

Multi-fidelity modeling of wind farm wakes based on a novel super-fidelity network

Rui Li, Jincheng Zhang, Xiaowei Zhao*

Intelligent Control & Smart Energy (ICSE) Research Group, School of Engineering, University of Warwick, Coventry, CV4 7AL, UK

ARTICLE INFO

Keywords:

Convolutional Neural Network
CFD simulation
Deep learning
Super-fidelity
Wake model

ABSTRACT

The modeling of wake effects plays an essential role in wind farm optimal design and operation. In this study, a novel deep learning method, called Super-Fidelity Network (SFNet), is proposed for wind farm wake modeling, which would be the first attempt to combine the advantages of both analytical models and numerical models through deep learning methods. Specifically, the low-fidelity flow fields generated by the analytical models serve as the prior information for predicting high-fidelity flow fields. Then the SFNet learns the mapping relationships between low-fidelity data and high-fidelity data, thereby predicting high-fidelity flow fields without resorting to huge computational resources. Numerical experiments demonstrate that the mean absolute error of the developed model is just 1.9% with respect to the freestream wind speed when compared with high-fidelity data, after trained on only 45 samples. In addition, the generalizability of the proposed SFNet in yaw angles, wind speeds and array column extensions is verified by a series of numerical experiments. Furthermore, the experimental results demonstrate that the trained model is able to predict the flow field of a wind farm consisting of 100 turbines within several seconds based on a standard desktop.

1. Introduction

As an important and promising low-carbon alternative to fossil fuels, wind energy has been experiencing rapid and continuous growth in recent years [1]. Generally, wind turbines are installed in large-scale arrays to form wind farms, thereby reducing the overall cost. However, the lower wind speed and higher turbulence intensity caused by upstream turbines, i.e. the wake effects, have a huge impact on the downstream turbines, e.g. considerably reducing their power generations and increasing their structural loads [2]. The incorrect estimation of the wake effects will greatly undermine the prediction accuracy of the wind farm energy yield. Therefore, massive efforts have been poured into wind farm wake modeling, from low-fidelity analytical wake models to high-fidelity numerical wake models.

The analytical models [3–5] are formulated analytically which can generate the flow field in real-time even for large-scale wind farms. Starting from the development of the one-dimensional models [3,4,6], the analytical models have been extended to two-dimensional [7–9] and three-dimensional [2,10,11] models. Meanwhile, more contributing factors are included in subsequent studies, including yaw effects [12–14], background flow fields [15] and terrain conditions [16]. Besides, the Supervisory Control and Data Acquisition (SCADA) data are also incorporated into the analytical models [17]. However, limited by model inadequacy and parameter uncertainty, the modeling

errors of analytical models are still significant. For example, many improved analytical models are still derived based on the Gaussian wake model. Even though the velocity in far-wake areas can be approximated quite well, the near-wake features are usually not accurately captured. By contrast, based on the Computational Fluid Dynamics (CFD), the numerical models solve the Navier–Stokes (NS) equations using numerical approaches. Reynolds-averaged Navier–Stokes (RANS) and Large Eddy Simulation (LES) have been carried out with the turbine rotors modeled by the actuator line method (ALM) [5,18,19] or the actuator disk method (ADM) [20–22]. Even though both far-wake and near-wake areas of the flows can be comprehensively simulated, a CFD simulation for a wind farm with tens of turbines requires tremendous computational resources due to the refined mesh needed to resolve the flow dynamics. For example, on a desktop workstation with an Intel Xeon CPU @3.60 GHz, it requires more than 100 h on 12 processors to simulate a single NREL Phase VI wind turbine [23]. And the required computational resources will rapidly explode with the increase of the scale of the wind farm (i.e. the number of wind turbines). To sum up, the analytical models with low fidelity are efficient but lack flow details, while numerical models with high fidelity can generate high-quality flow fields but the computational requirements are too high for engineering applications.

* Corresponding author.

E-mail addresses: rui.li.4@warwick.ac.uk (R. Li), jincheng.zhang.1@warwick.ac.uk (J. Zhang), xiaowei.zhao@warwick.ac.uk (X. Zhao).

<https://doi.org/10.1016/j.enconman.2022.116185>

Received 14 June 2022; Received in revised form 19 August 2022; Accepted 25 August 2022

Available online 6 September 2022

0196-8904/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Nomenclature**Abbreviations**

ABL	Atmospheric Boundary Layer
ADM	Actuator Disk Method
ALM	Actuator Line Method
ANN	Artificial Neural Network
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
Conv+BN	Convolution + BatchNorm operations
CV	Computer Vision
DL	Deep Learning
FEM	Feature Extraction Module
FFM	Feature Fusion Module
GAN	Generative Adversarial Network
GNN	Graph Neural Network
GPR	Gaussian Process Regression
HPC	High-Performance Computing
LAM	Linear Attention Mechanism
LES	Large Eddy Simulation
LSTM	Long-Short Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MLP	Multilayer Perceptron
NREL	National Renewable Energy Laboratory
NS	Navier–Stokes
RANS	Reynolds-Averaged Navier–Stokes
RMSE	Root Mean Square Error
SCADA	Supervisory Control and Data Acquisition
SCRTP	Scientific Computing Research Technology Platform
SF	Super-Fidelity
SFNet	Super-Fidelity Network
SOWFA	Simulator fOr Wind Farm Applications
SR	Super-Resolution
VAWT	Vertical-Axis Wind Turbine

Symbols

\mathcal{D}	The degradation mapping function
E	The expected loss
\mathcal{F}	The super-fidelity model
F_l	The low-fidelity input flow field
F_h	The high-fidelity input flow field
\hat{F}_h	The high-fidelity approximation
\mathcal{L}	The loss function
N	The number of flow field pairs
n	The number of flow field pairs in the test set
N_x	The numbers of pixels in x dimension
N_y	The numbers of pixels in y dimension
S_f	The freestream wind speed
Φ	The regularization term
λ	The tradeoff parameter of $\Phi(\cdot)$
θ	The parameters of the super-fidelity model
δ	The parameters of the degradation process

In order to narrow the gap between these two kinds of wake models, a series of Machine Learning (ML) methods especially Deep Learning (DL) methods have been proposed over the recent years. For example,

ML algorithms have been widely employed to improve the accuracy of wind farm power prediction with consideration of wake losses, including Long Short Term Memory (LSTM) network [24], Artificial Neural Network (ANN) [25,26], Convolutional Neural Network (CNN) [27] and Graph Neural Network (GNN) [28]. The impacts of atmospheric turbulence and stability measurements on wind farm power prediction were examined in [29]. Aside from power prediction, ML algorithms have also been introduced and applied in wind farm wake modeling. In [30], the comparison of three dimensionality reduction techniques for reducing the flow field dimension was conducted, while a neural network was adopted to forecast the reduced coefficients from the input parameters. Furthermore, based on the Generative Adversarial Network (GAN), a surrogate model trained by high-fidelity data was developed for the wake predictions in [31], which can generate stream-wise and spanwise velocity components simultaneously. In [32], by employing the ML method and RANS/ADM coupling approach, a novel framework for turbine wake predictions was proposed. The influence of Atmospheric Boundary Layer (ABL) flows on wake effects was included in [33] for the prediction of wake velocity. For cooperative yaw control, a double-layer machine learning framework was proposed in [34] using an ANN yawed wake model. The cumulative wake for a wind farm was analyzed by [35] based on Gaussian Process Regression (GPR) model. The local inflow information and wake expansion feature were extracted and their relationship was established by the random forest method in [17]. The random forest was also introduced in order to reconstruct the wake flow of the Vertical-Axis Wind Turbine (VAWT) [36].

The main limitation of the above methods is that the advantages of both analytical models and numerical models are not fully exploited. A brief summary of the main features of those methods is provided in Table 1. To be specific, if only utilizing the data of analytical models, even the most advanced machine learning method can only predict the low-fidelity flow fields. By contrast, if only utilizing the data of high-fidelity numerical models, the generalizability of machine learning methods will be seriously limited by the number of available data samples. Due to the computational requirement, it is technically infeasible to generate a large-scale high-fidelity dataset by numerical models, which is especially true for utility-scale wind farms with dozens of wind turbines. Therefore, the reliability of machine learning methods trained on a small-scale high-fidelity dataset is questionable, especially when generalizing it to an untrained scenario (e.g. a new wind speed) without any prior information. On the other hand, although the flow fields generated by the low-fidelity analytical models do not contain detailed flow features, the general features (such as the yaw effects and wind speeds) are well captured, which can serve as the prior information for high-fidelity flow fields. In summary, the analytical models can provide the basic status while the numerical models can supply precise details of the flow fields. Therefore, by fusing the information from both low-fidelity and high-fidelity models, a novel wind farm wake model with strong generalizability may be formed balancing the accuracy and the efficiency. On the basis of the above consideration, the issue of wake modeling is treated here from the Computer Vision (CV) perspective and defined as a Super-Fidelity (SF) task, i.e. an analog to the Super-Resolution (SR) task, which aims to fully exploit the information from both analytical models and numerical models.

The super-resolution task in CV aims to output high-resolution images from the input low-resolution images [39]. It has been widely applied in real-world applications [40]. Similarly, for the proposed super-fidelity task, the inputs are low-fidelity flow fields generated by analytical models while the outputs are high-fidelity flow fields generated by numerical models. Different from the super-resolution task, the input and output of the super-fidelity task are in the identical resolution but of different fidelities. That is to say, the super-fidelity task targets fidelity rather than resolution. Thus, a novel Super-Fidelity Network (SFNet) is proposed to model and address the defined super-fidelity task as an image-to-image task. To be specific, to maintain

Table 1
The main features of some recently developed wake models. Low-fidelity and High-fidelity represent the fidelity of the data used in the corresponding research.

