

Article

Not peer-reviewed version

Critical Sample Size Analysis on Uncertainty Aerodynamic Evaluation for Compressor Blade With Stagger Angle Errors

Haohao Wang , Limin Gao *, Baohai Wu

Posted Date: 9 November 2023

doi: 10.20944/preprints202311.0639.v1

Keywords: compressor blade; manufacturing uncertainty; stagger angle; uncertainty quantification; limited samples; polynomial chaos



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Critical Sample Size Analysis on Uncertainty Aerodynamic Evaluation for Compressor Blade with Stagger Angle Errors

Haohao Wang 1,2, Limin Gao 1,2,* and Baohai Wu 3

- ¹ School of Power and Energy, Northwestern Polytechnical University, Xi'an 710129, China
- ² National Key Laboratory of Aerodynamic Design and Research, Xi'an 710129, China
- School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710129, China
- * Correspondence: gaolm@nwpu.edu.cn

Abstract: Many probability-based uncertainty quantification (UQ) schemes require a large amount of sampled data to build credible probability density function (PDF) models for uncertain parameters. Unfortunately, the data collected in compressor blades of aero-engine are mostly limited due to expensive and time-consuming tests. In this paper, we develop a preconditionerbased data-driven polynomial chaos (PDDPC) method that can efficiently deal with uncertainty propagation of limited sampled data. The calculation accuracy of PDDPC method is closely related to the sample size of collected data. Therefore, the influence of sample size on PDDPC method is investigated using a nonlinear test function. Subsequently, we consider the real manufacturing errors of stagger angle for compressor blades. Under three different operating conditions, PDDPC method is applied to investigate the effect of stagger angle error on UQ results of multiple aerodynamic parameters of a two-dimensional compressor blade. The results show that as the sample size of measured data increases, UQ results of aerodynamic performance obtained by PDDPC method gradually converge. There exists a critical sample size that ensures accurate UQ analysis of compressor blades. The probability information contained in the machining error data is analyzed through Kullback-Leibler divergence, and the critical sample size is determined. The research results can make a valuable reference for the fast and cheap UQ analysis of compressor blades in practical engineering.

Keywords: compressor blade; manufacturing uncertainty; stagger angle; uncertainty quantification; limited samples; polynomial chaos

1. Introduction

Increasingly demanding target for high reliability and stable service performance of the advanced aviation compressor requires it to resist various sources of uncertainties [1,2]. Compressor blades are usually designed as three-dimensional twisted parts with complex curved structures to meet the requirements of aerodynamic performance. Geometric errors in compressor blades are one of the most important sources of uncertainty due to poor rigidity, tooling deflection, material residual stress and other factors [3,4]. The harsh operating environment such as a high adverse pressure gradient will magnify the influence of geometric uncertainties on the compressor performance. Garzon et al. [5] found that due to the inevitable geometric uncertainties, the average compressor efficiency can be reduced by 1%, accompanied by a dispersion of aerodynamic performance. With the recent development in computing power, quantitative evaluation of the influence of geometric uncertainties on compressor performance has attracted extensive attention [6–8].

Advanced uncertainty quantification (UQ) methods are essential tools in the process of uncertainty evaluation for aviation compressors. The most commonly used UQ methods in the field of turbomachinery include Monte Carlo method [9,10], surrogate model-based method [11], multifidelity methods [12,13], and polynomial chaos-based method [14–16]. These methods are

probability-based UQ schemes, and their uncertainty evaluation process can be generally divided into three steps. Firstly, it requires to build the probability density function (PDF) models of uncertain input parameters, where a large amount of sampled data is collected and analyzed as the underlying modeling basis. The constructed PDF model occupies the primary position since it determines the random input space for the whole UQ process. The second step is to combine the UQ algorithm with the response system such as a CFD solver to obtain the input-output mapping. The third step is post-processing and statistical analysis of the output results. The probability-based UQ methods require the PDF model as the random input for uncertainty analysis. Therefore, how to obtain an accurate and high-fidelity PDF model is an important prerequisite for obtaining reliable UQ results.

A credible PDF model usually requires a large amount of sampled data as the underlying support. Unfortunately, it is very difficult to extract or collect sufficient and available machining error data for compressor blades in the field of aero-engines. The reason for limited data can be generally attributed to three aspects. First, the profile detection of compressor blades using coordinate measuring machines or optical measurements is very expensive and time-consuming. Therefore, sampling inspection technology is often used to judge whether the manufactured blades are qualified in practical engineering. Second, it is difficult to obtain sufficient machining data of the same types from the limited compressor rows. Third, the aero-compressor blades often involve the problem of data information security. As a consequence, the available machining error data of compressor blades are mostly limited and scarce. In this context, it is impossible to obtain an accurate high-fidelity PDF model, which greatly weakens the availability of the aforementioned probability-based UQ methods. In most practical applications, many reports [17–19] either avoid such construction of PDF models completely or fall back to less accurate and lower-fidelity PDF models to conduct the UQ analysis. The credibility of model output will bear severe risk using an unreliable input PDF model [20,21]. Therefore, it is imperative to seek a novel UQ method suitable for limited sampled data for compressor blades from the ground up.

In the field of geo-sciences, Oladyshkin et al. [22,23] proposed a data-driven polynomial chaos (DDPC) method to solve the uncertainty of carbon dioxide storage when facing scarce permeability information. The idea of DDPC algorithm is to use the statistical moments of raw measured data to replace the PDF model for UQ analysis. Recently, Ahlfeld and Montomoli [24,25] combined the DDPC method with the Smolyak technique to solve high-dimensional UQ problems in turbomachinery. They applied this method to study the influence of manufacturing uncertainties on diffuser efficiency and turbine blades. By comparing with the low-fidelity PDF model obtained from 100 measured data, it is found that the DDPC method was more accurate and reliable. Guo et al. [26] used DDPC method to investigate the performance impact of manufacturing errors on a 2D compressor blade. However, their results are still based on the input PDF model using kernel density estimation, ignoring the problem of limited sampled data. Wang et al. [1] proposed a novel preconditioner-based DDPC (PDDPC) method, which can alleviate the ill-conditioned problem when using high-order statistical moments. With consideration of the limited sampled data (100 samples) of stochastic stagger angle errors, they used PDDPC method to conduct a robust optimization design for a compressor blade.

In the practical implementation, this DDPC-based UQ method is still rare in turbomachinery blades. First, there exists a poor numerical condition when using high-order statistical moments of sampled data [27]. Second, the calculation accuracy of DDPC method is closely related to the sample size of collected data. To the best of the authors' knowledge, there is no relevant literature to study the critical sample size that can be used for accurate UQ analysis. From the perspective of engineering, the processing cycle and detection time are very expensive for the compressor blades. Therefore, determining the critical sample size has great benefits for saving costs and shortening UQ cycle.

For turbomachinery blade, the stagger angle is an important parameter because it affects the relative inlet flow direction, the variations of blade passages and blade aeroelasticity [1,13]. Wang et al. [28] evaluated the influence of the uncertain stagger angle errors on the aerodynamic performance of a turbine blade, and discovered that the stagger angle errors can cause the fluctuation of the loss

2

coefficient up to 40%. Phan and He [13,29] found that the stagger angle variations can affect the blade loading, exit flow angle, blade aeroelasticity and aerodynamic loss for the turbine blade. Lu et al. [30] found that the mis-staggering can be amplified by aerodynamic loading and the designed peak efficiency cannot be achieved for a transonic fan blade. Suriyanarayanan et al. [31] studied the influence of stagger angle errors on the NASA rotor 67. Their results showed that the sinusoidal arrangement of stagger angle variations throughout the annulus is optimal compared to the linear and random types. These investigations have demonstrated that stagger angle errors have a significant impact on turbomachinery blades. Therefore, it is imperative to conduct the critical sample size analysis on uncertainty aerodynamic evaluation for compressor blades with stagger angle errors.

