**Electronic Supplementary Materials**
For https://doi.org/10.1631/jzus.A2500344

# Hierarchical learning method for array flow field prediction integrated with a deep neural network

Shanxun SUN[1], Zijiang XU[1], Zhuoheng WANG[1], Shuangshuang CUI[2], Ting HE[1], Yang CAI[1]

[1]*Energy and Electricity Research Center, Jinan University, Zhuhai 519070, China*
[2]*Guangdong-Hong Kong-Macao Greater Bay Area Meteorological Intelligent Equipment Research Center, Guangzhou 510530, China*

## Section S1

$$\begin{cases} e_u = \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} + v\dfrac{\partial u}{\partial y} + w\dfrac{\partial u}{\partial z} - \dfrac{\partial p}{\partial x} + \dfrac{1}{Re}\left(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} + \dfrac{\partial^2 u}{\partial z^2}\right) \\ e_v = \dfrac{\partial v}{\partial t} + u\dfrac{\partial v}{\partial x} + v\dfrac{\partial v}{\partial y} + w\dfrac{\partial v}{\partial z} - \dfrac{\partial p}{\partial y} + \dfrac{1}{Re}\left(\dfrac{\partial^2 v}{\partial x^2} + \dfrac{\partial^2 v}{\partial y^2} + \dfrac{\partial^2 v}{\partial z^2}\right) \\ e_w = \dfrac{\partial w}{\partial t} + u\dfrac{\partial w}{\partial x} + v\dfrac{\partial w}{\partial y} + w\dfrac{\partial w}{\partial z} - \dfrac{\partial p}{\partial z} + \dfrac{1}{Re}\left(\dfrac{\partial^2 w}{\partial x^2} + \dfrac{\partial^2 w}{\partial y^2} + \dfrac{\partial^2 w}{\partial z^2}\right) \\ e_{div} = \dfrac{\partial u}{\partial x} + \dfrac{\partial v}{\partial y} + \dfrac{\partial w}{\partial z} \end{cases} \tag{S1}$$

The Reynolds number $Re$ in wind farms can be calculated using the formula $U_\infty D/\nu_{eff}$, where $D$ represents the turbine rotor diameter, $U_\infty$ denotes the average free-stream wind speed, and $\nu_{eff}$ indicates the total effective viscosity, which is the sum of air's dynamic viscosity and turbulent viscosity. $f_x$, $f_y$ and $f_z$ are set s as external force terms in the inlet boundary conditions.

## Section S2

Table S1 lists the basic parameters of this wind turbine. These established parameters constitute a standardized framework for performance simulation and operational analysis of wind turbine arrays across diverse wind farm configurations.

**Table S1 Basic Parameters of NREL 5MW Wind Turbine**

| Name | Parameter |
|---|---|
| Rated power | 5MW |
| Rotor orientation | Upwind type |
| Number of blades | 3 |
| Blade diameter | 126m |
| Rated wind speed | 11.4m/s |
| Rated rotational speed | 12.1rpm |

**Section S3**

The turbine layout follows the staggered configuration established in the Vested experimental study. In offshore wind farms, vertical spacing between turbines is typically around the 7D range, where D represents the turbine rotor diameter, and horizontal spacing ranges between 3D to 5D (Wei et al., 2021). Therefore, as shown in Fig. S1, the longitudinal spacing of wind turbines in the wind farm simulation is set to 7D, and the lateral spacing is set to 4D, where D denotes the rotor diameter of the wind turbine.

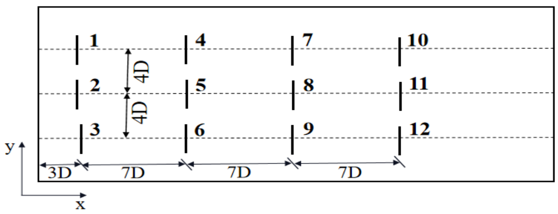


**Fig. S1 The layout of twelve wind turbines in a wind farm**

**Section S4**

The main simulation uses the same computational domain size as the precursor atmospheric boundary layer simulation. Wind turbines are integrated into this domain using an actuator line model with a three-tiered mesh refinement strategy, as illustrated in Fig. S2. Region I employs a background grid with a mesh size of 20m. Region II spans from upstream 1D to downstream 24.5D relative to the first row of wind turbines, featuring a mesh size of 10m. Region III further refines the grid from upstream 0.5D to downstream 24D relative to the first row of wind turbines, using a 5m mesh size to precisely capture detailed flow dynamics. This grid refinement approach aims to accurately simulate the flow dynamics within the wind farm, comprising approximately 40.61 million grids in total.

