Coarse-to-fine Physics-guided Mixture-of-experts for Urban Wind Field Reconstruction from a Single Radar Column

Anonymous authors
Paper under double-blind review

Abstract

Accurate low-altitude urban wind fields are essential for UAV safety, pollutant dispersion assessment, and wind analysis around buildings. However, reconstructing high-resolution 3D urban wind fields from a single vertical radar column is highly challenging due to extremely sparse observations. Existing methods often rely on distributed sensors or multiple LiDARs, while direct sparse-to-dense mapping entangles large-scale wind field completion with fine-scale structure recovery, and shared representations struggle to capture diverse local flow patterns. To address these challenges, we propose a coarse-to-fine urban wind field reconstruction framework that decomposes the task into low-resolution wind field completion and multi-stage super-resolution reconstruction, progressively bridging the large spatial gap between sparse observations and dense targets. We further introduce a physics-guided mixture-of-experts module with different physical priors that encourages different experts to specialize in distinct local flow patterns, including general air, building boundaries, vortices, high-speed regions, and ordinary areas. To mitigate the mismatch between ideal and predicted intermediate states during cascaded reconstruction, we develop a momentum-energy-guided state selection mechanism that prioritizes representative predicted states according to their momentum and kinetic-energy discrepancies. Our experiments achieve **0.1408 MAE** and **0.0404 MSE**, reducing these errors by **46.00%** and **70.19%** over the state-of-the-art method, respectively, demonstrating the effectiveness of our proposed method.Code is available at https://anonymous.4open.science/r/wind-reconstruction-anonymous-4C75/

1 Introduction

Accurate perception of low-altitude urban wind fields is fundamental to safe UAV operation, pollutant dispersion assessment, and building wind-environment analysis (Achermann et al., 2024; Blocken, 2014). Flow separation, reattachment, and interactions induced by building clusters produce pronounced spatial heterogeneity, including complex wakes, shear layers, and localized vortices (Tominaga et al., 2008; Blocken, 2014; Cao et al., 2025). Meanwhile, equipment cost, sensor placement, and line-of-sight occlusion limit practical observations to only a small portion of the spatial domain (Gao et al., 2024; Ke et al., 2026; Jiang et al., 2026). Reconstructing a complete, high-resolution urban wind field from extremely sparse observations provided by a single vertical radar column therefore remains highly challenging.

Deep generative approaches provide new tools for this problem (Hu et al., 2024; Cao et al., 2025; Zhou et al., 2026). Early studies used reduced-order methods such as POD-LSE to reconstruct urban wind fields from limited sensor measurements, but their representation of nonlinear flow structures was limited (Hu et al., 2022). Generative models such as GANs were subsequently introduced to improve the reconstruction of building wakes and local turbulent structures (Zhang et al., 2022; Hu et al., 2024). Gao et al. combined graph neural networks, autoencoders, and physical constraints to reconstruct high-resolution urban wind fields from sparse sensors, while Liu et al. used sparse measurements to correct

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

CFD predictions in a two-stage framework (Gao et al., 2024; Liu et al., 2024). Gao et al. (2024) further combined graph neural networks, autoencoders, and physical constraints to recover high-resolution urban wind fields from sparse sensor measurements, while Liu et al. (2024) adopted a two-stage framework that uses sparse field measurements to correct CFD-based predictions. More recently, Cao et al. (2025) enhanced urban canopy turbulence reconstruction through pattern-constrained generative learning. Ke et al. (2026) proposed GFNO, which maps irregular sparse observations onto latent grids before reconstruction with Fourier neural operators. Jiang et al. (2026) incorporated physics-informed learning to reconstruct continuous three-dimensional urban wind fields from sparse Doppler LiDAR observations.

However, existing methods still largely rely on distributed sensors or multiple LiDARs (Gao et al., 2024; Cao et al., 2025; Ke et al., 2026; Jiang et al., 2026). Reconstructing a high-resolution 3D urban wind field from a single extremely sparse radar remains challenging for two reasons. **First, a substantial spatial-scale gap exists between sparse observations and the high-resolution target.** Direct reconstruction requires simultaneous completion of large unobserved regions and recovery of fine-scale structures, coupling global information propagation with local detail recovery. **Second, urban wind fields exhibit pronounced spatial heterogeneity.** General air regions, ordinary areas, building boundaries, vortices, and high-speed regions follow distinct local flow characteristics that are difficult to model with a single shared representation.

To address the first challenge, we propose a **coarse-to-fine urban wind field reconstruction framework** that decomposes the task into low-resolution completion and multi-stage super-resolution reconstruction, separating global flow propagation from local structural refinement. To address the second challenge, we introduce a **physics-guided mixture-of-experts module**, in which local physical experts model general air regions, ordinary areas, building boundaries, vortices, and high-speed regions using different physical priors and complementary residual corrections. We further propose a **momentum–energy-guided state selection mechanism** that selects predicted intermediate states according to momentum and energy discrepancies between ideal and predicted states, improving adaptation to cascaded inputs and suppressing error accumulation.

Experimental results show that our method outperforms the state-of-the-art method, achieving **0.1408 MAE** and **0.0404 MSE**, with relative reductions of **46.00%** and **70.19%**, respectively. Ablation studies confirm consistent gains from each component.

The main contributions are:

- We propose a **coarse-to-fine urban wind field reconstruction framework** that progressively bridges the spatial gap between extremely sparse single-radar observations and high-resolution 3D targets.
- We introduce a **physics-guided mixture-of-experts module** for differentiated modeling of heterogeneous local flow structures.
- We propose a **momentum–energy-guided state selection mechanism** that selects intermediate states according to momentum and energy discrepancies, mitigating state-distribution mismatch and error accumulation in cascaded reconstruction.

2 Related Work

2.1 Physics-Informed Neural Networks

Physics-informed neural networks (PINNs) incorporate PDE residuals, boundary conditions, and observations into a unified objective for data-scarce forward and inverse problems (Raissi et al., 2019; Karniadakis et al., 2021). Variational and domain-decomposition methods, including VPINN, cPINN, and XPINN, improve the treatment of complex domains and multiscale physics (Kharazmi et al., 2021; Jagtap et al., 2020; Jagtap & Karniadakis, 2020). Adaptive weighting and curriculum strategies have also been introduced to alleviate gradient imbalance and optimization difficulties (Wang et al., 2021; Krishnapriyan et al., 2021).

