

## Surrogate models for twin-VAWT performance based on Kriging and artificial neural networks

Yaoran Chen<sup>a</sup>, Dan Zhang<sup>b,\*</sup>, Xiaowei Li<sup>b</sup>, Yan Peng<sup>a,b</sup>, Xiangyu Zhang<sup>b</sup>, Zhaolong Han<sup>c</sup>, Yong Cao<sup>c</sup>, Zhikun Dong<sup>d</sup>

<sup>a</sup> School of Artificial Intelligence, Shanghai University, Shanghai, 200444, China

<sup>b</sup> School of Mechatronic Engineering and Automation, Collaborative Innovation Center for the Marine Artificial Intelligence, Shanghai, 200444, China

<sup>c</sup> School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

<sup>d</sup> Department of Civil Engineering, The University of Hong Kong, Pokfulam, 999077, Hong Kong

### ARTICLE INFO

Handling Editor: Prof. A.I. Incecik

#### Keywords:

Twin vertical axis wind turbines  
Computational fluid dynamics  
Surrogate models  
Kriging  
Artificial neural network

### ABSTRACT

Though twin vertical axis wind turbines (VAWTs) have great potential in the application in oceanic energy harvest, their aerodynamic characteristics is quite complex, especially in the case of sufficient wake-blade interactions. At design stages, the prediction of their power performance usually relies on high-fidelity unsteady simulations based on computational fluid dynamics, whose time budget is high. In this paper, two surrogate models, i.e., Kriging and artificial neural networks (ANN), were adopted for the performance prediction of a twin-VAWT with a close staggered arrangement. Turbines' pitch angles and their averaged torques at the best tip speed ratio were taken as the input and output, respectively. The numerical study shows that both Kriging and ANN models can provide satisfactory predictions using only 22.45% of CFD observations as training set, and the  $R^2$  values for both upstream and downstream turbine models reach more than 0.99 and 0.98, respectively. Among them, the Kriging-based models appear to be more time-efficient and stable than those based on ANN under the moderate dataset in hand. In addition, the current sampling strategy was tested to be modest and robust through sensitivity analysis.

### 1. Introduction

Offshore wind energy is ample, renewable and clean with an accelerating growth in recent and upcoming decades (Guo et al., 2022). Twin vertical axis wind turbines (VAWT) with shared floating platforms are suitable and promising for wind energy utilization under marine environments (Jiang et al., 2020a). Their benefits include easy installation, high structural stability and larger power generation per area (i.e. higher energy density) (Jiang et al., 2020b). Despite these merits, the complex wake-blade interactions between rotors (Jin et al., 2020) require careful investigations on the aerodynamic performance at the design stages.

So far, the most widespread method for twin-VAWT analysis is the unsteady numerical simulation that relies on computational fluid dynamics (CFD) software. The remarkable efforts from Sahebzadeh et al. (2020) and Zanforlin et al. (Zanforlin and Nishino, 2016) have systematically revealed the potential rule of twin-VAWT performance at relative locations or in relative rotational directions, where wake-blade vortex interference and the blockage effects are concluded as the main

mechanism to change the performance of twin-VAWT. Other works from Chen et al. (2022) and Peng et al. (2020) also investigated more parameters of twin-turbine on its individual or overall performance, such as the solidity, phase lag and pitch angle. Apart from their scientific and engineering values, the aforesaid works were computationally expensive, and the computation load is becoming even higher in the foreseeable future. As an aggressive representative, in their recent work, Sahebzadeh et al. (2022) have tested totally 504 cases, with more than 1 million cells and an azimuthal increment of  $0.1^\circ$  per simulation. This makes people think how to effectively cut down the budget of computational resource while retaining acceptable precision.

One of the sophisticated solutions to save computational time is by using orthogonal tables (i.e. Taguchi tables). In the work of Peng et al. (2020), an L16 orthogonal table was designed to investigate the impact of 5 parameters (each parameter in 4 levels) on the twin-VAWT power output, where the ratio of signal to noise (i.e. S–N ratio) was used to analyze the sensitivity of all included parameters. The Taguchi method is also often used in other close areas that involve intense CFD

\* Corresponding author.

E-mail address: [dan.zhang@shu.edu.cn](mailto:dan.zhang@shu.edu.cn) (D. Zhang).

<https://doi.org/10.1016/j.oceaneng.2023.113947>

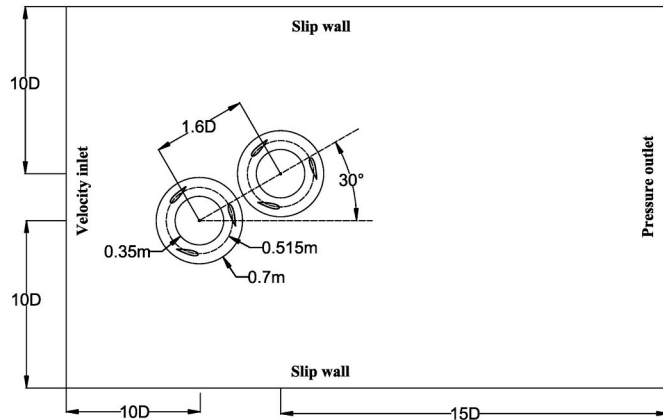
Received 29 October 2022; Received in revised form 14 January 2023; Accepted 11 February 2023

Available online 24 February 2023

0029-8018/© 2023 Elsevier Ltd. All rights reserved.

**Table 1**  
Settings of twin-VAWT model.

Diameter	1030 mm	Chord length	85.8 mm
Solidity	0.25	Profile	NACA0021
Blade span	1456.4 mm	Blade number	3
Tip speed ratio	2.65	Inlet speed	9 m/s
Relative distance	1648 mm	Relative angle	30 °
Phase difference	0 °	Rotational direction	anti-clockwise



**Fig. 1.** Twin-VAWT arrangement and the computational layout (Chen et al., 2022).

computations. In the work of Ma et al. (2022), another L16 table was adopted for the overall shape design of a hydrokinetic turbine, and learnt from the studies of Cheng et al. (2022), an L9 table was used to design and optimize the airfoil of VAWTs. However, such methods cannot quantitatively provide predictions (i.e. the response) for other input combinations, therefore they were often used to diagnose the most influential parameters, or offer more reasonable searching domains for additional optimization approaches (Shrestha and Manogharan, 2017).

In this regard, the surrogate models (SM) are more recommendable. Generally, the SMs are trained by limited data points that were expensive to obtain (such as CFD simulations), which can provide rapid response for other inputs once the model is established. So far, SMs have been successfully employed in various fields, including force coefficient prediction (Sun and Wang, 2019; Chen et al., 2021a) or surface pressure estimation (Zhao et al., 2021) of airfoils, flow field reconstruction for blunt bodies (Brunton et al., 2020; Erichson et al., 2020), and horizontal-axis or Savonius wind turbine power performance prediction (Biswas et al., 2016; Sun et al., 2020). Nevertheless, there were few articles including the applications of different SMs in the performance estimation of Darrieus VAWTs, and that of twin-VAWTs was even less.

Therefore, in this paper, two typical SMs were adopted and compared for twin-VAWT performance prediction. The models were based on Kriging algorithm and the artificial neural network (ANN). Instead of the auto-regression or response surface models, Kriging and ANN are typical among various statistical and inductive SMs, which have been successfully used to solve various multidimensional and nonlinear problems (Chen et al., 2021b; Ismail et al., 2019). The scenario of twin-VAWT is in a close staggered arrangement, where prominent vortex-blade interactions could happen, and the mechanism has been analyzed in authors' former publication by using CFD simulations (Chen et al., 2022). The inputs are twin turbines' pitch angles, and the outputs are the averaged moment of each turbine. More information can be found in Section 2: Data source.

The rest parts of this paper are arranged as follows. In Section 3, the method of Kriging and ANN will be presented; Then, the results and discussion of the SMs will be elaborated in Section 4. Finally, solid

conclusions will be drawn in Section 5.

## 2. Data source

The twin-VAWT data used in this paper was obtained from the unsteady CFD simulations. Detailed description of the CFD model was provided in authors' work (Chen et al., 2022), and a concise summary with additional mesh validation was provided in this section.

