

CRANFIELD UNIVERSITY

RAJESH SEKAR

UNDERSTANDING AND PREDICTING FLOW BEHAVIOUR OF
WATER AND AIR IN SERPENTINE PIPES USING MACHINE
LEARNING AND CFD MODELLING

THE SCHOOL OF WATER, ENERGY AND ENVIRONMENT
Msc By Research In Energy And Power

MSc by Research
Academic Year: 2022 - 2023

Supervisor: Dr Patrick Verdin
Associate Supervisor: Dr Liyun Lao
September 2023

CRANFIELD UNIVERSITY

SCHOOL OF WATER, ENERGY AND ENVIRONMENT
Msc By Research In Energy And Power

MSc By Research

Academic Year 2022 - 2023

RAJESH SEKAR

UNDERSTANDING AND PREDICTING FLOW BEHAVIOUR OF
WATER AND AIR IN SERPENTINE PIPES USING MACHINE
LEARNING AND CFD MODELLING

Supervisor: Dr Patrick Verdin
Associate Supervisor: Dr Liyun Lao
September 2023

This thesis is submitted in partial fulfilment of the requirements for
the degree of Master by Research

© Cranfield University 2023. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

ACADEMIC INTEGRITY DECLARATION

I declare that:

- The thesis submitted has been written by me alone.
- The thesis submitted has not been previously submitted to this university or any other.
- All content, including primary and/or secondary data, is true to the best of my knowledge.
- All quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

ABSTRACT

The flow behaviour of two-phase fluid (water and air) in serpentine pipes is complex and poorly understood. This study aims to address this research problem by using machine learning concepts to gain a deeper understanding of the flow behaviour and optimize the geometry of serpentine pipes. Experimental data from the Cranfield process engineering laboratory was used in this work, for a fixed pipe diameter and varying water and gas velocities. Regression models were developed and trained. The study aimed to accurately predict the pressure drop, void fraction and liquid film thickness in serpentine pipes in a timely manner with high accuracy.

In addition, Computational Fluid Dynamics (CFD) models were developed to predict the flow behaviour with two different geometries, and machine learning was applied to determine the best model for capturing the intermediate geometry flow behaviour. Results provide valuable insights into the behaviour of two-phase fluid in serpentine pipes. The use of machine learning in this research contributes to the field by offering a new approach for optimizing the geometry of serpentine pipes with improved accuracy and efficiency.

The findings demonstrate the potential for machine learning to play a role in improving our understanding of two-phase fluid flow in serpentine pipes. This research is expected to have potential future applications in various sectors, including automotive, electronics cooling systems, and industrial and chemical processing systems.

Keywords:

process engineering, multiphase, void fraction, liquid film thickness,

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my esteemed supervisor, Dr. Patrick Verdin, and my associate supervisor, Dr. Liyun Lao, for their unwavering guidance and support throughout my research journey in the realm of computer-based studies. Your expertise, encouragement, and mentorship have been instrumental in shaping the trajectory of my project and helping me navigate through various challenges.

Dr. Verdin, your profound knowledge and insightful guidance have provided me with a solid foundation in this intricate field. Your constructive feedback and thoughtful discussions have been invaluable in refining my research approach and expanding my horizons.

Dr. Lao, your guidance and expertise have been a constant source of inspiration. Your insightful suggestions and academic insights have enriched my understanding of computer-based research, enhancing the quality of my work.

I am immensely grateful for the collaborative environment that you both have cultivated. Your mentorship has facilitated interactions with fellow researchers, enabling me to engage in enlightening discussions and gain insights from diverse perspectives. These interactions have significantly contributed to my personal and academic growth.

I extend my heartfelt appreciation for the opportunity to explore the fascinating realm of computer-based studies under your supervision. Your mentorship has played a pivotal role in honing my research skills and deepening my understanding of the subject matter.

Once again, I express my sincere gratitude for your consistent support, guidance, and mentorship. Your collective efforts have been an invaluable asset in my research journey, and I am truly fortunate to have had the privilege of being under your supervision.

With sincere appreciation,

Rajesh Sekar

TABLE OF CONTENTS

ACADEMIC INTEGRITY DECLARATION	i
ABSTRACT	ii
ACKNOWLEDGEMENTS.....	iii
LIST OF FIGURES.....	vii
LIST OF TABLES	xii
LIST OF EQUATIONS.....	xiii
LIST OF ABBREVIATIONS	xiv
1 Introduction.....	1
1.1 Background and Motivation	1
1.2 Methodology	1
1.3 History of curved pipes	2
1.4 Multiphase Flow	4
1.5 Computational Fluid dynamics	6
1.6 Machine learning.....	7
1.7 Computational Fluid dynamics and machine learning in curved pipes.....	9
1.8 Structure of the Thesis	11
2 Aim and Objective	13
2.1 Aim.....	13
2.2 Objectives	13
The following objectives were set to achieve the aim of the project:.....	13
3 Literature Review	14
3.1 Overview.....	14
3.2 Flow regimes	14
3.3 Gravity influence	16
3.4 Computational Fluid dynamics.....	17
3.5 Machine learning techniques	18
3.6 Literature Review Papers.....	24
3.7 Literature findings	25
3.8 Literature findings summary.....	40
4 Experimental Data	42
4.1 Overview of Serpent rig	42
4.2 Flow conditions	43
4.2.1 Air.....	43
4.2.2 Water.....	43
4.3 Sensors and data acquisition system.....	43
4.4 Flow regime map	44
5 Machine Learning and Data Analysis	46
5.1 Overview	46
5.2 Regression models	46
5.3 Data Analysis.....	48

5.4 Hardware and code environment.....	50
5.5 Serpentine pipe data Pre-processing.....	50
5.5.1 Dataset – 01: Void fraction data	50
5.5.2 Dataset – 02: Liquid film thickness data	50
5.5.3 Step by step code used for data pre-processing:	51
5.5.4 Data pre-processing: Min Max Scalar function	52
5.5.5 Data pre-processing: Null error identification	62
5.5.6 Data pre-processing: Standard deviation	63
5.5.7 Data pre-processing: Histogram plots	65
5.5.8 Data pre-processing: pair plot	68
5.5.9 Data pre-processing: coefficient of the model	70
5.5.10 Data pre-processing: Correlation matrix heat map	73
5.6 Regression Models Implementation in this Study	76
6 Computational Fluid Dynamics and Machine learning.....	80
6.1 Introduction	80
6.2 Simulation requirement and approach	81
6.3 Geometries used in the study.	82
6.3.1 The main piping section.	83
6.3.2 Model 1 and Model 2 dimensions.....	83
6.3.3 Inlet and outlet sections.....	83
6.4 Boundary Conditions.....	84
6.5 Mathematical modelling and CFD	85
6.5.1 Assumptions.....	85
6.5.2 VOF theory	85
6.5.3 k-omega turbulence model equations	87
6.5.4 CFD model physics settings:.....	88
6.5.5 CFD model Solver settings:.....	89
6.6 Mesh Selection Study	90
6.6.1 Flow input data.....	90
6.6.2 Boundary conditions.....	91
6.6.3 Y+ criteria	91
6.6.4 Mesh and grid	93
6.6.5 Multi region mesh refinement study	97
6.7 CFD Model Validation	101
6.7.1 Case – 1 air 0.15 m/s & water 1m/s:	102
6.7.2 Case – 02 air 1 m/s & water 1m/s:	107
6.7.3 Case – 03 air 9 m/s & water 1m/s:	113
6.8 Summary of Validation Findings and Future Steps:.....	118
7 Results and Discussion	120
7.1 Flow characterisation using ML models.	120
7.2 Gas void data and Liquid Film Thickness validation and prediction.	121
7.2.1 Pressure, void fraction and liquid film thickness	122

7.2.2 ML's Research Findings Summary	143
7.3 CFD Simulation and ML-Based Geometry Optimization.....	149
7.3.1 Simulation results	150
7.3.2 ML Results	154
7.3.3 Validation – test data.....	156
7.3.4 Predictions	159
8 Conclusions.....	161
9 Future Work.....	165
REFERENCES.....	167
APPENDICES	175
CFD Results and Machine learning Code	175
A.1 CFD simulation void fraction average at downward, top bottom and middle.....	175
A.2 CFD simulation void fraction average at downward, top bottom and middle.....	179

LIST OF FIGURES

Figure 1-1: Streamlines show the development of flow in a tight bend (Zhang, Roberts and Benard, 2012)	3
Figure 1-2: Multiphase flow interface chart.....	5
Figure 3-1: Flow Regimes ((Islam Fetoui, 2005.)	14
Figure 3-2: Over fitting model complexity (Steven L. Brunton, 2022)	19
Figure 3-3: Over fitting model complexity (Steven L. Brunton, 2022)	20
Figure 3-4: Machine learning application in fluid dynamics (Sharma et al., 2023)	23
Figure 3-5: Different methods for embedding physics domain knowledge into a supervised learning framework (Sharma et al., 2023)	24
Figure 3-6: Statistical error in void fraction (Kendoush and Ai-Khatib, 1994) ..	31
Figure 3-7: Schematic illustration of the multiphase flow testing facility (Sharma et al., 2023).....	38
Figure 4-1 Serpentine test rig	42
Figure 4-2: vertical downward flow regime (Jiang and Rezkallah, 1993).....	45
Figure 5-1: Overview of machine approach learning flow chart.....	46
Figure 5-2: Libraries used in the code.	51
Figure 5-3: Code for reading and scaling data	52
Figure 5-4: Scaled value vs original value: superficial velocities and flow rates	54
Figure 5-5: Scaled value vs original value: input pressure and temperatures ..	56
Figure 5-6: Scaled value vs original value: output pressure at stations.....	57
Figure 5-7: Scaled value vs original value: void fraction at each station	57
Figure 5-8: Scaled value vs original value: film thickness at down CT(a).....	58
Figure 5-9: Scaled value vs original value: film thickness at down CM(a)	59
Figure 5-10: Scaled value vs original value: film thickness at down CB(a).....	59
Figure 5-11: Scaled value vs original value: film thickness at down CT(b).....	60
Figure 5-12: Scaled value vs original value: film thickness at down CM(b)	60
Figure 5-13: Scaled value vs original value: film thickness at down CB (b).....	61
Figure 5-14: Null error results.....	62

Figure 5-15: Standard deviation program.....	63
Figure 5-16: Input data standard deviation plot	64
Figure 5-17: Data set 1 output standard deviation plot.....	64
Figure 5-18: Data set 2 output standard deviation plot.....	65
Figure 5-19: Histogram code	66
Figure 5-20: Input histogram plot-superficial liquid (left) and gas (right) velocity	67
Figure 5-21: Input histogram plot-standard (left) and original (right) flow rates .	67
Figure 5-22: Input histogram plot gauge (left) and absolute pressure (right)....	68
Figure 5-23: Input histogram plot-inlet (left) and outlet (right) temperature.	68
Figure 5-24: Pair plot code	69
Figure 5-25:Input data pair plot	69
Figure 5-26: Code for coefficient calculation	70
Figure 5-27: Data set -01 coefficient	71
Figure 5-28: Data set -02 coefficient	72
Figure 5-29: Code for correlation matrix heat map	74
Figure 5-30: Data set 01 correlation matrix heat map	75
Figure 5-31: Data set 02 correlation matrix heat map	76
Figure 5-32: Import libraries used in the model.	77
Figure 5-33: Loading and preprocessing section.....	77
Figure 5-34: Regression models code.....	78
Figure 6-1: Simulation overview flow chart.....	81
Figure 6-2: CAD model representing the experimental pipe setup	82
Figure 6-3: Inlet section for liquid and gas.....	83
Figure 6-4: Trimmed mesh	93
Figure 6-5: Residuals plot (finer mesh, 8 prism layer)	95
Figure 6-6: Void fraction convergence plot (finer mesh, 8 prism layer)	95
Figure 6-7: Wall y+ plot (finer mesh, 8 prism layer).....	96
Figure 6-8: Mesh convergence plot.....	97
Figure 6-9: Volume fraction air plot (finer mesh, 8 prism layer)	99

Figure 6-10: Multi region mesh.....	100
Figure 6-11: Case 1- Void fraction plot downward pipe X section	102
Figure 6-12: Case 1 - Time average void fraction comparison	103
Figure 6-13: Case 1 – Down top time trace CFD vs experiment comparison .	104
Figure 6-14: Case 1 – Down middle time trace CFD vs experiment comparison	105
Figure 6-15: Case 1 – Down bottom time trace CFD vs experiment comparison	105
Figure 6-16: Case1 - Probability density function comparison	106
Figure 6-17: Case 2 Void fraction plot downward pipe X section	107
Figure 6-18: Case 2 - Time average void fraction comparison	108
Figure 6-19: Case 2 – Down top time trace CFD vs experiment comparison .	109
Figure 6-20: Case 2 – Down middle time trace CFD vs experiment comparison	110
Figure 6-21: Case 2 – Down bottom time trace CFD vs experiment comparison	111
Figure 6-22: Case 2 - Probability density function comparison.....	112
Figure 6-23: Case 3-Time average void fraction comparison.....	113
Figure 6-24: Case 3 – Down top time trace CFD vs experiment comparison .	114
Figure 6-25: Case 3 – Down top time trace CFD vs experiment comparison .	115
Figure 6-26: Case 3 – Down middle time trace CFD vs experiment comparison	116
Figure 6-27: Case 3 – Down bottom time trace CFD vs experiment comparison	117
Figure 6-28: Case 3 - Probability density function comparison.....	118
Figure 7-1: Sensor and reading location	122
Figure 7-2: Validation – pressure at downward pipe -top	125
Figure 7-3: Prediction – pressure at downward pipe -top	126
Figure 7-4: Validation – void fraction at downward pipe - top	126
Figure 7-5: Prediction – void fraction at downward pipe – top	127
Figure 7-6: Validation – liquid film thickness at downward pipe - top	127
Figure 7-7: Prediction – liquid film thickness at downward pipe – top	128