The motivation behind this article can be summarized as follows: first, we introduce a preconditioner-based data-driven polynomial chaos (PDDPC) method to ensure the computational robustness when using high-order statistical moments. Then, the influence of sample size of limited input data on PDDPC method is studied using a nonlinear test function. Finally, we consider the actual manufacturing error of stagger angle for compressor blades. The influence of sample size of stagger angle error on UQ results of multiple aerodynamic parameters is discussed under different operating conditions. The research results can provide a valuable reference for UQ analysis of turbomachinery blades.

2. Research object and numerical method

2.1. Compressor blade geometry and manufacturing uncertainty of stagger angle

In this work, the research object is a two-dimensional compressor blade, as shown in Figure 1. Its chord is 69.95mm, the design stagger angle (θ) is 26.58°, and the design incidence a is 2.5°. The main geometric and aerodynamic parameters are summarized in Table 1.

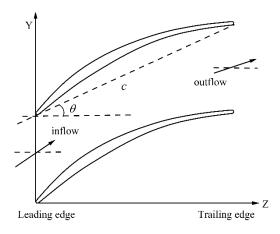


Figure 1. Compressor blade geometry.

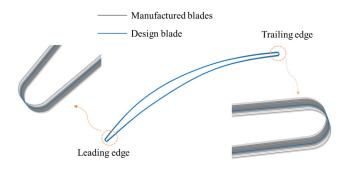
Table 1. Main geometric parameters of the compressor cascade.

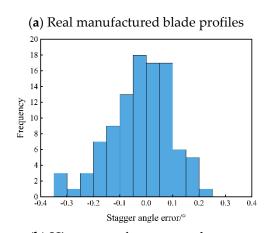
Parameters	Values
Chord length c (mm)	69.9
Camber angle (°)	39.61
Geometric inlet angle (°)	45.83
Maximum thickness/mm	3.51
Design incidence (°)	2.5
Stagger angle θ (°)	26.58

Due to inevitable torsional deformation, the actual stagger angles of manufactured blades always exhibit deviation from the initial design. The actual stagger angle consists of two parts: the initial design angle θ and the machining error $\Delta\theta$, see Equation (1).

$$\theta_{real} = \theta + \Delta\theta \tag{1}$$

In this work, the stagger angles originate from mid-height sections of 100 compressor blades based on coordinate measuring machine (CMM) measurements. CMM-based measurement techniques can provide the manufacturing error data of stagger angles between nominal and manufactured blades. Figure 2 presents the detailed measured results and the actual compressor blade profiles with torsional errors. As depicted in Figure 2, the machining errors of stagger angle have obvious uncertain and stochastic characteristics. In practical engineering, we cannot obtain more machining error data on blade stagger angle due to various constraint. From the perspective of probability theory [32], 100 sampled data cannot establish a high-fidelity PDF model. Therefore, we consider the data-driven polynomial chaos method to propagate uncertainty where the statistical moments of measured data are used as the random inputs.





(b) Histogram of stagger angle error

Figure 2. Measurement of stagger angle.

2.2. Numerical method and validation

Reynolds-Averaged Navier-Stokes (RANS) equations are used to analyze the aerodynamic performance of compressor blade using the commercial software NUMECA. The Spalart-Allmaras (SA) turbulence model [33] is adopted to enclose the governing equations. The second-order central difference scheme and four-step Runge-Kutta scheme are used for spatial discretization and time discretization, respectively.

Figure 3 shows the computational domain of compressor blade geometry. The inlet boundary conditions include total pressure, total temperature (300.0K), and airflow angle. The static pressure (101300 Pa) is imposed at the outlet boundary. The inlet Mach number is obtained by changing the inlet total pressure under various flow angles. Note that we use a 2D compressor blade case, so periodic boundary conditions are given along the pitch-wise direction. The Autogrid5 module is used to generate the O4H-type grid in the blade passage domain. The total gird number is approximate

 9.83×10^4 . The first layer grid-scale from the wall is set as 10^{-6} m to guarantee the dimensional distance y^+ to be less than 3 for meeting the requirement of the SA turbulence model.

The boundary conditions, number and structure of computational grids, turbulence model, and other calculation settings in this paper are consistent with our previous work [1]. To validate the reliability of the numerical simulation method, the predicted results of aerodynamic performance are compared with the experimental results in terms of the isentropic Mach number and exit loss coefficient. As can be seen from Figure 4, the predicted results agree well with the measured values, indicating the reliability of the chosen numerical method for predicting the aerodynamic performance.

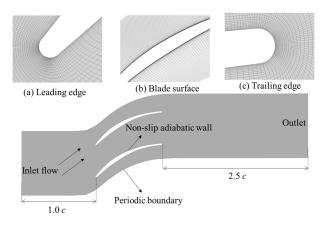


Figure 3. Computational mesh and domain of compressor blade passage.

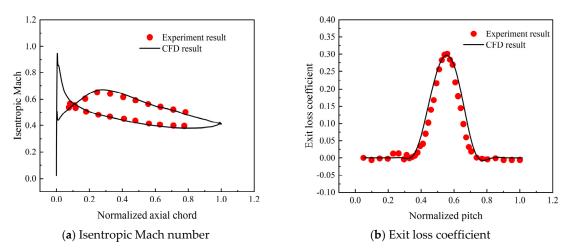


Figure 4. Comparison between experimental measurement and numerical results.

2.3. Key aerodynamic quantities of interest

The interested aerodynamic parameters for a 2D compressor blade include the total pressure loss coefficient ω , static pressure ratio π , isentropic Mach number Ma_{is} , and Mach number. In this paper, our goals lie in quantifying the performance variations associated with the uncertain stagger angle errors. These aerodynamic parameters are defined as

$$\omega = (P_{\rm in} - P_{\rm out})/(P_{\rm in} - P_{\rm l}) \tag{2}$$

$$\pi = P_2/P_1 \tag{3}$$

$$Ma_{is} = \sqrt{\frac{2}{\gamma - 1} \left(\left(\frac{P_{in}}{P} \right)^{(\gamma - 1)/\gamma} - 1 \right)}$$
 (4)

where P_{in} and P_{out} are the inlet total pressure and outlet pressure; P_1 , P_2 and P are the inlet static pressure, outlet static pressure and the local static pressure at location on the blade surface.

3. Uncertainty quantification method

3.1. Preconditioner-based data-driven polynomial chaos method

This part provides the necessary mathematical basis for the uncertainty propagation of limited measured data. We consider a random process in the probability space (Ω, Φ, Γ) with space of events Ω , σ -algebra and probability measure Γ . For the physical model $Y(\xi)$ with n-dimensional inputs $\xi \in \Omega$, if the ξ is random variable, the model output Y is also stochastic. According to the polynomial chaos theory [34], the model output Y can expand a series of orthogonal polynomials with order p:

$$Y = \sum_{i=0}^{Q} u_i \Psi(\xi), \quad Q = \frac{(n+p)!}{(n!p!)} - 1$$
 (5)

where u_i is the polynomial chaos coefficient, Q denotes the truncated expansion term, $\Psi(\xi)$ denotes the polynomial basis that is orthogonal with respect to the probability measure Γ .

The orthogonal basis $\Psi(\xi)$ is the core of mapping the relationship between random input and model output. For the traditional probability-based UQ methods, if we already know the PDF model of random input variable, the corresponding orthogonal basis can be easily obtained according to the Wiener-Askey scheme [35]. For example, the Hermite orthogonal polynomials can be used for modeling the effect of random variable described by continuous Gaussian probability distribution. Unfortunately, available sampled data is not enough to build an accurate and high-fidelity input PDF model in practical applications, and the correct orthogonal basis function also cannot be obtained.