### 2.1. CFD model and settings

Due to the lack of experimental or numerical references, though more appropriate for offshore wind energy harvest, multi-megawatt VAWT is not modelled in this work, and a typical small-scaled VAWT from past studies is adopted here. The geometry, layout and computational model of the twin-VAWT are the same as those of the previous publication (Chen et al., 2022). As shown in Table 1 and Fig. 1, two identical straight-blade turbines (in two-dimension) were installed in a close staggered arrangement, and each has a medium solidity of 0.25. The distance between rotor centres is 1.6D ( $D = 1.03$  m, denoting the turbine diameter), and the relative angle is 30 °. This layout may lead to prominent turbulence interferences between turbines. The profile of the blades is NACA0021, with a chord length of 85.8 mm. The variants are blade pitch angles for upstream and downstream turbine, respectively, ranging from  $-6^\circ$  to  $0^\circ$ , whose changes may pose perturbations on local flow field, and therefore influence the aerodynamic performance of VAWTs.

The layout and boundary conditions of the CFD model are as follows. The rotation domains are two concentric circles, where the upstream one is placed at 10D from the velocity inlet (9 m/s) and side boundaries. To ensure the development of the turbine wake, the zero-pressure outlet is set to be 15D away from the downstream turbine center. Finally, the boundary conditions of blade surfaces are non-slip walls, while the side boundaries are slip walls. The turbines rotate at the best tip speed ratio of 2.65.

The Shear Stress Transport (SST)  $k-\omega$  was adopted as the turbulence model (Menter, 2009), where unsteady Reynolds-Averaged Navier-Stokes (URANS) equations were solved through SIMPLE scheme. The second accuracy was taken for spatial and temporal discretization. The time step was set to 1 °, and the simulation terminated at the completion of more than 17 complete rotations. All cases were computed by using STAR CCM+ 13.06 running on supercomputing center in parallel, with more than 12 h per simulation.

### 2.2. Mesh convergence test

In the previous publication, the dependency of the mesh topology on single turbine (Chen et al., 2022) was tested through grid convergence index (GCI) (Roache, 1994). In this paper, to confirm the feasibility of the simulation, additional GCI test on twin-VAWT was performed.

The usage of GCI index for the mesh convergence test is summarized as follows. Assuming three mesh settings: "fine mesh", "medium mesh" and "coarse mesh", their grid numbers are  $N_1$ ,  $N_2$  and  $N_3$ , respectively. The enlarged factor is defined as follows:

$$r_{2,1} = (N_1/N_2)^{1/D_0} \quad (1)$$

$$r_{3,2} = (N_2/N_3)^{1/D_0} \quad (2)$$

where  $D_0$  refers to the dimension, which equals to 2 for current simulation. The asymptotic order  $p$  is then calculated through iteration process, of which the initialization is defined as:

$$p_0 = \left| \ln \left( \frac{P_3 - P_2}{P_2 - P_1} \right) \right| / \ln(r_{2,1}) \quad (3)$$

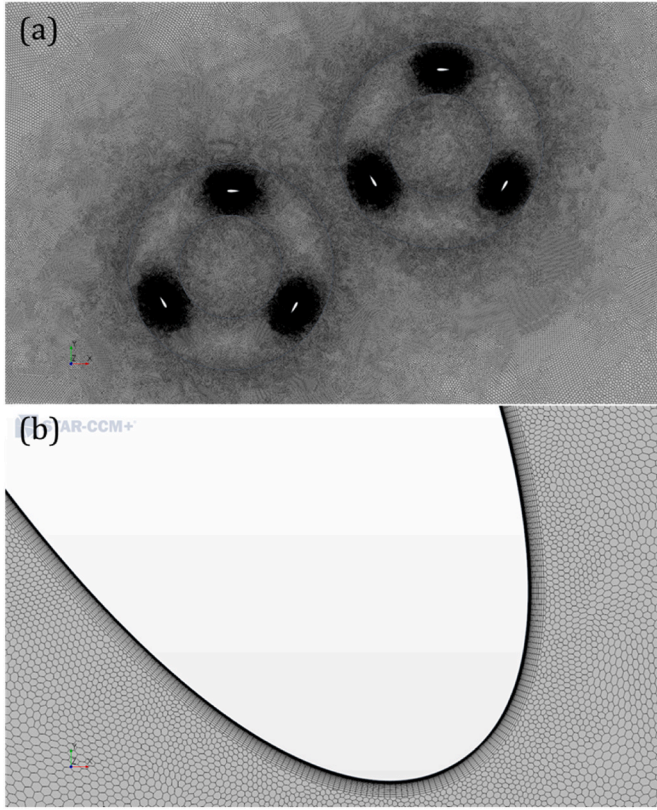


Fig. 2. Local mesh topology of fine mesh scheme: (a) twin-VAWT rotational domain; (b) blade vicinity.

where  $P_1, P_2$  and  $P_3$  stand for physical quantities for performance evaluation under three mesh schemes, respectively, which are set to the averaged turbine torques  $\bar{Q}$  when the computation becomes steady. The iteration process involves two reciprocating equations, expressed as follows.

$$p = \left| \ln \left( \frac{P_3 - P_2}{P_2 - P_1} \right) + q(p) \right| / \ln(r_{2,1}) \quad (4)$$

$$q(p) = \ln \left( \frac{r_{2,1}^p - \operatorname{sgn} \left( \frac{P_3 - P_2}{P_2 - P_1} \right)}{r_{3,2}^p - \operatorname{sgn} \left( \frac{P_3 - P_2}{P_2 - P_1} \right)} \right) \quad (5)$$

where  $\operatorname{sgn}(\cdot)$  refers to the signum. Finally, GCI and the asymptotic index

$\alpha$  can be written as follows.

$$GCI_{2,1} = 1.25 \times \left| \frac{P_1 - P_2}{P_1} \right| / (r_{2,1}^p - 1) \quad (6)$$

$$GCI_{3,2} = 1.25 \times \left| \frac{P_2 - P_3}{P_2} \right| / (r_{3,2}^p - 1) \quad (7)$$

$$\alpha = r_{2,1}^p \times \left( \frac{GCI_{2,1}}{GCI_{3,2}} \right) \quad (8)$$

As shown in Fig. 2, the polygon grid on the blade surfaces of the current fine mesh scheme is 0.15 mm, and the thickness of 28 prismatic layers near blades is 0.43 mm, thereby ensuring that the maximum value of  $y^+$  should be less than 1. An ideal factor  $r_0 = 1.25$  is selected to uniformly enlarge the grid size twice, forming the medium and coarse meshes. Under the validation by using  $0^\circ$  pitch cases, the GCI test results are listed in Table 1, where the total number of grids for fine, medium and coarse meshes are 951362, 791312 and 657620, respectively.

As can be seen, using the averaged torque as the criterion, the  $GCI_{3,2}$  values are larger than  $GCI_{2,1}$ , and the  $\alpha$  values approach to 1 for both upstream or downstream turbine. It is confirmed by the analysis that the refinements of the twin-VAWT meshes are in asymptotic range. In addition, better convergence can be found in upstream turbine, which could be partly due to the fact that the turbulence became more complex to depict as it developed downstream in this case.

### 3. Methods of surrogate models

To save the computational cost of CFD, surrogate models are used as the approximators. A small fraction of the twin-VAWT CFD cases is used as the training data and other cases are used as the test data to evaluate the accuracy. In this section, two typical SM methods are introduced to approximate the input-output function: the Kriging model (Section 3.2) and the neural network model (Section 3.3). In this paper, the inputs are pitching angle combinations of twin-VAWT blades, ranging from  $-6^\circ$  to  $0^\circ$  for both turbines. This combination is denoted as  $[\beta_1, \beta_2]$ , that is, the upstream turbine has a pitch of  $\beta_1$ , and the downstream turbine has a pitch of  $\beta_2$ . In this paper, separate SMs were built for both upstream and downstream turbines, where the outputs are their averaged rotor torque  $\bar{Q}_i$ . The expression can be written as follows:

$$\bar{Q}_i = \hat{f}(x, \omega | X) \quad (9)$$

where  $x = (\beta_1, \beta_2)$  refers to the input vector,  $\omega$  represents the parameter space from the model;  $X = (x^{(1)}, x^{(2)} \dots x^{(n)})^T$  denotes the available observations in hand, where  $n$  stands for the number of observations.