Figure 7-8: Validation – pressure at downward pipe - middle.....	128
Figure 7-9: Prediction – pressure at downward pipe – middle.....	129
Figure 7-10: Validation – void fraction at downward pipe - middle	129
Figure 7-11: Prediction – void fraction at downward pipe – middle	130
Figure 7-12: Validation – liquid film thickness at downward pipe - middle.....	130
Figure 7-13: Prediction – liquid film thickness at downward pipe – middle.....	131
Figure 7-14: Validation – pressure at downward pipe - bottom	131
Figure 7-15: Prediction – pressure at downward pipe – bottom	132
Figure 7-16: Validation – void fraction at downward pipe - bottom	132
Figure 7-17: Prediction – void fraction at downward pipe – bottom.....	133
Figure 7-18: Validation – liquid film thickness at downward pipe - bottom	133
Figure 7-19: Prediction – liquid film thickness at downward pipe – bottom	134
Figure 7-20: Validation – pressure at upward pipe – bottom	134
Figure 7-21: Prediction – pressure at upward pipe – bottom.....	135
Figure 7-22: Validation – void fraction at upward pipe - bottom.....	135
Figure 7-23: Prediction – void fraction at upward pipe – bottom.....	136
Figure 7-24: Validation – liquid film thickness at upward pipe - bottom	136
Figure 7-25: Prediction – liquid film thickness at upward pipe – bottom	137
Figure 7-26: Validation – pressure at upward pipe - middle	137
Figure 7-27: Prediction – pressure at upward pipe – middle	138
Figure 7-28: Validation – void fraction at upward pipe - middle.....	138
Figure 7-29: Prediction – void fraction at upward pipe – middle	139
Figure 7-30: Validation – liquid film thickness at upward pipe - middle	139
Figure 7-31: Prediction – liquid film thickness at upward pipe – middle	140
Figure 7-32: Validation – pressure at upward pipe - top.....	140
Figure 7-33: Prediction – pressure at upward pipe – top.....	141
Figure 7-34: Validation – void fraction at upward pipe - top	141
Figure 7-35: Prediction – void fraction at upward pipe – top	142
Figure 7-36: Validation – liquid film thickness at upward pipe – top	142

Figure 7-37: Prediction – liquid film thickness at upward pipe – top	143
Figure 7-38: Gas velocity 0.15m/s void fraction scalar plot for different pipe diameters.....	152
Figure 7-39: Gas velocity 1 m/s void fraction scalar plot for different pipe diameters.....	153
Figure 7-40: Gas velocity 9 m/s void fraction scalar plot for different pipe diameters.....	153
Figure 7-41: Validation 01 CFD vs ML comparison	157
Figure 7-42: Validation-02 CFD vs ML comparison.....	158
Figure 7-43: Validation-03 CFD vs ML comparison.....	158

LIST OF TABLES

Table 6-1: Inlet air and water diameter area.....	84
Table 6-2: Fluid properties	84
Table 6-3: Physics setting solver.....	88
Table 6-4: CFD solver settings	89
Table 6-5 : Experimental data used for the study.....	91
Table 6-6: Boundary conditions – inlet flow rates	91
Table 6-7: Mesh dependence study results.....	94
Table 6-8: Multi-region mesh size	100
Table 7-1: Inputs used for validation.	120
Table 7-2: Prediction input values	120
Table 7-3: Pressure data ML models performance matrix.....	143
Table 7-4: Void Fraction data ML models performance matrix.....	145
Table 7-5: Liquid Film Thickness data ML models performance matrix.....	147
Table 7-6: CFD simulation boundary conditions.....	150
Table 7-7: void fraction simulation results	151
Table 7-8: Mean Absolute Error values.....	155
Table 7-9: validation-01 CFD and ML void fraction results.....	156
Table 7-10: validation-02 CFD and ML void fraction results.....	157
Table 7-11: validation-0 3CFD and ML void fraction results.....	158
Table 7-12: ML predictions vs CFD data.....	160

LIST OF EQUATIONS

Equation 3-1 18

Equation 3-2 20

Equation 3-3 21

Equation 3-4 21

Equation 3-5 21

Equation 3-6 21

Equation 5-1 47

Equation 5-2 47

Equation 5-3 55

Equation 6-1 85

Equation 6-2 86

Equation 6-3 86

Equation 6-4 86

Equation 6-5 86

Equation 6-5 86

Equation 6-7 87

Equation 6-8 88

Equation 6-8 92

LIST OF ABBREVIATIONS

l	Length of the pipe (mm)
a	Radius of the pipe (mm)
A_1	Radius of the pipe (mm)
A_2	Radius of the pipe (mm)
c	Radius of the bend (mm)
CFD	Computational Fluid Dynamics
CAD	Computer Aided Design
u_{sl}	Superficial liquid velocity
u_{sg}	Superficial gas velocity
VOF	Volume of fluid

1 Introduction

1.1 Background and Motivation

Serpentine pipes are widely used in various industries for heat transfer, fluid flow, and energy exchange processes. These pipes, characterized by their complex geometry, play a crucial role in optimizing the performance and efficiency of systems such as heat exchangers, condensers, and refrigeration units (Vasiliev, 2005) . Understanding and accurately predicting the fluid dynamics within serpentine pipes is of paramount importance for system design, optimization, and operation.

Traditionally, the analysis of fluid flow in serpentine pipes has relied on Computational Fluid Dynamics (CFD) techniques (Lote, Vinod and Patwardhan, 2018). CFD methods utilize numerical algorithms and models to solve the governing equations of fluid motion, providing valuable insights into the behaviour of fluids in complex geometries. However, CFD simulations usually require substantial computational resources and can be time-consuming, especially for detailed analyses of large-scale systems or extensive parameter variations.

Recent advancements in machine learning (ML) have offered promising opportunities to enhance and complement traditional CFD approaches (Vinuesa and Brunton, 2022). ML techniques, such as artificial neural networks and deep learning algorithms, have demonstrated remarkable capabilities in pattern recognition, data analysis, and predictive modelling. Leveraging these techniques in conjunction with CFD simulations can significantly accelerate the analysis process and provide valuable insights into the intricate flow behaviour within serpentine pipes.

1.2 Methodology

The proposed research will employ a systematic methodology to accomplish the stated objectives. It will involve the following steps:

a) Review of existing literature on serpentine pipes, CFD methods, and machine learning techniques applied to fluid dynamics.

- b) Development and validation of a CFD model for simulating fluid flow in serpentine pipes, considering various boundary conditions and fluid properties.
- c) Collection and pre-processing of relevant data for training and testing machine learning models.
- d) Exploration and implementation of various machine learning algorithms, such as artificial neural networks, support vector machines, and random forests, for flow prediction and optimization.
- e) Integration of the CFD model with the trained machine learning models to enhance the accuracy and efficiency of flow analysis.
- f) Validation and benchmarking of the integrated framework using experimental data and comparisons with existing numerical models.
- g) Generation of practical guidelines and recommendations for the implementation of the proposed framework in industrial applications.

Through the above-described methodology, the research will provide valuable insights into the potential of integrating CFD and machine learning for serpentine pipe analysis and contribute to the advancement of engineering design and optimization.

1.3 History of curved pipes

Curved pipes have played a significant role throughout history, showcasing remarkable advancements in engineering, industry, and fluid transportation. These pipes have emerged as indispensable components across various sectors. This section will reach into the intriguing history of curved pipes, examining their evolution and enduring impact. Ancient civilizations, including the Romans, Greeks, and Egyptians, ingeniously utilized curved pipes constructed from materials such as clay, stone, and bronze for the transportation and distribution of water (Mays, Koutsoyiannis and Angelakis, 2007). These early societies recognized the need to overcome challenges in optimizing flow dynamics, leading to the development of curved pipe designs that allowed for efficient water movement.

Over time, the utilization of curved pipes expanded with the advent of the Industrial Revolution. As industries such as manufacturing, mining, and transportation flourished, the demand for more efficient fluid transportation systems grew exponentially. Curved pipes emerged as a viable solution, offering numerous advantages over their straight counterparts. Their unique design allowed for more compact installations, enabling efficient flow distribution and enhanced pressure management. During the 18th and 19th centuries, the significance of curved pipes extended to the development of steam engines (Frenken and Nuvolari, 2004). These groundbreaking inventions revolutionized industries and transportation. Curved pipes played a pivotal role in steam distribution, condensation, and heat exchange processes, facilitating the efficient transfer of thermal energy. By employing curved configurations, engineers were able to optimize the performance of steam-powered technologies, contributing to unprecedented advancements.

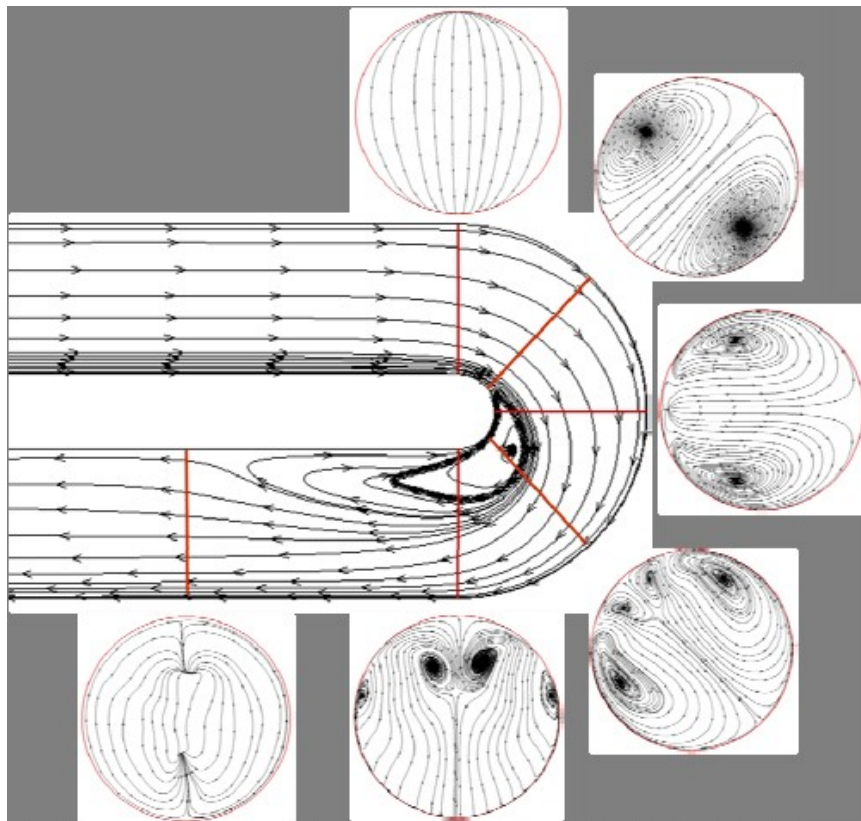


Figure 1-1: Streamlines show the development of flow in a tight bend (Zhang, Roberts and Benard, 2012)

The utilization of curved pipes further expanded in the 20th century, particularly in the field of heating, ventilation, and air conditioning (HVAC) systems (Faghri, 2012). The integration of curved pipe designs into heat exchangers, radiators, and cooling coils revolutionized the efficiency of heat transfer and temperature regulation. In the Figure 1-1 shows the stream lines flow in the bend pipe with flow variations across the bend and generation of secondary flow as well. This variation in the flow phenomenon needs to be understand to diverse its application. This application continues to be of paramount importance in modern HVAC systems, ensuring optimal thermal management and enhanced energy efficiency. As materials and manufacturing techniques advanced, the design and fabrication of curved pipes underwent notable transformations. Traditional materials such as copper and steel were joined by innovative alternatives like plastics and composite materials. The evolution of advanced manufacturing processes, including bending, welding, and casting, enabled the production of increasingly complex curved pipe configurations (Zhou et al., 2021). With the advent of computational fluid dynamics (CFD) and optimization techniques, the understanding and design of curved pipes reached unprecedented levels. Engineers could now analyze and simulate flow behaviour, optimizing pipe geometry and overall performance. This revolutionary approach empowered industries to achieve greater efficiency and reliability in fluid transportation systems.

In contemporary times, curved pipes continue to occupy a pivotal role across various industries. Their diverse applications span HVAC systems, automotive engineering, aerospace, chemical processing, and renewable energy. Curved pipes remain integral components of heat exchangers, condensers, fluid transportation systems, and other critical applications that demand efficient flow management and superior heat transfer capabilities.

