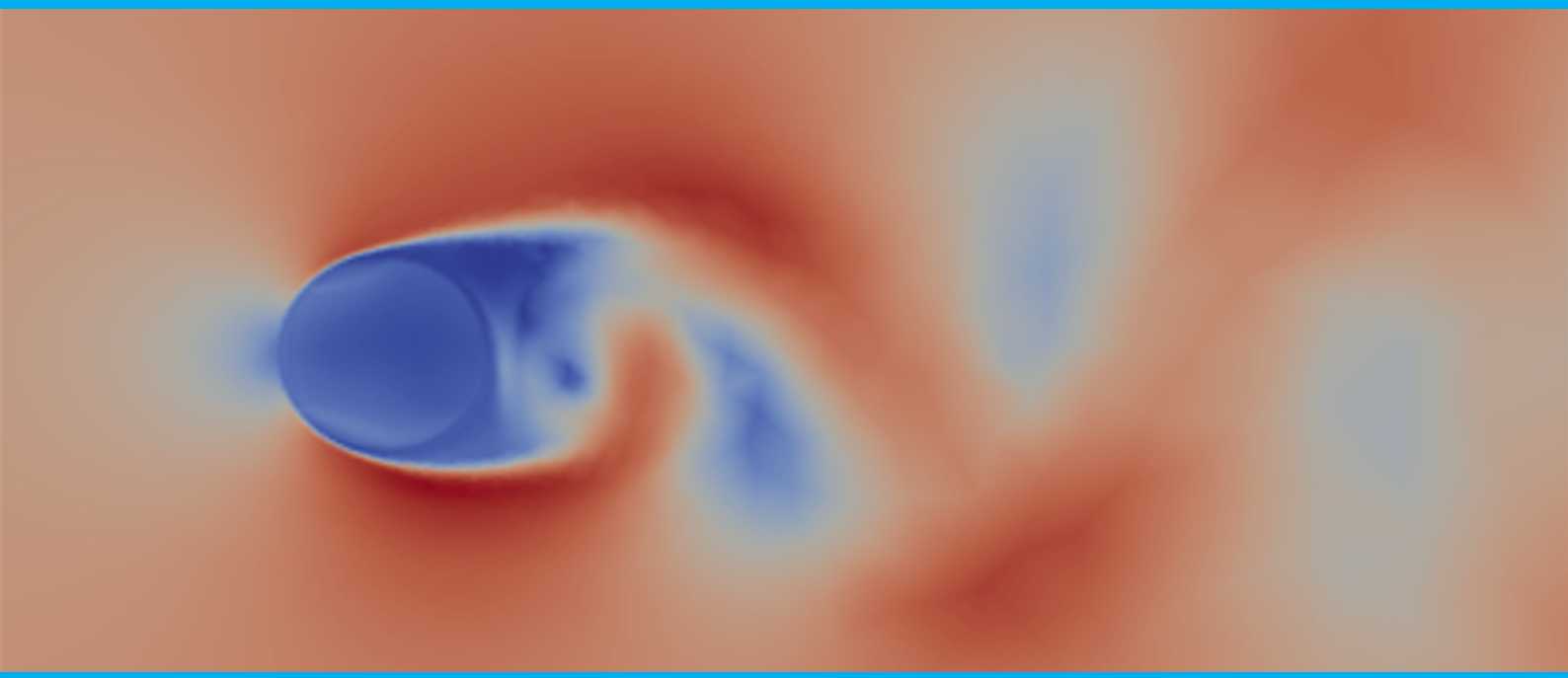
Exploratory research on application Multi-Level Multi-Fidelity Monte Carlo in fluid dynamics topics: study on flow past a porous cylinder

Master Thesis

Jan Modderman



TU Delft

Exploratory research on application Multi-Level Multi-Fidelity Monte Carlo in fluid dynamics topics: study on flow past a porous cylinder

MASTER THESIS

A thesis submitted to the Delft University of Technology in partial fulfillment
of the requirements for the degree of

Master of Science in Offshore & Dredging Engineering

by

Jan Modderman

15^{th} of October 2021

Supervisors: Prof. dr. ir. A. Metrikine
Dr. ir. J. O. Colomés Gené
Dr. ir. A. C. Viré
Ir. M. Vergassola

The work in this thesis was made in the
Offshore Engineering Group
Department of Hydraulic Engineering
Faculty of Civil Engineering & Geosciences
Delft University of Technology

Abstract

In offshore engineering complex simulation models are constructed for design optimization using Monte Carlo methods. These models incur large computational costs. Multi-Level Multi-Fidelity Monte Carlo is proposed as a method to reduce the computational cost of these simulations. In addition, research is conducted on the use of porous media as passive damping systems. Hence, an analysis on the effect of porosity on the vortex shedding frequency is conducted. This thesis is an exploratory investigation on the application of Multi-Level Multi-Fidelity Monte Carlo in fluid dynamics topics and its particular use for analysis of the effect of porosity on the vortex shedding frequency on a porous circular cylinder. Three case studies are conducted. Firstly, applying Multi-Level Multi-Fidelity Monte Carlo on a solid circular cylinder case, which is deemed as a successful application, based on the estimated quantity of interest, variance reduction and computational cost reduction. Furthermore, two parametric studies are conducted: 1) to discover empirical relationships (low-fidelity models) and 2) forward uncertainty propagation with Multi-Level Multi-Fidelity Monte Carlo using a uniform input distribution. Both parametric studies consist of a number of equally distributed points of porosity on a case setup of flow past a porous circular cylinder. The parametric studies use a frequency detection algorithm, which approximates the vortex shedding frequency using the frequency of lift force oscillation. The results of the first parametric study indicate there is a drop in vortex shedding frequency as experienced by the cylinder for increasing porosity. The hypothesis is that for increasing porosity the formation length of vortex shedding increases. Two empirical relationships are derived from the results by curve fitting the Strouhal number (dimensionless form of the vortex shedding frequency) versus porosity. These empirical relationships are incorporated in the Multi-Level Multi-Fidelity Monte Carlo method and applied to a similar parametric study on the effect of porosity on the vortex shedding frequency. The results indicate the presence of systemic errors in the high-fidelity model. The conjecture is that the major influence on these errors is due to the resolution of the frequency detection algorithm being too low. For this reason, no clear conclusion on the validity of the empirical relationships is obtained and further research is required.

"*Essentially, all models are wrong, but some are useful.*"

- **George E.P. Box, statistician**

Preface

I would like to sincerely express my gratitude to my daily supervisors Oriol Colomés-Gené and Marco Vergassola for their advice, feedback, guidance, patience and support during the research process. Spending a number of afternoons on getting a personal lecture on uncertainty, trying to install *Dakota* on *Cartesius*, joining in on meetings on Marco's research progress and working together to construct the geometry and mesh for a porous cylinder are memories that will stay with me. I would like to thank the other committee members Andrei Metrikine and Axelle Viré for providing interesting discussions and personal feedback. I wholeheartedly appreciate my friends and family for the endless time and support. Thank you all.

Contents

List of Abbreviations vi

List of Symbols vii

List of Figures viii

List of Tables xii

1 Introduction 1
- 1.1 Scope & research objectives . . . 1
- 1.2 Research strategy . . . 2
- 1.3 Thesis outline . . . 2

2 Literature research 3
- 2.1 Physics of fluid dynamics . . . 4
  - 2.1.1 Relevant equations of fluid motion. . . . 4
  - 2.1.2 Navier-Stokes for incompressible Newtonian fluids . . . 4
  - 2.1.3 Flow around a circular cylinder . . . 5
- 2.2 Numerical implementation . . . 8
  - 2.2.1 Finite Volume Method . . . 8
  - 2.2.2 Porosity modeling . . . 9
- 2.3 Uncertainty quantification . . . 10
  - 2.3.1 Introduction to parameter uncertainty. . . . 10
  - 2.3.2 Forward propagation of uncertainty . . . 11
  - 2.3.3 Multi-Level Multi-Fidelity Monte Carlo theory . . . 11
- 2.4 Multi-fidelity models . . . 17
  - 2.4.1 Definition . . . 17
  - 2.4.2 Low-fidelity models for flow around a circular cylinder . . . 17

3 Methodologies 20
- 3.1 Algorithms . . . 20
- 3.2 General methodologies . . . 22
  - 3.2.1 OpenFOAM . . . 22
  - 3.2.2 Gmsh . . . 22
  - 3.2.3 Dakota . . . 22
  - 3.2.4 Post processing in Python . . . 26

4 Experimental setup 27
- 4.1 Case study: benchmark cases . . . 28
  - 4.1.1 Case descriptions and validation. . . . 29
  - 4.1.2 Uncertain benchmark case in 2D and 3D . . . 32
  - 4.1.3 Multi-Level Multi-Fidelity Monte Carlo . . . 33
- 4.2 Parametric study on a porous cylinder . . . 34
  - 4.2.1 Case description . . . 34
  - 4.2.2 Parametric study. . . . 37
  - 4.2.3 Curve fitting . . . 38
- 4.3 Parametric study using Multi-Level Multi-Fidelity Monte Carlo . . . 38
  - 4.3.1 Case description . . . 39
  - 4.3.2 Parametric study. . . . 39
  - 4.3.3 Interval analysis using Multi-Level Multi-Fidelity Monte Carlo. . . . 40

5 Results: benchmark cases 42
5.1 Relative computational cost. . . . 42
5.2 Multi-Level Monte Carlo 43
5.2.1 Case in 2D 44
5.2.2 Case in 3D 46
5.3 Multi-Level Multi-Fidelity Monte Carlo 48
5.4 True value of estimated drag coefficient. . . . 53

6 Results: parametric study 55
6.1 Summarized results of the parametric study 55
6.1.1 Outlier case with $\beta = 0.15$ & $Re = 20100$ 56
6.1.2 Outlier case with $\beta = 0.30$ & $Re = 33500$ 57
6.1.3 Outlier case with $\beta = 0.25$ & $Re = 30150$ 59
6.1.4 Outlier case with $\beta = 0.20$ & $Re = 3350$ 60
6.1.5 Interpretation of the physics of flow past a porous cylinder 66
6.1.6 Remarks on vortex shedding frequency detection algorithm 66
6.1.7 Summarized results after inspection of outlier cases 66
6.2 Data-fitting 68
6.2.1 Gompertz function. . . . 68
6.2.2 Hill function 69
6.2.3 Weibull function 69
6.3 Low-fidelity models 70

7 Results: interval analysis using MLMFMC 72
7.1 Relative computational cost. . . . 72
7.2 Data points from the interval analysis. . . . 72
7.3 Pearson correlation coefficient 75

8 Conclusions & recommendations 77
8.1 Discussion 77
8.2 Conclusion 78
8.3 Recommendations 79

A Fit functions 80

B Data fit plots 81

C Correlation plots 88

Referenced documents 90

List of Abbreviations

CFD: Computational Fluid Dynamics

FFT: Fast Fourier Transform

FVM: Finite Volume Method

HF: High-Fidelity

HLHFMC: High-Level High-Fidelity Monte Carlo

HLMC: High-Level Monte Carlo

LF: Low-Fidelity

MCM: Monte Carlo Methods

MFCV: Multi-Fidelity Control Variates

MLMC: Multi-Level Monte Carlo

MLMFMC: Multi-Level Multi-Fidelity Monte Carlo

MSE: Mean Squared Error

NaN: Not-a-Number

PDF: Probability Distribution Function

List of Symbols

ρ: density for fluid dynamics or Pearson correlation coefficient for statistics

t: time

u: velocity vector

F: external force vector

p: pressure

μ: dynamic viscosity

D: cylinder diameter

L: characteristic length scale

ν: kinematic viscosity

Re: Reynolds number

f_{vs}: vortex shedding frequency

U: characteristic velocity

St: Strouhal number

F_d: drag force

F_l: lift force

C_d: drag coefficient

C_l: lift coefficient

$A_{\parallel}$: reference area parallel to the flow

$A_{\perp}$: reference area perpendicular to the flow

L_c: cylinder height

β: porosity (pore volume divided by total volume)

$\bar{U}$: mean inflow velocity

N_{pilot}: number of pilot samples

ϵ: tolerance

C_l: relative compuational cost of resolution level l

Y_l: difference function of resolution level l

N_l: optimal number of samples at resolution level l

Q: quantity of interest

l: resolution level from 0 to L

ρ_l: Pearson correlation coefficient at resolution level l

St_{cyl}: cylindrical Strouhal number (definition on page 62)

x^*: formation length

f_{lf}: frequency of lift force oscillation

List of Figures

2.1 Overview of the layout of the literature review. . . . 3
2.2 Flow around a circular cylinder denoted by purple flow lines with inflow from the left side. The cylinder diameter is denoted by D (blue double headed arrow) and the lift force F_l in the vertical direction (red arrow) and drag force F_d in the horizontal direction (green arrow). . . . 5
2.3 Several flow patterns for flow around a circular cylinder for varying ranges of Reynolds numbers (*Goharzadeh et al.* [11]). . . . 6
2.4 Relationship between the Reynolds and the Strouhal number for circular cylinders with smooth and rough surfaces (*Aderne* [12]). . . . 7
2.5 Relation between drag coefficient C_d and Reynolds number Re for a circular cylinder with a smooth surface, based on experimental data. (*Tritton* [13]). . . . 7
2.6 Relation between drag coefficient C_d and porosity β for a circular cylinder depicted by the circular dotted line. (*Steiros et al.* [14]). . . . 8
2.7 General format of a Monte Carlo Methods (MCM) algorithm. A range of possible input values is selected, from which N samples are taken using the probability density function (PDF). For each input value a model evaluation is then executed and the outputs from each evaluation are aggregated to obtain the results of the Monte Carlo Methods algorithm. . . . 11
2.8 The general idea of the Multi-Level Monte Carlo approach is substituting a single resolution of a single model in the Monte Carlo simulation by a group of multiple resolutions. . . . 13
2.9 The general idea of the Control Variates Monte Carlo approach is substituting a single model in the Monte Carlo simulation by a group of multi-fidelity models. . . . 14
2.10 The general idea of the Multi-Level Multi-Fidelity Monte Carlo Monte Carlo approach is substituting a single model in the Monte Carlo simulation by a group of multi-fidelity models with varying resolution levels for the high-fidelity model and low-fidelity models if applicable for the low-fidelity model. . . . 16
2.11 Eight empirical relations from theory between the Strouhal and Reynolds numbers. The names in between the brackets denote the data set on which the fit function was applied to. . . . 19

3.1 Scheme of the Multi-Level Monte Carlo (MLMC) algorithm. . . . 21
3.2 Scheme of the Multi-Level Multi-Fidelity Monte Carlo (MLMFMC) algorithm. . . . 22

4.1 Overview of the three experiments conducted in this study. First a benchmark case is used to implement the Multi-Level Multi-Fidelity Monte Carlo method. Then a parametric study using a porous circular cylinder is conducted in order to derive low-fidelity models. And, finally, these two sections are combined by creating a new case with a porous circular cylinder, to which the Multi-Level Multi-Fidelity is applied. . . . 28
4.2 Overview of the section regarding the implementation of MLMF using the benchmark cases. Starting with the case description as described in literature, the reconstructed cases in *OpenFOAM* and the validation of the reconstructed cases. Following with the implementation of the Multi-Level Monte Carlo method using the benchmark cases and lastly the implementation of the Multi-Level Multi-Fidelity method using the benchmark cases. . . . 28
4.3 Geometry of all 2D benchmark cases. Total size of rectangular plane is 2.2x0.41 m with a cylinder of 0.1 m diameter and no slip boundary conditions imposed on the top side, bottom side and the cylinder, zero gradient boundary conditions on the outlet and a spatial parabolic inflow condition on the inlet. . . . 29
4.4 The meshes of the 2D cases with lowest to highest resolution level from top to bottom respectively. 30
4.5 Geometry of all 3D benchmark cases. Total size of rectangular plane is 2.2x0.41x0.41 m with a cylinder of 0.41 m height and 0.1 m diameter and no slip boundary conditions imposed on the top side, bottom side, left side, right side and the cylinder, a zero gradient boundary condition on the outlet and a spatial parabolic inflow condition on the inlet. . . . 31

4.6 The meshes of the 3D cases with lowest to highest resolution level from top to bottom respectively. 32
4.7 Schematic overview of the Multi-Level Monte Carlo approach using a uniformly distributed mean inflow velocity ranging from 0.2 – 1.0 m/s to estimate the drag coefficient. 33
4.8 Schematic overview of the Multi-Level Multi-Fidelity Monte Carlo approach using a uniformly distributed mean inflow velocity ranging from 0.2 – 1.0 m/s to estimate the drag coefficient. . . 33
4.9 Overview of the section regarding the parametric study on a porous circular cylinder. Starting with the case description as described in literature. Following with the setup of the parametric study and lastly the curve fitting on the resulting data to obtain low-fidelity models which can be applied in cases for porous circular cylinders. 34
4.10 Case setup for the parametric study with a porous circular cylinder of diameter $D = 0.0335$ m and a wall thickness of 0.001 m. 35
4.11 Mesh of the case for the parametric study. 35
4.12 Resulting drag coefficient and Strouhal number versus porosity for the validation case with $Re = 6700$. 36
4.13 Flow patterns of validation study on porous circular cylinder for varyin porosity. 36
4.14 Schematic overview of the process of estimating the vortex shedding frequency. First the time signal is interpolated in order to have constant time step. Then the interpolated time signal is transformed to the frequency domain using the Fast Fourier Transform (FFT). Lastly, the frequency corresponding to the largest peak is found, which corresponds to the vortex shedding frequency. 38
4.15 Overview of the section regarding the interval analysis using Multi-Level Multi-Fidelity Monte Carlo methods on the parametric study. First the case description and parametric study are described. Lastly, the interval analysis using Multi-Level Multi-Fidelity Monte Carlo methods is described. 39
4.16 Geometry the 3D benchmark case with porous circular cylinder. Total size of rectangular plane is 2.2x0.41 m with a cylinder with a diameter of 0.1 m and no slip boundary conditions imposed on the top side, bottom side and the cylinder, zero gradient boundary conditions on the outlet and a uniform constant inflow condition on the inlet. 39
4.17 The Multi-Level Multi-Fidelity Monte Carlo method is implemented on the parametric study in order to reduce suspected epistemic uncertainty by estimating each point using interval analysis. 40
4.18 The meshes of the 3D cases with lowest to highest resolution level from top to bottom respectively. 41

5.1 Histogram of drag coefficient as a result of general Monte Carlo Methods for number of pilot samples $N_{pilot} = 10$ for three resolution levels, hence 50 total samples. Color coding is applied to demonstrate which resolution level corresponds to which result. 45
5.2 Histogram of the drag coefficient as a result of a Multi-Level Monte Carlo approach in 2D with three resolution levels, denoted by res0, res1 or res2. Color coding is applied to demonstrate which resolution level corresponds to which result. 45
5.3 Histogram as a result of the high level Monte Carlo run with 56 samples using a 2D case setup. . 46
5.4 Histogram of drag coefficient as a result of general Monte Carlo Methods for number of pilot samples $N_{pilot} = 10$ for three resolution levels, hence 50 total samples. Color coding is applied to demonstrate which resolution level corresponds to which result. 47
5.5 Histogram of the drag coefficient as a result of a Multi-Level Monte Carlo approach in 3D with three resolution levels, denoted by res0, res1 or res2. Color coding is applied to demonstrate which resolution level corresponds to which result. 47
5.6 Histogram of High-Level Monte Carlo simulation of 50 samples using a 3D case setup. 48
5.7 Histogram of drag coefficient as a result of general Monte Carlo Methods for number of pilot samples $N_{pilot} = 10$ for two model fidelities with three resolution levels each, hence 100 total samples. Color coding is applied to demonstrate which model fidelity and resolution level corresponds to which result. 49
5.8 Scatter plot of high-fidelity model (CFD 3D) vs low-fidelity model (CFD 2D) for varying resolution levels $l = 0, 1, 2$ with a best fit line in red. 50
5.9 Histogram of the drag coefficient as a result of a Multi-Level Multi-Fidelity approach with two model fidelities with three resolution levels each. The estimated drag coefficient is $C_d = 3.89$ with a standard deviation of $\sigma_{C_d} = 0.73$ 51

5.10 Histogram of the drag coefficient as a result of a Multi-Level Multi-Fidelity approach with two model fidelities, denoted by 2D or 3D, with three resolution levels each, denoted by res0, res1 or res2. . . . 51
5.11 Zoomed in histogram of the drag coefficient as a result of a Multi-Level Multi-Fidelity approach with two model fidelities, denoted by 2D or 3D, with three resolution levels each, denoted by res0, res1 or res2. . . . 52
5.12 Histogram of the drag coefficient as a result of a general Monte Carlo simulation with $N = 53$ samples using a high-resolution high-fidelity model. The estimated drag coefficient is $C_d = 4.27$ with a standard deviation of $\sigma_{C_d} = 0.79$ . . . 53
5.13 Scatter plot of all the estimated drag coefficients. . . . 53

6.1 Scatter plot of the parametric study. With porosity and Reynolds numbers and a colorbar which denotes the estimated Strouhal number. . . . 55
6.2 Surface plot of the parametric study with Strouhal number versus porosity and Reynolds numbers. . . . 56
6.3 Scatter plot of the parametric study with green circles indicating the outlier cases, which are inspected. . . . 56
6.4 Time signal and corresponding Fast Fourier Transform of the first outlier case $\beta = 0.15$ & $Re = 20100$. . . . 57
6.5 Flow pattern of the second outlier case ($\beta = 0.30$ & $Re = 33500$) at experimental time $t = 12.5$ s. . 58
6.6 Time signal of the lift coefficient for the second outlier case ($\beta = 0.30$ & $Re = 33500$). . . . 58
6.7 Frequency signal of the lift coefficient for the second outlier case ($\beta = 0.30$ & $Re = 33500$) with the red dashed line corresponding to the detected frequency peak. . . . 58
6.8 Flow pattern of the third outlier case ($\beta = 0.25$ & $Re = 30150$) at experimental time $t = 12.5$ s. . . 59
6.9 Time signal of the lift coefficient for the third outlier case ($\beta = 0.25$ & $Re = 30150$). . . . 59
6.10 Frequency signal of the lift coefficient for the third outlier case ($\beta = 0.25$ & $Re = 30150$) with the red dashed line corresponding to the detected frequency peak. . . . 60
6.11 Flow pattern of the fourth outlier case ($\beta = 0.20$ & $Re = 3350$) at experimental time $t = 12.5$ s. . . 60
6.12 Time signal of the lift coefficient for the fourth outlier case. . . . 61
6.13 Frequency signal of the lift coefficient for the fourth outlier case with the red dashed line corresponding to the detected frequency peak. . . . 61
6.14 Flow patterns with highlighted vortices at experimental time $t = 12.5$ s of varying porosity β at Reynolds number $Re = 3350$. . . . 62
6.15 Time signals of lift coefficient of varying porosity β at Reynolds number $Re = 3350$. **Important note:** the bottom image vertical axis has an added value of +1.868 for the lift coefficient, yet the ticks are presented in this form due to the small differences of the lift coefficient for this case. . 63
6.16 Results from a study by*Sun et al.* [49] on wake flow structures of screen cylinders using PIV. . . . 65
6.17 Results from a study by*Sun et al.* [50] on the development of wake flow structures behind screen cylinders. . . . 66
6.18 Scatter plot of the parametric study. With porosity and Reynolds numbers and a colorbar which denotes the estimated Strouhal number. . . . 67
6.19 Surface plot of the parametric study with Strouhal number versus porosity and Reynolds numbers. . . . 67
6.20 Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using a Gompertz fit function. . . . 69
6.21 Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using a Hill fit function. . . . 69
6.22 Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using a Hill fit function. . . . 70
6.23 Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using the calculated Gompertz model with $a = 0.26$, $b = -155.39$ and $c = 0.26$. . . . 71
6.24 Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using the calculated Hill model with $a = 0.26$, $b = 0.25$ and $c = -76.27$. . . . 71

7.1 Estimated cylindrical Strouhal number versus porosity using the Gompertz model as the low-fidelity model. 73
7.2 Estimated cylindrical Strouhal number versus porosity using the Hill model as the low-fidelity model. 73
7.3 Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.22$. 74
7.4 Difference functions of pilot samples at porosity $\beta = 0.30$. 74
7.5 Pearson correlation coefficient versus porosity using the Gompertz model as the low-fidelity model. 75
7.6 Pearson correlation coefficient versus porosity using the Hill model as the low-fidelity model. . 76

C.1 Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.28$. 88
C.2 Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.29$. 88
C.3 Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.30$. 89

List of Tables

2.1 Overview of fit parameters A and B as derived by *Roushan et al.* [40] 18

4.1 Overview of the relevant parameters for the three benchmarks cases in 2D 29
4.2 Overview of the results of validation for the three benchmarks cases in 2D 30
4.3 Overview of the relevant parameters for the three benchmarks cases in 3D 31
4.4 Overview of the results of validation for the three benchmarks cases in 3D 32

5.3 Estimated relative computational cost of each resolution level for the 2D benchmark case. . . . 42
5.4 Estimated relative computational cost of each resolution level for the 3D benchmark case. . . . 42
5.1 Computational times for 16 benchmark cases at three resolution levels with total summed value, maximum value, minimum value, multiplicative factor between minimum and maximum value and mean value for the 2D benchmark case. 43
5.5 Total estimated relative computational cost of each resolution level of the low-fidelity (2D) model and the high fidelity (3D) model for Multi-Level Multi-Fidelity Monte Carlo. 43
5.2 Computational times for 16 benchmark cases at three resolution levels with total summed value, maximum value, minimum value, multiplicative factor between minimum and maximum value and mean value for the 3D benchmark case. 44
5.6 Summary of input variables of the Multi-Level Monte Carlo simulation in 2D. Displaying the estimated relative computational cost of each resolution level for a Multi-Level Monte Carlo approach in 2D. The random variable is the mean inflow velocity U_m ranging from $U_m = 0.2$ to $U_m = 1.0$ m/s, samples are drawn from a uniform distribution. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$. 44
5.7 Summary of estimated relative computational cost, optimal number of samples and total cost of each resolution level of the low-fidelity (2D) model for a Multi-Level Monte Carlo approach. . 46
5.8 Summary of input variables of the Multi-Level Monte Carlo simulation in 3D. Displaying the estimated relative computational cost of each resolution level for a Multi-Level Monte Carlo approach in 3D. The random variable is the mean inflow velocity U_m ranging from $\bar{U} = 0.2$ to $\bar{U} = 1.0$ m/s, samples are drawn from a uniform distribution. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$. 46
5.9 Summary of estimated relative computational cost, optimal number of samples and total cost of each resolution level of the low-fidelity (2D) model for a Multi-Level Monte Carlo approach. . 48
5.10 Summary of input variables of the Multi-Level Multi-Fidelity simulation. Displaying the estimated relative computational cost of each resolution level of the low-fidelity (2D) model and the high-fidelity (3D) model for a Multi-Level Multi-Fidelity approach. The random variable is the mean inflow velocity U_m ranging from $\bar{U} = 0.2$ to $\bar{U} = 1.0$ m/s, samples are drawn from a uniform distribution. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$. 48
5.11 Summary of estimated relative computational cost, optimal number of samples and total cost of each resolution level of the low-fidelity (2D) model and the high-fidelity (3D) model for a Multi-Level Multi-Fidelity approach. 52
5.12 Summary of results of the first case study. 54

6.1 Summary of changes to the results of the parametric study due to outlier cases. 67

7.1 Total estimated relative computational cost of the low-fidelity model (equation) and the high fidelity (3D) model for Multi-Level Multi-Fidelity Monte Carlo. 72

1 Introduction

In the offshore environment increasingly complex simulation methods are being developed, coming closer and closer to approximating reality [1]. These methods are complex in nature, due to the uncertainties inherent to the offshore environment (e.g. uncertainty in wind and waves) and due to the coupling of models in order to correctly simulate the behaviour of an offshore structure. An example of a complex model is the floating offshore wind turbine [2], which in general is subject to separate models for mooring, hydrodynamics and aerodynamics. The downside of highly realistic models is the increase in computational resources required for simulation and subsequent increase in costs. Because of the inherent uncertainty in offshore, many simulations have to be run in order to correctly capture the reaction of the structure due to environmental forces. It is not known in advance how many simulations are required to be run for appropriate results.

This combination of high costs and unknown number of simulations for sufficient results requires a new approach of simulation. In this work the Multi-Level Multi-Fidelity Monte Carlo method is proposed as a solution to this problem. In order to study and subsequently apply the Multi-Level Multi-Fidelity Monte Carlo method [31] a simple fluid dynamics topic is chosen. To actively contribute to the research field of fluid dynamics, the case of flow past a porous circular cylinder is applied. Porous media have been the subject of research in coastal engineering for some time now, since there are many coastal applications in which porosity plays a clear role, e.g. mangroves [3] and breakwater rubble mounds [4]. However in the offshore environment research is also conducted in order to study the possible reduction of environmental forces by applying partially porous cylinders [5] [6] [7] [8]. Moreover, research is conducted on porous media for passive damping systems [16] [18]. Therefore, this work focuses on advanced Monte Carlo methods on the one hand and flow past a porous circular cylinder on the other.

1.1. Scope & research objectives

The aim of this research is to understand the process of applying Multi-Level Multi-Fidelity Monte Carlo on a fluid dynamics topic. The fluid dynamics topic is chosen in such a way that a contribution to research is made, whilst under the restriction that the fluid dynamics case is relatively simple. This simplicity is quantified as simple geometries and simple fluid flow effects. Initially the plan was to study the effect of porosity on the drag coefficient for a porous circular cylinder. However, after validation this plan was changed to focus on the effect of porosity on the vortex shedding frequency for a porous circular cylinder. Therefore, the main research question posed is:

How can Multi-Level Multi-Fidelity Monte Carlo be applied in fluid dynamics topics and what is the contribution of its application on the effect of porosity on the vortex shedding frequency for a porous circular cylinder?

Sub-questions

Furthermore the following sub-questions are addressed in this study to accomplish the main research goal.

1. What is a low-fidelity model and are there low-fidelity models available from literature for application in the parametric study?
2. What steps are taken when applying Multi-Level Multi-Fidelity Monte Carlo methods on a simple fluid dynamics case and how do the computational cost, estimated quantity of interest and estimated variance compare to general Monte Carlo methods?
3. What is the effect of varying porosity on the vortex shedding frequency of a porous circular cylinder and can an empirical relationship be discovered?
4. How can low-fidelity models be applied using Multi-Level Multi-Fidelity Monte Carlo and how do these reflect on the validity of this approach?

1.2. Research strategy

In order to answer the posed research questions, eight research objectives are set. These are the milestones of this work and serve as an overview of the research approach.

Research objectives

1. Literature review on fluid dynamics and CFD;
2. Literature review on parameter uncertainty and Monte Carlo methods;
3. Obtaining low-fidelity models that could qualify for application using Multi-Level Multi-Fidelity Monte Carlo from theory;
4. Application of Multi-Level Monte Carlo and Multi-Level Multi-Fidelity Monte Carlo on a benchmark case to validate algorithm functionality;
5. Literature review on porous cylinders and porosity modeling;
6. Setting up a parametric study to study the effect of porosity on the vortex shedding frequency;
7. Deriving low-fidelity models by curve-fitting data resulting from the parametric study for a porous circular cylinder;
8. Application of the Multi-Level Multi-Fidelity Monte Carlo approach for the case of a porous circular cylinder with varying low-fidelity models.

1.3. Thesis outline

In chapter 2 the background is presented on the two generalized categories of this study, namely the flow around a porous circular cylinder and uncertainty quantification using Multi-Level Multi-Fidelity Monte Carlo methods. In chapter 3 the implementation of algorithms, software used and relevant functions from software are summarized. In chapter 4 the experimental setup of all three numerical experiments conducted in this study are described: 1) application of Multi-Level Multi-Fidelity Monte Carlo, 2) parametric study on flow past a porous circular cylinder and 3) parametric study on flow past a porous circular cylinder using Multi-Level Multi-Fidelity Monte Carlo. In chapter 5 the results of the first experiment on application of the Multi-Level Multi-Fidelity Monte Carlo method are described and analysed. In chapter 6 the parametric study on flow past a porous cylinder is presented and an in-depth analysis is conducted. In chapter 7 a new -but comparable- parametric study is conducted using the third experimental setup in order to explore the possibilities of Multi-Level Multi-Fidelity Monte Carlo. In chapter 8 the key findings are summarized and recommendation are given on future research topics.

2

Literature research

The literature research part of this study is split up in two separate categories namely, the physical problem formulation in both the real and computational domain and the statistical aspect, which focuses on uncertainty quantification and means to reduce computational cost of simulations. In Figure 2.1 the two categories are denoted by the dashed line boxes. Both categories will unite when implementing the case studies.

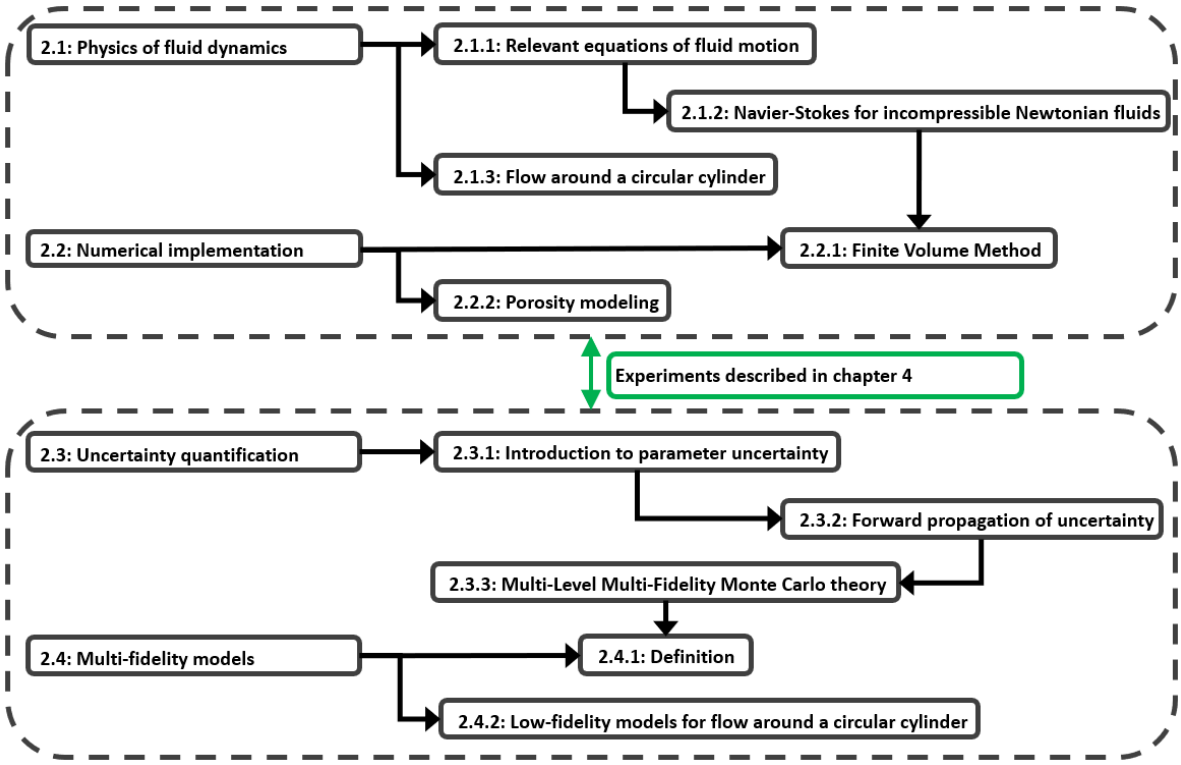


Figure 2.1: Overview of the layout of the literature review.

In section 2.1 the relevant physics of fluid dynamics are described, starting with the general equations for fluid dynamics and the assumptions which reduce those general equations to the incompressible Navier-Stokes and continuity equations in subsection 2.1.1 and subsection 2.1.2 respectively. In subsection 2.1.3 the flow around a circular cylinder is described, since this is the base case for the case studies conducted in this work. Therefore, it is important to have a proper notion on the drag and lift forces and the corresponding drag and lift coefficients for analysis in the studies. Furthermore a number of relevant dimensionless numbers - Reynolds and Strouhal number- and a short summary of a study on porous circular cylinders are described.

In section 2.2 the discretization of the physical problem is described. The finite volume method is briefly described with focus on the discretization of the Navier-Stokes equations for incompressible flows in subsection 2.2.1. Then, in subsection 2.2.2 a porosity model is described which is used in the parametric and final case study to simulate a thinly walled porous circular cylinder, instead of simulating the pores of the cylinder using a high resolution mesh in *OpenFOAM*.

In section 2.3 the process of uncertainty quantification related to this work is described. Starting with subsection 2.3.1 to introduce parameter uncertainty, including formal definitions of aleatoric and epistemic uncertainty. Next, in subsection 2.3.2 uncertainty quantification using forward uncertainty propagation is described. Concluding with subsection 2.3.3 on Multi-Level Multi-Fidelity Monte Carlo methods, in which regular Monte Carlo Methods are introduced. Furthermore two variance reduction techniques, namely: Multi-Level Monte Carlo and Control Variates are described separately and these are later combined to form the Multi-Level Multi-Fidelity Monte Carlo method.

In section 2.4 Multi-Fidelity models explained in more detail. In subsection 2.4.1 fidelity is defined and multi-fidelity models are properly defined as high-fidelity and low-fidelity models. Lastly, in subsection 2.4.2, a number of low-fidelity models are presented from literature which are applicable for the case of fluid flow around a circular cylinder.

2.1. Physics of fluid dynamics

The fundamental equations of fluid motion and the Navier-Stokes equation for incompressible Newtonian fluids are described, since these equations lie at the basis of any fluid dynamics topics. Furthermore, a section is dedicated to describing flow around a circular cylinder, as all case studies conducted in this work are in their essence studies of flow around a circular cylinder. This includes a description of flow patterns, relevant forces, like the drag and lift force, the corresponding dimensionless drag and lift coefficients and the dimensionless Reynolds and Strouhal numbers. Lastly, results from a recent published study on the effect of porosity on the drag coefficient for a porous circular cylinder are briefly summarized.

2.1.1. Relevant equations of fluid motion

Starting from a physics point of view the three most fundamental conservation principles for fluid dynamics are: the conservation of mass, conservation of linear momentum and conservation of energy. The conservation of energy equation is neglected in this work, since the focus is on incompressible flows. In that case the energy equation can be rewritten into a temperature equation which is completely decoupled from the conservation of mass and conservation of momentum equations, assuming that the viscosity of the fluid is temperature independent [9].

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{2.1}$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u}\mathbf{u}) = \mathbf{F} - \nabla p + \mu \nabla^2 \mathbf{u} \tag{2.2}$$

The equations of conservation of mass and conservation of momentum are presented in Equation 2.1 and Equation 2.2 respectively. Here, $\frac{\partial}{\partial t}$ is the time derivative and ∇ is the differential operator in Cartesian coordinates on three dimensional Euclidean space. Furthermore ρ is the water density; $\mathbf{u}$ is the three dimensional velocity vector; $\mathbf{F}$ is the force vector consisting of external forces; p is the pressure and μ is the dynamic viscosity.

2.1.2. Navier-Stokes for incompressible Newtonian fluids

The next step is to derive the Navier-Stokes equations for incompressible Newtonian fluids. For simplicity, the density and viscosity differences in water are assumed to be negligible. For this case, the density and viscosity are assumed to be constant, therefore Equation 2.1 and Equation 2.2 are reduced to the continuity equation and the Navier-Stokes equations for incompressible Newtonian fluids.

$$\begin{cases} \nabla \cdot \mathbf{u} = 0 \\ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = \mathbf{F} - \frac{\nabla p}{\rho} + \frac{\mu}{\rho}\nabla^2 \mathbf{u} \end{cases} \tag{2.3}$$

The upper equation in Equation 2.3 is the continuity equation and corresponds to the conservation of mass equation and the lower equation corresponds to the conservation of momentum equation. For the

momentum equation the first two terms on the left-hand side of the equation correspond to the inertial force. The pressure term and viscosity term on the right-hand side correspond to the pressure force and viscous force respectively [9].

2.1.3. Flow around a circular cylinder

A goal of flow analysis is to solve the Navier-Stokes equations and determine a relationship between the pressure and flow velocity, whilst subjected to certain geometric boundary conditions. Computational Fluid Dynamics (CFD) is a tool implemented to solve the Navier-Stokes equations numerically, which is applied in this study to the case of flow around a circular cylinder. The case is depicted in Figure 2.2 and it consists of a submerged circular cylinder placed in water with unidirectional flow.

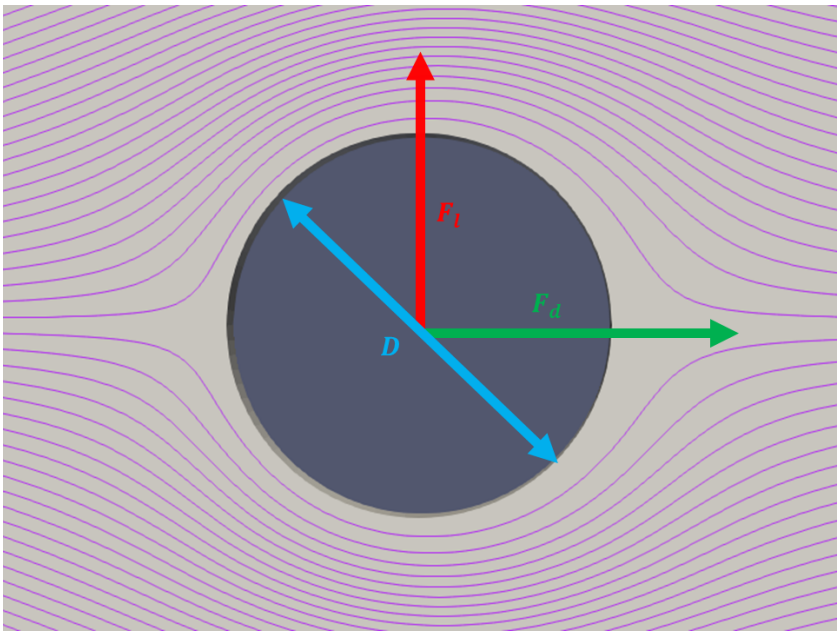


Figure 2.2: Flow around a circular cylinder denoted by purple flow lines with inflow from the left side. The cylinder diameter is denoted by D (blue double headed arrow) and the lift force F_l in the vertical direction (red arrow) and drag force F_d in the horizontal direction (green arrow).

The cylinder has diameter D and the drag and lift forces are denoted by F_d and F_l respectively. The lift force is the force that is perpendicular direction of incoming flow. The drag force is a force acting in the opposite direction of the relative motion of the object. Noteworthy, the lift force is always perpendicular with the drag force. In this case the cylinder is stationary and the fluid moves past the cylinder. The drag force differs from regular friction in the sense that the drag force has a dependency on the flow velocity. This dependency is either linear or quadratic, depending on the cylinder diameter, fluid viscosity, flow velocity and fluid density. Another way to summarize the aforementioned parameters in a singular variable is by using the Reynolds number.

Reynolds number

The Reynolds number (Re) is the ratio of the inertial forces over the viscous forces.

$$Re = \frac{UL}{\nu} = \frac{UD}{\nu} \tag{2.4}$$

The Reynolds number is given in Equation 2.4. Here U is the inflow velocity, L is a characteristic length scale, which is equal to the diameter of the cylinder D for the case of a circular cylinder and ν is the kinematic viscosity. In addition to the Reynolds number, the law of similarity states that for equal Reynolds number, the flow patterns should be the same. Hence when the geometry and the flow velocity are changed, but the Reynolds number remains constant, then the flow pattern should remain the same. The Reynolds number can be seen as a dimensionless flow velocity. A selection of the observed flow patterns are presented in Figure 2.3 [11].

