

Article

Surrogate Models for Performance Prediction of Axial Compressors Using through-Flow Approach

Xiaoxiong Wu ¹ , Bo Liu ^{1,*}, Nathan Ricks ² and Ghader Ghorbaniasl ²

¹ School of Power and Energy, Northwestern Polytechnical University, Xi'an 710129, China; Xiaoxiong.Wu@vub.be

² Department of Mechanical Engineering, Vrije Universiteit Brussel, 1050 Brussels, Belgium; Nathan.James.Ricks@vub.be (N.R.); Ghader.Ghorbaniasl@vub.be (G.G.)

* Correspondence: liubo704@nwpu.edu.cn

Received: 22 November 2019; Accepted: 22 December 2019; Published: 30 December 2019



Abstract: Two-dimensional design and analysis issues on the meridional surface, which is important in the preliminary design procedure of compressors, are highly dependent on the accuracy of empirical models, such as the prediction of total pressure loss model and turning flow angle. Most of the widely used models are derived or improved from experimental data of some specific cascades with low-loading and low-speed airfoil types. These models may work for most conventional compressors but are incapable of accurately estimating the performance for some specific cases like transonic compressors. The errors made by these models may mislead the final design results. Therefore, surrogate models are developed in this work to reduce the errors and replace the conventional empirical models in the through-flow calculation procedure. A group of experimental data considering a two-stage transonic compressor is used to generate the airfoils database for training the surrogate models. Sensitivity analysis is applied to select the most influential features. Two supervised learning approaches including support vector regression (SVR) and Gaussian process regression (GPR) are used to train the models with a Bayesian optimization algorithm to obtain the optimal hyper parameters. The trained models are integrated into the through-flow code based on streamline curvature method (SLC) to predict the overall performance and internal flow field of the transonic compressor on five rotational speed lines for validation. The predictions are compared with the experimental data and the results of conventional empirical models. The comparison shows that SVR and GPR respectively reduce the predicted error of empirical models by 62.2% and 55.2% for the total pressure ratio and 48.4% and 50.1% for adiabatic efficiency on average. This suggests that the surrogate models constitute an alternative way to predict the performance of airfoils in through-flow calculation where empirical models are inefficient.

Keywords: surrogate model; through-flow approach; axial compressors; empirical correlation; support vector regression; Gaussian process regression; total pressure ratio; deviation angle

1. Introduction

With the development of computation technology, the 3D computational fluid dynamics (CFD) approach has become a mainstream numerical simulation tool for compressors. However, during the preliminary design period, the 2D through-flow approach still plays an important role in design and analysis issues of compressors. Unlike the comprehensive but time-consuming 3D simulation, 2D approaches provide rapid and reliable estimation of the performance of compressors with few geometric parameters, before the geometry of blades is specified, and direct the subsequent steps. The popular through-flow approaches including streamline curvature method, stream function matrix method, or time marching method, to solve the simplified Navier-Stokes equations on the median flow surface

between adjacent blades with reasonable assumptions. The accuracy of these methods is determined by empirical correlations [1] which give the control equations a closed-form expression. In general, the empirical correlations originate from the qualitative boundary-layer theory and experimental data of some classical cascades, some of which are listed in [2]. They have a limited suitable operation extent while the accuracy becomes unreliable if they are applied to off-design conditions or high Mach number cases. Given that, many efforts have been devoted to improve the accuracy and application scope of the empirical correlations [3–5]. Besides the improvement works on the traditional correlations, Schmitz et al. proposed the thought of using surrogate models to predict the performance of airfoils [6], which inspired this work. They generated the airfoil database using S1-Solver MISES (a system for cascade analysis and design developed by MIT [7]), built the surrogate models by Bayesian neural networks, and integrated the Bayesian trained neural networks to S2-solver to evaluate a transonic 4.5-stage compressor [8]. In addition to the application in modeling the airfoil characteristics, the surrogate model technique has been widely used in other works: Nonlinear unsteady aerodynamic reduced-order models are developed for aeroelastic analysis of airfoils [9]; the convolutional neural network (CNN) is used to approximate the high-fidelity CFD models to predict the property of turbine vanes in an efficient and accurate way [10]; the adaptive machine learning framework, including functions of dynamic selection and self-tuning surrogate models, are developed to accelerate the optimization process of transonic compressors [11].

Surrogate models based on machine learning methods can solve the nonlinear issues such as the problem in through-flow calculations. This work aims to build surrogate models from experimental data of a certain transonic axial compressor to calculate total pressure loss and deviation angle of airfoils, instead of using conventional empirical models. The availability of predicting the performance and internal flow field by integrating the surrogate models with a through-flow program is also validated.

In brief, the procedure of building the surrogate models contains the following steps. A Kriging model is used to approximate the experimental data with all input parameters. Then, the most influential inputs are screened out based on the sensitivity analysis results of the Kriging model. The SVR or GPR models are trained with selected inputs while the hyper parameters of surrogate models are tuned by the Bayesian optimization algorithm. The experimental data is split into training and validation sets via a cross validation scheme, and the score of the validation set is set as optimization target during the optimization procedure. The final surrogate models are obtained by training the models on the whole data set with the optimal hyper parameters.

2. Blade Element Database

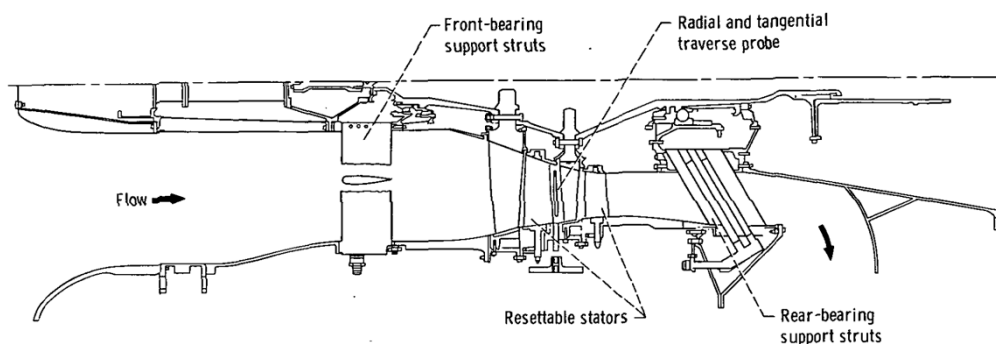
A reliable database, including the necessary parameters of blade elements, is an essential prerequisite to ensure the accuracy of the models. Empirical models are commonly derived from experimental data of 2D cascades within design conditions. Additional correlations for correction are required when empirical models are applied to estimate the performance of compressors since the internal flow field is more complex compared to cascades. However, this correction process inevitably introduces errors into the target values, which may mislead the prediction of overall performance. Thus, a set of experimental data of the real compressor is used directly to build the surrogate models. The experimental data are extracted from a two-stage axial transonic compressor developed by the NASA Lewis Research Center. The design parameters are briefly listed in Table 1 and can be referred to [12]. The experimental rotation speed and mass flow rate are corrected to the standard inlet condition at sea level by Equation (1). This compressor has a supersonic inflow condition in the tip zones of rotors with the existence of channel shock wave and subsonic flow field in the other zones. The airfoils of both rotors and stators were designed by the multiple-circular-arc (MCA) rule, which can represent the design concept of most the other transonic compressors developed in that age.

$$\begin{aligned} m_{correct} &= \frac{m\sqrt{T/288.15}}{P/101325} \\ \Omega_{correct} &= \frac{\Omega}{\sqrt{T/288.15}} \end{aligned} \quad (1)$$

Table 1. Main design parameters of the compressor.

Variable Name	Value
Overall Total Pressure Ratio	2.8
Overall Adiabatic Efficiency (%)	83.9
Tip Diameter (m)	0.787
Tip Speed (m/s)	441.96
Hub/Tip Ratio	0.4
Corrected Mass flow (kg/s)	83.5

The schematic of the test rig of this compressor is shown in Figure 1, and the other configuration information of the experiment can be found in [13]. Nine probes are distributed non-uniformly along the span of each blade row from 0.05 to 0.95 of span length, thus each test case contains 36 sets of blade element data for the 4-blade row, including measured aerodynamic parameters. There are 35 test cases conducted with rotational speed varying from 50% to 112% and 25 of them have detailed blade element data in the report. The 900 measurements of blade element data from the 25 test cases are used to construct the training samples for the next section. Based on the features of most empirical models, each sample includes 6 aerodynamic and 7 geometric parameters as input variables. The structure of the surrogate models is shown in Figure 2 in which the geometric and aerodynamic parameters are defined as: SPAN represents the fraction of blade height from hub; STAGGER is the angle between the chord line and axial direction; SOLIDITY is the ratio of chord length to pitch; T/C means the ratio of maximum thickness to chord length; A/C is the fractional distance of the maximum thickness location from the leading edge; CAM is the camber angle; LET/C is the ratio of the thickness of the leading edge to chord length; INC indicates the incidence angle between inflow angle and inlet metal angle; B1 is the inflow angle; M1 is the inflow Mach number; DF is the diffusion factor which represents the loading of the cascade; MVDR refers to the meridional velocity-density ratio as $\rho_2 V_{m2} / \rho_1 V_{m1}$ where 1 and 2 represent the inlet and outlet respectively; REYNOLD means the Reynolds number of the cascade. The total pressure loss ω and deviation angle δ are the outputs of the surrogate models. Figure 3 depicts the distribution of deviation angle and total pressure loss over incidence angle and Mach for the 900 samples in the blade element database. The variety parameters cover a wide working range of typical transonic compressors.

**Figure 1.** Schematic of two-stage fan test arrangement [13].

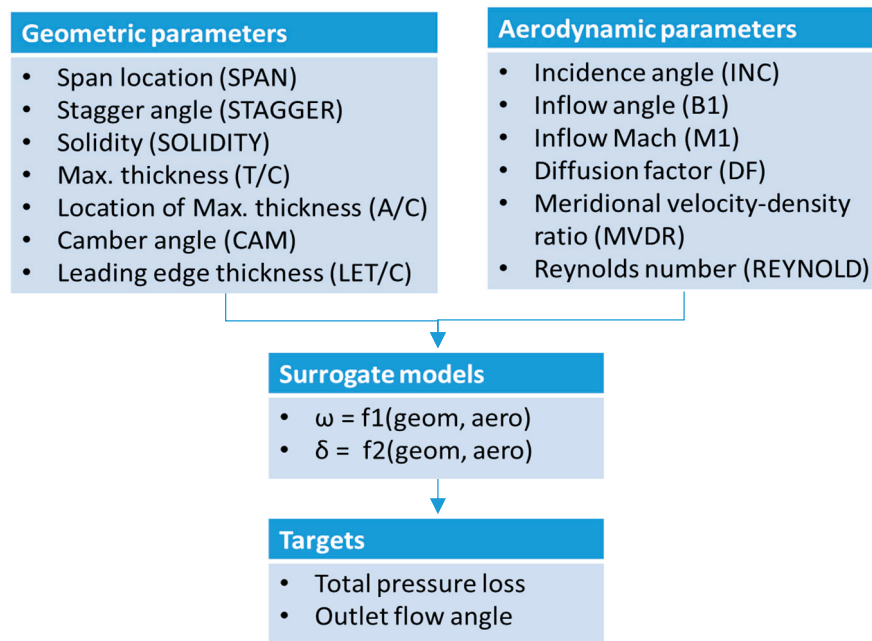


Figure 2. Structure of surrogate models.

