Explainability analysis of neural network-based turbulence modeling for transonic axial compressor rotor flows

Chutian Wu [a,b], Shizhao Wang [a,b], Xin-Lei Zhang [a,b,*], Guowei He [a,b,*]

[a] *State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing, 100190, China*
[b] *School of Engineering Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

ARTICLE INFO

ABSTRACT

Model–consistent training has become trending for data-driven turbulence modeling since it can improve model generalizability and reduce data requirements by involving the Reynolds–averaged Navier–Stokes (RANS) equation during model learning. Neural networks are often used for the Reynolds stress representation due to their great expressive power, while they lack interpretability for the causal relationship between model inputs and outputs. Some post–hoc methods have been used to explain the neural network by indicating input feature importance. However, for the model–consistent training, the model explainability involves the analysis of both the neural network inputs and outputs. That is, the effects of model output on the RANS predictions should also be explained in addition to the input feature analysis. In this work, we investigate the explainability of the model–consistent learned model for the internal flow prediction of NASA Rotor 37 at its peak efficiency operating condition. The neural–network–based corrections for the Spalart–Allmaras turbulence model are learned from various experimental data based on the ensemble Kalman method. The learned model can noticeably improve the velocity prediction near the shroud. The explainability of the trained neural network is analyzed in terms of the model correction and the input feature importance. Specifically, the learned model correction increases the local turbulence production in the vortex breakdown region due to non–equilibrium effects, which capture the blockage effects near the shroud. Besides, the ratio of production to destruction and the helicity are shown to have relatively high importance for accurately predicting the compressor rotor flows based on the Shapley additive explanations method.

© 2023 Elsevier Masson SAS. All rights reserved.

1. Introduction

The compressor is one of the crucial components in an aircraft engine, concerning the overall performance of the aero-engine. Accurate prediction of the internal compressor flows can provide prior assessments of the overall efficiency and further fuel consumptions of the aero-engine [1], which is of significant interest for the design of advanced engines. The Reynolds–averaged Navier–Stokes (RANS) method is the most widely used numerical approach for predicting compressor flows due to its computational efficiency. However, the accuracy of the RANS predictions in compressor flows is highly affected by the ability of turbulence models. Therefore, it is necessary to develop accurate turbulence models for predictions of compressor flows.

Over the past decades, various turbulence models [2–6], e.g., Spalart–Allmaras (SA) model and k-ε model, have been evaluated for compressor flows such as the NASA rotor 37 case [7–9], which is one canonical case for flows through a transonic compressor rotor. It has been recognized that the conventional eddy viscosity models often lead to significant discrepancies for mean flow separation and curvature [10]. Especially for compressor flows, the occurring boundary layer transition, shock waves, tip leakage vortices, and their interactions pose additional challenges for developing accurate turbulence models. Some efforts have been devoted to improving the RANS predictions of compressor flows by considering these physical phenomena. For instance, the modification with helicity and pressure gradient has been introduced in the SA model [11] to improve predictions of transonic axial compressor rotor flows, mainly in overall performance, e.g., the total pressure ratio. Further improvement is still worthy of investigation, particularly for accurate flow field predictions.

Machine learning methods [12–14] are increasingly used for developing turbulence models due to their capability of learning complex functional relationships from data. The symbolic expres-

sion [15,16], the tensor–based neural network [17], and the random forest [18,19] are proposed to represent the Reynolds stress. However, such models are often trained directly with the Reynolds stress data, which is difficult to obtain for compressor flows due to the enormous computational cost [20]. On the one hand, the compressor flows of high Reynolds numbers exhibit small spatial and temporal scales of turbulence requiring a high computational cost for resolution [20]. On the other hand, complex characteristics of turbomachinery flow [21,22,1,23,24], such as the blade-to-blade interaction, row-to-row interaction, laminar-turbulent transition, surge, and stall, pose additional computational burdens. The mesh refinement is required near blades, end-walls, and tip gaps, as well as a comprehensive 360-degree computational domain when necessary [25], to capture these flow phenomena. Moreover, the model is often trained in the *a priori* manner, i.e., independent of the CFD solvers. This can result in significant discrepancies when coupled with the RANS solver due to the inconsistency between the training and the prediction environments [26]. In view of these difficulties, the model–consistent training strategy [26,27] is proposed to learn data-driven turbulence models from indirect observation data such as velocity. It is achieved by involving the RANS solvers during the training process. Numerous methods have been proposed to perform the model–consistent learning including the ensemble Kalman method [28,29]. This method uses an ensemble of samples to approximate gradient and Hessian information to perform second-order optimization implicitly. Hence given its good training efficiency, it is practical for complex configurations [30,31] where the CFD calculation is computationally expensive.

The model–consistent training can improve the predictive ability of neural-network-based turbulence models, while the learned models still lack explainability. Some post–hoc methods such as the Shapley additive explanations (SHAP) analysis [32] have been introduced to indicate feature importance for neural-network-based turbulence models [33] that are learned with the *a priori* training. However, the model–consistent training involves the neural network and the RANS equations simultaneously. Hence, in addition to the input feature analysis of the neural network, predictive improvement in the RANS calculation should also be explained based on the learned model correction. This is in contrast to the *a priori* training method that only needs to investigate the effects of input features on the neural network outputs.

This work focuses on the explainability of data-driven turbulence modeling for the transonic flow of the axial compressor rotor. The neural network-based correction is introduced in the SA turbulence model with selected input features, including helicity. The ensemble Kalman method is used to learn the turbulence model from various observation data, such as the velocity measurements. The effects of the neural network model on flow prediction are investigated from the perspectives of the model inputs and outputs. Specifically, the explainability analysis of the input features is performed through the SHAP method. The pure and interaction effects of the input features on the output are investigated, highlighting the importance of the production–to–destruction ratio and the helicity for accurate predictions of rotor flows. On the other hand, the effect of the learned model correction field on flow prediction is also analyzed, which improves the velocity prediction by capturing the non–equilibrium effects in the vortex breakdown region.

The rest of the paper is outlined as follows. The ensemble Kalman-based turbulence modeling for rotor flows is introduced in Section 2. The test cases and corresponding results are presented and analyzed in Section 3. The explainability analysis of the learned model is discussed in Section 4. Finally, this paper is concluded in Section 5.

2. Framework of turbulence modeling based on ensemble Kalman method

2.1. RANS equations

The Reynolds–averaged Navier–Stokes equations for compressible flows in a rotating reference frame can be written as

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u_j)}{\partial x_j} = 0, \tag{1a}$$

$$\frac{\partial (\rho u_i)}{\partial t} + \frac{\partial \left(\rho u_i u_j + p \delta_{ij} \right)}{\partial x_j} = \frac{\partial}{\partial x_j} (\sigma_{ij} + \tau_{ij}) + \rho f_i, \tag{1b}$$

$$\frac{\partial (\rho E)}{\partial t} + \frac{\partial \left[(\rho E + p) u_j \right]}{\partial x_j} = \frac{\partial}{\partial x_j} \left[(\sigma_{ij} + \tau_{ij}) u_i - q_j - q_j^t \right]. \tag{1c}$$

In the formula above, u_i is the relative velocity, and the Coriolis force and centrifugal force are represented by $f_i = -2\varepsilon_{ijk}\Omega_j u_k + \frac{\partial}{\partial x_i}\left(\Omega^2 r^2/2\right)$, with $\Omega = \sqrt{\Omega_i \Omega_i}$ being the angular velocity magnitude, r is the radius to the rotating axis, and ε_{ijk} is the permutation tensor. The specific total energy E is defined as $E = \frac{1}{\gamma-1}p/\rho + \frac{1}{2}\left(u_i u_i - \Omega^2 r^2\right)$, where $\gamma = 1.4$ is the specific heat ratio of air. The equation of state $p = \rho R T$ relates the density ρ, the static pressure p, and the static temperature T, where $R = 287.03\,\mathrm{m}^2\mathrm{s}^2\mathrm{K}$ is the gas constant. The viscous stress tensor σ_{ij} and the heat flux q_j are defined as

$$\sigma_{ij} = \mu \left(2S_{ij} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij} \right), \quad q_j = -\kappa \frac{\partial T}{\partial x_j}. \tag{2}$$

The mean strain rate tensor is defined as

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right). \tag{3}$$

The thermal conductivity κ can be expressed as $\kappa = \mu c_p / Pr$ with Prandtl number $Pr = 0.72$ and the specific heat at constant pressure $c_p = \frac{\gamma}{\gamma-1}R$. The dynamic viscosity μ depends on the temperature described by Sutherland's law. In this work, the Reynolds stress τ_{ij} and turbulent heat flux q_j^t are expressed based on the Boussinesq hypothesis by

$$\tau_{ij} = \mu_t \left(2S_{ij} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij} \right), \quad q_j^t = -\frac{\mu_t c_p}{Pr_t}\frac{\partial T}{\partial x_j}, \tag{4}$$

where $Pr_t = 0.9$ is the turbulent Prandtl number and μ_t is the eddy viscosity estimated by the turbulence model.

The one-equation Spalart–Allmaras (SA) turbulence model [34] is commonly used to simulate the compressor rotor flows due to its high computational efficiency. The eddy viscosity in the SA model is estimated by

$$\mu_t = \rho \tilde{\nu} f_{v1}, \tag{5}$$

where f_{v1} is an intermediate function, and the working variable $\tilde{\nu}$ is solved by the transport equation as

$$\frac{\partial \tilde{\nu}}{\partial t} + u_j \frac{\partial \tilde{\nu}}{\partial x_j} = \mathcal{P} - \mathcal{D} + \mathcal{T}. \tag{6}$$

The right-hand terms of the equation above are production, destruction, and diffusion, respectively. Their detailed formulas are presented in Appendix A. The working variable $\tilde{\nu}$ transport equation is derived based on the equilibrium turbulence of the attached boundary layer flows [33], which often leads to significant predictive discrepancies for flow with strong non-equilibrium effects. To address this issue, a correction coefficient β is multiplied on the production term of the SA model as [35]

$$\frac{\partial \tilde{v}}{\partial t} + u_j \frac{\partial \tilde{v}}{\partial x_j} = \beta \mathcal{P} - \mathcal{D} + \mathcal{T}. \tag{7}$$

By doing so, the correction function β with respect to selected input features can be learned by minimizing the discrepancy between the model prediction and experimental data. The introduced coefficient β is dimensionless and can offer an appropriate initial value, i.e., $\beta = 1$, from the baseline SA model. The underlying mechanism of this correction is to consider the non-equilibrium effect as the conventional model assumes fixed ratios between the production and dissipation [36]. The introduction of $\beta(\boldsymbol{x})$ is equivalent to adding an extra source term $(\beta(\boldsymbol{x}) - 1)\mathcal{P}$, which changes the entire balance of the model, instead of merely a modification of the production term [37]. Note that one can also add corrections in the destruction terms, which have been conducted in Ref. [38]. Previous works [39,40,35,41,42,37,43] have demonstrated the efficacy of this corrective method to improve the accuracy of the RANS predictions. Hence we adopt this framework to enhance the turbulence models for transonic compressor flows.

2.2. Neural network-based turbulence model

Neural networks are often used to represent the Reynolds stress due to their great expressive power. We use a feedforward neural network to construct the correction β in the SA turbulence model. The neural network comprises 5 input features, i.e.,

$$q = \left\{ \frac{\mathcal{P}}{\mathcal{D}}, \quad \frac{\|\boldsymbol{S}\|}{\|\Omega\|}, \quad \delta, \quad \chi, \quad h \right\}. \tag{8}$$

The first four features are suggested and chosen by Holland et al. [44], while the fifth feature, i.e., helicity h, is added based on the practice of turbulence modeling [45] for compressor flows. Specifically, the feature q_1 is the ratio of production to destruction in the SA model, which indicates the non-equilibrium effects. Feature q_2 is the ratio of strain rate magnitude to vorticity magnitude, which measure the relative importance of shear strain and rotation. Feature q_3 is an adverse pressure gradient indicator as

$$\delta = \frac{2}{3} \frac{\mu_t \|\boldsymbol{S}\|}{\tau_w}, \tag{9}$$

which measures the ratio of the local shear stress to the shear stress at the nearest wall. Medida et al. [46] used this indicator to modify the SA turbulence model and showed predictive improvement for the flow over 2D airfoils at high angles of attack. Feature q_4 is the ratio of the working variable $\tilde{v}$ to molecular viscosity v. Feature q_5 is the local helicity h, defined by

$$h = \frac{\boldsymbol{u} \cdot \boldsymbol{\omega}}{\|\boldsymbol{u}\| \|\boldsymbol{\omega}\|}, \tag{10}$$

which indicates the cosine value of the angle between velocity and vorticity. Note that the helicity is not Galilean invariant since it involves a velocity vector. However, helicity is associated with the linkages and knottedness of vortex lines within fluid flows [47], which are typical flow structures in nonlinear energy transfer processes, i.e., energy backscatter. Moreover, it has been introduced to improve the SA model by considering such phenomena in compressor blade cascade [45] and compressor rotor [47]. Therefore, helicity is chosen as one of the input features in this work. The input features are summarized in Table 1. These raw input features are mapped to a Gaussian distribution by $q = F_q^{-1}(F_z(z))$, where z is the raw input feature, $F_z(z)$ is the cumulative distribution function of variable z, and $F_q(q)$ is the standard accumulative distribution. With the Gaussian scaling, all input features are brought in a similar distribution with zero mean and variance of one. The training results without the input feature of helicity

Table 1
Input features of neural network.