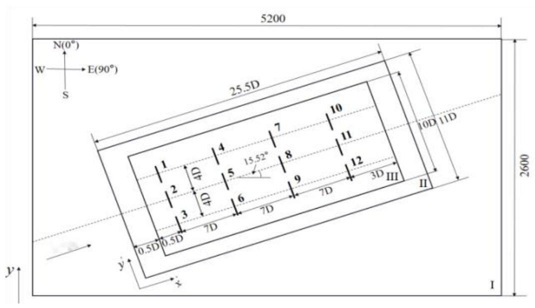


**Fig. S2 The configuration of the computational domain and grid sizes**

In the precursor atmospheric boundary layer (ABL) simulation, the east and north boundaries employ zero-pressure gradient conditions instead of periodic boundaries for the main simulation. The south and west boundaries of the main simulation use the precursor ABL flow field data as inlet boundary conditions. To ensure complete ABL inflow development, the total simulation time is set to 19,500 seconds, with the initial 18,000 seconds allocated for ABL inflow generation, during which the flow field is considered to reach quasi-equilibrium when the velocity at half the boundary layer height fluctuates around a stable mean value. The boundary layer height characterizes the interaction between the surface and the free atmosphere in the precursor ABL simulation. The inflow boundary information for the last 1,500 seconds is stored for use in the main simulation, with a time step of 0.05 seconds. Only the results from the final 20 seconds of the simulation are selected for analysis.

## Section S5

**Table S2 parameters of the integrated network framework**

| layer | Neuron numbers. | Activation function. |
|---|---|---|
| 1-3 | 256 | Hyperbolic tangent functions |
| 4-5 | 256 | Relu functions |
| 6-8 | 256 | Hyperbolic tangent functions |
| 9 | 4 | Linear activation function |

## Section S6

Fig. S3 validates the proposed hierarchical learning model. Panel (a) shows the simulated true flow field, while panel (b) presents the flow field predicted by the hierarchical learning method. After processing, the amount of training data is approximately one-third of the data volume input into the combined model described in Section III. This reduction is achieved by randomly selecting one-third of the input data for training, utilizing the constraint of the 3D NS equation residual term within the hierarchical learning model. For an initial wind speed of 11.4 m/s, the mean absolute errors (MAE) of the predicted wind field at the 11th second for the streamwise, spanwise, and vertical velocity components are 0.23, 0.14, and 0.14, respectively, with a mean absolute error of 0.22 for the resultant velocity. As shown in Fig. S3a and S3b, it is evident that the low-velocity regions downstream of the wind turbines, as well as the variations in high and low velocities and wind shear across the wind farm, are accurately predicted. This demonstrates that the hierarchical learning model developed in this study effectively balances local and global predictions, enabling real-time and accurate wake field prediction.

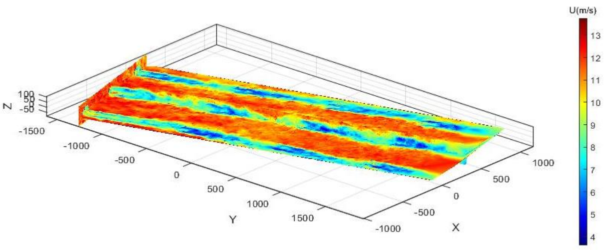


**Fig. S3(a) The real flow field**

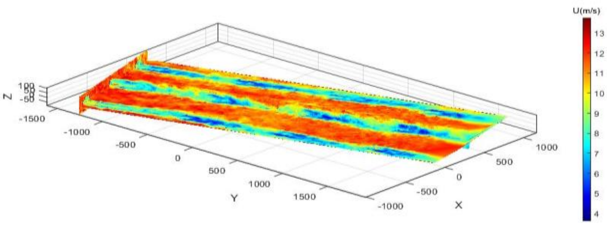


**Fig. S3(b) The predicted flow field**

## Section S7

Fig. S4 visualizes the streamwise velocity *u*, the spanwise velocity *v*, and the vertical velocity *w* at a specific moment in the wind farm. The velocity contour plots of each direction obtained from real simulations are compared. It is observed that whether for the streamwise velocity *u*, which has higher velocity values, or for the vertical velocity *w*, which has lower velocity values, the hierarchical learning model based on physics-informed neural networks (PINNs) can accurately predict the values. Additionally, in regions with larger velocity variations, the model can also accurately predict the transitions. In horizontal comparisons, our proposed hierarchical learning method accurately predicts the range and direction of the wind field velocities, and it also accurately forecasts the temporal changes in wind speed in both vertical and horizontal directions.