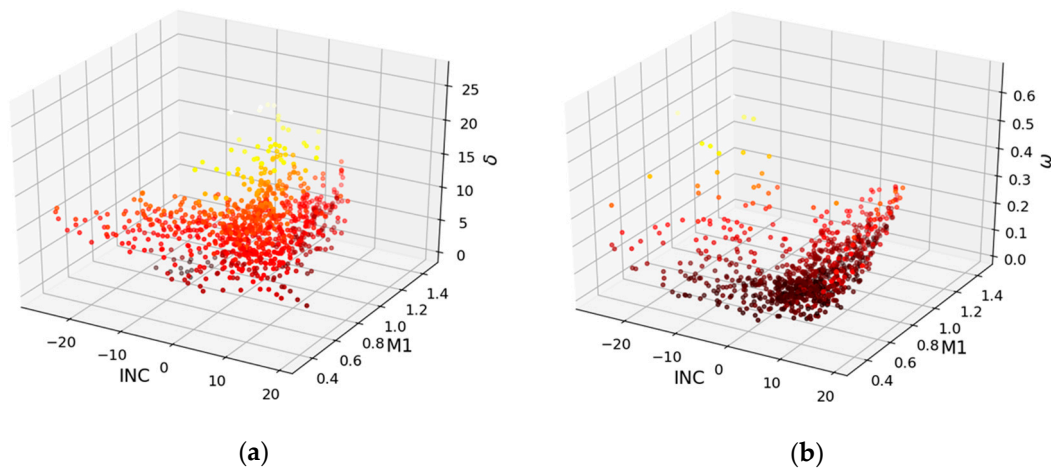


Figure 3. Distribution of experimental data: (a) The distribution of deviation angle; (b) The distribution of total pressure loss.

3. Method for Building Surrogate Models

The steps for building the surrogate models for total pressure loss and deviation angle (flowchart shown in Figure 4) are as follows:

1. Pre-process the experimental data: A Kriging approach is implemented to create a response surface model based on the blade element database with all the features as input parameters. Sensitivity analysis method is applied on the response surface to estimate the elementary effect of each input feature. The features with a greater effect on the objects are screened out as input parameters for training either the total pressure loss model or deviation angle model.
2. Train models and tune hyper parameters: support vector regression and Gaussian process regression are used to train the surrogate models, respectively, with the input parameters determined by Step 1. The optimal models are obtained by tuning the hyper parameters in these models using a Bayesian optimization approach. The coefficients of determination (R^2) of the cross-validation results are used as criteria to evaluate the models during the optimization process.
3. Post-process with validation: Validate the final models using the whole experimental database.

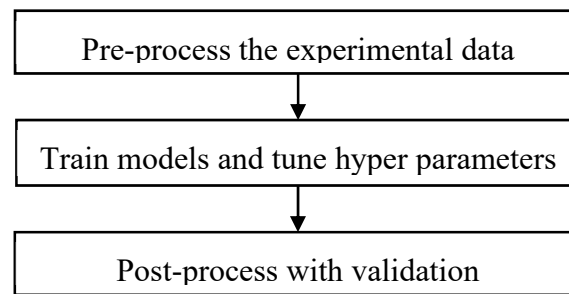


Figure 4. Procedure of models generation.

3.1. Features Selection

For the two targets of total pressure loss and deviation angle, either the mapping relationships or the elementary effects between input features and targets is different. The noise may be brought into the models and mislead the prediction if non-influential variables are selected as input. Therefore, the sensitivity analysis method is required to screen out the features which have important effects on the targets before training the surrogate models. An elementary effects (EE) method proposed by Morris [14] is applied here to select the necessary input parameters for the two models respectively. To simply exemplify the EE method, let the models have the form as Equation (2) with k input factors, where X represents for the geometric and aerodynamic parameters, and Y for the total pressure loss or deviation angle.

$$Y = f(X_1, X_2 \dots X_k) \quad (2)$$

Morris provides two sensitivity measures for each input factor: μ to assess the importance of the input factor on the output, and σ to describe the non-linear effects and interactions. These two measures are obtained from a design based on the construction of a series of trajectories in the space of the inputs, where inputs are randomly moved one-at-a-time (OAT). In this design, each model input is assumed to vary across p selected levels in the space of the input factors. The region of space Ω is thus a k -dimensional p -level grid. Each trajectory is composed of $k + 1$ points since input factors move one by one of a step Δ in $\{0, 1/(p-1), 2/(p-1), \dots, 1\}$ while all others remain fixed. Along each trajectory, the elementary effect for each input factor is defined as:

$$d_i(X) = \frac{Y(X_1, \dots, X_{i-1}, X_i + \Delta, X_{i+1} \dots X_k) - Y(X)}{\Delta} \quad (3)$$

where $X = (X_1, X_2, \dots, X_k)$ is any selected value in Ω such that the transformed point is still in Ω . Further, r elementary effects are estimated for each input $d_i(X^{(1)})$, $d_i(X^{(2)})$, \dots , $d_i(X^{(r)})$ by randomly sampling r points $X^{(1)}$, $X^{(2)}$, \dots , $X^{(r)}$. The two measurements μ and σ are respectively the mean and standard deviation of the distribution of the elementary effects of each input:

$$\mu_i = \frac{1}{r} \sum_{j=1}^r d_i(X^{(j)}) \quad (4)$$

$$\sigma_i = \sqrt{\frac{1}{r-1} \sum_{j=1}^r (d_i(X^{(j)}) - \mu_i)^2} \quad (5)$$

These two measures rank the input factors in order of influence on the target. The input variable with both low μ_i and σ_i can be confirmed as a non-influential factor. A low μ may occur when the

model is not monotonic and elementary effects with opposite sign cancel each other out. An improved method was proposed by Campolongo [15] to use the absolute value of $d_i(X^{(j)})$ to calculate μ :

$$\mu_i = \frac{1}{r} \sum_{j=1}^r |d_i(X^{(j)})| \tag{6}$$

A Kriging approach [16] is implemented to create a response surface model, giving all the geometric and aerodynamic factors as input based on the blade element database since the exact form of mapping function is unknown. The optimal model for sensitivity analysis is obtained by the tuning of hyper parameters using an optimization algorithm described in the next section. For this case, k is 13 while r and p are assigned to 1000 and 4, leading to 1000 trajectories and 4 levels on each. The distribution of elementary effect of each input factors are shown in Figure 5, with ranking based on importance. The results indicate that the influential features for each target are not consistent. The top factors where the sum of elementary effects occupy greater than 90% are selected as inputs. The final screened inputs for models of total pressure loss and deviation are listed in Table 2.

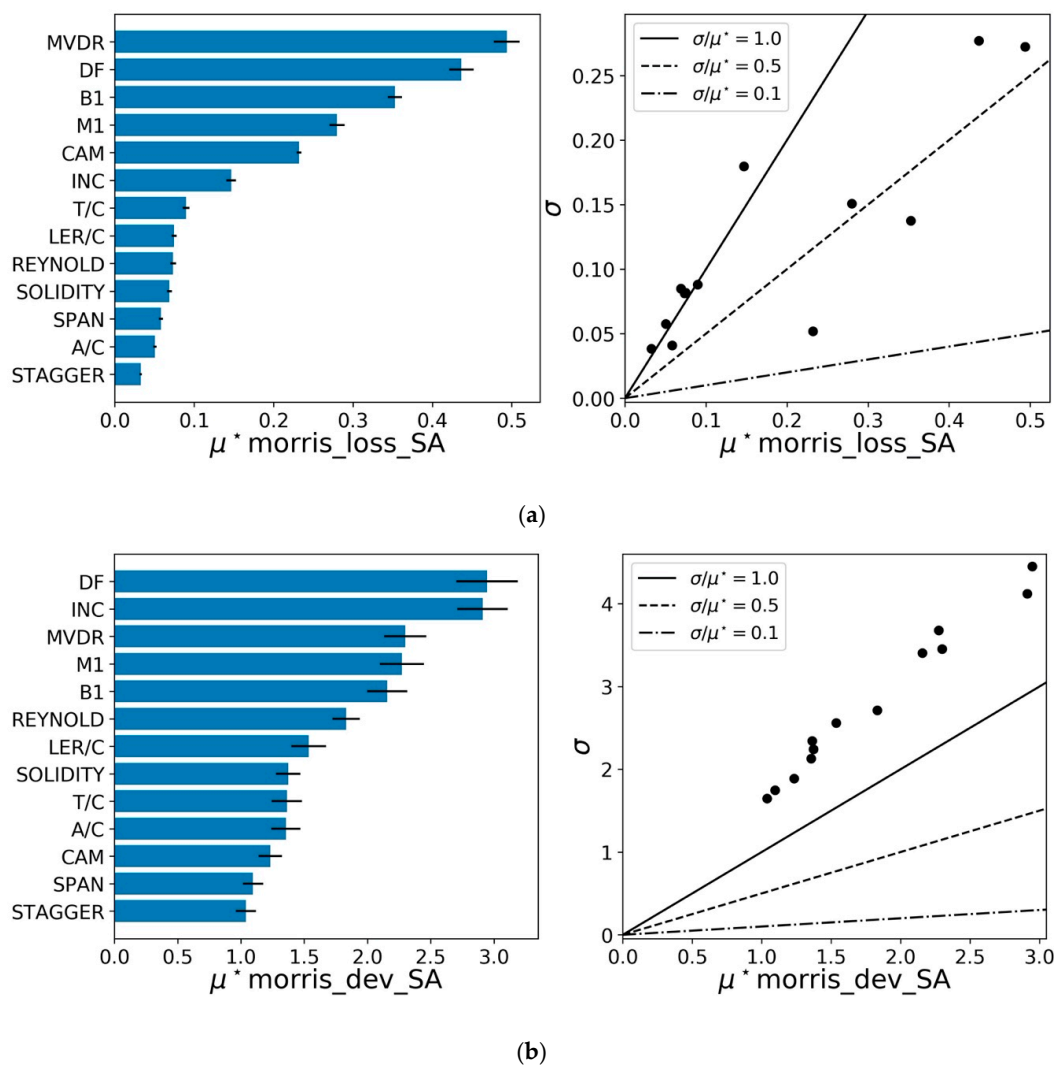


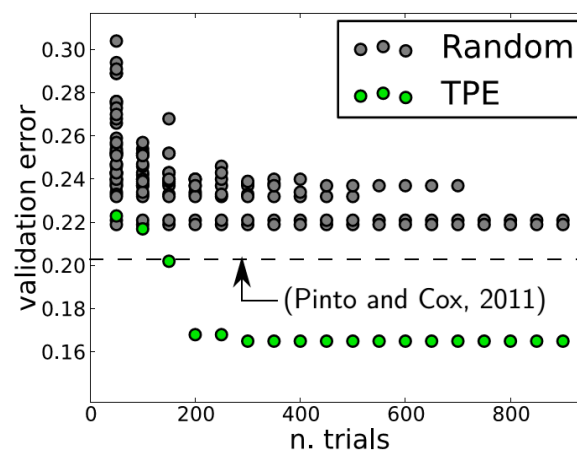
Figure 5. Sensitivity analysis of blade element database: (a) Elementary effect distributions of total pressure loss; (b) Elementary effect distributions of deviation angle.

