Sustainable Cities and Society 150 (2026) 107890

Contents lists available at ScienceDirect

Sustainable Cities and Society

journal homepage: www.elsevier.com/locate/scs

An iterative graph neural operator with propagation field for urban-scale wind field reconstruction

Minjin Kong [a], Jinwoo Choi [a], Samir Shrestha [b], Sobit Sapkota [b], Jun-Ki Choi [c], Taehoon Hong [d,*]

[a] Graduate Research Assistant and Ph.D. Student, Department of Architecture and Architectural Engineering, Yonsei University, Seoul, Republic of Korea
[b] Graduate Research Assistant, Department of Mechanical Engineering, University of Dayton, 300 College Park, Dayton, OH, USA
[c] Professor, Department of Mechanical Engineering, University of Dayton, 300 College Park, Dayton, OH, USA
[d] Underwood Distinguished Professor, Department of Architecture and Architectural Engineering, Yonsei University, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Computational fluid dynamics
Graph neural operator
Surrogate model
Propagated confidence Field
Flow field reconstruction

ABSTRACT

Accurately predicting climate and environmental variables is essential in a wide range of engineering applications, and Computational Fluid Dynamics (CFD) has been widely used as a key analytical tool in these fields. However, conventional CFD simulations have limitations due to high computational costs and sensitivity to boundary conditions, restricting their use in large-scale or real-time scenarios. While data-driven CFD surrogate models have emerged as promising alternatives, they still face challenges in terms of handling unstructured meshes and accommodating flexible sensor configurations. To address these limitations, this study proposes an iterative graph neural operator with a Propagated confidence Field for steady-state CFD surrogate modeling on unstructured meshes. Instead of predicting the flow field all at once, this model iteratively improves the solution by reflecting the convergence behavior of steady-state CFD. A Propagated confidence Field was introduced to represent the spatial range of information obtained from inlet and sensor nodes. Validation results show that the proposed model maintains stable reconstruction performance across sparse and flexible sensor configurations within the tested urban CFD system. Error metrics decrease as sensor density increases, while cosine similarity improves, indicating enhanced reconstruction accuracy. However, once sensor density exceeds a certain level, the effect of performance improvement decreases. This model maintains the flow direction and overall structure even when observational data is insufficient. The proposed framework provides a discretization-invariant approach capable of flexibly handling varying sensor configurations. Future work will focus on improving physical consistency through physics-informed extensions.

1. Introduction

Accurate demand prediction is essential for narrowing the gap between energy demand and supply (Aduda et al., 2016; Aliero et al., 2021; Azadeh et al., 2008; Goe & Gaustad, 2014; Koo et al., 2014; Suganthi & Samuel, 2012; L. Zhao et al., 2026). Forecasting energy demand for heating and cooling depends on heating and cooling loads, which are calculated based on differences between outdoor and indoor thermal environments; therefore, accurately measuring or predicting the thermal environment forms the basis of energy-demand prediction for heating and cooling (Chang et al., 2024; Guo et al., 2025; Kong et al., 2025; Schatz & Kucharik, 2016; Wi, 2017). Quantitative prediction of climatic and environmental variables has become an essential component in various engineering fields, with applications ranging from indoor environmental control to urban air-quality management (J. Liu & Niu, 2016; Toparlar et al., 2017). In particular, reliably estimating the distributions of physical variables such as temperature, humidity, wind direction, and wind speed has long been treated as an important research topic (Fan et al., 2024; Jie et al., 2026). Because these variables involve complex interactions and nonlinear characteristics, various numerical approaches have been continuously developed to quantify them.

At the center of these numerical approaches is the development of Computational Fluid Dynamics (CFD). CFD is a physics-based simulation method that directly solves governing equations numerically, and it has been widely used for the quantitative analysis of complex environmental

* Corresponding author at: Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, 03722, Republic of Korea.
E-mail addresses: min920606@yonsei.ac.kr (M. Kong), jinwoo818@yonsei.ac.kr (J. Choi), shresthas26@udayton.edu (S. Shrestha), sapkotas4@udayton.edu (S. Sapkota), jchoi1@udayton.edu (J.-K. Choi), hong7@yonsei.ac.kr (T. Hong).

https://doi.org/10.1016/j.scs.2026.107890
Received 4 May 2026; Received in revised form 2 September 2026; Accepted 3 September 2026
Available online 4 September 2026
2210-6707/

Abbreviation

CFD	Computational Fluid Dynamics
GNO	Graph Neural Operator
IGNO-PF	Iterative Graph Neural Operator with Propagation field
MAE	Mean Absolute Error
RMSE	Root Mean Square Error

variables, particularly flow fields. CFD simulation has become a powerful tool for precisely analyzing airflow distributions and heat-transfer phenomena across multiple spatial scales (Stamou & Katsiris, 2006; Toparlar et al., 2017). Owing to these characteristics, CFD enables quantitative comparison and evaluation of various design alternatives and operational scenarios before physical experiments or facility construction, thereby substantially improving the reliability of engineering decision-making (Baek et al., 2026; Choi, Hong, et al., 2026; Choi, Kong, et al., 2026; Toparlar et al., 2017).

Despite these advantages, CFD-based analysis still has structural limitations. Representative issues include very high computational cost and strong sensitivity to boundary-condition specification. In particular, conventional CFD methods face inherent constraints in applications that require both high accuracy and fast responsiveness over large spatial domains (Blocken, 2015; Hajdukiewicz et al., 2024). These limitations become even more pronounced in situations requiring real-time prediction or repeated simulation.

Against this background, rapidly advancing artificial intelligence (AI) and data-driven modeling techniques have attracted attention as promising alternatives that can complement or partially replace conventional CFD. Machine-learning-based CFD surrogate models have the potential to dramatically reduce computational cost by approximating iterative numerical procedures through data-driven computation (Calzolari & Liu, 2021; Chen et al., 2025; Kastner & Dogan, 2023; Kops et al., 2025; Mao et al., 2025; L. Zhao et al., 2024; Zhao et al., 2024). For example, Vinuesa & Brunton (2022) have comprehensively reviewed the various ways in which machine learning can accelerate CFD through data-driven surrogate modeling, turbulence-model improvement, and machine-learning-CFD hybrid strategies. They concluded that, in the near term, approaches that complement existing solvers rather than completely replacing them are likely to provide the most practical benefits. Jeon et al. (2024) proposed a residual-based physics-informed transfer learning framework to address error accumulation in long-term simulations and demonstrated that a hybrid approach alternating between machine-learning prediction and CFD computation enables fast yet reliable analysis.

Nevertheless, existing machine-learning-based CFD surrogate models still have several important limitations. First, Convolutional Neural Network-based grid-structured models assume regular grids, therefore, when applied to complex geometries represented by unstructured meshes, interpolation is inevitably required, which may lead to the loss of local flow characteristics. Second, Graph Neural Networks can naturally represent unstructured meshes, but their computational cost increases rapidly as graph size grows, making extension to large-scale three-dimensional problems difficult (Kovachki et al., 2023). Third, most deep-learning-based models do not guarantee discretization invariance. In other words, because they depend on specific sensor locations or sampling points, their applicability is limited when sensors are added, removed, or relocated in real-world environments (Bahmani et al., 2025; Gao et al., 2024). Furthermore, many previous studies have focused on predicting the next time step of transient flow fields, while relatively few studies have directly addressed steady-state CFD analysis (Gao et al., 2024; Giral et al., 2025; Huang et al., 2022; Z. Liu et al., 2023; Shao et al., 2023; S. Wu et al., 2024; Zhao et al., 2024).

Steady-state CFD analysis is generally performed by iteratively solving nonlinear residual equations (Wang et al., 2018), and it can be interpreted as a fixed-point iteration process that gradually converges through iterative algorithms such as pressure-velocity coupling schemes. CFD simulation can also be understood as a functional mapping that takes geometric configurations and boundary conditions as inputs and outputs the final climate field, namely the flow field.

Based on this perspective, this study proposes the concept of a Propagated confidence Field as a new framework for representing the reliability of information during the iterative prediction process. To address unstructured-mesh problems and discretization invariance, a Graph Neural Operator (GNO), which can learn mappings from input fields to output fields on irregular domains, was adopted as the base model (Wang et al., 2018). In the proposed approach, only geometric information, inlet boundary conditions, and sensor conditions are provided as initial inputs, and the neural operator is applied iteratively so that inlet and sensor information is gradually propagated throughout the computational domain. Through this iterative propagation process, the model is guided to converge to the final steady-state solution.

Unlike conventional GNN- or GNO-based CFD surrogate models that directly regress the flow field from prescribed input conditions, IGNO-PF is designed to reconstruct the flow field from sparse and flexibly located sensor observations through confidence-aware iterative propagation. Consequently, the proposed Iterative Graph Neural Operator with Propagated confidence Field (IGNO-PF) is designed to simultaneously approximate the physical process by which flow information is spatially transmitted and the iterative convergence process of steady-state CFD analysis. This provides a CFD surrogate-modeling paradigm that alleviates some limitations of existing direct-regression approaches while remaining efficient and scalable.

2. Method and materials

2.1. Iterative graph neural operator with propagated confidence field

2.1.1. Problem setup

This study considers the prediction of a velocity field over an irregular CFD mesh represented as a graph $G = (V, E)$. Each node $i \in V$ corresponds to a computational cell, and each edge $(i,j) \in E$ represents geometric adjacency between neighboring cells. The goal is to iteratively compute the velocity field $u_t \in \mathbb{R}^{N\times 3}$ over a fixed graph structure T times. The key idea of IGNO-PF is to incorporate the concept of p_t, the extent to which information is received from nodes that provide reliable information, into the iterative computation (refer to Fig. 1).

At each iteration $t \in \{0, \ldots, T\}$, the model performs direct state prediction:

$$u_{t+1} = f_\theta(G, x_t) = f_\theta(G, x^{base}, u_t, p_t), \tag{1}$$

where x_t denotes model input; x^{base} denotes static node features; u_t is the current velocity field; and p_t is a Propagated confidence Field.

Unlike residual formulations, the model directly outputs the next state u_{t+1}, while boundary conditions are enforced externally after each update.

2.1.2. Node representation

Each node i is associated with a feature vector x_i^{base} constructed from CFD preprocessing. This includes (i) geometry mask; (ii) inflow mask; (iii) outflow mask; (iv) sensor mask; (v) surface distance; (vi) surface normal. Although the original dataset contains coordinate information and inflow (sensor) velocity channels, these are excluded from the model input x^{base}. Since the proposed model uses the predicted velocity field u_{t+1} as the velocity input for the next rollout step, the inflow velocity components are imposed directly on u_{t+1} through boundary condition clamping. This design avoids redundant use of the inflow velocity information while ensuring that the prescribed inflow condition is consistently maintained during iterative propagation. Spatial co-

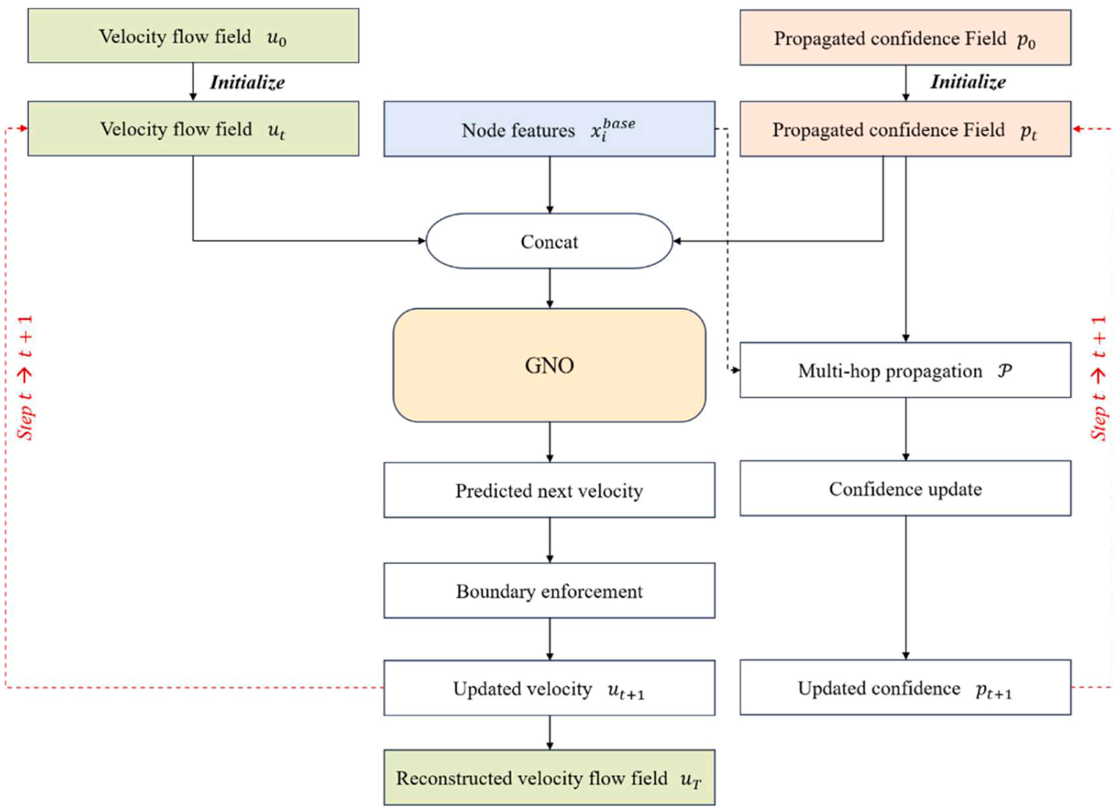

Fig. 1. Model architecture of IGNO-PF.

ordinates are used for calculating relative geometric information of G.

Thus, the model input at step t is defined as:

$$x_i^{(t)} = [x_i^{base}, u_i^{(t)}, p_i^{(t)}]. \quad (2)$$

This constitutes an integrated representation, where static features, dynamic state, and confidence are jointly embedded.

2.1.3. Edge representation and geometry encoding

Each edge (i,j) stores relative geometric information:

$$e_{ij} = \left[\Delta x_{ij}, \Delta y_{ij}, \Delta z_{ij}, d_{ij}\right], \quad r_i = (\boldsymbol{x}_i, \boldsymbol{y}_i, \boldsymbol{z}_i), \quad d_{ij} = \| r_j - r_i \|. \quad (3)$$

where r_i denotes position vector of node i; and d_{ij} denotes relative position vector between node i and j.

However, the model does not use these raw quantities directly. Instead, the edge attributes are transformed into a relative geometry encoding:

$$g_{ij} = \left[\frac{\Delta x_{ij}}{d_{ij}}, \frac{\Delta y_{ij}}{d_{ij}}, \frac{\Delta z_{ij}}{d_{ij}}, \log d_{ij}\right]. \quad (4)$$

This encoding captures direction through the unit displacement vector and scale through the logarithm of distance. No additional geometry terms, such as raw displacement or distance are used.

