Energy 239 (2022) 122456

Contents lists available at ScienceDirect

Energy

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

The mean wake model and its novel characteristic parameter of H-rotor VAWTs based on random forest method

Zhikun Dong [a], Yaoran Chen [a], Dai Zhou [a, b, c, *], Jie Su [a], Zhaolong Han [a, b, c, d], Yong Cao [a], Yan Bao [a, c], Feng Zhao [a], Rui Wang [a], Yongsheng Zhao [a, b], Yuwang Xu [a, b]

[a] School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China
[b] State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai, 200240, PR China
[c] Key Laboratory of Hydrodynamics of Ministry of Education, Shanghai Jiao Tong University, Shanghai, 200240, PR China
[d] Institute of Polar and Ocean Technology, Institute of Marine Equipment, Shanghai Jiao Tong University, Shanghai, 200240, China

ARTICLE INFO

Article history:
Received 18 May 2021
Received in revised form
11 October 2021
Accepted 25 October 2021
Available online 26 October 2021

Keywords:
Vertical axis wind turbine
Wake model
Wake analysis
Random forest
Feature importance

ABSTRACT

Using the random forest (RF) algorithm, this study presented a key parameter to characterize the mean wake of H-rotor VAWTs while modelling the wake. First, the RF algorithm was used to establish the regression relationship between the average wake velocity distribution and the rotor features. Next, the feature crosses method was combined with the RF algorithm to analyze the interaction and importance of the inputs. It was found that the normalized importance of a synthetic feature in wake modelling occupied a considerable significance, reaching 0.884 out of 1. The RF wake model with this parameter as the only input feature could successfully reconstruct the wake. It was found that this feature may reflect the ability of incident wind passing through the operating rotor and played a decisive role in the wake velocity distribution, including initial velocity deficit and wake recovery rate. The universality of this parameter was proved through cases analysis of wind turbines under different sizes and operating conditions. The study of the wake field is important for the modelling of the H-rotor VAWT wake field, and hence affects the optimal configuration of the wind farm.

1. Introduction

With the progress of science and technology and the reduction of development costs, the share of renewable energy in global energy supply is increasing [1]. According to International Energy Agency [2], wind energy accounts for a considerable proportion of this growth, and the absolute generation of wind power is expected to increase the most among all renewable energy sources. More large-scale wind turbines are installed in clusters in different layouts to form wind farms, especially in places with rich wind resources such as offshore areas [3]. In this development trend, VAWTs show great advantages compared with horizontal axis wind turbines (HAWTs) since they do not require yawing [4], are easy to upscale, and have the ability to improve the efficiency of power generation when closely arranged [5]. Among VAWTs, the H-rotor VAWTs are more potential due to their high power generation efficiency and simple structure [4].

For large-scale wind farms, the mutual interaction between wind turbines through wake becomes the key factor limiting the overall generation efficiency of wind farms [6]. The wake of wind turbine is characterized by higher turbulence levels and a large velocity deficit, especially when compared with the free stream wind [7]. Higher turbulence increases structural fatigue and extreme loads of turbines [8], while velocity deficits will result in a significant reduction in power generation for the downstream turbines, and the power losses can easily reach 40% under the full-wake conditions [9]. To minimize such losses, the layout of wind farm needs to be adjusted continuously during design, which requires rapid calculation of power efficiency. In practice, the efficiency is often calculated by superimposing the wake speed defects of a single wind turbine [10]. Therefore, studying the wake of a single wind turbine is the beginning of wind farm layout optimization.

According to the distance from the wind turbine, the wake can be divided into near wake region (2–4 rotor diameters) and far wake region (>4 rotor diameters). The flow distribution of the near

* Corresponding author. School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China.
E-mail address: zhoudai@sjtu.edu.cn (D. Zhou).

https://doi.org/10.1016/j.energy.2021.122456
0360-5442/

wake is highly complex and heterogeneous [11], while the far wake region has more general characteristics (e.g., average flow distribution [12]), since the effect of the detailed features of the rotor on the flow field is gradually reduced. In wind farms, the space each turbine is usually greater than 4 rotor diameters [13], which means the downstream wind turbines usually operate in the far wake region of the upstream turbines. Therefore, studying the average flow distribution of the far wake area is essential to predict and improve the long-term output power of wind farms.

Considerable research efforts have been devoted to the experimental, analytical and numerical studies of the wake of VAWTs. Herein, only the studies of the wake of the H-rotor VAWTs are considered. In recent years, scholars have focused on the measurement of near wake [14–17] and far wake [18–21] of H-rotor VAWTs. And some analytical models have been created to attempt to describe the wake characteristics of VAWTs, especially the mean velocity distribution [22,23]. These analytical models greatly reduce the complexity of wake prediction, but were not completely accurate. The types of numerical simulation used until now can be divided into two categories according to the different treatment of the rotor. Option (1) is modeling and meshing, a technique directly based on the geometry of the blade [24–26]. The actuator line technique (2) can be applied when the line carrying the body force is used to replace the real geometry of the rotor [28–30]. Method (1) can provide various detailed characteristics of the rotor and wake, but its computational cost is high. Since a wind farm usually contains many turbines, its optimal design is often a process of continuous iteration, which makes numerical simulation calculations too expensive.

In recent years, machine learning (ML), as a relatively new method, has been widely used in the research of VAWTs. Scholars in the field have applied the method to create new wind turbine prototypes [31,32], monitor the working state [33] and predict the performance of wind turbines [34–37]. In these studies, algorithms such as artificial neural networks (ANN) and RF have achieved very good results and could quick and accurately model any nonlinear relationship between input-output variables. This lays a foundation for the study of wake flow, since the velocity distribution of the wake flow is also a nonlinear function of the features of the turbine and its environmental features [12], including blade tip speed ratio and wind speed. There have been works available on ML modelling of HAWT wakes [38–41]. Ti et al. [41], by a simple single-layer ANN mapped the wind speed and turbulence intensity of the incoming wind to the velocity field and turbulence field of HAWT wake. However, unlike the HAWT's wake field which is more sensitive to the inlet wind, the wake of VAWT is expected to be highly dependent on the geometric and kinetic characteristics of the turbine [42]. As for the modelling of the wake field of H-rotor VAWTs using ML method, no relevant research has been found.

Wind farms should be properly arranged in layout to minimize the impact of wake on downstream wind turbines, which requires a good understanding of the effects of different characteristics of wind turbines on wake. The discovery of some simple laws and key parameters can greatly simplify the optimization process, which can be achieved through ML model interpretation [43]. By explaining the relationship between model variables and outputs, it is possible to explore the scientific principles contained in the “black box” model and make it more applicable, going beyond the limits of the training data set. However, the interpretation of ML models is a very complex and understudied area, and trying to recognize the interaction and importance of input features is a good step towards decoding the global interpretation of the model [44]. This study makes an attempt to identify potential feature interactions to sort out features that may be important in the model's decision strategy.

To summarize, previous methods for predicting the wake of VAWTs had their inherent drawbacks. Analytical models were timesaving, but not completely accurate, while numerical simulations were computationally too expensive to be applied in engineering. The ML method used in this article can fill this gap with the support of a certain amount of accurate data, ensuring both accuracy and low computational cost. In this paper, the RF algorithm is used to predict the isolated H-rotor VAWT wake field. The method was chosen because of its strong regression ability and also because of its capability to assess the importance of input features [45], which provides a way to figure out the ways the turbine features can affect the wake [46]. Combined with the feature crosses method, the decisive parameter of wake flow was successfully found and its physical significance was analyzed through a case study. In addition, the case study also proved that the parameter is generalizable, so it could be used in wind farm optimization and in the establishment of a new analytical wake model.

The research process of this paper is shown in Fig. 1, and the paper is organized as follows: Section 2 describes the numerical examples of VAWTs wake, which are performed to provide data for subsequent RF model training. In Section 3, the method of establishing the wake prediction model of H-rotor VAWTs based on RF is described. Besides, combining feature crosses and feature importance analysis, a simple framework for exploring feature interaction is proposed. In section 4, a RF wake model is constructed, and the validity of the model is proved by comparing with the CFD results. In section 5, the key feature is obtained, and its physical meaning is explained and its influence on wake field is analyzed using statistical data. Then, case studies of wind turbines under different geometries and operating conditions proved the universality of this parameter. Section 6 concludes the paper.

2. Data source

2.1. Data origin

The dataset required for RF model training is obtained from the resources provided by Tingey and Ning [27]. They studied the details of the numerical calculations thoroughly and made the computational models open source for future research. The two-dimensional Reynolds-average Navier-Stokes equations are used

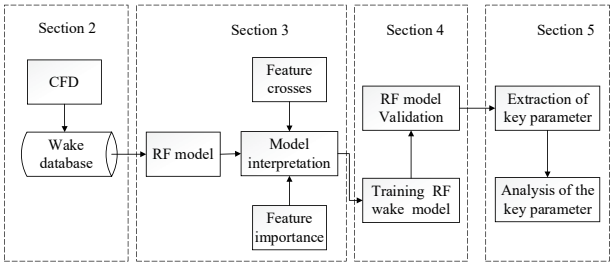


Fig. 1. Flow chart of the research on the H-rotor VAWTs mean wake.

to simulate the flow field of the H-rotor VAWT, since the energy exchange process in the wake of H-rotor VAWT basically occurs in the horizontal plane perpendicular to the axis of rotation [48]. This was also based on a review of the relevant literature [24,25,49,50], which shows that well established 2D models are sufficient to reveal majority of the flow physics surrounding VAWT, such as dynamic stall and wake effects, especially for turbines with high blade aspect ratios [51]. Trial calculations were performed using a full three-dimensional model, but due to its immense calculation time requirements and the huge number of examples required, it is impractical for present research.