Table 2. Inputs for either model.

Model	Inputs
ω	MVDR, DF, B1, M1, CAM, INC, T/C, LER/C, REYNOLD
δ	INC, DF, MVDR, M1, B1, REYNOLD, LER/C, A/C, CAM, T/C, SOLIDITY

3.2. Hyper Parameters Optimization

The performance of surrogate models is strongly related to the configuration of the internal hyper parameters, such as learning rates, kernel functions, or regularization factors, besides the reliable samples. It is necessary to tune the hyper parameters to make sure that the models have optimal performance. Various search strategies are thus developed to find the best hyper parameter. For the cases with just a few hyper parameters, a manual search can easily be used to find the optimal values by experienced experts, but it depends greatly on prior knowledge. The most common automatic search strategies include grid search and random search. Grid search is a simple strategy for the investigation of user-defined points. It may lead to either a miss of the optimal value if too few points are specified, or excessive resource consumption if every possible combination of hyper parameters is considered, especially for complicated cases. Random search [17] was developed to overcome the drawbacks of the grid search and successfully used in many applications. It requires fewer random samples than grid search, fewer computational resources, and performs more efficiently. Each search process of either grid or random search is independent to the others, since they are both uninformed search algorithms, thus the previous results of hyper parameters cannot guide the next evaluation. These algorithms can be applied to a variety of search problems because the target information is not needed, however, at the cost of search efficiency. Therefore, an informed search algorithm using Bayesian optimization with higher search efficiency [18] (shown in Figure 6, dash denotes hand-tuned) is used optimize the hyper parameters.

**Figure 6.** Performance of different search techniques on data sets LFW [18].

Bayesian optimization, based on the sequential model-based global optimization [18,19] (SMBO) algorithm that uses information from previous trials as advice for future exploration, has been applied in the modeling cases where the calculation of a fitness function is expensive [20,21]. The true fitness function of minimization problem $f : \chi \rightarrow \mathbb{R}$:

$$x^* = \operatorname{argmin}_{x \in \chi} f(x) \quad (7)$$

within domain $\chi \in \mathbb{R}^d$ is considered to solve the optimal hyper parameters of x^* . SMBO approximates f with surrogate models, usually as Gaussian Process (GP) or Tree-structured Parzen Estimator

(TPE) approach since the true form is unknown. The schematic of the SMBO algorithm is shown in Algorithm 1.

Algorithm 1 The pseudo-code of generic sequential model-based optimization.

```

1: Input :  $f, \chi, S, M$ 
2:  $D \leftarrow \text{InitialSamples}(f, \chi)$ 
3: for  $i \leftarrow |D|$  to  $T$  do
4:    $p(y|x, D) \leftarrow \text{FitModel}(M, D)$ 
5:    $x^* \leftarrow \text{argmax}_x S(x, p(y|x, D))$ 
6:    $y^* \leftarrow f(x^*) \triangleright \text{Expensive step}$ 
7:    $D \leftarrow D \cup (x^*, y^*)$ 
8: end for

```

Typically, a small set of samples from χ initialize the probabilistic regression model (or prior function) M as cheap surrogate for the expensive fitness function f . Then a new set of values x_i is obtained by optimizing the acquisition function S . A new evaluation of y_i corresponds to x_i , and is sequentially obtained by performing f directly and appended to the evaluation history set of the best locations and targets $D = \{(x_1, y_1), \dots, (x_i, y_i)\}$. The model M refits the updated record D and repeats the last steps for T times. The criterion of expected improvement (EI) [22] is chosen as S because it is intuitive and well performing. Expected improvement is the expectation for f when the evaluation of $f(x)$ negatively exceeds a specified threshold y^* :

$$EI_{y^*}(x) := \int_{-\infty}^{\infty} \max(y^* - y, 0) p_M(y|x) dy \quad (8)$$

EI is always non-negative, either zero for y more than y^* , or positive when y is less than y^* , and the minimum y corresponds to the maximum EI. Herein the higher budget TPE is selected to approximate f . Unlike GP that models the prior $p(y|x)$ directly, the TPE defines $p(x|y)$ using two non-parametric densities:

$$p(x|y) = \begin{cases} l(x) & \text{if } y < y^* \\ g(x) & \text{if } y > y^* \end{cases} \quad (9)$$

From Bayes theorem, we get:

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} \quad (10)$$

Considering $\gamma = p(y < y^*)$, the $p(x)$ can be expressed as follows:

$$p(x) = \int_{\mathbb{R}} p(x|y)p(y)dy = \gamma l(x) + (1 - \gamma)g(x) \quad (11)$$

Putting Equations (9)–(11) into (8), the EI based on the TPE algorithm is given by:

$$EI_{y^*}(x) = \frac{\gamma y^* l(x) - l(x) \int_{-\infty}^{y^*} p(y) dy}{\gamma l(x) + (1 - \gamma)g(x)} \propto \left(\gamma + \frac{g(x)}{l(x)}(1 - \gamma) \right)^{-1} \quad (12)$$

This expression indicates that x with high probability under $l(x)$ and low $g(x)$ points to the maximum EI. The x which gives the greatest EI is x^* , and the optimal hyper parameters are the best x^* in the observation history D . The open source optimization package Hyperopt developed by Bergstra et al. [18] is used as optimizer for the hyper parameters.

3.3. Modeling by Support Vector Regression

Support vector machine (SVM) [23] is a type of supervised learning method usually used for classification (SVC) [24] or regression (SVR) [25,26]. Its high efficiency in processing high dimensional data spaces, and ability to use only a subset of samples (support vectors) in the decision function, has allowed a wide application of SVM in solving real world problems. Attributed to the assistance of kernel functions (called kernel trick) by mapping the original sample space \mathbb{R} to a higher or infinite dimensional space \mathbb{R}^n (Hilbert space), while effectively avoiding the probable curse of dimensionality without complicating the calculation, SVM can easily handle the nonlinear sample data on higher dimensional space by using a linear method (optimal hyper plane). It's flexible to adjust SVM performance with versatile kernel functions for specific problem. Kernel functions such as linear, polynomial, RBF, and sigmoid are commonly used. In this work RBF is chosen as the kernel function with the form of:

$$k(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (13)$$

The hyper parameter γ denoting the effect range of RBF function should be optimized by SMBO. It works on the Euclidean distance $\|x - x'\|$ between x and x' . For the given training vectors $x_i \in \mathbb{R}^p, i = 1, \dots, n$ and the target vector $y \in \mathbb{R}^n$, the following target function of SVR needs to be solved:

$$\begin{aligned} \min_{w, b, \zeta, \zeta^*} & \frac{1}{2} w^T w + C \sum_{i=1}^n (\zeta_i + \zeta_i^*) \\ \text{s.t.} & y_i - w^T \phi(x_i) - b \leq \varepsilon + \zeta_i \\ & w^T \phi(x_i) + b - y_i \leq \varepsilon + \zeta_i^* \\ & \zeta_i, \zeta_i^* \geq 0, i = 1, \dots, n \end{aligned} \quad (14)$$

This is a convex quadratic program problem with linear constraints. An equivalent dual formulation for simplifying calculation is transformed by introducing the addition of Lagrange multipliers based on Lagrange duality. The dual function is:

$$\begin{aligned} \min_{\alpha, \alpha^*} & \frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \varepsilon e^T (\alpha + \alpha^*) - y^T (\alpha - \alpha^*) \\ \text{s.t.} & e^T (\alpha - \alpha^*) = 0 \\ & 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, \dots, n \end{aligned} \quad (15)$$

where e is the unit vector, penalty coefficient C and margin ε are the other two hyper parameters requiring optimization. C is dedicated to avoid over-fitting as the regularization approach, and ε is the margin to tolerate the noise. Q is a $n \times n$ semi-definite matrix given by:

$$Q_{n \times n} = \begin{bmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ \vdots & \ddots & \vdots \\ K(x_n, x_1) & \cdots & K(x_n, x_n) \end{bmatrix}, \quad \text{with } K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (16)$$

where K is the RBF kernel function, ϕ is the mapping to high dimensional space and n is the amount of samples. This dual optimization problem can be solved by sequential minimal optimization (SMO) algorithm [27] whose structure is briefly shown in Algorithm 2. The decision function (formulation of surrogate models) is as follows, where ρ indicates the interception of model.

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + \rho \quad (17)$$

During this modeling task, the SMBO algorithm works on the domain of hyper parameters (first 3 columns in Table 3) defined by range of exponential function $e^{(a,b)}$ for 100 iteration steps, where $[a, b]$ is a uniform distribution.

Algorithm 2 Pseudo-code of Sequential Minimal Optimization algorithm

-
- 1: Input : samples $x_i, y_i, i = 1, \dots, n$
 - 2: Output : Lagrange multipliers α_i
 - 3: Initialize $\alpha_i = 0$
 - 4: **while** not meet KKT condition **do**
 - 5: choose a pair of α_i and α_j by heuristic algorithm
 - 6: fix the other α and solve the optimization problem to get α_i and α_j
 - 7: express α_j by α_i
 - 8: **end while**
-

Table 3. Search domain and optimal values of hyper parameters.

Hyper Paras	Lower Bound	Upper Bound	DEV OPT	LOSS OPT
γ	e^{-7}	e^2	8.305×10^{-2}	3.741×10^{-2}
C	e^{-4}	e^4	4.024×10^1	1.342×10^0
ε	e^{-7}	e^2	9.386×10^{-2}	4.649×10^{-3}

The coefficient of determination R^2 is used as the score of fitness function f (y_i means the true target value, \bar{y} denotes the mean value of y_i and f_i is the prediction values by SVR). The target of optimization is to find an optimal set of hyper parameters maximizing the R^2 in which high value means good fit.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (18)$$

The surrogate models are usually over-fitted if the validation set is consistent with the trained one. The performance of models may deteriorate if the training data is simply split into two parts of training data and validation set, since the training information does not cover both sets. The cross validation [28] (CV) method is therefore developed to improve this situation. By splitting blade element database to 5 sets randomly of similar size, where one set is for validation, and the other four are for training sets with 5 loops. The mean of the 5 R^2 of validation sets is taken as the criteria for evaluation of models in this work.