Input	Physical significance	Description
q_1	$\mathcal{P}/\mathcal{D}$	ratio of production to destruction
q_2	$\|\boldsymbol{S}\|/\|\Omega\|$	ratio of strain magnitude to vorticity magnitude
q_3	δ	adverse pressure gradient indicator
q_4	χ	ratio of the SA variable to the kinematic viscosity
q_5	h	consistency of direction between velocity and vorticity

and production-to-destruction ratio are presented in Appendix B, which demonstrate the necessity of introducing the feature of helicity and production-to-destruction ratio.

2.3. Training neural-network using ensemble Kalman method

The ensemble Kalman method is adopted in this work to train the neural network weights due to its high efficiency and ease of implementation. Specifically, the ensemble Kalman method incorporates low-rank approximated Hessian information, leading to a second-order optimization. Moreover, the method is non-intrusive and derivative-free, without requiring extra effort in code redevelopment for CFD solvers.

The ensemble Kalman method is a statistical inference technique based on Monte Carlo sampling. The method draws random samples of weights and further uses the covariance of the samples to estimate the gradient of objective functions. The objective function to be minimized is defined as

$$J = \|w^{i+1} - w^i\|_{\mathsf{P}^i} + \|\mathcal{H}[w^{i+1}] - y\|_{\mathsf{R}^i},$$

where w is the neural network weight, P is the sample covariance, $\mathcal{H}$ represents the model operator that propagates the neural network weight to the observation quantity, R is the observation error covariance, and y is the observation data with random noise that conforms to a Gaussian distribution with mean zero and covariance matrix R. The method employs an ensemble of realizations to estimate sample statistics. The sample mean $\bar{\mathsf{W}}$ and covariance P are represented as

$$\bar{\mathsf{W}}^i = \frac{1}{M} \sum_{j=1}^{M} w_j^i, \tag{11}$$
$$\mathsf{P}^i = \frac{1}{M-1} (\mathsf{W}^i - \bar{\mathsf{W}}^i)(\mathsf{W}^i - \bar{\mathsf{W}}^i)^\top,$$

where M is the sample size, i is the iteration number, j is the sample number, and $\mathsf{W} = \{w_j\}_{j=1}^{M}$ is the ensemble of samples. According to the Gauss-Newton method, the neural network weight requires the gradient and Hessian information for iterative updates. The ensemble Kalman method employs the sample covariance to estimate the gradient and Hessian information [48]. In the i-th iteration step, the update scheme of each sample w_j can be formulated as

$$w_j^{i+1} = w_j^i + \mathsf{P}^i \mathsf{H}^\top (\mathsf{H} \mathsf{P}^i \mathsf{H}^\top + \mathsf{R}^i)^{-1}(y_j - \mathsf{H} w_j^i), \tag{12}$$

where H is the tangent linear observation operator. The readers refer to Ref. [29] for further details of the ensemble-based learning framework. Fig. 1 illustrates the procedure of ensemble Kalman-based turbulence modeling for compressor rotor flows. This procedure consists of the following steps:

(a) The weights of neural networks are pre-trained to generate the baseline value, i.e., $\beta = 1$. Further, initial samples of weights
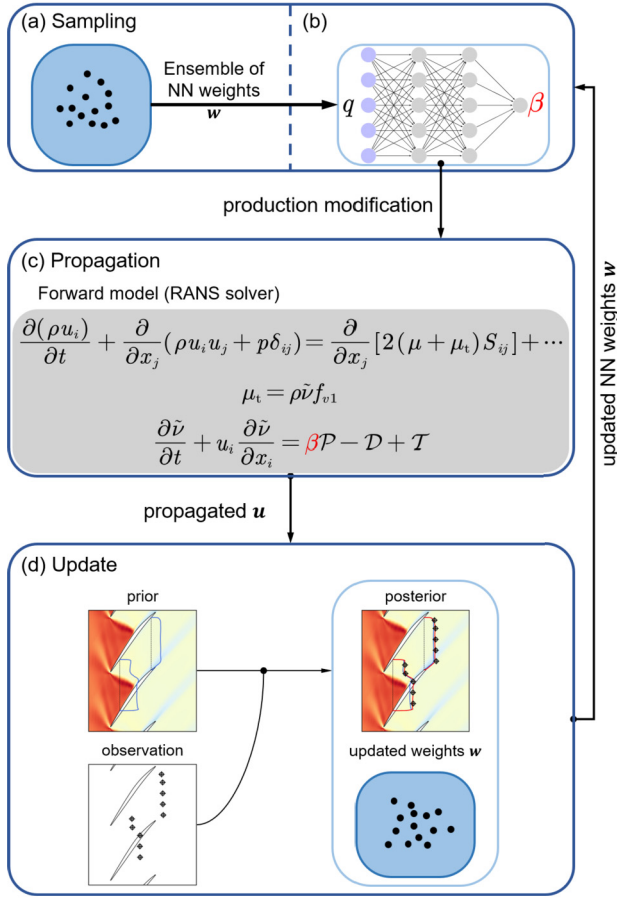
Fig. 1. Framework of the ensemble-based neural network training for the SA turbulence model. The framework consists of the following steps: (a) sampling the weights of the neural network; (b) propagate the model correction β to velocity by solving the RANS equations; (c) update the neural network weights by incorporating observation data.

are drawn around the pre-trained weights based on Gaussian distribution.

(b) The flow features q are extracted based on the predicted flow field, which serves as inputs to the neural network. Further, the inputs are propagated to the model correction β through the neural network.

(c) The predicted quantities, such as the Mach number, the mass flow rate, the overall total pressure ratio, and the overall adiabatic efficiency, are propagated from the model correction β by solving the RANS equations.

(d) The neural network weights are updated based on the ensemble-based statistical analysis of flow prediction and experimental observations.

Steps (b)-(d) are iterated until the data misfit is less than the observation noise level, or the maximum iteration number is reached.

2.4. Explainability for neural-network-based models

Intrinsic methods [49] are difficult to be used for the explainability analysis of neural networks due to high model complexity. A feasible approach is to perform the post–hoc analysis that determines the causal relationship between network inputs and outputs. The Shapley additive explanations (SHAP) method [50] is one widely used post–hoc approach to explain the feature contribution of neural networks. The idea behind the SHAP analysis comes from cooperative game theory. Each feature is treated as a player, and the model prediction is regarded as the payout. The SHAP analysis

aims to compute the contribution, i.e., importance value, of each feature to the prediction.

In our neural network model, the network output β can be expressed as a sum of contributions from each feature, i.e.,

$$\beta = \phi_0 + \sum_{i=1}^{5} \phi_i, \tag{13}$$

where ϕ_i is the contribution of input feature q_i, and ϕ_0 is the null-output of the model, i.e., the mean of outputs.

The SHAP values indicate feature contributions, which are computed by

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} \delta_S(i),$$
$$\delta_S(i) = f_{S \cup \{i\}} \left(x_{S \cup \{i\}} \right) - f_S \left(x_S \right), \tag{14}$$

where F is the set of all features, the summation iterates over all possible subsets S of F excluding q_i noted by $S \subseteq F \setminus \{i\}$, and the operator $|\cdot|$ means the number of elements in a set. The symbol $!$ represents the factorial of a non-negative integer. δ_S represents the marginal contribution of q_i as defined in Eq. (14), where x_S denotes the values of the input features in the set S. In the computation of δ_S, one model $f_{S \cup \{i\}}$ is trained with the feature present, and another model f_S is trained with the feature withheld. The factorial term in front of δ_S assigns each feature with a probability weight, representing the importance of the marginal contribution.

The SHAP interaction value $\phi_{i,j}$ indicates the interaction effect of two features q_i, q_j. It is calculated by [51]

$$\phi_{i,j} = \sum_{S \subseteq F \setminus \{i,j\}} \frac{|S|!(|F| - |S| - 2)!}{2(|F| - 1)!} \delta_S(i, j), \quad i \neq j,$$
$$\delta_S(i, j) = f_{S \cup \{i,j\}} \left(x_{S \cup \{i,j\}} \right) - f_{S \cup \{i\}} \left(x_{S \cup \{i\}} \right) - f_{S \cup \{j\}} \left(x_{S \cup \{j\}} \right)$$
$$+ f_S \left(x_S \right), \tag{15}$$

The pure effect of q_i is then defined as

$$\phi_{i,i} = \phi_i - \sum_{i \neq j} \phi_{i,j} \tag{16}$$

which subtracts all interaction values from the SHAP value.

The training algorithm of the ensemble Kalman method is implemented in the DAFI code [52]. The open–source solver MULTALL [53] is used to perform the CFD simulations for compressor rotor flows. The MULTALL code is a multistage, three-dimensional, steady viscous flow solver based on the time marching finite volume method, and the multigrid method is implemented to speed up calculation. The open-source SHAP library [54] is used to conduct the post–hoc analysis.

3. Training results for flows in NASA rotor 37

In this work, the ensemble Kalman method is used to improve the RANS prediction for internal flows in NASA Rotor 37. This transonic rotor is designed as an inlet stage for an eight-stage 20 : 1 pressure ratio advanced core compressor [7] and has been widely used for investigating compressor rotor flows [55,56]. The aerodynamic design parameters are listed in Table 2. The CFD simulation is performed to predict the flow field for NASA Rotor 37 at the near peak efficiency operating condition, where mass flow is 98% of the choking mass flow ($\dot{m}_{\text{choke}} = 20.93 \, \text{kg/s}$). This is the operating condition close to the design operating condition.

(a)   (b)
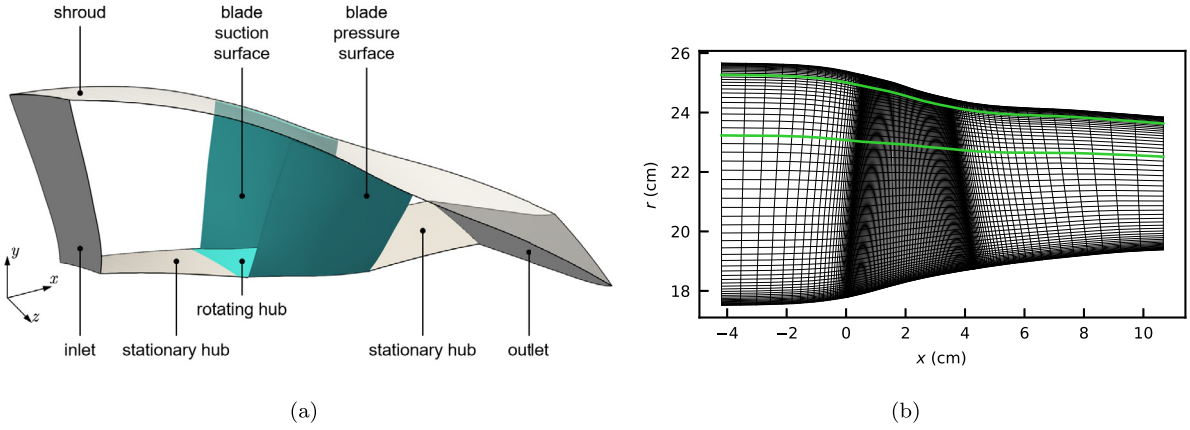
Fig. 2. Geometry of computational domain and mesh. Panel (a) shows the geometry of the computational domain of one blade sector with the inlet, outlet boundary, and walls. The white-colored is stationary, and the green-colored is rotating. Panel (b) shows the H-type mesh in meridional view with green lines highlighting the 70% and 95% span. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Table 2
Aerodynamic design parameters.

Number of blades	36
Tip diameter at leading edge	0.5074 m
Hub diameter at leading edge	0.3576 m
Rotational speed	17 188.7 rpm
Tip speed	454 m/s
Tip clearance	0.356 mm
Total pressure ratio	2.106
Total temperature ratio	1.270
Adiabatic efficiency	0.877
Mass flow rate	20.19 kg/s

3.1. Case set-up

An H-type structured mesh with about 0.51 million cells is used to discretize the computational domain based on our mesh sensitivity study. The numbers of cells in the pitchwise, streamwise, and spanwise directions are 45, 171, and 63, respectively. The near–wall mesh size in the spanwise and pitch-wise directions are scaled in viscous distance $\delta_\nu = 2.0 \times 10^{-6}$ m and set as $\Delta r_1^+ = 55.0$ and $\Delta (r\theta)_1^+ = 42.5$, respectively. Fig. 2 shows the computational domain and the mesh from a meridional view. The locations of the 70% and 95% stream surfaces are highlighted by the green lines, which are used for the following plots. The inlet is located 4.19 cm upstream from the leading edge of the blade section on the hub surface. The inlet boundary condition sets the total pressure and temperature profile measured by experiments along the spanwise direction. The axial velocity is imposed at the inlet. The outlet boundary condition adopts radial equilibrium assumption [57] with the static pressure at the hub. The walls are no-slip and assumed adiabatic. The rotational periodic boundary condition is used since the computational domain covers one blade sector.

