POLITECNICO DI TORINO
Repository ISTITUZIONALE

Hierarchical Dynamic Hypergraph Learning for Ultrashort-Term Wind Farm Cluster Power Forecasting With Input-Aware Meteorological Integration Method

Original
Hierarchical Dynamic Hypergraph Learning for Ultrashort-Term Wind Farm Cluster Power Forecasting With Input-Aware Meteorological Integration Method / Yang, M., Huang, Y., Wang, B.o., Xu, C., Chen, J., Hosseini Imani, M., Huang, T.. - In: IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS. - ISSN 1551-3203. - (2026), pp. 1-12. [10.1109/TII.2026.3716951]

Availability:
This version is available at: 11583/3015316 since: 2026-09-09T12:23:16Z

Publisher:
IEEE

Published
DOI:10.1109/TII.2026.3716951

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2026 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

17 November 2026

Hierarchical Dynamic Hypergraph Learning for Ultra-Short-Term Wind Farm Cluster Power Forecasting with Input-Aware Meteorological Integration Method

Mao Yang, Yutong Huang, Bo Wang, Chuanyu Xu, Jingsi Chen, Mahmood Hosseini Imani, Tao Huang

***Abstract*—Accurate multi-step ultra-short-term wind power forecasting for wind farm clusters is essential for secure and flexible power system operation. However, existing forecasting methods often rely on pairwise spatial correlations or static graph structures, which limits their ability to represent multi-scale spatial dependencies and transient fluctuation patterns among wind farms. In addition, the temporal mismatch between future numerical weather prediction (NWP) information and different forecasting horizons may further degrade multi-step forecasting accuracy. To address these challenges, this study proposes a hierarchical dynamic hypergraph forecasting framework with input-aware meteorological modeling for wind farm cluster power forecasting. Specifically, four types of hyperedges are constructed to describe local correlations, regional coordinated structures, global similarity patterns, and transient response relationships among wind farms. A multi-channel architecture is then designed to jointly extract historical power information, future NWP features, graph-based pairwise spatial representations, and hypergraph-based high-order spatial representations. Furthermore, an input-aware modeling strategy is developed to adaptively incorporate horizon-related meteorological information, thereby reducing temporal mismatch in multi-step forecasting. The proposed method is validated using real-world data from large-scale wind farm clusters in China. Experimental results averaged over multiple independent trials show that the proposed method achieves lower NRMSE and NMAE than representative transformer-based and graph-based benchmark models, with an average absolute NRMSE reduction of 1.5% over the benchmark models and a 13.1% relative NRMSE reduction compared with the strongest baseline under the tested setting. The results indicate that the proposed framework can improve multi-step forecasting accuracy and curve-tracking consistency.**[1]

***Index Terms*—wind power forecasting, multistep forecasting, graph convolution networks on hypergraphs, dynamic hypergraph generation mechanism, modeling method**

NOMENCLATURE

A. Abbreviations

AC	Accumulation Calculation
AGCNs	Adaptive Graph Convolution Networks
BiLSTM	Bidirectional Long Short-Term Memory
CWC	Coverage Width-based Criterion
DF	Direct Forecasting
DSTGCNs	Dynamic Spatio-Temporal Graph Convolutional Networks
DCD	Dynamic Cluster Division
ENDCC	Entropy Dynamic Correlation Coefficient
EGANs	Embedded Graph Attention Networks
FDCA	Frequency Domain Gain Channel Attention
GCNs	Graph Convolutional Networks
GDCNs	Graph Diffusion Convolution Networks
GNNs	Graph Neural Networks
HC	Holistic Calculation
HGCNs	Hypergraph Convolutional Networks
LIME	Local Interpretable Model-Agnostic Explanations
LG	Local and Global
LSTM	Long Short-Term Memory
NWP	Numerical Weather Prediction
MLP	Multi-Layer Perceptron
PCC	Pearson Correlation Coefficient
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval normalized averaged width
RF	Recursive Forecasting
STIG	Spatio-Temporal Information Gain
USTWPF	Ultra-Short-Term Wind Power Forecasting

B. Variables and Symbols

$A_{p,t}(i, j)$	Power-correlation adjacency matrix at time t
$A_{ws,t}(i, j)$	NWP-correlation adjacency matrix at time t
b	The current input batch
b_{att}	The corresponding bias term
B	Degree matrix of hyperedges
c	The index of the cluster center
d	Input feature dimension of each node
D	Degree matrix of graph nodes
$caps$	Installed capacity of the wind farm cluster
$center$	Cluster center
$C_{cluster}$	The learnable clustering center
f	Forecasting model in LIME analysis
f_{att}	Multi head attention mechanism
F_{power}	Features extracted from historical power channels
F_{nwp}	Features extracted from future NWP channels
F_{hyper}	Features extracted from hypergraph channels
$F_{input\text{-}aware}$	The aggregated feature representation
g	Surrogate model in LIME
h	Output feature dimension
H_{star}	Represents star-shaped hypergraph
H_{multi}	Represents multi-center coordinated hypergraph
$H_{cluster}$	Represents clustering hypergraph
$H_{dyattention}$	Represents dynamic-response hypergraph
i, j	Denote wind farms
k_{star}	The number of selected neighbors of star hypergraph

[1] This work was supported by the Smart Grid-National Science and Technology Major Project under Grant 2025ZD0805500. (Corresponding author: Yutong Huang.)

Y. Huang and M. Yang are with the Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin 132012, China (e-mail: 20102339@neepu.edu.cn, 1202300033@neepu.edu.cn);

B. Wang is with State Key Laboratory of Operation and Control of Renewable Energy & Storage Systems, China Electric Power Research Institute, Beijing 100912, China (e-mail: 2202200279@neepu.edu.cn);

C. Xu is with the State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: chuanyuxu@hust.edu.cn);

T. Huang, Jingsi Chen and M. H. Imani are with the Department of Energy of Politecnico di Torino, 10129 Torino, Italy. (e-mail: tao.huang@polito.it, jingsi.chen@polito.it, mahmood.hosseiniimani@polito.it).

k_{cluster}	The number of clusters
k_{dynamic}	The number of connected neighbors of dynamic-response hypergraph
K	Length of NWP wind speed sequence
l	Layer index
L	Length of historical wind power sequence
n_{center}	The number of cluster centers
N_t	Number of interval prediction samples
o	The attention scoring vector
$pcc_{\text{p},t}$, $pcc_{\text{ws},t}$	Pearson correlation coefficient of power or NWP at timestamp t
P_{predict}	Predicted wind power
P	Measured wind power
q_{th}	Quantile level
R	The variation range of the prediction interval
std	Standard deviation between two sequences
t	Current timestamp
$topk$	The selection of the k most similar neighbors
s_u	Denotes the compatibility score computed from the encoded features
S_{star}	The similarity matrix of star-shaped hyperedges
$S_{cluster}$	The clustering probability matrix
$S_{dynamic}$	The similarity matrix calculation of dynamic-response hyperedges
$X^{(l)}$	Node feature matrix at layer l
$X^{(l+1)}$	Updated node feature matrix at layer $l+1$
X_{power}	Denotes the historical power input features
X_{nwp}	Denotes the future NWP input features
U_t, L_t	Represent the upper boundary and lower boundary of the prediction interval, respectively
v_i, v_j	Denotes the wind power or wind speed sequence
$\bar{v}_i$, $\bar{v}_j$	Mean values of the power or NWP sequences of wind farms i and j
V	Number of wind farms
W	Learnable convolution kernel or weight matrix
W_{att}	The attention projection matrix
W_{encoder}	Encoder weight
WS	Wind speed sequence feature
Z_{hyper}	Represents the encoded mixed features
Z_{nwp}	Represents the features encoded from the m NWP inputs
Z_{p}	Denotes the encoded power features
σ , tanh, Softmax	Activation Function
α	The output weight matrix
μ	Represents the confidence level of the probability interval
η	The penalty parameter
π_x	Represents the locality of the local sample x to be explained

I. Introduction

A. Research Background & Motivation

According to the latest 2025 statistics from the Global Wind Energy Council, with 117 GW installed globally in 2024, wind energy is expanding into new regions and solidifying its role as a core pillar of the worldwide energy transition. China continues to lead in new installations globally [1]. As of December 31, 2024, China's cumulative grid-connected wind power capacity reached 521 GW generating 991,600 GWh of electricity [2]. In high wind-penetration modern power systems, the fluctuating, stochastic, and intermittent nature of wind energy introduces significant uncertainty [3]. With the ongoing expansion of wind power deployment, the challenges associated with maintaining forecasting accuracy have become increasingly prominent.

USTWPF refers to predicting wind power over a 0-4h horizon [4] and plays an important role in wind turbine control, dispatch scheduling, and power system operation [5]. With the increasing penetration of wind power, forecasting wind farm cluster output becomes more challenging because large-scale integration strengthens both stochastic fluctuations and spatial coupling among wind farms. Conventional forecasting models, which are often designed for individual wind farms or based on weak coupling assumptions, have limited ability to characterize the cooperative evolution of wind farm clusters. As a result, forecasting deviations may be amplified in ultra-short-term operation, where regulation time windows are limited and scheduling decisions require timely power information. Therefore, accurate USTWPF is essential for maintaining grid stability and supporting large-scale wind energy integration [6], [7].

Motivated by these challenges, this study focuses on USTWPF for large-scale wind farm clusters. The main objective is to address two key limitations of existing methods: insufficient modeling of dynamic wind interactions within the cluster and severe error accumulation in extended multi-step forecasting. Accordingly, this work seeks to improve the representation of evolving spatial dependencies and enhance forecasting robustness over longer horizons.

B. Related Work

Regarding USTWPF, considerable advanced research has been conducted [8], [9]. In practical power system dispatch, USTWPF for individual wind turbines [10] and single wind farms [11] is no longer sufficient. Therefore, wind farm cluster forecasting has become necessary for supporting grid connection and dispatch planning. Numerous forecasting methods targeting wind farm clusters [12], [13] have demonstrated their effectiveness, yet accurately characterizing the real-time evolution of dynamic wind processes remains challenging. In addition, methods incorporating dynamic wind farm cluster partitioning [14] can improve spatial representation to a certain extent, but their practical application is still affected by computational constraints and modeling complexity.

GNNs have attracted substantial attention in wind power forecasting because of their capability to model spatial dependencies among wind farms [15], [16]. Existing studies have constructed graph structures from different perspectives. For example, Ref. [17] incorporated spatiotemporal information gains into graph structures, Ref. [18] integrated causal information into GNNs, and Ref. [19] developed a graph convolutional model based on wind-direction-driven clustering. These studies demonstrate the effectiveness of graph-based models for wind farm cluster forecasting. However, most existing GNN-based methods mainly improve the construction of adjacency matrices. Since adjacency matrices essentially describe pairwise relationships between wind farms, they are limited in representing high-order interactions involving multiple wind farms. Moreover, the coordinated evolution of wind processes in large-scale wind farm clusters is not only reflected by local inter-station correlations, but also by regional coordination, global similarity, and transient response relationships. Therefore, relying only on pairwise graph structures is insufficient to fully characterize the multi-scale spatial dependencies of evolving wind fields.

To overcome the limitation of pairwise graph modeling, hypergraphs [20] provide a promising way to represent high-order relationships. Hypergraph construction based on clustering principles [21] enables models to learn region-

specific characteristics more effectively [22], achieving improved forecasting accuracy compared with conventional graph structures. Compared with standard graphs, hypergraphs can connect multiple wind farms within a single hyperedge, thereby capturing group-wise interactions and regional coupling patterns. Nevertheless, existing hypergraph-based forecasting methods still have limitations when applied to wind farm cluster forecasting. Most of them focus on constructing a single type of high-order relationship or a relatively static hypergraph structure, while the heterogeneous dependency patterns among wind farms are not explicitly distinguished. In particular, local neighboring interactions, regional coordinated variations, global similarity patterns, and transient response relationships have different physical and modeling meanings, but they are rarely represented in a unified hierarchical hypergraph structure. Therefore, this study constructs multiple types of hyperedges and integrates graph and hypergraph learning to jointly capture pairwise and high-order spatial dependencies in evolving wind processes.

Since USTWPF involves multi-step time-series forecasting, increasing the forecasting horizon inevitably introduces temporal lag and accumulated prediction errors. Existing studies have attempted to improve multi-step forecasting performance through model enhancement [23], [24] or by incorporating physical mechanisms into refined modeling frameworks. For example, Ref. [25] improved forecasting accuracy by constructing separate models for different scenarios and integrating multi-scenario outputs, while Ref. [26] utilized graph structures based on wind farm cluster aggregation to capture spatial dependencies. In addition, some studies have investigated multi-step forecasting from a methodological and interpretability perspective. Ref. [27] provided a systematic analysis of wind power forecasting methodologies and discussed suitable modeling strategies for multi-step prediction, while Ref. [28] introduced an interpretability-driven metric to evaluate feature learning capability across different horizons. Although these studies provide useful insights, most existing methods still focus on improving model structures, scenario representations, or interpretability evaluation. The temporal correspondence between future NWP inputs and different forecasting horizons has not been sufficiently considered. In multi-step forecasting, the meteorological information most relevant to each target horizon may differ, and direct concatenation or uniform fusion of NWP features may lead to temporal mismatch. Therefore, this work develops an input-aware meteorological modeling strategy to adaptively incorporate horizon-related NWP information and reduce accuracy degradation in extended forecasting tasks.

C. Discussion and Contributions

In summary, existing USTWPF research primarily focuses on model refinement and engineering optimization to improve forecasting accuracy. However, several fundamental challenges remain. Conventional methods struggle to effectively extract informative features from high dimensional and multi-source data in a manner that is aligned with each forecasting horizon, which leads to performance degradation in multi-step forecasting. In addition, when applied to large-scale wind farm clusters, existing approaches exhibit limited capability in capturing the real-time dynamic evolution of wind fields.

To address these challenges, this study proposes a hierarchical multi-hypergraph USTWPF framework for wind farm clusters. The novelty of this work does not lie in the isolated use of graph or hypergraph learning, but in the development of a physically motivated hierarchical dynamic hypergraph structure tailored to multi-step wind farm cluster forecasting. Specifically, four types of hyperedges are constructed to represent local correlations, regional structures, global coordination patterns, and transient response relationships, thereby capturing complementary spatial dependencies in evolving wind processes. A multi-channel architecture is designed to jointly extract historical power information, future NWP features, and hypergraph-based high-order spatial representations. These features are then fused to obtain a dynamic and comprehensive representation of wind farm cluster evolution. Furthermore, an input-aware modeling strategy is developed to adaptively incorporate the most relevant meteorological information for each forecasting horizon, thereby alleviating temporal misalignment between NWP inputs and prediction targets and reducing error accumulation in multi-step forecasting. The proposed approach is validated using real-world data from large-scale wind farm clusters in Jilin Province, China, and demonstrates improved forecasting performance.