1.4 Multiphase Flow

The multiphase flow refers to the simultaneous movement of multiple fluid phases, like oil, gas, drilling fluid, fracturing fluids, water and air within a single system. These fluids exhibit distinct fluid properties and flow characteristics.

Figure 1-2 shows the interfaces possibility chart, when multiple fluids coexist and interact at a phase interface within a shared system, their respective behaviours are influenced and may manifest differently.

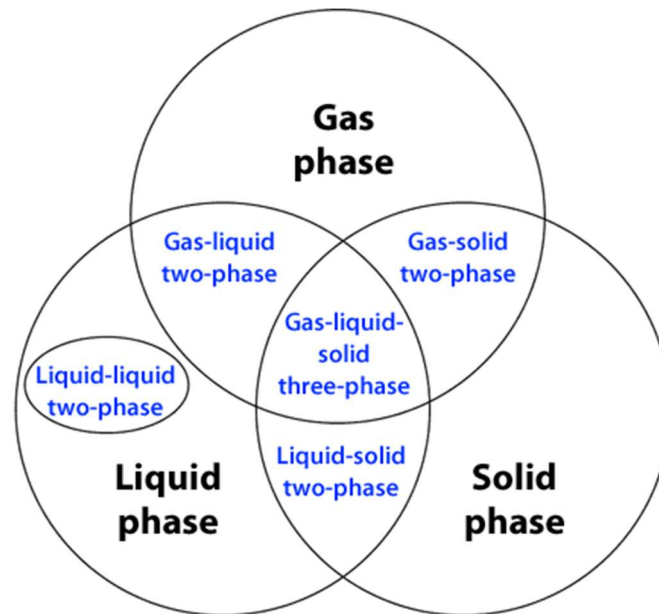


Figure 1-2: Multiphase flow interface chart

Compared to modern times, multiphase flow was less understood and could not be controlled as easily in ancient times. Ancient civilizations like the Egyptians, Greeks, and Romans possessed a fundamental understanding of fluid behaviour, but they had a limited comprehension of complex multiphase processes (Li et al., 2021). The construction of aqueducts and irrigation systems is a notable historical example of dealing with multiphase flow. In order to ensure effective water distribution and transportation, ancient engineers had to control the flow of water while taking into account its interactions with the atmosphere and solid objects. However, with the development of fluid mechanics and the Industrial Revolution, thorough scientific research and rigorous analyses of multiphase flow came into being much later. Fluid dynamics was greatly advanced by early thinkers like Daniel Bernoulli, Leonhard Euler, and Henri Pitot, creating the foundation for understanding the fluid flow phenomena. With the help of experimental techniques, advancement in computational methods the study of multiphase flow has evolved significantly over the years. Those advancement helps to solve the complex flow regimes better empirical correlations. For example, the

experimental technique such as high-speed imaging, laser-based diagnostics, and advances sensors have enabled the researchers to understand the complex multiphase flow characterises and to validate and to improve the numerical simulations accuracy.

1.5 Computational Fluid dynamics

Computational Fluid Dynamics (CFD) is a powerful and versatile numerical tool used to simulate and analyze fluid flow phenomena in various engineering applications. With the increasing demand for efficient and reliable solutions in industries such as oil and gas, chemical processing, aerospace, and environmental engineering, CFD has emerged as an indispensable tool for understanding and optimizing complex multiphase flow(Calay and Holdø, 2003).

Multiphase flow involves the simultaneous movement of multiple fluid phases, such as liquids, gases, and solids, interacting with each other in a shared space. The interactions can be highly dynamic, with complex phenomena like droplet and bubble formation, phase change, and interfacial instabilities. The accurate prediction of such behaviour is challenging due to the inherent complexities and non-linearities involved.

In the past, traditional experimental methods were predominantly used to study multiphase flow. However, these methods had limitations in capturing the intricate details of the flow due to practical constraints and high costs. CFD, on the other hand, offers a cost-effective and time-efficient alternative by numerically solving the governing equations of fluid dynamics and employing advanced numerical algorithms.

This unique capability of CFD enables engineers and researchers to gain in-depth insights into the flow phenomena, study scenarios that may be difficult or dangerous to reproduce in experiments, and optimize the design and performance of engineering systems (Andersson et al., 2011). Additionally, CFD allows for the exploration of a wide range of operating conditions and geometry configurations, facilitating a comprehensive understanding of the underlying physics.

In this pursuit, state-of-the-art CFD software leverages high-performance computing resources to efficiently solve the complex multiphase flow equations. These simulations involve discretizing the governing equations using finite difference, finite volume, or finite element methods, followed by iterative numerical solvers to converge to the desired solution. The accuracy and reliability of CFD simulations have significantly improved over the years, thanks to advancements in turbulence modelling, multiphase flow models, and grid generation techniques.

This introductory article aims to delve into the fundamentals of CFD as applied to solving complex multiphase flow problems. Governing equations, turbulence models, and numerical methods employed in CFD simulations will be described in this work accurately predict and visualize multiphase flow phenomena. Furthermore, challenges and opportunities in multiphase CFD will be discussed, highlighting its relevance and significance in modern engineering and scientific research.

Through this thesis, the potential of CFD will be demonstrated as a robust tool for unravelling the mysteries of multiphase flow and its potential to drive innovation and advancements in a wide range of industries.

1.6 Machine learning

Machine learning (ML) has emerged as a transformative approach to tackle challenging problems in various fields, and fluid dynamics is no exception. In recent years, the combination of machine learning and fluid dynamics has opened up new avenues for understanding and predicting complex multiphase flow phenomena with greater accuracy and efficiency (Vinuesa, Brunton and McKeon, 2023).

Fluid dynamics plays a critical role in numerous engineering and scientific applications, from designing efficient transportation systems to optimizing chemical processes. However, the behaviour of multiphase flows, where multiple fluid phases coexist and interact, presents significant complexities that traditional numerical methods struggle to handle effectively.

Machine learning, a subset of artificial intelligence, offers a data-driven and adaptive approach to model complex systems (A et al., 2020). By leveraging large datasets and powerful computational tools, ML algorithms can learn patterns and relationships hidden within the data, providing valuable insights and predictive capabilities. In the context of fluid dynamics, ML techniques offer a promising avenue to overcome the limitations of conventional numerical simulations and to explore new frontiers in understanding multiphase flow behaviour.

The marriage of machine learning and fluid dynamics offers several advantages. First, ML methods can augment the accuracy of traditional numerical simulations by reducing modelling errors and capturing complex physics that may be challenging to represent using classical approaches. Second, ML can significantly accelerate simulation times by providing surrogate models that approximate the solution space, enabling rapid exploration of various scenarios and design iterations. This expedites the optimization process and enhances engineering productivity.

One of the fundamental challenges in fluid dynamics, especially in multiphase flow, is turbulence modelling. Turbulent flows exhibit chaotic and unpredictable behaviour, making them notoriously difficult to model accurately. Machine learning techniques, such as artificial neural networks and deep learning, have shown promise in predicting turbulent flows with greater fidelity than conventional turbulence models.

Furthermore, ML-based data-driven models can capture the intricate interfacial dynamics between different fluid phases, such as droplet coalescence, bubble breakup, and phase change phenomena (Nguyen et al., 2013). These fine-grained predictions facilitate a deeper understanding of multiphase flow behaviour and enable engineers to optimize system designs for improved performance and efficiency.

In this work, the different machine learning methodologies employed in fluid dynamics, will be described with a focus on complex multiphase flows. The key concepts of ML, including supervised and unsupervised learning will be describing discuss how these methods are adapted and applied to fluid dynamics

problems. Additionally, the challenges and opportunities of integrating machine learning with fluid dynamics simulations will be described, and the most recent advances in this exciting field will be highlighted.

By embracing the synergies between machine learning and fluid dynamics, a future emerges where simulations become more accurate, efficient, and accessible, revolutionizing our understanding of multiphase flow and driving innovations across diverse industries.

1.7 Computational Fluid dynamics and machine learning in curved pipes.

The history of Computational Fluid Dynamics (CFD) in the context of serpentine pipes is envisioned, a testament to the remarkable progress made in understanding and predicting flow behaviour.

Early Approaches: In the early stages of serpentine pipe research, experimental methods and empirical correlations were the primary means of understanding flow behaviour. However, as computational capabilities advanced, researchers began exploring numerical methods to gain deeper insights into the complex flow phenomena occurring in serpentine pipes.

Emergence of CFD: The emergence of CFD in the 1970s provided a paradigm shift in fluid dynamics research. It offered the ability to simulate fluid flow, heat transfer, and other relevant phenomena within serpentine pipes using numerical techniques. Initially, the computational power was limited, and simplified models were used to approximate the flow behaviour. However, as computing power increased, CFD models became more accurate and capable of capturing intricate flow details.

Validation and Verification: As CFD models gained popularity, efforts were directed towards validating and verifying their accuracy. Researchers conducted benchmark experiments and compared the CFD predictions with experimental data to ensure the reliability of the numerical simulations. This iterative process led to the refinement and improvement of CFD models specifically tailored for serpentine pipe flows.

Advancements in Turbulence Modelling: Turbulence is a key aspect of fluid flow in serpentine pipes. Over the years, significant advancements have been made in turbulence modelling within CFD simulations. From early models like the Reynolds-Averaged Navier-Stokes (RANS) equations to more sophisticated approaches like Large Eddy Simulation (LES) and Direct Numerical Simulation (DNS), researchers have strived to accurately capture the complex turbulent flow characteristics in serpentine pipes.

Geometry Optimization: CFD has been instrumental in optimizing serpentine pipe geometry for improved performance. Through parametric studies and optimization algorithms, researchers have utilized CFD simulations to identify the optimal pipe configurations that minimize pressure drop, enhance heat transfer, and optimize flow distribution. These insights have resulted in improved design guidelines for serpentine pipes in various industrial applications.

Multiphase Flow Modelling: Serpentine pipes often encounter multiphase flow scenarios, such as gas-liquid or liquid-solid flows. CFD has been extended to model and analyze such complex multiphase flows within serpentine pipes. Advanced numerical techniques, such as Volume of Fluid (VOF) and Eulerian-Lagrange methods, have been employed to simulate and understand the intricate interactions between phases within the pipe.

Integration with Machine Learning: In recent years, the integration of CFD with machine learning techniques has emerged as a promising avenue. Machine learning algorithms have been applied to enhance CFD simulations, enabling faster computations, improved accuracy, and automated optimization processes. These advancements have further expanded the capabilities of CFD in analysing and predicting flow behaviour in serpentine pipes.

Future Perspectives: The future of CFD in serpentine pipe research is promising. With ongoing advancements in computing power, numerical algorithms, and experimental techniques, CFD simulations will continue to provide valuable insights into the flow behaviour of serpentine pipes. The integration of artificial intelligence, data-driven modelling, and high-fidelity simulations will further enhance our understanding and prediction capabilities.

In summary, the introduction of CFD has revolutionized the study of flow behaviour in serpentine pipes. From early developments to the present day, CFD has evolved as a powerful tool for analysing, optimizing, and predicting flow phenomena. Its integration with experimental validation, turbulence modelling, geometry optimization, multiphase flow simulations, and machine learning techniques has significantly advanced our understanding of serpentine pipe flows, contributing to enhanced performance and efficiency in various industrial applications.

1.8 Structure of the Thesis

The thesis is structured in a way to comprehensively investigate the application of (CFD) and machine learning techniques (MLT) in the analysis of two phase flow phenomena, specifically focusing on serpentine pipes. This introduction chapter provides a detailed overview of the research topic, highlighting the growing importance of understanding and predicting multiphase flow behaviour in serpentine pipes due to their significance in various process engineering applications. The motivation for employing both CFD and machine learning methods is outlined, emphasizing their potential to complement each other and enhance the accuracy and efficiency of multiphase flow simulations.

Chapter 2 describes the aim and objectives of the research, setting the groundwork for the investigation.

To establish a solid theoretical foundation, Chapter 3 provides an extensive literature review. It covers a wide range of topics, including the historical background of multiphase flow analysis, the significance of serpentine pipes in process engineering, traditional CFD techniques, and the recent advancements in machine learning methods applied to fluid dynamics. This literature review aids in identifying the gaps and challenges in the existing research, paving the way for innovative contributions in the subsequent chapters.

Chapter 4 focuses on the experimental background and data collection for the research. It describes the experimental setup and data sources used to obtain relevant multiphase flow data in serpentine pipes. The chapter provides crucial

information about the physical system under investigation, including flow parameters, boundary conditions, and fluid properties. This data is essential for both model development and validation.

Chapter 5 presents a detailed account of the data pre-processing procedures for machine learning applications. Data cleaning, feature selection, and normalization techniques are explained to ensure the quality and suitability of the dataset for machine learning algorithms. This chapter ensures that the data used for training and validation are accurate and representative of real-world conditions.

The developed CFD model is the focus of Chapter 6. The model architecture and governing equations are described, highlighting any modifications or enhancements made to incorporate multiphase flow characteristics. The validation process against benchmark cases is presented, demonstrating the accuracy and reliability of the model.