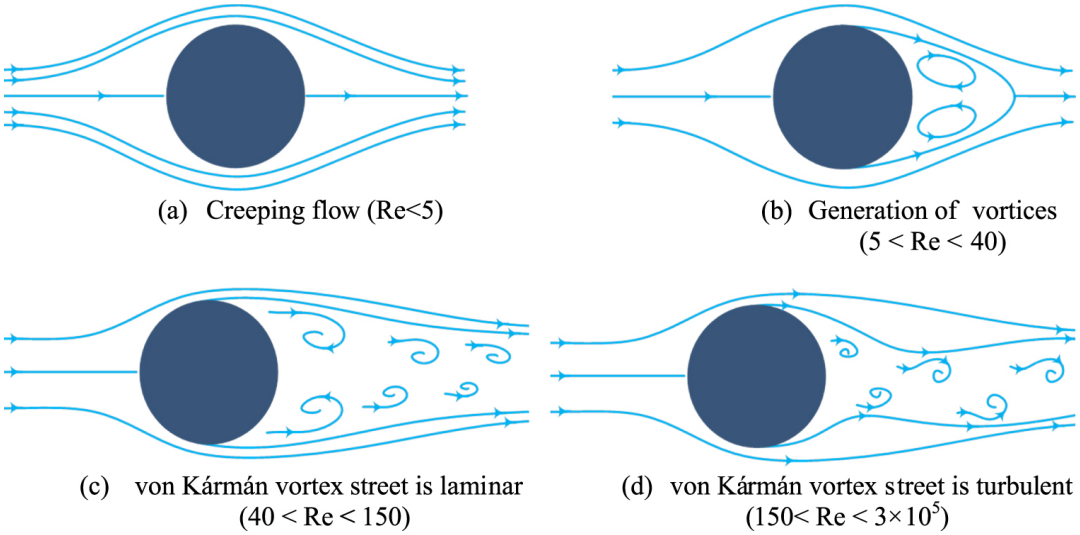


Figure 2.3: Several flow patterns for flow around a circular cylinder for varying ranges of Reynolds numbers (*Goharzadeh et al.* [11]).

At very low Reynolds number the flow is laminar and symmetrical (Figure 2.3 a). For increasing Reynolds number, two bound vortices will start to appear, which are initially stable and symmetrical (Figure 2.3 b). However at $Re = 40$ these vortices become instable and will start to detach from the back of the cylinder, which is called vortex shedding. This process of vortex shedding forms a so called Von Kármán vortex street in which the upper and lower vortex are shed from the back of the cylinder periodically and in an asymmetrical manner Figure 2.3 c & d). The frequency at which the vortex shedding occurs is of interest, since the vortex shedding causes lateral forces on the cylinder due to the asymmetrical flow pattern and thus asymmetrical pressure distribution. This could cause resonance in case the vortex shedding frequency and natural frequency of the cylinder are similar.

Strouhal number

The frequency of vortex shedding is also represented in a dimensionless form as the Strouhal number (St). The Strouhal number is used to describe oscillating flow mechanisms [9].

$$St = \frac{f_{vs} D}{U} \tag{2.5}$$

In Equation 2.5 the Stouhal number is given. Here, f_{vs} is the vortex shedding frequency, D is the cylinder diameter and U is the flow velocity. Research has been conducted towards a relationship between the Reynolds and the Strouhal number for flow around a circular cylinder with smooth and rough surfaces, as is depicted in Figure 2.4 [12]. Interestingly, the Strouhal number remains about constant around $St = 0.2$ for a range of Reynolds numbers: $1.5 \cdot 10^2 < Re < 10^5$.

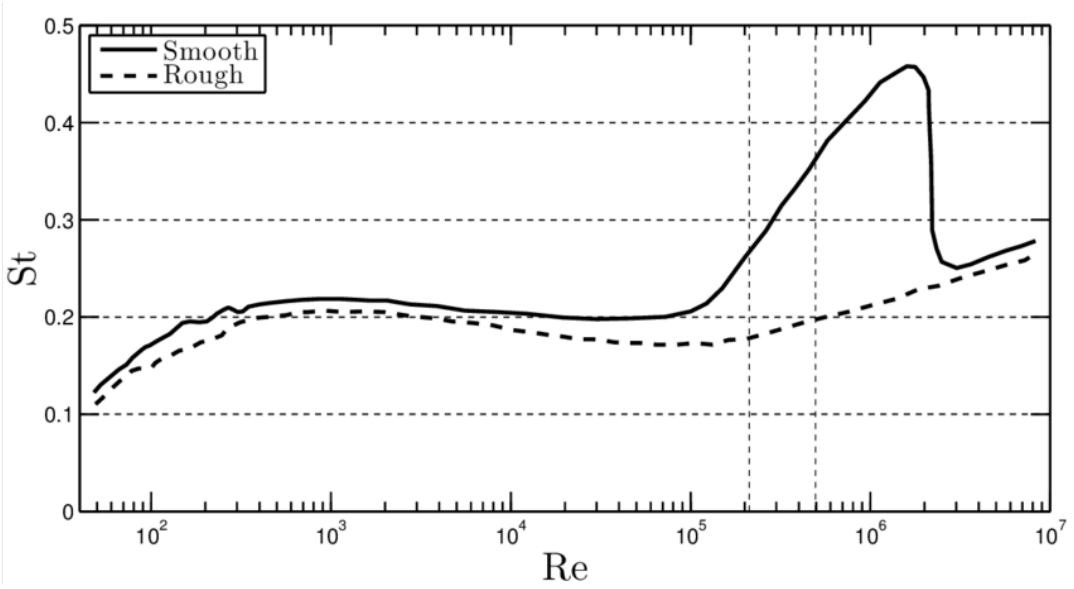


Figure 2.4: Relationship between the Reynolds and the Strouhal number for circular cylinders with smooth and rough surfaces (*Aderne* [12]).

Drag & lift coefficient

In general, the drag and lift force can be represented in a dimensionless form as the drag, C_d, and lift, C_l, coefficient respectively.

$$C_d = \frac{2F_d}{\rho A_{\parallel} U^2} \tag{2.6}$$

$$C_l = \frac{2F_l}{\rho A_{\perp} U^2} \tag{2.7}$$

In Equation 2.6 and Equation 2.7 A is the reference area with the subscript denoting whether it is in the direction parallel or perpendicular to the flow. For the case of a circular cylinder, the reference area is equal for both drag and lift coefficient and is calculated as: $A_{\parallel} = A_{\perp} = L_c D$. Here, L_c is the length or height of the cylinder and D is the cylinder diameter.

Furthermore the textbook by *Tritton* [13] presents the following figure, in which the relation between the drag coefficient C_d and the Reynolds number Re for a circular cylinder is depicted. The curve is based on experimental work (for a list of references of the experiments, see *Tritton* [13] chapter 3.4: drag on page 33).

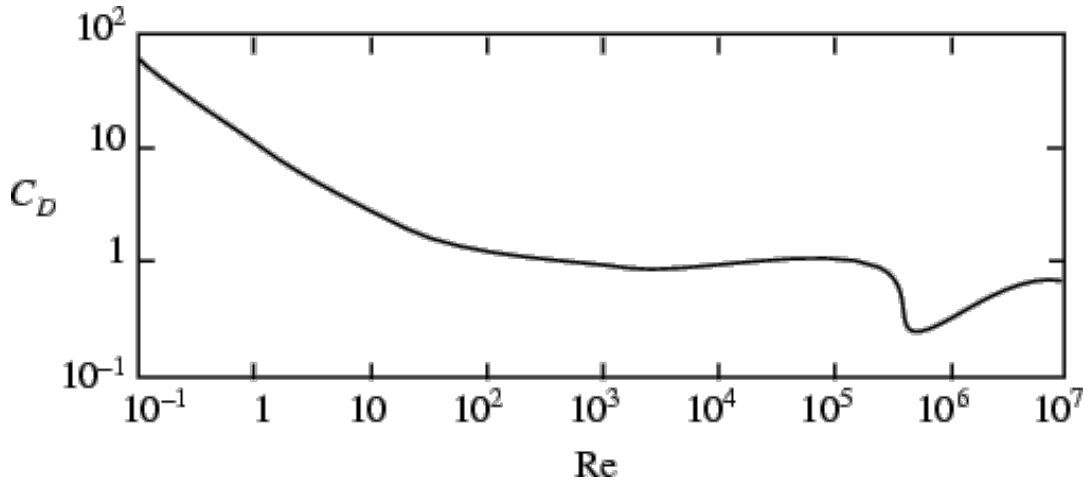


Figure 2.5: Relation between drag coefficient C_d and Reynolds number Re for a circular cylinder with a smooth surface, based on experimental data. (*Tritton* [13]).

For the region $Re < 1$ the drag coefficient is inversely proportional to the Reynolds number, from $1 < Re < 2000$ the drag coefficient decreases to about $C_d = 1$. Then for the region $2000 < Re < 3 \cdot 10^5$ the drag coefficient appears to remain constant around $C_d = 1$, after which a steep decrease in drag coefficient occurs to $C_d = 0.3$, which increases thereafter as the Reynolds number increases.

Porous cylinder and drag coefficient

Lastly, the results of a study on flow past a porous circular cylinder [14] are summarized in order to provide a validation case for the implementation in later sections. The effects of porosity and pore shape on the drag coefficient is studied for fully submerged square and circular cylinders with steady laminar flow in a water basin. The porosity is varied from 0 to 60 % and varying values for cylinder diameter and pore diameter were used. The results of this experiment for a circular cylinder are depicted in Figure 2.6 [14].

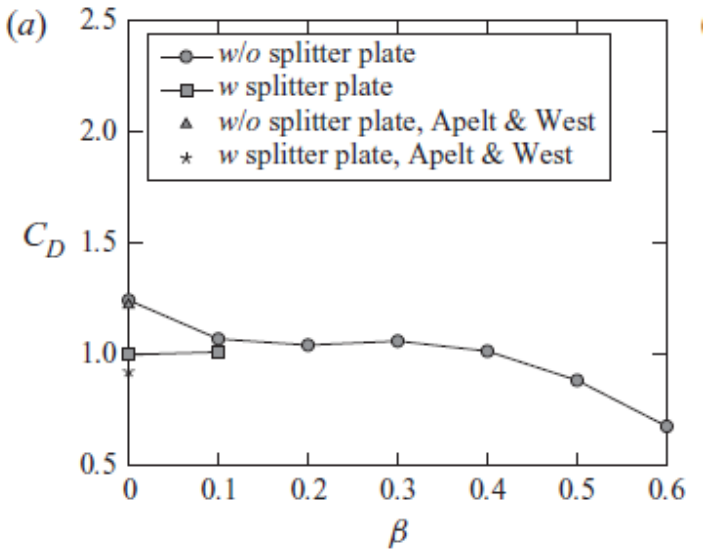


Figure 2.6: Relation between drag coefficient C_d and porosity β for a circular cylinder depicted by the circular dotted line. (*Steiros et al.* [14]).

The Reynolds number for this experiment is $Re = 6700$ with an inflow velocity of $U = 0.2$ m/s and cylinder diameter of $D = 0.0335$ m. Initially at porosity $\beta = 0$ the drag coefficient is at a maximum of $C_d = 1.3$, which decreases to a value of $C_d = 1$ for increasing porosity up to a porosity of $\beta = 0.4$, after which the drag coefficient decreases even more down to about $C_d = 0.7$.

2.2. Numerical implementation

The Navier-Stokes equations as stated before need to be discretized in order to be applicable in a virtual environment. There are several ways of achieving this discretization, however for CFD the norm is usually to apply the Finite Volume Method. Secondly, in preparation for the parametric and second case study, a porosity model is introduced for thinly walled porous circular cylinders.

2.2.1. Finite Volume Method

One of the most applied discretization methods used in CFD is the Finite Volume Method (FVM) [9]. At the basis of this discretization method lies the integral form of a conservation law. The goal is to divide the entire domain into many smaller domains named control volumes, then for each control volume the integral conservation law is approximated.

$$\rho \frac{\partial}{\partial t} \iiint_{V_p} \mathbf{u} dV + \rho \iiint_{V_p} \nabla \cdot (\mathbf{uu}) dV = \iiint_{V_p} \mathbf{F} dV - \iiint_{V_p} \nabla p dV + \mu \iiint_{V_p} \nabla^2 \mathbf{u} dV \tag{2.8}$$

Applying the integral form to the momentum equation part of Equation 2.3 results in Equation 2.8. The FVM applies Gauss' theorem to the integral conservation laws [9]. The definition of Gauss' theorem is presented in Equation 2.9. It states that the volume integral over the divergence of some vector field A is equal

to the closed surface integral of some vector field A. In other words, the net flux out of the volume is given by the sum of all sources of the vector field within that same volume.

$$\iiint_V (\nabla \cdot \mathbf{A}) dV = \oiint_{S(V)} \mathbf{A} \cdot d\mathbf{S} \tag{2.9}$$

$$\rho \frac{\partial}{\partial t} \iiint_{V_p} \mathbf{u} dV + \rho \oiint_{S(V_p)} \mathbf{u}(\mathbf{u} \cdot d\mathbf{S}) = \iiint_{V_p} \mathbf{F} dV - \oiint_{S(V_p)} p d\mathbf{S} + \mu \oiint_{S(V_p)} \nabla \mathbf{u} \cdot d\mathbf{S} \tag{2.10}$$

Applying Gauss' theorem to all terms with the nabla operator in Equation 2.8 results in Equation 2.10. The integral $\iiint_{V_p} dV$ corresponds to volume integrals over control volume V_p and $\oiint_{S(V_p)} d\mathbf{S}$ corresponds to closed surface integrals over the closed surface of the control volume $S(V_p)$. The next step in the discretization is the approximation of the integrals, derivatives, gradients and divergence.

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(m_i) \Delta x = \int_a^b f(x) dx \tag{2.11}$$

For the integrals the midpoint rule is often applied [9], which is given in Equation 2.11 with $\Delta x = \frac{b-a}{n}$. On the left-hand side it states that the sum of function $f(x)$ taken in each point $x = m_i$ is equal to the integral of function $f(x)$ when the number of points taken approaches infinity. For a full list of possible schemes for the temporal derivative, gradient and divergence the reader is referred to the *OpenFOAM* User Guide[1].

2.2.2. Porosity modeling

There are several possibilities for modeling porous media, the most intuitive being creating a mesh in which the case is realistically modeled one to one with the physical case (constructing each pore in the mesh). However, this leads to extreme local refinement in the mesh and high computational costs. In an effort to model porous media with lower computational costs, several porosity models have been proposed using general fluid modeling methods. These fluid modeling methods are the relatively novel Lagrangian method named Smoothed Particle Hydrodynamics (SPH) [15] and the relatively more established Eulerian methods based on Reynolds-Averaged Navier-Stokes (RANS) equations.

In this work the latter is implemented using the porosity model, as originally described by *del Jesus et al.* [16]. Moreover, the original porosity model has been implemented in OpenFOAM by *Higuera et al.* [4] in an external package named *olaFlow* [17]. Another implementation of this method in *OpenFOAM* for thin walled porous structures has been done by *Feichtner et al.* [18]. This porosity model for thin walled porous structures in *OpenFOAM* is implemented in this study.

The porosity model is based on the RANS equations with the additional step of volume averaging, which is called Volume-Averaged Reynolds-Averaged Navier-Stokes (VARANS). In other words: the porosity model implements an averaged pressure loss over the porous media by adding sink terms to the momentum equation. Furthermore, it incorporates porosity in the convective, diffusive, external force and pressure terms in the region of the porous medium.

$$\frac{\partial \mathbf{u}_D}{\partial t} + (\mathbf{u}_D \cdot \nabla) \frac{\mathbf{u}_D}{\beta} = \beta \mathbf{F} - \frac{\beta}{\rho} \nabla p + \frac{\beta \mu}{\rho} \nabla^2 \frac{\mathbf{u}_D}{\beta} - I(A, B, C) \tag{2.12}$$

$$I(A, B, C) = A \mathbf{u}_D + B \mathbf{u}_D |\mathbf{u}_D| + C \frac{\partial \mathbf{u}_D}{\partial t} \tag{2.13}$$

The rewritten momentum equation (from Equation 2.3) is presented in Equation 2.12. Here, $\mathbf{u}_D$ is the Darcy velocity, which is related to the intrinsic velocity in the pores by $\mathbf{u}_i = \frac{\mathbf{u}_D}{\beta}$, where β is the porosity and the pressure sink terms are summarized in the term $I(A, B, C)$. This is presented in Equation 2.13 and depends on coefficients A, B and C.

$$A = \psi \frac{(1-\beta)^3}{\beta^2} \frac{\mu}{D_{50}^2} \tag{2.14}$$

[1] A categorized list of schemes is located at `https://www.openfoam.com/documentation/guides/latest/doc/guide-schemes.html`. Last accessed on: 12-08-2021

$$B = \chi(1 + \frac{7.5}{\mathrm{KC}})\frac{1-\beta}{\beta^3}\frac{\rho}{D_{50}} \tag{2.15}$$

$$C = \frac{c}{\Delta x} \tag{2.16}$$

The coefficients A, B and C are defined in Equation 2.14, Equation 2.15 and Equation 2.16 and these terms correspond to a linear, quadratic and temporal sink terms respectively. The mean nominal diameter of the porous material is represented by D_{50}; KC is the Keulegan-Carpenter number, which relates the drag and inertial forces for bodies in oscillatory flow, c is an inertial coefficient and Δx is the wall thickness.

$$\psi = \frac{\beta^2}{l\Delta x(1-\beta)^3} \tag{2.17}$$

$$\chi = \frac{\beta^3 C_f}{17\Delta x(1-\beta)} \tag{2.18}$$

The coefficients ψ and χ are denoted in Equation 2.17 and Equation 2.18 respectively, where l is a length coefficient related to the hole size of the porous medium and C_f is a coefficient, given as $C_f = \frac{1-\beta}{\delta\beta^2}$ with an assumed value of $\delta = 0.5$. For this work $D_{50} = 1$, KC = 1, $\psi = 0$ and $c = 0$ as is done by *Feichtner et al.* [18].

2.3. Uncertainty quantification

According to *Gabriel Terejanu* uncertainty quantification is:

"The theoretical and computational fabric that connects the three pillars of science – theory, experimentation, and computation – through which uncertainties are characterized and informed to guide the scientific discovery and decision-making process.". [2]

Due to this broad definition there are many applications for uncertainty quantification. In order to work with uncertainty quantification, a general description on parameter uncertainty is presented in subsection 2.3.1. In subsection 2.3.2 forward propagation of uncertainty is described, since this is part of the uncertainty quantification process. Often, sampling is used for propagation of uncertainty and therefore Monte Carlo Methods and more advanced Monte Carlo implementations are explained in subsection 2.3.3.

2.3.1. Introduction to parameter uncertainty

In physics and engineering, uncertainty usually refers to a range of values in which the true value is somewhere present: this is also known as uncertainty of measurement. Other ways to express uncertainty are by error, mean squared error, standard uncertainty, Type A evaluation or Type B evaluation [19]. When introducing modeling uncertainties, it is favorable to categorize all uncertainties as either aleatory or epistemic, these are known as irreducible and reducible uncertainty respectively [20]. Aleatoric uncertainty is the intrinsic uncertainty due to the non-deterministic nature of the phenomenon, these phenomena can only be reduced to a probabilistic model. On the other hand, epistemic uncertainty is due to a lack of knowledge or incomplete information about the topic. It is possible to obtain more information about epistemic uncertainty from existing data, which can be used to decrease uncertainty used for simulation [20]. Epistemic uncertainty does not have to be modelled using a probabilistic model and it can for some cases be seen as knowing the range of values in which the true value is present. Thus it is favorable to separate the uncertainties into aleatory and epistemic categories for modeling, since epistemic uncertainty could be reduced using extra information. Furthermore, epistemic uncertainty could be introduced to a simulation due to coupling of components, modeling assumptions and modeling approximations [21].

[2] This quote is obtained from the webpage by Gabriel Terejanu -associate professor of computer science at the university of North Carolina Charlotte- active in the uncertainty quantification research group. The URL is located at: `https://www.uncertaintyquantification.org/` Last accessed on 25-08-2021

2.3.2. Forward propagation of uncertainty

According to *Geraci et al.* [22] the basic framework of uncertainty quantification is identifying and characterizing uncertain input parameters using forward propagation of uncertainty. Forward propagation of uncertainty is described as representing the uncertainty of the input parameters using a probability distributions. Then drawing a number of samples from the distributions of the input parameters and executing the simulations using those drawn samples. Lastly, the output quantities corresponding to those samples are analyzed in order to understand the statistics of these output quantities. Using probability distributions to represent uncertain input parameters suggests that the uncertainty being quantified or propagated is aleatoric. However, this does not have to be true. According to *Swiler et al.* [23] it is possible to propagate epistemic uncertainty. One of the methods implemented for propagation of epistemic uncertainty is interval analysis using sampling. Here, it is assumed that nothing is known about the epistemic uncertain input parameters, except that these parameters are enclosed in a certain interval. Uncertainty is introduced to this interval -usually using a uniform distribution- and the process of forward propagation of uncertainty is conducted. Drawbacks of this form of interval analysis are:

1. No cumulative distribution function of the output can be constructed; and
2. due to sampling the true bounds of epistemic uncertainty are often underestimated [23].

2.3.3. Multi-Level Multi-Fidelity Monte Carlo theory

The goal of this section is to describe and derive a Multi-Level Multi-Fidelity Monte Carlo (MLMFMC) approach. According to *Sin et al.* [24] Monte Carlo Methods (MCM) are becoming more accepted and established as a research approach. However, currently one of the main drawbacks is that the highly realistic models require a lot of computational resources for a single simulation. Hence a suggestion is made to implement an MLMFMC approach to reduce this drawback of MCM [25]. Firstly, there is a short description regarding MCM. Then both the Multi-Level Monte Carlo (MLMC) method and the Multi-Fidelity Control Variate (MFCV) method are described separately. In the last subsection, the MLMFMC approach is constructed by combining the MLMC and MFCV methods.

Monte Carlo Methods

The Monte Carlo Methods (MCM) are algorithms using repeated random sampling in order to solve mathematical problems [26]. Currently, MCM are used extensively for uncertainty quantification [27]. In general, an MCM algorithm behaves according to the process described in Figure 2.7.

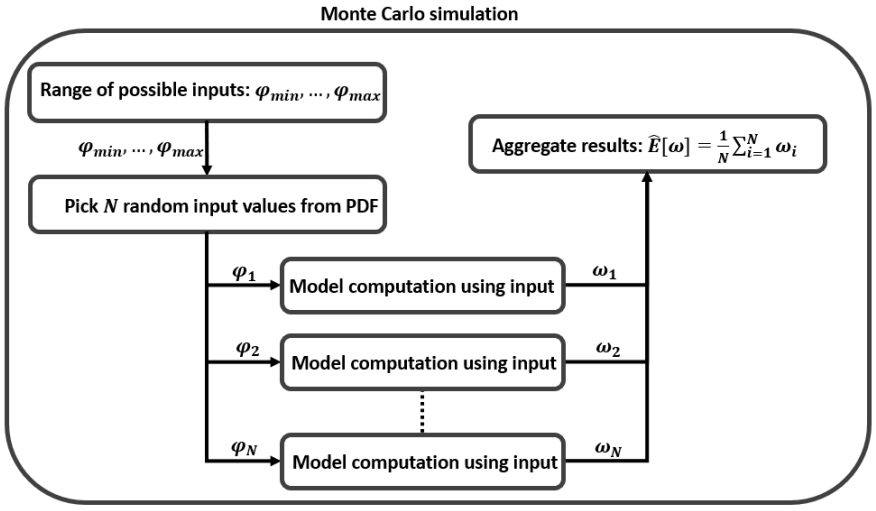


Figure 2.7: General format of a Monte Carlo Methods (MCM) algorithm. A range of possible input values is selected, from which N samples are taken using the probability density function (PDF). For each input value a model evaluation is then executed and the outputs from each evaluation are aggregated to obtain the results of the Monte Carlo Methods algorithm.

The first step consists of defining the range of possible inputs on the domain and assigning a Probability Distribution Function (PDF) to these ranges of inputs. Secondly, N random samples are selected from the

PDF. At the third step, for each input value an output is calculated, using a single model realization. At last, the results are aggregated in the fourth step in order to estimate the true output value [26]. A review as to when and why MCM should be used is described by *Kroese et al.* [28]. The main arguments for the usage of the MCM for this work are as follows:

1. MCM work well with high input dimensionality and highly complex and/or coupled systems; and
2. MCM are easy to implement, since a pre-defined number of samples are taken and for each sample a model simulation is performed.

In preparation for the next sections, the mathematical description of the Monte Carlo estimator is presented.

$$\mathbb{E}[Q] = \int_{\Omega} Q(\omega) p(\omega) d\omega \tag{2.19}$$

$$\hat{Q}_N^{MC} = \frac{1}{N} \sum_{i=1}^{N} Q(\omega^{(i)}) \tag{2.20}$$

The expected value $\mathbb{E}[Q]$ of a generic quantity of interest Q is determined using Equation 2.19. Here Ω is the sample space, $p(\omega)$ is the probability of event ω happening. Using MCM to estimate the expected value of Q results in Equation 2.20, where $\hat{Q}_N^{MC}$ is the Monte Carlo estimator of Q for N samples. Other relevant features of Monte Carlo estimators are the estimator variance, Mean Squared Error (MSE) and the estimator bias.

$$\mathbb{V}\text{ar}(\hat{Q}_N^{MC}) = \mathbb{E}[(\hat{Q}_N^{MC} - \mathbb{E}[\hat{Q}_N^{MC}])^2] \tag{2.21}$$

$$\text{MSE}(\hat{Q}_N^{MC}) = \mathbb{E}[(\hat{Q}_N^{MC} - Q)^2] = \mathbb{V}\text{ar}(\hat{Q}_N^{MC}) + \text{Bias}^2(\hat{Q}_N^{MC}, Q) \tag{2.22}$$

$$\text{Bias}(\hat{Q}_N^{MC}, Q) = \hat{Q}_N^{MC} - Q \tag{2.23}$$

The former corresponds to the spread of a set of points around their estimated mean value and the latter is the averages squared difference between the estimated and true values. Both features are represented in their mathematical form in Equation 2.21 and Equation 2.22 respectively. The MSE can be expanded and rewritten into the sum of estimator variance and the estimator bias squared. The estimator bias is the difference between the estimated value and the true value as given in Equation 2.23.

Multi-Level Monte Carlo

The goal of the Multi-Level Monte Carlo (MLMC) method is to estimate the value of a quantity of interest with a predefined target variance. This is achieved by redistributing samples to lower resolution realizations and reconstructing the estimator of the high resolution model using the differences in estimated quantities of interest between the resolution levels [30]. The general process is depicted in Figure 2.8.

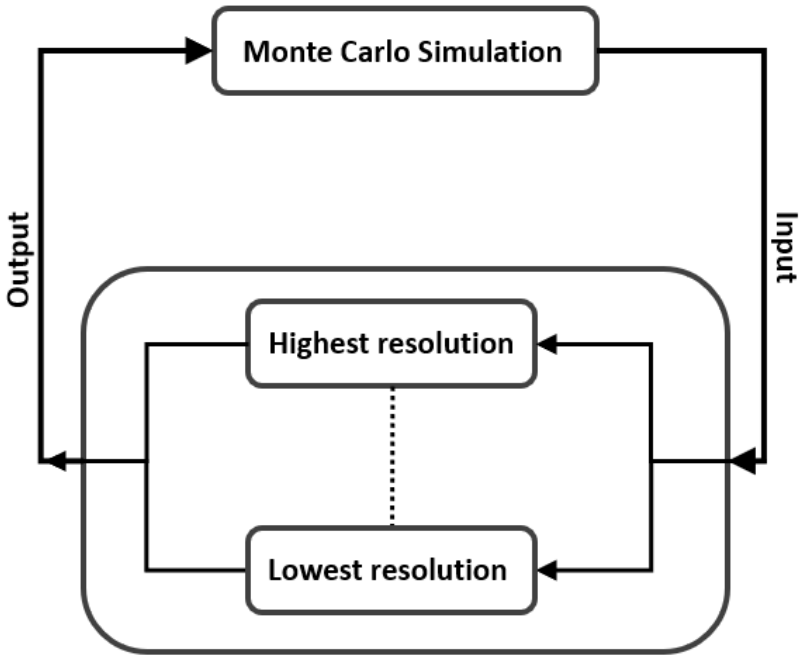


Figure 2.8: The general idea of the Multi-Level Monte Carlo approach is substituting a single resolution of a single model in the Monte Carlo simulation by a group of multiple resolutions.

Next, the mathematical derivation of the MLMC method is described. The algorithmic implementation of the MLMC approach is described in chapter 3.

$$Q_L = Q_0 + \sum_{l=1}^{L} [Q_l - Q_{l-1}] \tag{2.24}$$

The basis of the MLMC method is given by the telescoping sum identity [30] as is presented by Equation 2.24. The values of Q are determined at each resolution level $l = 0, 1, ..., L$ with L as the highest resolution level. The MCM estimator of Q is estimated by combining Equation 2.20 and Equation 2.24.

$$\hat{Q}_{N_L}^{MLMC} = \sum_{l=0}^{L} \hat{Y}_l = \sum_{l=0}^{L} \frac{1}{N_l} \sum_{i=1}^{N_l} Y_l^{(i)} \tag{2.25}$$

$$Y_l^{(i)} = \begin{cases} Q_0^{(i)}, & l = 0 \\ Q_l^{(i)} - Q_{l-1}^{(i)}, & l > 0 \end{cases} \tag{2.26}$$

The resulting estimator using MLMC is then calculated using Equation 2.25, where difference function estimator is $\hat{Y}_l = \frac{1}{N_l} \sum_{i=1}^{N_l}$. For clarity in later derivations, the difference function $Y_l^{(i)}$ is introduced. In Equation 2.25 N_l is the number of samples of the MC estimator at each resolution level.

$$\text{MSE}(\hat{Q}_{N_L}^{MLMC}) = \mathbb{V}\text{ar}(\hat{Q}_{N_L}^{MLMC}) + \text{Bias}^2(\hat{Q}_{N_L}^{MLMC}, Q) = \frac{\epsilon^2}{2} + \frac{\epsilon^2}{2} = \epsilon^2 \tag{2.27}$$

$$\mathbb{V}\text{ar}(\hat{Q}_{N_L}^{MLMC}) = \mathbb{E}[(\hat{Q}_{N_L}^{MLMC} - \mathbb{E}[\hat{Q}_{N_L}^{MLMC}])^2] = \sum_{l=0}^{L} \frac{1}{N_l} \mathbb{V}\text{ar}(Y_l) \tag{2.28}$$

$$C = \sum_{l=0}^{L} C_l N_l \tag{2.29}$$

The next step is to find an optimal number of samples for each level N_l by achieving a minimum total computational cost C for a set target variance ϵ^2 [31]. Here, ϵ^2 is equal to the MSE of the MLMC estimator. It is assumed that contribution between the variance and the bias term of the MSE is equal with respect to distribution of the computational burden [31], as is displayed in Equation 2.27. Hence, it is assumed that the

variance of the MLMC estimator (Equation 2.28) is equal to $\mathbb{V}\text{ar}(\hat{Q}_{N_L}^{MLMC}) = \frac{\epsilon^2}{2}$. The total computational cost is given by Equation 2.29. Here C_l is the computational cost at level l for each realization of $\hat{Y}_l$.

$$N_l = \frac{1}{\epsilon^2} \sum_{k=0}^{L} \sqrt{\mathbb{V}\text{ar}(\hat{Y}_k) C_k} \sqrt{\frac{\mathbb{V}\text{ar}(\hat{Y}_l)}{C_l}} \tag{2.30}$$

The optimal number of samples per level is determined using a minimization of the computational cost under variance constrain as is done by *Geraci et al.* [31]. The resulting optimal number of samples per level is determined using Equation 2.30.

Multi-Fidelity Control Variates

The goal of the Multi-Fidelity Control Variates (MFCV) is to reduce the variance of the high-fidelity model estimator by using the variance of the low-fidelity models and the covariance between the high- and low-fidelity models [25]. High- and low-fidelity models are explained in more detail in section 2.4. The general process of MFCV is depicted in Figure 2.9. The MCM simulation provides the random input values, which are then calculated using multi-fidelity models and a new estimator is constructed with -in theory- smaller variance.

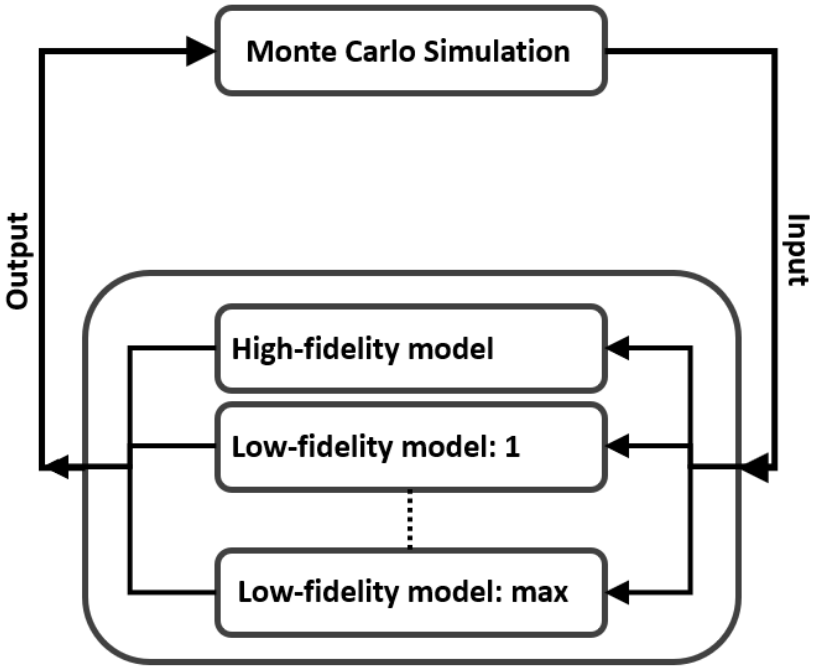


Figure 2.9: The general idea of the Control Variates Monte Carlo approach is substituting a single model in the Monte Carlo simulation by a group of multi-fidelity models.

$$Q^{HF,CV} = Q^{HF} + \alpha(Q^{LF} - \mathbb{E}[Q^{LF}]) \tag{2.31}$$

In control variate theory the quantity of interest resulting from a high-fidelity model $Q^{HF,CV}$ is described using Equation 2.31. The quantities of interest resulting from high-fidelity and low-fidelity models are denoted by HF and LF respectively and the variable α is yet undetermined. It is assumed that the expected value of each low-fidelity model $\mathbb{E}[Q^{LF}]$ is known before evaluating each low-fidelity model [25].

$$\hat{Q}_N^{HF,CV} = \hat{Q}_N^{HF} + \alpha(\hat{Q}_N^{LF} - \mathbb{E}[Q^{LF}]) \tag{2.32}$$

The same relation as Equation 2.31 holds for the estimators of the quantities of interest as presented in Equation 2.32. The additional subscript N denotes the number of samples taken. The next part of this section is dedicated to deriving a value for α for which the variance is minimized.

$$\frac{d\mathbb{V}\text{ar}(\hat{Q}_N^{HF,CV})}{d\alpha} = 0 \tag{2.33}$$

$$\mathbb{V}\text{ar}(\hat{Q}_N^{HF,CV}) = \mathbb{V}\text{ar}(\hat{Q}_N^{HF}) + 2\alpha\text{Cov}(\hat{Q}_N^{HF}, \hat{Q}_N^{HF}) + \alpha^2\mathbb{V}\text{ar}(\hat{Q}_N^{LF}) \tag{2.34}$$

$$\frac{d\mathbb{V}\text{ar}(\hat{Q}_N^{HF,CV})}{d\alpha} = 2\text{Cov}(\hat{Q}_N^{HF}, \hat{Q}_N^{HF}) + 2\alpha\mathbb{V}\text{ar}(\hat{Q}_N^{LF}) = 2\rho\sqrt{\mathbb{V}\text{ar}(\hat{Q}_N^{HF})}\sqrt{\mathbb{V}\text{ar}(\hat{Q}_N^{LF})} + 2\alpha\mathbb{V}\text{ar}(\hat{Q}_N^{LF}) = 0 \tag{2.35}$$

The minimization is stated in Equation 2.33, hence the first step is to obtain the variance of the MFCV estimator, which is depicted in Equation 2.34. Next the variance of the MFCV estimator is filled into the minimization equation, resulting in Equation 2.35. Here ρ is the Pearson correlation coefficient.

$$\alpha = -\rho\frac{\sqrt{\mathbb{V}\text{ar}(\hat{Q}_N^{HF})}}{\sqrt{\mathbb{V}\text{ar}(\hat{Q}_N^{LF})}} \tag{2.36}$$

$$\mathbb{V}\text{ar}(\hat{Q}_N^{HF,CV}) = (1-\rho^2)\mathbb{V}\text{ar}(\hat{Q}_N^{HF}) \tag{2.37}$$

Rewriting the filled in minimization (Equation 2.35) and solving for α results in a definition of α (Equation 2.36). Finally, in order to prove variance reduction, α is substituted into the variance of the MFCV estimator (Equation 2.34). This results in Equation 2.37, which proves the variance reduction, since for the Pearson correlation coefficient it holds that $-1 \leq \rho \leq 1$ or equivalently $0 \leq \rho^2 \leq 1$.

Multi-Fidelity Control Variates with estimated control means

The expected value of the low-fidelity models $\mathbb{E}[Q^{LF}]$ used here (see Equation 2.31) might not be known a priori. However this is required for control variate theory. Hence an additional step is added, called estimated control means [32]. This technique is also applied in the works of [31] [33] [34] [35]. The goal of this additional step is to estimate the expected value of a low-fidelity model using more than N samples. In the aforementioned works the number of samples taken for the low-fidelity models depends on the number of samples N, namely $(1+r)N$ samples for the low-fidelity model with r an unknown parameter that corresponds to the increase in samples for the low-fidelity model.

$$\mathbb{E}[Q^{LF}] \approx \mathbb{E}[\hat{Q}^{LF}] = \frac{1}{(1+r)N}\sum_{i=1}^{(1+r)N}\hat{Q}^{LF,(i)} \tag{2.38}$$

$$\mathbb{V}\text{ar}(\hat{Q}_N^{HF,CV}) = \mathbb{V}\text{ar}(\hat{Q}_N^{HF}) + 2\alpha\text{Cov}(\hat{Q}_N^{HF}, \hat{Q}_N^{HF}) + \alpha^2\mathbb{V}\text{ar}(\hat{Q}_N^{LF}) + \alpha_*^2\mathbb{V}\text{ar}(\hat{Q}_{rN}^{LF}) \tag{2.39}$$

$$\alpha_* = -\rho\sqrt{\frac{1}{1+r}}\frac{\sqrt{\mathbb{V}\text{ar}(\hat{Q}_N^{HF})}}{\sqrt{\mathbb{V}\text{ar}(\hat{Q}_N^{LF})}} \tag{2.40}$$

The expected value of a low-fidelity model is then estimated using Equation 2.38. Due to the estimation of the expected value of the low-fidelity model an additional bias term $\alpha_*^2\mathbb{V}\text{ar}(\hat{Q}_{rN}^{LF})$ is introduced to the variance of the MFCV estimator (Equation 2.34). This variance of the MFCV estimator with estimated control means is depicted in Equation 2.39. The α_* in the bias term is a modified α factor and is described by Equation 2.40.

$$\mathbb{V}\text{ar}(\hat{Q}_N^{HF,CV}) = \left(1 - \frac{r\rho^2}{1+r}\right)\mathbb{V}\text{ar}(\hat{Q}_N^{HF}) \tag{2.41}$$

The equation for the variance of the control variate theory with estimated control means is derived by filling in both α and α_* into the variance of the MFCV estimator with estimated control means (Equation 2.40), resulting in Equation 2.41. From literature [31] [33] [34] [35] an optimal value of the parameter r is determined when taking into account the cost of both high- and low-fidelity models. The relationship between the cost values is denoted using an equivalent cost C^{EQ} in Equation 2.42 with corresponding optimal parameter r given in Equation 2.43.

$$C^{EQ} = C^{HF} + C^{LF}(1+r) \tag{2.42}$$

$$r = -1 + \sqrt{\frac{\rho^2 C^{HF}}{1-\rho^2 C^{LF}}} \tag{2.43}$$

Multi-Level Multi-Fidelity Monte Carlo method

The MLMC and MFCV approaches are combined in order to achieve even greater variance reduction. The idea is to use the same approach as is described for MLMC with the additional step of MFCV added to the difference function (Equation 2.26). So for each resolution level the variance is reduced by using MFCV. The general idea of the MLMFMC approach is depicted in Figure 2.10.

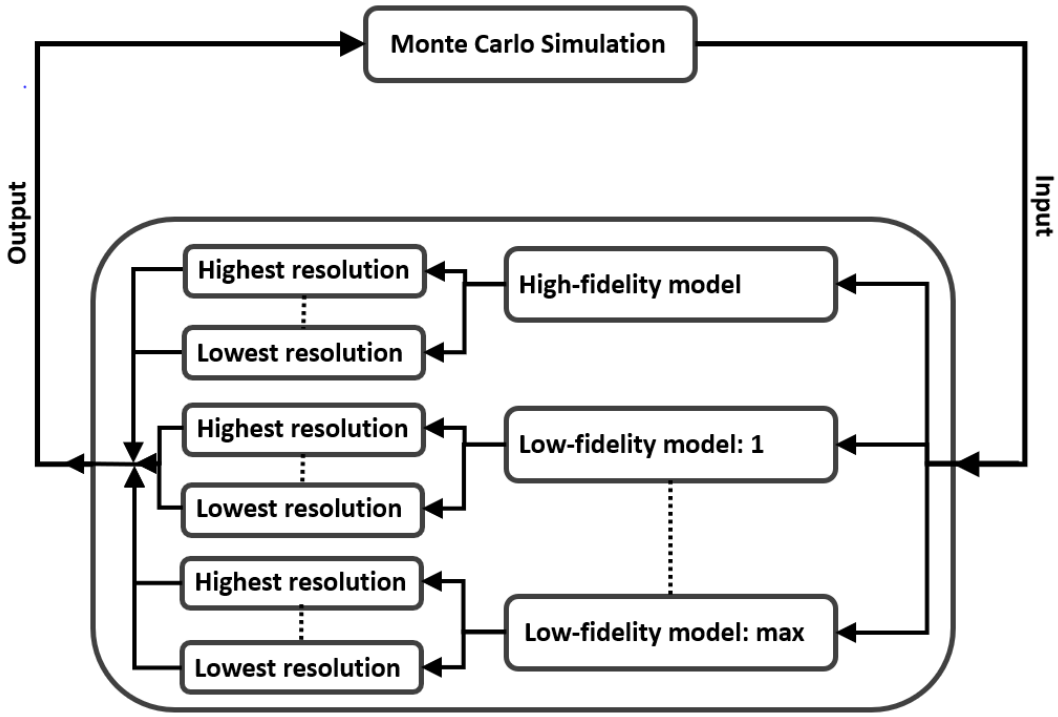


Figure 2.10: The general idea of the Multi-Level Multi-Fidelity Monte Carlo Monte Carlo approach is substituting a single model in the Monte Carlo simulation by a group of multi-fidelity models with varying resolution levels for the high-fidelity model and low-fidelity models if applicable for the low-fidelity model.

$$\mathbb{E}[Q^{HF,MLMFMC}] = \sum_{l=0}^{L} [\hat{Y}_l^{HF} + \alpha_l(\hat{Y}_l^{LF} - \hat{\mathbb{E}}[Y_l^{LF}])] \tag{2.44}$$

Combining the telescoping sum identity (Equation 2.24), the difference function (Equation 2.26) from MLMC with the base definition (Equation 2.31) from MFCV, results in Equation 2.44. Again the optimal number of samples per level is calculated, however now there is an additional term present due to the addition of MFCV. This extra term should force a large part of the computational burden towards the lower-fidelity model. The following equations are the solutions from the constrained optimization using Lagrange multipliers as derived by *Geraci et al.* [31]

$$r_l = -1 + \sqrt{\frac{\rho_l^2 C_l^{HF}}{1-\rho_l^2 C_l^{LF}}} \tag{2.45}$$

$$\Lambda_l = 1 - \frac{r_l \rho_l^2}{1+r_l} \tag{2.46}$$

$$N_l^{HF} = \frac{1}{\epsilon^2} \sum_{k=0}^{L} \sqrt{\frac{\mathbb{V}\text{ar}(Y_k^{HF}) C_k^{HF}}{(1-\rho_k^2)}} \Lambda_k \sqrt{\frac{\mathbb{V}\text{ar}(Y_l^{HF})(1-\rho_l^2)}{C_l^{HF}}} \tag{2.47}$$

$$N_l^{LF} = (1 + r_l) N_l^{HF} \tag{2.48}$$

Here r_l is the parameter from MFCV (Equation 2.43), ρ_l is the Pearson correlation coefficient and Λ_l is the additional term due to the introduction of the MFCV method. For all these cases the subscript l denotes the resolution level at which the variables are determined. The optimal number of samples for the high-fidelity model N_l^{HF} and for the low-fidelity model N_l^{LF} are denoted by Equation 2.47 and Equation 2.48 respectively. There are other optimization settings implemented to the MLMFMC in Dakota, although fully describing these would be out of the scope of this work and therefore a reference is made to the work of *Geraci et al.* [31].

