Data-driven optimization of building-integrated ducted openings for wind energy harvesting

Sensitivity analysis of metamodels

Kaseb, Z.; Montazeri, H.

DOI
10.1016/j.energy.2022.124814

Publication date
2022

Document Version
Final published version

Published in
Energy

Citation (APA)
Kaseb, Z., & Montazeri, H. (2022). Data-driven optimization of building-integrated ducted openings for wind energy harvesting: Sensitivity analysis of metamodels. *Energy*, *258*, Article 124814. https://doi.org/10.1016/j.energy.2022.124814

Energy 258 (2022) 124814

Contents lists available at ScienceDirect

Energy

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

Data-driven optimization of building-integrated ducted openings for wind energy harvesting: Sensitivity analysis of metamodels

Z. Kaseb [a], H. Montazeri [b,*]

[a] Delft University of Technology, Delft, the Netherlands
[b] Eindhoven University of Technology, Eindhoven, the Netherlands

ARTICLE INFO

Keywords:
Computational fluid dynamics (CFD)
Design of experiments (DOE)
Machine learning
Surrogate model
Diffuser-augmented wind turbine (DAWT)
Zero-energy building

ABSTRACT

Metamodels are developed and used for aerodynamic optimization of a ducted opening integrated into a high-rise building to maximize the amplification factor within the duct. The duct consists of a nozzle, a throat, and a diffuser. 211 high-resolution 3D RANS CFD simulations are performed to generate training and testing datasets. The space-filling design and Genetic algorithm are used for data sampling and optimization, respectively. The performance of five commonly-used metamodels is systematically investigated: Response Surface Methodology (RSM), Kriging (KG), Neural Network (NN), Support Vector Regression (SVR), and Genetic Aggregation Response Surface (GARS). The investigation is based on (i) detailed in-sample and out-of-sample evaluations of the metamodels, (ii) annual available power in the wind ($P_{available}$), and (iii) annual energy production (AEP) for a 3-bladed horizontal-axis wind turbine (HAWT) installed in the mid-throat for the optimum designs obtained by the metamodels. The results show that converging-diverging ducted openings can magnify the experienced wind speed by the turbine and enhance the available wind power. In addition, the use of different metamodels can lead to a variation of up to 153% in the estimated $P_{available}$. For a small dataset, crude yet still acceptable accuracy can be achieved for Genetic Aggregation Response Surface and Kriging at a very low computational time.

1. Introduction

1.1. State-of-the-art and research gaps

Urban wind energy harvesting systems offer a decentralized power generation close to where the generated power is needed and can reduce the power load and the cost of electricity transmission infrastructure. By eliminating the long-distance from end-users, decentralized power generation leads to reducing transmission losses, allowing for power generation with high efficiency [1,2]. This is essential for the implementation of sustainable development goals [3], such as realization of smart grid systems [4,5] and carbon emissions reduction [6]. However, the power performance of wind turbines in the urban environment is highly influenced by the complex wind flows characterized by rather low mean wind speed and high turbulence intensity [7]. As a result, for a considerable time fraction of the year, the mean wind speed is lower than the cut-in wind speed of the turbine. Consequently, the turbine fails to start and generate power [8]. In addition, because the wind direction frequently changes and the turbulence intensity is rather large, wind turbines undergo substantial fluctuations resulting in reduced generated power and also component fatigue. For these reasons, among others, urban wind energy potential has not yet been intensively exploited [9–11].

Wind energy in urban areas can be harvested through stand-alone wind turbines [12,13], building-mounted wind turbines on roofs [2], and building-integrated wind turbines incorporated into the architecture of buildings [6]. The integration of wind turbines into the architectural design of buildings, e.g., between parallel buildings and within ducted openings through high-rise buildings, can increase the power performance of the wind turbines [14,15]. Building-integrated ducted openings allow the turbine to experience significantly higher velocities than the freestream velocity due to flow acceleration inside the duct. This increases the capacity factor by enabling the turbine to remain operational within a larger number of hours per year and also enhances the available wind power by magnifying the experienced wind speed by the turbine. For example, earlier research has shown that converging-diverging ducted openings can increase the streamwise wind velocity inside the duct by up to 33% [16]. In addition, they can efficiently reduce turbulence intensity, which is essential to reduce the risk of fatigue damage of urban wind turbines, inconsistent power output,

* Corresponding author.
E-mail address: h.montazeri@tue.nl (H. Montazeri).

https://doi.org/10.1016/j.energy.2022.124814
Received 27 January 2022; Received in revised form 9 July 2022; Accepted 11 July 2022
Available online 14 July 2022

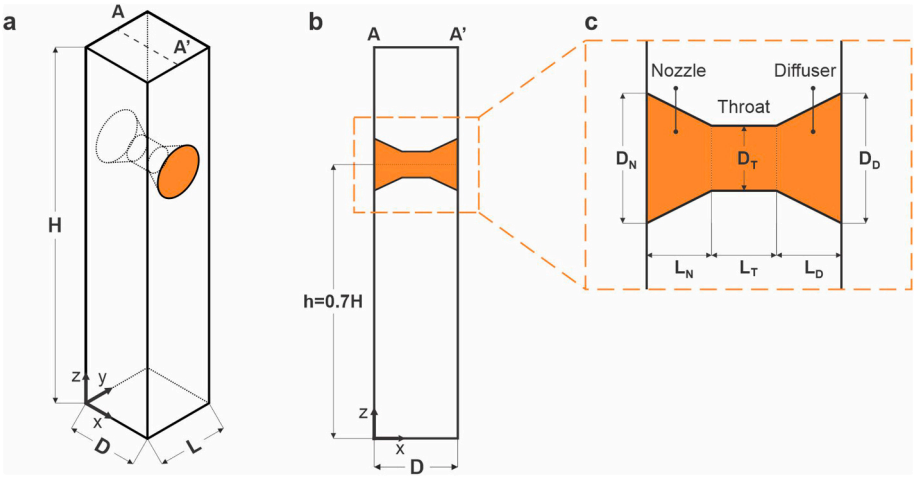

Fig. 1. (a) Perspective view and (b and c) side view of building and integrated ducted opening.

etc. Previous studies have also shown that building-integrated ducted openings can reduce vibration and noise pollution and reduce the risk of flicker and ice or wind turbine parts that could be shed and fall onto pedestrians and traffic [17,18]. Therefore, building-integrated ducted openings can be promising options for wind energy harvesting in low mean wind speed environments, such as urban areas.

However, the geometrical characteristics of ducted openings need to be aerodynamically optimized to adapt to the complex wind flow in the urban environment [19–21]. The power performance of building-integrated ducted wind turbines is influenced by several parameters that vary independently, e.g., wind resource, building characteristics and geometry, turbine technology, etc. Therefore, it would be computationally very expensive to perform direct optimization for which many alternatives need to be investigated. For example, CFD-based optimization of such complex systems would provide an unprecedented challenge in computational grid generation and excessive required computational resources for performing the large number of simulations, which are not always affordable [22]. This is especially the case for complex duct geometries, e.g., converging-diverging ducts. Therefore, advanced optimization approaches, such as metamodel-based design optimization, need to be used to obtain an aerodynamically optimized design of building-integrated ducted openings [23].

Metamodels (often called "data-driven models") are statistical models developed based on a limited number of simulations to predict the output results of simulations yet to be performed [24,25]. Therefore, the use of metamodels is computationally less expensive than performing direct CFD-based optimization. However, the accuracy of metamodels is of concern as it can be affected by the level of complexity and the number of input parameters for the design optimization, and also the "architecture" of the metamodel (e.g., Refs. [23,26,27]). Earlier studies have shown that the use of different metamodels for a given engineering application can considerably affect the accuracy [28,29]. For example, Paiva et al. [29] performed a comparative study to assess the performance of Response Surface Methodology (RSM), Kriging (KG), and Neural Network (NN) for three aircraft applications based on the corresponding out-of-sample errors. It was observed that KG and NN provide more accurate results where the number of input parameters is relatively large. In another study, Simpson et al. [28] compared the performance of RSM and KG for design optimization of an aerospike nozzle in order to improve the vehicle's aerodynamic and structural performance, e.g., the engine thrust to weight ratio, where the comparison was based on an out-of-sample evaluation. The results showed that for a small training dataset size and a small number of input parameters, RSM and KG yield different yet still comparable out-of-sample errors.

1.2. Objectives and novelty

To the best knowledge of the authors, a detailed evaluation of different metamodels for aerodynamic shape optimization of building-integrated converging-diverging ducted openings has not yet been performed. Therefore, in this study, the performance of different commonly-used metamodels is systematically investigated for aerodynamic shape optimization of such ducted openings integrated into high-rise buildings. Five metamodels are considered: Response Surface Methodology (RSM), Kriging (KG), Neural Network (NN), Support Vector Regression (SVR), and Genetic Aggregation Response Surface (GARS). The assessment is based on (i) detailed in-sample and out-of-sample evaluations of the metamodels, (ii) annual available power in the wind ($P_{available}$), and (iii) annual energy production (AEP) for a 3-bladed horizontal-axis wind turbine (HAWT) installed in the mid-throat for the optimum designs obtained by the metamodels. In total, 211 high-resolution 3D steady Reynolds-averaged Navier-Stokes (RANS) CFD simulations are performed to generate training and testing datasets. The CFD simulations are extensively validated with wind-tunnel measurements. The space-filling design and Genetic algorithm are used for data sampling and optimization, respectively.

The results of this study bring a better understanding of the aerodynamic design optimization of building-integrated ducted openings, contribute to increasing the penetration level and competitiveness of the renewable energy systems, and support the realization of a sustainable, zero- and plus-energy built environment. Given the complexity involved in wind energy performance prediction of such systems in the built environment, the data-driven optimization approach can maximize energy production and potentially lead to a higher level of successful deployments of urban wind energy systems. In addition, the performed sensitivity analysis of the metamodels in this study contributes to improving the prediction accuracy and is beneficial to the design of aerodynamically optimized building-integrated ducted openings.

1.3. Outline

The building-integrated ducted opening is described in Section 2. The methodology is presented in Section 3. The computational settings for CFD simulations are presented in Section 4. Section 5 presents detailed information about metamodeling. The results of the optimization and the sensitivity analysis of metamodels are provided in Section 6. A discussion on the limitations of the study and the main conclusions are presented in Sections 7 and 8, respectively.

Table 1
List of parameters and corresponding min. and max. values for metamodeling and optimization, and geometrical characteristics of the reference case.

Parameters	Minimum	Maximum	Values for reference case
D_N/D	0.3	0.9	0.9
D_T/D	0.3	0.9	0.3
D_D/D	0.3	0.9	0.6
L_N/D	0.2	0.4	0.3
L_D/D	0.2	0.4	0.3

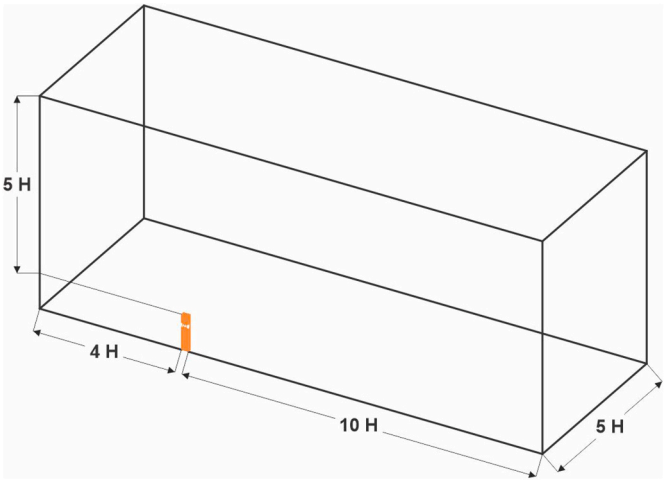

Fig. 2. Perspective view of computational domain (H = 150 m).

