second term L_{reg} is an L2 regularization term that helps to avoid an overfitting problem, and λ is a hyperparameter:

$$loss = \left\| Y_t - \hat{Y}_t \right\| + \lambda L_{reg}. \quad (7)$$

EXPERIMENTS

In this section, we evaluate the prediction performance of GCN-GRU model on five data sets.

Data Description

(1) Two Canadian climate wind speed data sets. One of them includes nine stations, and the other includes 19 stations. These real wind speed data are all recorded at 10 m height, and the stations are all located in Alberta. We chose the wind speed data from the historical climate data set, which can be downloaded from the Canada Environment and natural resources website. (https://climate.weather.gc.ca/historical_data/search_historic_data_e.html).

(2) Two Eastern wind integration data sets. One includes wind speed data from 10 sites, and the other contains the corresponding wind power data at the same places. The wind speed data was downloaded from the National Renewable Energy Laboratory (NREL) website (https://www.nrel.gov/grid/wind-integration-data.html), and the corresponding wind power data set was computed by the Renewable Energy Potential (reV) Model (Maclaurin et al., 2019) by a given set of latitude and longitude coordinates. Different from the Canada climate wind speed data, the NREL wind speed data are simulated by predictive models at 100 m height.

(3) One NRGStream wind power data set. We chose seven wind farms listed by the Alberta Electric System Operator (AESO) and use the hourly wind power data from 2018 to 2019.

In the experiments, the input data was normalized to the interval [0,1]. In addition, 80% of the data was used as the training set and the remaining 20% was used as the testing set. We predicted the wind speed of the next 1, 2, and 3 h. The detailed information of these data sets can be found in **Supplementary Materials**.

TABLE 6 | The average value of prediction results of all wind stations in NRGStream wind power experiments.

Data sets		NRGStream wind power data set		
Step	**Method**	**RMSE**	**MAE**	**Accuracy**
1 h	GCN-GRU	12.3121 ± 0.0284	8.5175 ± 0.0901	0.7756 ± 0.0013
	GRU	12.0442 ± 0.0571	8.133 ± 0.1196	0.7814 ± 0.0029
	ARIMA	12.4423	7.3886	0.7826
	SVR	11.9899	7.078	0.7885
2 h	GCN-GRU	15.7505 ± 0.122	11.5557 ± 0.1454	0.7248 ± 0.0037
	GRU	15.7717 ± 0.1568	11.2753 ± 0.1789	0.7234 ± 0.005
	ARIMA	16.3477	9.8913	0.718
	SVR	15.5462	9.2681	0.7281
3 h	GCN-GRU	18.5172 ± 0.2087	13.8651 ± 0.2239	0.6829 ± 0.0056
	GRU	18.6734 ± 0.3189	13.795 ± 0.373	0.6798 ± 0.0082
	ARIMA	19.624	12.3354	0.662
	SVR	18.1365	11.1136	0.6824

Evaluation Metrics

To evaluate the performance of the GCN model, we use three metrics of the difference between the real wind data Y_t and the prediction $\widehat{Y_t}$, including

(1) Root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}.$$

(2) Mean absolute error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_t - \hat{Y}_t\right|.$$

(3) Accuracy

$$Accuracy = 1 - \frac{\left\|Y_t - \hat{Y}_t\right\|_F}{\|Y\|_F}.$$

Parameter Settings

Our goal is to see whether the GCN can capture the environment information and improve the prediction result. As the GCN-GRU and GRU share the same GRU model, we use the same parameters for GCN-GRU and GRU, such as the amount of GRU unit, the learning rate, the training epoch, and batch size.

For the input layer, the training data set (80% of the overall data) is taken as input in the training process, and the remaining data is used as input in the testing process. The GCN-GRU model is trained using the Adam optimizer.