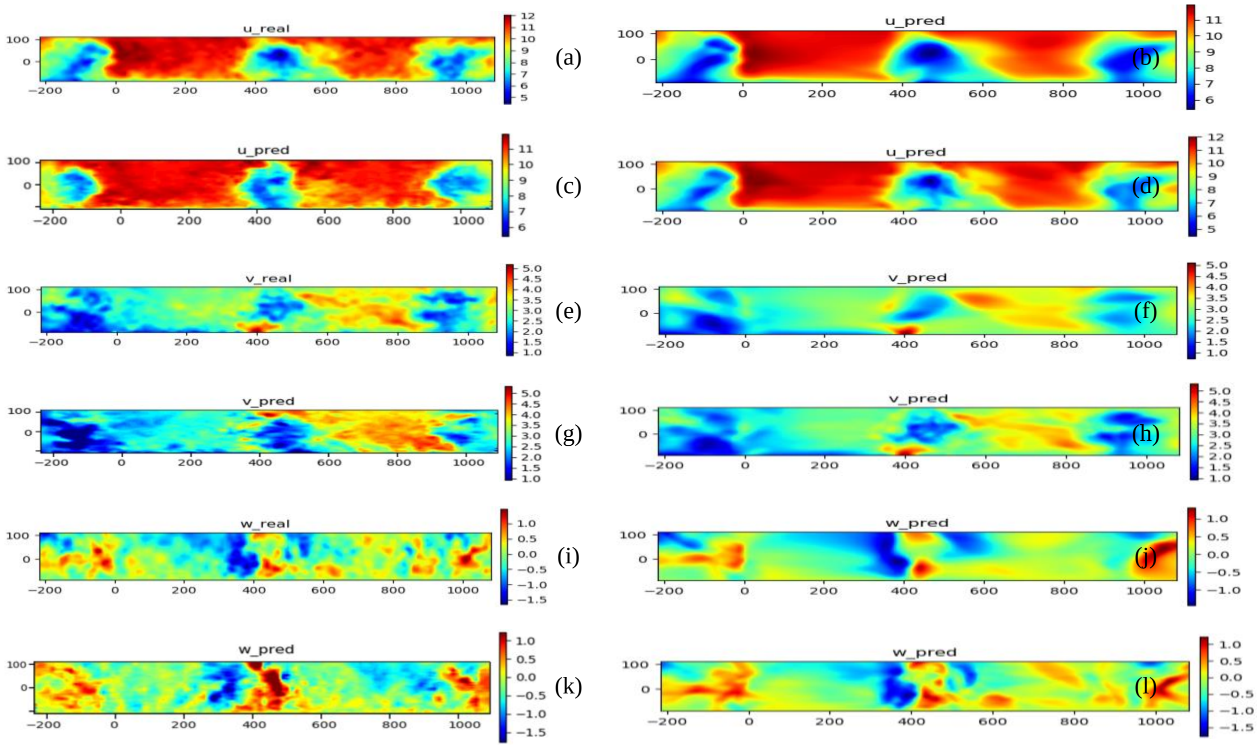


**Fig. S4 Visualization of the Flow Field Velocity at a Specific Moment in the Wind Farm**
**(a) Real Velocity *u*; (b) Velocity *u* Predicted by the MLP; (c) Velocity *u* Predicted by the CNN; (d) Velocity *u* Predicted by the Hierarchical Learning Model; (e) Real Velocity *v*; (f) Velocity *v* Predicted by the MLP; (g) Velocity *v* Predicted by the CNN; (h) Velocity *v* Predicted by the Hierarchical Learning Model; (i) Real Velocity *w*; (j) Velocity *w* Predicted by the MLP; (k) Velocity *w* Predicted by the CNN; (l) Velocity *w* Predicted by the Hierarchical Learning Model.**

## Section S8

In Table S3, $u_1$, $v_1$, and $w_1$ represent the wind speed values predicted by the proposed hierarchical learning method; $u_2$, $v_2$, and $w_2$ are the wind speed values predicted by the MLP; $u_3$, $v_3$, and $w_3$ represent the wind speed values predicted by the CNN. Analysis of the error metrics for each wind speed direction reveals that the proposed hierarchical learning method yields smaller error values for maximum error, minimum error, MAE, MRE, and RMSE compared to the MLP and CNN, indicating better prediction performance.