Reference	Main contribution	Wind turbine wake features		
		Low-fidelity	High-fidelity	Multiple turbines
[37]	Short-term wind speed prediction	Yes	No	Yes
[10]	3D wakes	No	Yes	No
[23]	Evaluation of three ML algorithms	No	Yes	No
[33]	Impact of ABL flows	No	Yes	No
[36]	Mean wake of H-rotor VAWTs	No	Yes	No
[38]	Stochastic expansion of CFD	No	Yes	Yes
[30]	Surrogate modeling	No	Yes	Yes
[31]	Spanwise velocity prediction	No	Yes	Yes

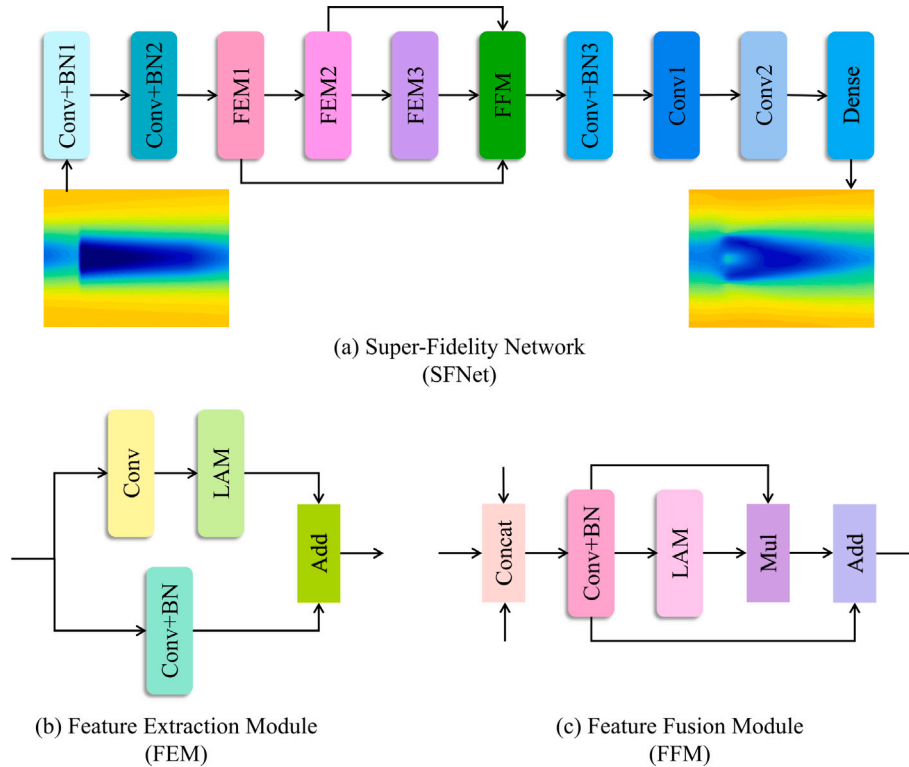


Fig. 1. An overview of the super-fidelity Network, (a) network architecture, (b) the Feature Extraction Module (FEM), and (c) the Feature Fusion Module (FFM). Note that Conv+BN means the Convolution + BatchNorm, LAM denotes the Linear Attention Mechanism, Conv signifies the Convolution layer, add indicates the add operation, Concat represents the Concatenate operation, and mul is the multiplication operation.

both the input and the output in the raw 2D format, the CNN is adopted as the basic layer to construct the proposed SFNet instead of the traditional Multilayer Perceptron (MLP). The former can process the input in the 2D format, thereby extracting the abundant spatial correlations among the pixels. The latter normally reshapes the input into the 1D format first, which would inevitably lead to the loss of spatial information. Meanwhile, the flow convection and diffusion are ubiquitous within wake flow fields. Thus, capturing such complicated and non-local relationships is clearly beyond what CNN is capable of, as CNN, by design, focuses on only local patterns. To model the long-range dependencies of the whole flow fields, the Linear Attention Mechanism (LAM) [41] is then introduced to build the SFNet. As a simplified dot-product attention mechanism, LAM has shown its great potential in the computer vision area. It can model the relationship of every pair of pixels in the input image with $O(N)$ complexity [42]. Therefore, the LAM is employed to enable the SFNet to extract the non-local information of the whole flow fields. The main contributions of this paper are summarized below:

- (1) To rapidly and accurately predict wake effects within wind farms, a novel super-fidelity task is defined to generate high-fidelity flow fields from low-fidelity inputs. The background is that high-fidelity data are usually finite and expensive while low-fidelity data are normally abundant and cheap. The target of the super-fidelity task is to design and train a model that can learn the mapping relationships using limited flow field pairs. Thereafter, the trained model can be generalized to those low-fidelity data without the corresponding high-fidelity pairs. In this way, the accuracy of the low-fidelity data can be enhanced, which is greatly useful for engineering applications such as wind farm layout optimization.
- (2) From the computer vision perspective, a benchmark named Super-Fidelity Network is proposed to address the super-fidelity task. The numerical results show that, compared with high-fidelity data, the Root Mean Square Error (RMSE) predicted by the proposed method is just 4.3% trained on only 9 flow field pairs. The RMSE can be further optimized to only 2.8% using 45 flow field pairs.

Table 2
The detailed setting of each layer in the proposed SFNet.

Name	Input size	Output size	Kernel	Channel	Stride	Padding
Conv+BN1	$1 \times 30 \times 50$	$32 \times 30 \times 50$	(3, 3)	32	(1, 1)	(1, 1)
Conv+BN2	$32 \times 30 \times 50$	$32 \times 30 \times 50$	(3, 3)	32	(1, 1)	(1, 1)
FEM	Conv	$32 \times 30 \times 50$	(1, 1)	32	(1, 1)	(0, 0)
	Conv+BN	$32 \times 30 \times 50$	(3, 3)	32	(1, 1)	(1, 1)
	LAM	$32 \times 30 \times 50$	–	32	–	–
FFM	Conv+BN	$96 \times 30 \times 50$	(3, 3)	96	(1, 1)	(1, 1)
	LAM	$96 \times 30 \times 50$	–	96	–	–
Conv+BN3	$96 \times 30 \times 50$	$32 \times 30 \times 50$	(3, 3)	32	(1, 1)	(1, 1)
Conv1	$32 \times 30 \times 50$	$8 \times 30 \times 50$	(3, 3)	8	(1, 1)	(1, 1)
Conv2	$8 \times 30 \times 50$	$1 \times 30 \times 50$	(1, 1)	1	(1, 1)	(0, 0)
Dense	1×1500	1×1500	–	–	–	–

- (3) The generalizability of the proposed SFNet in wind speeds, yaw angles and array column extensions is comprehensively evaluated and verified through simulation tests. Furthermore, the test on a wind farm with 10×10 turbines demonstrates that our method can maintain stability and robustness for large-scale wind farms.

The remaining part of this paper is organized as follows: the problem formalization, the dataset generation and the proposed SFNet are described in Section 2. Thereafter, the numerical experiments are reported and discussed in Section 3. The conclusions are finally drawn in Section 4.

2. Methodology

2.1. Problem formalization

The super-fidelity task formulated in this work aims at reconstructing the high-fidelity flow fields from the corresponding low-fidelity inputs. Generally, the low-fidelity flow field F_l can be modeled as the output of the following process (known as the degradation in the computer vision community):

$$F_l = \mathcal{D}(F_h; \delta) \quad (1)$$

where \mathcal{D} denotes a degradation mapping function, F_h represents the corresponding high-fidelity flow field and δ means the parameters of the degradation process. For the super-resolution task, the degradation process is normally unknown. But for super-fidelity, the degradation process is caused by the different principles of analytical wake models and numerical wake models. Specifically, the former generates the low-fidelity flow fields based on analytical formulations, while the latter generates the high-fidelity flow fields by solving the NS equations using numerical methods. Similar to the super-resolution task, only the low-fidelity flow fields generated by analytical wake models are available in the vast majority of cases. Then, the target of the super-fidelity task is to recover the approximation \hat{F}_h of the high-fidelity flow field F_h from the low-fidelity input, following:

$$\hat{F}_h = \mathcal{F}(F_l; \theta) \quad (2)$$

where \mathcal{F} is the super-fidelity model and θ represents the parameters of \mathcal{F} . Hence, given a super-fidelity model \mathcal{F} , the target is to narrow the gap between the approximation \hat{F}_h and the high-fidelity flow field F_h as close as possible by optimizing the parameters θ :

$$\begin{aligned} \theta^* &= \arg \min_{\theta} E(\theta), \\ E(\theta) &= \sum_{n=1}^N \mathcal{L}(F_h, \hat{F}_h) + \lambda \Phi(\theta), \\ \mathcal{L}(F_h, \hat{F}_h) &= \mathcal{L}(F_h, \mathcal{F}(F_l; \theta)), \end{aligned} \quad (3)$$

where $E(\theta)$ means the expected loss, the loss function $\mathcal{L}(F_h, \hat{F}_h)$ measures the disparity between the high-fidelity references and the predicted results, $\Phi(\theta)$ is the regularization term weighted by the tradeoff parameter λ , and N represents the number of flow field pairs.

2.2. Dataset of flow field pairs

As illustrated in the problem formalization, the dataset for the super-fidelity task includes two parts: the low-fidelity flow fields as the input and the high-fidelity flow fields as the reference.

For high-fidelity data, the LES flow solver SOWFA (Simulator for Wind Farm Applications) [43] developed by the National Renewable Energy Laboratory (NREL) is employed to solve the filtered NS equations. The simulation domain is $3000 \times 3000 \times 1000$ m, where the inflow wind comes from the southwest direction. For the mesh generation, the two-level mesh refinement is adopted, where the outer mesh dimension, inner mesh dimension and the dimension of the mesh in between are $12 \times 12 \times 12$ m, $3 \times 3 \times 3$ m and $6 \times 6 \times 6$ m, respectively. The total number of cells is 1.8×10^7 to guarantee a 3 m mesh size around the turbine rotors, thereby capturing the detailed turbine wake dynamics. To investigate flow fields both for freestream and upstream wake conditions, three NREL 5 MW baseline turbines [44] operating in a row are simulated, where the 2D mean velocity field around each turbine at the turbine hub height is extracted from the simulation data. In order to include a wide range of operating conditions, three freestream mean wind speeds at 8 m/s, 9 m/s and 10 m/s are considered where each inflow condition contains 30 simulations with different yaw angles in the range of $[-30^\circ, 30^\circ]$. Thus, 90 large eddy simulations have been carried out in total, thereby generating 270 turbine samples. More details about the simulation can be referred to Ref. [30].

For low-fidelity data, the Gaussian analytical wake model implemented in FLORIS [5] is adopted. Specifically, 270 corresponding low-fidelity samples are generated under the same configuration as SOWFA. To be specific, three NREL 5 MW baseline turbines operating in a row are simulated under 8 m/s, 9 m/s and 10 m/s wind speeds with the same yaw angles used for SOWFA. Other operating parameters are also set as the same as SOWFA, such as the air density (1.225), the freestream turbulence intensity (0.06) and the tip speed ratio (8.0).