Chapter 7 is dedicated to the results and discussions. It showcases the findings obtained from both the ML and CFD simulations. The results are thoroughly analysed, and comparisons are made with traditional CFD simulations to assess the added benefits of machine learning. The discussions leads into the implications of the findings, discussing the model's strengths and limitations in predicting complex multiphase flow phenomena.

Finally, Chapter 8 concludes the thesis, summarizing the key contributions of the research and their alignment with the stated objectives. It also provides an overview of the research's significance in advancing the understanding of multiphase flow in serpentine pipes. Additionally, potential avenues for future research are proposed, highlighting opportunities for further advancements in the integration of CFD and machine learning methodologies for studying multiphase flow in various engineering applications.

2 Aim and Objective

2.1 Aim

This aim of this work is to develop an integrated process by combining, experimental, numerical modelling and machine learning techniques to predict the two phase flow behaviour in a serpentine pipes efficiently and accurately.

2.2 Objectives

The following objectives were set to achieve the aim of the project:

Objective 1: A literature review is conducted to identify the current state of the research using the machine learning concept in serpentine pipes and the numerical modelling of the flow physics.

Objective 2: The experimental data available from the Cranfield process engineering lab are processed to train the machine learning models. For the given diameter of the pipe, pressure drop, gas void fraction and liquid film thickness is available for different superficial gas and water velocities.

Objective 3: Different regression models are trained, and the results are compared.

Objective 4: Three-dimensional models are created for different serpentine pipe diameters.

Objective 5: The base model is simulated via CFD and validated against experimental data.

Objective 6: CFD simulation data is used to train the machine learning algorithm to predict the intermediate data needed to optimize the geometry of the serpentine pipes with improved accuracy and efficiency.

3 Literature Review

3.1 Overview

The primary focus of the current investigation is to address specific objectives related to the study of two-phase vertical serpentine pipes. While a detailed literature review encompassing machine learning and CFD in this context is important, it falls outside the scope of this study. Consequently, the subsequent sections of this chapter concentrate on the topics directly relevant to the research goals. These sections explore the significance and practical applications of machine learning in understanding serpentine flow characteristics and optimizing geometry.

3.2 Flow regimes

Flow regimes in multiphase flow refer to distinct patterns or behaviours that fluids (such as gas and liquid) exhibit when flowing together in a pipe or other conduit. These regimes are defined by the relative proportions of the phases, their velocities, and the interactions between them.

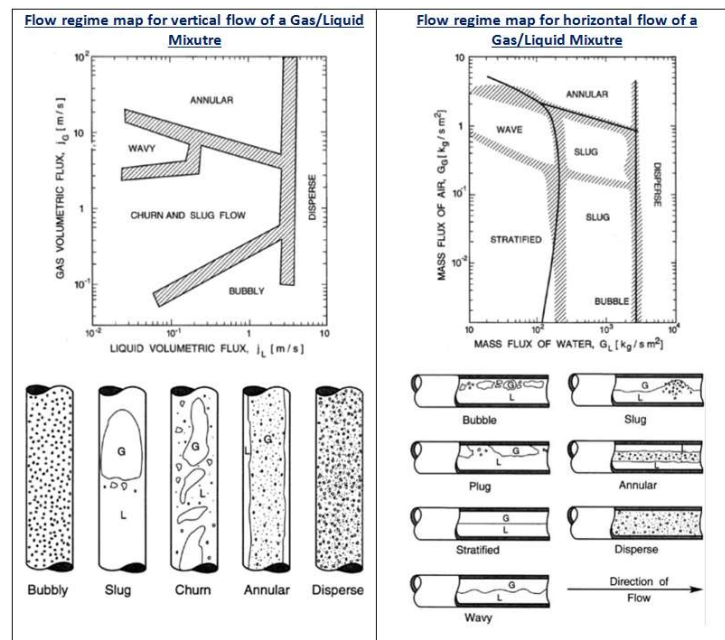


Figure 3-1: Flow Regimes ((Islam Fetoui, 2005.)

Understanding flow regimes is essential for various engineering applications, such as oil and gas production, chemical processes, and nuclear engineering. Below is an overview of some common multiphase flow regimes: Figure 3-1 (Islam Fetoui, 2005.) shows the flow regimes map.

Bubbly Flow: In this regime, small gas bubbles are dispersed in the liquid phase. The liquid phase typically dominates the flow, and the bubbles move with the flow, often rising to the top of the pipe (Shawkat, Ching and Shoukri, 2008). Bubbly flow is common in low-velocity conditions.

Slug Flow: In slug flow, gas bubbles form large, elongated slugs separated by liquid. These slugs move intermittently through the pipe. Slug flow (Kang, Kim and Kim, 2021) can lead to significant pressure fluctuations and can be challenging to control in certain systems.

Stratified Flow: In stratified flow, the two phases remain mostly separate, with the gas usually forming a layer on top of the liquid (Berthelsen and Ytrehus, 2005). The interface between the two phases is well-defined, and the flow can be stable under certain conditions.

Annular Flow: Annular flow occurs when a thin layer of one phase forms a film on the pipe's inner wall, with the other phase (usually gas) flowing in the core of the pipe (van Eckeveld et al., 2018). This regime is often characterized by a continuous liquid film along the pipe's perimeter.

Churn Flow: Churn flow is a combination of stratified and slug flows (Vaze and Banerjee, 2011). It involves the periodic formation and breakdown of liquid slugs within a continuous liquid film.

Wavy Flow: In wavy flow, the two phases travel in waves along the pipe's cross-section (Lini~t and Lopez, 1997). This regime is usually observed in pipes with small diameters or during transient conditions.

Droplet Flow: In this regime, small liquid droplets are dispersed in the gas phase (Cotton J. S. Ph.D., Shoukri M. Ph.D. and Chang J. S. Ph.D., 2001) .

Droplet flow is common in systems with high gas velocities, and the liquid droplets may contribute to heat and mass transfer.

Mist Flow: Similar to droplet flow, but with smaller and more uniformly dispersed liquid droplets in the gas phase (Jiang et al., 2020). Mist flow is common in certain industrial applications.

These flow regimes have different implications for heat transfer, pressure drop, and phase separation, making them important to consider when designing and analyzing multiphase flow systems. Researchers and engineers use experimental data and computational models to predict and control the flow regime transitions based on the operating conditions and equipment design.

3.3 Gravity influence

The influence of gravity on serpentine multiphase flow is a critical factor in various engineering applications. Gravity affects the behavior of both liquid and gas phases as they navigate through the intricate pathways of serpentine pipes or channels (Liu, Chen and Shi, 2012). Understanding this influence is essential for designing efficient systems in fields such as chemical engineering, heat exchangers, and fluid transport. Here's an overview of the key aspects of gravity's impact on serpentine multiphase flow:

Phase Separation: Gravity plays a significant role in phase separation within serpentine pipes (Liu, Chen and Shi, 2012). When there's a significant density difference between the liquid and gas phases, gravity can cause phase segregation. The denser phase tends to settle at the bottom of the pipe, while the lighter phase occupies the upper regions. This phenomenon can affect the distribution of phases in different sections of the serpentine path, leading to variations in flow behavior and heat transfer.

Slug Flow Formation: In serpentine channels with upward sections followed by downward sections (like U-shaped segments), gravity can influence slug flow formation (Cheah, Kevrekidis and Benziger, 2013). As the gas-liquid mixture moves through the channel, gravity can lead to the creation of slugs, where the liquid phase tends to accumulate in the downward segments, while the gas

occupies the upper parts. These slugs can have significant implications for pressure drop and heat transfer.

Flow Resistance: Gravity-induced head differences along the serpentine path can impact flow resistance. The liquid phase may experience increased pressure in the upward sections due to gravity opposing the flow, while the gas phase may experience less resistance. This uneven distribution of resistance can lead to variations in flow rates and pressure drops.

Phase Distribution: Gravity can influence the distribution of phases within each segment of the serpentine flow path. In downward segments, gravity can promote liquid accumulation, leading to stratified or annular flow patterns. In upward segments, gravity may work against the liquid phase, leading to more dispersed or bubbly flow patterns.

Flow Stability: Gravity can impact the overall stability of the multiphase flow in serpentine channels. The accumulation of liquid in certain segments can lead to flow instabilities, slug coalescence, or even flow reversal if the gas phase cannot overcome the gravitational forces.

To effectively model and design serpentine multiphase flow systems, it's crucial to account for gravity's influence.

3.4 Computational Fluid dynamics

CFD has emerged as a powerful tool for analyzing and simulating complex fluid flow phenomena, offering valuable insights into the behavior of multiphase systems (Calay and Holdø, 2003). Over the past several decades, CFD has played a pivotal role in advancing our understanding of two-phase flow, enabling researchers and engineers to study the intricate interactions between different fluid phases in various industrial and scientific applications.

The history of CFD in solving two-phase flow problems can be traced back to the mid-20th century when the foundations of fluid dynamics and numerical techniques were established. Early efforts focused on single-phase flows, but the growing demand for accurate predictions in multiphase environments prompted

researchers to extend CFD methods to handle the challenges posed by interfacial interactions, phase transitions, and complex flow regimes. The 1970s and 1980s witnessed significant developments in numerical algorithms and computing resources, leading to the expansion of CFD capabilities to tackle more intricate fluid flow problems, including those involving two-phase flow phenomena.

As computational resources continued to advance, the 1990s marked a turning point for CFD in two-phase flow research (Lun, Calay and Holdo, 1996). Researchers started to apply CFD methods to study multiphase flows in various engineering disciplines, such as chemical engineering, nuclear engineering, and thermal systems. This era saw the emergence of specialized CFD codes and techniques tailored to address the unique challenges posed by two-phase flow, such as capturing phase distribution, interfacial dynamics, and phase change phenomena.

The integration of CFD with experimental data and the development of validated models further solidified its role in solving two-phase flow problems. Today, CFD has become an indispensable tool for predicting and optimizing the behavior of multiphase systems, providing a deeper understanding of flow regimes, heat transfer, pressure drop, and phase distribution. The historical progression of CFD in tackling two-phase flow challenges underscores its importance in shaping our ability to analyze and design complex fluid systems, driving innovation across a wide range of industries.

3.5 Machine learning techniques

All machine learning methods are fundamentally based on the concept of optimization. This encompasses various aspects, including the framework for selecting models and regression techniques that strive to offer models that are both concise and interpretable when analyzing data. In the realm of regression techniques, the foundational approach is curve fitting, which involves shaping data to a curve. Additionally, more advanced methods, such as polynomial and exponential fitting, arise from the resolution of linear systems.

$$Ax = b$$

Equation 3-1

Within the realm of optimization strategies, a primary focus in data science is the ability to discern whether a given model has encountered the issues of overfitting or underfitting when dealing with the data. The utilization of cross-validation techniques is paramount in the evaluation of any proposed model. This involves the division of a provided dataset into distinct sets: a training set, a validation set, and a withhold set. The model is constructed using the training and validation data, subsequently being subjected to testing using the withhold set for the final assessment.

Prototypical behaviour of Over and underfitting of data:

- When dealing with overfitting, enhancing the model complexity or extending the training epochs (iterations) results in a more pronounced reduction of error on the training data, but this typically leads to an increase in the error observed on the withheld data shown in Figure 3-2 (Steven L. Brunton, 2022)

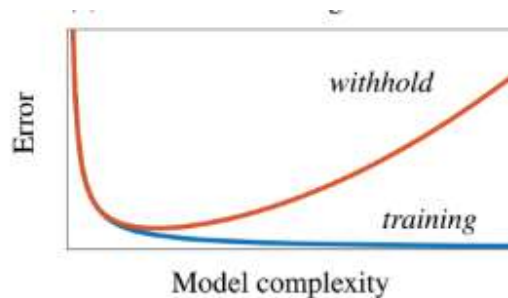


Figure 3-2: Over fitting model complexity (Steven L. Brunton, 2022)

In cases of underfitting, the constrained model complexity places a limit on the error performance as shown in Figure 3-3 (Steven L. Brunton, 2022). These fundamental graphical representations are prevalent in the field of data science and play a pivotal role in the evaluation of a model.

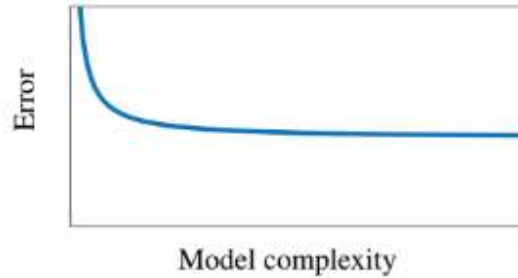


Figure 3-3: Over under fitting model complexity (Steven L. Brunton, 2022)

Classic Curve fitting: Curve fitting stands as a fundamental cornerstone within the realm of data science. In astronomical contexts, curve fitting demonstrated remarkable efficacy, especially when dealing with the straightforward elliptical orbits displayed by planets and comets. From a mathematical perspective, the approach will emphasize throughout this study is regression. Similar to curve fitting, regression seeks to infer the relationships among variables through an array of statistical techniques. In particular, this involves exploring the overarching correlation between independent variables (X), dependent variables (Y), and certain undisclosed parameters represented by β :

$$Y = f(X, \beta) \quad \text{Equation 3-2}$$

In this context, where the function $f(\cdot)$ represents the regression function, and the parameters β serve as indicators of how well this function aligns with the data. The success of the optimization process hinges on the precise definition of a function to be optimized, which is commonly referred to as the objective function.