2.4. Multi-fidelity models

In this section multi-fidelity models -which were mentioned in subsection 2.3.3- are explained in more detail. In this study multi-fidelity models consist of high-fidelity and low-fidelity models. Firstly, a formal definition is described, accompanied with several examples for the description of low-fidelity models. Secondly, an overview is presented for low-fidelity models for the circular cylinder cases as used in this thesis.

2.4.1. Definition

The definition of fidelity in the fields of modeling and simulation refers to the degree of a model to reproduce the behaviour of its real life target. Fidelity can be seen as a measure of realism of a model or the degree of similarity [36].

Low-fidelity models is a term used to describe models with a low measure of realism. The concept of low is rather abstract though and could lead to a discussion on what qualifies as a low measure of realism. To abstain from this discussion the following definition is deemed appropriate for this thesis: the low-fidelity models are described as all implemented models, which are not the high-fidelity model. The high-fidelity model is defined as the most realistic model, which for this work, is the 3D CFD simulation implemented in OpenFOAM. Ideally, low-fidelity models have the following two properties:

1. A general trend, which is also present in the real-world situation
2. Low computational cost with respect to the high-fidelity model

Lastly, the current description of low-fidelity models is fairly broad and could be split into three categories according to *Peherstorfer et al.* [25], two of which are relevant to this study: simplified models and data-fit models. Simplified models are models which are derived from a high-fidelity model by applying domain expertise and knowledge of implementation of a high-fidelity model. For example with respect to CFD, direct numerical simulation corresponds to a high-fidelity model and models including large eddy simulations, Reynolds-averaged Navier-Stokes or potential flow correspond to low-fidelity models. Data-fit models are based on input-output related data without necessarily understanding the inner workings of a model. Examples are empirical relationships based on experiments or numerical simulations.

2.4.2. Low-fidelity models for flow around a circular cylinder

As stated above, the goal is to obtain a number of empirical relationships from literature, which could be applied as low-fidelity models in the final experiment. Due to the limited literature available with respect to empirical relations between the Reynolds number and Strouhal number for the case of a porous circular cylinder, the case of a non-porous circular cylinder is considered. For the porous circular cylinder relations are derived using a parametric study in section 4.2. Since literature based on a non-porous circular cylinder is considered, the assumption is made that the introduction of porosity will not have a severely negative impact on the trend of empirical relationships for a circular cylinder. It is expected that the Pearson correlation coefficient will demonstrate the validity of this assumption for each low-fidelity model. The relevant empirical relationships for this study are relations between the Strouhal number and the Reynolds number.

Low-fidelity models: Strouhal number

The oldest relationship between the Strouhal number and Reynolds number is proposed by *Lord Rayleigh* [37].

$$St = a(1 - \frac{b}{Re}) \tag{2.49}$$

The fit coefficients a and b for this relation are also derived by *Roshko* [38] and the relationship is then given in Equation 2.50, which is applicable for the laminar Von Kármán vortex street range of Reynolds numbers.

$$St = 0.21(1 - \frac{21}{Re}), \qquad 40 < Re < 200 \qquad 3D \tag{2.50}$$

However, according to *Roushan et al.* [40] the relationship in Equation 2.49 is valid for a larger range of Reynolds numbers up to several thousand for a 2D case, since the vortex street remains laminar for a larger range of Reynolds numbers. The relationship for a 2D case is given in Equation 2.51.

$$St = 0.18(1 - \frac{28}{Re}), \qquad 40 < Re < 3000 \qquad 2D \tag{2.51}$$

Another relation is derived by *Williamson et al.* [39], which is presented in Equation 2.52. It should be noted that the upper boundary of $Re = 1200$ is not necessarily the upper boundary of validity of this empirical relation, since the authors suggested further research on the range of validity of Reynolds numbers.

$$\begin{cases} St = 0.2698 - \frac{1.0271}{\sqrt{Re}} & 40 < Re < 260 \qquad 3D \\ St = 0.2234 - \frac{0.3490}{\sqrt{Re}} & 260 < Re < 1200 \qquad 3D \end{cases} \tag{2.52}$$

In the study by *Roushan et al.* [40] a 2D empirical relation using this $\sqrt{Re}$ relation has also been presented and is given in Equation 2.53.

$$St = 0.21 - \frac{0.825}{\sqrt{Re}}, \qquad 40 < Re < 3000 \qquad 2D \tag{2.53}$$

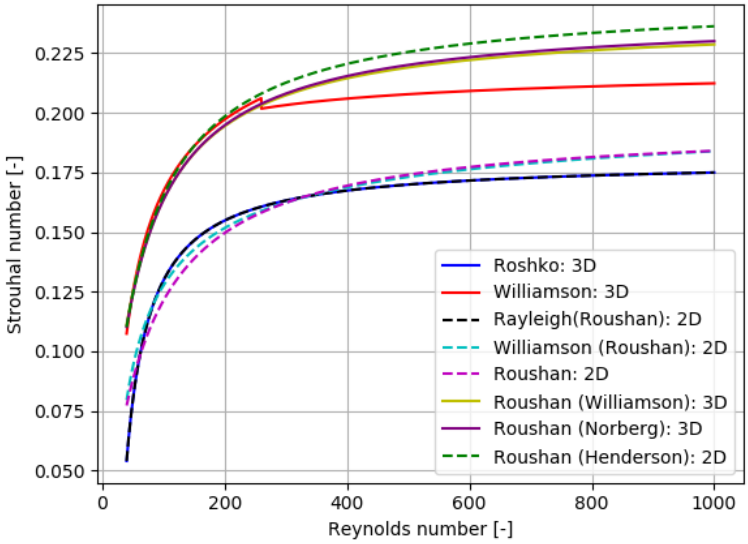
Furthermore the work by *Roushan et al.* [40] provides a new general fit function plus the fit coefficients based on their own experiment and that of three other studies. The general fit function is presented in Equation 2.54 and the corresponding fit coefficients A and B are presented in Table 2.1

$$St = \frac{1}{(A + \frac{B}{Re})} \tag{2.54}$$

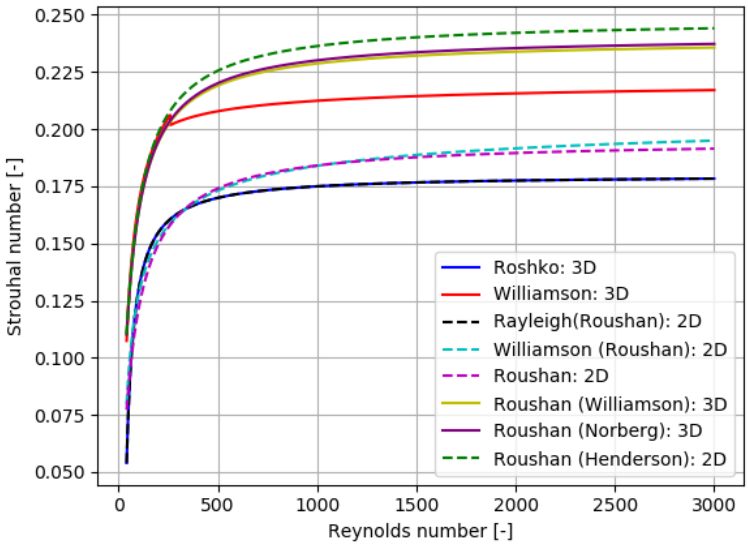
Experiments by	***A***	***B***	**range of** Re	**dimension**
Williamson et al. [39]	4.18	193	$40 < Re < 180$	3D
Norberg [41]	4.15	197	$40 < Re < 180$	3D
Henderson [42]	4.03	202	$40 < Re < 1000$	2D
Roushan et al. [40]	5.12	313	$40 < Re < 3000$	2D

Table 2.1: Overview of fit parameters A and B as derived by *Roushan et al.* [40]

In Figure 2.11 all eight described empirical relationships are depicted. The empirical relationships appear to be categorized into two groups for smaller Re as seen in Figure 2.11a, whereas for larger Re the fits tend to spread as seen in Figure 2.11b. Furthermore as is expected the Strouhal number remains about constant for increasing Reynolds number. Yet, depending on which empirical relationship is taken, this constant value differs on a relatively large range of $0.17 < St < 0.24$.



(a) Smaller range of Reynolds numbers



(b) Larger range of Reynolds numbers

Figure 2.11: Eight empirical relations from theory between the Strouhal and Reynolds numbers. The names in between the brackets denote the data set on which the fit function was applied to.

All theory described in this chapter serves as the basis for the algorithms described in chapter 3 and the experiments described in chapter 4.

3

Methodologies

In this chapter the implementation of algorithms for MLMC and MLMFMC are described. This includes schematics covering from the inner working of the algorithms from input to output and this first section is to be used as a reference in the next chapter when referring to MLMC and MLMFMC in both case studies.

Furthermore the software packages used in this work are described, including features and functions used, as well as external packages. For the fluid dynamics part of this thesis the open-source software *OpenFOAM* is used, with an additional package *olaFlow* for the porosity modeling. The meshes are constructed using *GMSH* and imported to *OpenFOAM*. For the statistics and implementation of MLMC and MLMFMC, the software package *Dakota* by Sandia Labs is used. Lastly, a short section regarding Fourier transforms and curve fitting is added in which the implemented functions are mentioned and explained briefly.

3.1. Algorithms

This section contains the schematic overviews of the MLMC algorithm and the MLMFMC algorithm. The purpose of these schematics is to portray the algorithm using general input and output variables, which can be substituted in the next chapters so only the inputs and outputs will have to be described in the next chapters.

Multi-Level Monte Carlo

The MLMC procedure is depicted in Figure 3.1. The bold black arrow on the left depicts the input parameters of the MLMC, ω is used for the single or multiple general Monte Carlo simulation input variables in the form of a probability density function, N_{pilot} is the number of pilot samples, ϵ is the tolerance which is related to the target variance ϵ^2 and C_l denotes the computational cost of the general Monte Carlo simulation at each resolution level l. The bold black arrow on the right depicts the output value which is the MLMC estimator at the highest resolution level. The working of the algorithm is as follows, first a number of samples equal to N_{pilot} are taken using general Monte Carlo simulation at each resolution level. The resulting sets of estimators denoted by $\{\hat{Q}_0\}, \{\hat{Q}_1\}, ..., \{\hat{Q}_L\}$ each have size equal to N_{pilot} and these sets are used to calculate the optimal number of samples at each resolution level N_l. The difference function Y_l for each level l is calculated, the resulting sets of differences are then used to calculate the optimal number of samples N_l The other parameters used to calculate N_l are ϵ and $C_0, C_1, ..., C_L$. Next a general Monte Carlo simulation is executed using N_l at each resolution level, from which finally the estimator of the MLMC at the highest resolution level is calculated.

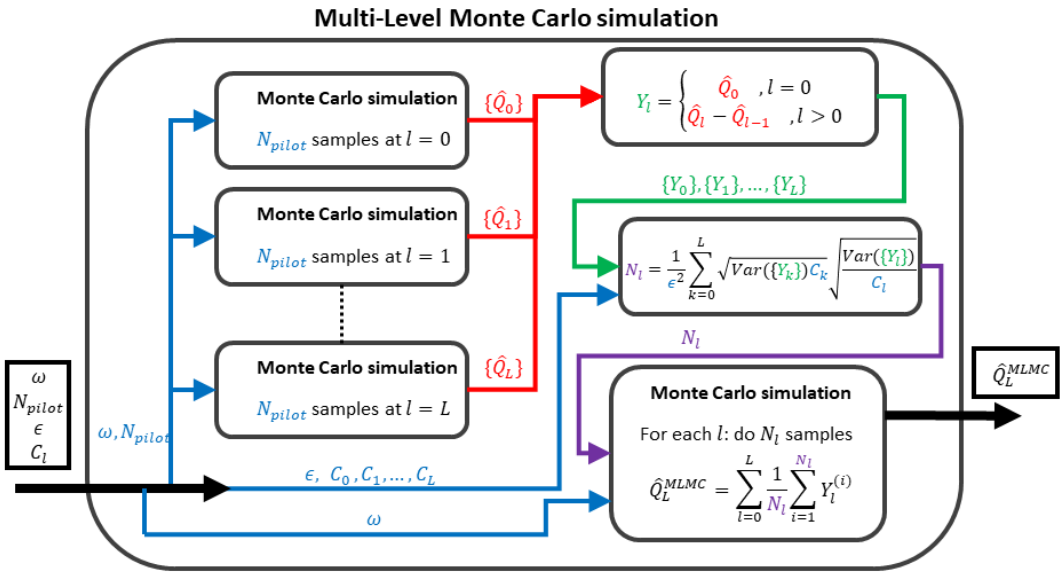


Figure 3.1: Scheme of the Multi-Level Monte Carlo (MLMC) algorithm.

The tolerance ϵ should be minimized, since it is related to the target variance and a smaller target variance means smaller errors in the quantities of interest. Pilot samples N_{pilot} are a set of samples taken to calculate the optimal number of samples per resolution level as can be seen in Equation 2.30 and Figure 3.1. There is no clear method for calculating the number of pilot samples and thus a trade-off has to be made since a high number of pilot samples will yield statistically more accurate quantities of interest, but the higher resolution samples have a higher computational cost. Lastly the computational cost of each resolution level is required, a possible way of assigning a cost to each resolution level is by using relative computational cost. Here the lowest resolution level gets a number assigned to it, in this case $C_0 = 1$ and the higher resolution computational costs are calculated with respect to this lowest computational cost. The computational costs here are based on the average execution time of the general Monte Carlo simulations at each resolution level in combination with the assumption that the lowest resolution has relative computational cost of $C_0 = 1$.

Multi-Level Multi-Fidelity Monte Carlo

A similar process as described for the MLMC is applied to the MLMFMC as depicted in Figure 3.2. Like the MLMC, a number of pilot samples are run first. The difference function between resolution levels is calculated, be mindful that for a single resolution low-fidelity model the difference function is equal to the model output, since this is the lowest resolution level. Moreover, for each resolution level higher than the lowest resolution level, the pilot sample is simulated at the current resolution level and for one resolution level lower, since the difference function needs to be calculated.

The main workload of the MLMFMC is conducted in the next steps. Firstly, the variance of the difference function of both high- and low-fidelity at each resolution level are calculated. Then the covariance is calculated between the model fidelities at each resolution level in order to calculate the Pearson correlation coefficient ρ_l (step 1 in the top right box). Then the term r_l (step 2) which is used in order to estimate the required number of samples of the low-fidelity model is calculated using the relative computational cost of the model fidelities and the Pearson correlation coefficient. Using the term r_l and the Pearson correlation coefficient the variance reduction factor Λ_l is calculated. In step 4 the term α_l is calculated using the Pearson correlation coefficient ρ_l and the variance of high- and low-fidelity models. The additional number of low-fidelity samples N_l^{LF} that are to be calculated in order to estimate the value of the low-fidelity difference function is calculated in step 5 and the optimal number of high-fidelity samples are calculated in step 6 using the relative computational costs, variances of the difference functions and the tolerance. Step 4, 5 and 6 have purple color coding since these are required for the final step in the bottom left box in which the estimator is constructed. The expected value of the difference function of the low-fidelity model $E[Y_l^{LF(i)}]$ (last term in the summation of the estimator) is estimated using the additional number of low-fidelity samples.

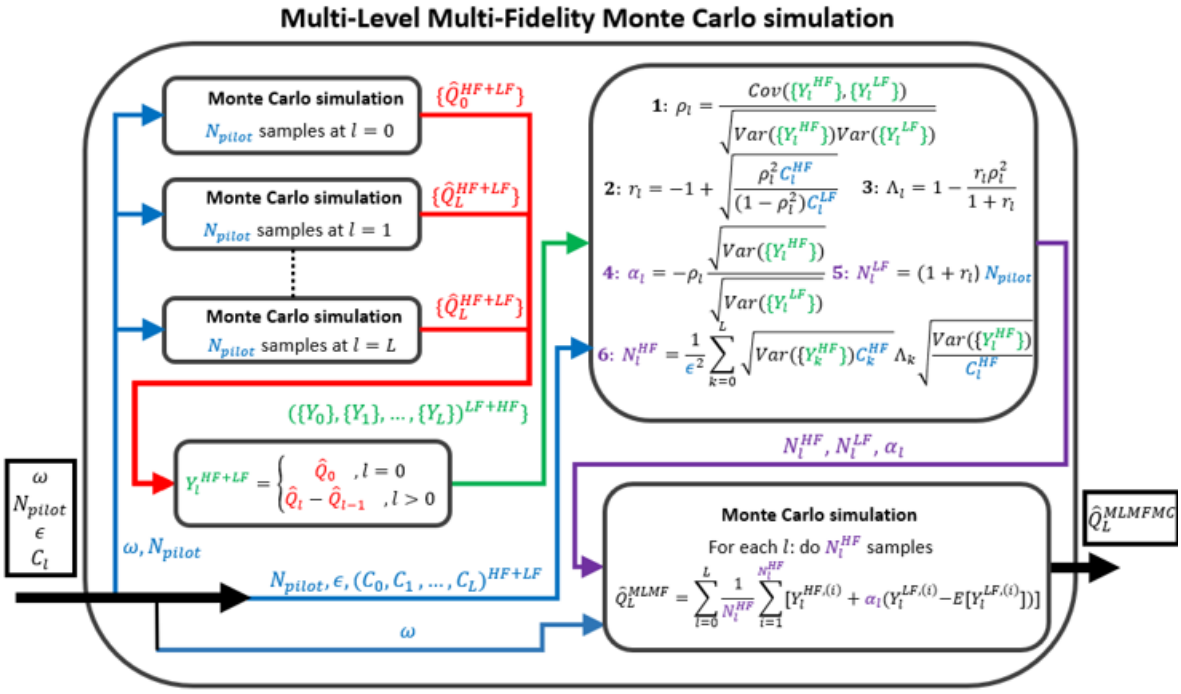


Figure 3.2: Scheme of the Multi-Level Multi-Fidelity Monte Carlo (MLMFMC) algorithm.

3.2. General methodologies

In this section, the software is presented from an implementation point of view. The fluid dynamics are solved using *OpenFOAM*, simulation control and statistics are applied using *Dakota*, post processing and visualization are done using *Python* and *Paraview* respectively and the mesh is constructed using *Gmsh*.

3.2.1. OpenFOAM

OpenFOAM is an open source C++ toolkit developed for solving partial differential equations [43]. It has many functions related specifically to the topic of solving fluid dynamics problems. For this study *OpenFOAM version 2012* is used, including the *olaFlow* toolkit [17]. For visualization of the results *Paraview* [44] is used. In this study the *pimpleFoam* (case study 1) and *olaFlow* (case study 2 and 3) solvers are used.

3.2.2. Gmsh

Gmsh is an open source finite element mesh generator [45]. In this study the geometry and mesh modules are used to construct and generate the mesh. The mesh is then converted for usage in *OpenFOAM* using the *gmshToFoam* function. No mesh convergence studies are conducted in this work, the meshes are checked visually for incorrect features and using the *checkMesh* function.

3.2.3. Dakota

Dakota is developed by Sandia National Laboratories as a tool used for iterative mathematical and statistical methods, which interface to computational models. In short, *Dakota* provides a practical way of analysing iterative parametric black-box simulations by providing answers to questions related to: sensitivity, uncertainty, optimization, calibration, verification and validation [46]. For this study, the goal is uncertainty quantification using Multi-Level Multi-Fidelity Monte Carlo. The building blocks of which are available in *Dakota*. In the next sections an overview is presented of the *Dakota* input file and interface driver file.

Input file

The *Dakota* input file consists of six code blocks in general, namely: environment, method, model, variables, interface and responses. For clarity, the code blocks of the input file are described separately below. Starting with the environment code block.

```
##############################################################################
environment
    tabular_data
        tabular_data_file = 'table_out.dat'
##############################################################################
```

The environment code block denotes the top level settings of the *Dakota* simulation. In this case only the output file is defined here as *table_out.dat*. Next, the method code block is defined as:

```
##############################################################################
method
    model_pointer = 'HIERARCH'
    multilevel_sampling
      seed = 1234
      pilot_samples = 10
      convergence_tolerance = .1
      allocation_target = variance
##############################################################################
```

Here, the model applied is *multilevel_sampling*, which corresponds to Multi-Level Monte Carlo. The HIERARCH *model_pointer* is a special case only applicable for Multi-Level Monte Carlo and Multi-Fidelity Control Variates. In case HIERARCH is used with multiple model blocks and varying resolution levels, the Multi-Level Multi-Fidelity Monte Carlo method is applied. The *seed* is the input value used for random number generation. The *pilot_samples* and *convergence_tolerance* are mentioned before in section 3.1 and correspond to N_{pilot} and ϵ respectively. The *allocation_target* specifies the quantity used for optimal sample allocation or redistribution, which is the variance for this case. Since, the Multi-Level Multi-Fidelity Monte Carlo approach in this study applies a high- and a single low-fidelity model, there are three code blocks for the model specification as displayed below.

```
##############################################################################
model
        id_model = 'HIERARCH'
        surrogate hierarchical
          ordered_model_fidelities = 'LF' 'HF'
##############################################################################

##############################################################################
model
        id_model = 'LF'
        variables_pointer = 'LF_VARS'
        simulation
          solution_level_control = 'key'
          solution_level_cost = 1.
##############################################################################

##############################################################################
model
        id_model = 'HF'
        variables_pointer = 'HF_VARS'
        simulation
          solution_level_control = 'key'
          solution_level_cost = 64800000. 174048000. 348288000.
##############################################################################
```

The upper code block describes the Multi-Level Multi-Fidelity Monte Carlo model as an overarching model which is denoted using the *surrogate* key word in combination with the *hierarchical* tag. The *ordered_model_fidelities* lists the models from low- to high-fidelity models from left to right. The *id_model* assigns a tag to each model block so these can be distinguished. The *variables_pointer* assures each model is assigned the correct input variables. The *solution_level_control* and *solution_level_cost* are the name tag of the estimated computational cost and the values of the computational costs respectively. Next the variables blocks are described, which assign variables to each model block using the pointers. There are two variables block, one for the high- and for the low-fidelity model.

```
################################################################################
variables
    id_variables = 'LF_VARS'
    uniform_uncertain = 1
        lower_bounds = 0.245
        upper_bounds = 0.255
        descriptors 'beta'
    discrete_state_set
        integer = 1
        num_set_values = 1
        set_values = 91050
        initial_state = 91050
        descriptors 'key'
################################################################################

################################################################################
variables
    id_variables = 'HF_VARS'
    uniform_uncertain = 1
        lower_bounds = 0.245
        upper_bounds = 0.255
        descriptors 'beta'
    discrete_state_set
        integer = 1
        num_set_values = 3
        set_values = 30 31 32
        initial_state = 30
        descriptors 'key'
################################################################################
```

The variables code blocks first denote to which model these variables are applied using the *id_variables* keyword. For this study, there is one uniform uncertain variable, which requires a lower and upper bound. The *descriptors* keyword assigns a name to the parameter. The level of each model are setup using *discrete_state_set*. For the high-fidelity model there are three resolution levels, hence there are three *set_values*. For the low-fidelity model there is only one.

```
################################################################################
interface
        fork
        asynchronous
        evaluation_concurrency = 48
        analysis_driver = 'simulator_script'
        parameters_file = 'params.in'
        results_file    = 'results.out'
        work_directory directory_tag
        copy_files = 'templatedir/*'
        named 'workdir' file_save  directory_save
        aprepro
################################################################################
```

The interface code block specifies asynchronous parallel execution of 48 processes using the tags *fork, asynchronous* and *evaluation_concurency*. The interface driver file is called using *analysis_driver* and *parameters_file* and *results_file* create the input parameters and results file for each case. The last number of commands state that a new working directory is constructed for each case, that all files in the *templatedir/* directory are copied to the working directory and that the input and output files are saved inside this working directory. Lastly, the responses code block describes the format of the output file.

```
#############################################################################
responses
        response_functions 1
        no_gradients
        no_hessians
#############################################################################
```

This code block states there is only 1 output value. The keywords of *no_gradients* and *no_hessians* are required and state that no gradient or Hessian information is used for this simulation. These are all the code blocks required to setup a simple case of Multi-Level Multi-Fidelity Monte Carlo method using *OpenFOAM*.

Interface driver file

The interface driver file is called in the input file in the interface code block using *analysis_driver*. For this study the job of the interface driver file is to setup the *OpenFOAM* case, run the case, do post processing and forward the simulation output back to *Dakota*. For this thesis, the interface driver file is setup using a pre processing, analysis and post processing section. The pre processing code sets up each case and fills in the correct input values to the files.

```
#############################################################################
key=$(head -n 4 $1 | cut -d"=" -f2 | cut -d"}" -f1 | tail -n 1 | xargs)
level="${key:1:2}"
fidelity="${key:0:1}"
#############################################################################
```

First the resolution level and model fidelity are extracted from the input file *params.in*. Then these values are used to setup the case for a low-fidelity model (*lowfidelity.py*) or high-fidelity model (3D *OpenFOAM* case).

```
#############################################################################
if [ $fidelity == "9" ]; then
        pwd
        cp ../lowfidelity.py .
        python3 lowfidelity.py
        mv tmp.txt $2
fi
if [[ $fidelity == "3" ]]; then
        meshdir="mesh${level}"
        casedir="casebase${fidelity}"
        dprepro $1 porosityDict.template porosityDict.in
        dprepro $1 U.template3 U.in
        dprepro $1 forceCoeffs.template3 forceCoeffs.in
        pwd
        cp -r ../${casedir}/* .
        cp -r ../meshes/${fidelity}/${meshdir}/polyMesh constant/
        cp ../meshes/${fidelity}/${meshdir}/porosityIndex 0/porosityIndex
        cp U.in 0/U
        cp forceCoeffs.in system/forceCoeffs
        cp porosityDict.in constant/porosityDict
        cp ../postprocess.py .
        cp ../OFjob.sh .
        decomposePar -fileHandler collated > log.decomposePar
fi
#############################################################################
```

The low-fidelity model simply copies a file named *lowfidelity.py*, executes it and exports it to the output file denoted by *$2*. The high-fidelity model first needs to setup the correct mesh and casebase variables. Then the *dprepro* function is used to fill in the variables from *params.in* to the correct *OpenFOAM* files. Lastly, the case is setup by copying all files to the working directory and decomposing the mesh. The analysis and post processing of the high-fidelity model is condcuted using this last code block.

```
##############################################################################
sbatch OFjob.sh > log.job
python3 postprocess.py
mv tmp.txt $2
##############################################################################
```

The *OFjob.sh* script adds a parallel *OpenFOAM* job to the queue and executes it. The *postprocess.py* conducts the post procecssing and lastly the output is moved to the output file denoted by *$2*. The script is written in the *Bash* programming language since the *Linux* programming environment on *Ubuntu 20.04 LTS* is used. This concludes the overview of the interface driver file.

3.2.4. Post processing in Python

Post processing of results from the *OpenFOAM* files is handled using Python. Most noteworthy is the package *SciPy* [47], which is used for the Fast Fourier Transform algorithm and provides a 2D curve fitting function.

Fast Fourier Transform

The Fast Fourier Transform (FFT) is an algorithm that computes the discrete Fourier transform of a signal. The Fourier transform is used to transform a signal in the temporal domain to the frequency domain or the other way around using the inverse Fourier transform. The FFT algorithm applied in this work is the FFT implemented with the package *SciPy* [47] using the following base equation:

$$X[k] = \sum_{n=0}^{N-1} e^{-2\pi jkn/N} x[n] \tag{3.1}$$

Interpolation

Since adjustable time stepping is enabled in *OpenFOAM*, the time signal has a non-constant time step. In order to apply the FFT, a constant time step is required and therefore the time signal is interpolated. The *interp1d* function is applied using linear interpolation.

Curve fitting

For curve fitting, the function *curve_fit* from the package *SciPy* [47] is applied. This uses non linear least squares method to fit a function to the data.

4

Experimental setup

In this chapter all three experiments conducted in this work are described. It is recommended to the reader to first read section 4.1 and section 4.2 and the results corresponding to these sections, which are described in chapter 5 and chapter 6 respectively. This is suggested due to the iterative nature of this study and the dependency of the final experiment on the results of the first two experiments.

In section 4.1 a case study is conducted regarding benchmark cases for flow around a circular cylinder from literature. The cases are reconstructed in *OpenFOAM* and validated. Thereafter the MLMC and MLMFMC approaches are implemented on the benchmark cases. Initially the MLMC method is implemented for both 2D and 3D cases, which result in an estimated value for the drag coefficient and an equivalent number of samples of high resolution samples to reach -in theory- the same estimated value. The resulting estimations of the drag coefficients are compared and the decrease in computational cost is calculated. In order to further understanding of the MLMC, the intermediary results are analysed. Next the MLMFMC approach is implemented on the 3D cases, again intermediary results are analysed, the results of the MLMFMC approach and general MCM are compared and the computational cost reduction is calculated.

In section 4.2 a parametric study on porous cylinders in the subcritical Reynolds numbers regime is conducted. The experimental setup is taken from literature and reconstructed in *OpenFOAM*, after which the case setup is validated. The validation is unsuccessful, since the simulated drag coefficient results do not correspond to the suggested literature values and curve shape. Therefore, flow patterns were analysed and the focus of the parametric study is changed to the Strouhal number, porosity and Reynolds number. The results of the parametric study are analysed using curve fitting to -ideally- obtain one or several empirical relationships.

In section 4.3 the final experiment is described, which combines the MLMFMC approach on the benchmark case setup with the parametric study on flow around a porous circular cylinder. The 3D case description of section 4.1 is modified to include a porous circular cylinder. And a new parametric study on a smaller range of parameters is set up, based on the results of the parametric study of section 4.2. The MLMFMC approach is used to estimate the results of the parametric study in order to -in theory- reduce the effects of epistemic uncertainty due to the complex nature of the simulation. Moreover, low-fidelity models are implemented and analysed using the correlation between high- and low-fidelity models.

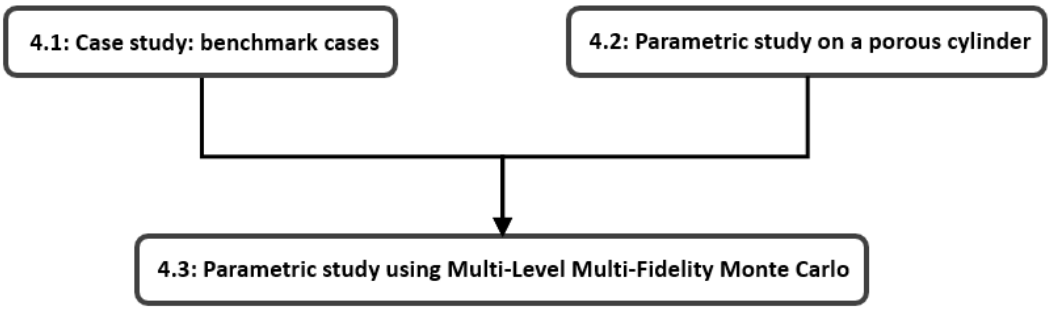


Figure 4.1: Overview of the three experiments conducted in this study. First a benchmark case is used to implement the Multi-Level Multi-Fidelity Monte Carlo method. Then a parametric study using a porous circular cylinder is conducted in order to derive low-fidelity models. And, finally, these two sections are combined by creating a new case with a porous circular cylinder, to which the Multi-Level Multi-Fidelity is applied.

4.1. Case study: benchmark cases

The goal of this experiment is to correctly setup the MLMFMC approach and corresponding infrastructure for fluid dynamics topics in both *Dakota* and *OpenFOAM*. Moreover, the intermediary results of the MLMFMC method are analysed in order to detect possible faults in implementation at an early stage. A fluid dynamics case using benchmark cases on flow past a circular cylinder in 2D and 3D, is selected from literature in order to validate the MLMFMC algorithm. Initially the benchmark cases are reconstructed in *OpenFOAM* and validated using literature. After validation, the benchmark cases are modified in order to create a 2D and 3D case for simple uncertainty propagation. This is achieved by assuming an uncertain inflow velocity. The results of this experiment with uncertainty propagation only have relevance in order to validate a correct implementation of the MLMFMC method. First the MLMC algorithm is implemented, as this is an intermediate step of the MLMFMC approach. The MLMC approach is implemented and validated in both 2D and 3D using the benchmark case with simple uncertainty. Then the MLMFMC method is implemented and validated for the 3D case, using the 2D case as a low-fidelity model. For both MLMC cases and the MLMFMC case the intermediary results are analysed. In subsection 4.1.1 the benchmark cases are described, reconstructed in *OpenFOAM* and validated. In subsection 4.1.2 the benchmark case with uncertainty is introduced. In subsection 4.1.3 the input parameters and the validation technique of both MLMC and MLMFMC methods are described. An overview of the sections of this first experiment is depicted in Figure 4.2.

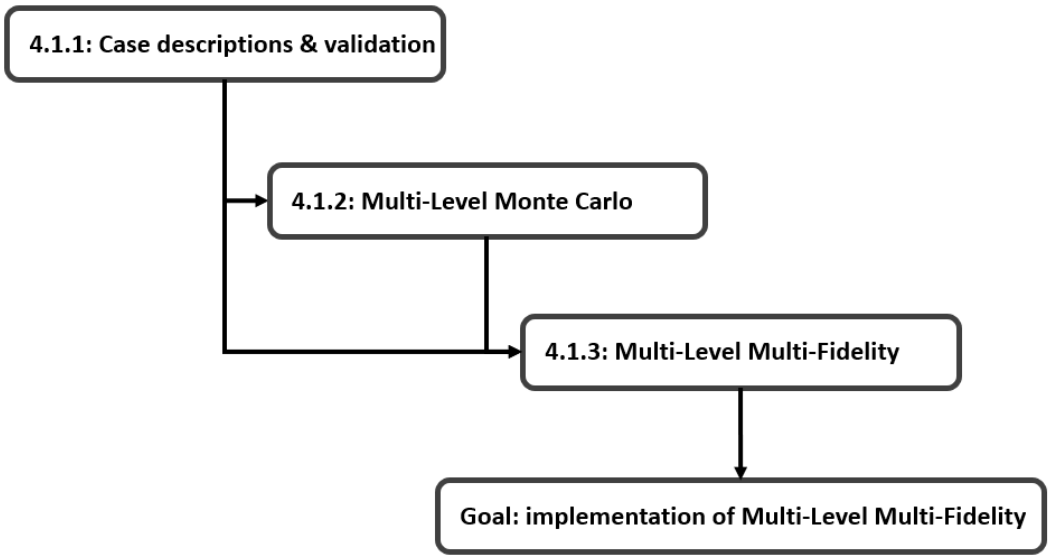


Figure 4.2: Overview of the section regarding the implementation of MLMF using the benchmark cases. Starting with the case description as described in literature, the reconstructed cases in *OpenFOAM* and the validation of the reconstructed cases. Following with the implementation of the Multi-Level Monte Carlo method using the benchmark cases and lastly the implementation of the Multi-Level Multi-Fidelity method using the benchmark cases.

4.1.1. Case descriptions and validation

The benchmark cases are described by *Schäfer et al.* [48] and are separated into six cases. Three similar cases in 2D and in 3D, in which the first case is steady flow and the second and the third cases are unsteady flow. The drag coefficient is calculated using *OpenFOAM*'s *forceCoeffsIncompressible* post processing tool.

2D cases

The geometry of the 2D cases consists of a rectangular plane of $(x, y) = (2.2, 0.41)$ m in which flow around a circular cross-section of a cylinder with a diameter of $D = 0.1$ m is considered as is depicted in Figure 4.3. Here $H_y = 0.41$ m is the height of the plane in the y-direction, U is the velocity in the x-direction and V is the velocity in the y-direction. The velocity boundary conditions are no slip boundary conditions on the sides, zero gradient on the outlet and there is a spatial parabolic inflow condition for U at the inlet as defined in the last column of Table 4.1 and $V = 0$ m/s for each case.

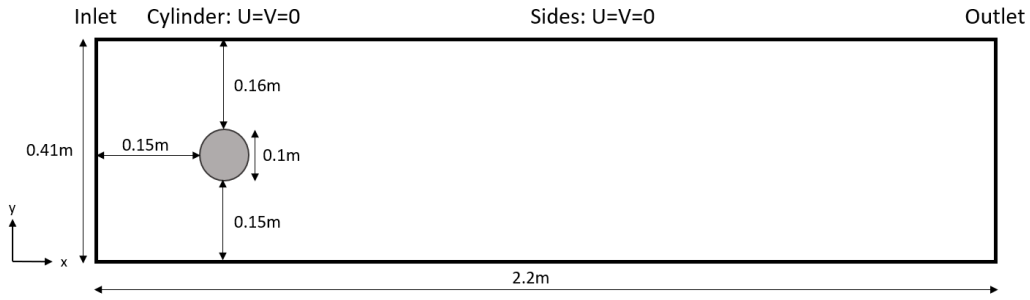


Figure 4.3: Geometry of all 2D benchmark cases. Total size of rectangular plane is 2.2x0.41 m with a cylinder of 0.1 m diameter and no slip boundary conditions imposed on the top side, bottom side and the cylinder, zero gradient boundary conditions on the outlet and a spatial parabolic inflow condition on the inlet.

As depicted in Table 4.1 case 2D-1 is a case with steady inflow conditions ($Re = 20$) around a cylinder, case 2D-2 is an exact copy of this case, however now with unsteady inflow conditions ($Re = 100$) around a cylinder. Case 2D-3 is a slight modification of case 2D-2, now there is a time dependency added to the inflow condition. In Table 4.1 the mean horizontal velocity component is given by $\bar{U}(t) = \frac{2}{3}U(0, H/2, t)$.

Case	Re **[-]**	U_m **[m/s]**	$\bar{U}$ **[m/s]**	**Inflow condition [m/s]**
2D-1	20	0.3	0.2	$U(y) = \frac{4U_m y(H-y)}{H^2}$ (4.1)
2D-2	100	1.5	1.0	$U(y) = \frac{4U_m y(H-y)}{H^2}$ (4.2)
2D-3	0-100	1.5	1.0	$U(y,t) = \frac{4U_m y(H-y)}{H^2} sin(\frac{\pi t}{8})$ (4.3)

Table 4.1: Overview of the relevant parameters for the three benchmarks cases in 2D

2D meshes

For implementation the benchmark cases are discretized and in preparation for the implementation of both MLMC and MLMFMC, three resolution levels are constructed. In Figure 4.4 the meshes for the benchmark cases in 2D are presented. The lowest resolution case is depicted at the top of the image, the middle resolution in the middle and the highest resolution is presented at the bottom of the image. In general the number of cells in each direction is doubled when the resolution level increases. The meshes are not tested using a mesh convergence study, but a mesh check is conducted using *OpenFOAM*'s *checkMesh* function.

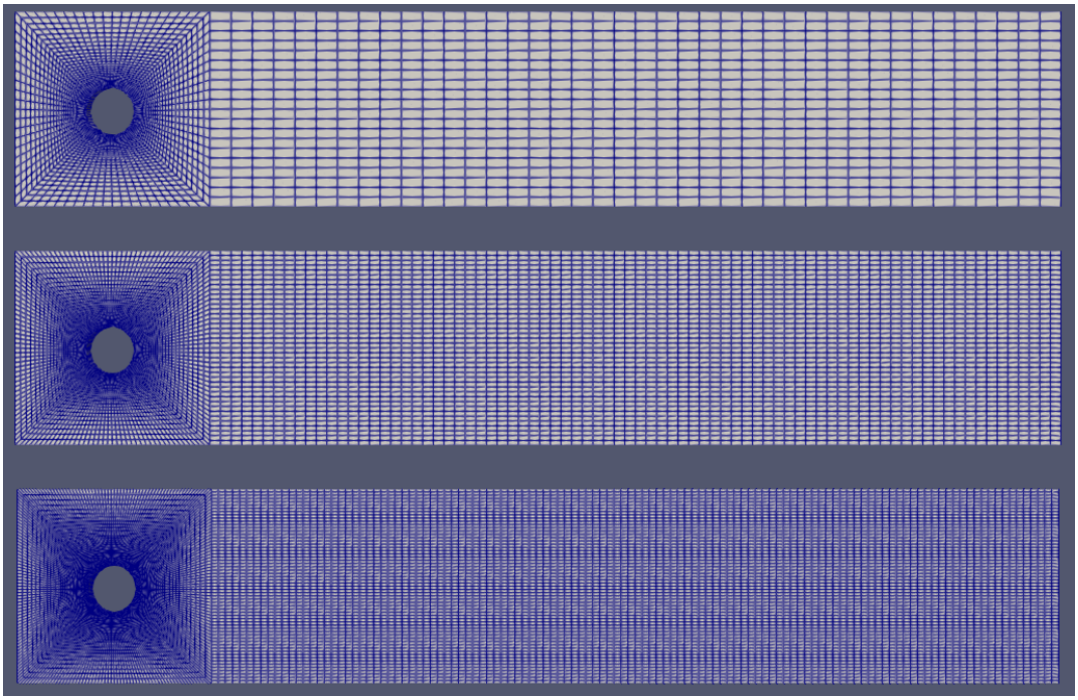


Figure 4.4: The meshes of the 2D cases with lowest to highest resolution level from top to bottom respectively.

Validation of 2D cases

The benchmark cases are validated using the results as summarized by *Schäfer et al.* [48], which are presented in the sixth column of Table 4.2. In the seventh column the drag coefficients resulting from the reconstructed benchmark cases are presented. For the fist case 2D-1, the drag coefficient is higher than the literature values for each case. However, the order of magnitude is the same and the difference is slight, therefore case 2D-1 is deemed to be validated and correctly implemented. The same reasoning is used for case 2D-3, since these values are also all higher than literature, but the differences are a tenth smaller than for the case of 2D-1. For case 2D-2 the calculated drag coefficients are smaller for resolution levels 0 and 1, whereas for resolution 2 the drag coefficient falls into the range of values from literature. Therefore all three cases are correctly implemented and validated in accordance with literature.