2.3. Super-fidelity network

To address the super-fidelity task formulated by Eq. (3), the Super-Fidelity Network is designed and built by Convolution + BatchNorm operations (Conv+BN), the Feature Extraction Modules (FEM) and the Feature Fusion Module (FFM). As shown in Fig. 1, the input low-fidelity flow field is first processed by Conv+BN1 and Conv+BN2. Then, the obtained feature maps are successively extracted by three FEMs, while the outputs of FEMs are concatenated and then fused by FFM. Thereafter, the fused feature maps are fed into the Conv+BN3, while the extracted features are further processed by two convolutional layers, i.e. the Conv1 and Conv2. Finally, a densely connected layer is attached to generate the high-fidelity approximation.

As an inherently chaotic flow system, the convection and diffusion phenomena are omnipresent in the turbulent wakes. Thus, the relationships between the local areas and global areas are both important for accurately predicting the wake effects. Especially, the correlations of local areas and global areas can be seen as the local spatial details and global contextual information for computer vision tasks respectively. Therefore, the convolutional layers alone cannot fully capture the non-local relationships, as the CNN mainly extracts local patterns and lacks the ability to model long-range context. In order to address this issue, the feature extraction module is designed which comprises a local branch and a global branch as shown in Fig. 1(b). Specifically, the local branch is a relatively simple structure with a convolutional layer and a batch norm layer. For the global branch, the input feature maps are fed into a convolutional layer and then processed by the

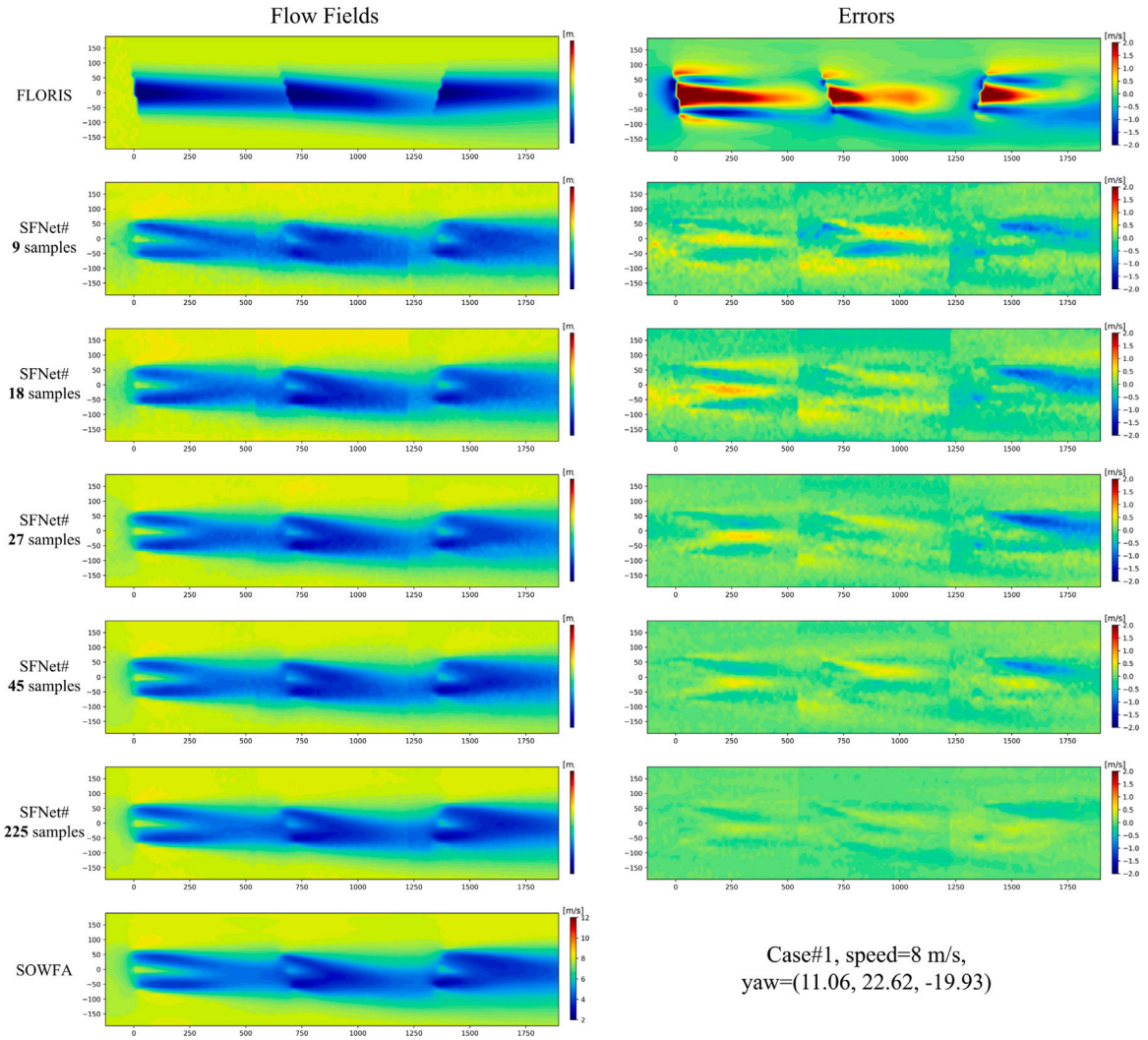


Fig. 2. The predicted flow fields and the corresponding error distributions using FLORIS and SFNet under different training samples, at an example case where the mean inflow wind speed is 8 m/s.

linear attention mechanism to extract the long-range and non-local relationships. The linear attention mechanism [41] is an improved version of the traditional dot-product attention mechanism [45], while the latter has been widely applied in vision-related and language-related tasks benefiting from its strong capabilities in capturing global dependencies [46]. However, the memory and computational costs of the dot-product attention mechanism increase quadratically with the size of the input over space and time, hugely hindering its potential in engineering applications. Therefore, in our previous work [41], we improved the attention mechanism based on the Taylor approximation and proposed the linear attention mechanism with linear complexity, whose effectiveness and efficiency have been verified in vision-related tasks [42]. The detailed mathematical explanation and calculation efficiency are given in [41,42]. Based on the above design, the local relationships can be extracted by the local branch while the global relationships can be captured by the global branch.

In the proposed SFNet, three feature extraction modules are attached successively, while the outputs are concatenated and aggregated by the feature fusion module. The structure of the feature fusion module can be seen in Fig. 1(c), where the concatenated feature maps are first processed by a Conv+BN block and extracted by the LAM operation. Thereafter, the output of the Conv+BN block is multiplied by the output of the LAM operation which is then added with the multiplied feature maps. The details of each component in the SFNet are provided in Table 2.

2.4. Model training

For model training, the Mean Squared Error (MSE) is selected as the loss function:

$$\mathcal{L}(F_h, \hat{F}_h) = \frac{1}{N_x \times N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (F_h^{i,j} - \hat{F}_h^{i,j})^2 \quad (4)$$

where $F_h^{i,j}$ and $\hat{F}_h^{i,j}$ indicate the value of the flow field at the position (i, j) obtained by SOWFA and the SFNet respectively, while N_x and N_y represent the numbers of pixels in x and y dimensions. By minimizing $\mathcal{L}(\cdot)$, the network is driven to approximate the high-fidelity flow fields as much as possible.

The model is optimized by the Adam optimizer with a batch size of 16 and a learning rate of 3×10^{-4} . To enhance the generalizability of the model, the flow field pairs are augmented by horizontal flip, vertical flip and random rotation from -10° to 10° . The probabilities to conduct those augmentation strategies for a pair are all set as 0.1.

2.5. Evaluation metrics

For all experiments, the performance of the super-fidelity results is measured by Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE):

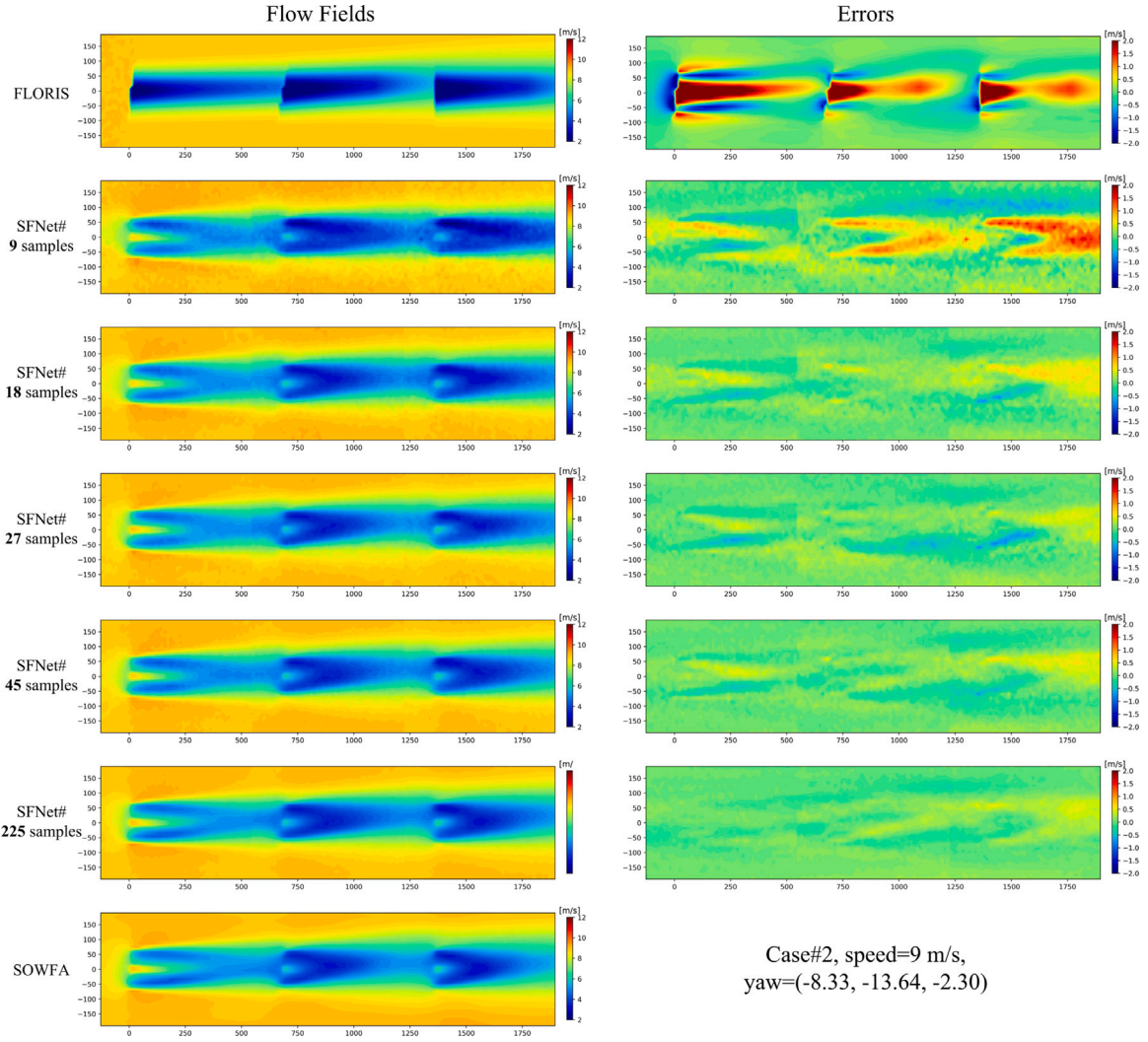


Fig. 3. The predicted flow fields and the corresponding error distributions using FLORIS and SFNet under different training samples, at an example case where the mean inflow wind speed is 9 m/s.

