



Physics-inspired and data-driven two-stage deep learning approach for wind field reconstruction with experimental validation



Yi Liu ^{a,b,1}, Ranpeng Wang ^{a,b,1}, Yin Gu ^{a,b,1}, Congjian Li ^c, Gangqiao Wang ^{d,*}

^a Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, 100084, China

^b School of Safety Science, Tsinghua University, Beijing, 100084, China

^c Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100094, China

^d Institute of Big Data Intelligent Management and Decision, College of Management, Shenzhen University, Shenzhen, 518055, China

ARTICLE INFO

Handling editor: Jesse L. The

Keywords:

Wind field reconstruction
Deep learning
CFD database
Physics-inspired neural networks
Data-driven error correction

ABSTRACT

Accurate and reliable wind forecasts for urban blocks play a pivotal role in the construction of zero-energy communities by guiding the selection and placement of wind turbines and the aerodynamic design optimization of ducted openings. While relatively accurate wind fields are available based on numerical methods, their heavy computational cost and discontinuity make it necessary to explore an interactive and end-to-end method. In this study, we develop a physics-inspired and data-driven two-stage deep learning approach that can reconstruct complex wind fields precisely. The proposed method integrates a physical feature extraction model of the flow field with a sparse measurement data-driven error correction approach. In particular, a well-designed and well-trained flow field feature extraction model (original model) can preserve salient features of CFD modelling, while data-driven error correction techniques may harvest the uncertainty features and fill the remaining gaps between the original model predictions and the measured data. The proposed method is verified by a measured dataset from a community in Beijing. Experimental validation illustrates that the proposed algorithm successfully accomplishes wind field reconstruction in complex terrains using sparse datasets. We show that the proposed two-stage strategy exhibits significantly improved prediction results over the purely original method, with an average accuracy improvement of 47.17% and a maximum accuracy improvement of 72.59%. Overall, the proposed method delivers the potential in accurate wind field construction and urban wind energy forecasting.

1. Introduction

The global greenhouse effect has intensified in recent years, prompting a concerted effort by the international community to achieve zero greenhouse gas emissions by 2050 [1]. Buildings are major contributors, consuming 30–40% of the world's energy and generating nearly 40% of global carbon dioxide emissions [2]. The development of zero-energy communities holds significant potential for mitigating global warming and has garnered global attention [3]. Notably, the European Union is actively promoting the Directive on Energy Performance of Buildings (EPBD), with the goal of decarbonizing the building stock and achieving net-zero energy buildings (Net ZEB) or nearly net-zero energy buildings (NZEBs) by 2050 [4]. Given that a substantial portion, approximately 75–90%, of the current building stock will still be in use by 2050, there is an imperative for the renovation of existing

buildings [5].

Over the past decades, extensive research has focused on retrofitting existing buildings to achieve zero carbon [6] or zero energy status [7]. Various renewable energy technologies, such as solar, geothermal, and wind energy [8], have been integrated into building systems to facilitate the transformation into zero energy structures [9]. Wind energy, in particular, has emerged as a mature technology compared to other renewables [10], making urban wind energy development an attractive option [11]. Harvesting urban wind energy involves installing wind turbines on existing building roofs [12] or as standalone structures within urban environments [13]. These small-scale wind energy systems offer decentralized, renewable, and sustainable energy for buildings, reducing energy costs by eliminating the inefficiencies and expenses associated with long-distance energy transmission [14].

The layout and selection of wind turbines [15] and the aerodynamic

* Corresponding author.

E-mail address: wanggq@szu.edu.cn (G. Wang).

¹ These authors contributed equally to this paper.

design optimization of air duct openings [16] in buildings are the key problems in engineering practice. Accurate and reliable wind forecasts for city blocks can provide effective guidance for this problem. However, the transient, random, and uncertain characteristics of the wind field, coupled with the complex geometry of urban blocks, makes accurate reconstruction and prediction of wind field distributions a highly complex and nonlinear process. Recently, numerous researches have been carried out to improve the prediction accuracy. Various reliable and efficient prediction methods have been developed, which can be broadly categorized into physical methods and data-based methods [17,18]. Generally, these physical methods [19–22] are usually developed to solve the partial differential equations that are associated with the physical modelling. However, the fundamental limitation of existing methods is their excessive computational cost and their inability to take into account uncertainties that exist under real conditions. This constraint not only limits the method from achieving dynamic simulation in an interactive manner, but also interferes with the end-to-end process of the engineering problem.

To overcome the shortcomings described above, data-based approaches including time series models [23], machine learning models [24–26] and hybrid models [27,28] have been proposed. In particular, data-driven deep learning (DL) methods have garnered widespread attention for their ability to extract and learn advanced features from data distributions via constructed deep neural networks, as well as their excellent performance in prediction accuracy [18]. With the development of various deep learning algorithms, several studies have been conducted to address the problem of wind field prediction using DL frameworks, such as the transformer neural networks using integrated spatio-temporal correlation [29], the transformer-based model with wavelet transform [30], and Bayesian framework-based bidirectional gated logic unit method [31].

Although previous DL methods have shown reasonable results for short-term prediction problems of wind speed or direction at given points, the optimal design of wind field layouts generally requires predicting the entire wind field in the presence of only sparsely dispersed sensor measurements available, which is not feasible to achieve using such supervised learning methods [32]. Furthermore, such simple data-driven methods may provide physically infeasible results, especially in the field of fluids, where the results should follow recognized physical rules.

These needs to address the limitations of data-driven methods have facilitated the emergence of physical-inspired neural networks, e.g., Sekar [33] developed a fast flow field prediction model of two-dimensional airfoil under different aerodynamic profiles. Deng et al. [34] accomplished a super-resolution reconstruction of the flow field using super-resolution generative adversarial network (SRGAN) and enhanced-SRGAN. Jin et al. [35] proposed a data-driven model for predicting the velocity field around a cylinder based on the close relationship among the Reynolds stresses in the wake, wake formation length and base pressure [36]. Such methods ensure that the trained network approximates the fluid-physical laws of the modelled flow field, such as the N–S equations, by taking full consideration of the fluid-physical mechanisms in the design of the network structure and the selection of the network input features and outputs [37]. While the aforementioned studies have demonstrated the feasibility of employing a physics-inspired neural network framework to address flow field prediction challenges, the related research remains at an early stage of development. Moreover, it is noteworthy that these methodologies typically rely on simulated datasets, sharing a common limitation with numerical simulation methods in their incapacity to comprehensively integrate the uncertainties inherent in real conditions. Consequently, there exists substantial potential for enhancing physics-inspired neural network methods and incorporate data error corrections to effectively tackle issues associated with wind reconstruction and short-term prediction in various complex terrains.

In this study, we develop a physics-inspired and data-driven two-

stage deep learning approach that can reconstruct complex wind fields precisely and efficiently, even in the presence of high-dimensional input uncertainties. The proposed method integrates a physical feature extraction model (original model) of the flow field with a sparse measurement data-driven error correction approach. Firstly, we have meticulously crafted the model during the stages of network input feature selection and network architecture design, and have carefully trained the model to preserve salient features of CFD modelling. Secondly, considering the uncertainties present in real conditions, an error correction method driven by sparse measured data has been proposed to fill the remaining gaps between the original model predictions and the measured data. Furthermore, utilizing a specific area in Beijing as a case study, we conducted experiments to perform performance testing on the technology proposed in this paper. These results are promising and show possibilities in wind speed reconstruction or long-term forecasting.

The remaining part of this paper is organized as follows: the fundamentals of the adopted CFD method and neural network model are described in Section 2. The physics-inspired and data-driven deep learning method proposed in Section 3. Section 4 gives the CFD dataset generation scheme, as well as the measured data acquisition and pre-processing process. Model validation is carried out in Section 5, demonstrating the wind field reconstruction results for a complex region. Finally, the conclusions are given in Section 6.

