

## Article

# Enhancing Sustainable Disaster Resilience: A Physics-Informed Spatial Attention Network for Wind Gust Forecast Correction at Sparse Stations

Mengyu Li <sup>†</sup>, Chi Yang <sup>†</sup>, Hao Huang and Xiaofang Liu <sup>\*</sup>

School of Computer Science and Engineering, Sichuan University of Science & Engineering, Yibin 644000, China; 324085406115@stu.suse.edu.cn (M.L.); 323085406228@stu.suse.edu.cn (C.Y.); 324085406108@stu.suse.edu.cn (H.H.)

<sup>\*</sup> Correspondence: lxf1969@163.com

<sup>†</sup> These authors contributed equally to this work.

## Abstract

Wind gusts pose an increasing threat to sustainable development, damaging resilient infrastructure (SDG 9), disrupting clean energy systems (SDG 7), and endangering community safety (SDG 11). However, the reliability of early warning systems remains limited by systematic biases in Numerical Weather Prediction (NWP) models and insufficient uncertainty quantification, particularly in regions with sparse monitoring networks. To address these challenges in sustainable disaster risk reduction, this study proposes a physics-informed deep learning framework—the Physics-Informed Spatial Attention Network (PISA-Net). The model integrates high-resolution WRF-UPP forecasts as a physical prior within a Transformer-based architecture, enabling effective bias correction and spatial dependency learning under data-sparse conditions. A hybrid probabilistic learning objective is employed to simultaneously improve deterministic gust predictions and provide calibrated uncertainty estimates. Evaluated on 61 extratropical cyclone events in the northeastern United States, PISA-Net substantially outperforms baseline NWP and conventional deep learning models, reducing the mean absolute error and root mean square error to 1.75 m/s and 2.26 m/s, respectively. In addition, the resulting 95% prediction intervals are well calibrated and offer reliable risk-based guidance. By improving both the accuracy and credibility of wind gust forecasts, PISA-Net provides a practical decision-support tool for infrastructure maintenance, wind farm operations, and public safety planning. This work demonstrates the potential of physics-informed deep learning to strengthen sustainable early warning systems in data-sparse regions.

**Keywords:** wind gust forecasting; bias correction; physics-informed deep learning; sustainable disaster risk reduction



Academic Editor: Luca Stabile

Received: 20 January 2026

Revised: 8 February 2026

Accepted: 10 February 2026

Published: 15 February 2026

**Copyright:** © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

## 1. Introduction

Wind gust, also known as “instantaneous maximum wind speed”, is a term used to describe sudden, strong winds, and in aviation flight weather routines, refers to the maximum value of instantaneous wind speeds greater than or equal to the average wind speed of 5 m/s within two or ten minutes [1,2]. Strong wind events from these storms pose a dual challenge: an immediate threat to life, property, and economic stability, and a significant barrier to sustainable development. They damage critical infrastructure and disrupt the wind energy sector, hindering the transition to renewable energy (SDG 7). In addition, cascading effects, such as fallen trees, power outages, and traffic paralysis,

undermine the socio-economic and environmental foundations necessary to build resilient and sustainable communities [3–5]. Therefore, high-precision and high-reliability gust forecasts are essential to issue effective warnings, optimize wind farm scheduling, and ensure the safety of social activities [6,7]. Currently, gust forecasts are mainly based on numerical weather prediction (NWP) models based on physical equations, such as the European Center for Medium-Range Weather Forecasts (ECMWF) model and the Weather Research and Forecasting (WRF) model [6,8]. However, the mechanism of gust generation involves complex physical interactions such as small-scale turbulence, topographic forcing, boundary layer processes, etc., and its spatial and temporal scales are often lower than the resolution of conventional observation networks and operational NWP models [9,10]. Therefore, even in high-resolution NWP models, gusts are difficult to accurately resolve and forecast due to their localized and sudden nature [7,11]. In addition, NWP models have their own uncertainties due to initial field errors, limitations of the physical parameterization scheme, and model approximations, which lead to significant Systematic Bias (SB) in their forecasts at specific locations, especially in complex terrain regions [6,12]. For example, the ECMWF model is difficult to provide accurate site-scale wind speed estimates due to its relatively coarse horizontal resolution (e.g., 9 km) [2,13].

Wind gust forecasting has evolved from traditional statistical methods to deep learning and machine learning approaches, with recent advances that integrate physical information [14–16]. Early studies, such as those by Cheng et al. [17], demonstrated that MOS outperforms methods like Kalman filtering, especially in stable weather conditions. The introduction of machine learning further improved forecasts, with Wang et al. [2] proposing a sparse Gaussian process-based MOS framework that improves multi-step probabilistic predictions. Kahl et al. [1] combined meteorologically stratified gust factors with MOS, exceeding climatological benchmarks. As deep learning gained traction, Schulz et al. [7] and Del et al. [18] showed that neural network-based post-processing methods, such as DRN, BQN, and MLP, effectively reduced biases and improved accuracy. Coburn et al. [19] found deep neural networks outperform traditional methods in forecasting gust probabilities and magnitudes, though underestimation of extreme events persists. More recently, Aslam et al. [20] introduced physics-informed vectors into deep learning models to improve wind speed forecasts.

Despite significant progress, existing gust forecasting methods still face notable limitations. The GNN-Transformer framework by Bentsen et al. [21] captures spatial dependencies but struggles with dynamic, non-uniform data and lacks an efficient multi-source forecast fusion mechanism. Leclerc et al. [3] post-processed regional gust probabilities using a convolutional network, but their approach relies on fixed zoning, making it unreliable for sparse sites. Potosnak et al. [22] incorporated physical equations for strong wind time series forecasting, yet their reliance on data-driven empirical factors undermines physical consistency and limits extrapolation to extreme events. Jahan et al. [6] used evidential neural networks to quantify uncertainty, improving accuracy, but still underestimated extreme gusts due to the scarcity of extreme samples. A key challenge across these methods is the insufficient quantification of uncertainty [23]. Traditional Bayesian methods, such as Gaussian process regression [24], are computationally intensive and unsuitable for sparse, dynamic data.

To address the critical challenges in gust forecasting—namely systematic biases in physical models, sparse observational coverage, and inadequate uncertainty quantification—this study proposes the Physics-Informed Spatial Attention Network (PISA-Net) to enhance sustainable disaster resilience and clean energy security. The model employs a residual learning strategy [25], using WRF physical forecasts as a prior benchmark and learning the residuals between the predictions  $\hat{Y}$  and the observations  $y$  to achieve physically consis-

tent bias correction. To accommodate spatially sparse and unevenly distributed station data, PISA-Net incorporates a masked modeling approach [26], dynamically masking non-observed sites to focus learning on available gust records. In addition, a hybrid probabilistic loss function is adopted to jointly optimize point predictions and uncertainty estimation, enabling the generation of reliable probabilistic forecasts and prediction intervals that support risk-informed decision-making for infrastructure protection, wind farm operations, and public safety planning. PISA-Net advances interpretable and actionable gust forecasting and contributes to climate adaptation and sustainable development under increasing weather extremes. The remainder of this paper is organized as follows: Section 1 introduces the background and sustainability implications; Section 2 describes the data sources and preprocessing; Section 3 outlines the model design, and training methodology of PISA-Net; Section 4 presents the experimental results and discusses their practical relevance; and Section 5 concludes with key findings and future research directions aligned with sustainable forecasting needs.