Least- Squares Fitting Methods: Consider classic least squares polynomial fitting for characterizing trends in data, the concept is straightforward and use a simple function to describe a trend by minimizing the sum square error between the selected function $f(\cdot)$

Classical curve fitting formulated as:

$$Ax = b \quad \text{Equation 3-3}$$

where A is a matrix, x is a vector of coefficients, and b is a vector of observed values.

Consider a set of n data points:

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots (x_n, y_n) \quad \text{Equation 3-4}$$

Where, numbers are the points in the set

Best fit line through these points. Approximate function

$$f(x) = \beta_1 x + \beta_2 \quad \text{Equation 3-5}$$

where the constants β_1 and β_2 are the parameters of the vector β .

Linear regression model

$$Y = f(A, \beta) = \beta_1 X + \beta_2 \quad \text{Equation 3-6}$$

Applying machine learning to study fluid mechanics: (Brunton, 2021) The application of machine learning in the study of fluid mechanics involves a systematic process divided into five stages:

1. Formulating the problem to be modeled.
2. Collecting and refining relevant data to inform the model.
3. Selecting an appropriate architecture to represent the model.
4. Designing a loss function to evaluate the model's performance.
5. Choosing and implementing an optimization algorithm to train the model.

Data science is the skillful practice of posing and answering questions using data. The specific subfield of machine learning focuses on leveraging historical data to construct models capable of automatically answering questions, often in real-

time, when presented with new data. The selection of a suitable system to model is crucial, with importance placed on it being both significant and feasible.

The nature of the problem, particularly how the model should work given specific inputs, leads to a classification of machine learning algorithms into three major categories as shown in Figure 3-4 (Sharma et al., 2023):

- supervised,
- unsupervised, and
- reinforcement learning.

Unsupervised learning is particularly challenging since it involves deriving structure solely from input data, commonly referred to as data mining. This approach results in the formation of clusters, which can then be used as proxy labels to classify new data. Additionally, if the data structure is assumed to be continuously adaptive, it is often considered as an embedding or dimensionality reduction task. Supervised learning tasks can be further classified into techniques like Principal Component Analysis (PCA) or Proper Orthogonal Decomposition (POD), which require continuous dimension reduction. Reinforcement learning, a significant branch of machine learning, involves agents learning to make control decisions in interactions with an environment to achieve high-level objectives (Hou et al., 2017). A classic example is learning to play games such as chess. Principal Component Analysis (PCA) is a linear technique that identifies the most significant orthogonal components, or "principal components," in the data. These components are ordered by the amount of variance they explain, allowing for dimensionality reduction while retaining as much information as possible.

Proper Orthogonal Decomposition (POD), also known as the Karhunen-Loève expansion, is a similar technique often used in fluid dynamics and applied mathematics (Cazemier, Verstappen and Veldman, 1998). It identifies the spatial modes of maximum energy (or variance) within a dataset. POD is particularly relevant in analyzing and reducing the dimensionality of fluid flow data, which can

be essential for understanding dominant flow structures or modes in complex systems.

Reinforcement learning, on the other hand, is a distinct branch of machine learning that involves an agent learning to make sequential decisions to maximize a cumulative reward signal. Unlike supervised learning, where the model is provided labeled examples, and unsupervised learning, where the model extracts patterns without explicit guidance, reinforcement learning learns through interaction with an environment. The agent takes actions, receives feedback (rewards), and uses this information to improve its decision-making over time. This type of learning is essential in various scenarios, such as game playing, robotics, and autonomous systems, where agents must learn to make effective decisions based on continuous interaction with their surroundings.

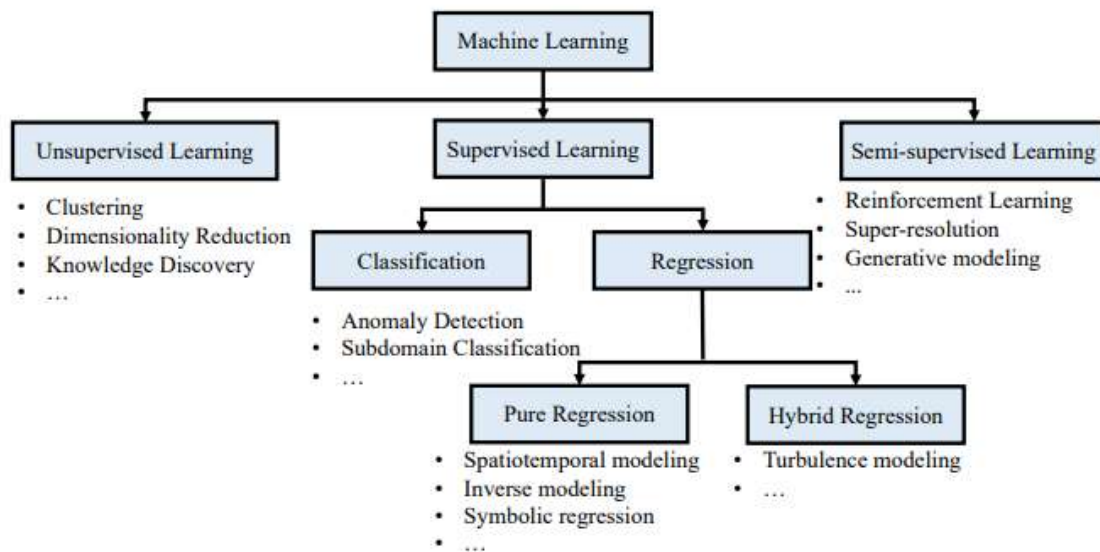


Figure 3-4: Machine learning application in fluid dynamics (Sharma et al., 2023)

Physics-Informed Machine Learning (PIML): Fluid mechanics represents a data-rich domain founded on fundamental physical principles that encapsulate conservation laws as shown in Figure 3-5 (Sharma et al., 2023). Sole reliance on data-driven methods might overlook these underlying theories, potentially disregarding essential physical constraints and conservation principles. Consequently, tasks within fluid mechanics serve as an ideal platform for innovating novel PIML algorithms. Although PIML advancements have brought

significant clarity to fluid flow modeling, numerous unanswered questions and opportunities persist. Each PIML model has thus far been showcased solely on specific problem configurations and datasets, leaving the benchmarking of existing PIML techniques as an ongoing challenge that must be addressed as machine learning within fluid mechanics continues to evolve.

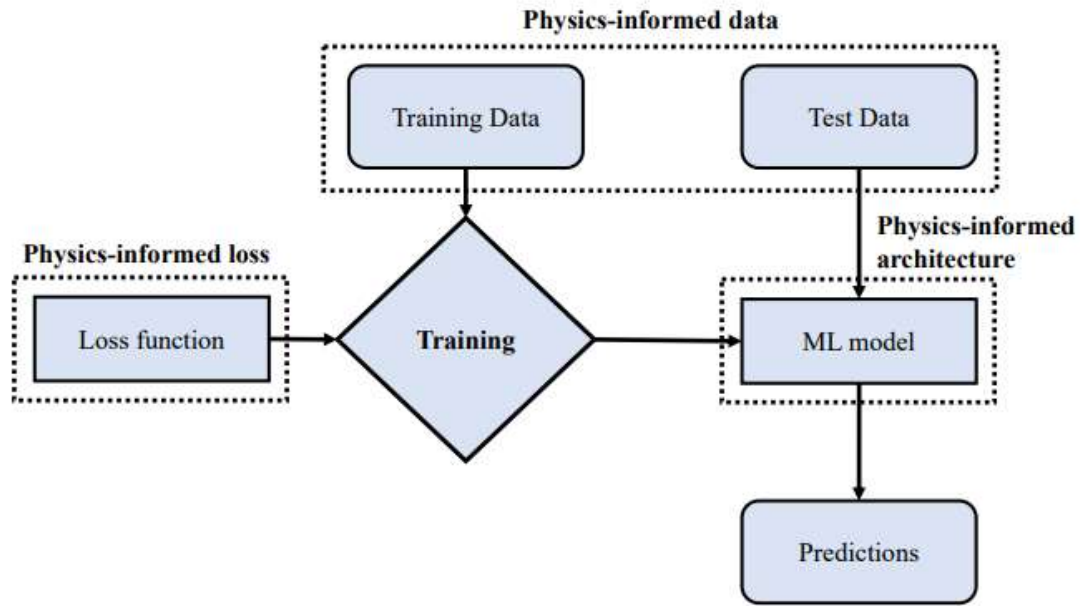


Figure 3-5: Different methods for embedding physics domain knowledge into a supervised learning framework (Sharma et al., 2023)

3.6 Literature Review Papers

This is the base paper of this research work (Yeung, 2009) the objective of this research being to investigate the impact of 180° bends on the behaviour of gas-liquid two-phase flows in large diameter downward and upward pipes, particularly focusing on the influence of serpentine configurations on flow structure, void distribution, and liquid film profiles along these sections. Experimental work has previously been performed in the Process System Laboratory at Cranfield University, and those experimental data were used for the current work. Such study is crucial due to its relevance to industrial processes, such as heat exchangers and oil refinery units, where serpentine patterns are commonly employed.

The research paper revealed that significant effects of both top and bottom bends on flow behaviour, with the extent of impact varying depending on the flow rates of air and water. For higher flow rates, the bends created flow maldistribution in adjacent straight sections, leading to the accumulation of water on the outer zone of the bend and air on the inner zone, resulting from the centrifugal force generated during flow negotiation around the bend.

In-depth analysis of film thickness development along the downward and upward sections demonstrated distinct liquid film behaviour near the bends, attributed to the centrifugal force exerted by the bends. Additionally, examination of the power spectral density (PSD) showed noticeable differences in shape between the PSD of sections adjacent to the bends and those located further away.

Furthermore, flow regime maps were constructed, indicating the presence of unstable flows, particularly along the upward section for low gas and water flow rates. A unique phenomenon was identified where the lower bend experienced periodic blockage by the liquid followed by expulsion due to accumulated air.

A comparison between the research findings and existing theoretical correlations from the literature showed some discrepancies. As a result, updated correlations were introduced, aligning well with the data obtained from the current study and previous works.

The contributions of this research lie in the detailed exploration of the effects of serpentine configurations on large diameter pipes, providing valuable insights for the design and operation of pipelines in industrial applications. The results presented in this thesis underscore the importance of understanding the intricate interactions between bends and flow patterns in order to mitigate flow-related challenges, such as maldistribution, coking, and overheating, thereby reducing operating costs and enhancing production efficiency.

3.7 Literature findings

The performance of the pipe largely depends on the flow rates, pipe radius (a), pipe length (l), curvature radius (c) and hydraulic characterises. The repeated curved sections produce secondary flow which benefits the mixing and heat

transfer with cost of high pressure drop. (Ciofalo and Di Liberto, 2015) found when the frictional velocity Reynold number exceeds the critical value the flow is unsteady, the curvature effect is captured by varying the Reynold number for different flow conditions, the correlations derived by normalising the straight pipe using multiplier. The heat transfer is linear function of Dean number and frictional are linear with the curvature ratio d (pipe radius to curvature radius) (Yang and Sung Fa Chang, 1993). Friction Coefficient and Nusselt Number would be used as a non-dimensional parameter to simplify the curvature effects and use the straight pipe results to optimise the geometry with the limitations of flow Reynold number. The pipes can be positioned horizontal and vertical based on the application, and it will influence the performance as well.

In (Ohn and Ustice, 1911.) 1911 experimentally studied the flow water in curved pipes, which is small diameter. The author observed that a slight sinuosity of the pipe lessened the flow in coiled pipes. In vertical pipes gravity plays a major role in the phase separation, pressure distribution and pressure drop due to its buoyancy effects. The liquid will holdup at the bottom and the air trap will be trapped at the top, after each bend the flow pattern will be different. The study says (Xin, Dong and Ebadian, 1996) when two phases or more than two phase flowing in the curved pipe the flow behaviour changes drastically in terms of the flow regimes

In study (Naphon and Wongwises, 2006a) revealed the important insights into the heat transfer and flow characteristics of various types of curved tubes, offering valuable information for enhancing our understanding of these complex systems. The investigation categorizes curved tubes into three distinct groups based on their curvature configurations. For helical coil tubes, a substantial body of theoretical and experimental research has been conducted, particularly concerning single-phase heat transfer and flow characteristics. However, there remains a significant gap in the knowledge regarding two-phase heat transfer behavior in this configuration.

Similarly, (Kaboodvand, Majidi and Jannesari, 2021) concludes the examination of spiral coil tubes has been limited, with only a handful of papers addressing certain aspects of this type of tube. This lack of comprehensive understanding underscores the need for further research in this area.

On the other hand, the study (Naphon and Wongwises, 2006b) identifies that other types of curved tubes have been more extensively studied, with a multitude of works focusing on single-phase and two-phase heat transfer characteristics.

Additionally, (Naphon and Wongwises, 2006b) concludes a limited number of papers have explored the effects of combined active and passive methods to enhance heat transfer rates and manage pressure drop. This observation suggests a potential avenue for future research, highlighting an area where the integration of innovative techniques could yield significant improvements.