The specific geometry and boundary conditions of the calculation model are shown in Table 1. The numerical simulation model in this paper is based on a 12 kW H-rotor vertical axis wind turbine designed by Uppsala University [47], and considering its rated wind speed of 12 m/s, an inlet wind speed range of 12–16 m/s was chosen for the experimental design. The selected inlet conditions, including the inlet wind speed as well as the turbulence intensity, are within the operating range of the referenced VAWT and have been studied in the relevant literature [50]. All the model settings, meshing and experimental validation are described in detail in the reference [52]. To ensure the reliability of the CFD calculation results, verification and validation of the CFD model need to be conducted before the numerical simulation.

2.2. Validation study

For the verification, a study on the convergence of the grid was conducted in the paper [52]. By reducing the basic unit of the CFD model by 1.4 times each time, 6 models with different grid sizes were calculated, and the number of units contained increased from about 400,000 units to 5 million units. By comparing the variations of turbine power generation efficiency with the number of grids, a suitable grid size is selected. To further quantitatively prove the independence of the grid, the general Richardson extrapolation [53] and the calculation of the grid convergence index (GCI) [54] are conducted in this paper, and the results are shown in Table 2. The convergence ratio (R) calculated by general Richardson's extrapolation method is 0.75, which is between 0 and 1, indicating that the convergence of the grid is monotonic. The GCI calculation shows that $GCI_{3,2}$ (1.93%) is larger than $GCI_{1,2}$ (0.19%), illustrating that the solution is convergent. The value of the asymptotic range α is close to 1, indicating that the mesh has reached the range of asymptotic convergence.

Then, two sets of independent validations were performed where the CFD results were compared with the experimental data of Castelli et al. [55] and Tescione et al. [15] on power coefficients *Cp* and wake velocity distributions, respectively.

For power coefficient, since there is no relevant validation in Ref. [52], we performed additional calculations and compared the current results for *Cp* with experimental data [55] and three-dimensional CFD validation data obtained by Su et al. [56], as shown in Fig. 2. It can be seen that the 2D numerical results have some error compared with the 3D model, but they still reflect the trend of the power coefficient with tip speed ratio very well. The 2D numerical results generally overestimate *Cp* of VAWTs, which is consistent with the description in Ref. [25], probably because the blade length and tower height are infinite in the 2D model, resulting in an overestimation of the blocking ratio and thus making the true inlet velocity larger than the defined value. In general, the numerical simulation applied in this paper can be used as a reliable simulation method in subsequent studies.

For wake velocity distributions, the numerical model is validated by comparing the wake velocity calculated by the CFD model with the experimental results. Based on the above meshing rules and model settings, the validation example in paper [52] was recalculated and compared with the results of the wind tunnel test [15]. As shown in Fig. 3, the speed values at 1D (diameter of rotor) and 1.75D behind the wind turbine are selected for comparison. There is a certain difference between the calculated minimum speed value and the experimental value by CFD, and the asymmetry of the speed distribution is not well predicted. These errors can be caused by using a 2D model for simplification in calculations [25]. However, in general the velocity values obtained from CFD calculations agree well with the experimental values, giving enough confidence in the results.

Table 1
Geometry and boundary condition settings of numerical examples.

Setting	Value	Boundary	Setting
Blade section	NACA 0021	Inlet	Velocity Inlet
Blade chord	[0.15–1.0] m	Outlet	Pressure Outlet
Diameter of turbine	6 m	Side	Velocity Inlet
Diameter of strut	0.125 m	Interface	Internal interface
Inlet wind speed	[12–16,0] m/s	Blade	No slip wall
Turbulence Intensity	15%	Strut	No slip wall

Table 2
Grid convergence calculation using General Richardson extrapolation and GCI method.

	$N_i(million)$	*Cp*	$GCI_{i+1,i}(\%)$	*R*	α
fine	1.42	0.438	0.19	0.75	0.99
medium	0.63	0.432	1.93	–	–
coarse	0.48	0.424	–	–	–

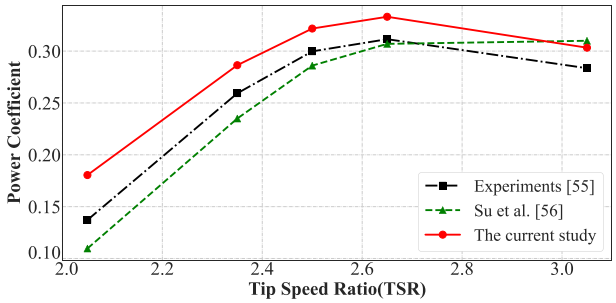


Fig. 2. Comparison of the current results with experimental and numerical data [55,56].

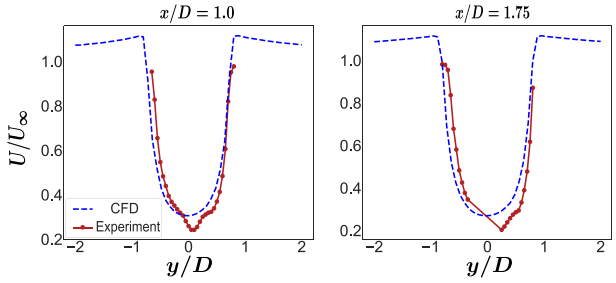


Fig. 3. The comparison between the wake velocity calculated by CFD and the experimental results [15].

2.3. Parameters' design for CFD cases

To ensure the validity of the flow field analysis, it is necessary to carefully select the input characteristics and their range of variation. Three dimensionless parameters were selected as the variables of the experimental design, namely tip speed ratio (TSR), solidity (σ) and Reynolds number (Re). They respectively represent geometrical characteristics, motion settings and the incident velocity of the VAWT system. They are defined by:

$$TSR = \frac{\omega R}{U_\infty} \tag{1}$$

$$\sigma = \frac{Nc}{R} \tag{2}$$

$$Re = \frac{\rho U_\infty D}{\mu} \tag{3}$$

where ω is the rotation speed, R is the rotor radius, U_∞ is the inlet wind speed, N is the number of blades, c is the blade chord length, ρ is the air density and μ is viscosity coefficient. They all have a significant impact on the power of the wind turbine, and therefore significantly affect the wake.

Combined with the actual operation of the wind turbine, the CFD simulations are performed at 23 TSRs (from 1.5 to 7.0), 5 solidities (from 0.15 to 1.0), and 4 Reynolds numbers (from 4.8 to 6.4 million), for a total of 460 simulations (shown in Fig. 4). In the simulation, the values of three dimensionless parameters are controlled by determining the chord length of the blade (c), the rotation speed (ω) and the wind speed of the inlet wind (U_∞). Although it is possible that the power is less than 0 under some of the parameters, which is impractical in actual engineering, these parameters are deliberately selected to make the research scope wider and more comprehensive, and easier to find the law of wake.

2.4. Dataset preparation

Before model training, the input and output variables should be determined first, and then the data should be extracted and transformed into a form that the model can understand. For the layout of the wind field, more attention is paid to the steady characteristics of the wake, so the wind speed mentioned later in the paper is all the time average of the streamwise component of velocity, denoted as U. To further improve the adaptability of the model, a more generalized variable is adopted in this paper, i.e. normalized velocity U/U_∞. As for the input features, when

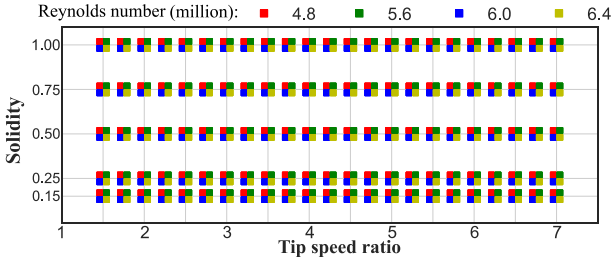


Fig. 4. The design of three parameters TSR, solidity and Re in the numerical experiment.

establishing the RF mean wake model and verifying its accuracy, three dimensionless parameters TSR, σ and Re were used. Since they not only reflect the main influencing factors, but also enhance the applicability of the model due to their dimensionless characteristics. Therefore, the RF-based H-rotor VAWT wake prediction model can be expressed as:

$$\frac{U}{U_\infty} = f(Re,\ TSR,\ \sigma) \tag{4}$$

The streamwise velocity of the wake in the inner domain is extracted from the calculation results of the numerical experiment as the output of the model training. When determining the scope of the inner domain, it should be ensured that its width is sufficient to cover the entire wake structure, and the wake speed has been recovered to a large extent within the length of the domain. If the internal domain is too large, it will contain too much wind field data from the surrounding environment that is not affected by the wind machine, which will not only increase the difficulty of training the ML model, but also affect its accuracy. Combined with the size of the numerical simulation model, the final size of the inner domain is 5D $\times$ 35D (30m $\times$210m) after several attempts. In some simulation results, the wake has not been fully developed, which is mainly caused by the diversity of input rotor parameters. But within the range of 35D, the wake has recovered to a considerable extent, and will not affect the research results, because our research is not based on a fully recovered wake.

A velocity value is extracted every 0.5 m in both the lateral and flow directions, that is, a matrix $\mathbf{V}$ of 61 $\times$ 421 is generated for each sample. Since the RF algorithm requires the output variable to be a one-dimensional matrix, the original data matrix $\mathbf{V}$ is reshaped from 61 $\times$ 421 to a 1 $\times$ 25681 data array. So, each sample $\mathbf{D}$ in the database can be represented as:

$$\mathbf{D} = [Re,\ TSR,\ \sigma;\ \boldsymbol{V}] \tag{5}$$

Although RF can use the out-of-bag error to measure its performance, to better demonstrate its generalization performance, a small number of samples (23 samples) are randomly selected as the test set.

3. Methodology

3.1. Description of RF wake model

RF is a kind of ensembled algorithm proposed by Breiman [45], which can be used for regression and classification. Since the purpose of this paper is to reconstruct the flow field, attention is focused on the RF regression.