2. Building-integrated ducted opening

A building with dimensions $L \times D \times H = 30 \times 30 \times 150\ m^3$ is considered (Fig. 1). The building has a ducted opening in which a wind turbine could be installed for wind energy harvesting. The duct is located at h = 0.7H, near the stagnation point on the windward façade if there would be no opening (e.g., Ref. [30]), and consists of three parts: nozzle, throat, and diffuser (Fig. 1b and c). The inlet opening of the nozzle and the outlet opening of the diffuser is located on the windward and leeward facades of the building, respectively. The building dimensions and the position of the ducted opening are assumed to be fixed. Five parameters determine the geometry of the ducted opening: the diameter of the nozzle inlet (D_N), the diameter of the throat (D_T), the diameter of the diffuser outlet (D_D), the length of the nozzle (L_N), and the length of the diffuser (L_D) (Fig. 1c). It should be noted that the length of the throat (L_T) can be determined using $L_T = L - (L_N + L_D)$ and is not considered as an input parameter. Table 1 shows the parameters and the corresponding maximum and minimum values used for metamodeling and optimization. The table also presents the geometrical characteristics of the "reference case" used for the grid-sensitivity analysis that will be presented in Section 4.1.

3. Methodology

Metamodels are developed and used for aerodynamic optimization of the duct to maximize the amplification factor inside the throat (Fig. 1). The amplification factor (AF) is defined as the ratio between the streamwise velocity in the middle of the throat and the undisturbed streamwise velocity at the height of the centerline of the ducted opening (U_{ref} = 4.19 m/s). The evaluation is based on 3D steady Reynolds-averaged Navier-Stokes (RANS) CFD simulations, extensively validated with wind-tunnel measurements. The design optimization application Ansys DesignXplorer is used to perform the optimization, which consists of three steps:

(i) *Sampling:* the space-filling design is used to generate samples uniformly distributed and as far apart as possible over the design space (e.g., Ref. [31]). The excellent performance of space-filling designs for computer experiments has been shown on several occasions in the literature (e.g. Ref. [25]). For the five input parameters presented in Table 1, this results in a training dataset of 162 different geometries.

(ii) *Metamodeling:* five commonly-used metamodels are considered: Response Surface Methodology (RSM), Kriging (KG), Neural Network (NN), Support Vector Regression (SVR), and Genetic Aggregation Response Surface (GARS). Detailed in-sample and out-of-sample evaluations are performed to assess the accuracy of the metamodels.

(iii) *Optimization:* Genetic algorithm [32,33] is used to find an optimum design that maximizes the amplification factor inside the duct. The population size, the maximum number of iterations, and the crossover and mutation rates are set at 100, 100, 0.9, and 0.1, respectively. The hyper-parameters are set based on the best practice guidelines (e.g., Ref. [34]). Note that for optimization, the diameter of the throat (D_T) is assumed to be constant D_T/D = 0.3. This assumption allows a fair comparison of the metamodels by assessing the annual energy production (AEP) considering a specific horizontal-axis wind turbine installed in the mid-throat of the optimized geometries. Therefore, four input parameters and a single objective are considered for optimization.

4. CFD simulations

4.1. Computational domain and grid

As the geometry of the building and the ducted opening, and the results of steady RANS simulations are symmetric in the lateral direction, only half of the computational domain is modeled (Fig. 2). The dimensions of the computational domain are based on the best practice guidelines for CFD simulations of urban flows [35,36]. The height of the domain is 5H, where H = 150 m is the building height. The upstream, downwind, and lateral domain lengths are 4H, 10H, and 5H, respectively.

The computational grid is generated using Fluent Meshing 19.3 and considering best practice guidelines for grid generation [35–38]. For all geometries, the grid is refined near the ground, the building, and duct surfaces where the average y^+ values are 589, 484, and 292, respectively. As standard wall functions are used in this study, these values ensure that the center point of the wall-adjacent cell is placed in the logarithmic layer. A maximum growth ratio of 1.1 controls the cells located in the immediate surroundings of the above mentioned regions. For the reference case, introduced in Table 1 and shown in Fig. 3, the grid consists of 2,456,780 polyhedral cells.

The grid resolution is resulted from a grid-sensitivity analysis performed for the reference case by coarsening and refining the grid with an overall linear factor of $\sqrt{2}$. The number of cells is 1,118,49 and 4,969,471 for the coarse and fine grids, respectively. Fig. 4 shows the dimensionless mean streamwise velocity (u/U_{ref}) along a horizontal and a vertical line in the middle of the ducted opening for the three grids where U_{ref} = 4.19 m/s is the undisturbed streamwise velocity at the height of the centerline of the ducted opening, i.e., h = 0.7H. For the horizontal line, the average absolute deviation between the coarse and the basic grids and the basic and fine grids are 0.42% and 1.43%, respectively. For the vertical line, these deviations are 1.70% and 2.23%, respectively. This indicates that the CFD results do not significantly change by further refining the grid. Therefore, the basic grid is retained for all simulations.

The grid convergence index (GCI) by Roache [39] is also calculated using the maximum streamwise velocity in the middle of the throat with a safety factor (F_s) of 1.25. The GCI^{coarse} and GCI^{fine} for the

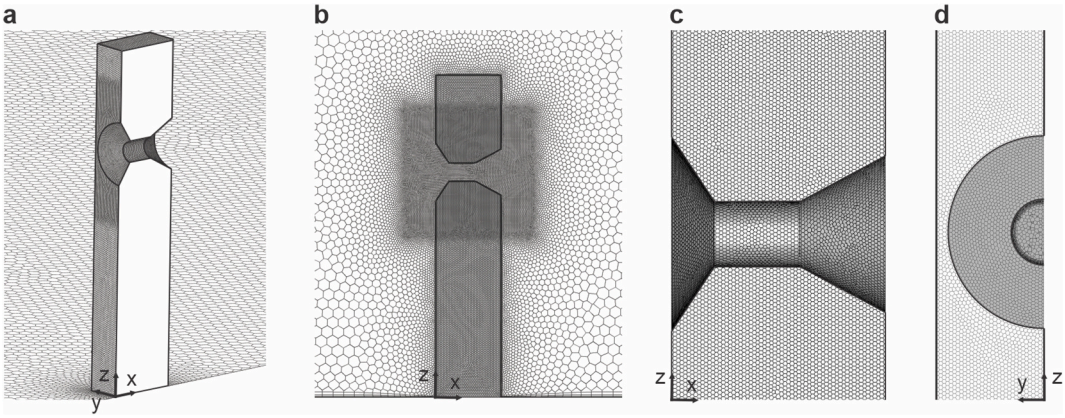

Fig. 3. Details of computational grid for reference case (2,456,780 cells) at (a) windward side and roof of building, ground and ducted opening surfaces, (b) symmetry plane and building walls, (c) ducted opening surfaces, and (d) nozzle surface and mid-throat plane.

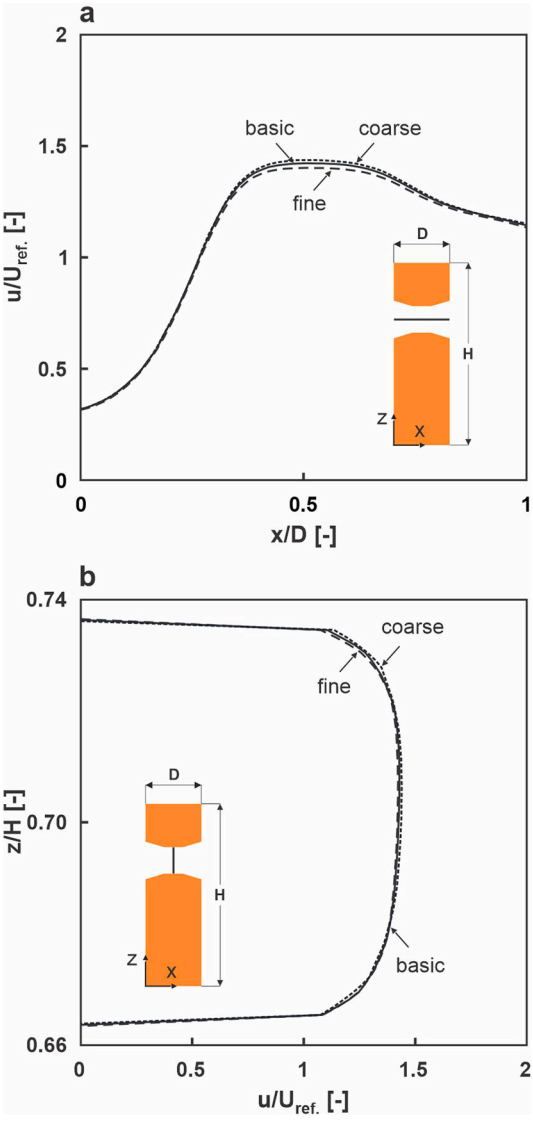

Fig. 4. Grid-sensitivity analysis: dimensionless mean streamwise velocity along (a) horizontal and (b) vertical lines in middle of ducted opening for three grids.

coarse-medium grid pair are 1.62% and 0.87%, respectively. As a large number of CFD computations are to be performed in the present study, the coarse-basic grid pair is used for this analysis.

4.2. Boundary conditions

At the inlet of the domain, neutral atmospheric boundary layer inflow profiles of mean velocity u (m/s) (Eq. (1)), turbulent kinetic energy TKE (m^2/s^2) (Eq. (2)), and turbulent dissipation rate ε (m^3/s^3) (Eq. (3)) are imposed [40]:

$$u(z) = \frac{u^*_{ABL}}{\kappa} \ln\left(\frac{z + z_0}{z_0}\right) \quad (1)$$

$$TKE(z) = 3.33\ u^{*2}_{ABL} \quad (2)$$

$$\varepsilon(z) = \frac{u^{*3}_{ABL}}{\kappa(z + z_0)} \quad (3)$$

where u^*_{ABL} is the atmospheric boundary layer friction velocity, $\kappa = 0.42$ is the von Karman constant, z is the height coordinate, and the reference velocity at $z = 10$ m is $u_{(10)} = 3$ m/s. It is assumed that the building is situated on a large grass-covered terrain with an aerodynamic roughness length $z_0 = 0.03$ m [41]. The ground surface and the building walls are considered as no-slip walls. The standard wall functions by Launder and Spalding [42] with roughness modification by Cebeci and Bradshaw [43] are used for both ground surface and building walls. The values of the roughness parameters, i.e., the sand-grain roughness height k_S and the roughness constant C_S, are determined using their consistent relationship with z_0 derived by Eq. (4) [44]. The k_S value is equal to the distance between the center point of the wall-adjacent cell and the wall throughout the domain ($k_S = 0.25$). Zero-gauge static pressure is applied at the outlet plane. Symmetry conditions are used at the top and lateral sides of the domain.

$$k_S = \frac{9.793\ z_0}{C_S} \quad (4)$$

4.3. Computational settings

The commercial CFD code ANSYS/Fluent 19.3 is used to perform the simulations. The 3D steady RANS equations are solved in combination with the Renormalization Group $k - \varepsilon$ turbulence model (RNG) [45,46] for closure. The turbulence model is selected based on a detailed validation study provided in Section 4.4. The SIMPLE algorithm is used for pressure-velocity coupling, and second-order discretization schemes are used for the RANS equations. A converged solution is assumed to be obtained when the scaled normalized residuals stabilize at a minimum of

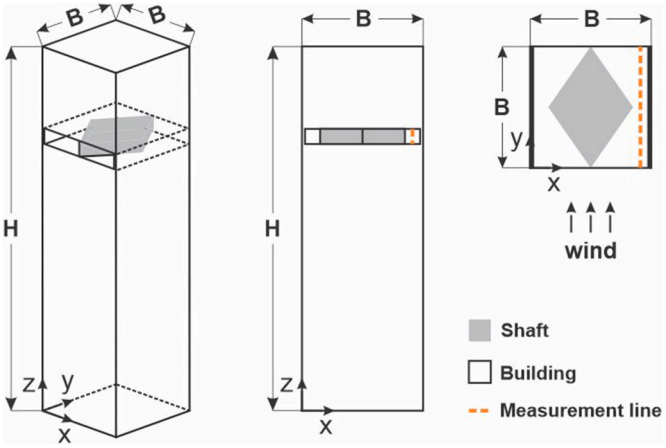

Fig. 5. CFD validation: schematic of reduced-scale building model used in wind-tunnel measurement [16].

