Master of Science

BAYESIAN INFERENCE OF γ-Re_θ TRANSITIONAL MODEL COEFFICIENT BASED ON PC-NIPC METHOD

The Graduate School

of University of Ulsan

School of Mechanical Engineering

Nguyen Hoai Thanh

BAYESIAN INFERENCE OF γ-Re_θ TRANSITIONAL MODEL COEFFICIENT BASED ON PC-NIPC METHOD

Supervisor: KyoungSik Chang

A Thesis

Submitted to

the Graduate School of University of Ulsan

In partial Fulfillment of the Requirements

for the Degree of

Master of Science

by

Nguyen Hoai Thanh

School of Mechanical Engineering

University of Ulsan, Republic of Korea

May 2019

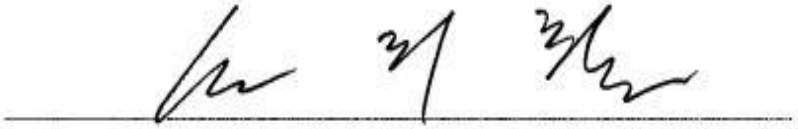
BAYESIAN INFERENCE OF γ-Re_θ TRANSITIONAL MODEL COEFFICIENT BASED ON PC-NIPC METHOD

This certifies that the masters thesis

of Nguyen Hoai Thanh is approved.

Committee Chairman Prof. Lee, Sang-Wook

Committee Member Prof. Chang, Kyoungsik



Committee Member Prof. Shin, Jichul

Department of Mechanical Engineering

University of Ulsan, Republic of Korea

August 2019

Acknowledgements

I would like to express my appreciation to my advisor, Prof. Chang Kyoungsik, for his guidance, advice and support during my study in University of Ulsan. All of my works in this dissertation cannot be accomplished without his support.

I would also like to thank to Prof. Lee Sang-Wook and Prof. Shin Jichul for serving on my graduate committee and providing suggestions and comments throughout this thesis.

I would like to thank my friends, my colleagues, all members of TTF Lab., for their help, guidance, and co-working during the time I study in University of Ulsan.

Above all, I would like to express my parents, my sister. Thank you for everything!

Ulsan, MAY 2019

NGUYEN HOAI THANH

Abstract

BAYESIAN INFERENCE OF γ-Re_θ TRANSITIONAL MODEL COEFFICIENT BASED ON PC-NIPC METHOD

In the present work, a comparative study of two major Non-Intrusive Polynomial Chaos methods, Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) and Non-Intrusive Spectral Projection (NISP), was conducted for the transitional $\gamma - Re_\theta$ transitional model. Three multiple model coefficients, $c_{a2,}$ $c_{e1,}$ and c_{e2}, were considered as multiple random inputs with the assumption of uniform distributions with ±10% deviation. The target transitional flows were one around a flat plate and Aerospatiale A-airfoil. Deterministic solutions were obtained by employing the open source software OpenFOAM. The results of two methods were compared to the results of Monte Carlo simulation with 500 runs. The order convergence of the mean value and the standard deviation (STD) were compared in terms of the quantities of interest, drag and lift coefficients. Further, the most effective model coefficient for each transitional flow can be found through the calculation of the Sobol index. And then we apply Bayesian Inference to demonstrate inverse problem to find the mean and stand deviation of the parameters constant in $\gamma - Re_\theta$ transitional model, and create the correlation matrix among the parameters with surrogate model which was made by Point-Collocation Non-intrusive Polynomial Chaos.

The Bayesian parameter calibration approach based on gPCE is integrated to the developed comprehensive framework of analyzing and identification. The gPCE is applied to the parameter calibration in two ways. The first one is using the gPC approximation as the surrogate model. The second fashion is expanding the recursive Bayesian estimator with the polynomial chaos basis. This technique, which is quite new, provides good results and has attractive properties.

Keywords: Point collocation; Spectral projection; Non-Intrusive polynomial chaos; CFD;Uncertainty quantification; Transition model.

Contents

List of Figures

List of Tables

Abbreviations

2D	Two dimensional
BC	Boundary Condition
gPC	Generalized Polynomial Chaos
gPCE	Generalized Polynomial Chaos Expansion
IC	Initial Condition
LHS	Latin Hypercube Sampling
MC	Monte-Carlo
MCMC	Markov-Chain-Monte-Carlo
CFD	Computational fluid dynamics
NIPC	Point-Collocation Non-Intrusive Polynomial
NISP	Non-intrusive Spectral Projection
PC	Polynomial chaos
PCE	Polynomial Chaos Expansion
PDF	Probability density function
UQ	Uncertainty quantificationa

Chapter 1. Introduction

This chapter introduces the concept of uncertainty in predictive science. The first section will explain the so-called path to knowledge that scientists and engineers use to make predictions of the physical reality surrounding us. It also describes the different kind of uncertainties, i.e. Consequently, the second section describes the thesis's objective. Finally, the last section outlines the structure of this document.

1.1 Introduction

Nowadays, Computational Fluid Dynamics (CFD) has become an indispensable tool in the design and development of engineering products. The target is to reduce the time and cost for the new design of a prototype via computer simulations. Recently, the rapid progress of computation performance and reduction of computational cost have enabled the adoption of the Uncertainty Quantification technique (UQ) [1-7], which requires many simulation cases. One important part of Uncertainty Quantification (UQ) is uncertainty propagation which is used to quantify the sensitivity and variation of the simulation results with consideration of single or multiple stochastic inputs such as the freestream velocity, model coefficients, and target geometry. A direct method for UQ is Monte Carlo simulation, but it is much expensive to obtain converged solutions. Therefore, the polynomial chaos (PC) approach developed by Wiener [8-9] is proposed as an alternative method. Such methods can be divided into two types: intrusive and non-intrusive approaches. The intrusive methods require modification of the existing code into a Galerkin projection-based flow solver [10-15]. This extra work can be a big burden for CFD engineers. Whereas a non-intrusive method that enables the use of established code without any modification is considered as a promising approach in UQ and has been applied to various CFD fields by many researchers. Hosder and Walters [16-17] applied Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) to selected computational fluid dynamics (CFD) problems such as aerodynamic flows around the airfoil/wing or in a nozzle. They considered the uncertainty of the freestream velocity, eddy viscosity ratio, and model coefficients of the adopted turbulence models. They demonstrated that Point-Collocation Non-Intrusive Polynomial (NIPC) improved the

accuracy of the polynomial chaos coefficient in NIPC approaches and reduced the computational expense by achieving the same level of the accuracy with a small number of sample points. Loeven [18] introduced a non-intrusive Probabilistic Collocation approach for efficient propagation of arbitrarily distributed parametric uncertainties. They also showed that the convergences depending on the order of approximation of the Probabilistic Collocation method and the Galerkin Polynomial Chaos method are the same. In the UQ technique for turbulence model coefficients, Platteeuw *et al*. [19] applied the Probabilistic Collocation method to the model coefficients of the standard k-ε turbulence model. They found the necessary order for proper quantification of drag coefficients around a flat plate and NACA 0012 airfoil. Schaefer *et al*. [20] studied uncertainty quantification of three turbulence models: the Spalart-Allmaras model, Wilcox k-ω model, and Menter's SST model for transonic wall-bounded flows. The significant model coefficients of three turbulence models were investigated through a sensitivity analysis.

The mechanism of transitional flow from laminar to turbulent flow is still a difficult problem in engineering and science fields. In particular, many researchers have developed various transition models based on Reynolds-Averaged Navier-Stokes (RANS) equations to apply to engineering problems to reduce the computational cost and time. Recently, Langtry and Menter [21-24] proposed a promising transition model based on local parameters of the intermittency γ and the momentum thickness Reynolds number Re_θ. This transition model has been applied to various engineering problems and the results such as the transition point and skin friction variation are superior to previous transition model results. However, in any transition or turbulence model, many assumptions and simplifications are required to close model coefficients with uncertainties. This factor is the motivation of the present work. We considered the uncertainties of the model coefficient of the $\gamma - R_\theta$ transition model and applied the UQ technique to the parameters.

In the present work, two non-intrusive methods, Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) and the Non-Intrusive Spectral Projection method, were

applied to the transitional flow around a flat plate and Aerospatiale A-Airfoil. The random inputs for UQ are the closure model coefficients of Menter's $\gamma - Re_\theta$ transition model and the quantities of the interests (QoIs) are the skin friction coefficient and drag/lift coefficients.

Even though the UQ methodologies used in the present work are not new one but an old-fashioned one in the UQ fields, we try to access the application capability of two UQ methods with somewhat transitional algorithm in the transitional flow simulation; the first one is NISP with the sampling of the quadrature rule and the second is Point-Collocation NIPC with the oversampling 2.

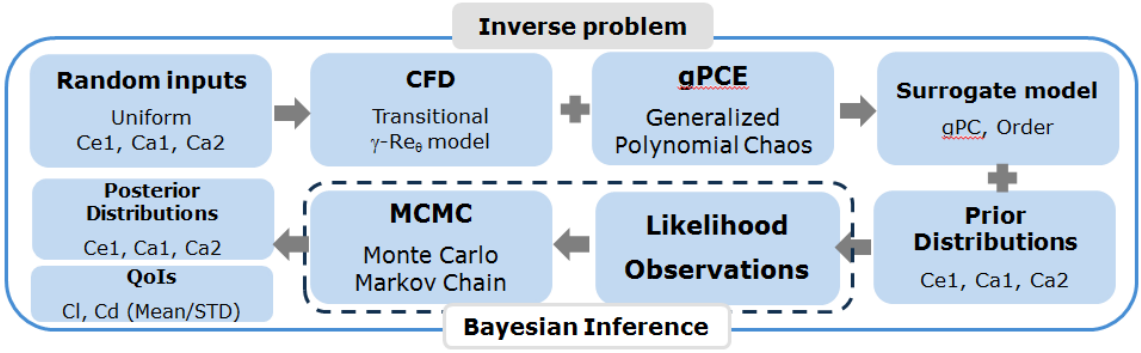


Fig. 1.1 *Process for Bayesian Inference and Overall simulation*

1.2 Thesis Objectives

The main purpose of the thesis is to investigate whether it is possible to quantify and to reduce the uncertainty of Menter's $\gamma - Re_\theta$ transition models using the Bayesian inference theorem. More specifically, we try to find out the constant parameter values for this model.

An underlying objective is to do determine the prior uncertainty by mean of the analyst's knowledge of the used model-instead of merely guessing, and the development of a methodical way of creating and verifying the correctness a surrogate model for the Menter's $\gamma - Re_\theta$ transition models.

1.3 Thesis Outline

The thesis comprises six chapters. After the introduction in this chapter, the theory problem definition which is the main contribution of the thesis is proposed in chapter 2. It introduce about the Polynomial Chaos, and the intrusive method, Bayesian inference, and Markov Chain Monte Carlo method. Next, in the chapter 3 we explain about the Menter's $\gamma - Re_\theta$ transition models and the uncertain parameters, is to define prior uncertainty interval for these parameter. This is followed by an extensive discussion on the determination of prior uncertainty interval for each of the considered model's uncertain parameters.

In chapter 4 we simulate for two cases over flat plate and Aerospatiale airfoil. In this chapter we a comparative study of two major Non-Intrusive Polynomial Chaos methods, Point-Collocation Non-Intrusive Polynomial Chaos (NIPC) and Non-Intrusive Spectral Projection (NISP), was conducted for the transitional γ-Re_θ transitional model. Three multiple model coefficients, Ca_2, Ce_1, and Ce_2, were considered with multiple random inputs with the assumption of uniform distributions with $\pm 10\%$ deviation. The target transitional flows were one around a flat plate and Aerospatiale A-airfoil. Deterministic solutions were obtained by employing the open source software OpenFOAM. The results of two methods were compared to the results of Monte Carlo simulation with 500 runs.

The next step, we inverse problem by using Bayesian inference via MCMC method which is explained in the chapter 5.In order to do that, we use the point collocation method to create surrogate models.

Finally, chapter 6 discusses the results of the Bayesian calibration for Menter's $\gamma - Re_\theta$ transition models. The Ce_1 is the most sensitive in parameters of transitional model. The correlation between input and output of model.

Chapter 2. Theory and Problem Definition

This chapter provides the theoretical basic. The mathematical model of PC and the Bayesian Inference, the Markov Chain Monte Carlo method are described first. The modeling of uncertainties in PC with Bayesian approach establishes a system with random variables.

2.1 Polynomial Chaos and generalized Polynomial Chaos

Consider an arbitrary real-valued random variable X = X (ω) according to some probability space (ω,ψ,p), with sample space $\Omega,\sigma-a\lg ebar,\psi,$ and probability measure P. For the stochastic formulation discussed here, it is required that the RVs are square-integrable and have a finite variance, i.e., X $\in$ L_2 (ω,ψ,p)= {X:E(X^2) < ∞}.

Let $\{\xi_i\}_{i=1}^{\infty}$ be a sequence of centered, normalized, mutually orthogonal Gaussian Variables. Let P_p denote the space of polynomial in $\{\xi_i\}_{i=1}^{\infty}$ having degree less or equal to $P \in N$ .Furthermore, let $\Xi_P \subset P_p$ be the set of polynomials that belong to P_p and are orthogonal P_{P-1}, and define P_p as the space spanned by Ξ_p. It yields:

$$P_p = P_{N-1} \oplus P_p \qquad\qquad L_2(\Omega,\Psi,P) = \overset{\infty}{\underset{i=0}{\oplus}} P_i \tag{2.1}$$

The subspace P_p of $L_2(\Omega,\Psi,P)$ is called the P-th Homogeneous Chaos, and Ξ_p is called the polynomial Chaos (PC) of order P. The PC expansion of a random variable X is:

$$X(\omega) = a_0\Xi_0 + \sum_{i_1=1}^{\infty} a_{i_1}\Xi_1\left(\xi_{i_1}(\omega)\right) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Xi_2\left(\xi_{i_1}(\omega),\xi_{i_2}(\omega)\right) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1}a_{i_2}a_{i_3}\Xi_3\left(\xi_{i_1}(\omega),\xi_{i_2}(\omega),\xi_{i_3}(\omega)\right)+\ldots \tag{2.2}$$

The Polynomial Chaos introduced by Wiener are constructed from the *Hermite Polynomial* $H_{k\ (.)}$. Cameron and martin proved the expression, in case of the PC constructed from the Hermite Polynomial, is convergent for any RVs X (ω) in the Hilbert space $L_2(\Omega)$.

2.1.2 Generalized Polynomial Chaos

Being a spectral polynomial this expansion has an exponential convergent rate. However, for non-Gaussian random variables, the expansion may exhibit low convergence rates and thus require a high truncation order. Therefore, Xiu and Karniadakis [12] employed the Askey-scheme to generalize the original Wiener's PC expansion to some common non-Gaussian measures, which replace the Hermite Polynomial by other orthogonal polynomials and named it as generalized Polynomial Chaos Expansion (gPCE). Xiu concluded that the Cameron-Martin theorem (Eq. (3.3)) can be generalized to the gPCE, because each type of polynomials forms the Askey scheme from a complete basis in the Hilbert space determined by their corresponding random vector ξ [11]. Table 3.1 shows the correspondences between the random variable distribution and orthogonal polynomial family.

The distribution established in the gPC expansion match the distribution used in the developed framework in this thesis well. As discussed in the the probability distribution functions, which can be derived from the typical information found in engineering are the normal distribution, the uniform distribution and the exponential distribution. The exponential distribution is a special case of the gamma distribution.

Table 2.1 *Correspondence between the type of generalized Polynomial Chaos their underlying random variables*

	pdf of RV $\xi(\omega)$	**gPC basis polynomials** $\Psi(\xi)$	**Support**
Continuous	Gaussian	Hermite	$(-\infty,\infty)$
	Gamma	Laguerre	$[0,\infty)$
	Beta	Jacobi	[a,b]
	Uniform	Legendre	[a,b]
Discrete	Poisson	Charlier	$\{0,1,2,\ldots\}$

	Binomial	Krawtchouk	{0,1,2,…,n}
	Negative Binomial	Meixner	{0,1,2,…}
	Hypergeometric	Hahn	{0,1,2,…,n}

2.1.3 Single random variable

The RV X(ω) can be expanded by one-dimensional gPC basis functions as:

$$X(\omega)=\sum_{K=0}^{\infty} a_k \psi_k\left(\xi(\omega)\right) \tag{2.3}$$

where $\psi_k(.)$ denotes the one-dimensional orthogonal polynomial with the degree of k, e.g. Hermite Polynomials H_k (.), Legendre Polynomials Le_k(.), or Laguerre Polynomials La_k(.).