$$MAE = \frac{1}{N_{test}} \sum_{n=1}^{N_{test}} |F_h - \hat{F}_h| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{n=1}^{N_{test}} (F_h - \hat{F}_h)^2} \quad (6)$$

where \hat{F}_h is the approximation predicted by the SFNet model, F_h is the reference high-fidelity flow field and N_{test} is the number of samples in the test set. The metrics are then normalized by the corresponding freestream wind speeds to obtain the relative errors:

$$MAE(\%) = \frac{MAE}{S_f} \times 100\% \quad (7)$$

$$RMSE(\%) = \frac{RMSE}{S_f} \times 100\% \quad (8)$$

where S_f means the freestream wind speed.

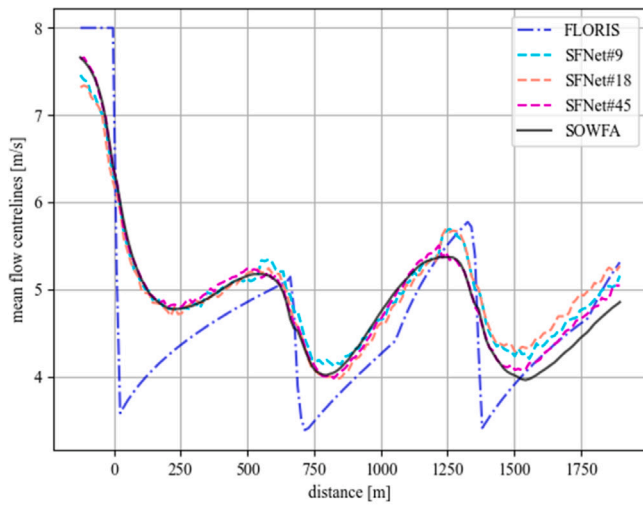
3. Results and discussions

To comprehensively analyze the performance of the proposed SFNet, the ablation study is first carried out to demonstrate the validity of each novel component in our model. Then, the prediction errors on

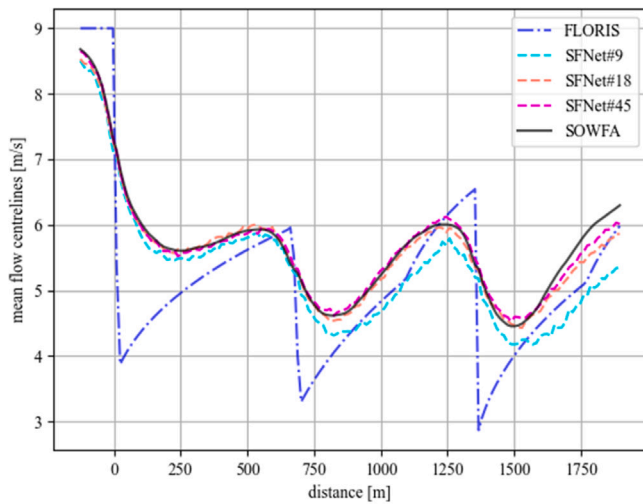
the test set using different training samples are calculated to evaluate the effectiveness of our model under different training scenarios. Thereafter, quantitative evaluations to verify the generalizability are carried out for wind speeds, yaw angles and array column extensions. Finally, a case study is carried out to demonstrate the ability of the proposed SFNet in simulating a large-scale wind farm.

3.1. Ablation study

To address the super-fidelity task, instead of directly adopting the traditional machine learning method such as MLP, the CNN-based SFNet is proposed with two novel modules, i.e. FEM and FFM. Intuitively, both the input and the output of the super-fidelity task are in 2D matrix format rather than 1D array format. With MLP, the samples need to be reshaped into 1D array format. In this way, the spatial correlations will be unavoidably eliminated during the reshaping operation. By contrast, with the convolutional operation, the intact spatial information will be retained and captured. But the initial design philosophy of the CNN determines what it can do is to extract the local spatial relationships within its limited receptive field. Considering the universality of the convection and diffusion phenomena in the wind farm, the local relationships are clearly not enough to model the spatial corrections of the complex wake effects. Thus, two novel modules are designed in the proposed SFNet, i.e. FEM and FFM, to further extract and fuse the feature maps.



(a) Case#1, 8 m/s



(b) Case#2, 9 m/s

Fig. 4. The mean flow centerline for two example cases, predicted by FLORIS, SFNet and SOWFA. SFNet#9, SFNet#18 and SFNet#45 represent the training scenarios using 9, 18 and 45 samples respectively.

Table 3

The ablation study about the FEM and FFM in the proposed SFNet.

Training samples	Model	MAE (m/s)	RMSE (m/s)
9	MLP	0.357	0.509
	Baseline1	0.327	0.471
	Baseline2	0.288	0.415
	SFNet	0.272	0.386
27	MLP	0.247	0.361
	Baseline1	0.233	0.338
	Baseline2	0.227	0.332
	SFNet	0.224	0.325
45	MLP	0.197	0.285
	Baseline1	0.192	0.281
	Baseline2	0.179	0.265
	SFNet	0.174	0.256

To demonstrate the advantage of CNN structure as well as the validity of FEM and FFM, a series of ablation studies are conducted. Specifically, the FFM and the global branch of FEM (the branch contains the LAM) are removed to build Baseline1. Then, only the FFM

of SFNet is removed to build Baseline2. Also, the MLP is implemented for comparison whose parameter count is roughly the same as the SFNet. The experiments are conducted under three conditions with different training samples which can be seen in Table 3. As can be seen, benefiting from the intact spatial information, even the basic baseline (Baseline1) can provide flow fields better than MLP. However, the performance gap will be narrowed with the increase in training samples. The above tendency illustrates that the increase of training samples can offset the defect of algorithms to a certain extent. As the long-range and non-local spatial relationships can be extracted by the global branch of FEM, the performance of Baseline2 is better than Baseline1. Similarly, the effectiveness of FFM can be verified by the comparison between SFNet and Baseline2.

3.2. Performance validation

To validate the effectiveness of the proposed method, the SFNet is trained using different numbers of samples. To be specific, as there are 30 simulations with 90 turbine samples for each wind speed, 1, 2, 3, 4, 5, 10, 15, 20, and 25 cases from each wind speed condition are selected for training and validation. Five cases from each wind condition (i.e. 45 samples in total), which are not used in the training process, are used for testing.

As can be seen from Table 4, trained on only 9 samples (including 1 for validation to prevent the over-fitting issue), the proposed SFNet can still deliver decent predictions with the MAE of only 3.0% and the RMSE of only 4.3%. Apparently, with the increase of training samples, the MAE and RMSE are reduced gradually. When 45 samples are used for training, the MAE and RMSE reach only 1.9% and 2.8% respectively.

To further illustrate the prediction performance, the results for two cases under different wind speeds and yaw angles are provided in Figs. 2 and 3, including the flow fields generated by the SFNet and the error distributions. As can be seen, not only the far-wake fields but also the near-wake features are successfully reconstructed by the SFNet, even when trained on only 9 samples. Specifically, the main features of the flow field including the wake deflection caused by the turbine yaw angle, the wake recovery in the streamwise direction and the upstream wake's impact on the downstream flow field are all accurately captured. Although the errors around the second and third turbines are higher than those of the first turbine due to the more complex wake conditions, this gap can be narrowed with the increase of training samples. By contrast, even though the Gaussian model in FLORIS provides a reasonable representation of far-wake regions, there exists a clear discrepancy in the near-wake flow fields for all three turbines.

The mean flow centerlines in Fig. 4 and velocity profiles in Fig. 5 further reveal the error distributions of the low-fidelity FLORIS data. Taking Case#1 in Fig. 4 as an example, the maximum discrepancy of the mean flow centerline reaches nearly 2.5 m/s for FLORIS data, while the maximum error of the SFNet does not exceed 0.5 m/s even if only trained by 9 samples. In Fig. 5, the velocity profiles near the wake areas of the low-fidelity FLORIS data are far away from the high-fidelity SOWFA data. By contrast, the developed SFNet trained on only 9 samples is able to predict the turbine wake flows at all streamwise locations for both near wake and far wake areas, vital for the accurate and reliable prediction of energy yields.

3.3. Generalizability

Although the proposed SFNet can clearly enhance the accuracy of low-fidelity flow fields when trained on only limited high-fidelity samples, the significance of such a model will be greatly reduced if it is only effective for those cases in the training set. Therefore, in this section, the generalizability of the SFNet in wind speeds, yaw angles and array column extensions, including both interpolations and extrapolations, are comprehensively demonstrated.