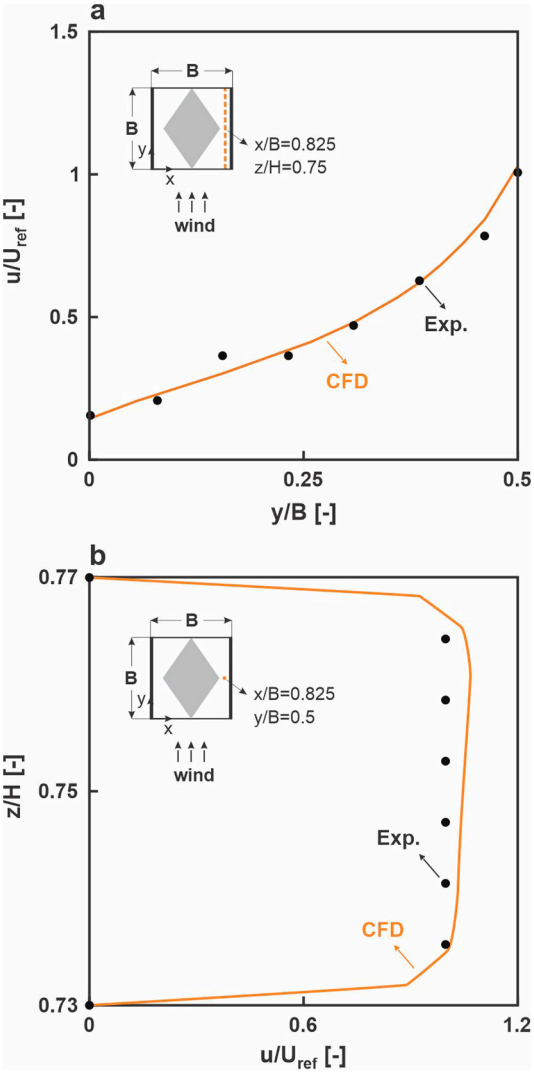

Fig. 6. CFD validation: comparison of wind-tunnel and CFD results of dimensionless streamwise wind velocity along (a) horizontal and (c) vertical lines located in middle of PIV measurement plane (x/B = 0.825).

10^{-5} for TKE, x, y, and z momentum, and 10^{-4} for continuity and ε.

4.4. Validation

Atmospheric boundary layer wind-tunnel experiments of amplification factor within a building-integrated ducted opening conducted by Hassanli et al. [16] are used for CFD validation. The building model was made of Perspex and had dimensions of B × B × H = 0.04 × 0.04 × 0.12 m^3. A through-building converging-diverging ducted opening was placed at 75% of the building height with the inlet and outlet openings on the windward and leeward facades, respectively (Fig. 5). The experiment was performed in a closed-loop wind tunnel with 20 m long and a cross-section of 2 × 3 m^2. The wind flow was parallel to the axis of the duct. Two-component Particle Image Velocimetry (PIV) was used to measure the time-averaged streamwise and vertical wind velocity components at a vertical plane inside the ducted opening.

For the CFD simulation, a computational model of the reduced-scale building model is made based on the best practice guidelines (e.g., Refs. [35,36,47]). The same computational setting and parameters as in Section 4.3 are used. Fig. 6a and b shows the measured and simulated dimensionless streamwise wind velocity in the PIV measurement plane (x/B = 0.825) along horizontal and vertical lines, respectively. A fairly good agreement between wind tunnel and CFD is achieved. The average absolute deviation from the experimental data for the streamwise wind velocity along the horizontal and vertical lines are about 6%, and 7%, respectively. Further information on possible reasons for the deviations can be found in Ref. [16].

Note that the selection of this wind-tunnel experiment was due to the lack of available high-quality experimental data of converging-diverging building-integrated ducted openings, the same as the one used in the present study. Nevertheless, the essential flow features and the level of flow complexity around the building and inside the ducted opening are relatively similar. Therefore, it can reasonably be assumed that the given combination of computational parameters and settings that provided accurate simulation results in the validation study will also provide accurate results for the full-scale buildings.

5. Metamodeling

A metamodel of computer simulations, represented as $y = f(x)$, can be given as $y = \hat{y} + e_g$, where e_g is the generalization error [48], i.e., the difference between the real output results directly obtained from the simulations (y) and those predicted by the metamodel ($\hat{y}$) [49]. Metamodels involve a large variety of statistical models, which can be categorized as parametric metamodels, non-parametric metamodels, and ensembles of parametric and non-parametric metamodels (e.g., Refs. [26,50]).

Response Surface Methodology (RSM), as a parametric metamodel, is widely used for different engineering applications. However, it is not basically suitable for approximating non-linear output results. In parametric metamodels, the input-output relationship is assumed to be known in the form of a low-degree polynomial where the samples are normally distributed, i.e., the deviation from the metamodel is assumed almost the same for all samples (Fig. 7a). Therefore, prediction based on parametric metamodels requires estimating a finite number of coefficients associated with polynomial regression models, and prediction highly depends on the selected polynomial [49,51]. In this study, a second-order polynomial is used, which provides information about the linear and quadratic interactions of the input parameters.

Kriging (KG), Neural Network (NN), and Support Vector Regression (SVR) are non-parametric metamodels (Fig. 7b–d). Unlike parametric metamodels, non-parametric metamodels are relatively flexible as prediction depends on the training dataset [50]. They are especially suitable in computer experiments where the input-output relationship is usually highly non-linear [52]. KG, also known as spatial correlation

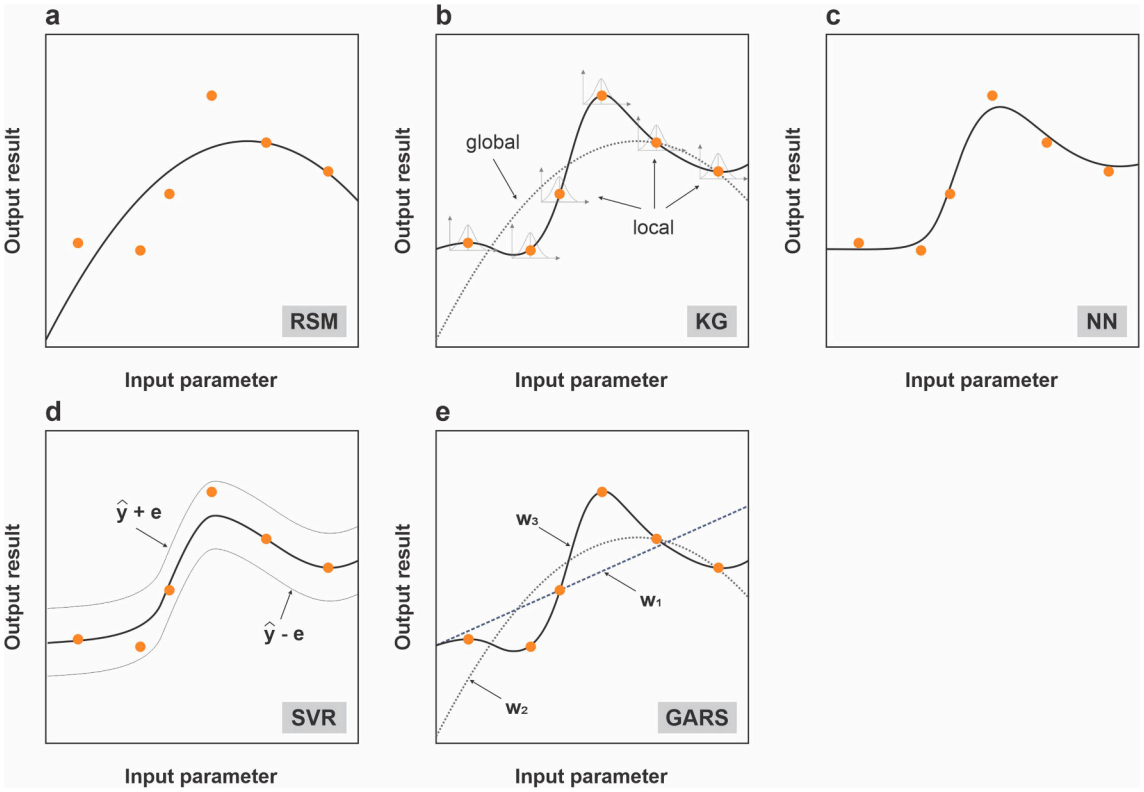

Fig. 7. Schematic of (a) parametric metamodels, (b–d) non-parametric metamodels, and (e) ensemble of three metamodels involving two parametric and one non-parametric in which the impacts on the predicted output results are based on weight factors (i.e., w_1, w_2, and w_3). Circles indicate samples.

modeling, is useful for predicting spatially correlated data. Among non-parametric metamodels, KG is one of the most flexible metamodels as it interpolates the training dataset and can be developed using a wide range of correlation functions. It consists of a global metamodel, usually a parametric metamodel, and local deviations [53,54]. In this study, the KG metamodel uses the Gaussian correlation function, which selects one correlation coefficient for each input parameter. NN mimics the functions and architecture of the nervous system (neurons). It has three layers of input, hidden and output to form a network where the hidden layer(s) connects the input and output layers. Each layer consists of several neurons, and the metamodel is developed by weighted connections between the neurons in the different layers. Each neuron in a hidden layer receives the output values of the neurons in the previous layer as its input value and transfers this value to the neuron(s) in the next hidden layer or output layer [55]. A transfer function, also known as the activation function, is used to map the output value of each neuron in the previous layer to the input value of the neurons in the current layer [56,57]. In this study, the NN architecture is as follows: an input layer with five neurons, each representing each of the five input parameters, a hidden layer with five neurons, and an output layer with one neuron representing the output result, i.e., the maximum amplification factor in the mid-throat. One hidden layer is used in line with recommendations by Ref. [58]. A sensitivity analysis is also performed to investigate the impact of the number of neurons in the hidden layer (N_n) on the accuracy of NN. The analysis is performed for ten cases with $N_n = 1$–10 based on in-sample and out-of-sample evaluations. The results show that $N_n = 5$ leads to the minimum in-sample (8%) and out-of-sample (15%) errors. Therefore, five neurons are considered in the hidden layer. In addition, the hyperbolic tangent is used as the transfer function for the hidden and output layers. SVR, also known as non-parametric regression, is a data classification method that finds the input-output relationship with an error margin (e) in a way that the predicted output results ($\hat{y}$) do not deviate from the real output results more than the error margin. SVR ignores those samples that create the maximum gradient in the input-output relationship, leading to a significant reduction in the computational time (e.g., Refs. [59,60]). In this study, the linear e-insensitive loss function is used to characterize the error margin.

Genetic Aggregation Response Surface (GARS) is an ensemble of parametric and non-parametric metamodels and consequently takes advantage of all involved metamodels. It employs a Genetic algorithm to generate a population of parametric and non-parametric metamodels (e.g., Polynomial Regression, KG, and SVR). Fig. 7e schematically shows an ensemble that involves two parametric and one non-parametric metamodels. An ensemble can be defined as $\hat{y}_{ens} = \sum_{i=1}^{n} w_i \hat{y}_i$ where n is the number of involved metamodels, w_i is the weight coefficients for each metamodel, and $\hat{y}_i$ demonstrates the individual metamodels. The weight coefficients satisfy $\sum_{n}^{i=1} w_i = 1$ and need to be determined such that the accuracy of the ensemble is maximized [61]. Detailed information about different techniques to determine weight coefficients (also known as weight factors) of involved metamodels can be found in Refs. [26,62]. Further information about the developed GARS in the present study will be provided in Section 6.1.