In wind-field reconstruction, PINNs have assimilated sparse LiDAR or mean-velocity observations with governing equations to recover atmospheric and separated flows (Chen et al., 2025; Patel et al., 2024). However, global residual constraints may produce competing gradients among distinct urban-flow regions. We instead use computable local physical priors to define expert responsibilities and embed physics into local feature representations.

2.2 Wind-Field Reconstruction and Prediction Models

Convolutional encoder–decoders capture multiscale spatial features, while super-resolution models recover turbulent structures from coarse grids (Ronneberger et al., 2015; Fukami et al., 2019). WindSeer and related methods enable rapid three-dimensional wind prediction from geometry and sparse observations in complex environments (Achermann et al., 2024; Gao et al., 2023; Zhou et al., 2026). Operator-learning models, such as FNO and Transolver, learn mappings between function spaces for PDEs and complex geometries (Li et al., 2021; Kovachki et al., 2023; Wu et al., 2024). Recent studies further explore geometric, patch-based, and probabilistic operators (Ke et al., 2026; Jiang et al., 2026; Zhang et al., 2026; Oommen et al., 2026). Most existing methods, however, directly predict a single resolution or assume ideal intermediate inputs. We separate low-resolution completion from multistage super-resolution reconstruction to progressively recover global structures and building-scale details.

2.3 Mixture-of-Experts Models

Mixture-of-experts (MoE) models route inputs to specialized subnetworks for conditional computation (Shazeer et al., 2017). GShard and Switch Transformer improve scalability and routing stability, while Expert Choice Routing and shared experts enhance capacity allocation and reduce redundancy (Lepikhin et al., 2021; Fedus et al., 2022; Zhou et al., 2022; Dai et al., 2024). Nevertheless, conventional expert specialization is learned implicitly, so expert semantics may drift or routing may concentrate in spatially continuous fields. Our physics-guided module combines prior physics experts with self-learning experts, enabling explicitly defined physical responsibilities and data-driven residual corrections to jointly model heterogeneous local flows.

3 Problem Definition

Given a dataset containing N samples, $\mathcal{D} = \{(\mathbf{L}_n, \mathbf{G}_n, \mathbf{U}_n)\}_{n=1}^{N}$, where $\mathbf{L}_n$ denotes the directly acquired sparse velocity observations from a single LiDAR together with their validity mask, $\mathbf{G}_n$ represents the building geometry and inflow conditions, and $\mathbf{U}_n$ denotes the ground-truth three-component wind field at the native target resolution. For each sample, a single LiDAR is deployed at a fixed horizontal location and provides local velocity observations at that location.

Given the sparse LiDAR observations $\mathbf{L}_n$ and the building-geometry condition $\mathbf{G}_n$, the three-dimensional wind-field reconstruction problem is formulated as

$$\widehat{\mathbf{U}}_n = F_{\mathrm{3DRecon}}\left(\mathbf{L}_n, \mathbf{G}_n\right), \tag{1}$$

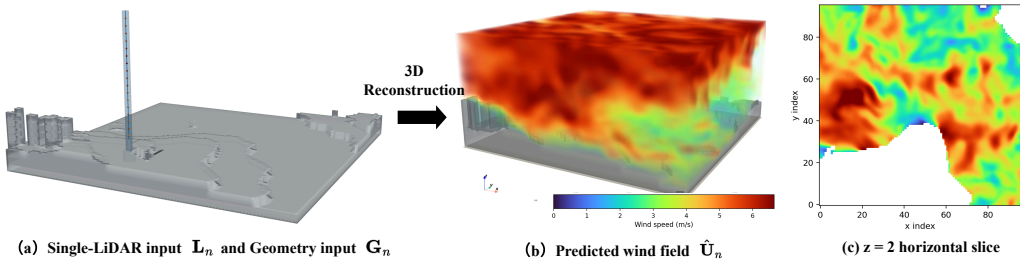
where F_{3DRecon} denotes the learned three-dimensional wind-field reconstruction model, and $\widehat{\mathbf{U}}_n$ is the complete three-dimensional wind field predicted by the model.

Let $\mathcal{A}_n$ denote the set of valid air voxels in sample n. Given the predicted wind field $\widehat{\mathbf{U}}_n$ and the ground-truth wind field $\mathbf{U}_n$, the voxel-wise L_1 reconstruction error over the air domain is defined as

$$\ell_{\mathrm{air}}^{(1)}\left(\widehat{\mathbf{U}}_n, \mathbf{U}_n; \mathcal{A}_n\right) = \frac{1}{3|\mathcal{A}_n|} \sum_{v \in \mathcal{A}_n} \left\| \widehat{\mathbf{U}}_{n,:,v} - \mathbf{U}_{n,:,v} \right\|_1 . \tag{2}$$

Based on this per-sample error, the global reconstruction loss over the dataset is defined as

$$\mathcal{L}_{\mathrm{MAE}} = \frac{1}{N} \sum_{n=1}^{N} \ell_{\mathrm{air}}^{(1)}\left(\widehat{\mathbf{U}}_n, \mathbf{U}_n; \mathcal{A}_n\right). \tag{3}$$



(a) Single-LiDAR input $\mathbf{L}_n$ and Geometry input $\mathbf{G}_n$ (b) Predicted wind field $\hat{\mathbf{U}}_n$ (c) z = 2 horizontal slice

Figure 1: Directly acquired sparse observations from a single LiDAR, the building-geometry input, the reconstructed three-dimensional wind field, and its horizontal slices.

This metric measures the mean absolute error between the predicted and ground-truth wind fields across all three velocity components within the air domain.

4 Method

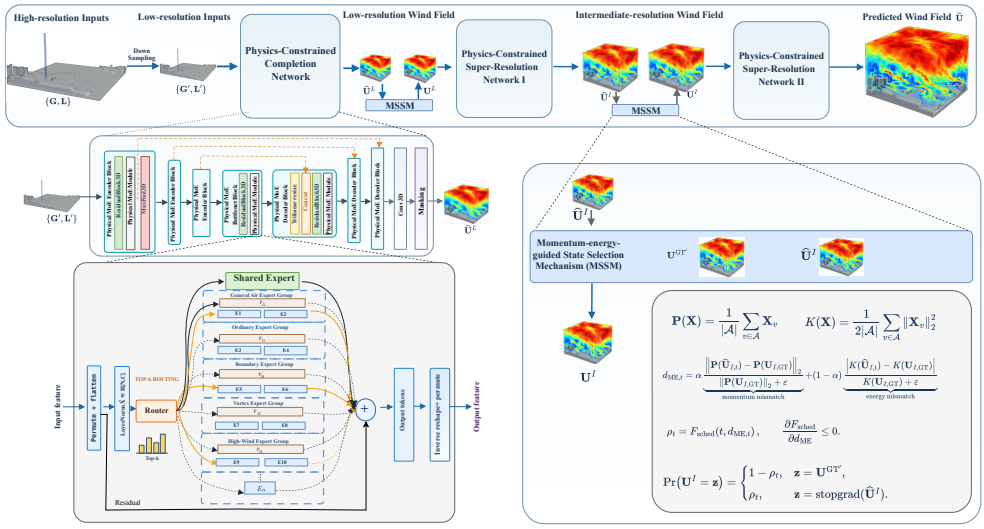


Figure 2: Coarse-to-fine urban wind field reconstruction framework: low-resolution wind field completion, multi-stage super-resolution reconstruction, the physics-guided mixture-of-experts module, and the momentum–energy-guided sampling mechanism.