The main contributions of this paper are summarized as follows:

(a) Four types of hyperedges, representing local correlations, regional feature divisions, global coordination patterns, and transient response relationships, are constructed to capture multi-scale physical dependencies within wind farm clusters. The fused hypergraph structure enables effective modeling of the dynamic evolution of the wind processes.

(b) A three-channel architecture is developed to process historical power data, future NWP inputs, and hypergraph representations. The extracted features are fused across multiple convolutional pathways to produce graph-level forecasts.

(c) An NWP input-aware multi-step modeling method is proposed, which incorporates the most relevant future atmospheric information at each horizon, effectively mitigating accuracy degradation as forecasting steps increase.

The remainder of this paper is organized as follows. Section II introduces the overall framework. Section III presents the experimental validation of the proposed method from several perspectives. Finally, Section IV concludes the paper.

II. PROPOSED METHOD

This section provides an overview of the proposed methodology. The overall architecture of the forecasting framework is illustrated in **Fig. 1**, which includes the dynamic graph structure of the forecasting model, the overall model architecture, and the input-aware modeling method.

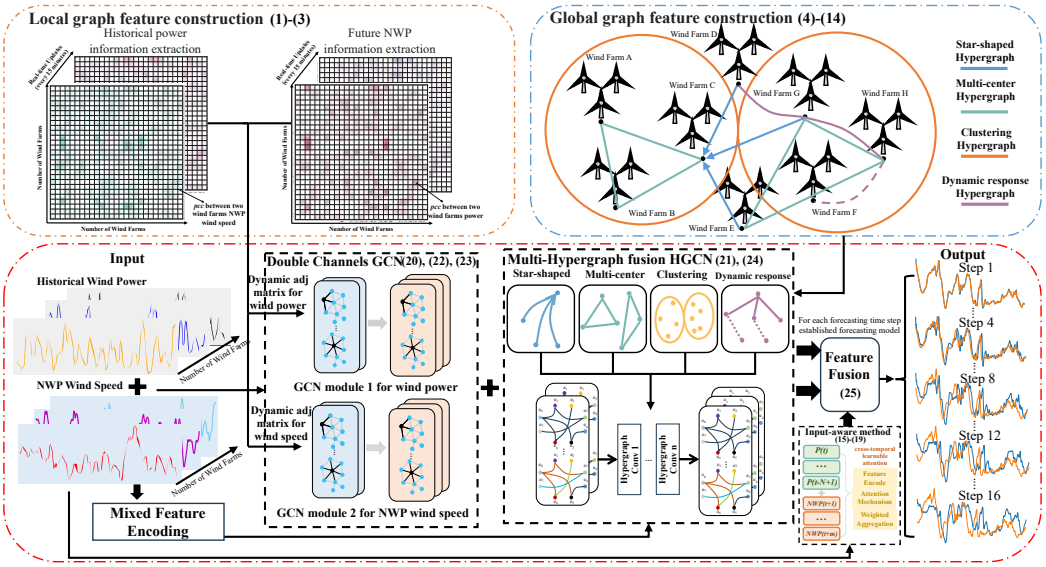


Fig. 1. Schematic of the proposed method

A. Overall Modeling Pipeline

The proposed framework consists of three major components: dynamic graph construction, hierarchical hypergraph learning, and input-aware meteorological integration. Given the historical power sequence and the future NWP sequence, two dynamic graph structures are first constructed based on power correlation and NWP correlation, respectively. These two graphs are used in two GCN channels to extract pairwise spatial dependencies from historical power and meteorological inputs. Meanwhile, the historical power and NWP features are jointly encoded to generate node embeddings, based on which four types of hyperedges are constructed to describe local, regional, global, and transient high-order relationships. The corresponding incidence matrices are integrated into a hierarchical hypergraph and used by the HGCN channel for high-order feature propagation. In parallel, the input-aware meteorological integration module adaptively aggregates the NWP features that are most relevant to each forecasting horizon. Finally, the GCN features, HGCN features, and input-aware meteorological features are fused and fed into the forecasting layer to directly predict the total power of the wind farm cluster.

B. Dynamic Graph Structure for GCN

To effectively extract historical wind power information and NWP features for USTWPF, dynamic graph structures are constructed for the GCN channel based on both historical and future inputs. Such graph structures enable the model to better capture the spatial dependencies among wind farms. Specifically, two graph structures are generated by computing the PCC every 15 minutes, and the corresponding formulations are given in (1) and (2).

$$A_{\mathrm{p},t}(i,j)=\begin{cases} e^{-\frac{pcc_{\mathrm{p},t}(i,j)}{\mathrm{std}^2}} & if\ pcc_{\mathrm{p},t}(i,j)\ge 0.75 \\ 0 & if\ pcc_{\mathrm{p},t}(i,j)<0.75 \end{cases} \tag{1}$$

$$A_{\mathrm{ws},t}(i,j)=\begin{cases} e^{-\frac{pcc_{\mathrm{ws},t}(i,j)}{\mathrm{std}^2}} & if\ pcc_{\mathrm{ws},t}(i,j)\ge 0.75 \\ 0 & if\ pcc_{\mathrm{ws},t}(i,j)<0.75 \end{cases} \tag{2}$$

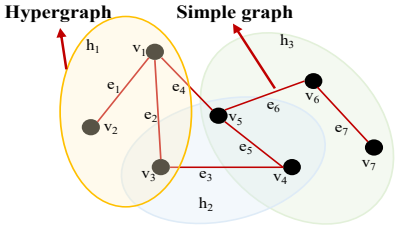


Fig. 2. Comparison of simple graph and hypergraph structure

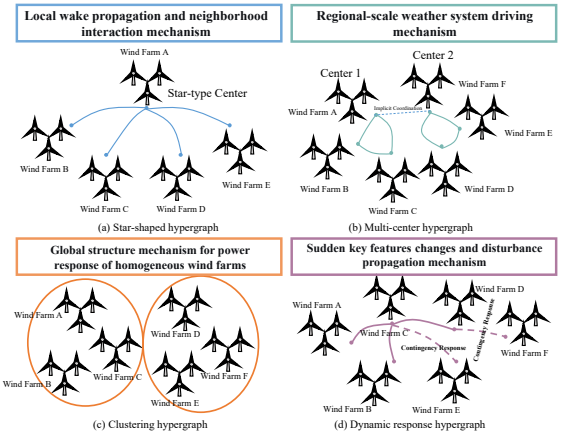


Fig. 3. Four schematic diagrams of hyperedge construction

$$pcc_t(i,j)=\frac{\sum_{t=1}^{n}\left[\left(v_{i,t}-\overline{v}_i\right)\left(v_{j,t}-\overline{v}_j\right)\right]}{\sqrt{\sum_{t=1}^{n}\left(v_{i,t}-\overline{v}\right)^2\sum_{t=1}^{n}\left(v_{j,t}-\overline{v}_j\right)^2}} \tag{3}$$

where t denotes the current timestamp, i and j denote two wind farms, $A_{\mathrm{ws},t}(i,j)$ and $A_{\mathrm{p},t}(i,j)$ are the adjacency matrices of the NWP wind-speed channel and the power channel, respectively, in the dual-channel GCN at timestamp t, *std* is the standard deviation of the corresponding sequence, $pcc_{\mathrm{p},t}$ and $pcc_{\mathrm{ws},t}$ are the correlation coefficients of power and wind speed, respectively, v_i and v_j denote the power or NWP value of i and j wind farms. $\overline{v}_i$ and $\overline{v}_j$ denote their corresponding mean values. (A fixed threshold of 0.75 is adopted here to preserve graph sparsity and structural stability during dynamic updating.)

C. Hypergraph Structure

Fig. 2 shows the difference between a hypergraph and a simple graph, where *h* denotes a hyperedge, *e* represents a simple graph edge, and *v* denotes a vertex. Unlike simple graphs, which can model only pairwise relationships, hypergraphs extend graph structures into high-order representations, allowing a more precise characterization of multivariate spatial correlations among wind farms. By generalizing vertex connectivity, hypergraphs encode richer relational information and capture high-order interactions within the cluster, thereby enabling higher-dimensional association modeling without loss of vertex information.

To better capture the correlations and interdependencies among wind farms within a cluster, a coordinated modeling strategy is adopted to characterize local, regional, global, and transient relationships. Based on this strategy, four types of hyperedges are constructed, namely star-shaped hyperedges, multi-center coordinated hyperedges, clustering hyperedges, and dynamic-response hyperedges, as shown in (4). Among them, the dynamic-response hyperedges are designed to capture key feature variations and abrupt fluctuations at each forecasting step.

$$H = \left[H_{\text{star}} \,\|\, H_{\text{multi}} \,\|\, H_{\text{cluster}} \,\|\, H_{\text{dyattention}} \right] \tag{4}$$

where H_{star} is a star-shaped hyperedge, H_{multi} is a multi-center coordinated hyperedge, H_{cluster} is a cluster-type hyperedge, $H_{\text{dyattention}}$ is a dynamic response hyperedge based on attention selection.

The structure diagram of each type of hyperedges is shown in **Fig. 3.** Detailed information for each hyperedge is presented below.

(a) *Star-shaped hypergraph*: This structure represents a topology in which a group of surrounding wind farms is connected through a single central node. Specifically, one wind farm is selected as the central node and linked to its most similar neighboring wind farms, thereby forming a star-like structure. This configuration reflects the coupling relationships among geographically proximate wind farms and characterizes the directional propagation of wake effects across sites. By capturing local fluctuation patterns and synchronizing nearby variations, the star-shaped hyperedge effectively extracts local spatial dependencies. In this way, it preserves the dominant local correlations arising from the physical continuity among wind farms. The construction process is given in (5) – (7). First, the mixed input features are encoded. Then, for each central node, the top-K most similar neighboring nodes are selected and connected to form the star-shaped hyperedge.

$$Z_{\text{hyper}} = \sigma\left(\left[X_{\text{power}} \,\|\, X_{\text{nwp}} \right] W_{\text{enconder}} \right) \tag{5}$$

$$S_{\text{star}} = Z_{\text{hyper}}^{(b)} \left(Z_{\text{hyper}}^{(b)} \right)^{\text{T}} \tag{6}$$

$$H_{\text{star}} = \{i\} \bigcup topk_j \left(S_{\text{star}}(i,j), k_{\text{star}} \right) \tag{7}$$

where Z_{hyper} represents the encoded mixed features, X_{power} and X_{nwp} denote the historical power and future NWP input features respectively, W_{encoder} is the encoder weight, S_{star} is the similarity matrix of star-shaped hyper-edges, b is the current input batch, k_{star} is the number of selected neighbors, *topk* is the selection of the k most similar neighbors, σ is the activation function.

(b) *Multi-center coordinated hypergraph*: This structure is constructed by selecting adjacent wind farms and forming hyperedges based on multiple regional centers. Each central wind farm represents a subregion and is connected to all wind farms within that region. Multiple central nodes coexist in parallel, resulting in a multi-center structure. This hypergraph captures the spatiotemporal dependencies among wind farms under different regional wind conditions and reflects the coordinated variations among groups of wind farms driven by similar atmospheric dynamics. It addresses regional power inconsistencies arising from localized weather differences across subregions, thereby revealing spatial gradient characteristics. Multi-center hyperedges thus effectively capture regionally coordinated variations and characterize the spatiotemporal propagation patterns of power changes. The computational procedure is defined in equations (8) and (9). First, the K-means++ algorithm is used to identify regional cluster centers. Subsequently, all wind farms within each region are connected to their corresponding center, forming the multi-center coordinated hyperedges.

$$center = \text{kmean}^{++} \left(Z_{\text{hyper}}^{(b)}, n_{center} \right) \tag{8}$$

$$H_{\text{multi}} = \left\{ j \mid c = arg\,\min_{c} \left\| Z_j - center_c \right\|_2 \right\} \tag{9}$$

where *center* is the cluster center, n_{center} is the number of cluster centers, c is the index of the cluster center.

(c) *Clustering hypergraph*: This structure is formed by connecting all wind farms within regions that exhibit similar power fluctuation patterns. Unlike the multi-center coordinated hyperedges, this structure does not rely on a designated central node. Instead, it provides a global representation of similarity-based connectivity across the entire wind farm cluster. The computation procedure is shown in (10) and (11): a clustering probability matrix is computed from the encoded features, based on this matrix, each node is assigned to the category with the highest probability, and nodes within the same category are grouped to construct the clustering hyperedges.

$$S_{\text{cluster}} = Z_{\text{hyper}}^{(b)} \left(C_{\text{cluster}} \right)^{\text{T}} \tag{10}$$

$$H_{\text{cluster}} = topk_j \left(\text{softmax} \left(S_{\text{cluster}}(j,c) \right), k_{\text{cluster}} \right) \tag{11}$$

where S_{cluster} is the clustering probability matrix, C_{cluster} is the learnable clustering center, Softmax is the activation function to compute the probability, k_{cluster} is the number of clusters.

(d) *Dynamic-response hypergraph*: This structure is generated using a multi-head attention mechanism. Each attention-derived node pair forms an independent hyperedge, which is updated dynamically according to instantaneous feature variations. This structure is designed to capture rapidly changing key features and infer the strength of relationships among wind farms based on current input conditions. The computational process is defined in (12) - (14). Multi-head attention is first applied to the encoded feature matrix of the current input batch to obtain an attention weight matrix; then, for each node, the top-*K* most relevant neighbors are selected to construct the dynamic hyperedges. In this study, the dynamic update of the hyperedges is achieved by recalculating the

TABLE I
PHYSICAL INTERPRETATION AND MODELING ROLE OF DIFFERENT HYPEREDGE TYPES

Edge type	Construction criterion	Physical meaning	Modeling role
Star-shaped	top-k similar neighbors around each center node	local neighboring interaction	captures local fluctuation propagation
Multi-center	K-means++ regional centers	regional coordinated variation	captures subregional wind regime differences
Clustering	learnable clustering probability	global similarity pattern	captures global group-wise dependency
Dynamic-response	multi-head attention top-k neighbors	transient response under abrupt variation	captures short-term dynamic changes

attention-based node relationships from the encoded current input features at each forecasting timestamp.

$$S_{\text{dynamic}} = \begin{cases} Z_{\text{hyper}}^{(i)} \left(Z_{\text{hyper}}^{(j)}\right)^{\text{T}} & i \neq j \\ 0 & i = j \end{cases} \tag{12}$$

$$\alpha_{(i,j)} = f_{\text{att}}\left(S_{\text{dynamic},ij}, W_{\text{att}}\right) \tag{13}$$

$$H_{\text{dyattention}} = \{i\} \bigcup topk_j\left(\alpha_{(i,j)}, k_{\text{dynamic}}\right) \tag{14}$$

where S_{dynamic} is the similarity matrix calculation (excluding self-loops), f_{att} is the multi-head attention mechanism used in the model for this study, W_{att} is the weight of the attention mechanism, α is the output weight matrix, and k_{dynamic} is the number of connected neighbors.