Figure 7 shows the hyper parameters optimization history records for SVR models of the deviation angle and total pressure loss. From observing the history of locations in Figure 7a,b, each hyper parameter has reached the optimal value (last 2 columns in Table 3) within the search domain corresponding to the highest R^2 of the validation set. The optimization history curve in Figure 7c indicates both models approached the optimal target within 40 iterations.

The trained models are validated by the whole blade element data set and the results shown in Figure 8. The abscissa denotes the experimental data while the ordinate means prediction values. Deviation angle model has a higher R^2 , very close to 1, as compared to the total pressure loss model, with $R^2 = 0.788$. It is observed that the loss model underestimated the total pressure loss in some cases with high values. The high total pressure loss indicates that the airfoil works on inferior flow conditions, like near stall or choke operations, which are far from the designed working region. For conditions in the normal working range, the loss model is deemed reliable. Furthermore, considering that the former has more importance than the latter in through-flow code, the R^2 of loss is acceptable. More work on the surrogate model should be developed in future to improve the loss prediction.

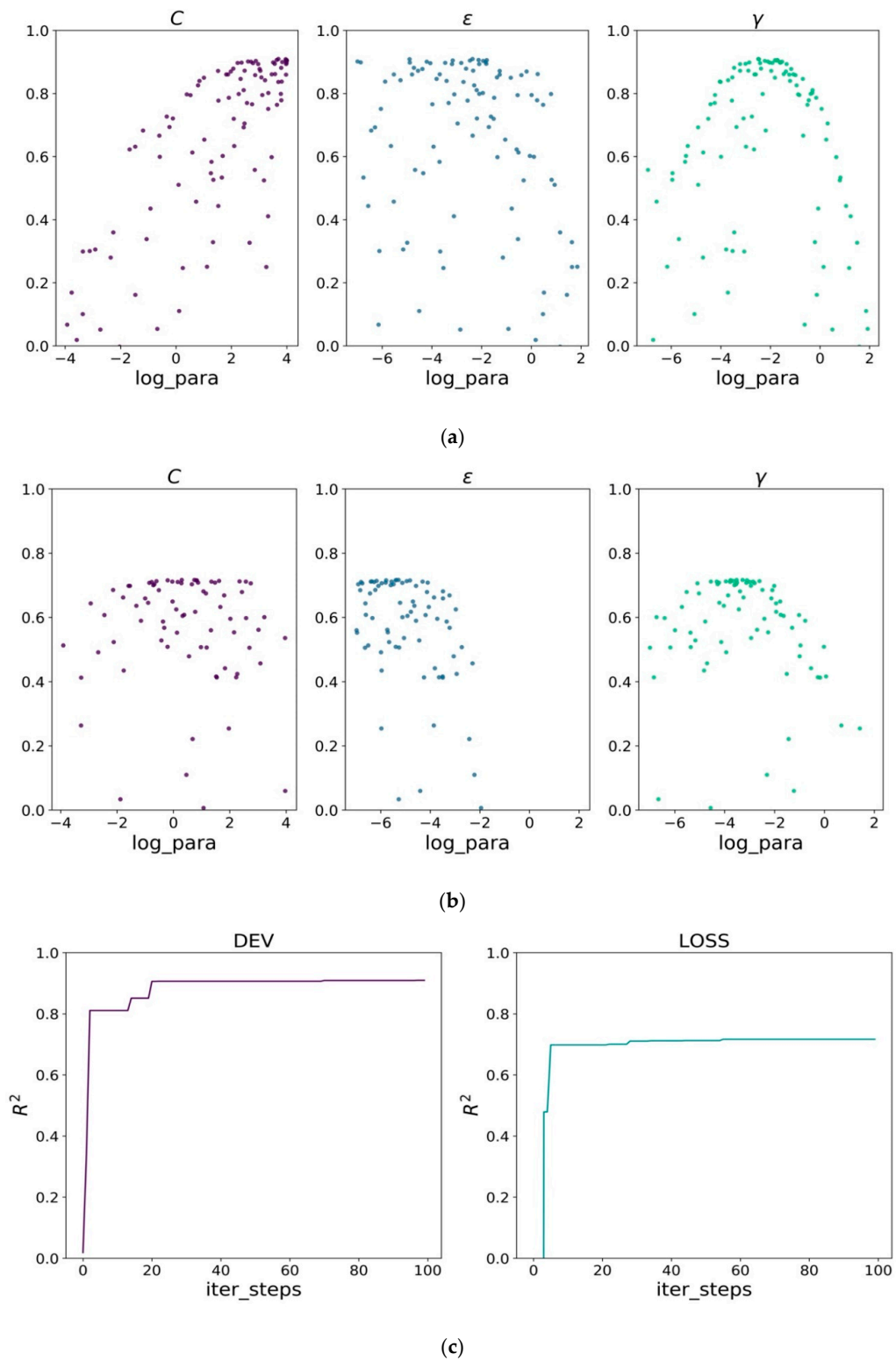


Figure 7. Optimization of training SVR models: (a) Hyper parameters search of deviation angle model; (b) Hyper parameters search of total pressure loss; (c) Optimization curve of models.

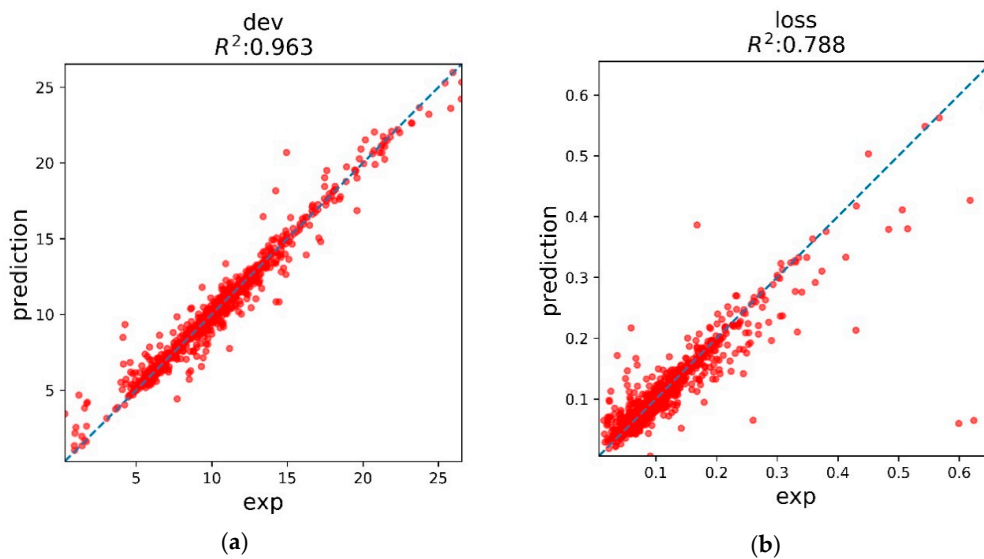


Figure 8. Validation of SVR models on the whole data set: (a) deviation angle; (b) total pressure loss.

3.4. Modeling by Gaussian Process Regression

Similar to SVR, Gaussian process regression [29] is another supervised learning method for regression and probabilistic classing problems. Its advantages include the prediction's interpolation of the observations, the prediction is probabilistic (Gaussian) so that the prediction can be improved in some concerned region based on the empirical confidence intervals, and can use various kernel functions, as in SVR. However, GPR loses efficiency in high dimensional spaces.

From the function view, Gaussian process (GP) is used to describe a distribution over functions considering Bayesian inference directly in function space. GP is a set of stochastic variables with any combination of which has multivariate Gaussian distributions [29]. Its properties depend completely on the mean function $m(x)$ and the covariance function $k(x, x')$:

$$\begin{cases} m(x) = E[f(x)] \\ k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))] \end{cases} \quad (19)$$

where $x, x' \in \mathbb{R}^d$ denotes the stochastic variables, and GP can be defined as:

$$f(x) \sim GP(m(x), k(x, x')) \quad (20)$$

For simplicity, the mean function of the variables is usually preprocessed to zero. Considering the regression problem, we have the generic model as follows:

$$y = f(x) + \varepsilon \quad (21)$$

where x denotes the input vectors, f is the function, while y is the observation target values polluted by noise ε . Further considering the noise subjected to standard normal distribution $\varepsilon \sim N(0, \sigma_n^2)$, the prior distribution of y and the joint prior distribution of y and prediction f are respectively shown:

$$\begin{aligned} y &\sim N(0, K(X, X) + \sigma_n^2 I_n) \\ \begin{bmatrix} y \\ f_* \end{bmatrix} &\sim N\left(0, \begin{bmatrix} K(X, X) + \sigma_n^2 I_n & K(X, x_*) \\ K(x_*, X) & k(x_*, x_*) \end{bmatrix}\right) \end{aligned} \quad (22)$$

In the formulas, $K(X, X)$ is a $n \times n$ definite and symmetric covariance matrix composed by element $k_{i,j} = k(x_i, x_j)$ measuring the correlation between x_i and x_j . $K(X, x_*)$ equaling to $K(x_*, X)^T$ indicates the

$n \times 1$ covariance matrix between test point x_* and input of training set X . $k(x_*, x_*)$ is the covariance of test point itself, and I_n is the identity matrix. The posterior of f_* can be calculated as:

$$f_*|X, y, x_* \sim N(\bar{f}_*, \text{cov}(f_*)) \quad (23)$$

where

$$\begin{aligned} \bar{f}_* &= K(x_*, X)[K(X, X) + \sigma_n^2 I_n]^{-1} y \\ \text{cov}(f_*) &= k(x_*, x_*) - K(x_*, X)[K(X, X) + \sigma_n^2 I_n]^{-1} K(X, x_*) \end{aligned}$$

where $\hat{\mu}_* = f_*$ and $\hat{\sigma}_{f_*}^2 = \text{cov}(f_*)$ are the mean and variance of the prediction corresponding to x_* , respectively. GPR has different selections of covariance function (kernel), where the most commonly used is the squared-exponential covariance function with the following form:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2}(x - x')^T M^{-1}(x - x')\right) \quad (24)$$

where $M = \text{diag}(l^2)$ and l is the so-called length-scale (analogic to γ in RBF), while σ_f^2 is the signal variance. The hyper parameters set $\theta = \{l, \sigma_f^2, \sigma_n^2\}$ can be achieved by the optimization of the gradient of negative log marginal likelihood $L(\theta) = -\log p(y|X, \theta)$ with quasi-Newton methods:

$$\begin{aligned} L(\theta) &= \frac{1}{2} y^T C^{-1} y + \frac{1}{2} \log|C| + \frac{n}{2} \log 2\pi \\ \frac{\partial L(\theta)}{\partial \theta} &= \frac{1}{2} \text{tr}\left((\alpha \alpha^T - C^{-1}) \frac{\partial C}{\partial \theta_i}\right) \end{aligned} \quad (25)$$

where

$$C = K_n + \sigma_n^2 I_n, \alpha = C^{-1} y$$

Alternatively, the SMBO algorithm is used here to optimize the hyper parameters for comparison with the SVR model. The search domain is listed in Table 4 (first three columns). The configuration of iteration steps, score criteria, and CV scheme are equivalent with the SVR model.

