Ref. Ares(2024)4544952 - 24/06/2024

GO-VIKING

Research and Innovation Action (RIA)

This project has received funding from the Euratom research and innovation programme 2021-2025 under Grant Agreement No 101059603

Start date : 2022-06-01 Duration : 48 Months

Development of UQ approaches for FIV analyses

Authors : Pr. Joris DEGROOTE (Ghent University), Henri Dolfen (UGent), Richard Dwight (TUDelft), Ali Eidi (TUDelft)

GO-VIKING - Contract Number: 101059603

Project officer: Angelgiorgio IORIZZO

Document title	Development of UQ approaches for FIV analyses
Author(s)	Pr. Joris DEGROOTE, Henri Dolfen (UGent), Richard Dwight (TUDelft), Ali Eidi (TUDelft)
Number of pages	77
Document type	Deliverable
Work Package	WP5
Document number	D5.3
Issued by	Ghent University
Date of completion	2024-06-24 14:43:08
Dissemination level	Public

Summary

Description of the implemented UQ approaches for FIV analysis

Approval

Date	By
2024-06-24 14:49:39	Dr. Papukchiev ANGEL (GRS)
2024-06-24 14:50:32	Dr. Papukchiev ANGEL (GRS)

Deliverable 5.3

Development of UQ approaches for FIV analyses

Version N°1.0
June / 2024

Henri Dolfen[1], Joris Degroote[1], Richard Dwight[2], Ali Eidi[2]

[1]Ghent University

[2]TU Delft

Funded by
the European Union

Disclaimer

The content of this report reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Document information

Grant Agreement / Proposal ID	101060826
Project Title	Gathering expertise On Vibration ImpaKt In Nuclear power Generation
Project Acronym	GO-VIKING
Scientific Coordinator	Angel Papukchiev, Angel.Papukchiev@grs.de, GRS Kevin Zwijsen, zwijsen@nrg.eu, NRG
Project starting date (duration)	1st June 2022 – 31st May 2026 (48 Months)
Related Work Package	WP5
Related Task(s)	Task 5.2
Lead Organisation	UGent
Contributing Partner(s)	UGent, TUDelft
Due Date	Month 25
Submission Date	
Dissemination level	PU - Public

History

Date	Version	Submitted by	Reviewed by	Comments
24/06/2024	1	Henri Dolfen	Angel Papukchiev	

Table of contents

1. PART 1: Parameter estimation and uncertainty quantification in RANS turbulence closure 11
1.1 Introduction 11
1.1.1 Motivation 11
1.1.1.1 Turbulence closure modelling 11
1.1.1.2 Probability in Computational Science and Engineering 12
1.2 Bayesian parameter estimation 13
1.2.1 Bayes' theorem 13
1.2.2 Parameter identification for a CFD problem 15
1.2.3 Application to regression: Model inadequacy 16
1.2.4 Computational considerations 18
1.2.4.1 Monte-Carlo methods 19
1.2.4.2 Surrogate modelling methods 20
1.2.4.3 Variational methods 20
1.3 RANS closure model coefficient 22
1.3.1 Classical identification of closure coefficients 22
1.3.1.1 Canonical flow I: Decay of isotropic turbulence 22
1.3.1.2 Canonical flow *II*: Free shear flows 23
1.3.1.3 Discussion 24
1.3.2 Bayesian calibration of closure model coefficients 25
1.3.2.1 RANS with $\boldsymbol{k-\varepsilon}$ closure Model 25
1.3.2.2 Experimental data for flat-plate boundary layers 26
1.3.2.3 Likelihood and model inadequacy 27
1.3.2.4 Priors on model coefficients and hyper-parameters 28
1.3.2.5 Calibration results 29
1.3.3 Bayesian model selection and model averaging 32
1.3.3.1 Theory of BMS and BMA 33
1.3.3.2 Application to turbulence modelling 34
1.3.4 Discussion 36
1.4 Bayesian estimation of Reynolds stress 36
1.4.1 Learning Reynolds-stress anisotropy from surface pressure 37
1.4.1.1 Turbulence production perturbation (TPP) 38
1.4.1.2 Anisotropy tensor perturbation (ATP) 39
1.4.1.3 Random matrix perturbation (RMP) 40
1.4.1.4 Test case: Separated flow behind a hump (2d) 41
1.5 Applications in the GO-VIKING project 44

1.6 Conclusion 45

2. PART 2: Application of non-intrusive uncertainty quantification techniques to evaluate the effect of stochastic deformation on the vibration characteristics of a tube bundle in axial flow46

2.1 Introduction 46

2.2 Methodology 48

2.2.1 Geometry parameterization 49

2.2.2 CFD model 50

2.2.3 CSM model 52

2.2.4 FSI coupling of models 53

2.2.5 Monte Carlo method 55

2.2.6 Non-intrusive Polynomial Chaos method 56

2.3 Results 58

2.3.1 Steady flow through deformed tube bundle 59

2.3.2 Stochastic modal properties of deformed tube bundle 63

2.4 Discussion 67

2.5 Conclusion 68

List of figures

Figure 1.1 Cartoon of Bayesian updating for an aerofoil. 1) Black pdfs at the left represent priors on the simulation inputs; when propagated through the code they result in the black pdfs in the forces on the right (the “prior predictive distribution”). 2) A noisy measurement of C_L represented by the red pdf; applying Bayes gives the posterior $\rho(\theta \mid y)$ on the inputs (red pdfs on the left). 3) Propagating this posterior through the code, gives new information about the unmeasured C_D (the “posterior predictive distribution”). 16

Figure 1.2 Regression of data with a straight-line, with and without model inadequacy. 18

Figure 1.3 Ratio $u'v'/k$ as a function of wall-distance for a zero-pressure gradient turbulent boundary layer from DNS at $Re_\tau = 1272$, data from (Schlatter & Örlü, 2010). 24

Figure 1.4 The experimental data from (Coles, 1969). Reproduced from (W. N. Edeling et al., 2014). 27

Figure 1.5 Joint posterior density on model coefficients, visualized as all pairs of 2d-marginal distributions- for a moderately adverse pressure gradient (flow 3300). Reproduced from (W. N. Edeling et al., 2014). 31

Figure 1.6 The marginal posterior distributions of the model coefficients and hyper-parameters for the 13 boundary layer cases. 32

Figure 1.7 Posterior probabilities of five turbulence models for 14 calibration datasets. Figure reproduced with permission from (W. Edeling et al., 2014). 35

Figure 1.8 BMA posterior predictive distribution for case 4400 (strong adverse pressure-gradient). Figure reproduced with permission from (W. Edeling et al., 2014). 35

Figure 1.9 Computation mesh and Particle Image Velocimetry (PIV) data for the NASA hump flow. 41

Figure 1.10 Pressure measurements (circles) and RANS predictions, for the baseline (SST) model, and the corrected models (left). Turbulence anisotropy along a line in the wake for the corrected models. 42

Figure 1.11 Perturbations to barycentric coordinates Δx, Δy for the ATP (top) and RMP (bottom) models. 43

Figure 1.12 Velocity profiles at various x-locations for the corrected models compared with experimental data. 44

Figure 2.1: Schematic of the flow domain with uncertain deformation. The undeformed computational domain is shown in Figure 2.2. The left figure shows the xy-plane and indicates the section plane in red, which is shown in the right figure (scaled down along the z-direction). This plane is the ez-plane and makes an angle θ with the xz-plane. The right figure shows the extent of deformation e_0 within the ez-plane. The dark blue dashed line indicates a cut in the middle of the axial length, where the maximal deformation happens. 50

Figure 2.2: The CFD mesh (blue) and CSM mesh (grey) along with the geometrical parameters (also see Table 1) and boundary conditions. 51

Figure 2.3: The y-displacement of the midle of the rod ($z = 0.7\ m$) as a function of time for the case without bow deformation ($e_0 = 0$). The four dashed vertical lines indicate the 100th and 200th time step (green) and the 500th and 1500th time step (blue). 55

Figure 2.4: Contour of (a) the pressure drop Δp and (b) the maximal velocity V_{max}, with 20 equidistant contour lines drawn. 59

Figure 2.5: Comparison of the output histogram for Δp between Monte Carlo and Polynomial Chaos results of (a) 1st order and (b) 3rd order. The histograms for the PCE are obtained by feeding the Monte Carlo sample to the obtained response surfaces. The values on the vertical axis are normalized such that the area under the histogram integrates to 1, corresponding to a probability density. 60

Figure 2.6: Mean squared error of PCE with respect to MC of (a) the pressure drop Δp and (b) the maximal velocity V_{max}. 61

Figure 2.7: Maximal velocity as a function of the displacement extent e_0 and coloured by the angle θ 62

Figure 2.8: Comparison of the output histogram for V_{max} between Monte Carlo and Polynomial Chaos results of (a) 6th order and (b) 8th order. The histograms for the PCE are obtained by feeding the MonteCarlo sample to the obtained response surfaces. The values on the vertical axis are normalized such that the area under the histogram integrates to 1, corresponding to a probability density. 63

Figure 2.9: Fitted (a) frequency f and (b) damping ratio ζ as a function of the beginning of the fitting interval, which is starting at time step 200 to 800, for all realizations in the 8th order pCE. The different colours indicate the different deformation extents e_0, the colour intensity the angles θ, both corresponding to the nodes of the quadrature defined using Legendre polynomials on the intervals given by Eq. 3 and 4. 64

Figure 2.10: R^2-value as a function of the beginning of the fitting interval, which is starting at time step 200 to 800, for all samples in the 8th order PCE. 64

Figure 2.11: Contour plots of the response surfaces for different polynomial orders. The Gauss–Legendre quadrature points are indicated using black dots, with the size of each dot representing the weight. 66

Figure 2.12: Output histogram from a Hammersley sample of size 10^7 from the PCE of (a) f and (b) ζ. 67

List of tables

Table 2.1: Geometrical data (also see Figure 2.2). 51

Table 2.2: Material properties of liquid lead-bismuth-eutectic and steel. 52

Table 2.3: Descriptive statistics resulting from Monte Carlo method. 61

Table 2.4: Descriptive statistics of the output quantities of interest, computed from the 8th order PCE 66

Abbreviations and Acronyms

Acronym	description
ALE	Arbitrary Lagrangian–Eulerian
ATP	Anisotropy Tensor Perturbation
BMA	Bayesian Model Averaging
BMS	Bayesian Model Selection
CDF	Cumulative Distribution Function
CFD	Computational Fluid Dynamics
CSE	Computational Science and Engineering
CSM	Computational Structural Mechanics
DNS	Direct Numerical Simulation
ELBO	Evidence Lower Bound
FIV	Flow-induced vibration
FSI	Fluid-structure interaction
GP	Gaussian Process
gPC	Generalized Polynomial Chaos
IQN-ILS	Interface quasi-Newton with inverse Jacobian from a Least Squares model
KDE	Kernel Density Estimation
KL	Kullback–Leibler
LBE	Lead-bismuth eutectic

LES	Large Eddy Simulation
LEVM	Linear Eddy Viscosity Model
MAP	Maximum A Posteriori
MC	Monte Carlo
MCMC	Markov-Chain Monte-Carlo
MSE	Mean squared error
NUTS	NO U-Turn Sampler
ODE	Ordinary Differential Equation
PC	Polynomial Chaos
PCE	Polynomial Chaos Expansion
PDE	Partial Differential Equation
PDF	Probability Density Function
PWR	Pressurized Water Reactor
QoI	Quantity of interest
TKE	Turbulent Kinetic Energy
TPP	Turbulence Production Perturbation
UQ	Uncertainty quantification
RANS	Reynolds-averaged Navier–Stokes
RMP	Random Matrix Perturbation
RST	Reynolds Stress Tensor

Funded by
the European Union

Summary

Uncertainty quantification methods have been researched by UGent and TUDelft. The partners have focused on dissimilar sources of uncertainty. TUDelft focused on quantifying the closure error in Reynolds-averaged Navier–Stokes (RANS) turbulence models, while UGent aims at quantifying uncertainties in simulation parameters, such as operating conditions and geometry. Therefore, this deliverable is presented in two parts.

The first part delves into parameter estimation and uncertainty quantification in RANS turbulence closure using the Bayesian framework. It covers the ability of this framework to characterize complex relationships between model inputs, outputs, and state variables, and addresses the issue of model inadequacy through statistical methods. The challenges of high computational costs associated with Bayesian methods, especially in high-dimensional parameter spaces like those in fluid-structure interaction simulations, are discussed. Various cost-management techniques, including variational methods, are explored. The report emphasizes the relevance of Bayesian approaches to data-driven modelling and highlights the importance of a statistical perspective in improving the accuracy and reliability of RANS models.

The second part focuses on the forward propagation of uncertainty in geometric parameters through a numerical model using the Monte Carlo and generalized Polynomial Chaos frameworks. The proposed methodology is tested on a wire-wrapped tube bundle in lead-bismuth eutectic flow. A stochastically parameterized bow deformation is propagated through a fluid-structure interaction simulation model, to obtain the uncertainty of output quantities of interest related to flow-induced vibration, namely the first mode frequency and damping.

Keywords

Uncertainty quantification, Computational fluid dynamics, Fluid-structure interactions, Bayes' theorem, RANS model closure, Flow-induced vibration, Tube bundles, Nuclear fuel bow

1. PART 1: Parameter estimation and uncertainty quantification in RANS turbulence closure

1.1 Introduction

1.1.1 Motivation

We explore parameter estimation and uncertainty quantification (UQ) in turbulent flow modelling. The practice of modelling acknowledges imperfection, as highlighted by George Box's famous quote:

"*All models are wrong, some are useful*" (1976).

This underscores that no model can perfectly predict a real physical system under all conditions; instead, models should capture the dominant effects relevant to specific circumstances. This allows us to understand the world better and design complex systems confidently. In Computational Science and Engineering (CSE), models often involve partial differential equations (PDEs), boundary and initial conditions, constitutive laws, and various parameters. None of these elements perfectly represent reality. UQ aims to quantify the discrepancy between the model and the real system, using probability to express this inherent uncertainty.

1.1.1.1 Turbulence closure modelling

Turbulence is a complex, multi-scale phenomenon where only statistical properties, like temporal or ensemble averages, are relevant for engineering. The Reynolds-averaged Navier-Stokes (RANS) framework predicts these statistics directly by solving averaged Navier-Stokes equations for the mean velocity field and deriving additional transport PDEs for higher-order turbulence statistics, such as turbulence kinetic energy (k) and dissipation rate (ε). These additional PDEs include terms calibrated to match empirical data, combining exact physical laws with approximate empirical models.

RANS models are computationally efficient for high-Reynolds numbers, complex geometries, and design processes, making them the preferred method for predicting turbulent flows in engineering. However, due to significant approximations, RANS can yield inaccurate predictions under certain conditions. Most RANS models have multiple free parameters that require calibration using experimental data or scale-resolving simulations. Traditionally, these parameters are calibrated through manual tuning or optimization against canonical flows, resulting in deterministic parameter values that predict these flows accurately but may poorly predict other flows.

Experienced practitioners recognize which flows are well-predicted (e.g., wall-bounded attached flow) and which are not (e.g., separated flow). Nevertheless, RANS models do not provide information about their accuracy to users. Some flows cannot be accurately predicted by any parameter values in the RANS closure, a limitation known as *model inadequacy*. This inadequacy stems from choices made in the closure term in the momentum equations and

the specific transport PDEs used. Ultimately, the RANS approximation itself, focusing solely on turbulence statistics, is the fundamental source of this inadequacy.

1.1.1.2 Probability in Computational Science and Engineering

Given that all models are inherently imperfect, we aim to understand the discrepancy between a model and reality. Essentially, we ask: How confident are we in the model's predictions? Or, what is our uncertainty? Depending on the answers, we can either proceed with our decision/design/analysis or discard the model and choose a more accurate (and likely costlier) one.

To formalize these questions, we need to represent our uncertainty, indicating possible ranges for parameters, states, and predictions. Using random variables and probability distributions is a natural choice. Alternatives include interval arithmetic, Dempster-Shaffer evidence theory (Dempster, 2008; Shafer, 1976), and fuzzy logics (Zadeh et al., 1996). We choose probability because:

- Multi-variate probability distributions can capture complex statistical dependencies induced by our physical system and computational model, which we can exploit using conditional probability.
- Probability theory, particularly Bayesian probability, ensures coherent wagering on events, avoiding certain loss against an optimal opponent (Fitelson, 2005).
- Probability generalizes classical logic, incorporating the possibility of error (Tao, 2013).

By representing inputs, parameters, states, and outputs of our simulations as random variables with associated probability densities, we can:

- Determine output uncertainty based on input or parameter uncertainty (*uncertainty propagation*), see PART 2.
- Infer parameter distributions from output measurements (*calibration*).
- Estimate state based on measurements (*data assimilation*).
- Identify statistical dependencies between variables (*sensitivity analysis*).

Probability represents two fundamental types of uncertainty:

- **Aleatoric uncertainty**: Inherent, irreducible randomness (e.g., coin toss outcomes), objectively agreed upon by all observers.
- **Epistemic uncertainty**: Observer's lack of knowledge (e.g., a hidden coin toss result), which can be reduced by acquiring more information.

In physics, both uncertainties are present, but in this part we focus on deterministic systems and their imperfect models, mainly concerned with epistemic uncertainty. This uncertainty in model accuracy can be reduced with data.

In the context of RANS closure modelling, approaching it probabilistically involves using empirical information to apply statistical methods. RANS models combine physical laws (nonlinear PDEs) and empirical models (also PDEs) applicable to various flows. This makes using statistical methods more complex and interesting compared to classical curve fitting.

The remainder of the report is structured as follows:

Section 1.2: We consider the application of Bayesian probability to parameter estimation. We develop statistical models and apply them to the model problem of linear regression, with particular attention to the issue of model inadequacy.

Section 1.3: We apply the same techniques to the problem of parameter calibration of RANS models against experimental data, directly observing the model inadequacy present in simple RANS models.

Section 1.4: Motivated by the observations in Section 1.3, we address model inadequacy through the concept of model-form uncertainty.

1.2 Bayesian parameter estimation

We aim to solve the following problem:

- **Problem 1:** Parameter Estimation

Given a known model $f: X \times Q \to \mathbb{R}$ that represents a physical process and takes two arguments: $\boldsymbol{x} \in X \subset \mathbb{R}$ independent variable often representing space, and $\theta \in Q \subset \mathbb{R}$, a model parameter. Additionally, we have $N \in \mathbb{N}$ observations $\boldsymbol{y} \in \mathbb{R}^N$ of the same physical process at known locations $\boldsymbol{x} \in \mathbb{R}^N$. The task is to determine the value of $\theta \in Q$ that best represents the observations when used in the model.

This problem, often referred to as a *calibration problem*, falls under the broader category of *inverse problems*, where the goal is to estimate a model input rather than an output. The primary challenge is its vague formulation, specifically the unclear definition of *best*. While a least-squares fit of the observations might be considered, its impact on the accuracy of the estimated θ and the model's predictive capability remains uncertain. The problem can be either underdetermined (with infinitely many solutions) or overdetermined (more common in parameter calibration). Additionally, the model's inherent imperfections mean that a least-squares fit does not distinguish between errors from model inadequacy and those from the choice of θ. To address these issues, we propose solving **Problem 1** by formulating a probability distribution on θ and using the Bayesian framework to identify this distribution.

1.2.1 Bayes' theorem

Bayes' theorem is a direct result of the definition of conditional probability:

$$\rho_{\Theta,Y}(\theta, y) = \rho_{\Theta|Y}(\theta|y) \cdot \rho_Y(y) = \rho_{Y|\theta}(y|\theta) \cdot \rho_\Theta(\theta)$$

It is often formulated as:

$$\rho_{\Theta|Y}(\theta \mid y) = \frac{\rho_{Y|\theta}(y \mid \theta) \cdot \rho_{\Theta}(\theta)}{\rho_Y(y)}, \quad \text{where } \rho_Y(y) > 0, \tag{1}$$

relating the conditional distribution of Θ given Y to that of Y given Θ. In Bayesian terminology:

- $\rho_{Y|\theta}(y \mid \theta)$ is the likelihood of y given θ.
- $\rho_{\Theta}(\theta)$ is the prior distribution of θ.
- $\rho_Y(y)$ is the evidence or marginal likelihood.
- $\rho_{\Theta|Y}(\theta \mid y)$ is the posterior distribution of θ given y.

For simplicity, we neglect the independent variable x from **Problem 1**, assuming the model $f(\cdot)$ predicts a single value instead of a function.

- **Example 1:** Approximate Function Inverse

Consider a function f with a single input $\theta \in \mathbb{R}$, returning a single output $y \in \mathbb{R}$, denoted as $y = f(\theta)$. We assume there is a true but unknown $\tilde{\theta} \in \mathbb{R}$ and represent our empirical uncertainty in this value with the random variable Θ, having the probability density function $\rho_{\Theta}(\theta)$. Given Θ, our uncertainty in the true value $\tilde{y}$ is represented as $Y = f \circ \Theta$ with pdf $\rho_Y(y)$.

To apply Bayes' theorem, we first specify our uncertainty in $\tilde{\theta}$ (in the absence of observations) using Θ as our prior. This prior incorporates all existing knowledge about $\tilde{\theta}$. When a noisy measurement $\hat{y}$ of y arrives, we characterize the measurement using a statistical model. For example, with additive Gaussian noise:

$$\hat{y} = \tilde{y} + \epsilon, \epsilon \sim \mathrm{N}(0, \hat{\sigma}^2).$$

This implies the measurement is unbiased (since $\mathbb{E}_{\epsilon} = 0$, so $\mathbb{E}_{\hat{y}} = y$). Given $y = f(\tilde{\theta})$, we have:

$$\hat{y} - f(\tilde{\theta}) = \epsilon \sim \mathcal{N}(0, \hat{\sigma}^2), \Longrightarrow \hat{Y} \sim \mathcal{N}(f(\tilde{\theta}), \hat{\sigma}^2). \tag{2}$$

Since $\tilde{\theta}$ is unknown, for any given $\theta \in \mathbb{R}$:

$$\hat{Y} \mid \theta \sim \mathcal{N}(f(\theta), \hat{\sigma}^2),$$

This provides the likelihood $\rho_{Y|\theta}(y \mid \theta)$ in Bayes' theorem. The denominator $\rho_Y(y)$ serves as a normalization factor independent of θ. Thus, the posterior distribution is:

$$\rho_{\theta|y}(\theta \mid y) \propto \rho_{Y|\theta}(y \mid \theta) \cdot \rho_{\theta}(\theta),$$

These posterior answers the question: "What do we know about $\tilde{\theta}$?" For an affine function f and a Gaussian prior, the posterior is also Gaussian, with its mean and standard deviation balancing between the prior and the observed data.

It is important to recognize that the statistical model $\hat{y} = \tilde{y} + \epsilon$ represents an objective physical process (measurement noise) and is inherently "wrong" (in the sense of George Box).

The random variable ϵ represents aleatoric uncertainty, and errors in this model lead to errors in the posterior distribution. In contrast, the prior represents purely epistemic uncertainty and cannot be "wrong" in the same sense.

More concerning is the assumption made by replacing $\tilde{y}$ with $f(\tilde{\theta})$ in the model, implying that the only difference between model predictions and $\hat{y}$ is measurement noise. This assumes that f is perfect, which is never true in reality. This issue, known as *model inadequacy*, will be addressed later.

1.2.2 Parameter identification for a CFD problem

To illustrate the Bayesian framework, consider a simple CFD problem involving a lifting aerofoil measured in a wind tunnel and simultaneously simulated in CFD code. Three uncertain parameters are identified as the main sources of discrepancy between the experimental and simulation results: angle-of-attack (α), Mach number (M_∞), and Reynolds number (Re). Collectively, these parameters are represented as $\boldsymbol{\theta} = (\alpha, M_\infty, Re)$.

After consulting with experimentalists and examining the tunnel properties, we assign prior random variables to these parameters to represent experimental uncertainty. These priors are independent, as we have no reason to assume correlation between the errors in α and Re. This results in a prior distribution Θ_0 with the pdf $\rho_0(\boldsymbol{\theta}) = \rho_0(\alpha, M_\infty, Re) = \rho_0(\alpha) \cdot \rho_0(M_\infty) \cdot \rho_0(Re)$. These priors are shown as black pdfs on the left side of Figure 1.1.

Using these priors and our simulation code $f: \mathbb{R}^3 \to \mathbb{R}^2$, we predict random variables $C = (C_L, C_D)$ for the lift (C_L) and drag (C_D). This gives us the prior predictive distribution, with marginals sketched as black pdfs on the right side of Figure 1.1, though they are dependent (since higher lift generally indicates higher drag). While this step is not strictly necessary for parameter calibration, it serves as a sanity check on the simulation and data. If the observation $\hat{y}$ in the next step is highly unlikely or impossible under the prior predictive distribution, the analysis is suspect, and modelling assumptions should be revisited.

Next, a noisy measurement $\hat{C}_L$ of $\tilde{C}_L$ is taken, with noise characterized as $\epsilon \sim \mathcal{N}(0, \sigma^2)$, where σ is estimated based on the measurement setup. This is represented by the red pdf for C_L on the right side of Figure 1.1. The statistical model is then $\hat{C}_L = C_L + \epsilon = f_1(\theta) + \epsilon$, and the likelihood has the pdf:

$$\rho(c_L \mid \boldsymbol{\theta}) \propto \exp\left[-\frac{1}{2}\left[\frac{(\hat{c}_L - f_1(\boldsymbol{\theta}))^2}{\hat{\sigma}^2}\right]\right],$$

Thus, the posterior distribution of the parameters is:

$$\rho(\boldsymbol{\theta} \mid C_L = \hat{c}_L) \propto \rho(\hat{c}_L \mid \boldsymbol{\theta}) \cdot \rho_0(\boldsymbol{\theta}),$$

Under this posterior distribution, the parameters $\boldsymbol{\theta}$ become dependent (unlike in the prior), and their marginals are represented as red pdfs on the left side of Figure 1.1. Key points include:

- The angle-of-attack α is strongly informed by the lift c_L, evidenced by the significantly reduced standard deviation compared to the prior.

- The Reynolds number Re is hardly informed (assuming attached flow), so its posterior is similar to the prior, indicating little information gain.

Quantifying the standard deviation of the parameters is a key advantage over a least-squares fit, which would provide best guesses for all $\boldsymbol{\theta}$ without indicating the irrelevance of Re.

Finally, we can propagate the posteriors of the parameters to predict forces, giving us the posterior predictive distribution. This provides uncertain estimates of the unmeasured drag c_D, informed by the lift measurements c_L. The posterior predictive distribution is:

$$\rho(c_D \mid C_L = \hat{c}_L) \propto \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \rho(c_D \mid \boldsymbol{\theta}) \cdot \rho(\boldsymbol{\theta} \mid \hat{c}_L) d\boldsymbol{\theta},$$

Similarly, the distribution $\rho(c_L \mid C_L = \hat{c}_L)$ can be compared to the observation $\hat{c}_L$. If the observation is implausible under this distribution, it indicates that the model might be suspect and potentially inconsistent.