For edge construction, i.e., local neighborhood, an adaptive radius–based neighbor search was employed using spatial coordinates to generate directed edges $(e_{ij} \in E)$, enabling effective aggregation of local flow information across irregular geometries. For each cell node, candidate neighbors were identified by querying a spatial index structure (cKDTree) built on the node coordinates. A node-specific search radius was then determined by scaling the distance to its k-th nearest neighbor, ensuring a consistent local connectivity level across varying spatial densities. Directed edges were formed by connecting each source node to all neighboring nodes located within this adaptive radius. To prevent excessive connectivity in dense regions, the number of neighbors per node was capped by selecting the closest nodes based on Euclidean distance. Although neighborhood sizes of 16 or 32 are commonly used in local graph-based models, a target neighborhood size of 50 was adopted in this study when determining the adaptive search radius (Y. Liu & Chen, 2025; Qi et al., 2025). This value does not represent a fixed degree or the average degree of the final graph; rather, it serves as a reference value for constructing sufficiently populated local neighborhoods. Because the final edges were generated using an adaptive radius-based search with a maximum-neighbor cap, the actual number of neighboring nodes varied depending on the local spatial density and geometry. This setting ensured sufficient receptive field coverage while remaining computationally feasible. Consequently, the local neighborhoods used by the IGNO-PF operator typically contained approximately 50 to 100 nodes, depending on the local node distribution. Edge attributes were constructed using relative geometric information, specifically the coordinate difference vector and the Euclidean distance between connected node pairs.

2.1.4. Relative-geometry neural operator

The model consists of an embedding layer, a stack of graph operator blocks, and a readout layer.

- Embedding

The input is embedded into a latent space:

$$h_i^{(0)} = \phi_{embed}(x_i^{(t)}). \quad (5)$$

- Graph operator block

Each layer performs geometry-conditioned message passing. For each edge (i, j), the message is computed as:

$$m_{ij}^{(\ell)} = K_\theta\left(g_{ij}\right) \odot V_\theta\left(h_j^{(\ell)}\right), \tag{6}$$

where K_θ is an MLP applied to the geometry encoding g_{ij}; V_θ is an MLP applied to the source-node feature; and $\odot$ denotes element-wise multiplication.

Messages are aggregated by summation:

$$\overline{m}_i^{(\ell)} = \sum_{j \in \mathcal{N}(i)} m_{ij}^{(\ell)}. \tag{7}$$

The node state is updated via:

$$h_i^{(\ell+1)} = h_i^{(\ell)} + \alpha \cdot U_\theta\left(\left[h_i^{(\ell)}, \overline{m}_i^{(\ell)}\right]\right), \tag{8}$$

where U_θ is an MLP; and α is a scaling factor.

Thus, the operator performs a geometry-parameterized transformation, rather than a generic edge-feature aggregation.

- Readout

After L layers (ℓ), the velocity field is predicted as:

$$u_{t+1,\ i} = \phi_{readout}\left(h_i^{(L)}\right). \tag{9}$$

2.1.5. Propagated confidence Field

A key component of the method is a scalar node-wise field $p_t \in [0,\ 1]^N$, representing propagation-based confidence. The field p_t provides a global propagation signal that informs the neural operator about the reach of source information. Therefore, it is not predicted by the neural operator, is not used as a message-passing gate, and is used only as an input feature. The Propagated confidence Field should not be interpreted as a physical flow variable, such as momentum, pressure, turbulent diffusion, or numerical residual. It should also not be interpreted as a statistically calibrated uncertainty measure or as a direct predictor of local reconstruction error. Instead, it is an auxiliary graph-based field that represents the spatial availability and propagation status of flow information propagated from inlet and sensor nodes during the iterative prediction process. Its multi-hop update, distance attenuation, and confidence boosting are therefore used to approximate the progression of information availability and iterative stabilization over an irregular computational graph, rather than to directly model a specific physical transport mechanism.

- Initialization

The initial field is defined from two types of source nodes, (i) inflow nodes; and (ii) sensor-observed nodes. These nodes are assigned:

$$p_i^{(0)} = 1, \tag{10}$$

while all other nodes start from 0.

Meanwhile, outflow nodes are excluded from the update and fixed as:

$$p_i^{(t)} = 0 \ \forall t\ . \tag{11}$$

- Multi-hop propagation $\mathcal{P}$

The confidence field is updated using a precomputed sparse matrix A, which encodes an H-hop neighborhood with distance-based weighting:

$$p_{t+1}^{raw} = A\, p_t. \tag{12}$$

The matrix A is constructed using (i) shortest-path distances up to H-hops; (ii) exponential distance decay; and (iii) row-wise normalization.

- Confidence boost

A mild nonlinear boost λ is applied:

$$p_{t+1} = p_{t+1}^{raw} + \lambda\, p_{t+1}^{raw}\left(1 - p_{t+1}^{raw}\right). \tag{13}$$

2.1.6. Iterative rollout with hard constraints

At each iteration, the following steps are executed. First, inflow nodes are clamped to prescribed velocities. Second, sensor nodes are clamped to observed values. Third, outflow node velocities are overwritten using nearest non-outflow nodes. The outflow-node velocity was not treated as reliable flow information. Its propagated confidence value was fixed to zero, as mentioned in Eq. (11), so that outflow nodes could not act as velocity-information sources during the iterative rollout. However, assigning zero velocity to outflow nodes would introduce an artificial zero-velocity state, although the actual outflow velocity is unknown rather than physically zero. Therefore, the velocity values at outflow nodes were filled using the nearest non-outflow nodes only to avoid artificial zero-value contamination, while their confidence remained zero. Fourth, the neural operator predicts u_{t+1}. Fifth, the confidence field is updated to p_{t+1}.

This yields the iterative system:

$$u_{t+1} = f_\theta\left(G,\ x^{base},\ \mathcal{B}(u_t),\ p_t\right), \qquad p_{t+1} = \mathcal{P}(p_t; G) \tag{14}$$

where $\mathcal{B}$ denotes boundary enforcement (from first step to third step); and $\mathcal{P}$ denotes multi-hop propagation.

Conceptually, in IGNO-PF, u_t can converge to u_{GT} only after p_t has converged to 1 across all nodes. Therefore, the number of iterative computations T must be sufficiently large to allow the p_t of all nodes to converge to 1 and subsequently allow the u_t to converge. In this study, considering the size of the local neighborhood and the number of target nodes, T was set to 90.

2.1.7. Training with targeted intermediate supervision

Training is performed using losses at selected rollout steps $\mathcal{T}$. At each target step $\tau \in \mathcal{T}$, a confidence-weighted loss is applied:

$$\mathcal{L}^{(\tau)} = \frac{\sum_i \left(p_i^{(\tau)}\right)^\gamma \|u_i^{(\tau)} - u_i^{\mathrm{GT}}\|^2}{\sum_i \left(p_i^{(\tau)}\right)^\gamma + \varepsilon}. \tag{15}$$

Gradients are computed only at these target steps, and truncated backpropagation is applied over short windows. The total training objective $\mathcal{L}^{(\mathcal{T})}$ is the average of these intermediate losses. As in previous studies on autoregressive inference, no clear theoretical criterion has yet been established for selecting target steps at which losses should be imposed during rollout (Brandstetter et al., 2022; McCabe et al., 2023; Yoo et al., 2026). Therefore, to mitigate autoregressive error accumulation caused by the train-inference mismatch commonly discussed in related studies, this study empirically defined $\mathcal{T}$ as temporal anchor set. Specifically, $\mathcal{T}$ were set to the first rollout step and to the steps corresponding to 10%, 25%, 40%, 60%, 75%, 90%, and 100% of T. The 10% and 25% steps were set relatively densely as early-horizon anchors to reduce the propagation of trajectory drift generated in the early rollout stage into the long-horizon prediction region. Similarly, the 75% and 90% steps were also set relatively densely as late-horizon anchors to directly constrain accumulated autoregressive error before reaching the final horizon. The 40% and 60% steps were set as intermediate anchors to maintain trajectory coverage between the early and late stages, but with lower density in consideration of the additional computational cost that would arise if dense supervision were imposed over the entire rollout interval.

2.2. CFD dataset preparation

In this study, CFD simulations were performed for a three-dimensional urban model to obtain high-resolution flow data for training the IGNO-PF architecture. Specifically, to construct a reliable training dataset, (i) computational domain and geometry setup, (ii) turbulence and physics model selection, and (iii) parametric CFD dataset generation were sequentially conducted.

The target area for the CFD simulation was Mapo-gu, Seoul, Republic of Korea, where the S-DoT (Smart Seoul Data of Things) project is being implemented to collect, distribute, and analyze various urban-phenomenon data, including wind direction and wind speed, with the aim of supporting data-driven urban policy and service development (refer to Fig. 2) (Seoul Metropolitan Government, 2026). Mapo-gu was selected as the target area because, through the S-DoT project, it has the largest number of wind-direction and wind-speed sensors installed within Seoul, making it suitable for securing the observational data required to develop IGNO-PF.

2.2.1. Computational domain and geometry setup

Computational domain was extended around the target urban area to ensure sufficient flow development and prevent backflow. The size of the analysis domain was determined based on the CFD guidelines for urban wind environment prediction proposed by the Architectural Institute of Japan (AIJ) and related previous studies (Lee & Kim, 2026; Shirzadi et al., 2020; Tominaga et al., 2008). To maintain the blockage effect below 3% and prevent flow distortion at the domain boundaries, the AIJ guideline recommends distances of at least $5H_{max}$ at the top and lateral boundaries, $10H_{max}$ at the inlet, and $15H_{max}$ at the outlet, where H_{max} denotes the highest building height in the domain.

In this study, to secure high-quality and detailed flow-field data for IGNO-PF training, the distances to the inlet, lateral, and outlet boundaries were uniformly extended to $15H_{max}$ (2,460 m), which is more conservative than the AIJ guideline recommendations. The top height of the domain was also set to $5H_{max}$ (820 m) from the ground to minimize numerical errors within the analysis domain and represent fully developed flow. During geometry preprocessing, Blender and SpaceClaim were used with integrated GIS data from OpenStreetMap (OSM) to simplify complex urban geometry and prevent analysis errors, thereby generating a watertight solid model without topological errors such as non-manifold geometry.

A mesh independence study was conducted using six meshes ranging from 6.2 to 22.0 million cells (refer to Table 1). The predicted drag coefficient converged as the mesh was refined, while the variation became significantly smaller beyond the medium mesh resolution (0.372, 0.379, 0.384, and 0.385 for Medium_1, Medium_2, Fine_1, and Fine_2, respectively). Considering the marginal improvement in accuracy relative to the substantial increase in computational cost, the Medium_1 mesh was selected for generating the CFD dataset.

2.2.2. Turbulence and physics model selection

In this study, a physics-based model was constructed to predict the wind-speed field in the target domain. The Reynolds-Averaged Navier-Stokes (RANS) equations were adopted as the governing equations (Shirzadi et al., 2020). The Realizable k-epsilon model was used as the turbulence model because it has been validated for predicting mean flow fields in urban canyons and is efficient for large-scale case computations, and the standard wall function was applied for near-wall flow analysis to ensure convergence.

Table 1
Mesh independence test results.

Mesh resolution	Number of cells	Drag coefficient
Coarse_1	6,189,442	0.192
Coarse_2	9,207,716	0.293
Medium_1	12,834,974	0.372
Medium_2	16,770,568	0.379
Fine_1	19,203,456	0.384
Fine_2	22,018,760	0.385

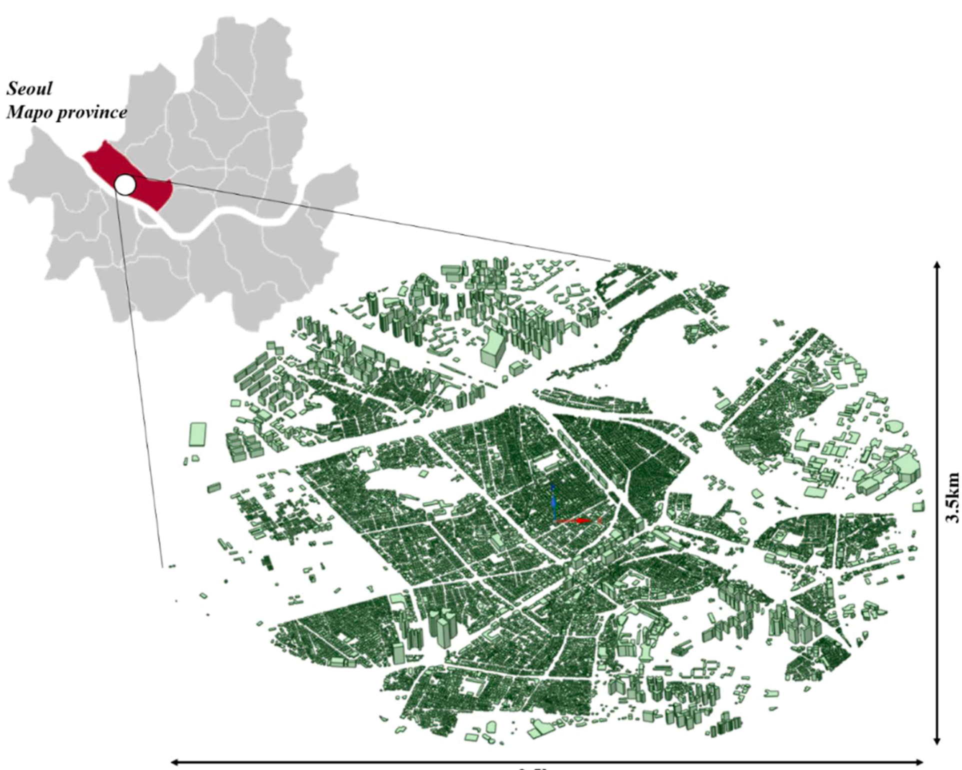

Fig. 2. Study area in Mapo province, Seoul, South Korea.

To validate the CFD simulation, the predicted wind speeds were compared with the measured wind speeds at the sensor locations. As summarized in Table 2, the CFD predictions showed reasonable agreement with the measurements, yielding an RMSE of 0.059 m/s and an R^2 value of 0.994.

2.2.3. Parametric CFD dataset generation

To construct a CFD dataset covering multiple wind-direction conditions, wind direction was selected as the input variable. Specifically, by referring to wind-direction and wind-attack-angle settings in related previous studies, the full range of inflow wind directions from 0° to 360° was discretized at 45° intervals, yielding eight conditions, and the inflow velocity (U_{ref}) was fixed at 1 m/s for each wind direction (Geng & Gou, 2025; Hågbo & Giljarhus, 2024; Hågbo et al., 2021; Kastner & Dogan, 2020, 2023; Peng et al., 2024). At the inlet boundary, an atmospheric boundary layer (ABL) profile was applied to reflect the variation of wind speed with height and to represent the actual urban wind environment.