To clarify the physical meaning of the proposed hypergraph construction, the four hyperedge types are not introduced as independent heuristic structures. Instead, they are designed to represent complementary dependency patterns in wind farm clusters. Their construction criteria, physical interpretations, and modeling roles are summarized in Table I. Therefore, the proposed hierarchical hypergraph construction provides a multi-scale representation of wind farm cluster interactions, where local, regional, global, and transient dependencies are modeled in a unified structure. They are dynamically updated every 15 minutes based on incoming inputs, enabling the model to more accurately characterize the complex and evolving dynamics of wind processes within large wind farm clusters. In this framework, the hierarchical structure is formed by integrating local, regional, global, and transient dependency patterns within a unified hypergraph representation. These dependency patterns correspond to different spatial interaction ranges and response characteristics of wind farm clusters.

D. Input Modeling Method

To better align meteorological forecasts with each forecasting horizon and alleviate lag and error accumulation in multi-step forecasting, an enhanced NWP input modeling strategy is developed. This strategy is motivated by previous studies on step-by-step forecasting and interpretability-driven analysis in wind power forecasting [27], [28]. Specifically, for each target forecasting step, the model takes as input not only the recent historical power sequence but also the most relevant

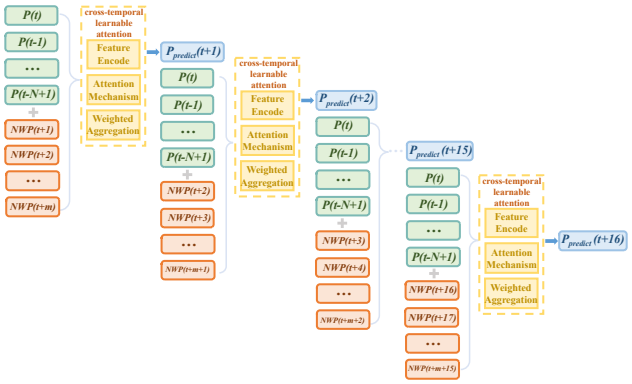


Fig. 4. The diagram of input-aware modeling method

sequence of future NWP preceding the target horizon. A cross-temporal attention mechanism is then introduced to compute learnable weights over these time points, so that the NWP representations can be aggregated adaptively. The aggregated features are finally fused with those extracted by the proposed model for power forecasting, as formulated in (15) to (19).

The underlying rationale is threefold:

(a) A step-wise modeling scheme is adopted to avoid recursive error propagation inherent in autoregressive strategies.

(b) A sequence of NWP features leading the target time step by m intervals is introduced to mitigate phase mismatch commonly observed in multi-step forecasting due to machine learning model characteristics.

(c) By learning attention weights over the m NWP inputs, the model adaptively emphasizes the most informative forecast feature, thereby reducing error amplification across horizons and improving both robustness and interpretability.

A schematic of the input-aware modeling method is provided in **Fig. 4**, where P denotes historical wind power, P_{predict} is the predicted power, t is the current time step, NWP is the future input features, N is the number of historical inputs, and m is the number of future NWP inputs.

$$Z_{\text{p}} = f_{\text{p}}\left(X_{\text{power},t}\right) \tag{15}$$

$$Z_{\text{nwp},u} = f_{\text{nwp}}\left(X_{\text{nwp},t+\text{m}-N+m}\right), \quad u = 1,2,\cdots,m \tag{16}$$

$$s_u = o^{\text{T}} \tanh\left(W_{\text{att}}\left[Z_{\text{p}} \middle\| Z_{\text{nwp},u}\right] + b_{\text{att}}\right) \tag{17}$$

$$\alpha_u = \frac{exp\left(s_u\right)}{\sum_{k=1}^{m} exp\left(s_k\right)} \tag{18}$$

$$F_{\text{input-aware}} = \sum_{u=1}^{m} \alpha_u Z_{\text{nwp},u} \tag{19}$$

where Z_{p} denotes the encoded power features, Z_{nwp} represents the features encoded from the m NWP inputs, and $F_{\text{input-aware}}$ is the aggregated feature representation, s_u denotes the compatibility score computed from the encoded features, o is the attention scoring vector, W_{att} is the attention projection matrix, b_{att} is the corresponding bias term, and *tanh* is the activation function.

The attention weights in the input-aware module are not manually specified heuristic coefficients. They are learned from the compatibility between the encoded historical power features and future NWP features, so that the model can adaptively emphasize the NWP inputs that are more relevant to each forecasting horizon.

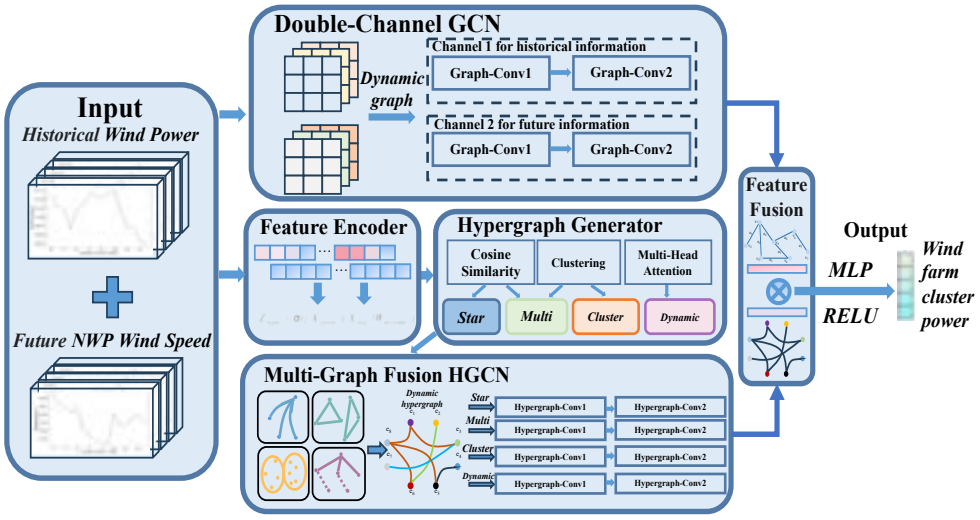


Fig. 5. The overall architecture of the forecasting model

E. Forecasting Model

To more effectively extract the complex dynamic spatiotemporal dependencies among wind farms within a cluster, we construct the forecasting model using both GCN and HGCN architectures. Their mathematical formulations are given in (20) and (21):

$$f_{\mathrm{GCN}}\left(X^{(l)},A\right)=\sigma\left(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{-\frac{1}{2}}X^{(l)}W^{(l)}\right) \tag{20}$$

$$f_{\mathrm{HGCN}}\left(X^{(l)},H\right)=\sigma\left(\hat{D}^{-\frac{1}{2}}HB^{-1}\hat{D}^{-\frac{1}{2}}X^{(l)}W^{(l)}\right) \tag{21}$$

where, $A\in \mathrm{R}^{N*N}$ is the adjacency matrix formed by the correlation information between wind farms, N is the number of wind farms. $\widehat{D}\in \mathrm{R}^{N*N}$ is the degree matrix for $\hat{A}$, $X^{(l)}\in \mathrm{R}^{N*d}$ is the feature of layer l, $X^{(l+1)}\in \mathrm{R}^{N*h}$ is the updated feature of layer l+1, d and h form the feature dimension of each node, representing the power or NWP data of a particular wind farm. $W\in \mathrm{R}^{d*h}$ is the parameter of the learnable convolutional kernel. H is the hypergraph. B is the degree matrix of the edge.

Building on these two models, this study enhances local–global collaborative modeling through the following components: **(a) Dual-GCN channels:** Two dynamically updated graph structures, introduced in Section B, are used to extract two complementary sources of information: historical wind power features through the power correlation graph and future NWP features through the NWP correlation graph. **(b) HGCN channel:** Four types of hypergraphs generated via multi-head attention, nearest-neighbor relationships, and clustering algorithms are used to extract high-order hypergraph features. **(c) Feature fusion module:** Features extracted by the GCN and HGCN channels are integrated and subsequently processed through fully connected layers to generate the final wind power forecasts.

After these enhancements, the improved mathematical formulations of the GCN and HGCN modules are provided in (22) – (25):

$$F_{\mathrm{power}}=\sigma\left(\sum_{x=1}^{L}\left(\sum_{y=1}^{V}\hat{D}^{-\frac{1}{2}}(p,p)\hat{A}_{p}(p,y)_{t-x+1}\hat{D}^{-\frac{1}{2}}(p,p)P_{t-x+1}(y)\right)W(x,q)\right) \tag{22}$$

$$F_{\mathrm{nwp}}=\sigma\left(\sum_{x=1}^{K}\left(\sum_{y=1}^{V}\hat{D}^{-\frac{1}{2}}(p,p)\hat{A}_{\mathrm{ws}}(p,y)_{t+x}\hat{D}^{-\frac{1}{2}}(y,y)WS_{t+x}(y)\right)W(x,q)\right) \tag{23}$$

$$F_{\mathrm{hyper}}=\sigma\left(\sum_{x=1}^{K+L}\left(\sum_{y=1}^{V}\hat{D}^{-\frac{1}{2}}(p,p)H(p,y)B^{-1}H^{T}(p,y)\hat{D}^{-\frac{1}{2}}(y,y)Z_{\mathrm{hyper}}(y)\right)W(x,q)\right) \tag{24}$$

$$P_{\mathrm{predict}}=\mathrm{MLP}\left(Concat\left(F_{\mathrm{power}},F_{\mathrm{nwp}},F_{\mathrm{hyper}},F_{\mathrm{input\text{-}aware}}\right)\right) \tag{25}$$

where, F_{power} and F_{nwp} are the GCN modules proposing power or NWP features, respectively, and the extracted time series lengths are L (wind power series) and K (NWP wind speed series), V is the number of wind farms, $\hat{A}_{\mathrm{p}}$ and $\hat{A}_{\mathrm{WS}}$ are the adjacency matrices updated in real time for power and wind speed, respectively, P and WS are the historical wind speeds or power series, and t is the current timestamp, p = 1, 2, 3, ⋯, d and q =1, 2, 3, ⋯, h. The spatial characteristics are adjusted by $\widehat{A}$, and W is used to extract the temporal characteristics. F_{hyper} is the hypergraph feature extraction module. MLP denotes the fully connected layer. The overall architecture of the model is shown in **Fig. 5.** For clarity, the main tensor dimensions are specified as follows. For a batch size of B, the dimensions of the historical power input X_{power} and future NWP input X_{NWP} are $B \times V \times L \times d_{\mathrm{p}}$ and $B \times V \times K \times d_{\mathrm{nwp}}$, respectively. Here, V is the number of wind farms, L is the historical power sequence length, K is the future NWP sequence length, and d_{p} and d_{nwp} are the corresponding input feature dimensions. The extracted features F_{power}, F_{nwp}, F_{hyper}, and $F_{\mathrm{input\text{-}aware}}$ are then projected to the same feature dimension and concatenated before the final MLP forecasting layer.

For model implementation, the four hyperedge construction strategies generate four incidence matrices. These incidence matrices are concatenated along the hyperedge dimension to form the integrated hierarchical hypergraph. The HGCN channel performs message passing on this integrated hypergraph to obtain the high-order spatial representation. Therefore, the hypergraph structure directly determines the high-order feature aggregation path among wind farms and is not used merely as an auxiliary conceptual representation. The hypergraph incidence matrices are dynamically regenerated for each forecasting timestamp or input batch according to the encoded input features, rather than being fixed throughout training or optimized as independent trainable parameters. The

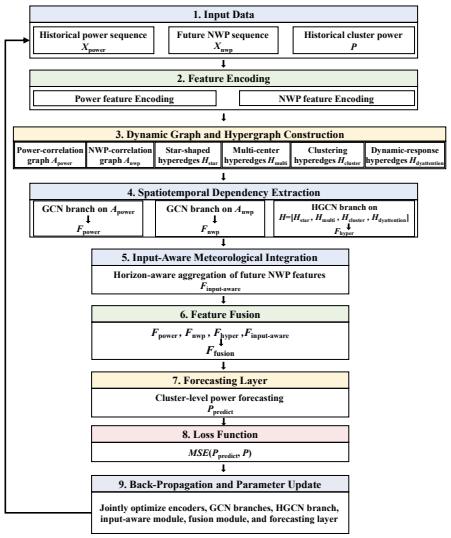
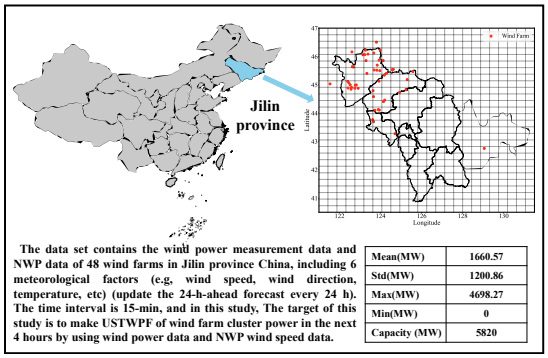


Fig. 6. Training architecture of the proposed framework

input-aware meteorological representation $F_{\text{input-aware}}$ is fused with F_{power}, F_{nwp}, and F_{hyper} in the final forecasting module, rather than being used to directly modify the hypergraph construction process.

The proposed forecasting framework is trained in an end-to-end manner. The GCN channels, HGCN channel, input-aware meteorological integration module, feature fusion module, and final forecasting layer are jointly optimized using the final cluster-level forecasting loss. No separate pre-training is performed for the graph or hypergraph branches. During each training iteration, the graph adjacency matrices and hypergraph incidence matrices are dynamically generated from the current input features and then used for graph-based and hypergraph-based feature propagation. The extracted representations are fused to obtain the final cluster-level prediction. The forecasting loss is calculated only between the predicted and measured cluster power. No additional meteorological feature regularization term or manually selected loss-weighting coefficient is introduced in the training objective. Gradients are back-propagated through all trainable neural modules, including the feature encoders, graph or hypergraph convolution layers, attention-based meteorological integration module, fusion module, and forecasting layer. Discrete structure selection operations, such as thresholding, top-k neighbor selection, and cluster assignment, are used only to determine connectivity patterns and do not introduce independent trainable parameters. The neural components that generate the corresponding feature embeddings or attention scores are optimized jointly with the forecasting objective. **Fig. 6** illustrates the training architecture of the proposed framework.

The node representations used for hyperedge generation are jointly encoded from the historical power and NWP variables. Accordingly, the hypergraph incidence matrices are reconstructed according to the power–meteorological states of the current input sample, enabling sample-dependent dynamic topology updating. It should be noted that the input-aware meteorological integration module does not directly modify the hyperedge connections or incidence matrices. Instead, it generates forecasting-horizon-specific meteorological features, which are fused with the hypergraph representations before the final forecasting layer.



The data set contains the wind power measurement data and NWP data of 48 wind farms in Jilin province China, including 6 meteorological factors (e.g, wind speed, wind direction, temperature, etc) (update the 24-h-ahead forecast every 24 h). The time interval is 15-min, and in this study, The target of this study is to make USTWPF of wind farm cluster power in the next 4 hours by using wind power data and NWP wind speed data.