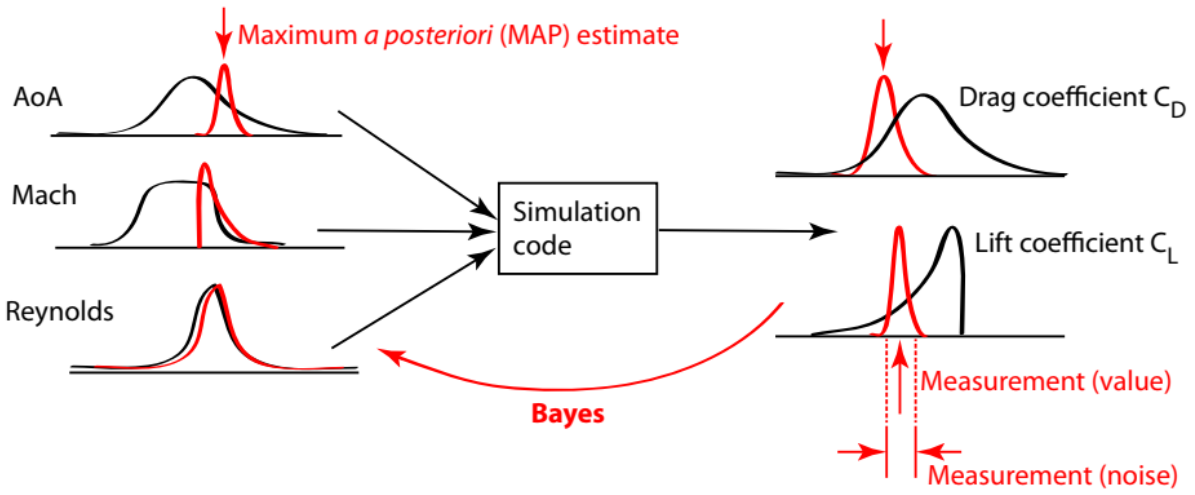


Figure 1.1 Cartoon of Bayesian updating for an aerofoil. 1) Black pdfs at the left represent priors on the simulation inputs; when propagated through the code they result in the black pdfs in the forces on the right (the "prior predictive distribution"). 2) A noisy measurement of C_L represented by the red pdf; applying Bayes gives the posterior $\rho(\boldsymbol{\theta} \mid \boldsymbol{y})$ on the inputs (red pdfs on the left). 3) Propagating this posterior through the code, gives new information about the unmeasured C_D (the "posterior predictive distribution").

1.2.3 Application to regression: Model inadequacy

In previous examples, our statistical model was of the form $y = f(\theta) + \epsilon$ with $\epsilon \sim \mathcal{N}(0, \sigma^2)$, capturing measurement error but assuming the model f is perfect. This is unrealistic, as models are inherently imperfect, affecting our posterior (Kennedy & O'Hagan, 2001).

To illustrate, consider regressing data $D \coloneqq \{(x_i, y_i) | i = 1, \ldots, N\}$ with an affine function $f(x; a, b) = ax + b$, where parameters $(a, b) =: \boldsymbol{\theta} \in \mathbb{R}^2$ must be estimated. Assuming non-informative priors $\rho(a, b) \propto 1$ and the statistical model:

$$y = f(x; a, b) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma^2), \tag{3}$$

which we call the *denial* model, the posterior is $\rho(a, b \mid D)$. Consider two possible ground truths:
A: $X \sim \mathcal{U}(0,1),\ \ Y = a^{\star}X + b^{\star} + \mathcal{N}(0, \sigma^2)$,

B: $X \sim \mathcal{U}(0,1),\ Y = a^{\star}X + b^{\star} + c^{\star}X^2 + \mathcal{N}(0, \sigma^2)$.

If A is true, the model (3) is perfect (given $a = a^*, b = b^*$), and we expect to recover estimates of a^*, b^* from the posterior. If B is true, the model is inadequate, and estimates of a^*, b^* will be influenced by $c^{\star}X^2$.

This effect is shown in Figure 1.2a and Figure 1.2b. In the former, with increasing samples, the posterior estimate (visualized by 50 samples from the posterior predictive distribution) approaches the truth, and the estimates of a, b converge to the exact values with reliable uncertainty estimates. In Figure 1.2b, model inadequacy causes a, b to converge to incorrect values, leading to false confidence in a regression model that does not correspond to the truth. This exemplifies how errors in the statistical model for the likelihood can lead to false conclusions, even with unlimited data, unlike poorly chosen priors which become irrelevant as data increases.

To address this, we explicitly account for model inadequacy by using a more general parameterization. Following Kennedy and O'Hagan (Kennedy & O'Hagan, 2001):

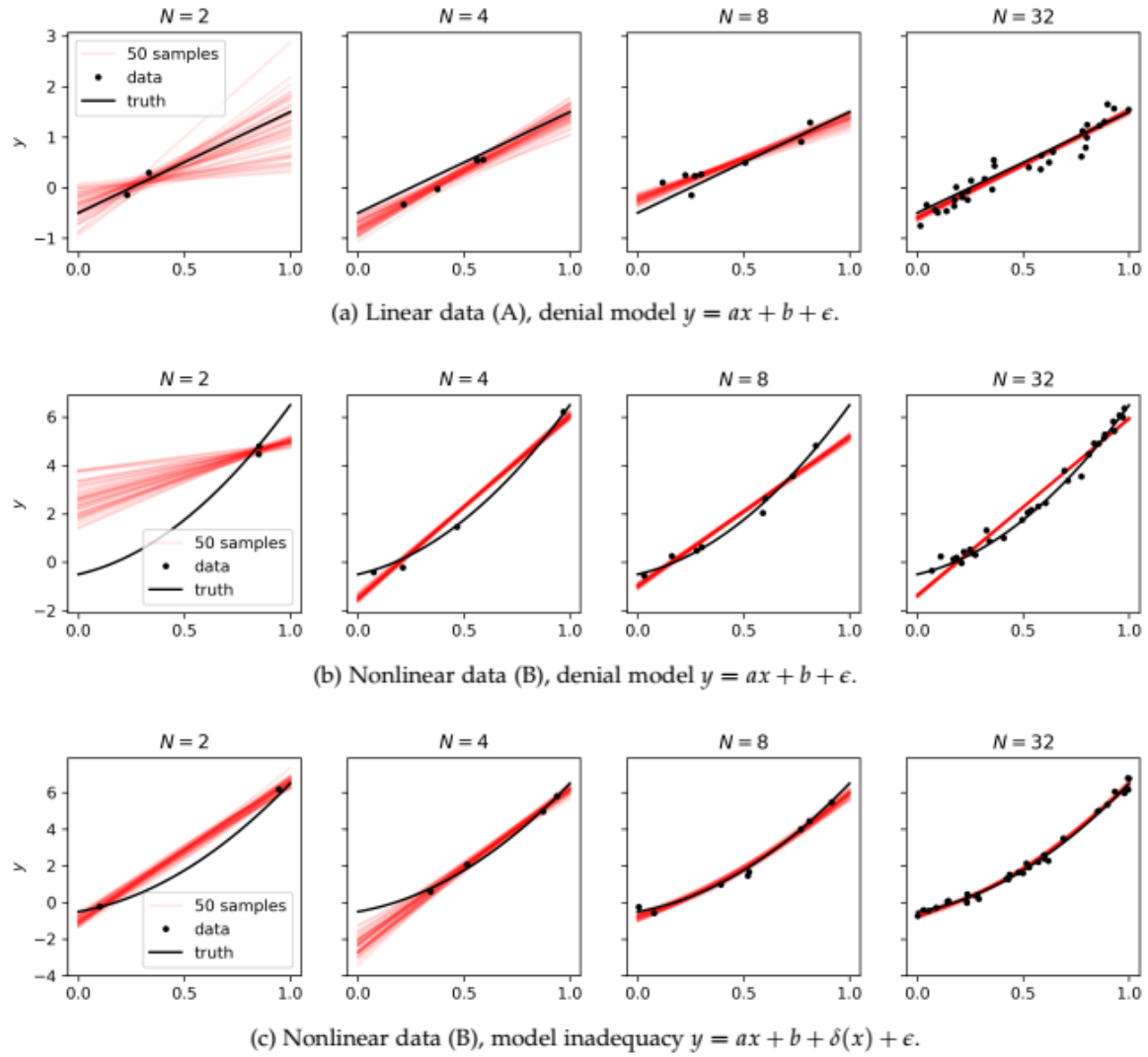
$$y = f(x; a, b) + \delta(x) + \epsilon, \qquad \delta \sim \mathcal{GP}\big(\mu_\delta(\cdot), r(\cdot,\cdot)\big), \qquad \epsilon \sim \mathcal{N}(0, \sigma^2), \tag{4}$$

where δ is a Gaussian process (GP) intended to account for systemic inadequacy in f. A Gaussian process has the following properties:

- Evaluated at any point x_i, $\delta(x_i) \sim \mathcal{N}(\mu_\delta(x_i), r(x_i, x_i))$.
- Evaluated at any two points x_i, x_j,$[\delta(x_i), \delta(x_j)] \sim \mathcal{N}([\mu_\delta(x_i), \mu_\delta(x_j)], \Sigma)$ where $[\Sigma]_{ij} = r(x_i, x_j)$.

The function μ_δ: $\mathbb{R} \to \mathbb{R}$ defines the process mean, and the covariance function r: $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defines the covariance and process variability.

With this model, the likelihood is similar to the *denial* model, but the posterior is now $\rho(a, b, \delta \mid \mathcal{D})$, using the data to estimate a,b, and δ. Applying this model to $\mathcal{D}_B$ yields Figure 1.2c, showing convergence to the true process as $N \to \infty$ with reliable uncertainty estimates. However, a,b may not converge to a^*,b^* because the model cannot distinguish between effects in f and δ. By setting the prior mean of δ to zero and a small prior variance, we specify a small model inadequacy, attributing most observed variability to f, and only the unexplained variability to δ.



(a) Linear data (A), denial model $y = ax + b + \epsilon$.

(b) Nonlinear data (B), denial model $y = ax + b + \epsilon$.

(c) Nonlinear data (B), model inadequacy $y = ax + b + \delta(x) + \epsilon$.

Figure 1.2 Regression of data with a straight-line, with and without model inadequacy.

1.2.4 Computational considerations

While we have discussed the statistical aspects of the Bayesian updating procedure, the computational aspects are also highly challenging. Bayesian methods were initially developed for models with trivial computational requirements, such as linear regression (Gelman et al., 1995), where analytic or Monte Carlo approaches are feasible. However, when dealing with CSE codes, the situation is more complex.

Unlike linear regression, CSE codes introduce a nonlinear dependence of the likelihood on the model parameters, making the likelihood and posterior distributions non-Gaussian and analytically intractable. Consequently, the posterior cannot be fully characterized by just a mean and covariance matrix.

Another challenge in evaluating posterior statistics is the high-dimensional coefficient space (e.g., six dimensions when calibrating the $k - \varepsilon$ model in the next section). The posterior pdf is likely to be nearly zero over most of this space, as poor matches between model predictions

and data result in low posterior probabilities. If the data is highly informative, the region of significantly nonzero probability may be very localized, making it challenging to find and estimate statistics in this region. This issue can be addressed by sampling from the posterior or solving optimization problems.

The three required operations are:

- Estimating the pdf or moments of $Y = f(\Theta_0)$ for a given prior Θ_0 ("uncertainty propagation");
- Obtaining samples or moments of the posterior $\rho(\theta \mid Y = \hat{y})$;
- Estimating the posterior predictive distribution $\rho(y \mid Y = \hat{y})$.

Often, approximations of these statistics are necessary for tractability- see e.g., (Edeling et al., 2018). The main computational approaches include Monte Carlo methods, surrogate modelling, and variational methods.

1.2.4.1 Monte-Carlo methods

Monte Carlo methods rely on generating random samples from a distribution to estimate statistics. For instance, to estimate the mean of $Y = f(\Theta_0)$ given samples $\theta_1, \dots, \theta_N \sim \Theta_0$, we use:

$$\mathbb{E}_\Theta Y := \int_{-\infty}^{\infty} f(\theta)\rho_0(\theta)d\theta \simeq \frac{1}{N}\sum_{i=1}^{N} f(\theta_i) + \mathcal{O}\left(\frac{\sigma_y}{\sqrt{N}}\right), \tag{5}$$

The error in this estimate is proportional to the standard deviation of Y, σ_y (Robert et al., 1999). These methods exhibit a slow but dimension-independent convergence rate of $\mathcal{O}(N^{-1/2})$. For moderate-dimensional inputs θ, Quasi-Monte-Carlo methods (Lemieux, 2009), which use low-discrepancy sequences, achieve errors of $\mathcal{O}((\log N)^M/N)$, where M is the dimension of θ. However, generating low-discrepancy sequences for dependent random variables can be challenging (Henderson et al., 2000). Another acceleration approach is Multi-level Monte Carlo (Giles, 2008), which uses a sequence of increasingly coarser grids and leverages the small variance of differences between successive levels to reduce error.

To estimate the pdf of Y, kernel-density estimation (KDE) (Scott, 2015) can be used. Starting with random samples $\theta_1, \dots, \theta_N$, the density is computed as:

$$\rho(y) \propto \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h}K\left(\frac{|y - f(\theta_i)|}{h}\right),$$

Here, h is a scale parameter (estimated by methods like Scott's rule (Scott, 2015)), and K is a positive kernel function, such as a Gaussian. This approach is similar to a generalized histogram and requires N evaluations of the CFD code, which is computationally intensive.

For obtaining samples from a distribution with a known pdf, such as the posterior $\rho(\theta \mid Y = \hat{y})$, Markov-Chain Monte Carlo (MCMC) methods are used (Brooks et al., 2011). Direct

sampling from the posterior is difficult due to the unknown cdf and the concentration of probability mass in a small region of the θ-space. MCMC constructs a continuous-state Markov chain with the target density as its stationary distribution. Sampling from this chain yields samples of the target distribution, but it is costly as each sample requires evaluation of the posterior density. Acceleration methods like Hamiltonian Monte Carlo (Brooks et al., 2011) and the No-U-Turn Sampler (NUTS) (Hoffman & Gelman, 2014) exist but still require thousands of evaluations.

Monte Carlo methods are generally impractical for Bayesian problems in expensive CFD codes and are better suited for use with cheap surrogates.

1.2.4.2 Surrogate modelling methods

For relatively low-dimensional spaces, up to about 20 dimensions, surrogate models of the CSE code response can be constructed based on evaluations of f. These surrogate models allow for the computation of statistics using Monte Carlo methods and the generation of samples using MCMC methods.

One approach is uncertainty propagation using polynomial-based methods such as Polynomial Chaos (PC) (Ghanem & Spanos, 1990) and probabilistic collocation (Loeven et al., 2007). In these methods, the model response in the stochastic space is represented by a polynomial, fitted either through Galerkin projection (in PC) or collocation. The polynomial basis is often chosen to match the input distribution (e.g., Hermite basis for Gaussian inputs), allowing the output distribution's statistics to be explicitly expressed in terms of the polynomial coefficients. These methods offer spectral convergence due to the polynomial bases but are challenging to apply to dependent distributions. Quadrature rules, such as those on simplices (Edeling et al., 2016; Witteveen et al., 2009) or sparse grids (Dwight et al., 2016), can also be used to evaluate expectations. This method is further elaborated upon in PART 2, section 2.2.6.

Another approach is Gaussian process regression (Kriging), which serves as a statistical surrogate model (Forrester et al., 2008). Gaussian processes are inherently suitable for the Bayesian framework and are widely used in the statistics community. They allow incorporating surrogate modelling error into predictions (Dwight & Han, 2009) and can utilize gradient information for improved dimensional scaling (de Baar et al., 2015). However, these methods still suffer from the curse of dimensionality, and dimensionality reduction is often necessary in practical applications (Anowar et al., 2021).

When surrogate models are used to generate samples or estimate posterior statistics, it is ideal to generate the surrogate using many code samples in regions of high posterior probability. Since these regions are not known *a priori*, adaptive sampling methods are required.

1.2.4.3 Variational methods

Variational methods overcome the curse of dimensionality by transforming the problem into an optimization task, which is then solved using gradient-based methods. The key advantages are that gradients of a loss function can be evaluated independently of the parameter-space dimension using techniques like adjoint methods (Peter & Dwight, 2010) or back-propagation,

and gradient-based optimization methods exhibit dimension-independent convergence (Nocedal & Wright, 1999). For PDE-based CSE models, this results in a PDE-constrained optimization problem (Hinze et al., 2008).

The simplest variational method is the *Maximum A Posteriori* (MAP) estimate. Given a posterior distribution $\rho(\theta \mid Y = \hat{y}) \propto \rho(\hat{y} \mid \theta) \cdot \rho_0(\theta)$, the MAP estimate is the value of θ that maximizes the posterior probability:

$$\theta_{MAP} := argmax_{\theta \in Q} \rho(\theta \mid \hat{y}) = argmax_{\theta \in Q} [log\, \rho(\hat{y} \mid \theta) + log\, \rho_0(\theta)\ ], \tag{6}$$

Working with log-probabilities avoids floating-point underflow issues. The evidence $\rho(y)$ is not needed as it is independent of θ. The derivatives $\nabla_\theta log\ \rho(\hat{y} \mid \theta)$ and $\nabla_\theta log\ \rho_0(\theta)$ are required, with the former necessitating derivatives of the CSE code and the latter being analytically derived for defined priors.

The main limitation of the MAP estimate is that it loses most of the statistical information in the posterior, such as variance and covariance of the estimated quantities. To regain some information, one can compute the Hessian of the posterior at the MAP estimate and use the implied quadratic response to fit a Gaussian, known as Laplace's approximation (Kass et al., 1991).

A more general approach is to approximate the posterior $\rho(\theta \mid \hat{y})$; with a variational distribution $q(\theta)$ with unknown hyperparameters. For a Gaussian distribution, this could be $q(\boldsymbol{\theta} \mid \boldsymbol{\mu}, \Sigma) \sim \mathcal{N}(\mu, \Sigma)$. The hyperparameters μ and Σ are estimated by minimizing the Kullback–Leibler (KL) divergence between $q(\boldsymbol{\theta} \mid \boldsymbol{\mu}, \Sigma)$ and the target distribution $\rho(\theta \mid \hat{y})$:

$$D_{KL}(q(\boldsymbol{\theta} \mid \mu, \Sigma) \mid\mid \rho(\boldsymbol{\theta} \mid \hat{y})) := \int q(\boldsymbol{\theta} \mid \mu, \Sigma)\, log \left(\frac{q(\boldsymbol{\theta} \mid \boldsymbol{\mu}, \Sigma)}{\rho(\boldsymbol{\theta} \mid \hat{y})} \right) d\boldsymbol{\theta} \tag{7}$$

$$= \mathbb{E}_{q(x \mid \mu, \Sigma)}\, log \left(\frac{q(\boldsymbol{\theta} \mid \boldsymbol{\mu}, \Sigma)}{\rho(\boldsymbol{\theta} \mid \hat{y})} \right) \tag{8}$$

The optimization problem is:

$$b_{opt}, \Sigma_{opt} := argmin_{\boldsymbol{\mu}, \Sigma} D_{KL}(q(x \mid \boldsymbol{\mu}, \Sigma) \mid\mid \rho(x \mid \hat{y})),$$

This challenge lies in estimating the integral in the KL divergence at every iteration of the optimizer. While Monte Carlo estimates can be used, introducing significant noise and increasing iteration costs, it is often preferable to maximize the evidence lower bound (ELBO), which minimizes an upper bound on the KL divergence (Kingma & Welling, 2013). This can be done using stochastic gradient ascent.

A sophisticated algorithm for this is Stein Variational Gradient Descent (Liu & Wang, 2016), which performs gradient descent on a set of particles representing independent samples of the posterior.

When gradients are unavailable, gradient-free optimizers can be used for low-dimensional parameter spaces. The Nelder-Mead simplex method (Powell, 1973) is recommended for up

to about 10 dimensions and is robust to noise. However, in this scenario, MCMC methods may be preferable.

1.3 RANS closure model coefficient

The Bayesian updating procedure from the previous section is applied to calibrating closure coefficients for RANS models. First, classical calibration methods are demonstrated in Section 1.3.1, where coefficients are identified one or two at a time using fundamental flows. Next, Section 1.3.2 presents an application of the Bayesian approach to the $k-\varepsilon$ model. Finally, Section 1.3.3 introduces Bayesian model selection and averaging, addressing the question of how to choose the best model (or an average model) from multiple candidates with different forms and coefficients for a specific problem. The solutions to these questions are formulated probabilistically.

1.3.1 Classical identification of closure coefficients

To highlight the contrast with the Bayesian approach, we present two examples of classical turbulence model calibration. Typically, this procedure involves:

- Identifying a *canonical flow* where the PDE model simplifies significantly.
- Solving the simplified system analytically.
- Fitting the solution to experimental or DNS data by tuning the closure coefficients.

A well-constructed model should consistently fit multiple fundamental flows, with the expectation that it will also capture the essential properties of more complex flows.

1.3.1.1 Canonical flow I: Decay of isotropic turbulence

Consider a flow where the mean velocity $\bar{u}$ is homogeneous (invariant to translations) and isotropic (invariant to rotations), meaning it is constant everywhere. In this scenario, the production of turbulence is zero due to the absence of velocity gradients, which can be seen from the exact turbulence kinetic energy (TKE) equation (Leschziner, 2015). Consequently, the kinetic energy in the turbulence scales $k^\star := \frac{1}{2}\overline{u_i' u_i'}$ decays. Empirical observations, supported by experiments and direct numerical simulation (DNS), indicate that this decay follows a power-law with an exponent close to 1.3 (Mohamed & Larue, 1990), i.e., $k^\star \propto t^{-1.3}$.

Now, consider applying the $k-\varepsilon$ model (Wilcox, 1998) to this flow. This model consists of two PDEs that approximate the turbulence kinetic energy $k \simeq k^\star$ and the dissipation rate $\varepsilon_0 \simeq \varepsilon^\star := \nu \overline{\frac{\partial u_i'}{\partial x_\iota}\frac{\partial u_i'}{\partial x_j}}$:

$$\frac{\partial k}{\partial t} + \overline{u_i}\frac{\partial k}{\partial x_i} = P_k - \varepsilon + \frac{\partial}{\partial x_i}\left[\left(\nu + \frac{\nu_T}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right], \tag{9a}$$

$$\frac{\partial \varepsilon}{\partial t} + \overline{u_i}\frac{\partial \varepsilon}{\partial x_i} = C_{\varepsilon 1} f_1 \frac{\varepsilon}{k} P_k - C_{\varepsilon 2} f_2 \frac{\varepsilon^2}{k} + E + \frac{\partial}{\partial x_i}\left[\left(\nu + \frac{\nu_T}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_i}\right], \tag{9b}$$

where the production term $P_k := \tau_{ij}\frac{\partial \overline{u_i}}{\partial x_j} \simeq \tau^{\star}_{ij}\frac{\partial \overline{u_i}}{\partial x_j} =: P^{\star}_k$, with a Boussinesq model for the Reynolds-stress tensor τ_{ij}:

$$\tau_{ij} := 2\nu_T S_{ij} - \frac{2}{3}k\delta_{ij}, \quad S_{ij} := \frac{1}{2}\left(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i}\right),$$

and the eddy viscosity defined as:

$$\nu_T := C_\mu f_\mu \frac{k^2}{\varepsilon},$$

where f_μ, f_1, and f_2 are blending functions, and the five closure coefficients are C_μ, $C_{\varepsilon 1}$, $C_{\varepsilon 2}$, σ_k and σ_ε.

Assuming isotropic homogeneous flow, all spatial derivatives disappear: $\frac{\partial \overline{u_i}}{\partial x_j} = 0$, $\frac{\partial k}{\partial x_j} = 0$ and $\frac{\partial \varepsilon}{\partial x_j} = 0$, as does turbulence production P_k. The equations then simplify to a system of ordinary differential equations (ODEs) with one remaining closure coefficient:

$$\frac{dk}{dt} = -\varepsilon \tag{10}$$

$$\frac{d\varepsilon}{dt} = -C_{\varepsilon 2}\frac{\varepsilon^2}{k}. \tag{11}$$

With the system of ODEs, the exact solutions are:

$$k(t) = k_0\left(\frac{t}{t_0}\right)^{-n}, \tag{12}$$

$$\varepsilon(t) = \varepsilon_0\left(\frac{t}{t_0}\right)^{-(n+1)}, \tag{13}$$

where t_0 is an arbitrary timescale, and the constraints are $\varepsilon_0 t_0 = nk_0$ and $nC_{\varepsilon 2} = n + 1$.

Since there is only one remaining parameter ($C_{\varepsilon 2}$) and the power-law solution of the $k - \varepsilon$ model matches the power-law observed in experiments, it is straightforward to fit them (e.g., using least-squares). For a decay rate $n \approx 1.3$, we get $C_{\varepsilon 2} \approx 1.77$. However, values in the range $C_{\varepsilon 2} \in [1.68, 1.92]$ are commonly used in practice (W. N. Edeling et al., 2014).

1.3.1.2 Canonical flow *II*: Free shear flows

In free-shear flows (e.g., shear layers, jets, and wakes), it is empirically observed that the production and dissipation of TKE approximately balance ($P^*_k \simeq \varepsilon_*$) in some regions of the flow. This balance is also approximately true in the law-of-the-wall region of boundary layers. The $k - \varepsilon$ model explicitly represents both production and dissipation, allowing the balance to be formulated as:

$$P_k = \nu_T \left(\frac{\partial \bar{u}}{\partial y}\right)^2 = C_\mu \frac{k^2}{\varepsilon}\left(\frac{\partial \bar{u}}{\partial y}\right)^2 \simeq \varepsilon. \tag{14}$$

Here, a coordinate system with y tangential to the shear layer is assumed, and it is assumed that the layer is thin enough to neglect other components of the velocity gradient (a modeling approximation). Using the Boussinesq model for τ_{ij}, we have $\tau^{\star}_{xy} := \overline{u'v'} \simeq \nu_T \frac{\partial \bar{u}}{\partial y}$, which together with the balance yields:

$$C_\mu = \left(\frac{\overline{u'v'}}{k}\right)^2.$$

DNS data represented in Figure 1.3 (Schlatter & Örlü, 2010) shows a plateau in the ratio $\overline{u'v'}/k$, with a value close to 0.3 (Van Der Laan et al., 2015). Based on this, $C_\mu \simeq 0.09$ is often considered reasonable.

However, the value of C_μ used in practice can vary significantly. For instance, when RANS is used for wind energy modelling (e.g., atmospheric boundary layers), C_μ is often chosen to be 0.03, irrespective of these considerations.

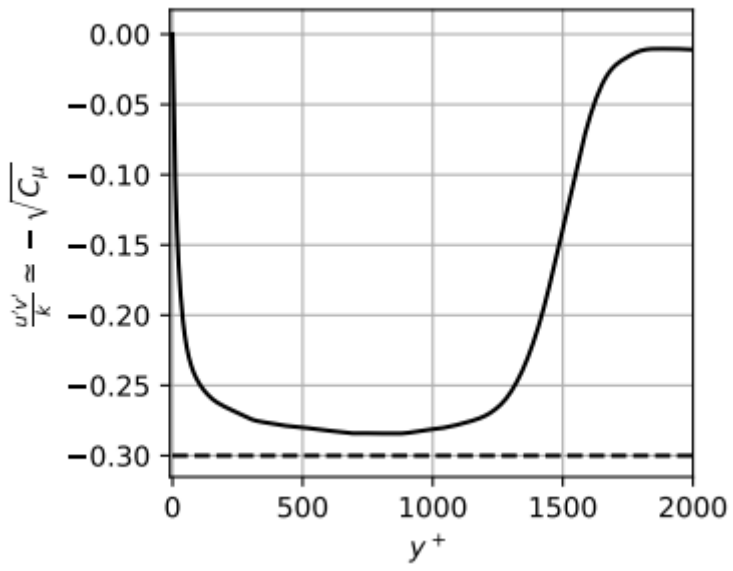


Figure 1.3 Ratio $u'v'/k$ as a function of wall-distance for a zero-pressure gradient turbulent boundary layer from DNS at $Re_\tau = 1272$, data from (Schlatter & Örlü, 2010).

1.3.1.3 Discussion

While calibrating against canonical flows has been highly successful in creating useful models, this approach does not generalize well to complex flows where many coefficients may be relevant. It also struggles with calibration against multiple, potentially incompatible datasets.