4.1 Coarse-to-Fine Urban Wind Field Reconstruction Framework

As shown in Fig. 2, we construct a coarse-to-fine urban wind field reconstruction framework comprising low-resolution completion and two-stage super-resolution reconstruction. The input $\mathbf{L}$ consists of directly acquired sparse velocity observations from a single radar and their validity mask, and $\mathbf{G}$ represents the building geometry and inflow conditions. First, the physics-constrained completion network F_{PCCN} directly predicts a complete low-resolution wind field from $\mathbf{L}$ and $\mathbf{G}$:

$$\widehat{\mathbf{U}}^{L} = F_{\mathrm{PCCN}}(\mathbf{L}, \mathbf{G}). \tag{4}$$

Subsequently, the two physics-constrained super-resolution networks $F^{(1)}_{\mathrm{PCSRN}}$ and $F^{(2)}_{\mathrm{PCSRN}}$ successively upscale the low-resolution wind field to the intermediate resolution and the native target resolution:

$$\widehat{\mathbf{U}}^{I} = F^{(1)}_{\mathrm{PCSRN}}(\widehat{\mathbf{U}}^{L}, \mathbf{G}), \qquad \widehat{\mathbf{U}} = F^{(2)}_{\mathrm{PCSRN}}(\widehat{\mathbf{U}}^{I}, \mathbf{G}). \tag{5}$$

Here, $\widehat{\mathbf{U}}^L$, $\widehat{\mathbf{U}}^I$, and $\widehat{\mathbf{U}}$ denote the low-resolution prediction, intermediate-resolution prediction, and final prediction, respectively. The three stages are supervised using the low-resolution ground-truth field $\mathbf{U}^L$, intermediate-resolution ground-truth field $\mathbf{U}^I$, and native target ground-truth field $\mathbf{U}$, respectively.Let $\mathcal{S} = \{L, I, F\}$ denote the three reconstruction scales, with $\widehat{\mathbf{U}}^F = \widehat{\mathbf{U}}$, $\mathbf{U}^F = \mathbf{U}$, and $\mathcal{A}_n^F = \mathcal{A}_n$. The air-domain reconstruction loss at stage $s \in \mathcal{S}$ is defined as

$$\mathcal{L}_{\mathrm{MAE}}^{(s)} = \frac{1}{N} \sum_{n=1}^{N} \ell_{\mathrm{air}}^{(1)} \left(\widehat{\mathbf{U}}_n^s, \mathbf{U}_n^s; \mathcal{A}_n^s \right), \qquad s \in \mathcal{S}. \tag{6}$$

The three networks are optimized separately in the order of low-resolution completion, first-stage super-resolution, and second-stage super-resolution. When $s = F$, Eq. (6) equals the global reconstruction loss; for any stage, $\mathcal{L}_{\mathrm{MAE}}^{(s)}$ is denoted by $\mathcal{L}_{\mathrm{MAE}}$ in its joint objective.

$\widehat{\mathbf{U}}^L$, $\widehat{\mathbf{U}}^I$, and $\widehat{\mathbf{U}}$ denote the low-resolution, intermediate-resolution, and final predictions, supervised by $\mathbf{U}^L$, $\mathbf{U}^I$, and $\mathbf{U}$, respectively, using Eq. (6). During training, the mechanism in Section 4.3 determines whether each network receives the corresponding ground-truth state or a stop-gradient prediction; during validation and inference, the super-resolution networks sequentially receive $\widehat{\mathbf{U}}^L$ and $\widehat{\mathbf{U}}^I$. The upstream network recovers global low-frequency structures, while the downstream networks refine velocity gradients, building wakes, and small-scale vortices, with multi-scale supervision preventing deferred errors. Measurements in $\mathbf{L}$ and their validity mask are imposed as constraints, solid-region velocities are set to zero, and no radar input is synthesized from $\mathbf{U}$, $\mathbf{U}^L$, or $\mathbf{U}^I$. All stages share the same topology and expert design but use independent parameters: F_{PCCN} recovers large-scale structures, while the two F_{PCSRN} networks restore fine-scale variations and provide intermediate states for physics-guided sampling.

4.2 Physics-Guided Mixture-of-Experts Module

Physics-guided mixture-of-experts modules are deployed at three encoder locations, the bottleneck, and three decoder locations to model diverse local flow conditions. For the general air domain, regular flows, building boundaries, vortices, and high-wind-speed regions, each module combines a prior physics expert with predefined responsibilities and physics-based supervision with a self-learning expert that captures residual, geometric, and local flow patterns not explicitly described by physical formulas.

General Air Expert. The General Air Expert operates over the valid air domain $\mathcal{A}$, where $\mathbf{u}$ and $\widehat{\mathbf{u}}$ denote the ground-truth and predicted three-dimensional velocities, respectively. It uses the continuity prior $\nabla \cdot \mathbf{u} = 0$ and matches the predicted divergence to the ground-truth divergence:

$$\begin{gathered} \nabla \cdot \mathbf{u} = 0, \\ \mathcal{L}_G = \lambda_{\mathrm{aff}} \mathcal{L}_{\mathrm{aff},G} + \lambda_{\mathrm{align}} \mathcal{L}_{\mathrm{align},G} + \lambda_{G,r} \left\langle \|\widehat{\mathbf{u}} - \mathbf{u}\|_1 \right\rangle_{\mathcal{A}} + \lambda_{G,d} \left\langle |\nabla \cdot \widehat{\mathbf{u}} - \nabla \cdot \mathbf{u}| \right\rangle_{\mathcal{A}}. \end{gathered} \tag{6}$$

Here, $\langle \cdot \rangle_{\mathcal{R}}$ denotes averaging over region $\mathcal{R}$. $\mathcal{L}_{\mathrm{aff},G}$ and $\mathcal{L}_{\mathrm{align},G}$ are the regional affinity and response-consistency losses, while the last two terms measure velocity reconstruction and divergence consistency over $\mathcal{A}$.