2.1.4 Multivariate random variable

Now the one-dimensional gPC basis is extended to the d-dimensional PC-basis. Given $\xi(\omega)=(\xi_1,\xi_2,...,\xi_d)$ a set of centered, normalized and mutually orthogonal Gaussian random variables, the PC expansion (2.3) has a form:

$$X(\omega)=a_0\Xi_0+\sum_{i_1=1}^{d} a_{i_1}\Xi_1\left(\xi_{i_1}(\omega)\right)+\sum_{i_1=1}^{d}\sum_{i_2=1}^{i_1} a_{i_1 i_2}\Xi_2\left(\xi_{i_1}(\omega),\xi_{i_2}(\omega)\right)+\sum_{i_1=1}^{d}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} a_{i_1}a_{i_2}a_{i_3}\Xi_3\left(\xi_{i_1}(\omega),\xi_{i_2}(\omega),\xi_{i_3}(\omega)\right)+... \tag{2.4}$$

To simplify the notation, the multi-index notation $i=(i_1,i_2,...,i_\alpha,...,i_d)$ with $|i|=\sum_{\alpha=1}^{d} i_\alpha$ is adopt. Let $i(\alpha)$ denote the $\alpha-$th element of the multi-index i. For construction gPC, the set of multi-indices $\rho(P)$ is defined as:

$$\rho(P):=\{i:|i|=p\} \tag{2.5}$$

The P-th Polynomial Chaos Ξ_p in equation (3.5) is constructed from one dimensional gPC basis functions as:

$$\Xi_p := \left\{ \bigcup_{i \in p(P)} \prod_{\alpha=1}^{d} \psi i(\alpha)(\xi_\alpha) \right\} \tag{2.6}$$

The expansion (2.4) can be rewritten in more compact form as

$$X(\omega) = \sum_{|i|=0}^{\infty} \beta_i \psi_i \left(\xi_1, \xi_2, ..., \xi_d \right) \tag{2.7}$$

Where β_i are the deterministic expansion coefficients and ψ_i are polynomials constructed by tensor products of the 1D-polynomials $\psi_{i(\alpha)}\left(\xi_\alpha\right)$.

$$\psi_i \left(\xi_1, \xi_2, ..., \xi_d \right) = \prod_{\alpha=1}^{d} \psi_{i(\alpha)} \left(\xi_\alpha \right) \tag{2.8}$$

Although using multi-index formulation is very clear, the single index is preferable to express gPC expansion. The multi-index can be converted to the single index. To do this, the lexicographic order is applied as also used by [12]. The relation between the single index k, the multi-index I and the orthogonal polynomial ψ_k is established in table 2.1 by using single index k, equation (2.7) is rewrite as

$$X(\omega) = \sum_{k=0}^{\infty} \beta_k \psi_k \left(\xi_1, \xi_2, ..., \xi_d \right) \tag{2.9}$$

with the spectral convergence property of the orthogonal projection, the expansion can be truncated to finite dimension. The finite-dimensional decomposition of in single-index form is:

$$X(\omega) \approx X_p(\omega) = \sum_{k=0}^{N_p} \beta_k \psi_k (\xi(\omega)) \tag{2.10}$$

where the basis dimension N_p is related to the dimension of the multivariate random variable d and the polynomial order P by

$$N_p + 1 = \frac{(P+d)!}{p!d!} \tag{2.11}$$

Table 2.2 *Single index, multi-index, and tensored polynomial*

$\|i\|$	$i=(i_1,i_2,...,i_\alpha,...,i_d)$	polynomials $\Psi(\xi)$	k
0	(0,0,…,0,0)	$\psi_0(\xi_1)\psi_0(\xi_2)...\psi_0(\xi_d)=1$	0
	(1,0,…,0,0)	$\psi_1(\xi_1)\psi_0(\xi_2)...\psi_0(\xi_d)=\psi_1(\xi_1)$	1
1	(0,1,…,0,0)	$\psi_0(\xi_1)\psi_1(\xi_2)...\psi_0(\xi_d)=\psi_1(\xi_2)$	2
	⋮	⋮	⋮
	(0,0,…,0,1)	$\psi_0(\xi_1)\psi_0(\xi_2)...\psi_1(\xi_d)=\psi_1(\xi_2)$	$\binom{d+1}{d}-1=d$
	(2,0,0,0,…,0)	$\psi_2(\xi_1)\psi_0(\xi_2)\psi_0(\xi_3)\psi_0(\xi_4)...\psi_0(\xi_d)$ $\psi_2(\xi_1)\psi_0(\xi_2)\psi_0(\xi_3)\psi_0(\xi_4)...\psi_0(\xi_d)$	d+1
	(1,1,0,0,…,0)	$\psi_1(\xi_1)\psi_1(\xi_2)\psi_0(\xi_3)\psi_0(\xi_4)...\psi_0(\xi_d)$	d+2
	(1,0,1,0,…,0)	$\psi_1(\xi_1)\psi_0(\xi_2)\psi_1(\xi_3)\psi_0(\xi_4)...\psi_0(\xi_d)$	
2	⋮	⋮	⋮
	(1,0,0,…,0,1)	$\psi_1(\xi_1)\psi_0(\xi_2)\psi_0(\xi_3)...\psi_0(\xi_{d-1})\psi_1(\xi_d)$	
	⋮	⋮	

	(0,0,0,…,0,2)	$\psi_0(\xi_1)\psi_0(\xi_2)\psi_0(\xi_3)...\psi_0(\xi_{d-1})\psi_2(\xi_d)$	$\binom{d+2}{d}-1$
	(3,0,…,0)	$\psi_2(\xi_1)\psi_0(\xi_2)\psi_0(\xi_3)\psi_0(\xi_4)...\psi_0(\xi_d)$	$\binom{d+2}{d}$
3	⋮	⋮	⋮

The polynomials $\psi_k(.)$ form an orthogonal and complete basis in L_2. With respect to inner product, the orthogonality is expressed as:

$$\langle \psi_i \mid \psi_j \rangle \equiv \int \psi_j(\xi)\psi_i(\xi)\, p(\xi)\, d\xi = \delta_{ij} \langle \psi_i \mid \psi_i \rangle = \delta_{ij} \left\| \psi_i^2 \right\| = \gamma_i \tag{2.12}$$

Because of the orthogonal property, the statistical values can be calculated directly from PCE coefficients. The mean of the RV X(ω) is given by

$$\bar{X} = E\left(X(\omega)\right) = E\left[\sum_{k=0}^{\infty} \beta_k \psi_k(\xi(\omega))\right]$$

$$= \sum_{k=0}^{\infty} \beta_k E[\psi_k(\xi(\omega)).1]$$

$$= \sum_{k=0}^{\infty} \beta_k \langle \psi_k \mid \psi_0 \rangle$$

$$\bar{X} = \beta_0. \tag{2.13}$$

Similarly, the variance of the RV $\sigma_X^2 = Var[X(\omega)]$ is:

$$Var[X(\omega)] = E\left[\left(X(\omega) - \bar{X}\right)^2\right]$$

$$= E\left[\left(\sum_{k=0}^{\infty}\beta_k\psi_k(\xi(\omega)) - \beta_0\right)^2\right]$$

$$= E\left[\left(\sum_{k=1}^{\infty}\beta_k\psi_k(\xi(\omega))\right)^2\right]$$

$$= \sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\beta_k\beta_l E[\psi_k(\xi(\omega))\psi_l(\xi(\omega))]$$

$$= \sum_{k=1}^{\infty}\beta_k^2\left\|\psi_k\right\|^2$$

$$\sigma_X^2 \approx \sum_{k=1}^{N_P}\beta_k^2\left\|\psi_k\right\|^2. \qquad (2.14)$$

2.1.4 Dependent random variables

The PCE requires the mutually independent random variables and the re-presentation of the joint density $p_{\xi}(\xi) = \Pi_{i=1}^{d} p\xi_i(\xi_i)$ for implementation as assumed in the course of this thesis. For correlated RV ξ_i, the Nataf transformation in a combination with a Cholesky decomposition can be employed to construct mutually independent Gaussian random variables, if marginal distributions and a correlation matrix are available. The reader is marginal distributions and a correlation matrix are available. The reader is referred to [LM10] for more detail about the PCE dependent RV.

2.1.5 PC expansion of random processes

The random process $\chi(r,\omega)$ in L_2can be separated into the deterministic and the stochastic parts with the stochastic discretization methods mentioned. The polynomial chaos expansion is mentioned in the section as one of the series expansion methods. The

PC expansion of the random process can be considered as the generalization of the PC expansion of the random vector, namely

$$\chi(r,\omega) = \sum_{k=0}^{\infty} x_k(r)\psi_k(\xi(\omega)) \approx \sum_{k=0}^{N_P} x_k(r)\psi_k(\xi(\omega)) \tag{2.15}$$

where the deterministic function $x_k(r)$ are called stochastic modes of the process. It is noted that the PC expansion contains more information than the second-order properties of random field $\chi(r,\omega)$. The knowledge of the correlation function is not sufficient to uniquely determine the set of coefficient $x_k(r)$ in the expansion (2.15).

For computational purposes, the stochastic modes are commonly discretized in the spatial domain with some basis function.

This leads equation (3.20) to

$$\chi\left(r,\omega\right) = \sum_{k=0}^{N_P} (x_k.\Phi(r))\psi_k(\xi(\omega)) \tag{2.16}$$

where $\Phi(r) = \left[\Phi_1(r),...,\Phi_{N_s}(r)\right]^T$ denotes the vector of shape functions and $X_k = \left[x_{k1},...,x_{kN_s}\right]^T$ represents corresponding coefficients. Equation (2.15) can be rewritten as:

$$\chi(r,\omega) = \sum_{k=0}^{N_p} (\mathrm{x}_k\,\psi_k(\xi(\omega))).\Phi(r) \tag{2.17}$$

Equation (2.17) shows the relation to the PC expansion of the random vector (2.15). This can be seen as the PC expansion with the method of point discretization. This approach is simple to compute and not restricted to Gaussian random fields. However, the method is only useful for medium to long correlation distances. One may need a fine mesh to fulfill the accuracy requirements in case of a small correlation length. A large number of discretization points N_s requires large computational efforts.

2.2 Non-Intrusive Method

This section introduces the theoretical background of the non-intrusive methods employed for the approximation of an output of a model involving random data, parameterized by a finite number of random variables $\xi(\theta)$ defined on a probability space (Θ, Σ, P), and probability density $p_\xi(\xi)$. We denote Ξ as the support of the density p_ξ [15,25] and s represents the model output, the quantities of interest. The approximation of the mapping is obtained in terms of the basis of a random functional.

$$s : \xi \in \Xi \mapsto s(\xi) = \sum_i s_i \Psi_i(\xi) \tag{2.18}$$

The non-intrusive methods rely on a set of deterministic model resolutions, corresponding to some specific values or realizations of ξ to construct the approximation s **(**ξ**).**

The spectral modes (α_k) of the random variable α^* can be obtained by solving the linear system of equations given above. The mean (μ_{α^*}) and the variance ($\delta^2_{\alpha^*}$) of the solution can be obtained, as shown below.

$$\mu_{\alpha^*} = \bar{\alpha}^*(x) = \int_\Xi \alpha^*(x,\xi)\, p(\xi)\, d\xi = \alpha_0(x) \tag{2.19}$$

$$\begin{aligned} \sigma^2_{\alpha^*} &= \mathrm{var}\left[\alpha^*(x,\xi)\right] = \int_\Xi \left(\alpha^*(x,\xi) - \bar{\alpha}^*_0(x)\right)^2 p(\xi)\, d\xi \\ &= \sum_{i=1}^{P} \left[\alpha_j^2(x) \left\langle \Psi_j^2(\xi) \right\rangle\right]. \end{aligned} \tag{2.20}$$

2.2.1 Non-Intrusive Spectral Projection

We consider a model where both the input and output are real valued scalar random variables on a probability space (Θ, Σ, P).

$$P_\xi(\xi) = \prod_{i=1}^{N} p_i(\xi_i) \tag{2.21}$$

Non-Intrusive Spectral Projection (NISP) [15] computes the projection coefficient of the random model output s (ξ) on a finite dimensional stochastic subspace $\mathcal{S}^p$ of $L_2\left(\Xi, p_\xi\right)$. The space equipped with the inner product $\langle,\rangle$,

$$\langle u,v\rangle = \int_\Xi u(\xi)v(\xi)p_\xi(\xi)d\xi\,, \tag{2.22}$$

is a Hilbert space. We denote (p+1) as the dimension of $\mathcal{S}^p$ and let $\left\{\Psi_0(\xi),...,\Psi_p(\xi)\right\}$ be an orthogonal basis of $\mathcal{S}^p$.

$$\left\langle \Psi_i,\Psi_j\right\rangle = \int...\int \Psi_i(\xi)\Psi_j(\xi)p_\xi(\xi)d\xi_1...d\xi_N = \left\langle \Psi_i,\Psi_j\right\rangle \delta_{ij}\,. \tag{2.23}$$

Non-Intrusive Spectral Projection provides a decomposition of $(s)^{\|}$ into a finite generalized Polynomial Chaos (gPC) basis $\left\{\Psi_0(\xi),...,\Psi_p(\xi)\right\}$. Thus, it is the truncated Polynomial Chaos Expansion (PCE) of the probability space. The spectral projection of the output is shown below.

$$s^{\|}(\xi) = \sum_{k=0}^{P} s_k\Psi_k(\xi) \in \mathcal{S}^p \tag{2.24}$$

Since $s^{\|}(\xi)$ is the projection of the output onto $\mathcal{S}^p$, for each index $k \in \{0,...,n\}$, the coefficients are given by

$$s_k = \frac{\left\langle s,\Psi_k\right\rangle}{\left\langle \Psi_k,\Psi_k\right\rangle}\,,\ k = 0,\ 1,\ \ldots, \tag{2.25}$$

such that the following is satisfied.

$$s - s^{\|} \perp \mathcal{S}^p \Leftrightarrow \left\langle s - s^{\|},\Psi_k\right\rangle = 0 \ \text{ for } \ 0 \le k \le P \tag{2.26}$$

When the generalized Polynomial Chaos (gPC) basis is selected, the values $\left\langle \Psi_k,\Psi_k\right\rangle$ are known analytically and therefore, only the computation of the scalar product remains:

$$\langle \Psi_k, \Psi_k \rangle s_k = \langle s, \Psi_k \rangle = \int_{\Xi} s(\xi_1, ..., \xi_N) \Psi_k p_1(\xi_1) ... p_N(\xi_N) d\xi_1 ... d\xi_N \,. \tag{2.27}$$

Here, $\Xi \subset R$ is the support of the probability density function p_ξ of the basic random variable ξ. The multidimensional integral of a function $f(\xi) \equiv s(\xi)\Psi_k(\xi)$ over Ξ with a non-negative weight p_ξ has a product form, which can be shortened as shown below.

$$\gamma\, f = \int_{\Xi} f(\xi) p_\xi(\xi) d\xi \tag{2.28}$$

When the dimensionality N of the integration is not too large, as is the case for many problems in practical engineering fields, deterministic cubature can be effectively used. A cubature is an approximation of the multidimensional integral $\gamma\, f$ as a discrete sum:

$$\gamma\, f \approx \sum_{i=1}^{N_Q} f(\xi^{(i)}) \mathrm{w}^{(i)} \,, \tag{2.29}$$

where $\xi^{(i)} \in \Xi$ and $w^{(i)} \in R$, i=1, … , N_Q, are the nodes and weights of the chosen quadrature formula. Various quadrature rules can be used by means of a non-negative weight function p_ξ. N_Q is the number of model resolutions.

In further sections, the set $\left\{\xi^{(i)}\right\}_{i=1}^{N+1}$ is called a sampling of ξ due to the very definition of the integral (2.29) and the quadrature node can be understood as particular realizations of a basic random variable ξ.

By combining equations (2.25), (2.27), and (2.29), it becomes clear how NISP relies on a set of deterministic model resolutions, corresponding to some specific realizations of ξ. Along this line, a deterministic simulation code can be used as a black-box, which associates each realization of the parameters to the corresponding model outputs.

2.2.2 Point-Collocation Non-intrusive Polynomial Approach

Unlike the Non-Intrusive Spectral Projection (NISP) method discussed above, this method does not aim to determine the projection of the model output s (ξ) on a

predefined stochastic subspace, but instead relies on the interpolation based on equation (2.24). Then, P+1 vectors $\left(\vec{\xi}=\left\{\xi_1,\xi_2,...,\xi_n\right\}_k, k=0,1,2,...,\mathrm{P}\right)$ are chosen in the random space for a given PC expansion with P+1 modes and the deterministic simulation code is evaluated at this point, which has a specific probability distribution. The discrete sum is taken over the number of output modes [15],

$$P+1=\frac{(n+p)!}{n!\,p!}, \tag{2.30}$$

where p and n are the order of polynomial chaos and the number of the random input dimension, respectively. The PC model obtained via equation (2.30) is formed by polynomials up to a maximum degree P (also called order of the PC model): this truncation strategy is referred to as a "total order expansion". A linear system of equations can be obtained [16-18], [25-27], as shown below.

$$\begin{bmatrix} \Psi_0\left(\vec{\xi_0}\right) & \Psi_1\left(\vec{\xi_0}\right) & \cdots & \Psi_p\left(\vec{\xi_0}\right) \\ \Psi_0\left(\vec{\xi_1}\right) & \Psi_1\left(\vec{\xi_1}\right) & \cdots & \Psi_p\left(\vec{\xi_1}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \Psi_0\left(\vec{\xi_p}\right) & \Psi_1\left(\vec{\xi_p}\right) & \cdots & \Psi_p\left(\vec{\xi_p}\right) \end{bmatrix} \begin{Bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha p \end{Bmatrix} = \begin{Bmatrix} \alpha^*\left(\vec{\xi_0}\right) \\ \alpha^*\left(\vec{\xi_0}\right) \\ \vdots \\ \alpha^*\left(\vec{\xi_0}\right) \end{Bmatrix} \tag{2.31}$$

We solved the linear system of equations (2.31), which requires P+1 deterministic function evaluations. If more than P+1 samples are chosen, then the over-determined system of equations should be solved using the Least Squares method. Hosder *et al*. [16] investigated the effects on the results by the number of collocation points in a systematic way through the introduction of a parameter, the oversampling rate n_p defined below.

$$n_p=\frac{number\ of\ samples}{P+1} \tag{2.32}$$

They used oversampling rates of n_p = 1, 2, 3, and 4 for the solution of a stochastic model problem with multiple uncertain variables to study the effect of the number of collocation points on the given accuracy of the polynomial chaos expansions. They

suggested that a number of collocation points that is corresponding to two times the minimum number required, which means that the oversampling rate n_p is 2, yields a better approximation to the statistic at each polynomial degree.