Experiment Results

In this part, we compare the performance of the GCN-GRU and GRU models and also show the prediction results with some other baseline methods, includin: (1) HA (Pinson, 2013),

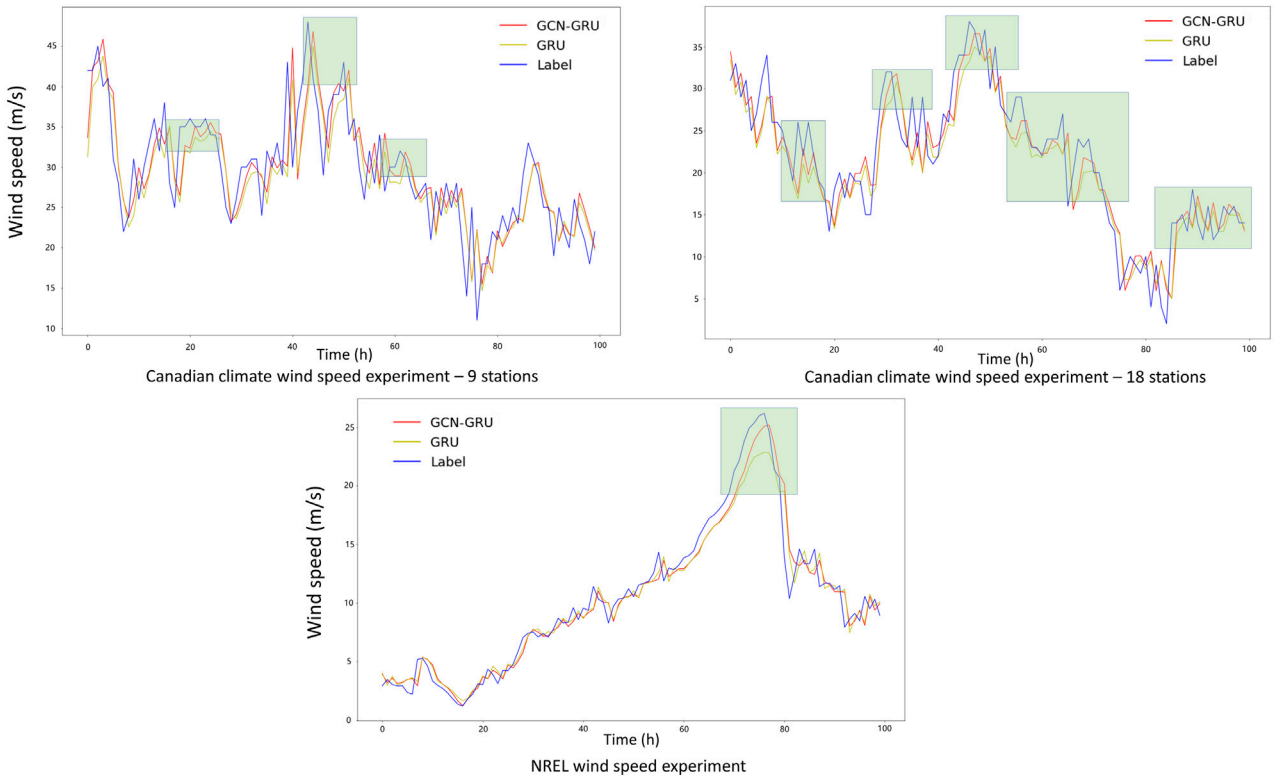

FIGURE 4 | Comparison of three wind speed prediction experiments between GCN-GRU and GRU; the light green areas are where GCN-GRU's prediction curves closer to the real curve, namely, GCN-GRU performs better than GRU in one-step wind speed prediction.