Building- Boundary Expert. The Building- Boundary Expert operates on region $\mathcal{B}$, consisting of air voxels adjacent to buildings. A viscous fluid satisfies the no-slip condition on a stationary solid surface, such that the wall velocity is zero. Accordingly, an air voxel i is assigned to $\mathcal{B}$ if its horizontal neighborhood $\mathcal{N}_{xy}(i)$ intersects the non-air domain $\mathcal{A}^c$. Let $\partial\Omega_s$ denote the building surface and ∇_B the boundary difference operator along the air–solid adjacency direction. After setting the solid-side velocity to zero, the discrete boundary gradient is

$$\nabla_B \mathbf{u}_i = \frac{\mathbf{u}_i - 0}{\Delta n_i}. \tag{7}$$

Here, Δn_i is the distance from the air-voxel center to the adjacent solid surface, and the difference describes the velocity transition from the stationary wall to the near-wall air voxel. The composite loss is

$$\mathbf{u}|_{\partial\Omega_s} = \mathbf{0}, \mathcal{L}_B = \mathcal{L}_{\text{self},B} + \lambda_{B,r} \langle |\widehat{\mathbf{u}} - \mathbf{u}|_1 \rangle_{\mathcal{B}} + \lambda_{B,g} \langle |\nabla_B \widehat{\mathbf{u}} - \nabla_B \mathbf{u}|_1 \rangle_{\mathcal{B}} . \tag{8}$$

Because discrete air voxels lie outside the solid surface, their velocities are not forced to zero; instead, the boundary velocity and adjacency-direction gradient are matched. The prior physics expert corrects near-wall velocity gradients, while the self-learning physics expert captures local variations caused by building geometry and inflow conditions.

Vortices Expert. The Vortices Expert operates on region $\mathcal{V}$, comprising non-boundary air voxels whose vorticity magnitude exceeds the sample-quantile threshold κ_ω. With the three-dimensional vorticity defined as $\boldsymbol{\omega} = \nabla \times \mathbf{u}$, its composite loss is

$$\begin{aligned} \boldsymbol{\omega} &= \nabla \times \mathbf{u}, \\ \mathcal{L}_V &= \mathcal{L}_{\text{self},V} + \lambda_{V,r} \langle \|\widehat{\mathbf{u}} - \mathbf{u}\|_1 \rangle_{\mathcal{V}} + \lambda_{V,\omega} \langle \|\nabla \times \widehat{\mathbf{u}} - \nabla \times \mathbf{u}\|_1 \rangle_{\mathcal{V}} . \end{aligned} \tag{9}$$

High-speed Expert. The High-speed Expert operates on region $\mathcal{H}$, where the velocity magnitude exceeds the sample-wise quantile threshold κ_s, typically covering narrow passages, corner-acceleration zones, and inflow-convergence regions. Because component-wise errors may not preserve velocity magnitude and transport direction, the expert constrains them separately. With ε denoting a numerical-stability constant, the magnitude, unit direction, directional error, and composite loss are defined as

$$s(\mathbf{u}) = |\mathbf{u}|_2, \qquad \mathbf{d}(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|_2 + \varepsilon}, \qquad D_{\text{dir}}(\widehat{\mathbf{u}}, \mathbf{u}) \quad = 1 - \frac{\widehat{\mathbf{u}}^\top \mathbf{u}}{|\widehat{\mathbf{u}}|_2 |\mathbf{u}|_2 + \varepsilon}. \tag{10}$$

$$\mathcal{L}_H = \mathcal{L}_{\text{self},H} + \lambda_{H,r} \langle |\widehat{\mathbf{u}} - \mathbf{u}|_1 \rangle_{\mathcal{H}} + \lambda_{H,s} \langle ||\widehat{\mathbf{u}}|_2 - |\mathbf{u}|_2| \rangle_{\mathcal{H}} + \lambda_{H,d} \langle D_{\text{dir}} (\widehat{\mathbf{u}}, \mathbf{u}) \rangle_{\mathcal{H}} . \tag{11}$$

This loss corrects magnitude attenuation and directional deflection in high-speed passages while capturing geometry- and inflow-induced acceleration patterns.

paragraphOrdinary Expert. The Ordinary Expert operates on region $\mathcal{O}$, excluding building-boundary, strong-vortex, and high-wind-speed voxels. Its three-dimensional velocity-consistency prior and composite loss are

$$\widehat{\mathbf{u}}_i \to \mathbf{u}_i = (u_{x,i}, u_{y,i}, u_{z,i})^\top, \quad i \in \mathcal{O}, \qquad \mathcal{L}_O = \mathcal{L}_{\text{self},O} + \lambda_{O,r} \langle \|\widehat{\mathbf{u}} - \mathbf{u}\|_1 \rangle_{\mathcal{O}} . \tag{12}$$

Here, $\mathcal{L}_{\text{self},O}$ comprises the regional affinity and response-alignment constraints.

After learning the regional responsibilities, the Router selects four experts for each voxel: one prior physics expert and three self-learning physics experts. For voxel feature $\mathbf{h}_i^\ell$ in module ℓ, the Router first computes normalized candidate scores:

$$\mathbf{s}_i^\ell = \text{softmax}\big(R^\ell(\mathbf{h}_i^\ell)\big) . \tag{13}$$

The experts are then selected separately from the prior and self-learning sets:

$$\mathcal{I}_i^\ell = \left\{ \underset{e \in \mathcal{E}_{\text{prior}}}{\arg\max}\, s_{i,e}^\ell \right\} \cup \text{TopK}\big(\{ s_{i,e}^\ell : e \in \mathcal{E}_{\text{self}} \}, 3\big) , \qquad |\mathcal{I}_i^\ell| = 4. \tag{14}$$

Here, $R^\ell(\cdot)$ is the learnable Router, $\mathcal{E}_{\text{prior}}$ and $\mathcal{E}_{\text{self}}$ are the prior- and self-learning-expert sets, and $\mathcal{I}_i^\ell$ is the selected index set. Regional supervision is applied during training, whereas inference selects experts using only input features. The selected expert outputs are weighted by the Router scores and added to the shared-channel output and backbone features:

$$\mathbf{h}_i^{\ell+1} = \mathbf{h}_i^\ell + E_\ell^{\text{shared}}\big(\mathbf{h}_i^\ell\big) + \sum_{e \in \mathcal{I}_i^\ell} s_{i,e}^\ell E_{\ell,e}\big(\mathbf{h}_i^\ell\big) . \tag{15}$$

The training objective combines the global reconstruction loss with the five regional losses:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{MAE}} + \mathcal{L}_G + \mathcal{L}_O + \mathcal{L}_B + \mathcal{L}_V + \mathcal{L}_H. \tag{16}$$

Here, $\mathcal{L}_{\text{MAE}}$ is the global air-domain reconstruction loss, and $\mathcal{L}_G$, $\mathcal{L}_O$, $\mathcal{L}_B$, $\mathcal{L}_V$, and $\mathcal{L}_H$ correspond to the General Air Expert, Ordinary Expert, Building Boundaries Expert, Vortices Expert, and High-Speed Expert, respectively. This objective jointly optimizes the Router, experts, and backbone. Validation and inference use neither ground-truth region labels nor additional physical-equation solving.