The open source, Chaospy [28] and Scilab [15] [29-30], was used for UQ analysis in the present work. The Chaospy [28] software package provides various stochastic tools for UQ analysis for the Point-Collocation NIPC method. Scilab software [15] was developed for spectral projection based on the Gaussian quadrature.

2.3 Bayesian Inference and Markov Chain Monte Carlo

2.3.1 Bayesian Theory

Bayesian statistics is a system for describing epistemological uncertainty using the mathematical language of probability. In the 'Bayesian paradigm,' degrees of belief in states of nature are specified; these are non-negative, and the total belief in all states of nature is fixed to be one. Bayesian statistical methods start with existing 'prior' beliefs, and update these using data to give 'posterior' beliefs, which may be used as the basis for inferential decisions. Bayes theorem allows one to formally incorporate prior knowledge into computing statistical probabilities.

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{p(\mathrm{B}|\mathrm{A})\,p(\mathrm{A})}{p(\mathrm{B})} \tag{2.33}$$

where $p(\cdot)$ denotes a probability distribution, and $p(\cdot|\cdot)$ a conditional distribution. When B represents data and A represents parameters in a statistical model, Bayes Theorem provides the basis for Bayesian inference. The 'prior' distribution $p(\mathrm{A})$ (epistemological uncertainty) is combined with 'likelihood' $p(\mathrm{B}|\mathrm{A})$ to provide a 'posterior' distribution $p(\mathrm{A}|\mathrm{B})$ (updated epistemological uncertainty): the likelihood is derived from an aleatory sampling model $p(\mathrm{B}|\mathrm{A})$ but considered as function of A for fixed B .

While an innocuous theory, practical use of the Bayesian approach requires consideration of complex practical issues, including the source of the prior distribution,

the choice of a likelihood function, computation and summary of the posterior distribution in high-dimensional problems, and making a convincing presentation of the analysis.

Bayes theorem can be thought of as a way of coherently updating our uncertainty in the light of new evidence. The use of a probability distribution as a 'language' to express our uncertainty is not an arbitrary choice: it can in fact be determined from deeper principles of logical reasoning or rational behavior; see Jaynes (2003) or Lindley (1953). In particular, De Finetti (1937) showed that making a qualitative assumptions of exchangeability of binary observations (i.e. that their joint distribution is unaffected by label-permutation) is equivalent to assuming they are each independent conditional on some unknown parameter θ , where θ has a prior distribution and is the limiting frequency with which the events occur.

2.3.2 Bayesian Inference

The basis for Bayesian inference is derived from Bayes' theorem. Here is Bayes' theorem, equation (2.16)

Replacing B with observations y, A with parameter set Θ, and probabilities Pr with densities p (or sometimes π or function f), results in the following

$$p(\Theta \mid y) = \frac{p(y|\Theta)\, p(\Theta)}{p(y)} \tag{2.34}$$

1. $p(\Theta)$ is the set of prior distributions for parameter set Θ, and uses probability as a means of quantifying uncertainty about Θ before taking the data into account.

2. $p(y|\Theta)$ is the likelihood or likelihood function, in which all variables are related in a full probability model.

3. $p(\Theta|y)$ is the joint posterior distribution that expresses uncertainty about parameter set Θ after taking both the prior and the data into account.

If parameter set Θ is partitioned into a single parameter of interest φ and the remaining parameters are considered nuisance parameters, then p(φ|y) is the marginal posterior distribution.

where $p(y)$ will be discussed below, $p(\Theta)$ is the set of prior distributions of parameter set Θ before y is observed, $p(y|\Theta)$ is the likelihood of y under a model, and $p(\Theta|y)$ is the joint posterior distribution, sometimes called the full posterior distribution, of parameter set Θ that expresses uncertainty about parameter set Θ after taking both the prior and data into account. Since there are usually multiple parameters, Θ represents a set of j parameters, and may be considered hereafter in this article as

$$\Theta = \theta_1, \ldots, \theta_j \tag{2.35}$$

The denominator

$$p(y) = \int p(y|\Theta)\, p(\Theta) d\Theta \tag{2.36}$$

Defines the "marginal likelihood" of y, or the "prior predictive distribution" of y, and may be set to an unknown constant c. The prior predictive distributions indicates what y should look like, given the model, before y has been observed. Only the set of prior probabilities and the model's likelihood function are used for the marginal likelihood of y. The presence of the marginal likelihood of y normalizes the joint posterior distribution, $p(\Theta|y)$, ensuring it is a proper distribution and integrates to one. By replacing p(y) with c, which is short for a 'constant of proportionality', the model-based formulation of Bayes' theorem becomes

$$p(\Theta \mid y) = \frac{p(y|\Theta)\, p(\Theta)}{c} \tag{2.37}$$

By removing c from the equation, the relationship changes from 'equals' (=) to 'proportional to' (∝)

$$p(\Theta \mid y) \propto p(y \mid \Theta)\, p(\Theta) \tag{2.38}$$

This form can be stated as the unnormalized joint posterior being proportional to the likelihood times the prior. However, the goal in model-based Bayesian inference is usually not to summarize the unnormalized joint posterior distribution, but to summarize the marginal distributions of the parameters. The full parameter set Θ can typically be partitioned into

$$\Theta = \{\Phi, \Lambda\} \tag{2.39}$$

where Φ is the sub-vector of interest, and Λ is the complementary sub-vector of Θ, often referred to as a vector of nuisance parameters. In a Bayesian framework, the presence of nuisance parameters does not pose any formal, theoretical problems. A nuisance parameter is a parameter that exists in the joint posterior distribution of a model, though it is not a parameter of interest. The marginal posterior distribution of φ, the parameter of interest, can simply be written as

$$p(\phi \mid y) = \int p(\phi, \Lambda | y) d\Lambda \tag{2.40}$$

In model-based Bayesian inference, Bayes' theorem is used to estimate the non-normalized joint posterior distribution, and finally the user can assess and make inferences from the marginal posterior distributions.

2.3.3 Prior Distributions

In Bayesian inference, a prior probability distribution, often called simply the prior, of an uncertain parameter θ or latent variable is a probability distribution that expresses uncertainty about θ before the data are considered. The parameters of a prior distribution are called hyper parameters, to distinguish them from the parameters (Θ) of the model. When applying Bayes' theorem, the prior is multiplied by the likelihood function and then normalized to estimate the posterior probability distribution, which is the conditional distribution of Θ given the data. Moreover, the prior distribution affects the posterior distribution. Prior probability distributions have traditionally belonged to one of two categories: informative priors and uninformative priors. Here, four categories of priors are presented according to information8 and the goal in the use of the prior. The

four categories are informative, weakly informative, least informative, and uninformative.

When prior information is available about θ, it should be included in the prior distribution of θ. For example, if the present model form is similar to a previous model form, and the present model is intended to be an updated version based on more current data, then the posterior distribution of θ from the previous model may be used as the prior distribution of θ for the present model. In this way, each version of a model is not starting from scratch, based only on the present data, but the cumulative effects of all data, past and present, can be taken into account. To ensure the current data do not overwhelm the prior, Ibrahim and Chen (2000) introduced the power prior. The power prior is a class of informative prior distribution that takes previous data and results into account. If the present data is very similar to the previous data, then the precision of the posterior distribution increases when including more and more information from previous models. If the present data differs considerably, then the posterior distribution of θ may be in the tails of the prior distribution for θ, so the prior distribution contributes less density in its tails. Hierarchical Bayes is also a popular way to combine data sets. Sometimes informative prior information is not simply ready to be used, such as when it resides in another person, as in an expert. In this case, their personal beliefs about the probability of the event must be elicited into the form of a proper probability density function. This process is called prior elicitation.

2.2.4 Likelihood

In order to complete the definition of a Bayesian model, both the prior distributions and the likelihood must be approximated or fully specified. The likelihood, likelihood function, or p(y|Θ), contains the available information provided by the sample. The likelihood is

$$p(y \mid \Theta) = \prod_{i=1}^{n} p(y_i|\Theta) \tag{2.41}$$

The data y affects the posterior distribution $p(\Theta|y)$ only through the likelihood function $p(y|\Theta)$. In this way, Bayesian inference obeys the likelihood principle, which states that for a given sample of data, any two probability models $p(y|\Theta)$ that have the same likelihood function yield the same inference for Θ. For more information on the likelihood principle, see this section.

In non-technical parlance, "likelihood"is usually a synonym for "probability", but in statistical usage there is a clear distinction: whereas "probability" allows us to predict unknown outcomes based on known parameters, "likelihood" allows us to estimate unknown parameters based on known outcomes.

In a sense, likelihood can be thought a reversed version of conditional probability. Reasoning forward from a given parameter θ, the conditional probability of y is the density p(y|θ). With θ as a parameter, here are relationships in expressions of the likelihood function

$$L(\theta \mid y) = p(y \mid \theta) = f(y \mid \theta) \tag{2.42}$$

where y is the observed outcome of an experiment, and the likelihood (L) of θ given y is equal to the density p(y|θ) or function f(y|θ). When viewed as a function of y with θ fixed, it is not a likelihood function L(θ|y), but merely a probability density function p(y|θ). When viewed as a function of θ with y fixed, it is a likelihood function and may be denoted as L (θ|y), p(y|θ), or f(y|θ) .

For example, in a Bayesian linear regression with an intercept and two independent variables, the model may be specified as

$$y_i \sim N(\mu_i, \sigma^2) \tag{2.43}$$

$$\mu_i = \beta_1 + \beta_2 X_{i,1} + \beta_3 X_{i,2} \tag{2.44}$$

The dependent variable y, indexed by i = 1, …, n, is stochastic, and normally-distributed according to the expectation vector μ, and variance σ^2 . Expectation vector μ

is an additive, linear function of a vector of regression parameters, β, and the design matrix X. Since y is normally-distributed, the probability density function (PDF) of a normal distribution will be used, and is usually denoted as

$$f(y) = \frac{1}{\sqrt{2\pi\sigma}} \exp[(-\frac{1}{2}\sigma^2)(y_i - \mu_i)^2]; \quad y \in (-\infty, \infty) \tag{2.45}$$

By considering a conditional distribution, the record-level likelihood in Bayesian notation is

$$p(y_i \mid \Theta) = \frac{1}{\sqrt{2\pi\sigma}} \exp[(-\frac{1}{2}\sigma^2)(y_i - \mu_i)^2]; \quad y \in (-\infty, \infty) \tag{2.46}$$

2.2.5 Markov Chain Monte Carlo (MCMC)

Markov chain Monte Carlo (MCMC) was invented soon after ordinary Monte Carlo at Los Alamos, one of the few places where computers were available at the time. Metropolis et al. (1953) simulated a liquid in equilibrium with its gas phase. The obvious way to find out about the thermodynamic equilibrium is to simulate the dynamics of the system, and let it run until it reaches equilibrium. The tour de force was their realization that they did not need to simulate the exact dynamics; they only needed to simulate some Markov chain having the same equilibrium distribution. Simulations following the scheme of Metropolis et al. (1953) are said to use the Metropolis algorithm. As computers became more widely available, the Metropolis algorithm was widely used by chemists and physicists, but it did not become widely known among statisticians until after 1990. Hastings (1970) generalized the Metropolis algorithm, and simulations following his scheme are said to use the Metropolis–Hastings algorithm. A special case of the Metropolis–Hastings algorithm was introduced by Geman and Geman (1984), apparently without knowledge of earlier work. Simulations following their scheme are said to use the Gibbs sampler. Much of Geman and Geman (1984) discusses optimization to find the posterior mode rather than simulation, and it took some time for it to be understood in the spatial statistics community that the Gibbs sampler simulated the posterior distribution, thus enabling full Bayesian inference of all

kinds. A methodology that was later seen to be very similar to the Gibbs sampler was introduced by Tanner and Wong (1987), again apparently without knowledge of earlier work. To this day, some refer to the Gibbs sampler as "data augmentation" following these authors. Gelfand and Smith (1990) made the wider Bayesian community aware of the Gibbs sampler, which up to that time had been known only in the spatial statistics community. Then it took off; as of this writing, a search for Gelfand and Smith (1990) on Google Scholar yields 4003 links to other works. It was rapidly realized that most Bayesian inference could be done by MCMC, whereas very little could be done without MCMC. It took a while for researchers to properly understand the theory of MCMC (Geyer, 1992; Tierney, 1994) and that all of the aforementioned work was a special case of the notion of MCMC. Green (1995) generalized the Metropolis–Hastings algorithm, as much as it can be generalized. Although this terminology is not widely used, we say that simulations following his scheme use the Metropolis–Hastings–Green algorithm. MCMC is not used only for Bayesian inference. Likelihood inference in cases where the likelihood cannot be calculated explicitly due to missing data or complex dependence can also use MCMC (Geyer, 1994, 1999; Geyer and Thompson, 1992, 1995, and references cited therein).

A sequence X_1, X_2, … of random elements of some set is a Markov chain if the conditional distribution of X_{n+1} given X_1, … , X_n depends on X_n only. The set in which the Xi take values is called the state space of the Markov chain.

A Markov chain has stationary transition probabilities if the conditional distribution of X_{n+1} given X_n does not depend on n. This is the main kind of Markov chain of interest in MCMC. Some kinds of adaptive MCMC have nonstationary transition probabilities. In this chapter we always assume stationary transition probabilities. The joint distribution of a Markov chain is determined by

- The marginal distribution of X_1, called the initial distribution
- The conditional distribution of X_{n+1} given X_n, called the transition probability distribution (because of the assumption of stationary transition probabilities, this does not depend on n).

People introduced to Markov chains through a typical course on stochastic processes have usually only seen examples where the state space is finite or countable. If the state space is finite, written $\{x_1, \ldots, x_n\}$, then the initial distribution can be associated with a vector $\lambda = (\lambda_1, \ldots, \lambda_n)$ defined by

$$P(X_1 = x_i) = \lambda_i, \qquad i = 1, \ldots, n, \tag{2.47}$$

and the transition probabilities can be associated with a matrix P having elements p_{ij} defined by

$$P(X_{n+1} = x_j \mid X_n = x_i) = p_{ij}, \qquad i = 1, \ldots, n \text{ and } j = 1, \ldots, n. \tag{2.48}$$

When the state space is countably infinite, we can think of an infinite vector and matrix. But most Markov chains of interest in MCMC have uncountable state space, and then we cannot think of the initial distribution as a vector or the transition probability distribution as a matrix. We must think of them as an unconditional probability distribution and a conditional probability distribution.

Chapter 3. Parametric Uncertainty of $\gamma - Re_\theta$ Transition Model

The first step in the application of the Bayesian inference technique to the calibration of uncertain transition model parameters, is to define an initial uncertainty interval for these parameters. The chapter starts by explaining how the transition equation are confronted with a so-called closure problem, resulting in the necessity of model containing closure coefficients which's values are uncertain. This followed by an extensive discussion on the determination of prior uncertainty interval for each of the considered model's uncertain parameters.

3.1 Introduction

This work describes the implementation of a newly published turbulent model into OpenFoam. The main advantage of this model is its capability to effectively take into account the important phenomenon "Laminar-Turbulent" transition in the boundary layer[1]. From physical point of view, initially a laminar boundary layer development which due to increased instabilities of the flow, loses its stability and the transition to a turbulent boundary layer occurs. Ordinary RANS modeling cannot capture this phenomenon and they usually result in completely turbulent boundary layer. In this model, two extra transport equations are solved and coupled to the $k - \omega SST$ model and control the production term of turbulent kinetic energy in the K equation. In the literature, this new model is referred as $\gamma - Re_\theta$ model. In order to demonstrate the behavior of this model a flat plate test case is simulated by the $\gamma - Re_\theta$ model base on k − ωSST model.

3.2 Description of Transition Model

In the section, the open source code OpenFOAM V.5.0 [31] was adopted for the deterministic flow solver. The steady incompressible Navier-Stokes governing equations were discretized spatially at the second order and the SIMPLE algorithm was adopted for continuity.