RF regression is composed of regression trees, which include root nodes, internal nodes and leaf nodes. The regression tree algorithm mainly includes two parts: (1) Tree generation. The root node contains all samples of the original data set D. Starting from the root node, each value of each feature is traversed. The value is used to split the original data set D into two sets: left set (sample of $\leq$ value) and right (sample of >value), each set is called an internal node. Calculate the difference between the predicted value and the true value in the two sets (generally using the mean square error), and find the value that minimizes the sum of the error values of both the left and right sets. In this case, the feature and value are the best segmentation features and the best segmentation values. Repeat the above steps for the left and right subsets until the

termination condition is reached, and the last set that does not split is called the leaf nodes. (2) Pruning. Some subtrees are cut off from the bottom of the generated regression tree to make the regression tree smaller, to improve the generalization of the model and make a more accurate prediction of the unknown data.

RF aggregates a group of unrelated regression trees, and the final output of the model is determined by each tree in the forest. The regression tree algorithm is relatively simple, but its prediction accuracy is low and unstable, and the small disturbance of the sample will cause a large change in the structure of the tree. These drawbacks can be improved by ensemble methods, so RF Algorithm came into being. These random trees are similar to the regression trees introduced above, but there are two key differences, leading to improve algorithm's performance while reducing the amount of calculation, as sketched in Fig. 5. First, samples and features were randomly sampled before each tree training, and second, the pruning process was eliminated.

RF can conveniently give the importance of the input features, which is of great help to the understanding of VAWT wake. For each regression tree, the corresponding out-of-bag (OOB, the samples which did not appear in the bootstrap sample used to establish the decision tree) data is selected to calculate the error. Then noise interference is randomly added to a feature X of all OOB samples, and the out-of-bag data error is calculated again. If the accuracy of the OOB data drops significantly, it indicates that this feature has a great impact on the prediction results of the sample, that is, the importance of this feature is relatively high.

3.2. Feature crosses

Feature crosses refers to learning the cross combination of two or more original features, which is helpful to express the nonlinear relationship between output and input, and the combination form includes addition, subtraction, multiplication, division, power, logarithm, etc. Therefore, the use of feature crosses when using extended linear models is an effective method for training large-scale data sets. In this paper, feature crosses is selected, combined with the importance analysis of RF, to show the nonlinear relationship of the features learned by RF.

The form of feature crosses is diverse and cannot be exhausted in the research process. Inspired by the dimensionless numbers (TSR, σ, Re etc.) in wind turbine research, the expression (6) is selected to construct the features.

$$\mathrm{F} = c^{\alpha}\omega^{\beta}U_{\infty}{}^{\gamma} \tag{6}$$

where F is the synthetic feature, α, β, and γ are all integers, and c, ω and U_{∞} are the original variables in the numerical experiment. These three variables all have different dimensions, so operations such as addition and subtraction are obviously not applicable, and formula (6) is the best way to construct new features through them. Given the diameter of the wind turbine, various parameters with physical meaning can be constructed through the formula (6), including Re $(\gamma = 1)$, TSR $(\beta = 1,\ \gamma = -1)$, σ $(\alpha = 1)$. And this is also the reason why the basic features are selected instead of three dimensionless numbers when constructing features.

To reduce the amount of calculation, further assumptions are made on the parameters in the formula, that is α, β, and γ are all integers. From the perspective of both dimensions and physical meaning, this assumption is reasonable. Let S be the sum of the absolute values of the exponents of the three fundamental quantities, i.e.,

$$\mathrm{S} = |\alpha| + |\beta| + |\gamma| \tag{7}$$

S represents the complexity of the synthetic features. And the dimension of the new feature is:

$$[L]^{\alpha} \cdot [T]^{-\beta} \cdot [L]^{\gamma}[T]^{-\gamma} = [L]^{\alpha+\gamma}[T]^{-\beta-\gamma} \tag{8}$$

where L and T are the dimensions of length and time respectively.

In this paper, combining feature crosses with feature importance analysis of RF, an analysis process is proposed as shown in the Fig. 6.

As can be seen from the Fig. 6, S is used as the control index to carry out feature crosses. Given S α, β and γ that satisfy $|\alpha| + |\beta| + |\gamma| = \mathrm{S}$ are assigned by traversal, and then new synthetic features are created by using equation (6). It should be noted that if $(\alpha_i, \beta_i, \gamma_i) = -(\alpha_j, \beta_j, \gamma_j)$, only one of them will be retained. The

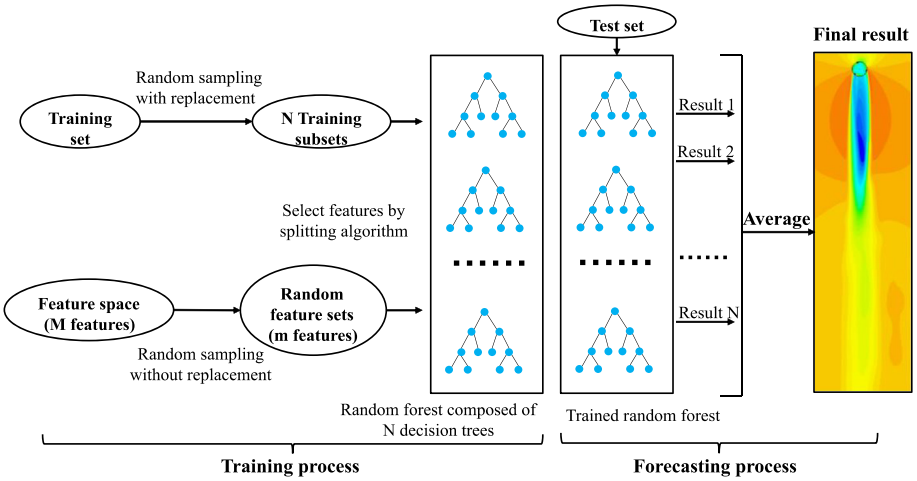


Fig. 5. Schematic diagram of general program for generating average wake prediction model of RF.

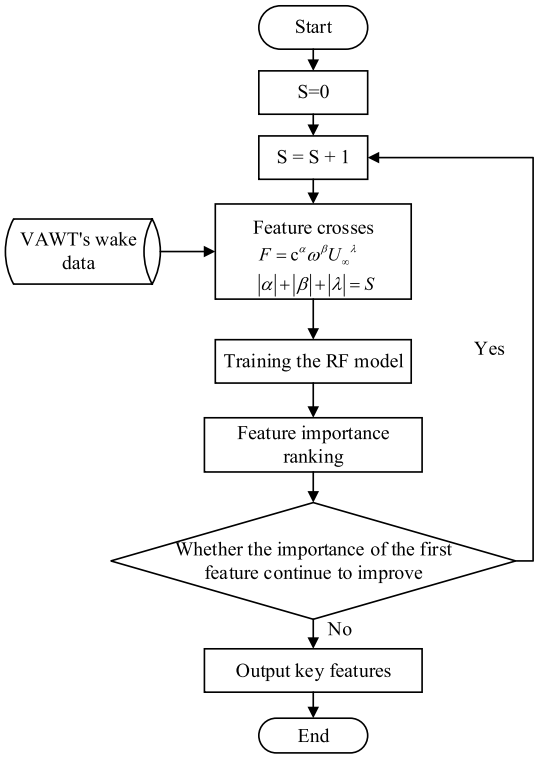


Fig. 6. Analysis flowchart of feature interaction and feature importance.

generated features are used to train the RF model, and the importance of these features is given and ranked in the process of realizing wake field regression Importance here is a measure of how each input feature contributes to the prediction results of the model, i.e., how a small change on a particular feature changes the predicted wake velocity value. At this point, this round of analysis is completed and then the value of S is then increased to generate more complex composite features, starting the next round of analysis. If the importance of that first-ranked feature decreases as S increases, it is considered that the key feature has been found and the entire analysis process stops. That is, the importance peaks at the key parameters found. More complex synthesis parameters generated afterwards have instead a reduced influence on the results due to the presence of redundant information.

During each round of analysis, features of the same complexity are compared together to avoid low-dimensional features being masked by higher-dimensional features. The comparison of the most important features at different complexity levels also facilitates the understanding of their physical meaning. The ultimate goal is to identify potential feature interactions to understand which features are important in the prediction process of the model, and to discover the general laws of H-rotor VAWT wakes while interpreting the model.

4. Training and validation of the RF mean wake model

The RF mean wake model is first trained. This is not only a new H-rotor VAWT wake modeling method using ML, but also a validation for the subsequent extraction of key parameters, since the model interpretation is based on the accuracy of the model.

4.1. Training and testing of RF model

In this paper, the number of samples and features of the database in this article are not so big, the performance of RF model is mainly controlled by two hyperparameters, the maximum depth and the number of trees. The cross-validation method is used to perform grid search on these two parameters. Fig. 7 presents the variations of 10-fold mean squared error (MSE) of training set and validation set with the number of trees and the maximum depth of

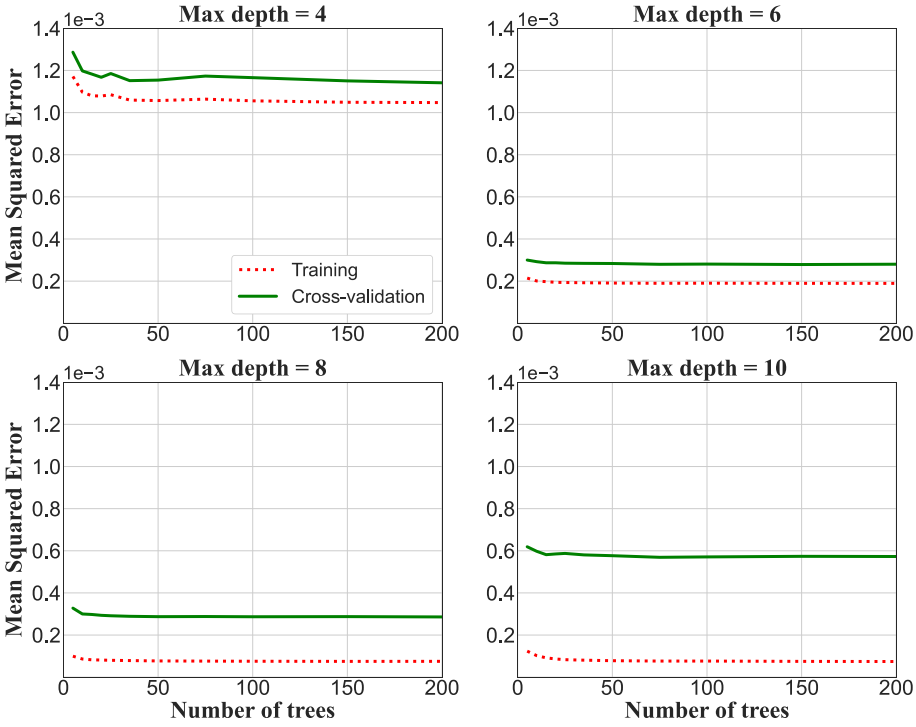


Fig. 7. Variations of 10-fold mean squared error of training set and validation set with hyperparameters of RF.