Mean(MW)	1660.57
Std(MW)	1200.86
Max(MW)	4698.27
Min(MW)	0
Capacity (MW)	5820

Fig. 7. Geographical distribution and summary statistics of the experimental dataset

TABLE II
HYPERPARAMETER CONFIGURATION

Module and Parameter Setting
GCN layer: 4 (Hidden layer: 32)
HGCN layer: 2 (Hidden layer: 32)
Number of attention heads: 4
Active functions of fully connected network: ReLU
Optimizer: Adam
Batch size: 32
Number of Epochs: 100
Learning Rate: 1×10^{-3}
Historical power input length L: 16
Future NWP sequence length K: 16
NWP lead-window size m: 3

III. CASE STUDY AND ANALYSIS

A. Datasets and Implementation Details

To evaluate the effectiveness of the proposed method, we use datasets from 48 representative wind farms in Jilin Province, China, whose geographical distribution and data description are shown in **Fig. 7**. The wind farm cluster has a total installed capacity of 5820 MW. The study uses 181 days of data from each wind farm, sampled every 15 min, yielding 17376 sampling points per site. Data from January to May are used for training, while data from June serve as the testing set. Sliding-window samples were generated separately after the chronological split to avoid boundary leakage. The model inputs include historical power measurements and NWP-derived wind speed forecasts. The selected 48 wind farms capture both power fluctuation patterns and spatial characteristics across multiple subregions, thereby enabling a comprehensive assessment of the proposed method's forecasting capability. The dataset has a consistent temporal resolution of 15 min, and all wind farms are synchronized at the same sampling timestamps. The data cover January to June; thus, the training set contains winter-to-spring operating conditions, while the testing set corresponds to early-summer operating conditions. The operational records contain no missing values, and the missing data ratio is therefore 0% for both power measurements and NWP inputs.

The experiments were conducted on an NVIDIA L20 GPU. For evaluation, the average performance of each method was calculated over 10 independent trials. Both the number of wind farm clusters and the number of the most relevant nodes were determined using the elbow method, and both were set to 7. The main hyperparameter settings of the proposed model are summarized in **Table II**.

To improve the general applicability of the case study, no extensive manual hyperparameter search was conducted. For the comparative models, the same training and testing splits, preprocessing procedures, input information, and evaluation metrics were used. The main training settings, including optimizer, batch size, learning rate, and training epochs, were kept consistent with Table II whenever applicable, while model-specific structural parameters were set according to the original references or standard recommended configurations. All trainable parameters were initialized using Xavier uniform initialization for linear and graph convolutional layers, while bias terms were initialized to zero. The model was trained using the Adam optimizer with a batch size of 32, an initial learning rate of 1×10^{-3}, and 100 training epochs. No additional learning-rate scheduling strategy was used in the main experiments. The training process was terminated after 100 epochs, and the model parameters from the final epoch were used for testing because the training loss became stable in the later training stage. In addition, the wind power data were normalized by installed capacity, while the NWP wind speed data were normalized using the maximum wind speed in the training set. All selected data were real operational measurements provided by the State Grid Corporation of China and contained no missing values. Furthermore, negative values and values exceeding the installed capacity were corrected to ensure data quality and experimental reliability.

This study evaluates the effectiveness of the proposed method with the following three predictive indicators: the NRMSE, NMAE, correlation coefficient (R). NRMSE and NMAE quantify the forecasting error after normalization by the installed capacity of the wind farm cluster, thereby improving the comparability across different datasets. Specifically, NRMSE emphasizes larger forecasting deviations, while NMAE reflects the average absolute error level. The R is used to measure the temporal consistency between the predicted and measured power series. Therefore, a model with lower NRMSE and NMAE and higher R is considered to have better forecasting performance. The expressions of these indices are given in (26)-(28):

$$NRMSE=\sqrt{\frac{\sum_{i=1}^{n}(P_i-P_{\text{predict},i})^2}{caps^2}\times\frac{1}{n}} \tag{26}$$

$$NMAE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{P_i-P_{\text{predict},i}}{caps}\right| \tag{27}$$

$$R=\frac{\sum_{i=1}^{n}[(P_i-\hat{P}_i)(P_{\text{predict},i}-\hat{P}_{\text{predict},i})]}{\sqrt{\sum_{i=1}^{n}(P_i-\hat{P}_i)^2\sum_{i=1}^{n}(P_{\text{predict},i}-\hat{P}_{\text{predict},i})^2}} \tag{28}$$

where P_i is the actual value, $P_{\text{predict},i}$ is the predicted value, $\hat{P}_i$ is the average of the actual value for the corresponding period, $\hat{P}_{\text{predict},i}$ is the average of the predicted value for the corresponding period, *caps* is the installed capacity of the corresponding wind farm cluster, and n is the sample size for the corresponding predicted period.

In order to verify the effectiveness of the probabilistic prediction method, it is evaluated based on the PICP, the PINAW and the CWC. The PICP index reflects the ratio of the measured power within the prediction interval and is used to evaluate the reliability of the probability interval:

$$PICP=\frac{1}{N}\sum_{i=1}^{N_t}k_i^{(a)} \tag{29}$$

where N_t is the number of prediction samples; α is the given confidence; k_i is used to indicate whether the i^{th} predicted value is within the probability interval, and if the prediction power is within its relative interval, then k_i=1, otherwise k_i=0. PINAW reflects the sharpness of the prediction interval. The narrow interval has more information and practical value than the wide interval, and the formula is as follows:

$$PINAW=\frac{1}{N_tR}\sum_{i=1}^{N_t}\left[U_t(x_i)-L_t(x_i)\right] \tag{30}$$

where, R is the variation range of the prediction interval, which is used to normalize the average bandwidth. U_t and L_t represent the upper boundary and lower boundary of the prediction interval, respectively.

In order to comprehensively consider PICP and PINAW, CWC is introduced, and its formula is defined as follows:

$$CWC=PINAW(1+\gamma e^{-\eta(PICP-\mu)}) \tag{31}$$

$$\gamma=\begin{cases}0 & PICP\geq\mu\\ 1 & PICP<\mu\end{cases} \tag{32}$$

where, μ represents the confidence level of the probability interval, and η is the penalty parameter. When η is smaller, the penalty that the PICP does not reach the confidence level is smaller, that is, the lower the importance. This paper is set to 0.1. If PICP reaches the confidence level, then CWC=PINAW.

To provide a clear link between the experimental protocol and the reported figures and tables, eight groups of experiments were designed under the unified settings described above.

(a) **Sensitivity analysis**: The effects of the number of hyperedge neighbors k, the number of clusters, the NWP lead-window size m, the graph relationship measure, and the graph construction threshold on forecasting performance are investigated.

(b) **Model comparison**: Cluster-level forecasting paradigms and representative graph-based forecasting models are compared under the same chronological data division, input information, training settings, and 16-step ultra-short-term forecasting task.

(c) **Ablation on graph construction**: Controlled graph and hypergraph variants are constructed by progressively introducing or combining power-correlation graphs, NWP-correlation graphs, and local, regional, global, and dynamic-response hyperedges. Their effects are evaluated using forecasting metrics and error distributions.

(d) **Comparison of input modeling strategies**: Recursive forecasting, direct forecasting, NWP-enhanced variants, horizon-specific modeling, and the proposed input-aware strategy are compared at forecasting steps 1, 4, 8, 12, and 16.

(e) **Model interpretability analysis**: LIME is used to analyze feature contributions at the 16th forecasting step for different input modeling and graph construction strategies.

(f) **Probabilistic forecasting extension**: The proposed framework is extended using quantile loss and compared with representative parametric and non-parametric probabilistic forecasting methods at confidence levels of 85%, 90%, and

TABLE III
COMPARISON OF DIFFERENT PARAMETERS

Parameters		NRMSE	NMAE	R
k	3	0.0609	0.0506	0.9283
	5	0.0589	0.0493	0.9321
	7	0.0580	0.0484	0.9345
	9	0.0582	0.0489	0.9332
	11	0.0597	0.0498	0.9306
Clusters	3	0.0604	0.0501	0.9294
	5	0.0589	0.0490	0.9327
	7	0.0580	0.0484	0.9345
	9	0.0587	0.0489	0.9330
m	2	0.0590	0.0491	0.9318
	3	0.0580	0.0484	0.9345
	4	0.0584	0.0487	0.9336
	5	0.0592	0.0494	0.9312

TABLE IV
SENSITIVITY ANALYSIS OF RELATIONSHIP MEASURES FOR GRAPH CONSTRUCTION

Relationship measure	NRMSE	NMAE	R
PCC	0.0580	0.0484	0.9345
Spearman	0.0586	0.0480	0.9324
Mutual information	0.0577	0.0482	0.9336
Kendall	0.0593	0.0490	0.9298
Granger causality	0.0580	0.0485	0.9339
Geographic distance	0.0602	0.0511	0.9254

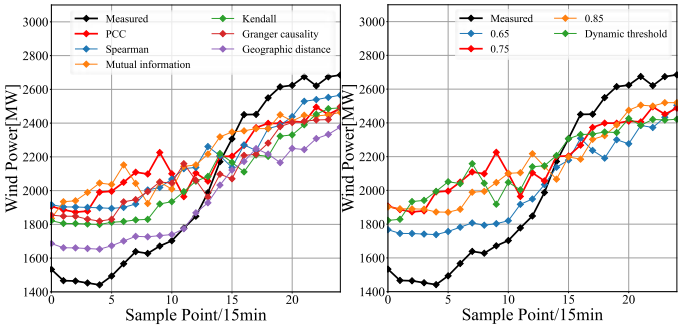


Fig. 8. The forecasts by different graph construction settings: (a) different relationship measures; (b) different threshold settings.

95%, using PICP, PINAW, and CWC.

(g) **Generalization experiments**: The framework is further evaluated on the Inner Mongolia and Gansu datasets under different wind farm cluster scales, NWP quality levels, temporal resolutions, and forecasting horizons.

(h) **Computational and limitation analysis**: Model parameters, FLOPs, offline training time, and online running time are compared across different cluster scales, followed by a discussion of computational cost, parameter dependence, and practical deployment limitations.

Unless otherwise stated, all experiments use the same chronological data division, preprocessing procedures, parameter configurations, and evaluation criteria described in this subsection. The reported results are averaged over 10 independent trials.

B. Result Analysis

***(a) Sensitivity analysis of threshold definitions*:** To further evaluate the robustness of the proposed method, a sensitivity analysis is conducted on several critical hyperparameters, including the number of hyperedge neighbors (k), the number of clusters, and the lead window size (m) of NWP inputs. Specifically, k is varied from 3 to 11, the number of clusters is varied from 3 to 9, and m is varied from 2 to 5. The corresponding forecasting performance in terms of NRMSE is illustrated in **Table III**.

TABLE V
SENSITIVITY ANALYSIS OF THRESHOLD OF GRAPH CONSTRUCTION SETTINGS

Threshold	NRMSE	NMAE	R
0.65	0.0601	0.0576	0.9210
0.75	0.0580	0.0484	0.9345
0.85	0.0582	0.0501	0.9322
Dynamic threshold	0.0578	0.0475	0.9388

The results show that the proposed method maintains stable performance across a wide range of parameter settings. Although optimal performance is achieved around k=7 and cluster number=7, the performance degradation remains marginal when deviating from these values. Similarly, the model demonstrates low sensitivity to the selection of m, indicating that the input-aware mechanism is robust to variations in temporal alignment.

As shown in **Table IV**, **Table V** and **Fig. 8**, the forecasting results remain relatively stable when different relationship measures and threshold settings are used. Although minor differences can be observed among PCC, Spearman correlation, mutual information, Kendall correlation, Granger causality, geographic distance, and different threshold configurations, the overall variation in NRMSE, NMAE, and R is limited. This indicates that the proposed framework is not highly dependent on a deliberately selected graph construction metric or threshold. Therefore, the main performance improvement cannot be attributed to the specific use of PCC or the threshold of 0.75, but mainly results from the hierarchical dynamic hypergraph learning and input-aware meteorological integration mechanisms.

These findings confirm that the proposed framework does not rely on fine-tuned hyperparameters and exhibits strong robustness in practical applications.

***(b) Model comparison*:** First, to verify the effectiveness of the proposed overall framework, the experiment compares the following forecasting models:

- **Transformer** [30] **+HC**: A holistic cluster-level forecasting baseline using Transformer [30], directly forecasting the total cluster power.
- **Transformer+AC**: An accumulation-based cluster-level forecasting baseline using Transformer [30], summing individual wind farm forecasts to obtain the total cluster power.
- **Transformer+DCD**: A Transformer-based model applying DCD from Ref. [29], modeling each sub-cluster every four hours and summing sub-cluster forecasting result for total power.
- **GCN** [15]: Standard GCN model retaining only the graph-convolutional spatial feature extraction, constructing the graph via Pearson correlation, and forecasting the total cluster power through a unified output layer.
- **Dual-Channel GCN** [31]: Models through fluctuation correlation computation and switching mechanisms.

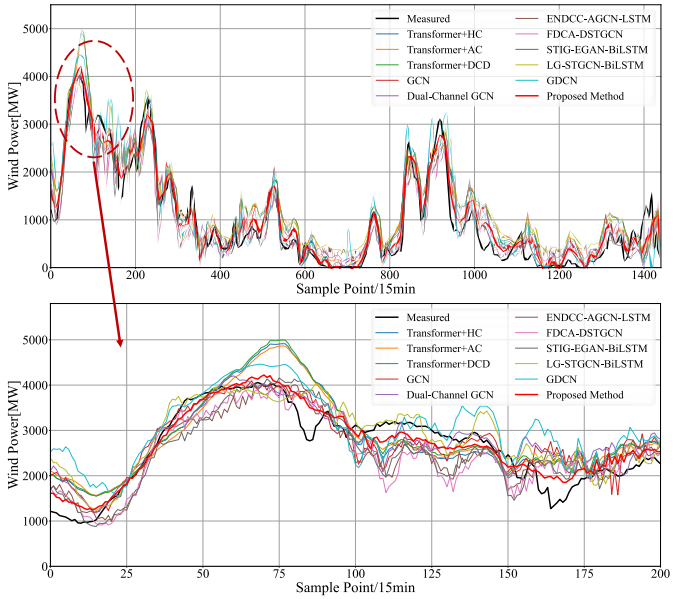


Fig. 9. The forecasts by different forecasting models

TABLE VI
COMPARISON OF DIFFERENT FORECASTING METHODS

Method	NRMSE	NMAE	R
Transformer+HC	0.0787	0.0657	0.8997
Transformer+AC	0.0811	0.0686	0.9017
Transformer+DCD	0.0808	0.0681	0.8978
GCN	0.0727	0.0590	0.9106
Dual-Channel GCN	0.0729	0.0605	0.9081
ENDCC-AGCN-LSTM	0.0706	0.0583	0.8974
FDCA-DSTGCN	0.0682	0.0554	0.9009
STIG-EGAN-BiLSTM	0.0668	0.0548	0.9078
LG-STGCN- BiLSTM	0.0684	0.0584	0.9042
GDCN	0.0672	0.0572	0.9011
Proposed Method	**0.0580**	**0.0484**	**0.9345**

- **ENDCC-AGCN-LSTM** [32]: Builds graphs via global dynamic correlation computation, improved in this study to update dynamically every 4 hours for USTWPF adaptation.
- **FDCA-DSTGCN** [19]: Constructs graphs based on prevailing wind direction correlation with frequency-domain feature extraction for cluster power forecasting.
- **STIG-EGAN-BILSTM** [17]: USTWPF model constructed via STIG.
- **LG-STGCN-BILSTM** [33]: To capture both local and global dependencies in meteorological data.
- **GDCN** [34]: Diffusion structure-based graph diffusion convolution model.