Table 4. Search domain for hyper parameters.

Hyper Paras	Lower Bound	Upper Bound	DEV OPT	LOSS OPT
σ_f^2	e^{-4}	e^4	9.605×10^0	5.248×10^1
l	e^{-4}	e^4	1.506×10^1	1.739×10^1
σ_n^2	e^{-7}	e^2	1.779×10^0	3.046×10^{-3}

Figure 9a,b describes the hyper parameters tuning procedure of training GPR models for both targets. As in the SVR cases, both models have reached the highest R^2 over optimal hyper parameters (last 2 columns in Table 4) within the search domain. The optimization record curve in Figure 9c shows that both models converge in 20 iterations, which is faster than SVR. From Figure 9a,b, it is observed that σ_n^2 affects the R^2 more than the other two. This indicates that the prior distribution of noise is the main influencing factor for the GPR model's performance.

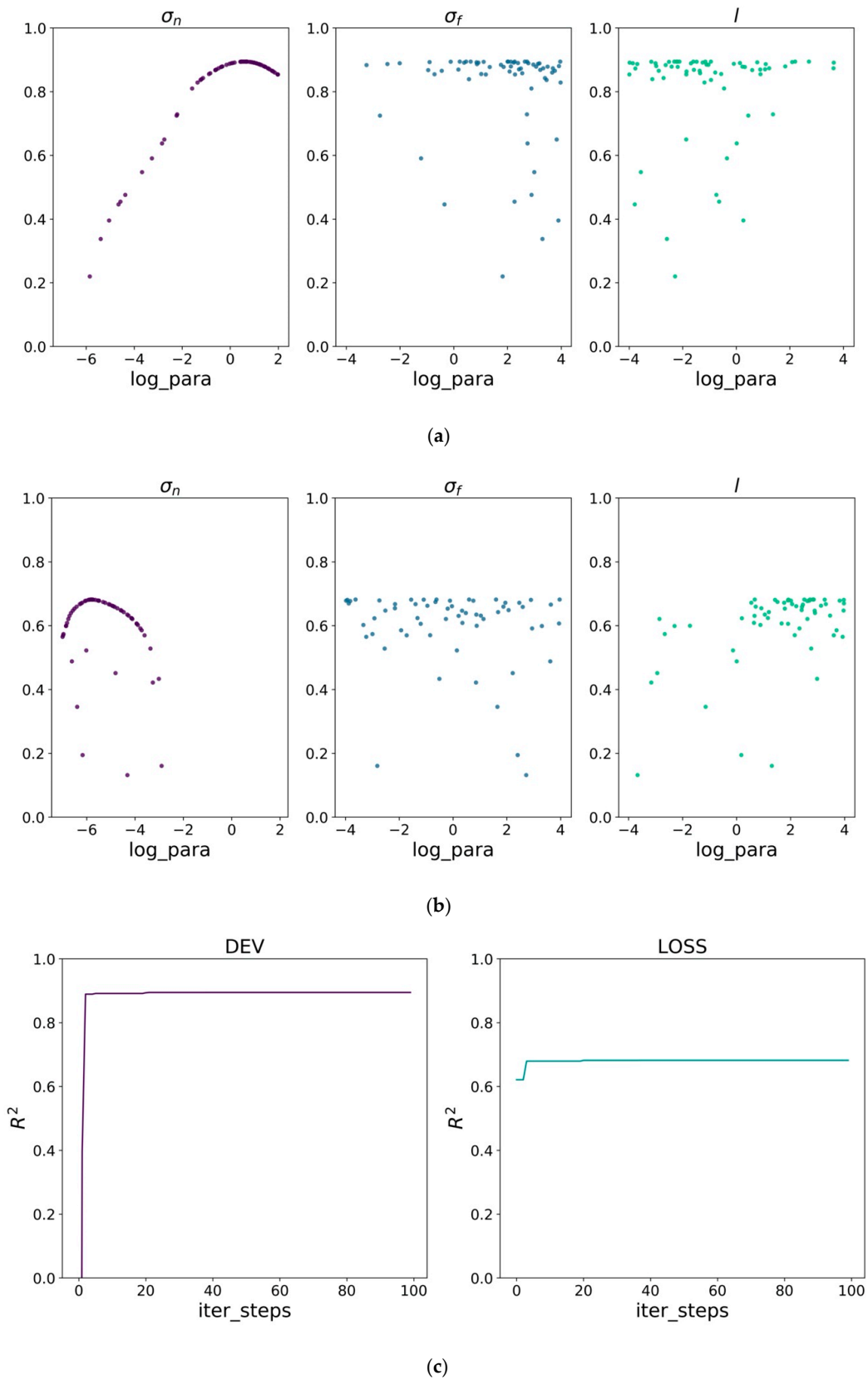


Figure 9. Optimization of training GPR models: (a) Hyper parameters search of deviation angle model; (b) Hyper parameters search of total pressure loss; (c) Optimization curve of models.

Figure 10 shows the comparison of prediction values of trained GPR models with experimental data. The R^2 of GPR validation results is close to the SVR results, which demonstrates that the two surrogate models perform nearly the same on the training data. In the next section, both models are applied in through-flow code to predict both design and off-design operation performance of the compressor.

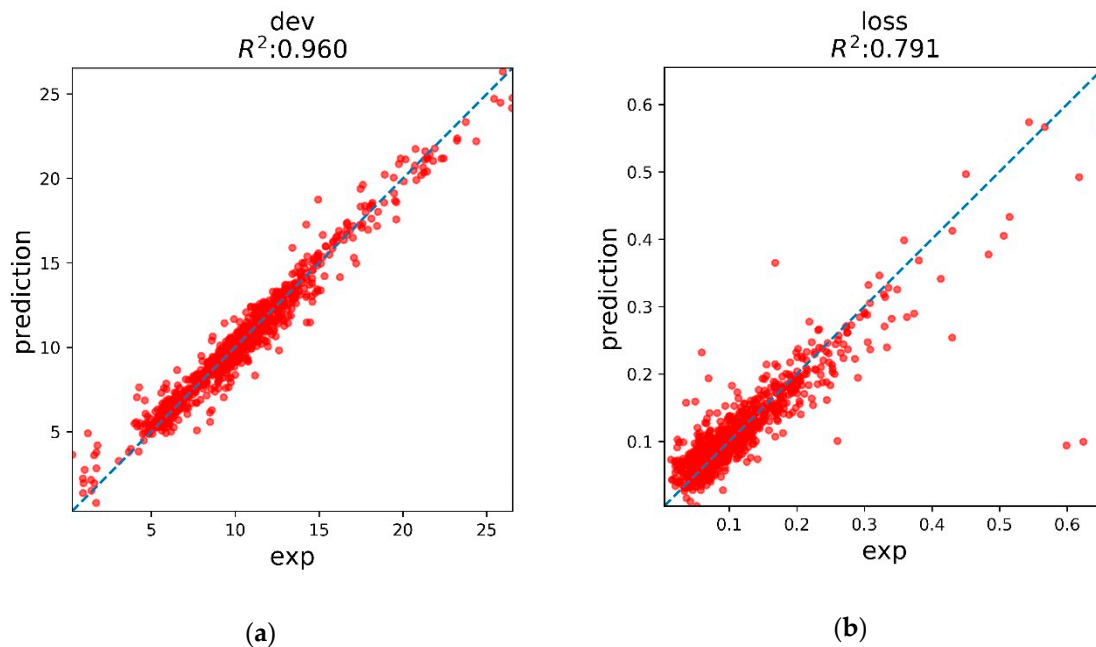


Figure 10. Validation of GPR models on whole data set; (a) deviation angle; (b) total pressure loss.

4. Results and Discussion

4.1. Streamline Curvature Approach Based through-Flow Program

The dominated through-flow method streamline curvature approach is implemented here to simulate the flow field in the axial compressor, as well as its overall performance. Based on the general through-flow theory [30,31] and actively developed since the 1960s [32], SLC has been successfully applied to the design and analysis of various types of turbomachinery. It simplifies the computational complexity greatly from real 3D flow simulations in blade channels to solve the Euler equations on 2D (or quasi-3D) flow surfaces (Figure 11) on the premise of some assumptions including axial symmetric flow, adiabatic, steady, etc. The governing formulas are composed of a radial equilibrium equation and massflow continuity equation, with a constraint condition that the meridional velocity should be subsonic:

$$\frac{\partial V_m}{\partial n} = \frac{1}{V_m} \left(\frac{\partial I}{\partial n} - T \frac{\partial S}{\partial n} - \frac{V_{\theta} - \omega r}{r} \frac{\partial r V_{\theta}}{\partial n} \right) - V_m \left(\frac{\cos(\varphi + \phi)}{r_m} - \frac{\sin(\varphi + \phi)}{V_m} \frac{\partial V_m}{\partial m} \right) \quad (26)$$

$$G = \int_{hub}^{tip} 2\pi r \rho V_m \cos(\varphi + \phi) dn$$

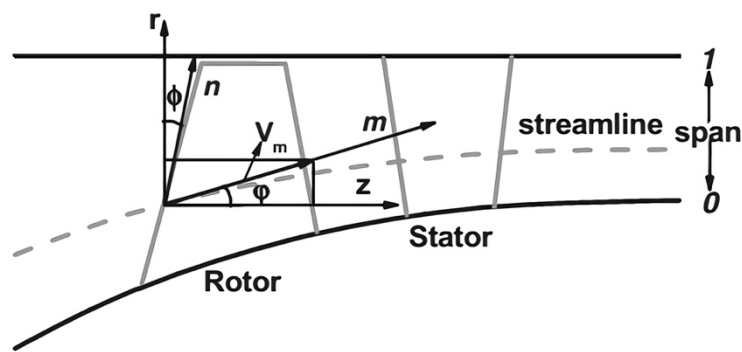


Figure 11. Schematic of computation domain.