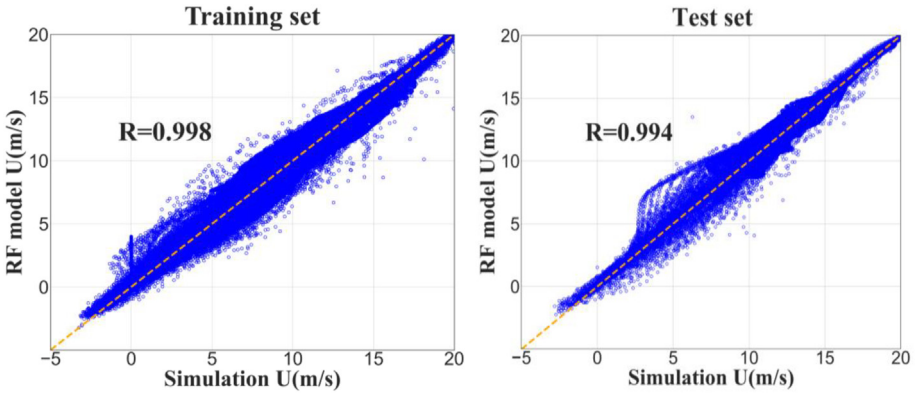


Fig. 8. Scatter plot of velocity U regressed by RF model against numerical experiment data in training set and test set.

trees for the dataset of wake velocity distribution. MSE is defined as

$$\mathrm{MSE} = \frac{1}{M}\sum_{t=1}^{M}(U_{RANS} - U_{RF})^2 \tag{9}$$

where M is the number of speed samples in each case, U_{RANS} is the velocity calculated by CFD and U_{RF} is the velocity predicted by RF model.

When the maximum depth of the tree changes from 4 to 6, both the training error and the validation error decrease significantly. But when the value changes to 8, the training error becomes smaller and the validation error remains basically unchanged. When it is further increased to 10, the error of the validation set does not decrease but increase, and obvious overfitting occurred. In addition, MSE remains stable overall when the number of trees exceeds 100. Therefore, for the data set in this article, the number of trees in the RF and the maximum depth of the trees are set to 100 and 6 respectively.

To evaluate the performance of the trained RF model, the prediction results of the RF were compared with the data measured from the original numerical experiment. As mentioned above, 23 samples were randomly selected as test set to demonstrate the prediction performance of RF model. Fig. 8 shows a scatter plot of the velocity U of the training set and the test set predicted by the RF model against the original numerical simulation data. It can be seen that the correlation coefficient R of the training set and the test set are both high, which shows that the training convergence and generalization performance of the model are good, and the RF model can well predict the wake of H-rotor VAWT.

4.2. Prediction of wake flow of H-rotor VAWTs using RF

To examine the VAWT wake prediction of RF model, contours of the wake velocity distribution of three cases selected in the test set was studied, i.e. Re $= 4.8\times10^6$, $\sigma = 1.0$, $TSR = 3.25$; Re $= 6.0\times10^6$, $\sigma = 0.25$, $TSR = 4.25$, $Re = 4.8\times10^6$, $\sigma = 0.15$, $TSR = 1.75$. These three examples are representative because they are essentially evenly distributed throughout the sample space. The relative error between RF prediction and RANS simulation results is defined as:

$$\mathrm{Error} = \frac{U_{RF} - U_{RANS}}{U_{RANS}} \times 100\% \tag{10}$$

The results of RF model and numerical simulation are shown and compared in Fig. 9. The RF model can well predict the spatial distribution of the mean velocity of H-rotor VAWT wake under different inflow conditions. Most of the errors are concentrated in the wake recovery area, where the velocity changes rapidly, making the velocity distribution more complicated. The results show that the RF model trained in this paper can accurately predict the average velocity distribution of a single H-rotor VAWT wake.

5. Acquisition and analysis of the key parameter $U_\infty/N\omega c$

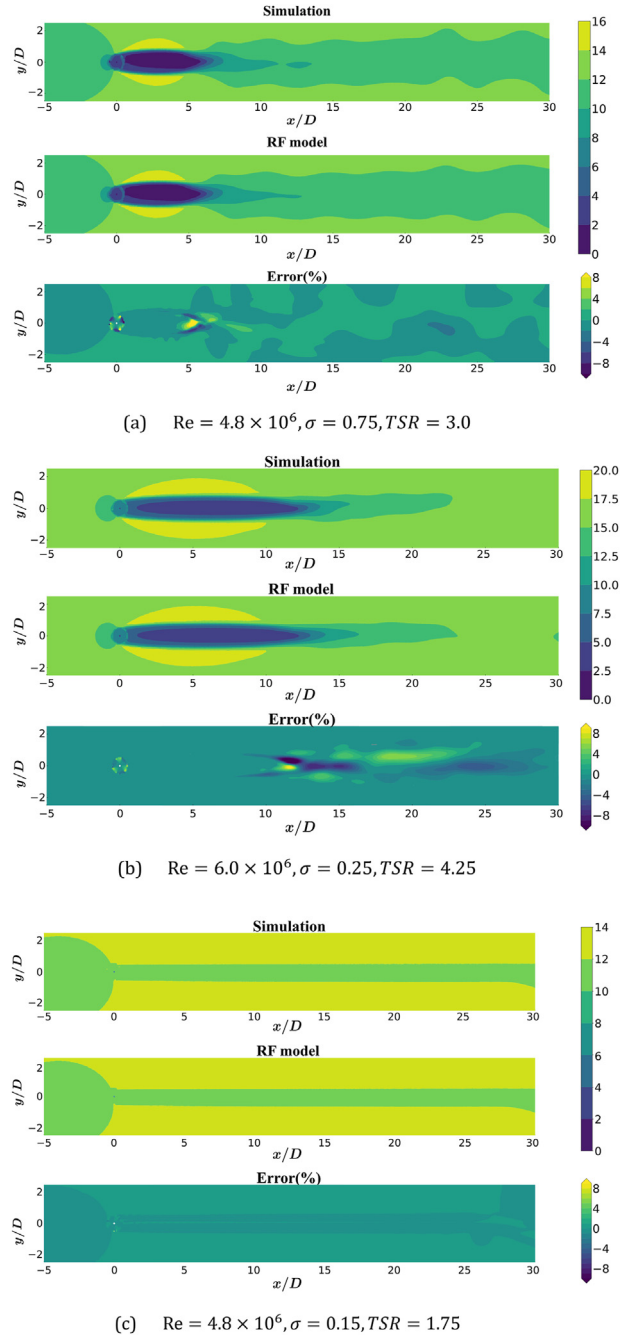
5.1. The extraction of the key parameter

This subsection demonstrated that a key parameter $U_\infty/N\omega c$ can be extracted by the following analysis.

To explain the model and understand which features are important in the prediction process, the key parameter is analyzed using the framework shown in Fig. 6. Fig. 10 shows the ranking of importance of features when S = 1,2,3,4. And when S = 2,3,4, only the top six features are selected for display due to the excessive number of synthetic features.

When S = 1, that is, when all input features are original features, the importance of ω and c is the highest, reaching 0.562 and 0.385, respectively. However, U_∞ plays a small role in the RF model. The main reason is that during the regression process, U_∞ was used to normalize the average velocity distribution of the wake, which means that more attention was paid to the velocity distribution mode of the wake region rather than the specific value. And the range of U_∞ is so small that the laws of motion of the fluid remain basically the same within this order of magnitude, which also results in its low importance. This is good news for the study of the wake flow of VAWTs, because in practical engineering, the operating speed range of H-rotor VAWTs is not particularly wide. Even for strong typhoons, the wind speed is usually around 50 m/s. According to the research, the change of the velocity distribution mode is not particularly obvious, which can greatly simplify the difficulty of the work in the wind farm layout.

When S = 2, the importance of the synthetic features is obviously differentiated, and the features obtained by simple multiplication of ω and c account for 0.852 importance, while other features have little influence on the regression results. When S = 3, a feature $U_\infty/\omega c$ with more importance appears, reaching 0.884. It's obvious that this feature is a perfect combination of the three original features, simple and effective. With the further increase of the value of S, the complexity of the synthetic features increases continuously, but no features of higher importance appear. To better find out the



(a) $\mathrm{Re} = 4.8 \times 10^6, \sigma = 0.75, TSR = 3.0$

(b) $\mathrm{Re} = 6.0 \times 10^6, \sigma = 0.25, TSR = 4.25$

(c) $\mathrm{Re} = 4.8 \times 10^6, \sigma = 0.15, TSR = 1.75$

Fig. 9. Comparison of average velocity distribution between RF model prediction and RANS simulation under different feature sets.

most important features in wake prediction, Table 3 shows the number of composite features, the accuracy of the model, the most important feature and its importance ratio when S is from 1 to 7.