Importantly, the study highlights the relative scarcity of research specifically focused on spirally coiled tubes. Given this dearth of information, there is a clear opportunity for the current research work to contribute to this important field. Investigating multiphase flow characteristics within serpentine pipes, building upon the existing knowledge gaps, and potentially exploring novel methods to enhance heat transfer and manage pressure drop could significantly advance our understanding of curved tube systems and lead to practical applications in various engineering contexts.

(AL-Dogail and Gajbhiye, 2021)'s work involves comprehensive experiments on two-phase flow in a horizontal pipe to assess the impact of density, viscosity, and surface tension on flow regime and pressure drop. A variety of conditions were tested, with the use of different liquids and gases at varying velocities. The experiments aimed to identify the flow patterns and transitions between them. The collected data was then used to create flow regime maps, analyze the boundaries, and compare the results with existing models. This dataset is crucial for understanding the effects of fluid properties on horizontal two-phase flow and serves as a valuable resource for multiphase flow modeling and validation in the field. Similar research studies for vertical pipes are missing.

The research from (Falcone et al., 2001) provides a comprehensive review and assessment of the development, evaluation, and use of Multiphase Flow Meters (MFM) systems in the Oil & Gas industry over the last decade. It highlights the challenges and limitations of various metering systems designed to measure the flow rates of gas, oil, and water in three-phase flows. Both established and novel technologies are considered within this context, and the operational and economic advantages or shortcomings of these technologies are evaluated from an operator's perspective.

The (Falcone et al., 2001) emphasizes the essential role of MFM systems in exploiting marginal fields and explores how these systems are becoming increasingly crucial in areas such as flow assurance, deep water developments, downhole/seabed separation systems, and wet gas fields. The work discusses the practical reliability, accuracy, and lessons learned from the available technology, acknowledging the ongoing need for further research and development to address the unsolved challenges in this field.

The conclusions highlight form (Falcone et al., 2001) the evolution of MFM as an established discipline in the Oil & Gas industry. While a generic solution to MFM remains elusive, specific methodologies are emerging for particular applications. The paper underscores the importance of operator involvement and collaboration with MFM manufacturers during the pre-engineering phase of MFM applications to ensure the quality of results. The anticipated increase in demand for accurate MFM solutions, especially in deepwater and marginal field developments, calls for continued technical/scientific contributions and stronger communication between the industry and governmental organizations worldwide. This work sheds light on the path forward for the development of the next generation of MFM devices to meet evolving industry needs.

The research carried out by (Ramprasad, Mallikarjuna and Bhatta, 2021) summarized here highlights the substantial impact of fluid flow in asymmetric convergent channels, particularly under the influence of magnetic fields, with significant implications for engineering and biological domains such as chemical and food industries, as well as blood flow in capillaries and arteries.

The study utilizes numerical methods to solve the governing equations for fluid and particle transport, including heat generation effects. Rigorous tests, including grid independence and result validation, have been conducted to ensure the reliability of the computational outcomes, which have demonstrated good agreement with published results. The investigation provides valuable insights into the behavior of various parameters on fluid and particle velocity, temperature profiles, skin friction, and Nusselt number through graphs, tables, and linear regression analysis. The study pioneers the exploration of the combined effects of cross-flow Reynolds number and magnetic fields in the context of convergent channels, shedding light on their significance in terms of drag coefficient, heat transfer rate, and overall flow characteristics. The findings reveal trends such as reduced fluid velocity with increased magnetic parameter, Grashof number, and Reynolds number, while an increase in fluid temperature is observed with higher Prandtl number and heat source parameter values.

(Mazumder, 2012) performed thorough CFD analysis of single-phase and two-phase flows in a 90-degree elbow with a 12.7 mm diameter. They investigated flow behavior at multiple locations, varying air and water velocities. Pressure and velocity profiles were assessed, revealing that pressure drop behaviors were consistent between single-phase and multiphase flows. Comparisons with empirical models showed good agreement, validating their approach.

The above findings offer valuable insights into flow dynamics in curved geometries. The current research work predicting water and air behavior in serpentine pipes using machine learning and CFD modeling can extend these insights to more complex systems. By leveraging machine learning, it can use develop predictive models for serpentine pipes, enhancing our understanding of curvature effects on flow behavior and pressure drop. Validating the models with CFD simulations will bolster the practical application of my thesis.

The work from (Kendoush and Ai-Khatib, 1994) close to the research, where the experimental setup involves a comprehensive system, including an air compressor, reservoir tank, air-water mixer, and a test section composed of a 3.8cm-I.D, 4-m-long perspex tube. Various measurements, such as absolute and

differential pressure readings, void fraction (α) through neutron emission, and water and air flow rates, are obtained. The flow regime map is created based on direct visual observations, utilizing phase superficial velocities of water and air as coordinates.

The study offers valuable insights into the transition boundaries between different flow regimes, aligning with previous research findings from (Dukler and Taitel, 1976). Notably, the comparison reveals agreements and discrepancies in the interpretation of flow patterns under varying conditions. The research emphasizes the significance of understanding flow regimes in vertical downward two-phase flow, especially for applications in nuclear, power, and petrochemical industries. Accurate prediction of flow patterns is essential for precise calculations of factors such as pressure drop, pumping power, heat transfer, and mass transfer, crucial for engineering design.

The study identifies potential areas for future research, such as examining the effect of tube diameter on flow pattern maps for larger-diameter tubes and developing innovative flow pattern sensors based on statistical techniques. The utilization of such sensors could eliminate the need for transparent tubes and low-pressure testing, offering practical advantages for industrial applications, particularly when employing neutron sources for void fraction measurement. This work addresses the statistical error found in the experiment as well.

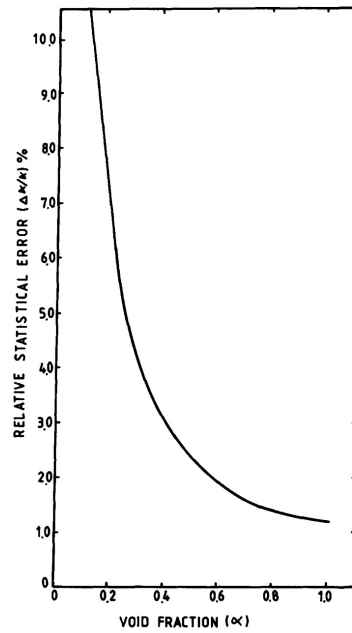


Figure 3-6: Statistical error in void fraction (Kendoush and Ai-Khatab, 1994)

(Roul and Dash, 2012) research investigates the pressure drops in two-phase flow within horizontal pipes, specifically focusing on air-water flows. The numerical approach employs the Eulerian-Eulerian model, simulating various operating conditions covering a range of gas and liquid superficial velocities. Local pressure drops are determined using extrapolation from computed upstream and downstream linearized pressure profiles at the orifice section. Validation of the numerical results against experimental data from the literature demonstrates a good agreement, confirming the reliability of the simulations. Figure 3-6 (Kendoush and Ai-Khatab, 1994) shows the liquid fraction error trend with minimal void fraction value the error in the calculation is statistically high.

The study also takes into account the effect of orifice geometry by considering different inner diameters, area ratios, and thicknesses of the orifice plates. The findings contribute valuable insights into the behavior of two-phase flow pressure losses through orifices, offering a comprehensive understanding of the impact of geometric variations on the system. Furthermore, the conclusions of the research are based on a separate experimental study of downward two-phase flow in an isothermal air-water system. This experimental analysis identifies three distinct flow regimes: bubbly, slug, and annular, demonstrating the complexity of flow

pattern transitions. Importantly, the study establishes that the slip ratio is less than unity, indicating the significance of interphase interactions. Overall, this research combines numerical simulations, experimental validation, and flow regime analysis to enhance our understanding of pressure drops in two-phase flow systems. It highlights the importance of considering orifice geometry and provides valuable insights into the behavior of different flow patterns, contributing to the advancement of knowledge in the field of fluid dynamics.

(Roul and Dash, 2012) provides the significance of heat transfer enhancement techniques for improving the performance of heat exchangers. These techniques are categorized into two groups: active and passive. Active techniques rely on external forces like electric fields or vibrations, while passive techniques involve fluid additives or specific surface geometries. Curved tubes are a widely utilized passive heat transfer enhancement method in various heat transfer applications. The review focuses on heat transfer and flow characteristics in single-phase and two-phase flow within curved tubes, categorizing them into three main groups based on tube curvature: helically coiled tubes, spirally coiled tubes, and other coiled tubes. For helical coil tubes, the review highlights numerous theoretical and experimental works related to single-phase heat transfer and both single-phase and two-phase flow characteristics. However, the study points out that there is limited reporting on two-phase heat transfer characteristics in this configuration. This work from (Srinivasan and Christensen, 1992) says the spiral coil tubes have relatively fewer published papers, with only one providing the correlation of in-tube heat transfer coefficients, and none discussing flow characteristics and pressure drop in detail. For other curved tubes, the review notes that there is extensive coverage of single-phase and two-phase heat transfer characteristics, with a significant number of studies on single-phase heat transfer. However, only one work has reported on two-phase flow characteristics in this category. Additionally, some papers have discussed the combined effects of active and passive methods on heat transfer rate and pressure drop enhancement. The review highlights the overall limited research on certain types of curved tubes, particularly spirally coiled tubes, indicating potential areas for further exploration and research.

The study by (Eyo et al., 2021) introduces a new method for objective flow regime identification using conductance data and kernel techniques in a horizontal annular test section. The approach achieves over 90% accuracy in real-time identification against visual observations by an expert. It offers 2-D visualizations of conductance data for tracking flow regime transitions and addresses the need for objective tools in opaque channels or uncertain regime transitions. This innovative method has broad implications for improving performance and safety in diverse industrial applications.

This research focuses on utilizing artificial intelligence (AI) models, including traditional machine learning algorithms and deep learning architectures, to classify two-phase flow patterns in pipes. (Zhang et al., 2011) concludes the flow patterns are influenced by various physical variables such as gas and liquid superficial velocities, angle of inclination, diameter, viscosity, density, and superficial tension. These patterns are critical in industrial applications, and accurately predicting them can optimize processes and enhance safety.

The primary objective of the comprehensive research study from (Hussein, Ahmed and Habeeb, 2022) was to delve into the intricate dynamics of two-phase flow, specifically focusing on the flow patterns of a liquid-liquid mixture within a rectangular channel. This research was uniquely designed to address the complex interactions that occur when two different fluids, kerosene and water in this case, coexist and flow together in a confined environment. Both experimental and numerical approaches were diligently employed to gain a thorough understanding of the phenomenon.

For the experimental phase, (Hussein, Ahmed and Habeeb, 2022) constructed specialized rig was meticulously constructed to create the necessary conditions for the study. The rectangular horizontal smooth channel was used as the primary conduit, and three different lengths were chosen for investigation: 1, 2, and 3 meters. Additionally, the channel heights varied, being set at 5, 7.5, and 10 cm, respectively. These dimensions were carefully selected to capture a comprehensive range of scenarios and interactions that can occur within such systems.

Two distinct inlet methods were employed, each contributing valuable insights into the flow behavior. In the first method, kerosene was introduced from the top of the water inlet, mimicking a particular flow configuration. The second method involved introducing kerosene from the channel's center, which resulted in a different arrangement with water both above and below the kerosene, thereby exploring diverse flow patterns. The flow conditions for all tests involved an input water velocity of 0.2 m/s and an input kerosene velocity of 0.1 m/s, creating a consistent framework for comparison. The numerical aspect of the study was conducted using advanced computational fluid dynamics (CFD) simulations facilitated by the ANSYS Fluent software package. The analysis incorporated two primary modeling techniques: the volume of fluid (VOF) model and the mixture multiphase flow model. These computational methods allowed for a detailed exploration of the flow patterns and interactions within the rectangular channel under various conditions. The study encompassed an extensive range of scenarios, with more than 226 different cases investigated. The findings from both experimental and numerical approaches were meticulously compared and analyzed. Notably, the results demonstrated a remarkable level of agreement between the experimental data and the numerical simulations, validating the accuracy and reliability of the study's methodologies. As a forward-looking direction, the study acknowledges the potential for further investigations. It identifies several intriguing avenues for future exploration. Specifically, the study proposes the examination of different inlet velocities for oil and water, enabling a more nuanced understanding of flow behavior. Additionally, the inclusion of ribbed channels as a new parameter for investigation promises to shed further light on the complexities of multi-phase flow dynamics in confined geometries.

The study by (Arteaga-Arteaga et al., 2021) uses a substantial dataset, primarily the Shoham database, which contains 5,675 samples representing six distinct flow patterns. In total, 12 databases, encompassing 9,029 data points, are used for training and testing machine learning models. The performance of these models is evaluated against an extensive set of metrics to ensure robust results. Among the machine learning algorithms evaluated, the Extra Trees model demonstrates the highest accuracy, achieving a remarkable 98.8% success rate

in classifying flow patterns. This outcome suggests that Extra Trees is a particularly suitable choice for this specific classification problem. The authors (Arteaga-Arteaga et al., 2021) emphasize that this work provides novel techniques and insights for using AI in predicting flow patterns, and it demonstrates the potential of AI as a powerful tool in industrial processes. The methodology developed here is seen as a valuable contribution to the field, and its potential application is highlighted. In terms of future work, the authors suggest expanding the methodology to incorporate even larger experimental datasets. This step aims to further validate and refine the proposed approach, making it more applicable to a wider range of real-world scenarios. The collaborative efforts of the authors, with detailed contributions from each, underscore the comprehensive nature of this study, ensuring its robustness and reliability.