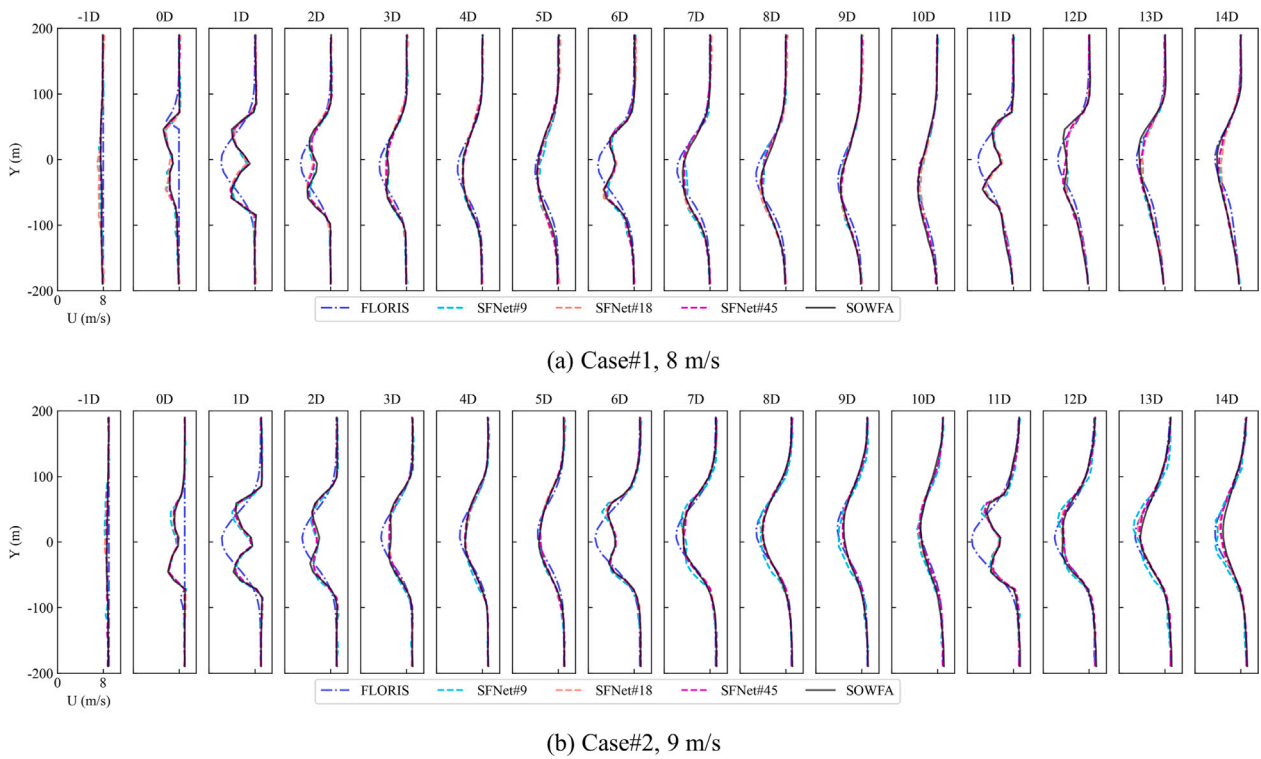


Fig. 5. The velocity profiles for two example cases, predicted by FLORIS, SFNet and SOWFA. SFNet#9, SFNet#18 and SFNet#45 represent the training scenarios using 9, 18 and 45 samples respectively.

Table 4
The results (including the prediction MAE and RMSE) using different numbers of training samples.

Training cases	Training samples	Test Samples	MAE (m/s)	RMSE (m/s)	MAE (%)	RMSE (%)
1	9	45	0.272	0.386	3.0	4.3
2	18	45	0.235	0.334	2.6	3.7
3	27	45	0.224	0.325	2.5	3.6
4	36	45	0.198	0.296	2.2	3.3
5	45	45	0.174	0.256	1.9	2.8
10	90	45	0.159	0.233	1.8	2.6
15	135	45	0.153	0.226	1.7	2.5
20	180	45	0.126	0.181	1.4	2.0
25	225	45	0.116	0.172	1.3	1.9

3.3.1. Generalizability in wind speeds

Three experiments are designed to demonstrate the generalizability of the SFNet in wind speeds. The samples within the whole dataset are divided into three groups according to their freestream mean wind speeds: 8 m/s, 9 m/s and 10 m/s. Thereafter, for each experiment, the proposed SFNet is trained and validated under limited samples only in two groups, and then tested by those samples from the remaining group. For example, for the first experiment, 15 samples from the 8 m/s group and 15 samples from the 9 m/s group are selected to train and validate the model. Then, the performance of the trained model is tested with all samples (90 samples in total) from the 10 m/s group. As those samples with 10 m/s freestream mean wind speed are totally unavailable during the training procedure, the generalizability of the proposed SFNet from 8 m/s and 9 m/s scenarios to 10 m/s scenario can be truly verified. The experiments using 8 m/s and 10 m/s groups to test the 9 m/s group, as well as using 9 m/s and 10 m/s groups to test the 8 m/s group, are conducted in a similar way.

As shown in Table 5, although the test samples are totally unseen during the training phase, both the MAE and RMSE are still at a relatively low level for all three scenarios, especially when considering the limited number of training samples (only 15 samples for each wind speed). Taking the second scenario as an example, trained by samples under 8 m/s and 10 m/s, the proposed SFNet can be generalized to

Table 5
The results on the generalizability of the proposed SFNet in wind speeds.

Training speeds	Test speed	MAE (m/s)	RMSE (m/s)	MAE (%)	RMSE (%)
8 m/s 9 m/s	10 m/s	0.214	0.318	2.1	3.2
8 m/s 10 m/s	9 m/s	0.187	0.278	2.1	3.1
9 m/s 10 m/s	8 m/s	0.185	0.272	2.3	3.4

those samples under 9 m/s with the MAE of only 2.1% and the RMSE of 3.1%. Apparently, the errors will locate in a similar or lower range when generalizing the SFNet to other wind speeds between 8 m/s and 10 m/s. The flow field results in Fig. 6 further illustrate the generalizability of the SFNet in wind speeds.

3.3.2. Generalizability in yaw angles

Two experiments are designed to demonstrate the generalizability of the SFNet in yaw angles. In the first experiment, 30 samples whose yaw angles are located within $[-20^\circ, 20^\circ]$ are taken as the training set, while those samples with yaw angles in $[-30^\circ, -20^\circ]$ or $[20^\circ, 30^\circ]$ are all used as the test set. Similarly, 30 samples whose yaw angles are larger than 10° or less than -10° are selected as the training set, while those samples with yaw angles in $[-10^\circ, 10^\circ]$ are all used as the test set. As the yaw angles in the training set and test set are distributed in

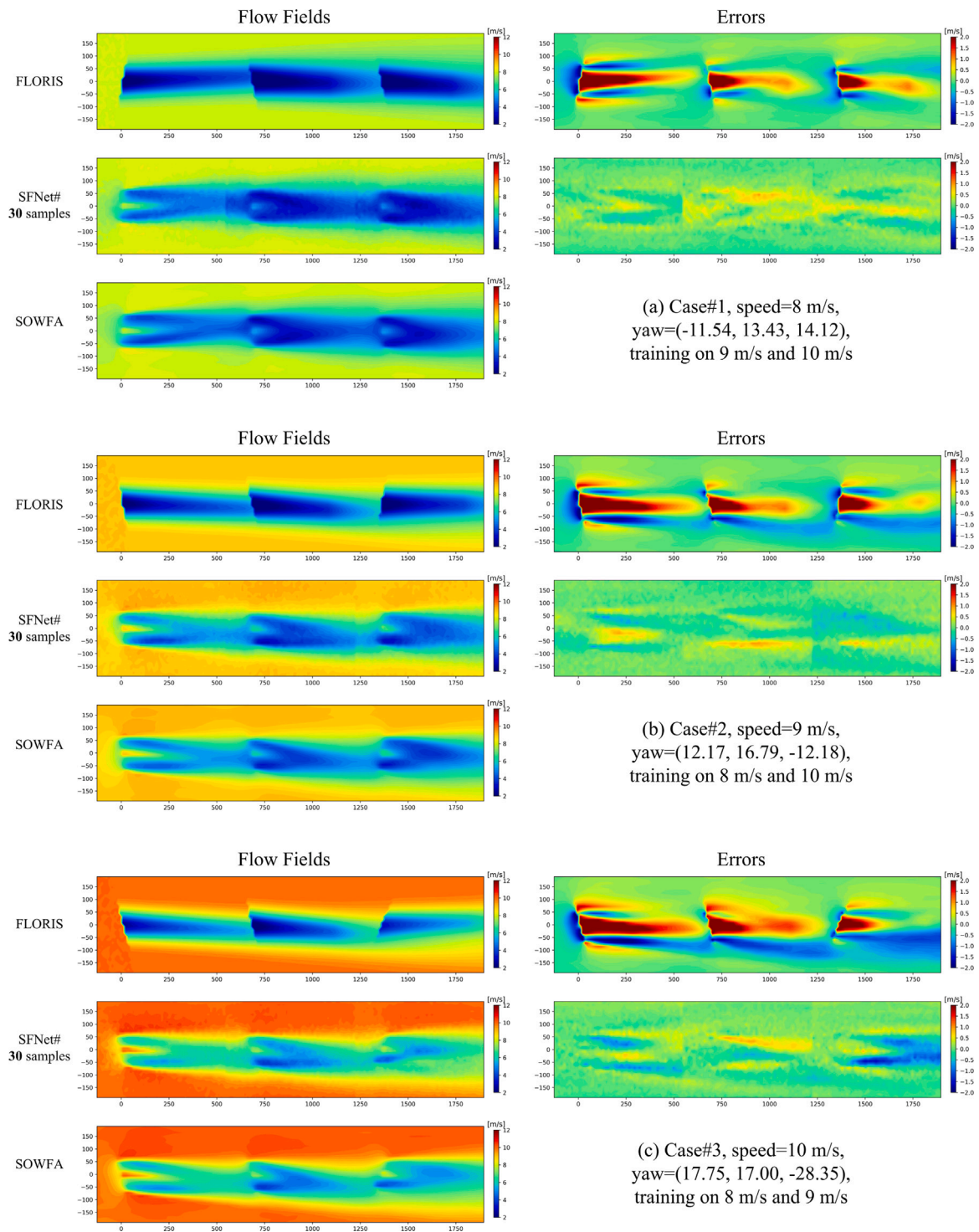


Fig. 6. The predicted flow fields and the corresponding error distributions using FLORIS and SFNet under different training settings of wind speeds, at three example cases.

Table 6
The results on the generalizability of the proposed SFNet in yaw angles.