2.3. Data generation for IGNO-PF

The objective of this study is to develop IGNO-PF, which is robust to unstructured meshes, invariant to discretization, and capable of handling non-fixed sensor arrangements. However, using the entire CFD dataset directly for IGNO-PF training is practically difficult because of the high memory requirements. Therefore, this study constructs the training data by sampling computationally manageable subdomains from the CFD dataset. Based on the eight CFD datasets described above, the training dataset was constructed as follows.

Subdomains were generated using arbitrary setpoints at ground level, and cylindrical domains were used to construct prediction regions without directional bias under 360° inflow wind-direction conditions (Gao et al., 2024). The height of the cylindrical domain was empirically determined based on the height of the sensor nodes that propagate key information, reflecting the characteristics of IGNO-PF designed to model information propagation. Because S-DoT sensors are installed at approximately 3.5±0.3m, the ceiling height was set to around 7 m to maintain comparable distances from the sensor to the ground and from the sensor to the ceiling (Seoul Metropolitan Government, 2026). The horizontal radius was empirically adjusted so that the number of mesh cells in the cylindrical domain did not exceed a computationally manageable scale while allowing IGNO-PF to learn prediction regions of various spatial scales, and in this experiment, an appropriate radius was found to lie between 50m and 150m. The nodes of each subdomain case were generated from the mesh cells contained in the corresponding cylinder, and the cell centroid of each mesh cell became an individual node constituting the IGNO-PF graph. Among all nodes in a subdomain case, nodes located within a distance equal to 10% of the cylinder radius from the cylinder boundary were defined as inlet and outlet nodes. Nodes located on the side opposite to the wind direction with respect to the setpoint were assigned as inlet nodes, and the remaining boundary nodes were assigned as outlet nodes. The nodes excluding inlet and outlet nodes were defined as target nodes to be predicted by IGNO-PF (refer to Fig. 3). In this experiment, 12 non-overlapping setpoints were selected for each of the eight CFD datasets with different wind directions, resulting in a total of 96 subdomain cases (8 CFD datasets × 12 setpoint selections). The maximum number of nodes in each subdomain case was limited to 24,000, which was set as the maximum allowable number considering the memory requirements and computational cost of IGNO-PF training in this experiment. Because the IGNO-PF operator estimates u_{t+1} based on local neighborhoods composed of 50 to 100 nodes, a single case can be interpreted not as one individual sample, but as a collection of many local neighborhood samples defined around each target node. Although these local neighborhoods substantially overlap, they form different input patterns depending on node locations and surrounding structural characteristics, therefore, a sufficiently large dataset is secured at the actual training-unit level.

Table 2
Measured and predicted wind speed at sensor location for CFD validation.

Sensor	Sensor location (Latitude / Longitude)	Measured speed	Predicted speed
#1	37.55589 / 126.91037	0.1	0.109
#2	37.55192 / 126.93068	0.1	0.111
#3	37.54747 / 126.93325	0.1	0.133
#4	37.56231 / 126.91081	0.4	0.421
#5	37.56483 / 126.90245	0.2	0.213
#6	37.55462 / 126.89596	2.4	2.298
#7	37.55538 / 126.92137	0.9	1.01

Next, to reflect the sparse and non-uniform sensor-placement characteristics of real cities, 10 random sensor-configuration variations were generated for each subdomain case, thereby constructing a total of 960 training datasets (96 subdomain cases × 10 sensor variations) (Hao et al., 2023; Kovachki et al., 2023). In each variation, between 0 and 5 target nodes located at a height of 3.5±0.3 m from ground level were randomly selected and assigned as sensor nodes. In this context, because IGNO-PF targets wind-direction and wind-speed information, when only the current S-DoT sensors that provide wind-direction and wind-speed information are considered, the number of sensors within a radius of 150m is approximately one. However, based on the overall sensor-placement status of the S-DoT project in Mapo-gu, up to five S-DoT sensors can be located within the same radius, and these sensors may provide wind-direction and wind-speed information in the future. Therefore, in this experiment, the number of sensor nodes was set between 0 and 5 in consideration of the scalability of IGNO-PF (Seoul Metropolitan Government, 2026).

Finally, for effective model training, 58 cases, corresponding to 6% of the 960 training datasets, were modified by replacing 1% to 10% of all target nodes with sensor nodes (i.e., low sensor node density dataset), and 38 cases, corresponding to 4%, were modified by replacing 10% to 40% of all target nodes with sensor nodes (i.e., high sensor node density dataset). This is related to the autoregressive error accumulation issue discussed above. Because of the characteristics of IGNO-PF, the model must learn the iterative computation process from the initial stage to the middle and late stages; however, no explicit ground truth exists for the intermediate stage after iterative computation has progressed to some extent. In theory, if a large number of datasets and sufficient training time were available, the model could be expected to generate its own intermediate-stage targets. However, because both time and computational cost are constrained, this study considered a strategy for providing indirect information about the intermediate stage. Motivated by related studies showing that neural operators can improve learning efficiency by mixing lower-resolution and higher-resolution datasets owing to discretization invariance, this study uses low sensor node density datasets in a role analogous to lower-resolution datasets and high sensor node density datasets in a role analogous to higher-resolution datasets. In other words, by regarding sensor nodes as nodes that have converged through iterative computation to values approximating the ground truth, datasets with low sensor node density were expected to substitute training data for the early stage of iterative computation, whereas datasets with high sensor node density were expected to substitute training data for the intermediate stage after some progress in iterative computation. Here, the sensor node density ranges and the dataset allocation ratios for each density range were empirically determined by referring to related multi-fidelity and multi-resolution studies (Kovachki et al., 2023; Li et al., 2024; X. Y. Liu et al., 2024; Tang et al., 2024). The training dataset was split by case, with 80% used for training and 20% for testing.

2.4. Evaluation metrics for IGNO-PF validation

For the validation of IGNO-PF, one validation case was constructed

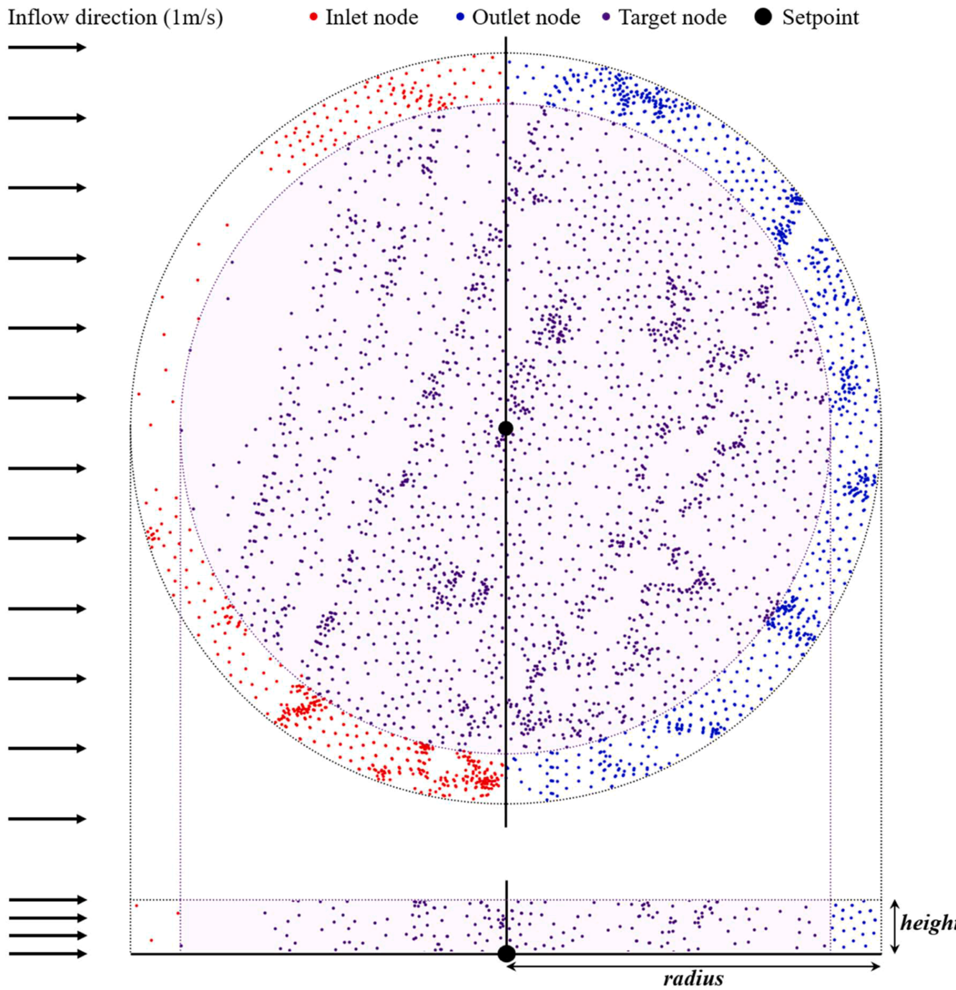

Fig. 3. Example of node extraction nodes from CFD mesh.

for each of the eight wind-direction conditions considered in this study. The validation setpoints were randomly selected from S-DoT sensors located in subdomains that were spatially separated from the training and test subdomains. The validation subdomains were also selected to be mutually non-overlapping. Thus, the validation dataset was designed to assess model performance across all eight wind-direction conditions while maintaining spatial separation among the training, test, and validation subdomains. Then, to evaluate the basic performance of IGNO-PF and conduct a sensitivity analysis with respect to sensor density, 9 scenarios were assigned to each validation case (refer to Table 3). First, scenario #1 serves as baseline and simulates the installation density per unit area of S-DoT sensors that provide wind-direction and wind-speed information. In scenarios #2 to #9, the ratios of sensor nodes to target nodes were set to 1%, 2%, 5%, 10%, 15%, 20%, 25%, and 30%, respectively. Because the number of target nodes differs across validation cases, the scenarios were defined based on the relative ratio of sensor nodes to target nodes rather than the absolute number of sensor nodes. In particular, 1% and 2% conditions were added to more finely evaluate performance changes under low-density sensing conditions, and the ratio was then increased at 5% intervals in the subsequent range (Cao et al., 2025; Du et al., 2026; H. Wu et al., 2023).

Table 3
Validation dataset scenario description.

Scenarios	Sensor description
#1	Node closest to the setpoint, i.e., S-DoT, is designated as the sensor node
#2	1% of all target nodes are replaced with sensor nodes
#3	2% of all target nodes are replaced with sensor nodes
#4	5% of all target nodes are replaced with sensor nodes
#5	10% of all target nodes are replaced with sensor nodes
#6	15% of all target nodes are replaced with sensor nodes
#7	20% of all target nodes are replaced with sensor nodes
#8	25% of all target nodes are replaced with sensor nodes
#9	30% of all target nodes are replaced with sensor nodes

The evaluation metrics in this study were selected to consider the vector, components (u, v, w), and direction of the wind flow. Root mean square error (RMSE) and relative L2 error (ε_{rel,L^2}) were used to evaluate the vector and component-wise errors. Mean absolute error end-point error (MAE (EPE)) was used to evaluate the positional error of the velocity vector itself, that is, the distance error between the predicted and observed vectors. Finally, cosine similarity ($\cos\theta$) was used to evaluate directional agreement.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2} \tag{16}$$

$$\varepsilon_{rel,L^2} = \sqrt{\frac{\sum_{i=1}^{N}\|P_i - O_i\|^2}{\sum_{i=1}^{N}\|O_i\|^2}} \tag{17}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\| P_i - O_i \| \tag{18}$$

$$\cos\theta = \frac{1}{N}\sum_{i=1}^{N}\frac{P_i \cdot O_i}{\| P_i\|\|O_i\| + \epsilon} \tag{19}$$

where P_i stands for predicted value of node i; O_i stands for observed value of node i.

In addition to the data-fidelity metrics described above, the physical plausibility of the predicted velocity fields was evaluated by comparing the divergence of the velocity field between the CFD reference and IGNO-PF prediction. The divergence of a velocity field represents local volumetric expansion or compression of the flow and can therefore be used as an indicator of artificial flow sources or sinks in the predicted flow field. Because the present dataset is defined on an irregular graph rather than a structured grid, the velocity divergence was estimated using a local least-squares approximation over the graph neighborhood defined in section 2.1.3.

For each evaluation node i, the local graph neighborhood $\mathcal{N}_i$ was used to estimate the local velocity-gradient tensor. For each neighboring node j in $\mathcal{N}_i$, the relative position and velocity vectors are defined as $\Delta r_{ij} = r_j - r_i$ and $\Delta u_{ij} = u_j - u_i$. The local velocity-gradient tensor was then obtained by solving the following weighted least-squares problem:

$$\mathrm{G}_i = \underset{\mathrm{G}}{\operatorname{argmin}}\sum_{j\in\mathcal{N}_i}\omega_{ij}\| \Delta u_{ij} - \mathrm{G}\Delta r_{ij} \|_2^2 , \omega_{ij} = \frac{1}{\| \Delta r_{ij}\| + \epsilon} \tag{20}$$

where G_i is the estimated local velocity-gradient tensor; and ω_{ij} is a distance-based weighting factor with ϵ to avoid division by zero.