6. Results

6.1. Metamodel accuracy

It is well known that the size of the training dataset can significantly affect the performance of a metamodel (e.g., Refs. [63,64]). Therefore, care is required in the selection of proper dataset size, which is highly correlated to the architecture of the metamodel, the number of input parameters, the complexity of the input-output relationship, etc. (e.g., Ref. [65]). In this study, a sensitivity analysis is performed based on

Table 2
DOE methods and number of generated samples for seven training dataset sizes.

Step	DOE method	No. of generated samples	Dataset size
1	Max. Min. Distance Space-Filling Design	27	27
2	Centered L2 Space-Filling Design	27	54
3	Maximum Entropy Space-Filling Design	27	81
4	Latin Hypercube Design	27	108
5	Max. Min. Distance Space-Filling Design	27	135
6	Latin Hypercube Design	27	162
7	Central Composite Design	27	189

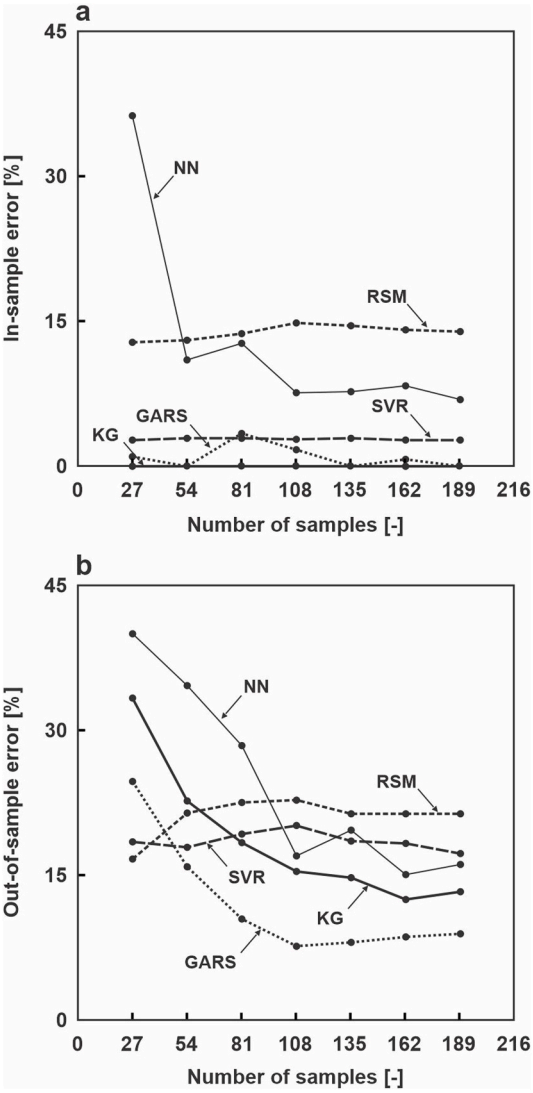

Fig. 8. Sensitivity analysis for impact of dataset size on (a) in-sample and (b) out-of-sample errors for five metamodels.

in-sample and out-of-sample errors to systematically investigate the impact of the training dataset size on the accuracy of the metamodels. Seven different design of experiments (DOE) methods are used to gradually generate unique samples in seven steps where each DOE method adequately covers the design space. In each step, new samples are added to the existing samples generated in the previous step(s). The size of each step is the same as the number of samples for the Central Composite Design DOE method for the five input parameters, i.e., $2^{(k-1)} + 2k + 1 = 2^4 + 2 \times 5 + 1 = 27$. Note that it could be possible to

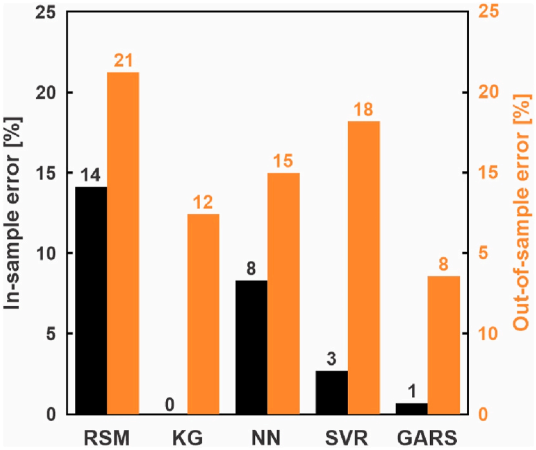

Fig. 9. In-sample and out-of-sample error for different metamodels using 162 samples (step 6 in Table 2).

generate samples using a given space-filling method with a larger number of samples. However, this could fail in full coverage of the design space at each step, leading to a poor generalization. Table 2 presents the different DOE methods used for the seven steps.

For the in-sample evaluation, the k-fold cross-validation is performed [66,67]. The training dataset is split into k = 10 sets of samples (folds) of which k-1 folds are used as the training set. The last fold is used as the testing set to validate the metamodel based on the square root of the average squared difference between the output results directly obtained from CFD simulations (AF_{CFD}) and those predicted by the metamodel ($AF_{metamodel}$), i.e., root mean square error (RMSE), as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(AF_{CFD} - AF_{metamodel})^2}{n}} \quad (5)$$

where n is the number of samples in the testing set. This process is repeated k times in which each fold is used once as the testing set. The in-sample error is computed using Eq. (6):

$$\text{in-sample error} = \frac{\sum_{i=1}^{k} RMSE_i}{k} \quad (6)$$

where $RMSE_i$ represents the RMSE obtained for each fold, and k is the number of folds.

For the out-of-sample evaluation, the Max. Min. Distance Space-Filling Design method and the Latin Hypercube Design method are used to generate 22 testing samples. For each of the methods, the testing samples are generated using the pure quadratic model sampling for the five input parameters, i.e., $2k + 1 = 2 \times 5 + 1 = 11$. The out-of-sample error is computed using Eq. (5) where n = 22, AF_{CFD} and $AF_{metamodel}$ are the output result directly obtained from CFD simulations, and predicted by the metamodel for the testing samples, respectively.

Fig. 8 presents the in-sample and out-of-sample errors for the seven dataset sizes. The most sensitive in-sample error to the training dataset size is observed for NN (Fig. 8a). In this case, by increasing the number of samples from 27 (step 1) to 54 (step 2) and 189 (step 7), the in-sample error reduces by 70% and 81%, respectively. In addition, for NN, KG, and GARS, a significant sensitivity of the out-of-sample error to the training dataset size can be seen for the rather small number of samples, i.e., steps 1–4 (Fig. 8b). A closer look at Fig. 8 reveals that 162 samples (step 6) can be considered sufficiently large for all the metamodels as a further increase in the size of the training dataset does not significantly reduce the in-sample and out-of-sample errors for the five metamodels. Therefore, the dataset with 162 samples is considered the training dataset for the remaining study.

Fig. 9 shows the in-sample and out-of-sample errors for the training

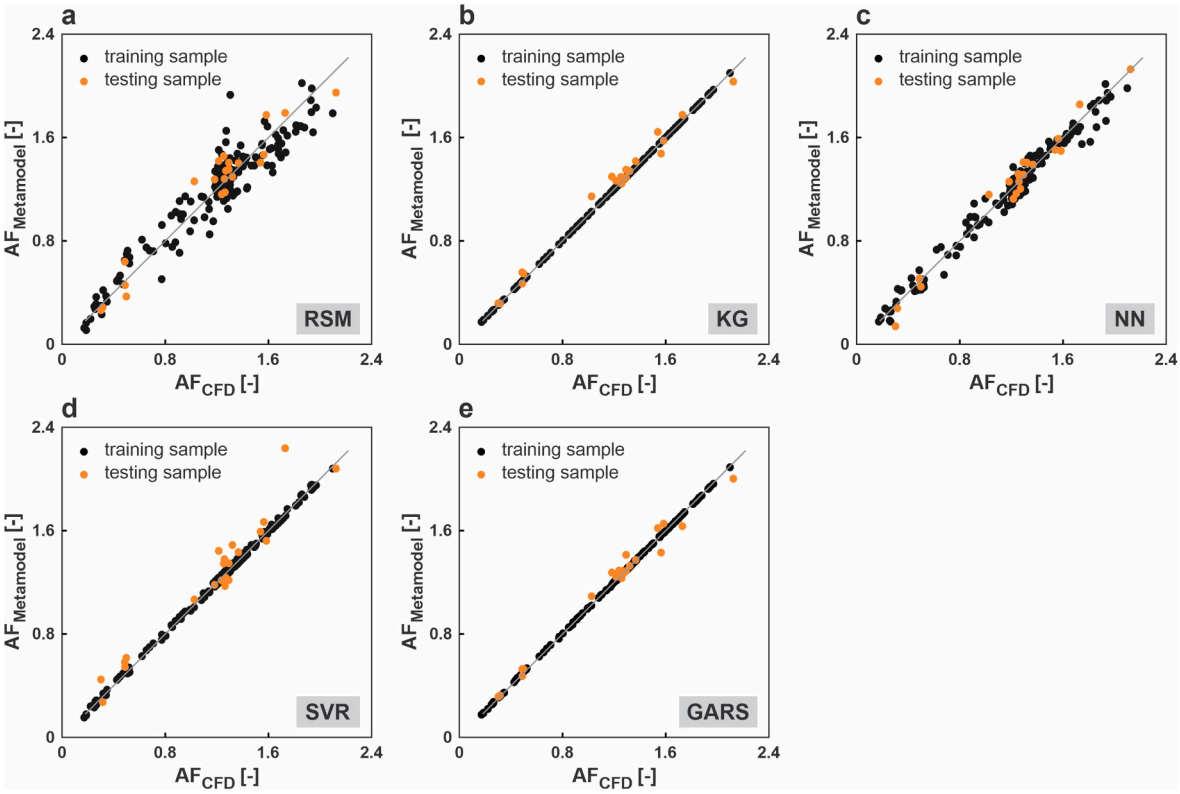

Fig. 10. Maximum amplification factor predicted by metamodels ($AF_{metamodel}$) versus values directly obtained using CFD simulations (AF_{CFD}) for training and testing datasets for different metamodels.

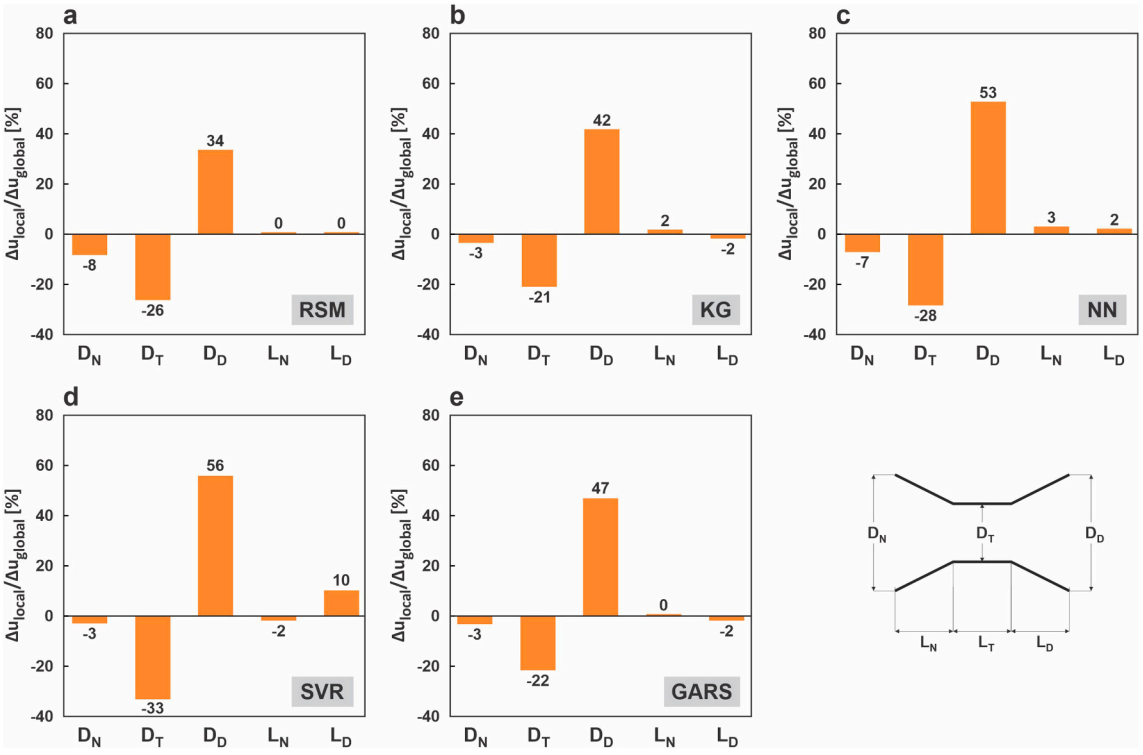

Fig. 11. Parameter importance measurement for different metamodels.