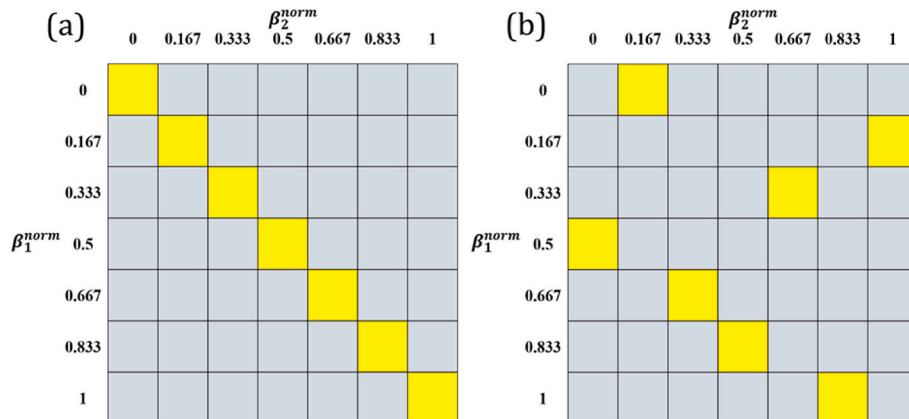


Fig. 3. Two-variable, seven-point LHS plans: (a) diagonal plan; (b) random plan.

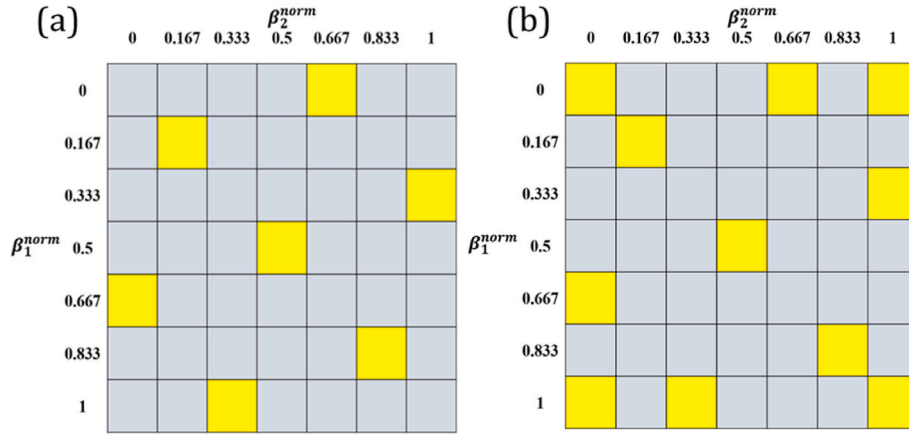


Fig. 4. Best filled sampling plans according to min-max criterion: (a) LHS plan; (b) LHS plan with corner values.

### 3.1. Sampling

The idea of SM is to infer the output based on limited observations. Therefore, the preliminary step is to formulate a sampling plan. Basically, there are two principles to follow: (1) The observations should be as less as possible; (2) The observations should also be highly representative among available spaces.

It is assumed for SMs that the relation between inputs and outputs should be smooth and continuous among spaces, and near inputs will lead to similar outputs. Moreover, based on the previous knowledge (Rezaeiha et al., 2017), pitch angle difference of more than 1° may have evident influence on turbine’s performance. Therefore, only integer numbers of pitch angles are addressed by sampling (i.e. pitch angle equals to -6°, -5°, -4° ... 0°). Those values are first normalized to the range of [0, 1] by using Formula (10):

$$\beta_i^{norm} = \frac{\beta_i - \beta_{i,min}}{\beta_{i,max} - \beta_{i,min}} \quad (10)$$

where  $\beta_{i,min} = -6^\circ$ , and  $\beta_{i,max} = 0^\circ$  in current cases.

Fig. 3 shows several Latin hypercube samplings (LHS) (Helton and Davis, 2003) plotted, where the yellow squares denote the sampled points, and the grey ones refer to the un-sampled points. This method can be used to uniformly spread the projections of the points on all axes (i.e. the multi-dimensional stratification). However, it cannot be ensured that the observations are space-filling based on this plan. For instance, as shown in Fig. 3(a), the simplest LHS layout is to set all points along the diagonal, but it can be clearly seen that they might not be representative in the entire space.

In view of this, the *max-min* criterion from Morris and Mitchell (1995) is used to evaluate the space-filling of the LHS, thereby maximizing the sorted distance of observation pair  $d$  while minimizing the number  $J$  of pairs with the corresponding distance in succession. The definition of Distance  $d$  is expressed as follows:

$$d_p(\mathbf{x}^{(a)}, \mathbf{x}^{(b)}) = \left( \sum_{j=1}^k |x_j^{(a)} - x_j^{(b)}|^p \right)^{1/p} \quad (11)$$

where  $p = 2$  refers to the norm of Euclidean distance;  $\mathbf{x}^{(a)}, \mathbf{x}^{(b)}$  stand for two observation vectors, respectively;  $j$  represents the counter of the observation pairs with the same distance of  $d_p$ .

In this paper, to find the best filled sampling plan, 10,000 random LHSs were compared and ranked through bubble sort. The best LHS is shown in Fig. 4(a) below, which intuitively and perfectly filled the whole input combinations. Apart from the above LHS, for more accurate regression around the boundaries, four pitch angle combinations at the corners were taken as the observation points. Therefore, the final

sampling plan (Fig. 4(b)) involves 11 cases in total, cutting down 77.55% of the computational cost compared with pure 49 CFD simulations. Then, in Section 4, the sensitivity of SMs in the aforementioned sampling plans will be further elaborated.

### 3.2. Kriging model

The Kriging model is a non-parametric model developed by Danie Krige first for usage in geostatistics (Krige, 1951) and it made its way to the approximation of computational experiments since the work of Sacks et al. (Sacks et al., 1989). Given the sampled observations  $X = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}\}^T$ , with the observed response  $\mathbf{y} = \{y^{(1)}, y^{(2)}, \dots, y^{(n)}\}^T$ , the target of Kriging model is to find an expression for a predicted value at a new point  $\mathbf{x}$ . As we believe the engineering function should be **smooth** and **continuous**, the sample data are correlated with each other. In Kriging model, the responses are from a stochastic process even though they are in fact from the deterministic computational code of CFD software. Here, we denote such random field as follows:

$$\mathbf{Y} = (Y(\mathbf{x}^{(1)}), Y(\mathbf{x}^{(2)}) \dots Y(\mathbf{x}^{(n)}))^T \quad (12)$$

whose mean is  $\mathbf{1}\mu$  and the random variables are correlated through the Kriging basic function as follows:

$$\text{cor}[Y(\mathbf{x}^{(i)}), Y(\mathbf{x}^{(l)})] = \exp \left( - \sum_{j=1}^k \theta_j |x_j^{(i)} - x_j^{(l)}|^{p_j} \right) \quad (13)$$

where when  $p_j$  was fixed to be 2 and  $\theta_j$  was constant, this become the well-known Gaussian radial basis. Then, the correlation matrix of all observations and a covariance matrix are expressed as:

$$\boldsymbol{\Psi} = \begin{bmatrix} \exp \left( - \sum_{j=1}^k \theta_j |x_j^{(1)} - x_j^{(1)}|^{p_j} \right) & \dots & \exp \left( - \sum_{j=1}^k \theta_j |x_j^{(1)} - x_j^{(n)}|^{p_j} \right) \\ \vdots & \ddots & \vdots \\ \exp \left( - \sum_{j=1}^k \theta_j |x_j^{(n)} - x_j^{(1)}|^{p_j} \right) & \dots & \exp \left( - \sum_{j=1}^k \theta_j |x_j^{(n)} - x_j^{(n)}|^{p_j} \right) \end{bmatrix} \quad (14)$$

$$\text{cov}(\mathbf{Y}, \mathbf{Y}) = \sigma^2 \boldsymbol{\Psi} \quad (15)$$

where  $k$  represents the dimension of inputs (equals to 2 in this case);  $\boldsymbol{\theta}$  and  $\mathbf{p}$  stand for the key parameter matrix, with the shape of  $(k, 1)$ ; and  $\sigma$  denotes the standard deviation. Hence the likelihood is expressed as:

$$L(\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(n)} | \boldsymbol{\mu}, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2} |\boldsymbol{\Psi}|^{1/2}} \exp \left( - \frac{(\mathbf{y} - \mathbf{1}\boldsymbol{\mu})^T \boldsymbol{\Psi}^{-1} (\mathbf{y} - \mathbf{1}\boldsymbol{\mu})}{2\sigma^2} \right) \quad (16)$$