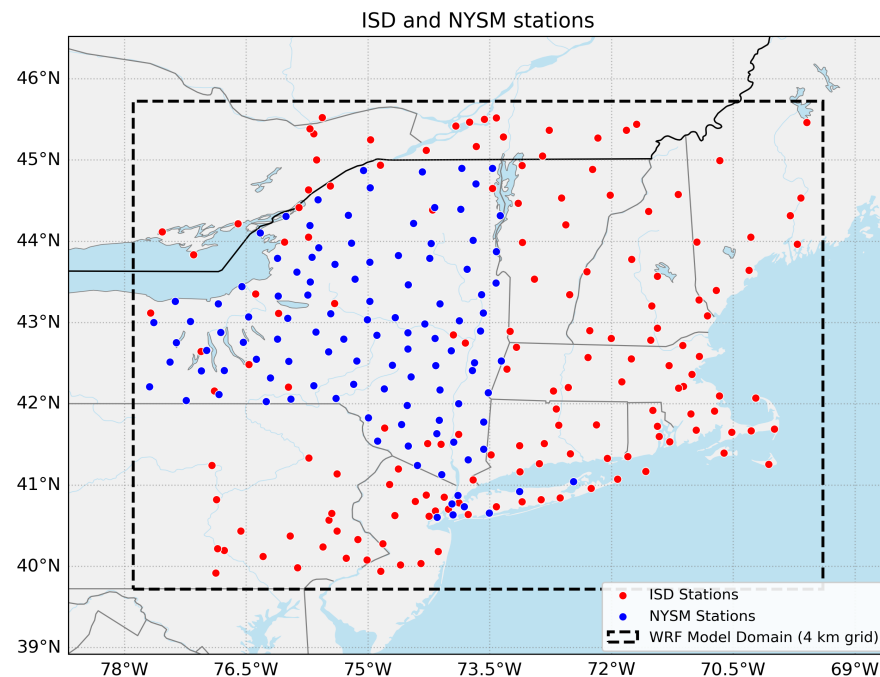
## 2. Data and Preprocessing

### 2.1. Data

The data used in this study are entirely based on the publicly available dataset of Jahan et al. [6], which contains hourly wind gust observations and corresponding WRF model outputs for 61 extratropical storm events over the northeastern United States during 2017–2021. Wind gust observations follow the 3 s sliding average definition recommended by the World Meteorological Organization (WMO) [27]. Observations originate from two sources: the National Oceanic and Atmospheric Administration's Integrated Surface Database (ISD) and the New York State Mesonet (NYSM) [28]. ISD aggregates measurements from multiple observing systems, including ASOS, AWOS, METAR, C-MAN, and buoy stations, and applies multi-level quality-control procedures. NYSM, a high-density surface network, provides 5-min-averaged gust records at standard stations and is used to construct the hourly gust series following the processing methodology described in the original dataset publication [6]. Specifically, for each hour, gust values at 10 min before, 5 min before, and at the top of the hour are extracted, and the maximum of the three is selected as the representative hourly gust. These values are subsequently merged with quality-controlled ISD observations. In total, 266 surface stations are included in this study.

The WRF Model is a widely used mesoscale numerical weather prediction system capable of simulating complex atmospheric processes from the convective scale to the regional scale [29,30]. Its Unified Post-Processor (UPP), on the other hand, is a standardized diagnostic and post-processing tool that transforms the raw grid-point variables of the WRF model outputs into more user-friendly meteorological elements, e.g., WRF-UPP (after post-processing, this step is referred to in the paper as WRF-UPP) can estimate the maximum gust potential by a parameterized scheme based on the downward propagation of the boundary-layer momentum, which can provide a basic product for operational forecasting and subsequent analysis.

The WRF simulations were originally performed using version 4.2.2 with a two-way nested configuration consisting of a 12-km outer domain (D01) and a 4-km inner domain (D02), as described by Jahan et al. [12] (Figure 1). Each storm was simulated for 60 h with a 12-h spin-up period removed, resulting in 48 h of hourly output. In this work, only variables from the D02 domain are extracted as predictors for model training and evaluation. Detailed physical parameterization schemes and initial-boundary condition settings are referred to in the original publication and are not repeated here.



**Figure 1.** Spatial distribution of meteorological observation stations in the study area. Red dots represent ISD sites (156) and blue dots represent NYSM sites (110). The black dashed rectangular box represents the inner simulation area of the WRF model (D02, 4 km resolution).

The predictor variables adopted in this study follow the original dataset construction and variable selection procedures described in that work [6], and no additional feature selection is performed in this study. This feature set was originally constructed to represent the key physical processes governing wind gust generation, including vertical momentum transport, boundary-layer turbulence, surface roughness effects, and atmospheric stability. Surface pressure and 2-m air temperature are included in the released dataset and are therefore preserved to represent synoptic-scale forcing and near-surface thermal influences. This physically motivated and vertically distributed predictor set allows the proposed model to capture multi-scale controls on surface gusts, from large-scale pressure gradients to boundary-layer mixing and downward momentum transport. In this study, we refer to the gust forecasts estimated from the UPP as WRF-UPP and denote them as  $v_{wrf}$ . Detailed descriptions of all variables are provided in Table 1.

**Table 1.** WRF-derived atmospheric variables used as input features for training PISA-Net.

| Features                | Description                         | Units |
|-------------------------|-------------------------------------|-------|
| WS_10 m                 | Wind speed at 10 m height           | m/s   |
| WS_850 mb               | Wind speed at 850 mb pressure level | m/s   |
| WS_950 mb               | Wind speed at 950 mb pressure level | m/s   |
| PBLH                    | Planetary boundary layer height     | km    |
| Ustar                   | Friction velocity                   | m/s   |
| WindDC (sin)            | Sine of surface wind direction      | –     |
| WindDC (cos)            | Cosine of surface wind direction    | –     |
| Terrain_height          | Terrain height                      | m     |
| Lapse_rate (sfc_1 km)   | Lapse rate (surface to 1 km height) | °C/km |
| Lapse_rate (sfc_2 km)   | Lapse rate (surface to 2 km height) | °C/km |
| PSFC                    | Surface pressure                    | Pa    |
| T2                      | 2-m temperature                     | °C    |
| Gust_WRF-UPP (Baseline) | WRF-UPP gust forecast               | m/s   |

## 2.2. Data Reorganization

To ensure the temporal validity of the model evaluation and prevent information leakage, this study strictly follows the principle of constructing the original dataset and adopts the strategy of dividing it by storm events. A total of 61 storms were divided into three mutually exclusive subsets in chronological order of occurrence: the training set (the first 42 storms, totaling 131,889 data), the validation set (the subsequent 8 storms, 44,585 data) and the test set (the last 11 storms, totaling 45,068 data). In the model training phase, all input features are Z-score normalized (Equation (1)) to improve the model training stability and convergence efficiency.

$$\hat{x} = \frac{x - \mu_{\text{train}}}{\sigma_{\text{train}}} \quad (1)$$

The symbols  $\mu_{\text{train}}$  and  $\sigma_{\text{train}}$  represent the mean and standard deviation of the training set features, respectively. It is important to note that both the validation set and the test set are normalized using the statistics of the training set to strictly avoid data leakage.

To adapt the PISA-Net model proposed in this paper, we have performed key data restructuring. At the same time, to accommodate the sparse distribution of meteorological stations and the dynamic nature of observations in real-world scenarios, we propose a spatiotemporal set representation method. Specifically, for a given forecast time  $t$ , we define the set  $\mathcal{S}_t$  as the dynamic set of stations that report valid gust observations at that moment. For each station  $i \in \mathcal{S}_t$ , we extract the corresponding 12 WRF-derived physical feature variables to form a feature vector, with  $v_{\text{wrf}}$  as the physical reference. Therefore, the data at each time  $t$  is organized as the following triplet, which serves as the basic unit for model processing:

$$X_t = [x_t^1, x_t^2, \dots, x_t^{N_t}] \in \mathbb{R}^{N_t \times 12} \quad (2)$$

$$v_{\text{wrf},t} = [v_{\text{wrf},t}^1, v_{\text{wrf},t}^2, \dots, v_{\text{wrf},t}^{N_t}]^T \in \mathbb{R}^{N_t} \quad (3)$$

$$y_t^{\text{obs}} = [y_t^{\text{obs},1}, y_t^{\text{obs},2}, \dots, y_t^{\text{obs},N_t}]^T \in \mathbb{R}^{N_t} \quad (4)$$

In the following:

- $N_t = |\mathcal{S}_t| \leq 266$  represents the number of valid stations at time  $t$ , which dynamically reflects the sparse distribution of stations over time;
- $x_t^i \in \mathbb{R}^{12}$  is the feature vector at station  $i$  at time  $t$  for the WRF model;
- $v_{\text{wrf},t}^i$  is the forecast value of the WRF model for station  $i$  at time  $t$ , as a physical reference;
- $y_t^{\text{obs},i}$  is the actual observed value of station  $i$  at time  $t$ , as the monitoring target.