In practical engineering, the measured data collected in manufactured compressor blades are mostly limited $(\xi_1, \xi_2, \cdots, \xi_N)$. Here, ξ_i is available sampled data and N is the sample size. The key part of data-driven polynomial chaos (DDPC) method is to use the statistical moments of sampled data to construct orthogonal basis $\Psi(\xi)$. The definition of kth statistical moment μ_k can be represented as:

$$\mu_k = \int \boldsymbol{\xi}^k d\Gamma(\boldsymbol{\xi}) \approx \frac{1}{N} \sum_{i=1}^N \xi_i^k.$$
 (6)

Next, we use the orthogonality of the basis function $\Psi(\xi)$. The orthogonal base $\Psi^k(\xi)$ of degree k can be written as:

$$\Psi^{(k)}(\xi) = \sum_{i=0}^{k} h_i^{(k)} \xi^i$$
 (7)

where $h_{(i)}^k$ is the constant term. The relationships between orthogonal base $\Psi^k(\xi)$ and lower degree $\Psi^l(\xi)$ ($l \leq k$) can be represented as:

$$\int_{\xi \in \Omega} h_0^{(0)} \left[\sum_{i=0}^k h_i^{(k)} \xi^i \right] d\Gamma(\xi) = 0;$$

$$\vdots$$

$$\int_{\xi \in \Omega} \left[\sum_{i=0}^{k-1} h_i^{(k-1)} \xi^i \right] \left[\sum_{i=0}^k h_i^{(k)} \xi^i \right] d\Gamma(\xi) = 0;$$

$$\int_{\xi \in \Omega} \left[\sum_{i=0}^k h_i^{(k)} \xi^i \right] \left[\sum_{i=0}^k h_i^{(k)} \xi^i \right] d\Gamma(\xi) = 1.$$
(8)

Here, we assume that the orthogonal functions are the standard form, so their inner product is 1 when the degree is the same. Alternatively, Eq. (8) can be further simplified as:

Now, we use the definition of statistical moment, Equation (9) can be rewritten as:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_0 & \mu_1 & \cdots & \mu_k \\ \mu_1 & \mu_2 & \cdots & \mu_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_k & \mu_{k+1} & \cdots & \mu_{2k} \end{bmatrix}; \boldsymbol{h} = \begin{bmatrix} h_0^{(k)} \\ h_1^{(k)} \\ \vdots \\ h_k^{(k)} \end{bmatrix}; \boldsymbol{c} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$(10)$$

where μ represents the statistical moment matrix and h is the unknown constant term. This makes it obvious that one can determine the orthogonal base $\Psi(\xi)$ by solving Equation (10). The statistical moment matrix μ is known as the Hankel matrix. A recent open literature [36] has provided abundant evidence that Hankle matrix is often ill-conditioned and has a poor numerical condition. Therefore, the solution system of Equation (10) is very unstable, which severely limits the flexibility and effectiveness of DDPC method.

To address this problem, we propose a novel preconditioner-based data-driven polynomial chaos (PDDPC) method to enhance the computational robustness and instability. The proposed preconditioner M should have the following properties: 1) the form of M is easy to derive; 2) M can alleviate the ill-conditioned problem. It can be found that the initial statistical moment matrix μ is positive definite and symmetric and its inverse matrix can be easily derived. Therefore, the diagonal inverse matrix of μ is proposed as the preconditioner M:

$$M = diag(\mu)^{-1} = diag(\mu_0, \mu_2, \dots, \mu_{2k})^{-1}$$
 (11)

Combined with the preconditioner M, Eq. (10) can be rewritten as:

$$\begin{bmatrix} 1 & \mu_0^{-1}\mu_1 & \cdots & \mu_0^{-1}\mu_k \\ \mu_2^{-1}\mu_1 & 1 & \cdots & \mu_2^{-1}\mu_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{2k}^{-1}\mu_k & \mu_{2k}^{-1}\mu_{k+1} & \cdots & 1 \end{bmatrix} \begin{bmatrix} h_0^{(k)} \\ h_1^{(k)} \\ \vdots \\ h_k^{(k)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mu_{2k}^{-1} \end{bmatrix}$$
(12)

The constructed preconditioner M can improve the ill-conditioned conditions, so the orthogonal base $\Psi(\xi)$ can be calculated smoothly. The polynomial chaos coefficient u_i can be solved using Galerkin projection method (see e.g., [37]).

Finally, the statistical results (i.e., mean value $\mu(Y)$ and the standard deviation $\sigma(Y)$) of the model output Y can be computed by Equation (13) and Equation (14). Figure 5 gives a schematic diagrem of the uncertainty propagation process of PDDPC method.

$$\mu(Y) = u_0 \tag{13}$$

$$\sigma^{2}(Y) = \sum_{i=1}^{Q} u_{i}^{2} \langle \Psi(\xi), \Psi(\xi) \rangle. \tag{14}$$

Figure 5. The schematic diagram of the PDDPC method.

3.2. Validation on PDDPC

Based on the self-developed PDDPC method, this part will analyze the influence of sample size on the model outputs through a nonlinear test function. According to our previous published literature [1], the computational robustness and accuracy of PDDPC have been validated by a series of tests using only 100 sampled data. Results show that when using the p = 3 order method, PDDPC has reached high calculation accuracy and good convergence behavior. Therefore, we still use the third-order PDDPC to study the impact of the sample size in this paper. The nonlinear test function is described as follows:

$$Y = 7\cos^2(X) + X^2 + 10X, X \sim Beta(4, 4),$$
 (15)

where random variable *X* is assumed to follow *Beta* (4,4) distribution.

According to Equation (6), it can be found that the sample size N will affect the statistical moments and the output results. In order to obtain the sensitivity of PDDPC method to the sample size N, we randomly generate 11 sets of sample data (N = 20, 30, 40, 50, 60, 80, 100, 150, 200, 300 and 1000). The analytical values of $\mu(Y)$ and $\sigma(Y)$ of test function are calculated using 106 Monte Carlo simulations as the benchmark. The relative error of calculation result is defined as:

$$Err = |Y^{\text{MC}} - Y^{\text{PDDPC}}|/|Y^{\text{MC}}| \tag{16}$$

Figure 6 shows the convergence rate in terms of $\mu(Y)$ and $\sigma(Y)$ estimated using PDDPC method under each sample size N. The relative errors of $\mu(Y)$ and $\sigma(Y)$ are given in Table 2. It can be observed that when N is less than 40, there is a large fluctuation in the calculation results of model outputs. When the sample size N exceeds 40, the prediction results of the mean and standard deviation of test case gradually converge. But it can also be observed that as the sample size N increases, the results do not converge monotonously. This can be attributed to the randomness of sampling. Fortunately, the results calculated by the PDDPC method still maintain a high level of accuracy. Although PDDPC method is a UQ scheme suitable for scarce sampled data, it still has a certain dependence on the sample size N. As the sample size N increases, the dependence of PDDPC gradually decreases. When the sample size N is greater than 40, PDDPC has a high calculation accuracy in predicting the statistical results in terms of the test function. Based on the results of the test function, we can infer that N = 40 is the critical sample size, which can be used for the correct propagation of uncertainty.

8



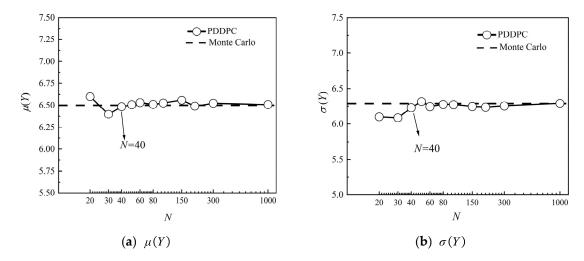


Figure 6. Convergence results of $\mu(Y)$ and $\sigma(Y)$ with different sample sizes.

 Table 2. Relative errors for test functions.

N (Sample size)	$Err-\mu(Y)\%$	$Err-\sigma(Y)\%$
20	1.53	1.13
30	1.59	0.68
40	0.25	0.16
50	0.09	0.18
60	0.46	0.28
80	0.15	0.60
100	0.38	0.21
150	0.91	0.69
200	0.16	0.91
300	0.35	0.37
1000	0.12	0.03

4. Results and Discussion

In this section, the influence of sample size of stagger angle error on uncertain aerodynamic evaluation of compressor blade is investigated under the condition of inlet Mach number of 0.7. Figure 7 shows the uncertainty propagation process of stagger angle errors where the PDDPC method is coupled with the CFD solver. In order to fully evaluate the effect of sample size, we consider three incidence operating conditions, i.e., blocking incidence condition ($i = -0.5^{\circ}$), design incidence ($i = 2.5^{\circ}$) and high positive incidence ($i = 7^{\circ}$) conditions (near stall).