Menter proposed the $\gamma - \mathrm{Re}_\theta$ transitional model [21] based on the fully turbulent SST k-**ω** model. It includes two transport equations for intermittency,$\boldsymbol{\gamma}$, and for the transported transition momentum thickness Reynolds number $\tilde{\mathrm{Re}}_{\theta t}$. The dimensionless intermittency and Reynolds number based on the momentum thickness are formulated as follows.

$$\rho\frac{D\gamma}{Dt}=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_t}\right)\frac{\partial\gamma}{\partial x_j}\right]+F_{length}c_{a1}\rho S\left(\gamma F_{onset}\right)^{0.5}(1-c_{e1}\gamma)+c_{a2}\rho\Omega\gamma F_{turb}(1-c_{e2}\gamma) \quad (3.1)$$

$$\rho\frac{D\tilde{\mathrm{R}}e_{\theta t}}{Dt}=\frac{\partial}{\partial x_j}\left[\sigma_{\theta t}(\mu+\mu_t)\frac{\partial\tilde{\mathrm{Re}}_{\theta t}}{\partial x_j}\right]+c_{\theta t}\frac{(\rho U)^2}{500\mu}(\mathrm{Re}_{\theta t}-\tilde{\mathrm{R}}e_{\theta t})(1-F_{\theta t}) \quad (3.2)$$

The modeled transport equations are controlled by the following functions:

$$F_{onset}=\max(F_{onset2}-F_{onset3},0),F_{turb}=\exp[-(R_T/4)^4], \quad (3.3)$$

$$F_{onset2}=\min[\max(F_{onset1},F_{onset1}^4),2],\mathrm{F}_{onset1}=\frac{\mathrm{Re}_v}{2.193\mathrm{Re}_{\theta c}}, \quad (3.4)$$

$$F_{onset3}=\max(1-(R_T/2.5)^3,0), \quad (3.5)$$

$$F_{\theta t}=\min\left\{\max\left[F_{wake}.\exp\left[-\left(\frac{U^2}{375\Omega v\tilde{R}e_{\theta t}}\right)^4\right],1-\left(\frac{c_{e2}\gamma-1}{c_{e2}-1}\right)^2\right],1\right\},F_{wake}=\exp\left[-\left(\frac{\mathrm{Re}_\omega}{10^5}\right)^2\right], \quad (3.6)$$

with the following parameters:

$$\mathrm{Re}_v=\frac{\rho Sy^2}{\mu},\mathrm{R}_T=\frac{\rho k}{\mu\omega},\mathrm{Re}_\omega=\frac{\rho\omega y^2}{\mu},\ \mathrm{Re}_v=\frac{\rho Sy^2}{\mu},\mathrm{R}_T=\frac{\rho k}{\mu\omega},\mathrm{Re}_\omega=\frac{\rho\omega y^2}{\mu} \quad (3.7)$$

The details of the intermittency equation are presented in a previous work [21]. Both $\mathrm{F}_{\mathrm{length}}$ and $Re_{\theta c}$ in the above equation are empirical correlations and functions of the transition momentum-thickness Reynolds number $Re_{\theta t}$, which is also a function of the turbulence intensity Tu and the streamwise pressure gradient λ_θ in the free flow. The

model constants are c_{a1}=2.0, c_{e1}=1.0, c_{a2}=0.06, c_{e2}=50.0, c_{α}=0.5, σ_{γ}=1.0, $\sigma_{\theta t}$=2.0 and $c_{\theta t}$=0.03.

For predicting separation induced transition, the following modification is utilized.

$$\gamma_{sep} = \min(2.0\max(0, \frac{\mathrm{Re}_v}{3.235R_{\theta c}} - 1)F_{reattach}, 2)\,\mathrm{F}_{\theta t} \tag{3.8}$$

$$F_{reattach} = e^{-\left(\frac{R_T}{20}\right)^4} \tag{3.9}$$

The effective value of γ is thus obtained as follows.

$$\gamma_{effective} = \max(\gamma, \gamma_{sep}) \tag{3.10}$$

The empirical correlation used in this model is based on the pressure gradient parameter and turbulent intensity defined as follows.

$$\lambda_0 = \frac{\theta^2}{v}\cdot\frac{dU}{ds}, \mathrm{K} = \frac{v}{U^2}\cdot\frac{dU}{ds}, Tu = \frac{100(2k/3)^{1/2}}{U}, \frac{dU}{ds} = \frac{\partial u_j}{\partial x_j}\frac{u_i u_j}{U^2} \tag{3.11}$$

Re_{θ} is defined as follows.

$$\mathrm{Re}_{\theta t} = \frac{\rho\theta_t U_0}{\mu} \tag{3.12}$$

The following correlation equations were defined by Langtry [21].

$$\mathrm{Re}_{\theta t} = \begin{cases} \left[1173.51 - 589.428Tu + \frac{0.2196}{Tu^2}\right]F(\lambda, Tu); & Tu \le 1.3 \\ 331.5[Tu - 0.5658]^{-0.671}F(\lambda, Tu); & Tu > 1.3 \end{cases} \tag{3.13}$$

$$F(\lambda, Tu) = \begin{cases} 1 + e^{-(2Tu/3)^{1.5}}\cdot(12.986\lambda + 123.66\lambda^2 + 405.689\lambda^3); & \lambda \le 0 \\ 1 + 0.275e^{-2Tu}\cdot(1 - e^{(-35\lambda)}); & \lambda > 0 \end{cases} \tag{3.14}$$

Chapter 4. Deterministic Simulations for γ-Re_θ Transitional Model

The Aerospatiale airfoil and the flat plate are considered for the simulation of transitional flows and QoIs are drag and lift coefficient over two geometry. First, the prior uncertainty intervals for the closure coefficients are studied with the uniform distributions of multiple random inputs and then most effective three model coefficients are selected based on the Sobol index. Point-Collocation Non-Intrusive Chaos (NIPC) method is adopted using gPC(generalized Polynomial Chaos) and the polynomial order is from 2^{nd} to 5^{th}.

4.1 Introduction

In the $\gamma - Re_\theta$ transition model, two transport equations with an extra nine model coefficients should be solved additionally based on the SST k-ω turbulent model. If a total of nine model coefficients are considered for multiple random inputs, tremendous computational time and cost are required. For example, when the order of the polynomial chaos p is 2 or 3, the required number of random inputs is 19,683 or 262,144 in the Non-Intrusive Spectral Projection (NISP) method. First, the sensitivity analysis of each model coefficient was conducted with $\pm 10\%$ deviation for the flat plate and then three most effective model coefficients (Ce1, Ca2, and Ce2) were able to be obtained for the multiple random inputs. The deviation and distribution of the model coefficients are set based on the references [16-18]. Table 1 presents the results of the sensitivity analysis which show the difference with the reference value of the deterministic solution. To carry this out, each model coefficient was assumed to have a uniform distribution with the fixed constant from the original model [21] and $\pm 10\%$ deviation. Then, we applied the UQ technique with the single random input to the target problem. We calculated the variance depending on each random model coefficient and determined the model coefficients that show the largest variance of the quantities of interest. Considering the simulation time and cost, three top model coefficients Ca_2, Ce_1, and Ce_2 are selected and these were applied to the NIPC and NISP methods with multiple random inputs. The following table shows the probability distributions of the three coefficients.

Table 4.1 *The sensitivity analysis of each model coefficient*

Coefficient	CD of lower boundary (-10%)	CD of upper boundary (10%)	Difference percent (%)
C_{a2}	0.00391	0.00410	4.60
C_{e1}	0.00373	0.00421	11.4
C_{e2}	0.00397	0.00406	2.22
Ca_1	0.00395	0.00402	1.74
C_{α}	0.00401	0.00405	0.99
δ_{γ}	0.00399	0.00401	0.49
$\sigma_{\theta t}$	0.00389	0.00392	0.76
$C_{\theta t}$	0.00402	0.00398	1.0

Table 4.2 *The uniform distributions of the three model coefficients ± 10% deviation*

Coefficient	Original Value	Lower Boundary (-10%)	Upper Boundary (10%)
C_{a2}	0.06	0.054	0.066
C_{e1}	1.0	0.9	1.1
C_{e2}	50.0	45.0	55.0

From the probability distributions of the three coefficients shown in Table 4.2, in our calculations, we used multi-dimensional Legendre polynomials which are orthogonal in the interval for each random dimension. Stochastic solutions to the target problems were obtained using three approaches: Monte Carlo with 500 samples, NIPC, and NISP methods. For NIPC, the oversampling rate n_p was set to 2, as recommended by Hosder *et al*. [16]. The Gaussian quadrature rule was adopted to select the sample points by the specified order, p+1 for NISP.

Table 4.3 *The total numbers of selected points for Monte Carlo, NIPC, and NISP for three chosen points (n is the dimension and p is the order of stochastic computation)*

n=3	Monte Carlo	Point-Collocation	Spectral Project
	MC	NIPC	NISP
Order		$2\frac{(n+p)!}{n!\,p!}$	$(p+1)^n$
2		20	27
3	500	40	64
4		70	125
5		112	216

The $\gamma - Re_\theta$ transition model introduced by Menter [21] was applied to two target problems: transitional flow over a flat plate and flow around Aerospatiale A-airfoil[32-34]. These are well-known validation problems for validation of CFD algorithms and transition/turbulence models. Experimental data are presented through the site of *turbulence modeling resources* by NASA [34]. The deterministic solutions of two transitional flows were validated by comparison with reference data.

4.2 Transitional Flow Over a Flat Plate

The first case is the transitional flow over a flat plate and the deterministic solution of this case was validated using the reference data of Langtry and Menter [21]. Figure 4.2 shows the computational mesh for the flat plate and the boundary conditions of the computational domain over the flat plate. The freestream turbulence intensity and eddy viscosity ratio values were set to and (Savill 1996) [33], respectively. A mesh with 43,000 cells was adopted and the mesh was clustered to the wall for in Table 4.6. The corresponding Reynolds number based on the free stream velocity (5.4 m/s) and length of the flat plate (0.3 m) is 1.0×10^6. Figure 4.3 shows the comparison of the skin friction profiles predicted by the deterministic solution with the $\gamma - Re_\theta$ transition model and experimental data [33-34]. The skin friction coefficient determined by the present deterministic solution is able to predict the transition point and developing turbulent

flow well. The slight early recovery from laminar to turbulent flow and underestimation of the skin friction right after transition agree well with previous simulation results [21].

Table 4.4 *Information regarding the inflow conditions and grid system for flow over a flat plate*

Inflow Condition	U_∞	5.4 m/s
	T_u	3.3%
	μ_t / μ	10.0
Grid System	(Lx,Ly)	(3.2c,1.0c)
	(Nx,Ny)	(230,100)
	Re_L	1e+6 (L=0.3 m)

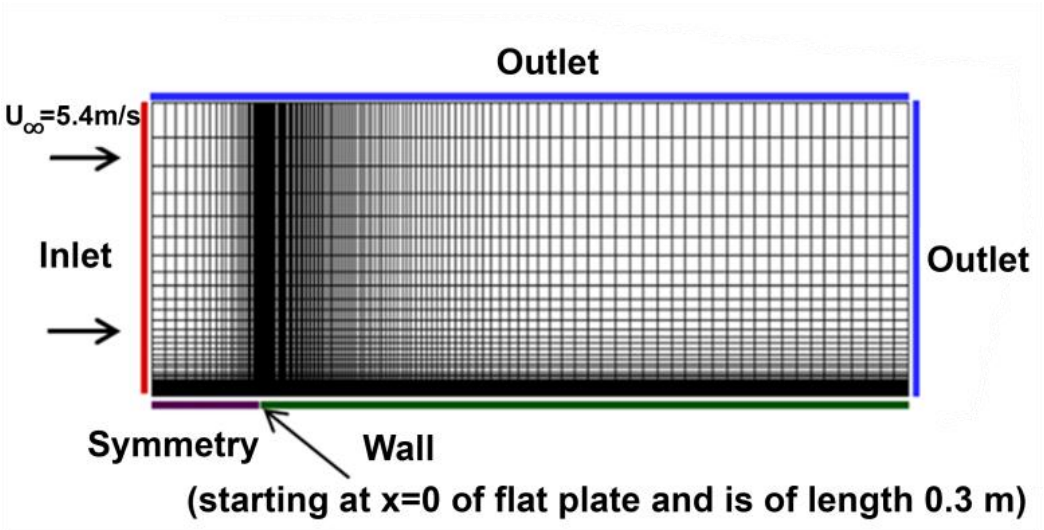


Fig. 4.1 *Computational mesh for the flat plate*

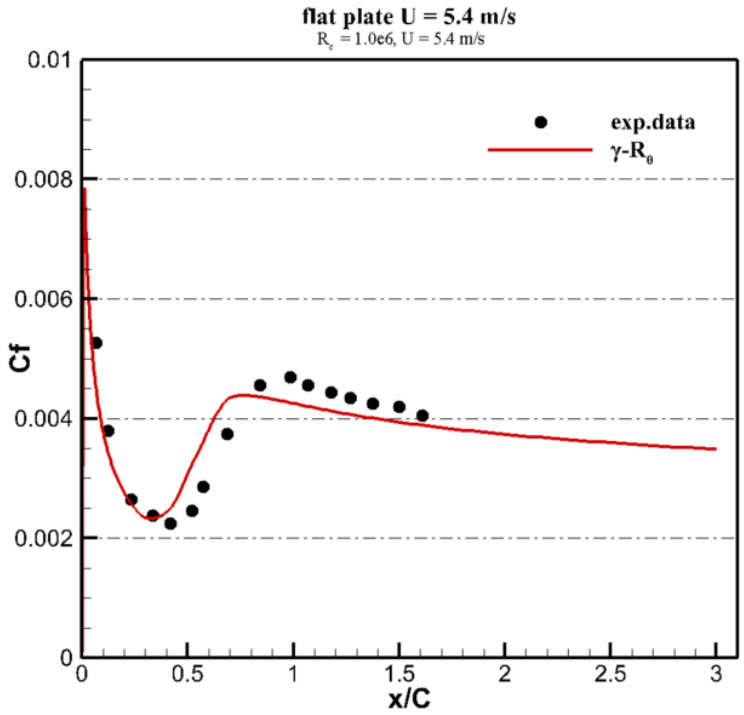


Fig. 4.2 *Skin friction coefficient distribution with the original formulation over the flat plate*

In this study, the drag coefficient is considered as the quantity of interest of UQ. Before discussing the stochastic results of UQ, the predicted drag coefficient obtained by the deterministic solution with the originally fixed model coefficients is 0.00401.

Table 4.5 shows the drag coefficients predicted by the uncertainty quantification technique with the assumptions of uniform distribution of three model coefficients, Ca_2, Ce_1, and Ce_2, and three methods: two methods are Non-Intrusive Polynomial Chaos methods (Point-Collocation NIPC and Non-Intrusive Spectral Projection) and a Monte Carlo method with 500 simulations. The mean value and standard deviation (STD) of the drag coefficient of all three methods are presented with respect to the polynomial order.

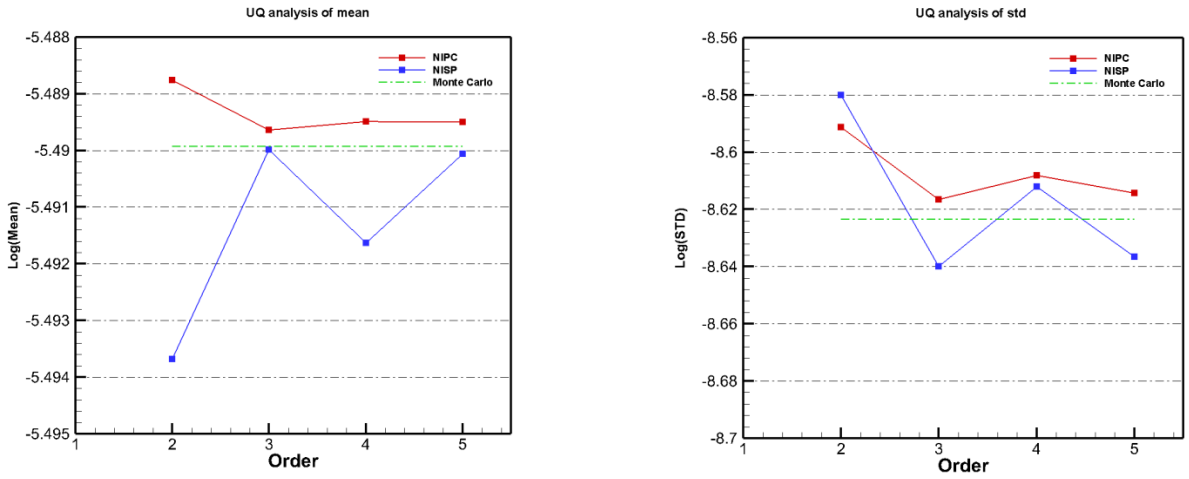
Table 4.5 *The mean and standard deviation of the drag coefficient values for flat plate*

Order	Mean_NIPC	STD_NIPC	Mean_NISP	STD_NISP	Mean_MC	STD_MC
2	0.00413298	0.0001857	0.00411269	0.0001878		
3	0.00412933	0.0001811	0.00412791	0.0001769	0.004128139	0.00017982
4	0.00412997	0.0001826	0.00412113	0.0001819		
5	0.00412992	0.0001815	0.00412758	0.0001775		

The mean drag coefficients approaches approximately 0.00413 which show 2.9% difference compared with the deterministic solution, 0.00401. The standard deviation, which is not able to be predicted by the deterministic simulation, is approximately 0.00018, which is 4.4% of the mean value.

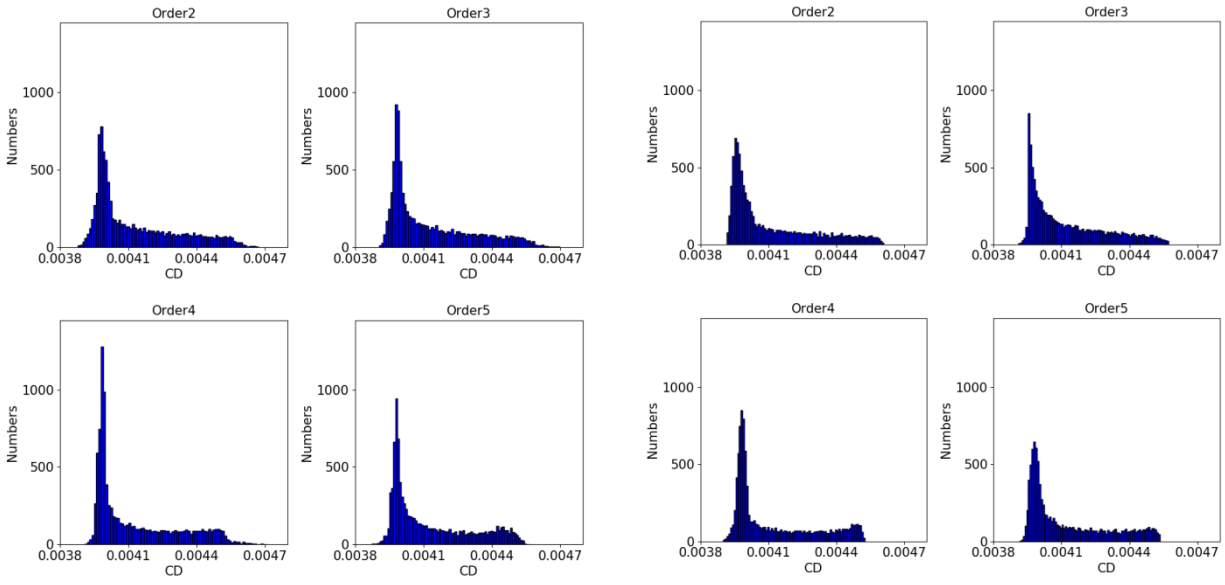
The percent errors of the mean and STD values are, respectively, 0.043% and 0.9% for Point-Collocation NIPC and 0.014% and 1.29% for NISP when the two methods with 5^{th} order are compared to those in the Monte Carlo method. Figure 4.3 shows the convergence of the mean and STD of the drag coefficient according to the polynomial order of the two methods. The long dashed line is the reference value, which corresponds to the Monte Carlo results. The results of Point-Collocation NIPC show faster convergence of the mean and STD than NISP and consistent values except for the case with 2^{nd} order. However, in NISP, the results of the odd number orders (3^{th} and 5^{th} orders) are close to the Monte Carlo results and the results of the even number orders (2^{nd} and 4^{th} orders) have some discrepancies compared to the reference values.