The above Equations (2)–(4) form a sparse observation representation with a spatiotemporal structure that allows the model to directly deal with the problem of uneven distribution of station data. Based on this representation, unlike the direct output of gust forecasts [6,31], PISA-Net is designed as a physically guided residual correction model whose core output is the corrected residuals based on the WRF forecast gust values  $m$  and their uncertainty estimates  $\sigma$  (in this study,  $\sigma$  is interpreted as aleatoric uncertainty associated with observational errors, sub-grid-scale variability, and unresolved physical processes):

$$(m, \sigma) = f_{\theta}(X_t, v_{\text{wrf},t}) \quad (5)$$

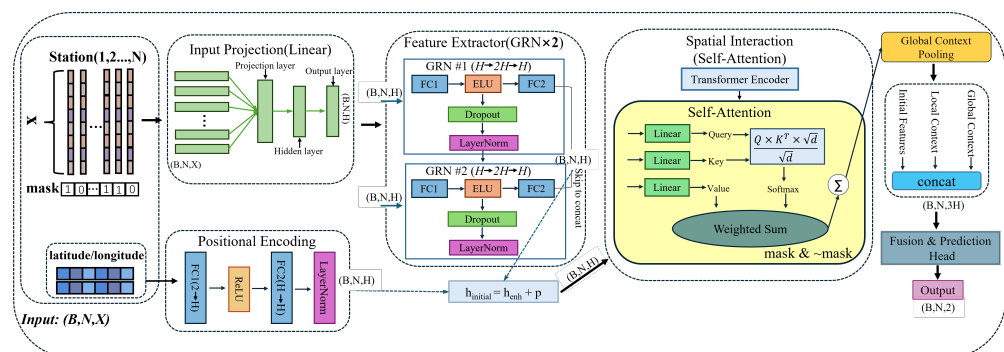
The final wind forecast value  $\hat{y}_t$  is obtained by adding  $m$  to  $v_{\text{wrf}}$ :

$$\hat{y}_t = v_{\text{wrf}} + m \quad (6)$$

### 3. Methods

#### 3.1. PISA-Net Model Architecture

In this study, we propose the PISA-Net model (Figure 2), which utilizes a “physically guided residual correction” strategy to learn residuals and uncertainties between the physical forecasts from WRF-UPP and the observations. This is achieved through a deep learning model that uses WRF-UPP’s physical forecasts as the baseline. The core architecture of the model includes components such as a gated residual network (GRN), position encoding, interactive spatial Transformer, and global context fusion. These modules enhance the model’s capability and generalization performance under sparse observation conditions by enabling hierarchical feature fusion and spatial-temporal embedding, ultimately enabling robust gust residual prediction and uncertainty quantification.



**Figure 2.** Schematic architecture of the proposed PISA-Net model for wind gust bias correction and uncertainty quantification.

##### 3.1.1. Feature Enhancement and Positional Encoding

After linear projection of the input features into a unified hidden space, the projected representation  $\mathbf{h}_{\text{proj}}$  is first obtained. Based on this representation, two stacked Gated Residual Networks (GRNs) are used to perform adaptive nonlinear feature enhancement, producing enhanced feature representation  $\mathbf{h}_{\text{enh}}$ . Subsequently, the geographic coordinate information is encoded to generate the positional embedding  $\mathbf{p}$ , which is injected into the enhanced features of GRN through the addition of elements. This results in the position-aware representation  $\mathbf{h}_{\text{initial}}$ , which serves as input to the subsequent spatial interaction modeling module.

$$\mathbf{h}_{\text{proj}} = \text{FC}_{\text{in}}(\mathbf{x}) \quad (7)$$

$$\mathbf{h}_{\text{enh}} = \text{GRN}_2(\text{GRN}_1(\mathbf{h}_{\text{proj}})) \quad (8)$$

$$\mathbf{p} = \text{FC}_2(\text{ReLU}(\text{FC}_1(\text{lon, lat}))) \quad (9)$$

$$\mathbf{h}_{\text{initial}} = \mathbf{h}_{\text{enh}} + \mathbf{p} \quad (10)$$

##### 3.1.2. Spatial Interaction

After obtaining the position-sensitive feature representation  $\mathbf{h}_{\text{initial}}$ , the model performs spatial interaction modeling to capture inter-station dependencies. A Transformer Encoder is applied to  $\mathbf{h}_{\text{initial}}$  to learn spatial correlations through masked self-attention, producing Transformer-enhanced features  $\mathbf{h}_{\text{trans}}$ . To effectively handle missing station data, a binary mask  $M \in \{0, 1\}^{B \times N}$  is introduced, where  $M_j = 1$  indicates a valid observation at station  $j$ , and  $M_j = 0$  denotes missing data. This mask is consistently used in both the

attention computation and global feature aggregation. Based on  $\mathbf{h}_{\text{trans}}$ , masked global pooling aggregates valid station features into a global contextual representation  $\mathbf{h}_{\text{global}}$ . Finally,  $\mathbf{h}_{\text{trans}}$ ,  $\mathbf{h}_{\text{initial}}$ , and the global broadcast context  $\mathbf{h}_{\text{global}}$  are fused to obtain the multi-scale spatial representation  $\mathbf{h}_{\text{fused}}$  for subsequent prediction.

$$\mathbf{h}_{\text{trans}} = \text{TransformerEncoder}(\mathbf{h}_{\text{initial}}), \quad \mathbf{h}_{\text{trans}} \in \mathbb{R}^{B \times N \times H} \quad (11)$$

$$\mathbf{h}_{\text{global}} = \sum_{j=1}^N M_j \cdot \mathbf{h}_{\text{trans},j}, \quad M \in \{0,1\}^{B \times N} \quad (12)$$

$$\mathbf{h}_{\text{fused}} = \text{LayerNorm}(\mathbf{h}_{\text{trans}} + \mathbf{h}_{\text{initial}} + \text{Repeat}(\mathbf{h}_{\text{global}}, N)) \quad (13)$$

### 3.1.3. Joint Optimization of Output and Probabilistic Prediction

Finally, the model further processes the features through the deep prediction network. The network consists of two fully connected layers, GELU activation functions, and a Dropout layer. The final output is the wind bias value  $m \in \mathbb{R}^{B \times N}$  and its uncertainty estimate  $\sigma \in \mathbb{R}^{B \times N}$ , calculated as:

$$m, \sigma = \text{Split}(\text{MLP}_{\text{pred}}(h_{\text{fused}})) \quad (14)$$

where  $\text{MLP}_{\text{pred}} : \mathbb{R}^H \rightarrow \mathbb{R}^2$  projects the input feature  $h_{\text{fused}}$  onto a two-dimensional output to predict the wind bias. The  $\text{Split}(\cdot)$  operation divides the output into two values representing the wind prediction and its uncertainty estimate. The wind bias value represents the model's prediction for the wind change, while the uncertainty estimate provides a measure of confidence in the prediction, helping users interpret the model's prediction reliability and spatial uncertainty.

The loss function guides the optimization of the deep learning model by adjusting its parameters during training. Most previous studies have relied on MAE or MSE as the default objective functions for gust prediction [32–34]. To simultaneously improve point prediction accuracy and achieve well-calibrated uncertainty estimates, we adopt a hybrid probabilistic loss function following established uncertainty-aware regression formulations in prior work [6]. This objective augments the mean squared error term with a Gaussian negative log-likelihood component to regularize predictive uncertainty. The overall loss function is defined as:

The loss function  $\mathcal{L}$  is defined as:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N \left[ (y_i - \hat{y}_i)^2 + \lambda \left( \log \sigma_i^2 + \frac{(y_i - \hat{y}_i)^2}{\sigma_i^2} \right) \right] \quad (15)$$

where  $\hat{y}_i$  denotes the final prediction,  $y_i$  is the observed value, and  $\sigma_i^2$  represents the predictive variance estimated by the model. The parameter  $\lambda$  controls the trade-off between point prediction accuracy and uncertainty calibration. A sensitivity analysis on the validation set (Table 2) shows that MAE and RMSE remain relatively stable for  $\lambda \in [0.01, 1.0]$ , whereas uncertainty-related metrics vary more substantially: larger  $\lambda$  values yield wider prediction intervals and higher coverage rates. Accordingly, we select  $\lambda = 0.1$  as a balanced configuration that preserves deterministic accuracy while maintaining reasonable uncertainty calibration. In Table 2, arrows indicate whether lower ( $\downarrow$ ) or higher ( $\uparrow$ ) values correspond to better performance, and boldface highlights the selected configuration.