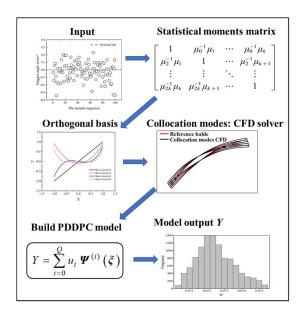
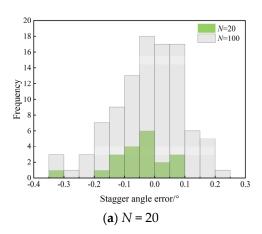
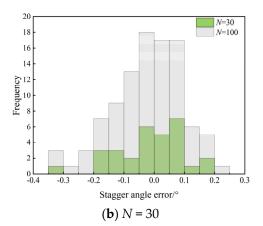


Figure 7. A schematic diagram of uncertainty propagation of stagger angle errors.

4.1. Measured data processing

The measured data of stagger angle errors in Section 2 are used as the input in UQ analysis process. In Section 3.2, the test function shows that when sample size N is greater than 40, the third-order PDDPC method can achieve a high calculation accuracy. In this section, we also use the third-order PDDPC method to study the influence of sample size on aerodynamic performance of compressor blade. Six groups (N = 20, 30, 40, 50, 60 and 80) of sample size are generated from the initial sample set (N = 100), respectively. Figure 8 shows the histograms for each sample size. These sampled data will be used as random inputs for PDDPC method. It can be observed that as the sample size N decreases, the probability information contained in the sampled data is further reduced. When the sample size N is 20 or 30, many measured error data have been lost compared with the initial sample set.





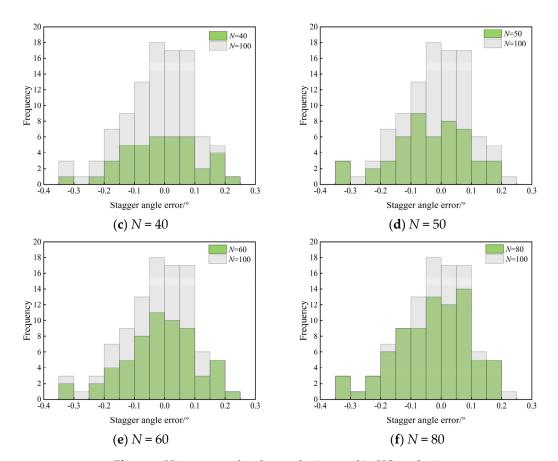


Figure 8. Histograms of each sample size used in UQ analysis.

4.2. Aerodynamic UQ results: loss coefficient $\,\omega\,$ and static pressure ratio $\,\pi\,$

Under the blocking condition of incidence i = -0.5°, Figure 9 shows the mean value $\mu(\omega)$ and standard deviation $\sigma(\omega)$ of the total pressure loss coefficient ω . It is worth noting that the previous open literature [1,25] has shown that 100 samples can be used for uncertainty propagation using the data-driven polynomial chaos method. Therefore, here we use the UQ results of 100 samples as the convergence reference values. It can be concluded from Figure 9a that the values of $\mu(\omega)$ gradually convergence with the increase of the stagger angle error samples. When the sample size N is 20 or 30, the values of $\mu(\omega)$ fluctuates greatly. When the sample size N is higher than 40, the values of $\mu(\omega)$ basically reach the convergence state. Figure 9b gives the convergence plot of $\sigma(\omega)$ with increasing data size N. It can be revealed that when the sample size is higher than 50, the results of $\sigma(\omega)$ can achieve a good convergence behavior.

For the static pressure ratio π of compressor blade, it can be found from Figure 10 that the when sample size exceeds 40, the values of $\mu(\pi)$ and $\sigma(\pi)$ estimated by PDDPC method can achieve aerodynamic results convergence. Under the operating condition of incidence i = -0.5°, the UQ results of the loss coefficient and static pressure ratio show that PDDPC method has satisfactory performance to address the problem of limited measured data. For the uncertainty evaluation of compressor blades, using fewer measured data can not only reduce the detection cost but also greatly speed up the uncertainty analysis cycle.

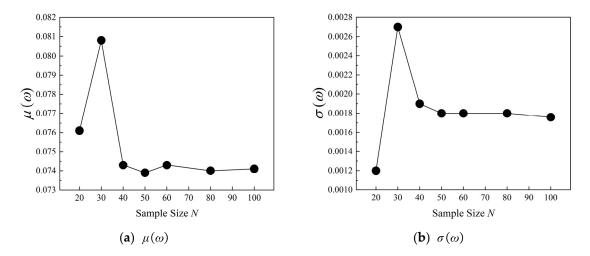


Figure 9. Convergence plots of $\mu(\omega)$ and $\sigma(\omega)$ under incidence i = -0.5°.

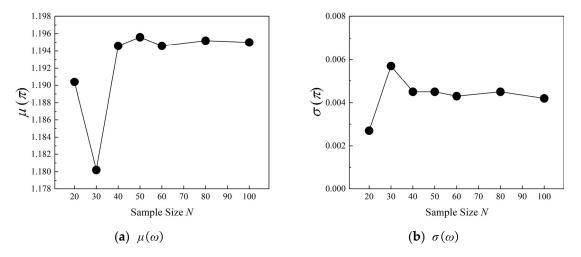
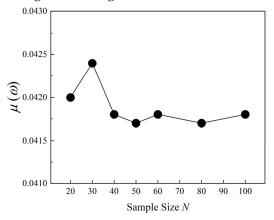


Figure 10. Convergence plots of $\mu(\pi)$ and $\sigma(\pi)$ under incidence i = -0.5°.

Figure 11 displays the convergence plots of $\mu(\omega)$ and $\sigma(\omega)$ under the design incidence i = 2.5° condition. It can be found that when sample size N exceeds 40, the values of $\mu(\omega)$ have good aerodynamic convergence behavior. Although the values of $\sigma(\omega)$ have obvious fluctuations with increasing sample size, these fluctuations are very small. It is worth mentioning that the loss coefficient ω of the compressor blade is less affected by the uncertain stagger angle errors at design incidence condition. Figure 12 shows the UQ results ($\mu(\pi)$ and $\sigma(\pi)$) of static pressure ratio. When the sample size N exceeds 40 and 50 respectively, the calculation results of $\mu(\pi)$ and $\sigma(\pi)$ can achieve good convergence rate.



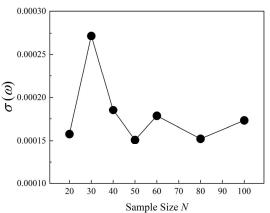




Figure 11. Convergence plots of $\mu(\omega)$ and $\sigma(\omega)$ under incidence $i = 2.5^{\circ}$.

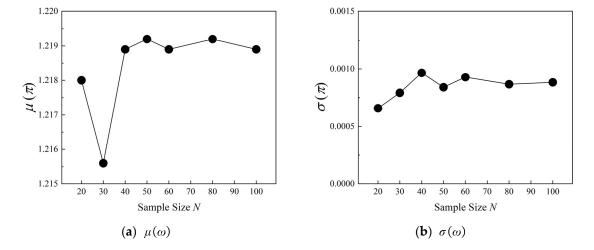


Figure 12. Convergence plots of $\mu(\pi)$ and $\sigma(\pi)$ under incidence $i = 2.5^{\circ}$.

Figure 13 and Figure 14 present the UQ results of ω and π obtained by PDDPC method under the operating conditions of incidence i = 7°. As can be seen from the convergence plots, when the sample size reaches 40, the UQ results of ω and π both exhibit a good convergence state. Table 3 gives the critical sample sizes for different aerodynamic parameters under various operating conditions. It can be found that for the PDDPC method, 40 or 50 can be regarded as the critical sample size for the aerodynamic parameters of loss coefficient and static pressure ratio. Generally speaking, the aerodynamic performance of compressor blades is more sensitive to the stagger angle error at high positive incidence condition. The convergent UQ results of aerodynamic parameters can provide more solid confidence for uncertainty analysis using fewer samples.