In the previous calibrations, the final step is parameter estimation, essentially a statistical process. For isotropic decay, this is relatively straightforward due to the clear power-law

(Mohamed & Larue, 1990) behaviour observed in data, allowing for accurate estimates using least-squares. However, even this problem presents complications, such as:

- Experimental data issues: noise, measurement error, and uncertain boundary conditions, especially in grid-turbulence experiments.
- DNS data issues: averaging error and finite domain size.
- Initial transients before the power-law are established, which must be excluded.
- Sensitivity of the asymptotic decay rate to Reynolds number and the initial spectrum (Yu et al., 2005).

These issues are manageable, and statistical modelling choices (e.g., data-set selection and transient removal) have a minor effect on the estimate of n. However, for free-shear flows, the variety of observed behaviour and the lack of a clear $u'v'/k$ plateau make a universal C_μ value less clear. Here, statistical modelling choices are more likely to influence results. Moreover, data suggests a range of true C_μ values, while the model specifies a single C_μ, highlighting model inadequacy.

Critically for practical engineering, these calibrations rely on simplifying equations to achieve an analytic solution, making re-calibration against practical cases difficult. It is also unclear how to calibrate when required quantities are not available in the calibration dataset (e.g., unmeasured Reynolds-stresses). Bayesian calibration approaches address these limitations and provide uncertainty information on the parameters.

1.3.2 Bayesian calibration of closure model coefficients

In this section, we will perform Bayesian calibration of the $k-\varepsilon$ model using experimental data from flat-plate boundary layers under various pressure gradients, following the methodology of W. N. Edeling et al. (2014). The pioneering work by Oliver and Moser (2011) laid the groundwork for this type of calibration.

1.3.2.1 RANS with $k-\varepsilon$ closure Model

The $k-\varepsilon$ model for a boundary layer reduces to:

$$\frac{\partial k}{\partial t}+\overline{u}\frac{\partial k}{\partial x_1}+\overline{v}\frac{\partial k}{\partial x_2}=\nu_T\left(\frac{\partial \overline{u}}{\partial x_2}\right)^2-\varepsilon+\frac{\partial}{\partial x_2}\left[\left(\nu+\frac{\nu_T}{\sigma_k}\right)\frac{\partial k}{\partial x_2}\right], \tag{15a}$$

$$\frac{\partial \varepsilon}{\partial t}+\overline{u}\frac{\partial \varepsilon}{\partial x_1}+\overline{v}\frac{\partial \varepsilon}{\partial x_2}=C_{\varepsilon 1}f_1\frac{\varepsilon}{k}\nu_T\left(\frac{\partial \overline{u}}{\partial x_2}\right)^2-C_{\varepsilon 2}f_2\frac{\varepsilon^2}{k}+E+\frac{\partial}{\partial x_2}\left[\left(\nu+\frac{\nu_T}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_2}\right] \tag{15b}$$

where the eddy viscosity is defined as (Wilcox, 1998):

$$\nu_T=C_\mu f_\mu\frac{k^2}{\varepsilon}.$$

In the $k-\varepsilon$ model, the true turbulence dissipation ε^* takes a nonzero, approximately constant value ε_0 at solid walls. The model solves for ε, the offset from this value ($\varepsilon^* \simeq \varepsilon_0 + \varepsilon$. The system (15a) - (15b) includes several closure coefficients and empirical damping functions that act directly on these coefficients. Without these damping functions, the $k-\varepsilon$ model would not accurately predict the viscous near-wall region (Wilcox, 1998). The Launder-Sharma $k-\varepsilon$ model (Launder & Sharma, 1974) specifies these damping functions as follows:

$$f_\mu = exp\left[-\frac{3.4}{1+\frac{Re_T}{50}}\right], \qquad f_1 = 1, \qquad (16)$$
$$f_2 = 1 - 0.3exp[-Re_T^2], \qquad \varepsilon_0 = 2\nu\left(\frac{\partial\sqrt{k}}{\partial x_2}\right)^2, \quad E = 2\nu\nu_T\left(\frac{\partial^2\overline{u}}{\partial x_2^2}\right)^2,$$

where $Re_T \equiv k^2/\varepsilon\nu$.

For the Launder-Sharma $k-\varepsilon$ model, the closure coefficients are:

$$C_\mu = 0.09,\ C_{\varepsilon 1} = 1.44,\ C_{\varepsilon 2} = 1.92,\ \sigma_k = 1.0,\ \sigma_\varepsilon = 1.3. \qquad (17)$$

We will obtain the joint posterior distribution for all five parameters.

1.3.2.2 Experimental data for flat-plate boundary layers

The data set comprises 13 selected experiments from the 1968 AFOSR-IFP-Stanford conference proceedings (Coles, 1969), which include various flat-plate turbulent boundary layer flows with different pressure gradients. These gradients were artificially applied by adjusting the shape of the upper wall of the test section. Mean velocity profiles were measured using hot-wire anemometry.

In the Wilcox (1993) study, the flows are qualitatively categorized as "favourable," "mildly adverse," "moderately adverse," and "strongly adverse," based on the velocity profile shapes compared to the zero-pressure gradient case. Most cases are near or in equilibrium conditions, with exceptions for some strongly adverse pressure gradients.

We will calibrate the parameters of the $k-\varepsilon$ model for each of the 13 cases individually. Despite the limited spectrum of turbulent flows represented in the dataset, significant variability in the calibrated parameter values will be observed.

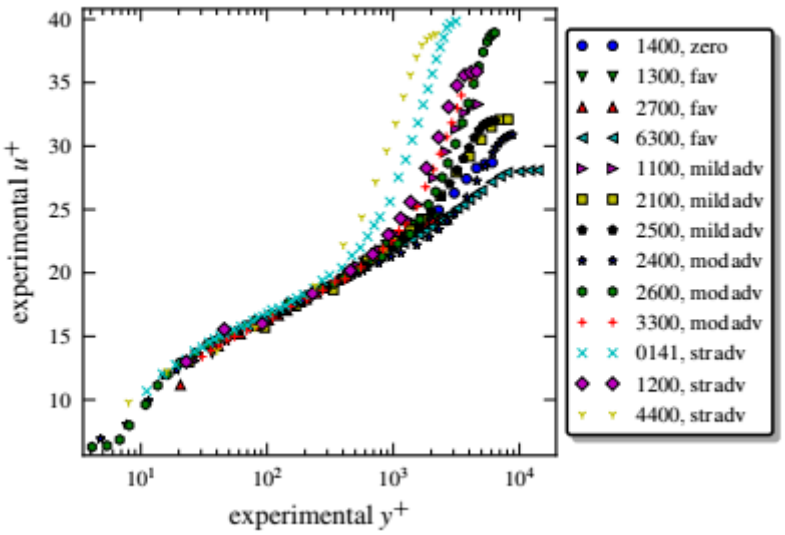


Figure 1.4 The experimental data from (Coles, 1969). Reproduced from (W. N. Edeling et al., 2014).

1.3.2.3 Likelihood and model inadequacy

We do not expect the $k-\varepsilon$ model to perfectly reproduce all boundary layer profiles, even with optimal parameter values. Therefore, similar to Section 1.2.3, we include a model inadequacy term in the statistical model. Since the velocity at the wall is zero, we use a multiplicative term to preserve this property, unlike the additive model inadequacy in Section 1.2.3. Given velocity-profile data for a single flow z, we model the data-to-simulation relationship as (W. N. Edeling et al., 2014):

$$z = \eta(\boldsymbol{y}^+) \cdot u^+(\boldsymbol{y}^+, \boldsymbol{t}; \boldsymbol{\theta}) + \epsilon \tag{18}$$

where y^+ are the wall-distances of the velocity measurements, and all quantities are in wall units. The term t refers to boundary conditions (including the imposed pressure gradient), which vary by case. We model the measurement error as:

$$\epsilon \sim \mathcal{N}(0, \sigma_z^2 I),$$

assuming any quantifiable bias has been removed from the measurements in post-processing, and that error at all measurement locations is identical and uncorrelated.

Model inadequacy is represented with a Gaussian process as in Section 1.2.3 (Kennedy & O'Hagan, 2001):

$$\eta(\boldsymbol{y}^+) \sim \mathcal{GP}(1, c_\eta)$$

with unit mean and covariance function chosen as:

$$c_\eta(y^+, y^{+'} \mid \gamma) \coloneqq \sigma^2 \, exp\left[-\left(\frac{y^+ - y^{+'}}{10^\alpha \ell}\right)^2\right],$$

where y^+ and $y^{+\prime}$are two measurement points along the velocity profile, and $\ell = 5.0$ is a user-specified length scale in wall units. The smoothness of the model inadequacy term is controlled by the correlation-length parameter α, which determines how rapidly the discrepancy between model predictions and reality can vary in y^+ units. The parameter σ represents the magnitude of the model inadequacy in velocity units. Defining the hyperparameters $\gamma := (\sigma, \alpha)$ directly would require prior knowledge of the model inadequacy, which is generally unavailable. Thus, we treat them as stochastic hyperparameters and estimate them simultaneously with the turbulence model parameters $\boldsymbol{\theta}$. The Bayesian update equation becomes:

$$\rho(\boldsymbol{\theta}, \gamma \mid \boldsymbol{z}) \propto \rho(\boldsymbol{z} \mid \boldsymbol{\theta}, \gamma) \cdot \rho_0(\boldsymbol{\theta}) \cdot \rho_0(\gamma), \tag{19}$$

requiring priors on the hyperparameters and two additional quantities to be estimated from the data, but otherwise, nothing changes.

Since $\eta(\cdot)$ is a Gaussian process, $\zeta(\boldsymbol{y}^+) := \eta(\boldsymbol{y}^+) \cdot u^+(\boldsymbol{y}^+)$ is also a Gaussian process with modified mean and covariance functions:

$$\zeta \mid \boldsymbol{\theta}, \gamma \sim \mathcal{GP}(\mu_\zeta, c_\zeta) \tag{20}$$

$$\mu_\zeta(y^+ \mid \boldsymbol{\theta}) = u^+(y^+, \boldsymbol{t}; \boldsymbol{\theta})$$

$$c_\zeta(y^+, y^{+\prime} \mid \boldsymbol{\theta}, \gamma) = u^+(y^+, \boldsymbol{t}; \boldsymbol{\theta}) \cdot c_\eta(y^+, y^{+\prime} \mid \gamma) \cdot u^+(y^{+\prime}, \boldsymbol{t}; \boldsymbol{\theta}).$$

Evaluating the Gaussian process at the observation locations $\boldsymbol{y}^+$ results in a multivariate Gaussian random variable $\boldsymbol{\zeta} \mid \boldsymbol{\theta}, \boldsymbol{\gamma}$:

$$\boldsymbol{\zeta} \mid \boldsymbol{\theta}, \boldsymbol{\gamma} \sim N\big(\boldsymbol{u}^+(\boldsymbol{\theta}), K_\zeta\big), \qquad [K_\zeta]_{ij} := u^+(y_i^+; \boldsymbol{\theta}) \cdot c_\eta(y^{+,i}, y^{+,j} | \gamma) \cdot u^+(y_j^+; \boldsymbol{\theta}),$$

i.e., all terms remain Gaussian. Consequently, the expression for the likelihood remains simple:

$$p(z \mid \boldsymbol{\theta}, \gamma) \propto \exp\left[-\frac{1}{2} d^T K^{-1} \boldsymbol{d}\right],$$

$$\boldsymbol{d} := z - u^+(\boldsymbol{y}^+; \boldsymbol{\theta})$$

$$K := \sigma_z^2 I + K_\zeta, \tag{21}$$

with $\boldsymbol{d}$ representing the data-misfit.

1.3.2.4 Priors on model coefficients and hyper-parameters

Priors for both the closure coefficients θ and hyperparameters γ are chosen as independent uniform distributions. The selection of interval endpoints is based on three factors:

- The range of coefficients recommended in the literature.

- The range of coefficients for which the solver remains stable.
- Avoidance of apparent truncation of the posterior at the edges of the prior domain.

Regarding the last factor, Bayes' rule (19) implies that values with zero probability under the prior will also have zero probability under the posterior, regardless of the data evidence. Therefore, zero priors should only be chosen where there is certainty of impossibility, such as negative values for quantities that must be positive. However, we slightly violate this principle to maintain solver stability. To ensure the prior is not unduly constraining the posterior, we later check for truncation at the prior domain edges.

As an implementation note, the value of $C_{\varepsilon 1}$ is constrained using:

$$C_{\varepsilon 1} = \frac{C_{\varepsilon 2}}{\mathcal{P}/\varepsilon} + \frac{\mathcal{P}/\varepsilon - 1}{\mathcal{P}/\varepsilon},$$

with a fixed $\frac{\mathcal{P}}{\varepsilon} = 2.09$ following Platteeuw et al. (2008). Additionally, instead of using σ_ε as a parameter, we use the Kármán constant κ and the relation:

$$\kappa^2 = \sigma_\varepsilon \sqrt{C_\mu}(C_{\varepsilon 2} - C_{\varepsilon 1}).$$

This approach imposes dependence between σ_ε and other coefficients in the prior, which is not equivalent to specifying a uniform prior on σ_ε.

The distributions for the remaining parameters are:

$$C_{\varepsilon 2} \sim \mathcal{U}(1.15,2.88), C_\mu \sim \mathcal{U}(0.054,0.135), \sigma_k \sim \mathcal{U}(0.450,1.15), \tag{22}$$

$$\kappa \sim \mathcal{U}(0.287,0.615), \sigma \sim \mathcal{U}(0,0.1), \log\ \alpha \sim \mathcal{U}(0,4), \tag{23}$$

all independently distributed.

1.3.2.5 Calibration results

The simulation code used is Wilcox's EDDYBL (Wilcox, 1998). Sampling from the posterior is performed directly with MCMC, implemented in PyMC (Abril-Pla et al., 2023). This approach is computationally feasible in this case since the simulation reduces to one spatial dimension. For more complex flows, a surrogate modelling approach would be practical due to the low dimension of the parameter space.

The posterior $\rho(\theta, \gamma \mid z)$ is a six-dimensional pdf for each of the 13 datasets. Visualization of a single pdf can be achieved by plotting all 1D and 2D marginals:

$$\rho(\boldsymbol{\theta}_i \mid \boldsymbol{z}) = \int \rho(\boldsymbol{\theta}, \gamma \mid \boldsymbol{z}) d\gamma d\theta_{\sim i}, \qquad \rho(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j \mid \boldsymbol{z}) = \int \rho(\boldsymbol{\theta}, \gamma \mid \boldsymbol{z}) d\gamma d\theta_{\sim i,j},$$

where the integrals are over indices other than ii and i,j respectively. Given samples from the posterior via MCMC, the marginal pdfs can be estimated by histograms or KDEs. A representative result is shown in Figure 1.5 for a moderately adverse pressure-gradient case. The plot shows the most likely values and variances of all parameters individually, as well as

dependencies between parameters, such as a weak negative correlation between $C_{\varepsilon 2}$ and C_{μ}. This weak dependency structure is partly due to the model parameters being designed to represent different aspects of turbulence.

From the 1D marginals, it is evident that $C_{\varepsilon 2}$ and κ are relatively well-identified from the data (small posterior variance), while C_{μ} and σ_k are less well-identified. This is only partially consistent with intuition: the Kármán constant κ refers directly to the gradient of velocity in the log-region of boundary layers and should be well-identified. However, from the discussion of canonical flows above, we would expect C_{μ} to be more important for the boundary layer than $C_{\varepsilon 2}$. The discrepancy might be related to higher model sensitivity to some parameters than others, or the different prior ranges chosen for the parameters. Nonetheless, the posterior provides specific information about the range of parameter values consistent with the data.

Comparing posteriors between cases using 1D marginals in Figure 1.6, we see significant case dependence, which is surprising given that all cases are flat-plates. Of all parameters, $C_{\varepsilon 1}$ is the best identified and most consistent, with all posteriors sharply peaked within the range of 1.3 to 1.5, followed by $C_{\varepsilon 2}$. In comparison, $C_{\varepsilon 2}$ is hardly informed beyond the prior. Most unsettling is that κ is clearly identified in most cases but with widely varying peak values, suggesting that this parameter (as applied to the $k-\varepsilon$ model) lacks universality. The hyperparameter σ is hardly identified in any case, indicating a lack of evidence about the magnitude of model inadequacy, perhaps due to the large number of parameters and limited data in each case. Conversely, the length-scale of the inadequacy is consistently estimated to be towards the upper bound.

Further analysis of these results can be found in (W. N. Edeling et al., 2014). It is clear that identifying universal parameters for $k-\varepsilon$ across the full range of flows considered is unlikely. The following section will explore Bayesian model selection and averaging as a means of addressing this issue.

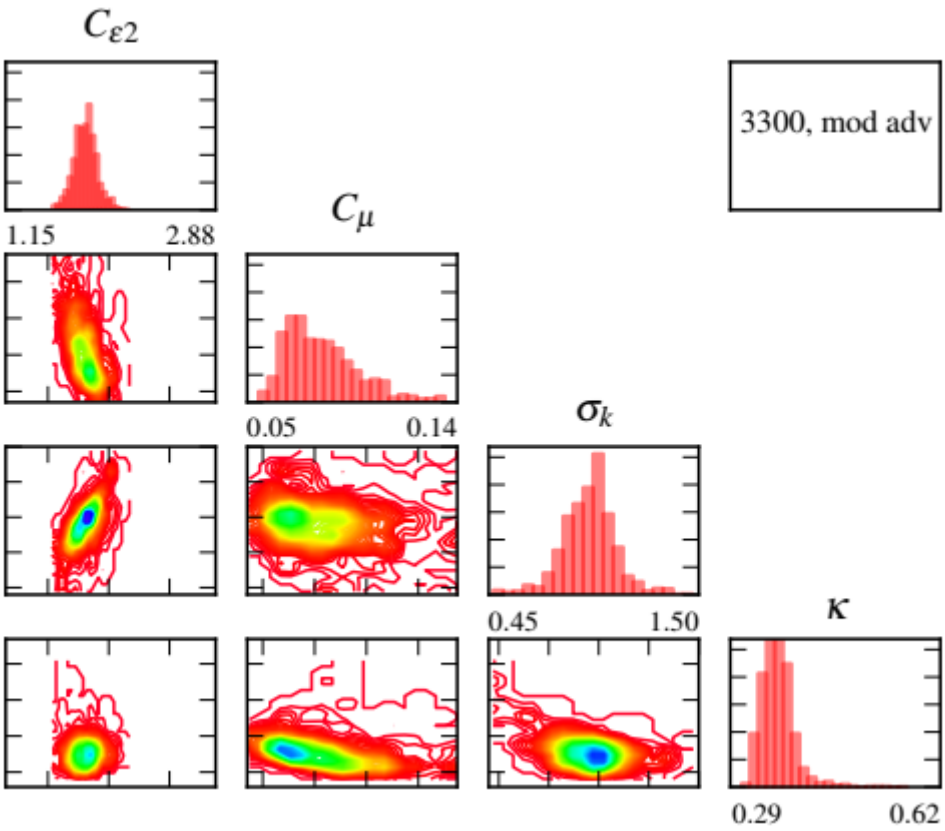


Figure 1.5 Joint posterior density on model coefficients, visualized as all pairs of 2d-marginal distributions- for a moderately adverse pressure gradient (flow 3300). Reproduced from (W. N. Edeling et al., 2014).

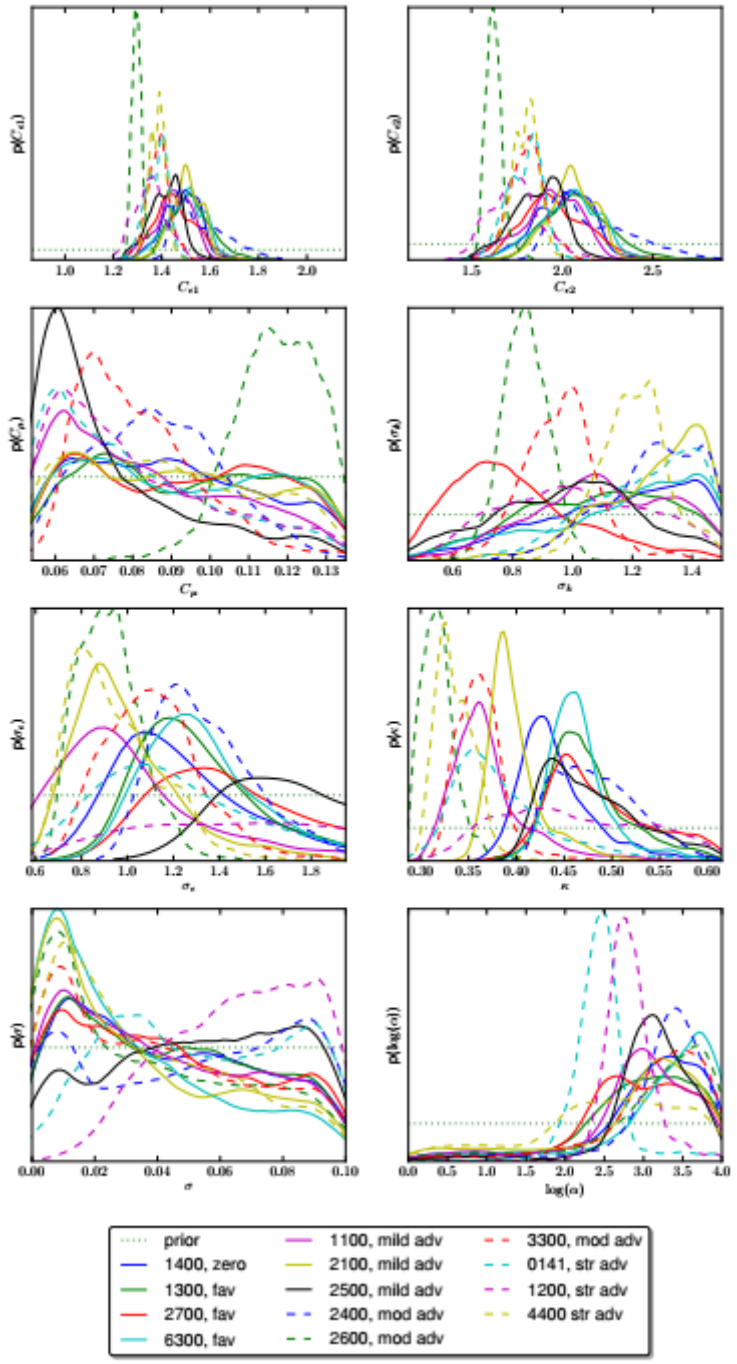


Figure 1.6 The marginal posterior distributions of the model coefficients and hyper-parameters for the 13 boundary layer cases.

1.3.3 Bayesian model selection and model averaging

Similar to other complex physical systems, there are many available models for RANS closure. Widely adopted models include $k-\varepsilon$, $k-\omega$ (Wilcox, 2008), $k-\omega$ SST (Menter, 1994), and Spalart-Allmaras (Spalart & Allmaras, 1992), all of which are linear eddy-viscosity models (LEVMs) based on the Boussinesq approximation. Each model has a unique set of coefficients,

and while some coefficients may share names across different models, their meanings and relationships to model predictions are usually distinct.

The observations from the previous section, indicating significant variability in optimal calibration coefficients from test case to test case, lead us to a natural question: Which model with which set of coefficients should we use for a particular test case? There are two scenarios to consider:

- When limited reference data is available for the specific or a similar case.
- When the flow considered is unseen, i.e., no reference data is available.

To address this problem, we use Bayesian statistics, applying Bayesian Model Selection (BMS) and Bayesian Model Averaging (BMA) to provide probabilistic answers to the question.

1.3.3.1 Theory of BMS and BMA

Both BMS and BMA are straightforward applications of the Bayesian calibration methodology. Consider a set of $Q \in \mathbb{N}$ candidate models $\boldsymbol{M} = \{M_1, M_2, \dots, M_Q\}$, each with its own vector of closure coefficients $\boldsymbol{\theta}^{(q)} \in \mathbb{R}^{N_q}$ for $q \in \{1, \dots, Q\}$, potentially with differing numbers of coefficients per model. Given data z, the posterior probability of each model conditioned on the data can be obtained using Bayes' theorem:

$$\mathbb{P}(M_q|z) = \frac{\rho(z|M_q)\mathbb{P}_0(M_q)}{\sum_{i=1}^{Q} \rho(z|M_i)\mathbb{P}_0(M_i)} \tag{24}$$

where $\mathbb{P}(\cdot)$ indicates the probability mass of a discrete random variable, and $\rho(\cdot)$ represents the probability density of a continuous random variable. Here, $\mathbb{P}_0(Mq)$ is the prior representing the subjective confidence in model M_q without considering the data, and $\rho(z \mid M_q)$ is the likelihood of observing data z given model M_q at any value of the parameters.

The likelihood $\rho(z \mid M_q)$ is defined in terms of $\rho(z \mid \boldsymbol{\theta}^{(q)}, M_q)$ as:

$$\rho(z|M_q) = \int \cdots \int \rho(z|\theta^{(q)}, M_q)\rho(\boldsymbol{\theta}^{(q)}|M_q)d\boldsymbol{\theta}^{(q)} \tag{25}$$

where the parameters of model M_q are integrated out. The term $\rho(z \mid \boldsymbol{\theta}_i, M_q)$ is the likelihood of data z conditioned on both the specific model M_q and its parameters. For example:

$$z = u^+\left(y^+; M_q, \boldsymbol{\theta}^{(q)}\right) + \epsilon$$

where the CSE code u^+ depends on the choice of model and corresponding coefficients. The term $\rho(\boldsymbol{\theta}^{(q)}|M_q)$ in (25) is the prior probability density of the parameters $\boldsymbol{\theta}^q$ conditioned on the given model M_q. This may be chosen uniformly within some interval or as the posterior from a previous calibration against a different flow and data set.

The main difference between this formulation and single-model calibration is the increased number of variables to identify (both M and $\boldsymbol{\theta}^q$ for all q), with one of these variables being

discrete. Adding more variables increases the data demands, and typically, we learn less about any given parameter with the same amount of data.

1.3.3.1.1 Bayesian Model Selection (BMS)

BMS involves computing $\mathbb{P}(M_q \mid z)$ and choosing a single model $\widehat{M} \in \boldsymbol{M}$ that satisfies:

$$\mathbb{P}(\widehat{M}|z) \geq P(M_q|z) \qquad \forall q \in \{1, \dots, Q\},$$

i.e., a maximum a posteriori estimate. The coefficients $\hat{\theta}$are chosen based on the posterior $\rho(\widehat{\boldsymbol{\theta}} \mid \boldsymbol{z}, \widehat{M})$, which can be computed from (25).

1.3.3.1.2 Bayesian Model Averaging (BMA)

BMA goes further by using a weighted linear combination of all models to predict unmeasured quantities. Let $\hat{z}$ be an unmeasured quantity related to z. For instance, $\boldsymbol{z}$ might be flow velocities measured within a limited window, and $\hat{z}$ the velocity at another unmeasured coordinate in the same flow. BMA models $\hat{z}$ as:

$$\rho(\hat{z} \mid z) = \sum_{q=1}^{Q} [\int \cdots \int \rho(\hat{z} \mid M_q, \boldsymbol{\theta}^{(q)}) \rho(\boldsymbol{\theta}^{(q)} \mid M_q, \boldsymbol{z}) d\theta^{(q)}] \cdot \mathbb{P}(M_q \mid \boldsymbol{z}), \tag{26}$$

which is the posterior predictive distribution of $\hat{z}$ given $\boldsymbol{z}$. This involves many evaluations of the CSE code. The computation can be simplified by replacing $\rho(\theta^{(q)}|M_q, z)$ with a Dirac-δ function at the MAP estimate, reducing the computational cost to Q CSE solves but neglecting parameter variability (Edeling et al., 2018).