Case	**Level**	Re [-]	U_m **[m/s]**	$\bar{U}$ **[m/s]**	**Drag coefficient [-]: literature | 2D CFD**	
2D-1	0	20	0.3	0.2	5.5700 - 5.5900 [48]	5.7761
	1					5.7663
	2					5.7662
2D-2	0	100	1.5	1.0	3.2200 -3.2400 [48]	3.1912
	1					3.2129
	2					3.2261
2D-3	0	20-100	1.5	1.0	2.9300 - 2.9700 [48]	2.9936
	1					2.9879
	2					2.9772

Table 4.2: Overview of the results of validation for the three benchmarks cases in 2D

3D cases

The 3D cases are similar to the 2D cases, hence the boundary conditions are almost equal: on the sides and on the cylinder there is a no slip boundary condition, the outlet has a zero gradient boundary condition and the inlet has a spatial parabolic boundary condition for U and $V = W = 0$ m/s. Here W is the velocity in the z-direction and U and V are as defined for the 2D case. Instead of a rectangular plane, there is now a rectangular box of $(x, y, z) = (2.5, 0.41, 0.41)$ m with a cylinder of height $H_z = 0.41$ m in the z-direction and diameter of $D = 0.1$ m as is depicted in Figure 4.5.

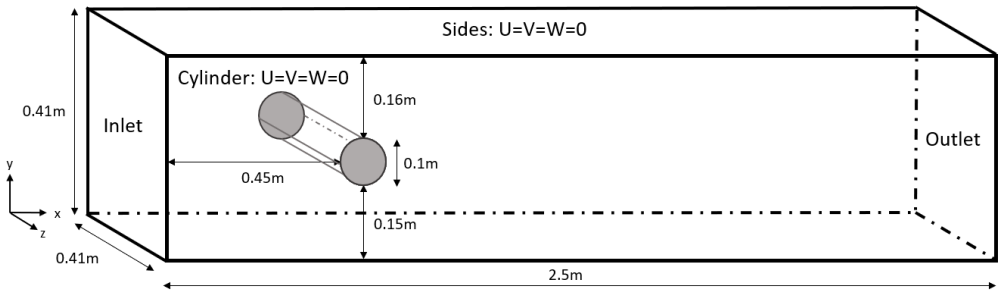


Figure 4.5: Geometry of all 3D benchmark cases. Total size of rectangular plane is 2.2x0.41x0.41 m with a cylinder of 0.41 m height and 0.1 m diameter and no slip boundary conditions imposed on the top side, bottom side, left side, right side and the cylinder, a zero gradient boundary condition on the outlet and a spatial parabolic inflow condition on the inlet.

As depicted in Table 4.3 case 3D-1 is a case with steady inflow conditions (Re = 20) around a cylinder, case 3D-2 is an exact copy of this case, however now with unsteady inflow conditions (Re = 100) around a cylinder. Case 3D-3 is a slight modification of case 3D-2, now there is a time dependency added to the inflow condition. In Table 4.3 the mean horizontal velocity component is given by $\bar{U}(t) = \frac{4}{9}U(0, H/2, H/2, t)$.

Case	Re **[-]**	U_m **[m/s]**	$\bar{U}$ **[m/s]**	**Inflow condition [m/s]**
3D-1	20	0.45	0.2	$U(y,z) = \frac{16U_m y(H-y)z(H-z)}{H^4}$ (4.4)
3D-2	100	2.25	1.0	$U(y,z) = \frac{16U_m y(H-y)z(H-z)}{H^4}$ (4.5)
3D-3	0-100	2.25	1.0	$U(y,z,t) = \frac{16U_m y(H-y)z(H-z)}{H^4} sin(\frac{\pi t}{8})$ (4.6)

Table 4.3: Overview of the relevant parameters for the three benchmarks cases in 3D

3D meshes

In Figure 4.6 the meshes for the benchmark cases in 3D are presented. The lowest resolution case is depicted at the top of the image, the middle resolution in the middle and the highest resolution is presented at the bottom of the image. In general the number of cells in each direction is doubled when the resolution level increases. The mesh is checked using *OpenFOAM*'s *checkMesh* function.

Validation of 3D cases

In Table 4.4 the results of the validation study conducted on the benchmark cases is presented. In the last column of the table, the drag coefficient correctly corresponds to the order of magnitude of the drag coefficient from literature. However, it appears that the drag coefficient at each case and at each resolution is slightly larger than the values from literature, with the exception of case 3D-1 at resolution the lowest resolution level, which falls into the range of literature values. The same behaviour is seen for the majority of the 2D cases as well. Yet the order of magnitude is similar and deemed sufficiently accurate. Therefore, it is concluded that the 3D cases are correctly reconstructed in *OpenFOAM* and the correctness of results is validated.

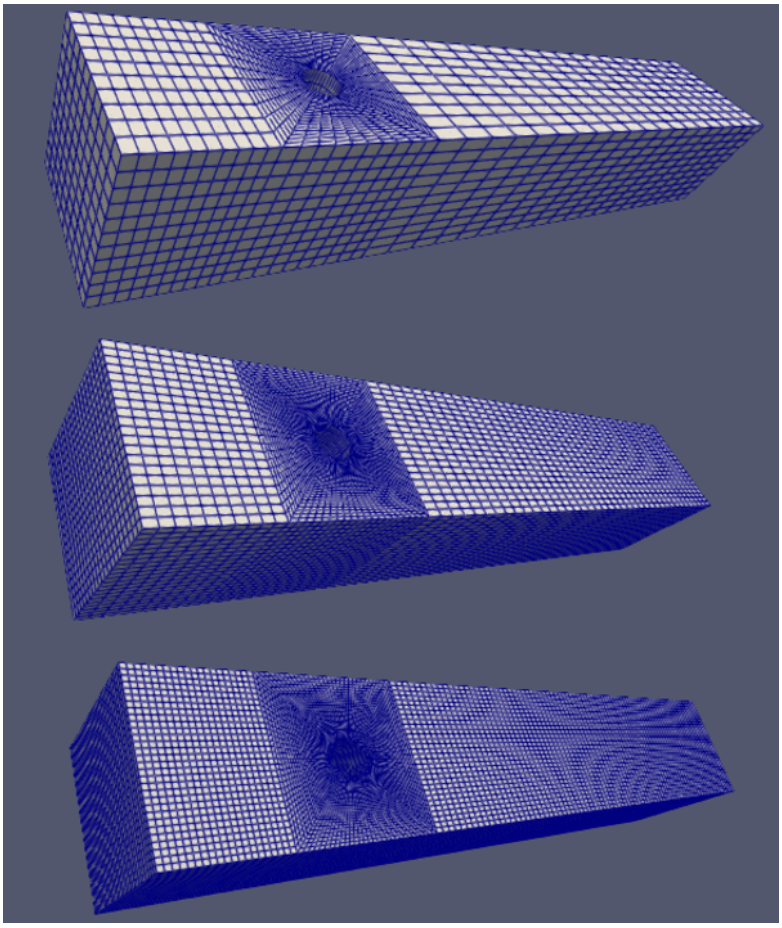


Figure 4.6: The meshes of the 3D cases with lowest to highest resolution level from top to bottom respectively.

Case	**Level**	Re [-]	U_m **[m/s]**	$\bar{U}$ **[m/s]**	**Drag coefficient [-]: literature | 2D CFD**	
3D-1	0	20	0.45	0.2	6.0500 - 6.2500 [48]	6.2278
	1					6.3084
	2					6.3335
3D-2	0	100	2.25	1.0	3.2900 - 3.3100 [48]	3.4318
	1					3.3803
	2					3.3843
3D-3	0	20-100	2.25	1.0	3.2000 - 3.3000 [48]	3.4251
	1					3.3710
	2					3.3789

Table 4.4: Overview of the results of validation for the three benchmarks cases in 3D

4.1.2. Uncertain benchmark case in 2D and 3D

Uncertainty has to be introduced into the benchmark cases in order to demonstrate validity of the MLMFMC method. For simplicity, the range of values of the mean inflow velocity is taken from the first two benchmark cases. With the first benchmark case as the lower boundary at $\bar{U} = 0.2$ m/s and the second benchmark case as the upper boundary at $\bar{U} = 1.0$ m/s with uniform distribution. It is assumed that the true value that is being estimated by the MLMFMC method should correspond to the drag coefficient at $\bar{U} = 0.6$ m/s due to the uniform distribution of the mean inflow velocity.

4.1.3. Multi-Level Multi-Fidelity Monte Carlo

The input parameters of both MLMC and MLMFMC algorithms are presented in section 3.1. Both methods require:

1. Uncertain input parameter (mean flow velocity); and
2. number of pilot samples N_{pilot}; and
3. target variance ϵ^2 or tolerance ϵ; and
4. relative computational cost of each resolution level C_l and model fidelity.

For this experiment the input parameters for both MLMC and MLMFMC are summarized in Figure 4.7 and Figure 4.8 respectively. The MLMC method is applied to a 2D and a 3D case and the MLMFMC approach is implemented with the 3D case as the high-fidelity method and the 2D case as the low-fidelity method.

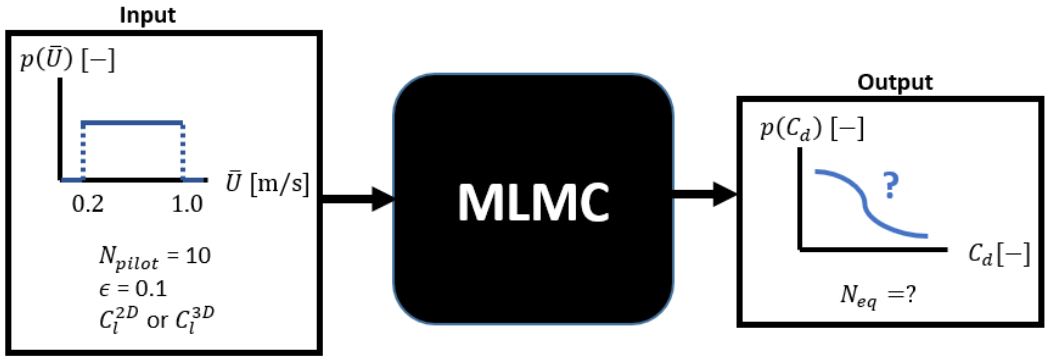


Figure 4.7: Schematic overview of the Multi-Level Monte Carlo approach using a uniformly distributed mean inflow velocity ranging from 0.2 – 1.0 m/s to estimate the drag coefficient.

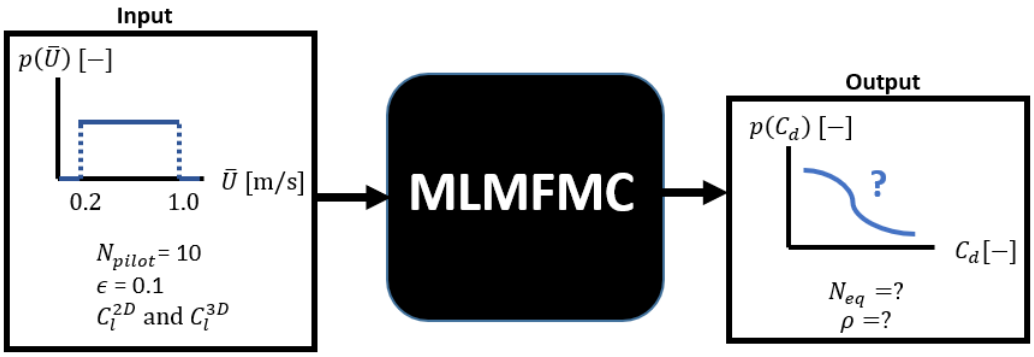


Figure 4.8: Schematic overview of the Multi-Level Multi-Fidelity Monte Carlo approach using a uniformly distributed mean inflow velocity ranging from 0.2 – 1.0 m/s to estimate the drag coefficient.

The uncertain input parameter is described in subsection 4.1.2. The number of pilot samples is set to $N_{pilot} = 10$. There is no way to find the optimal number of pilot samples, other than trial and error. Also increasing the number of pilot samples will increase the computational cost, since more high resolution cases will be run. On the other hand, the number of pilot samples needs to be large enough to correctly capture the preliminary statistics required for the MLMFMC. Hence a trade-off should be made when determining the number of pilot samples. It is assumed that a number of pilot samples equal to $N_{pilot} = 10$ is sufficient. Although another study could be done to investigate the optimal number of pilot samples. The tolerance is set to $\epsilon = 0.1$. The relative computational cost is approximated by running a number of cases and averaging the simulation time over these cases. It should be noted to the reader that the MLMC requires a relative computational cost of each resolution level for a single model fidelity, whereas the MLMFMC has costs for varying model fidelities. The computational cost of the lowest resolution lowest fidelity model is always set

to $C_0 = 1$. The output of both MLMC and MLMFMC methods is the estimated value of the drag coefficient. Furthermore, a number of intermediary results are available i.e. all samples, optimal number of samples per resolution level, number of equivalent samples using high resolution high-fidelity MCM and the Pearson correlation coefficient for MLMFMC.

The extra steps taken with the MLMFMC approach compared to the MLMC approach depend on the correlation between both model fidelities. In order to obtain accurate and statistically representative results using the low-fidelity model, there needs to be high correlation between the high- and low-fidelity model. There is no easy way to quantify what high correlation entails, especially at this stage of the study. Therefore the correlation is analysed.

Lastly, the number of equivalent samples N_{eq} of the High resolution Level High-Fidelity Monte Carlo (HLMC or HLHFMC) are used to estimate the drag coefficient. The estimated drag coefficient of HLMC is compared with that of MLMC and the estimated drag coefficient of HLHFMC is compared with that of MLMFMC in order to validate a correct implementation of the MLMFMC method.

4.2. Parametric study on a porous cylinder

The goal of the parametric study is to find a relationship for circular cylinders in the subcritical Reynolds numbers regime between the porosity and the influence on forces on the cylinder and the vortex shedding frequency. The subcritical Reynolds numbers regime is denoted by as $\mathrm{Re} < 2 \cdot 10^5$. The approach to find such a relation is by applying a parametric study with equally distributed input variables and then curve fitting the resulting output values. Ideally one or several functioning low-fidelity models, in the form of empirical relationships, are the result of this parametric study.

For the parametric study, the MLMC approach could be applied, since it also produce a large enough number of samples, from which a relation could be deduced. Furthermore, the MLMC approach would guarantee a certain level of uncertainty in the results. However, based on the previous case study it is likely that the number of cases that are to be executed for an MLMC approach is in the order of one thousand. Also including the increased computational cost for these cases, the decision has been made not apply the MLMC method.

The general overview of this section is depicted in Figure 4.9. First the case for a porous circular cylinder from literature is described, which is reconstructed in *OpenFOAM* and a validation is conducted. Then the specifics of the parametric study are described, after which a short section on curve fitting is added in order to derive low-fidelity models via data-fitting.

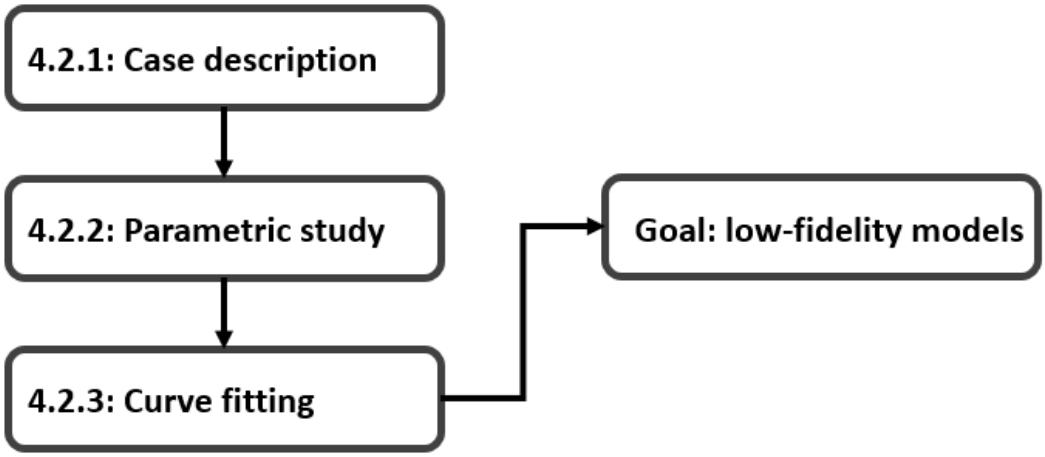


Figure 4.9: Overview of the section regarding the parametric study on a porous circular cylinder. Starting with the case description as described in literature. Following with the setup of the parametric study and lastly the curve fitting on the resulting data to obtain low-fidelity models which can be applied in cases for porous circular cylinders.

4.2.1. Case description

The case applied to the parametric study is taken from the paper by *Steiros et al.* [14], since there is a proposed relationship between drag coefficient and porosity. Moreover, by implementing this case using the porosity model as described by *Feichtner et al.* [18], there is the possibility to implement porosity without the use

of direct numerical simulation. Thereby avoiding the need of highly refined mesh near the cylinder, which might pose as a problem for the parametric study due to the number of cases required to be run. Next, the case setup with porosity model is validated using the results from *Steiros et al.*, which could indicate the validity of the proposed porosity model for this study.

The case setup consists of a rectangular box with dimensions $(x, y, x) = (1.45, 0.46, 0.27)$ m with a porous circular cylinder of radius $r = 0.01675$ m with wall thickness $\Delta x = 0.001$ m. The cylinder is located at a distance 0.21325 m from the inlet and both sides in the y-direction. The case setup is depicted in Figure 4.10.

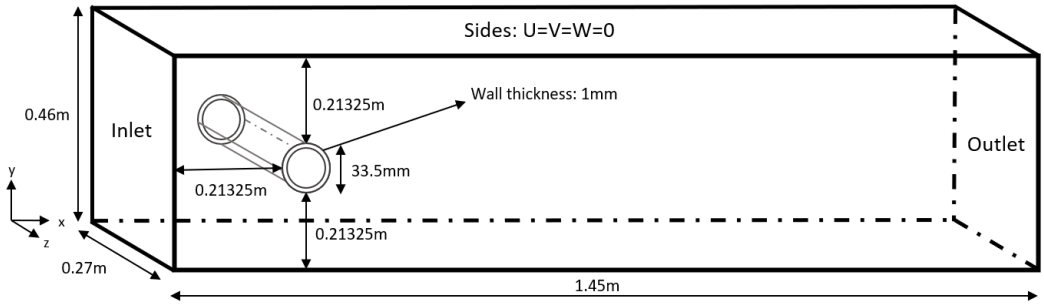
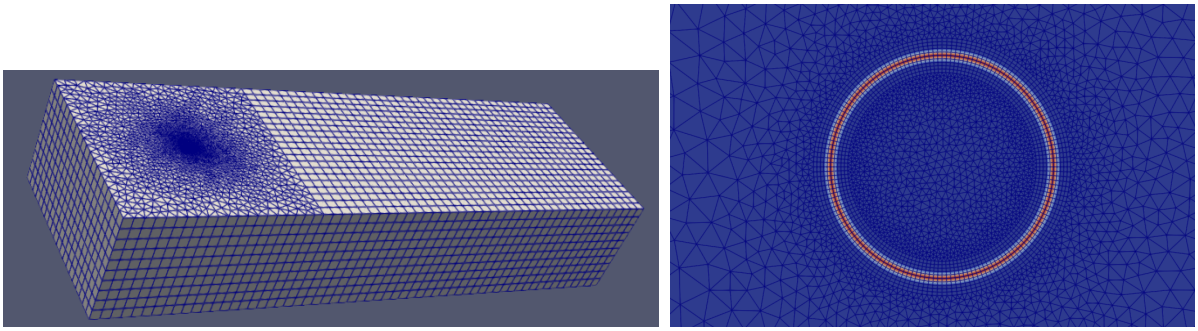


Figure 4.10: Case setup for the parametric study with a porous circular cylinder of diameter $D = 0.0335$ m and a wall thickness of 0.001 m.

The inlet has a constant uniform inflow velocity, all four sides have the no slip boundary condition and the outlet has a zero gradient boundary condition. The pressure boundary conditions are all zero gradient except for the outlet, which is set to a fixed value of zero. For the porosity model the factor χ from Equation 2.18 is calculated to be $\chi = 11.7647$. As described in the literature research, it is assumed that both $\psi = 0$ and $c = 0$.

Mesh

The mesh of the case description is presented below in Figure 4.11. The mesh has a total number of cells of 103780 and is checked using *OpenFOAM*'s *checkMesh* function.



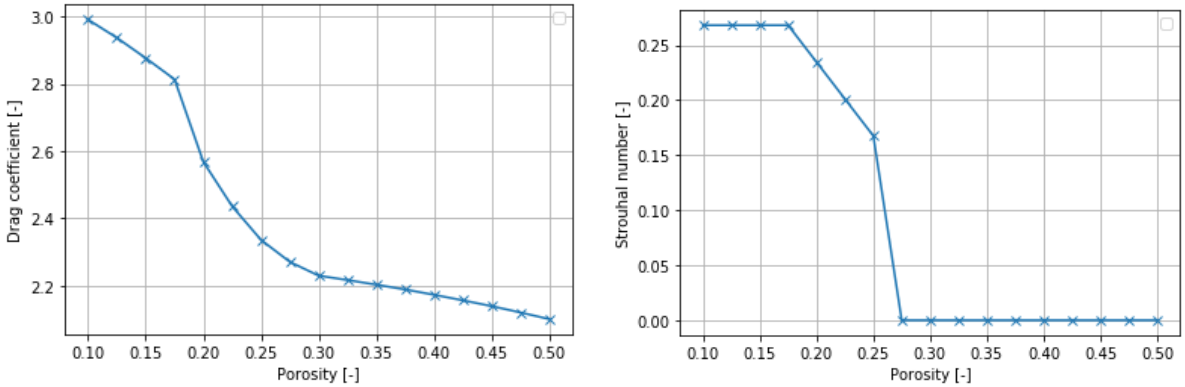
(a) Top view of the mesh.

(b) Section of the mesh on the cylinder (porous zone in red).

Figure 4.11: Mesh of the case for the parametric study.

Validation

The case setup is validated using the results presented in Figure 2.6, which displays the drag coefficient versus porosity. For non-porous media the drag coefficient is at a maximum of $C_d \approx 1.25$, then the drag coefficient drops down at 10% porosity to about $C_d \approx 1.1$, which remains stable up to 30% porosity, after which the drag coefficient decreases to $C_d \approx 0.9$ at 50% porosity and to $C_d \approx 0.7$ at 60% porosity. The resulting drag coefficient and Strouhal number for the case here, are presented in Figure 4.12a.

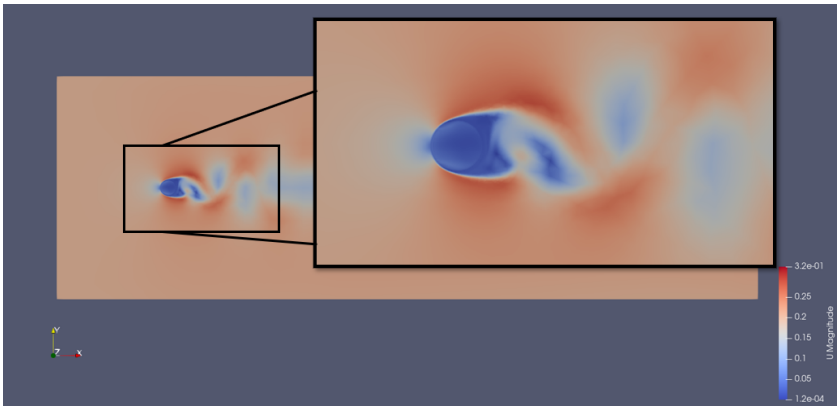


(a) Drag coefficient versus porosity.

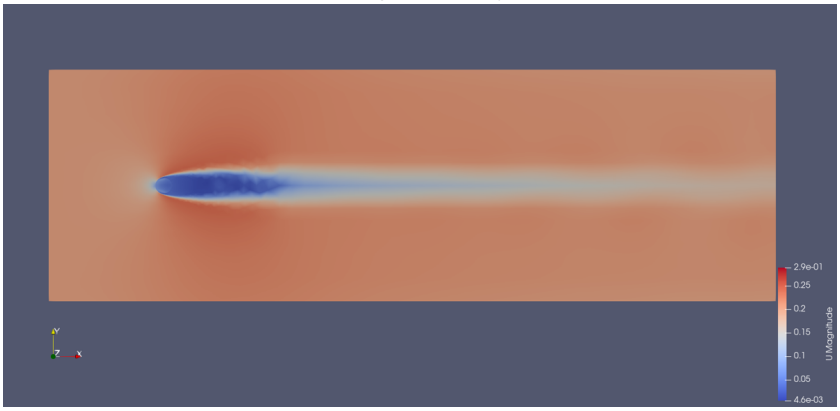
(b) Strouhal number versus porosity.

Figure 4.12: Resulting drag coefficient and Strouhal number versus porosity for the validation case with $Re = 6700$.

The drag coefficient has a shape and range of values that differ from the results from literature. The drag coefficient does decrease for increasing porosity in a similar way for the porosity values over 30%. It is possible that the drag coefficient is overestimated, since the thin walled assumption might not hold for this cylinder. However, even though the drag coefficient values do not correspond to the expected values from literature, the flow regime appears to be as expected for flow in the subcritical Reynolds numbers regime, meaning there is a von Kármán vortex street. To demonstrate, several images of developed flow for this case are depicted in Figure 4.13.



(a) Flow pattern of porous circular cylinder with inflow velocity $U = 0.2$ m/s and porosity of 20% at experiment time step $t = 11$ seconds. There is vortex shedding at this porosity.



(b) Flow pattern of porous circular cylinder with inflow velocity $U = 0.2$ m/s and porosity of 35% at experiment time step $t = 11$ seconds. There is no vortex shedding at this porosity.

Figure 4.13: Flow patterns of validation study on porous circular cylinder for varyin porosity.

Yet, at a porosity of 27.5%, this flow regime changes and no vortex shedding appears to be present. This cut-off of vortex shedding is also depicted in Figure 4.12b, where the Strouhal number appears to be slowly decreasing for increasing porosity with a sharp drop to $St = 0$ at a porosity of 27.5 %. For a non-porous cylinder the Strouhal number is expected to remain constant at $St = 0.2$ for the used subcritical Reynolds numbers regime. No further publications have been found on the subject of the relationship between the Strouhal number and porosity. Even though the estimated Strouhal number for a porous cylinder at $St = 0.27$ can not be validated, it is of the same order as for a non-porous cylinder. Therefore, the decision is made to accept the problems with the determination of the drag coefficient of this case as is and, instead focus on the vortex shedding frequency in the form of the Strouhal number.

4.2.2. Parametric study

As stated before, the parametric study consists of taking a number of equally distributed input values. There are two input values for each case, namely the inlet velocity and the porosity. The range of inlet velocities is $U = 0.1$ to $U = 1.0$ m/s with increments of $U = 0.1$ m/s, it is taken in such a way that the maximum possible Reynolds number still falls in the subcritical Reynolds number regime and also includes the single base case with $U = 0.2$ m/s as taken from literature [14]. The corresponding range of Reynolds numbers is $Re = 3350$ to $Re = 33500$ in increments of $Re = 3350$.

The porosity ranges from $\beta = 0.1$ to $\beta = 0.5$ with increments of $\beta = 0.025$. The lower boundary of this porosity range is set to 10%, because the computational cost increases extremely fast for low porosity, up to the point where the maximum allowed Courant number has to be reduced to $Co_{\max} = 0.1$ or even $Co_{\max} = 0.05$, which increases the computational cost of the current setup by a factor 4 or 8 respectively. The upper boundary is set with the idea of structural integrity in mind, based on the assumption that a porosity higher than 50% will not be realistically feasible to implement.

The parametric study is entirely implemented using *Dakota*, hence the entire process of case execution in *OpenFOAM* up to post processing and data extraction is automated. A total of 170 different input combinations are to be run, from which the vortex shedding frequency is obtained.

Depending on the input values, the flow develops at a minimum after an experimental time of about $t = 5$ seconds and at a maximum after an experimental time of about $t = 10$ seconds. Since it is deemed too difficult to automatically detect the force oscillations, which are expected in the developed flow, the decision is made to assume that for each of the 170 cases that the flow is developed after an experimental time of $t = 10$ seconds. Thus, the vortex shedding frequency is calculated in the time range after $t = 10$ up to $t = 15$ seconds.

Forces & coefficients

The total force and drag coefficient are calculated using *OpenFOAM*'s *forces* and *forceCoeffs* post processing tools, which use pressure integration over the boundary layer surface. These are required to in order to estimate the vortex shedding frequency.

Vortex shedding frequency

The process of estimating the vortex shedding frequency is depicted in Figure 4.14.

The vortex shedding frequency is obtained by applying the Fast Fourier Transform (FFT) to an interpolated time signal of the lift coefficient and finding the largest peak in the frequency spectrum. The FFT requires the time range to consist of a constant time step, which is not the case, since variable time stepping is implemented (using *OpenFOAM*'s *adjustableTimeStep* feature) to decrease computational cost. Hence the time spectrum with variable time step is fist interpolated to a time spectrum with constant time step of size $\Delta t = 0.0001$ s. The *forceCoeffs* post processing tool also exports the lift coefficient, which is used to estimate the vortex shedding frequency. The lift coefficient switches sign in case of oscillations and is averaged around $C_{\text{lift}} = 0$. The signal should be averaged around 0 in the temporal domain when applying FFT to avoid a static offset peak at frequency $f = 0$ Hz in the frequency domain. Lastly, the frequency of the largest peak in the FFT's frequency domain is selected as the vortex shedding frequency. Furthermore, a filter is added which estimates the amplitude of the drag coefficient. If the amplitude of the drag coefficient is smaller than 1% of the mean drag coefficient, it is assumed that there are no vortices or oscillations and thus the vortex shedding frequency $f_{vs} = 0$ Hz.

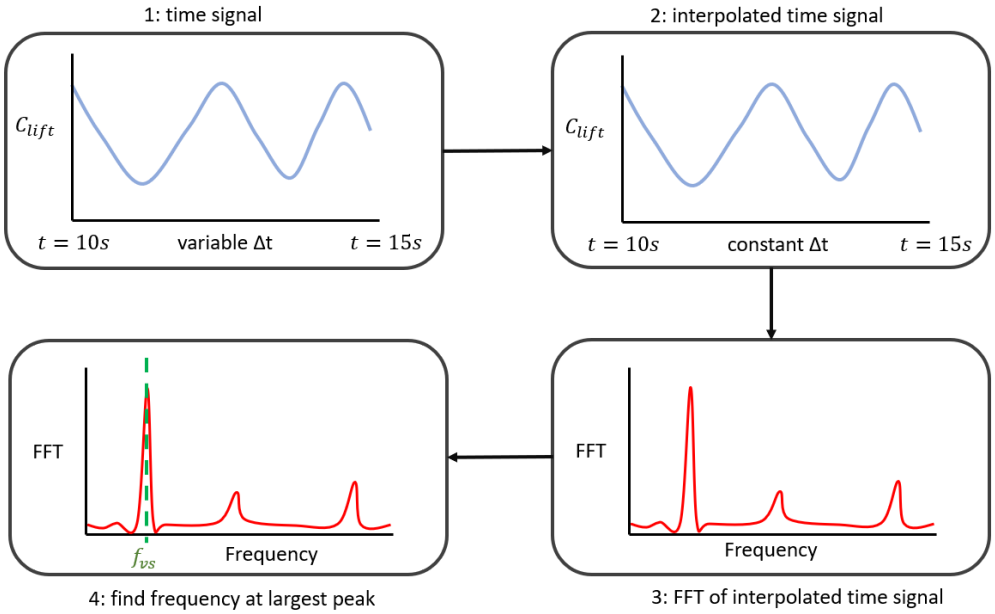


Figure 4.14: Schematic overview of the process of estimating the vortex shedding frequency. First the time signal is interpolated in order to have constant time step. Then the interpolated time signal is transformed to the frequency domain using the Fast Fourier Transform (FFT). Lastly, the frequency corresponding to the largest peak is found, which corresponds to the vortex shedding frequency.

4.2.3. Curve fitting

To finalize this section, curve fits are applied to two variations of the results by varying both inflow velocity and porosity and measure their effects on the vortex shedding frequency. Furthermore the inflow velocity and vortex shedding frequency can be rewritten into their respective non-dimensional forms, namely the Reynolds and Strouhal number. In Appendix A a table is presented with all possible curve fit functions that will be applied and ideally non-polynomial functions will fit the data and can be used as low-fidelity models.

4.3. Parametric study using Multi-Level Multi-Fidelity Monte Carlo

In this section, the final experiment of this thesis is described. Moreover, it combines both the MLMFMC and the parametric study on a porous circular cylinder. There are three goals of this experiment, namely:

1. Generating more data points to study the effects of porosity on the cylindrical Strouhal number; and
2. indicating validity of the proposed low-fidelity models for flow past a porous circular cylinder by analysis of the correlation between high- and low-fidelity models using Multi-Level Multi-Fidelity Monte Carlo; and
3. conducting an interval analysis using Multi-Level Multi-Fidelity Monte Carlo in order to estimate the bounds of the epistemic uncertainty of the high-fidelity model.

In order to accomplish these three goals, a new -but comparable- fluid dynamics case is proposed. The benchmark case for flow around a circular cylinder is modified to include porosity. Therefore the case is for flow around a porous circular cylinder. Moreover, a new parametric study is setup for a smaller range of porosity values at a singular Reynolds number in the subcritical Reynolds numbers regime. At each point of the new parametric study, an interval analysis is conducted using MLMFMC. Furthermore, at each point the correlation between high- and low-fidelity models is analysed. And lastly, inherent to the application of a parametric study, more data points are generated. In subsection 4.3.1 the proposed case description is described and constructed. In subsection 4.3.2 the proposed parametric study is described and in subsection 4.3.3 the interval analysis and settings of the MLMFMC are described. An overview of the subsections is depicted in Figure 4.15

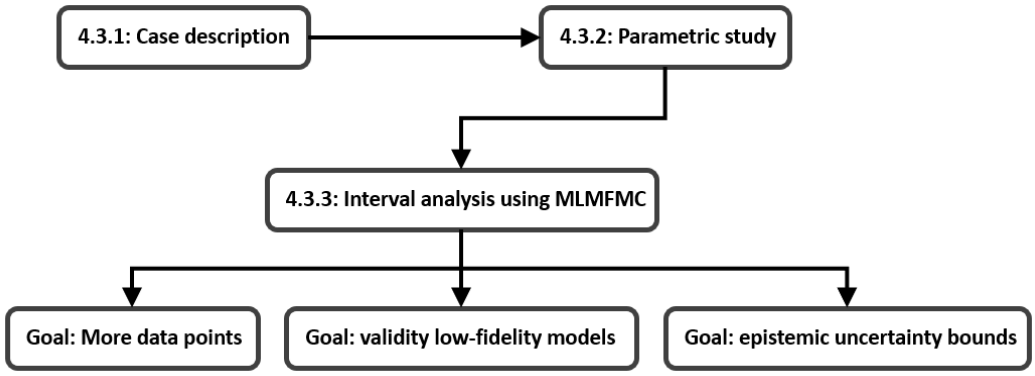


Figure 4.15: Overview of the section regarding the interval analysis using Multi-Level Multi-Fidelity Monte Carlo methods on the parametric study. First the case description and parametric study are described. Lastly, the interval analysis using Multi-Level Multi-Fidelity Monte Carlo methods is described.

4.3.1. Case description

The case setup is similar to that of the benchmark cases (subsection 4.1.1) with equal geometry. The difference being that the cylinder is porous and that the inflow condition is uniform constant instead of parabolic in space. Since the wall thickness of the porous cylinder may also influence the results, it is decided to maintain a constant cylinder diameter over wall thickness ratio as in the previous parametric study. Therefore, the wall thickness of this porous cylinder is: $D/t = 0.0335/0.001 = 0.1/t \longrightarrow t = 0.00299 \approx 0.003$ m. A schematic of the case geometry is depicted in Figure 4.16.

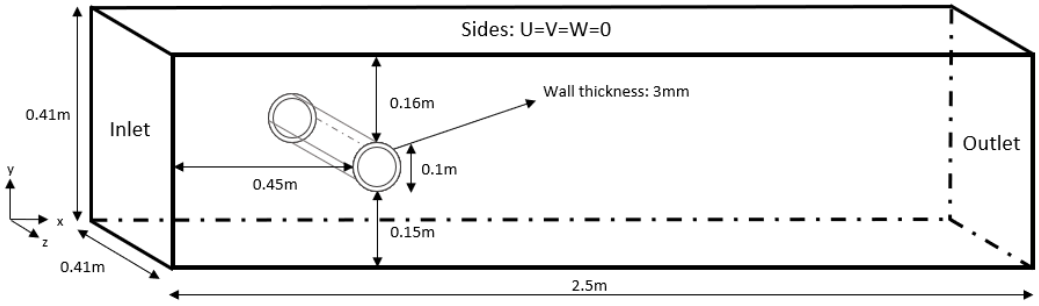


Figure 4.16: Geometry the 3D benchmark case with porous circular cylinder. Total size of rectangular plane is 2.2x0.41 m with a cylinder with a diameter of 0.1 m and no slip boundary conditions imposed on the top side, bottom side and the cylinder, zero gradient boundary conditions on the outlet and a uniform constant inflow condition on the inlet.

The boundary conditions for velocity are no slip on the four sides, zero gradient on the outlet and uniform constant on the inlet. For pressure, the boundary conditions on the sides and inlet are zero gradient and the outlet is uniform constant zero.

Mesh

The mesh is constructed for three resolution levels, which are presented from lowest to highest resolution from top to bottom in Figure 4.18. On the right side a close up is depicted of the porous area for each resolution level. No convergence study is conducted on the meshes, only *OpenFOAM*'s *checkMesh* function is applied.

4.3.2. Parametric study

The goal of this parametric study is to generate more data points on the effects of increasing porosity on the cylindrical Strouhal number. The results of the previous parametric study (section 4.2) indicate the largest change in cylindrical Strouhal number occurs for the porosity range of $0.20 < \beta < 0.30$. Therefore, the new parametric study consists of eleven equally distributed points in the range $0.20 < \beta < 0.30$ with increments of $\Delta\beta = 0.01$. The Reynolds is not varied for this parametric study and set to $Re = 10000$.

4.3.3. Interval analysis using Multi-Level Multi-Fidelity Monte Carlo

Each value of the parametric study is estimated using the MLMFMC method in order to reduce the suspected epistemic uncertainty. An overview of the process of MLMFMC is depicted in Figure 4.17.

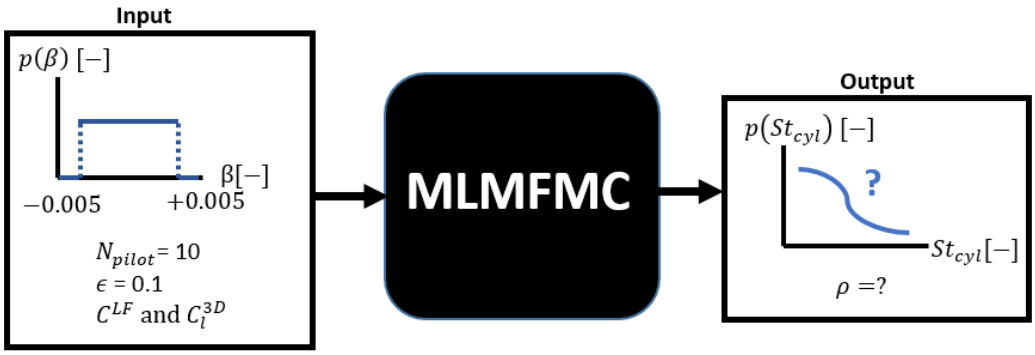


Figure 4.17: The Multi-Level Multi-Fidelity Monte Carlo method is implemented on the parametric study in order to reduce suspected epistemic uncertainty by estimating each point using interval analysis.

After analysing the results of the parametric study (section 4.2), several outlier cases were detected. Reasons for the outlier cases vary and these are most likely due to systemic errors in the vortex shedding frequency detection algorithm. Uncertainties of this kind are epistemic uncertainties. Furthermore, the full simulation process of the parametric study from input to output consists of many separate parts working together (e.g. *Dakota*, *OpenFOAM*, *olaFlow* and several *Python* and *Bash* scripts). In order to quantify the epistemic uncertainty, the MLMFMC method is combined with the parametric study in order to propagate epistemic uncertainty. In the paper by *Swiler et al.* [23] three methods of propagation of epistemic uncertainty are described: interval analysis, Dempster-Shafer Theory of Evidence and second order probability. The interval analysis is implemented, since it is assumed nothing is known about the uncertain input variables, except that these values lie in a certain range of values. This range of values coincides with that of the parametric study. An interval analysis using random sampling is implemented with MLMFMC and the intermediary results are analysed in order to estimate the boundaries of the epistemic uncertainty. Therefore, at each porosity value of the parametric study, an interval of $\beta \pm 0.005$ is assigned with uniform distribution.

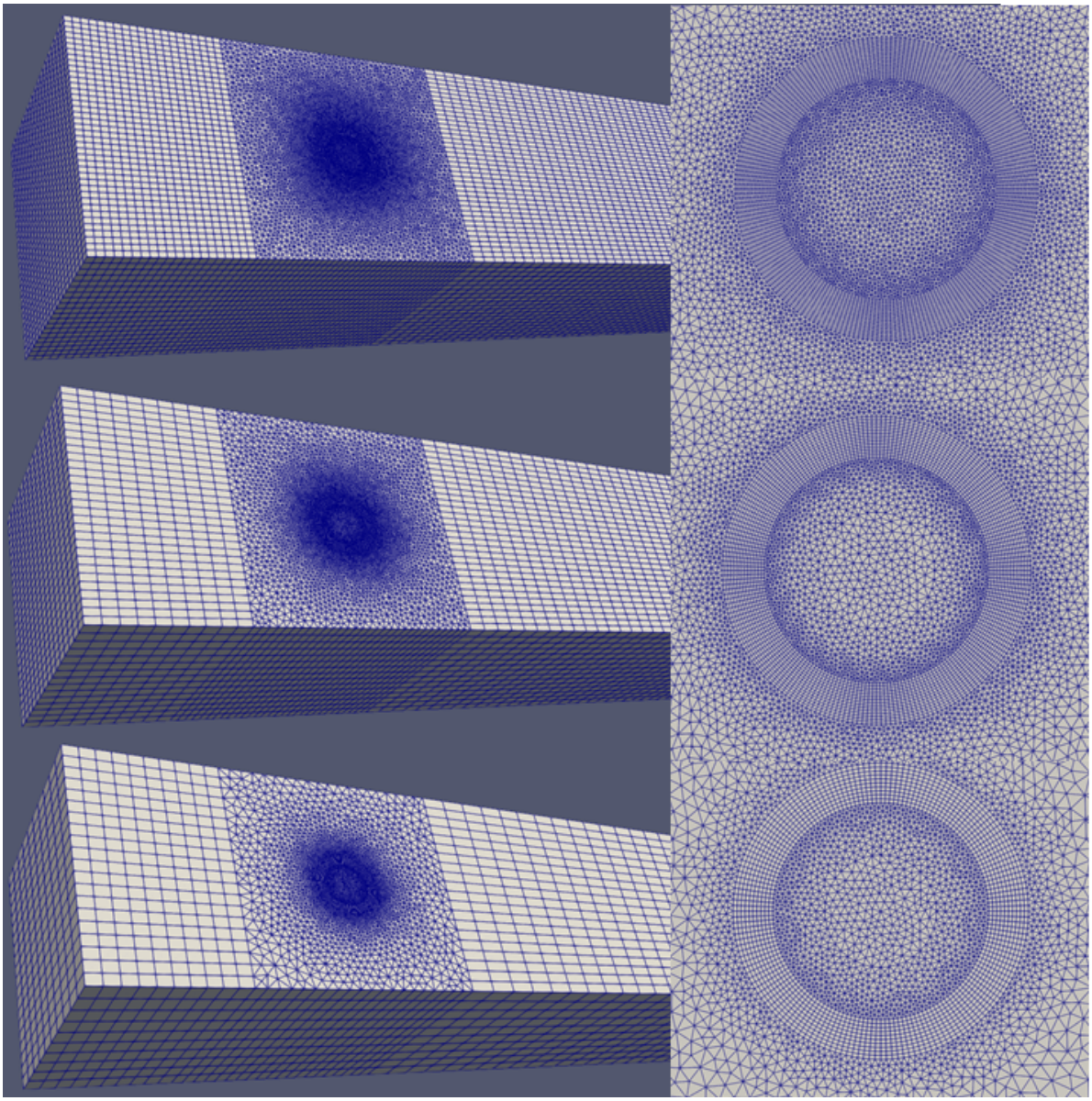


Figure 4.18: The meshes of the 3D cases with lowest to highest resolution level from top to bottom respectively.