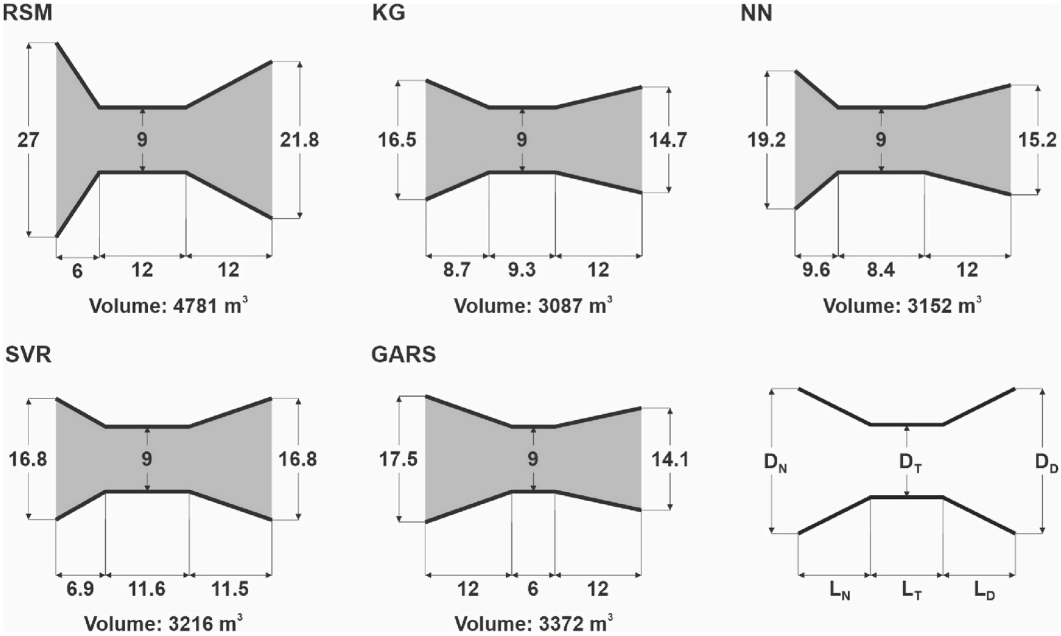

Fig. 12. Optimum designs obtained by different metamodels (dimensions are in meters). Wind flow is from left to right.

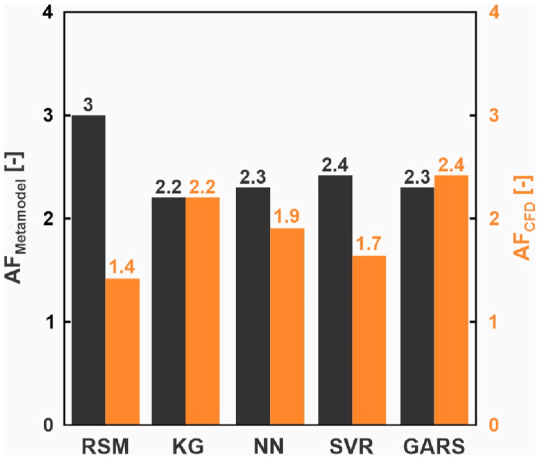

Fig. 13. Maximized amplification factor in mid-throat obtained from metamodels ($AF_{metamodel}$) and directly obtained by CFD simulations (AF_{CFD}) for different metamodels.

dataset with 162 samples (step 6) for the five metamodels. Fig. 10 shows the maximum amplification factors predicted by the metamodels ($AF_{metamodel}$) for 162 training and 22 testing samples versus those directly obtained by CFD simulations (AF_{CFD}). The following observations can be made:

- For the given dataset, the use of different metamodels can affect the out-of-sample error by up to 61%.
- The largest in-sample and out-of-sample errors are achieved for RSM (Fig. 9). For this metamodel, relatively large deviations between the $AF_{metamodel}$ and AF_{CFD} for the training and testing datasets are also observed (see Fig. 10a), indicating that RSM fails to accurately capture the complex input-output relationship.
- For SVR, a relatively large difference between the in-sample and out-of-sample errors can be seen, which indicates overfitting, i.e., the metamodel is perfectly developed over the training dataset while it is incapable of accurately predicting the testing dataset, thus it lacks generalization [68] (Fig. 9).
- The in-sample and out-of-sample errors associated with GARS are relatively small (Fig. 10e). For the dataset with 162 samples, it involves a polynomial regression and two KG metamodels with linear and quadratic correlation functions. The corresponding weights for the involved metamodels are 2%, 51%, and 47%, respectively.

6.2. Parameter importance measurement

Parameter importance measurement is performed to assess the relative importance of each input parameter in the prediction of the output result (e.g., Refs. [64,65]). It is done by varying an input parameter from its minimum to maximum values while the other input parameters are kept fixed at a given value. The following steps are taken to determine the importance of each input parameter [69,70]:

1. for i = 1 to m do
2. select one input parameter ($Var_{(i)}$)
3. assume constant values for the other input parameters ($Var_{(1,i-1)}$)
4. vary value of $Var_{(i)}$ from minimum to maximum
5. calculate output result ($AF_{metamodel}$) for each value of $Var_{(i)}$
6. determine the sensitivity of $AF_{metamodel}$ to each input parameter ($\Delta u_{local}/\Delta u_{global}$)
7. end do

where m is the number of input parameters, Δu_{local} is the difference between the largest and smallest $AF_{metamodel}$ when one input parameter varies, and Δu_{global} is the difference between the largest and smallest $AF_{metamodel}$ when all input parameters vary.

In the present study, for all the metamodels, the importance of each input parameter is determined based on the maximum amplification factor in the mid-throat ($AF_{metamodel}$). The results are shown in Fig. 11. Note that the negative percentages denote that the input parameters are reversely proportional to the output result. The following observations can be drawn:

- All the metamodels identify the diameter of the diffuser outlet (D_D) as the most important input parameter in predicting $AF_{metamodel}$. For example, for the most and the least accurate metamodels, i.e., GARS and RSM, the importance of D_D is 47% and 34%, respectively. Note that in converging-diverging ducts, D_D can significantly affect the

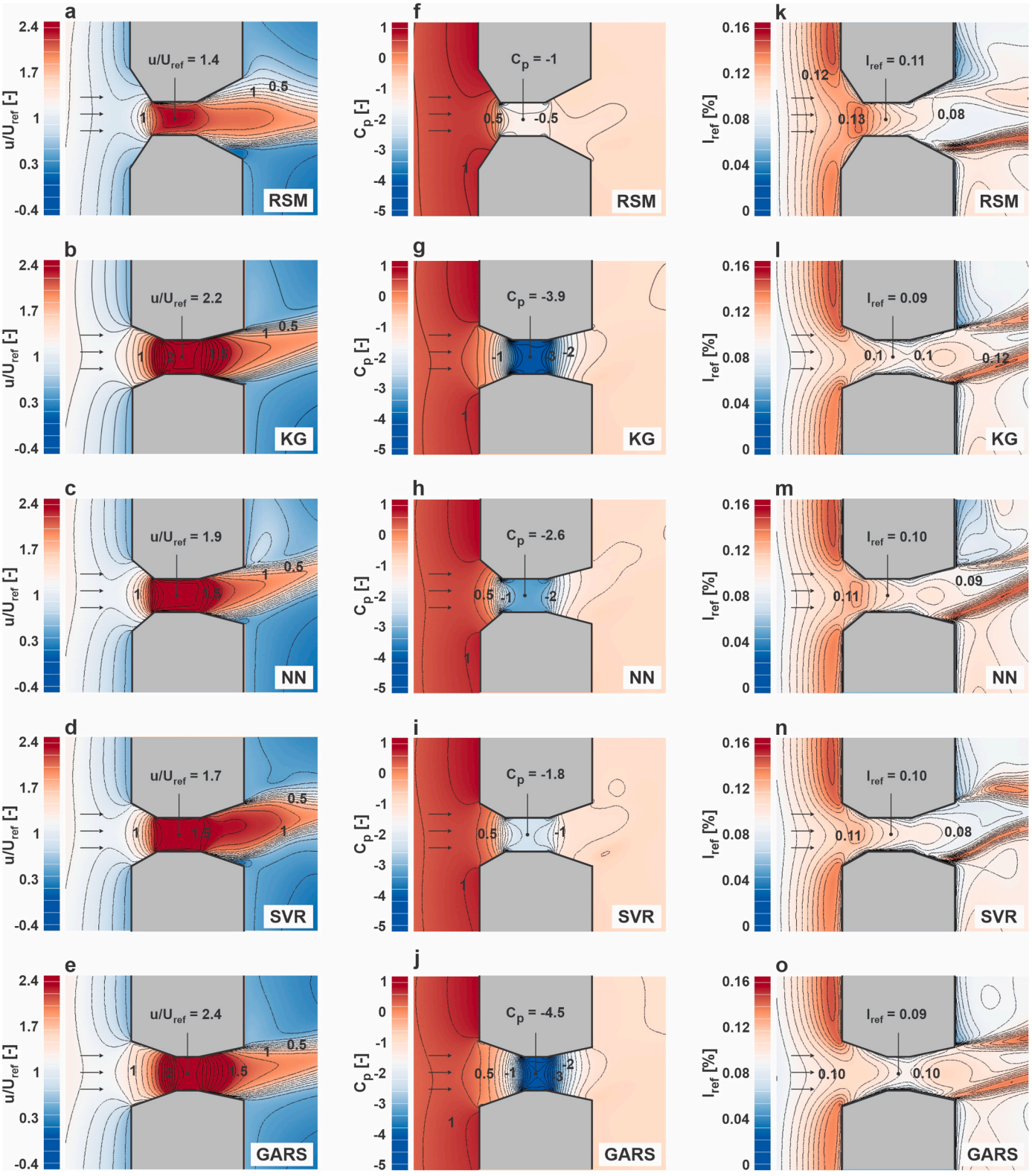

Fig. 14. Contour plots of dimensionless streamwise velocity, static pressure coefficient, and reference turbulence intensity in symmetrical plane for optimized duct geometry obtained by different metamodels. Wind flow is from left to right.

positive pressure gradient inside the diffuser (e.g., Ref. [19]), which could lead to the separation of the boundary layer from the walls and greater pressure losses [71].