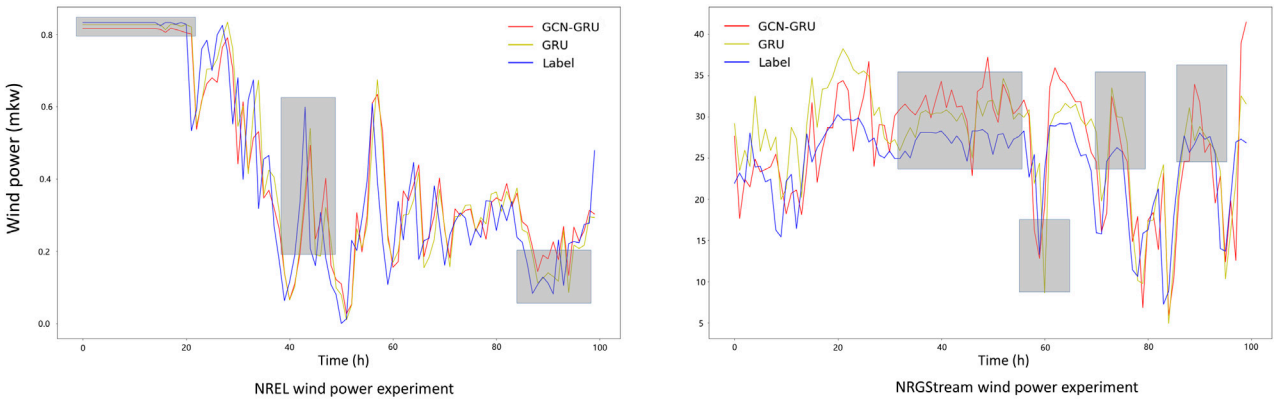

FIGURE 5 | Comparison of two wind power prediction experiments between GCN-GRU and GRU; the gray areas are where GRU's prediction curves closer to real curve, namely, GRU performs better than GCN-GRU in one-step wind power prediction.

which uses the average wind information in the historical periods as the prediction; (2) ARIMA (Stohl et al., 1995), which fits the observed time series into a parametric model to predict future wind data; and (3) the support vector regression model (SVR) (Smola and Schölkopf, 2004), which uses historical data to train the model and obtains the relationship between the input and output and then predicts by giving the future wind data. We use the linear kernel, and the penalty term is 0.001.

As the GCN-GRU model is also based on GRU, according to the experiments, we find that the GRU's prediction is not quite stable; that is, the result is a little different each time as the

neural network optimization based on stochastic gradient descent to find the global minimum is different each time. To show that the prediction improvement by the GCN is not incidental, we repeat each experiment 10 times to calculate the standard error and show the range of the prediction results. As prediction accuracy of each station is different, **Tables 1–3** give the result of a specific station whose prediction accuracy is best.

We can see that the neural network models, including the GCN-GRU and GRU models, achieve better prediction precision than other baselines, such as ARIMA, SVR, and HA. For example, for the 1-h wind speed prediction on the data set of "Canada climate wind speed experiment–nine stations," the RMSE of the GCN-GRU and the GRU models are approximately 5.4% and 4.1% lower than that of the ARIMA model, and the accuracy of these are improved by 1.4% and 1.1%. Compared with the SVR model, the RMSE of the GCN-GRU and the GRU models are reduced by 3.4% and 2.1%, and the accuracy is approximately 2.2% and 1.8% higher than that of the SVR model. The RMSE error of the GCN-GRU and GRU models are reduced by approximately 30.7% and 29.8% compared with the HA model, and the accuracies are approximately 14.2% and 13.9% higher than that of HA. Next, we focus on the comparisons between GCN-GRU's and GRU's results.

Tables 1, **2** show the prediction results of the GCN-GRU model, GRU model and other baseline methods for Canada climate wind speed data sets, NREL wind speed data, and NREL wind power data, respectively. By comparing the experiments of GCN-GRU and GRU in these four experiments, we can see that all the GCN-GRU results are better than the GRU model. It shows that GCN can capture some geographical information and improve the prediction performance.

In the nine-station Canada climate wind speed prediction, we can see that, comparing with GRU, there is 0.0831 improvement for RMSE evaluation and approximately 0.3% improvement in accuracy. In the nine-station Canada climate wind speed prediction, we can see that there is 0.3114 improvement for RMSE evaluation and approximately 1.8% improvement in accuracy. To explain the difference in the Canada climate nine- and 18-station experiments, although they are all wind prediction at 10 m height, we note that more wind stations can provide more surrounding information; that is why the 18 stations experiment's result is much better than nine stations.