Figure 4.5 shows the detailed probability density function of the drag coefficient which is generated from the calculated polynomial coefficients based on 10,000 sampling points by the Latin-Hypercube sampling method. In all cases, the peaks of the probability density function are biased toward lower values with the given interval but the value of the peak shows small differences depending on the polynomial order. As the order increases, the shape of this function has a similar pattern but there is a small difference towards higher drag coefficient values.



(**a**) (**b**)

Fig. 4.3 *The mean and standard deviations of the drag coefficient of the three methods for orders 2 to 5: (**a**) mean value and (**b**) standard deviation*



(**a**) (**b**)

Fig. 4.4 *PDF analysis of the drag coefficient: (**a**) Point-Collocation Non-Intrusive Polynomial*

4.3 Transitional Flow Over a A_Airfoil

In this chapter, the transitional flow around the Aerospatiale A-airfoil is considered under the conditions of the maximum lift angle of attack ($\alpha = 13.1^{\circ}$) and $\mathrm{Re}_c = 2.07 \times 10^6$ based on the freestream velocity and chord length. The information of the geometry and the experimental results for the Aerospatiale-A airfoil is available from the Office National d'Etudes et Recherches Ae1rospatiales F1 (ONERA F1) wind tunnel (Chaput 1997) [32]. The inflow conditions and grid system are summarized in Table 4.6. The turbulent intensity and eddy viscosity ratio were set to $T_u = 0.2\%$ and $\mu_t / \mu = 10$ [20], respectively. The length of the upstream part and downstream part are 20c and 25c and the height of the computational domain is 30c. At the upstream and side parts, the velocity inlet boundary condition is given and at the downstream, the Neumann boundary condition is set for the outlet. To satisfy $y^+ \le 1$ for correct resolving of the turbulent boundary layer, the distance of the first mesh off the wall is 10^{-5} c and the growth ratio is 1.2.

Figure 4.6 shows the skin friction coefficient distributions over the upper wall of the airfoil predicted by the deterministic solver with the $\gamma - \mathrm{Re}_{\theta}$ transition model as well as the comparison with experimental data [32]. The transition point is predicted to be between 10% and 15% of the chord length, which agrees well with other simulation results [35]. The behavior with developing turbulent flow after transition shows consistent results with experimental data.

Further, the coefficients of the drag and lift forces were calculated and compared to experimental data, F1 and F2 [32]. Unfortunately, there are large discrepancies between the experimental data, F1 and F2, especially for the drag coefficient. When the measurement sensitivity of this transitional flow is considered, the present simulation shows reasonable results like the previous simulation results [35-36]. It was confirmed that the present simulation adopts the adequate methodologies and grid system.

Table 4.6 *Information about the inflow conditions and grid system for flow over Aerospatiale A-airfoil*

Inflow Condition	U_∞	32.8 m/s
	T_u	0.2%
	μ_t / μ	10.0
	α	13.1°
Grid System	(Lx,Ly)	(45c,30c)
	(Nx,Ny)	(350,120)
	Re_c	2.07×10^6

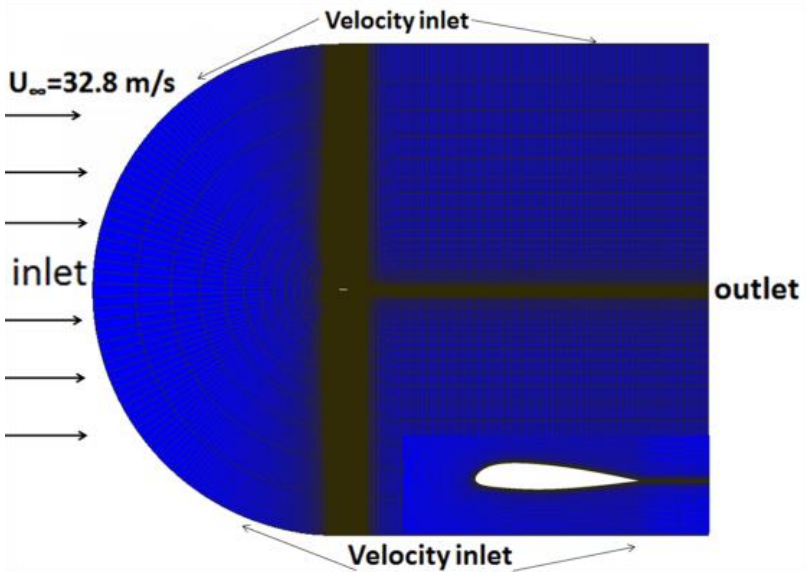


Fig. 4.5 *Computational mesh for Aerospatiale A-airfoil*

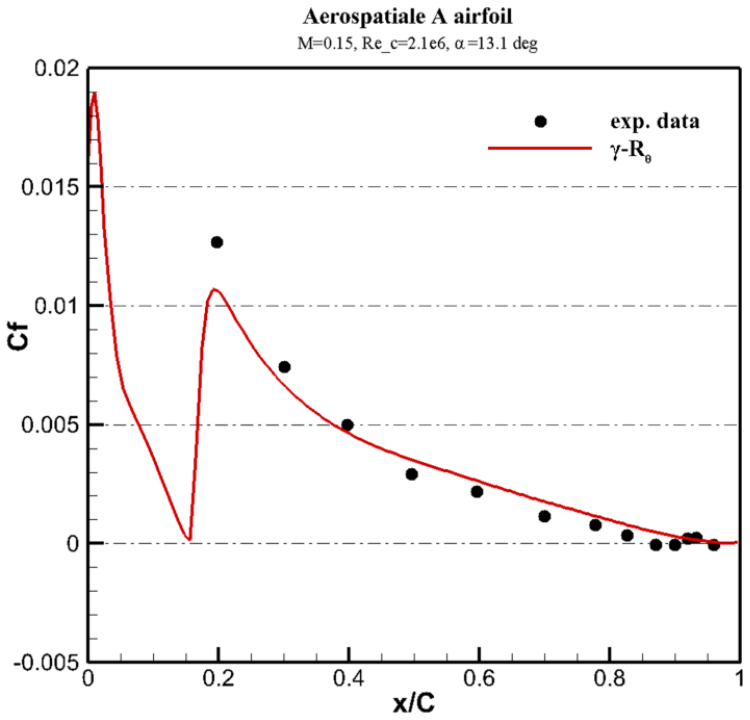


Fig. 4.6 *Skin friction coefficient distribution for the original formulation over Aerospatiale A-airfoil*

Table 4.7 *Lift and drag at*

Model	C_L	C_D
$\boldsymbol{\gamma - R_\theta}$	1.390	0.0206
F1 exp	1.562	0.0208
F2 exp	1.515	0.0308

The mean value and STD of the drag and lift coefficients of the Aerospatiale A-Airfoil were calculated using the Point-Collocation NIPC, NISP, and Monte Carlo methods. The number of samples of Monte Carlo was 500, which is the same as in the previous case. Table 8 shows the drag coefficient with respect to the polynomial order.

Table 4.8 *The mean and standard deviation of the drag coefficient (C_D) values for A-airfoil*

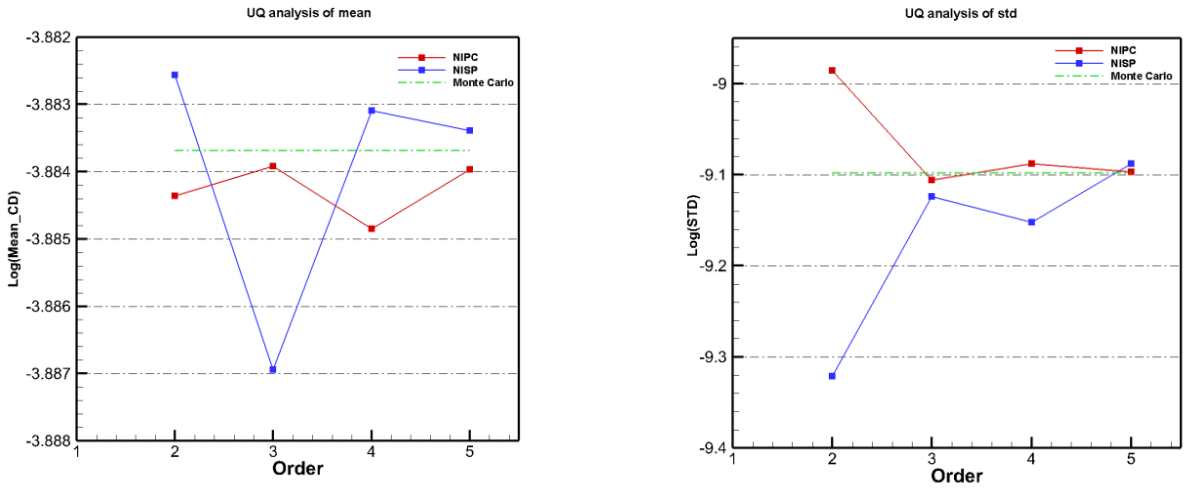
Order	Mean_NIPC	STD_NIPC	Mean_NISP	STD_NISP	Mean_MC	STD_MC
2	0.020561	0.0001252	0.020598	0.0000895	0.02057489	0.0001119
3	0.020570	0.000111	0.020508	0.0001090		
4	0.020551	0.000113	0.020587	0.0001060		
5	0.020569	0.000112	0.020581	0.0001130		

When the polynomial order is 5^{th}, the mean drag coefficient is predicted to be 0.020569 by PC-NIPC and 0.020581 by NISP. When these are compared with the deterministic solution of the original model coefficients, it can be shown that the drag coefficient is not biased as in the flat plate case.

The percent errors of the mean and STD values are 0.029% and 0.09% for Point-Collocation NIPC and 0.03% and 0.97% for NISP when the two methods with 5^{th} order are compared to the Monte Carlo method. When the convergence of the drag coefficient with respect to the polynomial order is checked, the overall pattern is similar to the previous case for the drag coefficient of the flat plate flow. The mean value of the drag coefficient already converged within 0.5% ($\Delta C_D \sim 0.0001$) at the 2^{nd} order, whereas the STD converged after the 2^{nd} polynomial order. As the polynomial order increases, the mean and STD approach the values obtained by the Monte Carlo method (long dashed line).

The distributions of the probability density functions of the drag coefficients are plotted with variation of the order in Figure 4.8. In this case, the distributions of the PDF reveal different patterns depending the adopted method and the order of the polynomial. In the case of Point-Collocation NIPC, the distribution of the case with 2^{nd} order is largely different than other order cases and has similar patterns starting from the 3^{rd} order. NISP has a similar trend as Point-Collocation and the results are similar starting

from the 4th-order case. Interestingly, there are two peaks in the distribution of the PDF in high-order cases for the two methods. The transitional flow simulation of Aerospatiale A-Airfoil is more sensitive to used grid, convergence criteria and the adopted scheme than the flat plate simulation. This means that the quantities of interest such as drag and lift coefficients are sensitive to the computational conditions (grid, convergence criteria and so on) and sampling points which are selected at the given order of polynomial. We presume that the sampling points are insufficient to resolve the exact probabilistic distribution in the lower order 2 and 3.



(**a**) (**b**)

Fig. 4.7 *The mean and standard deviations of the drag coefficient of the three methods for orders 2 to 5:* ***(a)*** *mean value and* ***(b)*** *standard deviation*

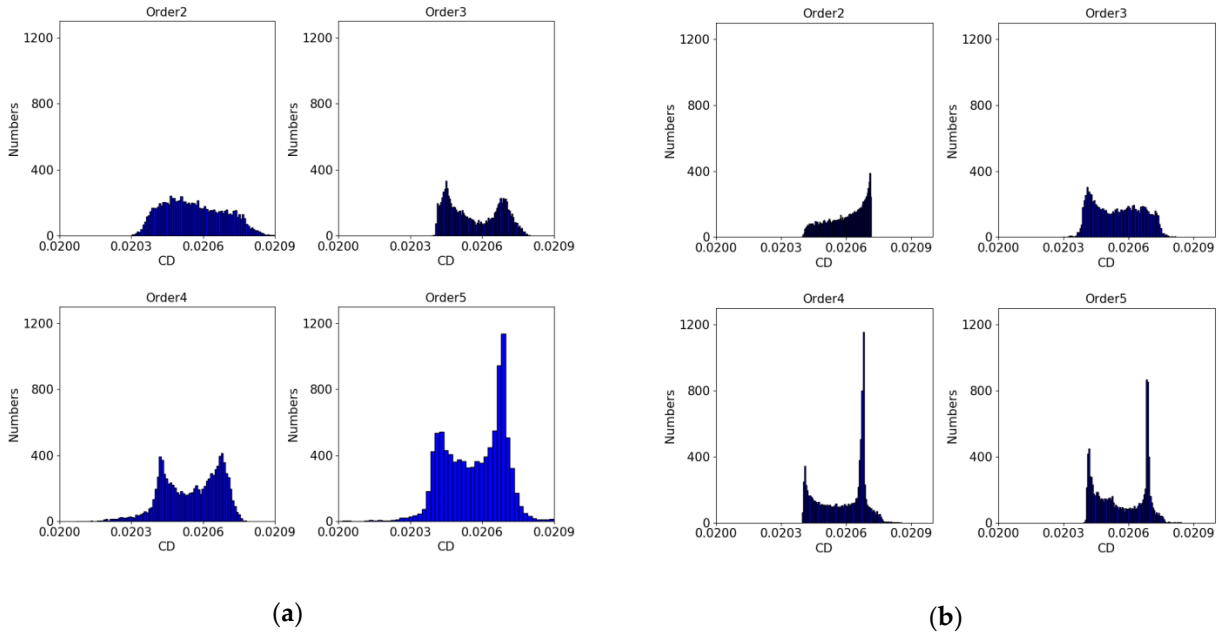


Fig. 4.8 *PDF analysis of the drag coefficient (CD): (**a**) Point-Collocation Non-Intrusive Polynomial and (**b**) Non-Intrusive Spectral Projection*

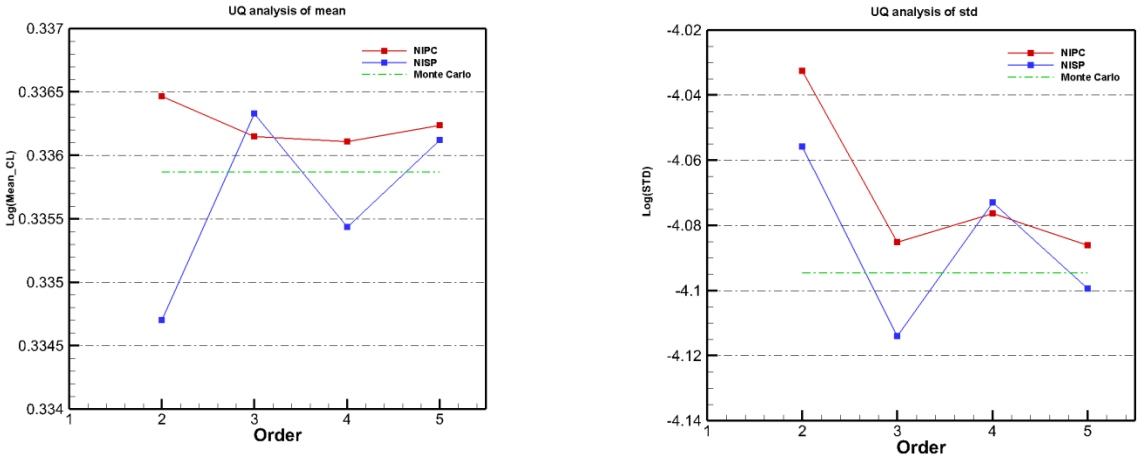
Table 4.9 shows the mean and STD of the lift coefficient obtained by the three methods. The mean values of all methods are predicted to be approximately 1.399 whereas the deterministic solution is 1.390. This difference can result because the distribution of PDF is biased toward the positive difference value, which is confirmed in Figure 4.10.

Table 4.9 *The mean and standard deviation values of the lift coefficient (C_L) for A-airfoil*

Order	Mean_NIPC	STD_NIPC	Mean_NISP	STD_NISP	Mean_MC	STD_MC
2	1.39999	0.017730	1.397522	0.017321	1.399155	0.016661
3	1.39955	0.016821	1.399802	0.016343		
4	1.39949	0.016969	1.398551	0.017026		
5	1.39967	0.016804	1.399507	0.016583		

Similar to the convergence of the drag coefficient, the mean value of the lift coefficient converges early from the 2nd order but the STD shows a large difference when the order is 2. At the 5th order, there is a 0.85% difference in Point-Collocation NIPC and 0.46%

in NISP. Even though the convergence rate of Point-Collocation NIPC is faster than that of NISP, the accuracy at the highest order (5^{th} order) is higher in NISP than Point-Collocation NIPC.