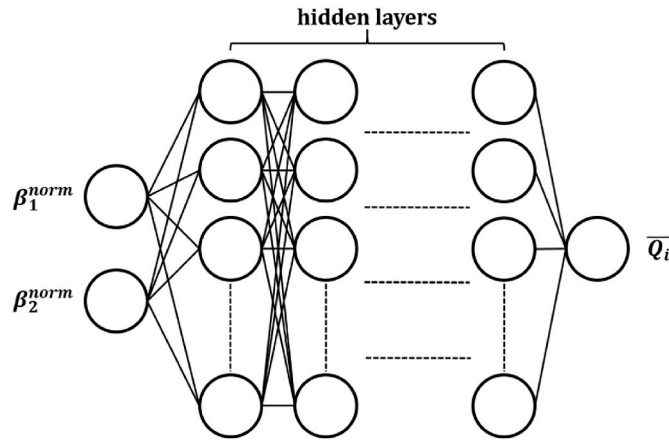


Fig. 5. Diagram of the ANN surrogate models.

By taking derivatives of the natural logarithm of (R-2-5) and setting to zero, the maximum likelihood estimates (MLEs) for  $\mu$  and  $\sigma^2$  can be obtained. Finally, the determination of appropriate  $\theta$  and  $p$  is conducted by maximizing the corresponding logarithmic function as follows, with constant terms removed:

$$\ln(L) \approx -\frac{n}{2} \ln(\hat{\sigma}^2) - \frac{1}{2} \ln|\Psi| \quad (17)$$

where  $\hat{\sigma}^2 = (y - \mathbf{1}\hat{\mu})^T \Psi^{-1} (y - \mathbf{1}\hat{\mu}) / n$  is the MLE of  $\sigma^2$ , using  $\hat{\mu} = \mathbf{1}^T \Psi^{-1} y / \mathbf{1}^T \Psi^{-1} \mathbf{1}$  as the MLE of  $\mu$ .

To maximize Formula (13), since its computational budget is low, the numerical optimization based on heuristic searching algorithms is used. In this paper, the particle swarm optimization (PSO) was adopted to find the solution to  $\theta$  and  $p$ .

### 3.3. Artificial neural network

Inspired by human neural structures, the ANN also has strong ability in establishing nonlinear mappings. Despite that the deep interpretability of ANN is still unclear, its universality as approximator for **any continuous function** has been mathematically proved (Hornik et al., 1989). As expressed by formula (18), in ANN model, each neuron receives  $n$  input signals from other neurons. Those signals are transmitted through weighted connection, and their summation is then handled by activation functions.

$$y_{ANN} = \sigma \left( \sum_{i=1}^n w x_{ANN} + b \right) \quad (18)$$

where  $x_{ANN}$  and  $y_{ANN}$  refer to the input and output of the cell, respectively;  $\sigma(\cdot)$  represents a nonlinear activation function, taking the rectified linear unit (ReLU) in this paper;  $w$  represents the weight vector; and  $b$  denotes the bias. Further, connecting many of such neurons in a certain hierarchy result in a neural network, as shown below:

The diagram of ANN is shown in Fig. 5. Dense layers with identical number of neurons were adopted to establish the model. The Adam

Table 3  
Hyperparameter settings of ANN models.

	Layer number	Cells per layer	Learning rate	Activation function	Batch size
Level 1	1	20	0.1	ReLU	2
Level 2	2	30	0.01	ReLU	2
Level 3	3	40	0.001	ReLU	2
Level 4	4	50	0.0001	ReLU	2
Upstream SM	Level 3	Level 1	Level 3	ReLU	2
Downstream SM	Level 1	Level 2	Level 3	ReLU	2

(Kingma and Ba, 2014) optimizer was used to train the model, taking the mean absolute error (MAE) as the loss function. The default epoch is set to 1000, and the early stop technique was applied, with the patience of 100.

Different from other algorithms, the hyperparameters have significant impact on the performance of ANN models. In this work, the hidden layers number, the cell (per layer) number and the learning rate were tuned through grid search (i.e., brute force). The activation function was fixed to be ReLU because of its training benefits (Brownlee, 2019). The batch size was fixed to be two as the number of training samples was small, and for similar reasons, the layer number and cells per layer were also in a small range. Detailed searching settings and the corresponding results are shown in Table 2, where at least 30 trails were conducted for each combination of levels.

## 4. Results and discussion

In this section, the results of SMs are compared with each other and also the ground truth obtained from CFD simulations. The discussion of the SM result involves accuracy, efficiency and the sensitivity to sampling plans.

### 4.1. Evaluation index

Given the true response from simulation results as  $y = (y^{(1)}, y^{(2)} \dots y^{(N)})^T$  where  $N$  refers to the total number of the dataset, and the corresponding predictions from SM are denoted as  $\hat{y} = (\hat{y}^{(1)}, \hat{y}^{(2)} \dots \hat{y}^{(N)})^T$ . Two criteria are used to quantify the performance of the SM predictions, which are  $R^2$  and the MAE loss value expressed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y^{(i)} - \hat{y}^{(i)})^2}{\sum_{i=1}^N (y^{(i)} - \bar{y})^2} \quad (19)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y^{(i)} - \hat{y}^{(i)}| \quad (20)$$

where  $\bar{y}$  refers to the mean of the true response. In general,  $R^2$  shows the consistency of the overall tendency, while MAE evaluates the difference between the model predictions and the ground truth.

Table 2  
The mesh dependency test through grid convergence index for upstream and downstream turbine of twin-VAWT system (0° pitch).

	Scheme	$N_i$	Blade surface size (mm)	$\bar{Q}_i$ (Nm)	$r_{i+1,i}$	$GCI_{i+1,i}$	$\alpha$
Upstream turbine	fine, i = 1	951362	0.15	3.7019	1.10	0.0516	0.98
	medium, i = 2	791312	0.1875	3.6330	1.10	0.0763	/
	coarse, i = 3	657620	0.234375	3.5325	/	/	/
Downstream turbine	fine, i = 1	951362	0.15	4.0171	1.10	0.0997	0.97
	medium, i = 2	791312	0.1875	3.8898	1.10	0.1438	/
	coarse, i = 3	657620	0.234375	3.7111	/	/	/

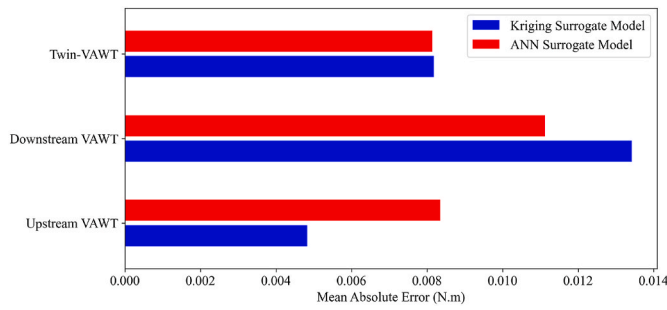


Fig. 6. Mean absolute errors for Kriging and ANN surrogate models on upstream, downstream and twin-VAWT.

4.2. Accuracy

Using the current sampling plan (Fig. 4(b)), the accuracy of Kriging and ANN models is discussed in this section.

First of all, the MAE loss for the average moments predictions of upstream turbine (UT), downstream turbine (DT) and the twin-VAWT are shown in Table 3 and Fig. 6, respectively below.