The experimental measurements [8] are used as observation data, including the Mach number, the mass flow rate, the overall total pressure ratio, and the overall adiabatic efficiency. The observation of Mach number is located at 70% span and 95% span for fixed axial locations corresponding to 20%, 40%, 65%, 90%, and 104% of rotor chord. The experiment [8] only provides the measurements of the Mach number at these two span positions. We use all the available measurements to ensure the predictive capability of the learned model at both span positions. The data points are uniformly distributed from the suction surface to the pressure surface, with a fixed interval being 0.025 rotor pitch. These locations are selected by Suder [8] to provide the flow field details near the shock/vortex interaction region, upstream and downstream of the shock impingement on the blade suction surface, and near the

blade trailing edge. It is noted that the experimental data at 104% chord is shifted so that the profiled wake is centered in the middle of the plot. The experimental data of Mach number are acquired by laser anemometer system, and the overall performance is acquired by aerodynamic probe measurements [8]. The integrity and reproducibility of experimental data have been validated in Suder's report [8].

The neural network consists of an input layer, two hidden layers, and an output layer. The input layer consists of 5 neurons, where each neuron represents an input feature. Each hidden layer consists of 5 neurons which are activated by the rectified linear (ReLU) function. The output layer consists of 1 neuron, representing the correction coefficient β. The hyper-parameters in neural networks are determined by following the criteria [58] that two hidden layers can represent functions with any kind of shape, and the number of hidden neurons should be between the size of the input layer and the size of the output layer. We also perform a sensitivity study on the neural network architecture as shown in Appendix C, which shows that the selected network architecture provides the best predictive improvement in the Mach number.

We draw 50 samples to train the neural network based on the ensemble Kalman method. The standard deviation of neural network weights is set as 0.6, and the standard deviation of the observation value is set as 0.1 in this work. The number of samples and the standard deviation value is set based on our sensitivity study to achieve a compromise between training accuracy and efficiency. An excessively large value of standard deviation can result in divergence of the CFD simulation, while an excessively small value can lead to slow convergence [59].

3.2. Training results in overall performance

The total pressure and total temperature at the streamwise locations are given by

$$\frac{\overline{p}_i}{p_{\text{ref}}} = \left[\frac{\int_{A_i} \left(\frac{p_{\text{stg}}}{p_{\text{ref}}} \right)^{\frac{\gamma-1}{\gamma}} \rho \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{s}}{\int_{A_i} \rho \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{s}} \right]^{\frac{\gamma}{\gamma-1}} , \quad \overline{T}_i = \frac{\int_{A_i} T_{\text{stg}} \rho \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{s}}{\int_{A_i} \rho \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{s}} , \quad (17)$$

where p_{stg} is the stagnation pressure, p_{ref} is the reference pressure as 101 325 Pa in this work, T_{stg} is the stagnation temperature, $\rho \boldsymbol{u}$ is the momentum flux, and A_i is a cross-channel area. The total pressure ratio π and the adiabatic efficiency η are defined by

Table 3
Overall performance.

	$\dot{m}/\dot{m}_{\text{choke}}$	π	η
Experimental	0.982	2.075	0.879
Baseline	0.996	2.130	0.874
Learned	0.985	2.115	0.865

Table 4
Summary of prediction error in radial total pressure ratio, radial total temperature ratio, radial adiabatic efficiency, and Mach number with the baseline and the learned models.

	baseline	learned
Radial total pressure ratio	2.73%	2.05%
Radial total temperature ratio	0.91%	0.97%
Radial adiabatic efficiency	1.62%	1.98%
Mach number	8.34%	7.28%

$$\pi = \frac{\overline{p}_2}{\overline{p}_1}, \quad \eta = \frac{\left(\frac{\overline{p}_2}{\overline{p}_1}\right)^{\frac{\gamma-1}{\gamma}} - 1}{\frac{\overline{T}_2}{\overline{T}_1} - 1}, \tag{18}$$

where subscripts 1 and 2 indicate the inlet and outlet cross-channel plane, respectively.

Our CFD simulations indicate that both the baseline and learned models can predict total pressure ratio and adiabatic efficiency close to experimental data. Compared to the baseline model, the learned model has better agreement with the experimental data, demonstrating the capability of the present ensemble-based turbulence modeling. Table 3 shows the overall performance of the total pressure ratio and adiabatic efficiency of the NASA Rotor 37 case with a comparison among the baseline model, the learned model, and the experimental data. The learned results show the mean of all samples. The learned model slightly improves the prediction of the flow rate and pressure ratio but deteriorates the prediction of the adiabatic efficiency compared to the baseline model. Such limited improvement in overall performance may be due to data imbalance since there are 330 data points for velocity, while there is only one data point for flow rate, pressure ratio, and adiabatic efficiency, respectively. The weight on the velocity observation should be reduced to further improve the prediction of the overall performance. However, we note that the present work focuses on the explainability of ensemble-based turbulence modeling in improving flow field predictions. The data balance strategies are of significant interest in learning from disparate data but are out of the scope of the present work, which will be investigated in the near future to improve the prediction of the overall performance and flow field simultaneously.

Fig. 3 presents the radial distribution of the total pressure ratio, total temperature ratio, and adiabatic efficiency for the NASA Rotor 37. The learned model provides better predictions in total pressure ratio and Mach number than the baseline model compared to experimental data. Moreover, the total pressure ratio is higher than the experimental values below the 40% span, which is also observed by other RANS or LES simulations [4,60,61]. The error is summarized in Table 4, quantitatively showing the comparable ability of the learned model in predicting the overall performance.

3.3. Training results in Mach number distributions

Fig. 4 compares the Mach number contours of the baseline model, the learned model, and experimental measurements at 95% span. Both the baseline and learned models can capture the vortex flow near the blade tip. Specifically, tip clearance flow passes over the rotor blade tip from the pressure surface to the suction surface over the front part of the blade chord, rolling into a tip vortex [4].

The tip vortex mitigates downstream and interacts with the normal shock wave at the leading edge, which causes the shape of the shock wave to be distorted [62]. The vortex breakdown occurs after the shock region, and the diffused vortexes move to the blade pressure surface and finally merge into the wake downstream [62]. It can be seen from Fig. 4(a)(b) that the learned models predict a more apparent region of low Mach number in the passage near the blade pressure surface side than the baseline model. The low-momentum flow on the pressure surface side indicates the blockage effects [8] due to the interaction between the shock wave and the tip leakage vortex. The difference in the predicted Mach number between the baseline and learned models can be clearly seen in Fig. 4(d). Note that there is a noticeable difference in Mach number near the bow shock wave between the baseline and learned models, likely due to variation of upcoming flow angle. Specifically, the flow angle $\alpha = \arctan(\|\boldsymbol{\Omega} \times \boldsymbol{r}\|/u_x)$ is the angle between relative velocity and x axial direction, where u_x is the relative velocity component in x direction. Hence the flow angle with the learned model differs from the baseline as the learned models predict a slightly different mass flow rate as shown in Table 3, leading to the variation of u_x [63].

Figs. 5(a)(b) show the contours of the Mach number at the 40% chord position. It can be seen that there is a low–speed region close to the shroud wall with both the baseline and learned models. The learned model reduces the Mach number near the shroud compared to the baseline model, as seen in Fig. 5(c), which is formed by the tip vortex rolled into the pressure surface side. Also, the learned model decreases the Mach number around the shock wave from hub to shroud, which indicates that the upcoming flow angle changes in the whole spanwise range.

The velocity profiles at various stations are plotted to better analyze the flow predictions with the learned model. Fig. 6 show the plots of the velocity profiles in Mach number from the suction surface to the pressure surface at different chord locations. The abscissa is normalized with the circumferential arc length at the local spanwise position.

At the 70% spanwise position, both the baseline and the learned models provide similar predictions of the Mach number at different chord positions. At 70% span and 20% chord, the Mach number indicates that the shock impingement on the blade suction surface occurs downstream. From 20% chord to 40% chord, there is a Mach number increment between the passage shock and the pressure surface. This suggests that the flow is accelerating near 40% chord on the pressure surface side of the passage. The Mach number at 65% chord indicates that this location is downstream of the shock impingement point on the blade suction surface. At 90% chord, the Mach number is almost constant across the pitch except near the blade surfaces. In addition, the blade suction surface boundary layers become noticeably thicker from 65% to 90% chord. The Mach number downstream of the blade trailing edge at 104% chord illustrates the width and depth of the rotor wake. At the 104% chord, the width of the wake can be approximated as the sum of the blade thickness at the 90% chord and the boundary layer thicknesses on both the pressure and suction surface sides.

The observation location at 95% span and 40% chord is downstream of the shock-vortex interaction and upstream of the shock impingement on the blade suction surface. It can be seen that the learned model effectively improves the prediction of the Mach number behind the shock wave compared to the baseline. The observation location of 95% span and 65% chord is downstream of the shock impingement on the blade suction surface. Noticeable drops in Mach number can be seen near the mid-pitch, forming two distinct regions with a difference of 0.25 Mach number. This can also be observed from the contour plot in Fig. 4. The learned model significantly improves the velocity prediction in the low-speed region compared to the baseline model. At the observation location
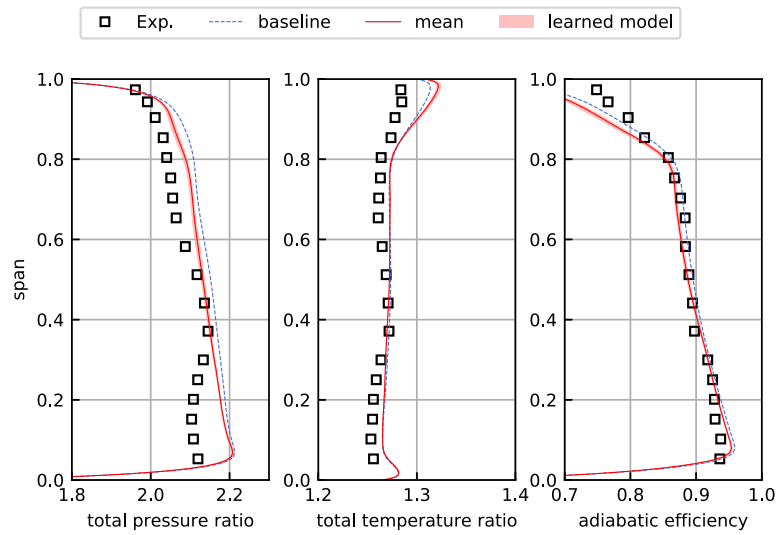
Fig. 3. Radial distribution of the total pressure ratio, total temperature ratio, and adiabatic efficiency with comparison among experiments (square), the baseline model (blue dashed line), and the learned model. The results with the learned model show all samples (red band) and sample mean (red line).
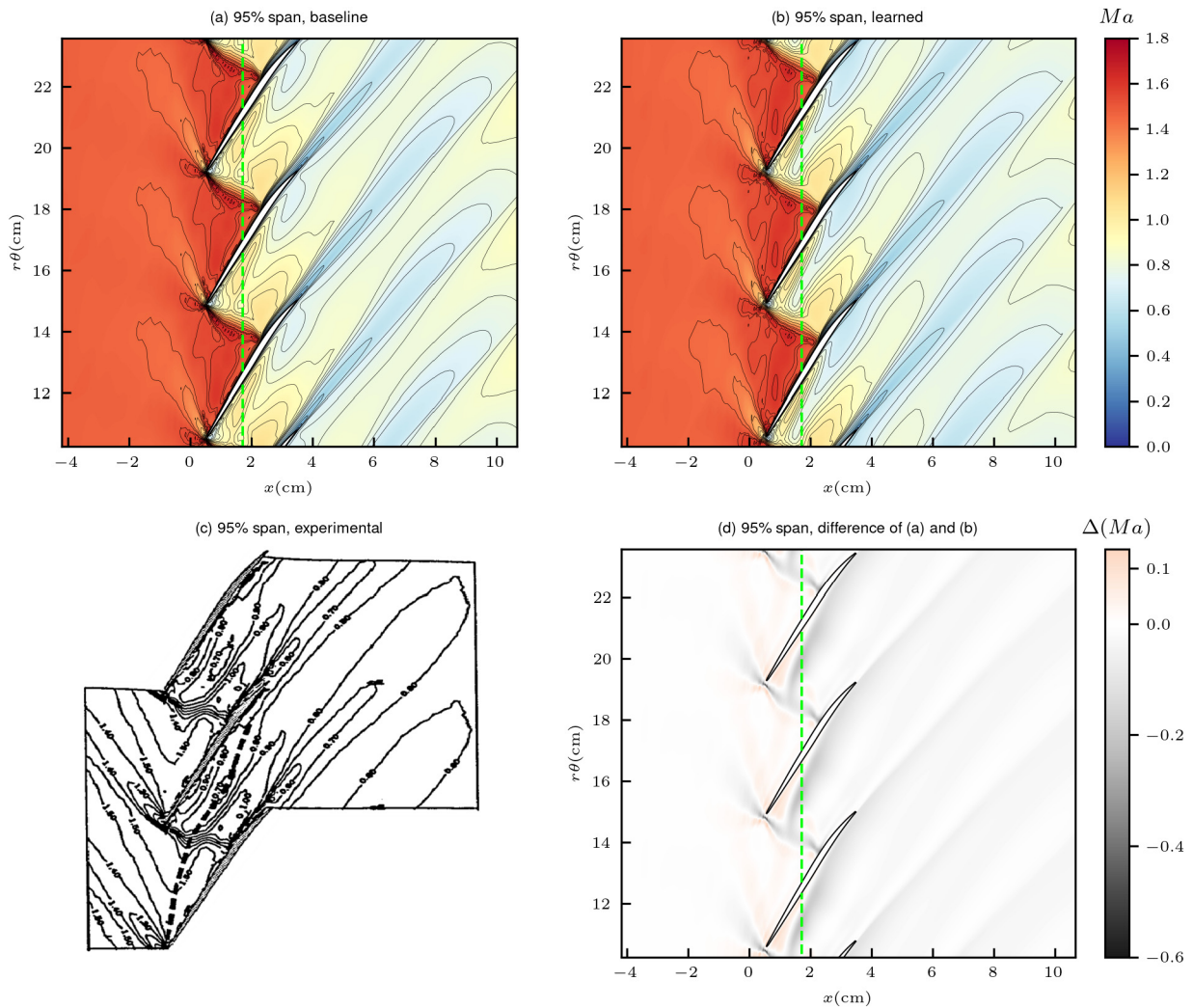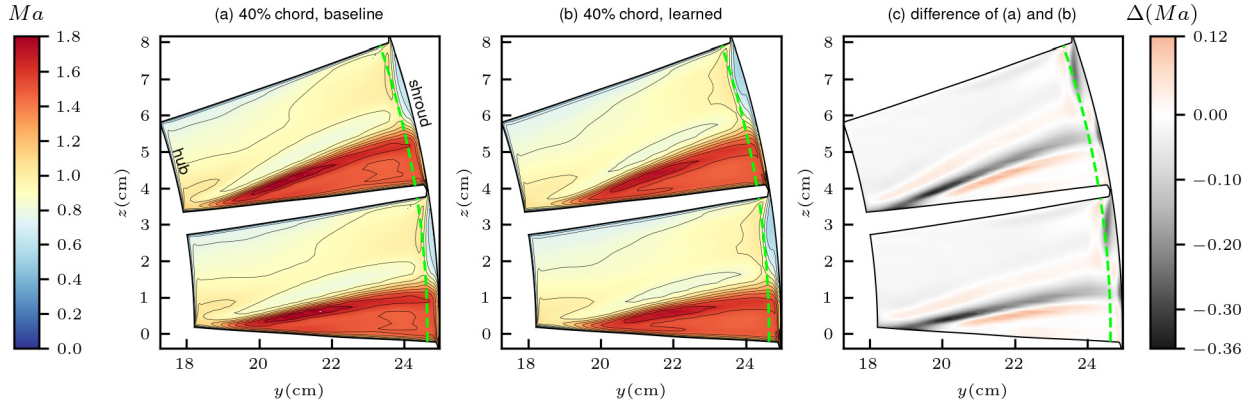