**Table 2.** Sensitivity analysis of the weighting parameter  $\lambda$  on validation data.

| $\lambda$   | MAE (m/s) ( $\downarrow$ ) | RMSE (m/s) ( $\downarrow$ ) | PICP (95%) ( $\uparrow$ ) | MPIW (Width) ( $\downarrow$ ) |
|-------------|----------------------------|-----------------------------|---------------------------|-------------------------------|
| 0.01        | 2.00                       | 2.58                        | 87.3%                     | 8.29                          |
| 0.05        | 1.97                       | 2.56                        | 89.6%                     | 8.51                          |
| <b>0.10</b> | <b>1.99</b>                | <b>2.59</b>                 | <b>89.4%</b>              | <b>8.45</b>                   |
| 0.50        | 1.98                       | 2.58                        | 90.1%                     | 8.63                          |
| 1.00        | 1.97                       | 2.57                        | 90.6%                     | 8.73                          |

We implemented PISA-Net and all baseline models using the PyTorch 2.7.0 framework. The experiments were conducted using station-based meteorological observations. All meteorological predictors and spatial coordinates were standardized using Z-score normalization. The network is built upon a Transformer architecture, in which geographic coordinates are concatenated with meteorological predictors and processed through a multi-head self-attention mechanism. The encoder consists of four stacked layers with four attention heads per layer, and a dropout rate of 0.1 is adopted to mitigate overfitting. Physical consistency is incorporated through a residual learning strategy. For optimization, the Adam optimizer was employed with an initial learning rate of  $1 \times 10^{-3}$ , which was adjusted during training using a cosine annealing scheduling strategy together with default momentum parameters. The model was trained for up to 100 epochs with a batch size of 32, and an early-stopping mechanism with a patience of 10 epochs was implemented based on the validation Gaussian negative log-likelihood loss to prevent overfitting. All experiments were conducted on a workstation equipped with a single NVIDIA RTX 4090 GPU (Nvidia, Santa Clara, CA, USA) with 24 GB of memory.

### 3.2. Evaluation Indicators

In this study, a variety of statistical metrics are used to assess the forecast performance of the model. Two main types of assessment methods are used: the first type focuses on forecast errors, using the mean absolute error (MAE), the root mean square error (RMSE), the center root mean square error (CRMSE) and the Pearson correlation coefficient, which are used to quantitatively measure the precision and consistency of the model in forecasting gusts, as shown in the following Equations (16)–(19). The second category is the Prediction Interval Coverage Probability (PICP), which is used to test the reliability of uncertainty estimates: the actual coverage effect of forecast uncertainty is assessed by constructing a prediction interval at a desired confidence level (e.g., 95%) and calculating the proportion of actual observations that fall within the interval. Together, these indicators systematically evaluate the comprehensive performance of the model from two dimensions: deterministic forecast accuracy and probabilistic forecast reliability.

The formulas for the first type of evaluation indicator are shown below.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (16)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (17)$$

$$\text{CRMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n [(\hat{y}_i - \bar{y}) - (y_i - \bar{y})]^2} \quad (18)$$

$$\text{Pearson} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (19)$$

where  $\hat{y}_i$  represents the  $i$ -th sample's final predicted wind value,  $y_i$  represents the  $i$ -th sample's actual observed wind value,  $\bar{y}$  represents the mean of the observed wind values, and  $n$  is the total number of samples.

To evaluate the construction and coverage performance of prediction intervals, we follow the procedure below.

First, based on the model outputs including the wind prediction mean  $\hat{y}_i$  and the corresponding uncertainty estimate  $\sigma$ , the prediction interval (PI) at a given confidence level is constructed as [35–37]

$$PI = \hat{y}_i \pm z\sigma, \quad (20)$$

where  $z$  denotes the  $z$ -score corresponding to the confidence level (e.g.,  $z = 1.96$  for a 95% confidence level). Accordingly, the lower and upper bounds of the prediction interval are defined as

$$L_{PI} = \hat{y}_i - z\sigma, \quad U_{PI} = \hat{y}_i + z\sigma. \quad (21)$$

Subsequently, the reliability of the prediction interval is evaluated using the PICP, which is defined as

$$PICP = \frac{1}{n} \sum_{i=1}^n C_i, \quad C_i = \begin{cases} 1, & \text{if } y_i \in [L_{PI}, U_{PI}], \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

Here,  $n$  denotes the total number of samples. The PICP directly reflects the empirical coverage ability of the prediction interval at a given confidence level. For example, under a 95% confidence level, a PICP value of 0.75 indicates that 75% of the observations fall within the interval  $\hat{y}_i \pm 1.96\sigma$ . In general, higher confidence levels correspond to wider prediction intervals and higher expected coverage rates.

## 4. Results and Analysis

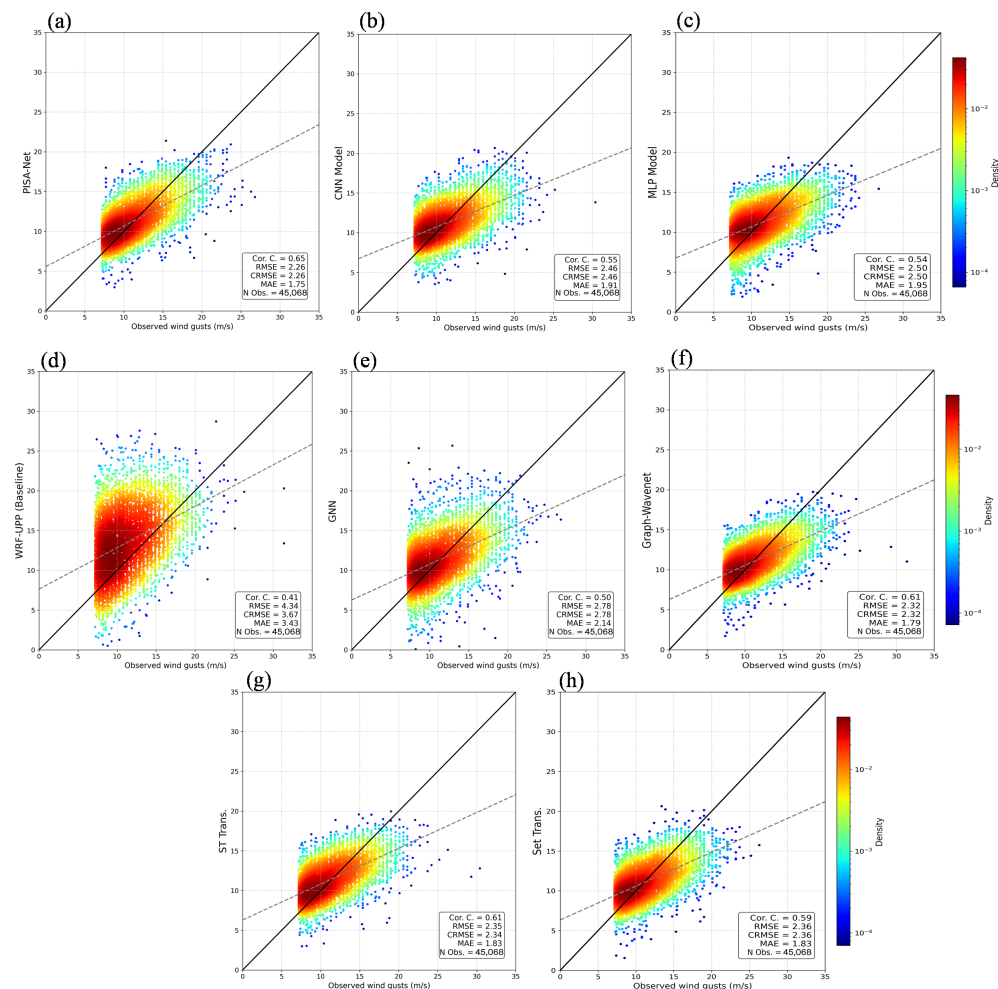
### 4.1. Baseline Model

In order to comprehensively evaluate the prediction performance of PISA-Net, we selected MLP [38], CNN [39], GNN [40], Set transformer [41] (hereafter referred to as Set Trans.), Spatio-temporal Transformer [42] (hereafter referred to as ST Trans.), and Graph WaveNet [43] as the baseline model, using WRF-UPP as the baseline value. As a generalized state-of-the-art architecture for dealing with irregular point set data, Set Trans. provides a rigorous mathematical benchmark for this study. ST Trans. provides a standard paradigm for modeling Spatio-temporal dependencies through the parallel Spatio-temporal attention mechanism. Graph WaveNet fuses graph diffusion convolution with temporal convolution to explicitly model the spatio-temporal propagation of meteorological elements, and represents the state-of-the-art of graph neural networks in the GNN. GNN as a representative benchmark model for spatial deep learning, aims to validate the necessity of explicitly modeling inter-site spatial correlations. All comparison models are trained and fine-tuned under a uniform experimental setup, including the same data segmentation, input features, training epoch number, and optimization strategy to ensure the fairness of the comparison [44].