As can be seen, both Kriging and ANN models outperform true simulation in predictions. On the whole, the MAE values of twin-VAWT torque are comparable regardless of using Kriging or ANN models, achieving a considerably low value around 0.008 N • m, that is, around 0.2% of the torque outputs. Moreover, in a vertical contrast, the relative error of UT predictions is significantly less than that of DT. This may be partly due to the fact that the impact of UT on DT is more complex than that of DT on UT. As for UT, the Kriging model has an impressive loss of 0.00482 N • m, and that of MAE of ANN is relatively larger. However, for DT predictions, the Kriging model shows inferior performance than ANN, and the MAE of Kriging dramatically becomes 0.01342 N • m, that is, 20% higher than its counterpart.

Similar trends can also be found in scatter plots, as shown in Fig. 7 where R<sup>2</sup> values are marked for each turbine with different models. As can be seen from Fig. 7(a), the consistency of UT torque prediction based on Kriging model is excellent, in another word, all dots are located extremely close to the diagonal line. The red dots in Fig. 7(b) based on

ANN model also show a good distribution, achieving a high R<sup>2</sup> value of 0.990. However, as shown in Fig. 7(c) and (d), some of the prediction values evidentially deviate from the dashed line, and compared with ANN model (R<sup>2</sup> = 0.989), the bias of Kriging model (R<sup>2</sup> = 0.980) is more significant.

Fig. 8 shows the 3D surfaces and 2D projections of the true CFD results and SM predictions for UT average torque, where the black or white dots denote the 11 observations provided for model training. It can be summarized from Fig. 8(a) and (d) that, when UT pitch remains unchanged, the power performance of UT continuously increases with increase of the DT pitch from -6° to 0°. Meanwhile, when DT pitch is fixed, the UT torque first rises until its pitch reaches about -3°, and then decreases as the pitch angle becomes larger. As can be seen from Fig. 8 (b) and (c), such trends can be generally well depicted through both SMS trained by using limited observations.

However, there are subtle details not captured when comparing Fig. 8(d), (e) and (f). As shown in Fig. 8(d), the local optimal UT pitch usually suffers a fluctuation with the decrease of DT pitches: it first locates at about -3°; then gradually moves to -2°; and finally moves back to -3°. Such movement performed quite differently in SM predictions. On the one hand, for Kriging predictions, the contours are layered and inerratic, and all local optimums are around -3°. This is because the predictions of Kriging model actually maximize the likelihood of the observed data and themselves, given parameters  $\theta$  and  $p$  that were found as Section 3.2 described (Sacks et al., 1989). Hence, such intrinsically stationary process will produce optimal unbiased estimation of the test samples, where the closer two points were separated, the closer correlation of their responses would be. Though the hierarchical rules are also clear in ANN predictions, evidential over-fittings are presented in contour lines at top-right part of Fig. 8(f). The reason may be that ANN model excessively learned from the responses of [-3°, -3°], [-2°, -6°] and [-1°, -1°], and the low average torque from [-2°, -6°] was incorrectly generalized to other pitch combinations that also had a UT pitch of -2°.

Similarly, Fig. 9 shows the 3D surfaces and contour plots of the results of DT average torque. It can be found from Fig. 9(a) that while DT pitch remains the same, the DT torque shows a tendency of decrease, the UT pitch increases, and the local optimal DT pitch slightly deviates to -2

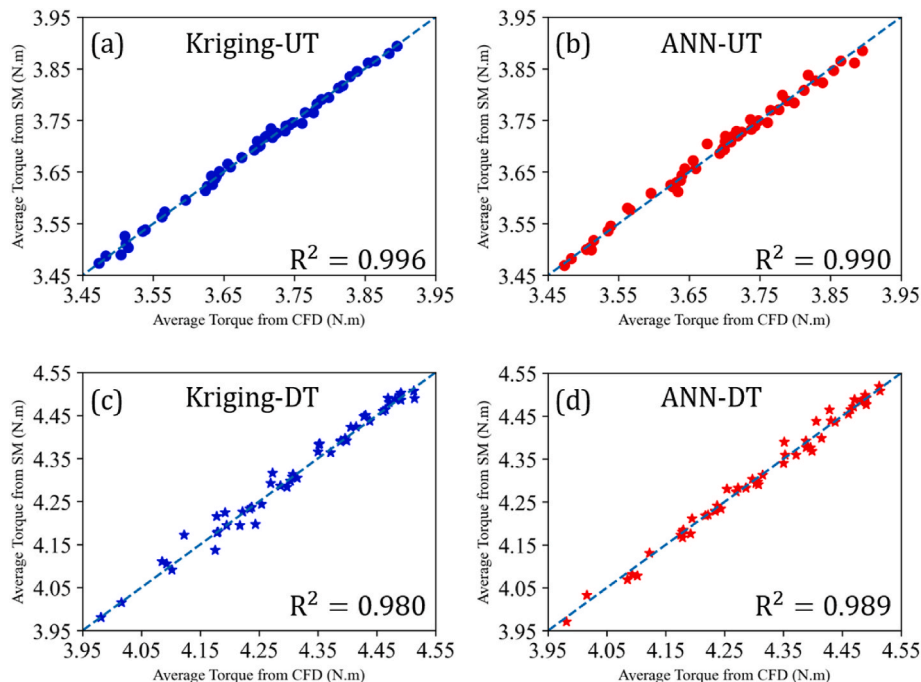


Fig. 7. Scatters plots and r-square values between ground truth and predictions from surrogate models: (a) Kriging, UT; (b) ANN, UT; (c) Kriging, DT; (d) ANN, DT.

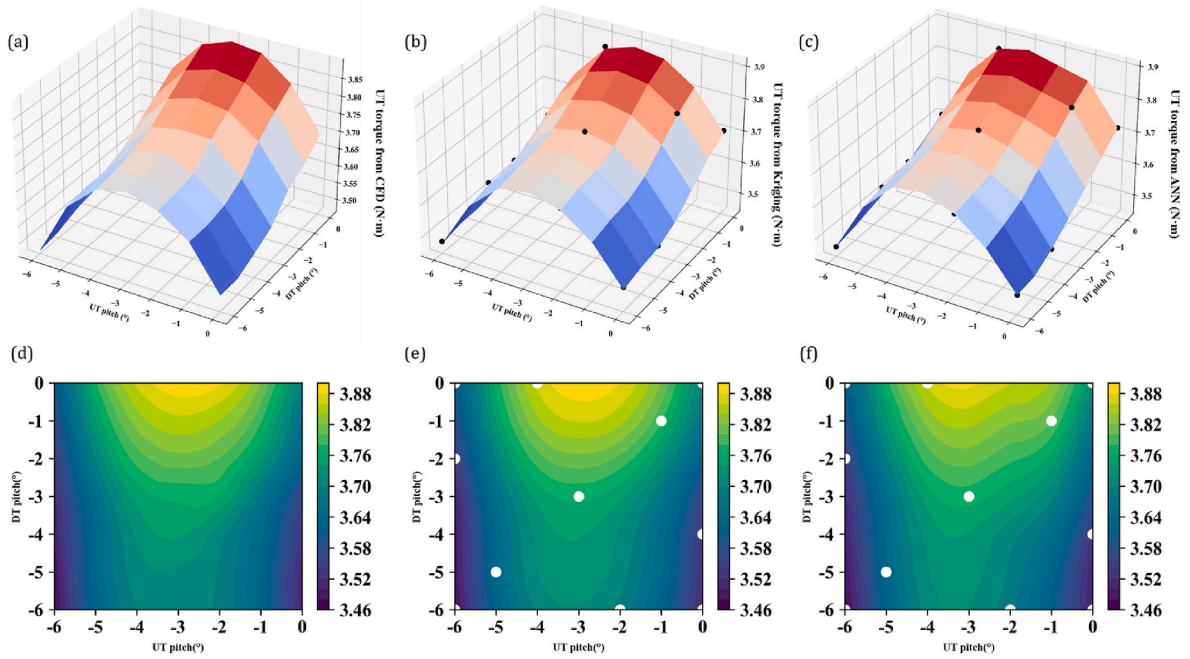


Fig. 8. The 3D surfaces and 2D contours of UT torque (a) CFD, 3D; (b) Kriging, 3D; (c) ANN, 3D; (d) CFD, 2D; (e) Kriging, 2D; (f) ANN, 2D.