The experiments described in this chapter in section 4.1, section 4.2 and section 4.3 serve as the basis for the results, as described in chapter 5, chapter 6 and chapter 7 respectively.

5

Results: benchmark cases

In this chapter the results of the first case study: benchmark cases are described. The goal of this chapter is to demonstrate a successful implementation of the MLMFMC approach using a high-fidelity 3D model and a low-fidelity 2D model implemented in *OpenFOAM*. Furthermore the estimator variance reduction and computational cost reduction of both MLMC and MLMFMC are demonstrated using benchmark cases of flow around a circular cylinder.

In section 5.1 the relative computational cost of both 2D and 3D cases are estimated using simulation time averaging. In the second section the MLMC is implemented on a 2D and 3D CFD case, demonstrating computational cost reduction and furthermore, the results are compared with the results of the respective equivalent number of high resolution general MCM simulations. In the third section the MLMF is implemented for a 3D CFD case, again demonstrating computational cost reduction and comparing the results with the results of the equivalent number of high resolution high-fidelity general MCM simulations. Lastly, the true value is estimated using the data points from all simulations and the results are summarized.

5.1. Relative computational cost

In this section the results of the process of estimating the relative computational cost are described for both 2D and 3D meshes. The computational costs are estimated using simulation time averages of the runs.

As can be seen in Table 5.1 and Table 5.2 there is a large range in computational time values at each resolution level. As expected for both 2D and 3D cases, the higher resolution levels will take more time than the lower resolution levels due to the increased number of cells. As a consequence the cell size is decreased as well, which influences the Courant number and thus indirectly forces the time step to decrease as well. The range differences in computational times at each resolution level could be attempted to be explained using the mean inflow velocity, since the mean inflow velocity ranges from 0.2 m/s to 1.0 m/s, this means that at a maximum the velocity is a factor 5 larger. And although this effect does not one to one explain the differences between the minimum and maximum values in the same column, it could play a significant role. Especially in the 2D cases this factor is larger ranging from 4.95 to 6.30, whereas in the 3D cases this factor ranges from 2.99 to 3.47.

	low resolution [-]	middle resolution [-]	high resolution [-]
computational cost	1	6.47	23.54

Table 5.3: Estimated relative computational cost of each resolution level for the 2D benchmark case.

	low resolution [-]	middle resolution [-]	high resolution [-]
computational cost	1	12.15	53.94

Table 5.4: Estimated relative computational cost of each resolution level for the 3D benchmark case.

Case number	low resolution [s]	middle resolution [s]	high resolution [s]
1	198.81	1332.59	4997.63
2	170.55	1149.84	4266.64
3	205.40	1370.87	5123.36
4	43.92	244.28	829.08
5	63.82	364.36	1242.66
6	129.37	876.69	3188.49
7	120.74	780.82	2805.19
8	172.45	1142.69	4275.71
9	51.57	290.12	980.96
10	42.91	245.81	829.47
11	134.88	920.65	3356.31
12	48.32	272.40	920.22
13	99.09	613.30	2127.55
14	101.83	621.40	2206.50
15	109.98	701.19	2497.84
16	212.59	1413.47	5221.65
sum	1906.23	12340.43	44869.26
max	212.59	1413.47	5221.65
min	42.91	244.28	829.08
factor	4.95	5.77	6.30
mean	119.14	771.28	2804.33

Table 5.1: Computational times for 16 benchmark cases at three resolution levels with total summed value, maximum value, minimum value, multiplicative factor between minimum and maximum value and mean value for the 2D benchmark case.

Using the mean computational time values from Table 5.1 and Table 5.2 the relative computational cost is calculated for the 2D and 3D benchmark case and presented in Table 5.3 and Table 5.4 respectively. The same data used for calculating the relative computational costs for the 2D and 3D benchmark cases for MLMC is used, however since the MLMF uses a low-fidelity model and a high fidelity model the relative computational cost is re-calculated with respect to lowest resolution of the low-fidelity model. Hence the lowest resolution of the 2D case will be assigned a relative computational cost of $C_0^{2D} = 1$ and the other relative computational costs are calculated with respect to this lowest computational cost, as is presented in Table 5.5.

	low resolution [-]	middle resolution [-]	high resolution [-]
2D	1	6.47	23.54
3D	2.24	27.19	120.69

Table 5.5: Total estimated relative computational cost of each resolution level of the low-fidelity (2D) model and the high fidelity (3D) model for Multi-Level Multi-Fidelity Monte Carlo.

5.2. Multi-Level Monte Carlo

In this section the results from two separate implementations of MLMC are described. The first case is MLMC in 2D and the second case is MLMC in 3D. The input parameters for an MLMC approach - which are presented in chapter 3 in Figure 3.1- are the computational cost per resolution level C_l, tolerance ϵ, the numer of pilot samples N_{pilot} and the random input variable $\omega = \bar{U}$, which is the mean inflow velocity for both MLMC simulations. For consistency the random variable $\bar{U} = 0.2$ to $\bar{U} = 1.0$ m/s, tolerance $\epsilon = 0.1$ and number of samples $N_{pilot} = 10$ are used for both 2D and 3D cases.

First the results from the 2D MLMC case are described. Starting with the resulting preliminary statistics which are used to calculate the optimal number of samples per resolution level. Then, the estimated value of the drag coefficient and estimator variance from the MLMC are discussed and the computational cost reduction is calculated. Lastly, a High-Level Monte Carlo (HLMC) simulation is executed and the resulting estimated value and estimator variance are discussed and compared to the MLMC. Secondly the results from the 3D MLMC are described, which follows the same procedure as that for the 2D MLMC.

Case number	low resolution [s]	middle resolution [s]	high resolution [s]
1	396.66	4902.49	22631.80
2	351.09	4246.78	18922.90
3	401.34	5002.52	23362.50
4	134.27	1626.02	6727.88
5	185.24	2258.19	9676.08
6	293.56	3602.81	16428.00
7	282.54	3502.94	14838.10
8	354.32	4249.79	19076.80
9	156.70	1902.91	7882.49
10	135.00	1621.88	6878.25
11	297.40	3657.15	16613.90
12	151.80	1779.79	7615.47
13	246.37	3009.67	13012.20
14	246.86	2986.80	13858.60
15	246.80	2864.73	12606.80
16	385.12	4617.68	19924.10
sum	4265.07	51832.15	230055.87
max	401.34	5002.52	23362.50
min	134.27	1621.88	6727.88
factor	2.99	3.08	3.47
mean	266.57	3239.51	14378.49

Table 5.2: Computational times for 16 benchmark cases at three resolution levels with total summed value, maximum value, minimum value, multiplicative factor between minimum and maximum value and mean value for the 3D benchmark case.

5.2.1. Case in 2D

As described before the MLMC approach requires four input parameters, for the case of 2D MLMC these are summarized in Table 5.6

Fidelity	**Level** l	**Cost** C_l **[-]**	**Random variable** $\bar{U}$ **[m/s]**	**Tolerance** ϵ **[-]**	**Samples** N_{pilot} **[-]**
2D	0	1	Uniform: 0.2 - 1.0	0.1	10
2D	1	6.47			
2D	2	23.54			

Table 5.6: Summary of input variables of the Multi-Level Monte Carlo simulation in 2D. Displaying the estimated relative computational cost of each resolution level for a Multi-Level Monte Carlo approach in 2D. The random variable is the mean inflow velocity U_m ranging from $U_m = 0.2$ to $U_m = 1.0$ m/s, samples are drawn from a uniform distribution. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$.

Preliminary statistics from pilot samples

The number of pilot samples is set to $N_{pilot} = 10$, the pilot samples are taken at each resolution level l. Moreover at each resolution level the difference function is calculated. Therefore at resolution level $l = 1$ and $l = 2$ twice the number of pilot samples are taken. In total, 50 samples are obtained. In Figure 5.1 a histogram is presented using a color map to depict which resolution level corresponds to which value of the drag coefficient.

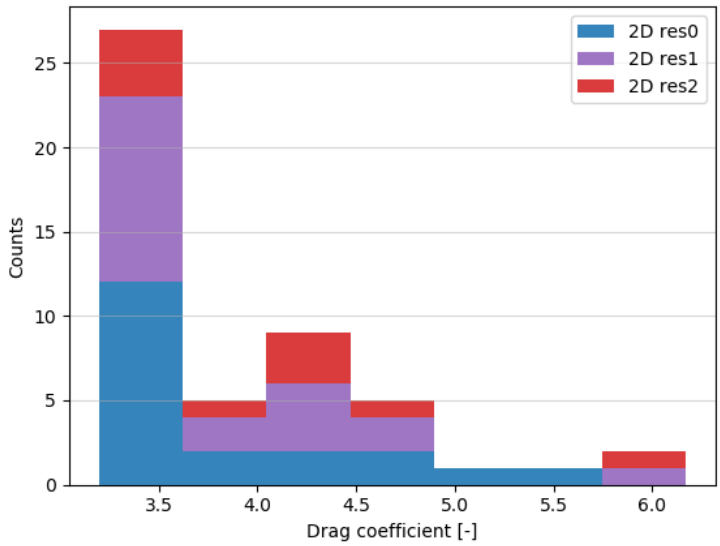


Figure 5.1: Histogram of drag coefficient as a result of general Monte Carlo Methods for number of pilot samples $N_{pilot} = 10$ for three resolution levels, hence 50 total samples. Color coding is applied to demonstrate which resolution level corresponds to which result.

Results from Multi-Level Monte Carlo

The MLMC simulation results in the following histogram in Figure 5.2, again depicting the counts versus the drag coefficient with the color coding used to denote which resolution level is responsible for which sample. Clearly, the samples are redistributed to the lower resolution levels. To summarize, the results of the MLMC are presented in Table 5.7. An interesting observation is that at the middle resolution level there are less samples taken than at the highest resolution level. This could indicate that the mesh differences lowest and middle resolution are too small and maybe a more refined mesh should be used.

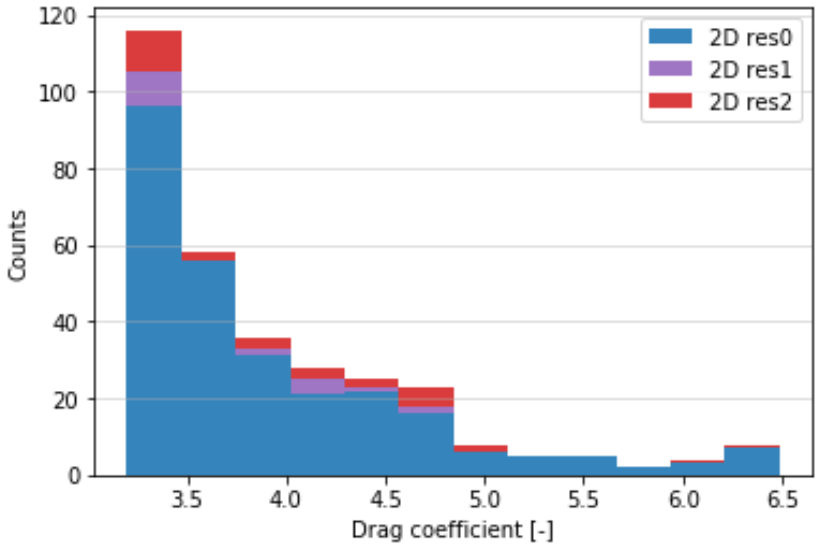


Figure 5.2: Histogram of the drag coefficient as a result of a Multi-Level Monte Carlo approach in 2D with three resolution levels, denoted by res0, res1 or res2. Color coding is applied to demonstrate which resolution level corresponds to which result.

Fidelity	**Resolution level** l	**Computational cost** C_l [-]	**Optimal samples** N_l [-]	**Total cost** [-]
2D	0	1	270	270.00
2D	1	6.47	18	116.46
2D	2	23.54	30	706.20
Total summed cost of MLMC:				1092.66
N_{eq} **samples according to MLMC:**			55.44 → 56	1318.24

Table 5.7: Summary of estimated relative computational cost, optimal number of samples and total cost of each resolution level of the low-fidelity (2D) model for a Multi-Level Monte Carlo approach.

The estimated drag coefficient is $C_d = 3.92$ with a standard deviation of $\sigma_{C_d} = 0.77$. The total cost of the MLMC is calculated to be 1092.66 and the cost of the number of equivalent high resolution samples is 1318.24. Therefore, the cost reduction of MLMC with respect to normal MCM is 17 %. In order to validate the estimated drag coefficient and standard deviation, the high resolution Monte Carlo simulation is run.

High-Level Monte Carlo

In Figure 5.3 the histogram of the estimated samples are presented. The estimated drag coefficient is $C_d = 4.28$, which is a relatively large difference with respect to the MLMC. The standard deviation is $\sigma_{C_d} = 0.77$ is equal to that of the MLMC.

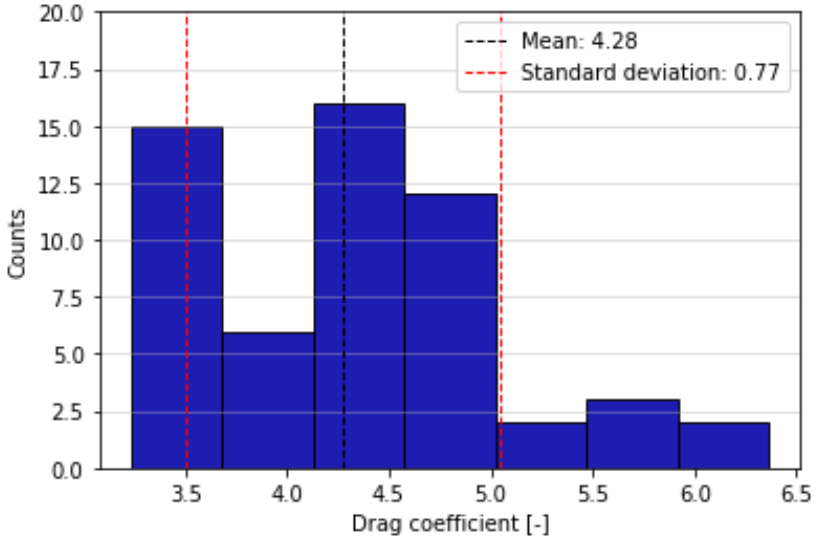


Figure 5.3: Histogram as a result of the high level Monte Carlo run with 56 samples using a 2D case setup.

5.2.2. Case in 3D

In this section, the MLMC is implemented for the 3D case, the MLMC input parameter setup is equal to that of the 2D MLMC with a difference in relative computational cost. The input parameters are summarized in Table 5.8.

Fidelity	**Level** l	**Cost** C_l [-]	**Random variable** $\tilde{U}$ **[m/s]**	**Tolerance** ϵ [-]	**Samples** N_{pilot} [-]
3D	0	1	Uniform: 0.2 - 1.0	0.1	10
3D	1	12.15			
3D	2	53.94			

Table 5.8: Summary of input variables of the Multi-Level Monte Carlo simulation in 3D. Displaying the estimated relative computational cost of each resolution level for a Multi-Level Monte Carlo approach in 3D. The random variable is the mean inflow velocity U_m ranging from $\tilde{U} = 0.2$ to $\tilde{U} = 1.0$ m/s, samples are drawn from a uniform distribution. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$.

Preliminary statistics from pilot samples

In Figure 5.4 the distribution of samples from the pilot samples is depicted. The shape is similar to that of the pilot samples of the MLMC in 2D. A total of 50 samples are taken and most of these are located at the lower values of the drag coefficient.

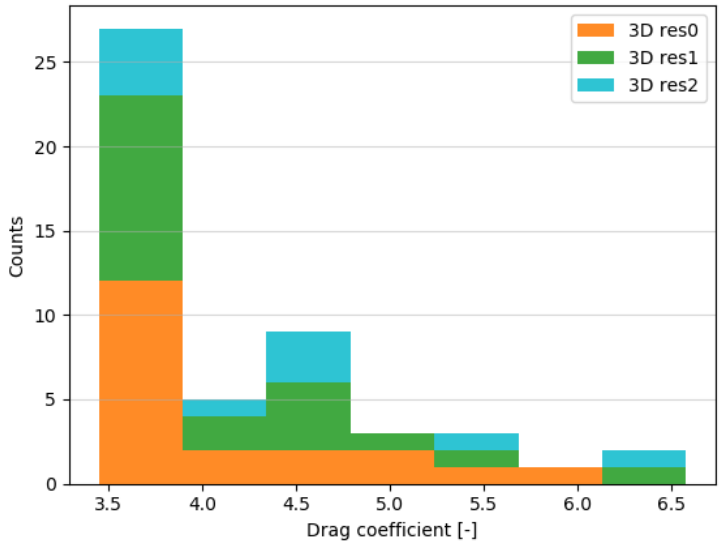


Figure 5.4: Histogram of drag coefficient as a result of general Monte Carlo Methods for number of pilot samples $N_{pilot} = 10$ for three resolution levels, hence 50 total samples. Color coding is applied to demonstrate which resolution level corresponds to which result.

Results from Multi-Level Monte Carlo

The MLMC results in the histogram in Figure 5.5. As expected, the computational cost is shifted towards the lowest resolution. The resulting samples per resolution level and the respective computational costs are presented in Table 5.9. The estimated drag coefficient is $C_d = 4.18$ with a standard deviation of $\sigma_{C_d} = 0.79$. Using the MLMC a computational cost reduction of 14.82% is achieved.

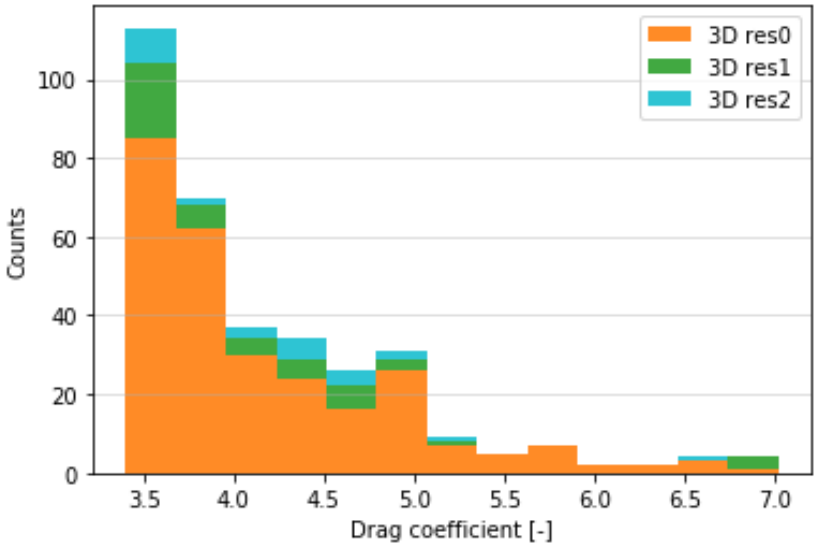


Figure 5.5: Histogram of the drag coefficient as a result of a Multi-Level Monte Carlo approach in 3D with three resolution levels, denoted by res0, res1 or res2. Color coding is applied to demonstrate which resolution level corresponds to which result.

Fidelity	Resolution level l	Computational cost C_l [-]	Optimal samples N_l [-]	Total cost [-]
3D	0	1	270	270
3D	1	12.15	47	571.05
3D	2	53.94	27	1456.38
Total summed cost of MLMC:				2297.43
N_{eq} **samples according to MLMC:**			49.55 → 50	2697

Table 5.9: Summary of estimated relative computational cost, optimal number of samples and total cost of each resolution level of the low-fidelity (2D) model for a Multi-Level Monte Carlo approach.

High-Level Monte Carlo

The High-Level Monte Carlo simulation for the 3D case requires 50 samples to be taken. The resulting histogram is depicted in Figure 5.6, with an estimated drag coefficient of $C_d = 4.29$ with a standard deviation of $\sigma_{C_d} = 0.81$.

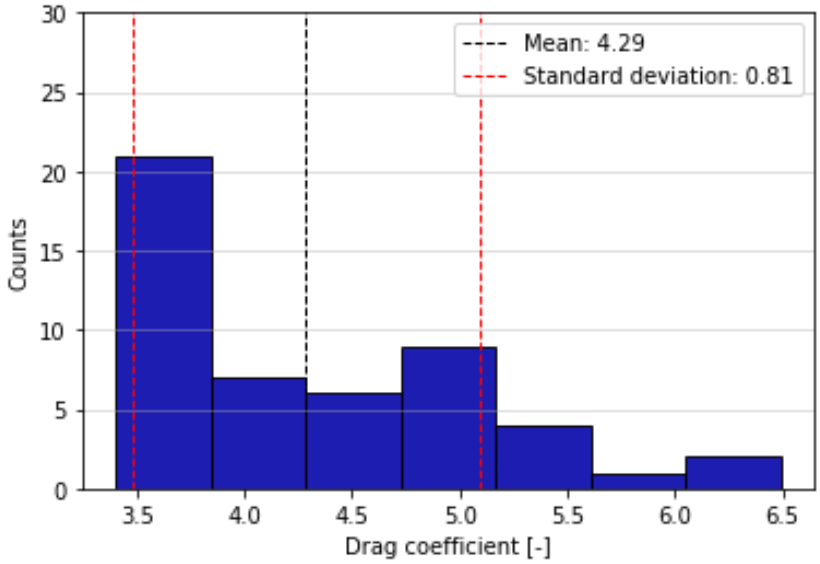


Figure 5.6: Histogram of High-Level Monte Carlo simulation of 50 samples using a 3D case setup.

5.3. Multi-Level Multi-Fidelity Monte Carlo

In this section the results from the process of implementing the MLMFMC approach are described. As depicted in Figure 3.2 the MLMFMC has four different input parameters, which are displayed in Table 5.10. The relative computational cost per model fidelity and per level are taken from section 5.1. One random variable is chosen, namely the mean inflow velocity, which is uniformly distributed between $U = 0.2$ and $U = 1.0$ m/s. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$.

Fidelity	Level l	Cost C_l [-]	Random variable $\tilde{U}$ [m/s]	Tolerance ϵ [-]	Samples N_{pilot} [-]
2D	0	1	Uniform: 0.2 - 1.0	0.1	10
2D	1	6.47			
2D	2	23.54			
3D	0	2.24			
3D	1	27.19			
3D	2	120.69			

Table 5.10: Summary of input variables of the Multi-Level Multi-Fidelity simulation. Displaying the estimated relative computational cost of each resolution level of the low-fidelity (2D) model and the high-fidelity (3D) model for a Multi-Level Multi-Fidelity approach. The random variable is the mean inflow velocity U_m ranging from $\tilde{U} = 0.2$ to $\tilde{U} = 1.0$ m/s, samples are drawn from a uniform distribution. The tolerance is set to $\epsilon = 0.1$ and the number of pilot samples is set to $N_{pilot} = 10$.

Firstly, the results from the general MCM for the pilot samples are discussed, including the correlation between the high- and low-fidelity models per resolution level. Afterwards, the results from the MLMFMC approach are discussed and the computation cost reduction is calculated. Lastly, the results of the MLMFMC approach are compared with the results from the High-Level High-Fidelity Monte Carlo (HLHFMC) simulation.

Preliminary statistics from pilot samples

Before the variance reduction aspect of MLMF is applied, several parameters ($\rho_l, r_l, \Lambda_l, \alpha_l$ and N_l, see Figure 3.2 top right corner) need to be calculated using estimated values from general MCM with a number of samples equal to the number of pilot samples $N_{pilot} = 10$. Since there are only two model fidelites, namely high- and low-fidelity with three resolution levels each, 60 samples are taken. Moreover, to calculate the difference between two resolution levels, the random input variables of both levels need to be the same, hence for resolution level 1 and 2, 10 additional samples have to be taken for both model fidelities. In total, 100 pilot samples are taken in order to setup the MLMF approach. The resulting histogram with corresponding color coding to depict resolution level and model fidelity is depicted in Figure 5.7.

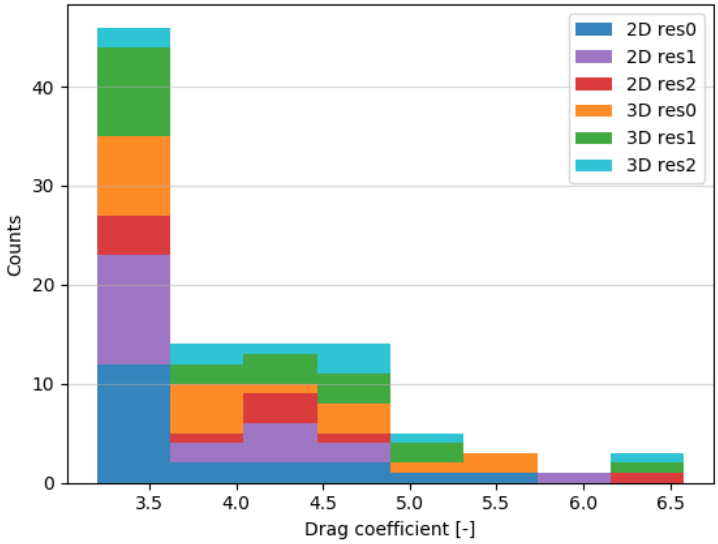
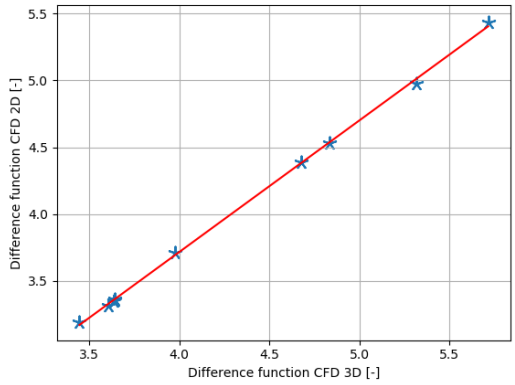


Figure 5.7: Histogram of drag coefficient as a result of general Monte Carlo Methods for number of pilot samples $N_{pilot} = 10$ for two model fidelities with three resolution levels each, hence 100 total samples. Color coding is applied to demonstrate which model fidelity and resolution level corresponds to which result.

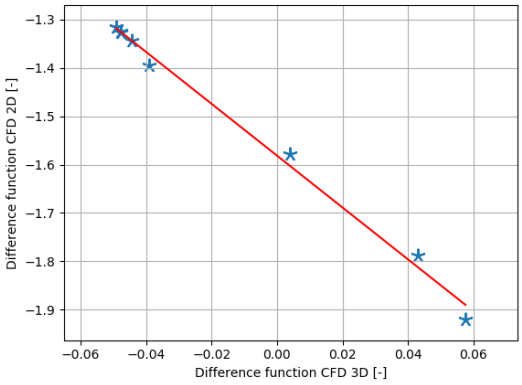
Correlation between models

The correlation between models is determined during the preliminary statistical analysis and corresponds to the Pearson correlation coefficient. In Figure 5.8 three scatter plots of the difference function of the low-fidelity model (CFD 2D disk) versus the high-fidelity model (CFD 3D) are depicted for each resolution level l. The red line is a best fit applied to the scatter plot, which indicates if there is positive or negative correlation. For the lowest and highest resolution level there is positive correlation. For the middle resolution the correlation is negative. This indicates that at the middle resolution level, the difference between the middle and lower resolution levels is negative. From the scatter plots the correlation matrices are calculated. As expected from the scatter plots there is high correlation. More specifically the Pearson correlation coefficients at each resolution level are calculated as:

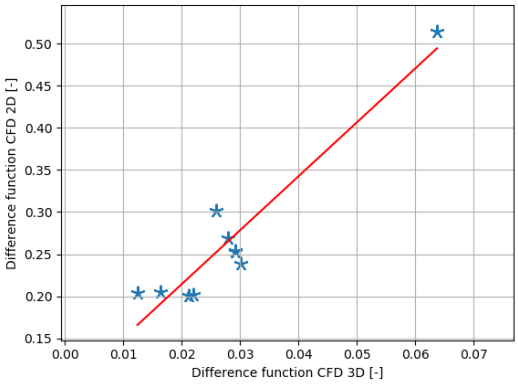
- $\rho_0 = 0.999690$
- $\rho_1 = -0.996981$
- $\rho_2 = 0.94865$



(a) Resolution level $l = 0$.



(b) Resolution level $l = 1$.



(c) Resolution level $l = 2$.

Figure 5.8: Scatter plot of high-fidelity model (CFD 3D) vs low-fidelity model (CFD 2D) for varying resolution levels $l = 0, 1, 2$ with a best fit line in red.

The Pearson correlation coefficient at the highest resolution appears to be the lowest. This is expected, since the highest resolution level should provide the most accurate results. So, at the highest resolution level the model fidelity differences should be more noticeable than at the lower resolution cases. The Pearson correlation coefficient at the lowest and middle resolution level lie closer together in absolute value, but the correlation coefficient at the middle resolution level is negative.

Results from Multi-Level Multi-Fidelity Monte Carlo & computational cost reduction

In Figure 5.9 the histogram of the drag coefficient resulting from the MLMFMC simulation is depicted with an estimated drag coefficient of $C_d = 3.89$ and standard deviation of $\sigma_{C_d} = 0.73$. Both the estimated value and the standard deviation of the drag coefficient are reduced with respect to the histogram of the pilot samples.

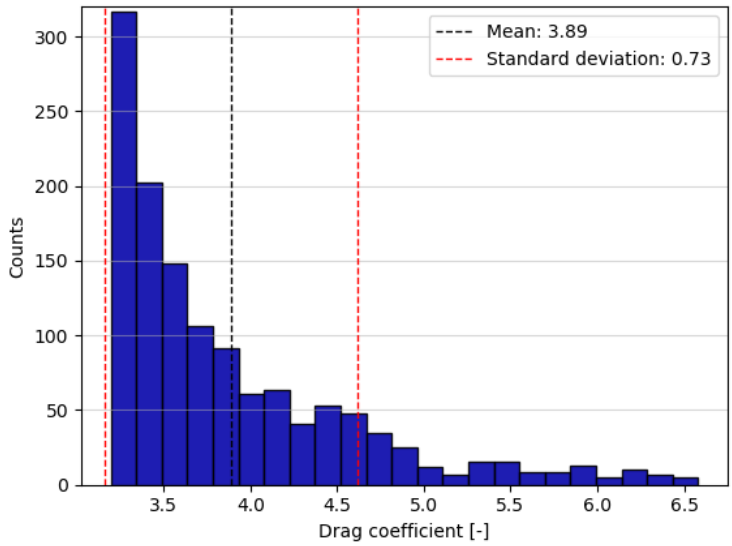


Figure 5.9: Histogram of the drag coefficient as a result of a Multi-Level Multi-Fidelity approach with two model fidelities with three resolution levels each. The estimated drag coefficient is $C_d = 3.89$ with a standard deviation of $\sigma_{C_d} = 0.73$

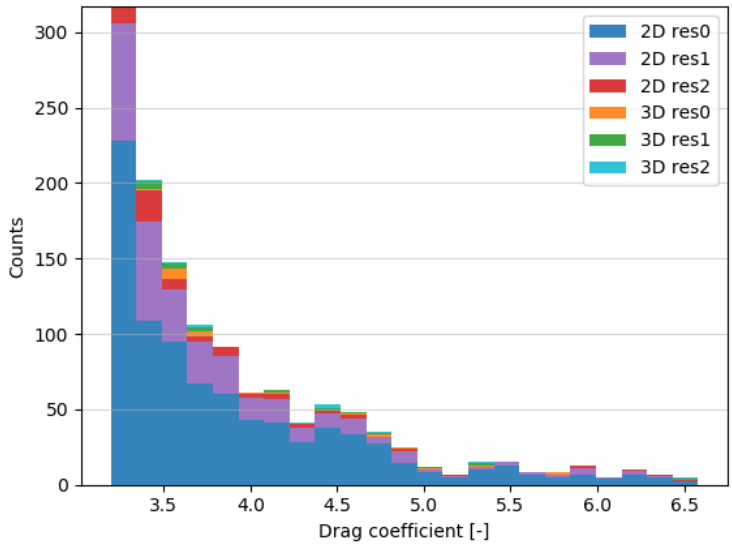


Figure 5.10: Histogram of the drag coefficient as a result of a Multi-Level Multi-Fidelity approach with two model fidelities, denoted by 2D or 3D, with three resolution levels each, denoted by res0, res1 or res2.

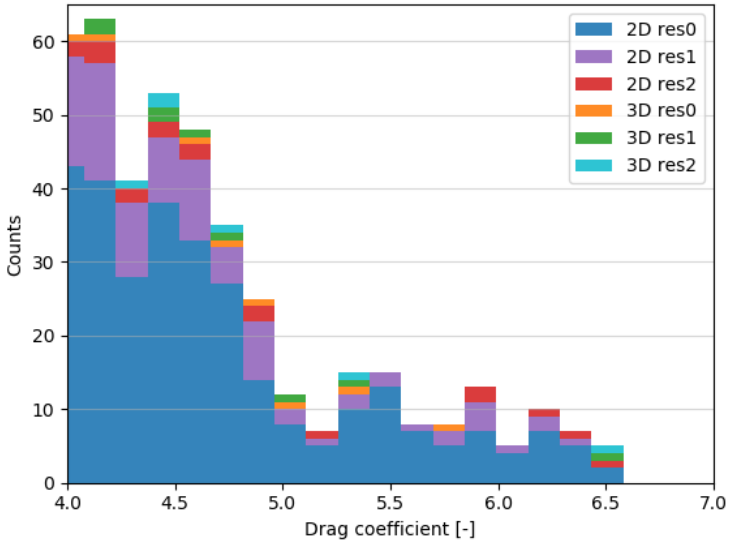


Figure 5.11: Zoomed in histogram of the drag coefficient as a result of a Multi-Level Multi-Fidelity approach with two model fidelities, denoted by 2D or 3D, with three resolution levels each, denoted by res0, res1 or res2.

Figure 5.9 is presented in both Figure 5.10 and Figure 5.11 with color coding to demonstrate how the computational burden is redistributed to the lower resolution and lower-fidelity model. The lowest resolution low-fidelity model denoted by a blue color, clearly corresponds to the largest number of samples taken. Followed by the middle resolution low-fidelity model in purple and the highest resolution low-fidelity model in red. Relatively, there are so little cases run by the high-fidelity model that these are quite difficult to see in Figure 5.10. For visibility Figure 5.11 is added, which is the same histogram with a lower bound on the drag coefficient (starting at $C_d = 4.0$) and an upper bound on number of counts (ending at Counts = 65). Here, the high-fidelity model samples from lowest to highest resolution are depicted by orange, green and cyan respectively. The number of samples of each model fidelity and resolution are summarized in Table 5.11. Furthermore the total summed cost of the MLMFMC approach and the number of equivalent high-fidelity samples N_{eq} are presented.

Fidelity	**Resolution level** l	**Computational cost** C_l [-]	**Optimal samples** N_l [-]	**Total cost** [-]
2D	0	1	601	601
2D	1	6.47	255	1649.85
2D	2	23.54	67	1577.18
3D	0	2.24	10	22.40
3D	1	27.19	10	271.90
3D	2	120.69	10	1206.90
Total summed cost of MLMFMC:				5329.23
N_{eq} **samples according to MLMFMC:**			52.31 → 53	6396.57

Table 5.11: Summary of estimated relative computational cost, optimal number of samples and total cost of each resolution level of the low-fidelity (2D) model and the high-fidelity (3D) model for a Multi-Level Multi-Fidelity approach.

For the high-fidelity model only the pilot samples are run, since for each resolution level l the optimal number of samples per level $N_l = 10$. This is unsurprising, since the correlation between the model fidelities is high and thus the MLMFMC approach deems the high-fidelity model cases to expensive to run with respect to the low-fidelity model cases. The total summed cost of the MLMFMC approach is 5329.23, whereas the total cost of the equivalent number of high-fidelity samples is 6396.57. This means that there is a decrease in computational cost of 17%. In order to compare results of the MLMFMC approach with the equivalent number of high-fidelity samples, a general MCM study is executed.

High-Level High-Fidelity Monte Carlo

The histogram in Figure 5.12 is determined using general MCM for the equivalent number of high-fidelity samples or High-Level High-Fidelity Monte Carlo (HLHFMC). The estimated drag coefficient $C_d = 4.27$ is larger than the estimated drag coefficient determined using the MLMFMC approach, the same holds for the standard deviation, which is $\sigma_{C_d} = 0.79$ for the HLHFMC. The shape of the histogram is similar to that of the MLMFMC approach with a peak at the beginning.

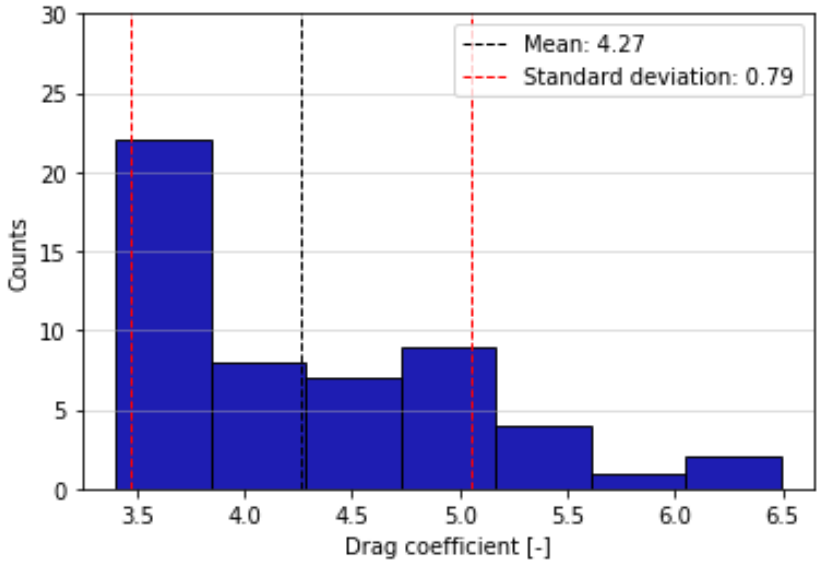


Figure 5.12: Histogram of the drag coefficient as a result of a general Monte Carlo simulation with $N = 53$ samples using a high-resolution high-fidelity model. The estimated drag coefficient is $C_d = 4.27$ with a standard deviation of $\sigma_{C_d} = 0.79$

5.4. True value of estimated drag coefficient

In order to get an idea of what the 'true' estimated drag coefficient is, all samples taken are plotted in Figure 5.13. In green the true value of the drag coefficient according to the 3D case, which is $C_d^{3D} = 3.89$ and in red for the 2D case $C_d^{2D} = 3.64$.

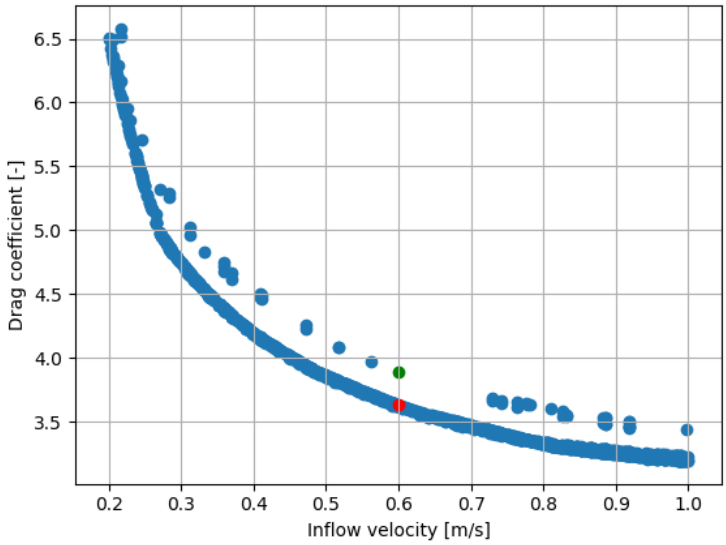


Figure 5.13: Scatter plot of all the estimated drag coefficients.

	MLMC 2D	**HLMC 2D**	**MLMC 3D**	**HLMC 3D**	**MLMFMC**	**HLHFMC**
C_d	3.92	4.28	4.18	4.29	3.89	4.27
σ_{C_d}	0.77	0.77	0.79	0.81	0.73	0.79

Table 5.12: Summary of results of the first case study.

To summarize, the results from both MLMC simulations and the MLMFMC with corresponding high resolution high-fidelity simulations are presented in Table 5.12. Every simulation appears to be overestimating the true value, except for the MLMFMC. This overestimation with respect to the true value is very apparent in all cases of high resolution and/or high-fidelity. The reason for this overestimation is, presumably due to the large range of inflow velocities (range of 0.8) and the corresponding large large of drag coefficients (range of around 4.5). The standard deviation remains relatively constant regardless of simulation method and is not reduced below 0.73. In hindsight, it might have been better to apply a different case setup. Instead of the large range of mean inflow velocities, it would possibly have been better to apply a uniform distribution on a small interval around a point of interest e.g. the first $\bar{U} = 0.2$ m/s or second $\bar{U} = 1.0$ m/s benchmark case. The MLMFMC does estimate the true value correctly, it should be noted though that the relative computational cost of MLMFMC with respect to MLMC in 2D is a factor five higher.

Concluding, the MLMFMC is correctly applied to a simple fluid dynamics case since the estimated value corresponds to the true value. The large irreducible standard deviation indicates that the range of input values used might not be optimal for the application of the MLMFMC and thus smaller ranges could be used in future studies using interval analysis for example. The computational cost reduction is demonstrated and results are compared with a general Monte Carlo simulation at high resolution and high-fidelity. The high resolution high-fidelity simulation overestimates the drag coefficient. It is presumed this overestimation is due to the large range of mean inflow velocities.

6

Results: parametric study

In this chapter the results of the parametric study as described in section 4.2 are described. First the behaviour of Strouhal number by varying porosity and Reynolds number is presented. In the next section the described results are fitted using a number of general fit functions. Lastly, a number of fits are selected to be implemented as low-fidelity models, which are used in the final case study in the next chapter.

6.1. Summarized results of the parametric study

The results of the parametric study are summarized in Figure 6.1 and Figure 6.2. During validation of the porosity model, the step function-like behaviour of the Strouhal versus the porosity was noticed with the step at porosity $\beta = 0.275$. Now, after the parametric study, it appears this behaviour is near constant for the range of Reynolds numbers applied with a Strouhal number of around $St = 0.26$.

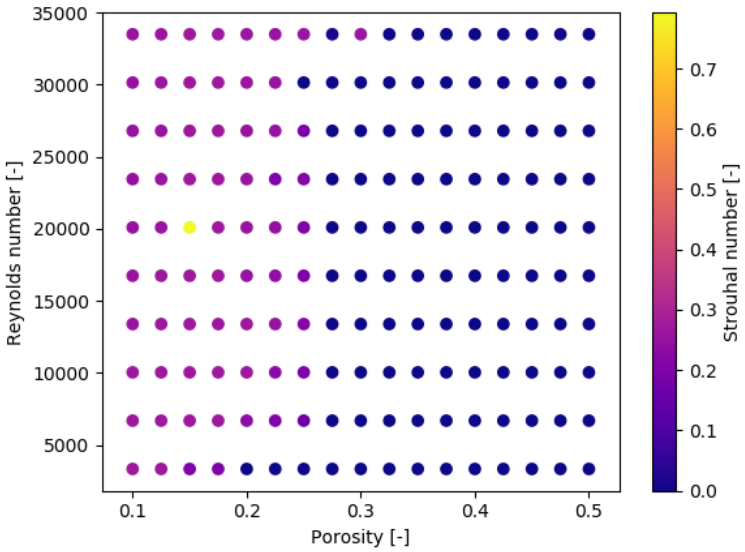


Figure 6.1: Scatter plot of the parametric study. With porosity and Reynolds numbers and a colorbar which denotes the estimated Strouhal number.