To evaluate the effectiveness of the proposed framework, Table VI reports the forecasting performance of all benchmark models, while Fig. 9 and Fig. 10 provide the corresponding trajectory comparison and overall statistical evaluation.

As shown in **Table VI**, the graph-based forecasting methods generally outperform the three cluster level forecasting paradigms. The NRMSE values of these three methods are 0.0787, 0.0811, and 0.0808, respectively, whereas the graph-based models reduce NRMSE to below 0.073 in all cases. This result indicates that explicitly modeling the spatial

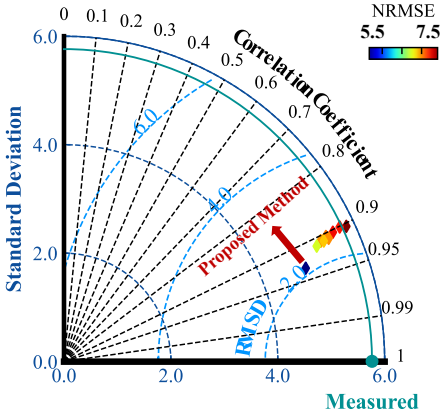


Fig. 10. Taylor diagrams of different forecasting models

TABLE VII
PAIRED t-TEST OF THE PROPOSED METHOD AGAINST REPRESENTATIVE BASELINES

Comparison	NRMSE p-value	NMAE p-value	R p-value
Proposed vs. GCN	4.99×10^{-13}	3.62×10^{-11}	2.43×10^{-11}
Proposed vs. FDCA-DSTGCN	1.45×10^{-10}	1.71×10^{-9}	2.44×10^{-12}
Proposed vs. GDCN	1.93×10^{-11}	3.21×10^{-11}	4.85×10^{-13}
Proposed vs STIG-EGAN-BiLSTM	3.10×10^{-10}	1.64×10^{-9}	2.44×10^{-12}

dependencies and inter farm correlations within wind farm clusters is more effective than relying only on cluster aggregation strategies.

Among all compared methods, the proposed framework achieves the best performance across all evaluation metrics, with an NRMSE of 0.0580, an NMAE of 0.0484, and an R of 0.9345. Compared with the strongest baseline STIG-EGAN-BiLSTM, whose NRMSE is 0.0668, the proposed method achieves a relative reduction of 13.1%. This confirms that the proposed method provides a stronger representation capability than existing graph-based methods.

The forecasts of different models are compared in **Fig. 9** over a shortened representative interval selected from the testing set under the 16-step ultra-short-term forecasting task. This interval is selected because it contains pronounced power fluctuations and peak-valley transitions, which can more clearly reflect the tracking capability of different forecasting models. As shown in the enlarged region, several benchmark models exhibit visible deviations from the measured curve, especially during rapid ramp-up, ramp-down, and local peak periods. In contrast, the proposed method remains closer to the measured power trajectory in both variation trend and fluctuation amplitude. This visual comparison indicates that the proposed model has stronger tracking ability under high-fluctuation conditions. The superiority shown by the curves is also consistent with the quantitative evaluation results reported in the corresponding table.

To further reduce the influence of random initialization and training fluctuations, the proposed method and the strongest baseline, STIG-EGAN-BiLSTM, were evaluated over 10 independent trials. For STIG-EGAN-BiLSTM, the NRMSE, NMAE, and R are 0.0668 ± 0.0008, 0.0548 ± 0.0006, and 0.9078 ± 0.0014, respectively. For the proposed method, the corresponding values are 0.0580 ± 0.0005, 0.0484 ± 0.0004, and 0.9345 ± 0.0010. The corresponding 95% confidence

	Star-shaped	*Muti-center*	*Clustering*	*Dynamic-response*
Method G	√	×	×	×
Method H	×	√	×	×
Method I	×	×	√	×
Method J	√	√	×	×
Method K	√	×	√	×
Method L	√	×	×	√
Method M	×	√	√	×
Method N	×	×	√	√
Method O	√	√	√	×
Method P	×	√	√	√
Method Q	√	√	×	√
Method R	√	×	√	√

Fig. 11. Schematic Diagram of Hyperedge Ablation Experiment Design

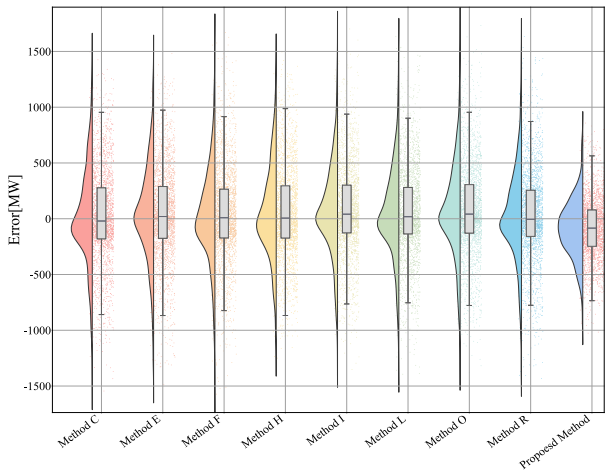


Fig. 12. Raincloud plot of different graph construction methods

intervals for STIG-EGAN-BiLSTM are [0.0662, 0.0674], [0.0544, 0.0552], and [0.9068, 0.9088] for NRMSE, NMAE, and R, respectively, while those of the proposed method are [0.0576, 0.0584], [0.0481, 0.0487], and [0.9338, 0.9352]. These results show that the proposed method achieves lower average forecasting errors and higher correlation with stable variations across repeated runs, suggesting that the improvement is not derived from a single training run. Furthermore, paired t-tests were conducted between the proposed method and representative baseline models based on the 10 independent trials. As shown in Table VII, all p-values for NRMSE, NMAE, and R are substantially below 0.05, confirming that the observed improvements are statistically significant rather than being caused by random initialization or training fluctuations.

The Taylor diagram in **Fig. 10** is used to jointly compare the correlation coefficient, normalized standard deviation, and centered normalized root-mean-square difference of different forecasting models. The measured wind power series is regarded as the reference point. The radial coordinate represents the normalized standard deviation, which describes the agreement between the fluctuation amplitude of the predicted and measured series. The angular coordinate represents the correlation coefficient, which reflects the temporal consistency of the predicted variation trend. The distance between each model point and the reference point corresponds to the centered normalized root-mean-square difference, which measures the deviation between the predicted and measured series after removing the mean bias.

Before constructing the Taylor diagram, both measured and predicted power series are normalized by the installed capacity of the wind farm cluster, so that all compared models are evaluated under the same physical scale. Then, the standard deviation of each predicted series is divided by the standard deviation of the measured series. Therefore, the reference point has a normalized standard deviation of 1. The centered root-mean-square difference is also calculated using the centered predicted and measured series and is normalized by the standard deviation of the measured series. This normalization ensures that the Taylor diagram compares the relative fluctuation amplitude, temporal correlation, and centered deviation of different models under a consistent scale.

As shown in Fig. 10, the model points are relatively clustered, indicating that several methods exhibit similar overall fluctuation characteristics in the Taylor space. Therefore, the Taylor diagram should be interpreted as a complementary visualization rather than the sole evidence of model superiority. Compared with the benchmark models, the proposed method is closer to the reference point, with a higher correlation coefficient, a normalized standard deviation closer to 1, and a smaller centered normalized root-mean-square difference. This suggests that the proposed method better captures the fluctuation amplitude and temporal variation trend of the measured power series.

This improvement can be mainly attributed to two aspects. First, the hierarchical hypergraph construction enables the model to jointly capture local, regional, global, and transient dependencies, thereby improving the representation of multi scale wind field evolution. Second, the input aware modeling strategy adaptively aligns meteorological inputs with different forecasting horizons, which helps reduce temporal phase mismatch and error accumulation in multi-step forecasting.

***(c) Ablation Study on Graph Construction*:** Several adjacency matrix construction methods are employed to benchmark the proposed method

- **Method A**: Graph construction based on the geographical distance, basic GCN.
- **Method B**: Graph construction based on the correlation of NWP at each station, basic GCN.
- **Method C**: Graph construction based on NWP and power correlation, Dual-Channel GCN, extracting future and historical information respectively.
- **Method D**: Graph construction based on power correlation and hypergraph constructed by clustering, GCN and HGCN.
- **Method E**: Graph construction based on NWP and power correlation, and hypergraph constructed by clustering, Dual-Channel GCN and HGCN.

The above five methods are used to verify that the introduction of hypergraph construction methods can effectively improve the accuracy.

- **Method F:** Graph construction based on attention-based dynamic generation of hyperedges.

The other comparative methods are illustrated in **Fig. 11** and are used to verify the effectiveness of the proposed hypergraph construction strategy in this paper. Methods A–R are controlled

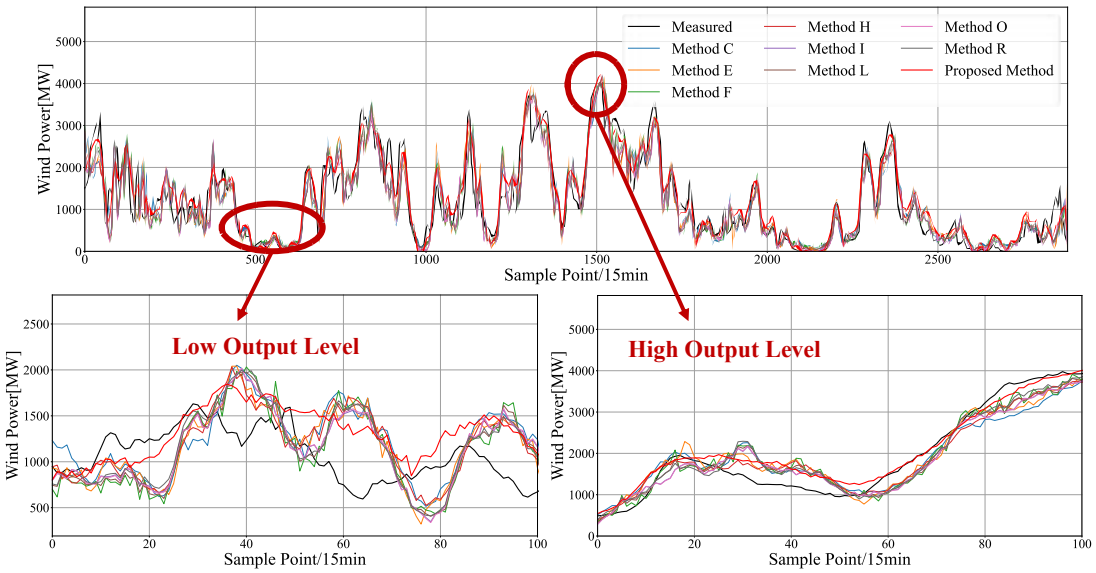


Fig. 13. The forecasts by different graph construction methods

TABLE VIII
COMPARISON OF DIFFERENT ADJACENCY MATRIX CONSTRUCTION FOR THE 16TH STEP

Method	NRMSE	NMAE	R
Method A	0.0709	0.0576	0.9035
Method B	0.0674	0.0557	0.9035
Method C	0.0662	0.0542	0.9047
Method D	0.0659	0.0545	0.9087
Method E	0.0659	0.0536	0.9077
Method F	0.0658	0.0529	0.9072
Method G	0.0657	0.0530	0.9076
Method H	0.0660	0.0534	0.9071
Method I	0.0658	0.0531	0.9074
Method J	0.0653	0.0535	0.9093
Method K	0.0651	0.0527	0.9102
Method L	0.0649	0.0523	0.9108
Method M	0.0652	0.0531	0.9098
Method N	0.0650	0.0525	0.9105
Method O	0.0643	0.0518	0.9118
Method P	0.0646	0.0521	0.9112
Method Q	0.0641	0.0514	0.9129
Method R	0.0639	0.0512	0.9136
Proposed Method	**0.0580**	**0.0484**	**0.9345**

variants constructed in this study for ablation analysis rather than independently proposed forecasting algorithms.

The quantitative results in **Table VIII** further support these observations. The dual adjacency construction based on both power and NWP correlations, namely Method C, outperforms the simpler graph construction methods Method A and Method B, reducing the NRMSE from 0.0709 and 0.0674 to 0.0662. This indicates that jointly modeling historical power and meteorological information is more effective than relying on a single graph construction strategy. **Fig. 12** presents the raincloud visualization of forecasting errors for different methods. The cloud shape reflects the overall error distribution, the scattered points indicate sample-level error dispersion, and the embedded boxplot shows the median, interquartile range, and outlier characteristics. A more compact cloud distribution and a narrower interquartile range indicate that the forecasting errors are more concentrated and that large deviations occur less frequently.

After introducing hypergraph structures, the forecasting performance is further improved. Methods D and E, which incorporate clustering based hypergraph construction, both achieve lower errors than Method C, demonstrating the advantage of modeling high order dependencies beyond pairwise graph relationships. Moreover, the results from Methods F to R show that the combination of multiple hyperedge types consistently improves forecasting performance over single hyperedge settings. In particular, Method O, which integrates star shaped, multi center, and clustering hyperedges, reduces the NRMSE to 0.0643, while Methods Q and R further improve it to 0.0641 and 0.0639 after introducing dynamic response information. This indicates that dynamic response hyperedges provide additional benefits beyond static multi scale hypergraph construction.

As shown in **Fig. 13**, the forecasts obtained by different graph construction strategies are compared within a representative high-fluctuation interval. The purpose of this figure is to show how different spatial dependency modeling strategies affect the ability to track the measured wind power curve. Compared with the other adjacency-matrix-based construction strategies, the proposed hierarchical hypergraph construction follows the measured curve more closely during rapid fluctuation and peak-valley transition periods. This indicates that the proposed graph-hypergraph structure can provide more effective spatial dependency representation for wind farm cluster forecasting. The curve-level observation is consistent with the quantitative results, where the proposed method achieves lower normalized forecasting errors.