Fig. 4. Mach number contours at 95% span. Panels (a) and (b) show the Mach number predicted with the baseline SA model and the learned model, respectively. Panel (c) shows the Mach number measured by experiments, and the black dashed line indicates the trajectory of tip leakage vortex [8]. Panel (d) shows the Mach number difference between the baseline model and the learned model, $\Delta(Ma) = Ma_{(\text{learned})} - Ma_{(\text{baseline})}$. The green dash line with the constant streamwise location in (a)(b)(d) indicates the 40% chord position.

Fig. 5. Mach number contours at 40% chord. Panels (a) and (b) show the Mach number predicted by the baseline model and the learned model, respectively. Panel (c) shows the Mach number difference between the baseline model and the learned model, $\Delta(Ma) = Ma_{(\text{learned})} - Ma_{(\text{baseline})}$. The green dash line with constant spanwise locations in (a)(b)(c) indicates the 95% span position.

of 95% span and 90% chord, the boundary layer on the suction surface side is thicker than that on the pressure surface side due to the shock-induced flow separation. The high–speed flow near the suction surface side can still be distinguished from the low–speed flow near the pressure surface side. However, the difference in Mach number is not as pronounced as that at 65% chord due to diffusion effects as the flow moves downstream. The 104% chord position is located in the wake, where noticeable velocity loss can be seen. The wake near the suction surface side is wider due to the boundary layer separation, leading to asymmetric velocity distribution. The baseline model underestimates the wake mixing and predicts a sharp drop, while the learned model improves the prediction of the velocity distribution near the suction surface side.

Note that the difference in Mach number before and after the shock wave can indicate the effect of the tip vortex on the flow [8]. Hence, the baseline model generally overpredicts Mach numbers at the observation location by underestimating the strength of the shock-leakage vortex interaction. In contrast, the neural network model can improve the prediction by appropriately capturing the interaction effect in this region. The relative error of the Mach number of the baseline and learned models is given in Table 4.

3.4. Generalizability tests

We validate the capability of the learned neural network model in predicting overall performance under other operating conditions. The static outlet pressure ratio at the hub is varied to obtain the operating conditions of Rotor 37 at different flow rates at the design speed. The prediction results are presented in Fig. 7, with a comparison to the baseline model. It is observed that the baseline model overestimates the total pressure ratio at the design speed, while the learned model makes better predictions compared to the experimental data. In terms of predictive efficiency, the differences between the learned and baseline models were negligible. The selected observation data primarily consists of Mach numbers, with the absence of thermodynamic variables like temperature. Consequently, the prediction of isentropic efficiency does not exhibit a noticeable improvement. These results demonstrate that the neural network models have a certain degree of generalization ability to the unseen operating conditions.

4. Explainability analysis of the learned model

In this section, the behavior of the learned neural network model is investigated to explain the predictive improvement with the ensemble Kalman–based turbulence modeling. The explainability in this work involves the learned model correction and

the input feature contribution. Specifically, the predictive improvement in Mach number should be explained based on the learned model correction. Also, the contribution of the input features to the learned model correction should be analyzed. In the following, the two aspects of explainability are investigated, respectively.

4.1. Explainability of learned model correction

We explain the improvements of the learned model in the Mach number prediction by analyzing the learned eddy viscosity. Fig. 8 compares the eddy viscosity predicted with the baseline and learned models. From Fig. 8(a) and (b), it can be seen that both the baseline and learned models provide the eddy viscosity with significantly large values downstream of the shock wave, near the shroud end wall and in the wake at 95% span. The interaction between the passage shock and the tip leakage flow leads to tip vortex breakdown and unsteady effects [4,9,64], which cause blockage effects to the mainstream. Compared with the baseline model, the learned model increases eddy viscosity significantly near the shroud on the pressure surface side and decreases that on the suction surface side, as shown in Fig. 8(c)(f). The baseline model may underestimate the eddy viscosity and further overestimate the Mach number near the shroud, as shown in Fig. 6. In contrast, the learned model increases the eddy viscosity after the interacting region between the shock and tip vortex. By doing so, the learned model enhances momentum and energy transfer. As such, the predicted velocity magnitude is reduced near the shroud and has a better agreement with the experimental data.

The learned eddy viscosity is further explained based on the neural network output β. Fig. 9 shows the sample mean of β at 95% span and 40% chord. The red color highlights the region with $\beta > 1$, and the blue color indicates that with $\beta < 1$. It can be observed that $\beta > 1$ covers the region after the shock wave to a certain distance downstream of the trailing edge and the wake, as well as near the shroud and blade, indicating enhanced production in these regions. Note that the model correction β has noticeably large values in the wake while the eddy viscosity varies slightly from the baseline. This is likely due to the relatively small production $\mathcal{P}$ in the wake, which can be seen by the contour plots in Fig. 10(a), and hence changes in β have little effect on eddy viscosity. Besides, the correction β is reduced in the region upstream of the shock wave, indicating the weakened production effect. Moreover, the flow near the shroud is very sensitive to β, and a slight increase of β can result in significant changes in eddy viscosity based on Figs. 8 and 9. Such high sensitivity can be due to the local strong strain rate and further large production term $\mathcal{P}$ in the $\tilde{\nu}$ transport equation, as seen in Fig. 10(b). Hence, small variations in β can result in significant changes in eddy viscosity by
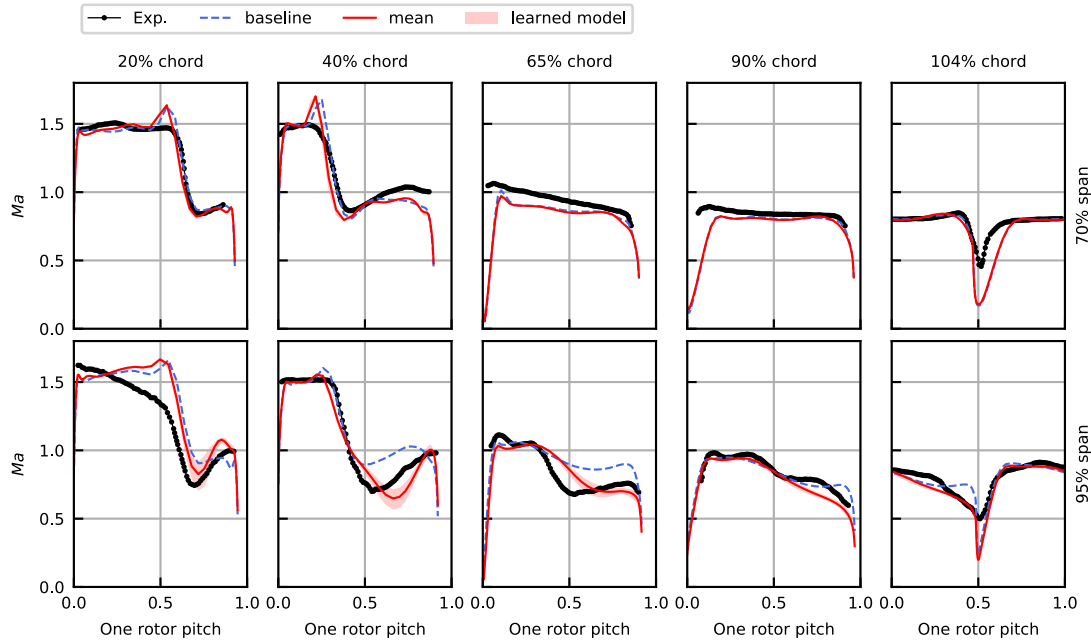
Fig. 6. Plots of Mach number at 70% span (upper panels) and 95% span (bottom panels) for fixed streamwise locations of 20%, 40%, 65%, 90% and 104% chord at the near peak efficiency operating condition. The comparison is conducted among the experiments (dotted lines), the baseline SA model (blue dashed lines), and the learned model (red lines). The red band indicates the coverage of the samples from the learned model.
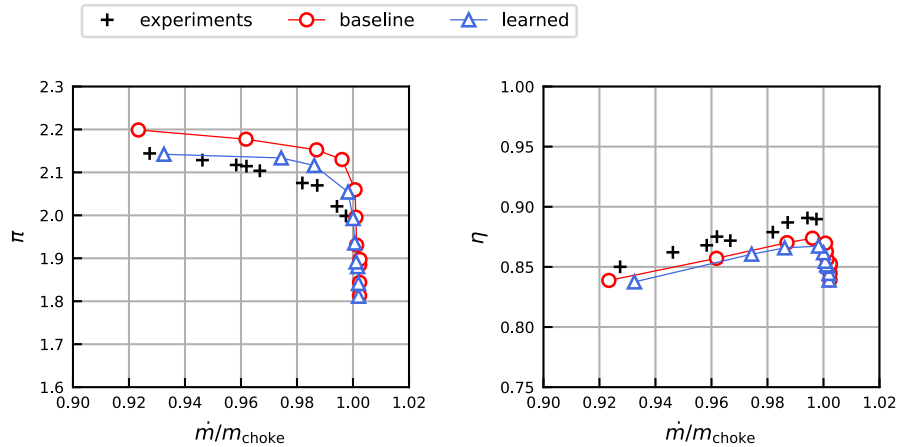

Fig. 7. Validation on the overall performance of Rotor 37: (left) the total pressure, (right) the adiabatic efficiency.

multiplying onto the production term. Besides, the neural network model increases the model correction β after the interacting region between the shock and tip leakage vortex, which augments eddy viscosity compared to the baseline model as shown in Fig. 8. The poor performance of the baseline model near the shroud can be due to the complicated interaction between the mainstream and the tip leakage vortices, which violates the equilibrium assumption in the conventional SA model. In contrast, the learned model improves the prediction by increasing the turbulent production in the interacting region with strong non-equilibrium effects.

4.2. Explainability of model input features

The individual contribution of each input feature is investigated to explain the neural network in this section. The SHAP method is used to perform post–hoc explainability analysis based on the sample mean of the model correction β and network inputs q. The SHAP value ϕ_i can determine the contribution of the input feature q_i to the neural network output β as illustrated in 2.4. The con-

tribution of each input feature to the learned model correction is discussed in the following.