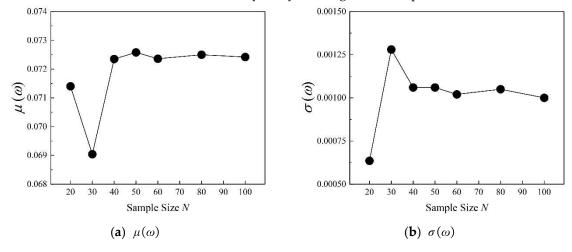


Figure 13. Convergence plots of $\mu(\omega)$ and $\sigma(\omega)$ under incidence $i = 7^{\circ}$.

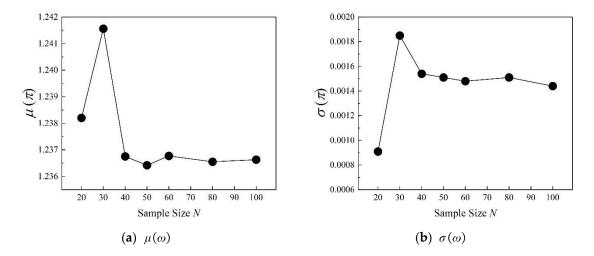


Figure 14. Convergence plots of $\mu(\pi)$ and $\sigma(\pi)$ under incidence $i = 7^{\circ}$.

Table 3. Critical sample sizes of aerodynamic parameters.

A 1	Critical sample size		
Aerodynamic parameters	<i>i</i> = -0.5°	$i = 2.5^{\circ}$	<i>i</i> = 7°
$\mu(\omega)$	40	40	40
$\sigma(\omega)$	50	40	40
$\mu(\pi)$	40	40	40
$\sigma(\pi)$	40	50	40

4.3. Aerodynamic UQ results: isentropic Mach number Ma_{is}

Figure 15 shows the mean value of the isentropic Mach number $\mu(Ma_{is})$ and the enlarged part at the incidence i = -0.5° condition. It can be clearly observed that the values of $\mu(Ma_{is})$ calculated by different sample sizes are essentially the same. When the sample size is 20 or 30, there are some differences on the acceleration section of blade suction side. The results of $\sigma(Ma_{is})$ are shown in Figure 16. For $\sigma(Ma_{is})$ on blade pressure side, the sensitive zones to uncertain stagger angle errors obtained by the lower sample size (N = 20 and 30) differ from other sample sizes. The predicted deviations occur mainly in the leading edge region and the back part of the blade. For $\sigma(Ma_{is})$ on blade suction side, the predicted deviations are mainly concentrated on the first 40% of normalized chord length when the sample size N is 20 or 30.

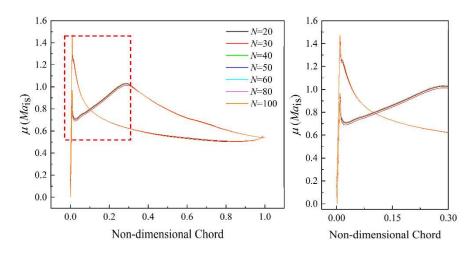


Figure 15. Results of $\mu(Ma_{is})$ with different sample sizes at incidence $i = -0.5^{\circ}$.

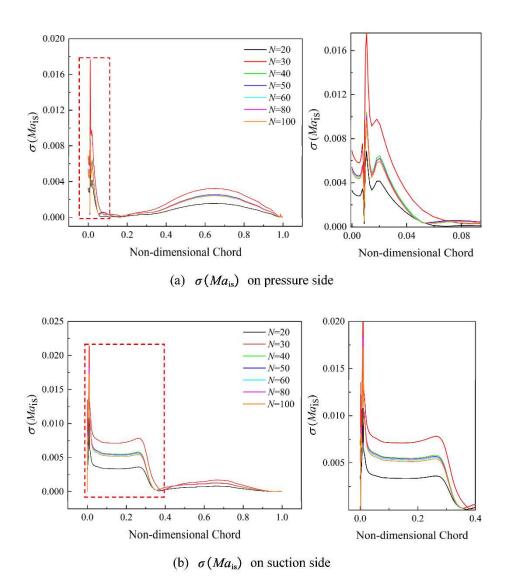


Figure 16. Results of $\sigma(Ma_{is})$ with different sample sizes at incidence $i = -0.5^{\circ}$.

At the design incidence condition, the influence of the sample size on the $\mu(Ma_{is})$ is similar to that of incidence i = -0.5°, as shown in Figure 17. Figure 18 shows the results of $\sigma(Ma_{is})$. It can be illustrated from Figure 18a that the influence of stagger angle error on the pressure side (PS) is mainly concentrated in the first 30% normalized chord length. When the sample size is lower than 40, there is an underestimation or overestimation of fluctuations of blade isentropic Mach number. A similar situation exists for the results of $\sigma(Ma_{is})$ on the suction side (SS).

Under the operation condition of incidence i = 7°, Figure 19 and Figure 20 show the convergence plots of $\mu(Ma_{is})$ and $\sigma(Ma_{is})$. As expected, the values of $\mu(Ma_{is})$ are less sensitive to the sample size of stagger angle errors. It can be seen from Figure 20 that although the trend of convergence plots for fewer samples (N = 20 or 30) is similar to the other sample sizes, there is a large prediction error. For compressor blade designers, the incorrect estimates of sensitive zones can cause a significant risk for blade robust design. Therefore, it is important to obtain a critical sample size that can be used for accurate UQ analysis. Table 4 gives the critical sample size of Ma_{is} under various operating conditions. We could observe that the $\mu(Ma_{is})$ is less sensitive to sample size of stagger angle error. Sample size N = 40 can be considered as the critical size.

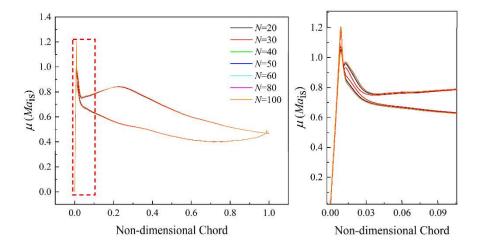


Figure 17. Results of $\mu(Ma_{is})$ with different sample sizes at incidence $i = 2.5^{\circ}$.

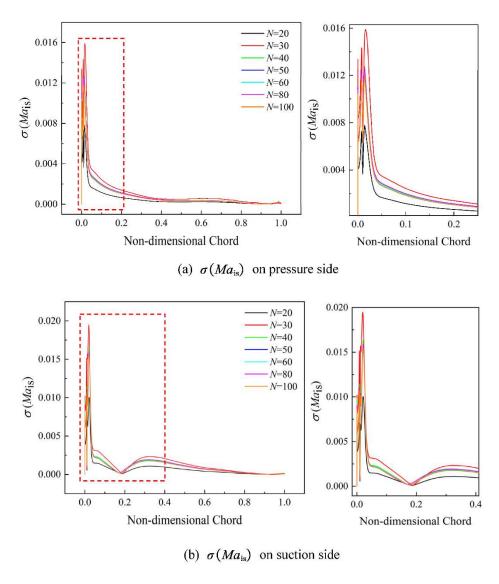


Figure 18. Results of $\sigma(Ma_{is})$ with different sample sizes at incidence $i = 2.5^{\circ}$.

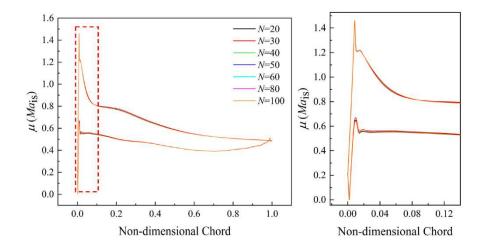


Figure 19. Results of $\mu(Ma_{is})$ with different sample sizes at incidence $i = 7^{\circ}$.