2. Related methods

This section briefly describes the fundamentals of the numerical simulation and machine learning methods used in this work, i.e., fluid flow model and neural network model.

2.1. Fluid flow model

Fluid flow adheres to the principles of mass conservation, momentum conservation, and energy conservation. The mass conservation equation delineates the temporal alteration of fluid mass, elucidating that the disparity between the mass entering and exiting a control volume per unit time equates to the change in mass within the control volume. The mass conservation equation is expressed as follows:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0 \quad (1)$$

The momentum conservation equation indicating that the sum of forces acting on a given system is equal to the rate of change of momentum of that system over time. The momentum conservation equation can be expressed as follows:

$$\left\{ \begin{array}{l} \rho \left(u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} \right) = \rho f_x - \frac{\partial p}{\partial x} + \mu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \\ \rho \left(u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} \right) = \rho f_y - \frac{\partial p}{\partial y} + \mu \left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \right) \\ \rho \left(u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} \right) = \rho f_z - \frac{\partial p}{\partial z} + \mu \left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \right) \end{array} \right. \quad (2)$$

The energy conservation equation delineates the dynamics of energy variation within fluid flow. The increase in internal energy of a fluid element is composed of three parts: energy entering the element due to heat conduction, energy generated by the element itself, and work done on the element by the surrounding fluid. The energy conservation equation is expressed as follows:

$$\lambda \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) = c_p \rho \left[\frac{\partial T}{\partial t} + \frac{\partial(uT)}{\partial x} + \frac{\partial(vT)}{\partial y} + \frac{\partial(wT)}{\partial z} \right] \quad (3)$$

Above these equations, u , v and w represent the component velocities of the wind in the x , y and z directions, respectively. ρ denotes the fluid

density, p stands for pressure, f represents the external force per unit volume acting on the element, μ is the viscosity coefficient, λ is the second viscosity coefficient, c_p is the specific heat capacity of the fluid, T is the fluid temperature.

The governing equations in practical problems are complex, making it challenging to obtain analytical solutions. Therefore, numerical methods are commonly employed to approximate the solutions of physical quantities. In this study, high-fidelity wind field data in complex terrains are obtained using the open-source Computational Fluid Dynamics (CFD) software OpenFOAM, as detailed in Section 4.1.

2.2. Neural network model

Neural networks possess formidable learning capabilities, enabling them to discern statistical patterns in vast datasets that are beyond the capacity of manual extraction. Employing neural networks for the wind field reconstruction involves utilizing known measurement data as input and the desired points for reconstruction as output. With extensive datasets, the network autonomously learns the characteristic parameters of the flow process. As an example, the procedure for reconstructing the flow field through the application of a backpropagation (BP) neural network, without explicitly addressing the hierarchical structure of the network, can be articulated as follows:

$$\tilde{\varphi}(x, t) = f \left(\sum_{i=1}^n w_i x_i + \theta \right) \quad (4)$$

Where x_i represents the known data, w_i and θ are the network parameters to be learned. The deviation between the output results and the actual values is given by:

$$\epsilon = \varphi(x, t) - \tilde{\varphi}(x, t) \quad (5)$$

During the learning phase, neural networks utilize optimization algorithms such as stochastic gradient descent to propagate the deviation ϵ layer by layer forward, achieving parameter learning and minimizing errors.

3. Physics-inspired and data-driven two-stage deep learning methods

In this section, we propose a two-stage deep learning approach for wind field reconstruction that integrates physics-inspired and data-driven techniques (see Fig. 1). Firstly, we establish a deep learning model for physical feature extraction (Origin model) based on neural network to capture the inherent physical mechanisms and structural features embedded in numerical simulation results, thus achieving rapid reconstruction on simulated datasets. Secondly, a data-driven empirical network model for error correction is proposed. This model utilizes monitoring data from limited points to construct a pseudo-real dataset, mitigating discrepancies between the original model calculations and actual measurements, thereby facilitating the accurate reconstruction of actual flow fields.

3.1. Neural network model for physical feature extraction

The deep learning model proposed in this section utilizes pre-generated wind field CFD datasets with wind speeds at N measurement points as inputs, and trains the network by supervised learning to preserve the salient features of CFD modelling. Subsequently, the inverse process calculates real-time wind speeds at each point throughout the field, facilitating prompt reconstruction of the wind field. Considering the implicit correlations between velocity components in different directions, we designed a single-column network structure, while inputting the value of the velocity component at each measurement point is used to reconstruct the flow field in a certain direction. In order to satisfy the objective of reconstructing the entire flow field with fewer measurement points, we opted for a BP neural network model. The model consists of three fully connected layers, and Tanh is employed as the activation function to further extract advanced features [38]. The network structure is depicted in Fig. 2.

Furthermore, we use the L_2 loss function to measure the error between the calculated and actual values. It is worth noting that each grid point undergoes a significant wind speed variation under different inflow conditions (u_x, u_y). Without proper control, utilizing the L_2 loss function directly for measuring deviations may result in equally weighting points with both high and low actual wind speeds throughout

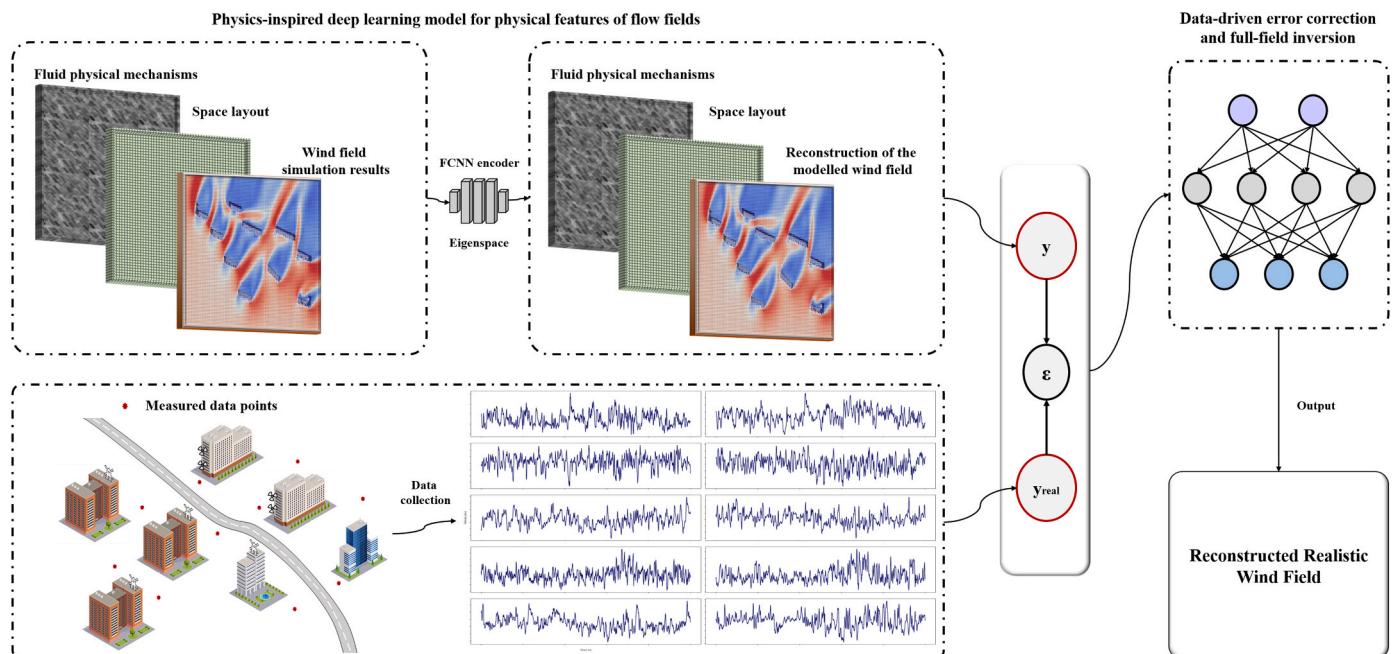


Fig. 1. Physics-inspired and data-driven two-stage flow field reconstruction model.