Training yaw angles	Test yaw angles	MAE (m/s)	RMSE (m/s)	MAE (%)	RMSE (%)
$[-20^\circ, 20^\circ]$	$[-30^\circ, 20^\circ]$ and $[20^\circ, 30^\circ]$	0.252	0.372	2.8	4.1
$[-30^\circ, -10^\circ]$ and $[10^\circ, 30^\circ]$	$[-10^\circ, 10^\circ]$	0.223	0.315	2.5	3.5

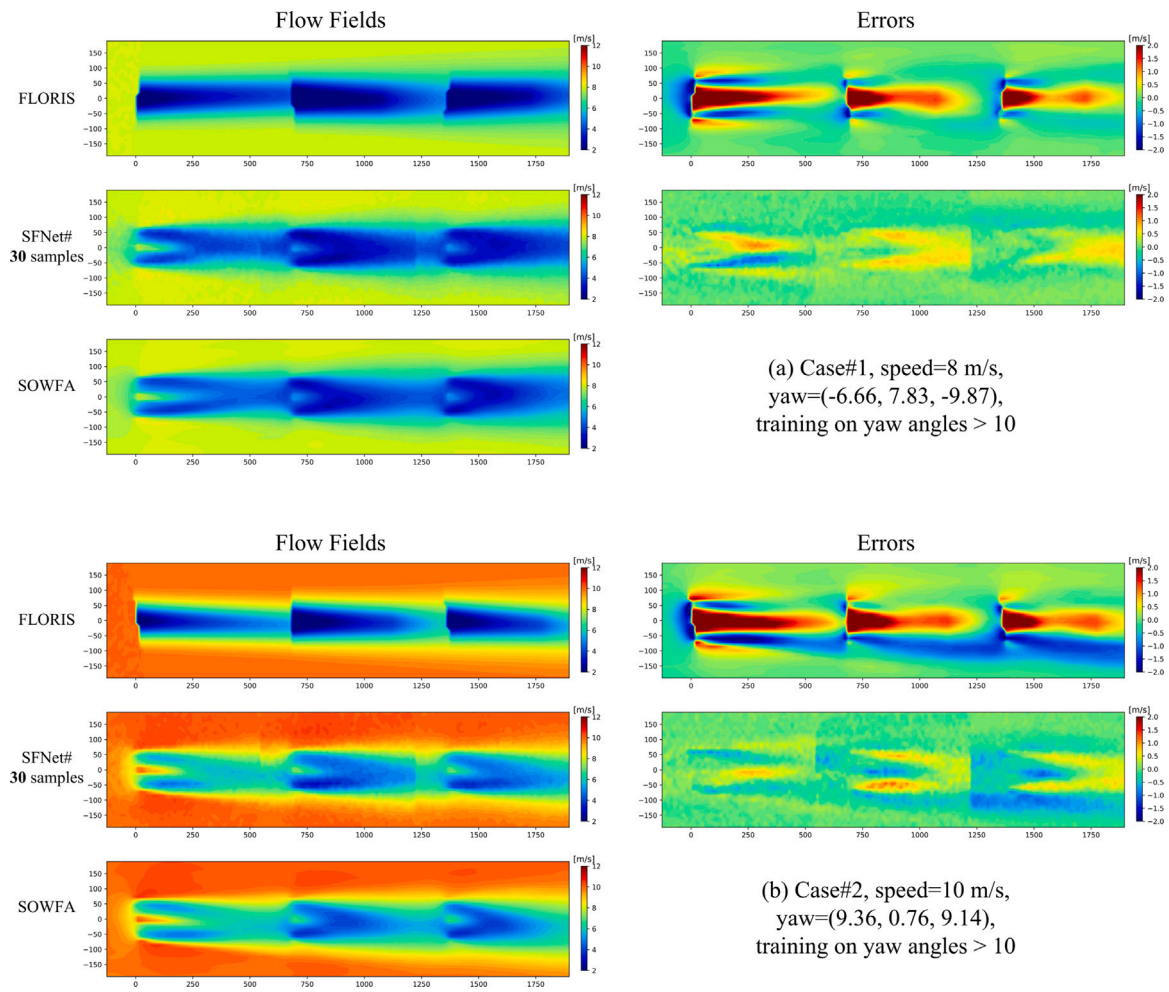


Fig. 7. The predicted flow fields and the corresponding error distributions using FLORIS and SFNet at two example cases, where the SFNet is trained on samples with yaw angles greater than 10° .

totally different ranges, the generalizability of the proposed SFNet in yaw angles can be verified.

As shown in Table 6, trained by a specific range of yaw angles, the proposed SFNet can be generalized to those samples with a different range of yaw angles with low prediction errors. Further, these two experiments demonstrate that the SFNet can be generalized to samples with larger or smaller yaw angles compared with the training samples, respectively. Obviously, the MAE and RMSE of the SFNet will reach a lower level when predicting the yaw angles located in the training yaw angle range, even though the training samples can only cover limited values of the potential yaw angles. Qualitative results in Figs. 7 and 8 further demonstrate the generalizability of the SFNet in yaw angles.

3.3.3. Generalizability in array column extensions

The whole simulation is carried out based on an array with three turbines operating in a row (a 1×3 array). Thus, the effectiveness of the model when extending to larger multiple columns (such as 1×5 or 1×10 arrays), i.e. the generalizability of the model in column extensions, is unverified and unwarranted. To demonstrate the generalizability of the SFNet in extending to multiple columns, the training samples are divided into three groups according to their columns: the first column, the second column and the third column. Then, 15 samples in the first column and 15 samples in the second column are selected to train and validate the model, while whole samples located in the third column are taken as the test set. That is to say, the model is trained by limited

Table 7

The results on the generalizability of the proposed SFNet in array column extensions.

Training turbines	Test turbine	MAE (m/s)	RMSE (m/s)	MAE (%)	RMSE (%)
1st 2nd	3rd	0.306	0.397	3.4	4.4

samples from a 1×2 array and then tested by the extended turbines on the third column. As those extended turbines are not in the training set, the generalizability of the SFNet in extending to more columns of turbines can thus be verified.

As shown in Table 7 and Fig. 9, the proposed SFNet can be generalized to the turbine which is not contained in the training set. Trained on only 30 samples from a 1×2 array, the errors on the extended third turbine are only 3.4% measured by MAE and 4.4% measured by RMSE. As for the extensions in the spanwise direction, it is straightforward as the wake interactions mainly take place in the streamwise direction. Therefore, after verifying the generalization performance, the prediction using the proposed SFNet for the utility-scale wind farms can thus be carried out.

3.4. Model predictions

To demonstrate the effectiveness of the developed SFNet for large-scale wind farm wake predictions, the low-fidelity flow fields of a 10×10 wind farm is simulated using FLORIS as the input. The yaw angles

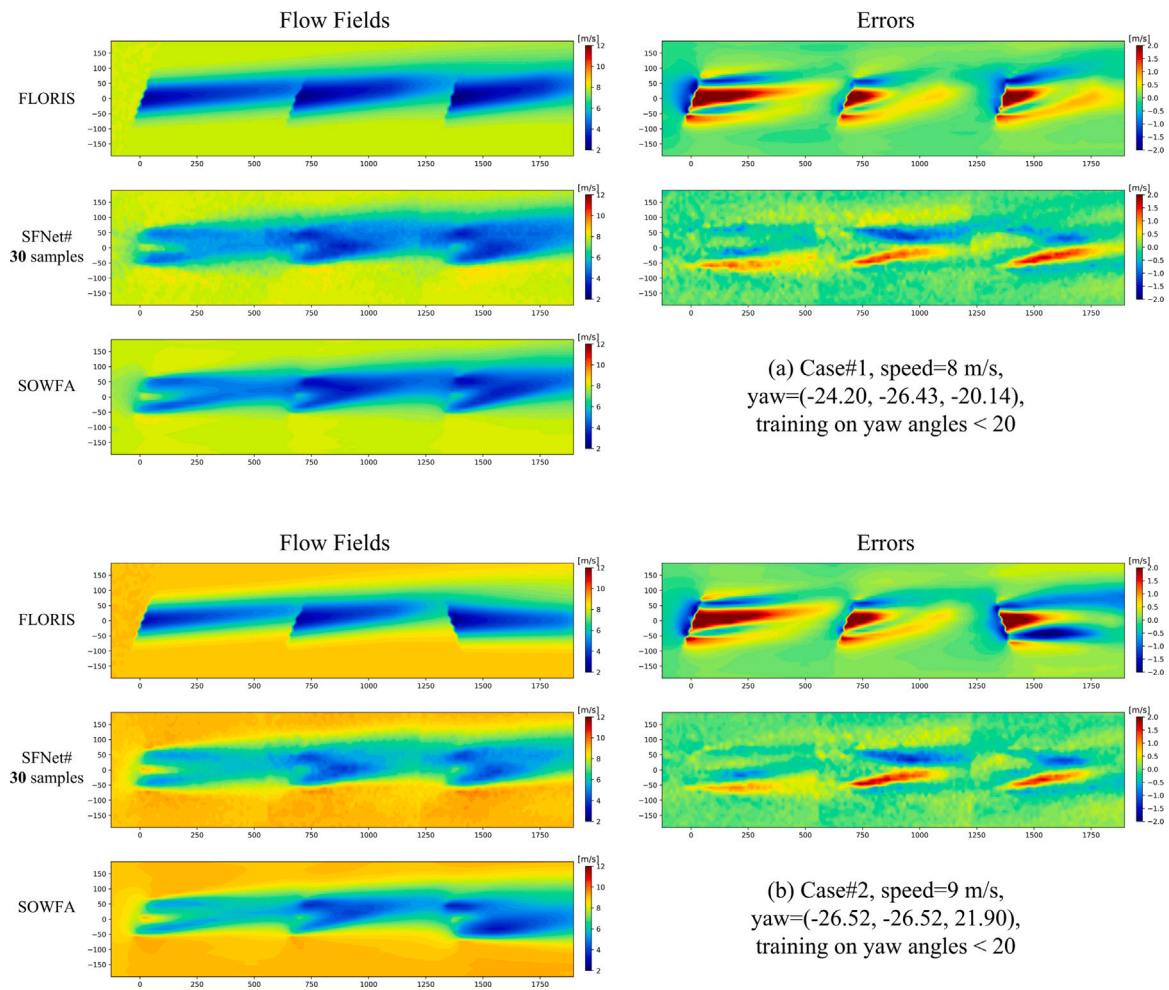


Fig. 8. The predicted flow fields and the corresponding error distributions using FLORIS and SFNet at two example cases, where the SFNet is trained on samples with yaw angles less than 20° .

are randomly selected between $[-20^\circ, 20^\circ]$ for each turbine with the 9 m/s freestream mean wind speed, while the SFNet trained on 45 samples is employed for prediction. As illustrated in Fig. 10, the near-wake features, wake interactions and yaw effects for all turbines are all successfully reconstructed. In particular, the whole low-fidelity data generation and model prediction procedure are conducted on a single Intel Core i7-7700 CPU and completed within several seconds. While on the other hand, if an LES model were used, massive computing resources would be required for such a large-scale wind farm. By contrast, what the SFNet needs is only a limited set of high-fidelity data samples and a low-fidelity analytical model. Then, the developed SFNet could deliver a decent flow fields prediction as rapid as the low-fidelity model while retaining the features captured in the high-fidelity model.