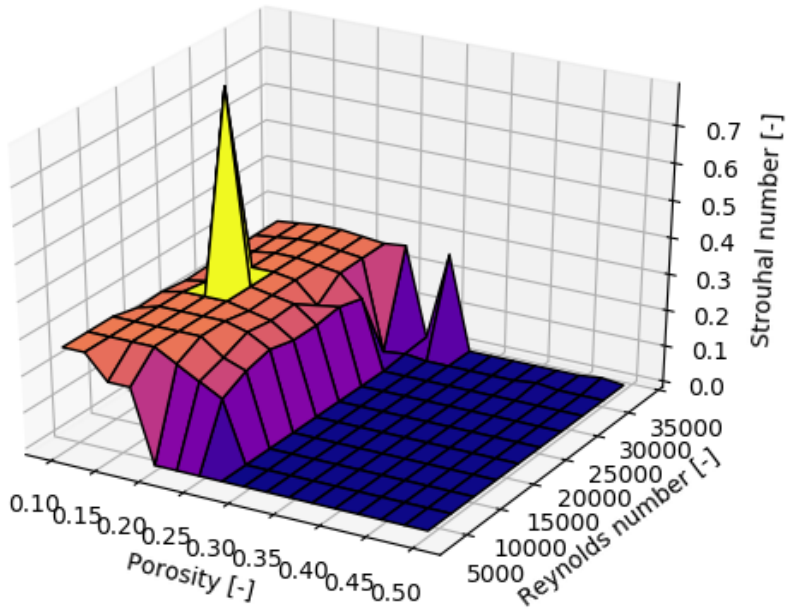


Figure 6.2: Surface plot of the parametric study with Strouhal number versus porosity and Reynolds numbers.

The only exceptions are at the lowest Reynolds number of $Re = 3350$ and the highest two Reynolds numbers at $Re = 30150$ and $Re = 33500$. Also noteworthy, there is a sudden jump at $Re = 33500$ for porosity $\beta = 0.3$. At porosity $\beta = 0.15$ and Reynolds number $Re = 20100$ there is a sharp increase in Strouhal number to $St = 0.79$. It is presumed this estimated value is incorrect and a closer inspection of this case is conducted. The outlier cases that are inspected are depicted in Figure 6.3 using green circles.

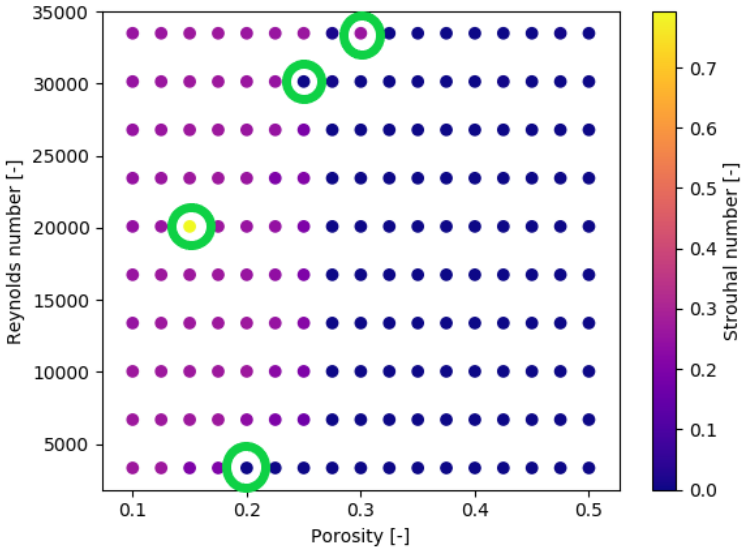
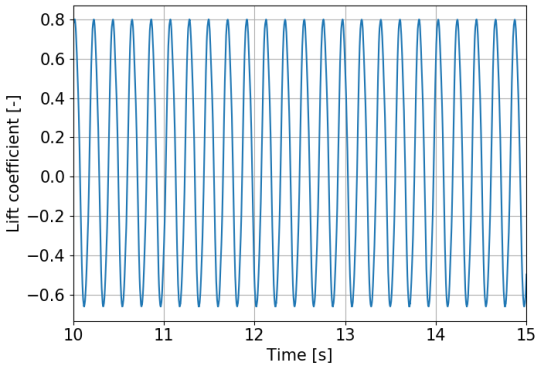


Figure 6.3: Scatter plot of the parametric study with green circles indicating the outlier cases, which are inspected.

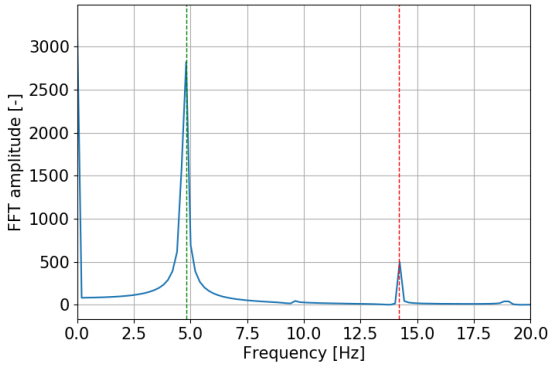
6.1.1. Outlier case with $\beta = 0.15$ & $Re = 20100$

In Figure 6.4 the time signal and corresponding Fast Fourier Transform (FFT) are depicted. By simply counting the peaks in the time signal, the vortex shedding frequency should be around $f_{vs} = 5$ Hz. According to the plot of the FFT the vortex shedding frequency is $f_{vs} = 4.81$ Hz (green dashed line). Unfortunately it appears

that the script used to automatically detect the peak corresponding to the vortex shedding frequency failed and instead it selected the second largest peak, which is denoted by the dashed red line. To confirm this: the frequency at the peak at the red dashed line is 14.21 Hz. Converting this frequency to the Strouhal number for this case means: $St = 14.21 \cdot 0.0335/0.6 = 0.79$, which means that this peak was indeed marked as the peak corresponding to the vortex shedding frequency. Fortunately, this is an outiler case being 1 out of 170 total cases where the peak detection failed. However, this is indication that a more sophisticated method of peak detection should be used and that a human factor should also be considered when interpreting results. Lastly, to determine the correct Strouhal number, the following calculation suffices: $St = 4.81 \cdot 0.0335/0.6 = 0.27$. This Strouhal number is expected when taking into account the values of the Strouhal number for surrounding cases as depicted in both Figure 6.1 and Figure 6.2.



(a) Time signal of the lift coefficient.



(b) Fast Fourier Transform of the time signal with the green dashed line at frequency of 4.81 Hz and the red line at frequency 14.22 Hz.

Figure 6.4: Time signal and corresponding Fast Fourier Transform of the first outlier case $\beta = 0.15$ & $Re = 20100$.

6.1.2. Outlier case with $\beta = 0.30$ & $Re = 33500$

For this case it is expected, based on the neighbouring cases that the vortex shedding frequency $f_{vs} = 0$. However, according to the results the Strouhal number is $St = 0.26$, which means a vortex shedding frequency $f_{vs} = 7.81$ Hz is detected. The order of magnitude of this result corresponds to that of the other results and it is therefore of great interest to inspect this result. In Figure 6.5 the flow velocity field at experimental time $t =$ 12.5 s is depicted. There is no sign of vortex shedding, further confirming the case that this is an outlier case. In Figure 6.6 the time signal of the lift coefficient is depicted. The lift coefficient should be averaged around zero for cases where vortex shedding occurs. For this case the lift coefficient is averaged close to zero, at $C_l = 0.00202$, but the irregularity of the time signal indicates that this is not a case with vortex shedding. Lastly, the FFT is depicted in Figure 6.7 with the red dashed line corresponding to the detected frequency peak. It should be noted that the frequency signal plot is zoomed in and the FFT amplitude axis is limited from 0 to 1 in order to display the detected peak. Based on the aforementioned observations and the frequency signal, it is concluded that this outlier case is due to the filter section of the vortex shedding frequency algorithm failing.

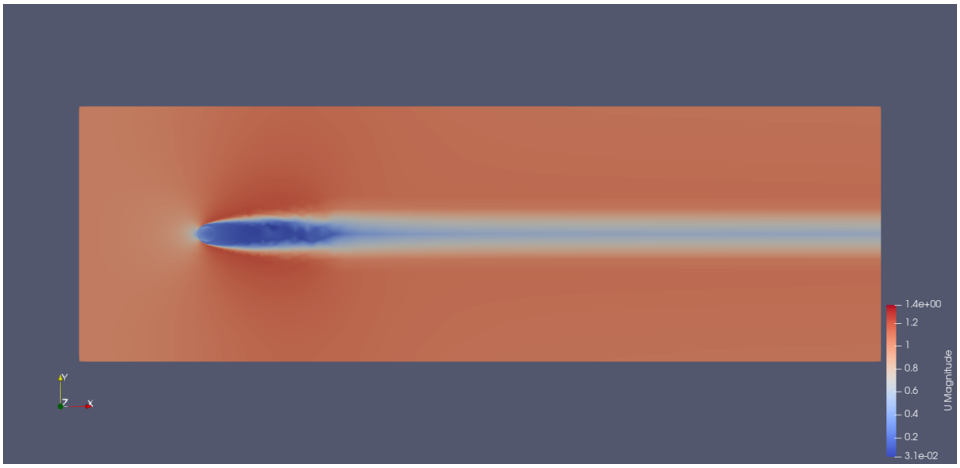


Figure 6.5: Flow pattern of the second outlier case ($\beta = 0.30$ & $Re = 33500$) at experimental time $t = 12.5$ s.

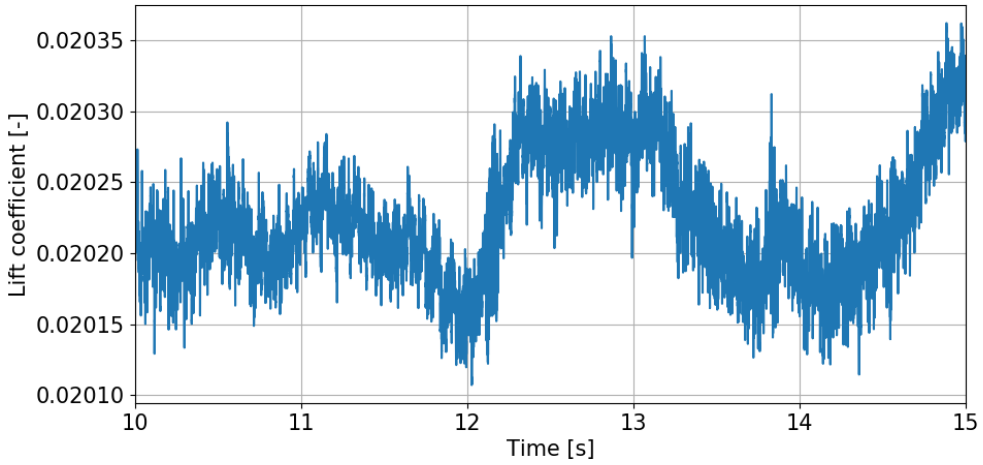


Figure 6.6: Time signal of the lift coefficient for the second outlier case ($\beta = 0.30$ & $Re = 33500$).

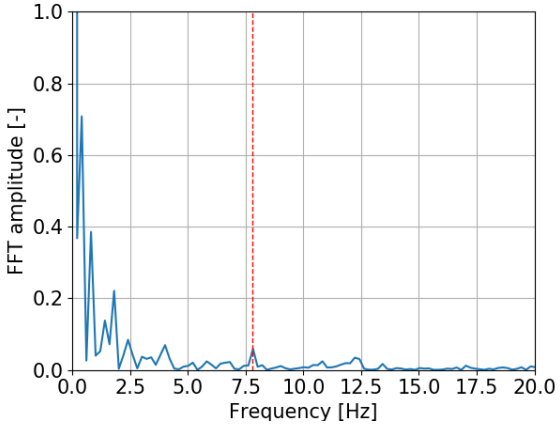


Figure 6.7: Frequency signal of the lift coefficient for the second outlier case ($\beta = 0.30$ & $Re = 33500$) with the red dashed line corresponding to the detected frequency peak.

6.1.3. Outlier case with $\beta = 0.25$ & $Re = 30150$

According to the results the Strouhal number is $St = 0$ for this case, meaning the vortex shedding frequency is $f_{vs} = 0$ Hz. Yet, the surrounding cases indicate that this is a possible false negative result. The flow pattern is depicted in Figure 6.8 and it shows vortex shedding is present, indicating that this case is a false negative. For further inspection, both time and frequency signals are depicted in Figure 6.9 and Figure 6.10 respectively. According to a manual reading of the frequency signal the vortex shedding frequency $f_{vs} = 5.61$ Hz, which corresponds to a Strouhal number of $St = 0.21$.

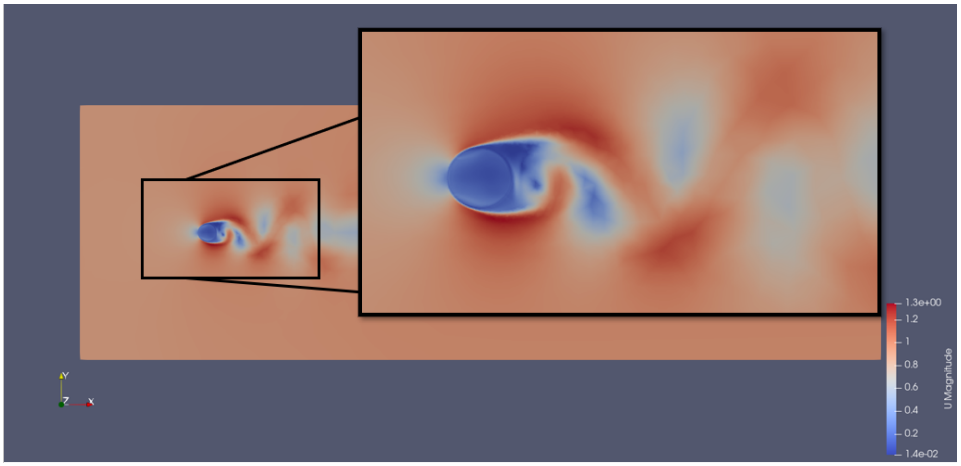


Figure 6.8: Flow pattern of the third outlier case ($\beta = 0.25$ & $Re = 30150$) at experimental time $t = 12.5$ s.

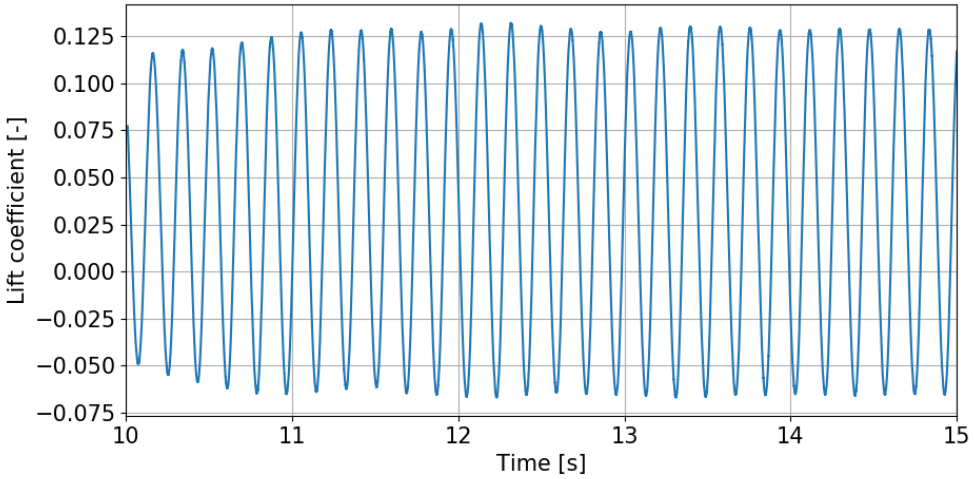


Figure 6.9: Time signal of the lift coefficient for the third outlier case ($\beta = 0.25$ & $Re = 30150$).

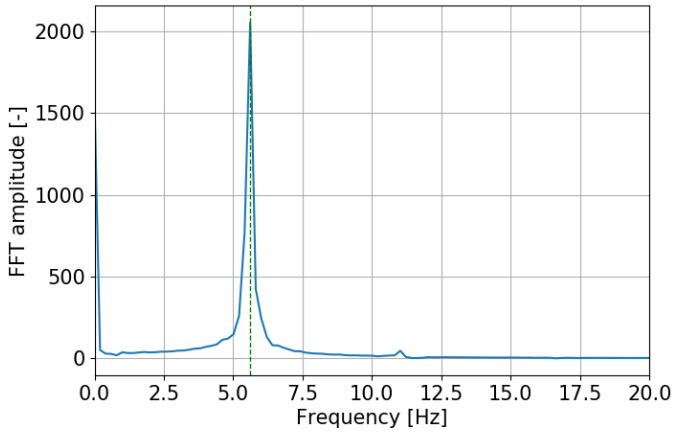


Figure 6.10: Frequency signal of the lift coefficient for the third outlier case ($\beta = 0.25$ & $Re = 30150$) with the red dashed line corresponding to the detected frequency peak.

6.1.4. Outlier case with $\beta = 0.20$ & $Re = 3350$

For this case the Strouhal number is determined to be $St = 0$, which is indeed possible, however the drop to zero frequency is at a lower porosity compared to all other cases. Hence it is an interesting case to inspect. The flow pattern is presented in Figure 6.11 and there is vortex shedding, since several vortices are highlighted on the image. Therefore there is an indication that this case is a false negative, the time signal and frequency signal are presented in Figure 6.12 and Figure 6.13 respectively. The increasing amplitude of the time signal in Figure 6.12 could indicate that the flow is not fully developed yet during the experimental time range of $t = 10$ to $t = 15$ s. Moreover, the lift coefficient is not averaged around zero, as is expected. The vortex shedding frequency detection algorithm fails on these points, as it assumes both fully developed flow and a lift coefficient averaged around zero. The lift coefficient could be higher in this case due to the increased distance at which the vortices shed off. For this case, it appears that the vortices shed at a distance of 3 times the cylinder diameter. Furthermore, an inspection is conducted on higher porosity cases to find the porosity at which vortex shedding stops, assuming vortex shedding does stop at a certain porosity.

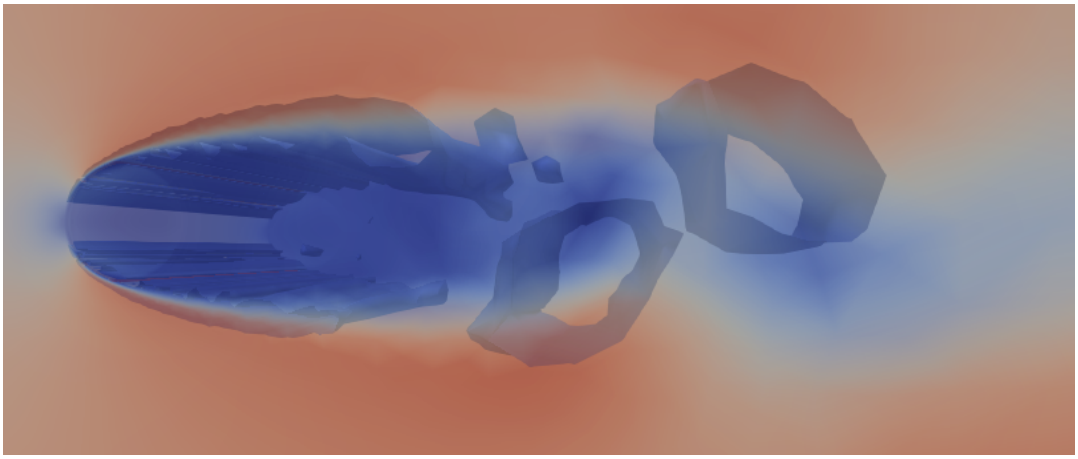


Figure 6.11: Flow pattern of the fourth outlier case ($\beta = 0.20$ & $Re = 3350$) at experimental time $t = 12.5$ s.

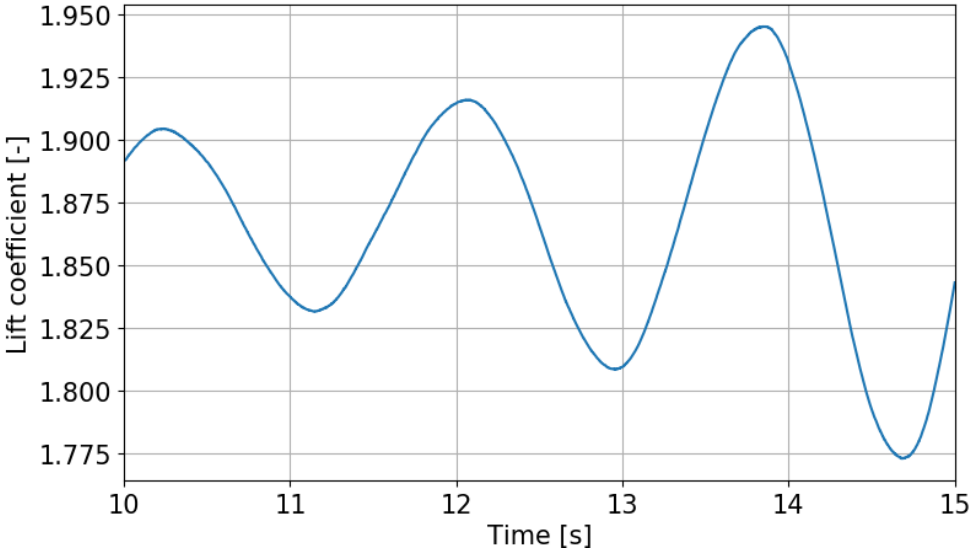


Figure 6.12: Time signal of the lift coefficient for the fourth outlier case.

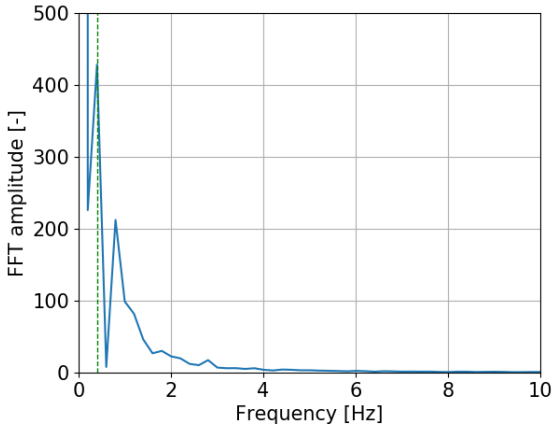
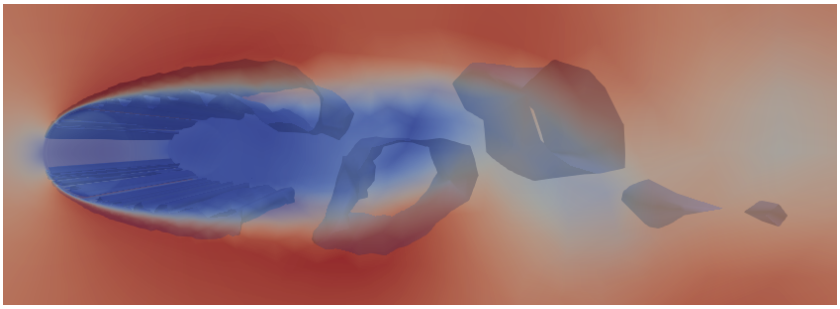


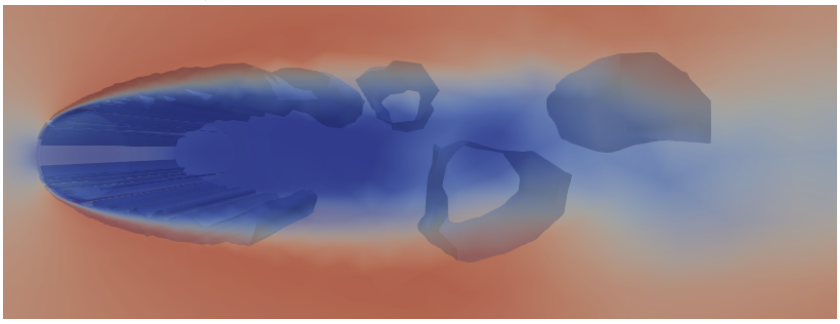
Figure 6.13: Frequency signal of the lift coefficient for the fourth outlier case with the red dashed line corresponding to the detected frequency peak.

Determining the Strouhal number at $Re = 3350$

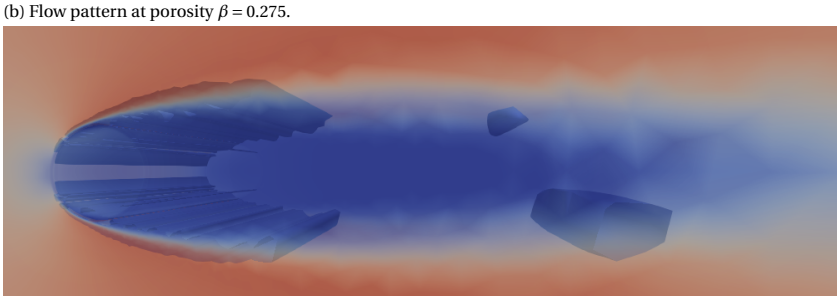
In Figure 6.14 several flow patterns are depicted for varying porosity levels $\beta = 0.225, 0.275$ and 0.325 at $Re = 3350$. In all porosity cases up to $\beta = 0.325$ vortex shedding is detected, which gives the suggestion that the vortex shedding frequency detection algorithm needs to be adjusted. Moreover, it appears that for increasing porosity the distance -at which vortex shedding occurs with respect to the cylinder- increases. In total the vortex shedding frequency is estimated for six cases, using the frequency signals by applying the FFT and by visual estimation of the vortex shedding frequency using the video of the flow pattern in *Paraview*.



(a) Flow pattern at porosity $\beta = 0.225$.



(b) Flow pattern at porosity $\beta = 0.275$.



(c) Flow pattern at porosity $\beta = 0.325$.

Figure 6.14: Flow patterns with highlighted vortices at experimental time $t = 12.5$ s of varying porosity β at Reynolds number $Re = 3350$.

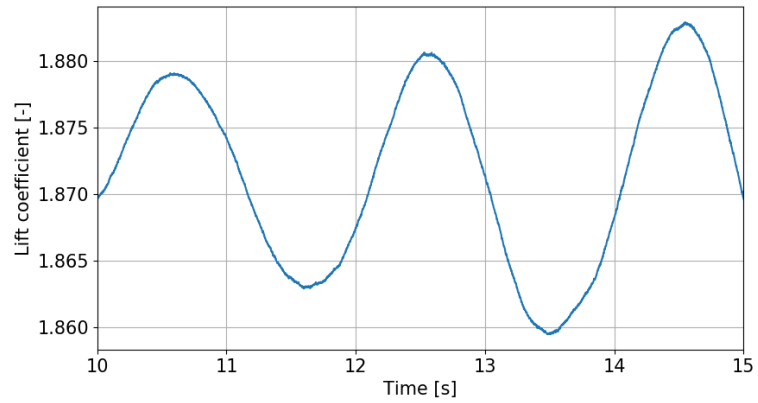
The interesting part is that the vortex shedding does occur and thus the Strouhal number should not be zero, since the vortices are seen in Figure 6.11. However, at this point there is a crossroads in the definition of the Strouhal number. Officially the Strouhal number is a dimensionless number used to describe oscillating flow mechanisms, which -initially- was vortex shedding. Yet in this implementation and for the case of a non-porous circular cylinder the vortex shedding frequency is of interest due to the possible occurrence of resonance when the vortex shedding frequency is close to the natural frequency of the cylinder. Therefore, vortex shedding frequency and frequency of lift force oscillation were the though of as the same and hence the Strouhal number was used for the frequency of lift force oscillation in non-dimensional form. It appears that the introduction of porosity to the problem of flow around a circular cylinder causes a divide to the formulation of the Strouhal number as simply a dimensionless form of the vortex shedding frequency. Thus, a formal definition is proposed:

From here on out,

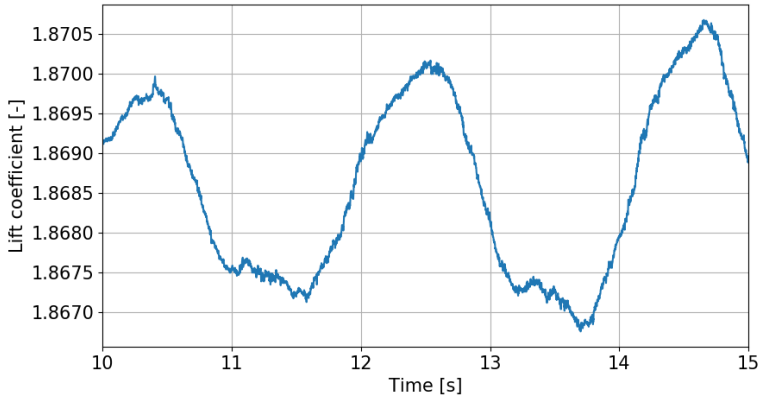
1. **the Strouhal number St refers to the vortex shedding frequency; and**
2. **the cylinderical Strouhal number St_{cyl} refers to the frequency with which the lift force on the cylinder oscillates.**

In an attempt to explain the divide between the two different Strouhal numbers, the time signals of the lift coefficients corresponding the to flow patterns from Figure 6.14 are depicted in Figure 6.15. Just like in Figure 6.12 the lift coefficient does not oscillate around $C_l = 0$ anymore, but instead oscillates around $C_l \approx 1.86$. The same is true for the cases depicted in Figure 6.15. The amplitudes of the time signals increase

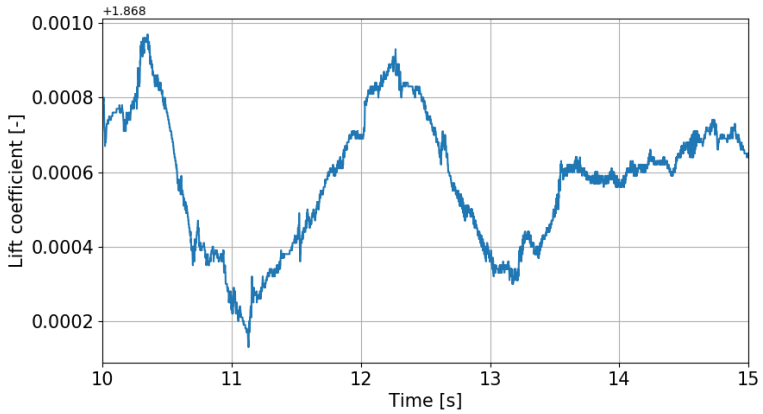
for the case of $\beta = 0.225$ and $\beta = 0.275$, which could indicate that the flow is not fully developed yet and a longer experimental time should be used for simulation. The time signal becomes less smooth for increasing porosity with clear differences between the top and middle plot in Figure 6.15. Also the amplitude of the signal is decreasing for increasing porosity. For the case of $\beta = 0.325$ the trend of increase in amplitude over time is not present, moreover the oscillations are becoming nearly constant, since the maximum lift coefficient from peak to valley is $C_l \approx 0.0008$ compared to a peak to valley lift coefficient at $\beta = 0.275$ of $C_l \approx 0.0035$.



(a) Flow pattern at porosity $\beta = 0.225$.



(b) Flow pattern at porosity $\beta = 0.275$.



(c) Flow pattern at porosity $\beta = 0.325$. Be mindful of the term +1.868 added to the ticks on the vertical axis.

Figure 6.15: Time signals of lift coefficient of varying porosity β at Reynolds number $Re = 3350$. **Important note:** the bottom image vertical axis has an added value of +1.868 for the lift coefficient, yet the ticks are presented in this form due to the small differences of the lift coefficient for this case.

A possible cause for the divide between Strouhal number and cylindrical Strouhal number could be found in the increasing length of the wake that is present behind the cylinder. Comparing Figure 6.8 with Figure 6.11 presents a clear difference in length of the wake behind the cylinder. In Figure 6.8 the vortex shedding occurs nearly directly behind the cylinder, whereas in Figure 6.11 the vortices do not appear until a length of at least 3, possibly 4 times the cylinder diameter is reached.

The idea is that due to the increased porosity the lift force at the top (and bottom) of the cylinder do not 'feel' the effects of vortex shedding anymore, because the vortex shedding occurs some distance away from the cylinder. Therefore, there could be some relation between the porosity and the distance between the back of the cylinder and the point at which vortex shedding occurs. Thus as a result of the parametric study a hypothesis is presented:

For the case of water flow past a porous circular cylinder: increasing porosity results in an increase in the distance between the back of the cylinder and the point at which vortex shedding occurs.

In order to properly support this hypothesis an in-depth study should be conducted on all results of the parametric study. However due to time constraints and the goal of the thesis this is not done. A quick review on literature is conducted in order to find an explanation of this behaviour and two papers were found that support the hypothesis, both of which are conducted by *Sun et al.* [49] [50]. To summarize, the results from [49] are presented in Figure 6.16 and the results from [50] are presented in Figure 6.17.

In Figure 6.16 the flow patterns of the first study are depicted. The top case is for a solid cylinder (a and f), where vortex shedding occurs at the back of the cylinder. In the images on the right side a length scale is depicted denoted by x^*, which is a length scale normalized using the cylinder diameter ($D = 0.0335$ m for the parametric study). The length scale is then derived as $x^* = \frac{x}{D}$ and corresponds to the proposed distance between the back of the cylinder and the point at which vortex shedding occurs. The results for larger porosity cylinders are presented for porosity $\beta = 0.37$ b and g, for porosity $\beta = 0.48$ c and h, for porosity $\beta = 0.61$ d and i and for porosity $\beta = 0.67$ e and j. The range of porosity values applied here is larger than the one inspected in the parametric study. But, the proposed hypothesis could be confirmed using this experimental data. Lastly, the length scale $x^* \approx 5$ for porosity $\beta = 0.37$ and $x^* \approx 2$ for the solid cylinder according to [49]. For the inspected cases of the parametric study -which are at porosity values in between those two mentioned values- this length scale is also between $2 < x^* < 5$.

A follow-up study [50] was conducted in which three conceptual models for vortex formation and decay processes in the wake are proposed. In Figure 6.17 these three models are presented. The first model is for a solid circular cylinder and depicts vortex shedding at the back of the cylinder at length $x^* \approx 2$. The second model is for porosity of $\beta = 0.48$ depicts the vortex shedding at length $x^* \approx 7$ and the third model is for porosity $\beta = 0.61$ with a vortex shedding length of $x^* \approx 18$

As a final note on the analysis on outlier cases of the parametric study, there are clear indications that support the experimental results of *Sun et al.* [49] [50]. Hence, it is strongly recommended to further pursue analysis of the results of this parametric study and to possibly set up the experiments of *Sun et al.* in a virtual environment to validate the results using CFD. Future improvements on the current setup of the parametric study are mostly allocated to the mesh and porosity model, as enumerated below.

1. For the mesh: apply a circular structured mesh instead of an unstructured mesh in order to properly display the propagation of vortices through the mesh, especially further away from the cylinder.

2. For the porosity model: setup a case in which the drag coefficient is properly calculated; and

3. use this setup to also study the effect of porosity on the drag coefficient, as well as on the development of wake vortices.

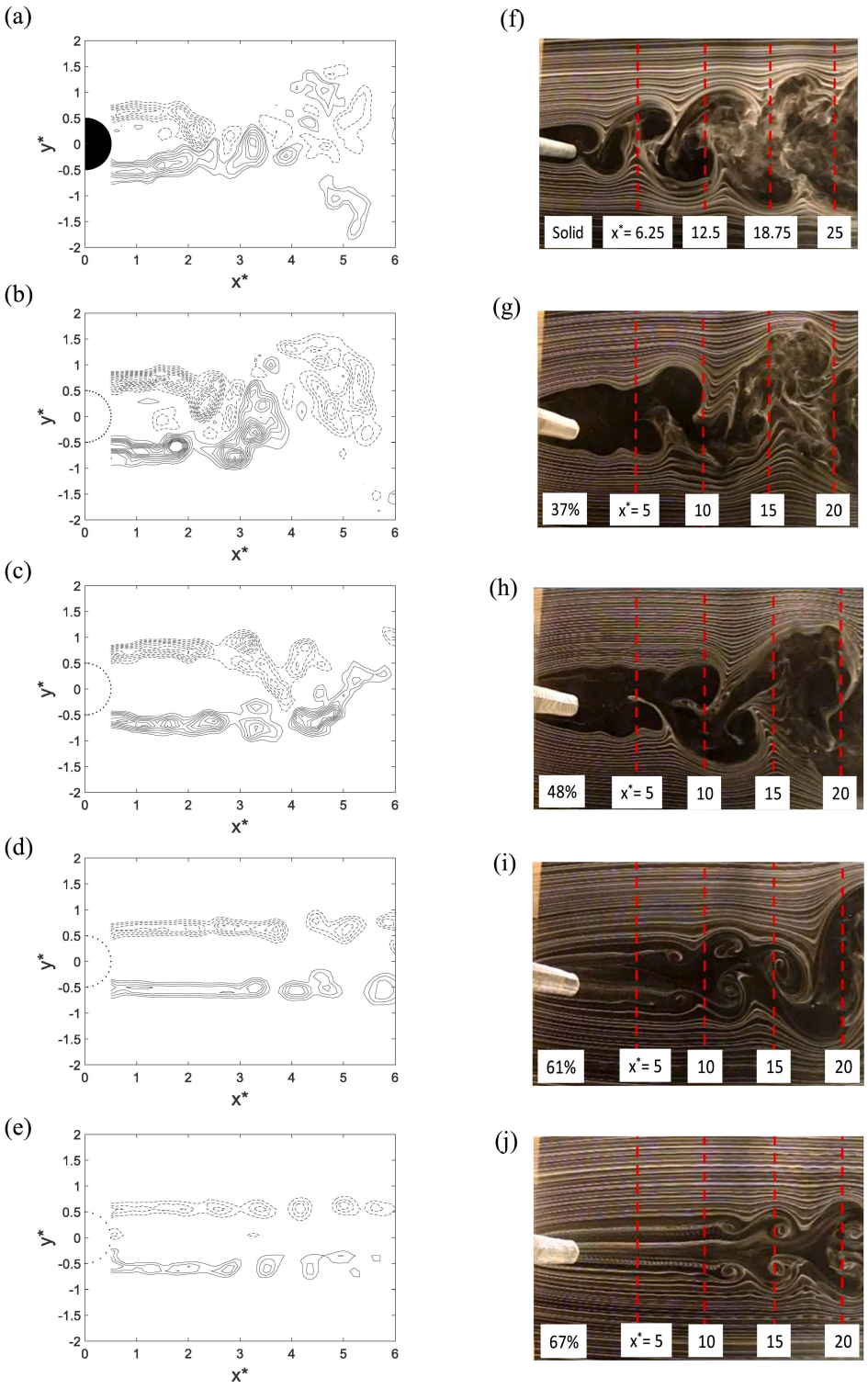


Figure 6.16: Results from a study by *Sun et al.* [49] on wake flow structures of screen cylinders using PIV.

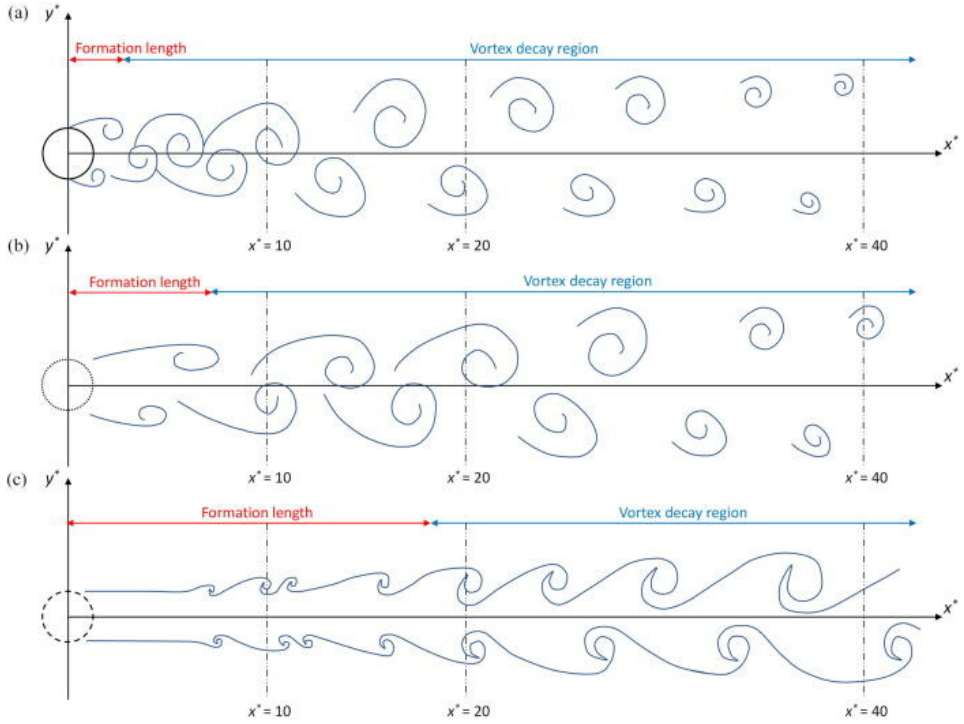


Figure 6.17: Results from a study by *Sun et al.* [50] on the development of wake flow structures behind screen cylinders.

6.1.5. Interpretation of the physics of flow past a porous cylinder

As stated before, the lift force as experienced on the cylinder decreases for increasing porosity. As depicted in Figure 6.15 the lift coefficient is not averaged around zero anymore at a porosity of 22.5% for the lowest Reynolds number. The way the lift force decreases or the shape of a curve of lift force versus porosity is not analysed in this study. In the results of the parametric study the frequency cuts off to zero at a certain porosity, which is expected since this is the point at which the cylinder does not experience the lift force anymore due to the increased wake formation length. The reason for the increased wake formation length is described as flow bleeding, causing a delay in the interaction between the shear layers. The vortices persist for a longer distance due to the weak interaction across the center line for a porous cylinder. The way the vortex street is developed initially is the same as that of a solid cylinder [50].

6.1.6. Remarks on vortex shedding frequency detection algorithm

The vortex shedding frequency detection algorithm is constructed in a simple and low level manner. Based on the results of the parametric study the algorithm performs fine for cases of high porosity $\beta \geq 0.4$ and for cases of low porosity $\beta \leq 0.2$. The cases for porosity $0.2 < \beta < 0.4$ the algorithm produced one confirmed false positive (subsection 6.1.2), one confirmed false negative (subsection 6.1.3) and one true positive, but with incorrect frequency value (subsection 6.1.1). Resulting in 3 out of 170 cases failure rate. Ideally, the following improvements or implementations should be added to this algorithm in order to minimize the need for human intervention.

1. Time signal analysis: determine when the flow is developed and detect oscillations;
2. Frequency signal filtering: more sophisticated method to determine which frequency peak corresponds to the vortex shedding frequency, instead of using a percentage of the drag coefficient as filter criteria.

6.1.7. Summarized results after inspection of outlier cases

In Table 6.1 the changes made to the data of the parametric study due to the determined outlier cases are summarized. A total of 3 out of 170 cases were changed. In the fourth and fifth column the re-determined frequency of lift force oscillation and cylindrical Strouhal number are presented.

Outlier case	**Reynolds number** Re **[-]**	**Porosity** β **[-]**	**Frequency** f_{lf} **[Hz]**	**Cylindrical Strouhal number** St_{cyl} **[-]**
1	20100	0.15	4.81	0.27
2	33500	0.30	0	0
3	30150	0.25	5.61	0.21

Table 6.1: Summary of changes to the results of the parametric study due to outlier cases.