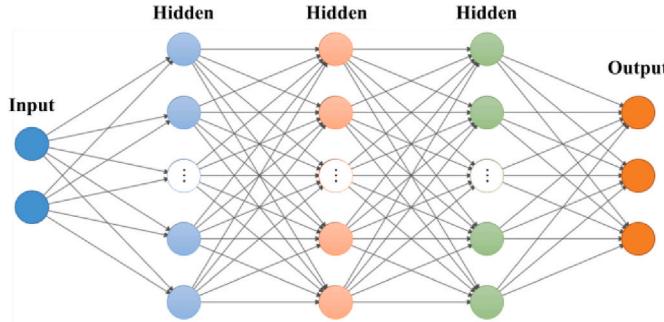


Fig. 2. Structure of fully connected network.

the learning process. Typically, points with lower wind speeds are distributed around buildings, which has a high practical value. Therefore, we designed the weighting coefficients to improve the sensitivity of the network to low wind speed points.

$$\text{Loss} = \frac{1}{N} \left(w_{ui} \cdot \sum_{i=1}^N (u_i - \hat{u}_i)^2 + w_{vi} \cdot \sum_{i=1}^N (v_i - \hat{v}_i)^2 \right) \quad (6)$$

3.2. Empirical model for error correction

The neural network models employed for flow field reconstruction adopt a point-to-point, fully supervised training approach. However, limited by the lack of actual measurement capabilities, it is difficult to obtain real measurement data for a wide range of wind fields. Namely, directly training and refining the network model across the entire flow field using data from a restricted number of points becomes a formidable task. Therefore, we propose an empirical neural network model for error correction. Each output of the deep learning model in Section 3.1 corresponds to the prediction of wind speed at each point. By combining Equation (4), the final predictions can be approximately considered as the sum of the product of the wind speed at the input point and the trainable point weight. It is evident that the wind speed around each point closely resembles the wind speed at the point itself, and this resemblance is inversely proportional to the distance. In other words, in the process of wind field reconstruction, the wind speed at the input point should contribute more to the prediction of wind speed at points closer to itself. Therefore, the multi-layer black-box neural network model in Section 3.1 can be simplified into a highly interpretable reconstruction model. The predicted wind speed at each point is the sum of the product of the input wind speed and the trainable point weight, along with the untrainable distance weight, as shown in Equation (7).

$$V_j = \sum_{i=1}^N V_i \times W_i \times W_{S_{ij}} \quad (7)$$

Where, V_j represents the predicted wind speed at point j , N is the number of input points, V_i is the wind speed at input point, W_i is the trainable point weight, $W_{S_{ij}}$ is the non-trainable relative distance weight. This empirical model greatly simplifies the original model in Section 3.1, allowing the reconstruction of flow field through training a few parameters.

3.3. Data-driven error correction method

Here, we combine the error correction empirical model designed in Section 3.2 and propose a data-driven error correction scheme (see Fig. 3) to extract errors resulting from the simplification of CFD modeling and other uncertainties under real conditions. The simplified empirical model can reconstruct the flow field using monitoring data from a limited number of points, thereby constructing a pseudo-monitoring dataset. This dataset can assist the neural network model in Section 3.1 in point-to-point supervised error learning. Because of the fewer number of trainable parameters in the empirical model, the empirical model needs to learn and train an independent set of parameters for each set of wind field data to ensure model accuracy. Although the empirical model can be trained directly on monitoring data, the limitations of data points may lead to overfitting. Therefore, in the training process of the empirical model, pre-training is first performed on the calculated values of the original model. Further, the monitoring data are used for secondary training, and the trained pseudo-monitoring dataset is fed back to the original model. This iterative process allows the proposed two-stage model to continuously approach the real values.

4. Numerical simulation and monitoring experiment

As discussed in Section 3, the reconstruction of wind speed distribution necessitates the creation of a Computational Fluid Dynamics (CFD) database and a monitoring dataset. In this paper, we take a community in Beijing as an example, and describe its simulation database generation scheme, sparse measured dataset measurement method, and measured data preprocessing method to validate the applicability of the present method under block complex terrain conditions.

4.1. Simulation dataset generation scheme

This study focuses on a specific community in Beijing, and a three-dimensional model of the community was created based on the three-dimensional reality map and CAD plan. The physical dimensions of

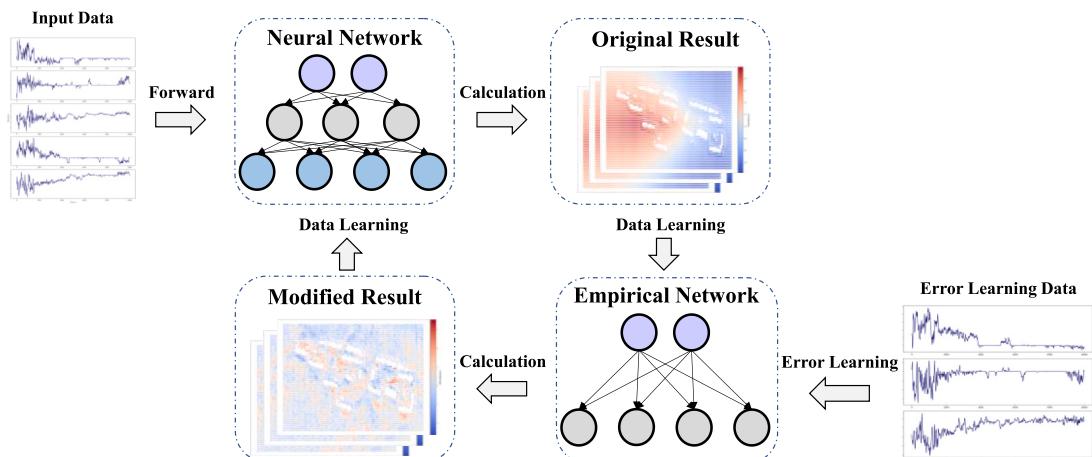


Fig. 3. Data-driven error correction method.

the community model are 600 m long, 300 m wide and 120 m high, with 12 buildings ranging in height from 13 m to 84 m and an average building height of 66 m. The grid layout and subdivision play a crucial role in the numerical calculations. To ensure accuracy, a maximum grid size of 5 m × 5 m × 5 m was used, with additional refinement applied to grids intersecting the geometric surfaces of buildings. The refinement level was set to 3 × 3 × 3. The total number of grids reached 1,040,349. The three-dimensional modelling of the community and the gridded model are shown in Fig. 4.

In reality, wind can blow from any direction at different speeds. By decomposing the inflow wind into (u_x, u_y) velocity components and setting different inflow surfaces and velocities, the simulation can mimic the randomness of wind in the real world. The inflow velocity ranges from 0 to 20 m/s with 0.2 m/s increments, allowing the simulation of wind fields within common wind speed ranges. To ensure reliable computational results, OpenFOAM was used as the numerical simulation software. OpenFOAM is an open-source CFD software that includes different solvers for specific fluid flow problems. For the numerical simulation of the community-scale wind field, we assumed an incompressible turbulent flow with Newtonian fluid properties. The k-Epsilon turbulence model was chosen for its reliability, good convergence, small memory footprint and fast computational speed. Simulation calculations for wind fields with different inflow wind speeds can be performed in batch mode using control scripts. The boundary conditions and solver settings for numerical simulations are shown in Tables 1 and 2.