4.3 Physics-Guided Sampling Mechanism

As shown in Fig. 2, we introduce a physics-guided sampling mechanism between the two physics-constrained super-resolution networks. This mechanism first calculates the momentum–energy discrepancy between the predicted and ideal intermediate states, then determines the sampling probability of the predicted state by combining this discrepancy with training progress, and finally selects the actual input to the second-stage network from the ideal and predicted states.

For the set of air voxels $\mathcal{A}$, the domain-averaged velocity momentum proxy and mean kinetic energy of state $\mathbf{X}$ are defined as

$$\mathbf{P}(\mathbf{X}) = \frac{1}{|\mathcal{A}|}\sum_{v\in\mathcal{A}} \mathbf{X}_v, \qquad K(\mathbf{X}) = \frac{1}{2|\mathcal{A}|}\sum_{v\in\mathcal{A}} \|\mathbf{X}_v\|_2^2 . \tag{17}$$

Based on these physical statistics, the momentum–energy discrepancy of the predicted intermediate state relative to the ideal intermediate state is defined as

$$d_{\mathrm{ME},t} = \alpha \frac{\left\|\mathbf{P}(\widehat{\mathbf{U}}_t^I) - \mathbf{P}(\mathbf{U}_{\mathrm{GT}}^I)\right\|_2}{\left\|\mathbf{P}(\mathbf{U}_{\mathrm{GT}}^I)\right\|_2 + \varepsilon} + (1-\alpha)\frac{\left|K(\widehat{\mathbf{U}}_t^I) - K(\mathbf{U}_{\mathrm{GT}}^I)\right|}{K(\mathbf{U}_{\mathrm{GT}}^I) + \varepsilon}. \tag{18}$$

Here, the first and second terms measure the momentum and mean-kinetic-energy discrepancies, respectively; $\alpha \in [0,1]$ weights the two terms, and $\varepsilon > 0$ ensures numerical stability. Equation (18) is evaluated for each sample, and the batch mean defines $d_{\mathrm{ME},t}$ at epoch t.

The training progress t and momentum–energy discrepancy $d_{\mathrm{ME},t}$ are input to the curriculum scheduler:

$$\rho_t = F_{\mathrm{sched}}(t, d_{\mathrm{ME},t}), \qquad \frac{\partial F_{\mathrm{sched}}(t, d_{\mathrm{ME},t})}{\partial d_{\mathrm{ME},t}} \le 0. \tag{19}$$

Here, $\rho_t \in [0,1]$ is the probability of using the predicted intermediate state $\widehat{\mathbf{U}}_t^I$ as the second-stage input. The actual input is sampled as

$$\Pr\big(\mathbf{U}_t^I = \mathbf{z}\big) = \begin{cases} 1-\rho_t, & \mathbf{z} = \mathbf{U}_{\mathrm{GT}}^I, \\ \rho_t, & \mathbf{z} = \mathrm{stopgrad}\Big(\widehat{\mathbf{U}}_t^I\Big). \end{cases} \tag{20}$$

The stop-gradient operation detaches the predicted state from the upstream computation graph. The sampled state $\mathbf{U}_t^I$ is then passed to the second-stage network in Eq. (5) to produce $\widehat{\mathbf{U}}_t$.

5 Experiments

5.1 Experimental Setup

Data and Splits The dataset contains 4,824 independent three-dimensional wind fields covering 24 inflow directions. The fixed training, validation, and test sets contain 3,864, 480, and 480 samples, respectively. Each sample is treated as an independent steady-state wind field, without temporal sequence modeling. The native ground-truth field has a resolution of $96 \times 96 \times 20$. During training, the wind fields are normalized by 30 m/s, and the physical units are restored for evaluation. Sparse observations along a single vertical radar column are read directly from the radar data, rather than generated by sampling the ground-truth field or through cross-scale coordinate mapping. The coarse-to-fine framework operates sequentially at resolutions of $24 \times 24 \times 20$, $48 \times 48 \times 20$, and the native target resolution. All methods use identical data splits, geometric inputs, sparse observations, and evaluation code.

Evaluation Metrics We report the mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and peak signal-to-noise ratio (PSNR). Lower values are better for the first three metrics, whereas a higher PSNR is better. The relative error reduction is calculated as $(E_{\mathrm{base}} - E_{\mathrm{model}})/E_{\mathrm{base}}$.

5.2 Comparison with Existing Methods

Table 1 compares conventional interpolation, reduced-order, and learning-based methods under a consistent single-radar observation protocol. The baselines are ordered chronologically by publication year, with methods from the same year ordered by MAE and Ours listed separately at the end; each method is accompanied by its original work or architectural source. Among the learning-based baselines, Mamba-UNet3D performs best across all four metrics, while PINN-DA-SA achieves an MAE of 0.323351, an RMSE of 0.453792, and a PSNR of 36.4053 dB. Our method achieves the best results on all four metrics.

Table 1: Three-dimensional wind-field reconstruction results at the target resolution.