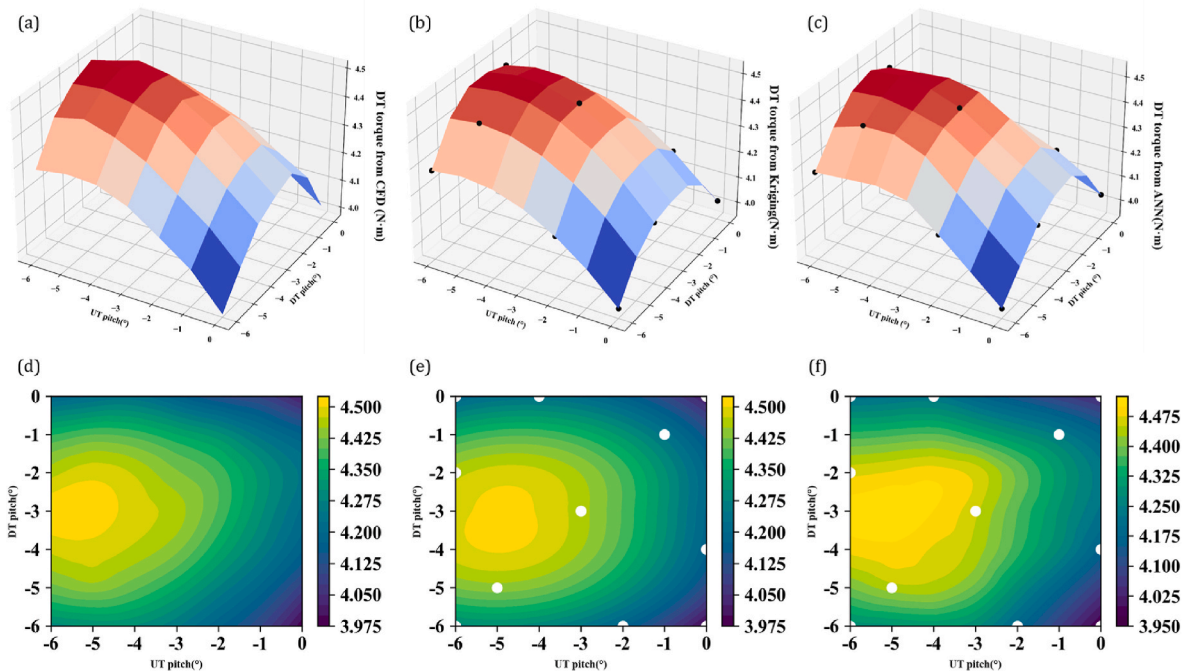


Fig. 9. The 3D surfaces and 2D contours of DT torque (a) CFD, 3D; (b) Kriging, 3D; (c) ANN, 3D; (d) CFD, 2D; (e) Kriging, 2D; (f) ANN, 2D.

° from  $-3^\circ$ . Noticeably, the prediction results from SMs have coessential patterns, as shown in Fig. 8. It can be seen from Fig. 9(b) and (e) that the Kriging results have a rather smooth surface and contour boundaries, but the contour lines of ANN models exhibit anomalous shapes, and spread successively. This indicates that the ANN model has a relatively high variance that was caused by overfitting. Such phenomenon was mainly due to the lack of training data, for instance, at the area between points  $[-6^\circ, -2^\circ]$  and  $[-3^\circ, -3^\circ]$ .

In addition, for Kriging models only, Fig. 10 shows the basic functions for UT and DT, respectively. As can be found, all  $p$  values are

approaching to 2 (quasi-Gaussian), indicating that the functions have a favourable smoothness as  $x_j^{(i)} - x_j \rightarrow 0$ . In the meantime, for both turbines, the  $\theta_j$  value of itself is larger than that of the other turbine (i.e.  $0.45 > 0.01$  and  $0.23 > 0.11$ ), showing that the change of the pitch angle of the self-turbine still poses larger influence on its power performance than the other turbine's pitch does. The aforementioned analysis can be clearly reflected in contours shown in Figs. 8 and 9.

Despite this, a quite small value of  $\theta_2 = 0.01$  can be found in UT model, while a comparatively moderate value of  $\theta_1 = 0.11$  is observed in DT model. This implies that the impact of DT pitch angles on UT

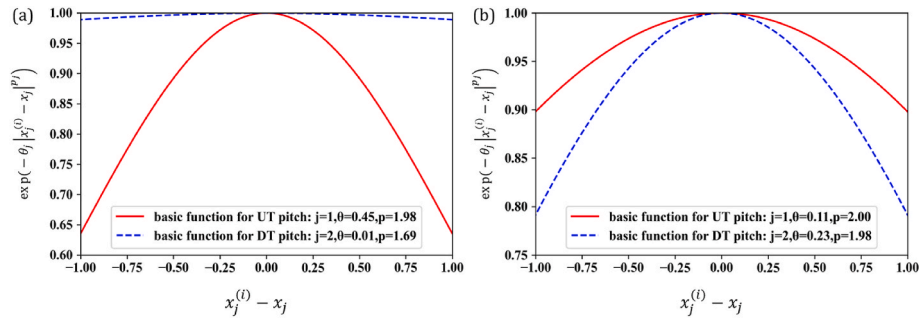


Fig. 10. The basic functions for UT or DT pitch on average torques of various turbine: (a) UT; (b) DT.

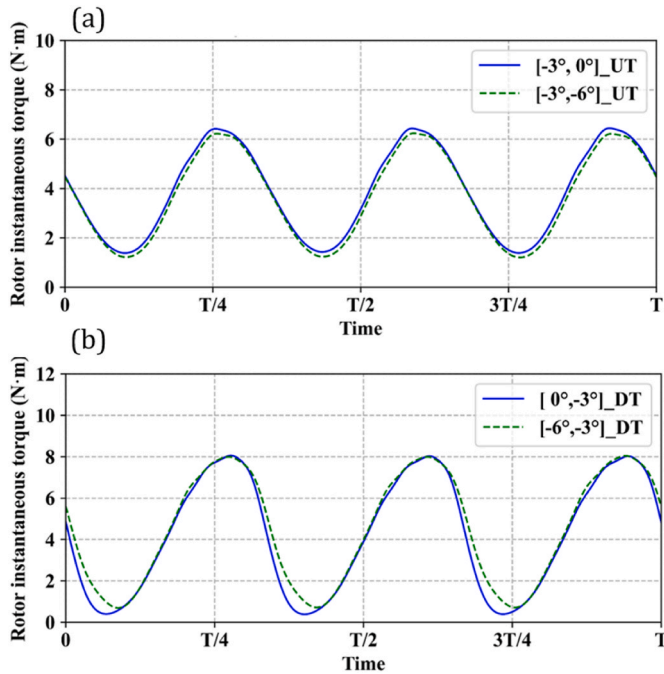


Fig. 11. Comparison of UT and DT torques between different pitch combinations: (a) UT, with pitches of  $[-3^\circ, 0^\circ]$  and  $[-3^\circ, -6^\circ]$ ; (b) DT, with pitches of  $[0^\circ, -3^\circ]$  and  $[-6^\circ, -3^\circ]$ .

Table 4 Mean absolute error of surrogate models on mean rotor torque of UT, DT and twin-VAWT system.

	UT	DT	Twin-VAWT
Kriging	0.00482 N • m	0.01342 N • m	0.00818 N • m
ANN	0.00835 N • m	0.01112 N • m	0.00814 N • m

performance is less than that of UT pitch angles on DT performance. And this can also be confirmed through CFD post-processing by extracting rotor instantaneous torques, as shown in Fig. 11. And the change of DT pitch distinctly alters the torque curves of UT (Fig. 11(b)) than the inverse (Fig. 11(a)). For more information of the fluid mechanism behind, please refer to the work of Chen et al. (2022)

### 4.3. Efficiency

The SM scripts of this paper were based on pykriging (Paulson and Ragkousis, 2015) and Keras (Gulli and Pal, 2017) packages in python 3.8 environment. The domestic server details are: Intel(R) Core(TM) i7-7500 CPU@2.70 GHz, 2.90 GHz, 16 GB RAM.