The divergence at node i was then calculated as the trace of the estimated velocity-gradient tensor:

$$\nabla\cdot u_i \approx \mathrm{tr}(\mathrm{G}_i) = \frac{\partial \mathrm{u}}{\partial \mathrm{x}} + \frac{\partial \mathrm{v}}{\partial \mathrm{y}} + \frac{\partial \mathrm{w}}{\partial \mathrm{z}} \tag{21}$$

The divergence was computed for both the CFD reference and IGNO-

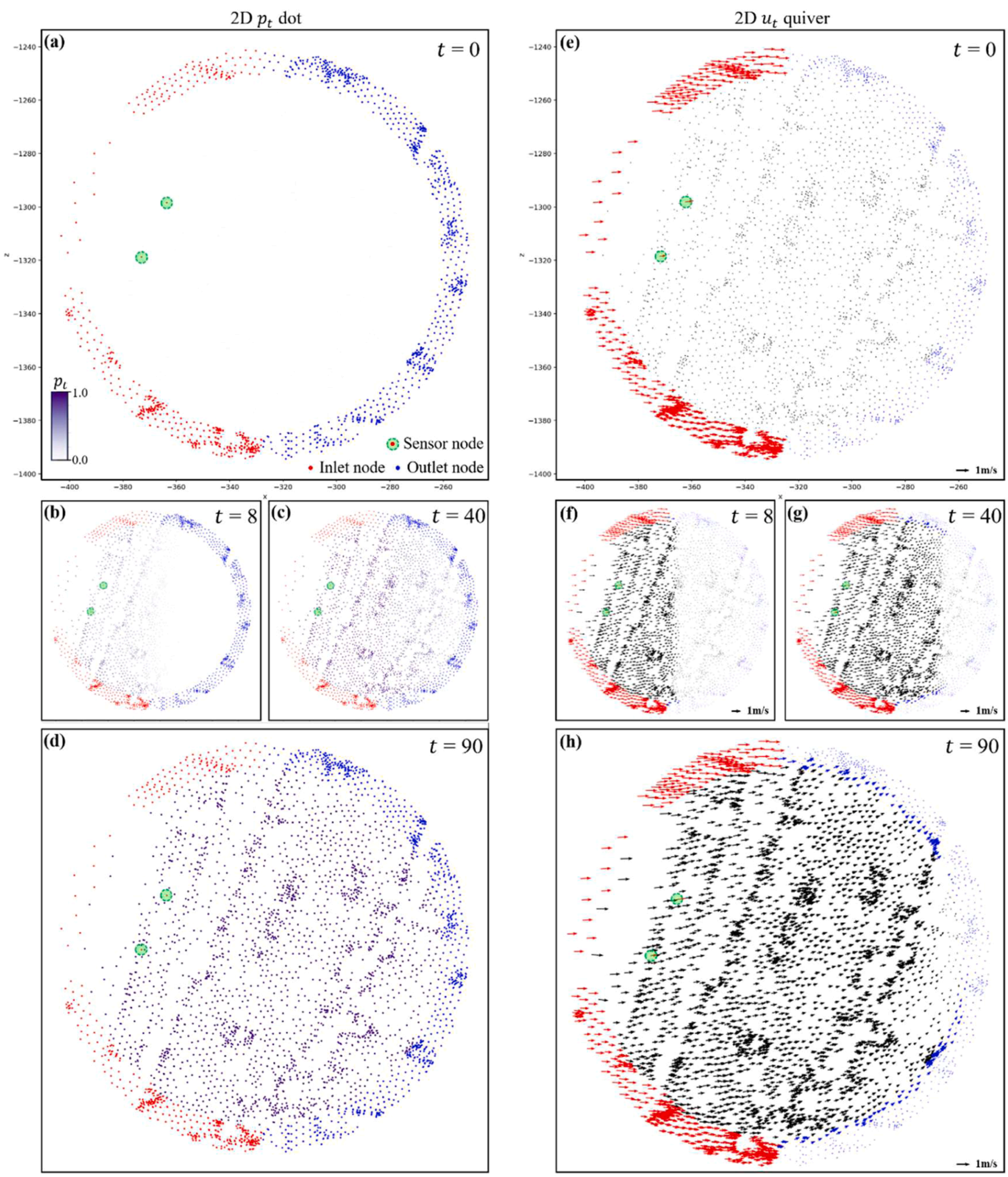

Fig. 4. Propagated confidence Field and velocity field results by step.

PF predicted fields over the evaluation nodes. Three divergence-based statistics, (i) mean absolute divergence; (ii) mean absolute divergence error; and (iii) divergence-error bias, were used for physical plausibility assessment. The mean absolute divergence compares the overall level of local volumetric imbalance between the CFD and predicted fields, while the latter two metrics quantify the magnitude and signed tendency of local source/sink discrepancy relative to the CFD field, respectively.

3. Results and discussion

In this study, IGNO-PF was trained using an NVIDIA RTX 3080 12 GB GPU with 12 GB of memory. The training required 51 hours to complete.

3.1. Trained model overview

Fig. 4 shows the Propagated confidence Field (p_t) and velocity field (u_t) estimated by the trained IGNO-PF for one test dataset with a radius of 75m and 18,500 target nodes. The y-axis value, corresponding to height, of p_t and u_t is omitted in this visualization.

In Fig. 4(a) to (d), p_t is shown in white when it is close to 0 and in purple when it is close to 1. As the iteration step t approaches T, namely 90, the purple points are observed to gradually spread from the red inlet nodes toward the blue outlet nodes. Because p_t of target nodes is updated using p_t values in the local neighborhood, and since only the inlet nodes and sensor nodes set p_0 value to 1, p_t of target nodes adjacent to the inlet nodes was updated from 0 to a value close to 1 with only 8 iterations (refer to Fig. 4(b)). By contrast, for target nodes located on the opposite side of the inlet nodes, more than 40 iterations were required before p_t was updated to values greater than 0 (refer to Fig. 4(c)). It can also be observed that u_t on the right is calculated and converges according to the level of diffusion of $p_t \approx 1$ (refer to Fig. 4(e) to (h)). When comparing u_8 quivers and u_{40} quivers of target nodes close to inlet nodes, u_t of those nodes converged to some extent with only 8 iterations of calculation, but for target nodes located on the opposite side, the degree of information propagation was insufficient with only 40 iterations of calculation, so some nodes were not computed at all. The core idea of IGNO-PF is to provide the Propagated confidence Field p_t as an additional input feature so that the model can use this information as needed. This example confirms that the model meaningfully uses p_t when estimating u_t.

Fig. 5 shows the predicted value (u_{90}), observed value (u_{GT}), and the differences between them for the test dataset used as an example in Fig. 4. Comparing the predicted quiver in Fig. 5(a) with the observed quiver in Fig. 5(b), the discrepancy in quiver direction and magnitude increases from the left side to the right side of the domain. This can also be confirmed from the difference quiver at $t = 90$ in Fig. 5(f). Target nodes close to source nodes that provide accurate information, namely inlet or sensor nodes, show results that are almost identical to the observed values because the source nodes are included in their local neighborhoods.

Naturally, as target nodes become farther from source nodes, they must make predictions using secondary and tertiary values derived from the source nodes, and thus their accuracy inevitably decreases. In particular, the predicted quivers of target nodes adjacent to outlet nodes are relatively biased northward compared with the observed quivers.

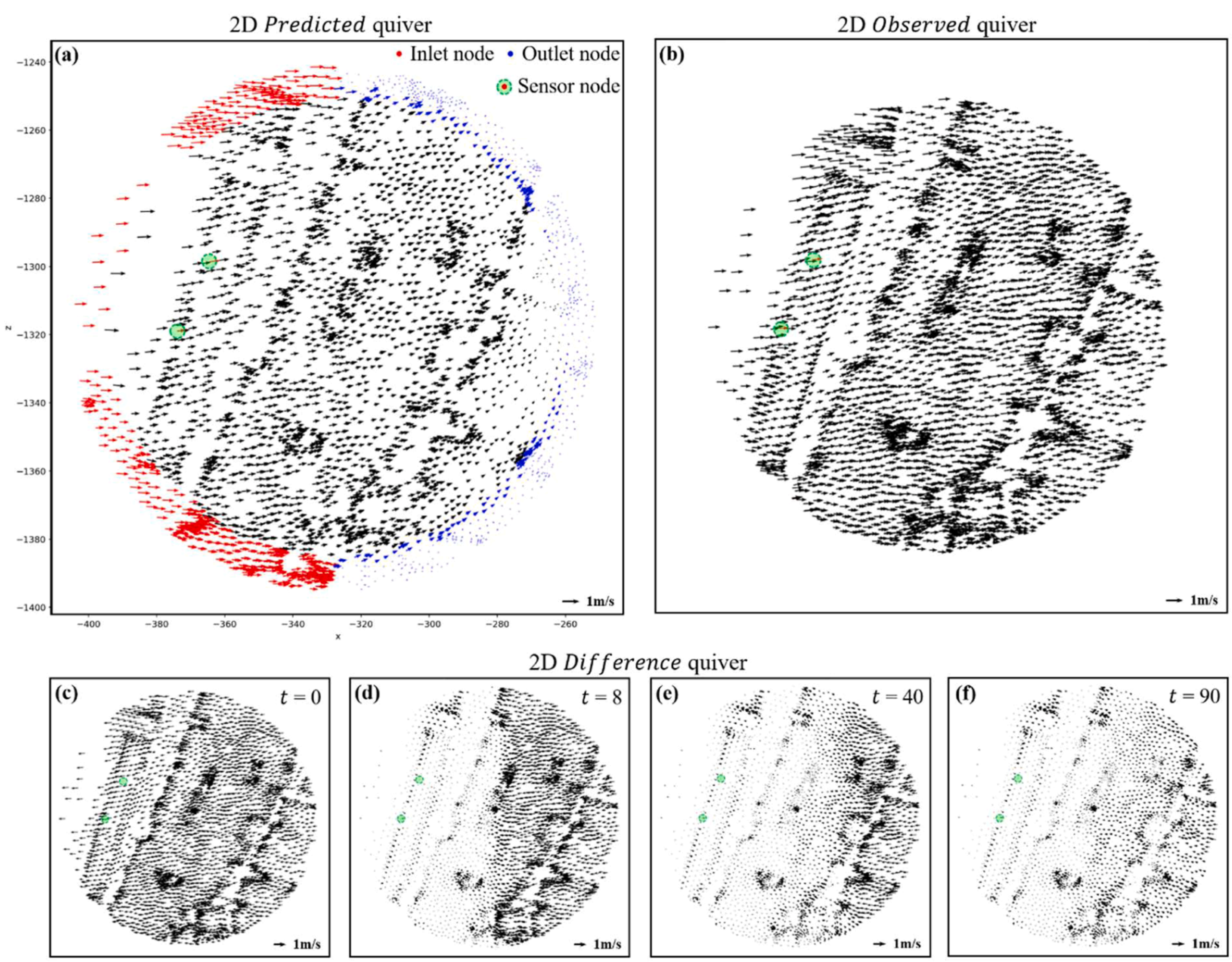

Fig. 5. Comparison between predicted and observed quiver.

Nevertheless, as the iterative computation proceeds, the error tends to (i) propagate from the inlet nodes toward the outlet nodes and (ii) gradually decrease (refer to Fig. 5(c) to (f)). This result indicates that the design intention of IGNO-PF, which aims to mimic the convergence behavior observed in steady-state CFD analysis, is appropriately reflected.

3.2. *Empirical validation of the proposed IGNO-PF concept*

To empirically support the proposed IGNO-PF concept, scenario-step similarity analysis, rollout behavior analysis, and ablation study were performed using the intermediate velocity fields generated during validation.

3.2.1. *Scenario-step similarity analysis*

More specifically, scenario-step similarity analysis compares predicted velocity fields across different sensor density scenarios and rollout steps (refer to Fig. 6). If high sensor density datasets indeed provide information analogous to later stage rollout states, as mentioned in section 2.3, then velocity fields from high density scenarios at early rollout steps should be similar to those from low density scenarios at later rollout steps.

The matrices in Fig. 6 present pairwise comparisons between the rollout steps of scenarios #1 to #4 and those of scenarios #2 to #6. The rollout steps considered are $t = 3$, 9, 18, 27, 36, 45, 54, 63, 72, 81, and the final step. Fig. 6(a) illustrates the relative L2 error, whereas Fig. 6(b) reports the cosine similarity for scenario-step similarity analysis. In both matrices, the lower-left region is highlighted in red, while the upper-right region becomes progressively closer to white. Because lower relative L2 errors and higher cosine similarity values are encoded in red, the lower-left region indicates relatively high similarity. This region represents comparisons between lower-density scenarios at later rollout steps and higher-density scenarios at earlier rollout steps. Therefore, the high similarity observed in this region indicates that changes in sensor density can produce velocity-field states analogous to different stages of the iterative rollout. This result provides empirical support for using low- and high-sensor-node-density datasets as indirect intermediate supervision in the proposed model.

3.2.2. *Rollout behavior analysis*

To examine whether the predictions of the trained model with $T = 90$ sufficiently converged before t reached final rollout step, a rollout behavior analysis was conducted. Specifically, velocity fields between successive rollout steps within the same scenario were compared, and the mean values of p_t were analyzed as an auxiliary indicator of the saturation of propagated source information during the iterative prediction process (refer to Fig. 7).

Fig. 7(a) shows the relative L2 error between velocity fields at successive rollout step intervals, such as from $t = 3$ to 9 and from $t = 9$ to 18, for the four scenarios. Fig. 7(b) shows the cosine similarity between velocity fields at the same successive rollout intervals, and Fig. 7(c) shows the mean value of p_t at each rollout step. It should be noted that

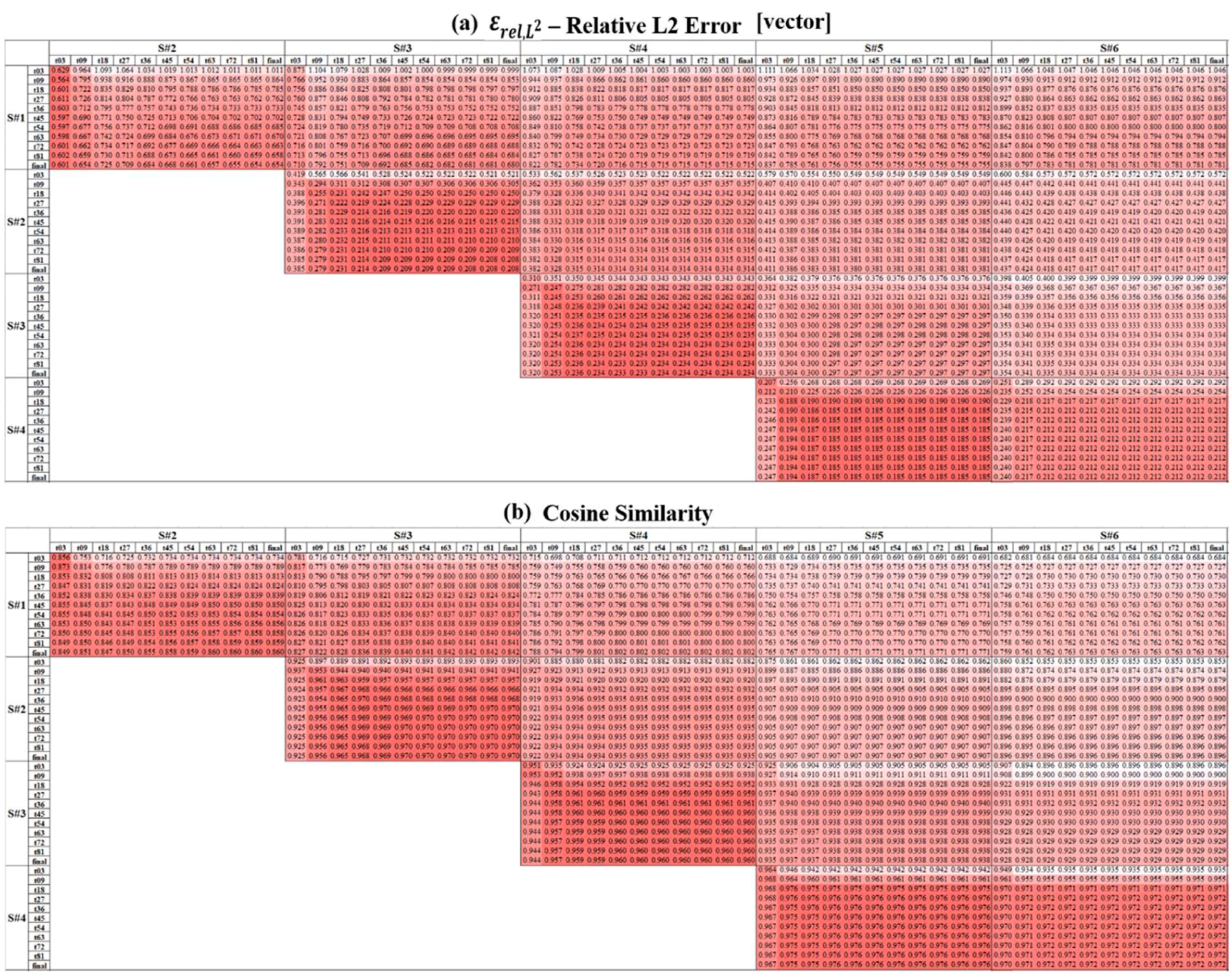

Fig. 6. Scenario-step similarity analysis results.