The computations are performed on the mesh grid formed by streamlines and quasi-orthogonal stations. The discrete governing equations are solved by finite difference with specified boundary conditions for loops until the location of streamlines do not change any more. A private SLC based through-flow program developed by the authors is applied for this work [33]. The structure of this program is shown in Figure 12. The entropy S and swirl rV_θ in the governing equations are updated by the total pressure loss and deviation angle formed below when solving the governing equations.

$$\omega = \frac{(P_{w2}^*)_{id} - (P_{w2}^*)}{P_{w1}^* - P_1} \tag{27}$$

$$\delta = \beta_2 - \beta_{2k}$$

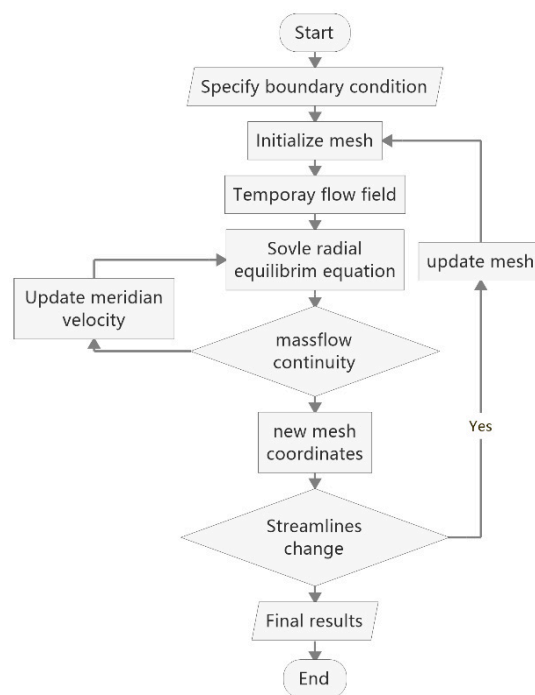


Figure 12. Structure of through-flow program.

4.2. Comparison with Empirical Models

With the development of SLC, versatile empirical correlations of loss and deviation angles are proposed and improved by countless experts and engineers who design turbomachinery by SLC in

practice. Some typical and widely used correlations are applied in this program to form the loss and deviation angle for comparison with surrogate models:

$$\begin{aligned}\delta &= \delta^* + \delta_i + \delta_{avdr} \\ \omega_{tot} &= \omega_{prof} + \omega_{sh} + \omega_{sec}\end{aligned}\quad (28)$$

The deviation angle δ contains the correlations of design angle δ^* [34], off-design correction δ_i [35] and AVDR correction δ_{avdr} [36] while the total pressure loss considers airfoil profile loss ω_{prof} [37], shock loss ω_{sh} [38], and secondary flow loss ω_{sec} [38]. These correlations have the similar inputs as the surrogate models.

The research object is the transonic compressor described at the beginning. The through-flow program coupling with either empirical models or surrogate models are respectively run for 5 rotational schemes (50%, 70%, 85%, 95%, 100% of design rotational 10,720 rpm), corresponding to the experimental cases. Figure 13 shows the meridional computation grids for this compressor, with 12 streamlines along the spanwise direction, where the first and last streamline coincide with hub and shroud. 18 quasi-orthogonal stations are distributed in the axial direction, with 8 stations coinciding with the rotor/stator blades leading and trailing edges, and the others for the inlet and outlet channel.

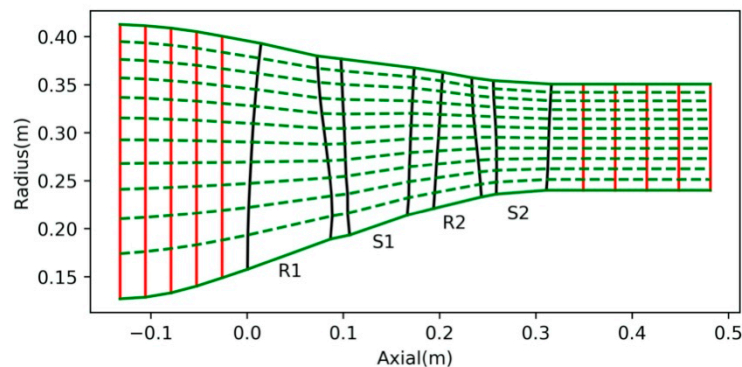


Figure 13. Computation grids for the two-stage transonic compressor.

Figure 14 shows the overall performance maps of total pressure ratio (Figure 14a) and adiabatic efficiency (Figure 14b) computed by different types of models for the five rotational speed schemes. The experimental characteristic lines are also plotted for reference. SVR, GPR, EMP, and EXP represent SVR models, GPR models, empirical correlations and experimental data, respectively. In general, the SLC method has an intrinsic defect that is incapable of predicting performance in choke conditions very well, since mass flow is the input for flow continuity equation and needs to be adjusted to predict different performance. The SLC approach performs well for the operations points of low rotational speed, where the flow in blade channel is not yet choked. However, for configurations where the Mach number in the throat area of the blade channel is close to 1, which means the blade channel has reached the maximum capability of through-flow, the mass of the compressor remains constant, although the performance still changes. This situation usually happens in high rotational speed regions, where SLC has a higher prediction error with the experiment as shown. Some alternative methods are developed to solve this problem, such as the implementation of dedicated models for the choke regions, localized choked flow treatment [39], or time marching through-flow method [40], which is out of the scope of this article. For the comparison of the different models, treatment of this defect is not taken into consideration in this work.

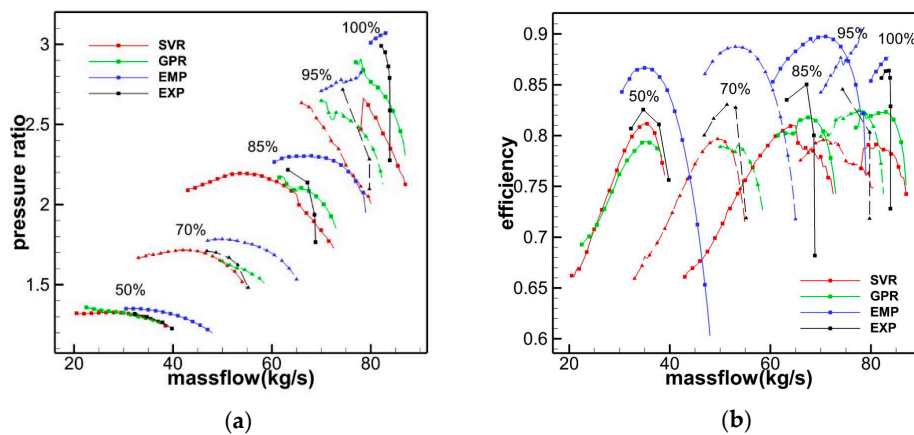


Figure 14. Comparison of experimental data and predicted performance by different models: (a) Total pressure ratio; (b) Adiabatic efficiency.

The root mean squared errors (RMSE) of the predicted overall performance of different models on the five characteristic lines compared to experimental data are listed in Table 5. The π and η indicate RSME of total pressure ratio, adiabatic efficiency of each model and the $\Delta\pi$ and $\Delta\eta$ means the percentage of increment of RSME for SVR and GPR compared to EMP. From the variation of predicted errors of each model over the whole characteristic lines, the accuracy of EMP, GPR, and SVR keeps improving with the reduction of rotational speeds. Regarding EMP, GPR, and SVR, the error of π decreases from 0.6038, 0.2430, and 0.4146 to 0.0626, 0.0089, and 0.0057, respectively while the error of η decrease from 0.1576, 0.0593, and 0.0679 to 0.0587, 0.0246, and 0.0163 respectively. The reduction of RSME indicates that all the models have better performance on conditions of low speed than higher. The reasons for surrogate models and empirical correlations are similar. The empirical correlations work well for low-loading airfoil types and low Mach number conditions because they are derived from data of some typical subsonic airfoils like NACA-65 series, but they are insufficient to simulate transonic compressors with high rotational speeds. Similarly, most training data are low speed samples, while a small part of high speed samples are located in rotor tip zones with high rotational speeds, which makes the surrogate model have a higher accuracy for the low speed cases.

Table 5. RMSE of different models compared to experimental data for whole rotational speeds and the percentage of error increment relative to EMP.

ROT	EMP		GPR				SVR			
	π	η	π	$\Delta\pi$	η	$\Delta\eta$	π	$\Delta\pi$	η	$\Delta\eta$
100%	0.4177	0.0668	0.2430	−41.81%	0.0491	−26.58%	0.4146	−0.73%	0.0679	1.66%
95%	0.6038	0.1576	0.2012	−66.68%	0.0593	−62.37%	0.2872	−52.43%	0.0501	−68.22%
85%	0.3334	0.1192	0.1661	−50.16%	0.0710	−40.39%	0.1306	−60.82%	0.0616	−48.33%
70%	0.1559	0.0907	0.0521	−66.62%	0.0412	−54.58%	0.0450	−71.13%	0.0334	−63.12%
50%	0.0626	0.0587	0.0089	−85.78%	0.0246	−58.10%	0.0057	−90.83%	0.0163	−72.30%

The increments of RMSE for GPR and SVR prove the notable improvement of predicted accuracy. Compared with the empirical correlations, the surrogate models both show better agreement with the experimental data than empirical models over all the characteristic lines. The accuracy improvement of surrogate models also increase as the rotational speeds are decreasing. Compared with empirical models, the increments of error of π change from −41.81% and −0.73% to −85.78% and −90.83%, while the increments of error of η change from −26.58% and 1.66% to −58.10% and −72.30% respectively for GPR and SVR. In addition, the GPR models have better performance compared to SVR for high rotational speed conditions, while vice versa for low speeds. The characteristic lines of EMP are smoother than the surrogate models since empirical models usually use parabolic curves to approximate the function relationship between total pressure loss/deviation angle and incidence angle increments.

In contrast, the surrogate models cannot get enough information from regions where training samples are sparse and lead to inaccurate predictions, which is the same reason why characteristic lines of high rotational speeds are rougher than the lower rotational speeds. This can be improved by adding more training samples.

4.3. Predicted Flow Field Analysis

For further comparison of the models, the operation point near the stalled margin of experimental data on the 50% rotational speed line (the left endpoint with massflow 32.3 kg/s) is chosen as the reference case, since all the models perform better at lower rpm. The total pressure ratio π , adiabatic efficiency η , and the prediction errors ERR of all models are listed in Table 6. Both GPR and SVR have a lower prediction error for both π and η than the empirical correlations. The source of the errors can be seen in the following spanwise distribution of predictions.

Table 6. Errors of total pressure ratio and adiabatic efficiency.

	EXP	GPR	SVR	EMP
π	1.317	1.306	1.311	1.352
ERR	-	0.84%	0.38%	3.13%
η	80.69%	78.19%	79.9%	86.12%
ERR	-	3.1%	2.19%	7.78%

Figure 15 shows the distribution of deviation angle and total pressure loss along blade span for blade rows. In Figure 15a, the measured deviation angle basically distributes like a bucket shape, with higher values near endwalls and lower values in the mid-span, although a small difference exists in the first stage due to the effects of part-span shroud of R1 at 61% span location [13]. The predictions of SVR and GPR models show a good agreement with the EXP data, while the empirical correlations overestimated in S1, and underestimated in the second stage. The underestimated deviation angle in rotors mainly results in a higher predicted total pressure ratio for empirical correlations, because for a certain airfoil, a lower deviation angle means higher turning flow angle as well as more energy provided. In Figure 15b, the SVR and GPR models also show a better performance than the empirical correlations for total pressure loss prediction. The underestimation mainly happens in the tip region of the first stage, where the actual total pressure loss is higher and makes the predicted adiabatic efficiency of empirical correlations much higher than the experimental data. In contrast, the SVR and GPR models accurately predicted the loss distribution for most span ranges, except the tip region of S1.