Table 5 Time used to train Kriging, ANN and CFD in the existing work (Chen et al., 2022).

	UT	DT
Kriging	6.12 s	7.01 s
ANN	21.42 s	13.15 s
CFD	12 h	12 h

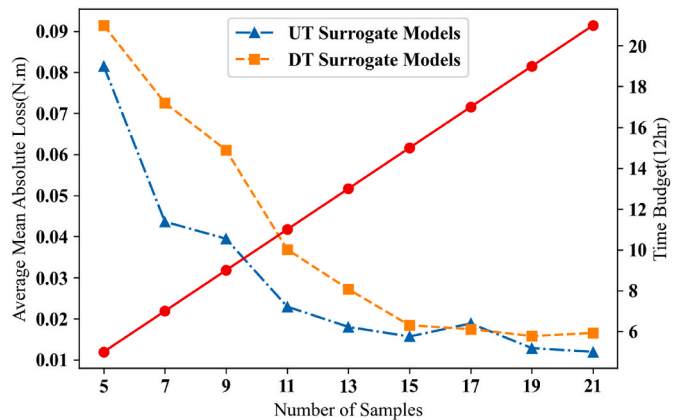


Fig. 12. The sensitivity of surrogate models to the number of samples ranging from 5 to 21.

Based on the current sampling plan for the existing work, the time to train Kriging and ANN models is listed in Table 4, and the prediction time is ignorable, which is therefore not discussed. Overall, the time for training both SMs is far less than that of CFD calculations (see Table 5; Chen et al., 2022).

As can be found, based on the existing two-dimensional small dataset, Kriging models are faster to train than ANNs. Moreover, it takes more time to train the UT-ANN model than DT, which is because that more trainable parameters appear when more layers are loaded (Table 2). However, it is worth noting that, for more complicated problems, as dataset expands, the budget of Kriging models will dramatically increase, in this case, the ANN-based models are more recommendable (Jiang et al., 2020c).

### 4.4. Sensitivity to sampling

The sensitivity of SMs to sampling plans will be elaborated from two aspects. First, the sensitivity to sampling numbers will be investigated. Then, the distribution of samples will be discussed, and the current best-filled LHS plan with corner points will be compared with other plans with the same sampling number.

To investigate the sensitivity of sample number, multiple random



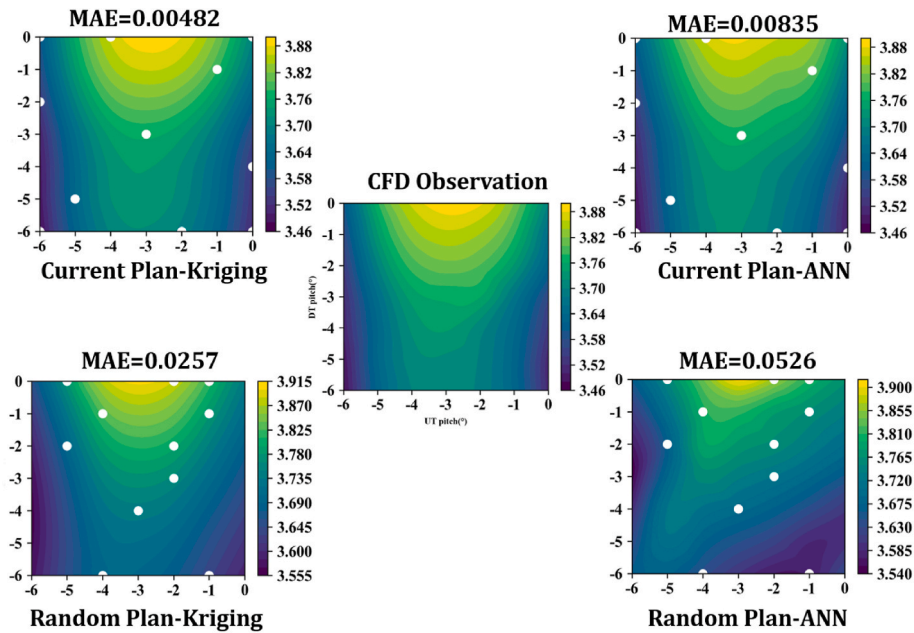


Fig. 13. Comparisons of the UT results of SMs using sampling plans by Kriging and ANN.

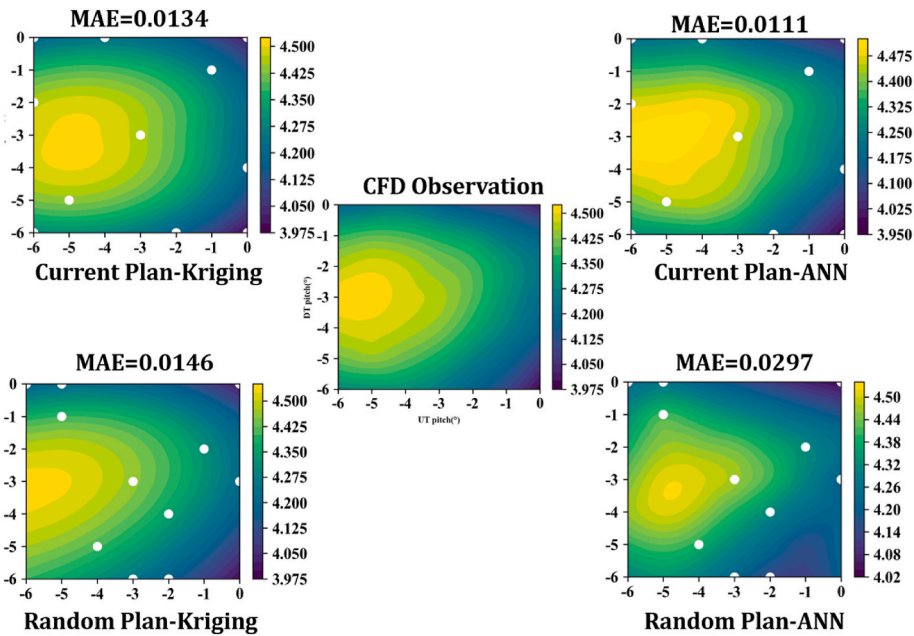


Fig. 14. Comparisons of the DT results of SMs using sampling plans by Kriging and ANN.

plans are prepared for each number ranging from 5 to 21 with both ANN and Kriging models, and the average MAE values are calculated and plotted, as shown in Fig. 12. As can be seen, when the sample numbers keep increasing, a convergence trend can be found in the average MAE of both UT and DT SMs. Meanwhile, the time budget to produce the observation points through CFD is linearly growing. As the intersection to balance two sides, in this case, the selection of totally 11 observations can not only ensure the performance of SMs in a convergence range, but also effectively control the time cost for training data generation.

Next, the influence of sample distribution is discussed, as shown in Figs. 13 and 14. First, regardless of UT or DT models, in general, the current sampling plan can better depict the true distributions, resulting in lower loss values. Second, the gap between ANN models with different samplings are much more evidential than that between Kriging ones,

where the MAE soars to 0.0526 N m for UT predictions, and doubles in DT predictions. Moreover, the contours of ANN still suffer overfitting problems where sampling points occur. Therefore, it can be concluded that, to deal with the current problem, ANN models are more sensitive to training data distribution than Kriging models.

It can also be found from Figs. 13 and 14 that the best plan could be different for different types of models or with different datasets. However, for engineering problems, designers cannot grasp the overall contours of the parameter space, therefore a universal solution should be provided. Based on the above analysis, the current sampling strategy (i. e. best-fill LHS and corner samples) conducted by using Kriging algorithm could be a promising and stable way to establish a rapid surrogate model.

## 5. Conclusion

In this paper, two types of surrogate models, i.e., Kriging and ANN, were employed for the average torque predictions of twin-VAWT with various pitch angle combinations. The data source was obtained from high-budget unsteady CFD simulations. Totally 11 out of 49 cases were sampled to train the surrogate models, saving 77.56% of the time consumption from CFD calculation. Through numerical experiments and analysis, the following conclusions can be drawn.