3.5. Discussions

As illustrated, the proposed SFNet has the ability to accurately generate high-fidelity flow fields based on the low-fidelity input even if trained on limited samples. Most importantly, the proposed pipeline for wake modeling is much more computational-friendly than LES models. To be specific, the high-fidelity flow fields generation procedure for each case requires about 1.13×10^4 CPU hours which can be completed within 44 h using local HPC clusters with 256 processors. Based on the comprehensive experiments, five training cases are sufficient to guarantee the performance of the proposed SFNet. Therefore, about

1.70×10^5 CPU hours are required to simulate the high-fidelity flow fields under three different wind speeds (five cases for each speed). Then, the high-fidelity wake interactions for the utility-scale wind farm, such as the 10×10 wind farm in Fig. 10, can be predicted by our pipeline within seconds using a standard desktop. In sharp contrast, if using an LES model to simulate such a large-scale wind farm under various operating conditions, the computational requirement will be enormous.

Considering both the accuracy and efficiency, the proposed SFNet has great potential for practical applications. Specifically, the flow field can be predicted by the proposed SFNet accurately for both far-wake and near-wake features. Thus, the applications such as the wind farm energy yield prediction which can be implemented with SOWFA, can also be implemented with the proposed SFNet but more efficiently. Taking the power generation prediction as an example, it is cubically related to the wind speed. Thus, a small prediction error of the flow fields will lead to a large error in power prediction. As the substantial experiments showed, the proposed SFNet can predict the flow fields much better than the analytical models and meanwhile significantly faster than the numerical models. In summary, the proposed SFNet combines both the great efficiency of the low-fidelity model and the high accuracy of the high-fidelity model.

On the other hand, there also exist some limitations in the proposed pipeline. First, the current pipeline focuses on bi-fidelity data, and cannot take advantage of the dataset with three or more levels of fidelities. In the future, the proposed SFNet will be extended to address this issue

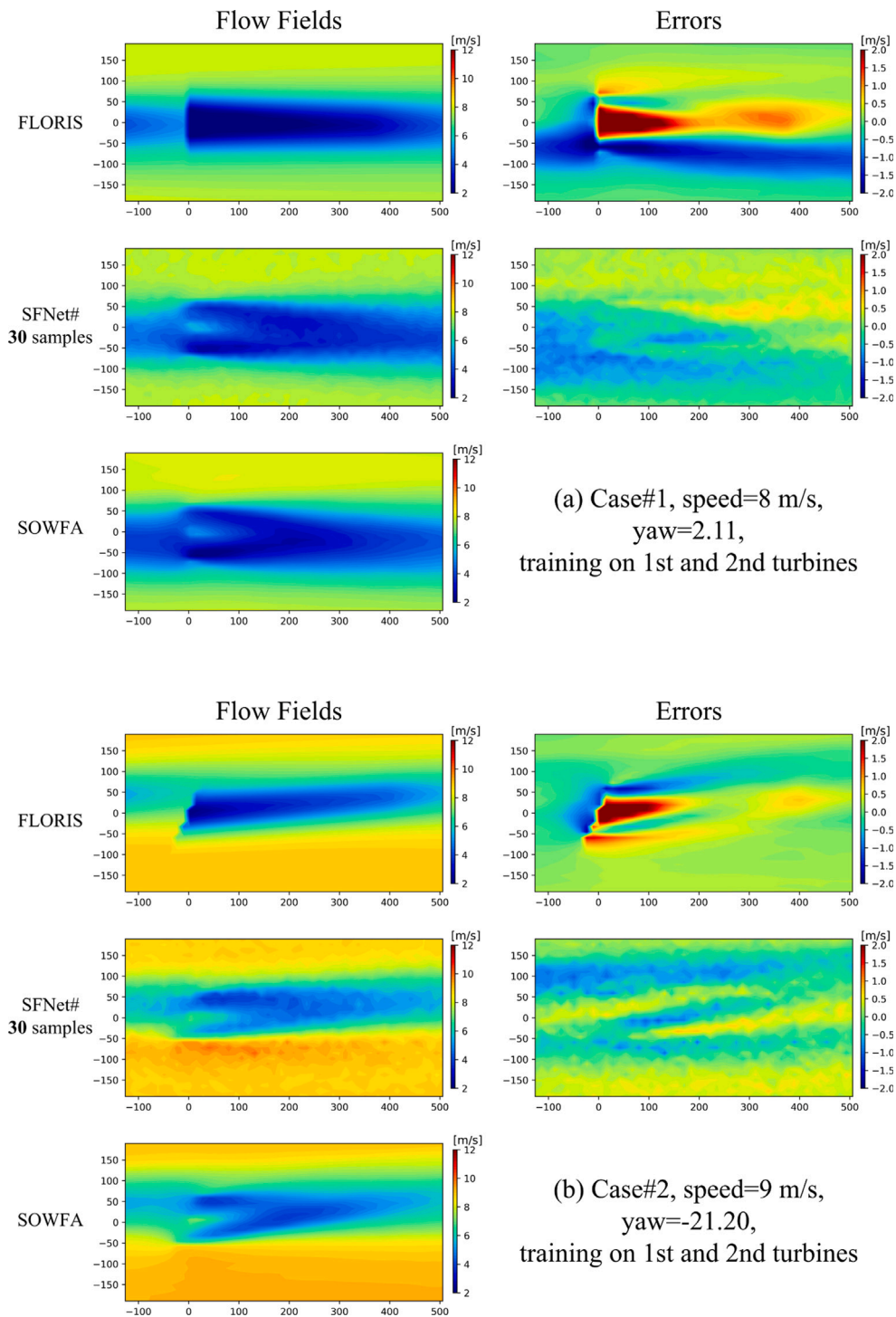


Fig. 9. The predicted flow fields and the corresponding error distributions using FLORIS and SFNet at two example cases, where the SFNet is trained on samples on the 1st and 2nd columns.

where the model will be trained to fuse information from various data sources. Second, limited by the dataset volume, the proposed SFNet is currently a two-dimensional wake model focusing on the flow fields at the turbine hub height. The application of the proposed SFNet in the three-dimensional scenario will be investigated in the future.

4. Conclusions

In this work, a novel Super-Fidelity Network was proposed for wind farm wake modeling, which, to the best of our knowledge, is

the first attempt to combine both low-fidelity analytical wake models and high-fidelity numerical models through deep learning methods. Taking the low-fidelity data generated by analytical models as the input, the developed SFNet is able to predict flow fields similarly as high-fidelity numerical models after trained on limited samples. The numerical experiments demonstrated that the proposed SFNet was able to significantly enhance the low-fidelity input even when trained and validated on only 9 samples. More importantly, the near-wake features, which are normally oversimplified by the analytical models, have been finely reconstructed by the developed SFNet.

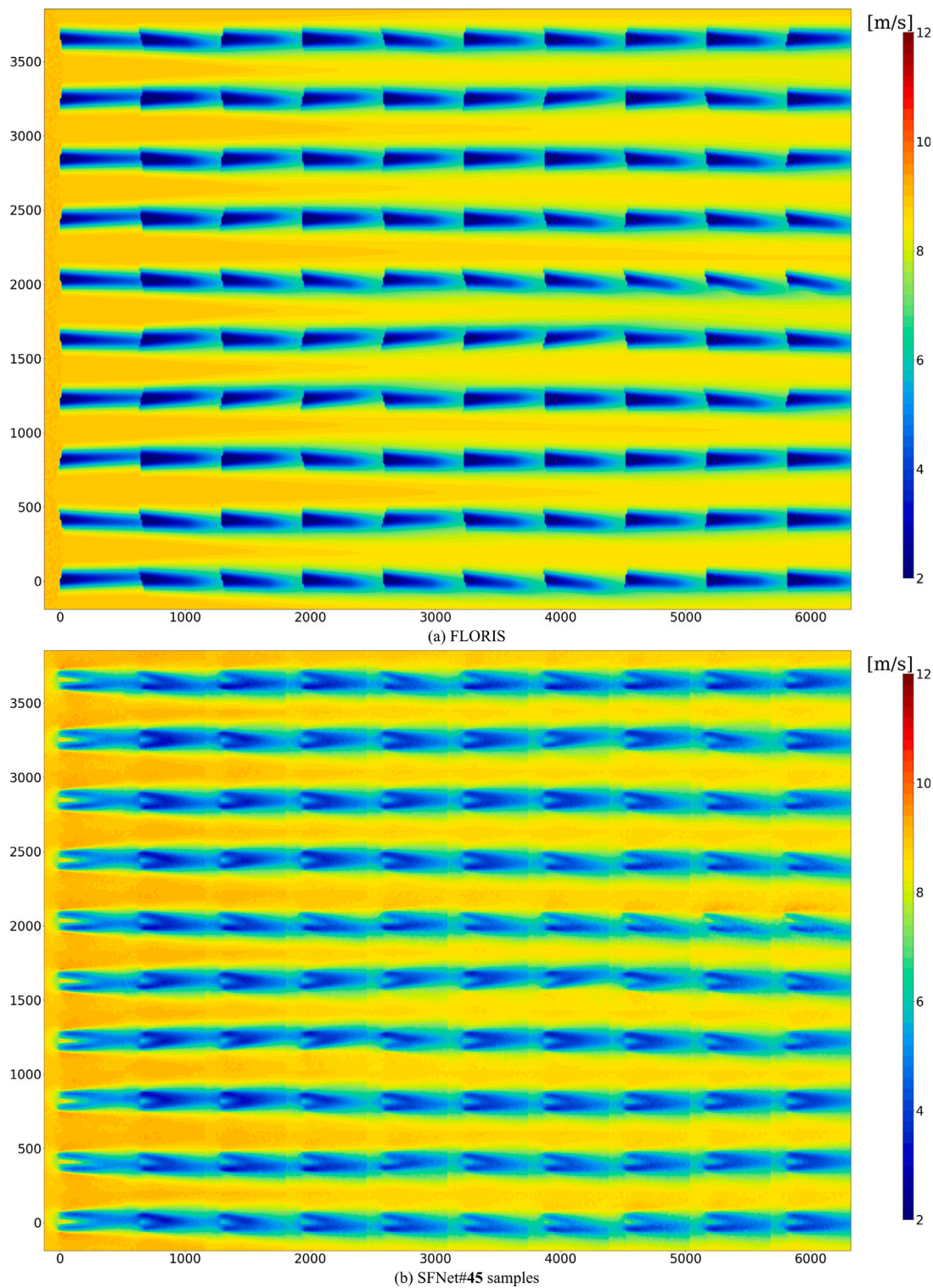


Fig. 10. The predicted flow fields around an 10×10 example wind farm by the FLORIS and the SFNet trained on 45 samples. The yaw angles are randomly selected between $[-20^\circ, 20^\circ]$ for each turbine and the mean wind speed is chosen as 9 m/s.