4.2. Real-time data monitoring scheme

In order to obtain accurate monitoring data, we used an ultrasonic anemometer to monitor the wind speed in the regional field. Considering the practical factors and the actual installation conditions, we strategically placed measuring equipment at ten different locations, as shown in Fig. 5. The anemometer is positioned at a height of 2.5 m above the ground, with a data acquisition frequency of once per second, a velocity accuracy of 0.1 m/s, and an angular accuracy of 1°. In addition, the anemometer is battery-powered, and data collection resumes only after recharging. We collected data for approximately five days over a period of four weeks. After adjusting the data based on collection time, we had over 450,000 records of wind speed data at ten points.

4.3. Data preprocessing method

The measured data require further processing before subsequent applications. The numerical results represent the steady state wind field distribution under fixed inflow conditions. In contrast, monitoring data is collected at a frequency of once per second in a transient field with varying inflow conditions. To facilitate the learning of relatively stable patterns by the network and to mitigate the increased learning difficulty caused by transient fluctuations, we used a window-averaging method to transform transient field data into pseudo-steady-state field data. First, the wind speed and wind direction data were converted into (u_x, u_y) velocity component data. Averaging was then performed over the length of the window. In this paper, a window length of 30 records is chosen to preserve the trend of the wind speed and to reduce the fluctuation. Monitoring points were matched to corresponding grid points

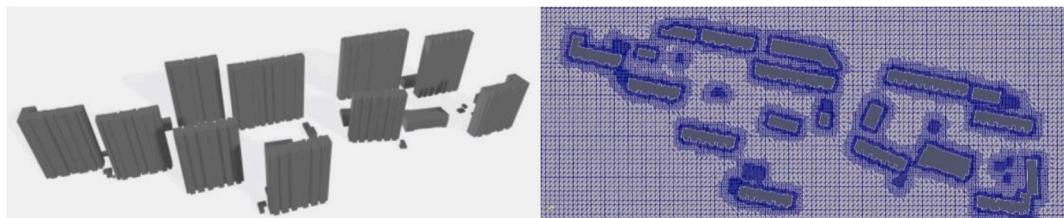


Fig. 4. Three-dimensional modelling and gridding of the community.

Table 1
Boundary condition settings.

Boundary surface	Boundary condition
Inlet surface	Velocity inlet
Outlet surface	Outlet
Top and surrounding boundary surfaces	Symmetric
Ground and building surfaces	No-slip

Table 2
Solution condition settings.

Turbulence model	Reynolds-averaged k-epsilon turbulence modeling
Computational model	simpleFoam
Coupling method	SIMPLE
Gradient	Gaussian linear
Discrete method	First order
Convergent residuals	10-3

using relative coordinates to establish the correspondence between monitoring and simulation data.

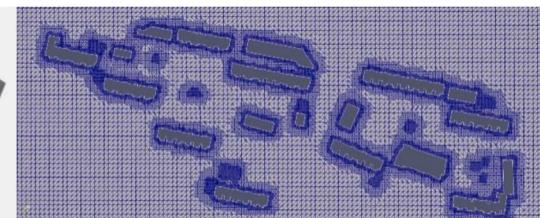
There are 10 monitoring sites included in the study. During the iterative training process, data from 5 points (5–9) were selected as inputs for both the physical feature extraction neural network model and the empirical model, while data from 3 points (1, 3, 10) were utilized for secondary training of the empirical model. The remaining data from 2 points (2, 4) were reserved for model accuracy validation. Measurement data from the validation points were not involved in any learning or training processes for any networks. The collected data, totaling over 450,000 records, was grouped into sets of tens of thousands. The average wind speed for input points (5) and validation points (2), as well as the wind speed distribution density graphs for the two validation points, are depicted in Fig. 6. The graphs highlight the significant variations in the measured wind field over different time periods, with the fourth set showing the maximum average wind speed and turbulence. This observation serves as robust evidence of the effectiveness and applicability of the proposed model.

5. Experimental validation

In this section, we validated the accuracy of the proposed algorithm based on experimental data from two validation points. Furthermore, we conducted a comparative analysis, evaluating the performance of the two-stage deep learning model proposed in this paper (Modified Model) against the neural network model for physical feature extraction (Original Model) and traditional spatial interpolation methods in reconstructing actual flow fields.

5.1. Model validation

Considering the actual wind speed range and to minimize disparities between simulated and actual data, we calculated a total of 4947 simulated flow field data within the inflow wind speed range ($u_x + u_y \leq 10$). The complete simulation results encompass all velocity components in the (x, y, z) directions over the entire grid. Due to no signif-



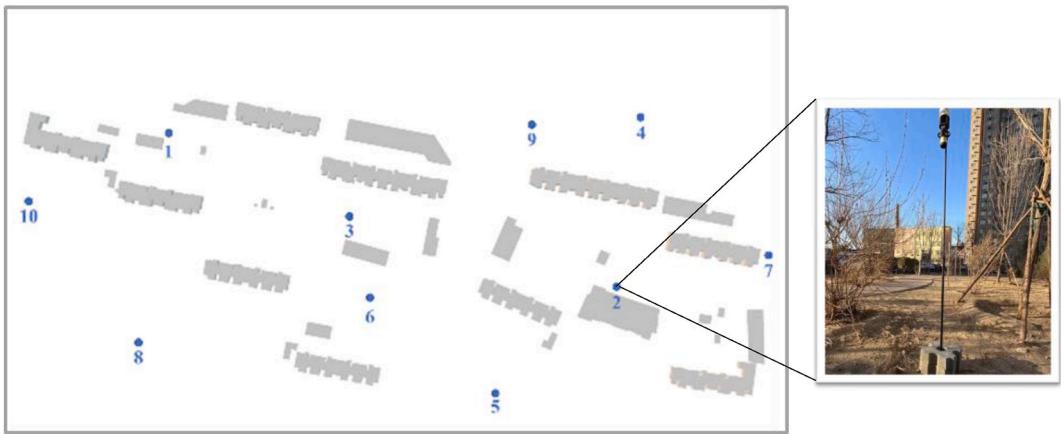


Fig. 5. Distribution of wind speed measurement points.

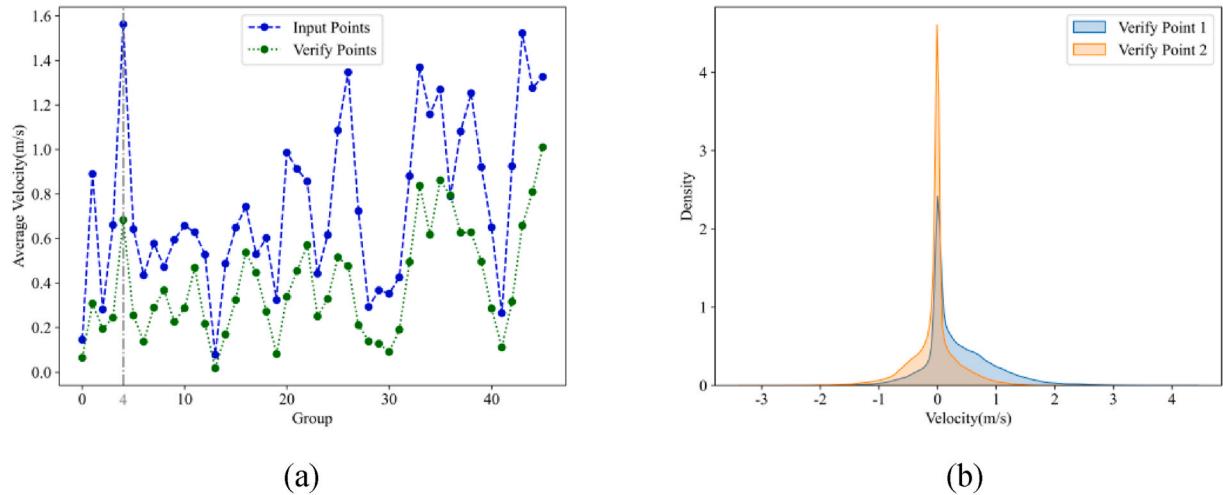


Fig. 6. Average wind speed and wind speed density for monitoring points: (a) average wind speed for input and validation points; (b) wind speed distribution density graphs for the two validation points.

icant difference in the reconstruction accuracy for different directions [39], the measured results in the x-direction were selected for validation. Traditionally, the mean relative error and mean absolute error are usually used as evaluation metrics for model validation in previous studies. Nevertheless, the wind speeds measured in this study frequently encompass velocities of 0 m/s, which presents a challenge in computing the relative error. Consequently, the mean absolute error is employed as the primary evaluation metric.