Method	MAE↓	MSE↓	RMSE↓	PSNR↑
IDW (Shepard, 1968)	0.760971	1.133841	1.064820	28.9969
RBF (Hardy, 1971)	1.853064	6.801311	2.607932	21.2165
Gappy-POD rank 8 (Everson & Sirovich, 1995)	0.797989	1.203275	1.096939	28.7388
Nearest (Amidror, 2002)	0.872074	1.585321	1.259095	27.5413
SwinUNet3D (Liu et al., 2021)	0.408411	0.326320	0.571244	34.4060
FNO3D (Li et al., 2021)	2.048149	8.115021	2.848688	20.4495
Mamba-UNet3D (Gu & Dao, 2023)	0.260785	0.135498	0.368101	38.2231
PINN-DA-SA (Patel et al., 2024)	0.323351	0.205927	0.453792	36.4053
Transolver 3D (Wu et al., 2024)	0.527788	0.539157	0.734273	32.2253
WindSeer 3D (Achermann et al., 2024)	1.777187	6.782193	2.604264	21.2287
GFNO-inspired (Ke et al., 2026)	0.283746	0.160229	0.400286	37.4950
Conditional Fourier (Jiang et al., 2026)	0.492129	0.472088	0.687086	32.8022
Patch-GTO (Zhang et al., 2026)	0.504193	0.484975	0.696401	32.6852
Gen4Turb EDM (Oommen et al., 2026)	0.581092	0.677346	0.823010	31.2343
Rooftop ViTAE (Zhou et al., 2026)	0.590223	0.800364	0.894630	30.5096
Ours	**0.140833**	**0.040397**	**0.200990**	**43.4789**

5.3 Ablation Study: Coarse-to-Fine Framework and Physics-Guided Mixture of Experts

Effects of the Framework and Modules. Table 2 compares direct target-resolution reconstruction, coarse-to-fine multi-stage reconstruction, the physics-guided mixture-of-experts module, and the state-feedback mechanism at $96\times96\times20$. Coarse-to-fine reconstruction reduces the MAE from 0.226916 to 0.216941, and the physics-guided mixture-of-experts module further reduces it to 0.195705. After adding state feedback, the model achieves an MAE of 0.140833, an RMSE of 0.200990, and a PSNR of 43.4789 dB.

Table 2: Results of the coarse-to-fine urban wind-field reconstruction framework at resolution 96.

Configuration	MAE↓	MSE↓	RMSE↓	PSNR↑
Direct reconstruction	0.226916	0.104494	0.323255	39.3515
Coarse-to-fine baseline	0.216941	0.094538	0.307471	39.7864
C2F with Physics-Aware MoE	0.195705	0.077315	0.278056	40.6598
C2F with Physics-Aware MoE +physical feedback	**0.140833**	**0.040397**	**0.200990**	**43.4789**

Overall and regional comparisons at the low-resolution completion stage. Table 3 compares the base backbone with the physics-guided mixture-of-experts module under the same $24 \times 24 \times 20$ test protocol, reporting four metrics over the overall domain, air domain, and three key local regions using identical masks and a fixed PSNR dynamic range of 30 m/s.

Table 3: Overall and regional metrics at the low-resolution completion stage.

Region	Model	MAE↓	MSE↓	RMSE↓	PSNR↑
Overall	Base backbone	0.208065	0.089055	0.298420	40.046
	Our module	**0.183896**	**0.070040**	**0.264651**	**41.089**
Air	Base backbone	0.226719	0.096828	0.311172	39.682
	Our module	**0.200319**	**0.076057**	**0.275785**	**40.731**
Boundary	Base backbone	0.134950	0.038542	0.196322	43.683
	Our module	**0.114138**	**0.028709**	**0.169438**	**44.962**
Vortex	Base backbone	0.282179	0.146292	0.382481	37.890
	Our module	**0.242626**	**0.109556**	**0.330992**	**39.146**
High-wind	Base backbone	0.220307	0.087580	0.295939	40.118
	Our module	**0.199709**	**0.072272**	**0.268834**	**40.953**

The physics-guided mixture-of-experts module reduces the overall MSE from 0.089055 to 0.070040 (21.35%) and increases PSNR by 1.043 dB. Boundary, Vortex, and High-wind MSE decreases by 25.51%, 25.11%, and 17.48%, respectively, confirming consistent global and local improvements.

Table 4 uses the complete local physics-expert model as the baseline, with "Without X expert" denoting removal of expert X. Its regional contribution is measured by $\Delta\text{MAE} = \text{MAE}_{\text{without}} - \text{MAE}_{\text{full}}$ and reported in both absolute (m/s) and percentage terms relative to the full model.

Table 4: Local MAE intervention diagnostics for the local physics experts.

Configuration	Air MAE↓	Ordinary MAE↓	Boundary MAE↓	Vortex MAE↓	High-wind MAE↓	Expert contribution ΔMAE (m/s; relative percentage)
Baseline (full local experts)	0.200319	0.193947	0.114138	0.242626	0.199709	–
Without General expert	0.236479	–	–	–	–	0.036160; 18.05%
Without Ordinary expert	–	0.231180	–	–	–	0.037233; 19.20%
Without Boundary expert	–	–	0.197462	–	–	0.083324; 73.00%
Without Vortex expert	–	–	–	0.296330	–	0.053704; 22.13%
Without High-wind expert	–	–	–	–	0.238002	0.038293; 19.17%

The contributions of all five local physics experts are positive. In particular, after removing the Boundary expert, the local MAE in the corresponding region increases by 0.083324 m/s, representing a 73.00% increase relative to the Baseline. The relative contributions of the remaining local physics experts range from 18.05% to 22.13%. These results demonstrate that each local physics expert plays a clear role in wind-field reconstruction within its corresponding region.

5.4 Ablation Study: Momentum–Energy-Guided Sampling

Table 5 lists six scheduling configurations—GT-only, prediction states only, Smoothstep, Gradient, Momentum, and our method—and consistently reports MAE, RMSE, PSNR, and SSIM.

Table 5: Comparison of physics-guided sampling methods at resolution 96.

Scheduling method	MAE↓	RMSE↓	PSNR↑	SSIM↑
GT-only	0.195705	0.278056	40.6598	0.926947
Pred-only	0.148492	0.212053	43.0135	0.956444
Smoothstep	0.147521	0.211325	43.0434	0.956832
Gradient	0.143006	0.206223	43.2557	0.959035
Momentum	0.144125	0.207009	43.2226	0.958423
Ours	**0.140833**	**0.200990**	**43.4789**	**0.960481**

Our method achieves an MAE of 0.140833, an RMSE of 0.200990, a PSNR of 43.4789, and an SSIM of 0.960481; among the current test results, Our method achieves the best performance across all four metrics among the evaluated scheduling methods.

6 Conclusion

This paper presents a coarse-to-fine framework for urban wind-field reconstruction, decomposing high-resolution reconstruction into low-resolution wind-field completion followed by multi-stage super-resolution reconstruction. The physics-guided mixture-of-experts module leverages physical priors to encourage different experts to learn specialized representations of heterogeneous local flow patterns. The momentum–energy-guided sampling mechanism adjusts the sampling probability of predicted states according to the momentum–energy discrepancy between the oracle and current predicted intermediate states, progressively exposing subsequent reconstruction stages to the model's own prediction errors during training. Together, these three components address large-scale observation gaps, the recovery of fine-grained structures around buildings, and intermediate-state mismatch in cascaded reconstruction.