Fig. 11 shows the SHAP values of the trained neural network. In Fig. 11(a), the input features are sorted in descending order of their average absolute SHAP value. The average absolute SHAP value reflects the overall importance of input features on the output, i.e., their global significance. It can be seen that $\mathcal{P}/\mathcal{D}$ (i.e., q_1) has the largest contribution to the output, which indicates the strong non-equilibrium effects for the NASA Rotor 37 case. It is followed by helicity (i.e., q_5) at nearly 0.2, demonstrating the significance of helicity for the compressor rotor flows. The importance of helicity is also validated by various numerical simulations [45,47,65], where helicity is introduced to improve the SA model for compressor flows. Features q_2 (i.e., $\|\boldsymbol{S}\|/\|\Omega\|$) and q_3 (i.e., δ) have the similar magnitude around 0.09, which is about 30%–50% of mean absolute SHAP value of q_1, q_5. Feature q_4 (i.e., χ) has the minimum global significance. The SHAP analysis indicates that the production–to–destruction ratio and the helicity are the two most important influencing factors for the model prediction. Fig. 11(b)
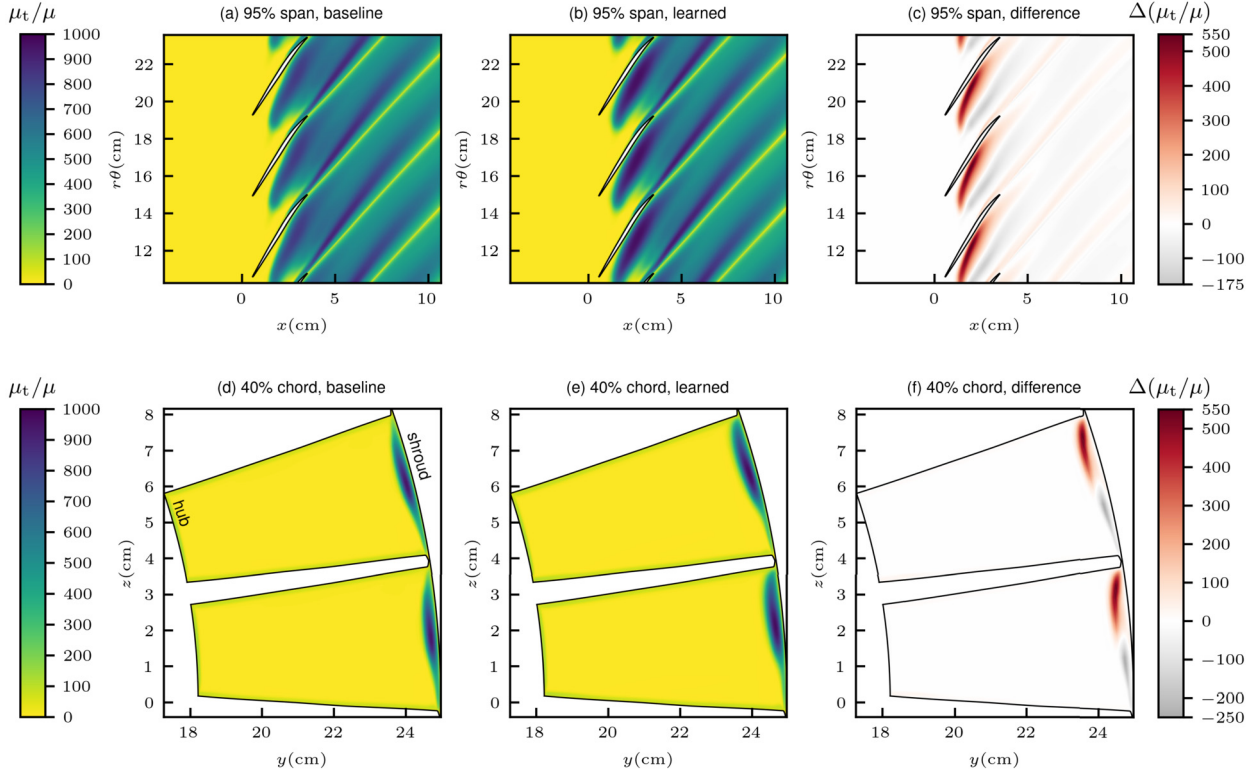
Fig. 8. Contour plots of the eddy viscosity. Panels (a, b,c) show the eddy viscosity at 95% span of the baseline model, the learned model, and the difference between them, respectively. Panels (d, e,f) show the eddy viscosity at 40% chord of the baseline model, the learned model, and the difference between them, respectively. The difference is calculated by $\Delta(\mu_t/\mu) = (\mu_t/\mu)_{\text{learned}} - (\mu_t/\mu)_{\text{baseline}}$.
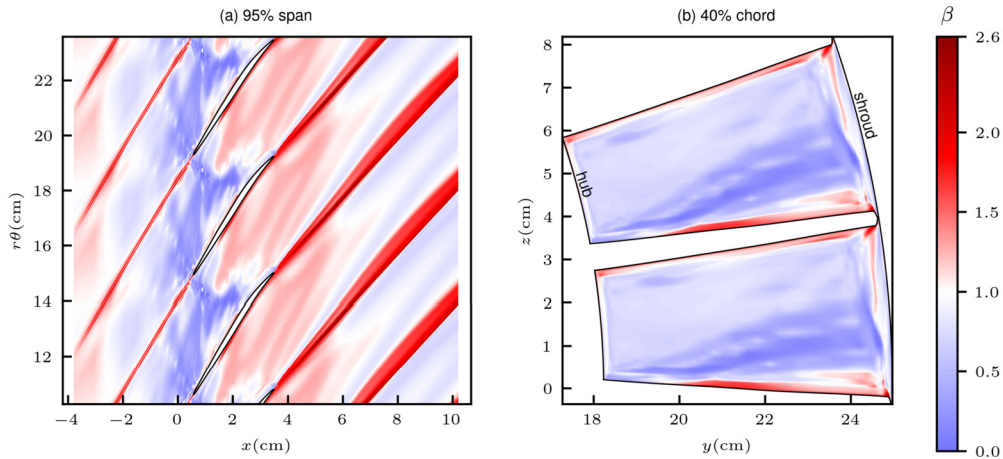

Fig. 9. Contour plots of the neural network output β at 95% span and 40% chord.

shows the bee swarm plot of the correlation between the feature input q_i and the corresponding SHAP value ϕ_i. It can be seen that the SHAP value ϕ_1 is inversely proportional to q_1, i.e., $\phi_1 \propto -q_1$, which means that the model correction β decreases as the ratio of the production to the destruction term increases. The SHAP value ϕ_5 is proportional to q_5, i.e., $\phi_5 \propto q_5$, which means that the network output β increases as the value of helicity increases. The SHAP values ϕ_2 and ϕ_3 are inversely proportional to q_2 and q_3, respectively, i.e., $\phi_2 \propto -q_2$ and $\phi_3 \propto -q_3$. That means the model correction β decreases as the adverse pressure gradient and strain rates increase. Also, it is noted that the negative contribution of strain rates to the model output is much larger than the positive contribution. The SHAP value ϕ_4 is proportional to the feature q_4, i.e., $\phi_4 \propto q_4$. However, the correlation between ϕ_4 and q_4 is much

less than other features, with most feature values having almost zero contribution to the output.

The spatial distribution of input feature values and SHAP values at the 95% span are plotted in Fig. 12 to illustrate the spatial relationship between input features and network output. The feature q_1, i.e., $\mathcal{P}/\mathcal{D}$, has distinguished differences before and after the shock wave. Specifically, the feature q_1 has relatively large values before the shock wave and small values after the shock wave. Its corresponding SHAP value is negatively correlated with the feature, as can be seen that the production term is reduced before the shock wave and increased after the shock wave. In the region with the shock wave, the input feature has relatively large values, while it has small feature values in the wake flows. Accordingly, the SHAP value of q_1 is small in the shock region and
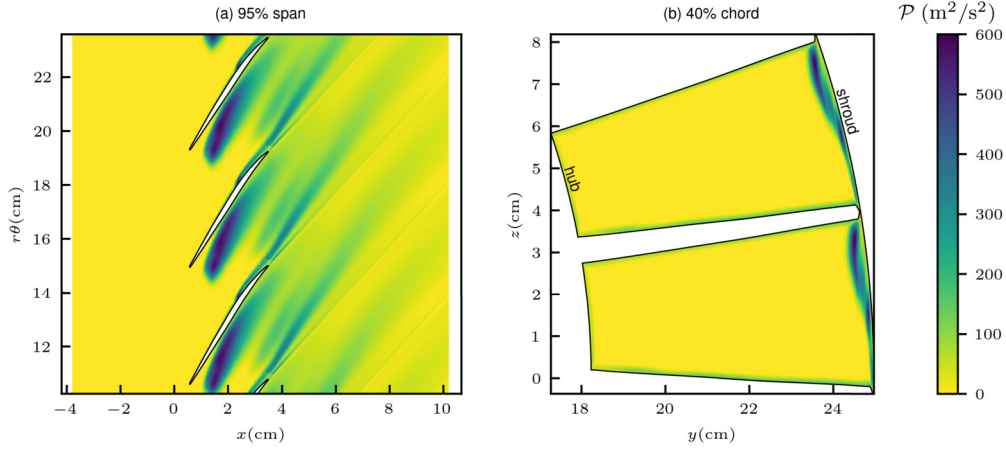
Fig. 10. Contour plots of the production term $\mathcal{P}$ predicted by the learned models at 95% span and 40% chord.
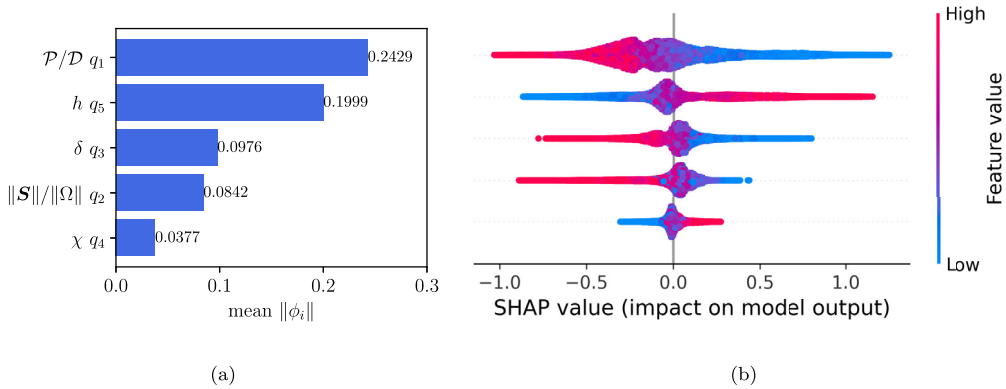

Fig. 11. SHAP analysis. (a) The bar chart shows the mean absolute SHAP value. (b) The bee swarm plot illustrates the details of the SHAP value. Each dot represents a CFD grid point, with horizontal coordinates indicating the SHAP value ϕ and the color bar indicating the value of inputs. Dots having the same SHAP value accumulate in the vertical direction, showing the distribution density of the SHAP value.

large in the wake. The feature q_2, i.e., $\|S\|/\|\Omega\|$, characterizes the ratio of strain rate and rotation rate, which has large values in regions with the shock wave, expansion wave, and near the center of the wake. These regions correspond to flow with strong strain rates. The corresponding SHAP values indicate that the production term of the baseline SA model should be reduced in these regions and increased in regions with lower strain rates. The feature q_3, i.e., δ, characterizes the adverse pressure gradient effects, which have a large value after the shock wave. That is likely due to the abrupt pressure rise in the shock wave region. The SHAP values indicate that the model correction needs to be reduced in regions with high adverse pressure gradients. The feature $\chi = \tilde{v}/v$ (q_4) is the normalized working variable in the SA model, which is directly related to eddy viscosity. Hence its plot is very similar to the contour of eddy viscosity in Fig. 8. The SHAP value also indicates a proportional correlation that the model correction term β increases in regions with a large χ. The feature h (i.e., q_5) characterizes the consistency of velocity and vorticity directions, which has a larger value at the shock wave and in the wake. The corresponding SHAP values indicate that the production term should be increased in these regions compared to the baseline model.

The SHAP value ϕ_i can be further divided into pure effects $\phi_{i,i}$ and interaction effects $\phi_{i,j}$. The pure effects $\phi_{i,i}$ indicate the effect of feature q_i on the output subtracting the effect with other features, and $\phi_{i,j}$ represents the joint impact of input features q_i and q_j on the output. Fig. 13(a) shows the mean absolute of the SHAP interaction value $\overline{|\phi_{i,j}|}$. This is a symmetric matrix, where the diagonal elements represent the overall contribution of the pure impact of feature q_i on the output, and the off-diagonal el-

ements represent the overall contribution of the joint impact of features q_i and q_j on the output. It can be seen that, for each input feature, the contribution of pure effects to the output is greater than the contribution of its joint impact with other features. The values of the diagonal elements $\phi_{i,i}$ are sorted in the same order as the SHAP values ϕ_i. The rank according to the overall impact on the output is in the following order: $\overline{|\phi_{1,1}|} > \overline{|\phi_{5,5}|} > \overline{|\phi_{3,3}|} > \overline{|\phi_{5,1}|} > \overline{|\phi_{2,2}|}$. In Fig. 13(b), the matrix from Fig. 13(a) is normalized by dividing each off-diagonal element with the corresponding diagonal element, i.e., $\overline{|\phi_{i,j}|}/\overline{|\phi_{i,i}|}$. The resulting matrix represents the relative impact on the output when input features q_j take joint effects with feature q_i. The largest normalized interaction SHAP value (off-diagonal) in each row of the matrix in Fig. 11(b) are in the first column and the last column. It can be seen that features $\mathcal{P}/\mathcal{D}$ (q_1) and $\|S\|/\|\Omega\|$ (q_2) have the greatest impact on the output when interacting with helicity h (q_5), and features δ (q_3), χ (q_4), h (q_5) have the greatest impact on the output when interacting with feature $\mathcal{P}/\mathcal{D}$ (q_1).