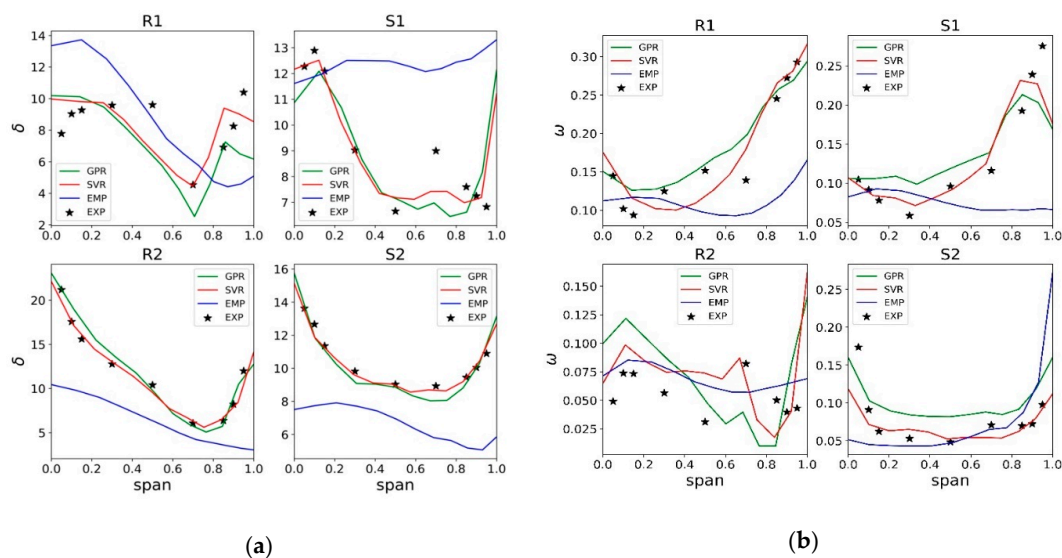


Figure 15. Spanwise distribution of predictions: (a) Deviation angle; (b) Total pressure loss.

The predicted errors relative to experimental data at hub, mid-span and tip blade span are listed in Table 7 for comparing the accuracy of EMP, GPR, and SVR models. For intuitive comparison, RMSE are also calculated for each blade row as the criteria. Although the EMP model produces a lower predicted error at a few local parts, GPR and SVR models have a lower RMSE for both deviation angle and loss predictions, except the loss predictions for R2.

Table 7. Prediction errors at hub (0.05) middle (0.5) and tip (0.95) span location and RMSE.

Row	SPAN	EMP Error		GPR Error		SVR Error	
		DEV	LOSS	DEV	LOSS	DEV	LOSS
R1	0.05	73.24%	−21.43%	30.63%	−1.98%	27.44%	6.67%
R1	0.5	−9.19%	−36.02%	−33.58%	5.32%	−30.31%	−22.08%
R1	0.95	−54.92%	−50.56%	−38.39%	−5.88%	−14.68%	−0.48%
RMSE		3.8738	0.0857	2.1004	0.0257	1.5869	0.0240
S1	0.05	−4.29%	−17.77%	−7.39%	1.05%	0.31%	−6.87%
S1	0.5	86.50%	−23.38%	7.38%	23.64%	7.96%	−5.37%
S1	0.95	90.38%	−75.58%	36.11%	−29.82%	24.69%	−23.67%
RMSE		4.0647	0.1020	1.2390	0.0371	0.8305	0.0277
R2	0.05	−52.01%	56.82%	0.16%	122.89%	−6.29%	63.11%
R2	0.5	−38.98%	103.71%	−5.97%	61.59%	−8.36%	136.88%
R2	0.95	−73.41%	54.87%	−6.32%	138.07%	−12.91%	94.17%
RMSE		6.3577	0.0223	1.1035	0.0428	0.7833	0.0274
S2	0.05	−44.06%	−72.24%	1.66%	−23.75%	−0.79%	−45.19%
S2	0.5	−23.03%	−3.02%	−1.98%	69.07%	−0.89%	9.18%
S2	0.95	−51.45%	71.84%	5.10%	37.98%	4.13%	−5.23%
RMSE		4.2813	0.0527	0.5064	0.0319	0.3260	0.0280

Furthermore, Figure 16 shows the predicted meridional contour of relative Mach number and total pressure (Pt) of different models comparing with measured data. For both variables, the contours of surrogate models show better agreement with experimental data than does the empirical correlations. As shown in Figure 16a, obvious discrepancy exists on the spanwise distribution of inlet relative Mach number for stators between EMP and EXP results. Based on the velocity triangle of the blade element section, the higher predicted inlet Mach number in the tip region of stators results from the underestimation of the deviation angle for upstream rotor blades, which corresponds to the analysis of EMP in Figure 15a. It's also easy to notice the difference between EMP and EXP on the both stators from Figure 16b. For the EMP result, the higher Pt at S1 tip inlet develops from the underestimation of deviation angle at R1 outlet (Figure 15a), while the high Pt region through S1 channel is attributed to the underestimation of total pressure loss at S1 tip (Figure 15b). Two notable high Pt zones are located near the endwalls of the tip and hub after R2 due to the underestimation of R2 deviation angle (Figure 15a). Table 8 lists the RMSE of Mach number and Pt corresponding to Figure 16 on both leading edge (LE) and trailing edge (TE) of each blade rows. The predicted accuracy of surrogate models on Ma and Pt has a notable improvement, which is up to 49.24% and 73.13% for GPR, and 71.62% and 87.19% for SVR.

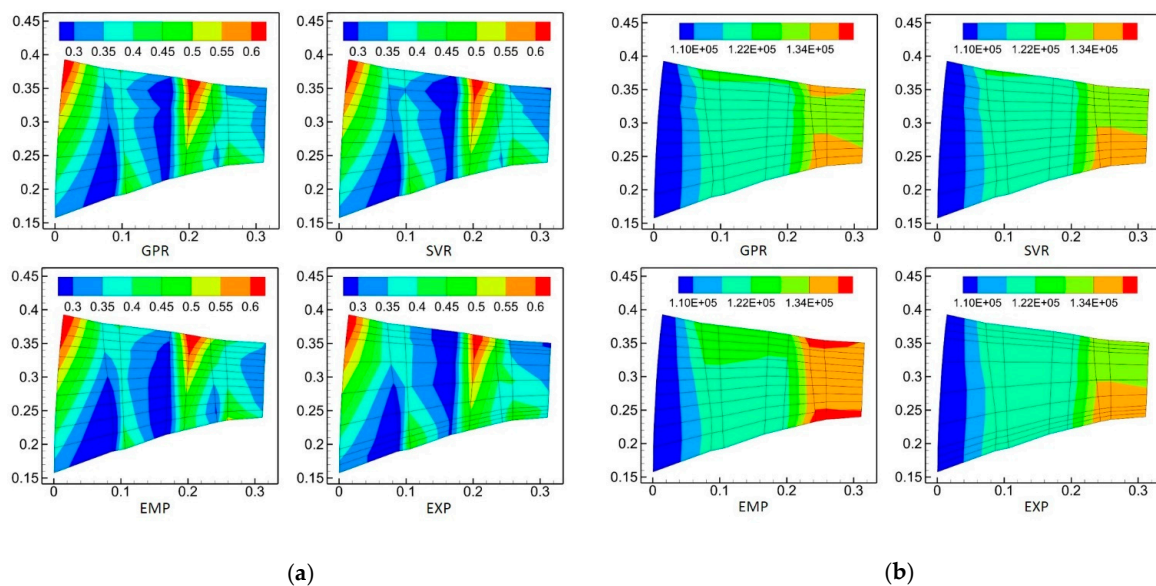


Figure 16. Meridional contour of different predictions: (a) Relative Mach number; (b) Total pressure.

Table 8. RMSE of relative Mach number and total pressure compared to experimental data and the percentage of error increment relative to EMP.

Row	Edge	EMP		GPR				SVR			
		Ma	Pt	Ma	Δ Ma	Pt	Δ Pt	Ma	Δ Ma	Pt	Δ Pt
R1	LE	0.009	0	0.008	−19.85%	0		0.007	−22.56%	0	
	TE	0.025	3526	0.013	−45.70%	1948	−44.77%	0.013	−46.45%	1145	−67.52%
S1	LE	0.028	3529	0.025	−13.08%	1954	−44.64%	0.016	−43.51%	1148	−67.47%
	TE	0.050	4473	0.033	−34.42%	2288	−48.86%	0.022	−56.66%	1425	−68.14%
R2	LE	0.025	4501	0.013	−49.24%	2303	−48.84%	0.007	−71.62%	1428	−68.26%
	TE	0.039	6648	0.020	−48.97%	2062	−68.98%	0.012	−69.37%	856	−87.13%
S2	LE	0.046	6648	0.028	−40.28%	2057	−69.06%	0.015	−66.74%	852	−87.19%
	TE	0.034	6218	0.029	−15.29%	1671	−73.13%	0.014	−59.10%	1048	−83.15%

In summary, the comparison suggests that the data-driven surrogate model provides an alternative approach to get a reliable prediction of compressor performance when typical empirical correlations are incapable.

5. Conclusions and Outlook

An alternative way of using surrogate models to predict the total pressure loss and deviation angle was developed in this work. Experimental data measured along the blade spans of a two-stage axial transonic compressor over a wide operation range constituted the samples database for the training models. The most influential inputs were identified from seven geometric and six aerodynamic features using a sensitivity analysis approach for both surrogate models. Two supervised learning methods, support vector regression and Gaussian process regression, were applied to train the surrogate models. The SMBO algorithm was implemented to tune the hyper parameters with cross validation scheme to avoid an over-fit situation. Both models obtained the optimal hyper parameters within 100 iteration steps. The validation showed that the SVR and GPR models have similar performance on the sample database. The optimal models were then integrated into a through-flow program developed using the streamline curvature method to predict the performance and flow field of the compressor. Some typical empirical correlations were also introduced as reference. The results of surrogate models showed good agreement with the measured data, whereas the problem of underestimation or overestimation

exists for the empirical correlations. Compared to empirical correlations, the surrogate models notably reduced the predicted errors of performance over all the operations. GPR reduced the RMSE of π and η up to 85.78% and 58.10%, and SVR up to 90.83% and 72.30%, respectively. The comparison suggested that the empirical correlations could cause considerable errors for the cases beyond the scope of application. The surrogate models based on experimental data might be a substitute method to achieve a reliable prediction.