As can be seen from the MSE of the model, the model is accurate enough that the analysis of feature importance can be trusted. And obviously, the most important feature is $U_\infty/\omega c$, which well shows the interaction between the three original variables and plays a crucial role in the wake prediction of H-rotor VAWTs.

5.2. RF mean wake model prediction with $U_\infty/\omega c$ as the only input feature

To prove the important role of the key parameters found in the flow field, the RF model with it as the only input feature was trained, that is, to establish the regression relationship shown in Equation (11).

$$\frac{U}{U_\infty} = f\left(\frac{U_\infty}{\omega c}\right) \tag{11}$$

As shown in Fig. 11, two cases from the test set shown in Fig. 9 were again selected to compare the results calculated by the RF model with key parameter as input feature and CFD. It can be seen that the RF model can still reconstruct the wake field well. Compared with Fig. 9, it can be found that the accuracy of this RF model is slightly lower than that of the original RF model, but it makes sense to sacrifice some accuracy for interpretability. The key parameter $U_\infty/\omega c$ can indeed be a good summary of the influence of the three parameters *Re*, *TSR* and σ on the wake field, which has a decisive effect on the wake field.

Next, this key parameter was further analyzed from the physical aspect.

5.3. The physical meaning of $U_\infty/\omega c$

Through Equation (8), the dimension of this parameter can be obtained as:

$$[L]^{-1} \cdot [T]^1 \cdot [L]^1 [T]^{-1} = [L]^{-1+1} [T]^{-1+1} = 1 \tag{12}$$

So, this is a new dimensionless parameter. It can be found that:

$$TSR \times \sigma = \frac{\omega R}{U_\infty} \times \frac{Nc}{R} = \frac{\omega Nc}{U_\infty} = N \times \left(\frac{U_\infty}{\omega c}\right)^{-1} \tag{13}$$

This new parameter is partially modified by adding the number of blades N into the expression, i.e. $\omega Nc/U_\infty$. This new parameter can be explored qualitatively. To a greater extent, $\sigma = Nc/R$ represents the static solidity, that is, the ability of the turbine to block the wind in a static state. But in the operation process of the wind turbine, if the rotation speed ω is very large, even if σ is small, the wind turbine will still have a significant blocking effect on the incident wind, making it difficult for the incident wind to pass through the wind turbine rotor. At this time, the turbine rotor behaves like a solid cylinder in the flow. Therefore, ωNc combines the motion characteristics and geometric characteristics of wind turbines and is a concept of "dynamic solidity", which coincides with the viewpoint in Ref. [21]. Correspondingly, the inverse of ωNc measures to some extent the ability of the incident wind to pass through the rotor of the wind turbine, which is also obviously related to the velocity of the incident wind. The faster the incoming wind travels, the easier it will pass through the rotor. Therefore, by multiplying $1/\omega Nc$ and U_∞, this new parameter is obtained, which measures the ability of the incident wind to pass through the moving wind turbine rotor.

5.4. Effect of $U_\infty/N\omega c$ on wake flow field behind wind turbines

As shown in the black line in Fig. 12, the velocity distribution directly behind the wind turbine was subsequently extracted and normalized it by the inlet wind velocity. Fig. 12 is a velocity

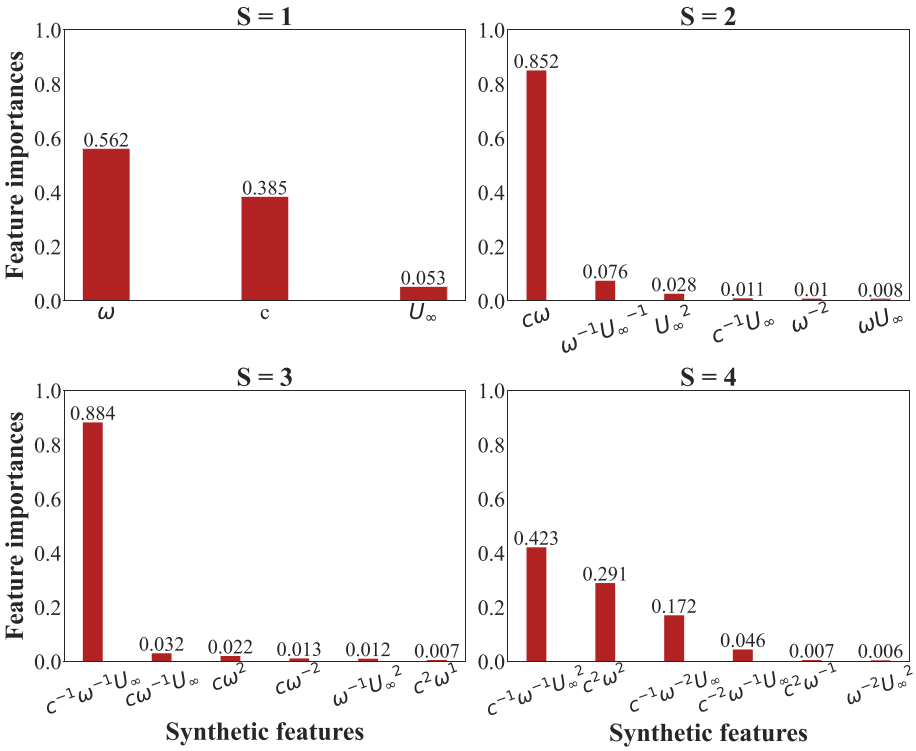


Fig. 10. Ranking of the importance of synthetic features of different complexity in RF algorithm.

Table 3
Running results of the framework when.S ≤ 7

S	Number of synthetic features	MSE	Most important feature	Importance ratio
1	3	0.00064	ω	0.562
2	9	0.00056	$c\omega$	0.852
3	19	0.00044	$U_\infty\omega^{-1}c^{-1}$	0.884
4	33	0.00047	$U_\infty{}^2\omega^{-1}c^{-1}$	0.423
5	51	0.00050	$U_\infty\omega^{-2}c^{-2}$	0.669
6	73	0.00047	$U_\infty{}^2\omega^{-2}c^{-2}$	0.649
7	99	0.00049	$U_\infty{}^3\omega^{-2}c^{-2}$	0.459

distribution diagram randomly selected from the database, and it can be seen that the velocity distribution of the black line can reflect the recovery of the wake to a large extent.

Fig. 13 plots the speed recovery of wind turbines under 20 different $U_\infty/N\omega c$, which basically cover the scope of $U_\infty/\ N\omega c$ involved in this paper. It can be clearly seen that the velocity distribution mode behind the wind turbine presents obvious regularity with the change of $U_\infty/N\omega c$. When $U_\infty/N\omega c$ is large, the velocity deficit behind the wind turbine is small, and the velocity presents a flat distribution. Along the direction indicated by the arrow, $U_\infty/N\omega c$ gradually decreases, the ability of the wind to pass through the wind machine is weakened, and the wind speed decreases seriously. With the decrease of $U_\infty/N\omega c$, the value of the minimum speed behind the wind turbine also decreases, and the position where it appears gradually moves forward and toward the wind turbine. A surprising result is that the deficit of wake velocity recovers more quickly as $U_\infty/N\omega c$ decreases, as shown in Fig. 13 by a steeper upward slope of the velocity curve. This may be related to the way the energy enters the wake. When the wake velocity is low, there is a bigger difference between the velocity of the wake and that of the surrounding fluid, so more energy will be injected into the wake of the wind turbine to make the velocity deficit recover faster.

5.5. Effect of $U_\infty/N\omega c$ on velocity deficit

Next, the relationship between the key parameter and the minimum velocity value in the wake flow field of the wind turbine is investigated. Fig. 14 and Fig. 15 respectively show the scatter plots of the minimum velocity value and its position against $U_\infty/N\omega c$.

Fig. 14 (a) shows the scatter plot between the minimum velocity value and the corresponding new parameter $U_\infty/N\omega c$ for 460 calculation examples. It can be seen that there is a significant correlation between them. Since $U_\infty/N\omega c$ of the samples is mainly distributed between 0 and 1, leading to the distribution of points is unbalanced and the rule is not obvious. Therefore, taking the reciprocal of $U_\infty/N\omega c$ as the abscissa, Fig. 14 (b) can be obtained, where the points are more evenly distributed. Combining the two pictures, there is a clear correlation between the minimum speed value after the wind turbine and $U_\infty/N\omega c$. As $U_\infty/N\omega c$ increases, the

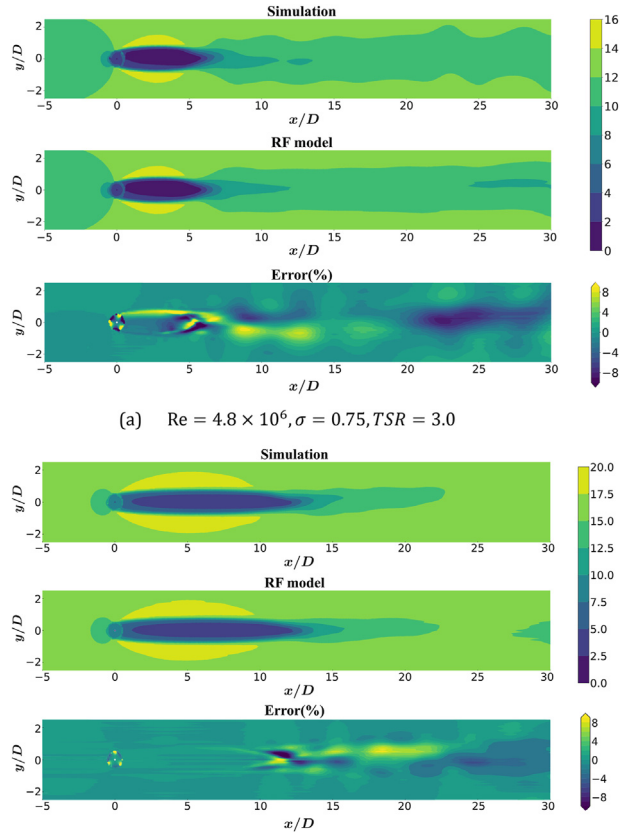


Fig. 11. Comparison of average velocity distribution between RF model prediction and RANS simulation.