By comparing the experiment results of NREL wind speed with the Canada climate wind speed experiment, we can see that the accuracy is much higher for all models. It mainly because the NREL data is about the wind at 100 m height, and the change of the wind is more stable than the wind of the Canada climate at 10 m height, which makes the model easier to predict. From these four experiments, we can see that the GCN model can improve the prediction accuracy.

However, we find that the prediction accuracy of all these models gets lower in NREL wind power prediction when we compare with NREL wind speed prediction. What is more, the GCN-GRU model almost achieves same prediction accuracy with the GRU model in 1-h step prediction. The reason for these unsatisfactory results mainly is that the NREL data sets are simulated by mathematic models other than real recorded data. Thus, we do another experiment based on the NRGStream data set, which records the real wind power from wind farms in Alberta, Canada.

According to the NRGStream wind power experiment, we can find that the GCN-GRU model has even worse performance in 1 h prediction. In the simulated and real wind power experiments, we can see that the GCN does not perform as well as in wind speed prediction. It is mainly because the wind power data is not so smooth as the wind speed data due to the operation intervention of wind farms, such as when the wind speed is higher than the limitation of the turbines; the wind power data will always be the highest value other than keeping increasing when the wind speed increases. If the wind speed is less than the specific speed that the turbines need, the wind power data is zero directly other than a small value like the wind speed. What is more, some wind farms may directly shut down wind turbines to protect them when the wind is extremely strong while others nearby may keep running. The character of wind power data leads to the lower prediction accuracy.

What is more, for the sake of showing the improvement of prediction results by GCN is not occasional, we also calculate the average value of prediction results of all wind stations in wind speed and power forecasting experiments. From **Tables 4–6**, we can see that GCN not only improves the prediction result for a specific station, it also owns the ability to enhance prediction accuracy for all wind stations although it still has a drawback when predict the wind power directly.

Comparison of the Prediction Between GCN-GRU and GRU

Figures 4, **5** are the prediction comparisons between GCN-GRU and GRU in wind speed and wind power prediction. In **Figure 4**, we can see that the change of wind speed is smooth although the curve fluctuates strongly. However, the interventions of wind turbines are obvious in both the NREL and NRGStream wind power, there exist many horizontal lines in the wind power data, and these interventions may lead to the bad performance of GCN in wind power predictions.

CONCLUSION

In this paper, a GCN model is proposed to improve precision of wind prediction; we evaluate the GCN-GRU model not only on 10 and 100 m height wind speed prediction, but also through wind power prediction. To show the performance of the GCN-GRU model, the final results are based on a collection of 10 experiments for each data set; the forecasting results are quite stable in different prediction steps. Although the GCN-GRU model does not achieve better performance than the GRU model in short-term

NRGStream wind power prediction, the results of five different wind data sets prove GCN's ability to capture spatial influence and improve the prediction results. Finally, our study offers new findings and may add considerable value to the energy management field; our finding confirms that the GCN model has positive influence on wind prediction. What is more, the failure of GCN in short-term wind power prediction also tells us that, if we want to achieve a higher accuracy in wind power prediction, it is better to combine the wind speed prediction with the wind power data.

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

Conceived and designed the experiments: ZL TW Analyzed the data: ZL Wrote the paper: ZL TW.

FUNDING

ZL is supported by the CSC Scholarship offered by the China Scholarship Council.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenvs.2022.836050/full#supplementary-material

REFERENCES

Ahmadian, A., Sedghi, M., Aliakbar-Golkar, M., Fowler, M., and Elkamel, A. (2016). Two-layer Optimization Methodology for Wind Distributed Generation Planning Considering Plug-In Electric Vehicles Uncertainty: A Flexible Active-Reactive Power Approach. *Energ. Convers. Management* 124, 231–246. doi:10.1016/j.enconman.2016.07.025

Ahmed, M. S., and Cook, A. R. (1979). *Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques*. No. 722. Washington: Transportation Research Board.

Berg, R. V. D., Kipf, T. N., and Welling, M. (2017). *Graph Convolutional Matrix Completion*. arXiv [Preprint]. Available at: https://arxiv.org/abs/1706.02263.