For predicting unmeasured flows (with no z available for the new flow), (26) can be applied with model probabilities $\mathbb{P}(M_q \mid \boldsymbol{z})$and parameter distributions $\rho(\boldsymbol{\theta}^{(q)}|M_q, \boldsymbol{z})$ from a similar flow for which $\boldsymbol{z}$ is available, enabling genuine predictions with uncertainty estimates (W. Edeling et al., 2014).

1.3.3.2 Application to turbulence modelling

W. Edeling et al. (2014) applied BMS and BMA to the same set of flat-plate boundary-layer (BL) flows discussed in Section 1.3.2.2. This study considered five turbulence models ($k-\omega$, $k-\varepsilon$, the stress–ω model, Baldwin-Lomax, and Spalart-Allmaras) applied to each of the 14 BL cases. Model probabilities were computed using equation (24), with MCMC used to sample from the posteriors.

Each case is termed a *scenario*, with boundary conditions and material properties summarized by S_j for $j \in \{1, \dots, 14\}$. Model probabilities $\mathbb{P}(M_i|S_j, z_j)$are conditioned on both data and scenario and are plotted in Figure 1.7. Despite clustering the scenarios into favourable/adverse cases, no clear pattern emerges in the best choice of model per case. Often, a single model dominates, reducing other model probabilities to nearly zero. This result likely stems from all models fitting the BL well, making model probabilities sensitive to small differences.

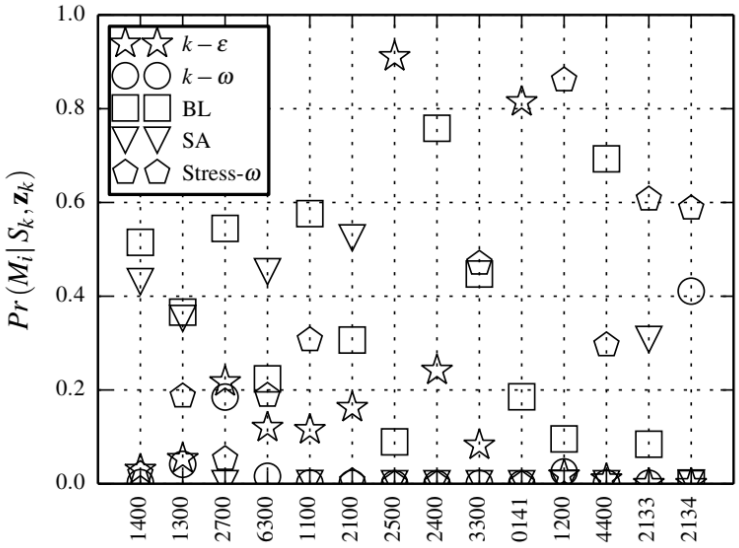
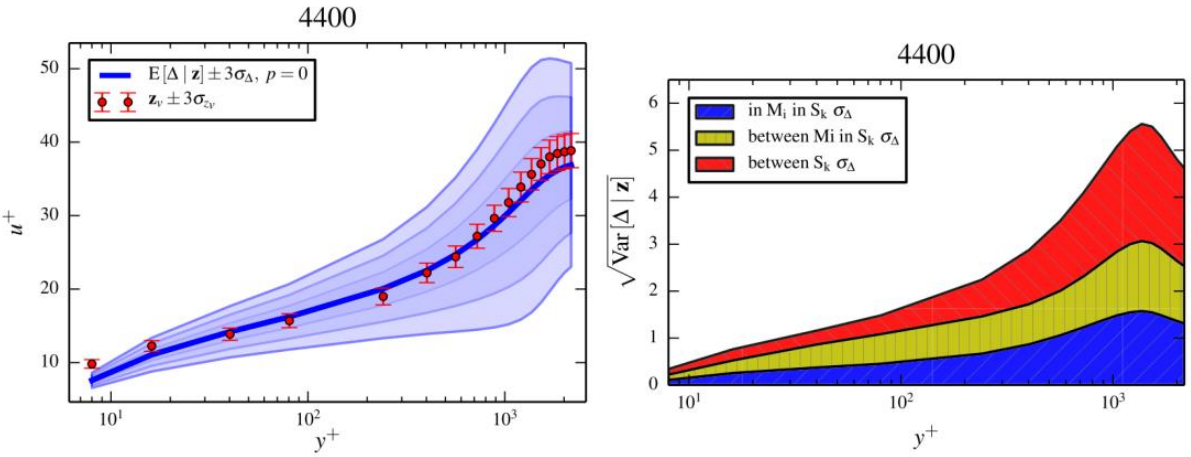


Figure 1.7 Posterior probabilities of five turbulence models for 14 calibration datasets. Figure reproduced with permission from (W. Edeling et al., 2014).

These results do not support confident selection of a single model, although this might differ with a more complex set of flows. However, predictions under uncertainty can still be made using multiple models. With five models calibrated for 14 scenarios, predictions can account for the full range of models. Figure 1.8a shows the BMA prediction for an adverse pressure-gradient case (4400) excluded from training. While the mean predictions align with experimental results, the variance is excessively large.



(a) Prediction with standard-deviation.

(b) Uncertainty breakdown.

Figure 1.8 BMA posterior predictive distribution for case 4400 (strong adverse pressure-gradient). Figure reproduced with permission from (W. Edeling et al., 2014).

Figure 1.8b decomposes this variance into three parts: (a) due to coefficient variability in a single model for a single scenario (blue), (b) due to variability between model predictions for a single scenario (yellow), and (c) due to variability in predictions between scenarios (red). All three sources contribute roughly equally to the total variance. This assumes that all training scenarios are equally relevant to the prediction scenario. Reducing uncertainty can be achieved by selecting only those training scenarios that correlate well with the prediction.

However, under the current statistical and modelling assumptions, the confidence level in RANS models remains empirically low.

1.3.4 Discussion

In this section, we have demonstrated the effectiveness of both the classical approach of selecting and fitting canonical flows and the Bayesian approach of explicit statistical modelling and uncertainty evaluation. The Bayesian method incorporates information about data accuracy and makes explicit statistical choices about the relationship between the CFD model and the data, including model inadequacy. After applying Bayes' rule, we obtained a joint posterior distribution of the model coefficients (C_μ, $C_{\varepsilon 2}$, etc.), characterizing our knowledge of these parameters given the data and our priors. This joint density was visualized via marginal distributions and propagated through the model to obtain the posterior predictive distribution, which provided uncertainties in model predictions for unseen cases.

The outlined procedure can be applied to flows of any complexity, providing a probabilistic characterization of model accuracy. The meaningfulness of this result depends significantly on the validity of the statistical modelling, both in the prior and likelihood, and particularly on the impact of neglecting model inadequacy, which can lead to serious misrepresentations of accuracy. The downside is the high computational cost, requiring many runs of the CFD solver during both the calibration and posterior predictive phases.

For a given turbulence closure model—or models in the case of BMA and BMS—this represents the best calibration approach. However, fundamental structural assumptions in RANS models, such as the Boussinesq approximation and the specific model form, are not addressed by this procedure. These issues will be the focus of the next section.

1.4 Bayesian estimation of Reynolds stress

So far, we have only adjusted model parameters (calibration), while acknowledging that the models themselves are inadequate to fully describe the physics of turbulence. Notably, all closure models discussed so far rely on the Boussinesq hypothesis, which neglects turbulence anisotropy. Consequently, even when performing model selection with Bayesian Model Selection (BMS) and averaging with Bayesian Model Averaging (BMA), we are choosing or averaging among models with a common source of inadequacy. This inadequacy is not well represented by the η term in equation (18), which treats inadequacy as a post-hoc correction to the model's predictions. In reality, the inadequacy originates within the closure model at the point where the Reynolds stress tensor (RST) is modelled. Mis-predictions of observed quantities result from this RST modelling, mediated by the transport equations.

Therefore, our goals are to (i) construct a statistical model that represents the inadequacy at its source, and (ii) generalize the model parameterization to identify corrections to the Boussinesq hypothesis. In this section, we will infer information about local turbulence anisotropy from surface pressure measurements and predict unmeasured quantities (velocity profiles) for the same flow. The main challenge is that we are now attempting to infer local quantities (spatially-varying stress corrections) instead of global coefficients. This involves inferring fields in the sense of field-inversion (Parish & Duraisamy, 2016), making it much more challenging to extract information from the posterior. Additional challenges include

constructing a prior for the tensor-valued correction to the anisotropy and the limited information that surface pressure provides about turbulence.

1.4.1 Learning Reynolds-stress anisotropy from surface pressure

In this section, we closely follow the work of Belligoli et al. (2021), focusing on data assimilation. Our goal is not to create a new or calibrated closure model but to estimate model error and unmeasured flow quantities for a specific steady-state flow with limited measurement observations. We are only concerned with the errors the baseline turbulence closure makes in this particular flow.

We assume knowledge of $\boldsymbol{z} \in \mathbb{R}^P$, experimentally obtained observations of true quantities $\boldsymbol{z}^* \in \mathbb{R}^P$, with independent Gaussian noise of known variance and zero bias:

$$\boldsymbol{z} = \boldsymbol{z}^\star + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_{exp}^2 I),$$

assuming any known bias has been removed from the data. Generalizing this noise model to correlated and spatially varying noise is straightforward (Sciacchitano, 2019). We assume the full (compressible) flow state is characterized by the variables $\boldsymbol{w} = [\boldsymbol{\rho}, \boldsymbol{\rho u}, \boldsymbol{\rho E}, \boldsymbol{k}, \boldsymbol{\omega}] \in \widehat{\mathcal{W}}$ defined on a computational mesh of size $N \in \mathbb{N}$. Given these, the observation operator $H: \widehat{\mathcal{W}} \to \mathbb{R}^P$ extracts the observed quantities by interpolation on the mesh, typically introducing negligible additional error. Our simulation's prediction of $\boldsymbol{z}^*$ is formulated as:

$$\boldsymbol{z} = H\boldsymbol{w} + \boldsymbol{\epsilon}, \ \ s.t. \quad \mathcal{R}(\boldsymbol{w}; \boldsymbol{\theta}) = 0, \tag{27}$$

where $\mathcal{R}(\cdot;\cdot) = 0$ represents the PDE constraint on the state given the model parameters θ. These parameters may include boundary conditions, material properties, etc., but we treat these as given and deterministic, using $\boldsymbol{\theta}$ to describe only the parameterization of the turbulence model.

Unlike previous sections, we now consider $\boldsymbol{\theta}$ defined at every mesh point, allowing a much more general parameterization than closure coefficients. For instance, $\boldsymbol{\theta}$ could describe the tensor-valued Reynolds stress at each mesh point, circumventing the Boussinesq assumption. In a continuous setting, the random variable describing uncertainty in $\boldsymbol{\theta}$ is a random process $\{\boldsymbol{\Theta}(x)\}_x$, which can be discretized on the computational mesh to give a high-dimensional random variable $\boldsymbol{\Theta}: \Omega \to \mathbb{R}^N$ (for a scalar process). Three particular model parameterizations are considered in the following subsections.

If θ represents Reynolds stresses, then the posterior distribution $\boldsymbol{\Theta} \mid z$ is of interest. However, our primary interest is usually in other unmeasured flow quantities $\hat{z} \in \mathbb{R}$. Our predictive model for these is:

$$\hat{z} = \widehat{H} w, \ \ s.t. \quad \mathcal{R}(w; \boldsymbol{\Theta} \mid z) = 0, \tag{28}$$

where $\widehat{H}$ extracts the desired unmeasured quantity from the state. This prediction is stochastic, with the distribution of $\hat{z}$ reflecting the uncertainty represented by the posterior $\boldsymbol{\Theta} \mid z$, which in turn represents the state of knowledge about the turbulence model.

The assumption in the statistical model (27) is that with the correct choice of $\boldsymbol{\theta} \in \mathbb{R}^{M\times N}$, z^* will be recovered exactly. Sufficient generality in the closure model parameterization is needed for this to be plausible. We do not use a Gaussian-process $\eta(\cdot)$ to represent model inadequacy as in (18) because such a term refers directly to the error made by the observation operator $\widehat{H}$, which is assumed to be negligible.

The problem formulation here is similar to the parameter calibration problem in Section 1.3.2, but with a much larger number of parameters to identify: five in the case of $k-\varepsilon$ calibration versus $\mathcal{O}(N)$ here. To find the MAP estimate, we face a high-dimensional PDE-constrained optimization problem, solvable with adjoint methods extensively developed in CFD for aerodynamic design (Peter & Dwight, 2010). However, restricting ourselves to the MAP estimate loses the uncertainty information in the posterior, which may be recovered using more general methods of variational Bayes (Nakajima et al., 2019).

We will consider three different parameterizations of Menter's $k-\omega$ SST model in the following sections.

1.4.1.1 Turbulence production perturbation (TPP)

Following Parish and Duraisamy (2016) and Singh and Duraisamy (2016), the transport equations for a general $k-\omega$ closure can be schematically represented as:

$$\frac{Dk}{Dt} = \mathcal{P}_k - \mathcal{D}_k + \mathcal{T}_k \tag{29}$$

$$\frac{D\omega}{Dt} = \theta_{TPP}(x)\mathcal{P}_\omega - \mathcal{D}_\omega + \mathcal{T}_\omega \tag{30}$$

where $\mathcal{P}$, $\mathcal{D}$, and $\mathcal{T}$ are the production, destruction, and cross-production/diffusion terms, respectively. These terms are modelled as functions of the available quantities: k, ω, and mean-velocity gradients. The spatial field $\theta_{\mathrm{TPP}}\colon \mathbb{R}^3 \to \mathbb{R}$ introduced to modify the balance between production and destruction, which can be a source of error in cases like incipiently separating flow. The modification is applied to the ω-equation because it is largely empirical, unlike the k-equation, which has a stronger theoretical basis.

As a prior, we assume:

$$\theta_{\mathrm{TPP}} \sim \mathcal{N}(1, \sigma_1^2 I),$$

so that in the absence of data, the original SST model is the most probable model, with no assumed correlation between mesh points a priori. The MAP estimate is found by solving:

$$\theta_{\mathrm{TPP}}^{\mathrm{MAP}} \coloneqq \mathrm{argmax}_\theta \rho(\,\theta \mid z\,) = \mathrm{argmin}_\theta \left[\frac{1}{2\sigma_{\mathrm{exp}}^2 P}\|Hw(\theta) - z\|^2 + \frac{1}{2\sigma_1^2 N}\|\theta - 1\|^2\right],$$

where the likelihood contributes to the data-misfit term and the prior contributes to the regularization. The relative sizes of $\sigma_1^2$and σ_{exp}^2 control the level of Tikhonov regularization.

The second equality follows from using log-likelihood and log-prior pdfs, which is necessary to avoid values of the objective function (and its gradients) that are extremely close to zero.

1.4.1.2 Anisotropy tensor perturbation (ATP)

The previous correction using TPP is still constrained by the Boussinesq hypothesis. To achieve greater generality, it is necessary to correct the model of the RST $\boldsymbol{\tau} = \tau_{ij} := \overline{u'_i u'_j}$ in the Reynolds-averaged Navier-Stokes equations:

$$\frac{DU_i}{Dt} = -\frac{1}{\rho}\nabla_i p + \nu\nabla^2 U_i - \nabla_j \tau_{ij}, \tag{31}$$

where U_i is the mean velocity and u'_i the fluctuating part. It is useful to decompose τ_{ij} into factors separating its amplitude, shape, and orientation (Eidi et al., 2021; Emory et al., 2013). The Reynolds stress tensor is a symmetric positive semi-definite tensor whose anisotropic component is:

$$\boldsymbol{b} := \frac{\boldsymbol{\tau}}{2k} - \frac{1}{3}\boldsymbol{I},$$

where k characterizes the magnitude of $\boldsymbol{\tau}$, and the symmetric tensor $\boldsymbol{b}$ is known as the normalized anisotropy tensor. A turbulence model is called realizable if its model for $\boldsymbol{\tau}$ is always positive semi-definite.

We aim to parameterize $\boldsymbol{b}$ with $\boldsymbol{\theta}$ and then use a k-equation to compute $\boldsymbol{\tau}$ in (31). Starting from the eigendecomposition:

$$b = X\Lambda X^T,$$

where $X \in \mathbb{R}^{3\times3}$ is the orthonormal matrix of eigenvectors of $\boldsymbol{b}$ and $\boldsymbol{\Lambda}$ is the diagonal matrix with real eigenvalues λ_i. By construction, $tr(\boldsymbol{b}) = 0$, so $\sum\lambda_i = 0$, meaning there are only two independent invariants of $\boldsymbol{b}$. These invariants can be represented using the barycentric map of Banerjee et al. (2007):

$$x_B = C_{1c}x_{1c} + C_{2c}x_{2c} + C_{3c}x_{3c}$$
$$y_B = C_{1c}y_{1c} + C_{2c}y_{2c} + C_{3c}y_{3c}$$

with $C_{1c} = \lambda_1 - \lambda_2$, $C_{2c} = 2(\lambda_2 - \lambda_3)$, and $C_{3c} = 3\lambda_3 + 1$ describing the character of the turbulence anisotropy (1-component, 2-component, or 3-component), and x_{ic}, y_{ic} are coordinates of the corners of a triangle in (x, y) space. Perturbations of these coordinates $(\Delta x, \Delta y)$ are our first parameters $\boldsymbol{\theta}$.

To modify the orientation of $\boldsymbol{b}$, we define a rotation Q via the unit quaternion:

$$h = h_r + h_i\boldsymbol{i} + h_j\boldsymbol{j} + h_k\boldsymbol{k}$$
$$= cos\,\frac{\phi}{2} + n_1 sin\,\frac{\phi}{2}\boldsymbol{i} + n_2 sin\,\frac{\phi}{2}\boldsymbol{j} + n_3 sin\,\frac{\phi}{2}\boldsymbol{k},$$

for a unit vector n and rotation angle ϕ, which make up the remaining components of θ:

$$\theta_{ATP} := [\Delta x, \Delta y, \phi, \boldsymbol{n}].$$

We assume the priors of each random field of θ_{ATP} are uncorrelated processes with zero mean and specified variance (e.g., $\Delta x \sim \mathcal{N}(0, \sigma_2^2)$), given which the MAP estimate is:

$$\theta_{\mathrm{ATP}}^{\mathrm{MAP}} = \mathrm{argmin}_{\theta}\left[\frac{1}{2\sigma_{\mathrm{exp}}^2 P}\|Hw(\theta) - z\|^2 + \frac{1}{2\sigma_1^2 N}\|\theta\|^2\right].$$

1.4.1.3 Random matrix perturbation (RMP)

We consider a final parameterization of the Reynolds stress tensor, addressing two deficits of the previous section: 1. The arbitrary choice of perturbation of b, and 2. The lack of guarantee that the result is realizable. Ideally, we would define a probability distribution directly on the space of admissible matrices. This is achieved with the random matrix approach (Soize, 2005; Xiao et al., 2017), which models the RST as a random matrix defined on the set $\mathbb{M}^{+0}$ of positive semidefinite matrices.

To define a distribution on this set, we use the maximum entropy principle, which provides the most noncommittal distribution (maximizing entropy) while satisfying given constraints. For the set $\mathbb{M}^{+0}$, the entropy of a pdf $\rho(\tau)$ is:

$$S(\rho) = -\int_{\mathbb{M}_0^+} \rho(\boldsymbol{\tau}) log\ \rho(\boldsymbol{\tau}) d\boldsymbol{\tau}$$

which we maximize subject to a given mean $\mathbb{E}(\boldsymbol{\tau}) = \bar{\boldsymbol{\tau}}$ and a specified dispersion δ satisfying $0 < \delta < \sqrt{2}/2$, which is analogous to a measure of variance (Soize, 2005).

In practice, it is convenient to work with a normalized positive definite random matrix whose mean is the identity, i.e., $\mathbb{E}\boldsymbol{G} = I$. We write $\tau = L_\tau^T \boldsymbol{G} L_\tau$ where $\bar{\tau} = L_\tau^T \boldsymbol{G} L_\tau$ is the Cholesky factorization of $\bar{\boldsymbol{\tau}}$. Enforcing the maximum entropy principle on $\boldsymbol{G}$ gives $\boldsymbol{G} = \boldsymbol{L}^T \boldsymbol{L}$ with $\boldsymbol{L}$ upper-triangular and:

- The off-diagonal elements of $\boldsymbol{L}$ are:

$$\boldsymbol{L}_{ij} = \frac{\delta}{2} w_{ij}, \qquad w_{ij} \sim \mathcal{N}(0,1), \qquad \forall i \neq j.$$

- The diagonal elements of L are:

$$\boldsymbol{L}_{ii} = \frac{\delta}{2}\sqrt{2u_i}, \qquad u_i \sim \Gamma(\alpha_i, 1),$$

where u_i are gamma-distributed, with pdfs:

$$\rho(u_i) = \mathbb{1}_{\mathbb{R}^+}(u_i)\frac{1}{\Gamma(\alpha_i)} u^{\alpha_i - 1} e^{-u}, \text{for } i \in \{1,2,3\},$$

where $\alpha_i := \frac{d+1}{\delta^2} + \frac{1-i}{2}$, and $\mathbb{1}_{\mathbb{R}^+}(u) = 1$ if $u \geq 0$ and 0 otherwise.

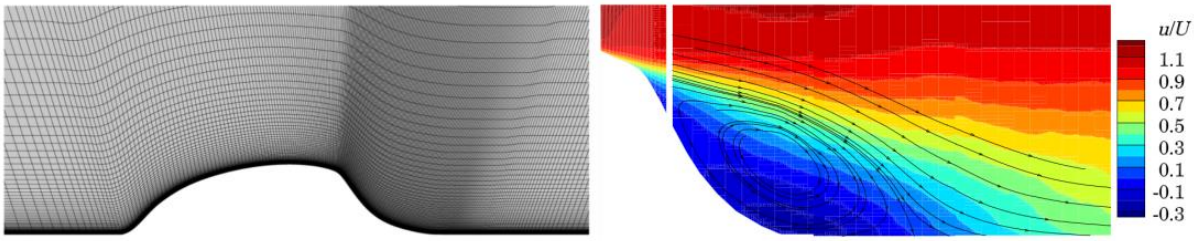
Thus, the parameters of our Bayesian calibration problem are $\theta_{\mathrm{RMP}} = [\boldsymbol{L}_{11}, \boldsymbol{L}_{12}, \boldsymbol{L}_{13}, \boldsymbol{L}_{22}, \boldsymbol{L}_{23}, \boldsymbol{L}_{33}]$, the six components of $\boldsymbol{L}$, with the prior given by the maximum entropy distribution with mean $\bar{\tau}$ taken as the Boussinesq prediction of τ at each mesh point, and user-specified dispersion δ. The MAP estimate is obtained as $\theta_{\mathrm{RMP}}^{\mathrm{MAP}} = \mathrm{argmax}\, \mathcal{J}(\theta)$ with:

$$\mathcal{J}(\theta) = \frac{1}{2\sigma_{\exp}^2 P}\|Hw(\theta) - \boldsymbol{z}\|^2 + \sum_{j=1}^{N}[\frac{\boldsymbol{L}_{11}^2 + \boldsymbol{L}_{12}^2 + \boldsymbol{L}_{13}^2 + \boldsymbol{L}_{22}^2 + \boldsymbol{L}_{23}^2 + \boldsymbol{L}_{33}^2}{2\sigma_\delta^2}$$
$$- \log\left(\mathbb{1}_{\mathbb{R}^+}(\boldsymbol{L}_{11,j})\boldsymbol{L}_{11,j}^{s_1-1} \cdot \mathbb{1}_{\mathbb{R}^+}(\boldsymbol{L}_{22,j})\boldsymbol{L}_{22,j}^{s_2-1} \cdot \mathbb{1}_{\mathbb{R}^+}(\boldsymbol{L}_{33,j})\boldsymbol{L}_{33,j}^{s_3-1}\right),$$

where $s_i = \frac{4}{\delta^2} + 1 - i$. Note that the presence of the gamma distributions in the prior results in a unique form of the regularization term, particularly penalizing non-realizable tensors with $\log\ 0 = -\infty$

1.4.1.4 Test case: Separated flow behind a hump (2d)

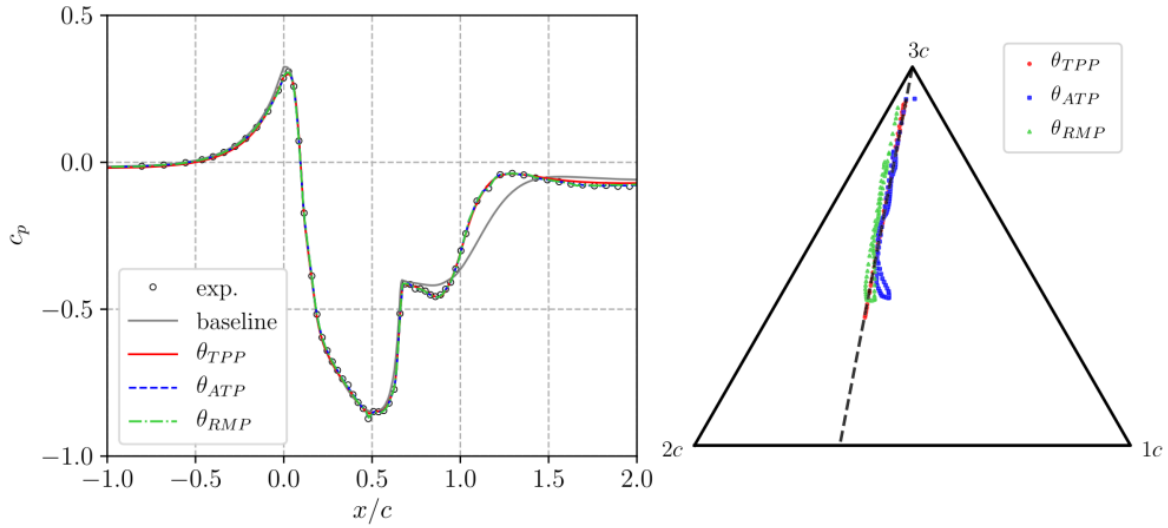
To illustrate the above approach, we consider separated flow behind the NASA hump with a chord-based Reynolds number of 9.36×10^5 (Greenblatt et al., 2006). The wall-resolved mesh is depicted in Figure 1.9a, and the experimental pressure coefficients on the surface of the hump are used as training data (Figure 1.10a). The value of σ_{exp} is set to 10^{-3} as specified in the experiment by Greenblatt et al. (Greenblatt et al., 2006) PIV measurements and hotwire velocity profile measurements are also available (Figure 1.9b), and will be held back as test data.



(a) NASA Hump mesh (b) PIV data

Figure 1.9 Computation mesh and Particle Image Velocimetry (PIV) data for the NASA hump flow.

RANS estimates of the surface pressure coefficient over the hump are shown in Figure 1.10a. The baseline SST model overpredicts the pressure trough at $x/c \approx 1.3$ (where $x/c = 0$ corresponds to the hump's leading edge). All three corrected models effectively reproduce the training data, capturing subtle features such as the pressure kinks at $x/c = 0.5$ and $x/c = 1.2$ which are likely experimental artifacts. This suggests that all three parameterizations are sufficiently general, including the simplest TTP, which just adjusts turbulence dissipation locally.



(a) Wall pressure coefficient

(b) Turbulence anisotropy at $x/c = 0.8$

Figure 1.10 Pressure measurements (circles) and RANS predictions, for the baseline (SST) model, and the corrected models (left). Turbulence anisotropy along a line in the wake for the corrected models.