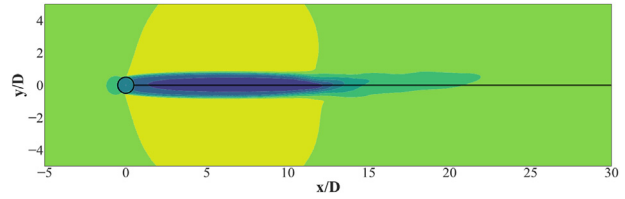


Fig. 12. Schematic diagram of wind turbine wake velocity extraction. The black circle indicates the location of the wind turbine, and the black line corresponds to the location of speed extraction.

minimum speed value increases and gradually approaches the speed of the incident wind. This also confirms the analysis of the physical meaning of $U_\infty/N\omega c$. Similarly, the ordinates of Fig. 15 (a) and (b) are the distance between the position where the minimum speed occurs and the turbine, and the abscissas are $U_\infty/N\omega c$ and $N\omega c/U_\infty$, respectively. Obviously, there is also a correlation between the position of the minimum speed and the key parameters.

It can seen that the relationship between them can be roughly divided into three segments by the blue and green lines in the figure. In Fig. 14, when $U_\infty/N\omega c$ is less than 0.45, using linear fitting, the slope of the line obtained is about 0.90 and R^2 (coefficient of determination) is 0.92. When $U_\infty/N\omega c$ is between 0.45 and 2.66, the slope of the line obtained by linear fitting is about 0.39, and R^2 is 0.91. In Fig. 15, when $U_\infty/N\omega c$ is small, the position where the

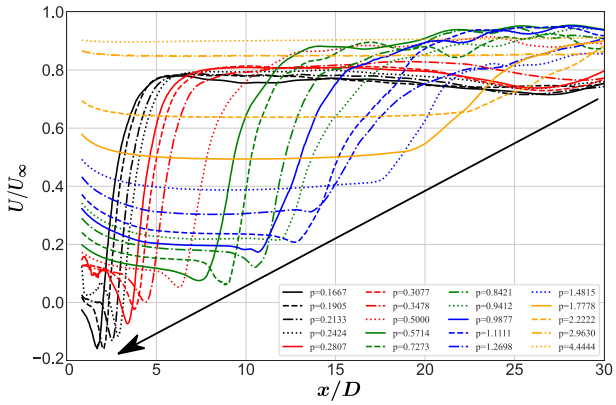


Fig. 13. Recovery of wake velocity directly behind the wind turbine under different $U_\infty/N\omega c$. The p in the legend represents $U_\infty/N\omega c$, and the black arrow represents the direction in which $U_\infty/N\omega c$ decreases.

minimum velocity appears has an approximate linear relationship with $U_\infty/N\omega c$. As $U_\infty/N\omega c$ increases, the line bifurcates, and the location of the bifurcation is just near the blue line. When $U_\infty/N\omega c$ continues to increase and exceeds the position of the green line, the velocity deficit behind the wind turbine is very small due to the excessive wind passing capacity. It can be seen from Fig. 13 that the velocity distribution at this time is very smooth, and the value and position of the minimum velocity may have a large error when extracting, which does not have much reference value. But there is no doubt that with the increase of the passing ability of the wind, the wake velocity gradually converges to the wind speed of the incident wind.

5.6. Effect of $U_\infty/N\omega c$ on velocity recovery

Finally, the impact of this parameter on the recovery speed of the wake velocity deficit was analyzed. The distance from the wind turbine where the wake velocity recovers to different degrees is extracted, and the diameter of the wind turbine is used to normalized it. Fig. 16 respectively shows the relationship between the distance from the wind turbine and $U_\infty/N\omega c$ when the wake velocity recovers to 0.4, 0.5, 0.6 and 0.7 times the incident wind velocity. Obviously, there is a strong correlation between the distance required for recovery and $U_\infty/N\omega c$, especially when the value of $U_\infty/N\omega c$ is small. The quadratic function was used to fit the relationship between them, and the results were shown in Fig. 16. The R^2 of the obtained results were all greater than 0.97. This law is also applicable to other different recovery degrees. It can be seen that $U_\infty/N\omega c$ plays a decisive role in controlling the wake recovery speed.

5.7. Case study on the generality of $U_\infty/N\omega c$

To examine the generality of this key parameter on H-rotor VAWTs' mean wake analysis, the flow fields of wind turbines under different geometry and operating conditions was used to verify it. This validation study was performed based on the experiment of Castelli et al. [55], the geometric characteristics and operating conditions of the wind turbine are shown in Table 4. By adjusting the speed of the wind turbine, $U_\infty/N\omega c$ is equal to different values

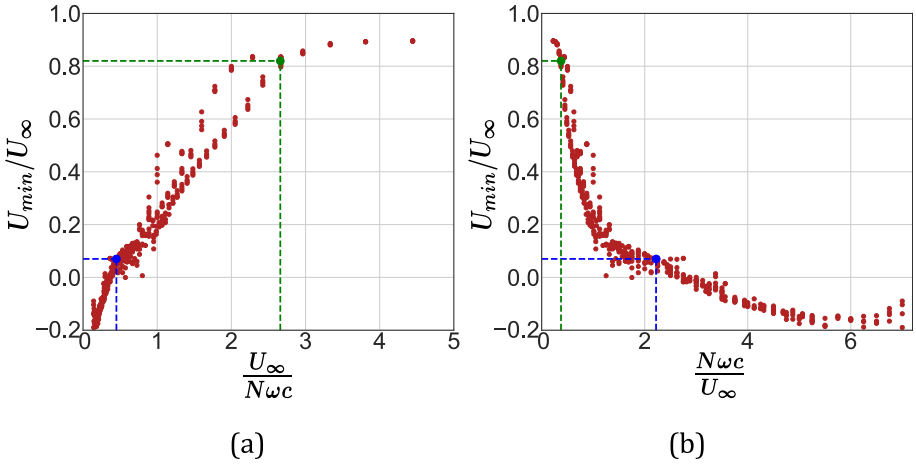


Fig. 14. (a)Scatter plot of the normalized absolute minimum velocity in wake flow field against $U_\infty/N\omega c$. (b)Scatter plot of the normalized absolute minimum velocity directly in wake flow field against $(U_\infty/N\omega c)^{-1}$. In the two figures, the blue dot represents the point where $U_\infty/N\omega c = 0.45$ and $U_{min}/U_\infty = 0.07$, and the green dot represents the point where $U_\infty/N\omega c = 2.66$ and $U_{min}/U_\infty = 0.82$. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

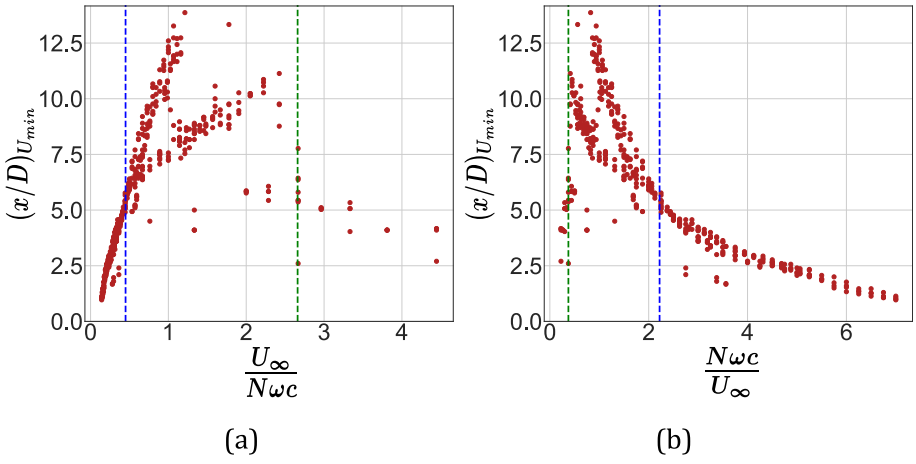


Fig. 15. (a) Scatter plot of the of the position of the minimum velocity in wake flow field against $U_\infty/N\omega c$. (b) Scatter plot of the of the position of the minimum velocity in wake flow field against $(U_\infty/N\omega c)^{-1}$.

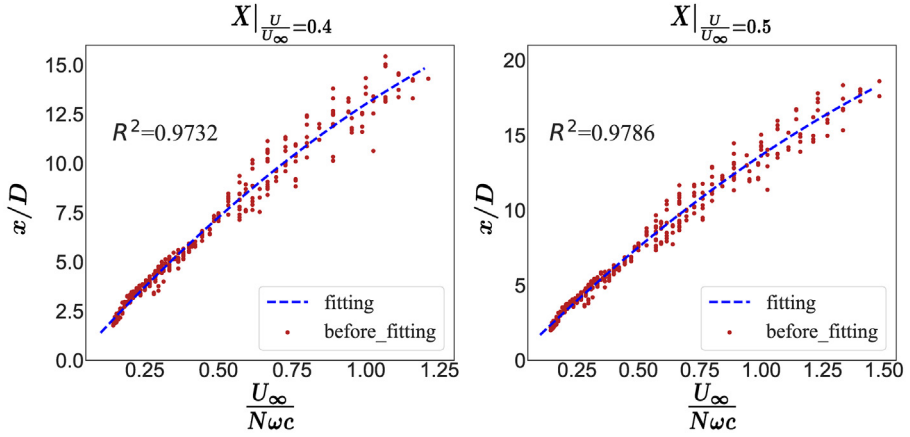
between 0.1 and 1.5 respectively. After the wake field is calculated, the relationship between $U_\infty/N\omega c$ and the wake recovery velocity, which is of most interest, is plotted, as shown in Fig. 17.