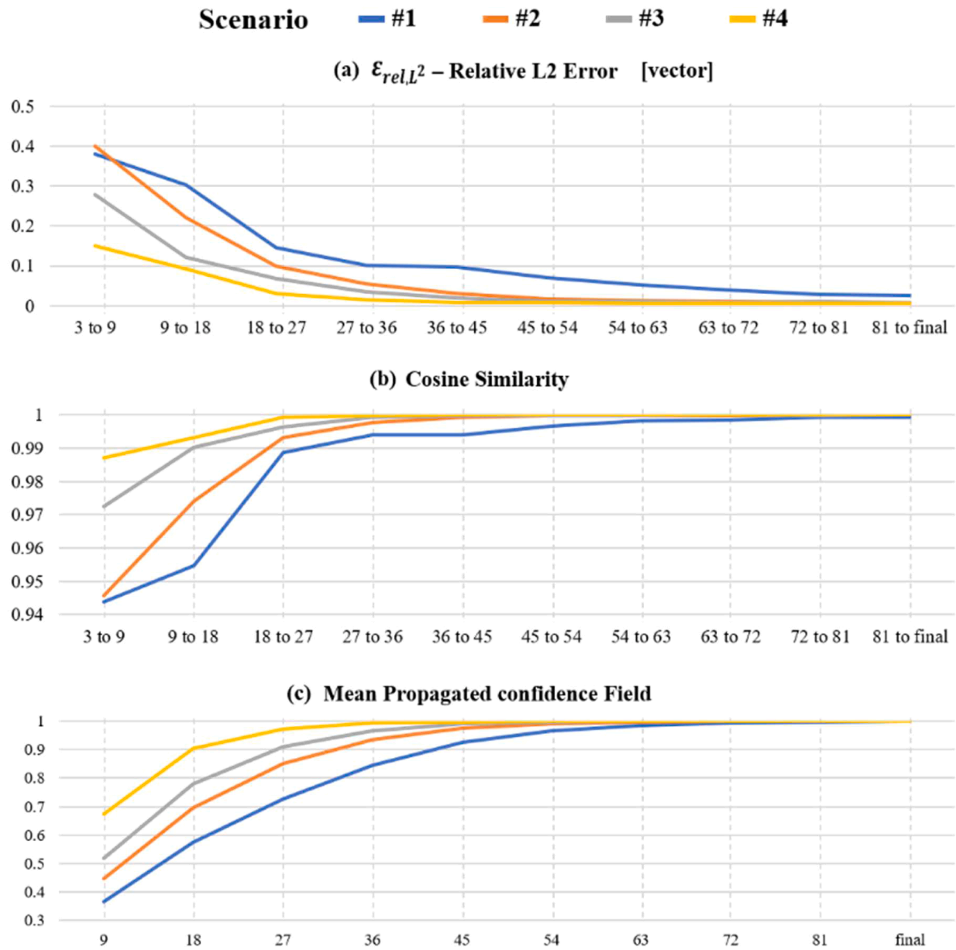

Fig. 7. Rollout behavior analysis results.

the rollout-step metrics used here are not identical to the residuals of a conventional CFD solver. In CFD simulations, residuals are computed from the discretized governing equations, whereas the present metrics quantify the stabilization of the predicted velocity field between successive neural-operator rollout steps. Therefore, this analysis is intended to evaluate the fixed-point-like convergence behavior of the IGNO-PF rollout, rather than to claim direct equivalence with CFD residual convergence. As the sensor density increases from scenario #1 to scenario #4, the iterative rollout process converges more rapidly. Notably, even in scenario #1, which has the lowest sensor density, the relative L2 error decreases continuously and reaches a near-plateau by the $t = 72$ to 81 interval, indicating that the velocity field becomes stabilized before the final rollout step. Similarly, the cosine similarity and the mean p_t approach plateau values before $t = 63$, suggesting that both the velocity-field update and the propagated source-information availability become stabilized even under the sparsest sensor configuration. Overall, the decreasing relative L2 difference, increasing cosine similarity, and saturation of the mean p_t indicate that successive velocity-field updates become progressively smaller as the rollout proceeds. It should be emphasized that this rollout-step analysis was conducted using intermediate rollout outputs generated from a single trained model with $T = 90$. Therefore, it should not be interpreted as a full sensitivity analysis based on independently retrained models with different T values. Although a full sensitivity analysis over multiple T values would require retraining and validating the model for each iteration setting, the rollout behavior analysis provides empirical support for the selected value of $T = 90$. The results show that the velocity field had already stabilized within the $T = 90$ rollout budget. Because the inference cost increases approximately linearly with the number of rollout steps, $T = 90$ was selected as a practical balance between convergence stability and computational cost.

To further examine whether the decreasing rollout update was simply caused by iterative over-smoothing, the spatial standard deviation of the velocity magnitude was additionally computed at each rollout step (refer to Table 4). Specifically, the velocity magnitude was first calculated at each evaluation node, and then its spatial standard deviation over all evaluation nodes was normalized by the corresponding value of the CFD reference field. If the iterative rollout progressively smoothed out the predicted flow field, this normalized value would be expected to continuously decrease as the rollout step increased.

However, the normalized speed spatial standard deviation remained close to the CFD reference level after the early rollout steps and reached a plateau of approximately 0.97. This indicates that the spatial contrast of the predicted velocity field was not progressively collapsed during the

Table 4
Rollout-step analysis of prediction stabilization and spatial-variance preservation.

Step	ε_{rel,L^2} [a]	Cosine similarity	Normalized speed spatial std [b]
$t = 18$	0.341	0.927	0.968
$t = 45$	0.342	0.927	0.971
$t = 63$	0.343	0.927	0.970
$t = 90$	0.343	0.926	0.970

Note: ε_{rel,L^2} [a] stands for Relative L2 Error; Normalized speed spatial std [b] denotes the spatial standard deviation of the predicted velocity magnitude divided by that of the CFD reference field rollout. In addition, the relative L2 error and cosine similarity also remained nearly unchanged after the plateau stage, suggesting that the late-stage rollout preserved the overall spatial variability while maintaining prediction stability. Therefore, the decreasing difference between successive rollout steps is unlikely to be explained solely by iterative over-smoothing. Nevertheless, this analysis does not fully resolve the reconstruction quality of wake-region or high-frequency local structures, which remains an important topic for future work.

3.2.3. Ablation study

The ablation study was conducted to evaluate the individual contributions of the key components suggested in the study, i.e., i) Propagated confidence Field, ii) iterative rollout, and iii) low and high sensor node-density datasets, by comparing performances of the proposed model with five ablated variants (refer to Fig. 8). Among the ablated variants, A2 was designed as a controlled standard-GNO baseline. A2 uses the same graph construction, geometric edge encoding, input variables, training-validation split, and GNO backbone as the proposed IGNO-PF, but removes both the iterative rollout process and the Propagated confidence Field. Therefore, A2 performs single-step direct regression from the available inlet and sensor information to the final velocity field, corresponding to the conventional use of a GNO-based CFD surrogate model under the same experimental setting.

Fig. 8(a) to 8(c) compare RMSE, relative L2 error, MAE (EPE), respectively, for the proposed model and five ablated variants across the nine sensor density scenarios. Fig. 8(d) compares the cosine similarity under the same scenarios. As highlighted by the red square markers, the proposed model achieved the best performance across all evaluation metrics except cosine similarity. In contrast, A1 without the Propagated confidence Field, highlighted by the black circle markers, exhibited degraded performance across all metrics, with the lowest cosine similarity among all models. The A2 variant without iterative rollout (i.e., standard GNO model), highlighted by the gray triangle markers, exhibited the worst performance in scenario #1, where only extremely sparse sensor observations were available. This is because predicting the wind flow is difficult when no valid wind-flow information exists within the edge neighborhood of the target node. Although the propagated information in A1 may be unreliable, iterative rollout is still able to propagate information throughout the graph, enabling A1 to achieve a reasonable level of performance even in scenario #1. In contrast, A2 cannot propagate sparse source information beyond the local neighborhood used for direct prediction, resulting in the poorest performance under the sparsest sensor configuration. From scenario #2 onward, where the sensor node-density reaches 1%, valid wind flow information is more likely to exist within the edge neighborhood of each target node. Under these conditions, A2, which directly predicts the results from a small amount of reliable information, can outperform A1, which must generate predictions without distinguishing between a small amount of reliable information and a large amount of uncertain information. However, beginning from scenario #5, where a large amount of reliable information becomes available, the influence of uncertain propagated information is substantially reduced. Consequently, A1 is able to produce more accurate predictions than A2 by propagating reliable information through iterative rollout. Although A2 is the only variant that outperforms the proposed model in terms of cosine similarity, the remaining evaluation metrics suggest a different interpretation. Taken together, these results imply that A2 may be biased toward matching the flow direction at the expense of accurately reconstructing the velocity magnitude. Nevertheless, the present comparison does not constitute an exhaustive benchmark against independently published state-of-the-art flow-field reconstruction models. Because published models often differ in mesh structure, boundary conditions, sensor configuration, input variables, and dataset construction, broader benchmarking against external models under strictly matched conditions remains an important direction for future work.

The A3 variant without both low and high sensor node-density datasets exhibited the poorest overall performance across almost all scenarios, demonstrating the importance of training with datasets spanning a wide range of sensor densities. In contrast, A4 without the high sensor node-density datasets and A5 without the low sensor node-density datasets showed relatively degraded performance in the scenarios corresponding to their omitted training datasets. Specifically, A4 performed worse in scenarios #6 to #9, whereas A5 showed inferior performance in scenarios #2 to #4. These results indicate that incorporating both low and high sensor node-density datasets improves the generalization capability of the proposed model across a wide range of sensor density conditions.

To further substantiate the computational efficiency of the proposed

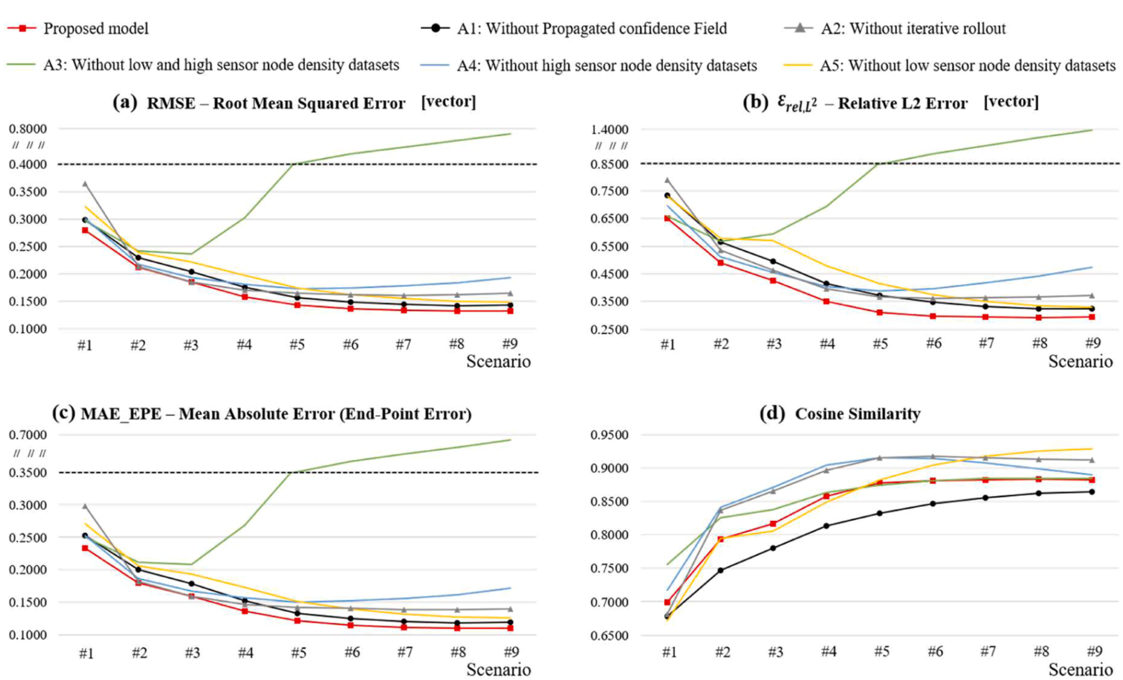

Fig. 8. Ablation study results.

surrogate model, the average computational time required for one validation sample was compared among the CFD simulation, the proposed IGNO-PF model, and the A2 variant, which corresponds to a standard GNO model without the iterative propagation process. Each steady-state CFD case was solved on a single core of a 12th Gen Intel i7-12700KF CPU at 3.61 GHz with 32 GB memory, requiring approximately 28,800 sec for the full CFD domain with 12,834,974 cells. As an approximate order-of-magnitude reference, the equivalent computational time for a sample with fewer than 24,000 cells was estimated to be approximately 50.0 sec by linearly scaling the full-domain CFD runtime by the number of cells. However, this estimate should be interpreted with caution because CFD runtime is affected not only by the number of cells but also by mesh quality, convergence criteria, boundary conditions, turbulence model, solver settings, and parallel efficiency.

The neural-operator inference for IGNO-PF and the A2 variant was performed on an NVIDIA GeForce RTX 2060 GPU with 6 GB of dedicated GPU memory. IGNO-PF required approximately 5.0 sec per sample for inference, whereas the A2 variant required approximately 0.35 sec per sample because it performs direct prediction without iterative propagation. It should also be noted that the CFD computation and neural-operator inference were performed on different hardware architectures, namely CPU and GPU, respectively. Therefore, the comparison is not intended as a strict hardware-normalized benchmark, but rather as an indicative comparison of computational cost under the computational settings used in this study. Within this limited comparison, IGNO-PF was approximately one order of magnitude faster than the estimated CFD simulation time for an equivalent sample size, while the A2 variant was faster because it does not perform iterative refinement. These results suggest the practical computational advantage of IGNO-PF as a CFD surrogate model, while also highlighting the trade-off between inference speed and iterative refinement.