Overall, the ablation results demonstrate that different hyperedge types capture distinct but complementary dependencies among wind farms. With the progressive integration of local, regional, global, and transient-response information, both NRMSE and NMAE are gradually reduced, indicating lower overall forecasting deviations and improved average prediction accuracy. Meanwhile, the continuous increase in R shows that the predicted power variations become more consistent with the measured temporal trajectory. The

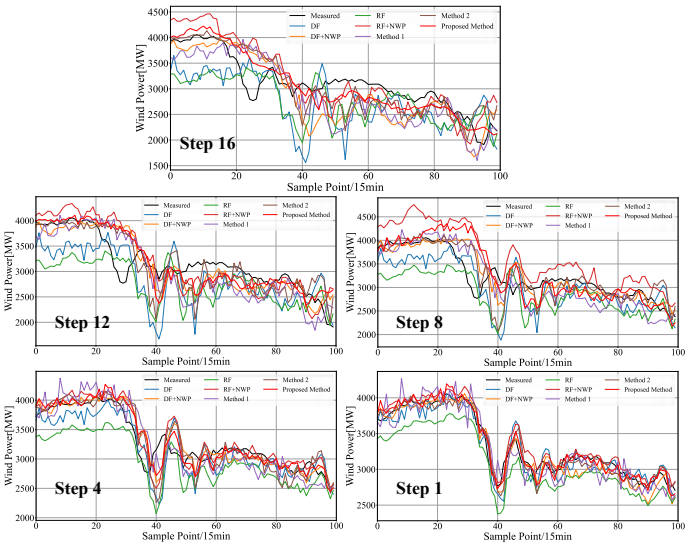


Fig. 14. The forecasts by input modeling methods for different steps

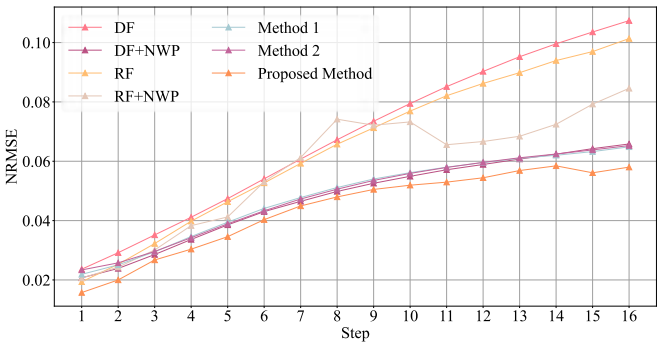


Fig. 15. The NRMSE for different input modeling methods

TABLE IX
COMPARISON OF NRMSE AT STEPS 1, 4, 8, 12, AND 16

Method	Step1	Step4	Step8	Step12	Step16
RF	0.0236	0.0411	0.0672	0.0903	0.1074
RF+NWP	0.0207	0.0336	0.0498	0.0589	0.0658
DF	0.0194	0.0398	0.0657	0.0862	0.1013
DF+NWP	0.0207	0.0383	0.0742	0.0667	0.0846
Method 1	0.0219	0.0346	0.0511	0.0595	0.0649
Method 2	0.0234	0.0342	0.0506	0.0596	0.0653
Proposed Method	**0.0158**	**0.0304**	**0.0470**	**0.0524**	**0.0580**

results do not indicate that any single hyperedge structure consistently dominates across all evaluation metrics. Instead, the final improvement mainly arises from the coordinated aggregation of multiple complementary dependencies within the hierarchical dynamic hypergraph. Quantitatively, hierarchical graph–hypergraph aggregation reduces the NRMSE by approximately 2.87%, while the subsequent incorporation of dynamic-response hyperedges provides a further reduction of approximately 0.62%.

***(d) Input modeling methods*:** To evaluate the effectiveness of the proposed input modeling method, the following comparison strategies are designed:

- **RF:** A recursive forecasting strategy in which the predicted value at each step is fed back as the input for the subsequent step.
- **RF+NWP:** The recursive forecasting strategy further augmented with NWP inputs.
- **DF:** A direct forecasting strategy based on a multi-input multi-output framework, where a single model is trained to simultaneously generate forecasts for all target horizons.
- **DF+NWP:** The direct forecasting strategy further augmented with NWP inputs.
- **Method 1:** An independent modeling strategy for each forecasting horizon, using only the NWP features corresponding to the target forecasting time.
- **Method 2:** An independent modeling strategy for each forecasting horizon, using all NWP features associated with the full set of target horizons.

RF and DF are implemented following the settings in Ref. [27], while RF+NWP and DF+NWP further incorporate NWP features into RF and DF, respectively. Method 1 and Method 2 are two input variants developed in this study, in which the concatenated NWP features are either shared across all forecasting horizons or provided only for the corresponding target horizon.

As shown in **Fig. 14**, the forecasting trajectories of different input modeling strategies at steps 1, 4, 8, 12, and 16 are used to illustrate how the prediction deviation changes as the forecasting horizon increases. The RF and DF strategies rely mainly on historical power information, and their tracking ability becomes limited when the forecasting step increases. After NWP information is introduced, the predicted curves become closer to the measured power trajectory, but phase mismatch and enlarged deviations still appear at longer horizons. Method 1 and Method 2 alleviate this problem to some extent by constructing separate models for different forecasting horizons. By contrast, the proposed input-aware modeling strategy maintains better consistency with the measured curve across different steps, especially around peak and valley variations. This indicates that the proposed strategy can more effectively use horizon-related NWP information and reduce multi-step forecasting deviation.

Together with the quantitative results in **Fig. 15** and **Table IX**, these observations show that the proposed method achieves consistently lower NRMSE across different forecasting horizons. As the prediction step increases, the errors of all methods generally increase, reflecting the growing difficulty of multi-step ultra-short-term forecasting. However, the proposed method maintains a lower overall error level and a smoother error-growth trend than the benchmark methods, especially at longer forecasting horizons. This indicates that the hierarchical dynamic hypergraph structure and input-aware meteorological integration can more effectively capture horizon-related spatial-temporal dependencies and reduce multi-step error accumulation. Therefore, the advantage of the proposed method is reflected not only in a single forecasting step, but also in the overall horizon-wise performance comparison.

The comparisons with RF, RF+NWP, DF, DF+NWP, Method 1, and Method 2 independently quantify the contribution of the input-aware meteorological integration module. The proposed strategy achieves more stable improvements across different forecasting horizons, with its advantage becoming more evident at longer forecasting steps.

***(e) Model interpretability analysis*:** In order to further demonstrate the effectiveness of the modeling method proposed, the mechanism for trusting the model on input features is implemented using LIME [28], which is designed to explain the

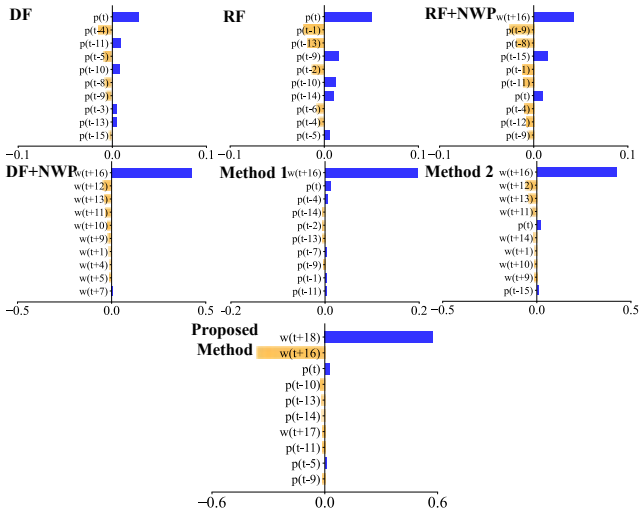


Fig. 16. Trust level of input features in the 16th step

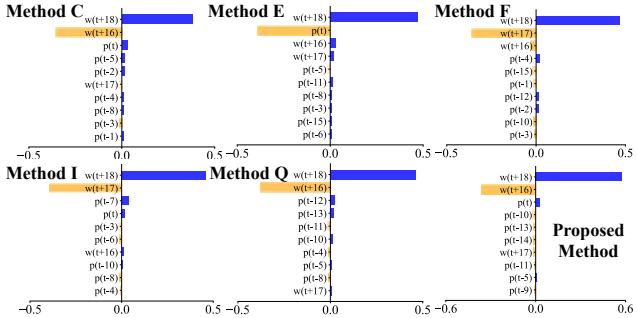


Fig. 17. Trust level of input features from different graph construction

intrinsic logic that leads to a particular forecasting and to identify supporting and opposing forecasting factors. LIME is compatible with any supervised algorithm and its implementation is achieved by building a simple, highly transparent alternative model that explains the mapping relationship of the model to the local samples. In this study, a linear regression model is chosen as the surrogate model and its mathematical expression is shown in (33).

$$\text{explanation}(x) = \arg\min_{g \in G} L(f, g, \pi_x) + \Omega(g) \tag{33}$$

where f *is* the model for wind power forecasting and g is the proxy model for f. π_x represents the locality of the local sample x to be explained. The trust degree of each forecasting modeling method in the 16 steps of forecasting is analysed by the LIME model.

In the LIME analysis, the feature attribution results at the 16th forecasting step for each input modeling strategy are shown in **Fig. 16**. The proposed input-aware modeling method assigns higher importance to NWP features, indicating that it leverages meteorological information more effectively and aligns better with the underlying physical mechanisms. Among all compared methods, the proposed method achieves the highest feature credibility, which contributes to the observed reduction in NRMSE and further demonstrates its superior predictive capability.

Simultaneously, we analyzed the contribution of inputs in the graph construction method, as shown in **Fig. 17**. The results demonstrate that the proposed method assigns higher contribution weights to input features than the other hypergraph designs, enabling the forecasting model to learn feature information more effectively. In particular, different hypergraph types lead to different NWP contribution patterns, reflecting their distinct roles in modeling wind process evolution. Combined with the forecasting accuracy comparison, these results further validate the effectiveness and interpretability of the proposed hypergraph design.

The sensitivity analysis of the lead window size m and the LIME-based feature attribution results further indicate that the proposed input-aware module can adaptively use horizon-related NWP information rather than relying on a fixed meteorological weighting rule.

(f) Effectiveness of Extending to Probabilistic Forecasting: In power systems dominated by renewable energy generation, point forecasting alone is insufficient to comprehensively characterize forecasting uncertainty. Therefore, the proposed method is further extended to a probabilistic forecasting framework to further validate its effectiveness. Probabilistic forecasting is based on quantile loss, which is formulated as in (34).

$$L_{q_{th}} = \begin{cases} q_{th}(P_i - P_{\text{predict},i}) & if\ P \geq P_i^{'} \\ (1-q_{th})(P_i - P_{\text{predict},i}) & if\ P < P_i^{'} \end{cases} \tag{34}$$

where q denotes the corresponding quantile. The following methods are selected for comparative analysis:

- **Method S [35]:** Parametric probabilistic forecasting method based on Gaussian mixture distribution
- **Method T [36]:** Improved kernel density estimation-based probabilistic forecasting method
- **Method U [37]:** Quantile regression probabilistic forecasting method
- **Method V [38]:** Diffusion learning generative probabilistic forecasting method
- **Method W [39]:** Machine learning probabilistic forecasting method using quantile regression forests
- **Method X [31]:** Probabilistic forecasting method based on fluctuation correlation graph construction
- **Method Y [40]:** Probabilistic forecasting method based on multi-model collaborative heterogeneity

Compared with these methods, the proposed method combines quantile-based probabilistic forecasting with hierarchical dynamic hypergraph learning and input-aware meteorological integration, enabling the probabilistic forecasts to benefit from multi-scale spatial dependency modeling and horizon-related NWP feature utilization.

As shown in **Table X**, the proposed method also achieves strong performance when extended to probabilistic forecasting. Under the three confidence levels of 85%, 90%, and 95%, the proposed method attains PICP values of 86.74%, 90.52%, and 95.02%, respectively, indicating that the prediction intervals can effectively cover the actual wind power trajectory. At the same time, the corresponding PINAW values are 29.66%, 32.65%, and 39.24%, showing that the proposed method maintains relatively narrow interval widths while preserving high coverage reliability. This demonstrates that the proposed method achieves a favorable balance between reliability and sharpness.

TABLE X
COMPARISON OF DIFFERENT PROBABILITY FORECASTING METHODS FOR THE 16TH STEP

Method	Confidence	PICP (%)	PINAW (%)	CWC (%)
Method S	85%	75.16	24.94	91.85
	90%	79.36	28.62	111.80
	95%	84.71	34.13	129.95
Method T	85%	82.45	30.26	69.41
	90%	86.73	35.35	84.55
	95%	92.35	44.81	103.41
Method U	85%	77.63	26.89	83.28
	90%	83.84	30.94	88.41
	95%	91.73	37.93	90.73
Method V	85%	83.54	**26.91**	56.98
	90%	87.50	**30.96**	68.63
	95%	92.18	37.95	82.13
Method W	85%	72.72	24.69	108.85
	90%	76.00	27.68	140.34
	95%	86.34	37.38	126.50
Method X	85%	78.81	33.41	221.95
	90%	85.71	36.03	99.09
	95%	92.61	31.20	84.36
Method Y	85%	77.63	29.84	98.76
	90%	84.09	35.10	98.45
	95%	**95.61**	45.01	45.01
Proposed Method	85%	**86.74**	29.66	**29.66**
	90%	**90.52**	32.65	**32.65**
	95%	95.02	**39.24**	**39.24**

From the perspective of the comprehensive metric CWC, the proposed method yields the lowest values among all compared methods at all three confidence levels, namely 29.66%, 32.65%, and 39.24%. This result indicates that the proposed framework provides the best overall probabilistic forecasting performance.

Compared with parameter distribution methods and kernel density estimation methods, quantile regression based and machine learning based approaches generally achieve better interval estimation performance. In addition, recently developed heterogeneous modeling and generative probabilistic forecasting methods further improve the overall quality of interval prediction and achieve relatively low CWC values. Nevertheless, the proposed method still consistently obtains the minimum CWC under all confidence levels, confirming its superiority in jointly optimizing interval coverage and interval width.

Fig. 18 further illustrates this advantage. As shown in the figure, the prediction intervals of the proposed method provide more effective coverage of the actual power curve while remaining more compact than those of the benchmark methods In particular, the proposed method better tracks the fluctuation trend of measured wind power during peak valley transitions and rapid variation periods. These results further verify that the proposed framework can be effectively extended from deterministic forecasting to probabilistic forecasting, thereby providing more comprehensive uncertainty information for decision makers.