There is an obvious relationship between the distance required for wake recovery and $U_\infty/N\omega c$, and the quadratic function fitting also achieves a good effect. In fact, $U_\infty/N\omega c$ is small currently, and the relationship between the two is nearly linear. Therefore, this key parameter is also applicable to wind turbines with different geometric characteristics. In the layout of a wind farm, it is possible to first calculate some examples in combination with the local environment and specific wind turbine conditions, and then infer the wake field under a wider range of operating conditions. And it is also feasible to use this parameter to put forward the wake analysis model of H-rotor VAWT.

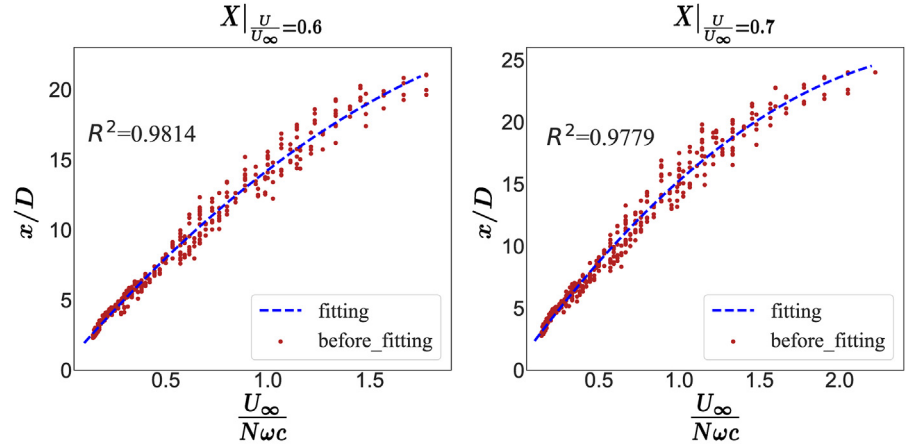
6. Conclusions

The RF model was used to regress from the rotor features to the mean distribution of the streamwise component of the velocity in wake field. On this basis, the importance of rotor features on the wake velocity distribution was analyzed by combining the feature crosses method. The conclusions are summarized below:

1) As the correlation coefficient R between the true and predicted velocity values in the training and test set are 0.998 and 0.994, respectively, the RF model could establish a good internal relationship between the spatial wake field and the turbine features. These features include tip speed ratio, solidity and Reynolds number.



(a) Fitting curve: $\frac{x}{D} = -0.24 + 16.75 \times \frac{U_\infty}{N\omega c} - 3.5 \times \frac{U_\infty}{N\omega c}^2$; (b) Fitting curve: $\frac{x}{D} = 0.11 + 16.16 \times \frac{U_\infty}{N\omega c} - 2.6 \times \frac{U_\infty}{N\omega c}^2$



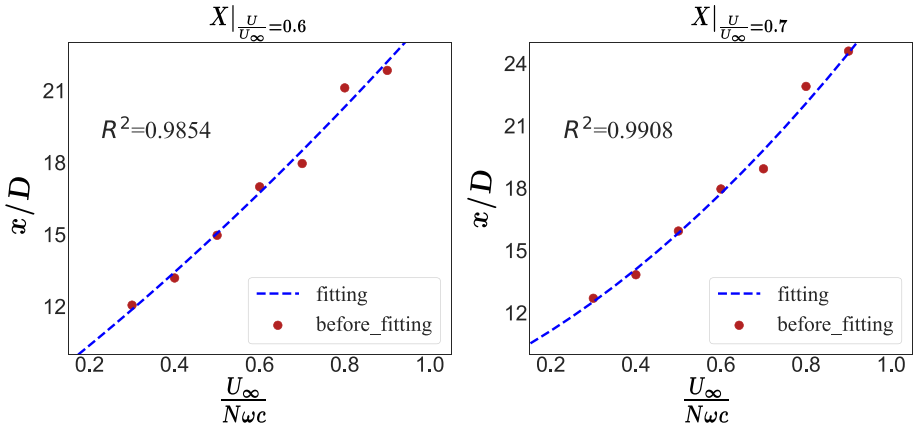
(c) Fitting curve: $\frac{x}{D} = 0.28 + 16.87 \times \frac{U_\infty}{N\omega c} - 2.88 \times \frac{U_\infty}{N\omega c}^2$; (d) $\frac{x}{D} = 0.61 + 17.82 \times \frac{U_\infty}{N\omega c} - 3.16 \times \frac{U_\infty}{N\omega c}^2$

Fig. 16. Scatter plot of $U_\infty/N\omega c$ and the distance from the wind turbine when the wake velocity recovers to 0.4, 0.5, 0.6 and 0.7 times the incident wind velocity.

Table 4
Geometry and operation condition settings of validation examples.

Setting	Value	Setting	Value
Blade section	NACA 0021	Number of blades	3
Blade chord	85.8 mm	Inlet wind speed	15 m/s
Diameter of turbine	1030 mm	Turbulence Intensity	5%

2) The synthetic parameter $U_\infty/N\omega c$ plays a key role in the formation of the wake field, and its importance has reached 0.899 (according to the importance analysis of the RF model). Based on this analysis, this parameter represents the ability of the incident wind to pass through the operating wind turbine.
3) $U_\infty/N\omega c$ plays a decisive role in the velocity distribution model behind the wind turbine, including the velocity deficit and the recovery rate of the deficit. A piecewise linear function was used to fit the relationship between the velocity deficit and $U_\infty/N\omega c$, achieving a good performance with R^2 greater than 0.9. The R^2 obtained when the quadratic function was fitted to the relationship between $U_\infty/N\omega c$ and the distance required for the recovery of the wake velocity to different degrees was greater than 0.97.
4) The universality of this parameter has been proved via a series of case studies of wind turbines with different measures and under different operating conditions. In this case, the relationship between $U_\infty/N\omega c$ and the distance to the turbine when the wake velocity returns to 0.6 and 0.7 times the incident wind



(a) Fitting curve: $\frac{x}{D} = 7.59 + 13.29 \times \frac{U_\infty}{N\omega c} + 3.27 \times \frac{U_\infty}{N\omega c}^2$; (b) Fitting curve: $\frac{x}{D} = 8.86 + 9.66 \times \frac{U_\infty}{N\omega c} + 8.57 \times \frac{U_\infty}{N\omega c}^2$

Fig. 17. Scatter plot of $U_\infty/N\omega c$ and the distance from the wind turbine when the wake velocity recovers to 0.6 and 0.7 times the incident wind velocity.

speed ($X|_{\frac{U}{U_\infty}=0.6,\ 0.7}$) was fitted using a quadratic function with R^2 of 0.9854 and 0.9908, respectively.

Overall, the key parameter found in this paper deepens people's understanding of the wake of H-rotor VAWTs, and provides new ideas for the modeling of the average wake field, which can be applied to the optimal layout of wind farms.

7. Limitations and future work

To our knowledge, this is the first time that a ML approach has been used to reconstruct the wake flow of VAWTs. Results of this paper have proved the great potential of the ML method in the prediction of the wake field. However, in the research process, the problem had to be partially simplified to reflect the main target of investigation, which inevitably makes the current model limited in the following aspects:

1) The variables of model input: Although TSR, σ and Reynolds number are critical parameters to describe an operating H-rotor VAWT, there are some other factors that can have a great impact on H-rotor VAWT wake, e.g., the turbulence intensity.
2) There is a lack of experimental validation of the results here presented. The meaning and potential application of the proposed key parameter on H-rotor VAWT wake has not been validated by wind tunnel or field measurements.
3) The applicability of the current ML model has not been fully comprehended. The model has been trained with the data on hand and it is unclear whether new data can be adapted into it or if it can act as a pre-trained model for other H-rotor VAWTs under different working scenarios.

This paper demonstrates the possibility of applying the RF model to a 2D wake flow model of H-rotor VAWTs, and it is evident that the same approach can be extended to a 3D case, since only the dimensionality of the features is increased, but the nature of the problem remains the same. However, obtaining the dataset needed to train the model can be challenging. The computational cost of 3D numerical simulations is high and the model needs to consider more input parameters at this point, e.g., height-to-diameter aspect ratio. However, there are some methods that can assist the construction of a 3D RF wake model: 1) Using orthogonal experimental design or Gaussian process regression to reduce the number of numerical simulation cases. 2) Using experimental data instead of numerical simulations. This guarantees that accurate data can be obtained quickly. 3) In order to reduce the difficulty of the training process, different models can be trained separately at several heights, and then integrated into the final one.

In addition to attempting to model the 3D wake, the work in this paper leaves room for improvement. The wake model of a single H-rotor VAWT obtained in this paper can be combined with the wake superposition model to quickly calculate the flow field and power efficiency of wind farms, thereby contributing to the optimization of the layout. Besides, the research in this paper is only limited to the H-rotor lift-driven VAWTs, though the similar method can be tried to be applied to the research of other wind turbines. It is also suggested that machine learning models are applied to study the unsteady characteristics of the wake field.

Credit author statement

Zhikun Dong: Conceptualization, Methodology, Software, Data curation. **Yaoran Chen:** Methodology, Revision. **Dai Zhou:** Conceptualization, Supervision. **Jie Su:** Revision. **Zhaolong Han:** Revision, Supervision. **Yong Cao:** Revision, Supervision. **Yan Bao:** Supervision. **Feng Zhao:** Methodology. **Rui Wang:** Methodology. **Yongsheng Zhao:** Supervision. **Yuwang Xu:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The financial supports from Innovation Program of Shanghai Municipal Education Commission(No.2019-01-07-00-02-E00066), the National Natural Science Foundation of China (Nos. 42076210, 52088102, 52122110, U19B2013, 51909159, 51809170), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning(No. TP2017013), "Shuguang

Program" supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission (No.19SG10), Program of Shanghai Academic/Technology Research Leader (No.19XD1402000), Shanghai Science and Technology Program (No.19JC1412800, 19JC1412800), The Oceanic Interdisciplinary Program of Shanghai Jiao Tong University (SL2020PT201) are gratefully acknowledged.