The corrective fields resulting from ATP and RMP are visualized in Figure 1.11, where both corrections are plotted in terms of eigenvalue shifts in the barycentric map for comparison purposes. Corrections in the orientation of $\boldsymbol{b}$ were negligible in both ATP and RMP and are not plotted. In this figure, data from some velocity profiles in the wake are used in addition to surface pressure due to the extremely small corrections using surface pressure alone. Corrections to turbulence anisotropy in the separated shear layer are most visible both immediately after separation and during the wake recovery phase. Visualizing the anisotropy correction in the barycentric map (Figure 1.10b), and assimilating pressure-data only, shows that the corrective models give Reynolds stress anisotropy predictions very close to the baseline SST prediction, with minor modifications. This demonstrates the sensitivity of the mean flow to small perturbations of the RST, especially at high Reynolds numbers. Consequently, in none of the scenarios considered is specific information about the RST obtained, and the calibrated parameters θ act only as model coefficients, not physically meaningful quantities.

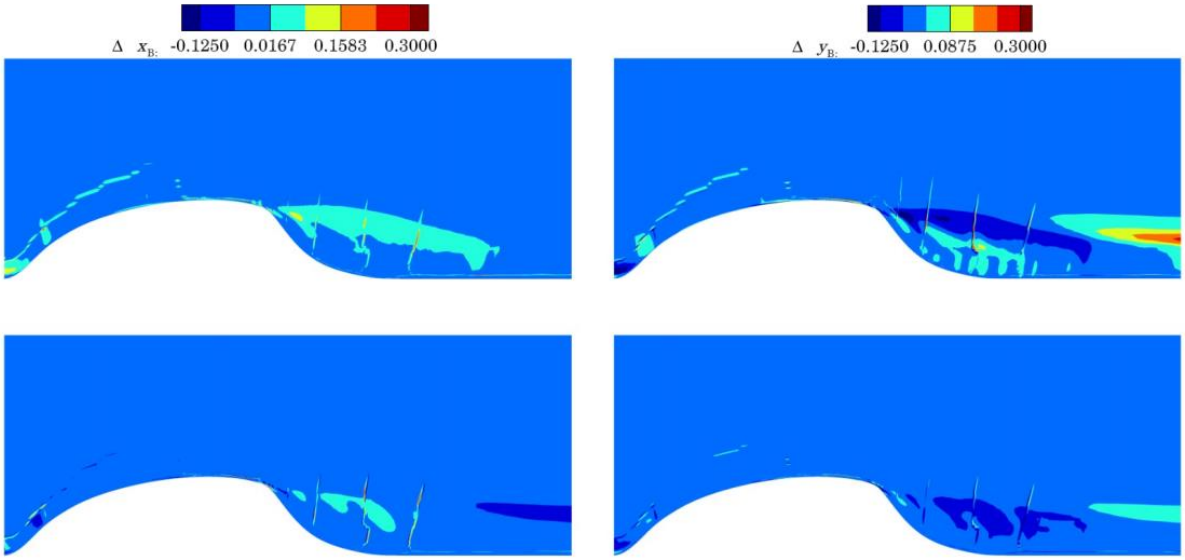


Figure 1.11 Perturbations to barycentric coordinates Δx, Δy for the ATP (top) and RMP (bottom) models.

The value of the method is demonstrated in the predictions of unmeasured quantities. In Figure 1.12, we examine predicted velocity profiles (no velocity data was used for training). In all cases, the velocity profiles are significantly improved compared to the baseline SST model, and the choice of parameterization (TPP, ATP, or RMP) makes little difference. These results suggest that the pressure distribution provides relatively little direct information about the turbulence but does indicate separation and reattachment locations. These quantities have been matched by the calibration with appropriate choices of θ, improving velocity profiles by accurately predicting the separation bubble extent.

Ultimately, this exercise demonstrates the importance of having access to measurement data relevant to (or correlated with) the quantities one wishes to estimate. Estimating turbulence characteristics purely from pressure data is likely not feasible, regardless of the model parameterization framework. However, deriving a better model that more accurately predicts mean-flow quantities may be feasible.

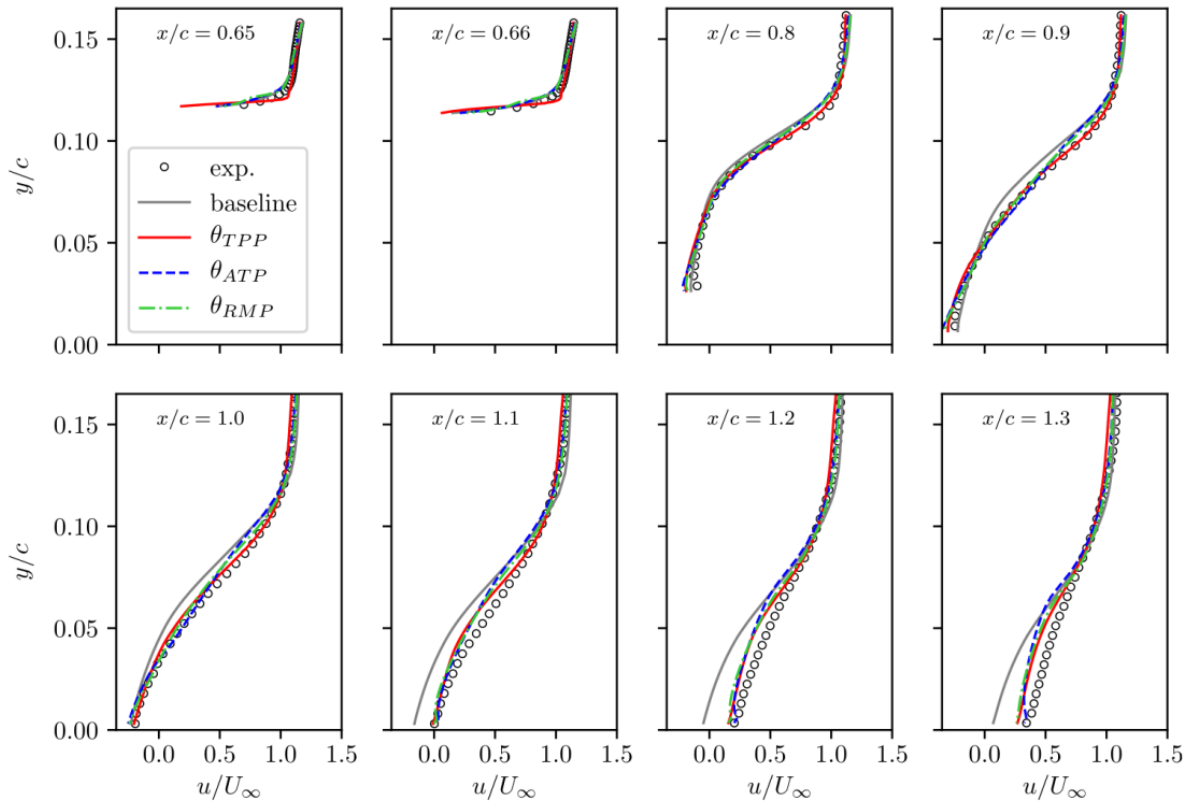


Figure 1.12 Velocity profiles at various x-locations for the corrected models compared with experimental data.

1.5 Applications in the GO-VIKING project

In the rapidly advancing field of nuclear energy, ensuring the safety and efficiency of nuclear reactors is crucial, especially given the increasing global energy demands and rising temperatures. A critical component of nuclear safety is the integrity of fuel rods, which can be compromised by turbulence-induced vibrations (TIV) caused by the axial flow of coolant liquids. TIV can lead to significant structural damage, including fatigue and wear, posing serious risks to reactor safety.

Despite the longstanding attention this phenomenon has received since the 1950s, traditional semi-empirical methods provide limited predictive accuracy and are not effective for multi-rod scenarios. Recent advancements have focused on Fluid-structure interaction (FSI) simulations as a robust tool for studying fuel rods' behaviour under TIV effects, utilizing the increased computational power available today. While DNS and large eddy simulation (LES) offer greater accuracy, their high computational demands make them impractical for complex FSI studies. Consequently, unsteady RANS (URANS) simulations are often preferred, despite their tendency to underpredict vibration displacement amplitudes.

Although URANS simulations offer promising avenues for faster, more reliable, and safer studies of physical phenomena, these simulations must be thoroughly validated to ensure their accuracy and reliability. By using advanced UQ techniques, quantifying the uncertainties in simulation parameters and propagating them through the models leads to more accurate and reliable predictions. This approach will help bridge the gap between idealized numerical

models and the complex, real-world behaviour of fuel assemblies, leading to safer and more efficient designs in nuclear power generation.

As we have discussed in the previous sections, we are going to elaborate on UQ in FSI simulations from a turbulence point of view. Our approach involves applying UQ methods to turbulence simulation to account for model inadequacy and improve the reliability of TIV predictions. By using Bayesian calibration techniques, we aim to quantify the uncertainties in RANS model parameters and propagate these uncertainties through the simulation to obtain probabilistic predictions of quantities of interest. This will enable us to provide more accurate and reliable predictions of FIV in fuel rods, considering both the uncertainties in the model constants and potential model form.

In our future work, we will study both fluid-only cases and FSI cases, following previous studies in the GO-VIKING project where experimental data are available for validation. We will conduct sensitivity analyses and UQ for turbulence model closures in the numerical simulations. This comprehensive approach will ensure that our simulations are robust, accurate, and validated against experimental observations, further advancing the safety and efficiency of nuclear reactors.

1.6 Conclusion

In this part, we have explored the Bayesian framework as it pertains to calibration and inverse problems in the context of CSE codes, particularly RANS turbulence modelling. The major value of this framework lies in its ability to precisely characterize relationships between arbitrary inputs, outputs, and state variables of complex computational models. For instance, the framework easily extends to multiple models (as seen in Bayesian Model Selection/Averaging in Section 1.3.3) and time-dependent problems. We have highlighted the risks of model inadequacy, how they can lead to erroneous conclusions, and how these risks can be mitigated through statistical modelling (Section 1.2.3).

The power of the Bayesian framework comes from representing relationships between variables in terms of joint probability density functions (pdfs). However, this generality comes with significant computational costs. Managing these costs, especially for expensive CSE codes and high-dimensional parameter spaces (e.g., FSI simulations), is the major challenge in applying Bayesian methods. We have discussed various approaches to handle this, notably the large class of variational methods (Section 1.2.4.3). Often, a statistical simplification is necessary, which may result in discarding some or all uncertainty information. Despite this, we see Bayes as an unusually effective means of reasoning about CSE systems, making otherwise arbitrary modelling choices explicit and well-motivated. The application to RANS modelling is timely, given the current interest in data-driven modelling (Ling et al., 2016; Parish & Duraisamy, 2016; Schmelzer et al., 2020; Steiner et al., 2022; Wu et al., 2018). We believe the best approach to these questions is necessarily statistical.

2. PART 2: Application of non-intrusive uncertainty quantification techniques to evaluate the effect of stochastic deformation on the vibration characteristics of a tube bundle in axial flow

2.1 Introduction

The accurate prediction of flow-induced vibrations (FIV) of cylindrical structures in axial flow is an important aspect of design evaluation of nuclear power plants. Misunderstood FIV can lead to unexpected failure (Banyay et al., 2019; Ferrari et al., 2018) due to damage mechanisms like fretting (Christon et al., 2016), posing a risk regarding safety and economic operation.

Computational Fluid Dynamics (CFD) methods allow prediction of thermal hydraulic performance, and FIV can be investigated by coupling flow simulations to a Computational Structural Mechanics (CSM) model in a fluid-structure interaction (FSI) simulation. Notable examples of research predicting vibration characteristics of a tube subject to axial flow is the pioneering work of De Ridder et al. (2013) who succeeded in capturing added-mass and -damping effects by exciting a free vibration and fitting the response to a characteristic vibration equation. Similar simulation strategies were explored by ter Hofstede et al. (2017) and De Santis and Shams (2017). The works mentioned in this paragraph can be considered as the starting point of current study.

So far, most numerical research limited itself to perfect geometries and operating conditions as they are designed, while in reality these may differ from the drawing table. Fuel assemblies are known to deform throughout their residence in the reactor (Pettigrew et al., 1998), mainly due to mechanisms as irradiation swell and thermal creep. In addition, also pre-stress can be introduced during manufacturing, for example in the case of wire-wrapped fuel assemblies (Roelofs et al., 2019). In order to meet the promise of being able to predict FIV in a realistic reactor environment, further developments are needed (Roelofs, 2020; Roelofs et al., 2019).

Peculiar about deformations is that they are (quantitatively) rather unpredictable (Wanninger, 2018) due to the complex environment and a multiplicity of causes and influences (Boltax et al., 1978; de Lambert et al., 2021; Leturcq et al., 2022; Roelofs et al., 2019; Sosnovsky et al., 2015), such as spatial distribution of coolant temperature and irradiation dose, material properties, fuel residence time and power cycle history (Bottcher & Hofman, 1977; Sakai et al., 1978). Especially deflection amplitude and direction were found to be sensitive to uncertain parameters (Wanninger, 2018). Therefore, Wanninger et al. (2018) argues that taking into account uncertainties is important for fuel assembly bow predictions. Multiple authors (Banyay et al., 2019, 2020; Liu et al., 2020) identify the presence of many aleatoric and epistemic uncertainties in nuclear power plants and stress the importance of uncertainty quantification (UQ) for the design and operation of reactor components. The Nuclear Energy Agency assigns high priority to a consolidated UQ method to use CFD in a first step safety analysis (NEA, 2016). Because of this, a stochastic approach to capture the effect of deformations on the vibration behaviour is proposed here, for which a UQ method is used. A

random fuel pin deformation is assumed, expressed as a deflection amplitude and direction, and this randomness is propagated through an FSI model to output quantities of interest (QoI). These are in this case the vibration characteristics, i.e. frequency and damping ratio of the first mode.

In addition to introducing uncertainties and changes to neutronics (de Lambert et al., 2019) and heat transfer, fuel assembly bow correlates with fretting damage as observed in PWR (Christon et al., 2016; Wanninger, 2018), although moderately (Rubiolo & Young, 2009). Fretting is one of the most important consequences of flow-induced vibrations, but its characterization is beyond the scope of current work, given the substantial computational cost that is associated with obtaining the turbulent forcing function (Dolfen, De Ridder, et al., 2022). This work focuses only on the vibration characteristics quantified as frequency and damping ratio, which are affected by the presence of the fluid due to added-mass and -damping effects. Determining those is also one of the steps in the multi-step methodology to predict fretting due to FIV formulated by Dolfen, De Ridder, et al. (2022). According to Wanninger et al. (2018) bow may result in different vibration characteristics. This further motivates the investigation of the influence of bow on FIV in this work.

Next to increasing confidence in simulation results of a realistic reactor environment, also a second reason to resort to UQ methods can be given. Geometrical or operational uncertainties are present in experiments (Roelofs et al., 2019) and can convolute numerical validations, as was suggested by Dolfen et al. (2019) and Bertocchi et al. (2020). This is especially the case because FIV experiments are typically complicated to construct. Taking into account these uncertainties and propagating them through the numerical model can help the comparison with experimental results by predicting the amount of deviation that is to be expected. Liu et al. (2020) managed to establish tolerance limits for experimental results of a natural convection lead-bismuth eutectic (LBE) test-loop with their coupled CFD-system code framework by using a UQ method.

The fuel bundle assumed for this research is loosely based on the geometry of MYRRHA (Abderrahim, 2020), a liquid metal fast reactor prototype design using LBE as coolant. The wire-wrapped fuel assemblies of this kind of reactors are found to exhibit a spiralling undulation of which the pitch is comparable to the helical pitch of the wire (Katsuyama et al., 2007; Sosnovsky et al., 2015; Uwaba et al., 2017). In the scope of this research, it is chosen to simplify the geometry by omitting the wire and assuming a C-shape bow with random amplitude and direction as stated above. C-shape bow of a single rod is not likely to occur in reality, but this simplification is assumed nevertheless to limit complexity induced in the meshing, as the meshes need to be able to absorb the parameterized deformation and additionally to deform while simulating the vibration. This is deemed acceptable, as the objective is here to develop a methodology in the first place. Robust mesh deformation is a research topic on its own, which could be included in future work. Promising results have already been obtained with the overset technique (Dolfen, Van Hauwermeiren, et al., 2022).

A straightforward UQ method is the crude Monte Carlo method. It relies on sampling the stochastic input and computing the stochastic output quantities of interest using numerical models, after which the created data can be used to compile descriptive statistics. The method is very robust since it only requires the individual sample points to be computable. However, a large number of evaluations is needed and the associated computational cost is

consequently high (Geffray & Macián-Juan, 2015; Liu et al., 2020). This is a general challenge for uncertainty quantification, especially when using expensive computational models. Reducing computational cost is a research topic on its own, so here another existing methodology is selected and tested for the FSI case with stochastic bow under investigation.

The generalized Polynomial Chaos (gPC) method is chosen, since it is seen as a highly mature method for *uncertainty propagation* (NEA, 2016). It involves Polynomial Chaos Expansions (PCE), based upon which methods have been introduced in the field of computational physics by Ghanem and Spanos (1991). PCE relies on expanding random processes in terms of Hermite polynomials. Later it was enhanced by generalizing them for other types of polynomials/random variables by Xiu and Karniadakis (2002) in the gPC method. The latter form is well-established and selected for this work. This method is briefly explained in Section 2.2.6.

The objective of this work is to use the gPC method on a simple FIV case, being a flexible rod with stochastic stress-free/permanent deformation surrounded by a rigid hexagonal bundle, to explore its capabilities for the experimental cases within the GO-VIKING consortium, such as AMOVI and TITAN. To this end the feasibility and applicability of the gPC method for the FIV case of interest is first tested on a similar but steady CFD case. Using steady CFD reduces the computational cost, allowing to create a Monte Carlo (MC) benchmark to validate the PCE method against. In the FSI case only this validated Polynomial Chaos (PC) code/framework is used, to predict the stochastic vibration characteristics affected by the random bow.

The ability to assign uncertainty ranges to relevant output quantities of interest, e.g. vibration frequency, damping and amplitudes aligns well with the GO-VIKING objectives, where state-of-the-art numerical simulations are compared to different FIV experiments. This allows a fairer judgement of the uncertainties in numerical predictions.

In Section 2.2 the research methodology and the different employed methods are explained and the details of the case under investigation are described. The results are discussed in Section 2.3.

The simulation data and pre- and post-processing scripts are available for download on the Zenodo platform (Dolfen et al., 2023).

2.2 Methodology

The geometry is schematically shown in Figure 2.1. The nominal central rod geometry is given an uncertain bow deformation (C-shape), represented by two random variables θ and e_0, which will be described in Section 2.2.1. The goal is to obtain the response of the natural frequency and damping ratio of the first flexural mode to this stochastic input.

As an accurate prediction of frequency and damping ratio requires a computationally expensive two-way coupled FSI simulation, using a crude Monte Carlo method is computationally not feasible. Therefore, the generalized Polynomial Chaos method is chosen for the FSI simulations, which is more efficient for low-dimensional stochastic input and smooth input-output behaviour. Its usage is first validated for different output quantities of interest extracted using cheaper steady CFD simulations, allowing usage of the Monte Carlo

method to serve as a benchmark. As most of the occurring flow physics are retained, this validation is assumed to be representative for the FSI problem too. Subsequently, the Polynomial Chaos method is applied on the two-way coupled unsteady FSI simulation used to extract the modal characteristics.

In Section 2.2.1, details about the parameterization of the stochastic deformation are provided. Subsequently the CFD (Section 2.2.2) and CSM (Section 2.2.3) model are discussed, as well as the FSI coupling and the procedure to extract the modal characteristics (Section 2.2.4). Finally, the Monte Carlo (Section 2.2.5) and Polynomial Chaos (Section 2.2.6) methodology are explained.

2.2.1 Geometry parameterization

A C-shape deformation of the centreline is assumed, following Eq. (1), which is a general expression for the first flexural mode of a clamped-clamped beam (Kelly, 2000), modelling the fuel rod bow. It is important to note that, while being modelled using the flexural mode shape, the bow represents the equilibrium configuration merely caused by thermal creep and irradiation creep, but not a vibration. Bow is a quasi-static phenomenon, which slowly evolves over time. As it exhibits much slower dynamics than the time scales of the flow-induced vibrations, it is assumed as a constant deformation for the purpose of this work. Its influence on the high-frequency vibrations is of interest, which it exerts via its effect on the flow patterns.

$$e(z) = C[cosh(\kappa z) - cos(\kappa z) - \frac{cosh(\kappa L) - cos(\kappa L)}{sinh(\kappa L) - sin(\kappa L)}(sinh(\kappa z) - sin(\kappa z)], \qquad (1)$$

with κ the solution of:

$$cos(\kappa L) - \frac{1}{cosh(\kappa L)} = 0. \qquad (2)$$

L represents the length of the tube and z the axial co-ordinate, while C is a scaling constant. Two independently uniformly distributed uncertain parameters are defined, θ and e_0, both schematically depicted in Figure 2.1. First, the deformation is assumed to occur in a plane containing the z-axis, called the ez-plane, of which the angle with the xz-plane is defined as the random variable θ. Second, the extent[1] of deformation is defined as the random variable e_0, i.e. the maximal deviation from the nominal (straight) centreline. In Eq. (1), C is chosen such that $\max(e(z)) = e_0$.

[1] It is chosen to call the bow amplitude 'extent' in this deliverable, to avoid possible confusion with the amplitude of the vibration of the tube, which is superposed on the stress-free bow.

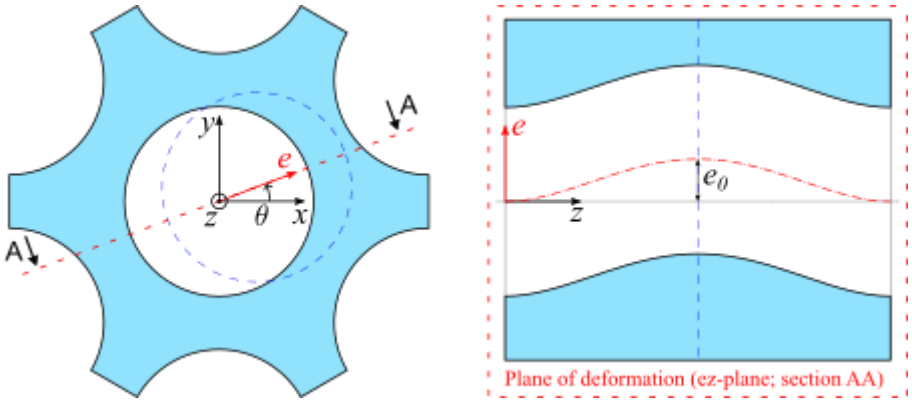


Figure 2.1: Schematic of the flow domain with uncertain deformation. The undeformed computational domain is shown in Figure 2.2. The left figure shows the xy-plane and indicates the section plane in red, which is shown in the right figure (scaled down along the z-direction). This plane is the ez-plane and makes an angle θ with the xz-plane. The right figure shows the extent of deformation e_0 within the ez-plane. The dark blue dashed line indicates a cut in the middle of the axial length, where the maximal deformation happens.

The two random variables have a uniform distribution as given below.

$$\theta \sim U(0°, 30°) \tag{3}$$

$$e_0 \sim U(0\ m, 0.001\ m) \tag{4}$$

The angle θ varies between $0°$ and $30°$, but due to the symmetries present in the nominal geometry this actually models every possible angle. The results can be periodically extended and mirrored to cover the response to the entire range of $0°$ to $360°$. The deformation extent varies between 0 and 0.001 m, the latter value being more than half of the gap width (1.85×10^{-3} m), as can be inferred from Table 2.1.

2.2.2 CFD model

A structured CFD mesh was created, which is visualized in Figure 2.2 together with the structural mesh and the geometric parameters are specified in Table 2.1. It consists of 700 divisions in axial direction, 60 circumferential divisions and 14 from one rod to another, resulting in a total of 882000 hexahedral cells. The finite volume solver Ansys Fluent is employed to solve the Reynolds-averaged Navier–Stokes (RANS) equations with the k-ω SST turbulence model (Menter, 1994) to close the system. The mesh is constructed such that a dimensionless wall distance $30 < y^+ < 100$ was obtained, implying the use of wall functions. It is worth mentioning that, due to the deformations and associated velocity changes, the y^+-values change somewhat for the different simulations, such that for some of the sample points a small region with $y^+ < 30$ exists. Altogether (spatially and all sample data combined), only 4% of the observed y^+-values had a value below 30.

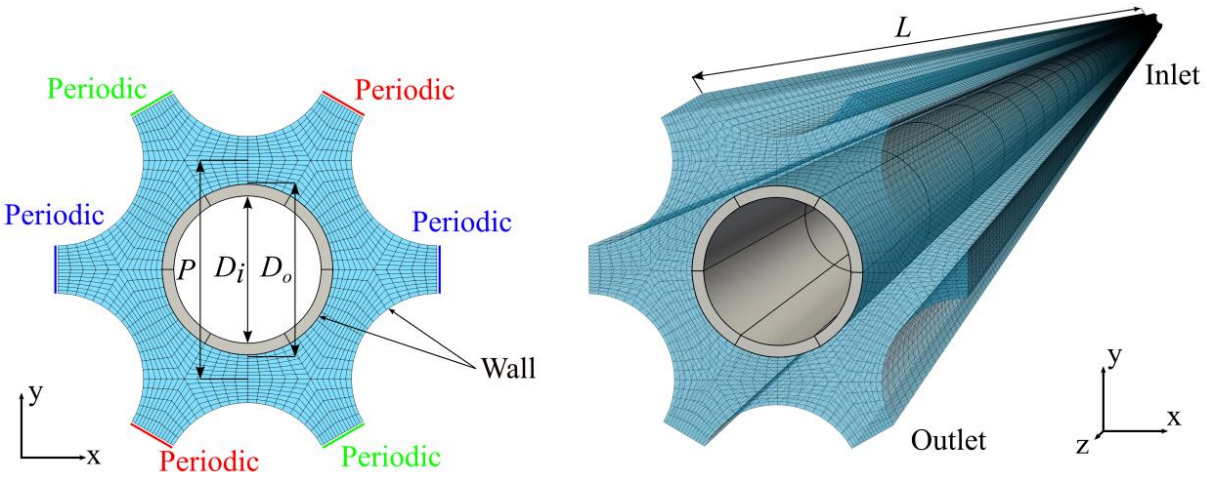


Figure 2.2: The CFD mesh (blue) and CSM mesh (grey) along with the geometrical parameters (also see Table 2.1) and boundary conditions.

To limit the computational cost, only parts of 6 neighbouring rods are modelled. Periodic boundary conditions were used to connect the sides of the domains, as indicated in Figure 2.2, virtually extending the domain to an infinite bundle. This means that fluid leaving at one side enters at the opposite side. This strategy was also adopted in Merzari and Ninokata (2011); Merzari et al. (2008); Ninokata et al. (2009); Shams, Mikuž, et al. (2018); Shams, Roelofs, et al. (2018), among others. The walls are modelled using no-slip boundary conditions. The Pb-Bi enters through a uniform velocity inlet at $1.9\ m/s$, which is around the nominal velocity for MYRRHA (Kennedy et al., 2020) and leaves via a constant pressure outlet. Based on the hydraulic diameter ($D_h = 5.33 \times 10^{-3}\ m$), this results in a Reynolds number of 6.13×10^4, corresponding to a fully turbulent condition. The turbulence intensity at the inlet is assumed to be 5% and the characteristic length scale $0.001\ m$. These values are equal to the ones in De Ridder et al. (2017); it was however shown in De Ridder et al. (2013) that they are of limited influence. The fluid properties are the same as in De Ridder et al. (2013) and can be found in Table 2.2.