(g) Generalization experiments: The above experiments are validated on only one dataset, which is insufficient to demonstrate the generalization capability of the proposed method. Therefore, this study further selects a wind power cluster consisting of 120 wind farms in Inner Mongolia Autonomous Region, China, and another wind power cluster consisting of 97 wind farms in Gansu Province, China, to verify

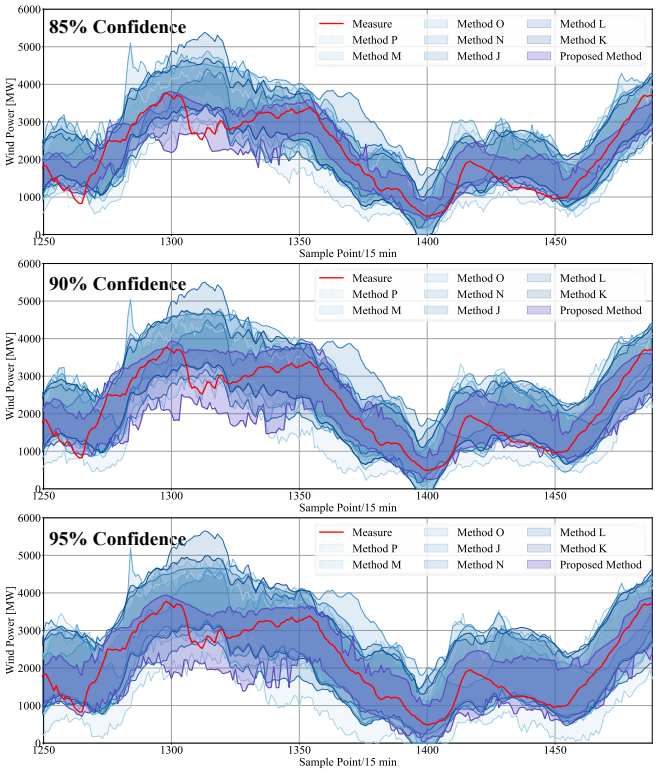
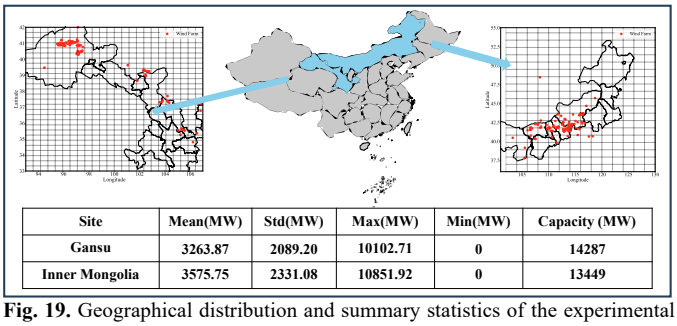


Fig. 18. The forecasts from different probabilistic forecasting methods



Site	Mean(MW)	Std(MW)	Max(MW)	Min(MW)	Capacity (MW)
Gansu	3263.87	2089.20	10102.71	0	14287
Inner Mongolia	3575.75	2331.08	10851.92	0	13449

Fig. 19. Geographical distribution and summary statistics of the experimental dataset

the generalization performance of the proposed method. The distribution and capacity information of the wind farms are shown in **Fig. 19.** For the Inner Mongolia dataset, the data cover an entire year, with the first 11 months used as the training set and the last month used as the test set. For the Gansu dataset, the data span 6 months, with the first 5 months used for training and the last month used for testing. The temporal resolution of both datasets is 15min. The normalization and data preprocessing procedures are the same as those used for the Jilin dataset.

For Inner Mongolia Autonomous Region, we selected two types of NWP data with different quality levels for the experiments, and the corresponding forecasting accuracy is presented in Table XI.

Compared with the baseline models and two advanced methods, the proposed method consistently improves forecasting accuracy under different NWP quality conditions. As shown in **Table XI**, when the NRMSE of the NWP data is 0.0827, the proposed method outperforms all comparative methods. In particular, compared with the strongest baseline, STIG-EGAN-BiLSTM, the proposed method further reduces

TABLE XI
COMPARISON OF DIFFERENT NWP ACCURACY FOR THE 16TH STEP OF OTHER DATASET

Method	NWP Accuracy (NRMSE)	NRMSE	NMAE	R
Transformer + HC	0.0827	0.0629	0.0523	0.9100
FDCA-DSTGCN		0.0601	0.0494	0.9204
STIG-EGAN-BiLSTM		0.0534	0.0468	0.9215
Proposed Method		**0.0489**	**0.0413**	**0.9421**
Transformer + HC	0.1104	0.1184	0.1022	0.8872
FDCA-DSTGCN		0.1052	0.0854	0.8945
STIG-EGAN-BiLSTM		0.0954	0.0800	0.9121
Proposed Method		**0.0809**	**0.0745**	**0.9204**

TABLE XII
COMPARISON OF DIFFERENT FORECASTING METHOD FOR THE DIFFERENT TIME HORIZON OF OTHER DATASET

Horizon	Method	NRMSE (15min)	NRMSE (30min)	NRMSE (1h)
1h ahead	STIG-EGAN-BiLSTM	0.0386	0.0218	0.0120
	Method 1	0.0435	0.0249	0.0141
	Method 2	0.0409	0.0231	0.0129
	Proposed Method	**0.0352**	**0.0194**	**0.0105**
2h ahead	STIG-EGAN-BiLSTM	0.0634	0.0381	0.0219
	Method 1	0.0684	0.0416	0.0248
	Method 2	0.0658	0.0398	0.0230
	Proposed Method	**0.0598**	**0.0352**	**0.0194**
3h ahead	STIG-EGAN-BiLSTM	0.0771	0.0521	0.0305
	Method 1	0.0826	0.0567	0.0336
	Method 2	0.0799	0.0548	0.0318
	Proposed Method	**0.0733**	**0.0489**	**0.0275**
4h ahead	STIG-EGAN-BiLSTM	0.0848	0.0630	0.0387
	Method 1	0.0903	0.0675	0.0421
	Method 2	0.0876	0.0656	0.0404
	Proposed Method	**0.0809**	**0.0598**	**0.0352**
5h ahead	STIG-EGAN-BiLSTM	0.0881	0.0709	0.0460
	Method 1	0.0939	0.0751	0.0497
	Method 2	0.0912	0.0733	0.0479
	Proposed Method	**0.0839**	**0.0678**	**0.0423**
6h ahead	STIG-EGAN-BiLSTM	0.0872	0.0764	0.0526
	Method 1	0.0931	0.0806	0.0568
	Method 2	0.0905	0.0787	0.0547
	Proposed Method	**0.0820**	**0.0733**	**0.0489**

the NRMSE from 0.0534 to 0.0489. This indicates that the proposed framework can effectively exploit higher-quality meteorological inputs to achieve superior forecasting accuracy. When the NWP accuracy decreases to 0.1104, the forecasting errors of all methods increase, indicating that the forecasting task becomes more challenging under lower-quality meteorological inputs. Nevertheless, the proposed method still achieves the best overall performance. Compared with STIG-EGAN-BiLSTM, it reduces the NRMSE by 0.0145 and improves the correlation coefficient from 0.9121 to 0.9204. These results demonstrate that the proposed framework remains effective even when the quality of NWP inputs deteriorates. Overall, the proposed method consistently outperforms the benchmark models under different NWP quality conditions. This confirms that the hierarchical dynamic hypergraph construction and input-aware modeling strategy provide robust representation capability and strong generalization performance in practical forecasting scenarios.

For the dataset from Gansu Province, we further validate the effectiveness of the proposed method by designing experiments

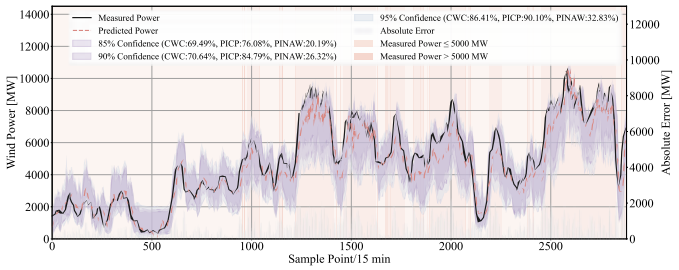


Fig. 20. The forecasts for the 16th-step forecasting results on the Inner Mongolia dataset

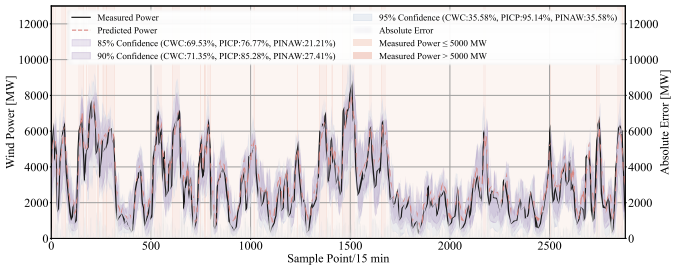


Fig. 21. The forecasts for the 16th-step forecasting results on the Gansu dataset

with different temporal resolutions and longer multi-step forecasting horizons. Specifically, the temporal resolutions are set to 15min, 30min, and 1h, while the forecasting horizon is extended to 6h.

As shown in **Table XII**, the forecasting errors of all methods generally increase as the prediction horizon extends from 1h ahead to 6h ahead, which reflects the cumulative effect of multi-step forecasting errors. However, the rate of increase becomes relatively moderate at longer horizons, especially from 4h ahead to 6h ahead, indicating that the error growth gradually stabilizes. This provides additional evidence of the long horizon error accumulation characteristics of the forecasting task. Meanwhile, the results also reveal clear sensitivity to the prediction time span. For the same target horizon, the NRMSE values under the 30min and 1h resolutions are consistently lower than those under the 15min resolution, since coarser temporal resolutions correspond to fewer forecasting steps and therefore lower prediction difficulty.

Combined with **Fig. 20** and **Fig. 21**, the results further show that the proposed method achieves superior forecasting performance on both the Inner Mongolia and Gansu datasets in point forecasting and probabilistic forecasting tasks. In point forecasting, the predicted curves of the proposed method closely track the variation trend of the actual power and maintain relatively small absolute errors during both high-output and low-output periods. In probabilistic forecasting, the proposed method generates more effective prediction intervals, attaining high coverage rates with relatively narrow interval widths. Moreover, among all compared methods, the proposed framework consistently achieves the best performance under different forecasting horizons and temporal resolutions. These findings further verify its strong generalization capability and indicate that the proposed framework can effectively suppress long-horizon error accumulation while preserving robust performance across different forecasting scenarios.

TABLE XIII
COMPUTATIONAL COST COMPARISON OF REPRESENTATIVE MODELS

Method	Parameters (M)	FLOPs (M)	Training Time	Running Time
Jilin Province				
GCN	0.01	58.43	0.5h	4s
Dual-Channel GCN	0.01	96.27	3h	6s
STIG-EGAN-BiLSTM	0.03	186.54	7h	9s
Proposed Method	0.01	341.42	15h	15s
Gansu Province				
GCN	0.01	132.86	2h	7s
Dual-Channel GCN	0.01	218.74	7h	11s
STIG-EGAN-BiLSTM	0.03	421.35	16h	17s
Proposed Method	0.01	742.68	28h	28s
Inner Mongolia				
GCN	0.01	167.42	2h	9s
Dual-Channel GCN	0.01	276.59	9h	14s
STIG-EGAN-BiLSTM	0.03	538.76	19h	22s
Proposed Method	0.01	926.31	34h	36s

***(h) Limitation analysis*:** Although the proposed framework achieves superior forecasting accuracy, several limitations should be noted.

First, as shown in **Table XIII**, the main practical limitation of the proposed method lies in its relatively high training cost. Since the graph and hypergraph convolution kernels are shared across wind farm nodes, the number of trainable parameters remains nearly unchanged across different regional datasets under the same model configuration. However, as the number of wind farms increases, the graph adjacency matrices, hypergraph incidence matrices, and candidate node-pair relationships become larger, resulting in higher FLOPs, storage requirements, training time, and online running time. For the Jilin dataset with 48 wind farms, the proposed method contains only 0.01M trainable parameters, while its FLOPs reach 341.42M due to dynamic hyperedge generation, multi-head attention-based candidate pair construction, and joint graph-hypergraph learning. When extended to the larger Gansu and Inner Mongolia datasets, the computational cost further increases.

Although the proposed method requires higher training cost than conventional GCN-based models, its online inference time remains well below the 15min update interval required for ultra-short-term forecasting. Therefore, the proposed framework is still feasible for practical deployment, while improving training efficiency through lightweight hyperedge generation and sparse hypergraph sampling remains an important direction for future work.

Second, the current hypergraph construction strategy still depends on predefined structural settings, especially the number of hyperedge neighbors and the number of clusters. Although the sensitivity analysis demonstrates that the proposed method is robust within a reasonable parameter range, these quantities are still determined by heuristic rules rather than being learned in a fully adaptive manner. Therefore, the setting of hypergraph structure remains another limitation of the present study. It should be noted that the robustness analysis in this study focuses on high-fluctuation forecasting intervals, different NWP quality levels, and cross-regional generalization. Dedicated sensor-noise injection and missing-NWP-input scenarios are not included and will be considered in future work.

Third, although recent pretrained weather foundation models provide a promising direction for large scale atmospheric representation learning, they were not included in the present comparison because their research boundary, data form, and modeling paradigm differ from those of this study. Specifically, such models are typically developed based on large scale gridded meteorological fields and upstream atmospheric representation learning, whereas this work focuses on wind farm cluster level forecasting based on operational power measurements and synchronized NWP features. Therefore, they were not incorporated into the current research framework for direct comparison. Although the proposed method improves ultra-short-term forecasting accuracy by enhancing spatial dependency learning through hierarchical dynamic hypergraph modeling, it does not introduce a new dispatch strategy or grid-balancing control method. The operational value of the improved forecasts, such as their influence on reserve allocation, dispatch cost, and grid-balancing performance, will be further investigated by coupling the forecasting model with power-system operation and dispatch models in future work.

Future research will therefore focus on three directions: improving training efficiency through lightweight dynamic hypergraph generation and sparse hyperedge sampling, reducing manual dependence in hypergraph structure design through adaptive hyperedge pruning and data driven hypergraph updating, and integrating pretrained meteorological representations into the proposed cluster level forecasting framework to further enhance its forecasting capability.

IV. CONCLUSION

This study proposed a USTWPF framework for large-scale wind farm cluster power forecasting. The proposed method uses GCNs to extract local historical power and future meteorological features, and constructs a hierarchical dynamic hypergraph to model local, regional, global, and dynamic-response relationships among wind farms. In addition, an input-aware meteorological integration strategy is introduced to adaptively incorporate horizon-related NWP information for multi-step forecasting.

The experimental results show that the proposed framework achieves the best overall forecasting performance among the compared methods under the tested settings. In the main case study, the proposed method reduces the average NRMSE and NMAE by 1.5% and 1.2%, respectively, and improves the correlation coefficient R by 3.15%. Compared with the strongest baseline, STIG-EGAN-BiLSTM, the proposed method achieves a 13.1% relative reduction in NRMSE. Further experiments on graph construction ablation, multi-step forecasting, probabilistic forecasting extension, different NWP quality conditions, and longer forecasting horizons verify the effectiveness of the proposed spatial dependency modeling and input-aware meteorological integration strategies. Additional dataset experiments also show promising transferability under the tested regional wind farm cluster conditions, while broader generalization still requires further validation.