### 4.2. Comparison of Results

The scatter distributions and corresponding statistical metrics of the predictions from different models on the test set with respect to the observed gusts are shown in Figure 3, and the quantitative results are summarized in Table 3 (The arrows and boldface follow the same conventions as in Table 2, while underlined values indicate the reference results used to compute the IMP). The WRF-UPP baseline exhibits pronounced systematic overestimation,

with a low prediction–observation correlation (0.41) and large errors (RMSE of 4.34 m/s and MAE of 3.43 m/s). The data-driven models substantially outperform the numerical baseline. However, MLP and CNN still show clear underestimation in the high-wind-speed range, a well-known limitation of deep learning models when learning from imbalanced samples at extreme values. The GNN, Graph WaveNet, ST Trans., and Set Trans. further improve both the overall fitting and extreme-value representation, yielding higher correlations (up to 0.61) and lower errors compared with the conventional neural networks. Among all compared approaches, the proposed PISA-Net achieves the best overall performance, with the highest correlation coefficient (0.65) and the lowest RMSE (2.26 m/s) and MAE (1.75 m/s). In addition, relative to MLP and CNN, the RMSE is decreased by 9.6% and 8.1%, and the MAE by 10.3% and 8.4%, respectively. Although slight underestimation remains for very high gust values, PISA-Net demonstrates superior capability in fitting the overall distributions. These improvements indicate that the proposed model can more effectively extract regional meteorological features and enhance prediction consistency through spatial interaction modeling and dynamic ensemble processing.

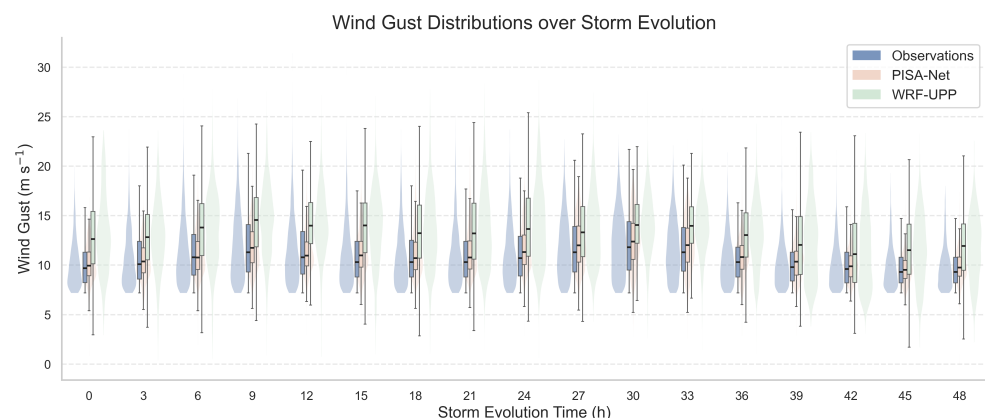


**Figure 3.** (a–h), respectively, present the joint probability distributions of predicted and observed wind-gust speeds from the PISA-Net, CNN, MLP, WRF-UPP baseline, GNN, Graph-Wavenet, ST Trans., and Set Trans. models. Color shading denotes the density of samples. The corresponding evaluation metrics, including Cor, RMSE, CRMSE, and MAE, are provided in each subplot.

**Table 3.** Comparative performance of PISA-Net and comparison models

| Model           | RMSE ( $\downarrow$ ) | CRMSE ( $\downarrow$ ) | MAE ( $\downarrow$ ) | Pearson ( $\uparrow$ ) |
|-----------------|-----------------------|------------------------|----------------------|------------------------|
| MLP             | 2.50                  | 2.50                   | 1.95                 | 0.54                   |
| CNN             | 2.46                  | 2.46                   | 1.91                 | 0.55                   |
| GNN             | 2.78                  | 2.77                   | 2.14                 | 0.49                   |
| Set Trans.      | 2.36                  | 2.36                   | 1.83                 | 0.59                   |
| ST Trans.       | 2.34                  | 2.34                   | 1.83                 | 0.61                   |
| Graph           | 2.32                  | 2.32                   | 1.79                 | 0.61                   |
| WaveNet         | 2.32                  | 2.32                   | 1.79                 | 0.61                   |
| WRF-UPP         | 4.34                  | 3.67                   | 3.43                 | 0.41                   |
| <b>PISA-Net</b> | <b>2.26</b>           | <b>2.26</b>            | <b>1.75</b>          | <b>0.65</b>            |
| IMP (%)         | 2.59                  | 2.59                   | 2.23                 | 6.56                   |

The results demonstrate that PISA-Net is more effective in capturing regional meteorological features and enhancing forecast consistency through spatial interaction modeling and dynamic ensemble processing. To further validate these findings and evaluate the robustness of the models from a distributional perspective, we perform a box-plot analysis of the predicted distributions for all samples in the test set (Figure 4) [45]. Wind gusts are governed by sub-grid-scale turbulence, boundary-layer dynamics, and topographic effects that are difficult to resolve at kilometer-scale model resolutions. It has been indicated in the literature [6] that WRF-UPP exhibits systematic misalignment with observations during intense storms, particularly for extreme gusts exceeding 25 m/s. Moreover, complex terrain has been shown to amplify both gust magnitudes and forecast uncertainty, further challenging physics-based models. These factors collectively explain the pronounced dispersion and bias of WRF-UPP in the high-wind regimes observed in Figure 4.

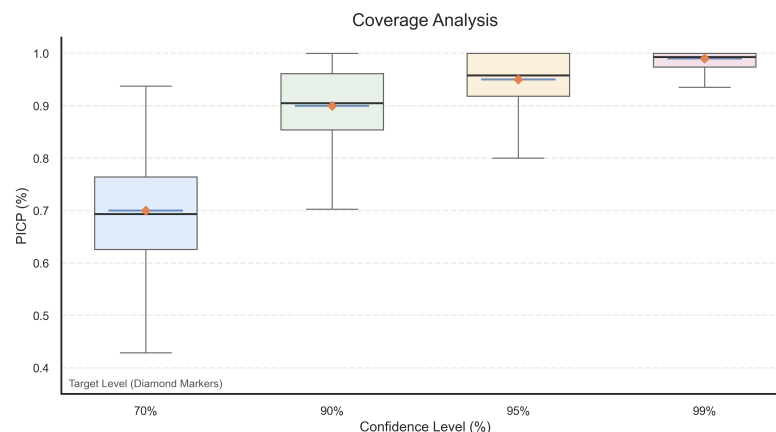


**Figure 4.** Box plots of wind gusts at 3 h intervals during the 48 h test-storm period, comparing observations, PISA-Net corrected forecasts, and the raw WRF-UPP baseline. Hour 0 denotes storm onset and hour 48 denotes the end of the storm period. For each time bin, box plots are constructed using samples from all stations across the 11 test storms.

In contrast, the box distributions of PISA-Net at various storm times are closest to the observations. Its median is consistent with the observed values, and its interquartile range reasonably reflects the time-varying dispersion of gust speeds. PISA-Net also better captures the high-wind tails during the storm intensification phase while avoiding the systematic bias exhibited by WRF-UPP. These results indicate that PISA-Net not only outperforms the baselines in terms of statistical error metrics, but also demonstrates higher reliability and stability in representing the gust distribution structure and its temporal evolution.

### 4.3. Uncertainty Analysis

PISA-Net demonstrates strong overall performance in quantifying uncertainty in gust forecasts. The behavior of the prediction intervals under different confidence levels is first examined. When a 70% confidence level (approximately twice the standard deviation) is adopted, the median PICP is only 0.68 and the distribution is highly discrete, indicating a substantial risk of under-coverage when using the “mean  $\pm$  uncertainty” interval at this level, which limits its usefulness for operational decision-making. As the confidence level is increased to 95%, the median PICP rises markedly to 0.96 and the distribution becomes much more concentrated, suggesting that the corresponding prediction intervals provide substantially improved coverage while maintaining a reasonable width (Figure 5).