Parameter	Value
Tube outer diameter D_o	$6.55 \times 10^{-3}\ m$
Tube inner diameter D_i	$5.65 \times 10^{-3}\ m$
Tube-to-tube pitch P	$8.4 \times 10^{-3}\ m$
Domain length L	$1.4\ m$

Table 2.1: Geometrical data (also see Figure 2.2).

Property	Value
Fluid density ρ_f	$10291\ kg/m^3$
Fluid viscosity μ_f	$0.0017\ Pa\ s$
Solid density ρ_s	$7850\ kg/m^3$
Young's elasticity modulus E	$200\ GPa$
Poisson's ratio ν	0.3

Table 2.2: Material properties of liquid lead-bismuth-eutectic and steel.

Second order schemes were used for all discretizations, and the SIMPLE algorithm was used to obtain the solution. Convergence was accepted when scaled residuals for all flow equations dropped below 10^{-6} and the relative change of the pressure drop and maximal velocity in the domain between successive iterations was smaller than 10^{-8}.

The stochastic deformation was applied to the mesh using a smoothing method, starting from a straight cylinder. This means that the nodes of the mesh are displaced while preserving their connectivity such that the wall ends up in the desired position and the mesh quality is as much as possible preserved, but no new nodes are created. The smoothing method used in Fluent models the mesh as a linearly elastic solid to calculate the new location of the nodes and a smoothing method parameter value of -0.4 was found to work well for the case under investigation. A good orthogonal quality was retained for all sample points. The deformation was applied on a converged case, retaining the data present in the cells and using it as an initial condition for the following simulation of the deformed geometry. When used in the FSI model, the flow was again allowed to converge in a steady manner after deformation, followed by 100 time steps of $1\ s$ and then the start of the FSI procedure (with a smaller time step), see Section 2.2.4.

2.2.3 CSM model

For the structure, the finite element method was applied, implemented in the software package Dassault Systèmes Abaqus. The geometry and mesh are also displayed in Figure 2.2 and the geometric and steel properties can be found in Table 2.1 and Table 2.2 respectively. No material damping was considered since the hydrodynamic damping is dominant. The mesh consists of 50 axial divisions and 6 around the circumference, totalling 300 second order elements. This is the same mesh as used in De Ridder et al. (2013), where no significant influence was found from refining the mesh. The structure is clamped at both extremities[2].

[2] Although the model was based on the one of De Ridder et al. (2013), axial sliding was allowed at the downstream extremity in their work, which substantially lowers the natural frequencies compared to a fully clamped condition. A direct comparison of the frequencies is consequently not possible.

Second order time discretization is again used, more specifically the Hilber–Hughes–Taylor scheme.

The implementation of the stochastic deformation of the structural mesh is less involved than for the CFD mesh. Each node is given a transverse displacement to obtain the C-shape deformation according to Eq. (1) and the stochastic parameters (e_0, θ), while the axial coordinate does not change. This means that axial sections are translated as a whole while connectivity is maintained.

Note that this is a geometric deformation, inducing no stresses as the deformed geometry acts as initial unstressed state in the FSI simulations. It is caused by thermal creep and irradiation creep effects, which are assumed here to cause a stochastic, stress-free deformation of the structure. These effects can cause some pre-stress and axial length change, but this is not taken into account here.

2.2.4 FSI coupling of models

The frequency and damping ratio of the deformed tube are determined by exciting a free vibration and fitting it to a characteristic modal equation, as will be explained below.

To this end, the previously discussed CFD- and CSM-model are used in a two-way coupled FSI simulation using the partitioned approach (Degroote, 2013). This implies that the two solvers can be considered black boxes, only communicating the solution at the mutual interface. Pressure and traction forces are communicated from the flow solver to the structural solver, and the structural solver in turn computes the displacements at the interface, which are provided to the flow solver to deform its domain boundary. Since the meshes are non-conformal at the interface, the data needs to be interpolated, for which a radial basis mapper is applied.

Because of the high density of LBE, an implicit coupling procedure is chosen, meaning that the exchange of solution data is conducted iteratively until equilibrium between the fluid and the structure is reached. Both a relative and absolute convergence criterion are used, requiring that the ℓ^2-norm of the displacement residual array either drops with a factor 10^3, or that its absolute value drops below $10^{-10}\ m$ before advancing to the next time step. The iteration limit, which is put at 20, is only occasionally reached. The stability and speed of convergence of the coupling iterations is enhanced using the IQN-ILS coupling algorithm (interface quasi-Newton with inverse Jacobian from a Least Squares model) (Degroote et al., 2009), by considering the system's displacement residual as function of the displacement a root-finding problem and applying a quasi-Newton method. The required inverse Jacobian is modelled from preceding coupling iterations, and in this case, also from previous time steps. This algorithm modifies the displacement between the structure solver and flow solver, such that spurious modes that impede convergence are attenuated. The partitioned coupling using IQN-ILS is implemented in the open-source tool CoCoNuT[3] (Delaissé et al., 2023).

As the coupling overhead in an FSI simulation is typically negligible compared to the cost associated with each of the solvers, it was chosen to implement CoCoNuT in Python. This makes the tool lightweight and portable. A modular approach is taken, separating the

[3] https://pyfsi.github.io/coconut

different functionalities in different modules that communicate data with each other. Examples are the coupling algorithms (Gauss–Seidel, Aitken relaxation, IQN-ILS), mappers between the fluid and structure mesh (linear interpolation, quadratic interpolation, radial basis functions) and solver wrappers. For a specific FSI simulation, the modules one wants to use are specified via a JSON settings file. The code is open source, and the modular structure makes it possible to adjust or extend the code to one's own needs, without having to make changes everywhere in the code, by changing one module or adding a custom module.

The FSI simulation requires the CFD domain to accommodate for the displacements at its boundary. The arbitrary Lagrangian–Eulerian (ALE) method is used and the mesh deformation itself is again determined using the "linearly elastic solid" smoothing method. As the vibrational displacements are much smaller than the stochastic deformation defining the geometry, the smoothing problem does not pose any issues.

A time step of $5 \times 10^{-4}\ s$ is used, as was done in De Ridder et al. (2013), who found that the hydrodynamic damping and natural frequency are almost insensitive to the time step size.

To simulate the free vibration of the rod, it is first given a displacement in positive y-direction by applying a force for some time, displacing the structure, and then releasing it. As the first mode is of interest here, it is convenient to choose the spatial force distribution such that the displacement just before removing the force *approximately* resembles the shape of the first flexural mode. According to beam theory, distributed force and displacement have a fourth derivative relationship. Approximating the bowed tube as a straight clamped-clamped beam, the following force spatial profile is derived:

$$\begin{cases} F_x(z,t) = 0 \\ F_y(z,t) = \alpha(t)\left[cosh(\kappa z) - cos(\kappa z) - \dfrac{cosh(\kappa L) - cos(\kappa L)}{sinh(\kappa L) - sin(\kappa L)}(sinh(\kappa z) - sin(\kappa z))\right] \\ F_z(z,t) = 0 \end{cases} \quad (5)$$

The parameter κ is determined again from Eq. (2). The time profile $\alpha(t)$ (see Eq. (5)) is chosen to ramp up over 100 time steps, then stay constant for 100 more time steps after which it immediately drops to 0, effectively releasing the structure into a free vibration. This profile also serves as scaling factor, which is chosen such that a maximal displacement of about $10^{-5}\ m$ is obtained.

$$\alpha(t) = \begin{cases} \dfrac{5000\ N \cdot t}{0.05\ s}, & if\ t \le 0.05\ s \\ 5000\ N, & if\ 0.05\ s < t < 0.1\ s \\ 0, & if\ t \ge 0.1\ s \end{cases} \quad (6)$$

It is emphasized that this displacement is *superposed* on the stochastic displacement defined above, causing a non-equilibrium situation (internal stresses are now present) starting a free vibration upon release of the acting force.

The resulting decaying vibration is plotted in Figure 2.3 for the middle of the straight ($e_0 = 0$) fuel rod. The 100th and 200th time step are indicated, as they demarcate the boundaries in the time profile $\alpha(t)$ (Eq. (6)). One can see that mainly a single mode—being the first bending mode—is excited using the previously discussed procedure, which is advantageous because it

simplifies curve-fitting a characteristic vibration response. This is done to derive the modal frequency and damping, as proposed by De Ridder et al. (2013) and De Santis and Shams (2017). This characteristic is given by Eq. (7), with $u(t)$ the displacement of the center ($z = 0.7\ m$), a the amplitude, c the viscous damping in the fluid, ω the natural frequency and ϕ the phase. Also an offset displacement u_o is allowed, since due to the bow present in the central rod, a non-zero equilibrium displacement results.

$$u(t) = a \cdot e^{-c \cdot t}\, sin(\omega \cdot t + \phi) + u_0 \tag{7}$$

These parameters are found using a non-linear least squares fitting with the obtained displacement data.

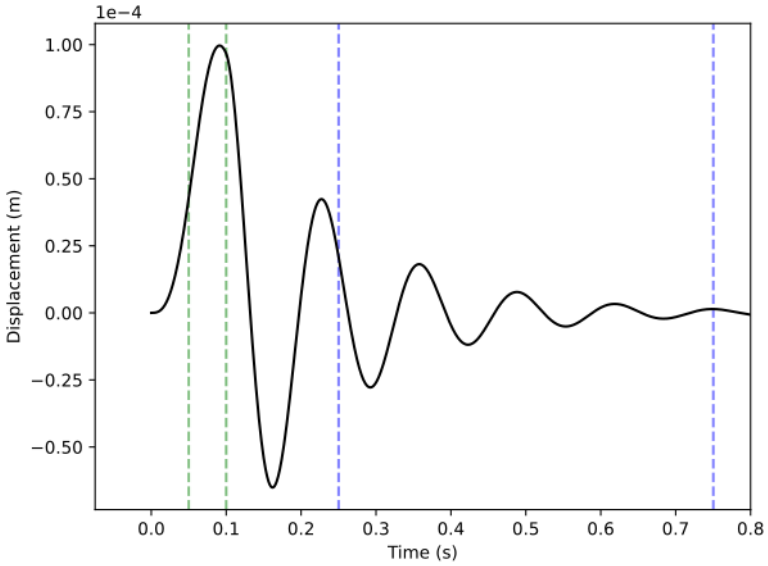


Figure 2.3: The y-displacement of the midle of the rod ($z = 0.7\ m$) as a function of time for the case without bow deformation ($e_0 = 0$). The four dashed vertical lines indicate the 100th and 200th time step (green) and the 500th and 1500th time step (blue).

The output quantities of interest are the natural frequency f and damping ratio ζ, which are further post-processed from the curve-fitted ω and c (see Eq. (7)) as given by the following equations.

$$f = \frac{\sqrt{\omega^2 + c^2}}{2\pi} \tag{8}$$

$$\zeta = \frac{c}{2\pi f} \tag{9}$$

2.2.5 Monte Carlo method

Monte Carlo (also see Section 1.2.4.1 in PART 1) refers to a broad family of methods using (pseudo-)random numbers to evaluate mathematical expressions (in a broad sense), with a high number of applications in diverse research areas. Monte Carlo methods have been used for decades and remain popular to date, owing to their ease of implementation and robustness.

In this case, where the goal is uncertainty quantification, it comes down to the propagation of the input uncertainty through the system to obtain information on the uncertain output quantities of interest. This means performing a pseudo-random sampling of the stochastic variables (θ, e_0) representing the input uncertainty and evaluating the model for each sample value. An output sampling is obtained, allowing to estimate its descriptive statistics, correlations, probability laws, histogram, etc.

Advantages are that the methods are per definition non-intrusive, hence their ease of implementation as the system's computational model can be considered a black-box, and that they impose no requirements on the mathematical behaviour of the response function as long as point-wise evaluation is possible. These properties make MC a sound candidate for validating the Polynomial Chaos method, described in the next section. A disadvantage is the slow convergence of the statistics, for example the mean converges at a rate of $\mathcal{O}(1/\sqrt{m})$, with m the sample size. This is the reason the Polynomial Chaos method is preferred for the more expensive FSI simulation needed to quantify the vibrational behaviour.

Monte Carlo methods have already been used for a long time and they are often considered a viable choice for many applications. To remediate some of their disadvantages, they are actively researched, and new variants and adaptations appear (Giles, 2015; Haji-Ali et al., 2016; Kroese et al., 2011; Kuo & Nuyens, 2016; Robbe et al., 2019). However, for this work its simplest form, the crude Monte Carlo method described above, is used to compare to a non-intrusive Polynomial Chaos method, as discussed in the next section.

2.2.6 Non-intrusive Polynomial Chaos method

Instead of relying on local information, as is done in the Monte Carlo approach, benefits can be gained from reconstructing the stochastic response, as is done with PCE-based methods. The core idea is that it is possible to represent random variables by a polynomial expansion of other random variables. This method is briefly discussed in this section. For a more detailed introduction the reader is referred to the books of Le Maître and Knio (2010) and Xiu (2010).

For expansion of a generic stochastic outcome $s(\vec{x}, t, \vec{\xi})$ in polynomials Ψ_k of the random vector $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_d)$ (the components of $\vec{\xi}$ assumed mutually independent for the purpose of this work), one can write (Xiu & Karniadakis, 2002):

$$s(\vec{x}, t, \vec{\xi}) = \sum_{k=0}^{\infty} s_k(\vec{x}, t)\Psi_k(\vec{\xi}) \approx \sum_{k=0}^{P} s_k(\vec{x}, t)\Psi_k(\vec{\xi}). \tag{10}$$

The expansion can be seen to have a deterministic part and a stochastic part. The first one is defined by the coefficients $s_k(\vec{x}, t)$, also called the stochastic modes of the expansion, with $\vec{x}$ representing a spatial coordinate and t representing time. The second part is one of the polynomials $\Psi_k(\vec{\xi})$. It can be seen that the approach is not restricted to random variables, $s(\vec{x}, t, \vec{\xi})$ can be a random field or stochastic process as well. The number of terms $P + 1$ is related to the expansion order, i.e. the largest polynomial order o present in the sum, as follows (Ghanem & Spanos, 1991):

$$P + 1 = \binom{o + d}{o} = \frac{(o + d)!}{o!\, d!}, \tag{11}$$

with d the number of dimensions of the stochastic vector $\vec{\xi}$.

The coefficients/modes $s_k(\vec{x},t)$ are multiplied by the polynomials Ψ_k of the random variable $\vec{\xi}$, which are orthogonal according to the following inner product:

$$\langle \Psi_i(\vec{\xi}), \Psi_j(\vec{\xi}) \rangle = \int_{\Xi} \Psi_i(\vec{\xi}) \Psi_j(\vec{\xi}) f_p(\vec{\xi}) d\vec{\xi} = \delta_{ij} \tag{12}$$

with Ξ the support (which can be infinite) of weight function $f_p(\vec{\xi})$.

$s(\vec{x},t,\vec{\xi})$ is approximated by its spectral decomposition, much akin to Fourier decomposition. The set of polynomials is directly related to the weight function $f_p(\vec{\xi})$. A notable example are the polynomials belonging in the so-called Askey scheme, which are orthogonal with respect to a weight function directly corresponding to common probability density functions. This makes them particularly interesting for representing stochastic quantities (Xiu & Karniadakis, 2002). For example, the standard normal distribution corresponds to Hermite polynomials, the beta distribution to Jacobi polynomials, the gamma distribution to Laguerre polynomials and the uniform distribution to Legendre polynomials. A PCE approach where the weight function is matched to the probability density is referred to as the generalized Polynomial Chaos (gPC) method as described in the work of (Xiu & Karniadakis, 2002).

The Polynomial Chaos representation of a stochastic quantity has several advantages. One can recognize an expectation operator in Eq. (12): $\langle \Psi_i(\vec{\xi}), \Psi_j(\vec{\xi}) \rangle = \mathbb{E}[\Psi_i(\vec{\xi}), \Psi_j(\vec{\xi})]$. This property indicates that the descriptive statistics of $s(\vec{x},t,\vec{\xi})$ are accessible through the inner product, even though calculating high-order moments can become intricate (Lefebvre, 2021). Besides that, the expansion also serves as a cheap surrogate model for the input-output relation $s: \vec{\xi} \mapsto s(\vec{\xi})$.

The determination of the coefficients $s_k(\vec{x},t)$ can be done in multiple ways. Non-intrusive methods are preferred over intrusive methods for this work. In the latter all stochastic quantities in the model's governing equations are replaced by their PCE representation (Eq. (10)), followed by a projection of the stochastic equations on each of the polynomial basis functions following Eq. (12), yielding a $P+1$ times larger set of equations to be solved for the stochastic modes. A drawback is that one needs access to the code of the computational model and modifying its governing equations can be rather involved, for example when projecting terms that are non-linear in the stochastic variables.

Different non-intrusive approaches exist, but they all determine the coefficients based on a sampling of the stochastic output, thus the model is run in a deterministic way for different realizations of the stochastic input. Non-intrusive spectral projection (Le Maître & Knio, 2010) relies on the orthogonal projection of the output quantities of interest $s(\vec{x},t,\vec{\xi})$ on the chosen orthogonal basis using Eq. (12). One obtains

$$s_k = \frac{\langle s, \Psi_k \rangle}{\langle \Psi_k, \Psi_k \rangle}, \tag{13}$$

of which the denominator is the analytically known norm of the polynomials. The numerator (from Eq. (12)) can usually not be determined analytically. Instead, one can resort to

numerical quadrature. A logical choice is Gauss quadrature, given the polynomial in the integrand.

When constructing a d-dimensional cubature rule by tensorizing the 1D rule, the required number of samples equals

$$(o+1)^d \tag{14}$$

This method suffers from the notorious *curse of dimensionality*, as the number of samples grows exponentially with the dimensionality of the stochastic input. The tensor product approach is often only deemed feasible for low dimensional input ($d < 5$) (Xiu, 2010). In higher-dimensional cases it is better to approximate the integral by using a Monte Carlo-based method or sparse cubature rules.

A general advantage of non-intrusive methods is that the different deterministic evaluations can be run in parallel, with optimal efficiency as they are completely separate computations, as is also the case for the Monte Carlo method. Such an efficiency would be exceedingly difficult to obtain by parallelizing the coupled set of equations of the intrusive PCE method. This property has extensively been exploited to execute the sample evaluations for this work in a reasonable amount of time. The non-intrusive PCE method as well as the employed cubature rules are all implemented in the open-source Python software package Chaospy[4] of Feinberg and Langtangen (2015).

In this work, a two-dimensional stochastic input is used $\vec{\xi} = (\xi_1, \xi_2) = (\theta, e_0)$ and since both have a uniform distribution, Legendre-chaos is employed. The different output quantities of interest generally will not have a uniform distribution (in which case a 1st order approximation would suffice). As the distribution of the QoI is a priori unknown, basing the choice of the polynomial basis on the stochastic input is a best guess. One hopes to retain good convergence properties as the polynomial order increases to make up for the non-linearity introduced when propagating the uncertainty through the model to obtain the stochastic response. The expansion coefficients are determined using Gauss–Legendre quadrature.

The performance of the chosen PCE approach is first compared to a result obtained with the Monte Carlo method, on a similar case where the same flow physics are acting, while being computationally sufficiently cheap to be run for a large Monte Carlo sampling. After that, the method is used on the full FSI problem described earlier in this section. The results are discussed in the next section.

2.3 Results

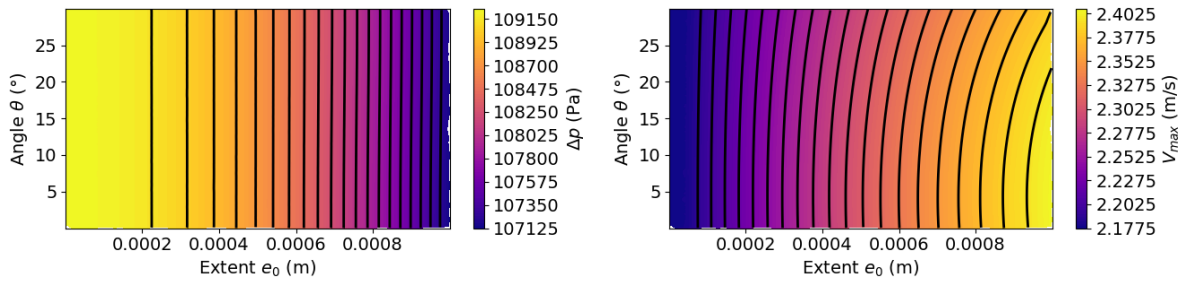
In this section the results of the two cases are discussed. First, in Section 2.3.1, a steady flow through the rigid (but stochastically deformed) structure is considered. The stochastic response to (θ, e_0) is propagated using both the Monte Carlo and gPC method with different expansion orders to verify the convergence behaviour of the latter. Then, in Section 2.3.2 the gPC method is used to investigate the stochastic frequency and damping ratio of the flexible tube.

[4] https://chaospy.readthedocs.io

2.3.1 Steady flow through deformed tube bundle

The flow domain presented in Figure 2.2 was used and deformed according to the discussion of Section 2.2.1. The central tube was kept rigid for this first step and a steady flow simulation was performed for each evaluation. Different output quantities of interest were considered and in the discussion below the pressure drop Δp and maximal velocity in the domain V_{max} will be analysed. As mentioned in Section 2.2.2, these were also subject to a strict convergence criterion for the CFD solver. The pressure drop is a relevant hydraulic parameter and Roelofs (2020) states that it should be investigated how this is affected by deformations and, even though vibrational behaviour is typically investigated as a function of bulk/inlet velocity (De Ridder et al., 2015; Divaret et al., 2014; Paidoussis, 1966a, 1966b) the maximal velocity is deemed a relevant local influence. Bulk velocity has an influence on frequency (De Ridder et al., 2015) and damping (Divaret et al., 2014; Moussou et al., 2017).

A crude Monte Carlo sample of 10000 evaluations is compared to the PCE method varying from a 1st to 9th order expansion, using the stochastic input (θ, e_0). This number of evaluations is deemed appropriately balanced in the trade-off between computational load and statistical convergence. Similar sample sizes were found in literature (Hosder et al., 2010). One evaluation required 234 seconds on average on a 28-core machine, where sample points with larger deformation took more time. The highest (9th) order PCE required model evaluations at 100 different quadrature nodes, following Eq. (14).



(a) (b)

Figure 2.4: Contour of (a) the pressure drop Δp and (b) the maximal velocity V_{max}, with 20 equidistant contour lines drawn.

The response of the quantities of interest (Δp and V_{max}) to the stochastic input (θ and e_0) are shown in Figure 2.4. These contour plots were created by interpolating a Delaunay triangulation of the collection of data points. They can also be generated from the Polynomial Chaos Expansion, as it serves as a surrogate for the response surface. For a sufficiently high order no difference can be discerned with the naked eye, which can serve as an intuitive notion of its convergence and correctness.

The pressure drop (Figure 2.4a) on the one hand almost only correlates with the deformation extent e_0 with which it can be observed to decrease, while it is unaffected by the angle θ. This decrease is caused by the increase of the subchannel size on one side of the central rod, which was confirmed with a test simulation on the rod bundle with the central rod given an eccentricity instead of a bowed shape. A decreased pressure drop was also found for bowed PWR fuel assemblies by Puragliesi et al. (2019).

The maximal velocity (Figure 2.4b) on the other hand does show an interaction with the angle, especially for larger deformations. The largest V_{max} is found for angles closer to 0°, i.e. bowing in the subchannels' direction. This kind of deformation increases the size of the subchannel on the concave side of the deformed rod the most. The subchannels already have the largest hydraulic flow-through area for undeformed rods. This kind of deformation increases this area even more, leading to larger velocities in their centre as the distance from the walls is greater. No angle effect is present for low to zero deflection, as this parameter loses meaning when the deformation vanishes.

Since a uniform inlet distribution was used, the area between the contour lines in Figure 2.4 directly relates to the histogram of the output quantity of interest. For example, Δp has a left-tailed histogram, as is clearly shown in Figure 2.5. This is a result of the flatness of the response for small e_0, causing a large probability density of higher pressure drops, and next to it a large zone of lower pressure drops with lower probability density, hence a rather skewed (asymmetric) histogram. The figure shows besides a 50-bin histogram of the Monte Carlo results also the same data generated using the spectral surrogate. Additionally, a 1000-bin, large sample is generated from the surrogate using a Hammersley low-discrepancy sequence quasi-random sampling technique (Hammersley, 1960; Kroese et al., 2011), in order to visualize the probability density of the output. It can be seen in Figure 2.5a that there is a large discrepancy for the 1st order approximation of the pressure drop, which mostly preserves the uniform shape of the inlet distribution. In Figure 2.5b excellent agreement is reached for a 3rd order expansion (using 16 quadrature nodes).

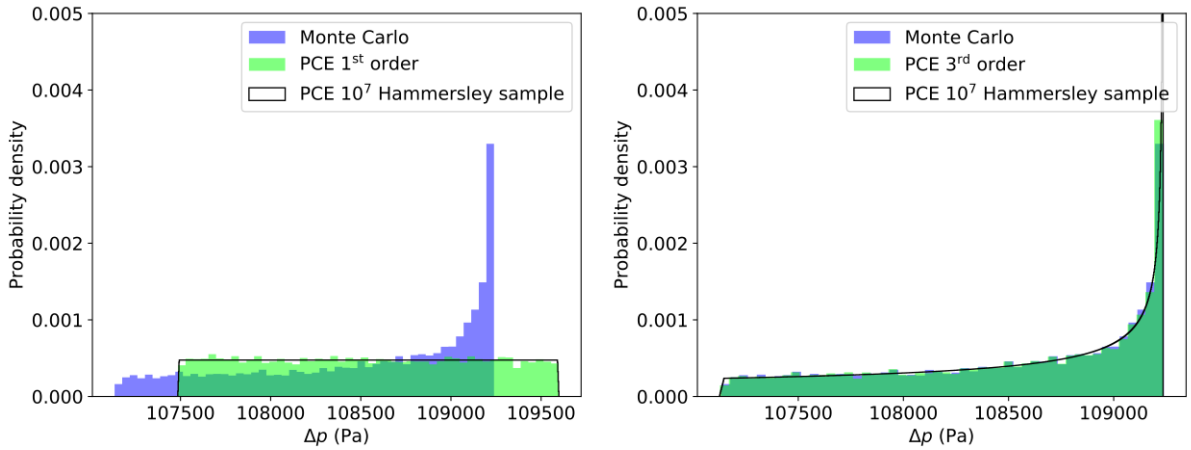
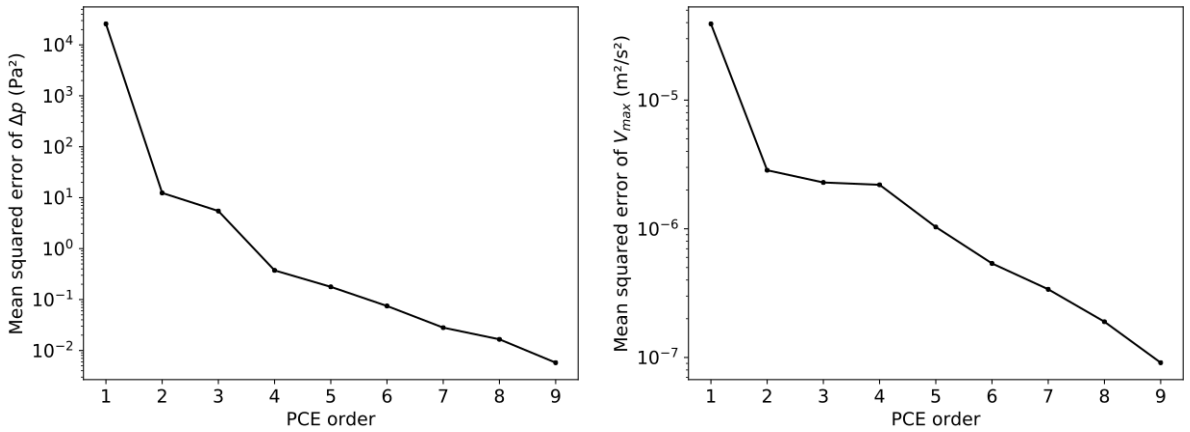


Figure 2.5: Comparison of the output histogram for Δp between Monte Carlo and Polynomial Chaos results of (a) 1st order and (b) 3rd order. The histograms for the PCE are obtained by feeding the Monte Carlo sample to the obtained response surfaces. The values on the vertical axis are normalized such that the area under the histogram integrates to 1, corresponding to a probability density.