Fig. 14 shows the relationship between $\phi_{i,i}$ and the corresponding input feature q_i, where each point represents a CFD grid point, and the solid blue line is a fitting curve to the data points. The relationship between the diagonal SHAP values and the input features is similar to that shown in Fig. 11. That is, $\phi_{1,1}$ is proportional to $-q_1$; $\phi_{2,2}$ is proportional to q_2 when $q_2 > 0$; $\phi_{3,3}$ is proportional to $-q_3$; $\phi_{4,4}$ is proportional to q_4; $\phi_{5,5}$ is proportional to q_5. In terms of the pure effects of each feature, the output is most sensitive to the feature $\mathcal{P}/\mathcal{D}$ since the fitting curve has the largest absolute value of the slope. The feature $\|S\|/\|\Omega\|$ has
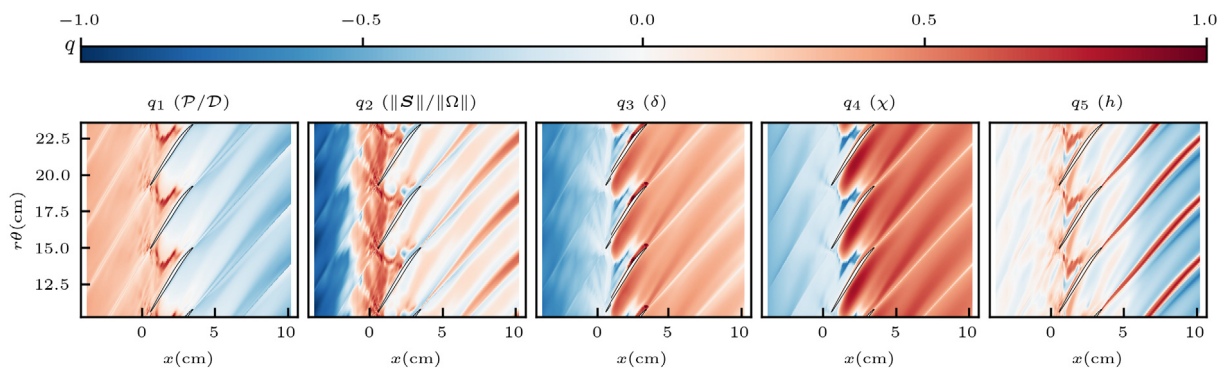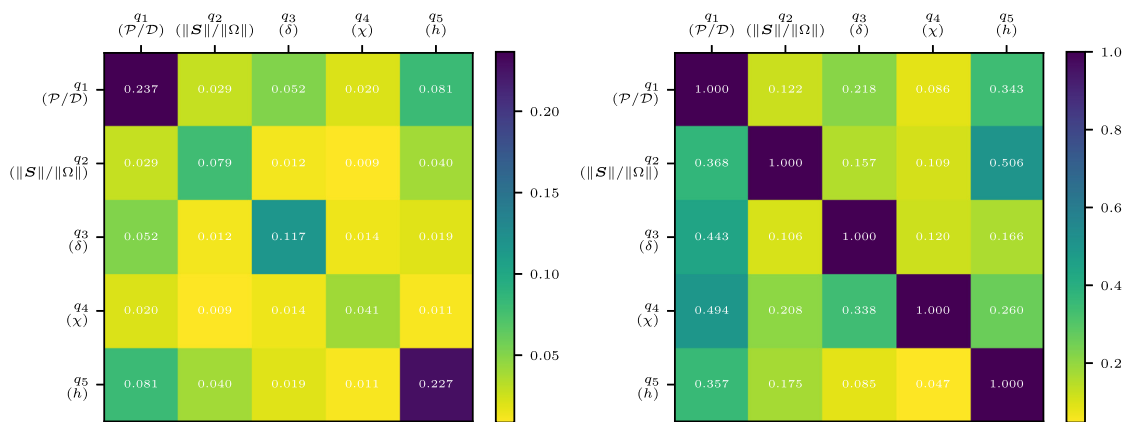
(a) Input features
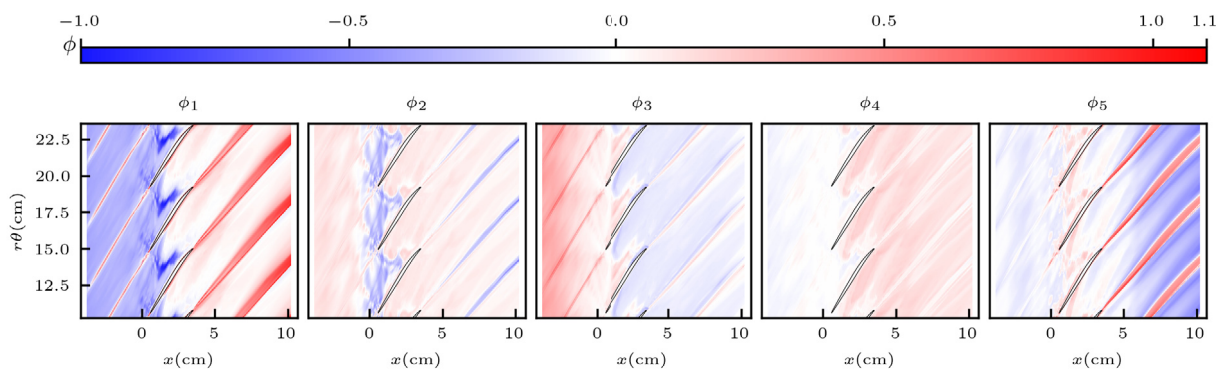
(b) SHAP values

Fig. 12. Contours of (a) input features (scaled) and (b) SHAP value at 95% span.

(a)

(b)

Fig. 13. Heatmap of (a) mean of absolute SHAP interaction values and (b) normalized mean of absolute SHAP interaction values.
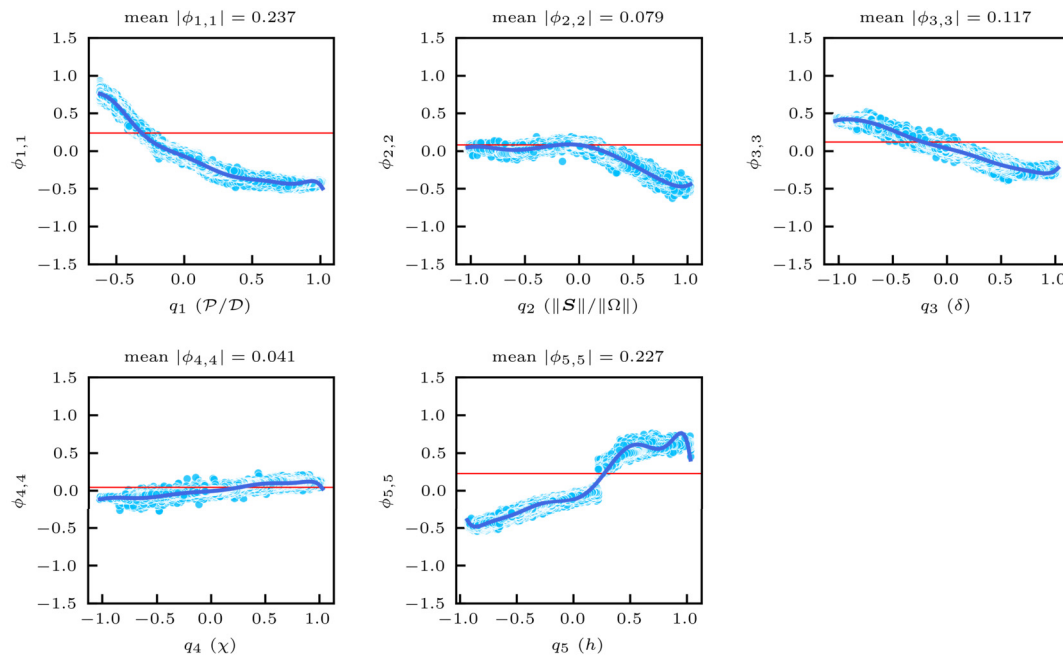
Fig. 14. Dialog SHAP interaction value $\phi_{i,i}$ on corresponding input feature q_i.

negative effects on the output when $q_2 > 0$, which indicates that this feature has noticeable impacts on the model output when the ratio of the strain rate tensor magnitude to the vorticity exceeds a certain threshold. The adverse pressure gradient indicator (i.e., q_3) negatively correlates with the neural network output, which has a relatively moderate slope compared to q_1. The feature χ (i.e., q_4) has a very small impact on the output, showing only a weak positive correlation. The pure effect of helicity h (q_5) on the neural network output is relatively complex. When h is small (i.e., $q_5 < 0.25$), its pure effect on the output increases monotonically and reduces the network output. However, in the range of $q_5 > 0.25$, the pure effect of helicity is not monotonic with two peaks.

Conclusively, we can gain insights into the black–box neural network model through explainability analysis of the input features and model correction, as summarized in Fig. 15. The input features in descending order according to their global significance on model output are $\mathcal{P}/\mathcal{D}$, h, δ, $\|\boldsymbol{S}\|/\|\Omega\|$, χ. It is found that the two most important features in the SA model correction are the ratio of the production term to the destruction term and the helicity. The former measures non-equilibrium effects, while the latter affects the energy backscatter in turbulence. The red arrows indicate larger values in feature importance, and the blue arrows indicate smaller values. Features $\mathcal{P}/\mathcal{D}$, $\|\boldsymbol{S}\|/\|\Omega\|$, and δ have a negative correlation with the model output, while χ and h have a positive correlation. In the region after the tip leakage vortex interacts with the shock wave, model output has $\beta > 1$. The correction on the SA turbulence model captures the non-equilibrium effect in the vortex breakdown region by increasing the eddy viscosity. As such, the momentum and energy transfer are enhanced, and the Mach number in the blockage is well predicted compared to the baseline model.

5. Conclusions

In this work, the explainability of the ensemble Kalman–based turbulence modeling is investigated for the transonic axial compressor rotor flows. A neural network-based model correction is trained in a model–consistent manner, which improves the flow field prediction by incorporating various experimental data, includ-ing the velocity. The learned neural network is analyzed to explain the predictive improvement based on the model correction and eddy viscosity fields. Moreover, the input features are analyzed to indicate their relative contributions to the learned model corrections based on the SHAP method. The numerical results show that the learned neural network significantly improves the Mach number prediction near the shroud by capturing the blockage effects in the vortex breakdown region. Such blockage effects are underestimated with the baseline model, likely due to the equilibrium assumption. In contrast, the learned model can consider the non-equilibrium effects to improve the flow prediction by increasing local production with the neural network-based model correction. Besides, our post–hoc explainability analysis of the neural network shows that the production–to–destruction ratio and the helicity are the two most important features of the learned neural network. This can guide the input feature selection for developing accurate turbulence models in future investigations.
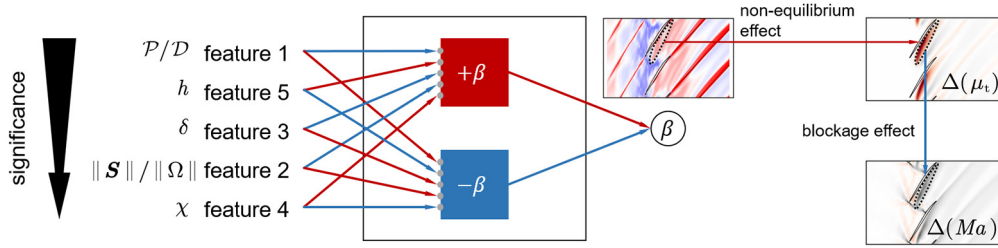
Fig. 15. Illustration of the neural network explanation in the input features and the model output.

Appendix A. Spalart–Allmaras turbulence model

The production, destruction, and transportation terms in the SA turbulence model read

$$\mathcal{P} = c_{b1}(1 - f_{t2})\tilde{S}\tilde{\nu},$$

$$\mathcal{D} = \left(c_{w1}f_w - \frac{c_{b1}}{\kappa^2}f_{t2}\right)\left(\frac{\tilde{\nu}}{d}\right)^2, \tag{A.1}$$

$$\mathcal{T} = \frac{1}{\sigma}\left(\nabla \cdot \left((\nu + \tilde{\nu})\nabla\tilde{\nu}\right) + c_{b2}\left(\nabla\tilde{\nu}\right)^2\right),$$

in which

$$\tilde{S} = S + \frac{\tilde{\nu}}{\kappa^2 d^2}f_{v2}, \tag{A.2}$$

where S is the magnitude of the vorticity, and d is the nearest distance to the wall. Functions used in the model are defined by

$$f_{v1} = \frac{\chi^3}{\chi^3 + c_{v1}^3}, \quad f_{v2} = 1 - \frac{\chi}{1 + \chi f_{v1}},$$

$$f_w = g\left[\frac{1 + c_{w3}^6}{g^6 + c_{w3}^6}\right]^{1/6}, \tag{A.3}$$

where $\chi \equiv \tilde{\nu}/\nu$, $g = r + c_{w2}\left(r^6 - r\right)$, $r \equiv \tilde{\nu}/(\tilde{S}\kappa^2 d^2)$. The constants $c_{b1}, c_{b2}, \kappa, \sigma, c_{w1} = c_{b1}/\kappa + (1 + c_{b2})/\sigma, c_{w2}, c_{w3}, c_{v1}$ are specified. The boundary condition on the wall is expressed by setting $\tilde{\nu} = 0$.