- The diameter of the throat (D_T) is recognized as the second most important input parameter for all the metamodels. For example, the importance of D_T is −22% and −26% for GARS and RSM, respectively.
- All metamodels show insignificant sensitivity of the $AF_{metamodel}$ to the diameter of the nozzle inlet (D_N), the length of the nozzle (L_N), and the length of the diffuser (L_D). For GARS, for example, the

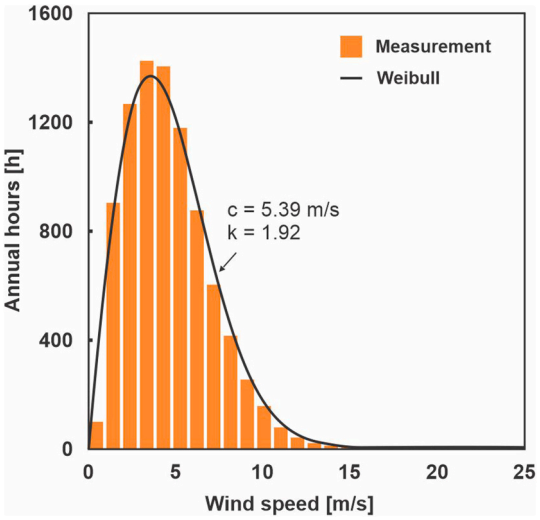

Fig. 15. Average annual mean wind speed frequency distribution for the city of Rotterdam, the Netherlands (2011-2020), obtained from Royal Netherlands Meteorological Institute (KNMI).

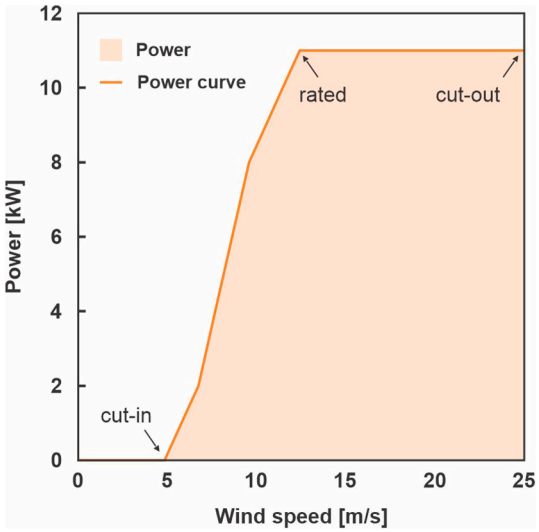

Fig. 16. Power curve of selected wind turbine (Hummer H8. 10 kW).

Table 3
Geometrical and operational characteristics of selected wind turbine (Hummer H8. 10 kW).

Type	Horizontal axis wind turbine (HAWT)
Power	
Rated power [kW]	10
Cut-in wind speed [m/s]	5
Rated wind speed [m/s]	11
Cut-out wind speed [m/s]	25
Survival wind speed [m/s]	50
Rotor	
Number of blades [−]	3
Diameter [m]	8.2
Swept area [m^2]	50.3
Rotor max. speed [U/min]	100
Tip speed [m/s]	43
Material	Fiberglass reinforced composite
Gear box	
Type	None-direct drive
Generator	
Type	Permanent magnet alternator (SCF technology)
Number [−]	1
Generator max. speed [U/min]	100
Voltage [v]	400

importance of these three parameters is about -3%, 0%, and -2%, respectively.

It should be noted that the significance of each parameter in this analysis depends on the following factors:

- The range of input parameters. For example, in this study, the outlet diameter of the diffuser can vary between 9 and 27 m, restricted with regard to the building dimensions. In contrast, the length of the diffuser varies in a more limited range, i.e., 6–12 m. The use of different ranges would affect the obtained significance for the parameters.
- The fixed values for the other parameters. In this study, the importance of each parameter is determined at the mean value of the other input parameters computed as the sum of the maximum and minimum values, presented in Table 1, divided by two. If other values were selected, the outcome would be different.

While the results of this analysis are case-specific, they can give insights into the relative importance of different input parameters on the output results.

6.3. Optimum design

Fig. 12 shows the optimum designs for the five metamodels. Fig. 13 presents the maximum amplification factors in the mid-throat obtained by the metamodels ($AF_{metamodel}$) and those directly obtained by CFD simulation (AF_{CFD}) for the optimum designs. The following observations can be made:

- The use of the different metamodels can lead to different optimum designs with a 36% variation in the $AF_{metamodel}$ where the maximum and minimum $AF_{metamodel}$ are achieved for RSM ($AF_{metamodel} = 3$) and KG ($AF_{metamodel} = 2.2$), respectively.
- The maximum and minimum diameter of the diffuser outlet (D_D) is predicted by the least and the most accurate metamodels, i.e., RSM ($D_D/D_T = 2.42$) and GARS ($D_D/D_T = 1.56$), respectively. Note that the pressure inside the diffuser scales in proportion to the D_D, and consequently the D_D can significantly affect the streamwise flow velocity in the duct [72].
- The minimum and maximum volume of the ducted opening is obtained for KG (3087 m^3) and RSM (4781 m^3), respectively. It should be noted that smaller ducted openings are usually preferred to reduce structural and wind loads (e.g., Refs. [71,73]).
- The deviation between the $AF_{metamodel}$ and the AF_{CFD} is more pronounced for the less accurate metamodels. This deviation is 114% and 41% for RSM and SVR, respectively. The main reason for the large deviations is the underfitting occured for RSM and overfitting occured for SVR. This deviation, however, is 0% for KG and 4% for GARS for which relatively small in-sample and out-of-sample errors are obtained (see Fig. 9).

In order to better understand the impact of the optimized geometries obtained by the five metamodels on the flow features, the following parameters are investigated in detail in the middle plane:

1. Dimensionless streamwise velocity (u/U_{ref}) where $U_{ref} = 4.19$ m/s is the streamwise velocity in the undisturbed flow at the height of the centerline of the ducted opening (h = 105 m).
2. Static pressure coefficient (C_p) computed as Eq. (7):

$$C_p = \frac{P - P_0}{0.5\rho U_{ref}^2} \tag{7}$$

Table 4
Available wind power and wind energy production for optimum design obtained by KG ($AF_{metamodel} = 2.2$) using annual mean wind speed frequency distribution of Rotterdam.

bin	$U_{(10)}$ [m/s]	$U_{(105)}$ [m/s]	$U_{(105)} \times AF_{metamodel}$ [m/s]	hr_i [h]	$(0.5\rho A u_i^3)$ [kW/h]	Annual available wind power per bin [kW]	P_i [kW/h]	Annual energy production per bin [kW]
1	0	0	0	100	0	0	0	0
2	1	1.4	3.1	904	0.5	461	0	0
3	2	2.8	6.2	1267	4.1	5182	2.4	3041
4	3	4.2	9.3	1425	13.7	19579	8.3	11827
5	4	5.6	12.4	1404	32.7	45883	11	15444
6	5	7.0	15.4	1179	63.7	75114	11	12969
7	6	8.4	18.5	876	109.9	96316	11	9636
8	7	9.8	21.6	604	175.2	105803	11	6644
9	8	11.2	24.7	416	261.1	108638	11	4576
10	9	12.6	27.8	254	372.3	94562	11	2794
11	10	14.0	30.9	157	510.2	80100	11	1727
12	11	15.4	33.9	79	678.5	53603	11	869
13	12	16.8	37.1	43	881.7	37915	11	473
14	13	18.2	40.1	22	1120.3	24647	11	242
15	14	19.7	43.2	12	1400.3	16804	11	132
16	15	21.0	46.3	7	1721.3	12049	11	77
17	16	22.5	49.4	2	2087.9	4176	11	22
18	17	23.9	52.5	0	2506.1	0	11	0
19	18	25.3	55.6	0	2973.5	0	11	0
20	19	26.7	58.7	0	3501.0	0	11	0
21	20	28.1	61.8	0	4081.5	0	11	0
						$P_{available}$=780832		**AEP =70472**

where P is the mean static pressure, $P_0 = 0$ Pa is the reference static pressure and $\rho = 1.225$ kg/m^3 is the air density.

3. Reference turbulence intensity based on the International Electrotechnical Commission (IEC) Standard 61400-1 (I_{ref}) [74]. The IEC Standard 61400-1 defines different classes of wind turbines for which a limitation on the maximum turbulence intensity is determined to control diverse effects of fluctuating loads. For high (Class A), medium (Class B), and low (Class C) turbulence intensity, I_{ref} = 0.16, 0.14, and 0.12, respectively. The I_{ref} refers to the expected value of hub-height turbulence intensity at a 10-min average wind speed of 15 m/s, derived by Eq. (8):

$$I_{ref} = \frac{TI}{(0.75 + \frac{5.6}{U_{ref}})} \tag{8}$$

where TI is the turbulence intensity and is calculated using Eq. (9):

$$TI = \frac{\sqrt{\frac{2}{3}k}}{U} \tag{9}$$

where k is turbulence kinetic energy directly obtained from RANS simulations, and U is the mean velocity. Note that the lower value of I_{ref} denotes the lower fluctuating loads that is necessary within the lifetime of a turbine.

The results are presented in Fig. 14. Accordingly, low turbulence intensity inside the duct is achieved based on the IEC Standard 61400-1 (Fig. 14k-o) for all the optimum designs. For all the metamodels, the turbulence level inside the duct is less than the value corresponding to $I_{ref} = 0.12$.

6.4. Wind energy performance

In order to give insight into the impact of different metamodels on the wind energy performance of the ducted openings, two parameters are computed: (i) The annual available wind power ($P_{available}$) in the mid-throat of the optimized ducted openings obtained by the five metamodels, and (ii) the annual energy production (AEP) for a given horizontal-axis wind turbine installed in the mid-throat of the five optimized ducted openings. For this assessment, it is assumed that the building is located in the city of Rotterdam, the Netherlands, and the corresponding mean wind speed distribution at 10 m height ($u_{(10)}$) is obtained from the Royal Netherlands Meteorological Institute (KNMI). The data is measured at 10-min intervals over 10 years (2011–2020). Fig. 15 shows the average annual mean wind speed frequency distribution for Rotterdam where the mean wind speed ranges from 0 to 20 m/s with 21 bins of 1 m/s. A Weibull distribution function is also fitted to the wind data to determine the scale (c) and shape (k) parameters where the corresponding values are c = 5.39 m/s and k = 1.92.

The $P_{available}$ can be estimated using Eq. (10):

$$P_{available} = \sum_{i=1}^{n=21} hr_i \left(0.5\rho A u_i^3\right) \tag{10}$$

where n is the number of the mean wind speed bins, hr_i the total hours associated with the i^{th} mean wind speed bin, $\rho = 1.225$ kg/m^3 the air density, A = 28.3 m^2 the cross-sectional area of the throat, and u the mean wind speed in the mid-throat given by Eq. (11):

$$u = u_{(105)} AF_{metamodel} \tag{11}$$

where $u_{(105)}$, wind speed at 105 m height, is estimated using the logarithmic law described by Eq. (1) and $u_{(10)}$ obtained from the KNMI. $AF_{metamodel}$ is the maximum amplification factor in the mid-throat (see Fig. 13).

A small-scale (urban) 3-bladed Horizontal Axis Wind Turbine (HAWT) with a rated power of 10 kW and a rotor diameter of 8.2 m is assumed to be installed in the mid-throat. Table 3 presents the geometrical and operational characteristics of the selected wind turbine. The turbine power curve, showing the turbine output power versus mean wind speed, is used to estimate the AEP using Eq. (12):

$$AEP = \sum_{i=1}^{n=21} hr_i P_i \tag{12}$$

where n is the number of the mean wind speed bins (i.e., 21), hr_i the total hours associated with the i^{th} mean wind speed bin, and P_i the corresponding turbine output power obtained from the turbine power curve developed by the manufacturer (Fig. 16).