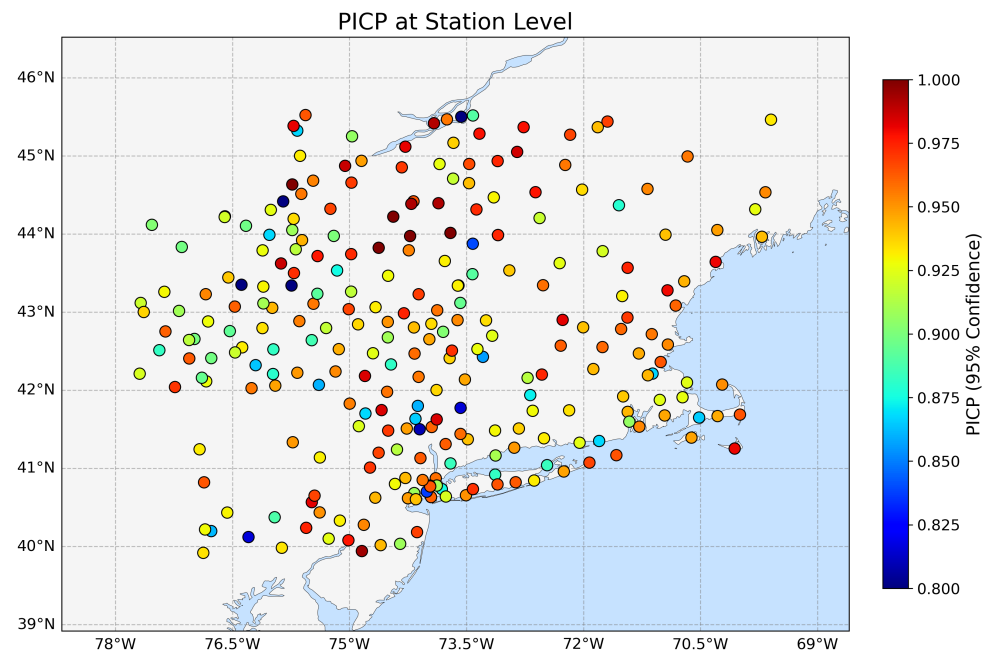


**Figure 5.** Boxplots of PICP at four confidence levels (70%, 90%, 95%, and 99%). The median PICP at each level is close to the nominal confidence level, indicating good probabilistic calibration.

A station-wise analysis at the 95% confidence level further reveals spatial heterogeneity in the uncertainty estimates. Approximately 67% of the stations (179 stations) achieve PICP values reaching or exceeding 0.95, indicating that the constructed prediction intervals successfully encompass most observed gust events. Nevertheless, about 33% of the stations still exhibit PICP values below 0.95, and 4.9% fall below 0.85 (Table 4 and Figure 6). These underperforming stations are primarily located in regions with complex terrain, such as the White Mountains and the Catskills, implying that uncertainty quantification remains challenging in mountainous areas characterized by strong spatial variability and sharp gust gradients. This limitation is likely related to the under-representation of extreme gust samples in the training dataset. From an operational perspective, these results suggest that adopting a 95% confidence level to construct prediction intervals (defined as “mean  $\pm$  1.96  $\times$  uncertainty”) offers a practical balance between coverage reliability and interval width. Such intervals can more effectively communicate forecast uncertainty and provide robust guidance for decision-making in applications such as energy management, transportation safety, and emergency response.

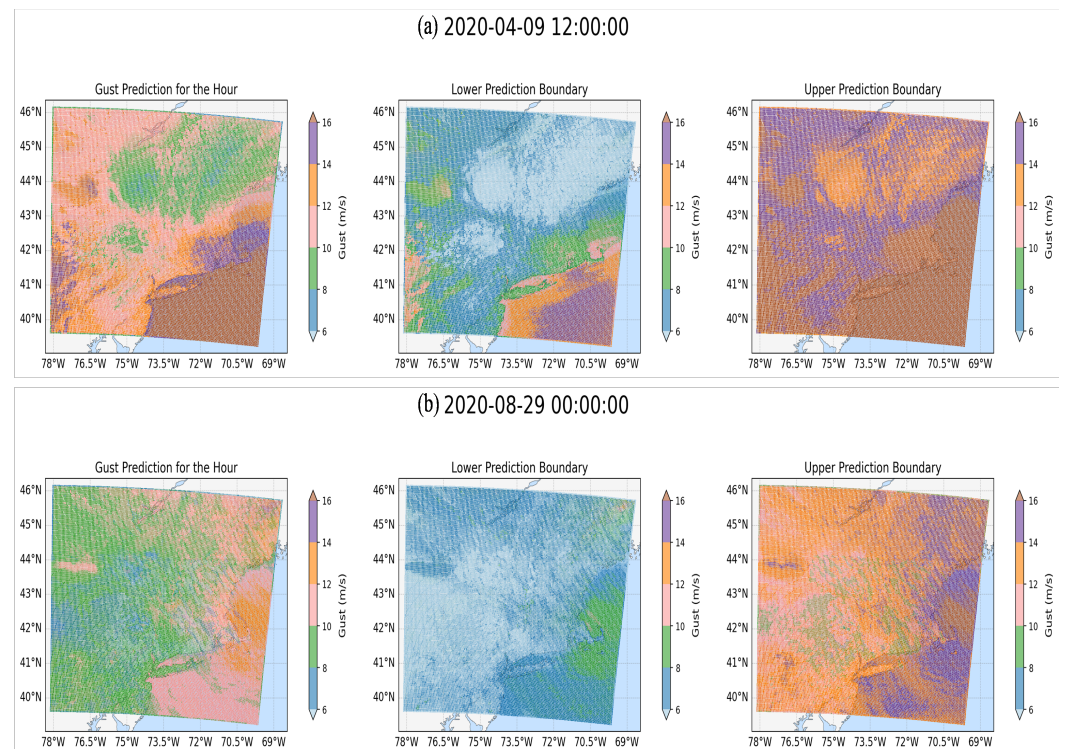
**Table 4.** PICP distribution statistics at the 95% confidence level.

| PICP Range                     | Number of Stations | Proportion | Description        |
|--------------------------------|--------------------|------------|--------------------|
| 1.0                            | 7                  | 2.6%       | Perfect Level      |
| $0.95 \leq \text{PICP} < 1.0$  | 178                | 66.9%      | Good Level         |
| $0.90 \leq \text{PICP} < 0.95$ | 51                 | 6.4%       | Insufficient Level |
| $0.85 \leq \text{PICP} < 0.90$ | 17                 | 6.4%       | Poor Level         |
| $\text{PICP} < 0.85$           | 13                 | 4.9%       | Very Poor Level    |



**Figure 6.** Spatial distribution of the PICP at the station level under the 95% confidence level. Each marker represents an observation station within the study area, and colors indicate the PICP values computed from the 11 test storms.

The prediction results of the PISA-Net model and its spatial distribution of uncertainty are shown in Figure 7, which presents the predicted gusts for the two storms of 9 April 2020 and 29 August 2020 and their corresponding upper and lower bounds of the prediction intervals, respectively. In the 9 April 2020 storm (left column), the spatial distribution of model-predicted gusts shows strong regional differences, while the range of the prediction intervals is significantly wider in some regions, indicating higher prediction uncertainty in these regions. For example, in parts of Connecticut, the forecast gusts are about 9–11 m/s, while their corresponding lower and upper bounds reach 5–7 m/s and 12–15 m/s, respectively, reflecting a larger range of possible fluctuations. In contrast, the overall range of prediction intervals is narrower and the spatial distribution is more consistent in the prediction results for the 29 August 2020 storm (right column), indicating that the model predictions for that storm have a high confidence level. The comparison of the two storms shows that the width of the prediction intervals is closely related to the intensity of the storm and the gradient of the spatial distribution, which is consistent with the previous analysis on the uncertainty influencing factors. Such spatial visualization results not only help meteorological operators to visually identify possible high-risk areas in the forecast, but also provide important information support for decision-making processes, such as disaster warning, energy dispatch and traffic control. Through the spatial representation of forecast intervals, decision makers can more effectively assess the credibility of forecasts and formulate more targeted countermeasures on this basis.