After changing the results of these 3 outlier cases, the data is presented in a surface and scatter plot depicted in Figure 6.18 and Figure 6.19 respectively.

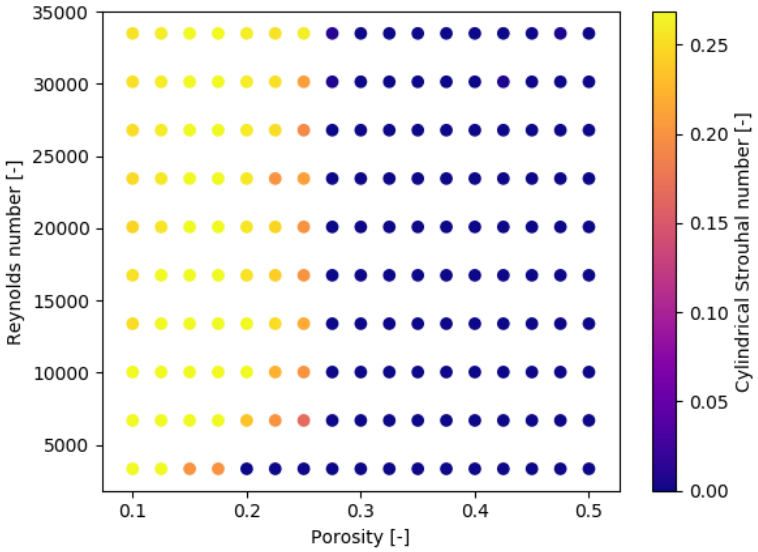


Figure 6.18: Scatter plot of the parametric study. With porosity and Reynolds numbers and a colorbar which denotes the estimated Strouhal number.

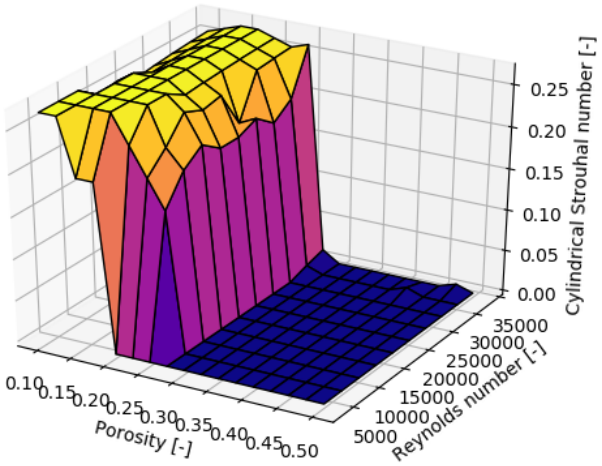


Figure 6.19: Surface plot of the parametric study with Strouhal number versus porosity and Reynolds numbers.

It appears that at low porosity $\beta = 0.1$ to $\beta = 0.125$ and for Reynolds numbers $Re > 10000$ the cylindrical Strouhal number increases very slightly from $St_{cyl} = 0.22$ to $St_{cyl} = 0.26$ and remains constant to around porosity $\beta = 0.2$, after which a decrease in cylindrical Strouhal number is detected with the steepest decreasing slope between porosity $\beta = 0.25$ and $\beta = 0.275$. For the case of Reynolds number $Re = 3350$ the cylindrical Strouhal number is already at a maximum for porosity $beta = 0.1$ and the steep decreasing slope occurs between porosity $\beta = 0.175$ and $\beta = 0.2$. Presumably, at low Reynolds numbers the behaviour as depicted in Figure 6.19 occurs at lower porosity. However, this is just speculation and a thorough investigation should be conducted for low Reynods number and low porosity for a proper description of the behaviour.

6.2. Data-fitting

In this section, the data from the parametric study as described in subsection 6.1.7 are data-fitted in order to find possible empirical relationships between the cylindrical Strouhal number, Reynolds number and/or porosity. A full list of functions to which the data is fitted, is presented in Appendix A.

According to theory the Strouhal number remains constant at $St \approx 0.21$ for the subcritical Reynolds numbers regime. The same behaviour is noticed in Figure 6.19 and Figure 6.18. Although now, the interest lies on the cylindrical Strouhal number, which is equal to the Strouhal number at porosity $\beta = 0$. The cylindrical Strouhal number remains about constant per value of porosity with a maximum oscillation of $St_{cyl} = 0.04$ before the cut-off porosity is reached. The cut-off porosity of $\beta_{co} = 0.175$ is reached fastest for lowest Reynods number $Re = 3350$. For the other Reynolds numbers, this cut-off porosity is $\beta_{co} = 0.25$. The results of the parametric study are data-fitted using 2D curve fits, for the case of cylindrical Strouhal number versus porosity the fit function should have behaviour similar to that of the step function. The fit functions with similar behaviour are:

- Gompertz function: $y(x,a,b,c) = a\exp(-\exp(-b(x-c)))$
- Hill function: $y(x,a,b,c) = \frac{ax^c}{b^c+x^c}$
- Weibull function: $y(x,a,b,c,d) = a-(a-b)\exp(-c|x|^d)$

Other functions were fitted as well, but by visual evaluation these were deemed unfit due to the following arguments. These fits would create non-physical results (e.g. cubic function results in a negative cylindrical Strouhal number for porosity $\beta > 0.3$), the data could not be fitted to the fit function or the functions provided very poor goodness of fit. It should be noted that no goodness of fit is calculated, but only a visual evaluation is conducted since most of the fit functions implemented here are non-linear. Lastly, it is possible to fit a 9^{th} order polynomial to the data, but the resulting fit will most likely not be relevant as a low-fidelity model due to overfitting. In the next subsections the results of the three fit functions are presented using 3D scatter plots. All 2D scatter plots including individual fit functions are presented in Appendix B.

6.2.1. Gompertz function

In Figure 6.20 the scatter plot from the parametric study is curve fitted for cylindrical Strouhal number versus porosity at each value of Reynolds number. None of the curve fits failed and no non-physical behaviour is detected. The Gompertz function has the steepest slope at Reynolds number $Re = 33500$ and the slopes at Reynolds numbers $Re = 3350$ and $Re = 6700$ are the least steep. The other seven functions have almost identical shapes with little differences.

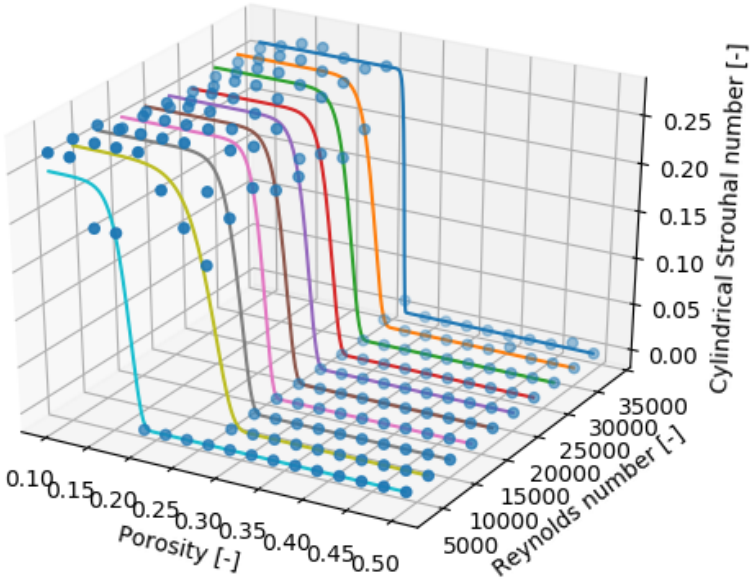


Figure 6.20: Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using a Gompertz fit function.

6.2.2. Hill function

In Figure 6.21 the scatter plot is curve fitted using the Hill function, the same behaviour as described for the Gompertz function is noticed. The main difference between the Gompertz function and the Hill function is the shape of the curve at the beginning of the steep slope decrease. The slope of the Hill function is consistently steeper, than that of the Gompertz function.

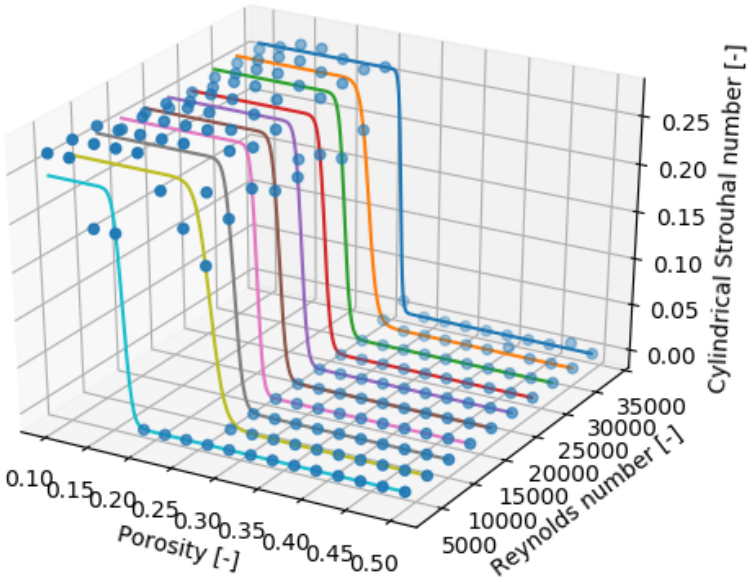


Figure 6.21: Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using a Hill fit function.

6.2.3. Weibull function

In Figure 6.22 the Weibull function is fitted to the results of the parametric study. The shapes of the curve fits correspond well with those of the Gompertz function. However, as depicted, the curve fit failed at two Reynolds numbers, namely at $Re = 13400$ and at $Re = 33500$.

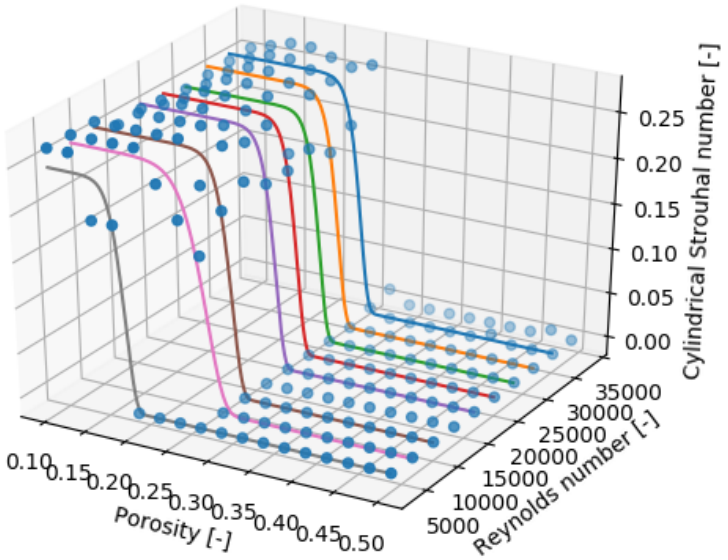


Figure 6.22: Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using a Hill fit function.

6.3. Low-fidelity models

This section is dedicated to the transformation from a number of individual fit functions to a number of low-fidelity models. The approach conducted here is simple non-weighted averaging of the fit parameters of the ten individual fit functions, followed by a visual analysis of goodness of fit.

The Weibull fit function is not regarded as a suitable candidate for a low-fidelity model for several reasons. The results of the Weibull fit function are incredibly similar to those of the Gompertz fit function, for two out of ten cases the fit function failed and the Weibull fit function requires more fit coefficients, than the Gompertz fit function. Therefore only the Gompertz and Hill fit functions are considered for low-fidelity models. In the next subsection the two proposed low-fidelity models are presented and 3D scatter plots with the low-fidelity models are depicted. The corresponding fit coefficients are stated and a visual goodness of fit is described. In appendix Appendix B the 2D scatter plots with low-fidelity models are presented for each Reynolds number case.

Gompertz model

In Figure 6.23 the Gompertz model is plotted in the 3D scatter plot of the parametric study with the following fit coefficients: $a = 0.26$, $b = -155.39$ and $c = 0.26$. For most cases the Gompertz model corresponds well with the results in the scatter plot, again the goodness of fit is worse for the cases of Reynolds numbers $Re = 3350$, $Re = 6700$ and $Re = 33500$.

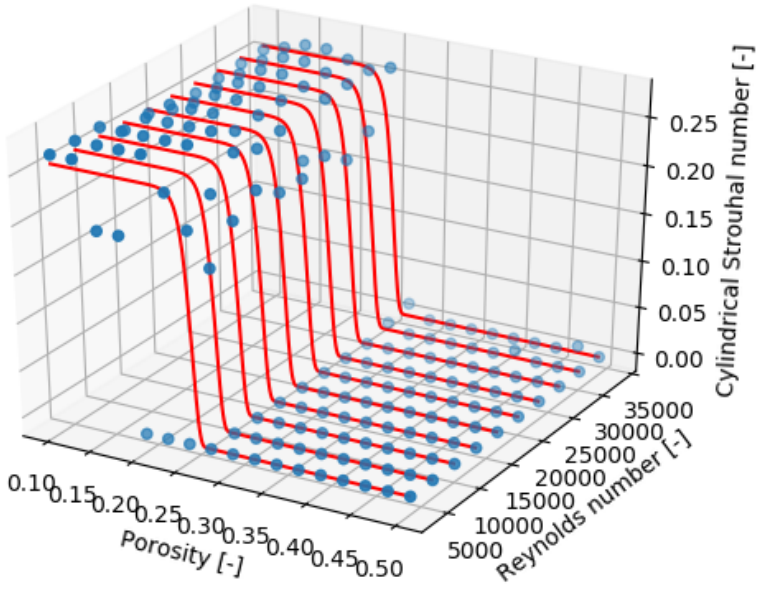


Figure 6.23: Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using the calculated Gompertz model with $a = 0.26$, $b = -155.39$ and $c = 0.26$.

Hill model

In Figure 6.24 the Hill model is depicted in the 3D scatter plot of the parametric study with the following fit coefficients: $a = 0.26$, $b = 0.25$ and $c = -76.27$. The differences between the Gompertz and Hill model are -visually- almost negligible. The only visual difference is the shape of the curve at porosity $\beta = 0.2$.

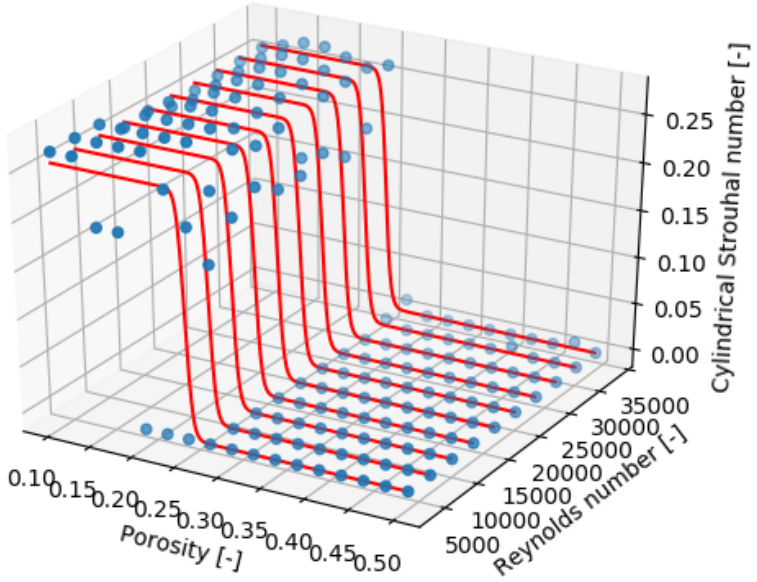


Figure 6.24: Scatter plot of the results of the parametric study with a curve fit of cylindrical Strouhal number versus porosity for each Reynolds number using the calculated Hill model with $a = 0.26$, $b = 0.25$ and $c = -76.27$. .

As a short evaluation on the fit coefficients, it is assumed these are constant over the Reynolds numbers. This is not correct as can be seen in the results, but the differences are small and for a low-fidelity model this is deemed sufficient and will be evaluated in the next case study. For further research, the data could be fitted using surface fitting. Another possibility is to assume a dependency of the fit coefficients on the Reynolds number, since the cut-off in frequency is more influenced by porosity than the Reynolds number.

7

Results: interval analysis using MLMFMC

The results of the final experiment are described in this chapter. In section 7.1 the computational costs used for the low-fidelity models and high-fidelity model are summarized. Next the results of the parametric study with interval analysis are presented, followed by an analysis of the correlation between high- and low-fidelity models. The case setup is different from that of chapter 6 in order to assess validity of the low-fidelity models in a general sense without overfitting. Lastly, a short description is added on the determination of bounds of epistemic uncertainty of the high-fidelity model.

7.1. Relative computational cost

In contrast to the relative computational cost estimation in section 5.1, the computational cost here is not derived on time averaging after running a number of cases. This decision was made for the reason that these cases are much more computationally expensive. Therefore it was deemed a better investment of time and computational resources to estimate the computational cost using an interpolation of the number of cores used and the number of cells in the mesh. For the case of 3D model fidelity at the lowest resolution level an average run time of $t = 60$ s on 48 cores is estimated based on the results of chapter 6. The low-fidelity model based on fits or simple equations are estimated to take an average run time of $t = 0.0005$ s. This corresponds to a relative computational cost of $C^{LF} = 1$. The relative computational costs are summarized in Table 7.1.

	low resolution [-]	middle resolution [-]	high resolution [-]
LF (equation)	1	-	-
HF (3D)	$64.800 \cdot 10^6$	$174.048 \cdot 10^6$	$348.288 \cdot 10^6$

Table 7.1: Total estimated relative computational cost of the low-fidelity model (equation) and the high fidelity (3D) model for Multi-Level Multi-Fidelity Monte Carlo.

7.2. Data points from the interval analysis

First the results from the interval analysis using MLMFMC are presented. The results are obtained using MLMFMC with varying low-fidelity models, namely the Gompertz and Hill model as described in section 6.3. The estimated cylindrical Strouhal numbers with the corresponding error bars are plotted in Figure 7.1 and Figure 7.2 for Gompertz and Hill respectively. On first sight the results look equal, which is almost true. The differences in the estimated value and standard deviation are in the third and fourth decimal respectively. Therefore, the following analysis applies to both plots.

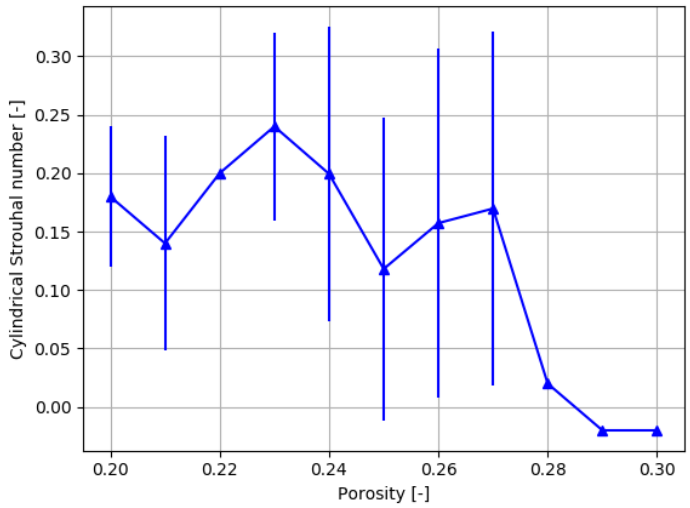


Figure 7.1: Estimated cylindrical Strouhal number versus porosity using the Gompertz model as the low-fidelity model.

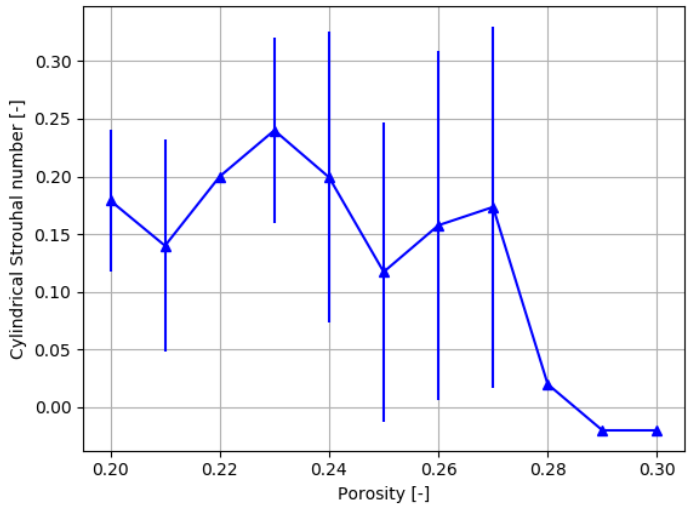
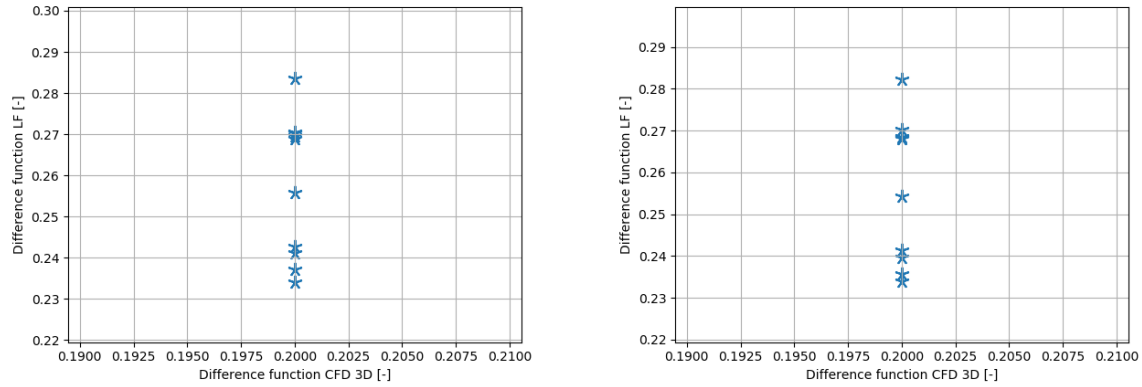


Figure 7.2: Estimated cylindrical Strouhal number versus porosity using the Hill model as the low-fidelity model.

The points that immediately stand out are the points for which there is no error bar present. These points are located at $\beta = 0.22$, $\beta = 0.28$, $\beta = 0.29$ and $\beta = 0.30$. According to the MLMFMC the variance is zero for these cases, hence no error bars are present. For further investigation, the correlation between model fidelities is plotted in Figure 7.3. The high- and low-fidelity models are clearly uncorrelated for both low-fidelity models. Furthermore, it is noted that the difference of the high-fidelity model appears to be constant at $Y^{HF} = 0.2$. The same behaviour is detected for the other porosity values, the figures for these are depicted in Appendix C. Since the values of the difference function for the high-fidelity model are so closely grouped further research should be conducted on the frequency detection algorithm. It is suspected that the frequency resolution is too low to pick up on small differences in frequency, since all samples of the high-fidelity model result in exactly the same cylindrical Strouhal number, which results in a single value of the difference function and thus no variance.

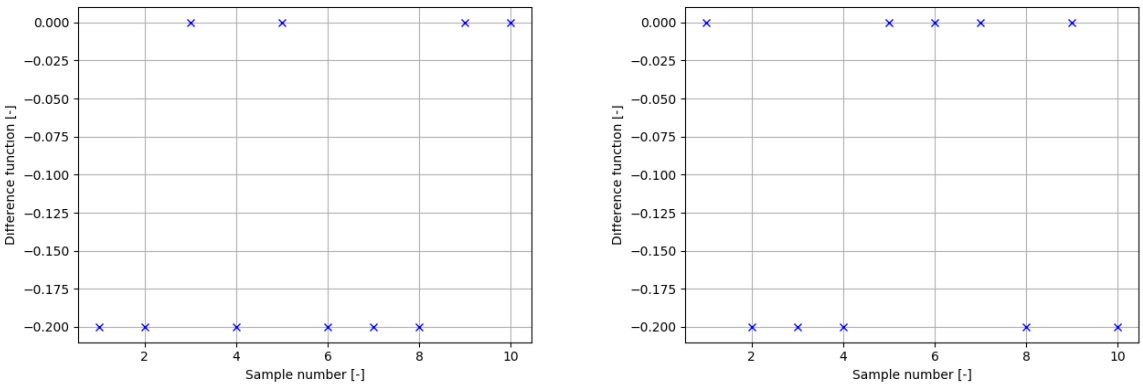


(a) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Gompertz model.

(b) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Hill model.

Figure 7.3: Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.22$.

Another observation for these points without error bars are the points at $\beta = 0.29$ and $\beta = 0.30$. These values have a negative cylindrical Strouhal number. This makes no sense physically, although it could be explained using the difference function. In case the majority of the difference function terms are negative, the estimator constructed on these differences is also negative. To follow up this explanation, the output files are analysed and after inspection this explanation seems to be correct. To clarify, in Figure 7.4a and Figure 7.4b the difference function of the pilot samples for porosity $\beta = 0.30$ are depicted. These plots show that the difference function is either 0 or -0.2, which means the total sum of the difference function is negative as well. Moreover, it is apparent here that the high-fidelity model has binary behaviour between 0 and 0.2, since the difference function is either 0 or -0.2. From the results it is noted that the differences between resolutions are either 0.2 - 0.2, 0 - 0 or 0 - 0.2. This further solidifies the idea that the frequency resolution is too low.



(a) Difference function of 10 pilot samples for porosity $\beta = 0.30$ for the difference between the lowest and middle resolution.

(b) Difference function of 10 pilot samples for porosity $\beta = 0.30$ for the difference between the middle and highest resolution.

Figure 7.4: Difference functions of pilot samples at porosity $\beta = 0.30$.

It could very well be for the case of $\beta \geq 0.3$ that the cylindrical Strouhal number truly is $St_{cyl} = 0$. And, that the only reason there is a negative difference function is due to the epistemic uncertainty introduced by the frequency detection algorithm. Although, no definite conclusion on the true value of the cylindrical Strouhal number can be drawn from these results. It is also possible that the cut-off behaviour is so extreme due to the low frequency resolution. The third observation is the large standard deviations of all points with error bars. With the largest standard deviation being $\sigma_{St_{cyl}} = 0.13$ at $\beta = 0.25$. These deviations are presumably so large, since there are large variations of cylindrical Strouhal numbers for the selected range of porosity values. Moreover, due to the binary-like behaviour of the high-fidelity model this variance should remain relatively

large. The final observation is on the shape of the curve. There is no clear explanation or idea as to why it has this shape with two peaks. It could be that the number of samples taken was too little. Another possibility is that in this range of values the systemic error due to the frequency detection algorithm is at a maximum. Although, no conclusion like that can be deduced from an interval analysis.

7.3. Pearson correlation coefficient

The Pearson correlation coefficient is analysed in order to investigate the validity of the low-fidelity Gompertz and Hill models. In Figure 7.5 and Figure 7.6 the Pearson correlation coefficient is plotted versus the porosity for the Gompertz and Hill model respectively. For the Gompertz model no Pearson correlation coefficient is calculated at $\beta = 0.3$ (NaN is encountered). For the Hill model NaN is encountered at $\beta = 0.20$, $\beta = 0.21$, $\beta = 0.22$, $\beta = 0.28$, $\beta = 0.29$ and $\beta = 0.30$. It is unclear why this is though, it could be that a division of zero by zero occurs when the variance of the difference function of either fidelity model is zero. However, due to the black box usage of *Dakota*, it is hard to pinpoint exactly why a NaN occurs. For some cases it can be presumed that, since the variance of the high-fidelity model is zero for the cases with no error bars in Figure 7.1 and Figure 7.2, that this means that at porosity values $\beta = 0.22$, $\beta = 0.28$, $\beta = 0.29$ and $\beta = 0.30$ the variance of the high-fidelity model is zero. This would result in a division of zero by zero and thus resulting in a NaN. Yet, in Figure 7.5 the Pearson correlation coefficient does exist for these values with $\beta = 0.30$ as the exception. A thorough investigation on why these values are NaN could be conducted, although it is considered out of the scope of this research. Moreover, for the Hill model in Figure 7.6, there are also values of the Pearson correlation coefficient equal to NaN for $\beta = 0.20$ and $\beta = 0.21$. After an inspection of the pilot samples it is concluded that the high-fidelity model variance is equal to zero here. The same high-fidelity model pilot samples are used for the Gompertz model, but there the Pearson correlation coefficient is very close to zero and not NaN. Therefore, it is unclear why the Pearson correlation coefficient is NaN or very close to zero for similar data.

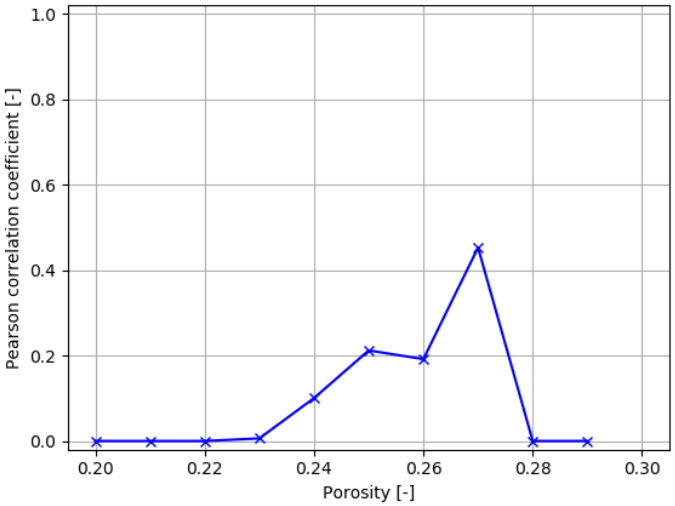


Figure 7.5: Pearson correlation coefficient versus porosity using the Gompertz model as the low-fidelity model.

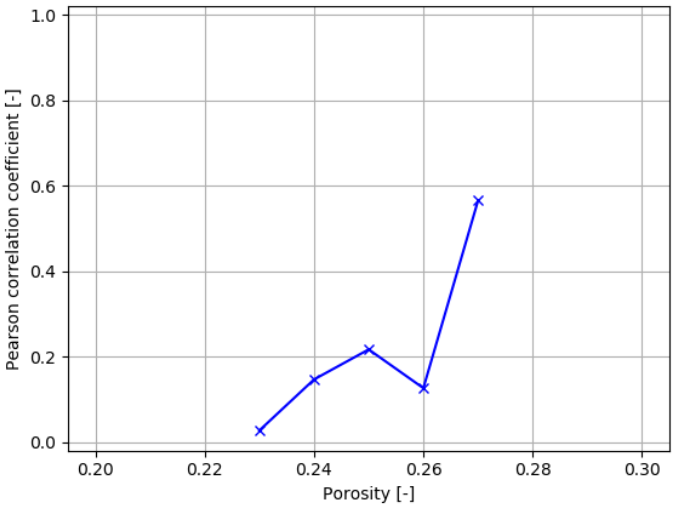


Figure 7.6: Pearson correlation coefficient versus porosity using the Hill model as the low-fidelity model.

On the other hand, there are two peaks in Figure 7.5 and Figure 7.6. One at $\beta = 0.25$ and the other at $\beta = 0.27$. The peak at $\beta = 0.25$ is about equal for both the Gompertz and Hill model at $\rho = 0.21$. The highest Pearson correlation coefficient for the Gompertz model is $\rho = 0.45$ at $\beta = 0.27$. For the Hill model the maximum Pearson correlation coefficient is also found at $\beta = 0.27$ with $\rho = 0.58$. This means the correlation between the high- and low-fidelity model is relatively low compared to the results of the first case study in chapter 5, which were at a minimum $\rho = 0.94$. The range between $\beta = 0.24$ and $\beta = 0.27$ could indicate that the cut-off occurs, since the high-fidelity model samples are not constant and the Pearson correlation coefficient is not NaN or zero. Concluding, it is not possible to state that either low-fidelity model correctly captures the physical behaviour, since an issue with the frequency detection algorithm resolution renders a number of the results as possibly incorrect. There is a range of porosity values for which the Pearson correlation coefficient does exist or is larger than zero and the Pearson correlation coefficient is relatively low compared to results from the previous case study (chapter 5). It is presumed that the epistemic uncertainty in the high-fidelity model is present in the entire range of porosity values and no clear bounds are determined.

8

Conclusions & recommendations

Firstly, the general summarized discussion of the experiments and literature research are described. Secondly, a conclusion is presented in order to answer the main research question. Thirdly, a list with recommendations on future research is presented, also reflecting on technical obstacles and what could be improved upon for this study.

8.1. Discussion

To answer the main research question, presented in chapter 1, four sub-questions are addressed. The first one is covered in section 2.4, the second, third and fourth research questions are addressed in chapter 5, 6 and 7 respectively. The discussion of these is provided below.

- *What is a low-fidelity model and are there low-fidelity models available from literature for application in the parametric study?*

  This sub-question is addressed in the literature review, introducing a formal definition of low-fidelity models as a model that follows a general trend of realism and has low computational cost. Furthermore, a list of empirical relationships is presented between the Strouhal number and Reynolds number. No empirical relationships with respect to the Strouhal number and porosity are found. Therefore, these relationships are derived from the results of the parametric study (chapter 6).

- *What steps are taken when applying Multi-Level Multi-Fidelity Monte Carlo methods on a simple fluid dynamics case and how do the computational cost, estimated quantity of interest and estimated variance compare to general Monte Carlo methods?*

  Chapter 5 is dedicated to the application and validation of Multi-level Multi-Fidelity Monte Carlo on a fluid dynamics case. From literature a study on benchmark cases is used in order to validate the correct application of Multi-Level Multi-Fidelity Monte Carlo using a 3D CFD case as the high-fidelity model and a 2D CFD case as the low-fidelity model. In short, Multi-Level Multi-Fidelity Monte Carlo calculates an optimal number of samples per resolution level using a pre-defined number of samples and then reconstructs the estimator using results from varying model fidelities at varying resolution levels. The pre-defined number of samples are called pilot samples. First, general Monte Carlo is applied using the pilot samples for each model fidelity and resolution level. Moreover, for each resolution level two samples are taken for each pilot sample: one sample for the higher resolution level and one for the lower resolution level. This is important since the difference function is calculated between resolution levels. If the input values of the two resolution levels differ, then the difference function is statistically irrelevant. The optimal number of samples per resolution levels are calculated using the results of the pilot samples. Thereafter, additional samples are taken for the optimal number of samples per resolution level. The results of the pilot samples and calculated optimal number of samples are then combined in order to reconstruct the estimated value. The estimator variance and computational cost of the Multi-Level Multi-Fidelity Monte Carlo should be reduced with respect to general Monte Carlo. The application is considered successful, since the estimated quantity of interest corresponds to the expected value, the variance and the computational cost are reduced with respect to general Monte Carlo.

- *What is the effect of increasing porosity on the vortex shedding frequency of a porous circular cylinder and can an empirical relationship be discovered?*

  In chapter 6 the results of the parametric study are described, an in-depth analysis is done in order to formulate an explanation for the observations. The effect of increasing porosity on the vortex shedding frequency is conducted using the Strouhal number. According to literature, for a non-porous circular cylinder, the Strouhal number should remain constant around $St = 0.2$ for the subcritical Reynolds numbers regime. According to the results of the parametric study a somewhat similar behaviour is noticed. The Strouhal number remains constant or increases slightly to $St = 0.26$, but remains about constant around that value. The most interesting observation is the cut-off in Strouhal number at porosity of around 30%. The Strouhal number has a step function-like behaviour, where for porosities under 30% it remains about constant and around 30% it drops to $St = 0$ very fast. The suggested physical explanation for this observation is based on the distance between the back of the cylinder and point at which vortex shedding occurs. It appears that this distance increases for increasing porosity. Therefore, the Strouhal number -which is calculated using the oscillations of the lift force as experienced by the cylinder- differs from the Strouhal number of actual vortex shedding. For increasing porosity, a point is reached at which the cylinder does not experience or 'feel' the lateral forces due to vortex shedding anymore. Lastly, two empirical relationships are proposed from the results which capture the step function-like behaviour of the Strouhal number for increasing porosity.

- *How can low-fidelity models be applied using Multi-Level Multi-Fidelity Monte Carlo and how do these reflect on the validity of this approach?*

  In chapter 7, Multi-Level Multi-Fidelity Monte Carlo and a parametric study on the effect of porosity on the Strouhal number are combined in order to indicate validity of the two proposed low-fidelity models. From the results, no clear indication of validity can be concluded. No boundaries of the range in which epistemic uncertainty is prevalent, can conclusively be determined. Since the Multi-Level Multi-Fidelity Monte Carlo approach cannot be effectively applied for epistemic uncertainty propagation for the porosity values at which the Strouhal number is zero. This is because the variance of the high-fidelity model is near zero at these values, resulting in no correlation between model fidelities. Furthermore, a critical assessment is conducted for the high-fidelity model, since there are some important observations regarding systemic errors present in this last experiment. Lastly, at this moment there are no clear final conclusions based on this final experiment. However, the observed trends of the effects of systemic error in the high-fidelity model on Multi-Level Multi-Fidelity Monte Carlo warrants further investigation.

8.2. Conclusion

The sub-questions are used to answer the following main research question:

How can Multi-Level Multi-Fidelity Monte Carlo be applied in fluid dynamics topics and what is the contribution of its application on the effect of porosity on the vortex shedding frequency for a porous circular cylinder?

The process of applying Multi-Level Multi-Fidelity Monte Carlo to fluid dynamics topics is summarized in the second sub-question. By applying the Multi-Level Multi-Fidelity Monte Carlo method to the case of flow past a porous circular cylinder, the goal is to implement the low-fidelity models derived in chapter 6 using the MLMFMC in order to get an impression validity of these low-fidelity models on a case with different geometries than that of chapter 6. However, the impression was that it could not be validated using the MLMFMC, due to the presumed systemic errors present in the high-fidelity model. The first indication of this systemic error was found in the analysis of the results in chapter 6. The bounds in which this error was most likely to occur, is in the range of porosity values between 20 and 30%. Since this is the range of porosity values in which the change in Strouhal number is largest. In order to obtain conclusive results, it is especially necessary to reduce the systemic error in this range of porosity values, albeit the systemic errors should be reduced in the entire range of values as well. The suspected way to achieve this is to increase the resolution of the frequency range in the Fourier transform, as this is determined to be the largest influence on the systemic error.

8.3. Recommendations

Based on this work the following recommendations are made for further research. These recommendations are comprised of points for future research and obstacles encountered during this study.

- ***Dakota***
  Even though *Dakota* is a relatively easy to use the black-box like features can be a burden to deal with when trying to analyze intermediary results. Also the user guide available online by *Sandia Labs* is sometimes unclear, not in-depth enough or data is missing or unavailable.

- ***olaFlow* toolbox & porosity modeling**
  The drag coefficient was not validated and therefore the porosity model should be validated. Since the force coefficients are calculated using pressure integration handled by *OpenFOAM*, hence it is likely that the forces are incorrectly calculated using the setup of the porosity model used in this study. Starting from the basics of the porosity model, it is quite likely the assumptions used here, were incorrect. The model assumptions are on thin walled structures and for wave-structure interaction. Meaning, the linear and temporal term were neglected and only the quadratic term was considered. For the case setup used in this study, it might have been better to assume the quadratic and temporal terms should be neglected and use the linear term, due to the absence of waves.

- **Study effect of porosity on vortex formation length by validation of experiments & new (numerical) experiments**
  Very little publications [49] [50] were found that report on the effect of porosity on the vortex formation length. Moreover, with the current setup available it would not be a large step to reconstruct the case setup and to validate the results of the papers by *Sun et al.* using numerical experiments. This can either be done using the porosity model or even constructing the porous region in the mesh. After validation of the porosity model, the parametric study (chapter 6) can be re-done in order to validate the results of this study. Another possibility is to study the effect of introducing waves to the cylinder or study the effect of the increased formation length on varying setups of multiple cylinders.

- **Frequency detection algorithm**
  The frequency detection algorithm caused some problems, especially in the last case study. Therefore, the frequency resolution should be increased and case studies 2 and 3 could be done again. Moreover, in a general sense of an automatic frequency detection algorithm the following points are of interest for further research:

  - Automatically detected when a flow pattern is fully developed
  - Find a way to effectively filter out incorrect frequency peaks in the frequency spectrum.

- **Validate proposed empirical relationships/low-fidelity models**
  Two empirical relationships are proposed with similar behaviour, perhaps it is interesting to validate these relationships using experiments. Although it is also likely that the currently selected fit coefficients over- or underestimate the true behaviour, so it is instead suggested to remember the Gompertz and Hill fit functions when a similar parametric study is conducted.

- **Mesh**
  The meshes used in this study are not checked for convergence or optimized in any way. It is recommended to do a mesh convergence study in the future and maybe use a structured mesh in order to have better control on the smallest cell that is created, especially near the porous region of the cylinder.

- **Low-fidelity models & uncertainty problems**
  Even though low-fidelity models are described in the literature review, some obstacles were encountered during the last case study and the definition presented in the literature might not be adequate. Moreover, when applying methods like MLMFMC more time should be spend on defining the uncertainty problem. In this thesis, quite some mistakes have been made in setting up a 'correct' uncertainty problem due to inexperience with the topic. Hence, it is suggested to invest more time in setting up the uncertainty problem, as well as finding suitable low-fidelity models for the specific case of applying MLMFMC.

A

Fit functions

The fit functions applied here are taken from a book by NCSS Statistical Software chapter 351 (page 351-2) available at `https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Curve_Fitting-General.pdf` (last accessed on: 11-06-2021).