The proposed method effectively captures the correlations and dependencies between the three-dimensional wind speeds. Both high wind speeds in the streamwise ($u$) and lower wind speeds in the vertical direction ($w$) are accurately predicted with minimal errors. This demonstrates that the proposed method successfully captures both the global and local details of the flow field evolution, enabling real-time reconstruction and accurate prediction of the wake field.

**Table S3 Velocity Component Prediction Errors of three Methods**

| | *ME* | *me* | *MAE* | *MRE* | *RMSE* |
|---|---|---|---|---|---|
| $u_1$ | -0.0721 | 0.0479 | 0.151 | 0.0190 | 0.204 |
| $u_2$ | -0.124 | 0.967 | 0.245 | 0.0259 | 0.323 |
| $u_3$ | 1.511 | -1.479 | 0.305 | 0.0362 | 0.419 |
| $v_1$ | 0.0864 | 0.0749 | 0.128 | 0.0451 | 0.169 |
| $v_2$ | -0.124 | -0.119 | 0.210 | 0.0741 | 0.272 |

| | | | | | |
|---|---|---|---|---|---|
| $v_3$ | 1.137 | -1.157 | 0.217 | 0.1253 | 0.286 |
| $w_1$ | -0.275 | -0.0676 | 0.146 | 0.298 | 0.188 |
| $w_2$ | -0.540 | 0.204 | 0.201 | 0.546 | 0.262 |
| $w_3$ | 1.021 | -0.984 | 0.179 | 1.634 | 0.242 |

The method exhibits outstanding performance, balancing global and local features compared to traditional wake prediction methods. It also significantly reduces algorithm complexity compared to existing integrated neural network frameworks, helping to lower computational costs and speed up processing. Compared to traditional wake prediction methods, the proposed hierarchical learning model, based on an integrated framework, makes more efficient use of wind farm data through layered processing and data handling, achieving higher prediction accuracy and avoiding premature convergence to local optima, thus providing superior wake field prediction results.

**Section S9**

The wind direction prediction errors across the four strategically selected regions are summarized in Tables S4. A consistent and conclusive pattern emerges from these results: the proposed hierarchical learning method ($dt_1$) achieves a decisive superiority over the conventional MLP ($dt_2$) and CNN ($dt_3$) benchmarks across nearly all evaluation metrics. This is unequivocally demonstrated by its significantly smaller MAE, MRE, and RMSE values in every scenario.

**Table S4 Wind Direction Prediction Errors at Different Positions**

| Position | | | *MAE* | *MRE* | *RMSE* |
|---|---|---|---|---|---|
| Row | The second row of wind turbines | $dt_1$ | 6.40 | 0.0794 | 9.17 |
| | | $dt_2$ | 13.26 | 0.1625 | 18.18 |
| | | $dt_3$ | 11.66 | 0.1438 | 15.10 |
| | The third row of wind turbines | $dt_1$ | 8.79 | 0.1096 | 13.69 |
| | | $dt_2$ | 15.59 | 0.2433 | 25.91 |
| | | $dt_3$ | 14.47 | 0.1708 | 20.53 |
| Column | The third column of wind turbines | $dt_1$ | 4.08 | 0.0556 | 8.20 |
| | | $dt_2$ | 10.94 | 0.1845 | 19.20 |
| | | $dt_3$ | 9.87 | 0.1276 | 15.75 |
| | The fourth column of wind turbines | $dt_1$ | 3.15 | 0.0361 | 4.29 |
| | | $dt_2$ | 9.11 | 0.0997 | 11.11 |
| | | $dt_3$ | 7.71 | 0.0879 | 10.07 |

The substantially larger errors of the MLP and CNN, particularly in metrics like MAE and RMSE which reflect overall accuracy, underscore their fundamental limitations. As purely data-driven models, they struggle to capture the intricate spatiotemporal physics that govern wind direction evolution in a wake field. The MLP, lacking inherent spatial inductive bias, fails to model the complex vortex advection and large-scale coherent structures. The CNN, while capable of recognizing spatial patterns, remains a black-box interpolator that cannot generalize beyond its training data distribution or incorporate fundamental physical laws. In contrast, the proposed hierarchical framework explicitly embeds the governing Navier-Stokes equations as a soft constraint. This physics-informed foundation guides the model towards physically consistent solutions, enabling it to extrapolate more reliably from sparse data and accurately capture the dynamic evolution of wind direction, thereby validating its enhanced generalization capability and robustness across diverse wake interaction regimes.