Initially, we validated the performance of the original model on the simulation dataset. From the complete simulated dataset, 952 flow fields with integer inflow wind speeds were randomly selected as the test dataset. On the test set, the model exhibited a velocity mean relative error of 15.7% and a mean absolute error of 0.208 m/s. These results indicate that the proposed neural network model for physical feature extraction performs well in simulating wind field reconstruction. An example of the reconstructed wind field for a northwest wind is shown in Fig. 7.

Furthermore, we selected the time period with the maximum fluctuations in the measured data (the fourth set, as shown in Fig. 6) and compared the performance of the modified model with the original model in reconstructing the actual flow field (see Fig. 8). The results indicate that during the highest fluctuation period, calculations of the modified model are still very close to the measured data at the two validation points (velocity error distribution map shown in Fig. 9). This is significantly better than the performance of the original model. In

other words, the proposed method has successfully captured errors arising from modeling simplifications and other uncertainties under real physical conditions to some extent.

5.2. Model performance analysis

In this section, we compared the performance of the modified model and the original model, along with traditional spatial interpolation methods (e.g., nearest interpolation, linear interpolation, cubic spline interpolation) in reconstructing the real wind field at the two validation points. There are more than 450,000 pieces of data. In order to present the results more directly, every 10,000 pieces of data were classified into one group for the accuracy test. As shown in Fig. 8, over the entire time period, the average absolute errors between the computed results of the original model and the measured values at the two validation points were 0.905 m/s and 0.798 m/s, respectively. For the modified model, the average absolute errors at the two validation points were 0.53 m/s and 0.368 m/s, representing an accuracy improvement of 41.4% and 53.8%, and an average improvement of 47.6%. Clearly, this further demonstrates that our proposed two-stage algorithm can more effectively capture advanced features of the actual flow field compared to the neural network model for physical feature extraction solely on simulated data.

Additionally, the research revealed that the accuracy of the original model decreases with the increase in the average wind speed of the

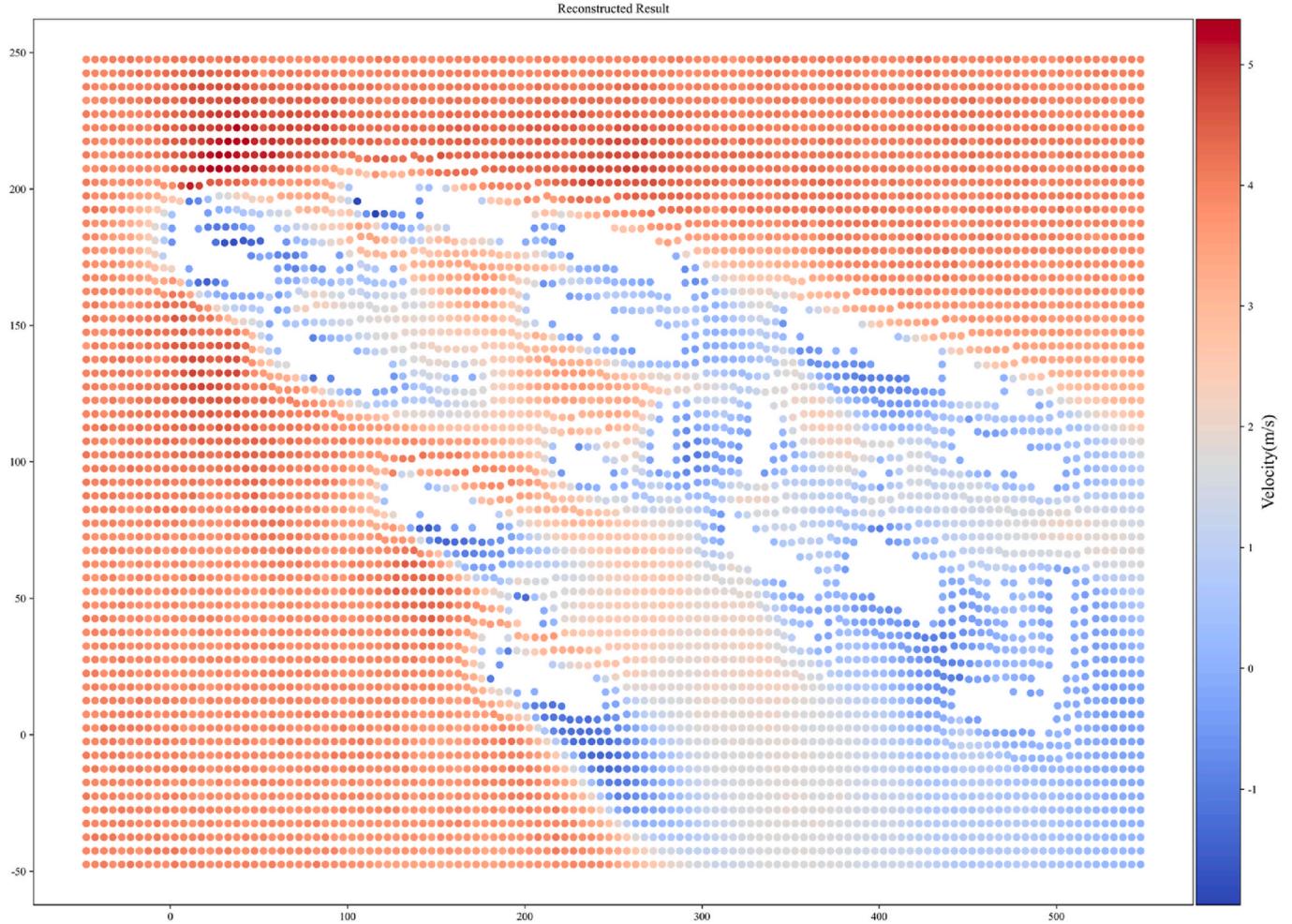


Fig. 7. Example of wind field reconstruction using the neural network model for physical feature extraction.

measured wind field. During the most fluctuating with the maximum average wind speed period (as shown in Fig. 6), the model exhibited an average absolute error of 2.451 m/s and 1.661 m/s at the two validation points, respectively (see Fig. 10). In other words, as the actual average wind speed increases, and the fluctuation of point-wise wind speeds becomes higher, the accuracy of the model's calculations decreases.

The potential reason lies in the fact that the original model is trained on data from steady-state wind fields, where the changes are uniform and stable. In contrast, actual scenarios are subject to intermittent and fluctuating winds, leading to greater variability in measured speed. Although this study has employed window averaging to reduce the fluctuations in measured data, it is still challenging to fully compensate this difference. In situations with significant wind speed variations, the original model's calculations are more prone to significant bias from extreme inputs. As a result, there is a high correlation between model accuracy and average wind speed. The modified model effectively avoids this phenomenon, as shown in Fig. 10. During the period of most dramatic wind speed changes, the model had average absolute errors of 0.671 m/s and 0.608 m/s at the two validation points, representing an improvement in accuracy of 72.5% and 63.3% respectively. This suggests that the proposed method has successfully captured the intermittent and fluctuating nature of actual incoming winds to some extent.