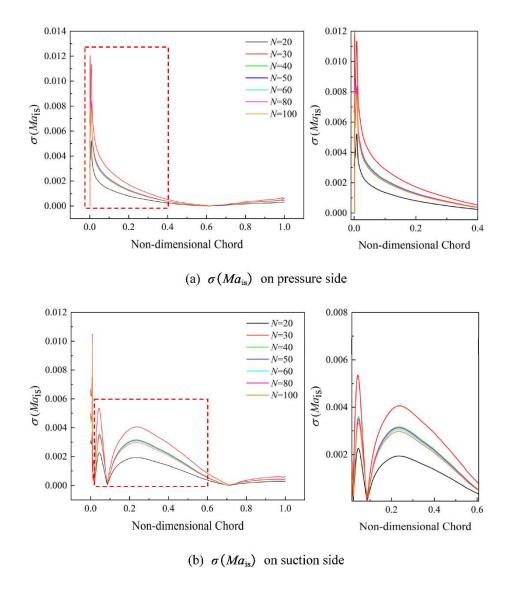


Figure 20. Results of $\sigma(Ma_{is})$ with different sample sizes at incidence $i = 7^{\circ}$.

Table 4. Critical sample sizes of *Ma*_{is}.

	<i>i</i> = -0.5°	<i>i</i> = 2.5°	<i>i</i> = 7°
$\mu(\mathit{Ma}_{\mathrm{is}})$			
$\sigma(Ma_{is})$ on PS	40	40	40
$\sigma(Ma_{\rm is})$ on SS	40	40	40

4.4. Aerodynamic UQ results: Mach contour

In this section, we consider the influence of sample size of stagger angle errors on Mach contour at different incidence conditions. At the incidence $i = -0.5^{\circ}$ condition, the contour of $\mu(\text{Mach})$ is shown in Figure 21. We could observe that the trends of μ (Mach) calculated by PDDPC method are broadly consistent when the sample size N is greater than 20. For the contour of $\sigma(Mach)$, it can be clearly found from Figure 22 that the flow field calculated by the sample size of 20 and 30 is quite different from the convergent flow field. Figure 23 and Figure 24 display the contour of $\mu(Mach)$ and $\sigma(\text{Mach})$ at the design incidence condition. As shown in Figure 23, it can be seen that the contour of μ (Mach) calculated by the sample size of 20 and 30 o differ from the convergent results (see the acceleration area dotted by rectangular dotted line). As can be seen from Figure 24, the contour of $\sigma(Mach)$ basically reaches the convergence state as the sample size exceeds 40. Figure 25 and Figure 26 display the contour of $\mu(Mach)$ and $\sigma(Mach)$ at incidence $i = 7^{\circ}$ condition. The contour of $\mu(Mach)$ is less sensitive to sample size of stagger angle errors. From Figure 26, it can be revealed that when sample size reaches 40, the contour of $\sigma(Mach)$ has been consistent with those of larger sample sizes. Table 5 gives the critical sample sizes of Mach at various incidence conditions. Based on these results, the blade designers can make reasonable choice about the sample size of manufactured blades in the uncertainty analysis stage, thus accelerating the UQ cycle and reducing costs.

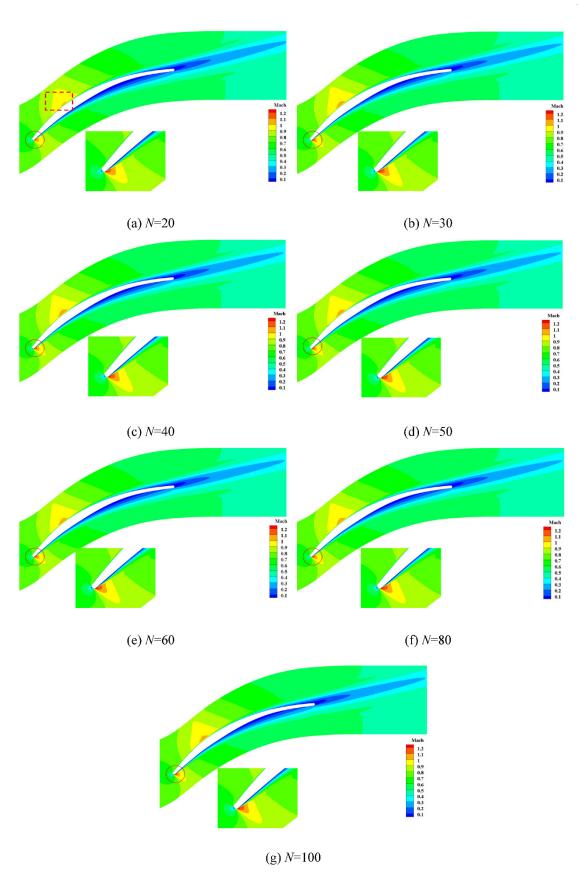


Figure 21. μ (Mach) contour with different sample sizes at incidence i = -0.5°.

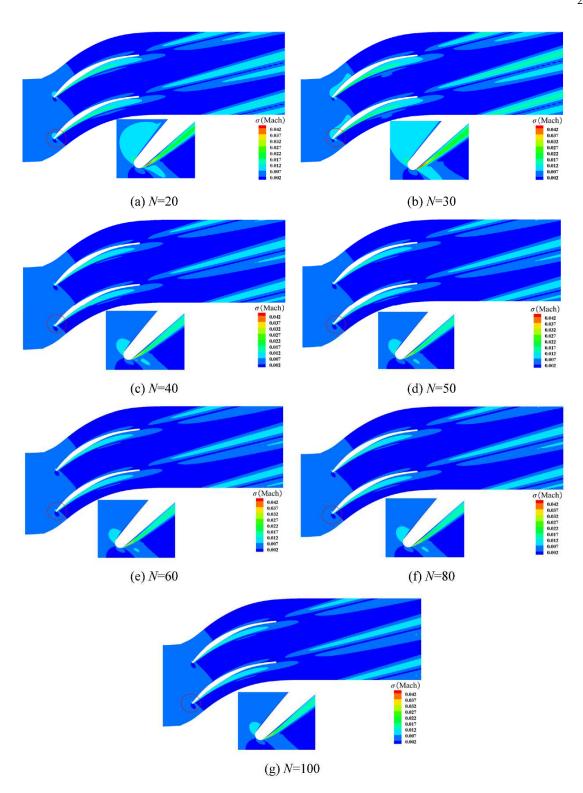


Figure 22. $\sigma(\text{Mach})$ contour with different sample sizes at incidence $i = -0.5^{\circ}$.



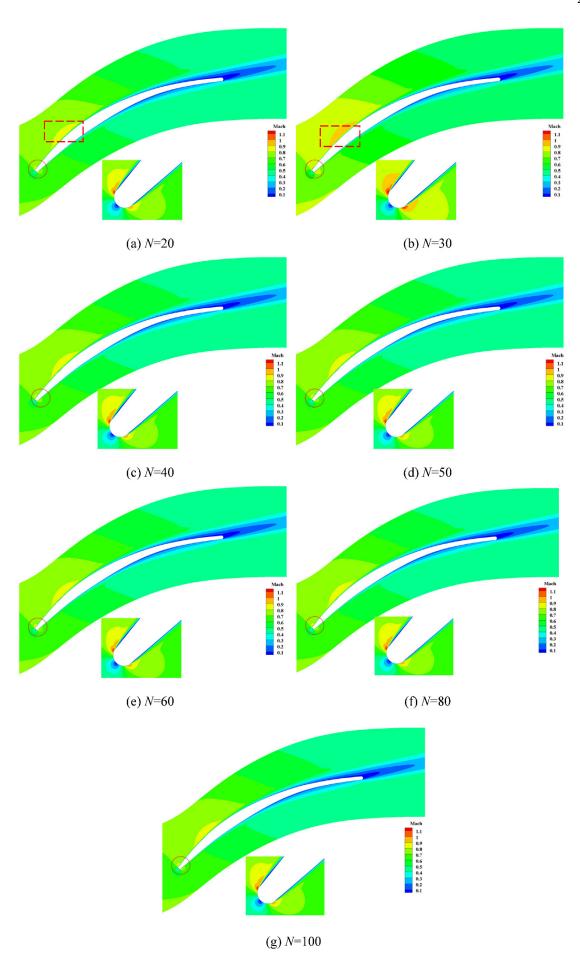


Figure 23. μ (Mach) contour with different sample sizes at incidence i = 2.5°.