(**a**) (**b**)

Fig. 4.9 *The mean and standard deviations of the lift coefficient of the three methods for orders 2 to 5: (**a**) mean value and (**b**) standard deviation*

In the lift coefficient, similar patterns of the PDF were obtained depending on the method and the order. Even though the peak of PDF is located at a lower value, the flat distributions of the lift coefficients range to the value of 1.43 or 1.44 in every case. This broad distribution towards a higher lift coefficient makes the predicted mean lift coefficient move to the higher value than the deterministic solution.

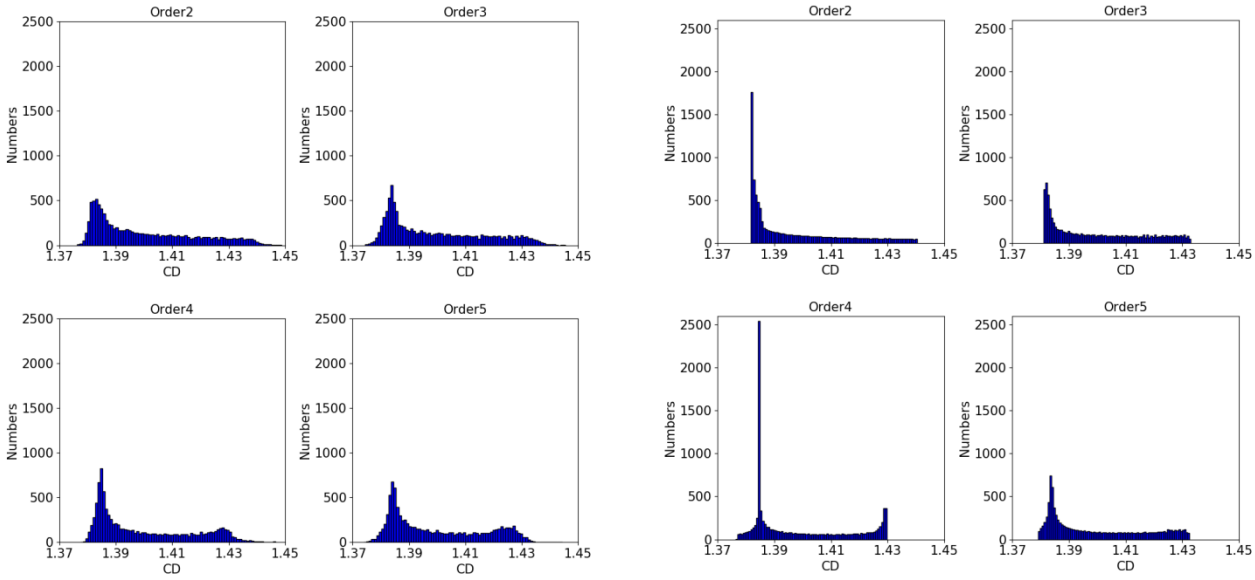


Fig. 4.10 *PDF analysis of the lift coefficient (CL): (**a**) Point-Collocation Non-Intrusive Polynomial and (**b**) Non-Intrusive Spectral Projection*

4.4 Comparison of the two methods

Figure 4.11 describes the distribution of the PDF of Point-Collocation NIPC and NISP for the 5^{th}-order polynomial and Monte Carlo method with 500 sampling points. In the case of flat plate, the peak values of the drag coefficients, 0.0041 agree well in three methods; NISP, PC-NIPC and Monte Carlo. This peak is located at a higher value than that (0.00401) of the deterministic solution. In the drag coefficients of Aerospatiale A-Airfoil simulation, there are two peaks in the PDF and the left peak is lower than the right one in every methods. The predicted peak values in NISP and PC-NIPC methods are approximately 0.0206 with 0.03% error with that of Monte Carlo method. The pattern of the distribution of the lift coefficient in A-Airfoil simulation is similar to that of the drag coefficient in the flat plate simulation. The peak value of NISP show smaller difference with 0.46% than that of PC-NIPC, 0.86% when compared with that of Monte Carlo's result. Even though the number of samplings, 500 in the present work is not sufficient to get the converged solution of MC, the overall pattern distribution and the number of the peak are consistent between MC and two adopted methods. When the

number of sample points for the UQ technique is considered, even though Point-Collocation NIPC requires a smaller number of samples than NISP (Table 4.3), similar accuracy and PDF as the mean and STD of the quantities of interest can be obtained. In particular, the number of random variable increases and the Point-Collocation NIPC is useful from a computational cost and time point of view.

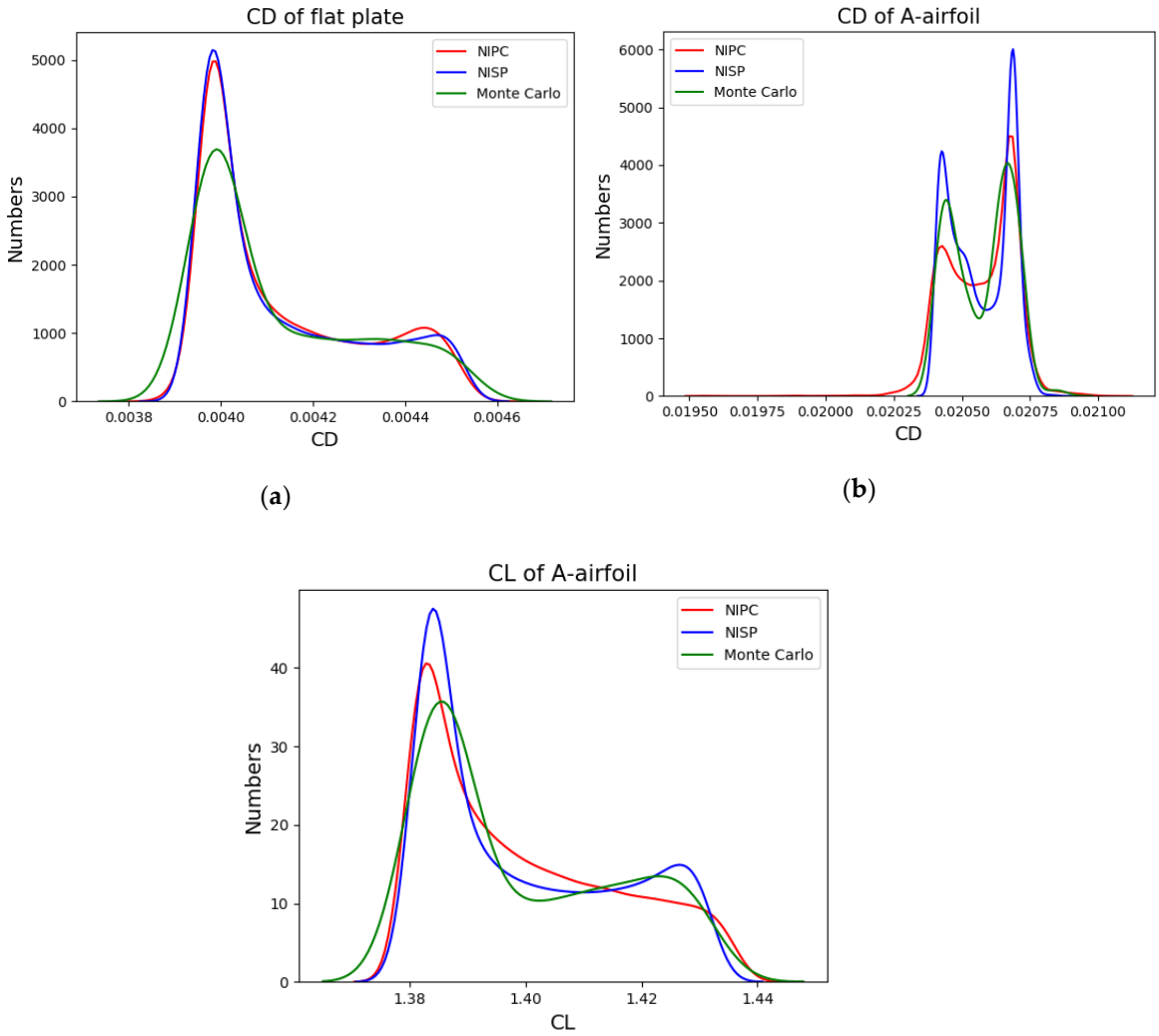


Fig. 4.11 *The distributions of the PDF at the 5th order of NIPC, NISP, and Monte Carlo (CL): (a) CD of flat plate, (b) CD of A-airfoil, and (c) CL of A-airfoil*

Sensitivity analysis of three model coefficients, c_{a2}, c_{e1}, and c_{e2}, was carried out using the Sobol index of the two methods in Tables 4.10 and 4.11. The Sobol indices provides quantitative information regarding how much effect the corresponding

random parameters have on quantities of interest. In the Point-Collocation NIPC and NISP methods, the most effective model coefficient is calculated as C_{e1}, whose Sobol index is approximately over 90% in all outputs of both the flat plate and A-airfoil cases. The model coefficient C_{e1} is included in the source term of the transport equation of the intermittency and is related to the F_{legnth} and F_{onset} terms. The other two variables, C_{e2} and C_{a2}, show similar effects on the results of less than 5% of Sobol indices. This means that more careful calibration of the model coefficient C_{e1} is required compared to other coefficients for better prediction of transitional flow.

Table 4.10 Sobol index of three model coefficients for Point-Collocation NIPC

Sobol Index	**Ca_2**	**Ce_1**	**Ce_2**
C_D Flat plate	0.01934	0.983	0.01955
C_D A-airfoil	0.32256	0.8963	0.15859
C_L A-airfoil	0.05347	0.9793	0.03839

Table 4.11 Sobol index of three model coefficients for NISP

Sobol Index	**Ca_2**	**Ce_1**	**Ce_2**
CD Flat plate	0.0034	0.9945	0.0022
CD A-airfoil	0.0287	0.9554	0.0281
CL A-airfoil	0.00024	0.9999	0.000234

Figure 4.12 displays the distribution of the skin friction coefficient from Point-Collocation NIPC at the 5^{th} order. It is confirmed that the change of the model coefficients has a significant effect on the variation of the skin friction coefficient at the location of the flat plate. There is a small STD before the transition point (i.e., laminar flow region) and then a large STD after the transition (i.e., developing turbulent flow region). This means that the model closure coefficients have a large effect on the turbulent flow region of the deterministic solution.

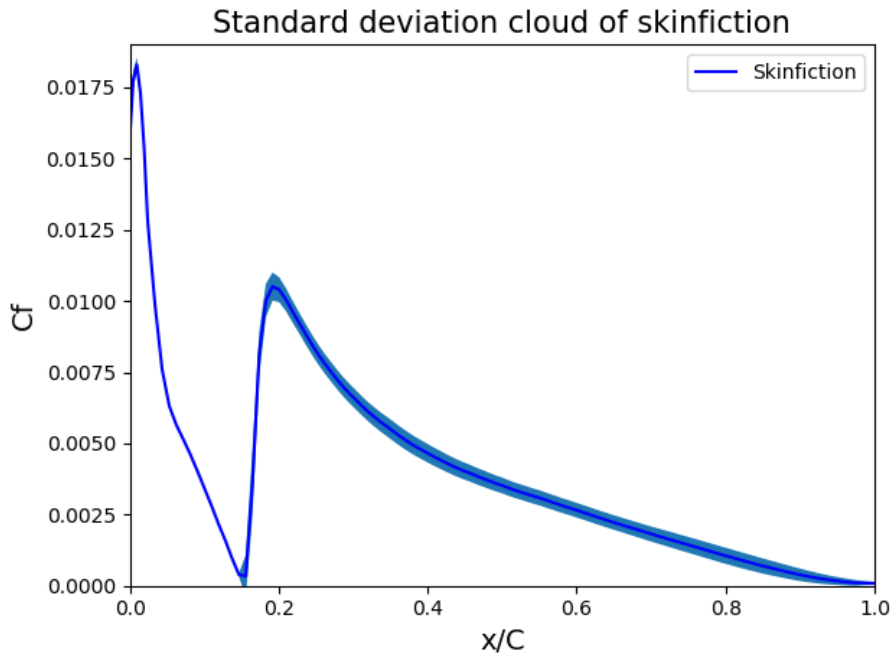


Fig. 4.12 *Variation of the skin friction coefficient of the upper surface of the A-airfoil for the Point-Collocation NIPC method*

In this study, we compared Point-Collocation NIPC and NISP methods for the $\gamma - \mathrm{Re}_{\theta}$ transitional model [21]. Three model coefficients, $c_{a2,}$ $c_{e1,}$ and c_{e2}, were considered as multiple random inputs with the assumption of a uniform distribution with ± 10% deviation. The quantities of interest are the drag and lift coefficients in transitional flow around a flat plate and Aerospatiale A-Airfoil.

The Point-Collocation NIPC has the advantage of having a small number of samples for uncertainty quantification analysis, even though the oversampling ratio is set to 2, followed by the NISP method. It was confirmed that the Point-Collocation NIPC will be useful when multiple random inputs are considered. Convergence with respect to the polynomial order is faster in Point-Collocation NIPC than NISP. Also, the mean value of the output can be obtained with a lower order than STD. This means that a higher polynomial order is necessary to obtain a converged STD. At the highest polynomial order of 5 evaluated in this study, NISP demonstrated slightly more accurate results when compared to the Monte Carlo simulation. However, the accuracy of NISP depends on whether it is an even or odd polynomial order. Through analysis of the Sobol index,

the most effective model coefficient of the $\gamma - Re_\theta$ transitional model is c_{e1}, which is related to F_{length} and F_{oneset} in the intermittency equation, by approximately over 90%.

In summary, when multiple random inputs are considered, Point-Collocation NIPC is able to provide reasonable stochastic values such as the mean and STD with a smaller number of samples than NISP. If only one or two variables are regarded as random inputs, NISP seems to be better at a high order.

Chapter 5. Bayesian Calibrated Uncertainty Parameters

The next step is to study propagation of the uncertain parameters that are determined in Chapter 4. In order to do that, we use the point collocation Non-Intrusive Polynomial method to create surrogate models. In this chapter we just applied on the flat plate simulation and A-Airfoil simulation to create a surrogate model by using the point collocation Non-Intrusive Polynomial method.

5.1 gPCE as surrogate Point Collocation model.

The main difficulty of solving the Bayesian inference in inverse problems is the extensive computational effort. In order to compute the posterior distribution, the forward model in the likelihood function has to be solved many times. Several approaches to overcome this obstacle have been proposed in literature lately. Some studies apply Model Order Reduction (MOR) to forward model to reduce the computation. Another approach is using surrogate models such as Gaussian Process emulation. As shown in chapter 3, the gPC approximation can be used as a surrogate model. Marzouk applied the gPC approximation to the Bayesian inverse problem. The posterior distribution can be approximated by using the gPC in the evaluation of the likelihood function. It also has been proven in the paper that the approximated posterior distribution converges to the exact posterior the Kullback-Leibler divergence measure.

The gPC surrogate can be considered as a polynomial surface response model of parameters. By considering the parameters with their specified ranges as independently uniformly distribution, the model can be approximate by

$$y^{M} \approx \hat{y}^{M}(\mathrm{q}) = \sum_{K=0}^{N_P} \beta_k \omega_k(q) \tag{5.1}$$

where gPC basis function Ψ_k are constructed from 1D-Legendre polynomial $\mathrm{Le_k}$ due to the uniform distribution of RVs. In this thesis the application of gPC as surrogate model to the parameter calibration problem is proposed.

Then the gPC approximation is used as the forward model in the evaluation of the likelihood function as:

$$L(q) = \left(\frac{1}{2\pi \det(C_W)} \right)^{n/2} \exp\left(-\frac{1}{2} (\hat{y}^M(q) - y)^T C_w^{-1} (\hat{y}^M(q) - y) \right) \tag{5.2}$$

This formulation can be applied to both batch and recursive method to approximate the posterior distribution. This approach is very useful in case that the original model y^M requires an extensive computation. The limitation of this method is that the gPC approximation is valid only in defined parameter space. The sampling outside the defined parameter space could lead to erroneous result. To avoid that problem, one should use the define parameter space as a prior of the inverse problem, or a truncated Gaussian distribution [15] in case of a prior normal distribution.

5.2 Bayesian Calibrated Uncertainty Parameters.

On the correlation matrix we are also considering p values that represent the sensitive of coefficients. It was called Pearson correlation coefficient. In statistics, the Pearson correlation coefficient also referred to as Pearson's r, the Pearson product-moment correlation coefficient or the bivariate correlation, is a measure of the linear correlation between two variables X and Y. According to the Cauchy–Schwarz inequality it has a value between +1 and −1, where 1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation. It was developed by Karl Pearson from a related idea introduced by Francis Galton in the 1880s and for which the mathematical formula was derived and published by Auguste Bravais in 1844. With surrogate model is with p =1 strong positive relationship, p= 0 not linearly correlated and p = -1 strong negative relationship.

$$\rho_{x,y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{5.3}$$

where:

- COV is the covariance
- σ_X is the standard deviation X

- σ_Y is the standard deviation of Y

the formula for ρ can be expressed in terms of mean and expectation. Since

$$\text{cov}(X, Y) = E\left[(X - \mu_x)(\text{Y} - \mu_y)\right] \tag{5.4}$$

the formula for p can also be written as:

$$\text{cov}(X, Y) = \frac{E\left[(X - \mu_x)(\text{Y} - \mu_y)\right]}{\sigma_X \sigma_Y} \tag{5.5}$$

In the sampling process sigma value that is noise values with normal distribution. It is also represented as the error value between surrogate model which was made by point collocation Non-intrusive Polynomial Chaos method and likelihood function that was estimated by Bayesian inference.