Reproducibility Statement

The main text and appendices specify the fixed temporal split, sensor mapping, input channels, normalization procedure, and physics-based differencing scheme. Training configurations, evaluation scripts, checkpoint manifests, random seeds, and result hashes will be made available through the anonymous GitHub repository. The test set is evaluated only after the experimental configuration has been frozen.

AI Use Statement

Generative AI was used only for limited editorial assistance, including proofreading, formatting checks, citation consistency checks, and revision of non-technical prose. It was not used to develop the research methodology, design or conduct experiments, generate experimental results, or make scientific conclusions. All AI-assisted content was reviewed and verified by the authors, who take full responsibility for the final manuscript.

References

Florian Achermann, Thomas Stastny, Bogdan Danciu, Andrey Kolobov, Jen Jen Chung, Roland Siegwart, and Nicholas Lawrance. Windseer: Real-time volumetric wind prediction over complex terrain aboard a small uncrewed aerial vehicle. *Nature Communications*, 15:3507, 2024. doi: 10.1038/s41467-024-47778-4.

Isaac Amidror. Scattered data interpolation methods for electronic imaging systems: A survey. *Journal of Electronic Imaging*, 11(2):157–176, 2002. doi: 10.1117/1.1455013.

Bert Blocken. 50 years of computational wind engineering: Past, present and future. *Journal of Wind Engineering and Industrial Aerodynamics*, 129:69–102, 2014. doi: 10.1016/j.jweia.2014.03.008.

Yong Cao, Peixing Xie, Guoshuo Huang, Wei Wang, Wen-Li Chen, Gang Hu, and Shuyang Cao. A pattern-constrained deep learning model for urban canopy turbulence reconstruction from sparse sensor data. *Building and Environment*, 285:113535, 2025. doi: 10.1016/j.buildenv.2025.113535.

Yuanqing Chen, Ding Wang, Dachuan Feng, Geng Tian, Vikrant Gupta, Renjing Cao, Minping Wan, and Shiyi Chen. Three-dimensional spatiotemporal wind field reconstruction based on LiDAR and multi-scale PINN. *Applied Energy*, 377:124577, 2025. doi: 10.1016/j.apenergy.2024.124577.

Damai Dai, Chengqi Deng, Chenggang Zhao, R. X. Xu, Huazuo Gao, Deli Chen, Jiashi Li, Wangding Zeng, Xingkai Yu, Y. Wu, Zhenda Xie, Y. K. Li, Panpan Huang, Fuli Luo, Chong Ruan, Zhifang Sui, and Wenfeng Liang. Deepseekmoe: Towards ultimate expert specialization in mixture-of-experts language models. *arXiv preprint arXiv:2401.06066*, 2024.

Richard Everson and Lawrence Sirovich. Karhunen–loève procedure for gappy data. *Journal of the Optical Society of America A*, 12(8):1657–1664, 1995. doi: 10.1364/JOSAA.12.001657.

William Fedus, Barret Zoph, and Noam Shazeer. Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity. *Journal of Machine Learning Research*, 23(120):1–39, 2022.

Kai Fukami, Koji Fukagata, and Kunihiko Taira. Super-resolution reconstruction of turbulent flows with machine learning. *Journal of Fluid Mechanics*, 870:106–120, 2019. doi: 10.1017/jfm.2019.238.

Huanxiang Gao, Gang Hu, Wenli Chen, Chao Li, Yiqing Xiao, and K. C. S. Kwok. Urban wind field real-time reconstruction based on optimal sparse sensors and deep learning

model. Conference abstract, 16th International Conference on Wind Engineering (ICWE 2023), Florence, Italy, 2023. URL https://web.aimgroupinternational.com/2023/icwe/papers/ICWE2023_AbstractSubmission-121_2022-12-27%2009_38_48.pdf.

Huanxiang Gao, Gang Hu, Dongqin Zhang, Wenjun Jiang, K. T. Tse, K. C. S. Kwok, and Ahsan Kareem. Urban wind field prediction based on sparse sensors and physics-informed graph-assisted auto-encoder. *Computer-Aided Civil and Infrastructure Engineering*, 39 (10):1409–1430, 2024. doi: 10.1111/mice.13147.

Albert Gu and Tri Dao. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*, 2023. doi: 10.48550/arXiv.2312.00752.

Rolland L. Hardy. Multiquadric equations of topography and other irregular surfaces. *Journal of Geophysical Research*, 76(8):1905–1915, 1971. doi: 10.1029/JB076i008p01905.

Chaoyi Hu, Hongyuan Jia, and Hideki Kikumoto. Estimation of airflow distribution in cubic building group model using POD-LSE and limited sensors. *Building and Environment*, 221:109324, 2022. doi: 10.1016/j.buildenv.2022.109324.

Chaoyi Hu, Hideki Kikumoto, Bingchao Zhang, and Hongyuan Jia. Fast estimation of airflow distribution in an urban model using generative adversarial networks with limited sensing data. *Building and Environment*, 249:111120, 2024. doi: 10.1016/j.buildenv.2023.111120.

Ameya D. Jagtap and George Em Karniadakis. Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. *Communications in Computational Physics*, 28(5):2002–2041, 2020. doi: 10.4208/cicp.OA-2020-0164.

Ameya D. Jagtap, Ehsan Kharazmi, and George Em Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020. doi: 10.1016/j.cma.2020.113028.

Wenjun Jiang, Taofeng Gu, Jize Zhang, Gang Hu, Wen Li Chen, and Hui Li. Physics-informed 3d wind field reconstruction in the urban canopy from LiDAR observations. *Building and Environment*, 301:114818, 2026. doi: 10.1016/j.buildenv.2026.114818.

George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3:422–440, 2021. doi: 10.1038/s42254-021-00314-5.

Yanyu Ke, Yifan Wang, Gang Hu, and K. T. Tse. GFNO: A generalized fourier neural operator framework for wind flow field reconstruction from sparse observations. *Building and Environment*, 304:115143, 2026. doi: 10.1016/j.buildenv.2026.115143.

Ehsan Kharazmi, Zhongqiang Zhang, and George Em Karniadakis. hp-VPINNs: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering*, 374:113547, 2021. doi: 10.1016/j.cma.2020.113547.

Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to pdes. *Journal of Machine Learning Research*, 24(89): 1–97, 2023.