**Figure 7.** Spatial maps of wind gust predictions and corresponding 95% prediction intervals for two representative storm cases. The left, middle, and right panels show the predicted gusts, lower bounds, and upper bounds, respectively.

#### 4.4. Ablation Study

In this experiment, we conducted ablation studies on three architectural components of PISA-Net: (1) the masking mechanism for sparse stations (Masking Mechanism); (2) the Transformer-based spatial interaction module (Spatial Interaction); (3) the internal feature-level shortcut connection that propagates the initial station representations to the final prediction head (Residual Connection). It should be emphasized that the WRF-UPP physical forecast is retained in all ablation settings, as the core objective of PISA-Net is residual bias correction rather than direct wind-gust prediction. The baseline configuration corresponds to the full model that incorporates all three modules, while the other variants are obtained by removing one architectural component at a time. Table 5 summarizes the quantitative results under different configurations.

**Table 5.** Ablation study of different PISA-Net configurations and their impacts on gust prediction performance.

| Configuration       | PISA-Net | — Mask | — Spatial | — Residual |
|---------------------|----------|--------|-----------|------------|
| Spatial Interaction | ✓        | ✓      | ×         | ✓          |
| Masking Mechanism   | ✓        | ×      | ✓         | ✓          |
| Residual Connection | ✓        | ✓      | ✓         | ×          |
| MAE                 | 1.75     | 1.87   | 1.83      | 1.78       |
| RMSE                | 2.26     | 2.40   | 2.34      | 2.31       |
| Pearson             | 0.65     | 0.57   | 0.62      | 0.61       |

Notes: “—” indicates that the corresponding module is removed from the full PISA-Net architecture. ✓ and × denote the presence and absence of a component, respectively.

The results show that none of the reduced models achieves optimal performance, while the complete architecture consistently delivers the best scores across all metrics. The full PISA-Net attains the lowest RMSE (2.26 m/s) and MAE (1.75 m/s), together with the highest Pearson correlation (0.65). Removing the masking mechanism causes the largest degradation, with RMSE increasing to 2.40 m/s and the correlation dropping to 0.57. This highlights its importance for handling irregular station distributions during spatial aggregation. Excluding the spatial interaction module also leads to a clear performance decline (RMSE = 2.34 m/s; MAE = 1.83 m/s), confirming the necessity of explicit spatial modeling. Removing the residual connection has a smaller but still noticeable effect (RMSE = 2.31 m/s), indicating its contribution to stable optimization and deep feature propagation. Overall, these results demonstrate that the strong performance of PISA-Net arises from the complementary roles of data masking, spatial interaction, and feature-level shortcut connections.

## 5. Conclusions

In this paper, we propose PISA-Net, a novel deep learning model for wind gust forecasting at sparse observation stations. The model effectively addresses key challenges, including systematic bias, irregular station distributions, and insufficient uncertainty quantification. Extensive experiments on 61 extratropical cyclone events demonstrate that PISA-Net achieves the highest correlation coefficient (0.65), together with the lowest RMSE (2.26 m/s) and an MAE of 1.75 m/s. Compared with the WRF-UPP baseline, PISA-Net provides substantial bias correction. Relative to multilayer perceptron (MLP) and convolutional neural network (CNN) models, it reduces RMSE by 9.6% and 8.1% and MAE by 10.3% and 8.4%, respectively. In addition, PISA-Net is suitable for industrial deployment, enabling integration with operational forecasting systems and supporting risk-aware decisions across energy, infrastructure, and public warning applications.

Despite these encouraging results, several challenges remain. Performance degradation in complex terrain and during the most extreme gust events indicates that further improvement is required. Future research will focus on embedding stronger physical constraints to enhance generalization in poorly instrumented regions, as well as extending the framework through multi-task learning to jointly forecast gusts and related meteorological variables for comprehensive hazard assessment.

Overall, this work provides a novel and effective technical pathway for accurate and reliable wind gust post-processing. By achieving substantial improvements in both deterministic accuracy and probabilistic reliability, PISA-Net demonstrates broad application potential in operational forecasting, renewable energy management, and mitigation of weather-related risks.

**Author Contributions:** Conceptualization, M.L., C.Y. and X.L.; methodology, M.L., C.Y. and H.H.; data curation, M.L.; writing—original draft, M.L. ; writing—review and editing, M.L. and C.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (Grant No. 42471437).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used in this study are publicly accessible.

**Acknowledgments:** We are grateful to the anonymous reviewers for their precious opinions and suggestions, whose professional reviews conspicuously enhanced the quality of this paper. Furthermore,

our appreciation goes to the editorial team for their painstaking efforts and guidance throughout the publication process.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Abbreviations

The following abbreviations are used in this manuscript:

|          |                                                   |
|----------|---------------------------------------------------|
| PISA-Net | Physics-Informed Spatial Attention Network        |
| NWP      | Numerical Weather Prediction                      |
| WRF      | Weather Research Forecasting                      |
| UPP      | Unified Post Processor                            |
| ECMWF    | European Center for Medium-Range Weather Forecast |
| PICP     | Prediction Interval Coverage Probability          |
| SDG 7    | Affordable and Clean Energy                       |
| SDG 9    | Industry, Innovation and Infrastructure           |
| SDG 11   | Sustainable Cities and Communities                |

## References

1. Kahl, J.D. Forecasting peak wind gusts using meteorologically stratified gust factors and MOS guidance. *Weather Forecast.* **2020**, *35*, 1129–1143. [\[CrossRef\]](#)
2. Wang, H.; Zhang, Y.M.; Mao, J.X. Sparse Gaussian process regression for multi-step ahead forecasting of wind gusts combining numerical weather predictions and on-site measurements. *J. Wind. Eng. Ind. Aerodyn.* **2022**, *220*, 104873. [\[CrossRef\]](#)
3. Leclerc, A.; Koch, E.; Feldmann, M.; Nerini, D.; Beucler, T. Improving predictions of convective storm wind gusts through statistical post-processing of neural weather models. *npj Nat. Hazards* **2025**, *2*, 100. [\[CrossRef\]](#)
4. Schoen, J.M.; Ashley, W.S. A climatology of fatal convective wind events by storm type. *Weather Forecast.* **2011**, *26*, 109–121. [\[CrossRef\]](#)
5. SDG Fund. *Sustainable Development Goals; The Energy Progress Report, Tracking SDG*; SDG Fund: New York, NY, USA, 2019; Volume 7, pp. 2–6.
6. Jahan, I.; Schreck, J.S.; Gagne, D.J.; Becker, C.; Astitha, M. Uncertainty Quantification of Wind Gust Predictions in the Northeast United States: An Evidential Neural Network and Explainable Artificial Intelligence Approach. *Environ. Model. Softw.* **2025**, *193*, 106595. [\[CrossRef\]](#)
7. Schulz, B.; Lerch, S. Machine learning methods for postprocessing ensemble forecasts of wind gusts: A systematic comparison. *Mon. Weather Rev.* **2022**, *150*, 235–257. [\[CrossRef\]](#)
8. Lopes, F.M.; Silva, H.G.; Salgado, R.; Cavaco, A.; Canhoto, P.; Collares-Pereira, M. Short-term forecasts of GHI and DNI for solar energy systems operation: Assessment of the ECMWF integrated forecasting system in southern Portugal. *Sol. Energy* **2018**, *170*, 14–30. [\[CrossRef\]](#)
9. Fovell, R.G.; Cao, Y. 5A. 2 Wind and Gust Forecasting in Complex Terrain. In Proceedings of the 15th Annual WRF Users' Workshop, Boulder, CO, USA, 23–27 June 2014.
10. Liu, G.; Wang, X.; Wu, Q.; Fang, D.; Wu, Z.; Liu, H.; Lyu, M. Effect of urbanization on gust wind speed in summer over an urban area in Eastern China. *Environ. Res. Lett.* **2023**, *18*, 074025. [\[CrossRef\]](#)
11. Qiu, L.; Howland, M.F. Quantitative evaluation of reanalysis and numerical weather prediction models for wind resource assessment in offshore environments. *J. Geophys. Res. Atmos.* **2025**, *130*, e2025JD043490. [\[CrossRef\]](#)
12. Jahan, I.; Cerrai, D.; Astitha, M. Storm gust prediction with the integration of machine learning algorithms and WRF model variables for the Northeast United States. *Artif. Intell. Earth Syst.* **2024**, *3*, e230047. [\[CrossRef\]](#)
13. Frnda, J.; Durica, M.; Rozhon, J.; Vojtekova, M.; Nedoma, J.; Martinek, R. ECMWF short-term prediction accuracy improvement by deep learning. *Sci. Rep.* **2022**, *12*, 7898. [\[CrossRef\]](#)
14. Hakimi, A.; Ghafarian, P. Simultaneous prediction of 10m and 100m wind speeds using a model inspired by brain emotional learning. *Sci. Rep.* **2025**, *15*, 4304. [\[CrossRef\]](#)
15. Zhou, K.; Zheng, Y.; Li, B.; Dong, W.; Zhang, X. Forecasting different types of convective weather: A deep learning approach. *J. Meteorol. Res.* **2019**, *33*, 797–809. [\[CrossRef\]](#)
16. Zhong, X.; Du, F.; Chen, L.; Wang, Z.; Li, H. Investigating transformer-based models for spatial downscaling and correcting biases of near-surface temperature and wind-speed forecasts. *Q. J. R. Meteorol. Soc.* **2024**, *150*, 275–289. [\[CrossRef\]](#)