Furthermore, this study compared the performance of the modified model with traditional spatial interpolation methods in reconstructing the actual flow field. Since interpolation methods can only effectively calculate wind field values for points within the input region (see Fig. 6), comparison could only be carried out on validation point 1. As shown in

Fig. 11, the results indicate that the modified model significantly outperforms all interpolation methods. Additionally, we found that the nearest interpolation method has a better accuracy than linear and cubic spline interpolation methods. This indirectly demonstrates the effectiveness of the distance-weighting method employed in constructing the empirical neural network model. We also compared the computational efficiency of the modified model with interpolation methods. Since the modified model only involves fast forward calculations during application, the reconstruction of 10,000 wind fields takes only 2 s. However, using the fastest calculation speed of the nearest interpolation method, the reconstruction of 10,000 wind fields would take 11.8 s. Clearly, the computational efficiency of the modified model far exceeds that of interpolation methods, and the computational efficiency of numerical calculation methods cannot be compared either.

6. Conclusion

In this study, we propose a physics-inspired and data-driven two-stage deep learning method. The application of this method to wind field reconstruction in a typical city block in Beijing has been successful, affirming its applicability and effectiveness. The main conclusions of this research are as follows.

- (1) A two-stage deep learning framework for wind field reconstruction is presented. This method combines physical feature extraction model of the flow field (origin model) with a sparse measurement data-driven error correction approach, aiming to

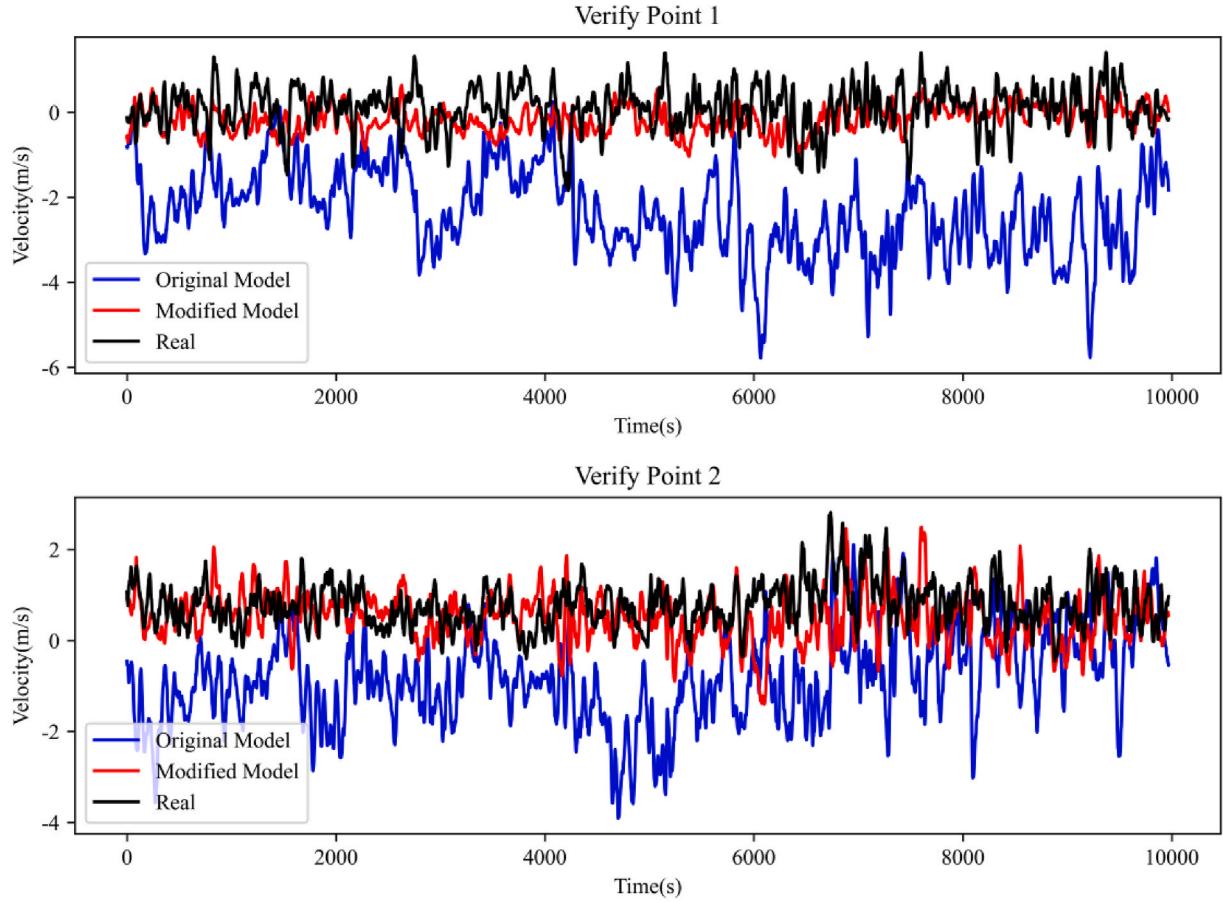


Fig. 8. Comparison of the calculated results of the proposed method in this paper with the experimental measurements.

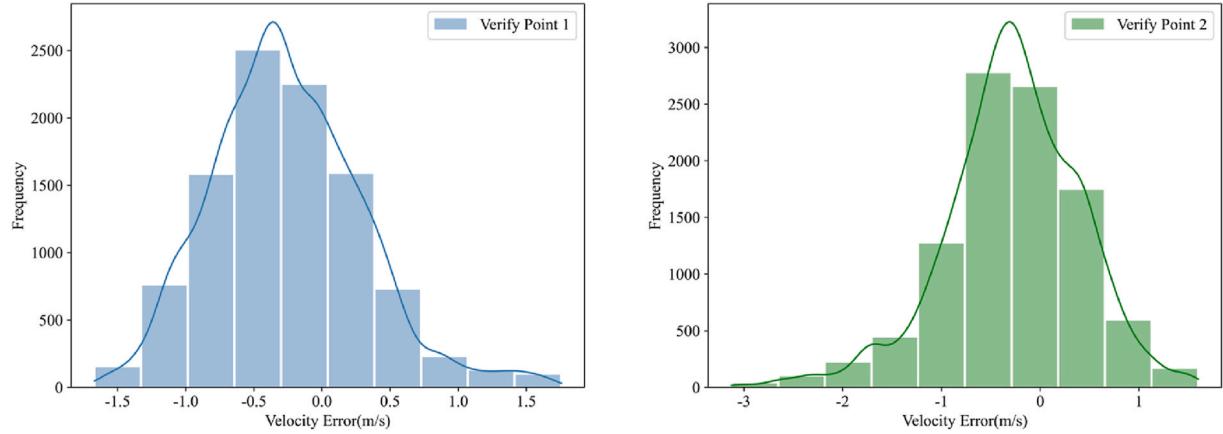


Fig. 9. Distribution of velocity errors of the modified model at the two validation points.

preserve significant features of Computational Fluid Dynamics (CFD) modeling and uncertainties under real conditions, thus achieving fast reconstruction of complex wind fields.