1. Both Kriging and ANN models can offer accurate predictions of average turbine torques with unknown pitch combinations for both UT and DT. All  $R^2$  values of the surrogate models are more than 0.98, and the UT predictions are more accurate than DT predictions for both models.
2. With the current small dataset, Kriging models are more time-efficient, and more stable than ANN models. Moreover, the parameter  $\theta$  of the basic functions of Kriging can provide helpful information to reveal the importance of input variables on the response. In another word, the impact of UT pitch angles on DT performance is more significant than its inverse.
3. Through sensitivity analysis, the current sampling number and distribution formed by the best-fill LHS and corner observation points are effective and robust.

To summarize, this paper provided a rapid, accurate and stable method for twin-VAWT performance prediction. In the future, hybrid method could be considered to integrate the advantages of Kriging and ANN models for more complicated problems, such as multi-VAWT interactions.

## CRedit authorship contribution statement

**Yaoran Chen:** Investigation, Methodology, Software, Writing – review & editing, Data curation. **Dan Zhang:** Methodology, Writing – review & editing. **Xiaowei Li:** Methodology, Writing – review & editing. **Yan Peng:** Conceptualization, Supervision. **Xiangyu Zhang:** Writing – review & editing. **Zhaolong Han:** Supervision, Writing – review & editing. **Yong Cao:** Investigation, Methodology. **Zhikun Dong:** Investigation.

## 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.

## Acknowledgements

We acknowledge the financial support from the National Natural Science Foundation of China, China (Nos. 61973208, 52122110, 62225308, 51879160) and the Kuaisu-Fuchi Project (No. 80912020201). This work was also sponsored by the "Shuguang Program" (18SG36) supported by Shanghai Education Development Foundation and Shanghai Municipal Education Commission. We would also like to acknowledge the dedicated help from Zhenglong Wu and Linrong Chen.

## References

- Biswas, A., Sarkar, S., Gupta, R., 2016. Application of artificial neural network for performance evaluation of vertical axis wind turbine rotor. *Int. J. Ambient Energy* 37 (2), 209–218.
- Brownlee, J., 2019. A gentle introduction to the rectified linear unit (ReLU). *Machine learning mastery* 6.
- Brunton, S.L., Noack, B.R., Koumoutsakos, P., 2020. Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* 52, 477–508.
- Chen, Y., Dong, Z., Su, J., Wang, Y., Han, Z., Zhou, D., Bao, Y., 2021a. Framework of airfoil max lift-to-drag ratio prediction using hybrid feature mining and Gaussian process regression. *Energy Convers. Manag.* 243, 114339.
- Chen, Y., Wang, Y., Dong, Z., Su, J., Bao, Y., 2021b. 2-d regional short-term wind speed forecast based on cnn-lstm deep learning model. *Energy Convers. Manag.* 244 (4), 114451.
- Chen, Y., Kuang, L., Su, J., Zhou, D., Cao, Y., Chen, H., et al., 2022. Investigation of pitch angles on the aerodynamics of twin-VAWT under staggered arrangement. *Ocean Eng.* 254, 111385.
- Cheng, B., Du, J., Yao, Y., 2022. Power prediction formula for blade design and optimization of dual Darrieus wind turbines based on taguchi method and genetic expression programming model. *Renew. Energy* 192, 583–605.
- Erichson, N.B., Mathelin, L., Yao, Z., Brunton, S.L., Mahoney, M.W., Kutz, J.N., 2020. Shallow neural networks for fluid flow reconstruction with limited sensors. *Proc. Royal Soc. A* 476 (2238), 20200097.
- Gulli, A., Pal, S., 2017. *Deep Learning with Keras*. Packt Publishing Ltd.
- Guo, Y., Wang, H., Lian, J., 2022. Review of integrated installation technologies for offshore wind turbines: current progress and future development trends. *Energy Convers. Manag.* 255, 115319.
- Helton, J.C., Davis, F.J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Saf.* 81 (1), 23–69.
- Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. *Neural Network.* 2 (5), 359–366.
- Ismail, H.Y., Shirazian, S., Skořetska, I., Mynko, O., Ghanim, B., Leahy, J.J., et al., 2019. ANN-Kriging hybrid model for predicting carbon and inorganic phosphorus recovery in hydrothermal carbonization. *Waste Manag.* 85, 242–252.
- Jiang, Y., Zhao, P., Zou, L., Zong, Z., Wang, K., 2020a. Two-dimensional computational fluid dynamics study on the performance of twin vertical axis wind turbine with deflector. *J. Energy Resour. Technol.* 142 (8), 081303.
- Jiang, Y., Zhao, P., Stoesser, T., Wang, K., Zou, L., 2020b. Experimental and numerical investigation of twin vertical axis wind turbines with a deflector. *Energy Convers. Manag.* 209, 112588.
- Jiang, P., Zhou, Q., Shao, X., 2020c. *Surrogate Model-Based Engineering Design and Optimization*. Springer, Berlin/Heidelberg, Germany.
- Jin, G., Zong, Z., Jiang, Y., Zou, L., 2020. Aerodynamic analysis of side-by-side placed twin vertical-axis wind turbines. *Ocean Eng.* 209, 107296.
- Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization arXiv preprint arXiv:1412.6980.
- Krige, D.G., 1951. A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. S. Afr. Inst. Min. Metall* 52 (6), 119–139.
- Ma, Y., Zhu, Y., Zhang, A., Hu, C., Liu, S., Li, Z., 2022. Hydrodynamic performance of vertical axis hydrokinetic turbine based on Taguchi method. *Renew. Energy* 186, 573–584.
- Menter, F.R., 2009. Review of the shear-stress transport turbulence model experience from an industrial perspective. *Int. J. Comput. Fluid Dynam.* 23 (4), 305–316.
- Morris, M.D., Mitchell, T.J., 1995. Exploratory designs for computational experiments. *J. Stat. Plann. Inference* 43 (3), 381–402.
- Paulson, C., Ragkousis, G., 2015. pyKriging: a Python Kriging Toolkit.
- Peng, H.Y., Han, Z.D., Liu, H.J., Lin, K., Lam, H.F., 2020. Assessment and optimization of the power performance of twin vertical axis wind turbines via numerical simulations. *Renew. Energy* 147, 43–54.
- Rezaeiha, A., Kalkman, I., Blocken, B., 2017. Effect of pitch angle on power performance and aerodynamics of a vertical axis wind turbine. *Appl. Energy* 197, 132–150.
- Roache, P.J., 1994. *Perspective: a Method for Uniform Reporting of Grid Refinement Studies*.
- Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. *Stat. Sci.* 4 (4), 409–423.
- Sahebzadeh, S., Rezaeiha, A., Montazeri, H., 2020. Towards optimal layout design of vertical-axis wind-turbine farms: double rotor arrangements. *Energy Convers. Manag.* 226, 113527.
- Sahebzadeh, S., Rezaeiha, A., Montazeri, H., 2022. Vertical-axis wind-turbine farm design: impact of rotor setting and relative arrangement on aerodynamic performance of double rotor arrays. *Energy Rep.* 8, 5793–5819.
- Shrestha, S., Manogharan, G., 2017. Optimization of binder jetting using Taguchi method. *JOM (J. Occup. Med.)* 69 (3), 491–497.
- Sun, G., Wang, S., 2019. A review of the artificial neural network surrogate modeling in aerodynamic design. *Proc. IME G J. Aero. Eng.* 233 (16), 5863–5872.
- Sun, H., Qiu, C., Lu, L., Gao, X., Chen, J., Yang, H., 2020. Wind turbine power modelling and optimization using artificial neural network with wind field experimental data. *Appl. Energy* 280, 115880.
- Zanforlin, S., Nishino, T., 2016. Fluid dynamic mechanisms of enhanced power generation by closely spaced vertical axis wind turbines. *Renew. Energy* 99, 1213–1226.
- Zhao, X., Du, L., Peng, X., Deng, Z., Zhang, W., 2021. Research on refined reconstruction method of airfoil pressure based on compressed sensing. *Theor. Appl. Mech. Lett.* 11 (2), 100223.