3.3. Model validation results

The evaluation results for validation cases that were not used for training or testing are presented below (refer to Table 5 and Fig. 9).

Overall, across the predefined validation scenarios, RMSE, relative L2 error, and MAE (EPE) consistently decreased as the sensor-node ratio and spatial coverage increased, whereas cosine similarity tended to increase. This indicates that, as the amount and spatial availability of observational information provided as input increase, the model can reconstruct the flow field more accurately, which is consistent with the general characteristics of data-driven surrogate models.

However, this performance improvement did not continue linearly. Instead, the improvement effect clearly diminished beyond a certain level of sensor density. For example, the error reduction was relatively large from scenarios #1 to #2, whereas after scenario #5, the improvement in performance was very limited despite the addition of more sensors. This can be interpreted as indicating that, after the model has obtained a sufficient level of spatial information, additional sensor information becomes redundant or does not substantially complement the spatial patterns already learned by the model. In other words, the model can effectively reconstruct the main flow structures even at relatively low sensor densities, and further increases in information appear to contribute mainly to fine-tuning the performance. It should be noted that the sensor-density scenarios used in this study simultaneously change the number of sensor nodes and their spatial coverage. Therefore, the observed performance improvement and diminishing returns should be interpreted as the combined effect of increased sensor-node number and expanded spatial coverage, rather than as the isolated effect of sensor count alone.

In addition to the mean values, Table 5 also reports the standard deviations of the vector-based metrics across validation samples to quantify sample-to-sample variability caused by differences in sensor placement and subdomain spatial structure. The standard deviations were relatively large under sparse sensor conditions, particularly in scenarios #1 to #4, whereas they decreased after scenario #5. This indicates that increased sensor-node density not only improves the average prediction accuracy but also reduces the variability of model performance across validation cases.

From the perspective of information transfer by Propagated confidence Field, limitations were observed when only a small number of sensors were placed near the center. This tendency was particularly evident in RMSE_vec, ε_{rel,L^2}_vec, MAE (EPE) (refer to Fig. 9(a) to (c)). In scenario #1, because sensors were confined to the central region, information transfer across the entire domain was limited, resulting in the highest errors. By contrast, from scenario #5 onward, sensors were sufficiently distributed across the spatial domain, allowing the global structure of the flow field to be effectively captured and resulting in lower error levels. This suggests that not only the absolute number of sensors but also their spatial distribution is an important factor. The larger errors observed in scenarios #1 and #2 also indicate a failure

Table 5
Summary of evaluation metrics by validation scenarios.

Scenario	RMSE[a]				ε_{rel,L^2} [b]				MAE (EPE)[c] ±std	Cosine similarity ±std
	u[d]	v[e]	w[f]	vec[g] ±std	u	v	w	vec ±std		
#1	0.183	0.068	0.192	0.281 ±0.142	0.545	0.997	0.823	0.652 ±0.249	0.233 ±0.120	0.700 ±0.228
#2	0.139	0.068	0.141	0.212 ±0.120	0.410	1.002	0.639	0.490 ±0.199	0.180 ±0.103	0.793 ±0.166
#3	0.123	0.068	0.118	0.185 ±0.105	0.370	1.003	0.551	0.425 ±0.167	0.160 ±0.092	0.817 ±0.144
#4	0.107	0.068	0.095	0.159 ±0.092	0.321	1.005	0.429	0.349 ±0.101	0.137 ±0.080	0.857 ±0.076
#5	0.095	0.068	0.083	0.143 ±0.081	0.292	1.008	0.365	0.311 ±0.074	0.122 ±0.069	0.877 ±0.052
#6	0.089	0.068	0.077	0.136 ±0.076	0.277	1.011	0.342	0.298 ±0.075	0.115 ±0.064	0.881 ±0.048
#7	0.087	0.068	0.075	0.134 ±0.074	0.273	1.015	0.334	0.294 ±0.076	0.112 ±0.062	0.882 ±0.051
#8	0.085	0.068	0.074	0.132 ±0.073	0.269	1.021	0.330	0.291 ±0.079	0.111 ±0.060	0.883 ±0.056
#9	0.085	0.068	0.074	0.132 ±0.073	0.271	1.029	0.333	0.294 ±0.087	0.111 ±0.060	0.882 ±0.056

Note: RMSE[a] stands for Root Mean Squared Error; ε_{rel,L^2} [b] stands for Relative L2 Error; MAE (EPE)[c] stands for Mean Absolute Error (End-Point Error); u[d] stands for x-direction velocity component; v[e] stands for y-direction velocity component; w[f] stands for z-direction velocity component; vec[g] stands for vector.

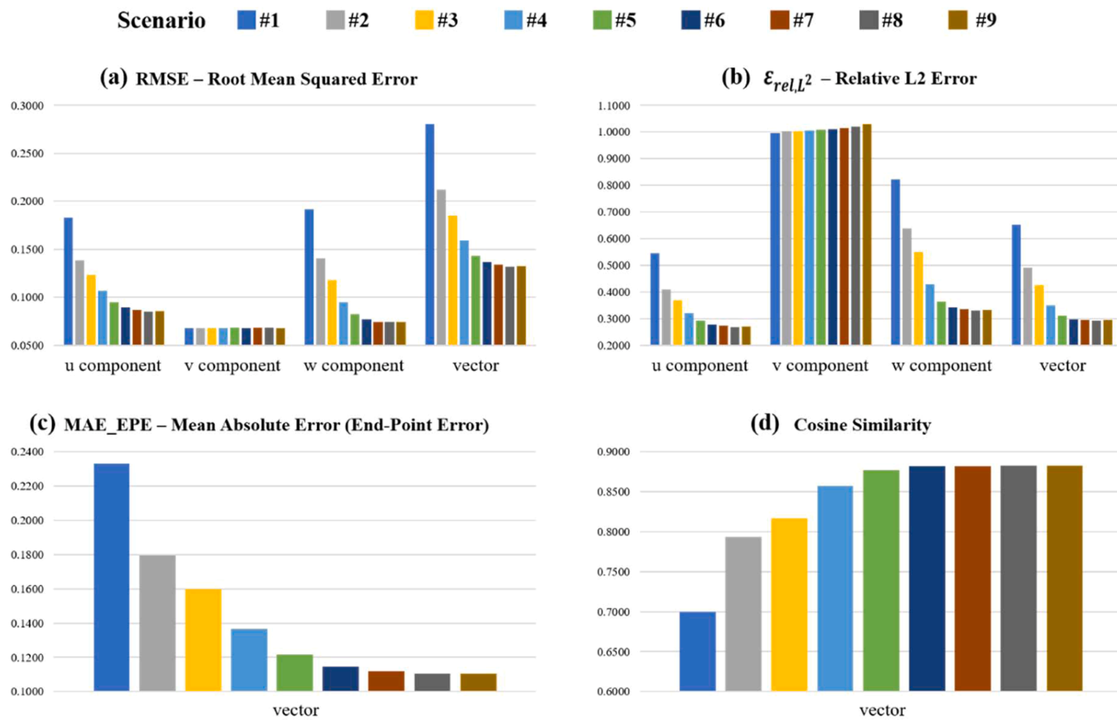

Fig. 9. Performance comparison across validation scenarios.

mode of IGNO-PF under extremely sparse and spatially concentrated sensor conditions, where some regions may not receive sufficient propagated information during the rollout process. However, the present study did not separately quantify spatial failure modes such as the low-confidence area ratio, the relationship between prediction error and distance to the nearest sensor, or the dependence of error on local building density and graph connectivity. These analyses should be conducted in future work to better characterize the limitations of IGNO-PF under poorly connected sensor configurations.

Examining the velocity components, the vertical component v showed the smallest RMSE, and its error variation with changes in sensor density was limited. This can be interpreted as the model predicting this component relatively stable because the absolute magnitude of v is smaller than that of the horizontal components and its spatial variation is relatively smooth. By contrast, the relative L2 error of the v component was the largest, because small absolute errors are reflected more strongly when the reference component magnitude is small. Therefore, although the v component may be evaluated unfavorably in terms of relative error, it can be regarded as the easiest component to predict from the perspective of absolute error. In contrast, the horizontal components u and w constitute the main flow structures in the urban wind environment and responded more sensitively to changes in sensor density. This is because horizontal flow is strongly affected by building arrangements, shielding effects, channeling effects, and recirculation regions, resulting in relatively large spatial variability. In addition, vector-based metrics reflect the combined errors of these directional components, therefore, they showed the largest errors in the initial scenarios and the clearest performance improvement as the number of sensors increased. This indicates that the reduction in errors of the horizontal components, particularly u and w, drove the overall reduction in vector error.

Finally, quantitatively comparing the absolute accuracy of the proposed model with those of existing studies is limited because the application conditions, such as mesh structure, boundary conditions, and dataset configuration, differ. Therefore, it is more appropriate to interpret the present results in terms of relative performance changes and trends associated with changes in sensor configuration rather than the absolute performance of the model. From this perspective, the proposed model maintained stable prediction performance even under limited sensor information and consistently reproduced both the performance improvement associated with increasing sensor density and its limitation, namely diminishing returns. In addition, the fact that cosine similarity remained above approximately 0.7 in all scenarios and above 0.8 in most cases indicates that the predicted flow vectors had high overall directional agreement with the reference field (refer to Fig. 9 (d)). Cosine similarity measures directional agreement between vectors, and values closer to 1 indicate that two vector fields share the same directional structure. Therefore, these results suggest that the proposed model not only reduces magnitude errors but also reasonably reproduces the major flow directions and spatial structures of the actual flow field.

It should be noted that the evaluation metrics reported in Table 5 are domain-aggregated quantities. Therefore, they may not fully capture localized reconstruction errors in complex urban-flow regions, such as street canyons, building wakes, near-wall regions, and low-velocity recirculation or stagnation zones. In particular, cosine similarity should be interpreted with caution in regions where the reference velocity magnitude is close to zero, because the physical meaning of directional agreement becomes less stable under near-stagnation conditions. To partially address this issue, an additional stratified error analysis was conducted by dividing the evaluation nodes in each validation sample into three equal-sized groups based on the CFD reference wind-speed magnitude.

Specifically, the evaluation nodes in each validation sample were divided into low, medium, and high speed tertiles according to the CFD reference wind-speed magnitude, and the evaluation metrics were computed separately for each group (refer to Table 6). As shown in Table 6, the low speed group showed a relative L2 error of 0.716, which was substantially higher than those of the medium and high speed groups. In contrast, the RMSE and MAE (EPE) values in the low speed group were comparable to those in the medium speed group. This indicates that the large relative error in the low speed group was mainly caused by the small magnitude of the reference velocity, rather than by excessively large absolute errors. In addition, cosine similarity was lowest in the low speed group, whereas it increased to 0.921 and 0.949

Table 6
Stratified error analysis based on CFD reference wind-speed tertiles.

CFD wind-speed tertile	RMSE_vec[a] ±std	ε_{rel,L^2}_vec[b] ±std	MAE (EPE)[c] ±std	Cosine similarity ±std
Low-speed	0.140 ±0.072	0.716 ±0.341	0.120 ±0.062	0.773 ±0.142
Medium-speed	0.140 ±0.083	0.364 ±0.217	0.119 ±0.075	0.921 ±0.097
High-speed	0.160 ±0.121	0.265 ±0.179	0.134 ±0.108	0.949 ±0.104

Note: RMSE_vec[a] stands for Root Mean Squared Error of vector; ε_{rel,L^2}_vec[b] stands for Relative L2 Error of vector; MAE (EPE)[c] stands for Mean Absolute Error (End-Point Error).

in the medium and high speed groups, respectively. These results support the reviewer's concern that directional metrics can be less stable in near-stagnation regions. Therefore, cosine similarity was interpreted mainly as an indicator of overall directional agreement in regions where the reference velocity magnitude is sufficiently large, rather than as a uniformly reliable local metric across all velocity regimes. Overall, the stratified analysis confirms that IGNO-PF more reliably reconstructs the main flow structures in medium and high speed regions, while localized prediction in low speed or recirculation-prone regions remains a limitation requiring further analysis.

Table 7 presents the divergence-based physical plausibility assessment for scenario #1, which corresponds to the sparsest sensor condition and thus represents the most challenging validation setting. The mean absolute divergence of the IGNO-PF predictions was generally comparable to that of the CFD reference fields across the eight validation cases. In several cases, including validation cases C2, C3, C4, C5, C6, and C7, the predicted fields even showed lower mean absolute divergence than the CFD fields. This indicates that the proposed model did not amplify the overall level of local volumetric imbalance relative to the CFD reference solution. The mean absolute divergence error ranged from 0.005397 to 0.057962, while the 95th percentile ranged from 0.016609 to 0.195825. Although relatively larger divergence discrepancies were observed in validation cases C3 and C4, these cases also showed relatively large divergence levels in the CFD reference fields, suggesting that the increased divergence error was partly associated with more complex local flow variation in the reference solutions. Importantly, the divergence-error bias remained small compared with the corresponding absolute divergence-error statistics. This indicates that the prediction errors did not consistently accumulate in a single signed direction. Overall, the divergence analysis shows that IGNO-PF did not systematically generate source-dominant or sink-dominant flow fields, even under the sparsest sensor condition. Together with the data-fidelity metrics and cosine similarity analysis, these results support the physical plausibility of the predicted velocity fields and indicate that the proposed model can reproduce the major flow structures without producing excessive artificial fluid sources or sinks.

To further examine whether the physical plausibility observed in scenario #1 was maintained across different sensor-density conditions, the divergence-based assessment was extended to all validation scenarios, as summarized in Table 8. The mean absolute divergence of the CFD reference fields remained nearly constant across scenarios, ranging from 0.022280 to 0.022335, because the same CFD reference fields were used while only the sensor configuration was changed. In contrast, the mean absolute divergence of the IGNO-PF predictions ranged from 0.012864 to 0.015465, showing values comparable to, or lower than, those of the CFD reference fields across all scenarios. This indicates that the predicted velocity fields did not amplify the overall level of local volumetric imbalance relative to the CFD reference solutions. The mean absolute divergence error also decreased as the sensor density increased, from 0.028517 in scenario #1 to 0.020304 in scenario #9. A similar trend was observed for the 95th percentile of the absolute divergence error, which decreased from 0.094300 to 0.068037. This suggests that increasing sensor density not only improved the overall velocity prediction accuracy, but also reduced local divergence discrepancies in the predicted flow fields. In addition, the divergence-error bias gradually approached zero as the scenario number increased, decreasing from 0.006284 in scenario #1 to 0.001890 in scenario #9. This indicates that the predicted fields did not show a systematic tendency toward either source-dominant or sink-dominant behavior.

Overall, the scenario-wise divergence analysis demonstrates that the physical plausibility of the IGNO-PF predictions was not limited to the sparsest sensor condition. Rather, the predicted fields maintained divergence characteristics comparable to the CFD reference fields across all sensor-density scenarios, while the divergence-error magnitude and signed bias generally decreased as more sensor information became available. These results support that IGNO-PF does not systematically generate physically unrealistic source- or sink-like flow structures across the validation scenarios.