- (2) Using a community in Beijing as a case study, we have established a high-fidelity Computational Fluid Dynamics (CFD) database and experimentally validated the efficacy of the proposed two-stage method. The findings indicate that the method effectively accomplishes wind field reconstruction in complex terrains using sparse datasets. The average absolute errors between the model results and measured values at two validation points are 0.53 m/s and 0.368 m/s, respectively. Furthermore, we also show that the

proposed two-stage strategy exhibits significantly improved prediction results over the purely original method, with an average accuracy improvement of 47.17% and a maximum accuracy improvement of 72.59%, indicating its potential applications value in the selection and placement of wind turbines and the aerodynamic design optimization of ducted openings.

- (3) The proposed algorithm exhibits good performance in simulating high wind speed fields with significant fluctuations.
- (4) The proposed algorithm shows distinct advantages in both computational efficiency and accuracy compared to traditional

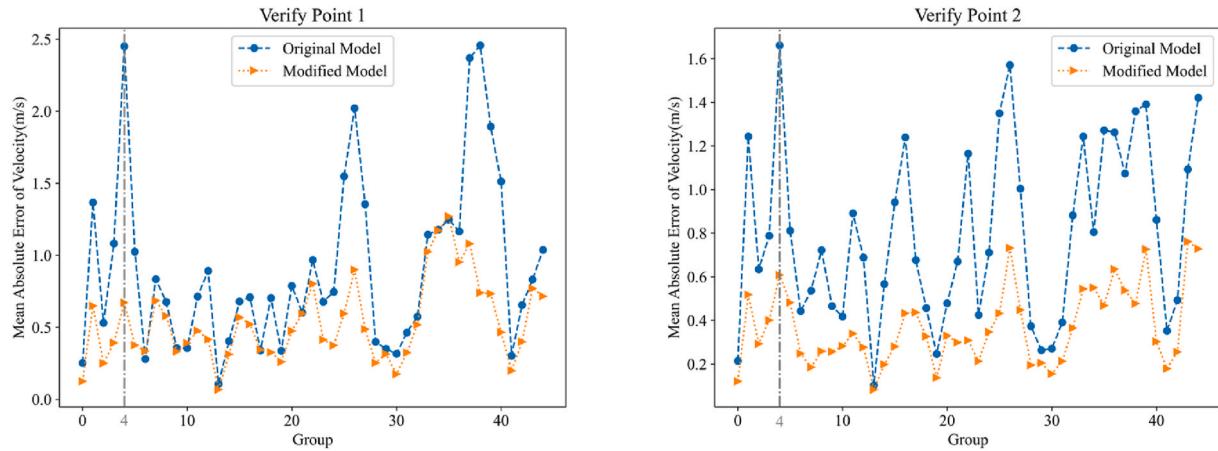


Fig. 10. Comparison of velocity errors between original and modified models over all groups.

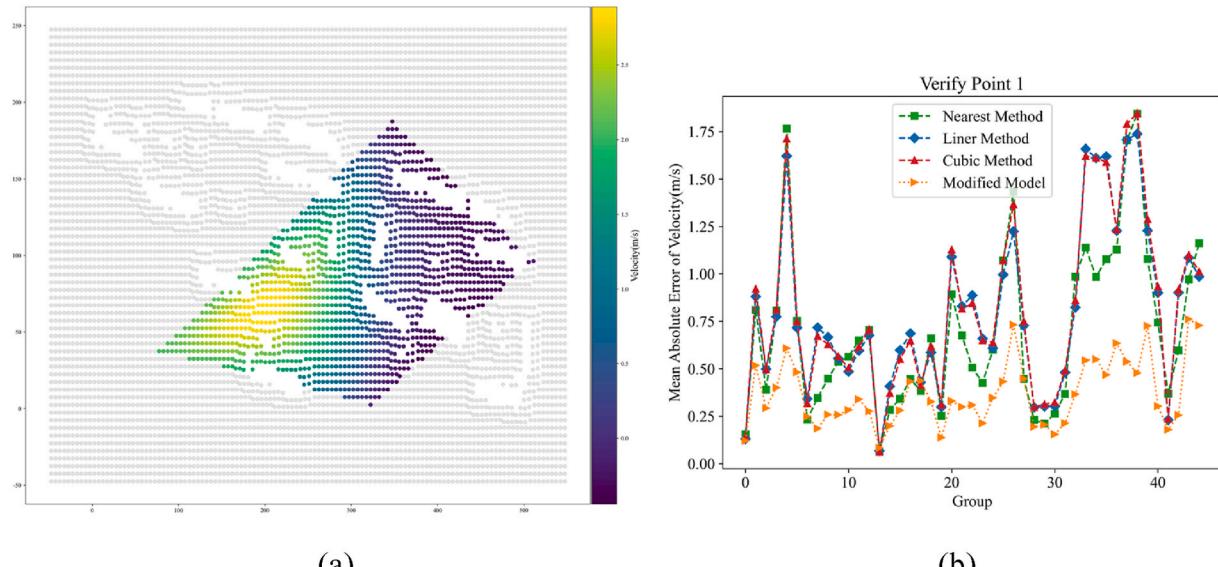


Fig. 11. Comparison of the performance of the modified model with the traditional spatial interpolation method: (a) schematic representation of the results from the spatial interpolation method; (b) velocity errors of different spatial interpolation methods.

spatial interpolation methods and numerical calculation methods.

In summary, this study introduces an innovative method for wind field reconstruction. Nevertheless, our validation efforts have thus far been confined to the urban block scenario. Notably, the proposed methodology holds promise for broader application in wind farm prediction, offering insights into optimal siting strategies and turbine placement. Future investigations are warranted to further explore and validate the method's potential across diverse contexts.

Funding

Funded by National Key R&D Program of China (No.2021YFC3001500) and National Natural Science Foundation of China (No.72174102, No. 72334003, No.72004141).

CRediT authorship contribution statement

Yi Liu: Writing – review & editing, Writing – original draft, Conceptualization. **Ranpeng Wang:** Writing – review & editing, Writing

– original draft. **Yin Gu:** Writing – review & editing, Writing – original draft. **Congjian Li:** Data curation. **Gangqiao Wang:** Writing – review & editing, Writing – original draft.

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.

Data availability

Data will be made available on request.

References

- [1] Ohene E, Chan APC, Darko A, Nani G. Navigating toward net zero by 2050: drivers, barriers, and strategies for net zero carbon buildings in an emerging market. *Build Environ* 2023;242:110472. <https://doi.org/10.1016/j.buildenv.2023.110472>.
- [2] International Energy Agency. *Transition to sustainable buildings: strategies and opportunities to 2050*. 2013. Paris.
- [3] Wang N, Phelan PE, Gonzalez J, Harris C, Henze GP, Hutchinson R, et al. Ten questions concerning future buildings beyond zero energy and carbon neutrality.