This chapter discusses the result of the Bayesian calibration for the transitional model for the flat plate and the Aerospatiale airfoil. In this chapter we just show the result for order 2 polynomial. The most sensitive coefficients are considered and other coefficients are fixed. The result for the γ-R_θ transitional model proved to be successful as you will see later on this chapter.

In the Figures 5.3, 5.6 and 5.9 the correlation value (p) of the Ce_1 nearest 1 value, it shows that the Ce_1 strong positive relationship with drag and lift coefficient over flat plate and Aerospatiale airfoil.

Figures 5.2, 5.5, and 5.8 show the mean and variance of the coefficient in the $\gamma - Re_\theta$ transitional model. The mean value of Ce_1 is 0.9, Ce_2 is 50.01, Ca_2 is 0.06. It proved these values approximately with original function which was shown in chapter 3.Figures 5.1, 5.4, and 5.7 illustrate MCMC (Markov Chain Monte Carlo) processes in the Bayesian inference with 10000 sampling therein it run simulation with 4 chains. These show the convergence process of MCMC.

In the Table 5.1, 5.2 and 5.3 the MCMC results are summarized over flat plate and Areospatial A-airfoil cases. Mean and standard describe of each coefficient. In table we also show the mean and stand deviation of each coefficient also detail the mean and stand deviation at 25%, 50%, 75% position and the maximum and minimum value of coefficient are shown.

The Figures 5.10, 5.11, 5.12 describes the result when we only change the Ce_1 value with error $\pm 10\%$ and keep the other coefficients constant. In the chapter 4 Ce_1 was the most effective coefficient in $\gamma - Re_\theta$ transition model. Once again we could prove Ce_1 have strong positive relation with the drag and lift coefficient over simulation of flat plate and Aerospatiale airfoil.

Generally we observe that uncertainty of the results is reduced at places where the density of the simulation data is highest. In addition, when the simulation uncertainty is lower, this also improves the uncertainty via the simulation results. The Bayesian inference results also turn out to be much more accurate than the deterministic results obtained using the default closure coefficient. The variance between the model's results has also reduced as predicted in chapter 4.

Aerospatiale airfoil for Cd:

Table 5.1 Result after Bayesian calibration

	Ca2	Ce1	Ce2	sigma	Cd
count	40000	40000	40000	40000	40000
mean	0.060041	1.004326	50.47536	0.00012	0.020681
std	0.003262	0.023254	2.851611	0.000021	0.000034
min	0.048007	0.080264	39.52427	0.000064	0.02057
25%	0.057833	0.994983	48.52861	0.00104	0.020678
50%	0.060033	1.00726	50.5385	0.000117	0.020678
75%	0.062245	1.018089	52.53308	0.000132	0.020701
max	0.073603	1.100643	58.94249	0.000309	0.020925

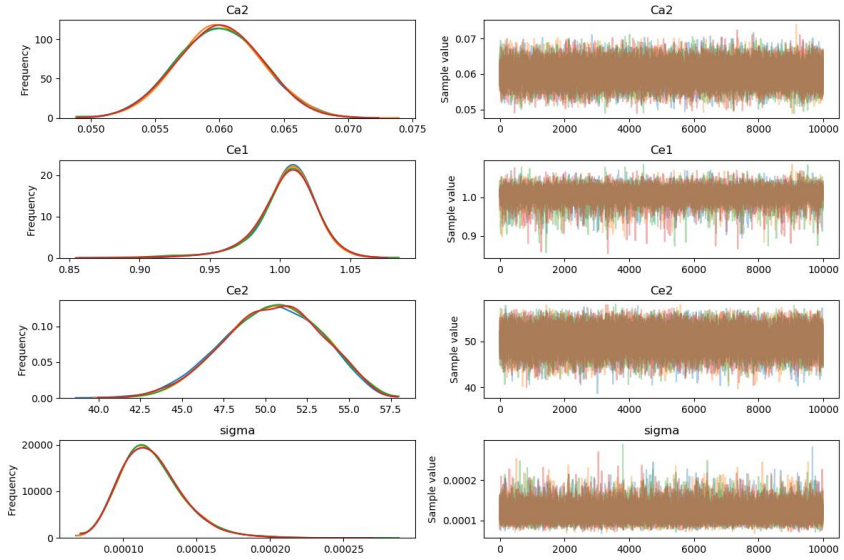


Fig. 5.1 *Process MCMC in the Bayesian inference on A-airfoil case*

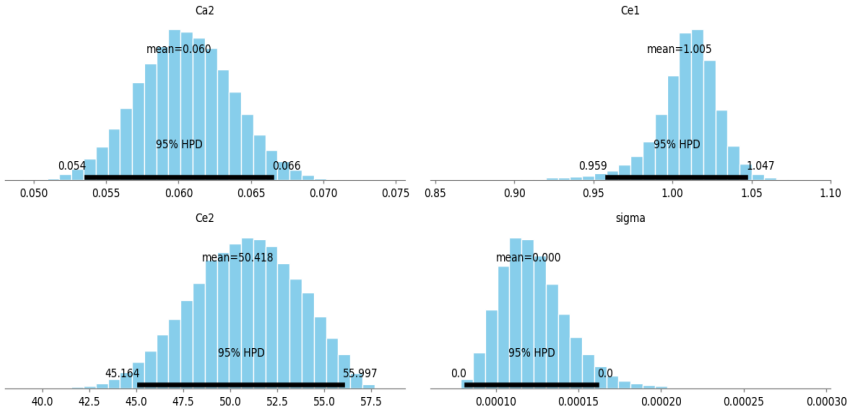


Fig. 5.2 *Histogram after Bayesian calibration on A-airfoil case*

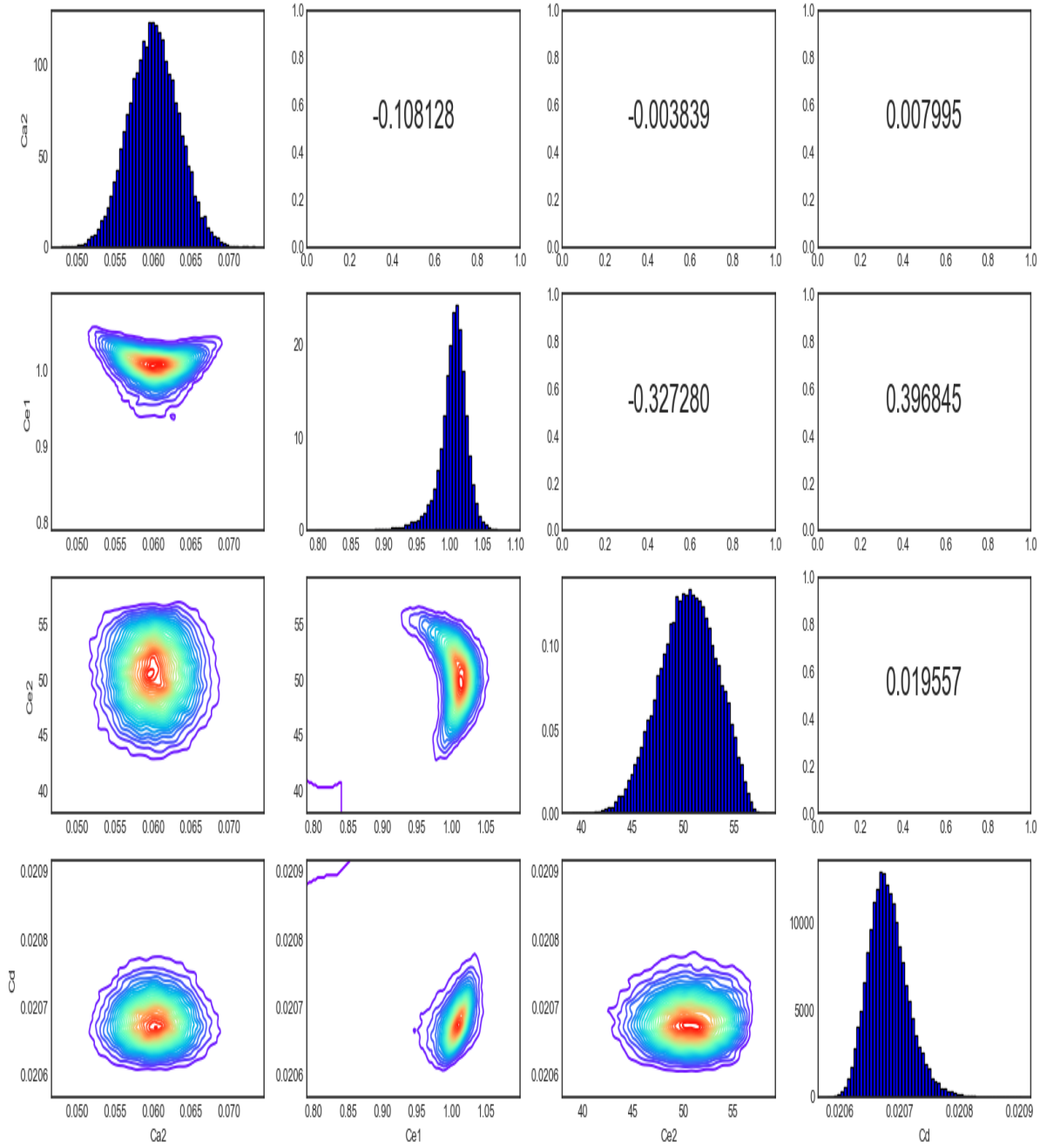


Fig. 5.3 *Histogram and correlation matrix after Bayesian calibration on A-airfoil case*

Aerospatiale airfoil for Cl:

Table 5.2 Result after Bayesian calibration

	Ca2	Ce1	Ce2	sigma	Cd
count	40000	40000	40000	40000	40000
mean	0.060169	0.988967	50.00378	0.017879	1.416592
std	0.003537	0.022222	2.774998	0.003189	0.004998
min	0.046061	0.938057	39.35955	0.008924	1.399862
25%	0.057797	0.978376	48.12376	0.015628	1.413102
50%	0.060136	0.98667	50.00745	0.017436	1.41613
75%	0.062521	0.995554	51.8816	0.019629	1.419603
max	0.074566	1.2101	61.06354	0.045417	1.450592

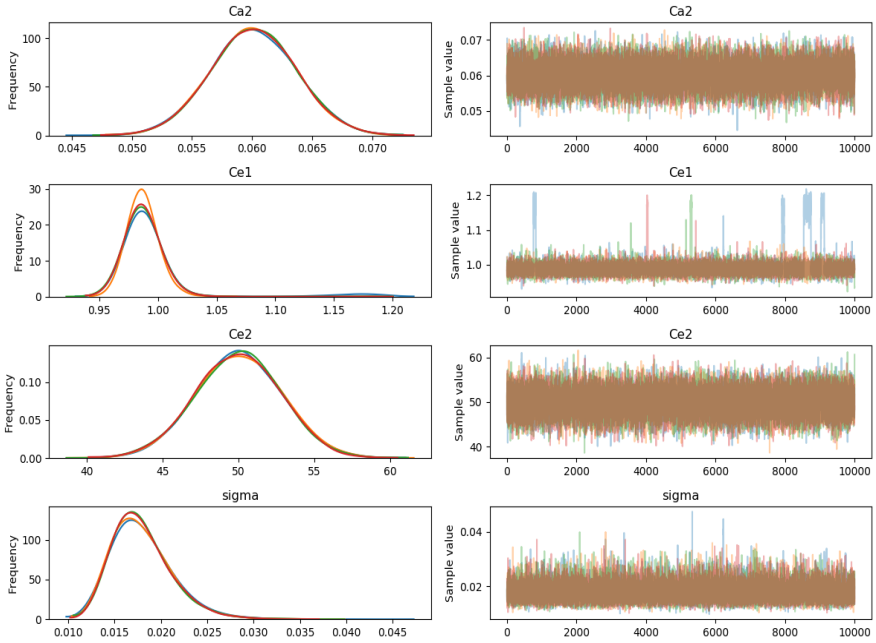


Fig. 5.4 *Process MCMC in the Bayesian inference on A-airfoil case*

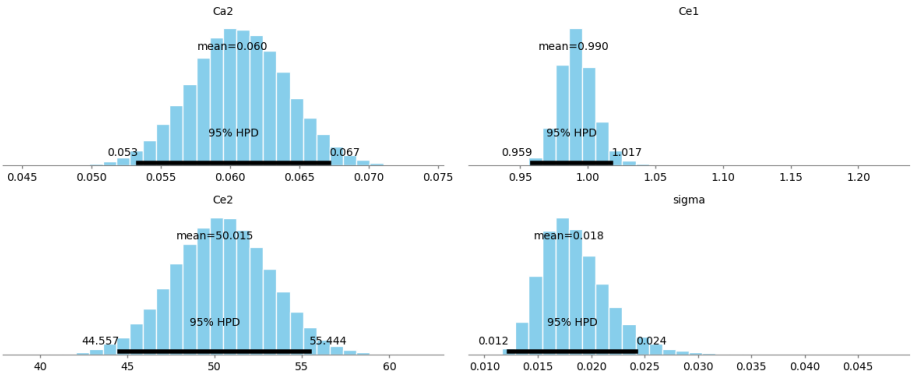


Fig. 5.5 *Histogram after Bayesian calibration on A-airfoil case*

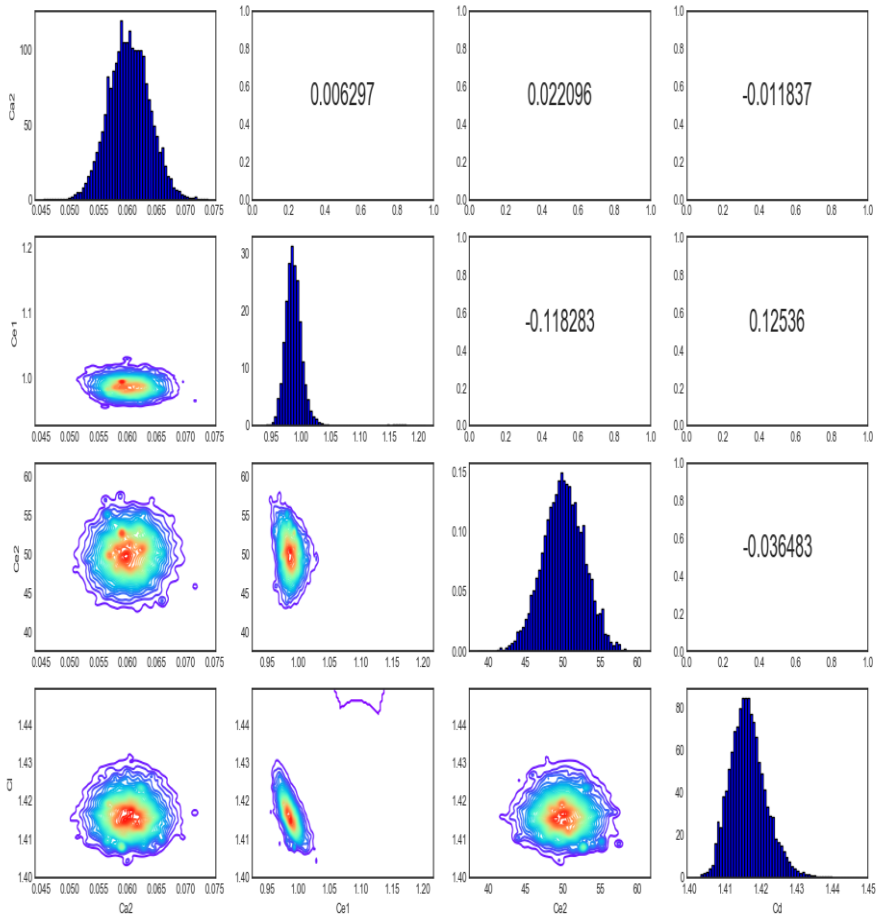


Fig. 5.6 *Histogram and correlation matrix after Bayesian calibration*

Flat plate for Cd:

Table 5.3 Result after Bayesian calibration

	Ca2	Ce1	Ce2	sigma	Cd
count	40000	40000	40000	40000	40000
mean	0.060052	0.990955	49.76604	0.000185	0.004301
std	0.003402	0.037206	2.823036	0.000033	0.000052
min	0.04708	0.926955	38.64756	0.000102	0.004144
25%	0.057729	0.973676	47.8736	0.000162	0.004265
50%	0.060074	0.982195	49.79495	0.000181	0.004297
75%	0.062353	0.992343	51.68134	0.000204	0.004332
max	0.073538	1.167536	63.91814	0.000546	0.004599