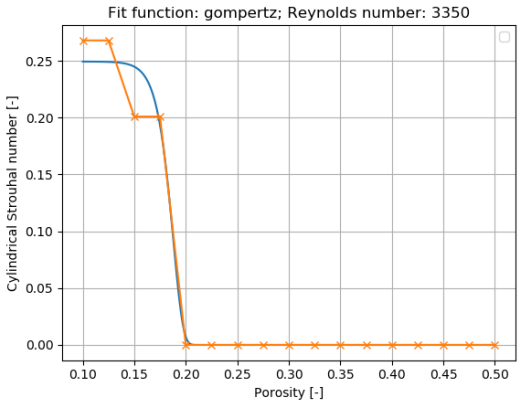
simple linear	$y = ax + b$
quadratic	$y = ax^2 + bx + c$
cubic	$y = ax^3 + bx^2 + cx + d$
polyratio11	$y = \frac{a+bx}{1+cx}$
polyratio22	$y = \frac{a+bx+cx^2}{1+dx+fx^2}$
polyratio33	$y = \frac{a+bx+cx^2+dx^3}{1+fx+gx^2+hx^3}$
polyratio44	$y = \frac{a+bx+cx^2+dx^3+fx^4}{1+gx+hx^2+ix^3+jx^4}$
Michaelis-Menten	$y = \frac{ax}{b+x}$
reciprocal	$y = \frac{1}{a+bx}$
Blaesdalen-Elder	$y = (a+bx)^{\frac{-1}{c}}$
Farazdaghi	$y = \frac{1}{a+bx^c}$
Holliday	$y = \frac{1}{a+bx+cx^2}$
exponential1	$y = e^{a(x-b)}$
monomolecular	$y = y = a(1-e^{-b(x-c)})$
parameterlogistic3	$y = \frac{a}{1+be^{-cx}}$
parameterlogistic4	$y = d + \frac{a-d}{1+be^{-cx}}$
Gompertz	$y = ae^{-e^{-b(x-c)}}$
Weibull	$y = a - (a-b)e^{-c\lvert x\rvert^d}$
Morgan-Mercer-Floding	$y = a - \frac{a-b}{1+c\lvert x\rvert^d}$
Richards	$y = a(1+(b-1)e^{-c(x-d)})^{\frac{1}{1-b}}$
logarithmic	$y = bln(x-a)$
power	$y = a(1-b^x)$
powerpower	$y = ax^{bx^c}$
exponential sum	$y = ae^{-bx} + ce^{-dx}$
exponential type 1	$y = ax^b e^{-cx}$
exponential type 2	$y = (a+bx)e^{-cx} + d$
normal	$y = a + be^{-c(x-d)^2}$
log normal	$y = a + \frac{b}{x}e^{-c(ln(\lvert x\rvert - d))^2}$
exponential2	$y = ae^{-bx}$
Michaelis-Menten 2	$y = \frac{ax}{b+x} + \frac{cx}{d+x}$
Michaelis-Menten 3	$y = \frac{ax}{b+x} + \frac{cx}{d+x} + \frac{fx}{g+x}$
Hill	$y = ax^{\frac{c}{b^c+x^c}}$

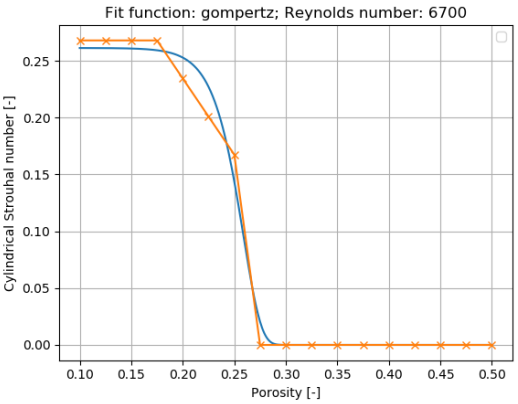
B

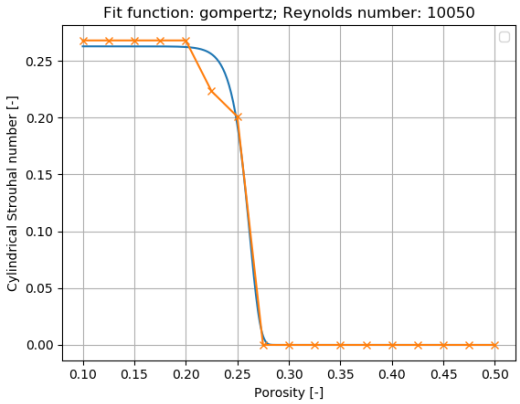
Data fit plots

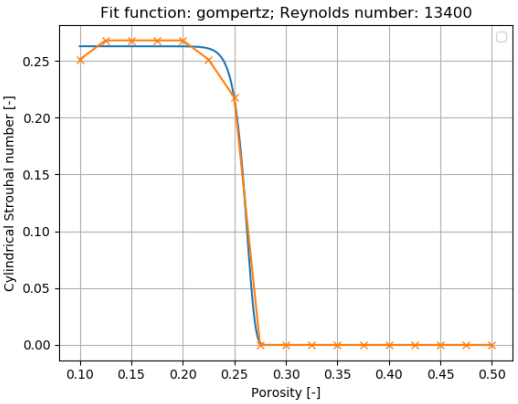
In the first two sections the Gompertz and Hill fit function are fitted to each case. In the last two sections all plots using a generalized Gompertz or Hill fit function are depicted using averaged fit coefficients.

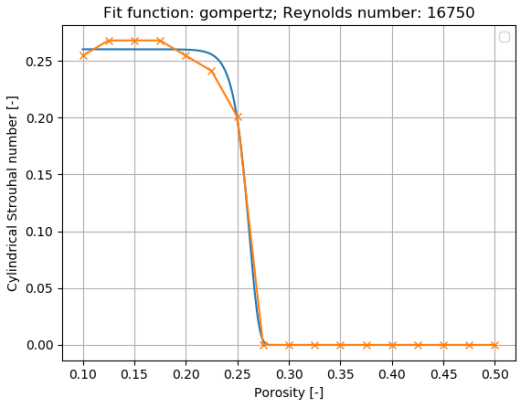
Gompertz fit function: individual fits

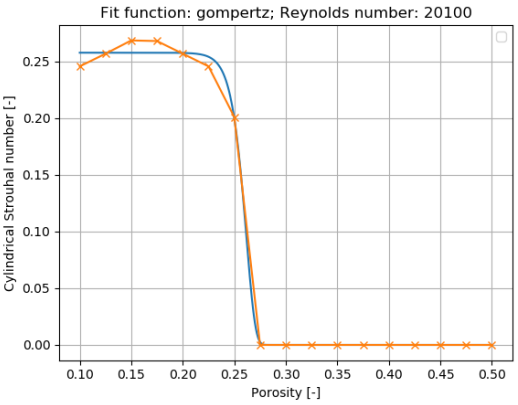


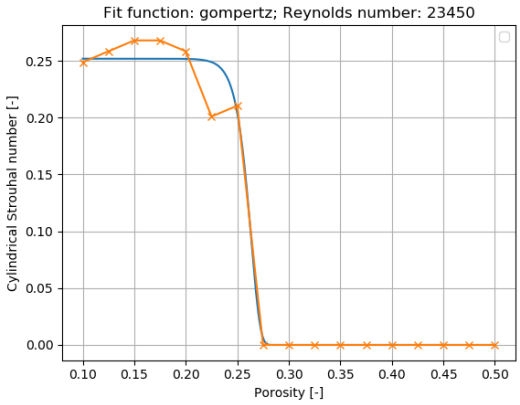


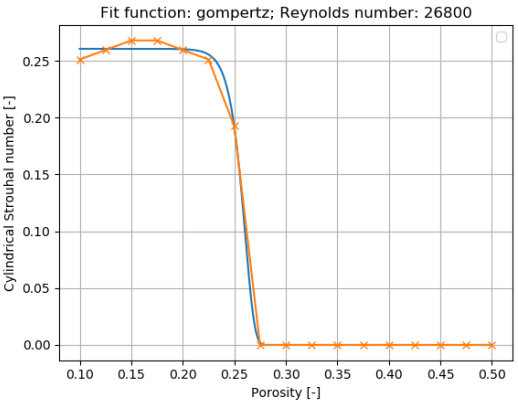


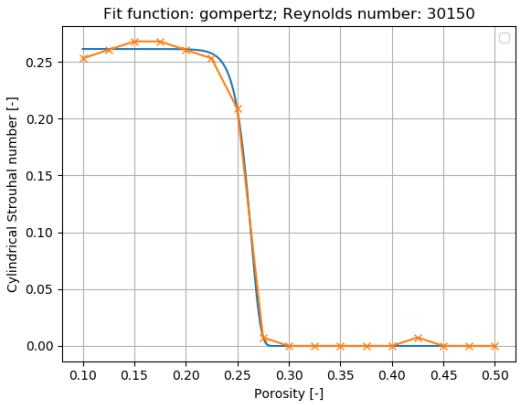


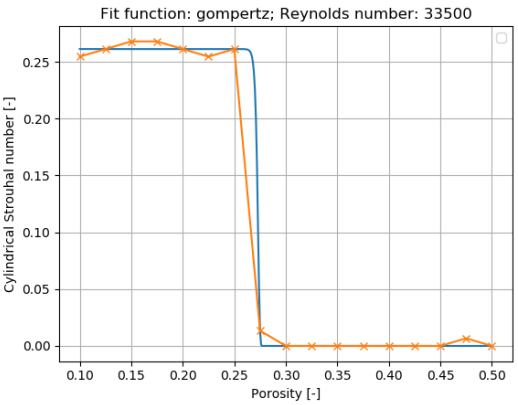

Fit function: gompertz; Reynolds number: 16750
Cylindrical Strouhal number [-]
Porosity [-]


Fit function: gompertz; Reynolds number: 20100
Cylindrical Strouhal number [-]
Porosity [-]

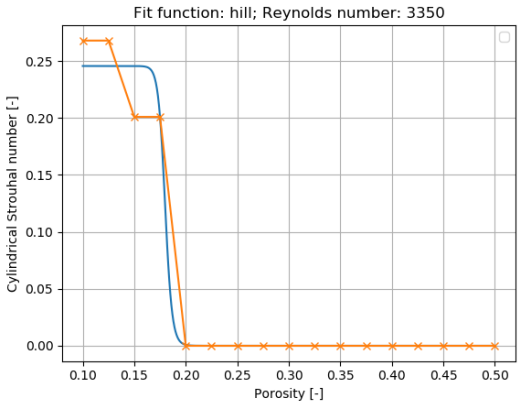

Fit function: gompertz; Reynolds number: 23450
Cylindrical Strouhal number [-]
Porosity [-]

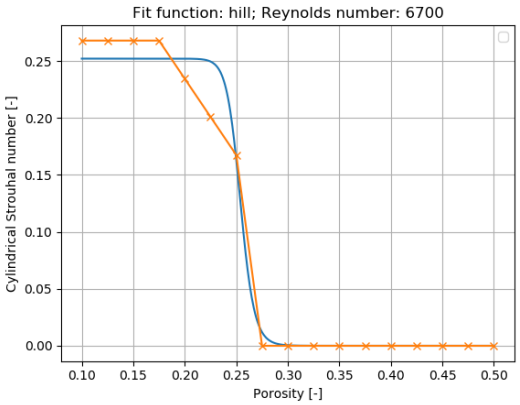

Fit function: gompertz; Reynolds number: 26800
Cylindrical Strouhal number [-]
Porosity [-]

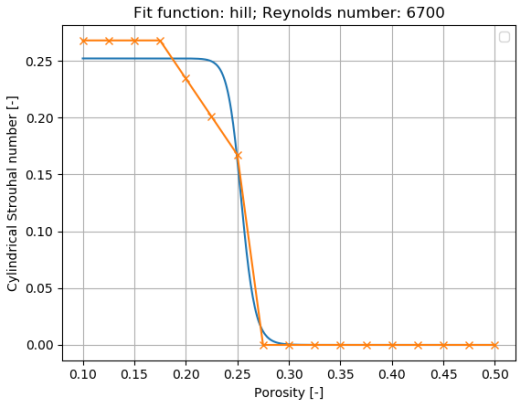

Fit function: gompertz; Reynolds number: 30150
Cylindrical Strouhal number [-]
Porosity [-]

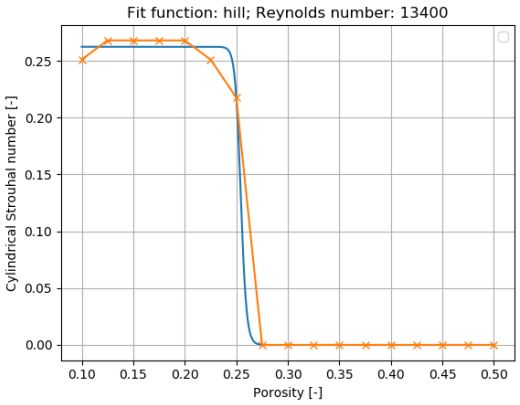

Fit function: gompertz; Reynolds number: 33500
Cylindrical Strouhal number [-]
Porosity [-]

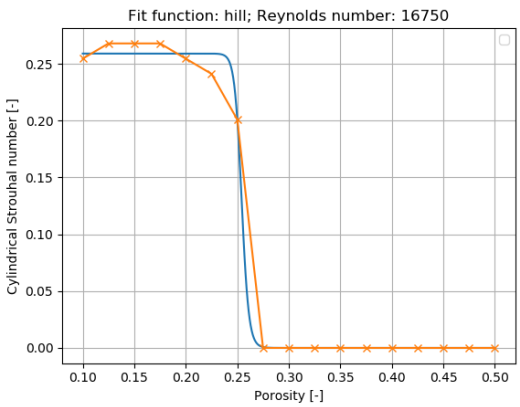
Hill fit function: individual fits

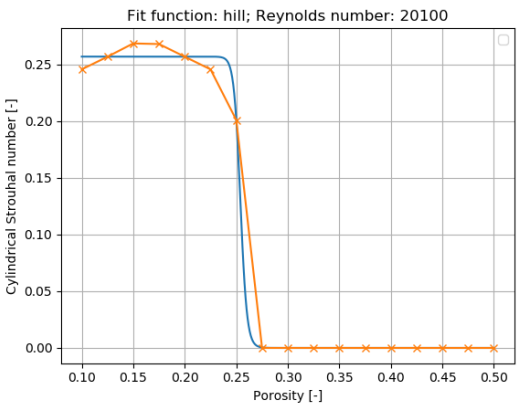


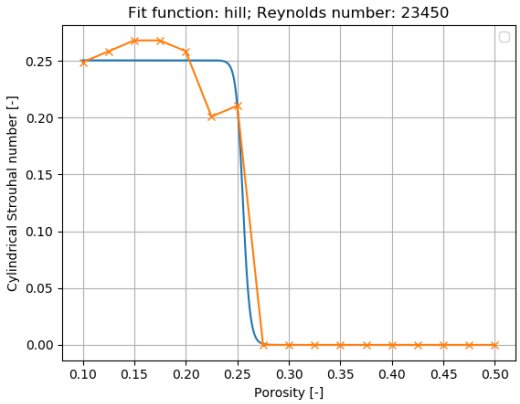


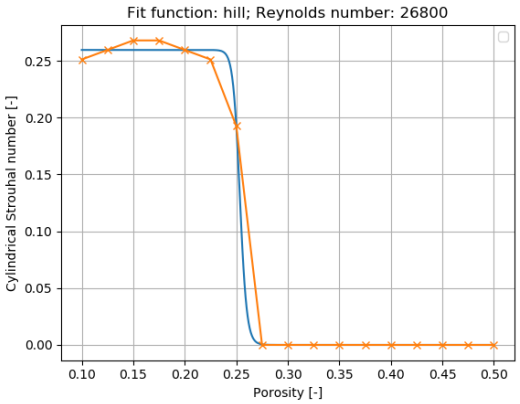


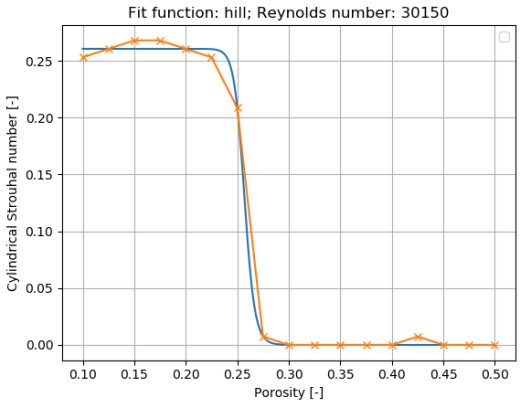


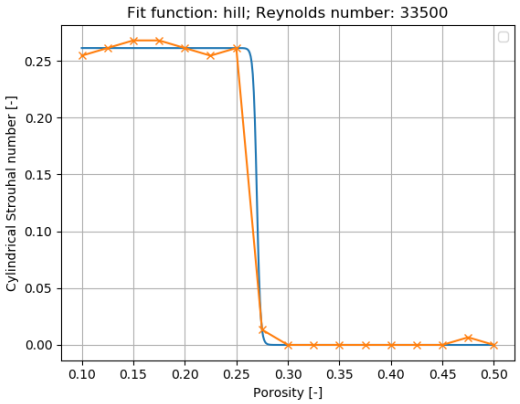




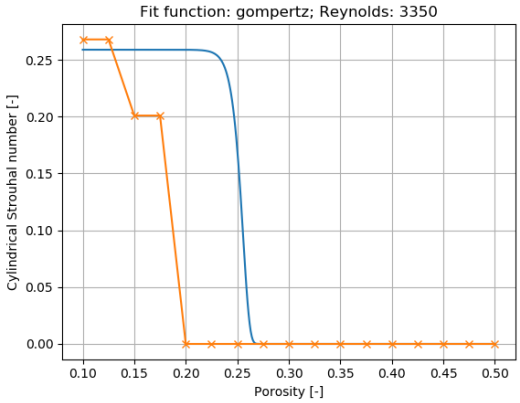


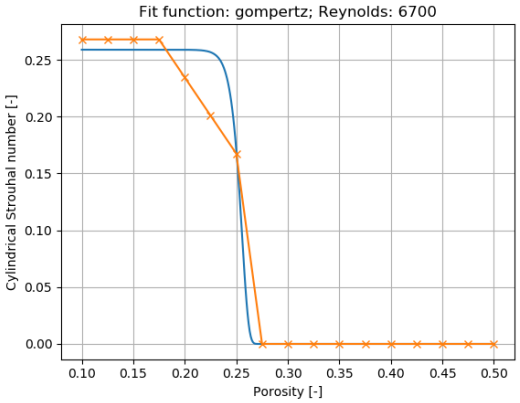


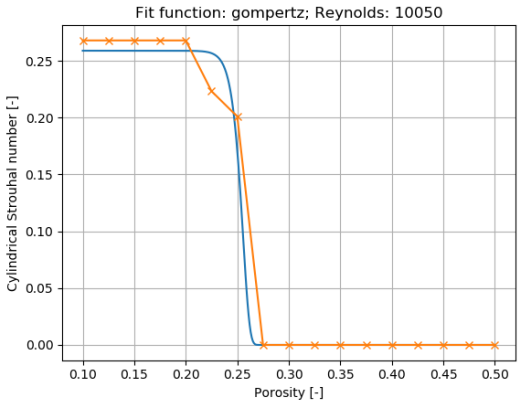


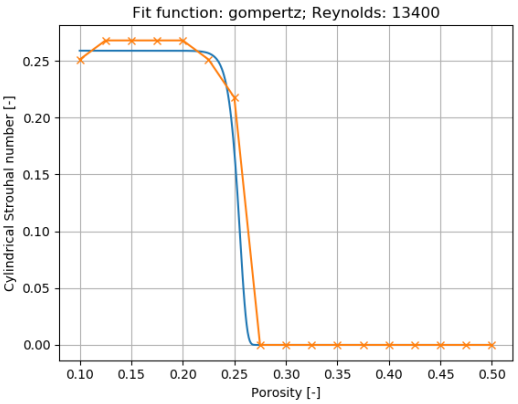


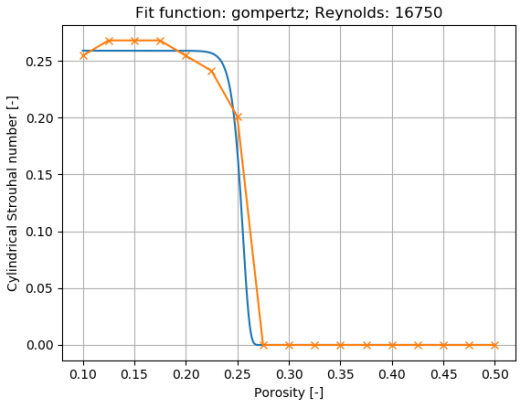
Gompertz fit function: generalized fit

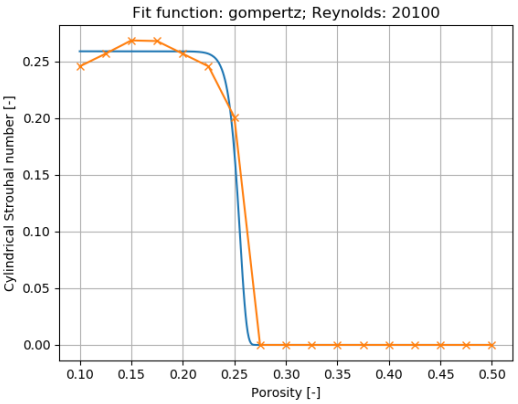


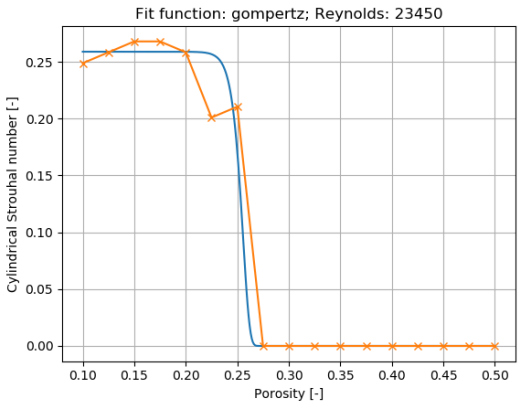


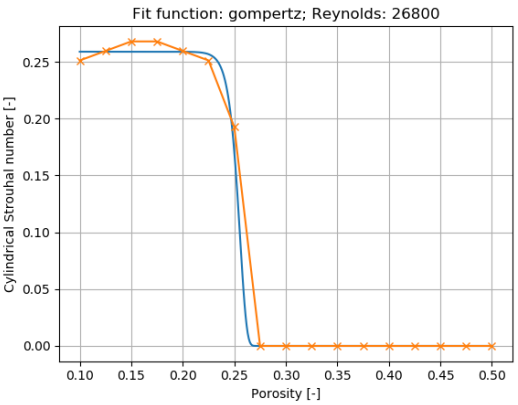

Fit function: gompertz; Reynolds: 10050
Cylindrical Strouhal number [-]
Porosity [-]

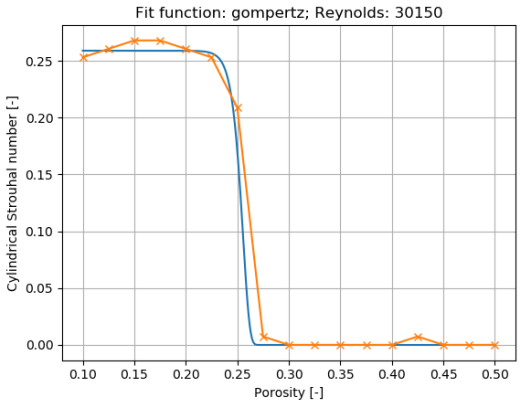

Fit function: gompertz; Reynolds: 13400
Cylindrical Strouhal number [-]
Porosity [-]

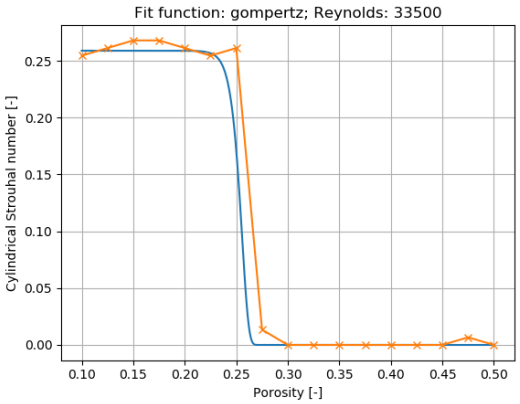

Fit function: gompertz; Reynolds: 16750
Cylindrical Strouhal number [-]
Porosity [-]


Fit function: gompertz; Reynolds: 20100
Cylindrical Strouhal number [-]
Porosity [-]

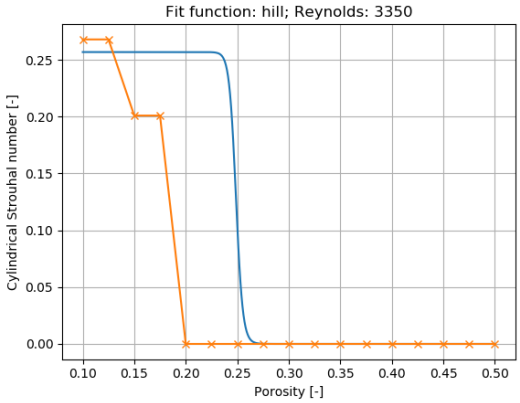

Fit function: gompertz; Reynolds: 23450
Cylindrical Strouhal number [-]
Porosity [-]

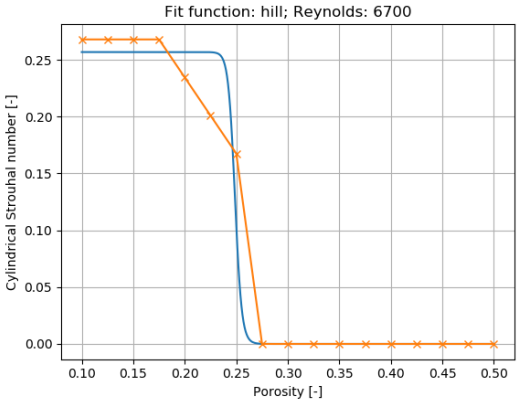

Fit function: gompertz; Reynolds: 26800
Cylindrical Strouhal number [-]
Porosity [-]

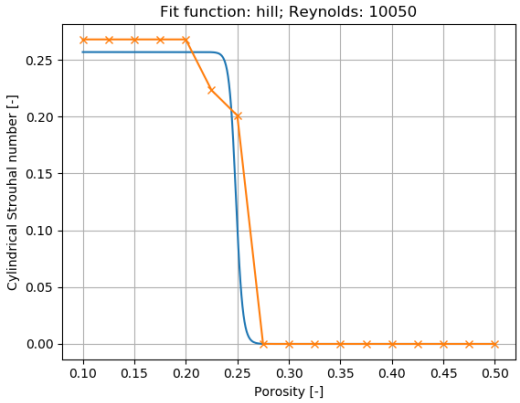


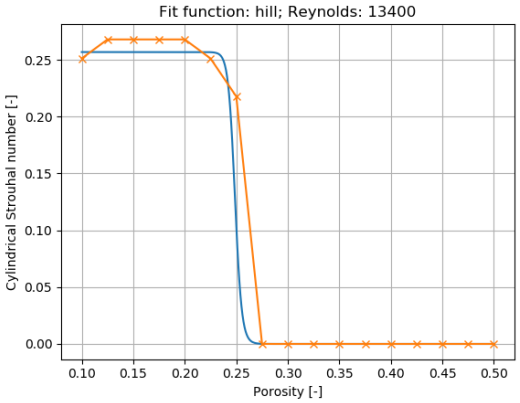


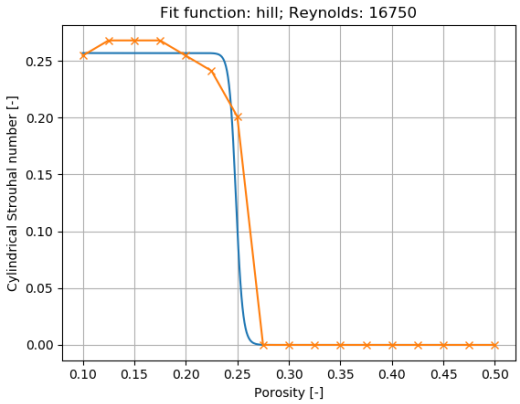
Hill fit function: generalized fit

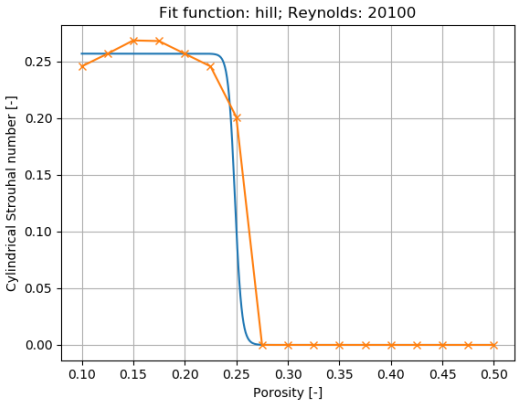


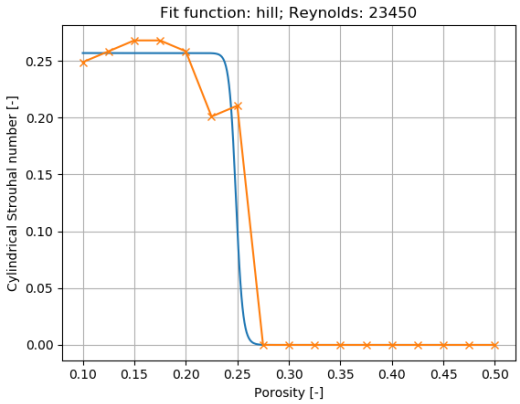


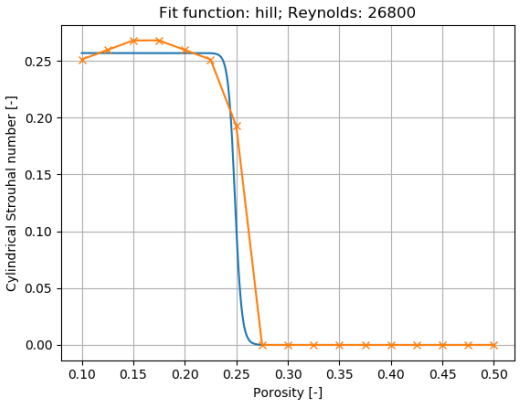


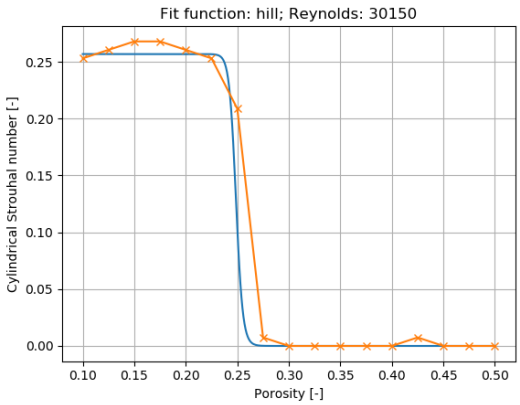


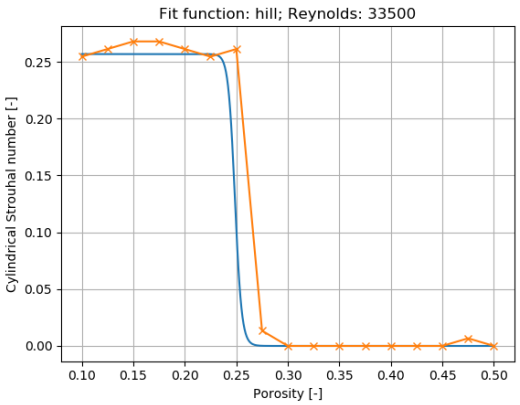

Fit function: hill; Reynolds: 16750
Cylindrical Strouhal number [-]
Porosity [-]


Fit function: hill; Reynolds: 20100
Cylindrical Strouhal number [-]
Porosity [-]


Fit function: hill; Reynolds: 23450
Cylindrical Strouhal number [-]
Porosity [-]

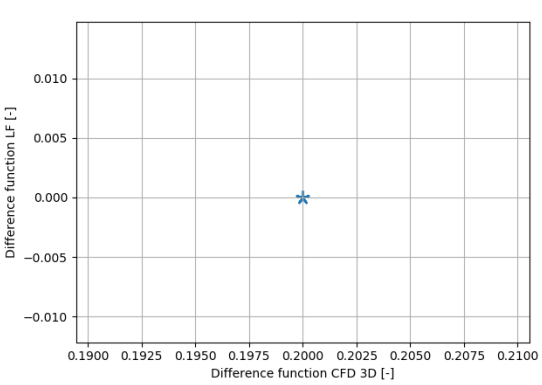

Fit function: hill; Reynolds: 26800
Cylindrical Strouhal number [-]
Porosity [-]


Fit function: hill; Reynolds: 30150
Cylindrical Strouhal number [-]
Porosity [-]

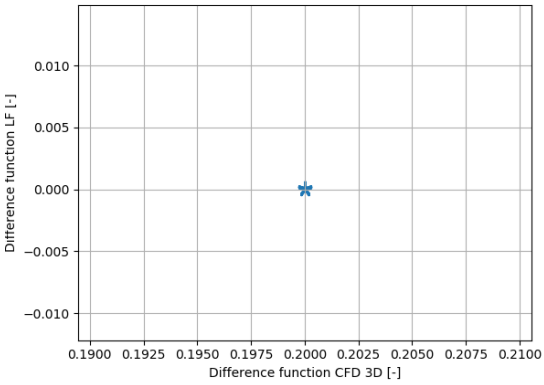

Fit function: hill; Reynolds: 33500
Cylindrical Strouhal number [-]
Porosity [-]

C

Correlation plots

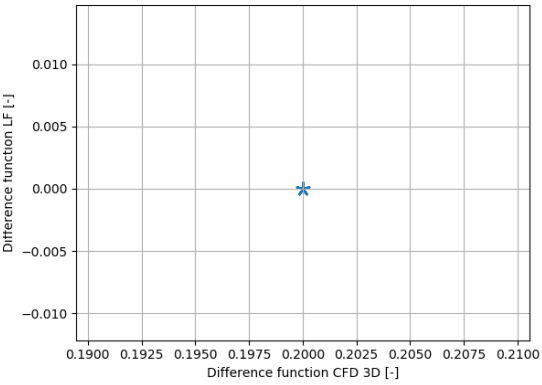


(a) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Gompertz model.

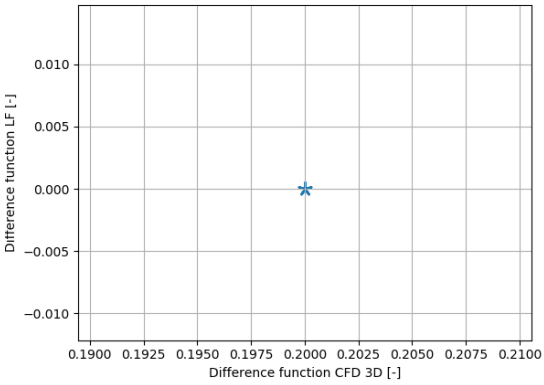


(b) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Hill model.

Figure C.1: Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.28$.

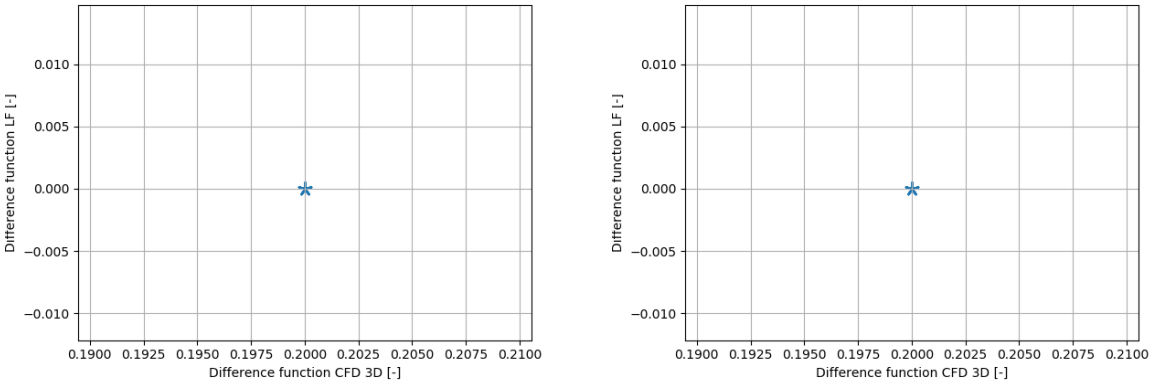


(a) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Gompertz model.



(b) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Hill model.

Figure C.2: Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.29$.



(a) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Gompertz model.

(b) Correlation plot of high-fidelity model 3D CFD with low-fidelity model Hill model.

Figure C.3: Correlation plots of the difference function between high-fidelity 3D CFD and low-fidelity models Gompertz (left) and Hill (right) for $\beta = 0.30$.

Referenced documents

[1] Wang Zhi-Kui, Tsai, Gwo-Chung, Chen Yi-Bao. "*One-Way Fluid-Structure Interaction Simulation of an Offshore Wind Turbine*". May 2014.

[2] Jonkman, Jason M. "*Dynamics Modeling and Loads Analysis of an Offshore Floating Wind Turbine*". National Renewable Energy Laboratory, November 2007, `https://www.nrel.gov/docs/fy08osti/41958.pdf`. NREL/TP-500-41958.

[3] Pudjaprasetya, S. R. "*Modelling and simulation of waves in three-layer porous media*". November 2013.

[4] Higuera, Pablo, Lara, Javier L., Losada, Inigo J. "*Three-dimensional interaction of waves and porous coastal structures using OpenFOAM®. Part I: Formulation and validation*". October 2013.

[5] Zhao, Fen-fang, Bao, Wei-guang, Kinoshita, Takeshi, Itakura, Hiroshi. "*Theoretical and Experimental Study on a Porous Cylinder Floating in Waves*". February 2011.

[6] Park, Min-Su, Jeong, Youn-Ju, You, Young-Jun. "*Numerical analysis of an offshore platform with partial porous cylinders due to wave excitation forces and seismic forces*". October 2012.

[7] Park, Min-Su, Jeong, Youn-Ju, You, Young-Jun. "*Water Wave Interaction by Dual Cylindrical Cylinders with Partial Porous Area*". April 2017.

[8] Zhao, Fen-fang, Kinoshita, Takeshi, Bao, Wei-guang, Huang, Liu-yi, Liang, Zhen-lin, Wan, Rong. "*Interaction between waves and an array of floating porous circular cylinders*". August 2012.

[9] Hirsch, Charles. "*Numerical computation of internal & external flows*". 2007.

[10] Kornbleuth, Marc ."*Studying the Viscous Flow Around a Cylinder Using OpenFoam*".

[11] Goharzadeh, Afshin, Molki, Arman. "*Measurement of fluid velocity development behind a circular cylinder using particle image velocimetry (PIV)*". November 2014.

[12] Aderne, Christopher. "*Investigation of the pressure and velocity distribution around cylinders in an airstream with PIV measurements and CFD*". October 2013.

[13] Tritton, David J. "*Physical fluid dynamics 2nd edition*". 1988.

[14] Steiros, Konstantinos, Kokmanian, Katherine, Bempedelis, Nikolaos, Hultmark, Marcus. "*The effect of porosity on the drag of cylinders*". August 2020.

[15] Dalrymple, Robert A., Rogers, Benedict D. "*Numerical modeling of water waves with the SPH method*". February 2006.

[16] del Jesus, Manuel, Lara, Javier L., Losada, Inigo J. "*Three-dimensional interaction of waves and porous coastal structures Part I: Numerical model formulation*". March 2012.

[17] Higuera, Pablo. "*OlaFlow CFD*". Available: `https://olaflow.github.io/`

[18] Feichtner, Anna, Mackay, Ed, Tabor, Gavin, Thies, Philipp R. "*Modelling wave interaction with thin porous structures using OpenFOAM*". September 2019.

[19] Taylor, Barry N., Kuyatt, Chris E. "*Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*". 1994.

[20] der Kiureghian, Armen , Ditlevsen, Ove. "*Aleatory or epistemic? Does it matter?*". March 2009.

[21] Urbina, Angel, Mahadevan, Sankaran, Paez, Thomas L."*Quantification of margins and uncertainties of complex systems in the presence of aleatoric and epistemic uncertainty*". September 2011.

[22] Geraci, Gianluca, Swiler, Laura P., Crussell, Jonathan, . Debusschere, Bert J. "*EXPLORATION OF MULTIFIDELITY APPROACHES FOR UNCERTAINTY QUANTIFICATION IN NETWORK APPLICATIONS*". June 2019.

[23] Swiler, Laura P., Paez, Thomas L., Mayes, Randall L., Eldred, Michael S."*Epistemic Uncertainty in the Calculation of Margins*". May 2009.

[24] Sin, Gurkan, Espuña, Antonio. "*Applications of Monte Carlo Method in Chemical, Biochemical and Environmental Engineering*". May 2020.

[25] Peherstorfer, Benjamin, Willcox, Karen, Gunzburger, Max. "*Survey of multifidelity methods in uncertainty propagation, inference and optimization*". June 2018.

[26] Kalos, Malvin H., Whitlock, Paula A. "*Monte Carlo Methods*". 2008.

[27] Zhang, Jiaxin. "*Modern Monte Carlo methods for efficient uncertainty quantification and propagation: A survey*". October 2019.

[28] Kroese, Dirk P., Brereton, Tim, Taimre, Thomas, Botev, Zdravko I. "*Why the Monte Carlo method is so important today*". June 2014.

[29] Jonkman, Jason. "*Dynamics of Offshore Floating Wind Turbines—Model Development and Verification*". June 2009.

[30] Giles, Michael B. "*Multilevel Monte Carlo methods*". 2018.

[31] Geraci, Gianluca, Eldred, Michael S., Iaccarino, Gianluca. "*A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications*". January 2017.

[32] Pasupathy, Raghu, Schmeiser, Bruce W., Taaffe, Michael R., Wang, Jin. "*Control-Variate Estimation Using Estimated Control Means*". April 2008.

[33] Geraci, Gianluca, Eldred, Michael S., Gorodetsky, Alex A., Jakeman, John D. "*Recent advancements in Multilevel-Multifidelity techniques for forward UQ in the DARPA Sequoia project*". January 2019.

[34] Gorodetsky, Alex A., Geraci, Gianluca, Eldred, Michael S., Jakeman, John D. "*A generalized approximate control variate framework for multifidelity uncertainty quantification*". January 2020.

[35] Maniaci, David C., Frankel, Ari L., Geraci, Gianluca, Blaylock, Myra L., Eldred, Michael S. "*MULTILEVEL UNCERTAINTY QUANTIFICATION OF A WIND TURBINE LARGE EDDY SIMULATION MODEL*". June 2018.

[36] Hays, Robert T., Singer, Michael J. "*Simulation fidelity in training system design: Bridging the gap between reality and training.*". 1989.

[37] Lord Rayleigh O.M. F.R.S. "*XLVIII. Æolian tones*". 1915.

[38] Roshko, Anatol. "*On the wake and drag of bluff bodies*". February 1955.

[39] Williamson, Chas H. K., Brown, Garry L. "*a series in* $1/\sqrt{Re}$ *to represent the Strouhal-Reynolds number relationship of the cylinder wake*". September 1998.

[40] Roushan, Pedram, Wu, Xiao-lun. "*Structure-Based Interpretation of the Strouhal-Reynolds Number Relationship*". February 2005.

[41] Norberg, Christoffer. "*An experimental investigation of the flow around a circular cylinder: influence of aspect ratio*". January 1994.

[42] Henderson, Ronald D. "*Nonlinear dynamics and pattern formation in turbulent wake transition*". December 1997.

[43] Weller, Henry G., Tabor, Gavin, Jasak, Hrvoje, Fureby, Christer. "*A tensorial approach to computational continuum mechanics using object-oriented techniques*". December 1998.

[44] Ahrens, James, Geveci, Berk, Law, Charles. "*ParaView: An End-User Tool for Large Data Visualization*". 2005.

[45] Geuzaine, Christophe, Remacle, Jean-François."*Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities*". 2009.

[46] Adams, Brian M., Bohnhoff, William J., Dalbey, Keith R., Ebeida, Mohamed S., Eddy, John P., Eldred, Michael S., Hooper, Russell W., Hough, Patricia D., Hu, Kenneth T., Jakeman, John D., Khalil, Mohammad, Maupin, Kathryn A., Monschke, Jason A., Ridgway, Elliott M., Rushdi, Ahmad A., Seidl, D. Thomas, Stephens, J. Adam, Swiler, Laura P., & Winokur, Justin G. "*Dakota, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 6.13 User's Manual*". November 2020.

[47] Virtanen, Pauli, Gommers, Ralf, Oliphant, Travis E., Haberland, Matt, Reddy, Tyler, Cournapeau, David, Burovski, Evgeni, Peterson, Pearu, Weckesser, Warren, Bright, Jonathan, van der Walt, Stéfan J., Brett, Matthew, Wilson, Joshua, Millman, K. Jarrod, Mayorov, Nikolay, Nelson, Andrew R. J., Jones, Eric, Kern, Robert, Larson, Eric, Carey, C J, Polat, İlhan, Feng, Yu, Moore, Eric W., VanderPlas, Jake, Laxalde, Denis, Perktold, Josef, Cimrman, Robert, Henriksen, Ian, Quintero, E. A., Harris, Charles R., Archibald, Anne M., Ribeiro, Antônio H., Pedregosa, Fabian, van Mulbregt, Paul, SciPy 1.0 Contributors. "*SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*". 2020.

[48] Schäfer, Michael, Turek, Stefan. "*Benchmark Computations of Laminar Flow Around a Cylinder*". 1996.

[49] Sun, Chenlin, Azmi, Azlin M., Zhou, Tongming, Zhu, Hongjun, Zang, Zhipeng. "*Experimental study on wake flow structures of screen cylinders using PIV*". October 2020.

[50] Sun, Chenlin, Azmi, Azlin M., Leontini, Justin, Zhou, Tongming, Zhu, Hongjun. "*Experimental study on the development of wake vortices behind screen cylinders*". August 2021.