Bruna, J., Zaremba, W., Szlam, A., and LeCun, Y. (2013). *Spectral Networks and Locally Connected Networks on Graphs*. arXiv [Preprint]. Available at: https://arxiv.org/abs/1312.6203.

Cho, K., Van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*. arXiv [Preprint]. Available at: https://arxiv.org/abs/1409.1259.

Choi, S.-D., Baek, S.-Y., Chang, Y.-S., Wania, F., Ikonomou, M. G., Yoon, Y.-J., et al. (2008). Passive Air Sampling of Polychlorinated Biphenyls and Organochlorine Pesticides at the Korean Arctic and Antarctic Research Stations: Implications for Long-Range Transport and Local Pollution. *Environ. Sci. Technol.* 42 (19), 7125–7131. doi:10.1021/es801004p

Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. arXiv [Preprint]. Available at: https://arxiv.org/abs/1412.3555.

Defferrard, M., Bresson, X., and Vandergheynst, P. (2016). Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. *Adv. Neural Inf. Process. Syst.* 29, 3844–3852.

Dowell, J., Weiss, S., Hill, D., and Infield, D. (2014). Short-term Spatio-Temporal Prediction of Wind Speed and Direction. *Wind Energy* 17 (12), 1945–1955. doi:10.1002/we.1682

Duvenaud, D., Maclaurin, D., Aguilera-Iparraguirre, J., Gómez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., et al. (2015). *Convolutional Networks on Graphs for Learning Molecular Fingerprints*, Advances in neural information processing systems.

Gneiting, T., Genton, M., and Guttorp, P. (2006). Geostatistical Space-Time Models, Stationarity, Separability, and Full Symmetry. *Monogr. Stat. Appl. Probab.* 107, 151–175. doi:10.1201/9781420011050.ch4

He, M., Yang, L., Zhang, J., and Vittal, V. (2014). A Spatio-Temporal Analysis Approach for Short-Term Forecast of Wind Farm Generation. *IEEE Trans. Power Syst.* 29 (4), 1611–1622. doi:10.1109/tpwrs.2014.2299767

Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput.* 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735

Hong, K. R., Qiu, L. S., Yang, D. X., and Jiang, M. (2021). Spatio-Temporal Evolution and Correlation Analysis of Urban Land Use Patterns and Air Quality in Pearl River Delta, China. *Front. Environ. Sci.* 472. doi:10.3389/fenvs.2021.698383

Khodayar, M., and Wang, J. (2018). Spatio-temporal Graph Deep Neural Network for Short-Term Wind Speed Forecasting. *IEEE Trans. Sustainable Energ.* 10 (2), 670–681. doi:10.1109/TSTE.2018.2844102

Kipf, T. N., and Welling, M. (2016). *Semi-supervised Classification with Graph Convolutional Networks*. arXiv [Preprint]. Available at: https://arxiv.org/abs/1609.02907.

Kisvari, A., Lin, Z., and Liu, X. (2021). Wind Power Forecasting - A Data-Driven Method along with Gated Recurrent Neural Network. *Renew. Energ.* 163, 1895–1909. doi:10.1016/j.renene.2020.10.119

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based Learning Applied to Document Recognition. *Proc. IEEE* 86 (11), 2278–2324. doi:10.1109/5.726791

Liu, J., and Guan, W. (2004). A Summary of Traffic Flow Forecasting Methods [J]. *J. Highw. transportation Res. Dev.* 3, 82–85.

Liu, Y., Guan, L., Hou, C., Han, H., Liu, Z., Sun, Y., et al. (2019). Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform. *Appl. Sci.* 9 (6), 1108. doi:10.3390/app9061108

Liu, Y., Liu, Y., and Yang, C. (2020). Modulation Recognition with Graph Convolutional Network. *IEEE Wireless Commun. Lett.* 9 (5), 624–627. doi:10.1109/lwc.2019.2963828

Maclaurin, G. J., Grue, N. W., Lopez, A. J., and Heimiller, D. M. (2019). *The Renewable Energy Potential (Rev) Model: A Geospatial Platform for Technical Potential and Supply Curve Modeling*. No. NREL/TP-6A20-73067. Golden, CO (United States): National Renewable Energy Lab.(NREL.