Aditi S. Krishnapriyan, Amir Gholami, Shandian Zhe, Robert M. Kirby, and Michael W. Mahoney. Characterizing possible failure modes in physics-informed neural networks. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Dmitry Lepikhin, HyoukJoong Lee, Yuanzhong Xu, Dehao Chen, Orhan Firat, Yanping Huang, Maxim Krikun, Noam Shazeer, and Zhifeng Chen. GShard: Scaling giant models with conditional computation and automatic sharding. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=qrwe7XHTmYb.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=c8P9NQVtmnO.

Yi Liu, Ranpeng Wang, Yin Gu, Congjian Li, and Gangqiao Wang. Physics-inspired and data-driven two-stage deep learning approach for wind field reconstruction with experimental validation. *Energy*, 298:131230, 2024. doi: 10.1016/j.energy.2024.131230.

Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 10012–10022, 2021.

Vivek Oommen, Siavash Khodakarami, Aniruddha Bora, Zhicheng Wang, and George Em Karniadakis. Learning turbulent flows with generative models for super resolution and sparse flow reconstruction. *Nature Communications*, 17:3707, 2026. doi: 10.1038/s41467-026-70145-4.

Yusuf Patel, Vincent Mons, Olivier Marquet, and Georgios Rigas. Turbulence model augmented physics-informed neural networks for mean-flow reconstruction. *Physical Review Fluids*, 9:034605, 2024. doi: 10.1103/PhysRevFluids.9.034605.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, 2015.

Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc V. Le, Geoffrey E. Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. In *International Conference on Learning Representations*, 2017.

Donald Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM National Conference*, pp. 517–524, 1968. doi: 10.1145/800186.810616.

Yoshihide Tominaga, Akashi Mochida, Ryuichiro Yoshie, Hiroto Kataoka, Tsuyoshi Nozu, Masaru Yoshikawa, and Taichi Shirasawa. AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(10–11):1749–1761, 2008. doi: 10.1016/j.jweia.2008.02.058.

Sifan Wang, Yujun Teng, and Paris Perdikaris. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5):A3055–A3081, 2021. doi: 10.1137/20M1318043.

Haixu Wu, Huakun Luo, Haowen Wang, Jianmin Wang, and Mingsheng Long. Transolver: A fast transformer solver for PDEs on general geometries. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 53681–53705, 2024. URL https://proceedings.mlr.press/v235/wu24r.html.

Bingchao Zhang, Ryozo Ooka, Hideki Kikumoto, Chaoyi Hu, and K. T. Tim Tse. Towards real-time prediction of velocity field around a building using generative adversarial networks based on the surface pressure from sparse sensor networks. *Journal of Wind Engineering and Industrial Aerodynamics*, 231:105243, 2022. doi: 10.1016/j.jweia.2022.105243.

Yujia Zhang, Jiaxi Qi, Ruiyan Chen, Yong Liu, Yuzhou Zhang, Lyulin Kuang, Rita Zhang, and Shengze Cai. Transformer-based neural operators for 3d wind field prediction over complex mountainous terrain. *Communications Physics*, 2026. doi: 10.1038/s42005-026-02770-w.

Yanqi Zhou, Tao Lei, Hanxiao Liu, Nan Du, Yanping Huang, Vincent Y. Zhao, Andrew M. Dai, Zhifeng Chen, Quoc V. Le, and James Laudon. Mixture-of-experts with expert choice routing. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Yihang Zhou, Chao Lin, Hideki Kikumoto, Ryozo Ooka, and Sibo Cheng. Rooftop wind field reconstruction using sparse sensors: From deterministic to generative learning methods. *Building and Environment*, 296:114480, 2026. doi: 10.1016/j.buildenv.2026.114480.

A Supplementary Materials

A.1 Quantitative Results

To complement the quantitative results in Table 1, Figure 3 presents the ground-truth wind field, predictions from different methods, and absolute wind-speed errors for the same fixed test sample and height slice under a unified color scale. The consistent test case and display range enable direct comparison of the ability of different methods to recover local velocity structures.

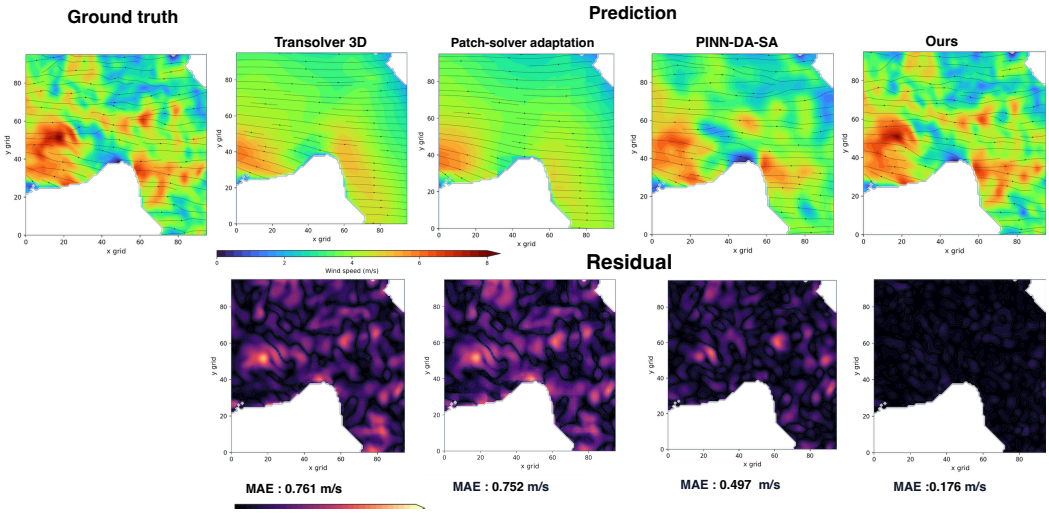


Figure 3: Ground-truth wind field, predictions from different methods, and absolute wind-speed errors on a fixed test sample.

A.2 Implementation and Comparisons

The physics-guided mixture-of-experts module is deployed at seven network locations, each with an always-active shared pathway. For each voxel, the router selects one prior physics expert and three self-learning physics experts from the candidate pool. Given the same input features, these experts provide complementary local residual corrections for general air regions, regular flows, building boundaries, vortices, and high-wind-speed regions.

The ablation studies compare direct prediction with coarse-to-fine reconstruction, standard residual blocks with the complete mixture-of-experts module, and ideal-state training with cascaded training. During validation and testing, all cascaded models use the predicted states generated by the frozen upstream networks as inputs.