Despite the improved forecasting accuracy, several limitations remain. The proposed method focuses on forecasting performance improvement and does not introduce a new dispatch strategy or grid-balancing control method. In

addition, the hierarchical dynamic hypergraph increases computational cost, and some hypergraph construction parameters still require empirical selection or sensitivity analysis. Future work will focus on lightweight dynamic hypergraph generation, adaptive hyperedge updating and pruning, and validation under more diverse wind farm layouts, terrain conditions, NWP sources, seasonal patterns, and extreme meteorological environments.

REFERENCES

[1] Global Wind Energy Council. Global Wind Report 2025. 2025, https://gwec.net/global-wind-report-2025/.

[2] National Energy Administration. The grid-connected operation situation of renewable energy in 2024. 2025, https://www.nea.gov.cn/20250221/e10f363cabe3458aaf78ba4558970054/c.html.

[3] M. A. Hossain, E. Gray, J. Lu, et al., "Optimized Forecasting Model to Improve the Accuracy of Very Short-Term Wind Power Prediction." *IEEE Transactions on Industrial Informatics*, vol. 19, no. 10, pp. 10145-10159, Oct. 2023, https://doi.org/10.1109/TII.2022.3230726.

[4] State Grid Corporation of China. Function specification of wind power forecasting system: NB/T 31046-2022 [2023-05-04]. https://www.doc88.com/p-35729449244944.htm.

[5] N. Zhu, Y. Wang, K. Yuan, et al., "Enhancing Wind Power Forecasting at Local Peak Points: A Novel Seq2LPP Model." *IEEE Transactions on Industrial Informatics*, vol. 21, no. 4, pp. 3286-3295, Apr. 2025, https://doi.org/10.1109/TII.2024.3523581.

[6] Z. Chen, C. Du, B. Zhang, et al., "Multistep Joint Probabilistic Forecasting of Offshore Wind Power: A Confidence-Triggered Clustering Missing-Data Tolerant Model." *IEEE Transactions on Industrial Informatics*, vol. 21, no. 12, pp. 9802-9812, Dec. 2025, https://doi.org/10.1109/TII.2025.3609054.

[7] M. Yang, Y. Huang, C. Xu, et al., "Review of several key processes in wind power forecasting: Mathematical formulations, scientific problems, and logical relations," *Applied Energy*, vol. 377, Art. no. 124361, Jan. 2025, https://doi.org/10.1016/j.apenergy.2024.124631.

[8] L. Liu, J Liu, Y Ye, et al. "Ultra-short-term wind power forecasting based on deep Bayesian model with uncertainty." *Renewable Energy*. vol.205, pp.598-607, Mar. 2023, https://doi.org/10.1016/j.renene.2023.01.038.

[9] L. Wang, J. Gao, Y. Li, et al., "A Method for Ultra-Short-Term Wind Power Forecasting of Large-Scale Wind Farms Based on Adaptive Spatiotemporal Graph Convolution." *Renewable Energy*, vol. 249, Art. no. 123321, Aug. 2025, https://doi.org/10.1016/j.renene.2025.123321.

[10] L. Yang, L. Wang, Z. Zheng et al., "A Continual Learning-Based Framework for Developing a Single Wind Turbine Cybertwin Adaptively Serving Multiple Modeling Tasks." *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4912-4921, Jul. 2022, https://doi.org/10.1109/TII.2021.3130721.

[11] T. Niu, J. Wang, P. Du et al., "WPFSAD: Wind Power Forecasting System Integrating Dual-Stage Attention and Deep Learning," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 11, pp. 11252-11264, Nov. 2023, https://doi.org/10.1109/TII.2023.3245196.

[12] J. Wang, M. Kou, R. Li, et al., "Ultra-Short-Term Wind Power Forecasting Jointly Driven by Anomaly Detection, Clustering and Graph Convolutional Recurrent Neural Networks." *Advanced Engineering Informatics*, vol. 65, Art. no. 103137, Jan. 2025, https://doi.org/10.1016/j.aei.2025.103137.

[13] P. Arora, S. M. J. Jalali, S. Ahmadian, et al., "Probabilistic Wind Power Forecasting Using Optimized Deep Auto-Regressive Recurrent Neural Networks," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 3, pp. 2814-2825, Mar. 2023. https://doi.org/10.1109/TII.2022.3160696.

[14] C. Liu, X. Zhang, S. Mei, et al., "Local-pattern-aware forecast of regional wind power: Adaptive partition and long-short-term matching." *Energy Conversion and Management*, vol.231, Art. no. 113799, Mar. 2021, https://doi.org/10.1016/j.enconman.2020.113799.

[15] Y. Wang, T. Zhao, M. Cui, et al., "Representation Then Augmentation: Wide Graph Clustering Network with Multi-Order Filter Fusion and Double-Level Contrastive Learning," *IEEE/CAA Journal of Automatica Sinica*, vol. 13, no. 2, pp. 421-435, Feb. 2026, https://doi.org/10.1109/JAS.2025.125564.

[16] Y. Song, D. Tang, J. Yu, et al., "Short-Term Forecasting Based on Graph Convolution Networks and Multiresolution Convolution Neural Networks for Wind Power." *IEEE Transactions on Industrial Informatics*, vol. 19, no. 2, pp. 1691-1702, Feb. 2023, https://doi.org/10.1109/TII.2022.3176821.

[17] M. Yang, Y. Guo, and F. Fan. "Ultra-Short-Term Prediction of Wind Farm Cluster Power Based on Embedded Graph Structure Learning with Spatiotemporal Information Gain." *IEEE Transactions on Sustainable Energy,* vol. 16, no. 1, pp. 308–322, Sept. 2024, https://doi.org/10.1109/tste.2024.3455759.

[18] Z. Li, Y. Lin, Y. Zhao, et al., "A Spatiotemporal Directed Graph Convolution Network for Ultra-Short-Term Wind Power Prediction." *IEEE Transactions on Sustainable Energy*, vol. 14, no. 1, pp. 39–54, Jan. 2023, https://doi.org/10.1109/TSTE.2022.3198816.

[19] M. Yang, J. Niu, B. Wang, et al., "FDCA-DSTGCN: A Wind Farm Cluster Power Day-ahead Prediction Model Based on Frequency Domain Information Gain and Dynamic Trend Sensing" *IEEE Transactions on Sustainable Energy*, vol.16, no. 4, pp. 2782-2796, Oct. 2025, https://doi.org/10.1109/TSTE.2025.3566592.

[20] Y. Gao, Z. Zhang, H. Lin, et al., "Hypergraph Learning: Methods and Practices." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 5, pp. 2548-2566, May. 2022, https://doi.org/10.1109/tpami.2020.3039374.

[21] N. Yadati, M. Nimishakavi, P. Yadav, et al., "HyperGCN: A New Method of Training Graph Convolutional Networks on Hypergraphs." *ArXiv*, Jan. 2018, https://doi.org/10.48550/arxiv.1809.02589.

[22] W. Zhong, P. Li, C. Tu. "Dynamic Hypergraph Wind Power Ultra Short Term Prediction Method Based on Wake Correlation." *Proceedings of the CSEE* (in Chinese), vol. 29, pp. 1-18.

[23] X. Wu, H. Zhan, J. Hu, et al., "Non-Stationary GNNCrossformer: Transformer with Graph Information for Non-Stationary Multivariate Spatio-Temporal Wind Power Data Forecasting." *Applied Energy*, vol. 377, Sept. 2024, pp. 124492, https://doi.org/10.1016/j.apenergy.2024.124492.

[24] Y. Zhao, H. Liao, S. Pan, et al., "Interpretable Multi-Graph Convolution Network Integrating Spatial-Temporal Attention and Dynamic Combination for Wind Power Forecasting." *Expert Systems with Applications*, vol. 255, Dec. 2024, pp. 124766, https://doi.org/10.1016/j.eswa.2024.124766.

[25] Y. Cui, C. Cheng, Y. Zhu, et al., "A Novel Two-Step Framework for Day-Ahead Wind Power Probabilistic Forecasting Considering Uncertainties and Ramp Events." *International Journal of Electrical Power & Energy Systems*, vol. 169, May 2025, pp. 110780, https://doi.org/10.1016/j.ijepes.2025.110780.

[26] M. Yang, R. Jiang, B Wang, et al., "Multi-Channel Attention Mechanism Graph Convolutional Network Considering Cumulative Effect and Temporal Causality for Day-Ahead Wind Power Prediction." *Energy*, Jun. 2025, pp. 137023, https://doi.org/10.1016/j.energy.2025.137023.

[27] J. Wang, Y Song, F. Liu, et al., "Analysis and Application of Forecasting Models in Wind Power Integration: A Review of Multi-Step-Ahead Wind Speed Forecasting Models." *Renewable and Sustainable Energy Reviews*, vol. 60, Jul. 2016, pp. 960–981, https://doi.org/10.1016/j.rser.2016.01.114.

[28] M. Yang, C. Xu, Y. Bai, et.al., "Investigating black-box model for wind power forecasting using local interpretable model-agnostic explanations algorithm: Why should a model be trusted?" *CSEE Journal of Power and Energy Systems*, vol. 11, no. 1, pp.227-242, Jan 2025, https://doi.org/10.17775/CSEEJPES.2021.07470.

[29] M Yang, T. Peng, W. Zhang et al., "Ultra-short Term Wind Power Prediction Based on Two-dimensional Coordinate Dynamic Division of Prediction Information". *Proceedings of the CSEE* (in Chinese), 2022,42(24):8854-8864.

[30] S. Mo, H. Wang, B. Li, et al., "Powerformer: A temporal-based transformer model for wind power forecasting," *Energy Reports*, vol. 11, pp. 736–744, Jun. 2024, https://doi.org/10.1016/j.egyr.2023.12.030.

[31] M. Yang, Y. Huang, Y. Guo, et al., "Ultra-short-term wind farm cluster power prediction based on FC-GCN and trend-aware switching mechanism," *Energy*, vol. 290, Art. no. 130238, Mar. 2024, https://doi.org/10.1016/j.energy.2024.130238.

[32] M. Yang, C. Ju, Y. Huang, et al., "Short-Term Power Forecasting of Wind Farm Cluster Based on Global Information Adaptive Perceptual Graph Convolution Network," *IEEE Transactions on Sustainable Energy*, vol. 15, no. 3, pp. 2063-2076, Jul. 2024, https://doi.org/10.1109/TSTE.2024.3397877.

[33] H. Liu, J. Huang, J. Hu, et al., "Enhancing wind power prediction accuracy: A novel method integrating seasonal temporal factors and advanced spatio-temporal feature extraction." *Energy*, vol. 336, Art. no. 138514, Nov. 2025, https://doi.org/10.1016/j.energy.2025.138514.

[34] X. Zhang, J. Wang, Y. Wang, et al., "MS-CGDM: Multi-scale conditional graph diffusion model for extreme weather source-load scenario generation." *Energy*, vol. 336, Art. no. 138492, Nov. 2025, https://doi.org/10.1016/j.energy.2025.138492.

[35] H. Zhang, Y. Liu, J. Yan, et al., "Improved deep mixture density network for regional wind power probabilistic forecasting," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2549–2560, Jul. 2020, https://doi.org/10.1109/TPWRS.2020.2971607.

[36] W. Dong, H. Sun, J. Tan, et al., "Regional wind power probabilistic forecasting based on an improved kernel density estimation, regular vine copulas, and ensemble learning." *Energy*, vol. 238, Art. no.122045, Jan. 2022, https://doi.org/10.1016/j.energy.2021.122045.

[37] J. Zhu, Y. He, X. Yang, et al., "Ultra-short-term wind power probabilistic forecasting based on an evolutionary non-crossing multi-output quantile regression deep neural network," *Energy Conversion and Management*, vol. 301, Art. no. 118062, Feb. 2024, https://doi.org/10.1016/j.enconman.2024.118062.

[38] L. Yin, Y. Liu. "Bidirectional encoder representations from transformer fusion quantum dual-stage attention bidirectional gated recurrent unit and diffusion method for short-term wind power prediction." *Engineering Applications of Artificial Intelligence*, vol. 169, Art. no. 114095, Apr. 2026, https://doi.org/10.1016/j.engappai.2026.114095.

[39] J. Jonkers, D. N. Avendano, G. V. Wallendael, et al., "A novel day-ahead regional and probabilistic wind power forecasting framework using deep CNNs and conformalized regression forests." *Applied Energy*, vol. 361, Art. no. 122990, May. 2024, https://doi.org/10.1016/j.apenergy.2024.122900.

[40] Y. Chen, J. Wang, and Z. Li. "Joint probability prediction of multi-site wind power based on multi-model collaborative heterogeneity." *Energy*, vol. 340, Art. no. 139236, Nov. 2025, https://doi.org/10.1016/j.energy.2025.139236.



Mao Yang received the Ph.D. degree in control theory and control engineering from Jilin University, Changchun, China, in 2010. Since 2018, he has been a Professor with the School of Electrical Engineering, Northeast Electric Power University, Jilin, China. His research interests include the areas of load forecasting and micro-grid operation management with an emphasis on wind power/photovoltaic power prediction.



Yutong Huang received the M.Eng. degree in Electrical Engineering from Northeast Electric Power University, in China, in 2023. He is currently pursuing a Ph.D. degree in Electrical Engineering at the Northeast Electric Power University, China. His research interests include wind power / photovoltaic power prediction at multiple Spatiotemporal scales



Bo Wang is a Professor of Engineering and Research Director with the China Electric Power Research Institute, Beijing, China. His research interests include electrical meteorology, renewable energy generation, and renewable power forecasting. He has participated as a project leader or core member in more than 50 national, local-government, and State Grid Corporation of China science and technology projects.



Chuanyu Xu received the B.S. degree in new energy science and engineering and the M.S. degree in electrical engineering from Northeast Electric Power University, Jilin, China, in 2019 and 2022, respectively. He is currently working toward the Ph.D. degree in electrical engineering with the Huazhong University of Science and Technology, Wuhan, China. His research interests include renewable energy power forecasting and scenario generation.



Jingsi Chen received her M.Sc. degree in ICT Engineering for Smart Societies from Politecnico di Torino, Italy, in 2024. She is currently pursuing the Ph.D. degree in the Department of Energy, Politecnico di Torino. Her research interests focus on electricity market, optimization and the intersection of artificial intelligence and power system.



Mahmood Hosseini Imani received the Ph.D. degree in Electrical Engineering from the Politecnico di Torino, Turin, Italy, in 2022. He is currently an Assistant Professor with the Politecnico di Torino. His research interests include electricity markets, power system operation and optimization, renewable energy integration, energy sustainability, data analytics, machine learning, demand response, and microgrids.



Tao Huang received the Ph.D. degree from the Politecnico di Torino, Turin, Italy, in 2011. He is currently a Researcher and Professor with the Department of Energy, Politecnico di Torino, Turin, Italy, and XiHua University, Chengdu, China, respectively. His research interests include critical infrastructure protection, vulnerability assessment, electricity markets, smart grids, etc.