This paper focuses on predicting multiphase flow rates using machine learning with multi-modal time series sensing data. The data consists of time series differential pressure data generated from a Venturi tube, as well as pressure and temperature data. Three machine learning methods, Deep Neural Network (DNN), Support Vector Machine (SVM), and Convolutional Neural Network (CNN), are implemented and compared in terms of their performance.

The experiments are conducted in a laboratory-scale flow facility, collecting multi-modal multi-phase flow sensing data under various flow conditions. The moving average of the collected instantaneous sensing data is used to train the DNN, SVM, and CNN models. The results show that both DNN and SVM methods can predict liquid and gas flow rates accurately under different flow conditions, such as varying water-to-liquid ratio, gas volume fraction, and flow regimes.

The study specifically investigates the performance of these three machine learning methods for three-phase flow rate prediction (oil/gas/water) based on time series sensing data. Notably, these methods do not require prior flow pattern information. The combination of four sensing parameters, including two differential pressures, standard pressure, and temperature, is used to train the networks in real-time on a pilot-scale multiphase flow facility under diverse flow conditions.

The quantitative analysis reveals that SVM yields the best estimation results when compared to DNN and CNN. The performance of DNN is similar to SVM but slightly degraded. Both DNN and SVM outperform CNN in predicting both liquid and gas flow rates, and the paper provides insights into the potential reasons for this performance difference. This work offers valuable contributions to the field of multiphase flow prediction and highlights the advantages and potential challenges of each machine learning method in this specific context.

In this paper states (Brunton, Noack and Koumoutsakos, 2019) field of fluid mechanics is rapidly advancing, fueled by the vast amounts of data generated from experiments, field measurements, and large-scale simulations across multiple spatiotemporal scales. Machine Learning (ML) provides a robust toolkit for extracting valuable insights from this data, enabling the translation of raw information into a deeper understanding of underlying fluid mechanics. Additionally, ML algorithms have the potential to enhance domain expertise and automate tasks related to flow control and optimization. This comprehensive review delves into the historical context, current developments, and emerging prospects of ML in the realm of fluid mechanics. It explains fundamental ML methodologies and explore their applications in understanding, modeling, optimizing, and controlling fluid flows. A careful examination of the strengths and limitations of these techniques is approached from the perspective of scientific inquiry, considering data as an integral component of modeling, experiments, and simulations. The transformative potential of ML as a potent information-processing framework is highlighted, demonstrating its capacity to augment and potentially revolutionize ongoing research endeavors and industrial applications in fluid mechanics.

The inclusion of Machine Learning (ML) introduces powerful information-processing algorithms crucial for the modeling, optimization, and control of fluid systems. Effective problem-solving in this field necessitates expertise in both ML techniques and a deep understanding of fluid mechanics.

Traditionally, fluid mechanics has grappled with vast datasets, leveraging ML for understanding, prediction, optimization, and control of flow dynamics. Presently,

the rapid progress in ML capabilities offers unprecedented opportunities, with fluid mechanics beginning to unlock the full potential of these robust methods.

Many intricate tasks in fluid mechanics, including reduced-order modeling, shape optimization, and feedback control, can be reformulated as optimization and regression challenges, in which ML plays a significant role. It has the potential to enhance optimization performance, reduce convergence time, facilitate dimensionality reduction, and identify low-dimensional manifolds and discrete flow regimes, significantly enriching our understanding.

Flow control strategies, traditionally following a linear progression from understanding to modeling and then to control, are being influenced by the more flexible ML paradigm. This approach allows for iterative processes between data-driven insights and first-principle approaches, ushering in greater flexibility and innovation in fluid mechanics research.

(Sharma et al., 2023) explores the application of machine learning techniques, particularly for regression problems in multiphase flow scenarios. Three distinct machine learning methods were investigated: Deep Neural Network (DNN), Support Vector Machine (SVM), and Convolutional Neural Network (CNN). The analysis demonstrated that both DNN and SVM methods can achieve favorable prediction accuracy for liquid and gas flow rates across diverse flow conditions. These conditions included variations in water-to-liquid ratio, different gas volume fractions, and diverse flow regimes.

Figure 3-7 (Sharma et al., 2023) shows the schematic illustration of the multiphase flow testing facility. Deep Neural Network (DNN), a powerful supervised machine learning approach, has been extensively used to address various regression problems. Support Vector Machine (SVM), a standard supervised learning model, is typically employed for categorization tasks. Convolutional Neural Network (CNN), known for its capabilities in solving regression problems, has attracted significant attention. The study specifically examined the performance of these machine learning methods in predicting three-phase flow rates (oil/gas/water) based on real-time time series sensing data. The methods utilized a sequence of four sensing parameters (two

differential pressures, standard pressure, and temperature) acquired from a pilot-scale multiphase flow system under diverse flow conditions.

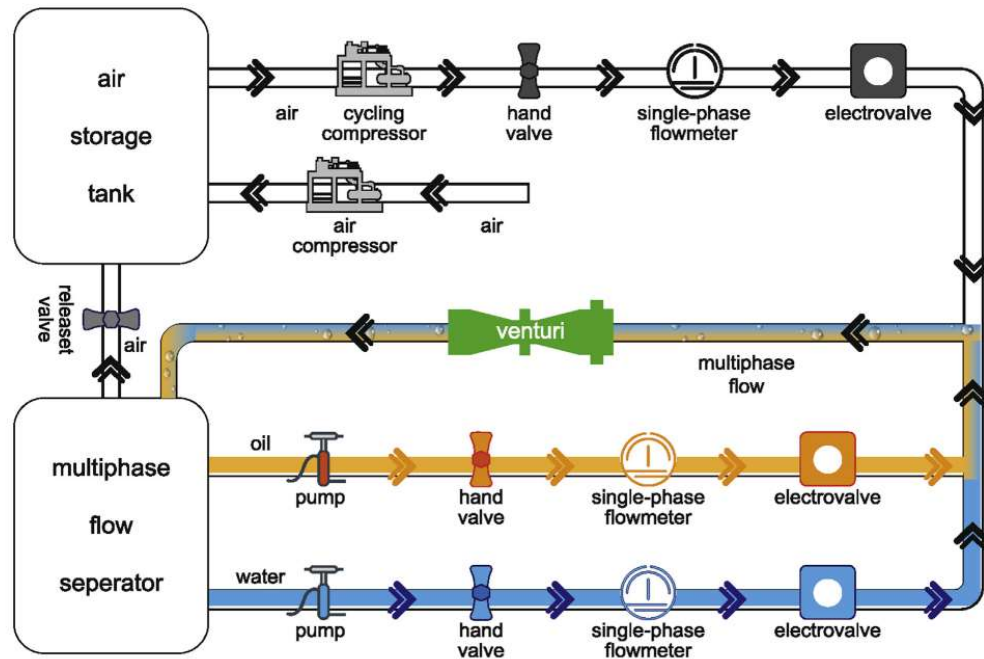


Figure 3-7: Schematic illustration of the multiphase flow testing facility (Sharma et al., 2023)

Notably, the proposed methods did not require prior knowledge of the flow pattern. The results indicated that all three approaches produced reasonable predictions for both liquid and gas flow rates. Quantitative analysis highlighted that SVM outperformed DNN and CNN, providing the most accurate estimation solutions. Although DNN exhibited slightly lower performance compared to SVM, it still performed well. Additionally, both DNN and SVM outperformed CNN in predicting both liquid and gas flow rates. These findings underscore the potential of machine learning, particularly SVM, in predicting flow behavior, paving the way for further advancements in understanding and optimizing complex fluid systems.

The Fukushima accidents underscored the need for passive safety systems (Seo and Lee, 2016.), particularly in scenarios like station blackouts (SBO). Thermosyphons, wickless heat pipes, offer a passive safety solution by leveraging the principles of evaporation and condensation. Despite their potential, few studies focus on heat pipe and thermosyphon simulation. This

study delves into the computational fluid dynamics (CFD) modeling of a 1m thermosyphon under vacuum pressure conditions, a common scenario. The study also includes a transparent thermosyphon experiment to validate hydraulic characteristics. The research aims to contribute to the understanding of thermal and hydraulic behavior in thermosyphons, with potential applications in nuclear and heat pipe engineering.

The geometry used for simulation consists of a 1000mm closed pipe with sections for the evaporator, adiabatic, and condenser, each of the same length. Boundary conditions involve heat flux input to the evaporator's outer wall, cooling of the condenser section, and a defined operating pressure. The CFD simulation employs a two-phase approach with water vapor as the primary phase and liquid water as the secondary phase. A transient analysis with appropriate time steps is conducted, maintaining the Courant number under 2. The simulation uses the SIMPLE algorithm for pressure-velocity coupling, employing first-order upwind schemes for momentum and energy, and Ged-Reconstruct for volume fraction.

Experimental apparatus includes a heat pipe section, heating, and cooling sections. A transparent heat pipe, a vacuum pump, and a working fluid injection component are part of the heat pipe section. Heating is accomplished by a transparent heater connected to a DC power supplier, and the cooling section connects to a chiller for controlling cooling water flow rate and temperature. Experimental steps involve achieving steady-state conditions while monitoring surface temperature and flow patterns. Flow pattern recordings are captured by a high-speed camera.

Simulation results demonstrate that the thermosyphon simulation model reaches a steady state, showcasing a transition from bubble boiling to cap bubble and slug bubbles, ultimately leading to churn flow. Experimental results from the transparent heat pipe align with the simulation, showing similar hydraulic characteristics. Further work involves verifying the thermal characteristics by comparing surface temperatures with experimental data. The study holds

potential for applications in thermal and nuclear engineering, offering valuable insights into the behavior of thermosyphons.

This review focuses on the work (AL-Dogail, 2003) application of Artificial Intelligence (AI) to predict the pressure drop in multi-phase flow in a horizontal pipe. The study gathered extensive experimental data by varying fluid properties such as density, viscosity, and surface tension using a two-phase air/liquid system. Machine learning techniques were employed to develop accurate models for predicting the pressure drop, and dimensionless parameters were utilized to enhance the models' applicability across various design and operational conditions. The results demonstrated that AI methods can achieve high accuracy in predicting pressure drop with minimal inputs, showcasing the potential of AI in understanding complex phenomena like multi-phase flow. The integration of AI with computational fluid dynamics is a promising area of development, particularly for addressing intricate interactions between different phases in fluid flow.

3.8 Literature findings summary

The above research paper served as a reference point to identify potential research opportunities utilizing machine learning regression models for a data-driven approach. This was established based on the comprehensive literature review that was undertaken. It is evident that the utilization of regression models, including linear, polynomial, and potentially other advanced regression techniques, hold significant potential for the analysis of serpentine data. Several research papers have highlighted the efficacy of regression models in capturing the complex relationships and patterns within serpentine data. For instance, in the study by Ramprasad, Mallikarjuna and Bhatta (2021) a linear regression model was successfully employed to predict key flow parameters in an asymmetric non-uniform channel, demonstrating the ability of such models to capture non-linear trends. This finding highlights the importance of exploring various regression methods.

Furthermore, the application of the Volume of Fluid (VOF) method in CFD analysis offers a robust approach for studying the behaviour of fluids in serpentine

pipes. Research by Hussein, Ahmed and Habeeb (2022) has shown that the VOF method effectively captures the intricate interfacial dynamics and phase interactions in complex flow geometries, making it particularly suitable for investigating fluid behavior in serpentine channels. The VOF method's ability to handle multi-phase flows and predict flow patterns within serpentine structures aligns well with the objectives of analyzing serpentine data in this context.

By combining the capabilities of regression models and the VOF method, there is potential for a comprehensive analysis of serpentine data. The literature suggests that this synergistic approach can provide valuable insights into flow patterns, pressure distributions, and other critical parameters in serpentine systems.

In conclusion, the findings from the literature review strongly support the use of regression models and the VOF method as powerful tools for analyzing serpentine data. The versatile nature of regression models allows for the capture of both linear and non-linear relationships, while the VOF method enables detailed simulations of fluid behavior in complex serpentine geometries. Leveraging these techniques in tandem holds great potential for advancing the understanding of serpentine flow dynamics and optimizing the design and operation of serpentine systems in various applications. Further research in this direction, exploring diverse regression models and refining the VOF method's application, will contribute to enhancing knowledge in this important field.

4 Experimental Data

Experimental data were obtained from previous experiments carried out at the process systems engineering laboratory at Cranfield University, UK (Yeung, 2009). An overview of the experimental rig is provided in this section. Void fractions and liquid film thicknesses were obtained separately for different liquid and gas flow conditions.