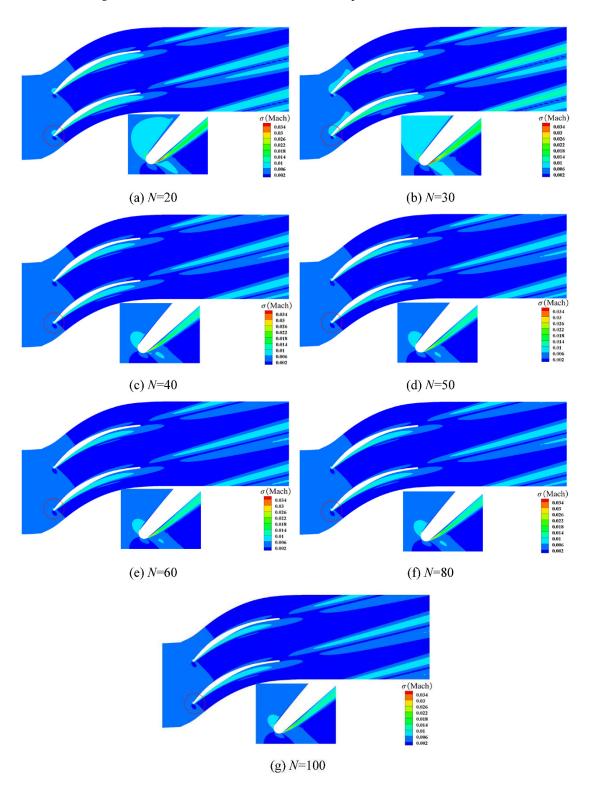


Figure 24. $\sigma(\text{Mach})$ contour with different sample sizes at incidence $i = 2.5^{\circ}$.

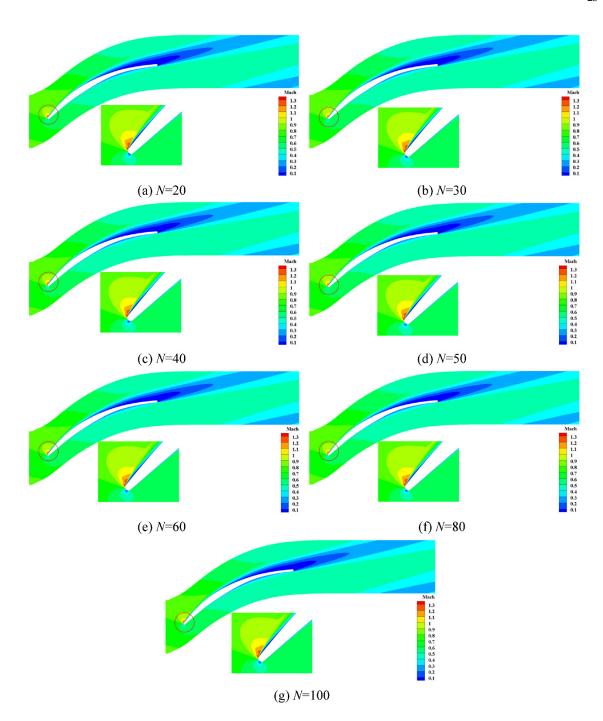


Figure 25. μ (Mach) contour with different sample sizes at incidence $i = 7^{\circ}$.



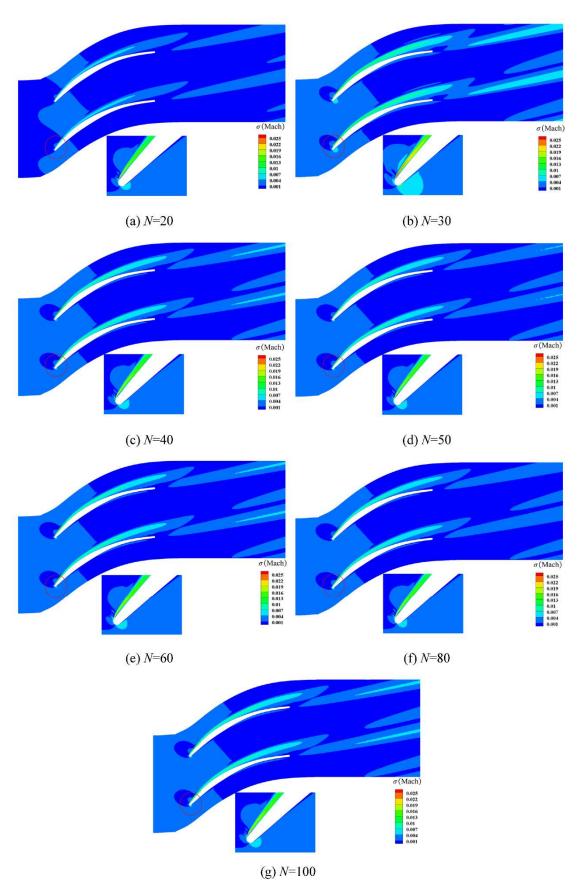


Figure 26. $\sigma(\text{Mach})$ contour with different sample sizes at incidence $i = 7^{\circ}$.

Table 5. Critical sample sizes of Mach.

Aerodynamic parameters -		Critical sample siz	ze
	<i>i</i> =-0.5°	<i>i</i> =2.5°	<i>i</i> =7°
$\mu(Mach)$	30	40	
$\sigma({ m Mach})$	40	40	40

4.5. Kullback-Leibler divergence analysis

As the sample size of measured data decreases beyond the critical size, the PDDPC method is unable to accurately propagate the uncertainty even with high order accuracy. In fact, different sample sizes contain different probability information. Here we use the Kullback-Leibler (KL) divergence [38] to quantify the difference in probability information between 100 sampled data (initial dataset) and other sample sizes (N=20, 30, 40, 50, 60, and 80) of stagger angle error data. These measured data have been shown in Section 4.1.

Given two sets of sampled data: $X = \left\{ \xi^{(i)} \right\}_{i=1}^{N_{h1}}$ and $Y = \left\{ \xi^{(j)} \right\}_{j=1}^{N_{h2}}$, the sampled data is divided into H groups. The probability that the sampled data falls within each group is p = p(X) and q = q(Y), respectively. The KL divergence can be computed by:

$$KL(p||q) = \sum_{i=1}^{H} p(X) \cdot (\log p(X) - \log q(Y))$$
 (17)

As can be seen in Figure 27, when the two sets of measured data are the same, the KL result is 0, indicating that there is no difference in probability information. As the sample size decreases, the probability information loss increases. When the sample size is lower than 40, the probability information loss increases sharply. Although the KL results for sample sizes 40 and 50 are close, combined with our analysis of the convergence behavior of aerodynamic parameters at different operating conditions, a better convergence state can be achieved when the sample size reaches 50. Therefore, from a more conservative point of view, we choose 50 as the critical sample size in this work with consideration of uncertain stagger angle errors.

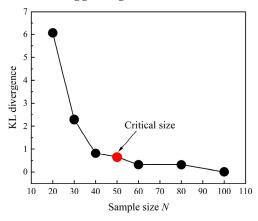


Figure 27. KL results of different sample sizes.

5. Conclusion

(1) A UQ scheme dedicated to dealing with limited samples is proposed based on a preconditioner-based data-driven polynomial chaos (PDDPC) method. The influence of sample size on the calculation accuracy of PDDPC is investigated using a nonlinear test function. Results show that the calculation accuracy of PDDPC method is dependent on the sample size of input data. When the sample size N is greater than 40, the results of $\mu(Y)$ and $\sigma(Y)$ of test function gradually converge to the benchmark values calculated by the Monte Carlo method.

conditions (blocking condition, design incidence condition and high incidence condition).