- Build Environ 2017;119:169–82. <https://doi.org/10.1016/j.buildenv.2017.04.006>.
- [4] Economidou M, Todeschi V, Bertoldi P, D'Agostino D, Zangheri P, Castellazzi L. Review of 50 years of EU energy efficiency policies for buildings. Energy Build 2020;225:110322. <https://doi.org/10.1016/j.enbuild.2020.110322>.
- [5] Zeyen E, Hagenmeyer V, Brown T. Mitigating heat demand peaks in buildings in a highly renewable European energy system. Energy 2021;231:120784. <https://doi.org/10.1016/j.energy.2021.120784>.
- [6] Yu FW, Ho WT. Tactics for carbon neutral office buildings in Hong Kong. J Clean Prod 2021;326:129369. <https://doi.org/10.1016/j.jclepro.2021.129369>.
- [7] Rose J, Thomsen KE, Mørck OC, Gutierrez MSM, Jensen SØ. Refurbishing blocks of flats to very low or nearly zero energy level—technical and financial results plus co-benefits. Energy Build 2019;184:1–7. <https://doi.org/10.1016/j.enbuild.2018.11.051>.
- [8] Liu Y, Xue S, Guo X, Zhang B, Sun X, Zhang Q, et al. Towards the goal of zero-carbon building retrofitting with variant application degrees of low-carbon technologies: mitigation potential and cost-benefit analysis for a kindergarten in Beijing. J Clean Prod 2023;393:136316. <https://doi.org/10.1016/j.jclepro.2023.136316>.
- [9] Liu Z, Liu Y, He B-J, Xu W, Jin G, Zhang X. Application and suitability analysis of the key technologies in nearly zero energy buildings in China. Renew Sustain Energy Rev 2019;101:329–45. <https://doi.org/10.1016/j.rser.2018.11.023>.
- [10] Ishugah TF, Li Y, Wang RZ, Kiplagat JK. Advances in wind energy resource exploitation in urban environment: a review. Renew Sustain Energy Rev 2014;37:613–26. <https://doi.org/10.1016/j.rser.2014.05.053>.
- [11] Stathopoulos T, Alrawashdeh H, Al-Quraan A, Blocken B, Dilimulati A, Paraschivou M, et al. Urban wind energy: some views on potential and challenges. J Wind Eng Ind Aerod 2018;179:146–57. <https://doi.org/10.1016/j.jweia.2018.05.018>.
- [12] Baldazzi F, Bianchini A, Carnevale EA, Ferrari L, Magnani S. Feasibility analysis of a Darrieus vertical-axis wind turbine installation in the rooftop of a building. Appl Energy 2012;97:921–9. <https://doi.org/10.1016/j.apenergy.2011.12.008>.
- [13] Kc A, Whale J, Urmee T. Urban wind conditions and small wind turbines in the built environment: a review. Renew Energy 2019;131:268–83. <https://doi.org/10.1016/j.renene.2018.07.050>.
- [14] Tasneem Z, Al Noman A, Das SK, Saha DK, Islam MdR, Ali MdF, et al. An analytical review on the evaluation of wind resource and wind turbine for urban application: prospect and challenges. Developments in the Built Environment 2020;4:100033. <https://doi.org/10.1016/j.dibe.2020.100033>.
- [15] Calautti K, Johnstone C. State-of-the-art review of micro to small-scale wind energy harvesting technologies for building integration. Energy Convers Manag X 2023; 20:100457. <https://doi.org/10.1016/j.ecmx.2023.100457>.
- [16] Alanis Ruiz C, Kalkman I, Blocken B. Aerodynamic design optimization of ducted openings through high-rise buildings for wind energy harvesting. Build Environ 2021;202:108028. <https://doi.org/10.1016/j.buildenv.2021.108028>.
- [17] Yuan X, Tan Q, Lei X, Yuan Y, Wu X. Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine. Energy 2017;129:122–37. <https://doi.org/10.1016/j.energy.2017.04.094>.
- [18] Zhang Z, Ye L, Qin H, Liu Y, Wang C, Yu X, et al. Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression. Appl Energy 2019;247:270–84. <https://doi.org/10.1016/j.apenergy.2019.04.047>.
- [19] Stathopoulos C, Kaperoni A, Galanis G, Kallos G. Wind power prediction based on numerical and statistical models. J Wind Eng Ind Aerod 2013;112:25–38. <https://doi.org/10.1016/j.jweia.2012.09.004>.
- [20] Milanese M, Tornese L, Colangelo G, Laforgia D, de Risi A. Numerical method for wind energy analysis applied to Apulia Region, Italy. Energy 2017;128:1–10. <https://doi.org/10.1016/j.energy.2017.03.170>.
- [21] Li P, Guan X, Wu J, Zhou X. Modeling dynamic spatial correlations of geographically distributed wind farms and constructing ellipsoidal uncertainty sets for optimization-based generation scheduling. IEEE Trans Sustain Energy 2015;6:1594–605. <https://doi.org/10.1109/TSTE.2015.2457917>.
- [22] Song MX, Chen K, He ZY, Zhang X. Wake flow model of wind turbine using particle simulation. Renew Energy 2012;41:185–90. <https://doi.org/10.1016/j.renene.2011.10.016>.
- [23] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88:1405–14. <https://doi.org/10.1016/j.apenergy.2010.10.031>.
- [24] Zhang C, Wei H, Zhao X, Liu T, Zhang K. A Gaussian process regression based hybrid approach for short-term wind speed prediction. Energy Convers Manag 2016;126:1084–92. <https://doi.org/10.1016/j.enconman.2016.08.086>.
- [25] Chen K, Yu J. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. Appl Energy 2014;113:690–705. <https://doi.org/10.1016/j.apenergy.2013.08.025>.
- [26] Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Appl Energy 2010;87:2313–20. <https://doi.org/10.1016/j.apenergy.2009.12.013>.
- [27] Liu H, Tian H, Li Y. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Appl Energy 2012;98:415–24. <https://doi.org/10.1016/j.apenergy.2012.04.001>.
- [28] Khosravi A, Koury RNN, Machado L, Pabon JJG. Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. Sustain Energy Technol Assessments 2018;25:146–60. <https://doi.org/10.1016/j.seta.2018.01.001>.
- [29] Sun S, Liu Y, Li Q, Wang T, Chu F. Short-term multi-step wind power forecasting based on spatio-temporal correlations and transformer neural networks. Energy Convers Manag 2023;283:116916. <https://doi.org/10.1016/j.enconman.2023.116916>.
- [30] Nascimento EGS, de Melo TAC, Moreira DM. A transformer-based deep neural network with wavelet transform for forecasting wind speed and wind energy. Energy 2023;278:127678. <https://doi.org/10.1016/j.energy.2023.127678>.
- [31] Liu L, Liu J, Ye Y, Liu H, Chen K, Li D, et al. Ultra-short-term wind power forecasting based on deep Bayesian model with uncertainty. Renew Energy 2023; 205:598–607. <https://doi.org/10.1016/j.renene.2023.01.038>.
- [32] Zhang J, Zhao X. Digital twin of wind farms via physics-informed deep learning. Energy Convers Manag 2023;293:117507. <https://doi.org/10.1016/j.enconman.2023.117507>.
- [33] Sekar V, Jiang Q, Shu C, Khoo BC. Fast flow field prediction over airfoils using deep learning approach. Phys Fluids 2019;31:057103. <https://doi.org/10.1063/1.5094943>.
- [34] Deng Z, He C, Liu Y, Kim KC. Super-resolution reconstruction of turbulent velocity fields using a generative adversarial network-based artificial intelligence framework. Phys Fluids 2019;31:125111. <https://doi.org/10.1063/1.5127031>.
- [35] Jin X, Cheng P, Chen W-L, Li H. Prediction model of velocity field around circular cylinder over various Reynolds numbers by fusion convolutional neural networks based on pressure on the cylinder. Phys Fluids 2018;30:047105. <https://doi.org/10.1063/1.5024595>.
- [36] Roshko A. Perspectives on bluff body aerodynamics. J Wind Eng Ind Aerod 1993; 49:79–100. [https://doi.org/10.1016/0167-6105\(93\)90007-B](https://doi.org/10.1016/0167-6105(93)90007-B).
- [37] Jin X, Lai M, Li H. Physics-enhanced deep learning methods for modelling and simulating flow fields. Chin J Theor Appl Mech 2021;53:2616–29.
- [38] Li C, Gao H, Liu Y. Fast reconstruction of a wind field based on numerical simulation and machine learning. Tsinghua Sci Technol 2023;63:882–7. <https://doi.org/10.16511/j.cnki.qhdxxb.2023.22.014>.
- [39] Zhang G, Zheng X, Liu S, Chen M. Three-dimensional wind field reconstruction using tucker decomposition with optimal sensor placement. Energy 2022;260:125098. <https://doi.org/10.1016/j.energy.2022.125098>.