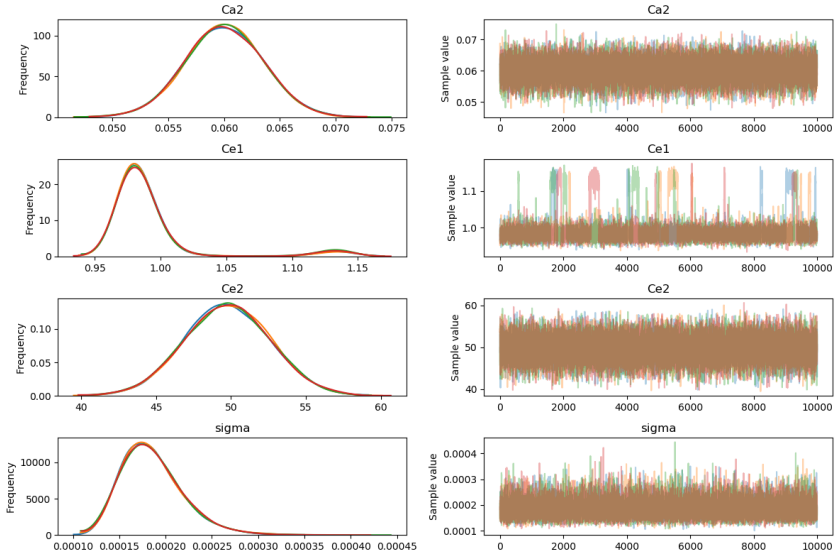


Fig. 5.7 *Process MCMC in the Bayesian inference on flat plate case*

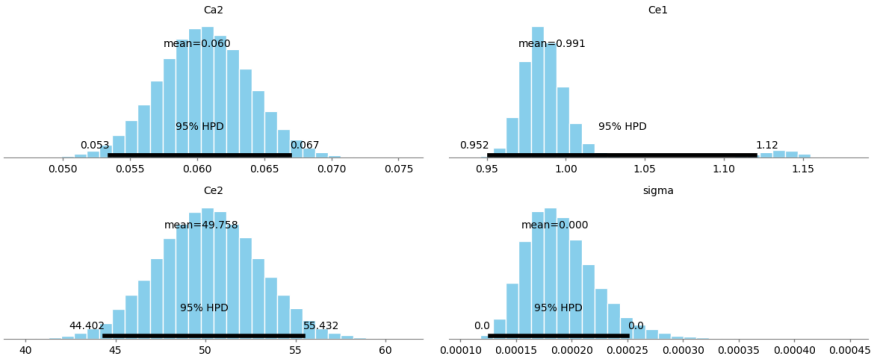


Fig. 5.8 *Histogram after Bayesian calibration on flat plate case*

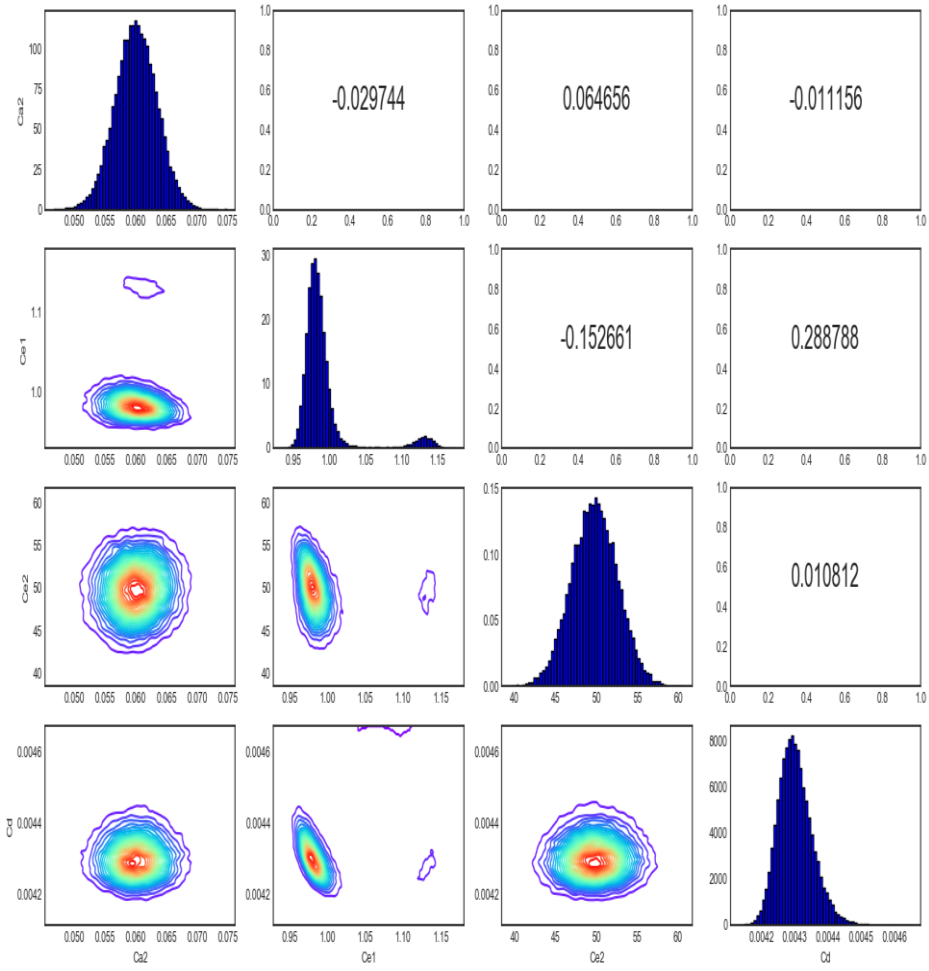


Fig. 5.9 *Histogram and correlation matrix after Bayesian calibration on flat plate*

Single variable:

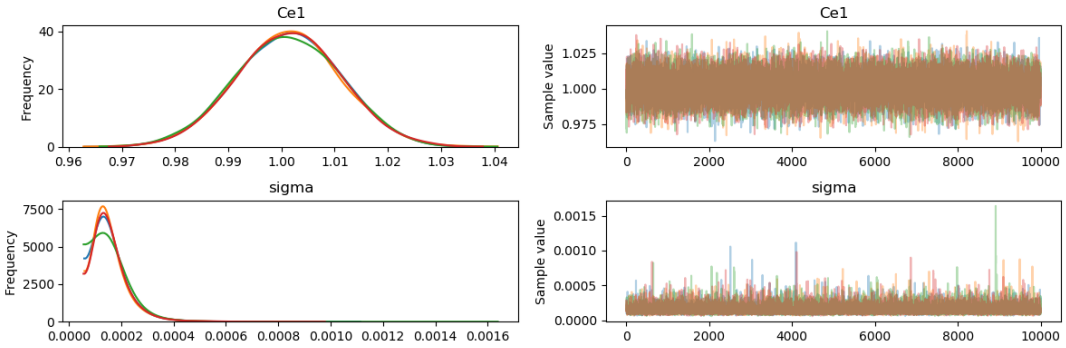


Fig. 5.10 *Process MCMC in the Bayesian inference*

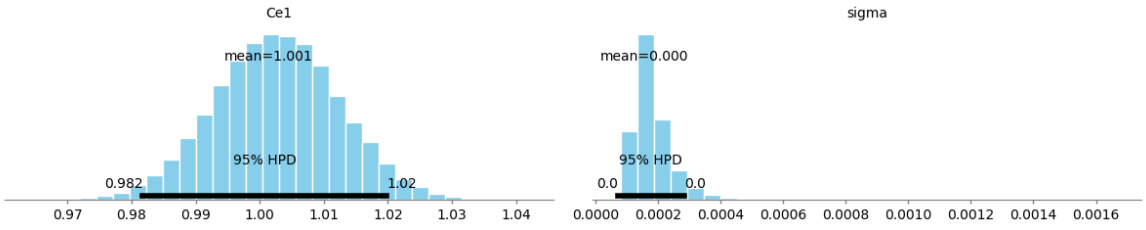


Fig. 5.11 *Histogram after Bayesian calibration*

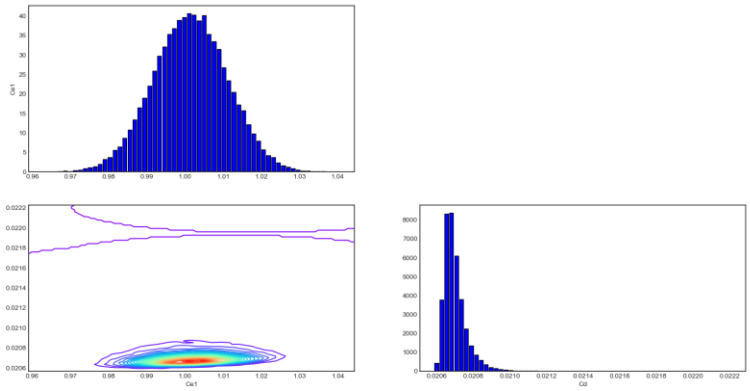


Fig. 5.12 *Histogram and correlation matrix after Bayesian calibration*

Chapter 6. Conclusion and Recommendations

This chapter draws main conclusions based on the work presented in this thesis. In particular, a summary of the thesis is provided together with a discussion on the key findings, the significance of the work, the future work and outlook.

6.1 Conclusions

In this study, we compared Point-Collocation NIPC and NISP methods for the $\gamma - \mathrm{Re}_\theta$ transitional model [21]. Three model coefficients, $c_{a2,}$ $c_{e1,}$ and c_{e2}, were considered as multiple random inputs with the assumption of a uniform distribution with ± 10% deviation. The quantities of interest are the drag and lift coefficients in transitional flow around a flat plate and Aerospatiale A-Airfoil.

The Point-Collocation NIPC has the advantage of having a small number of samples for uncertainty quantification analysis, even though the oversampling ratio is set to 2, compared with the NISP method. It was confirmed that the Point-Collocation NIPC will be useful when multiple random inputs are considered. Convergence with respect to the polynomial order is faster in Point-Collocation NIPC than NISP. Also, the mean value of the output can be obtained with a lower order than STD. This means that a higher polynomial order is necessary to obtain a converged STD. At the highest polynomial order of 5 evaluated in this study, NISP demonstrated slightly more accurate results when compared to the Monte Carlo simulation. However, the accuracy of NISP depends on whether it is an even or odd polynomial order. Through analysis of the Sobol index, the most effective model coefficient of the $\gamma - \mathrm{Re}_\theta$ transitional model is c_{e1}, which is related to F_{length} and F_{oneset} in the intermittency equation, by approximately over 90%. When multiple random inputs are considered, Point-Collocation NIPC is able to provide reasonable stochastic values such as the mean and STD with a smaller number of samples than NISP. If only one or two variables are regarded as random inputs, NISP seems to be better at a high order.

The UQ approaches are applied in this thesis to analyze and to identify transional flow simulation system. As the UQ approaches require extensive computation, the

generalized polynomial chao (gPC) is able to reduce the computational effort massively. The gPC can provide accurate pdf approximations with much lower computation effort. Sensitivity analysis is used to quantify the relative contribution of the input parameters to the system responses. The exploitation of polynomial chaos expansion (PCE), namely using as a surrogate model, and using in the linear Bayesian update are presented. The calibration methods with the help of gPC approximation can provide the results very similar to the results without the approximation. This means that the gPC approximation can reduce the computational effort of the Bayesian inverse computation.

Generally we observe that the prior uncertainty is significantly reduced through the Bayesian calibration. Hence, the main objective of this thesis has been achieved. Note that the best results can be obtained at places where the density of the simulation data is highest. In addition, when the simulation uncertainty is lower this also improves the uncertainty of the simulation results. It is therefore very important to be in the possession of good simulation data in order to apply the technique successfully. The availability of good simulation data is something which can definitely be improved.The Bayesian calibrated results also prove to be much more accurate than the deterministic results obtained using the default closure coefficient. The variance between the different order's results has also reduced as was estimated in chapter 5.

6.2 Future works

The further works may need to include the following two aspects:

- Apply UQ to heat or fluid flow simulation.

This work was supported by the National Research Foundation of Korea grant funded by the Korea Government (NRF-2016R1D1A1B03934121).

REFERENCES

1. Ghanem, R.; Spanos, P.D. Polynomial chaos in stochastic finite elements. J. App. Mechanics 1990, 57, 197–202.

2. Ghanem, R. Stochastic finite elements with multiple random non-Gaussian properties. J. Eng. Mechanics 1999, 26–40.

3. Ghanem, R.G. Ingredient for a general purpose stochastic finite element formulation. Compu. Meth. App. Mecha. Eng. 1999, 168, 19–34.

4. Mathelin, L.; Hussaini, M.Y.; Zang, T.A.; Bataille, F. Uncertainty propagation for turbulent, compressible nozzle flow using stochastic methods. AIAA J. 2004, 42, 1669–1676.

5. Debusschere, B.J.; Naijm, H.N.; Pebay, P.P.; Knio, O.M.; Ghanem, R.G.; Maitre, O.P.L. Numerical challenges in the use of polynomial chaos representations for stochastic processes. SIAM J. Sci. Compu. 2004, 26, 698–719.

6. Reagan, M.; Najm, H.N.; Ghanem, R.G.; Knio, O.M. Uncertainty quantification in reacting flow simulations through non-intrusive spectral projection. Combustion and Flame 2003, 132, 545–555.

7. Mathelin, L; Hussaini, M.Y.; Zang, T.A. Stochastic approaches to uncertainty quantification in CFD simulations. Numerical Algorithms 2005, 38, 209–236.

8. Wiener, N. The homogeneous chaos. Am. J. Math. 1938, 60, 897–936.

9. Wiener, N.; Wintner, A. The discrete chaos, Am. J. Math. 1943, 65, 279–298.

10. Knio, O.M.; Maitre, O.P.Le. Uncertainty propagation in CFD using polynomial chaos decomposition. F. Dyna. Rese. 2006, 38, 616–640.

11. Xiu, D.; Karniadakis, G.E. The Wiener-Askey polynomial chaos for stochastic differential equation. J. Sci. Comput. 2000, 24, 619–644

12. Xiu, D.; Karniadakis, G.E. Modeling uncertainty in flow simulations via generalized polynomial chaos. J. Comput. Phys. 2003, 187, 137–167.

13. Cho, J.R.; Chung, M.K. A k-ε-γ equation turbulence model. J. Flu. Mechan. 1992, 237, 301–332.

14. Mayle, R.E. The role of laminar-Turbulence modeling of by-pass transition. ASME J. Turbomach. 1991, 113, 509–537.

15. Xiu, D. Numerical Methods for Stochastic Computations: A Spectral Method Approach; Princeton University Press: Princeton, NJ, USA, 2010.

16. Hosder, S.; Walters, R.; Balch, M. Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables, 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Honolulu, Hawaii, USA, 23–26 April 2007; AIAA 2007-1939.

17. Hosder, S.; Walters, R.; Perez, R. A Non-Intrusive Polynomial Chaos Method for Uncertainty Propagation in CFD Simulations, 44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, USA, 9–12 January 2006; AIAA 2006-891.

18. Loeven, G.J.A.; Witteveen, J.A.S.; Bijl, H. Probabilistic Collocation: An Efficient Non-Intrusive Approach For Arbitrarily Distributed Parametric Uncertainties, 45th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, USA, 8–11 January 2007; AIAA 2007-317.

19. Platteeuw, P.D.A.; Loeven, G.J.A.; Bijl, H. Uncertainty quantification applied to the k-epsilon model of turbulence using the probabilistic collocation model, 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, April, 2008.

20. Schaefer, J.; West, T.; Hosder S.; Rumsey, C.; Carlson, J.; Kleb, W. Uncertainty quantification of turbulence model coefficients for transonic wall-bounded flow, AIAA J. 2017, 55, 195–213.

21. Langtry, R.B.; Menter, F.R. Correlation-based transition modeling for unstructured parallelized computational fluid dynamics Codes. AIAA J. 2009, 47, 2894–2906.

22. Menter, F.R. Two-equation eddy-viscosity turbulence models for engineering application. AIAA J. 1994, 32, 1598–1605.

23. Drela, M.; Giles, M.B. Viscous-inviscid analysis of transonic and low Reynolds number airfoils. AIAA J. 1987, 25, 1347–1355.

24. Rumsey, C.L.; Gatski, T.B.; Ying, S.X.; Bertelru, A. Prediction of high-lift flows using turbulent closure models. AIAA J. 1998, 36, 765–774.

25. Sullivan, T.J. Introduction to Uncertainty Quantification; Springer: Cham, Switzerland, 2015; Volume 63.

26. Bijl, H.; Lucor, D.; Mishra, S.; Schwab, C. Implementation of intrusive polynomial chaos in CFD codes and application to 3D Navier-Stokes. In Uncertainty Quantification in Computational Fluid Dynamics; Springer: Switzerland, 2013.

27. Perez, R.; Walters, R. An Implicit Compact Polynomial Chaos Formulation for the Euler Equations, 43rd AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, USA, 10–13 January, 2005, AIAA 2005-1406.

28. Feinberg, J.; Langtangen, H.P. Chaospy: An open source tool for designing methods of uncertainty quantification. J. Comput. Sci. 2015, 11, 46–57.

29. NISP details. http://atoms.scilab.org/toolboxes/NISP/2.1

30. Fu, S.; Wang, L. Progress in turbulence/transition modeling. Adv. Mech. 2007, 37, 409–16.

31. Zarmehri, A. CFD with OpenSource Software: Transitional Turbulence Model Tutorial; Jorgensen, N.G.; Andric, J. Developed for OpenFoam-2.0.x, 19 November 2012.

32. Schmidt, S.; Thiele, F. Detached eddy simulation of flow around A-airfoil. Flow, Turbu. Combu. 2003, 71, 261–278.

33. Savill, A.M. One-point closures applied to transition. In Turbulence and Transition Modeling; Hallback, M. Dordrecht: Kluwer, 1996, pp. 233–268.

34. Rumsey, C.L. Turbulence Modeling Resource. http://turbmodels.larc.nasa.gov/.

35. Ke, J.; Edwards, J.R. Numerical simulations of turbulent flow over airfoils near and during static stall. J. Aircraft 2017, 54, 1960–1978.

36. Goldberg, U.C.; Batten, P.; Peroomian, O.; Chakravathy, S. The R-□ transition prediction model. Inter. J. Compu. Flu. Dynamics 2015, 29, 26–39.