References

[1] REN21. Renewables 2020—global status report. A comprehensive annual overview of the state of renewable energy. Paris: REN21 Secretariat; 2020.
[2] IEA. Global energy review 2020. Paris: International Energy Agency; 2020.
[3] Bazilevs Y, Korobenko A, Deng X, Yan J, Kinzel M, Dabiri JO. Fluid–structure interaction modeling of vertical-axis wind turbines. J Appl Mech 2014;81(8).
[4] Borg M, Shires A, Collu M. Offshore floating vertical axis wind turbines, dynamics modelling state of the art. Part I: aerodynamics. Renew Sustain Energy Rev 2014;39:1214–25.
[5] Dabiri, John O. Potential order-of-magnitude enhancement of wind farm power density via counter-rotating vertical-axis wind turbine arrays. J Renew Sustain Energy 2011;3(4):73.
[6] Iungo GV, Porté-Agel F. Volumetric lidar scanning of wind turbine wakes under convective and neutral atmospheric stability regimes. J Atmos Ocean Technol 2014;31(10):2035–48.
[7] Vermeer LJ, Sørensen JN, Crespo A. Wind turbine wake aerodynamics. Prog Aero Sci 2003;39(6–7):467–510.
[8] Thomsen K, Sørensen P. Fatigue loads for wind turbines operating in wakes. J Wind Eng Ind Aerod 1999;80(1–2):121–36.
[9] Sanderse B, Van der Pijl SP, Koren B. Review of computational fluid dynamics for wind turbine wake aerodynamics. Wind Energy 2011;14(7):799–819.
[10] Zong H, Porté-Agel F. A momentum-conserving wake superposition method for wind farm power prediction. J Fluid Mech 2020;889.
[11] Crespo A, Hernandez J, Frandsen S. Survey of modelling methods for wind turbine wakes and wind farms. Wind Energy: Int J Prog Appl Wind Power Conversion Technol 1999;2(1):1–24.
[12] Porté-Agel F, Bastankhah M, Shamsoddin S. Wind-turbine and wind-farm flows: a review. Boundary-Layer Meteorol 2020;174(1):1–59.
[13] Dabiri JO. Potential order-of-magnitude enhancement of wind farm power density via counter-rotating vertical-axis wind turbine arrays. J Renew Sustain Energy 2011;3(4):043104.
[14] Battisti L, Zanne L, Dell'Anna S, Dossena V, Persico G, Paradiso B. Aerodynamic measurements on a vertical axis wind turbine in a large scale wind tunnel. J Energy Resour Technol 2011;133(3).
[15] Tescione G, Ragni D, He C, Ferreira CS, Van Bussel GJW. Near wake flow analysis of a vertical axis wind turbine by stereoscopic particle image velocimetry. Renew Energy 2014;70:47–61.
[16] Bachant P, Wosnik M. Characterising the near-wake of a cross-flow turbine. J Turbul 2015;16(4):392–410.
[17] Araya DB, Dabiri JO. A comparison of wake measurements in motor-driven and flow-driven turbine experiments. Exp Fluid 2015;56(7):1–15.
[18] Rolin V, Porté-Agel F. Wind-tunnel study of the wake behind a vertical axis wind turbine in a boundary layer flow using stereoscopic particle image velocimetry. J Phys Conference 2015;625:012012.
[19] Ryan KJ, Coletti F, Elkins CJ, Dabiri JO, Eaton JK. Three-dimensional flow field around and downstream of a subscale model rotating vertical axis wind turbine. Exp Fluid 2016;57(3):38.
[20] Rolin VF, Porté-Agel F. Experimental investigation of vertical-axis wind-turbine wakes in boundary layer flow. Renew Energy 2018;118:1–13.
[21] Araya DB, Colonius T, Dabiri JO. Transition to bluff-body dynamics in the wake of vertical-axis wind turbines. J Fluid Mech 2017;813:346.
[22] Whittlesey RW, Liska S, Dabiri JO. Fish schooling as a basis for vertical axis wind turbine farm design. Bioinspiration Biomimetics 2010;5(3):035005.
[23] Lam HF, Peng HY. Development of a wake model for Darrieus-type straight-bladed vertical axis wind turbines and its application to micro-siting problems. Renew Energy 2017;114:830–42.
[24] Danao LA, Edwards J, Eboibi O, Howell R. A numerical investigation into the influence of unsteady wind on the performance and aerodynamics of a vertical axis wind turbine. Appl Energy 2014;116:111–24.
[25] Lam HF, Peng HY. Study of wake characteristics of a vertical axis wind turbine by two-and three-dimensional computational fluid dynamics simulations. Renew Energy 2016;90:386–98.
[26] Ghasemian M, Ashrafi ZN, Sedaghat A. A review on computational fluid dynamic simulation techniques for Darrieus vertical axis wind turbines. Energy Convers Manag 2017;149:87–100.
[27] Tingey EB, Ning A. Development of a parameterized reduced-order vertical-axis wind turbine wake model. Wind Eng 2020;44(5):494–508.
[28] Shamsoddin S, Porté-Agel F. Large eddy simulation of vertical axis wind turbine wakes. Energies 2014;7(2):890–912.
[29] Shamsoddin S, Porté-Agel F. Large-eddy simulation of atmospheric boundary-layer flow through a wind farm sited on topography. Boundary-Layer Meteorol 2017;163(1):1–17.
[30] Shamsoddin S, Porté-Agel F. Effect of aspect ratio on vertical-axis wind turbine wakes. J Fluid Mech 2020:889.
[31] Preen RJ, Bull L. Towards the evolution of vertical-axis wind turbines using supershapes. Evolutionary Intelligence 2014;7(3):155–67.
[32] Sekar V, Zhang M, Shu C, Khoo BC. Inverse design of airfoil using a deep convolutional neural network. AIAA J 2019;57(3):993–1003.
[33] Göçmen T, Van der Laan P, Réthoré PE, Diaz AP, Larsen GC, Ott S. Wind turbine wake models developed at the technical university of Denmark: a review. Renew Sustain Energy Rev 2016;60:752–69.
[34] Debnath BK, Das R. Prediction of performance coefficients of a three-bucket Savonius rotor using artificial neural network. J Renew Sustain Energy 2010;2(4):043107.
[35] Sargolzaei J, Kianifar A. Neuro–fuzzy modeling tools for estimation of torque in Savonius rotor wind turbine. Adv Eng Software 2010;41(4):619–26.
[36] Clifton A, Kilcher L, Lundquist JK, Fleming P. Using machine learning to predict wind turbine power output. Environ Res Lett 2013;8(2):024009.
[37] Biswas A, Sarkar S, Gupta R. Application of artificial neural network for performance evaluation of vertical axis wind turbine rotor. Int J Ambient Energy 2016;37(2):209–18.
[38] Zhang J, Zhao X. A novel dynamic wind farm wake model based on deep learning. Appl Energy 2020;277:115552.
[39] 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.
[40] Ali N, Cal RB. Data-driven modeling of the wake behind a wind turbine array. J Renew Sustain Energy 2020;12(3):033304.
[41] Ti Z, Deng XW, Yang H. Wake modeling of wind turbines using machine learning. Appl Energy 2020;257:114025.
[42] Boudreau M, Dumas G. Comparison of the wake recovery of the axial-flow and cross-flow turbine concepts. J Wind Eng Ind Aerod 2017;165:137–52.
[43] Miller T. Explanation in artificial intelligence: insights from the social sciences. Artif Intell 2019;267:1–38.
[44] Molnar C. Interpretable machine learning: a guide for making black box models explainable. 2018 [Online]. Jun. 6, 2018. Available: https://christophm.github.io/interpretable-ml-book/.
[45] Breiman L. Bagging predictors. Mach Learn 1996;24(2):123–40.
[46] Kuz'min VE, Polishchuk PG, Artemenko AG, Andronati SA. Interpretation of QSAR models based on random forest methods. Molecular informatics 2011;30(6-7):593–603.
[47] Kjellin J, Bülow F, Eriksson S, Deglaire P, Leijon M, Bernhoff H. Power coefficient measurement on a 12 kW straight bladed vertical axis wind turbine. Renew Energy 2011;36(11):3050–3.
[48] Simão Ferreira CJ. The near wake of the VAWT: 2D and 3D views of the VAWT aerodynamics. Delft University of Technology; 2009.
[49] Balduzzi F, Bianchini A, Maleci R, Ferrara G, Ferrari L. Critical issues in the CFD simulation of Darrieus wind turbines. Renew Energy 2016;85:419–35.
[50] Rezaeiha A, Montazeri H, Blocken B. Characterization of aerodynamic performance of vertical axis wind turbines: impact of operational parameters. Energy Convers Manag 2018;169:45–77.
[51] Hand B, Cashman A. Conceptual design of a large-scale floating offshore vertical axis wind turbine. Energy Procedia 2017;142:83–8.
[52] Tingey EB. The development of a vertical-Axis wind turbine wake model for use in wind farm layout optimization with noise level constraints. Brigham Young University; 2017.
[53] Richardson LF. IX. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. Philos Trans R Soc Lond - Ser A Contain Pap a Math or Phys Character 1911;210(459–470):307–57.
[54] Roache PJ. Perspective: a method for uniform reporting of grid refinement studies. 1994.
[55] Raciti Castelli M, Ardizzon G, Battisti L, Benini E, Pavesi G. Modeling strategy and numerical validation for a Darrieus vertical axis micro-wind turbine. ASME Int Mech Eng Congress Exposition 2010;44441:409–18.
[56] Su J, Chen Y, Han Z, Zhou D, Bao Y, Zhao Y. Investigation of V-shaped blade for the performance improvement of vertical axis wind turbines. Appl Energy 2020;260:114326.