Mu, L., Su, J., Mo, X., Peng, N., Xu, Y., Wang, M., et al. (2021). The Temporal-Spatial Variations and Potential Causes of Dust Events in Xinjiang Basin During 1960–2015. *Front. Environ. Sci.* 360. doi:10.3389/fenvs.2021.727844

Niu, Z., Yu, Z., Tang, W., Wu, Q., and Reformat, M. (2020). Wind Power Forecasting Using Attention-Based Gated Recurrent Unit Network. *Energy* 196, 117081. doi:10.1016/j.energy.2020.117081

Paramati, S. R., Shahzad, U., and Doğan, B. (2022). The Role of Environmental Technology for Energy Demand and Energy Efficiency: Evidence from OECD Countries. *Renew. Sustainable Energ. Rev.* 153, 111735. doi:10.1016/j.rser.2021.111735

Pinson, P. (2013). Wind Energy: Forecasting Challenges for its Operational Management. *Stat. Sci.* 28 (4), 564–585. doi:10.1214/13-sts445

Shahzad, U., Schneider, N., and Ben Jebli, M. (2021). How Coal and Geothermal Energies Interact with Industrial Development and Carbon Emissions? An Autoregressive Distributed Lags Approach to the Philippines. *Resour. Pol.* 74, 102342. doi:10.1016/j.resourpol.2021.102342

Skogland, T. (1984). Wild Reindeer Foraging-Niche Organization. *Ecography* 7 (4), 345–379. doi:10.1111/j.1600-0587.1984.tb01138.x

Smola, A. J., and Schölkopf, B. (2004). A Tutorial on Support Vector Regression. *Stat. Comput.* 14 (3), 199–222. doi:10.1023/b:stco.0000035301.49549.88

Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H. (1995). Interpolation Errors in Wind fields as a Function of Spatial and Temporal Resolution and Their Impact on Different Types of kinematic Trajectories. *J. Appl. Meteorol.* 34 (10), 2149–2165. doi:10.1175/1520-0450(1995)034<2149:ieiwfa>2.0.co;2

Wizelius, T. (2015). *Developing Wind Power Projects: Theory and Practice*. London: Routledge.

Yan, J., and Ouyang, T. (2019). Advanced Wind Power Prediction Based on Data-Driven Error Correction. *Energ. Convers. Manag.* 180, 302–311. doi:10.1016/j.enconman.2018.10.108

Yu, M., Zhang, Z., Li, X., Yu, J., Gao, J., Liu, Z., et al. (2020). Superposition Graph Neural Network for Offshore Wind Power Prediction. *Future Generation Computer Syst.* 113, 145–157. doi:10.1016/j.future.2020.06.024

Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., et al. (2019). LSTM-EFG for Wind Power Forecasting Based on Sequential Correlation Features. *Future Generation Computer Syst.* 93, 33–42. doi:10.1016/j.future.2018.09.054

Yuan, X., Chen, C., Jiang, M., and Yuan, Y. (2019). Prediction Interval of Wind Power Using Parameter Optimized Beta Distribution Based LSTM Model. *Appl. Soft Comput.* 82, 105550. doi:10.1016/j.asoc.2019.105550

Zha, X., Ouyang, T., Qin, L., Xiong, Y., and Huang, H. (2016). Selection of Time Window for Wind Power Ramp Prediction Based on Risk Model. *Energ. Convers. Management* 126, 748–758. doi:10.1016/j.enconman.2016.08.064

Zhang, Z., Zhang, X., Gong, D., Kim, S.-J., Mao, R., and Zhao, X. (2016). Possible Influence of Atmospheric Circulations on winter Haze Pollution in the Beijing-Tianjin-Hebei Region, Northern China. *Atmos. Chem. Phys.* 16 (2), 561–571. doi:10.5194/acp-16-561-2016

Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., et al. (2019). T-gcn: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Trans. Intell. Transportation Syst.* 21 (9), 3848–3858.

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.

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.