	Average	Standard deviation	Skewness
Angle θ (theoretical)	$15°$	$8.660°$	0
Angle θ	$14.95°$	$8.726°$	-2.026×10^{-3}
Extent e_0 (theoretical)	$5.000 \times 10^{-4}\ m$	$2.887 \times 10^{-4}\ m$	0
Extent e_0	$5.043 \times 10^{-4}\ m$	$2.888 \times 10^{-4}\ m$	-1.588×10^{-2}
Pressure drop Δp	$1.085 \times 10^{5}\ Pa$	$6.306 \times 10^{2}\ Pa\ (0.58\%)$	-6.264×10^{-1}
Maximal velocity V_{max}	$2.298\ m/s$	$6.737 \times 10^{-2}\ m/s\ (2.9\%)$	-2.236×10^{-1}

Table 2.3: Descriptive statistics resulting from Monte Carlo method.

In Table 2.3 the statistics computed from the data gathered in the Monte Carlo sampling are summarized. The pressure drop is found to be left-skewed (as deduced earlier from the contour plots and asymmetry of the histograms in Figure 2.4 and Figure 2.5), while the skewness value of the maximal velocity is closer to zero. It can be seen that the statistics of the input random variables θ and e_0 are not matching their theoretical values beyond the second digit, indicating the slow convergence of this method. Also, the confidence intervals on the statistics are still rather wide with the sample size of 10000 (not shown). Therefore, it is decided to compare the Polynomial Chaos result using the mean squared error (MSE).



(a) (b)

Figure 2.6: Mean squared error of PCE with respect to MC of (a) the pressure drop Δp and (b) the maximal velocity V_{max}.

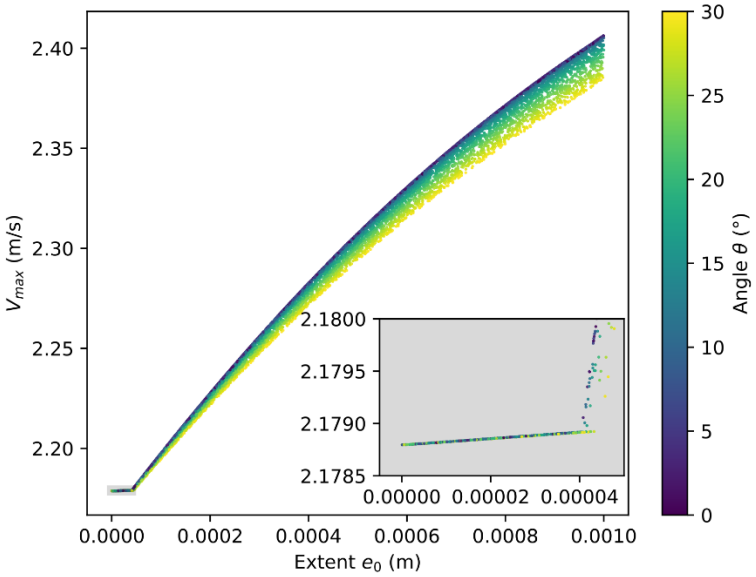


Figure 2.7: Maximal velocity as a function of the displacement extent e_0 and coloured by the angle θ.

The MSE is calculated as the mean of the squared differences between the Monte Carlo data and the PCE evaluated at the same locations. Figure 2.6 shows the evolution of the MSE of the output quantities of interest as the polynomial order increases, which is found to be monotonic each time. The convergence of the V_{max} seems to stall between 2nd and 4th order and it appears slower overall than the pressure drop when looking at the order of magnitude decrease of the MSE. This is related to the leftmost part of the response (left of the first contour line in Figure 2.4b), where it is rather flat as a function of e_0, after which a sharp increase follows that diminishes in slope afterwards. Figure 2.7 shows this in more detail: it can be seen that for e_0 below $4 \times 10^{-5}\ m$ the slope is very shallow, after which the behaviour completely changes. The response is actually a combination of two curves, because the location where the maximal velocity is found switches around $e_0 = 4 \times 10^{-5}\ m$. The maximal velocity for a deformation extent e_0 below $4 \times 10^{-5}\ m$ is found close to the inlet and is associated with the redistribution of the uniformly entering flow. When the deformation extent is above that value, the highest velocities are found at about 67% of the axial length, just downstream of the maximal deformation, due the flow redistribution caused by the C-shape deformation.

The highly non-linear response is difficult to fit with a low-order polynomial. These kind of responses are potentially better fitted with a piece-wise polynomial or non-polynomial basis (Le Maître & Knio, 2010), given no prior information on the response is available (in which case considering the velocities at the two locations as two different QoI's would be the most efficient strategy). This is however beyond the scope of the current work, as smooth response surfaces are expected for the vibration characteristics.

Besides the polynomial order used, there is also the aspect of the quadrature rule. Below the 5th order, no quadrature nodes are located in the region $e_0 < 4 \times 10^{-5}\ m$ using Gauss–Legendre quadrature, meaning that up to this order the PCE model is unaware of this

information. It can indeed be seen (Figure 2.6b) that beyond the 5th order the MSE starts decreasing again after the stalled region between 2nd and 4th order.

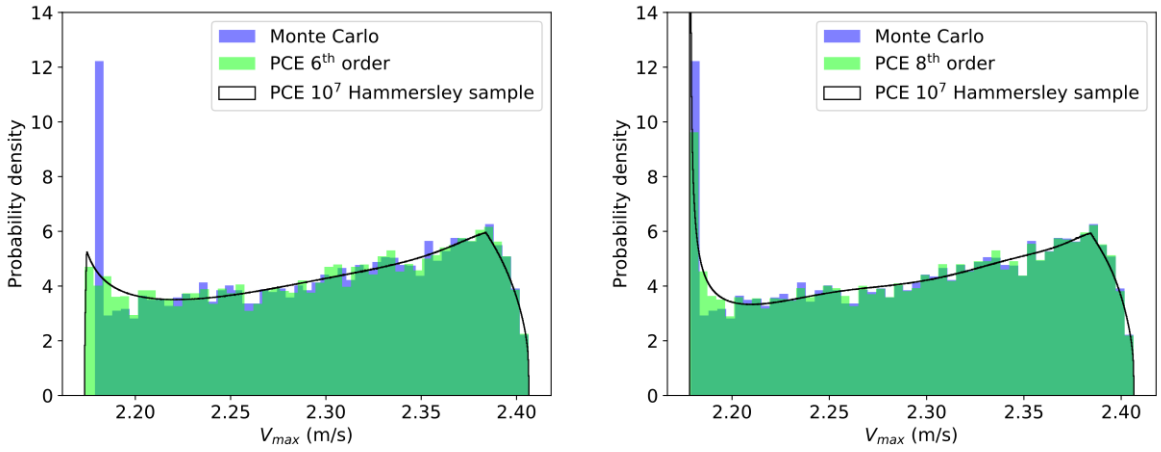


Figure 2.8: Comparison of the output histogram for V_{max} between Monte Carlo and Polynomial Chaos results of (a) 6th order and (b) 8th order. The histograms for the PCE are obtained by feeding the MonteCarlo sample to the obtained response surfaces. The values on the vertical axis are normalized such that the area under the histogram integrates to 1, corresponding to a probability density.

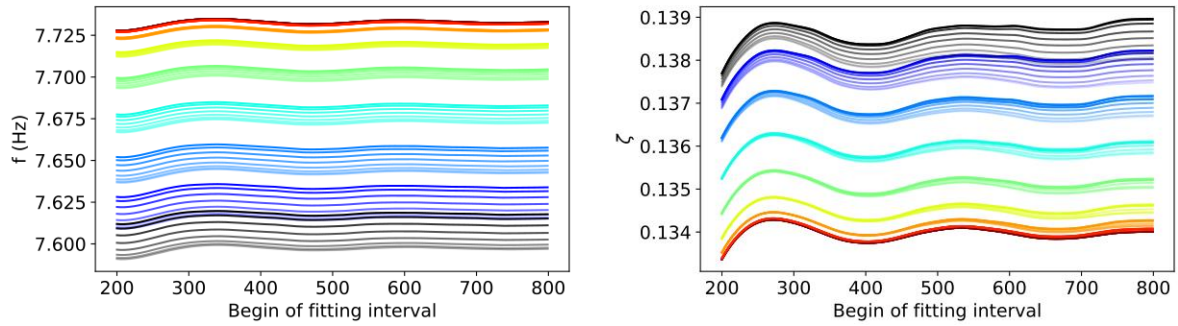
The flat part of the response (Figure 2.4b and Figure 2.7) causes a peak in the histogram, as can be seen in Figure 2.8. This peak is mostly missing for a 6th order approximation, while it is covered well using an 8th order PCE. A visually comparable agreement between histograms from MC and PCE as the latter result (Figure 2.8b) was already found at a 3rd order expansion for the pressure drop (Figure 2.5b). Some bi-modality is introduced by the peak on the left, which is also the reason for the more moderate skewness value observed in Table 2.1. The data right to the cusp $V_{max} \approx 2.38\, m/s$ in the histogram corresponds to the bottom-right region contained in the highest two contour levels in Figure 2.4b.

From the results it is concluded that the PCE converges fast with a smooth, well-chosen QoI with only a small number of sample points needed. Good results were obtained for orders below 9, which was the highest tested PCE order, requiring 100 evaluations (Eq. (14)) to determine the 55 coefficients for the terms in the expansion (Eq. (11)). The numerical complexity is mainly attributed to the creation and fitting of the Gauss quadrature. Once the samples are obtained, it takes less than $1\ s$ to construct the PCE, which is negligible compared to the time it takes to acquire the samples. A smooth input-output relation is expected for the frequency and damping ratio, being the QoI's in case of the flexible rod, such that no convergence issues like with V_{max} are expected. These results are presented next.

2.3.2 Stochastic modal properties of deformed tube bundle

The methodology explained in Section 2.2.4 was used to obtain the results below. The bowed central tube is excited in mainly its first mode by applying a distributed force for a brief time period in an FSI simulation, yielding a free vibration (e.g. Figure 2.3). The frequency and damping ratio were subsequently obtained from fitting the vibration response to an equation

expressing a decaying oscillation (Eq. (7)). A single evaluation of the FSI problem takes about 33 hours on 3×28 cores, such that Monte Carlo methods are computationally not feasible and solely PCE was relied on instead. Still multiple expansion orders were used ($p \in \{2, 4, 6, 8\}$) such that the convergence can be assessed by comparing them. The expansion orders correspond respectively to 9, 25, 49 and 81 sample points, following from Eq. (14).



(a) (b)

Figure 2.9: Fitted (a) frequency f and (b) damping ratio ζ as a function of the beginning of the fitting interval, which is starting at time step 200 to 800, for all realizations in the 8th order pCE. The different colours indicate the different deformation extents e_0, the colour intensity the angles θ, both corresponding to the nodes of the quadrature defined using Legendre polynomials on the intervals given by Eq. (3) and (4).

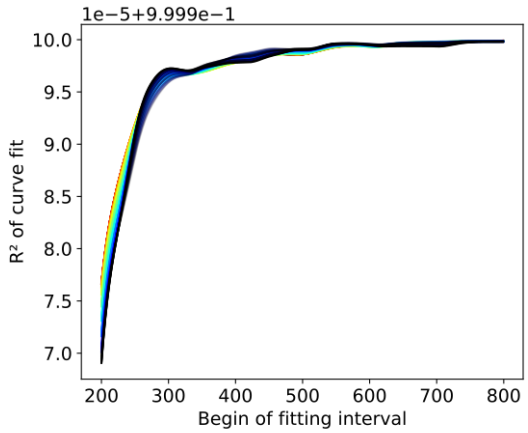
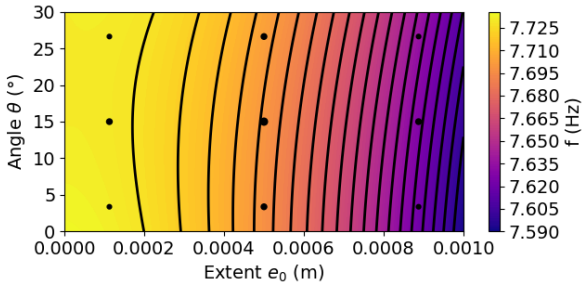


Figure 2.10: R^2-value as a function of the beginning of the fitting interval, which is starting at time step 200 to 800, for all samples in the 8th order PCE.

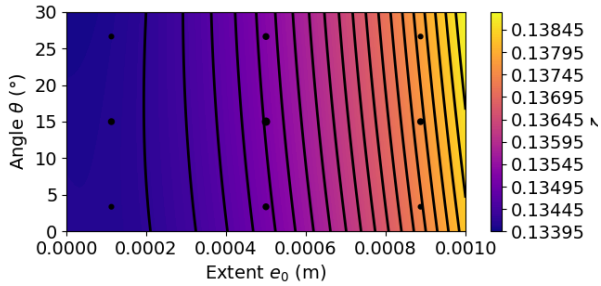
The simulations were halted after 1500 time steps, which also is the end of the data contained in the fitting window. In Figure 2.9, the fitted frequency and damping ratio can be seen as a function of the beginning of the fitting window with the end being fixed on 1500 for the 81 FSI evaluations of the 8th order PCE. The angle θ has little influence for low deformation extent e_0, while it disperses the results when the deformation becomes more substantial. This was also observed for the pressure drop and maximal velocity in the previous section. The frequency (Figure 2.9a) is rather insensitive to the beginning of the fitting interval, while the damping ratio (Figure 2.9b) varies a bit more. It is decided to fit the data in the interval $[500, 1500]$ i.e. $[0.25\ s, 0.75\ s]$, since time step 500 is already well beyond 200, where the

free vibration begins and still some effect of the initial load could be present. This can also be seen on a plot of the R^2-value of the least squares fit, see Figure 2.10, which flattens out when the beginning is put beyond 300 time steps. This means that the proposed model (Eq. (7)) matches the vibration data more closely beyond this point than before. A similar plot can be created where the end time of the window is varied with the beginning fixed (not shown). There, the sensitivity was even less. The fitting window is indicated on Figure 2.3 using blue dashed lines.

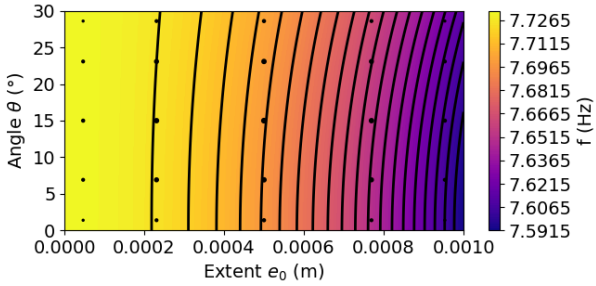
Figure 2.11 shows contour plots of the PCE models of all different orders considered, for the frequency and damping ratio. It becomes difficult to observe visually any difference beyond the 4th order for both, indicating convergence of the PCE with increasing polynomial order. The frequency can be seen to decrease with increasing deformation extent e_0, which is hypothesized to be effectuated by an augmented added-mass effect, due to the closer proximity of the walls. The same reasoning applies to the damping ratio, of which the increase is attributed to increased added-damping. It is known that confinement effects decrease the eigenfrequencies while increasing damping (Chung & Chen, 1977; De Ridder et al., 2017). A slight interaction with the deformation angle θ can be deduced as well, where for $\theta \to 30°$ the frequency is slightly less diminished, while the opposite is true for ζ.



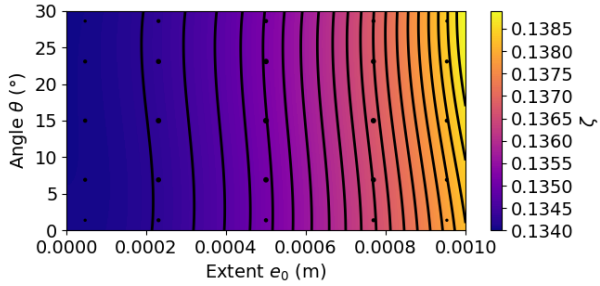
(a) 2nd order PCE of f



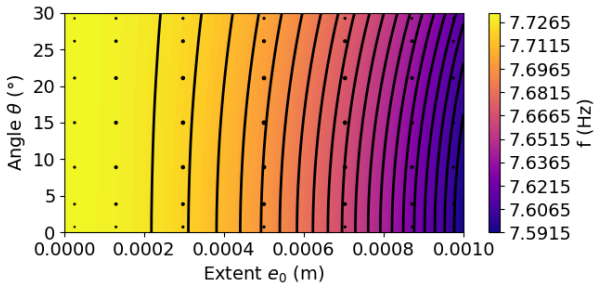
(b) 2nd order PCE of ζ

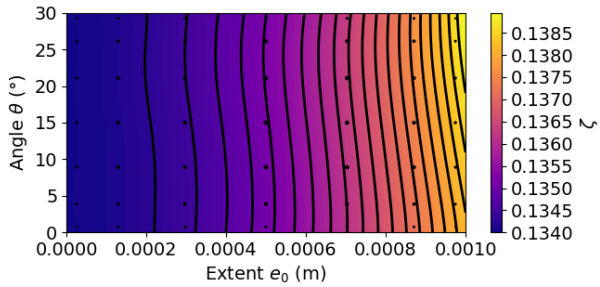


(c) 4th order PCE of f

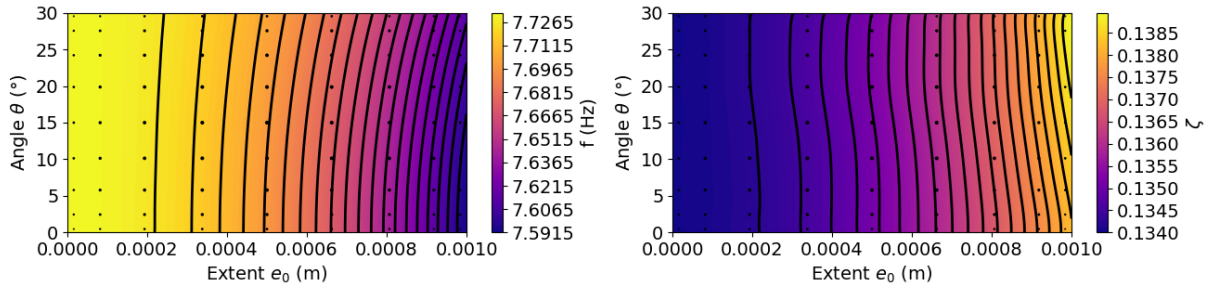


(d) 4th order PCE of ζ





(e) 6th order PCE of f (f) 6th order PCE of ζ



(g) 8th order PCE of f (f) 8th order PCE of ζ

Figure 2.11: Contour plots of the response surfaces for different polynomial orders. The Gauss–Legendre quadrature points are indicated using black dots, with the size of each dot representing the weight.

Table 2.4 summarizes the descriptive statistics of f and ζ, derived from an 8th order PCE. The convergence with respect to expansion order is fast for both the frequency and damping ratio (data not shown). The skewness converged the slowest. The difference between subsequent orders weighted by the 8th order result is well below 1% when going from 4th to 6th order. For the lower order statistics this is already the case when comparing 2nd to 4th order PCE. It is seen that the standard deviation is small for both quantities of interest, compared to the mean value. This indicates that the sensitivity of the vibration characteristics to this specific C-shape deformation is rather low for the flow rate under investigation. For other flow rates or different boundary conditions, a different outcome is deemed possible. For example, De Ridder et al. (2015) found that the onset of fluid-elastic instability is rather sensitive to misalignment. Their case was a static instability rather than a vibration, namely buckling, but a similar simulation approach was used. An a priori present C-shape bow could have a similar effect as misalignment since an asymmetric geometry with respect to the flow is believed to be the root cause. Investigating this is however beyond the scope of this work.

	Mean	**Standard deviation**	**Skewness**
$f\ (Hz)$	7.6885	0.0394 (0.51%)	−0.6877
$\zeta\ (-)$	0.1355	0.0013 (0.99%)	0.6781

Table 2.4: Descriptive statistics of the output quantities of interest, computed from the 8th order PCE.

Finally, using the same method as in Section 2.3.1, histograms can be constructed from the spectral surrogates. They are displayed for the different expansion orders in Figure 2.12. The discrepancy between the different orders is small, especially beyond the 4th order. The flatter part of the stochastic responses (for low e_0 in Figure 2.11) causes a large peak in the probability density for both quantities of interest. The frequency is observed to be left-tailed and the damping ratio right-tailed, which can also be inferred from the negative respectively

positive skewness value in Table 2.4. When looking at the complete range of the histogram, again a rather low influence of the deformation on the results can be inferred.

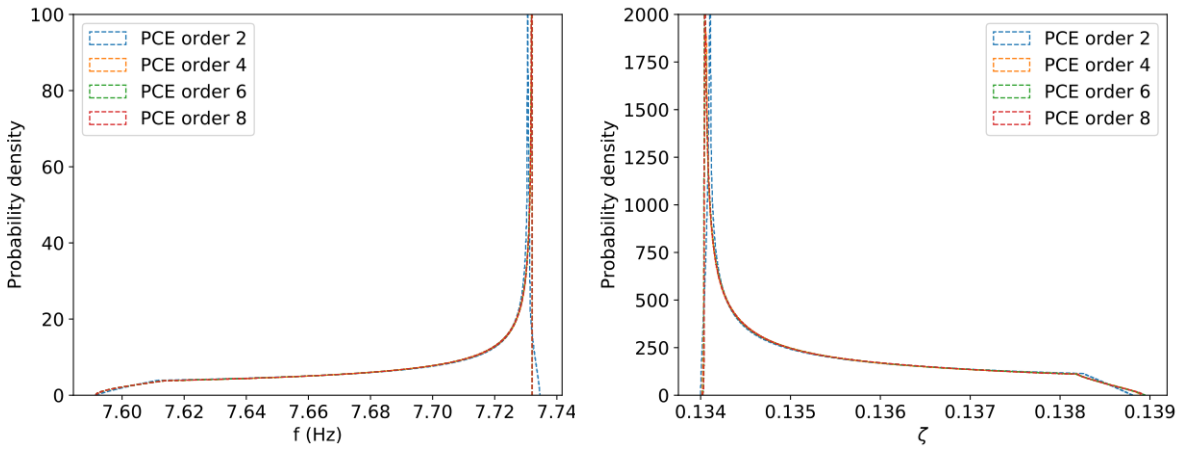


Figure 2.12: Output histogram from a Hammersley sample of size 10^7 from the PCE of (a) f and (b) ζ.

It can be concluded that a fast convergence as a function of PCE order was obtained for both the frequency and damping ratio, owing to a smooth stochastic response. The obtained coefficients of variation were in the order of 1%.

2.4 Discussion

In this work, a step towards determining the effect of deformations on the vibration characteristic of nuclear fuel was made by considering the deformation as a stochastic quantity. A uniform random deformation was assumed, since no exact input distributions are known and this is also a typical choice in literature (Banyay et al., 2020; Geffray & Macián-Juan, 2015). MYRRHA inspired the investigated case, but it is believed that the developed methodology can be ported without substantial issues to PWR or other type of reactors. The results would be affected, among other things, by the difference in array layout and flow rate.

In future work the uncertainty on the deformation needs to be quantified on its own and is also believed to be reducible to some extent (Faure et al., 2023). Reducing the uncertainties in number and range as much as possible before resorting to UQ is beneficial in terms of realism as well as computational efficiency. For a better specification of the deformation uncertainty, a model for the deformation based on irradiation dose, residence time and thermal creep could be added, like the one presented by Ye et al. (2020). This model could be combined with the FSI simulations, or work as an earlier step in a cascade of simulations. This means not assuming an uncertain deformation, but propagating more high-level uncertainties as material properties, manufacturing tolerances and operating conditions (e.g. temperature and irradiation). One should keep in mind that the fuel assemblies are only a small part of a complete system. Considering coupling at different scales (Toti et al., 2018) in a more holistic approach could prove beneficial. Still, knowledge on input uncertainty can remain limited, so

UQ methods that do not require full information on probability distributions but instead deal flexibly with the data that is present (Oladyshkin & Nowak, 2012) could be of interest.

Furthermore, in reality wire-wrappers involving many contact positions are proposed in designs like MYRRHA. This will result in a more complex output behaviour and probably also a higher sensitivity, as changing support conditions have a rather high influence on flow-induced vibrations and fretting (Dolfen, De Ridder, et al., 2022; Hassan et al., 2003). Of ultimate interest is not the vibration characteristic exactly, but predictions (in terms of probability) of excessive wear, unsafe and/or uneconomical behaviour. Considering a more complete model of the nuclear core will substantially increase the computational cost, so the efficiency of the UQ method will be of paramount importance when adding complexity in future work.

Another aspect to pay attention to is the mathematical behaviour of the model response to the uncertain inputs. In cases where strong non-linearities act, like the contact problem when fretting is of interest, the smoothness of the input-output relation can be reduced. This poses a challenge for the gPC method, which tries to represent this relation using smooth polynomials and would not achieve convergence or only at an unreasonably large order. This kind of issue was already encountered in the behaviour of V_{max} (see Section 2.3.1), so a piece-wise polynomial or non-polynomial basis (Le Maître & Knio, 2010) may be of interest to tackle this. It is noted however that non-linearity within a model does not necessarily imply a non-smooth response. In the work of Dolfen, De Ridder, et al. (2022), contact between different fuel rods was simulated, exhibiting strong non-linear properties due to the switching behaviour between support-active and non-active and the sticking and slipping nature of the contact. The relation between the predicted wear and friction coefficient was however smooth as soon as non-zero friction was considered, the friction coefficient being one of the most important uncertainties present in that work.

2.5 Conclusion

The effect of fuel bundle deformations on its vibration behaviour was investigated with a stochastic approach, exploring the feasibility of a shift from fully deterministic simulations to stochastic simulations, where the output quantities of interest are described in terms of probability. This is achieved by propagating a stochastic deformation through the computational models.

To this end the Polynomial Chaos method was used, after testing it first on a steady case without FSI, such that creating a Monte Carlo benchmark was feasible. From the comparison it was concluded that valid results were obtained using PCE and that the method converges sufficiently fast in terms of polynomial order.

Fast convergence was again observed for the freely vibrating tube with stochastic deformation. The FSI problem is representative for other FSI problems in this research field in terms of computational costs and physical behaviour, thus this research indicates that stochastic approaches are feasible for a small number of stochastic inputs.

For the case itself, the deformation extent and especially the angle have only a limited influence on frequency and damping for the C-shape bow in this lay-out. This was previously

not known. This could change when considering a more complete system, since important simplifications were assumed for this work.

When applying the methodology to more complex multi-physics cases, e.g. involving multi-phase flow, contact with the supporting structures or seismic scenarios, the smoothness of the input-output relation is an important aspect to consider and if needed a UQ method suitable for non-smooth behaviour can be selected. A second important aspect is the computational cost, which can be tackled on the level of a single evaluation, by making the used models more efficient, or on the level of the number of evaluations, which is determined by the UQ methodology.