Meanwhile, the f_{t2} related terms are ignored in the simulation. The SA-helicity model [45] is proposed to take account of the energy backscatter by modifying

$$\tilde{S} = \left(1 + C_{h1}h^{C_{h2}}\right)\omega + \frac{\tilde{\nu}}{\kappa^2 d^2}f_{v2}, \tag{A.4}$$

where h is the helicity density, and $C_{h1} = 0.71$, $C_{h2} = 0.6$ are constants.

We compare the prediction of the original SA model, the SA-helicity model, and the learned model in the Mach number at 70% span and 95% span. The results are presented in Fig. A.1, which show that the SA-helicity model does not achieve significant improvement compared to the SA model. We note that it is not a comprehensive comparison between the two turbulence models, and the results are obtained based on our particular implementation in the Multall code with the H-type mesh.

In the present work, we use the original SA model as the baseline model to exclude the influence of helicity. By doing so, we can validate the capability of machine learning in identifying important features, e.g., helicity, for turbulence modeling of compressor flows.

Appendix B. Neural network model with partial input features

We train the neural networks with partial input features to validate the effects of specific features, i.e., the production-destruction ratio and helicity, on the flow prediction. The used method and
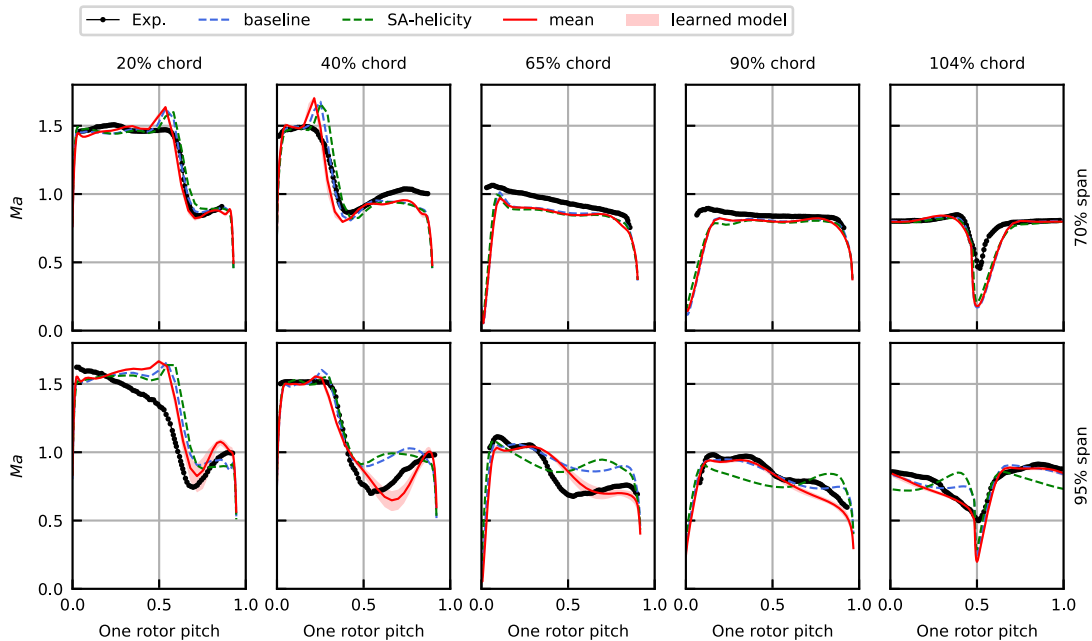

Fig. A.1. Plots of Mach number at 70% span (upper panels) and 95% span (bottom panels) for fixed streamwise locations of 20%, 40%, 65%, 90% and 104% chord at the near peak efficiency operating condition. The comparison is conducted among the experiments (dotted lines), the baseline SA model (blue dashed lines), the SA-helicity model (green dashed lines), and the learned model (red lines). The red band indicates the coverage of the samples from the learned model.
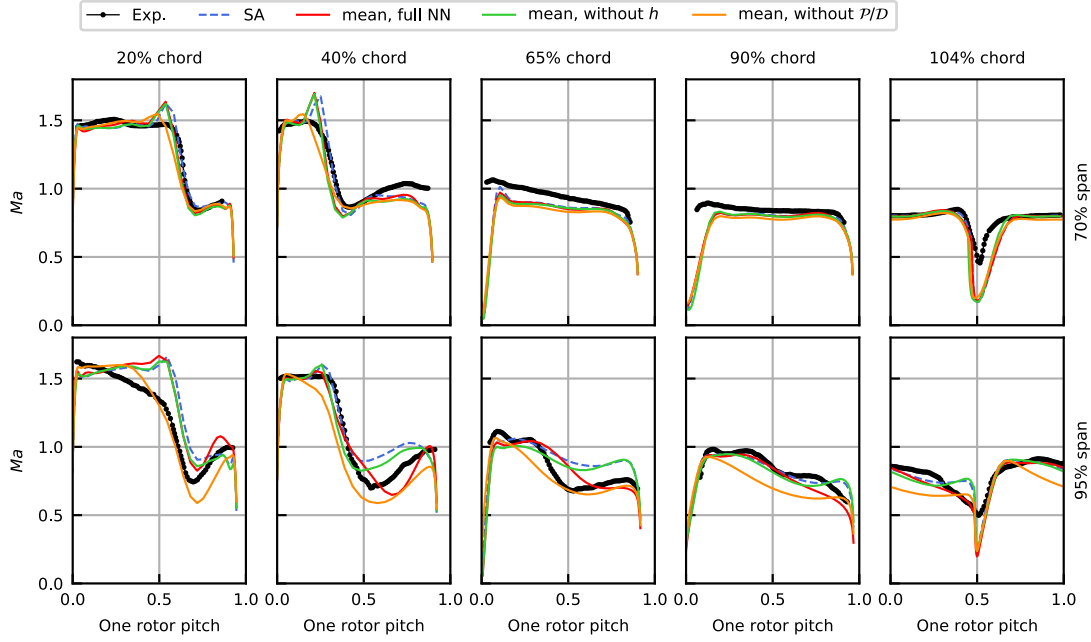
Fig. B.1. Plots of Mach number at 70% span (upper panels) and 95% span (bottom panels) for fixed streamwise locations of 20%, 40%, 65%, 90% and 104% chord at the near peak efficiency operating condition. The comparison is conducted among the experiments (dotted lines), the baseline SA model (blue dashed lines), the mean of the learned model with all five inputs (red lines), the mean of the learned model without helicity as input (green lines), and the mean of the learned model without $\mathcal{P}/\mathcal{D}$ as input (orange lines).
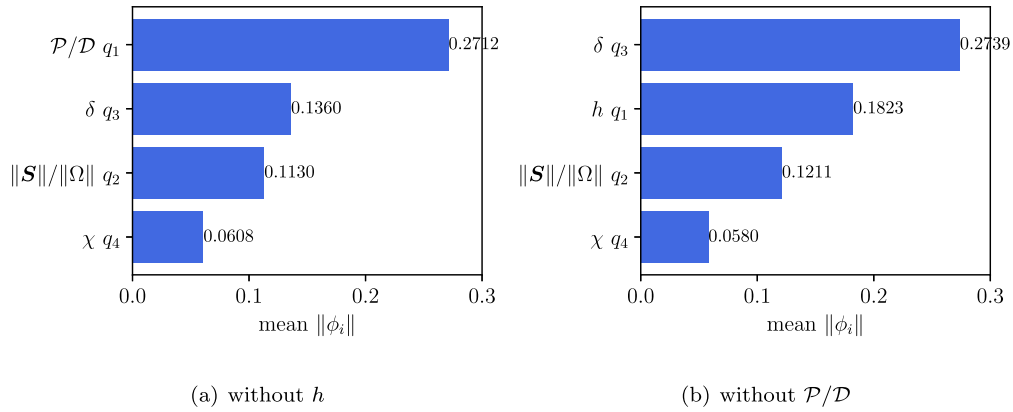

Fig. B.2. The bar chart of SHAP analysis. The chart shows the mean absolute SHAP value of neural network input features.

case setup are consistent with the case with full input features as illustrated in Section 3. The main results and discussion are presented in the following.

A neural network without the input of helicity is trained to highlight the importance of helicity as an input. The input features q only include the four features in Eq. (8), i.e., $\mathcal{P}/\mathcal{D}$, $\|\boldsymbol{S}\|/\|\Omega\|$, δ, χ. The predicted Mach number is illustrated in Fig. B.1. Compared to the full neural network with the input of helicity, the model has limited improvements in Mach number prediction. It can be seen that the neural network model predicts a higher Mach number than the model with input of helicity in the subsonic region, e.g., the 40% chord and 95% span, as well as 65% chord and 95% span. In 104% chord and 95% span, the model without inputs of helicity predicts a Mach number distribution in the wake center closer to the result of the baseline model. The apparent discrepancy in Mach number predictions shows the importance of helicity for improving the SA turbulence model.

Also, the neural network without the input of $\mathcal{P}/\mathcal{D}$ is trained to validate the importance of $\mathcal{P}/\mathcal{D}$ as an input. The input features q only include the four features in Eq. (8), i.e., $\|\boldsymbol{S}\|/\|\Omega\|$, δ, χ, h. The

predicted Mach number is illustrated in Fig. B.1. Compared to the full neural network with the input of $\mathcal{P}/\mathcal{D}$, the model has limited improvements in Mach number prediction. It can be seen that the neural network model without the input of $\mathcal{P}/\mathcal{D}$ predicts a lower Mach number over the entire rotor pitch than the full model, e.g., the 40% chord and 95% span, 65% chord and 95% span, as well as 90% chord and 95% span. The apparent discrepancy in Mach number predictions shows the importance of $\mathcal{P}/\mathcal{D}$ for improving the SA turbulence model.

To enhance the validation of SHAP analysis, the mean absolute SHAP value of the four features are illustrated in Fig. B.2. It shows that for the case without the input of h, the feature importance is consistent with the case with helicity as shown in Fig. 11. That is, the feature that contributes the most is $\mathcal{P}/\mathcal{D}$, while the feature that contributes the least is χ. The features δ, $\|\boldsymbol{S}\|/\|\Omega\|$ still have a similar SHAP value to the case with helicity as input. For the case without the input of $\mathcal{P}/\mathcal{D}$, the feature δ exhibits the greatest influence on the output, surpassing helicity h in terms of contribution. Meanwhile, the features $\|\boldsymbol{S}\|/\|\Omega\|$ and χ keep their positions as the least significant features, with unchanged relative order. The

Table C.5

Summary of prediction error in radial total pressure ratio, radial total temperature ratio, radial adiabatic efficiency, and Mach number with the baseline and different neural networks.

	baseline	NN 1	NN 2	NN3
Number of hidden layers		2	2	4
Number of neurons in each hidden layer		5	10	10
Radial total pressure ratio	2.73%	2.05%	3.78%	3.78%
Radial total temperature ratio	0.91%	0.97%	1.37%	1.37%
Radial adiabatic efficiency	1.62%	1.98%	1.50%	1.49%
Mach number	8.34%	7.28%	8.35%	8.35%

SHAP value of δ gets increased likely due to the strong connection between δ and $\mathcal{P}/\mathcal{D}$. The feature δ indicates the adverse pressure gradient, which can also represent the non–equilibrium effects as the feature $\mathcal{P}/\mathcal{D}$. The contribution of the feature δ is increased to capture the non–equilibrium effects with the exclusion of $\mathcal{P}/\mathcal{D}$. Also, it can be seen from Fig. 13(a) that $\overline{|\phi_{3,1}|}$, i.e., the interaction between the feature δ and the feature $\mathcal{P}/\mathcal{D}$, has the highest interaction value, excluding the diagonals in the third row.

Appendix C. Sensitivity study of neural network hyperparameters

We employ different neural network architectures to investigate the effect of the hyperparameters on the predictive performance of the neural network. Three network architectures are tested: (1) 2 hidden layers with 5 neurons per layer (NN1); (2) 2 hidden layers with 10 neurons per layer (NN2); and (3) 4 hidden layers with 10 neurons per layer (NN3). The predictive error is summarized in Table C.5, quantitatively showing the comparable ability of different neural networks in predicting the radial performance and Mach number. The results demonstrate that neural network NN1 exhibits the best prediction of Mach number improvement.

References

[1] R.D. Sandberg, V. Michelassi, Fluid dynamics of axial turbomachinery: blade- and stage-level simulations and models, Annu. Rev. Fluid Mech. 54 (1) (2022) 255–285, https://doi.org/10.1146/annurev-fluid-031221-105530.

[2] C. Hah, A.J. Wennerstrom, Three-dimensional flowfields inside a transonic compressor with swept blades, J. Turbomach. 113 (2) (1991) 241–250, https://doi.org/10.1115/1.2929092.

[3] A. Shabbir, J. Zhu, M. Celestina, Assessment of Three Turbulence Models in a Compressor Rotor, vol. 78729, American Society of Mechanical Engineers, 1996.