Table 4 presents the available wind power and wind energy

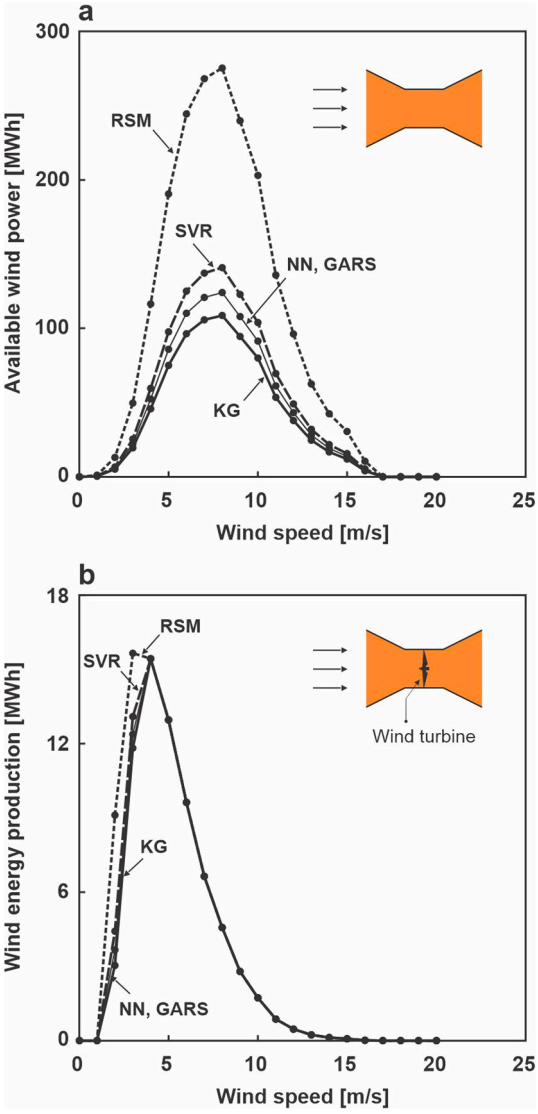

Fig. 17. (a) Available wind power and (b) wind energy production as a function of mean wind speed for each optimum design obtained by different metamodels.

Table 5
$P_{available}$ and AEP for different metamodels.

Metamodel	$P_{available}$ [MW]	AEP [MW]
RSM	1980.0	80.4
KG	780.8	70.5
NN	892.4	71.7
SVR	1013.8	73.1
GARS	892.4	71.7

production as a function of mean wind speed for the optimum design obtained by KG with $AF_{metamodel} = 2.2$. Fig. 17a and b shows the available wind power and the wind energy production as a function of mean wind speed for the optimum designs obtained by the five metamodels, respectively. Table 5 shows the $P_{available}$ and AEP estimated by the five metamodels. The following observations can be made:

- The use of the different metamodels can lead to large variations in the predicted $P_{available}$ (~153%) and AEP (~14%).
- The $P_{available}$ and AEP predicted by RSM and SVR is considerably larger than those of the other metamodels. This overestimation is correlated to the deficiencies of RSM and SVR in accurately predicting the $AF_{metamodel}$ (see Fig. 13). Therefore, the $P_{available}$ and AEP values obtained for the least accurate metamodels, i.e., RSM and SVR, should be treated with caution.

7. Discussion

The main limitations of this study need to be mentioned:

- A wide range of functions and hyper-parameters are usually available to develop a metamodel. These functions and hyper-parameters that have to be set by the user can significantly affect the accuracy of metamodels. For example, different correlation functions can be used to develop Kriging (KG), e.g., triangular, exponential, Gaussian, cubic, etc. [75]. Future research can focus on the impact of different functions and hyper-parameters on metamodel-based optimization of wind energy systems.
- In this study, the performance of the metamodels is evaluated based on in-sample and out-of-sample evaluations. Other criteria, such as robustness and interpretability, can also provide insight into the performance of the metamodels and can be investigated in future studies (e.g., Ref. [27]).
- In this study, the performance of different metamodels is evaluated for the given design of experiments (DOE) and optimization methods. However, earlier studies have shown that using different sampling (e.g., DOE) and/or optimization methods can considerably affect the prediction accuracy but also the optimization results. In addition, the combined impacts of different methods in each step can be explored.
- In this study, the only evaluated parametric metamodel is a second-order polynomial. However, earlier studies have shown that the metamodel accuracy can considerably increase by using higher-order polynomials [76] while they are computationally more expensive. It is of interest to further evaluate higher-order polynomials to investigate the accuracy of parametric metamodels against non-parametric metamodels for complex non-linear applications.
- This study focuses on a single-objective multivariable optimization. Nevertheless, the insights provided in the present study can still be useful in the multi-objective optimization of complex computer experiments. Special attention should be given to the following items:
  (i) Earlier studies have shown that increasing the number of objectives can provide further challenges to reaching the global optima [48]. This issue can be addressed by converting a multi-objective optimization to a single-objective optimization where proportions are identified for objectives, and their preferences are encoded as weights.
  (ii) The impact of different hyper-parameters, such as the population size (the number of generated geometries in each period), the number of iterations, the crossover and mutation rates, should be taken into account, as suggested in Refs. [34,77].
  (iii) Hybrid algorithms, i.e., combinations of more than two optimization algorithms, are expected to leverage the individual algorithms (e.g., Ref. [78]). For example, a hybrid of Genetic algorithm and Particle Swarm Optimization performs more efficiently than each of the two algorithms [79].

8. Conclusion

Metamodel-based optimization is performed to improve the aerodynamic performance of a building-integrated ducted opening. The performance of five commonly-used metamodels is evaluated by (i) detailed in-sample and out-of-sample evaluations of the metamodels, (ii) annual available power in the wind ($P_{available}$), and (iii) annual energy

production (AEP). The following conclusions are made:

- The use of different metamodels for a given dataset can affect the out-of-sample error by up to 61%.
- The second-order polynomial is not suitable for complex computer experiments, such as CFD simulations. In this case, increasing the dataset size rarely improves the metamodel accuracy.
- For a small dataset, crude yet still acceptable accuracy can be achieved for Genetic Aggregation Response Surface (GARS) and Kriging (KG) at a very low computational time.
- Non-parametric metamodels, such as Neural Network (NN) and Support Vector Regression (SVR), rely on the dataset size. Therefore, these metamodels should be treated with caution when the dataset size is limited.
- Among the studied geometrical characteristics for the building-integrated ducted opening, the diameter of the diffuser (D_D) and the diameter of the throat (D_T) are found to significantly impact the maximum amplification factor in the mid-throat ($AF_{metamodel}$) for all the metamodels while the diameter of the nozzle (D_N), the length of the nozzle (L_N), and the length of the diffuser (L_D) have a marginal impact.
- For the optimum designs obtained by different metamodels, the $AF_{metamodel}$ values experience a variation of about 36%.
- Using different metamodels can lead to significant variations in the predicted annual available power in the wind ($P_{available}$) and annual energy production (AEP) by up to 153% and 14%, respectively.

Credit authors statement

Zeynab Kaseb: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - Original Draft, Visualization. **Hamid Montazeri:** Conceptualization, Methodology, Writing - Review & Editing, Supervision, Project administration.

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

No data was used for the research described in the article.

Acknowledgement

The authors gratefully acknowledge the partnership with ANSYS CFD. The authors also acknowledge Nedpower SWH company, represented by ing. R. Kroezemann, for helpful discussions.

References

[1] El-Khattam W, Salama MMA. Distributed generation technologies, definitions and benefits. Elec Power Syst Res 2004;71:119–28. https://doi.org/10.1016/j.epsr.2004.01.006.
[2] Rezaeiha A, Montazeri H, Blocken B. A framework for preliminary large-scale urban wind energy potential assessment: roof-mounted wind turbines. Energy Convers Manag 2020;214:112770. https://doi.org/10.1016/j.enconman.2020.112770.
[3] Gusmão Caiado RG, Leal Filho W, Quelhas OLG, Luiz de Mattos Nascimento D, Ávila LV. A literature-based review on potentials and constraints in the implementation of the sustainable development goals. J Clean Prod 2018;198:1276–88. https://doi.org/10.1016/j.jclepro.2018.07.102.
[4] Zhang L, Gari N, Hmurcik LV. Energy management in a microgrid with distributed energy resources. Energy Convers Manag 2014;78:297–305. https://doi.org/10.1016/j.enconman.2013.10.065.
[5] Tasneem Z, Al Noman A, Das SK, Saha DK, Islam MR, Ali MF, R Badal MF, Ahamed MH, Moyeen SI, Alam F. An analytical review on the evaluation of wind resource and wind turbine for urban application: prospect and challenges, Developments in the Built Environment, vol. 4; 2020, 100033. https://doi.org/10.1016/j.dibe.2020.100033.
[6] Alanis Ruiz C, Kalkman I, Blocken B. Aerodynamic design optimization of ducted openings through high-rise buildings for wind energy harvesting. Build Environ 2021;202:108028. https://doi.org/10.1016/j.buildenv.2021.108028.
[7] Beller C. Urban wind energy - state of the art 2009. 2009.
[8] Nfaoui H. Wind energy potential. In: Comprehensive renewable energy. Elsevier; 2012. p. 73–92. https://doi.org/10.1016/B978-0-08-087872-0.00204-3.
[9] Ishugah TF, Li Y, Wang RZ, Kiplagat JK. Advances in wind energy resource exploitation in urban environment: a review. Renew Sustain Energy Rev 2014;37:613–26. https://doi.org/10.1016/j.rser.2014.05.053.
[10] Dilimulati A, Stathopoulos T, Paraschivoiu M. Wind turbine designs for urban applications: a case study of shrouded diffuser casing for turbines. J Wind Eng Ind Aerod 2018;175:179–92. https://doi.org/10.1016/j.jweia.2018.01.003.
[11] Kc A, Whale J, Urmee T. Urban wind conditions and small wind turbines in the built environment: a review. Renew Energy 2019;131:268–83. https://doi.org/10.1016/j.renene.2018.07.050.
[12] Sahebzadeh S, Rezaeiha A, Montazeri H. Towards optimal layout design of vertical-axis wind-turbine farms: double rotor arrangements. Energy Convers Manag 2020;226:113527. https://doi.org/10.1016/j.enconman.2020.113527.
[13] Sahebzadeh S, Rezaeiha A, Montazeri H. Vertical-axis wind-turbine farm design: impact of rotor setting and relative arrangement on aerodynamic performance of double rotor arrays. Energy Rep 2022;8:5793–819. https://doi.org/10.1016/j.egyr.2022.04.030.
[14] Hyams MA. Wind energy in the built environment, metropolitan sustainability: understanding and improving the urban environment. 2012. p. 457–99. https://doi.org/10.1533/9780857096463.3.457.
[15] Blanch MJ. Wind energy technologies for use in the built environment. Wind Eng 2002;26:125–43. https://doi.org/10.1260/030952402762056054.
[16] Hassanli S, Chauhan K, Zhao M, Kwok KCS. Application of through-building openings for wind energy harvesting in built environment. J Wind Eng Ind Aerod 2019;184:445–55. https://doi.org/10.1016/j.jweia.2018.11.030.
[17] Nunes MM, Brasil Junior ACP, Oliveira TF. Systematic review of diffuser-augmented horizontal-axis turbines. Renew Sustain Energy Rev 2020;133:110075. https://doi.org/10.1016/j.rser.2020.110075.
[18] Khamlaj TA, Rumpfkeil MP. Theoretical analysis of shrouded horizontal axis wind turbines. Energies 2017;10:1–19. https://doi.org/10.3390/en10010038.
[19] Sorribes-Palmer F, Sanz-Andres A, Ayuso L, Sant R, Franchini S. Mixed CFD-1D wind turbine diffuser design optimization. Renew Energy 2017;105:386–99. https://doi.org/10.1016/j.renene.2016.12.065.
[20] Stathopoulos T, Alrawashdeh H, Al-Quraan A, Blocken B, Dilimulati A, Paraschivoiu M, Pilay P. Urban wind energy: some views on potential and challenges. J Wind Eng Ind Aerod 2018;179:146–57. https://doi.org/10.1016/j.jweia.2018.05.018.
[21] Alsailani M, Montazeri H, Rezaeiha A. Towards optimal aerodynamic design of wind catchers: impact of geometrical characteristics. Renew Energy 2021;168:1344–63. https://doi.org/10.1016/j.renene.2020.12.053.
[22] Thévenin D, Janiga G. Optimization and computational fluid dynamics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. https://doi.org/10.1007/978-3-540-72153-6.
[23] Wang GG, Shan S. Review of metamodeling techniques in support of engineering design optimization. J Mech Des 2007;129:370–80. https://doi.org/10.1115/1.2429697.
[24] Box GEP, Draper N. Empirical model building and Response surfaces. first ed. New York: John Wiley & Sons; 1987.
[25] Fang K-T, Li R, Sudjianto A. Design and modeling for computer experiments. first ed. Chapman & Hall/CRC; 2006.
[26] Jiang P, Zhou Q, Shao X. Surrogate model-based engineering design and optimization. first ed. Singapore: Springer Singapore; 2020. https://doi.org/10.1007/978-981-15-0731-1.
[27] Li YF, Ng SH, Xie M, Goh TN. A systematic comparison of metamodeling techniques for simulation optimization in Decision Support Systems. Appl Soft Comput 2010;10:1257–73. https://doi.org/10.1016/j.asoc.2009.11.034.
[28] Simpson T, Mistree F, Korte J, Mauery T. Comparison of response surface and kriging models for multidisciplinary design optimization. In: 7th AIAA/USAF/NASA/ISSMO symposium on multidisciplinary analysis and optimization. Reston, Virigina, Virigina: American Institute of Aeronautics and Astronautics; 1998. p. 381–91. https://doi.org/10.2514/6.1998-4755.
[29] Paiva R, Crawford C, Suleman A, Carvalho A. A comparison of surrogate models in the framework of an MDO tool for wing design. In: 50th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics, and materials conference. Reston, Virigina: American Institute of Aeronautics and Astronautics; 2009. https://doi.org/10.2514/6.2009-2203.
[30] Montazeri H, Blocken B. Extension of generalized forced convective heat transfer coefficient expressions for isolated buildings taking into account oblique wind directions. Build Environ 2018;140:194–208. https://doi.org/10.1016/j.buildenv.2018.05.027.
[31] Joseph VR. Space-filling designs for computer experiments: a review. Qual Eng 2016;28:28–35. https://doi.org/10.1080/08982112.2015.1100447.
[32] Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments. Stat Sci 1989;4:409–23. https://doi.org/10.1214/ss/1177012413.
[33] Townsend A. Genetic Algorithms_a tutorial. Citeseer; 2003.
[34] Haupt RL. Optimum population size and mutation rate for a simple real genetic algorithm that optimizes array factors. In: IEEE antennas and propagation society international symposium. Transmitting waves of progress to the next millennium. 2000 digest. Held in conjunction with: USNC/URSI national radio science meeting