Compared with the existing machine learning wake modeling methods, the clear advantage of the SFNet was the strong generalizability. In our pipeline, the robust and consistent performance of the SFNet for untrained scenarios was guaranteed by the prior information provided by the low-fidelity input. A series of numerical experiments demonstrated

the stability and generalizability of the developed method for extending to the unseen range of wind speeds, yaw angles and array columns, even if trained on a very limited number of samples. Moreover, trained on only 45 samples, the SFNet is able to generate the flow fields for a 10×10 wind farm successfully.

The future work may involve the use of the proposed SFNet wake model for wind farm layout optimization. Another interesting direction is to extend the SFNet to dynamic wind farm wake modeling.

CRedit authorship contribution statement

Rui Li: Formal analysis, Conceptualization, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft. **Jincheng Zhang:** Formal analysis, Investigation, Conceptualization, Data curation, Project administration, Software, Visualization, Writing – review & editing. **Xiaowei Zhao:** Conceptualization, Funding acquisition, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

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.

Acknowledgments

This work has received funding from the UK Engineering and Physical Sciences Research Council (grant number: EP/S000747/1). The authors also acknowledge the Scientific Computing Research Technology Platform (SCRTP) at the University of Warwick for providing High-Performance Computing resources.

References

- Zhang J, Zhao X. Three-dimensional spatiotemporal wind field reconstruction based on physics-informed deep learning. *Appl Energy* 2021;300:117390.
- He R, Yang H, Sun H, Gao X. A novel three-dimensional wake model based on anisotropic Gaussian distribution for wind turbine wakes. *Appl Energy* 2021;296:117059.
- Jensen N. A note on wind generator interaction. *Risø-M, Risø National Laboratory*; 1983, no. 2411.
- Katic I, Højstrup J, Jensen NO. A simple model for cluster efficiency. In: *European wind energy association conference and exhibition*, Vol. 1. 1986, p. 407–10.
- Gebraad PM, Teeuwisse F, Van Wingerden J, Fleming PA, Ruben S, Marden J, Pao L. Wind plant power optimization through yaw control using a parametric model for wake effects—a CFD simulation study. *Wind Energy* 2016;19(1):95–114.
- Frandsen S, Barthelmie R, Pryor S, Rathmann O, Larsen S, Højstrup J, Thøgersen M. Analytical modelling of wind speed deficit in large offshore wind farms. *Wind Energy: Int J Prog Appl Wind Power Convers Technol* 2006;9(1–2):39–53.
- Gao X, Yang H, Lu L. Optimization of wind turbine layout position in a wind farm using a newly-developed two-dimensional wake model. *Appl Energy* 2016;174:192–200.
- Ge M, Wu Y, Liu Y, Li Q. A two-dimensional model based on the expansion of physical wake boundary for wind-turbine wakes. *Appl Energy* 2019;233:975–84.
- Tian L, Zhu W, Shen W, Zhao N, Shen Z. Development and validation of a new two-dimensional wake model for wind turbine wakes. *J Wind Eng Ind Aerodyn* 2015;137:90–9.
- Sun H, Yang H. Study on an innovative three-dimensional wind turbine wake model. *Appl Energy* 2018;226:483–93.
- Sun H, Yang H. Numerical investigation of the average wind speed of a single wind turbine and development of a novel three-dimensional multiple wind turbine wake model. *Renew Energy* 2020;147:192–203.
- Shapiro CR, Gayme DF, Meneveau C. Modelling yawed wind turbine wakes: a lifting line approach. *J Fluid Mech* 2018;841:R1.
- Lopez D, Kuo J, Li N. A novel wake model for yawed wind turbines. *Energy* 2019;178:158–67.
- Dou B, Guala M, Lei L, Zeng P. Wake model for horizontal-axis wind and hydrokinetic turbines in yawed conditions. *Appl Energy* 2019;242:1383–95.
- Brognia R, Feng J, Sørensen JN, Shen WZ, Porté-Agel F. A new wake model and comparison of eight algorithms for layout optimization of wind farms in complex terrain. *Appl Energy* 2020;259:114189.
- Reddy SR. Wind farm layout optimization (WindFLO): An advanced framework for fast wind farm analysis and optimization. *Appl Energy* 2020;269:115090.
- Nai-Zhi G, Ming-Ming Z, Bo L. A data-driven analytical model for wind turbine wakes using machine learning method. *Energy Convers Manage* 2022;252:115130.
- Lu H, Porté-Agel F. Large-eddy simulation of a very large wind farm in a stable atmospheric boundary layer. *Phys Fluids* 2011;23(6):065101.
- Churchfield MJ, Lee S, Michalakes J, Moriarty PJ. A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics. *J Turbul* 2012;13:N14.
- Nilsson K, Ivanell S, Hansen KS, Mikkelsen R, Sørensen JN, Breton S-P, Henningson D. Large-eddy simulations of the Lillgrund wind farm. *Wind Energy* 2015;18(3):449–67.
- Calaf M, Meneveau C, Meyers J. Large eddy simulation study of fully developed wind-turbine array boundary layers. *Phys Fluids* 2010;22(1):015110.
- Meyers J, Meneveau C. Large eddy simulations of large wind-turbine arrays in the atmospheric boundary layer. In: *48th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition*. 2010, p. 827.
- Purohit S, Ng E, Kabir IFSA. Evaluation of three potential machine learning algorithms for predicting the velocity and turbulence intensity of a wind turbine wake. *Renew Energy* 2022;184:405–20.
- Göçmen T, Giebel G. Data-driven wake modelling for reduced uncertainties in short-term possible power estimation. In: *Journal of physics: conference series*, Vol. 1037. IOP Publishing; 2018, 072002.
- Ti Z, Deng XW, Zhang M. Artificial neural networks based wake model for power prediction of wind farm. *Renew Energy* 2021;172:618–31.
- Sun H, Qiu C, Lu L, Gao X, Chen J, Yang H. Wind turbine power modelling and optimization using artificial neural network with wind field experimental data. *Appl Energy* 2020;280:115880.
- Yildiz C, Acikgoz H, Korkmaz D, Budak U. An improved residual-based convolutional neural network for very short-term wind power forecasting. *Energy Convers Manage* 2021;228:113731.
- Park J, Park J. Physics-induced graph neural network: An application to wind-farm power estimation. *Energy* 2019;187:115883.
- Optis M, Perr-Sauer J. The importance of atmospheric turbulence and stability in machine-learning models of wind farm power production. *Renew Sustain Energy Rev* 2019;112:27–41.
- Zhang J, Zhao X. Machine-learning-based surrogate modeling of aerodynamic flow around distributed structures. *AIAA J* 2021;59(3):868–79.
- Zhang J, Zhao X. Wind farm wake modeling based on deep convolutional conditional generative adversarial network. *Energy* 2022;238:121747.
- Ti Z, Deng XW, Yang H. Wake modeling of wind turbines using machine learning. *Appl Energy* 2020;257:114025.
- Kabir IFSA, Safiyullah F, Ng E, Tam VW. New analytical wake models based on artificial intelligence and rivaling the benchmark full-rotor CFD predictions under both uniform and ABL inflows. *Energy* 2020;193:116761.
- Yang S, Deng X, Ti Z, Yan B, Yang Q. Cooperative yaw control of wind farm using a double-layer machine learning framework. *Renew Energy* 2022;193:519–37.
- Barasa M, Li X, Zhang Y, Xu W. The balance effects of momentum deficit and thrust in cumulative wake models. *Energy* 2022;246:123399.
- Dong Z, Chen Y, Zhou D, Su J, Han Z, Cao Y, Bao Y, Zhao F, Wang R, Zhao Y, et al. The mean wake model and its novel characteristic parameter of H-rotor VAWTs based on random forest method. *Energy* 2022;239:122456.
- Dhiman HS, Deb D, Foley AM. Bilateral Gaussian wake model formulation for wind farms: a forecasting based approach. *Renew Sustain Energy Rev* 2020;127:109873.
- Richmond M, Sobey A, Pandit R, Kolios A. Stochastic assessment of aerodynamics within offshore wind farms based on machine-learning. *Renew Energy* 2020;161:650–61.
- Chen H, He X, Qing L, Wu Y, Ren C, Sheriff RE, Zhu C. Real-world single image super-resolution: A brief review. *Inf Fusion* 2022;79:124–45.
- Wang Z, Chen J, Hoi SC. Deep learning for image super-resolution: A survey. *IEEE Trans Pattern Anal Mach Intell* 2020;43(10):3365–87.
- Li R, Zheng S, Duan C, Su J, Zhang C. Multistage attention resU-Net for semantic segmentation of fine-resolution remote sensing images. *IEEE Geosci Remote Sens Lett* 2021;19:1–5.
- Li R, Zheng S, Zhang C, Duan C, Wang L, Atkinson PM. ABCNet: Attentive bilateral contextual network for efficient semantic segmentation of fine-resolution remotely sensed imagery. *ISPRS J Photogramm Remote Sens* 2021;181:84–98.
- Churchfield M, Lee S. *NWTC information portal (SOWFA)*. National Renewable Energy Laboratory; 2014, URL <https://nwtc.nrel.gov/SOWFA>.
- Jonkman J, Butterfield S, Musial W, Scott G. Definition of a 5-MW reference wind turbine for offshore system development. *Tech. Rep., National Renewable Energy Lab.(NREL)*, Golden, CO (United States); 2009.
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. *Adv Neural Inf Process Syst* 2017;30.
- Li R, Wang L, Zhang C, Duan C, Zheng S. A2-FPN for semantic segmentation of fine-resolution remotely sensed images. *Int J Remote Sens* 2022;43(3):1131–55.