In conclusion, the surrogate models in this work have the following advantages compared to the typical empirical correlations:

1. The models depend directly on the experimental database of compressors, instead of 2D cascade data used by empirical correlation. This means that the training data include all the information about the complicated flow field in the blade channel of the compressor. There is no need to build additional models to correct the 3D effects on total pressure loss or deviation angle which is usually used for typical empirical correlations. This can avoid introducing errors in the correction correlations.
2. Reference conditions are not needed in surrogate models. For empirical correlations, a definition of airfoils called reference or design condition should be determined at first, since all the off-design performance is based on the reference condition. Taking profile loss as an example, the profile loss is comprised of the basic value in reference conditions, and a correction value when the actual incidence angle deviates from the reference. Usually, the quadratic function is used to approximate the relationship between corrected value and the increment value of the incidence angle, while the reference condition needs to be predicted by another correlation, the result of which will directly affects accuracy of the subsequent prediction. The empirical correlations divide the nonlinear issue to some subset of linear issues for simplicity, at cost of accuracy. However, this is not an issue for the surrogate models because, firstly, the training samples nearly cover the whole operation range of the elementary airfoils, and secondly, the machine learning methods are born to settle the nonlinear issue.
3. Surrogate models need fewer manual operations than empirical correlations. The empirical correlations are constituted by redundant coefficients deduced from semi-experience and semi-theory. Manual adjustment of the correlations is required in specific cases to reduce errors. In contrast, surrogate models are generated by an automatic process and work as a black box in the through-flow program, so that only inputs and outputs are concerned. The hyper parameters can be obtained by an optimization algorithm. The characteristics of a black box make it friendly to the user who has little experience in modifying the empirical correlations. The only factor to be determined is which machine learning method to use.
4. The surrogate models have more flexible choices. The empirical correlations have seen limited change to the structure of the functions, despite the long period of development. It's difficult to make great improvement on the prototypes of correlations, without making changes. Essentially, the function relationship between inputs and outputs is highly nonlinear. The diversity of machine learning methods introduces more possibility into the approximation of the nonlinear relationship by surrogate models.

Besides the advantages, some drawbacks also exist and need to be improved in the future:

1. The distribution of the samples is not even, that makes the predicted characteristic lines fluctuate. The sparse regions in the sample space should be filled with more data, especially for the transonic/supersonic samples.
2. The scale of the database is small, so the trained surrogate models do not currently work well for other compressors. The database will be extended, however it's difficult to obtain enough experimental data of advanced compressors from open literature. Given that, samples based on CFD with appropriate modification by experimental data is an alternative way to extend the database.

3. The SVR and GPR models perform well on the small data set in this work, but are inefficient for large databases because the computational complexity of them is $O(n^3)$, where n is the samples amount. Some measures will be taken to solve this problem in the extended samples.
4. Traditional models have a wide applicability for conventional compressors since the physical knowledge is considered while surrogate models cannot reflect the physical knowledge explicitly. Machine learning methods, which have the capability of nonlinear approximation, can be applied to improve the accuracy of the traditional models, which reflect the theoretical knowledge, on the cases out of the operational scope of traditional models.

It should be noted that the same compressors are used for training and testing because of the limited database, but the over-fit situation is well controlled by the cross validation scheme because no serious oscillation appears in the prediction of overall performance. In addition, the surrogate models may perform better than traditional models on individual special compressors but cannot compete with traditional models on the other conventional cases. Only 36 airfoils are used to build the surrogate models while the traditional models have been deduced from a wider set of data. The improvement of applicability to different geometries remains the main task for surrogate models in subsequent work.

Author Contributions: Conceptualization and methodology, X.W. and B.L.; resources, G.G.; data curation and formal analysis, X.W. and N.R.; software, X.W. and B.L.; Supervision, B.L. and G.G.; writing—original draft preparation, X.W.; writing—review and editing, N.R. and G.G.; funding acquisition, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 51676162 and 51790512, National Science and Technology Major Project, grant number 2017-II-0001-0013 and China Scholarship Council, grant number 201806290159.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hirsch, C.; Denton, J. Through flow calculations in axial turbomachines. *AGARD Advis. Rep.* **1981**, *175*, 111.
2. Cetin, M.; Uecer, A.; Hirsch, C.; Serovy, G. *Application of Modified Loss and Deviation Correlations to Transonic Axial Compressors*; Advisory Group for Aerospace Research and Development: Neuilly-Sur-Seine, France, 1987.
3. Fottner, L. Improved blade profile loss and deviation angle models for advanced transonic compressor bladings: Part II—A model for supersonic flow. *J. Turbomach.* **1996**, *117*, 81.
4. Fottner, L. Improved blade profile loss and deviation angle models for advanced transonic compressor bladings: Part I—A model for subsonic flow. *J. Turbomach.* **1996**, *118*, 73.
5. Gallimore, S. Spanwise Mixing in Multistage Axial Flow Compressors: Part II—Throughflow Calculations Including Mixing. *J. Turbomach.* **1986**, *108*, 10–16. [[CrossRef](#)]
6. Schmitz, A.; Aulich, M.; Nicke, E. Novel approach for loss and flow-turning prediction using optimized surrogate models in two-dimensional compressor design. In Proceedings of the ASME 2011 Turbo Expo: Turbine Technical Conference and Exposition, Vancouver, BC, Canada, 6–10 June 2011; pp. 1103–1114.
7. Drela, M.; Youngren, H. *A User's Guide to MISES 2.53*; Massachusetts Institute of Technology: Cambridge, MA, USA, 1998.
8. Schmitz, A.; Aulich, M.; Schönweitz, D.; Nicke, E. Novel performance prediction of a transonic 4.5 stage compressor. In Proceedings of the ASME Turbo Expo 2012: Turbine Technical Conference and Exposition, Copenhagen, Denmark, 11–15 June 2012; pp. 2123–2134.
9. Kou, J.; Zhang, W. A hybrid reduced-order framework for complex aeroelastic simulations. *Aerosp. Sci. Technol.* **2019**, *84*, 880–894. [[CrossRef](#)]
10. Jin, Y.; Li, S.; Jung, O. Prediction of Flow Properties on Turbine Vane Airfoil Surface from 3D Geometry with Convolutional Neural Network. In Proceedings of the ASME Turbo Expo 2019: Turbomachinery Technical Conference and Exposition, Phoenix, AZ, USA, 17–21 June 2019.
11. Joly, M.; Sarkar, S.; Mehta, D. Machine Learning Enabled Adaptive Optimization of a Transonic Compressor Rotor with Precompression. *J. Turbomach.* **2019**, *141*, 051011. [[CrossRef](#)]
12. Messenger, H.; Kennedy, E. *Two-Stage Fan I: Aerodynamic and Mechanical Design*; NASA: Washington, DC, USA, 1972.

13. Ruggeri, R.S.; Benser, W.A. *Performance of a Highly Loaded Two Stage Axial-Flow Fan*; NASA: Washington, DC, USA, 1974.
14. Morris, M.D. Factorial sampling plans for preliminary computational experiments. *Technometrics* **1991**, *33*, 161–174. [[CrossRef](#)]
15. Campolongo, F.; Cariboni, J.; Saltelli, A. An effective screening design for sensitivity analysis of large models. *Environ. Model. Softw.* **2007**, *22*, 1509–1518. [[CrossRef](#)]
16. Forrester, A.; Sobester, A.; Keane, A. *Engineering Design via Surrogate Modelling: A Practical Guide*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
17. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
18. Bergstra, J.; Yamins, D.; Cox, D.D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In Proceedings of the 30th International Conference on Machine Learning (ICML 2013), Atlanta, Georgia, 16–21 June 2013.
19. Bergstra, J.S.; Bardenet, R.; Bengio, Y.; Kégl, B. Algorithms for hyper-parameter optimization. In Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; pp. 2546–2554.
20. Hutter, F. Automated Configuration of Algorithms for Solving Hard Computational Problems. Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada, 2009.
21. Hutter, F.; Hoos, H.H.; Leyton-Brown, K. Sequential model-based optimization for general algorithm configuration. In *Proceedings of the International Conference on Learning and Intelligent Optimization*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 507–523.
22. Jones, D.R. A taxonomy of global optimization methods based on response surfaces. *J. Glob. Optim.* **2001**, *21*, 345–383. [[CrossRef](#)]
23. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
24. Hsu, C.-W.; Chang, C.-C.; Lin, C.-J. *A Practical Guide to Support Vector Classification*; National Taiwan University: Taipei, Taiwan, 2003.
25. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
26. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.J.; Vapnik, V. Support vector regression machines. In Proceedings of the Advances in Neural Information Processing Systems (NIPS 1996), Denver, CO, USA, 2–5 December 1996; MIT Press: Cambridge, MA, USA, 1997; pp. 155–161.
27. Platt, J.C. Using analytic QP and sparseness to speed training of support vector machines. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 1999; pp. 557–563.
28. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI* **1995**, *14*, 1137–1145.
29. Rasmussen, C.E. Gaussian processes in machine learning. In *Summer School on Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 63–71.
30. Wu, C.-H. *A General Thought-Flow Theory of Fluid Flow with Subsonic or Supersonic Velocity in Turbomachines of Arbitrary Hub and Casing Shapes*; National Aeronautics and Space Administration: Washington, DC, USA, 1951.
31. Wu, C.-H. *A General Theory of Three-Dimensional Flow in Subsonic and Supersonic Turbomachines of Axial-, Radial, and Mixed-Flow Types*; National Aeronautics and Space Administration: Washington, DC, USA, 1952.
32. Novak, R. Streamline curvature computing procedures for fluid-flow problems. *J. Eng. Power* **1967**, *89*, 478–490. [[CrossRef](#)]
33. Wu, X.; Liu, B.; Lei, S.; Zhang, G.; Mao, X. Development of an improved streamline curvature approach for transonic axial compressor. In Proceedings of the ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition, Seoul, Korea, 13–17 June 2016.
34. Carter, A.D. *The Low Speed Performance of Related Aerofoils in Cascades*; HM Stationery Office: Richmond, UK, 1950; Volume 29.
35. Carmody, R.; Creveling, H. *Axial Flow Compressor Computer Program for Calculating off Design Performance/Program 4*; NASA: Washington, DC, USA, 1968.
36. Pachidis, V. Gas Turbine Advanced Performance Simulation. Ph.D. Thesis, Cranfield University, Cranfield, UK, 2006.

37. Koch, C.; Smith, L. Loss Sources and Magnitudes in Axial-Flow Compressors. *J. Eng. Power* **1976**, *98*, 411–424. [[CrossRef](#)]
38. Boyer, K.M.; O'Brien, W.F. An improved streamline curvature approach for off-design analysis of transonic axial compression systems. *J. Turbomach.* **2003**, *125*, 475–481. [[CrossRef](#)]
39. Tiwari, P.; Stein, A.; Lin, Y.-L. Dual-solution and choked flow treatment in a streamline curvature throughflow solver. *J. Turbomach.* **2013**, *135*, 041004. [[CrossRef](#)]
40. Yang, C.; Wu, H.; Yang, J.; Ferlauto, M. Time-marching throughflow analysis of multistage axial compressors based on a novel inviscid blade force model. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **2019**, *233*, 0954410019840588. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).