(cat. No.00CH37118). IEEE; 2000. p. 1034–7. https://doi.org/10.1109/APS.2000.875398.
[35] Tominaga Y, Mochida A, Yoshie R, Kataoka H, Nozu T, Yoshikawa M, Shirasawa T. AIJ guidelines for practical applications of CFD to pedestrian wind environment around buildings. J Wind Eng Ind Aerod 2008;96:1749–61. https://doi.org/10.1016/j.jweia.2008.02.058.
[36] Franke J, Hellsten A, Schlunzen KH, Carissimo B. The COST 732 Best Practice Guideline for CFD simulation of flows in the urban environment: a summary. Int J Environ Pollut 2011;44:419. https://doi.org/10.1504/IJEP.2011.038443.
[37] Tucker PG, Mosquera A. NAFEMS introduction to grid & mesh generation for CFD. NAFEMS CFD Working Group; 2001.
[38] Casey M, Wintergerste T. Quality and trust in industrial CFD - best practice guidelines. Brussels: ERCOFTAC; 2000.
[39] Roache PJ. Quantification of uncertainty in computational fluid dynamics. Annu Rev Fluid Mech 1997;29:123–60. https://doi.org/10.1146/annurev.fluid.29.1.123.
[40] Richards P, Hoxey R. Appropriate boundary conditions for computational wind engineering models using the k-ε turbulence model. J Wind Eng Ind Aerod 1993;46–47:145–53. https://doi.org/10.1016/0167-6105(93)90124-7.
[41] Wieringa J. Updating the Davenport roughness classification. J Wind Eng Ind Aerod 1992;41:357–68. https://doi.org/10.1016/0167-6105(92)90434-C.
[42] Launder BE, Spalding DB. The numerical computation of turbulent flows. Comput Methods Appl Mech Eng 1974;3:269–89. https://doi.org/10.1016/0045-7825(74)90029-2.
[43] Cebeci T, Bradshaw P. Momentum transfer in boundary layers. Washington DC: Hemisphere Publishing/McGraw-Hill; 1977. p. 319–21.
[44] Blocken B, Stathopoulos T, Carmeliet J. CFD simulation of the atmospheric boundary layer: wall function problems. Atmos Environ 2007;41:238–52. https://doi.org/10.1016/j.atmosenv.2006.08.019.
[45] Yakhot V, Orszag SA, Thangam S, Gatski TB, Speziale CG. Development of turbulence models for shear flows by a double expansion technique. Phys Fluid Fluid Dynam 1992;4:1510–20. https://doi.org/10.1063/1.858424.
[46] Choudhury D. Introduction to the renormalization Group method and turbulence modeling. Lebanon: Fluent Inc; 1993.
[47] Tamura T, Nozawa K, Kondo K. AIJ guide for numerical prediction of wind loads on buildings. J Wind Eng Ind Aerod 2008;96:1974–84.
[48] Kochenderfer MJ, Wheeler TA. Algorithms for optimization. first ed. The MIT Press; 2019. https://mitpress.mit.edu/books/algorithms-optimization.
[49] Simpson TW, Peplinski JD, Koch PN, Allen JK. On the use of statistics in design and the implications for deterministic computer experiments. In: Proceedings of DETC'97 1997 ASME design engineering technical conferences; 1997. p. 1–14. California.
[50] Kianifar MR, Campean F. Performance evaluation of metamodelling methods for engineering problems: towards a practitioner guide. Struct Multidiscip Optim 2020;61:159–86. https://doi.org/10.1007/s00158-019-02352-1.
[51] Myers RH, Montgomery DC, Anderson-Cook CM. Response surface methodology: process and product optimization using designed experiments. fourth ed. Wiley; 2016.
[52] Whittaker E. On a new method of graduation. Edinburgh Mathematical Society; 1923. p. 63–75.
[53] Cressie N. The origins of kriging. Math Geol 1990;22:239–52. https://doi.org/10.1007/BF00889887.
[54] Krige DG. A statistical approach to some basic mine valuation problems on the Witwatersrand. J Chem Metall Min Soc S Afr 1951;52:119–39.
[55] Zhang Y, Zhu R, Chen Z, Gao J, Xia D. Evaluating and selecting features via information theoretic lower bounds of feature inner correlations for high-dimensional data. Eur J Oper Res 2021;290:235–47. https://doi.org/10.1016/j.ejor.2020.09.028.
[56] Gatti C. Design of experiments for reinforcement learning. Cham: Springer International Publishing; 2015. https://doi.org/10.1007/978-3-319-12197-0.
[57] Hagan MT, Demuth HB, Beale MH. Neural network design. Boston: PWS Publishing; 1996.
[58] Stathakis D. How many hidden layers and nodes? Int J Rem Sens 2009;30:2133–47. https://doi.org/10.1080/01431160802549278.
[59] Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004;14:199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.
[60] Cortes C, Vapnik V. Support vector networks. Mach Learn 1995;20:273–97. https://doi.org/10.1007/BF00994018.
[61] Wang S, Jian G, Xiao J, Wen J, Zhang Z. Optimization investigation on configuration parameters of spiral-wound heat exchanger using Genetic Aggregation response surface and Multi-Objective Genetic Algorithm. Appl Therm Eng 2017;119:603–9. https://doi.org/10.1016/j.applthermaleng.2017.03.100.
[62] Acar E. Various approaches for constructing an ensemble of metamodels using local measures. Struct Multidiscip Optim 2010;42:879–96. https://doi.org/10.1007/s00158-010-0520-z.
[63] Haussler D. Decision theoretic generalizations of the PAC model for neural net and other learning applications. Inf Comput 1992;100:78–150.
[64] Villa-Vialaneix N, Follador M, Ratto M, Leip A. A comparison of eight metamodeling techniques for the simulation of N 2O fluxes and N leaching from corn crops. Environ Model Software 2012;34:51–66. https://doi.org/10.1016/j.envsoft.2011.05.003.
[65] Østergård T, Jensen RL, Maagaard SE. A comparison of six metamodeling techniques applied to building performance simulations. Appl Energy 2018;211:89–103. https://doi.org/10.1016/j.apenergy.2017.10.102.
[66] Bhavsar P, Safro I, Bouaynaya N, Polikar R, Dera D. Machine learning in transportation data analytics. In: Data analytics for intelligent transportation systems. Elsevier; 2017. p. 283–307. https://doi.org/10.1016/B978-0-12-809715-1.00012-2.
[67] Viana FAC, Haftka RT, Steffen V. Multiple surrogates: how cross-validation errors can help us to obtain the best predictor. Struct Multidiscip Optim 2009;39:439–57. https://doi.org/10.1007/s00158-008-0338-0.
[68] Harrell FE. Regression modeling strategies. New York, NY: Springer New York; 2001. https://doi.org/10.1007/978-1-4757-3462-1.
[69] Latunde T, Bamigbola OM. Parameter estimation and sensitivity analysis of an optimal control model for capital asset management. Advances in Fuzzy Systems 2018:1–11. https://doi.org/10.1155/2018/4756520 (2018).
[70] Lenhart T, Eckhardt K, Fohrer N, Frede H-G. Comparison of two different approaches of sensitivity analysis. Phys Chem Earth, Parts A/B/C 2002;27:645–54. https://doi.org/10.1016/S1474-7065(02)00049-9.
[71] Bekele N, Bogale W. Parametric study of a diffuser for horizontal axis wind turbine power augmentation. AIMS Energy 2019;7:841–56. https://doi.org/10.3934/energy.2019.6.841.
[72] Idelchik IE. Handbook of hydraulic resistance. J Pressure Vessel Technol 1987;109:260–1. https://doi.org/10.1115/1.3264907. second ed.
[73] Leloudas SN, Lygidakis GN, Eskantar AI, Nikolos IK. A robust methodology for the design optimization of diffuser augmented wind turbine shrouds. Renew Energy 2020;150:722–42. https://doi.org/10.1016/j.renene.2019.12.098.
[74] IEC 61400-1. Wind turbines Part 1 - design requirements for small wind turbines. International Electrotechnical Commision; 2014.
[75] Xie W, Nelson B, Staum J. The influence of correlation functions on stochastic kriging metamodels. In: Proceedings of the 2010 winter simulation conference. IEEE; 2010. p. 1067–78. https://doi.org/10.1109/WSC.2010.5679083.
[76] Friedman LW, Pressman I. The metamodel in simulation analysis: can it Be trusted? J Oper Res Soc 1988;39:939–48. https://doi.org/10.1057/jors.1988.160.
[77] Kaseb Z, Hafezi M, Tahbaz M, Delfani S. A framework for pedestrian-level wind conditions improvement in urban areas: CFD simulation and optimization. Build Environ 2020;184:107191. https://doi.org/10.1016/j.buildenv.2020.107191.
[78] Lu Y, Benlic U, Wu Q. A highly effective hybrid evolutionary algorithm for the covering salesman problem. Inf Sci 2021;564:144–62. https://doi.org/10.1016/j.ins.2021.02.053.
[79] Kaseb Z, Rahbar M. Towards CFD-based optimization of urban wind conditions: comparison of Genetic algorithm, Particle Swarm Optimization, and a hybrid algorithm. Sustain Cities Soc 2022;77:103565. https://doi.org/10.1016/j.scs.2021.103565.