Funded by
the European Union

Bibliography

Abderrahim, H. A. (2020). Realization of a new large research infrastructure in Belgium: MYRRHA contribution for closing the nuclear fuel cycle making nuclear energy sustainable. *EPJ Web Conf.*, *246*, 00012. https://doi.org/10.1051/epjconf/202024600012

Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C. J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C. C., & Martin, O. A. (2023). PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, *9*, e1516.

Anowar, F., Sadaoui, S., & Selim, B. (2021). Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne). *Computer Science Review*, *40*, 100378.

Banerjee, S., Krahl, R., Durst, F., & Zenger, C. (2007). Presentation of anisotropy properties of turbulence, invariants versus eigenvalue approaches. *Journal of Turbulence*(8), N32.

Banyay, G. A., Shields, M. D., & Brigham, J. C. (2019). Efficient global sensitivity analysis for flow-induced vibration of a nuclear reactor assembly using Kriging surrogates. *Nuclear Engineering and Design*, *341*, 1-15. https://doi.org/10.1016/j.nucengdes.2018.10.013

Banyay, G. A., Shields, M. D., & Brigham, J. C. (2020). Efficient global sensitivity analysis of structural vibration for a nuclear reactor system subject to nonstationary loading. *Nuclear Engineering and Design*, *361*, 110544. https://doi.org/10.1016/j.nucengdes.2020.110544

Belligoli, Z., Dwight, R. P., & Eitelberg, G. (2021). Reconstruction of turbulent flows at high Reynolds numbers using data assimilation techniques. *AIAA Journal*, *59*(3), 855-867.

Bertocchi, F., Rohde, M., De Santis, D., Shams, A., Dolfen, H., Degroote, J., & Vierendeels, J. (2020). Fluid-structure interaction of a 7-rods bundle: Benchmarking numerical simulations with experimental data. *Nuclear Engineering and Design*, *356*, 110394. https://doi.org/10.1016/j.nucengdes.2019.110394

Boltax, A., Foster, J. P., Kalinowski, J. E., & Swenson, D. V. (1978). *Design applications of irradiation creep and swelling data*. http://inis.iaea.org/search/search.aspx?orig_q=RN:10441859

Bottcher, J. H., & Hofman, G. L. (1977). *Subassembly bowing experience in EBR-II*. http://inis.iaea.org/search/search.aspx?orig_q=RN:09368244

Brooks, S., Gelman, A., Jones, G., & Meng, X.-L. (2011). *Handbook of markov chain monte carlo*. CRC press.

Christon, M. A., Lu, R., Bakosi, J., Nadiga, B. T., Karoutas, Z., & Berndt, M. (2016). Large-eddy simulation, fuel rod vibration and grid-to-rod fretting in pressurized water reactors. *Journal of Computational Physics*, *322*, 142-161. https://doi.org/10.1016/j.jcp.2016.06.042

Chung, H., & Chen, S.-s. (1977). Vibration of a Group of Circular Cylinders in a Confined Fluid. *Journal of Applied Mechanics*, *44*(2), 213-217. https://doi.org/10.1115/1.3424026

Coles, D. (1969). Computation of turbulent boundary layers. 1968 AFOSR-IFP-Stanford Conference,

de Baar, J., Scholcz, T., & Dwight, R. (2015). Exploiting adjoint derivatives in high-dimensional metamodels. *AIAA Journal*, *53*(5), 1391-1395.

de Lambert, S., Campioni, G., Faucher, V., Leturcq, B., & Cardolaccia, J. (2019). Modeling the consequences of fuel assembly bowing on PWR core neutronics using a Monte-Carlo code. *Annals of Nuclear Energy*, *134*, 330-341. https://doi.org/10.1016/j.anucene.2019.06.017

de Lambert, S., Cardolaccia, J., Faucher, V., Thomine, O., Leturcq, B., & Campioni, G. (2021). Semi-analytical modeling of the flow redistribution upstream from the mixing grids in a context of nuclear fuel assembly bow. *Nuclear Engineering and Design*, *371*, 110940. https://doi.org/10.1016/j.nucengdes.2020.110940

De Ridder, J., Degroote, J., Van Tichelen, K., Schuurmans, P., & Vierendeels, J. (2013). Modal characteristics of a flexible cylinder in turbulent axial flow from numerical simulations. *Journal of Fluids and Structures*, *43*, 110-123. https://doi.org/10.1016/j.jfluidstructs.2013.09.001

De Ridder, J., Degroote, J., Van Tichelen, K., & Vierendeels, J. (2017). Predicting modal characteristics of a cluster of cylinders in axial flow: From potential flow solutions to coupled CFD–CSM calculations. *Journal of Fluids and Structures*, *74*, 90-110. https://doi.org/10.1016/j.jfluidstructs.2017.07.006

De Ridder, J., Doaré, O., Degroote, J., Van Tichelen, K., Schuurmans, P., & Vierendeels, J. (2015). Simulating the fluid forces and fluid-elastic instabilities of a clamped–clamped cylinder in turbulent axial flow. *Journal of Fluids and Structures*, *55*, 139-154. https://doi.org/10.1016/j.jfluidstructs.2015.03.001

De Santis, D., & Shams, A. (2017). Numerical modeling of flow induced vibration of nuclear fuel rods. *Nuclear Engineering and Design*, *320*, 44-56. https://doi.org/10.1016/j.nucengdes.2017.05.013

Degroote, J. (2013). Partitioned Simulation of Fluid-Structure Interaction. *Archives of Computational Methods in Engineering*, *20*(3), 185-238. https://doi.org/10.1007/s11831-013-9085-5

Degroote, J., Bathe, K.-J., & Vierendeels, J. (2009). Performance of a new partitioned procedure versus a monolithic procedure in fluid–structure interaction. *Computers & Structures*, *87*(11), 793-801. https://doi.org/10.1016/j.compstruc.2008.11.013

Delaissé, N., Demeester, T., Haelterman, R., & Degroote, J. (2023). Quasi-Newton Methods for Partitioned Simulation of Fluid–Structure Interaction Reviewed in the Generalized Broyden Framework. *Archives of Computational Methods in Engineering*. https://doi.org/10.1007/s11831-023-09907-y

Dempster, A. P. (2008). Upper and lower probabilities induced by a multivalued mapping. In *Classic works of the Dempster-Shafer theory of belief functions* (pp. 57-72). Springer.

Divaret, L., Cadot, O., Moussou, P., & Doaré, O. (2014). Normal forces exerted upon a long cylinder oscillating in an axial flow. *Journal of Fluid Mechanics*, *752*, 649-669. https://doi.org/10.1017/jfm.2014.342

Dolfen, H., Bertocchi, F., Rohde, M., & Degroote, J. (2019). Vibrations in a 7-rod bundle subject to axial flow: Simulations and experiments. *Nuclear Engineering and Design*, *353*, 110227. https://doi.org/10.1016/j.nucengdes.2019.110227

Dolfen, H., De Ridder, J., Brockmeyer, L., Merzari, E., Kennedy, G., Van Tichelen, K., & Degroote, J. (2022). A multi-stage approach of simulating turbulence-induced vibrations in a wire-wrapped tube bundle for fretting wear prediction. *Journal of Fluids and Structures*, *109*, 103460. https://doi.org/10.1016/j.jfluidstructs.2021.103460

Dolfen, H., Van Hauwermeiren, D., Bral, A., & Degroote, J. (2022, 2022-06-05/2022-06-09). Fluid-structure interaction simulation of a wire-wrapped tube array using overset grids. 8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress 2022), Oslo, Norway.

Funded by
the European Union

Dolfen, H., Vandewalle, S., & Degroote, J. (2023). *Effect of stochastic deformation on the vibration characteristics of a tube bundle in axial flow: code and data* Zenodo. https://doi.org/10.5281/zenodo.7966359

Dwight, R., & Han, Z.-H. (2009). Efficient uncertainty quantification using gradient-enhanced kriging. 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 17th AIAA/ASME/AHS Adaptive Structures Conference 11th AIAA No,

Dwight, R. P., Desmedt, S. G., & Omrani, P. S. (2016). *Sobol indices for dimension adaptivity in sparse grids*. Springer.

Edeling, W., Cinnella, P., & Dwight, R. P. (2014). Predictive RANS simulations via Bayesian model-scenario averaging. *Journal of Computational Physics*, *275*, 65-91.

Edeling, W. N., Cinnella, P., Dwight, R. P., & Bijl, H. (2014). Bayesian estimates of parameter variability in the k–ε turbulence model. *Journal of Computational Physics*, *258*, 73-94.

Edeling, W. N., Dwight, R. P., & Cinnella, P. (2016). Simplex-stochastic collocation method with improved scalability. *Journal of Computational Physics*, *310*, 301-328.

Edeling, W. N., Schmelzer, M., Dwight, R. P., & Cinnella, P. (2018). Bayesian predictions of Reynolds-averaged Navier–Stokes uncertainties using maximum a posteriori estimates. *AIAA Journal*, *56*(5), 2018-2029.

Eidi, A., Ghiassi, R., Yang, X., & Abkar, M. (2021). Model-form uncertainty quantification in RANS simulations of wakes and power losses in wind farms. *Renewable Energy*, *179*, 2212-2223.

Emory, M., Larsson, J., & Iaccarino, G. (2013). Modeling of structural uncertainties in Reynolds-averaged Navier-Stokes closures. *Physics of Fluids*, *25*(11).

Faure, Q., Delipei, G., Petruzzi, A., Avramova, M., & Ivanov, K. (2023). Fuel performance uncertainty quantification and sensitivity analysis in the presence of epistemic and aleatoric sources of uncertainties [Original Research]. *Frontiers in Energy Research*, *11*. https://doi.org/10.3389/fenrg.2023.1112978

Feinberg, J., & Langtangen, H. P. (2015). Chaospy: An open source tool for designing methods of uncertainty quantification. *Journal of Computational Science*, *11*, 46-57. https://doi.org/10.1016/j.jocs.2015.08.008

Ferrari, G., Balasubramanian, P., Le Guisquet, S., Piccagli, L., Karazis, K., Painter, B., & Amabili, M. (2018). Non-linear vibrations of nuclear fuel rods. *Nuclear Engineering and Design*, *338*, 269-283. https://doi.org/10.1016/j.nucengdes.2018.08.013

Fitelson, B. (2005). Bayesian Epistemology. In: JSTOR.

Forrester, A., Sobester, A., & Keane, A. (2008). *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons.

Geffray, C., & Macián-Juan, R. (2015). Multi-scale uncertainty and sensitivity analysis of the TALL-3D experiment. *Nuclear Engineering and Design*, *290*, 154-163. https://doi.org/10.1016/j.nucengdes.2014.12.002

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman and Hall/CRC.

Ghanem, R., & Spanos, P. D. (1990). Polynomial chaos in stochastic finite elements.

Ghanem, R. G., & Spanos, P. D. (1991). Stochastic Finite Element Method: Response Statistics. In *Stochastic Finite Elements: A Spectral Approach* (pp. 101-119). Springer New York. https://doi.org/10.1007/978-1-4612-3094-6_4

Giles, M. B. (2008). Multilevel monte carlo path simulation. *Operations research*, *56*(3), 607-617.

Giles, M. B. (2015). Multilevel Monte Carlo methods. *Acta Numerica*, *24*, 259-328. https://doi.org/10.1017/S096249291500001X

Greenblatt, D., Paschal, K. B., Yao, C.-S., Harris, J., Schaeffler, N. W., & Washburn, A. E. (2006). Experimental investigation of separation control part 1: baseline and steady suction. *AIAA Journal*, *44*(12), 2820-2830.

Haji-Ali, A.-L., Nobile, F., & Tempone, R. (2016). Multi-index Monte Carlo: when sparsity meets sampling. *Numerische Mathematik*, *132*(4), 767-806. https://doi.org/10.1007/s00211-015-0734-5

Hammersley, J. M. (1960). Monte Carlo methods for sulving multivariable problems. *Annals of the New York Academy of Sciences*, *86*(3), 844-874. https://doi.org/10.1111/j.1749-6632.1960.tb42846.x

Hassan, M. A., Weaver, D. S., & Dokainish, M. A. (2003). The effects of support geometry on the turbulence response of loosely supported heat exchanger tubes. *Journal of Fluids and Structures*, *18*(5), 529-554. https://doi.org/10.1016/j.jfluidstructs.2003.08.011

Henderson, S. G., Chiera, B. A., & Cooke, R. M. (2000). Generating" dependent" quasi-random numbers. 2000 Winter Simulation Conference Proceedings (Cat. No. 00CH37165),

Hinze, M., Pinnau, R., Ulbrich, M., & Ulbrich, S. (2008). *Optimization with PDE constraints* (Vol. 23). Springer Science & Business Media.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, *15*(1), 1593-1623.

Hosder, S., Walters, R. W., & Balch, M. (2010). Point-Collocation Nonintrusive Polynomial Chaos Method for Stochastic Computational Fluid Dynamics. *AIAA Journal*, *48*(12), 2721-2730. https://doi.org/10.2514/1.39389

Kass, R. E., Tierney, L., & Kadane, J. B. (1991). Laplace's method in Bayesian analysis. *Contemporary Mathematics*, *115*, 89-99.

Katsuyama, K., Nagamine, T., Matsumoto, S.-i., & Sato, S. (2007). High energy X-ray CT study on the central void formations and the fuel pin deformations of FBR fuel assemblies. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, *255*(2), 365-372. https://doi.org/10.1016/j.nimb.2006.12.087

Kelly, S. G. (2000). *Fundamentals of Mechanical Vibrations*. McGraw-Hill. https://books.google.be/books?id=yVjvAAAAMAAJ

Kennedy, G., Van Tichelen, K., Pacio, J., Di Piazza, I., & Uitslag-Doolaard, H. (2020). Thermal-Hydraulic Experimental Testing of the MYRRHA Wire-Wrapped Fuel Assembly. *Nuclear Technology*, *206*(2), 179-190. https://doi.org/10.1080/00295450.2019.1620539

Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(3), 425-464. https://doi.org/https://doi.org/10.1111/1467-9868.00294

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Kroese, D. P., Taimre, T., & Botev, Z. I. (2011). *Handbook of Monte Carlo Methods*. Wiley. https://books.google.be/books?id=-j3bmyGXKUIC

Kuo, F. Y., & Nuyens, D. (2016). Application of Quasi-Monte Carlo Methods to Elliptic PDEs with Random Diffusion Coefficients: A Survey of Analysis and Implementation. *Foundations of Computational Mathematics*, *16*(6), 1631-1696. https://doi.org/10.1007/s10208-016-9329-5

Launder, B. E., & Sharma, B. I. (1974). Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc. *Letters in heat and mass transfer*, *1*(2), 131-137.

Le Maître, O., & Knio, O. M. (2010). *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*. Springer Netherlands. https://books.google.be/books?id=lOFCquL6SxYC

Lefebvre, T. (2021). On Moment Estimation From Polynomial Chaos Expansion Models. *IEEE Control Systems Letters*, *5*(5), 1519-1524. https://doi.org/10.1109/LCSYS.2020.3040851

Lemieux, C. (2009). *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer New York. https://books.google.nl/books?id=wj5OyydZ5bkC

Leschziner, M. (2015). *Statistical turbulence modelling for fluid dynamics-demystified: an introductory text for graduate engineering students*. World Scientific.

Leturcq, B., Le Tallec, P., Pascal, S., Fandeur, O., & Lamorte, N. (2022). A new reduced order model to represent the creep induced fuel assembly bow in PWR cores. *Nuclear Engineering and Design*, *394*, 111828. https://doi.org/10.1016/j.nucengdes.2022.111828

Ling, J., Kurzawski, A., & Templeton, J. (2016). Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. *Journal of Fluid Mechanics*, *807*, 155-166.

Liu, C., Papukchiev, A., & Macián-Juan, R. (2020). Uncertainty and sensitivity analysis of coupled multiscale simulations in the context of the SESAME EU-Project. *International Journal of Advanced Nuclear Reactor Design and Technology*, *2*, 117-130. https://doi.org/10.1016/j.jandt.2020.11.002

Liu, Q., & Wang, D. (2016). Stein variational gradient descent: A general purpose bayesian inference algorithm. *Advances in neural information processing systems*, *29*.

Loeven, G., Witteveen, J., & Bijl, H. (2007). Probabilistic collocation: an efficient non-intrusive approach for arbitrarily distributed parametric uncertainties. 45th AIAA Aerospace Sciences Meeting and Exhibit,

Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA Journal*, *32*(8), 1598-1605.

Merzari, E., & Ninokata, H. (2011). Proper orthogonal decomposition of the flow in a tight lattice rod-bundle. *Nuclear Engineering and Design*, *241*(11), 4621-4632. https://doi.org/10.1016/j.nucengdes.2010.12.005

Merzari, E., Ninokata, H., & Baglietto, E. (2008). Numerical simulation of flows in tight-lattice fuel bundles. *Nuclear Engineering and Design*, *238*(7), 1703-1719. https://doi.org/10.1016/j.nucengdes.2008.01.001

Mohamed, M. S., & Larue, J. C. (1990). The decay power law in grid-generated turbulence. *Journal of Fluid Mechanics*, *219*, 195-214.

Moussou, P., Guilloux, A., Boccaccio, E., & Ricciardi, G. (2017). Fluid Damping in Fuel Assemblies. ASME 2017 Pressure Vessels and Piping Conference,

Nakajima, S., Watanabe, K., & Sugiyama, M. (2019). *Variational Bayesian learning theory*. Cambridge University Press.

NEA. (2016). *Review of Uncertainty Methods for Computational Fluid Dynamics Application to Nuclear Reactor Thermal Hydraulics*. OECD Publishing.

Ninokata, H., Merzari, E., & Khakim, A. (2009). Analysis of low Reynolds number turbulent flow phenomena in nuclear fuel pin subassemblies of tight lattice configuration. *Nuclear Engineering and Design*, *239*(5), 855-866. https://doi.org/10.1016/j.nucengdes.2008.10.030

Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. Springer.

Oladyshkin, S., & Nowak, W. (2012). Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, *106*, 179-190. https://doi.org/10.1016/j.ress.2012.05.002

Oliver, T. A., & Moser, R. D. (2011). Bayesian uncertainty quantification applied to RANS turbulence models. Journal of Physics: Conference Series,

Paidoussis, M. P. (1966a). Dynamics of flexible slender cylinders in axial flow Part 1. Theory. *Journal of Fluid Mechanics*, *26*(4), 717-736. https://doi.org/10.1017/S0022112066001484

Paidoussis, M. P. (1966b). Dynamics of flexible slender cylinders in axial flow Part 2. Experiments. *Journal of Fluid Mechanics*, *26*(4), 737-751. https://doi.org/10.1017/S0022112066001496

Parish, E. J., & Duraisamy, K. (2016). A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, *305*, 758-774.

Peter, J. E., & Dwight, R. P. (2010). Numerical sensitivity analysis for aerodynamic optimization: A survey of approaches. *Computers & Fluids*, *39*(3), 373-391.

Pettigrew, M. J., Taylor, C. E., Fisher, N. J., Yetisir, M., & Smith, B. A. W. (1998). Flow-induced vibration: recent findings and open questions. *Nuclear Engineering and Design*, *185*(2), 249-276. https://doi.org/10.1016/S0029-5493(98)00238-6

Platteeuw, P., Loeven, G., & Bijl, H. (2008). Uncertainty quantification applied to the k-epsilon model of turbulence using the probabilistic collocation method. 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, 16th AIAA/ASME/AHS Adaptive Structures Conference, 10th AIAA Non-Deterministic Approaches Conference, 9th AIAA Gossamer Spacecraft Forum, 4th AIAA Multidisciplinary Design Optimization Specialists Conference,

Powell, M. J. (1973). On search directions for minimization algorithms. *Mathematical programming*, *4*, 193-201.

Puragliesi, R., Mukin, R., Clifford, I., Ferroukhi, H., & Seidl, M. (2019, August 18-23). Comparison of Computational Fluid Dynamics and Subchannel Numerical Solutions of Fuel Assemblies Characterised by Bowing. 18th International Topical Meeting on Nuclear Reactor Thermal Hydraulics (NURETH-18) Portland, OR.

Robbe, P., Nuyens, D., & Vandewalle, S. (2019). Recycling Samples in the Multigrid Multilevel (Quasi-)Monte Carlo Method. *SIAM Journal on Scientific Computing*, *41*(5), S37-S60. https://doi.org/10.1137/18m1194031

Robert, C. P., Casella, G., & Casella, G. (1999). *Monte Carlo statistical methods* (Vol. 2). Springer.

Roelofs, F. (2020). Liquid metal thermal hydraulics: State-of-the-art and future perspectives. *Nuclear Engineering and Design*, *362*, 110590. https://doi.org/10.1016/j.nucengdes.2020.110590

Roelofs, F., Dovizio, D., Uitslag-Doolaard, H., De Santis, D., Mathur, A., Mikuz, B., & Shams, A. (2019). Core thermal hydraulic CFD support for liquid metal reactors. *Nuclear Engineering and Design*, *355*, 110322. https://doi.org/10.1016/j.nucengdes.2019.110322

Rubiolo, P. R., & Young, M. Y. (2009). On the factors affecting the fretting-wear risk of PWR fuel assemblies. *Nuclear Engineering and Design*, *239*(1), 68-79. https://doi.org/10.1016/j.nucengdes.2008.08.021

Sakai, K., Okubo, Y., & Hishida, H. (1978). Three-dimensional deflection analyses of wire-spaced fuel pin bundles under temperature and hydrodynamic force fields with irradiation effects. *Nuclear Engineering and Design*, *48*(2), 595-610. https://doi.org/10.1016/0029-5493(78)90101-2

Schlatter, P., & Örlü, R. (2010). Assessment of direct numerical simulation data of turbulent boundary layers. *Journal of Fluid Mechanics*, *659*, 116-126.

Schmelzer, M., Dwight, R. P., & Cinnella, P. (2020). Discovery of algebraic Reynolds-stress models using sparse symbolic regression. *Flow, Turbulence and Combustion*, *104*, 579-603.

Sciacchitano, A. (2019). Uncertainty quantification in particle image velocimetry. *Measurement Science and Technology*, *30*(9), 092001.

Scott, D. W. (2015). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.

Shafer, G. (1976). *A mathematical theory of evidence* (Vol. 42). Princeton university press.

Shams, A., Mikuž, B., & Roelofs, F. (2018). Numerical prediction of flow and heat transfer in a loosely spaced bare rod bundle. *International Journal of Heat and Fluid Flow*, *73*, 42-62. https://doi.org/10.1016/j.ijheatfluidflow.2018.07.006

Shams, A., Roelofs, F., Baglietto, E., & Komen, E. M. J. (2018). High fidelity numerical simulations of an infinite wire-wrapped fuel assembly. *Nuclear Engineering and Design*, *335*, 441-459. https://doi.org/10.1016/j.nucengdes.2018.06.012

Singh, A. P., & Duraisamy, K. (2016). Using field inversion to quantify functional errors in turbulence closures. *Physics of Fluids*, *28*(4).

Soize, C. (2005). Random matrix theory for modeling uncertainties in computational mechanics. *Computer methods in applied mechanics and engineering*, *194*(12-16), 1333-1366.

Sosnovsky, E., Baglietto, E., Keijers, S., Van Tichelen, K., De Souza, T. C., Doolaard, H., & Roelofs, F. (2015, August 30-September 4). CFD simulations to determine the effects of deformations on liquid metal cooled wire wrapped fuel assemblies. 16th International Topical Meeting on Nuclear Reactor Thermal Hydraulics 2015 (NURETH-16), Chicago, IL.

Spalart, P., & Allmaras, S. (1992). A one-equation turbulence model for aerodynamic flows. 30th aerospace sciences meeting and exhibit,

Steiner, J., Viré, A., & Dwight, R. P. (2022). Classifying regions of high model error within a data-driven RANS closure: Application to wind turbine wakes. *Flow, Turbulence and Combustion*, *109*(3), 545-570.

Tao, T. (2013). *Compactness and contradiction*. American Mathematical Soc.

ter Hofstede, E., Kottapalli, S., & Shams, A. (2017). Numerical prediction of flow induced vibrations in nuclear reactor applications. *Nuclear Engineering and Design*, *319*, 81-90. https://doi.org/10.1016/j.nucengdes.2017.04.026

Toti, A., Vierendeels, J., & Belloni, F. (2018). Extension and application on a pool-type test facility of a system thermal-hydraulic/CFD coupling method for transient flow analyses. *Nuclear Engineering and Design*, *331*, 83-96. https://doi.org/10.1016/j.nucengdes.2018.02.004

Uwaba, T., Ohshima, H., & Ito, M. (2017). Analyses of deformation and thermal-hydraulics within a wire-wrapped fuel subassembly in a liquid metal fast reactor by the coupled code system. *Nuclear Engineering and Design*, *317*, 133-145. https://doi.org/10.1016/j.nucengdes.2017.03.026

Van Der Laan, M. P., Sørensen, N. N., Réthoré, P. E., Mann, J., Kelly, M. C., Troldborg, N., Schepers, J. G., & Machefaux, E. (2015). An improved k-ϵ model applied to a wind turbine wake in atmospheric turbulence. *Wind Energy*, *18*(5), 889-907.

Wanninger, A. (2018). *Mechanical Analysis of the Bow Deformation of Fuel Assemblies in a Pressurized Water Reactor Core* Technische Universität München].

Wanninger, A., Seidl, M., & Macián-Juan, R. (2018). Mechanical analysis of the bow deformation of a row of fuel assemblies in a PWR core. *Nuclear Engineering and Technology*, *50*(2), 297-305. https://doi.org/10.1016/j.net.2017.12.009

Wilcox, D. C. (1993). Comparison of two-equation turbulence models for boundary layers with pressure gradient. *AIAA Journal*, *31*(8), 1414-1421.

Wilcox, D. C. (1998). *Turbulence modeling for CFD* (Vol. 2). DCW industries La Canada, CA.

Wilcox, D. C. (2008). Formulation of the kw turbulence model revisited. *AIAA Journal*, *46*(11), 2823-2838.

Witteveen, J. A., Loeven, A., & Bijl, H. (2009). An adaptive stochastic finite elements approach based on Newton–Cotes quadrature in simplex elements. *Computers & Fluids*, *38*(6), 1270-1288.

Wu, J.-L., Xiao, H., & Paterson, E. (2018). Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework. *Physical Review Fluids*, *3*(7), 074602.

Xiao, H., Wang, J.-X., & Ghanem, R. G. (2017). A random matrix approach for quantifying model-form uncertainties in turbulence modeling. *Computer Methods in Applied Mechanics and Engineering*, *313*, 941-965. https://doi.org/https://doi.org/10.1016/j.cma.2016.10.025

Xiu, D. (2010). *Numerical Methods for Stochastic Computations*

A Spectral Method Approach. Princeton University Press. https://doi.org/10.2307/j.ctv7h0skv

Xiu, D., & Karniadakis, G. E. (2002). The Wiener--Askey Polynomial Chaos for Stochastic Differential Equations. *SIAM Journal on Scientific Computing*, *24*(2), 619-644. https://doi.org/10.1137/s1064827501387826

Ye, X., Gao, F., Chen, X., Liu, X., & Wei, Y. G. (2020). Full-scale finite element analysis of deformation and contact of a wire-wrapped fuel bundle subject to realistic thermal and irradiation conditions. *Nuclear Engineering and Design*, *364*, 110676. https://doi.org/10.1016/j.nucengdes.2020.110676

Yu, H., Girimaji, S. S., & Luo, L.-S. (2005). DNS and LES of decaying isotropic turbulence with and without frame rotation using lattice Boltzmann method. *Journal of Computational Physics*, *209*(2), 599-616.

Zadeh, L. A., Klir, G. J., & Yuan, B. (1996). *Fuzzy sets, fuzzy logic, and fuzzy systems: selected papers* (Vol. 6). World scientific.