[4] R.V. Chima, Calculation of tip clearance effects in a transonic compressor rotor, J. Turbomach. 120 (1) (1998) 131–140, https://doi.org/10.1115/1.2841374.

[5] R. Chima, SWIFT code assessment for two similar transonic compressors, in: 47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, 2009, p. 1058.

[6] B. Tartinville, C. Hirsch, Rotor 37, in: W. Haase, B. Aupoix, U. Bunge, D. Schwamborn (Eds.), FLOMANIA — A European Initiative on Flow Physics Modelling, Springer, Berlin, Heidelberg, 2006, pp. 193–202.

[7] L. Reid, R.D. Moore, Design and overall performance of four highly loaded, high speed inlet stages for an advanced high-pressure-ratio core compressor, Tech. Rep, NASA Lewis Research Center, Cleveland, OH, United States, 1978.

[8] K.L. Suder, Experimental Investigation of the Flow Field in a Transonic, Axial Flow Compressor with Respect to the Development of Blockage and Loss, Case Western Reserve University, 1996.

[9] J. Dunham, CFD validation for propulsion system components (la validation CFD des organes des propulseurs), Tech. Rep, Advisory Group for Aerospace Research and Development Neuilly-sur-Seine, France, 1998.

[10] D.C. Wilcox, et al., Turbulence Modeling for CFD, vol. 2, DCW Industries La, Canada, CA, 1998.

[11] X. He, F. Zhao, M. Vahdati, Evaluation of Spalart-Allmaras turbulence model forms for a transonic axial compressor, GPPS Paper GPPS (2020) 0013.

[12] Y. Zhao, H.D. Akolekar, J. Weatheritt, V. Michelassi, R.D. Sandberg, RANS turbulence model development using CFD-driven machine learning, J. Comput. Phys. 411 (2020) 109413, https://doi.org/10.1016/j.jcp.2020.109413.

[13] S.L. Brunton, Applying machine learning to study fluid mechanics, Acta Mechanica Sinica 37 (12) (2021) 1718–1726, https://doi.org/10.1007/s10409-021-01143-6.

[14] L. Zhu, W. Zhang, X. Sun, Y. Liu, X. Yuan, Turbulence closure for high Reynolds number airfoil flows by deep neural networks, Aerosp. Sci. Technol. 110 (2021) 106452, https://doi.org/10.1016/j.ast.2020.106452.

[15] J. Weatheritt, R. Sandberg, A novel evolutionary algorithm applied to algebraic modifications of the rans stress-strain relationship, J. Comput. Phys. 325 (2016) 22–37, https://doi.org/10.1016/j.jcp.2016.08.015.

[16] M. Schmelzer, R.P. Dwight, P. Cinnella, Discovery of algebraic Reynolds-stress models using sparse symbolic regression, Flow Turbul. Combust. 104 (2) (2020) 579–603, https://doi.org/10.1007/s10494-019-00089-x.

[17] J. Ling, A. Kurzawski, J. Templeton, Reynolds averaged turbulence modelling using deep neural networks with embedded invariance, J. Fluid Mech. 807 (2016) 155–166, https://doi.org/10.1017/jfm.2016.615.

[18] J.-X. Wang, J.-L. Wu, H. Xiao, Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data, Phys. Rev. Fluids 2 (2017) 034603, https://doi.org/10.1103/PhysRevFluids.2.034603.

[19] J.-L. Wu, C.A. Michelén-Ströfer, H. Xiao, Physics-informed covariance kernel for model-form uncertainty quantification with application to turbulent flows, Comput. Fluids 193 (2019) 104292, https://doi.org/10.1016/j.compfluid.2019.104292.

[20] X.I.A. Yang, K.P. Griffin, Grid-point and time-step requirements for direct numerical simulation and large-eddy simulation, Phys. Fluids 33 (1) (2021) 015108, https://doi.org/10.1063/5.0036515.

[21] N. Gourdain, F. Sicot, F. Duchaine, L. Gicquel, Large eddy simulation of flows in industrial compressors: a path from 2015 to 2035, Philos. Trans. R. Soc. Lond. A, Math. Phys. Eng. Sci. 372 (2022) (2014) 20130323.

[22] P. Tucker, Computation of unsteady turbomachinery flows: part 1—progress and challenges, Prog. Aerosp. Sci. 47 (7) (2011) 522–545, https://doi.org/10.1016/j.paerosci.2011.06.004.

[23] Z. Li, X. Yang, Large-eddy simulation on the similarity between wakes of wind turbines with different yaw angles, J. Fluid Mech. 921 (2021) A11.

[24] Z. Li, G. Dong, X. Yang, Onset of wake meandering for a floating offshore wind turbine under side-to-side motion, J. Fluid Mech. 934 (2022) A29.

[25] R.D. Sandberg, V. Michelassi, The current state of high-fidelity simulations for main gas path turbomachinery components and their industrial impact, Flow Turbul. Combust. 102 (4) (2019) 797–848.

[26] K. Duraisamy, Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence, Phys. Rev. Fluids 6 (2021) 050504, https://doi.org/10.1103/PhysRevFluids.6.050504.

[27] C.A. Michelén-Ströfer, H. Xiao, End-to-end differentiable learning of turbulence models from indirect observations, Theor. Appl. Mech. Lett. 11 (4) (2021) 100280, https://doi.org/10.1016/j.taml.2021.100280.

[28] C.A. Michelén-Ströfer, X.-L. Zhang, H. Xiao, Ensemble gradient for learning turbulence models from indirect observations, Commun. Comput. Phys. 30 (5) (2021) 1269–1289, https://doi.org/10.4208/cicp.OA-2021-0082.

[29] X.-L. Zhang, H. Xiao, X. Luo, G. He, Ensemble Kalman method for learning turbulence models from indirect observation data, J. Fluid Mech. 949 (2022) A26, https://doi.org/10.1017/jfm.2022.744.

[30] Y. Liu, X.-L. Zhang, G. He, Learning neural-network-based turbulence models for external transonic flows using ensemble Kalman method, AIAA J. (2023) 1–15.

[31] Z. Wang, W. Zhang, A unified method of data assimilation and turbulence modeling for separated flows at high Reynolds numbers, Phys. Fluids 35 (2) (2023) 025124, https://doi.org/10.1063/5.0136420.

[32] A. Adadi, M. Berrada, Peeking inside the black-box: a survey on explainable artificial intelligence (XAI), IEEE Access 6 (2018) 52138–52160, https://doi.org/10.1109/ACCESS.2018.2870052.

[33] X. He, J. Tan, G. Rigas, M. Vahdati, On the explainability of machine-learning-assisted turbulence modeling for transonic flows, Int. J. Heat Fluid Flow 97 (2022) 109038, https://doi.org/10.1016/j.ijheatfluidflow.2022.109038.

[34] P. Spalart, S. Allmaras, A one-equation turbulence model for aerodynamic flows, in: 30th Aerospace Sciences Meeting and Exhibit, 1992, p. 439.

[35] A.P. Singh, K. Duraisamy, Using field inversion to quantify functional errors in turbulence closures, Phys. Fluids 28 (4) (2016) 045110, https://doi.org/10.1063/1.4947045.

[36] C. Yan, H. Li, Y. Zhang, H. Chen, Data-driven turbulence modeling in separated flows considering physical mechanism analysis, Int. J. Heat Fluid Flow 96 (2022) 109004, https://doi.org/10.1016/j.ijheatfluidflow.2022.109004.

[37] A.P. Singh, K. Duraisamy, Z.J. Zhang, Augmentation of turbulence models using field inversion and machine learning, in: 55th AIAA Aerospace Sciences Meeting, 2017, p. 0993.

[38] K. Duraisamy, Z.J. Zhang, A.P. Singh, New approaches in turbulence and transition modeling using data-driven techniques, in: 53rd AIAA Aerospace Sciences Meeting, 2015, p. 1284.

[39] E.J. Parish, K. Duraisamy, A paradigm for data-driven predictive modeling using field inversion and machine learning, J. Comput. Phys. 305 (2016) 758–774, https://doi.org/10.1016/j.jcp.2015.11.012.

[40] A.P. Singh, K. Duraisamy, Using field inversion to quantify functional errors in turbulence closures, Phys. Fluids 28 (4) (2016) 045110, https://doi.org/10.1063/1.4947045.

[41] A.P. Singh, S. Medida, K. Duraisamy, Machine-learning-augmented predictive modeling of turbulent separated flows over airfoils, AIAA J. 55 (7) (2017) 2215–2227, https://doi.org/10.2514/1.J055595.

[42] A.P. Singh, K. Duraisamy, S. Pan, Characterizing and improving predictive accuracy in shock-turbulent boundary layer interactions using data-driven models, in: 55th AIAA Aerospace Sciences Meeting, 2017, p. 0314.

[43] J.R. Holland, J.D. Baeder, K. Duraisamy, Towards integrated field inversion and machine learning with embedded neural networks for RANS modeling, in: AIAA Scitech 2019 Forum, 2019, p. 1884.

[44] J.R. Holland, J.D. Baeder, K. Duraisamy, Field inversion and machine learning with embedded neural networks: physics-consistent neural network training, in: AIAA Aviation 2019 Forum, 2019, p. 3200.

[45] Y. Liu, L. Lu, L. Fang, F. Gao, Modification of Spalart-Allmaras model with consideration of turbulence energy backscatter using velocity helicity, Phys. Lett. A 375 (24) (2011) 2377–2381, https://doi.org/10.1016/j.physleta.2011.05.023.

[46] S. Medida, Correlation-based transition modeling for external aerodynamic flows, Ph.D. thesis, University of Maryland, College Park, 2014.

[47] B. Cui, X. Wang, R. Wang, Z. Xiao, Numerical investigation of transonic axial compressor rotor flows using an improved transition-sensitized turbulence model, Phys. Fluids 33 (3) (2021) 035149, https://doi.org/10.1063/5.0043633.

[48] X. Luo, A.S. Stordal, R.J. Lorentzen, G. Nævdal, Iterative ensemble smoother as an approximate solution to a regularized minimum-average-cost problem: theory and applications, SPE J. 20 (05) (2015) 962–982.

[49] C. Molnar, Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, Lulu.com, 2020.

[50] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, Curran Associates Inc., Red, Hook, NY, USA, 2017, pp. 4768–4777.

[51] K. Fujimoto, I. Kojadinovic, J.-L. Marichal, Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices, Games Econ. Behav. 55 (1) (2006) 72–99, https://doi.org/10.1016/j.geb.2005.03.002.

[52] C.A. Michelén-Ströfer, X.-L. Zhang, H. Xiao, DAFI: an open-source framework for ensemble-based data assimilation and field inversion, Commun. Comput. Phys. 29 (5) (2021) 1583–1622.

[53] J.D. Denton, Multall—an open source, computational fluid dynamics based, turbomachinery design system, J. Turbomach. 139 (12) (2017) 121001, https://doi.org/10.1115/1.4037819.

[54] https://github.com/slundberg/shap.

[55] W. Cui, X. Xiang, Q. Zhao, J. Xu, The effect of sweep on flowfields of a highly loaded transonic rotor, Aerosp. Sci. Technol. 58 (2016) 71–81, https://doi.org/10.1016/j.ast.2016.08.002.

[56] Z. Li, Y. Ju, C. Zhang, Quasi-wall-resolved large eddy simulation of transitional flow in a transonic compressor rotor, Aerosp. Sci. Technol. 126 (2022) 107620, https://doi.org/10.1016/j.ast.2022.107620.

[57] I.A. Johnsen, R.O. Bullock, Aerodynamic Design of Axial-Flow Compressors, Scientific and Technical Information Division, vol. 36, National Aeronautics and Space Administration, 1965.

[58] J. Heaton, Introduction to Neural Networks with Java, Heaton Research, Inc., 2008.

[59] X.-L. Zhang, H. Xiao, X. Luo, G. He, Combining direct and indirect sparse data for learning generalizable turbulence models, J. Comput. Phys. 489 (2023) 112272.

[60] C. Hah, Large eddy simulation of transonic flow field in NASA rotor 37, in: 47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, 2009, p. 1061.

[61] J. Joo, G. Medic, D. Philips, S. Bose, Large-eddy simulation of a compressor rotor, in: Proceedings of the Summer Program, Center for Turbulence Research, Stanford, CA, USA, 2014, p. 467.

[62] K. Yamada, M. Furukawa, T. Nakano, M. Inoue, K. Funazaki, Unsteady Three-Dimensional Flow Phenomena Due to Breakdown of Tip Leakage Vortex in a Transonic Axial Compressor Rotor, Turbo Expo: Power for Land, Sea, and Air, vol. 41707, 2004, pp. 515–526.

[63] H. Zhang, Y. Chen, Y. Lv, Development and validation of a combustion large-eddy-simulation solver based on fully compressible formulation and tabulated chemistry, Aerosp. Sci. Technol. 127 (2022) 107693, https://doi.org/10.1016/j.ast.2022.107693.

[64] K. Yamada, M. Furukawa, M. Inoue, K.-i. Funazaki, Numerical analysis of tip leakage flow field in a transonic axial compressor rotor, IGTC paper 95 (2003) 2003.

[65] W. Li, Y. Liu, Numerical investigation of corner separation flow using Spalart-Allmaras model with various modifications, Aerosp. Sci. Technol. 127 (2022) 107682, https://doi.org/10.1016/j.ast.2022.107682.