Fig. 10 provides a validation case that intuitively illustrates the effect of the spatial distribution of sensors on prediction accuracy. Fig. 10(e) to (h) are enlarged views of the same regions shown in Fig. 10(a) to (d), respectively. In Fig. 10(a) and (e), because of insufficient information, the predicted quivers of the target nodes located on the left side differ from the observed quivers, and it can be confirmed that computation does not proceed beyond a certain region. Fig. 10(b) and (f) show that adding only a few sensor nodes near the center enables computation in regions that were not computed in Fig. 10(a) and (e). Fig. 10(c) and (g), where more sensor nodes were added, produced results that were almost identical to the observed field.

Several important implications can be drawn from these results. First, when sensors are sparsely placed (refer to Fig. 10(a), (e)), the spatial imbalance of input information prevents flow information from being sufficiently propagated to certain regions. As a result, predictions become unstable or, in some regions, are effectively not computed. This indicates that, because the proposed model has a structure that

Table 7
Divergence-based physical plausibility assessment for scenario #1.

Validation case	Mean absolute divergence of each field		Mean absolute divergence error		Divergence-error bias
	CFD simulation	IGNO-PF	Total	95th percentile	
C1	0.003004	0.004259	0.005397	0.016609	-0.000893
C2	0.033458	0.012939	0.038623	0.137890	-0.004703
C3	0.037489	0.020003	0.050334	0.161008	0.018137
C4	0.041138	0.025667	0.057962	0.195825	0.023699
C5	0.020102	0.012228	0.027820	0.085017	0.004019
C6	0.008731	0.007525	0.014396	0.051266	0.002448
C7	0.015034	0.006576	0.018083	0.061157	0.001564
C8	0.019614	0.017455	0.027912	0.082656	0.013976

Table 8
Divergence-based physical plausibility assessment by scenarios.

Scenario	Mean absolute divergence of each field		Mean absolute divergence error		Divergence-error bias
	CFD simulation	IGNO-PF	Total	95th percentile	
#1	0.022321	0.012864	0.028517	0.094300	0.006284
#2	0.022317	0.014306	0.028824	0.094143	0.007998
#3	0.022329	0.014416	0.028184	0.092463	0.007934
#4	0.022282	0.014147	0.025938	0.085977	0.007080
#5	0.022267	0.013759	0.023438	0.079190	0.004711
#6	0.022335	0.014300	0.022248	0.074788	0.003449
#7	0.022334	0.014396	0.021426	0.072444	0.002960
#8	0.022316	0.015309	0.020821	0.070169	0.001984
#9	0.022280	0.015465	0.020304	0.068037	0.001890

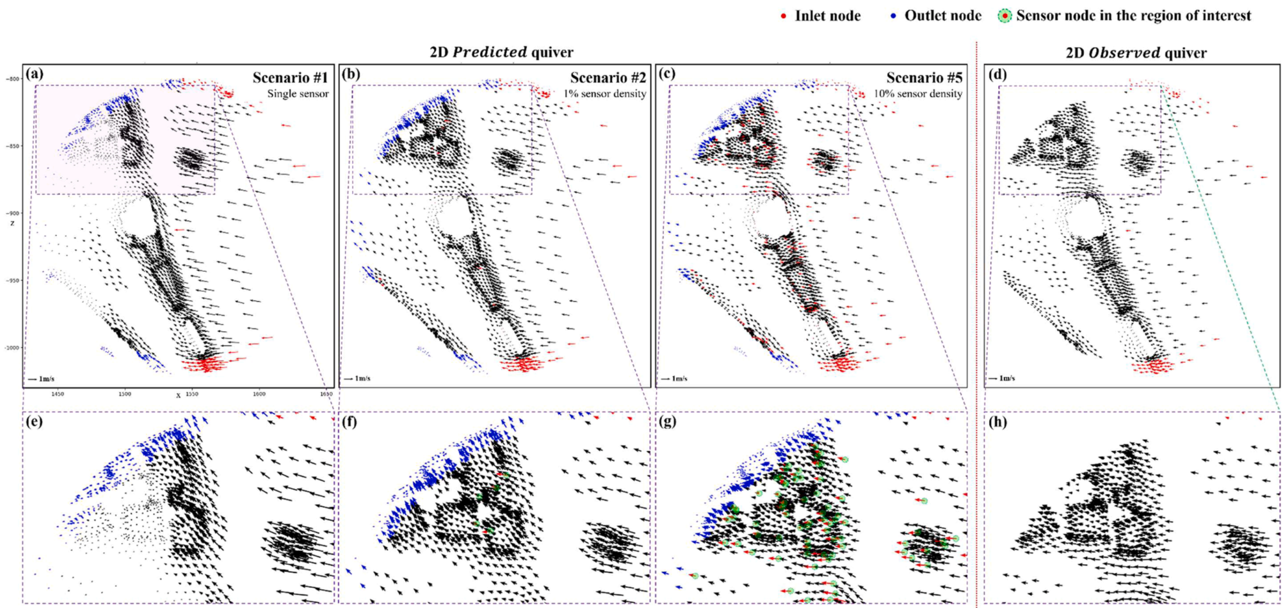

Fig. 10. Improvement of the predicted quiver based on the spatial distribution of sensors.

gradually diffuses spatial information based on the input observations, insufficient initial information prevents its influence from being adequately transmitted throughout the domain.

By contrast, when a small number of sensors are added and information connectivity is secured in the central region (refer to Fig. 10(b), (f)), the flow field is continuously reconstructed even in regions where prediction was previously not performed. This can be interpreted not simply as a result of an increased number of sensors, but as the formation of a global information-transfer path as spatially disconnected information flows become connected. In other words, the model performance depends strongly not only on the absolute number of sensors but also on their positions and connectivity within the domain.

In addition, when sensors are sufficiently distributed (refer to Fig. 10 (c), (g)), the predicted flow field is reproduced at a level almost identical to the observed values, demonstrating that the proposed model can precisely reconstruct the actual flow structure when sufficient input information is provided. However, when considered together with the preceding results, this performance improvement tends to saturate beyond a certain level of sensor density, which is consistent with the diminishing returns observed in the quantitative metrics.

Table 9 presents a post-hoc comparison designed to partially separate the effects of sensor number and spatial coverage under the sparsest sensing condition. Compared with scenario #1, the #1-Center configuration increased the number of sensor nodes by adding six sensors within 3 m of the original S-DoT sensor, while maintaining nearly the same spatial coverage. However, this configuration showed almost identical performance to scenario #1. This indicates that simply increasing the number of sensor nodes near the same location does not substantially improve reconstruction accuracy. The #1-Inlet configuration also did not improve the prediction accuracy. This can be interpreted as indicating that, because inflow information is already provided through the prescribed inlet boundary condition and propagated into the domain during the iterative rollout, additional sensors located near the inlet side provide limited complementary information for full-field reconstruction. In contrast, the #1-Outlet configuration showed the best performance among the five sparse-sensor configurations. Because regions farther from the inlet boundary are more weakly constrained by the prescribed inflow information and are more affected by accumulated reconstruction uncertainty, additional outlet-side sensors can provide more complementary information for reconstructing the downstream flow field. These results suggest that the prediction performance of

Table 9
Post-hoc comparison of sensor number and spatial coverage effects.

Scenario	Additional sensor configuration	RMSE_vec[a] ±std	ε_{rel,L^2}_vec[b] ±std	MAE (EPE)[c] ±std	Cosine similarity ±std
#1	S-DoT only	0.281 ±0.142	0.652 ±0.249	0.233 ±0.120	0.700 ±0.228
#1-Center	6 sensors within 3 m of S-DoT	0.282 ±0.142	0.651 ±0.247	0.233 ±0.120	0.702 ±0.225
#1-Global	6 sensors globally distributed	0.280 ±0.141	0.646 ±0.245	0.232 ±0.120	0.706 ±0.220
#1-Inlet	6 sensors located inlet-side	0.283 ±0.143	0.655 ±0.248	0.235 ±0.121	0.698 ±0.227
#1-Outlet	6 sensors located outlet-side	0.276 ±0.138	0.640 ±0.245	0.228 ±0.116	0.708 ±0.222

Note: RMSE_vec[a] stands for Root Mean Squared Error of vector; ε_{rel,L^2}_vec[b] stands for Relative L2 Error of vector; MAE (EPE)[c] stands for Mean Absolute Error (End-Point Error).

IGNO-PF is affected not only by the number of sensor nodes but also by whether the added sensors provide complementary spatial information in regions where reconstruction uncertainty is relatively high. Nevertheless, because this analysis was conducted as a post-hoc comparison using selected sparse-sensor configurations, it should be interpreted as partial evidence rather than a complete factorial sensitivity analysis. A fully controlled sensor-placement experiment remains necessary to rigorously isolate the individual effects of sensor count and spatial coverage.

Consequently, this experiment demonstrates that the proposed model has an information-propagation-based structure and that the spatial distribution of sensors is a key factor determining both the predictable region and the prediction accuracy. In particular, under limited sensor conditions, strategic placement that enables global information transfer may be more effective than simply increasing the number of sensors.

4. Conclusion

This study proposed an iterative Graph Neural Operator with Propagated confidence Field (IGNO-PF) framework for steady-state CFD surrogate modeling on unstructured meshes. The proposed approach is characterized by explicitly introducing an iterative update structure, unlike conventional single-step regression methods, based on the observation that steady-state CFD analysis can be interpreted as a process that converges to a fixed point through iterative numerical computation. As a result, the flow field is not determined by a single prediction; rather, it is designed to be gradually refined through iterative operations and to converge to the final solution.

The key component of the proposed model is the Propagated confidence Field, which serves as a global indicator of the extent to which information from source nodes such as inlet and sensor nodes has diffused within the computational domain. This field is used as an input feature without additional physical constraints or explicit gating, allowing the model to adaptively regulate its use of spatial information during the iterative computation process. Through this design, the proposed framework can simultaneously approximate the spatial transmission process of flow information and the convergence behavior of steady-state analysis.

Evaluation results across various validation scenarios showed that IGNO-PF maintained stable prediction performance under different sensor-density and sensor-placement conditions. Even when the number of sensors was limited or the sensors were spatially non-uniformly distributed, the model generally reproduce the overall flow structure. Although the absolute error levels varied depending on sensor density and placement, cosine similarity remained above approximately 0.7 in all scenarios and above 0.8 in most cases, confirming that the predicted flow vectors had high directional agreement with the reference solution. In addition, the divergence-based assessment supported the physical plausibility of the predicted velocity fields, while the stratified analysis based on CFD reference wind-speed tertiles showed that prediction errors and directional metrics should be interpreted with respect to the local magnitude of the reference flow.

This study also confirmed that prediction performance depends strongly not only on the absolute number of sensors but also on their spatial distribution. Sensor arrangements that enable the formation of information-propagation paths play a key role in global flow-field reconstruction, and this is closely related to the structural characteristics of the proposed model. Meanwhile, although overall prediction performance improved as sensor density increased, the magnitude of improvement gradually decreased beyond a certain level, which can be interpreted as indicating that additional sensor information makes only a limited contribution after the major flow structures have already been sufficiently reflected.

The proposed IGNO-PF framework offers several practical advantages over existing data-driven CFD surrogate models. First, it can directly process unstructured meshes without requiring separate mesh conversion or interpolation, and its low dependence on the discretization scheme enables flexible responses to various sensor-placement conditions. These characteristics enhance its applicability in real environments where sensor locations and numbers may change dynamically. However, the present validation assumed that the sensor observations were available and reliable, and did not explicitly consider sensor noise, systematic bias, missing observations, or outlier contamination.

Future studies should extend the proposed framework in three main directions. First, the physical consistency and robustness of the Propagated confidence Field should be further improved. In this study, the field was constructed using a predefined heuristic formulation to represent the availability and propagation status of source information; therefore, it should not be interpreted as a physical flow variable or a statistically calibrated uncertainty measure. Future work should explore data-driven or physics-informed propagation mechanisms and evaluate the robustness of IGNO-PF under imperfect sensing conditions, including sensor noise, bias, missing observations, and outlier contamination.

Second, more systematic sensitivity and ablation analyses are needed to quantify the effects of rollout iterations, neighborhood construction, propagation-matrix design, confidence boosting, and intermediate-supervision strategies on convergence stability, computational cost, and prediction accuracy. Such analyses would further clarify the trade-offs between iterative refinement, computational efficiency, and model complexity.

Third, the generalizability of IGNO-PF should be evaluated under broader CFD conditions. Although the validation cases covered all eight wind-direction conditions and were spatially separated from the training and test subdomains, the present dataset was generated from a single urban area and the same CFD mesh system. Therefore, the current results should be interpreted as spatially independent validation within the same urban CFD system, rather than unrestricted generalization to different urban morphologies, mesh densities, or mesh-partitioning strategies. Future work should extend the CFD dataset to include diverse urban morphologies, mesh configurations, and controlled validation settings for rotational consistency, including consistently rotated geometry, inlet directions, sensor layouts, coordinate systems, and velocity vectors. In addition, controlled experiments should further examine sensor-placement effects, localized failure modes in street canyons, building wakes, near-wall and low-speed regions, and the influence of local urban morphology and graph connectivity on reconstruction performance.

CRediT authorship contribution statement

Minjin Kong: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Jinwoo Choi:** Visualization, Validation, Investigation, Formal analysis, Data curation. **Samir Shrestha:** Visualization, Validation, Formal analysis. **Sobit Sapkota:** Visualization, Validation, Formal analysis. **Jun-Ki Choi:** Writing – review & editing, Validation, Supervision. **Taehoon Hong:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing 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.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT; Ministry of Science and ICT) (NRF-2021R1A3B1076769).

Data availability

Data will be made available on request.

References

Aduda, K. O., Labeodan, T., Zeiler, W., Boxem, G., & Zhao, Y. (2016). Demand side flexibility: Potentials and building performance implications. *Sustainable Cities and Society, 22*. https://doi.org/10.1016/j.scs.2016.02.011

Aliero, M. S., Qureshi, K. N., Pasha, M. F., & Jeon, G. (2021). Smart home energy management systems in internet of things networks for green cities demands and services. In *Environmental Technology and Innovation* (Vol. 22). https://doi.org/10.1016/j.eti.2021.101443.