17. Cheng, W.Y.Y.; Steenburgh, W.J. Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States. *Weather Forecast.* **2007**, *22*, 1304–1318. [[CrossRef](#)]
18. Del, I.V.; Starchenko, A.V. Correction of Numerical Forecasts of Wind Gusts Using Artificial Neural Networks and Observations. *Atmos. Ocean. Opt.* **2025**, *38*, 473–480. [[CrossRef](#)]
19. Coburn, J.; Pryor, S.C. Do machine learning approaches offer skill improvement for short-term forecasting of wind gust occurrence and magnitude? *Weather Forecast.* **2022**, *37*, 525–543. [[CrossRef](#)]
20. Aslam, L.; Zou, R.; Awan, E.; Butt, S.A. Integrating Physics-Informed Vectors for Improved Wind Speed Forecasting with Neural Networks. In Proceedings of the 14th Asian Control Conference (ASCC), Dalian, China, 5–8 July 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 1902–1907.
21. Bentsen, L.Ø.; Warakagoda, N.D.; Stenbro, R.; Engelstad, P. Spatio-temporal wind speed forecasting using graph networks and novel Transformer architectures. *Appl. Energy* **2023**, *333*, 120565. [[CrossRef](#)]
22. Potosnak, W.; Challu, C.; Olivares, K.G.; Miller, J.K.; Dubrawski, A. Severe Wind Event Prediction with Multi-variate Physics-Informed Deep Learning. In Proceedings of the ICLR 2024 Workshop: Tackling Climate Change with Machine Learning, Vienna, Austria, 7–11 May 2024.
23. El Rafei, M.; Sherwood, S.; Evans, J.P.; Ji, F. Analysis of extreme wind gusts using a high-resolution Australian Regional Reanalysis. *Weather Clim. Extrem.* **2023**, *39*, 100537. [[CrossRef](#)]
24. Cai, H.; Jia, X.; Feng, J.; Li, W.; Hsu, Y.-M.; Lee, J. Gaussian process regression for numerical wind speed prediction enhancement. *Renew. Energy* **2020**, *146*, 2112–2123. [[CrossRef](#)]
25. Eskandari, H.; Imani, M.; Moghaddam, M.P. A deep residual network integrating entropy-based wavelet packet ensemble model for short-term electrical load forecasting. *Energy* **2025**, *314*, 134168. [[CrossRef](#)]
26. Hondru, V.; Croitoru, F.A.; Minaee, S.; Ionescu, R.T.; Sebe, N. Masked image modeling: A survey. *Int. J. Comput. Vis.* **2025**, *133*, 7154–7200. [[CrossRef](#)]
27. Sheridan, P. Current gust forecasting techniques, developments and challenges. *Adv. Sci. Res.* **2018**, *15*, 159–172. [[CrossRef](#)]
28. Brotzge, J.A.; Wang, J.; Thorncroft, C.D.; Joseph, E.; Bain, N.; Bassill, N.; Farruggio, N.; Freedman, J.M.; Hemker, K., Jr.; Johnston, D.; et al. A technical overview of the New York State Mesonet standard network. *J. Atmos. Ocean. Technol.* **2020**, *37*, 1827–1845. [[CrossRef](#)]
29. Chisale, S.W.; Lee, H.S.; Calvo, M.A.S.; Jeong, J.S.; Aljber, M.; Williams, Z.; Cabrera, J.S. Advanced solar energy potential assessment in Malawi: Utilizing high-resolution WRF model and GIS to identify optimal sites for solar PV generation. *Renew. Energy* **2025**, *239*, 122084. [[CrossRef](#)]
30. Merino, A.; García-Ortega, E.; Navarro, A.; Sánchez, J.L.; Tapiador, F.J. WRF hourly evaluation for extreme precipitation events. *Atmos. Res.* **2022**, *274*, 106215. [[CrossRef](#)]
31. Amini, A.; Schwarting, W.; Soleimany, A.; Rus, D. Deep evidential regression. In Proceedings of the 34th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 6–12 December 2020; Volume 33, pp. 14927–14937.
32. Tran, Q.K.; Song, S. Computer vision in precipitation nowcasting: Applying image quality assessment metrics for training deep neural networks. *Atmosphere* **2019**, *10*, 244. [[CrossRef](#)]
33. Veillette, M.; Samsi, S.; Mattioli, C. Sevir: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. In Proceedings of the 34th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 6–12 December 2020; Volume 33, pp. 22009–22019.
34. Franch, G.; Nerini, D.; Pendesini, M.; Coviello, L.; Jurman, G.; Furlanello, C. Precipitation nowcasting with orographic enhanced stacked generalization: Improving deep learning predictions on extreme events. *Atmosphere* **2020**, *11*, 267. [[CrossRef](#)]
35. Sluijterman, L.; Cator, E.; Heskes, T. How to evaluate uncertainty estimates in machine learning for regression? *Neural Netw.* **2024**, *173*, 106203. [[CrossRef](#)] [[PubMed](#)]
36. Coskun, A. Prediction interval: A powerful statistical tool for monitoring patients and analytical systems. *Biochem. Med.* **2024**, *34*, 020101. [[CrossRef](#)]
37. Ramachandran, K.M.; Tsokos, C.P. *Mathematical Statistics with Applications in R*; Academic Press: Cambridge, MA, USA, 2020.
38. Naskath, J.; Sivakamasundari, G.; Begum, A.A.S. A study on different deep learning algorithms used in deep neural nets: MLP SOM and DBN. *Wirel. Pers. Commun.* **2023**, *128*, 2913–2936. [[CrossRef](#)] [[PubMed](#)]
39. Wang, A.; Chen, H.; Lin, Z.; Han, J.; Ding, G. Repvit: Revisiting mobile cnn from vit perspective. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 17–21 June 2024; pp. 15909–15920.
40. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P.S. A Comprehensive Survey on Graph Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4–24. [[CrossRef](#)] [[PubMed](#)]
41. Lee, J.; Lee, Y.; Kim, J.; Kosiorek, A.R.; Teh, Y.W. Set Transformer: A Framework for Attention-based Permutation-Invariant Neural Networks. In Proceedings of the 36th International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019.
42. Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Li, H. T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Trans. Intell. Transp. Syst.* **2019**, *21*, 3848–3858. [[CrossRef](#)]

43. Wu, Z.; Pan, S.; Long, G.; Jiang, J.; Zhang, C. Graph WaveNet for Deep Spatial-Temporal Graph Modeling. *arXiv* **2019**, arXiv:1906.00121.
44. Li, S.; Chen, M.; Yi, L.; Lu, Q.; Yang, H. MIESTC: A Multivariable Spatio-Temporal Model for Accurate Short-Term Wind Speed Forecasting. *Atmosphere* **2025**, *16*, 67. [[CrossRef](#)]
45. Williams, J.; Hill, R.R.; Pignatiello, J.J., Jr.; Chicken, E. Wavelet analysis of variance box plot. *J. Appl. Stat.* **2022**, *49*, 3536–3563. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.