- (3) When the sample size of measured data of stagger angle error is greater than 50, the values of $\mu(\omega)$, $\sigma(\omega)$, $\mu(\pi)$ and $\sigma(\pi)$ can reach a good convergence state regardless of operating conditions. The sample size has little effect on $\mu(Ma_{is})$ and $\mu(Mach)$. Its main effect is centered on $\sigma(Ma_{is})$ and $\sigma(Mach)$. When the sample size is 20 or 30, the sensitive zones to stagger angle errors calculated by PDDPC have large prediction errors. The incorrect estimates of sensitive zones can cause a significant risk for blade robust design. With the sample size exceeds 40, the estimated results of $\sigma(Ma_{is})$ and $\sigma(Mach)$ are generally consistent. The UQ results of multiple aerodynamic parameters show that PDDPC method has satisfactory performance to address the problem of limited measured data; there exists a critical sample size that can be used for rapid UQ analysis of compressor blades
- (4) The KL divergence is used to analyze the probability information loss between the initial sample size (N=100) and other sample sizes (N=20, 30, 40, 50, 60, and 80) of stagger angle error data. Results show that when the sample size of error data is lower than 40, the probability information loss increases sharply. Combined with the aerodynamic UQ convergence results, the sample size of 50 can be regarded as the critical size to accurately propagate the uncertainty of stagger angle errors.

Acknowledgements: The authors gratefully acknowledge the support of the National Natural Science Foundation of China (grant number 51790512 and 52175436), the Innovation Foundation for Doctoral Dissertation of Northwestern Polytechnical University (grant number. CX2021075).

Conflicts of Interest: The authors do not have any possible conflicts of interest

References

- 1. Wang H, Gao L, Yang G, et al., A data-driven robust design optimization method and its application in compressor blade, Physics of Fluids, 2023, 35(6):066114.
- 2. Dow EA, Wang Q. The Implications of Tolerance Optimization on Compressor Blade Design, Journal of Turbomachinery, 2015, 137(10).
- 3. Zhao B, Ding W, Shan Z, et al., Collaborative manufacturing technologies of structure shape and surface integrity for complex thin-walled components of aero-engine: Status, challenge and tendency, Chinese Journal of Aeronautics, 2023, 36(7):1-24.
- 4. Wang P, Li S, Zhang D, et al., The machining error control of blade shape based on multivariate statistical process control, Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2014, 229(11):1912-24.
- 5. Garzon VE, Darmofal DL. Impact of Geometric Variability on Axial Compressor Performance, Journal of Turbomachinery, 2003, 125(4):692-703.
- 6. Xu S, Zhang Q, Wang D, et al. Uncertainty Quantification of Compressor Map Using the Monte Carlo Approach Accelerated by an Adjoint-Based Nonlinear Method, 2023.
- 7. Li M, Yu X, Meng D, et al. A New Approach for Deviation Modeling in Compressors: Sensitivity-Correlated Principal Component Analysis, 2023.
- 8. Ma C, Gao L, Wang H, et al., Influence of leading edge with real manufacturing error on aerodynamic performance of high subsonic compressor cascades, Chinese Journal of Aeronautics, 2020.
- 9. Yao W, Chen X, Luo W, et al., Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles, Progress in Aerospace Sciences, 2011, 47(6):450-79.
- 10. Pilko A, Ferraro M, Scanlan J. Quantifying Specific Operation Airborne Collision Risk through Monte Carlo Simulation, 2023.
- 11. Wang J, Wang B, Yang H, et al., Compressor geometric uncertainty quantification under conditions from near choke to near stall, Chinese Journal of Aeronautics, 2023, 36(3):16-29.
- 12. Mohanamuraly P, Müller J-D. An adjoint-assisted multilevel multifidelity method for uncertainty quantification and its application to turbomachinery manufacturing variability, International Journal for Numerical Methods in Engineering, 2021, 122(9):2179-204.

26

- 13. Phan HM. Modeling of a turbine bladerow with stagger angle variation using the multi-fidelity influence superposition method, Aerospace Science and Technology, 2022, 121:107318.
- 14. Xiu D, Karniadakis GE. The Wiener--Askey Polynomial Chaos for Stochastic Differential Equations, SIAM Journal on Scientific Computing, 2002, 24(2):619-44.
- 15. Zhang K, Li J, Zeng F, et al., Uncertainty Analysis of Parameters in SST Turbulence Model for Shock Wave-Boundary Layer Interaction, 2022, 9(2):55.
- 16. Shi W, Chen P, Li X, et al., Uncertainty Quantification of the Effects of Small Manufacturing Deviations on Film Cooling: A Fan-Shaped Hole, 2019, 6(4):46.
- 17. Lange A, Voigt M, Vogeler K, et al., Impact of Manufacturing Variability on Multistage High-Pressure Compressor Performance, Journal of Engineering for Gas Turbines and Power, 2012, 134(11).
- 18. Xia Z, Luo J, Liu F. Performance impact of flow and geometric variations for a turbine blade using an adaptive NIPC method, Aerospace Science and Technology, 2019, 90:127-39.
- 19. Salehi S, Raisee M, Cervantes MJ, et al., An efficient multifidelity ℓ1-minimization method for sparse polynomial chaos, Computer Methods in Applied Mechanics and Engineering, 2018, 334:183-207.
- 20. Kiureghian AD, Liu PL. Structural Reliability under Incomplete Probability Information, Journal of Engineering Mechanics, 1986, 112(1):85-104.
- 21. Red-Horse JR, Benjamin AS. A probabilistic approach to uncertainty quantification with limited information, Reliability Engineering & System Safety, 2004, 85(1):183-90.
- 22. Oladyshkin S, Class H, Helmig R, et al., A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations, Advances in Water Resources, 2011, 34(11):1508-18.
- 23. Oladyshkin S, Nowak W. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion, Reliability Engineering & System Safety, 2012, 106:179-90.
- 24. Ahlfeld R, Belkouchi B, Montomoli F. SAMBA: Sparse Approximation of Moment-Based Arbitrary Polynomial Chaos, Journal of Computational Physics, 2016, 320:1-16.
- 25. Ahlfeld R, Montomoli F. A Single Formulation for Uncertainty Propagation in Turbomachinery: SAMBA PC, Journal of Turbomachinery, 2017, 139(11).
- 26. Guo Z, Chu W, Zhang H. A data-driven non-intrusive polynomial chaos for performance impact of high subsonic compressor cascades with stagger angle and profile errors, Aerospace Science and Technology, 2022, 129:107802.
- 27. Guo L, Liu Y, Zhou T. Data-driven polynomial chaos expansions: A weighted least-square approximation, Journal of Computational Physics, 2019, 381:129-45.
- 28. Kun W, Fu C, Jianyang Y, et al., Nested sparse-grid Stochastic Collocation Method for uncertainty quantification of blade stagger angle, Energy, 2020, 201:117583.
- 29. Phan HM, He L. Efficient modeling of mistuned blade aeroelasticity using fully-coupled two-scale method, Journal of Fluids and Structures, 2022, 115:103777.
- 30. Lu Y, Green J, Stapelfeldt SC, et al., Effect of Geometric Variability on Running Shape and Performance of a Transonic Fan, Journal of Turbomachinery, 2019, 141(10).
- 31. Suriyanarayanan V, Rendu Q, Vahdati M, et al., Effect of Manufacturing Tolerance in Flow Past a Compressor Blade, Journal of Turbomachinery, 2021, 144(4).
- 32. Chuang-Stein C. Sample size and the probability of a successful trial, Pharm Stat, 2006, 5(4):305-9.
- 33. He X, Zhao F, Vahdati M. Uncertainty Quantification of Spalart–Allmaras Turbulence Model Coefficients for Compressor Stall, Journal of Turbomachinery, 2021, 143(8).
- 34. Wiener N. The Homogeneous Chaos, Journal of Computational Physics, 1938, 60(4):897-936.
- 35. Xiu D, Karniadakis GE. Modeling uncertainty in flow simulations via generalized polynomial chaos, Journal of Computational Physics, 2003, 187(1):137-67.
- 36. Zheng M, Wan X, Karniadakis GE. Adaptive multi-element polynomial chaos with discrete measure: Algorithms and application to SPDEs, Applied Numerical Mathematics, 2015, 90:91-110.
- 37. Gao L, Ma C, Cai M, et al., Influence of uncertain inflow conditions on a subsonic compressor cascade based on wind tunnel experiment, Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2022, 236(15):8285-99.
- 38. Martín F, Moreno L, Garrido S, et al. Kullback-Leibler Divergence-Based Differential Evolution Markov Chain Filter for Global Localization of Mobile Robots, Sensors, 2015, 15(9):23431-58.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.