Azadeh, A., Ghaderi, S. F., & Sohrabkhani, S. (2008). Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. *Energy Conversion and Management,* (8), 49. https://doi.org/10.1016/j.enconman.2008.01.035

Baek, M., Choi, S., Kim, H., Jang, H., Kim, J., Hong, T., & Hong, J. (2026). CFD numerical setting combinations for greenhouse natural ventilation across diurnal time points using multi-criteria analysis. *Computers and Electronics in Agriculture, 247*, Article 111689. https://doi.org/10.1016/J.COMPAG.2026.111689

Bahmani, B., Goswami, S., Kevrekidis, I. G., & Shields, M. D. (2025). A resolution independent neural operator. *Computer Methods in Applied Mechanics and Engineering, 444*. https://doi.org/10.1016/j.cma.2025.118113

Blocken, B. (2015). Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. *Building and Environment, 91*. https://doi.org/10.1016/j.buildenv.2015.02.015

Brandstetter, J., Worrall, D. E., & Welling, M. (2022). Message passing neural pde solvers. In *ICLR 2022 - 10th International Conference on Learning Representations*.

Calzolari, G., & Liu, W. (2021). Deep learning to replace, improve, or aid CFD analysis in built environment applications: A review. In *Building and Environment* (Vol. 206). https://doi.org/10.1016/j.buildenv.2021.108315.

Cao, Y., Xie, P., Huang, G., Wang, W., Chen, W. L., Hu, G., & Cao, S. (2025). A Pattern-constrained deep learning model for urban canopy turbulence reconstruction from sparse sensor data. *Building and Environment, 285*. https://doi.org/10.1016/j.buildenv.2025.113535

Chang, C., Chen, C. Y., & Lin, T. P. (2024). Combining geographic information and climate data to develop urban building energy prediction models in Taichung, Taiwan. *Sustainable Cities and Society, 117*. https://doi.org/10.1016/j.scs.2024.105949

Chen, T., Li, R., Hu, X., Zhang, B., Liu, Y., Wang, L. (Leon), & Gao, N. (2025). Machine learning as CFD surrogate models for rapid prediction of building-related physical fields: A review of methods and state-of-the-art. In *Building and Environment* (Vol. 285). https://doi.org/10.1016/j.buildenv.2025.113667.

Choi, J., Hong, J., & Hong, T. (2026). Quantitative health risk assessment of benzene exposure from painting activities in urban construction using CFD: Seasonal air quality impacts. *Renewable and Sustainable Energy Reviews, 225*, Article 116149. https://doi.org/10.1016/J.RSER.2025.116149

Choi, J., Kong, M., Choi, D., Seo, S., Koo, C., & Hong, T. (2026). A systematic review on the integrations of CFD and artificial intelligence for the future perspectives of built environment: Multi-scale approaches from room to city. *Sustainable Cities and Society, 141*, Article 107278. https://doi.org/10.1016/J.SCS.2026.107278

Du, H., Ni, Y. Q., Chen, Z. W., Zeng, Y. J., & Chen, X. Y. (2026). ST-SwinUNet: Real-time image-to-image reconstruction and multi-step forecasting of multi-height urban wind fields. *Building and Environment, 298*, Article 114593. https://doi.org/10.1016/J.BUILDENV.2026.114593

Fan, C., Zou, B., Li, J., Wang, M., Liao, Y., & Zhou, X. (2024). Exploring the relationship between air temperature and urban morphology factors using machine learning under local climate zones. *Case Studies in Thermal Engineering, 55*. https://doi.org/10.1016/j.csite.2024.104151

Gao, H., Hu, G., Zhang, D., Jiang, W., Tse, K. T., Kwok, K. C. S., & Kareem, A. (2024). 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. https://doi.org/10.1111/MICE.13147

Geng, X., & Gou, Z. (2025). Automated framework for semantic segmentation and CFD-ready geometric reconstruction of urban elements from UAV data. *Sustainable Cities and Society, 134*. https://doi.org/10.1016/j.scs.2025.106948

Giral, F., Manzano, Á., Gómez, I., Koumoutsakos, P., & Clainche, S. Le (2025). Generative urban flow modeling: from geometry to airflow with graph diffusion. https://arxiv.org/pdf/2512.14725.

Goe, M., & Gaustad, G. (2014). Strengthening the case for recycling photovoltaics: An energy payback analysis. *Applied Energy*. https://doi.org/10.1016/j.apenergy.2014.01.036

Guo, M., Choi, Y., Cheong, S. M., & O'Neill, Z. (2025). Current and future residential electricity demand using large-scale smart meter data in a changing climate. *Sustainable Cities and Society, 130*. https://doi.org/10.1016/j.scs.2025.106623

Hågbo, T. O., & Giljarhus, K. E. T. (2024). Sensitivity of urban morphology and the number of CFD simulated wind directions on pedestrian wind comfort and safety assessments. *Building and Environment, 253*. https://doi.org/10.1016/j.buildenv.2024.111310

Hågbo, T. O., Giljarhus, K. E. T., & Hjertager, B. H. (2021). Influence of geometry acquisition method on pedestrian wind simulations. *Journal of Wind Engineering and Industrial Aerodynamics*, 215. https://doi.org/10.1016/j.jweia.2021.104665

Hajdukiewicz, M., González Gallero, F. J., Mannion, P., Loomans, M. G. L. C., & Keane, M. M. (2024). A narrative review to credible computational fluid dynamics models of naturally ventilated built environments. In *Renewable and Sustainable Energy Reviews* (Vol. 198). https://doi.org/10.1016/j.rser.2024.114404.

Hao, Z., Wang, Z., Su, H., Ying, C., Dong, Y., Liu, S., Cheng, Z., Song, J., & Zhu, J. (2023). GNOT: A general neural operator transformer for operator learning. *Proceedings of Machine Learning Research, 202*.

Huang, Z., Li, T., Huang, K., Ke, H., Lin, M., & Wang, Q. (2022). Predictions of flow and temperature fields in a T-junction based on dynamic mode decomposition and deep learning. *Energy*, 261. https://doi.org/10.1016/j.energy.2022.125228

Jeon, J., Lee, J., Vinuesa, R., & Kim, S. J. (2024). Residual-based physics-informed transfer learning: A hybrid method for accelerating long-term CFD simulations via deep learning. *International Journal of Heat and Mass Transfer*, 220. https://doi.org/10.1016/j.ijheatmasstransfer.2023.124900

Jie, P., Li, G., Wang, J., Santamouris, M., Du, S., Zhu, S., Yue, H., Feng, G., Hu, M., & Shi, X. (2026). WRF-based hybrid model for high-resolution prediction of urban thermal-wind environments under present and future climate scenarios. *Sustainable Cities and Society, 142*, Article 107315. https://doi.org/10.1016/J.SCS.2026.107315

Kastner, P., & Dogan, T. (2020). A cylindrical meshing methodology for annual urban computational fluid dynamics simulations. *Journal of Building Performance Simulation, 13*(1). https://doi.org/10.1080/19401493.2019.1692906

Kastner, P., & Dogan, T. (2023). A GAN-based surrogate model for instantaneous urban wind flow prediction. *Building and Environment, 242*. https://doi.org/10.1016/j.buildenv.2023.110384

Kong, M., Jung, S., Hong, T., Lee, J., & Jeong, K. (2025). Establishment of real-time individual room-level heating and cooling energy benchmark via thermal environment characteristics of the room: Case study in cooling season. *Journal of Building Engineering, 108*, Article 112943. https://doi.org/10.1016/J.JOBE.2025.112943

Koo, C., Park, S., Hong, T., & Park, H. S. (2014). An estimation model for the heating and cooling demand of a residential building with a different envelope design using the finite element method. *Applied Energy, 115*. https://doi.org/10.1016/j.apenergy.2013.11.014

Kops, R. B., Papa, R., Sêcco, N. R., & Malatesta, V. (2025). CFD-based surrogate modelling and optimization of the entrainment efficiency of supersonic air ejectors with temperature constraint. *Thermal Science and Engineering Progress, 67*. https://doi.org/10.1016/j.tsep.2025.104051

Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2023). Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research, 24*.

Lee, H. J., & Kim, J. J. (2026). Enhancing geometric fidelity and wind-field prediction accuracy in complex urban environments using a Cartesian cut-cell CFD model. *Sustainable Cities and Society, 140*, Article 107247. https://doi.org/10.1016/J.SCS.2026.107247

Li, S., Yu, X., Xing, W., Kirby, R. M., Narayan, A., & Zhe, S. (2024). Multi-resolution active learning of fourier neural operators. *Proceedings of Machine Learning Research, 238*.

Liu, J., & Niu, J. (2016). CFD simulation of the wind environment around an isolated high-rise building: An evaluation of SRANS, LES and DES models. *Building and Environment, 96*. https://doi.org/10.1016/j.buildenv.2015.11.007

Liu, X. Y., Zhu, M., Lu, L., Sun, H., & Wang, J. X. (2024). Multi-resolution partial differential equations preserved learning framework for spatiotemporal dynamics. *Communications Physics, 7*(1). https://doi.org/10.1038/s42005-024-01521-z

Liu, Y., & Chen, Y. (2025). DragSolver: A multi-scale transformer for real-world automotive drag coefficient estimation. *Proceedings of Machine Learning Research, 267*.

Liu, Z., Zhang, S., Shao, X., & Wu, Z. (2023). Accurate and efficient urban wind prediction at city-scale with memory-scalable graph neural network. *Sustainable Cities and Society*, 99. https://doi.org/10.1016/j.scs.2023.104935

Mao, R., Lan, Y., Liang, L., Yu, T., Mu, M., Leng, W., & Long, Z. (2025). Rapid CFD prediction based on machine learning surrogate model in built environment: A review. In *Fluids* (Vol. 10, Number 8). https://doi.org/10.3390/fluids10080193.

McCabe, M., Harrington, P., Subramanian, S., & Brown, J. (2023). Towards stability of autoregressive neural operators. *Transactions on Machine Learning Research, 2023*.

Peng, W., Qin, S., Yang, S., Wang, J., Liu, X., & Wang, L.(Leon) (2024). Fourier neural operator for real-time simulation of 3D dynamic urban microclimate. *Building and Environment, 248*. https://doi.org/10.1016/j.buildenv.2023.111063

Qi, X., Yang, M., & Tong, G. (2025). Predicting drag coefficients of vehicle geometries: A PointNet++ point cloud surrogate model approach. *AIP Advances, 15*(10). https://doi.org/10.1063/5.0300798

Schatz, J., & Kucharik, C. J. (2016). Urban heat island effects on growing seasons and heating and cooling degree days in Madison, Wisconsin USA. *International Journal of Climatology, 36*(15). https://doi.org/10.1002/joc.4675

Seoul Metropolitan Government. (2026). Smart Seoul Data of Things (S-DoT) Environmental Information. https://data.seoul.go.kr/dataList/OA-15969/S/1/datasetView.do.

Shao, X., Liu, Z., Zhang, S., Zhao, Z., & Hu, C. (2023). PIGNN-CFD: A physics-informed graph neural network for rapid predicting urban wind field defined on unstructured mesh. *Building and Environment, 232*. https://doi.org/10.1016/j.buildenv.2023.110056

Shirzadi, M., Tominaga, Y., & Mirzaei, P. A. (2020). Experimental and steady-RANS CFD modelling of cross-ventilation in moderately-dense urban areas. *Sustainable Cities and Society, 52*, Article 101849. https://doi.org/10.1016/J.SCS.2019.101849

Stamou, A., & Katsiris, I. (2006). Verification of a CFD model for indoor airflow and heat transfer. *Building and Environment, 41*(9). https://doi.org/10.1016/j.buildenv.2005.06.029

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting - A review. In *Renewable and Sustainable Energy Reviews* (Vol. 16, Number 2). https://doi.org/10.1016/j.rser.2011.08.014.

Tang, H., Kong, Q., & Morris, J. P. (2024). Multi-fidelity Fourier neural operator for fast modeling of large-scale geological carbon storage. *Journal of Hydrology, 629*. https://doi.org/10.1016/j.jhydrol.2024.130641

Tominaga, Y., Mochida, A., Yoshie, R., Kataoka, H., Nozu, T., Yoshikawa, M., & Shirasawa, T. (2008). AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. *Journal of Wind Engineering and Industrial Aerodynamics, 96*(10–11). https://doi.org/10.1016/j.jweia.2008.02.058

Toparlar, Y., Blocken, B., Maiheu, B., & van Heijst, G. J. F. (2017). A review on the CFD analysis of urban microclimate. In *Renewable and Sustainable Energy Reviews* (Vol. 80). https://doi.org/10.1016/j.rser.2017.05.248.

Vinuesa, R., & Brunton, S. L. (2022). Enhancing computational fluid dynamics with machine learning. *Nature Computational Science, 2*(6). https://doi.org/10.1038/s43588-022-00264-7

Wang, H., Wang, H., Gao, F., Zhou, P., & Zhai, Z. (John). (2018). Literature review on pressure–velocity decoupling algorithms applied to built-environment CFD

simulation. In *Building and Environment* (Vol. 143). https://doi.org/10.1016/j.buildenv.2018.07.046.
Wi, Y. M. (2017). A study on the baseline load estimation method using heating degree days and cooling degree days adjustment. *Transactions of the Korean Institute of Electrical Engineers, 66*(5). https://doi.org/10.5370/KIEE.2017.66.5.745
Wu, H., Chen, Y., Xie, P., Zhou, D., Tamura, T., Zhang, K., Cao, S., & Cao, Y. (2023). Sparse-measurement-based peak wind pressure evaluation by super-resolution convolutional neural networks. *Journal of Wind Engineering and Industrial Aerodynamics, 242*. https://doi.org/10.1016/j.jweia.2023.105574
Wu, S., Wang, H., & Luo, K. H. (2024). A robust autoregressive long-term spatiotemporal forecasting framework for surrogate-based turbulent combustion modeling via deep learning. *Energy and AI, 15*. https://doi.org/10.1016/j.egyai.2023.100333
Yoo, S., Koo, J., Choi, D., & Sung, M. (2026). DiffusionRollout: Uncertainty-Aware Rollout Planning in Long-Horizon PDE Solving. https://arxiv.org/pdf/2602.13616.
Zhao, L., Abuduwayiti, A., Zhu, Y., Liu, K., Zhang, R., Xu, C., & Xu, X. (2026). Climate-driven changes in city-scale building energy demand under urban-adjusted future climate scenarios. *Sustainable Cities and Society, 143*, Article 107369. https://doi.org/10.1016/J.SCS.2026.107369
Zhao, L., Zhou, Q., Li, M., & Wang, Z. (2024). Evaluating different CFD surrogate modelling approaches for fast and accurate indoor environment simulation. *Journal of Building Engineering, 95*. https://doi.org/10.1016/j.jobe.2024.110221
Zhao, R., Liu, S., Liu, J., Jiang, N., & Chen, Q. (2024). A two-stage CFD-GNN approach for efficient steady-state prediction of urban airflow and airborne contaminant dispersion. *Sustainable Cities and Society, 112*. https://doi.org/10.1016/j.scs.2024.105607
Zhao, Z., Li, H., & Wang, S. (2024). Machine learning-based surrogate models for fast impact assessment of a new building on urban local microclimate at design stage. *Building and Environment, 266*. https://doi.org/10.1016/j.buildenv.2024.112142