4.1 Overview of Serpentine rig

The Figure 4-1 shows the test rig which consists of a fluid supply system, metering section, flow separation system and the test section. The diameter of the pipe considered is 101.6mm, the flow loop is 20 meter in length overall. There are four straight pipes connected by 180° bend pipes (bend 1, bend 2 and bend 3). The radius of the bend is 203 mm. The middle two vertical sections are used to collect the raw data, located at 6 meters high. The pipe connected between bend 1 & bend 3 is named the downward pipe. The pipe connected between bend 2 and bend 3 is named the upward pipe.

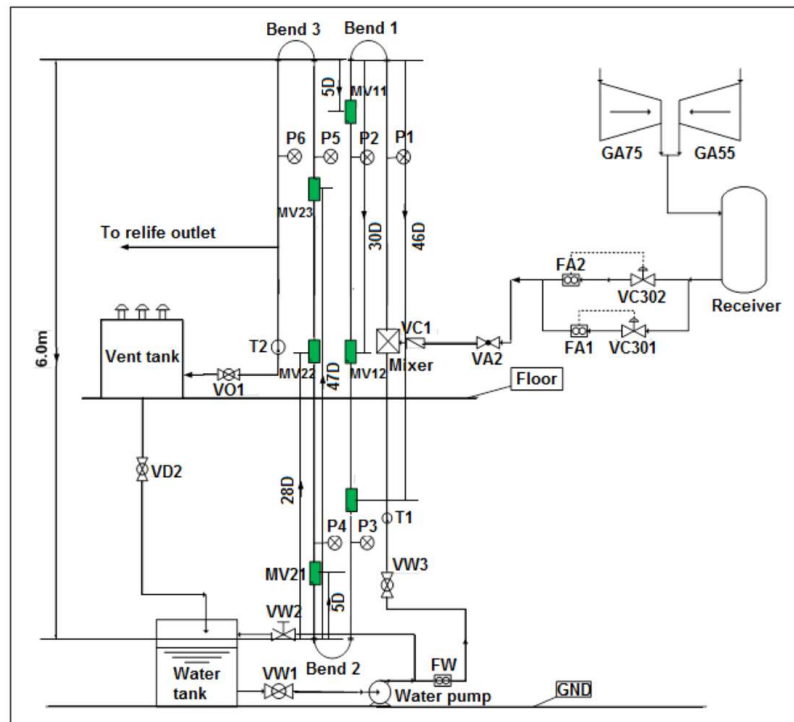


Figure 4-1 Serpentine test rig

Measurement and visual observation (MV) equipment is strategically positioned for specific purposes. The MV setup includes visual observation points along the pipe to visually assess and measure the thickness of the liquid film using film sensors placed at various circumferential locations. Additionally, a wire mesh sensor is employed to measure the gas void fraction. Furthermore, pressure and temperature sensors are installed along the test section to measure pressure and temperature data.

4.2 Flow conditions

4.2.1 Air

Air was supplied from two compressors connected in parallel, ensuring a robust maximum air flow rate of 1200 Sm³/h. To mitigate any undesirable pressure fluctuations, a dedicated 8 m³ volume station was employed to store the compressed air. Prior to utilization, the compressed air was meticulously guided through a series of three filters and a dedicated cooler, eliminating any potential debris and moisture present in the air stream. Furthermore, the air flow was measured and monitored within the precise range of 0 to 1000 Sm³/h, ensuring accurate control and monitoring of the system.

4.2.2 Water

Water was supplied from a water tank, boasting a volumetric capacity of 1.2 m³. The water flow rate was rated at 10 litres per second, operating under a pressure of 6 bar. To ensure accurate monitoring and control of the flow, a high-precision 100 mm electro-magnetic flow meter was employed. Throughout the experimental process, the flow velocity was thoughtfully varied within a wide range, ranging from 0.07 m/s to 1.5 m/s, enabling comprehensive analysis and evaluation.

4.3 Sensors and data acquisition system

In this comprehensive experiment, two air flow meters were thoughtfully incorporated to cover a wide range of air flow rates. The first meter exhibits a range of 0-150 Sm³/h, while the second meter caters to higher flow rates, ranging

from 150-4250 Sm³/h. Similarly, for accurate water flow measurement, a dedicated water flow meter was employed, featuring a range between 0.06-16 litres per second. To capture essential parameters, various sensors were strategically positioned within the experimental setup. A pressure sensor was calibrated to measure pressures spanning from 0-100 pig (pressure in gauge). Furthermore, a temperature sensor was utilized to measure fluid temperatures ranging from 0-100 degrees Celsius. To gain insights into the void fraction of the fluid, a wire mesh sensor was employed, allowing for a measurement within the range of 0-100%. Additionally, the experimental setup involved the use of liquid film thickness probes, ensuring measurements within the range of 0-3 mm. For efficient data acquisition and management, a robust system was implemented. The data acquisition system comprised of the DeltaV, Labview, and WMS systems. The DeltaV system was primarily responsible for recording fluid flow rates, while the Labview system was utilized to capture pressures, film thickness, and fluid temperature at various locations within the Serpent rig's test section. The Labview system also provided an invaluable real-time view of the time traces for these parameters. Lastly, the WMS (Wire Mesh Sensor) data acquisition system played a pivotal role in measuring the phase distribution across the pipe's cross-section, ensuring comprehensive data collection and analysis for this specific aspect of the experiment.

4.4 Flow regime map

From the experimental data collection, the superficial water velocity 1 m/s and gas velocity between 0.2 m/s to 18 m/s is used for this research work, which comes under the identified as slug flow using the reference flow regime illustration map Figure 4-2 the flow regime map was in terms of the liquid and gas superficial. This illustrates the significant influence of pipe diameter and fluid composition on flow patterns.

Studying slug flow in multiphase systems is crucial for comprehending the behavior of gas-liquid mixtures within pipelines or conduits (Jiang and Rezkallah, 1993). Understanding slug flow aids in designing efficient systems, predicting pressure fluctuations, ensuring safety, and optimizing heat and mass transfer

processes. By investigating slug flow, engineers and researchers enhance their ability to model, simulate, and innovate in diverse fields, ultimately leading to improved process reliability and efficiency.

Machine learning study: All available experimental data is used, but the slug flow will be analysed in detail.

Geometry optimization using CFD: Only the slug flow region only will be used for the study.

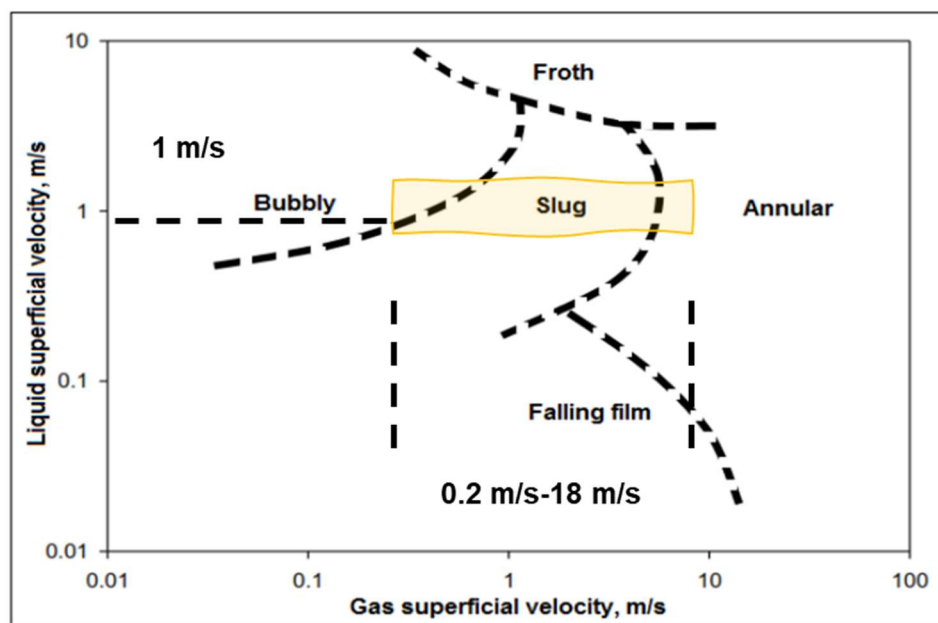


Figure 4-2: vertical downward flow regime (Jiang and Rezkallah, 1993)

5 Machine Learning and Data Analysis

5.1 Overview

In this chapter overlook into the data pre-processing performed to get clean data required for the for-regression model analysis. The Figure 5-1 shows the overview of the flow process, the serpentine pipe flow data gathered from the Cranfield university process engineering lab is received is used. The chapter 4 explains about the experimental overview the descriptive analysis is performed for those experimental data.

The experimental run for different superficial gas and liquid velocities, which is considered as input, the data set gas been separated into two, one set of data used to study the gas void fraction and other set of data used to get the liquid film thickness.

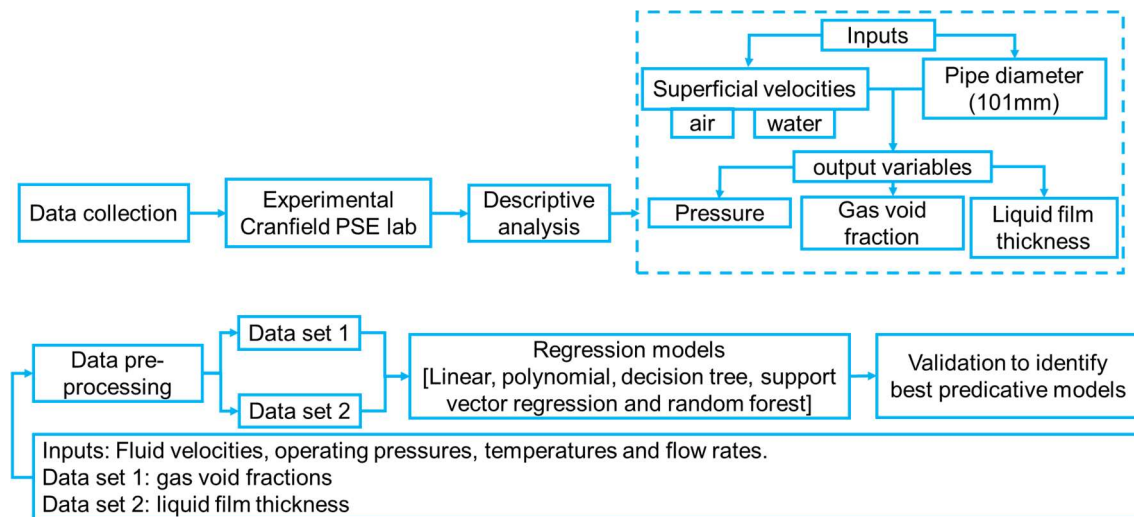


Figure 5-1: Overview of machine approach learning flow chart.

5.2 Regression models

Machine learning algorithms have gained attention in fluid engineering due to their outstanding capacity to analyse the intricate data patterns and make precise predictions. Among the various machine learning techniques based on detail research these algorithms are chosen to use in this study namely,

- linear regression,
- polynomial regression,
- random forest,
- support vector and
- decision tree

Linear regression (Hastie Trevor and Friedman, 2001) is a commonly employed algorithm for modelling the correlation between a dependent variable and one or more independent variables. It presupposes a linear relationship and endeavours to determine the best-fitting line that minimizes the sum of squared errors. The equation for linear regression can be expressed as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad \text{Equation 5-1}$$

where y is the dependent variable, $x_1, x_2, \dots, x_n$ are the independent variables, $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ are the regression coefficients, and ε is the error term.

Polynomial regression (Ostertagová, 2012) is an extension of linear regression that accommodates nonlinear relationships between the dependent and independent variables. This method employs higher-order polynomials to model the relationship. The equation for polynomial regression can be expressed as.

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_n x^n + \varepsilon \quad \text{Equation 5-2}$$

where x represents the independent variable, y represents the dependent variable, $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ are the regression coefficients, $x^2, x^3, \dots, x^n$ are the higher-order terms, and ε is the error term.

The Random Forest algorithm (Kulkarni and Sinha, 2014) is an ensemble learning method that combines multiple decision trees to make predictions. Each tree is constructed using a random subset of features and data samples. The final prediction is obtained by aggregating the predictions of individual trees. Random Forest is capable of handling nonlinear relationships, capturing complex interactions, and effectively managing high-dimensional data.

The Support Vector Decision Tree algorithm combines the strengths of Support Vector Machines (SVM) and decision trees (Bennett and Blue, 1998). It utilizes SVM as the splitting criterion during the construction of decision trees. This approach aims to identify the optimal hyperplane that maximally separates the data points belonging to different classes.

The thesis aims to investigate and compare the performance of Linear Regression, Polynomial Regression, Random Forest, and Support Vector Decision Tree algorithms for regression tasks. The primary focus will be on analyzing their predictive accuracy, ability to handle nonlinear relationships, interpretability, and computational efficiency.

5.3 Data Analysis

Data analysis and understating the experimental data is very important in the machine learning, it will help to get the clean data for model training by identifying and fixing various data issues. By performing the below data analysis and refining tasks, the quality and reliability of machine learning models can be significantly improved, leading to more accurate and meaningful predictions or classifications(Chicco Davide AND Oneto, 2022). The following are some common data analysis and cleaning tasks in machine learning:

- a) The dataset must be analysed to identify the missing values, zero values in features. The decision to deal with these missing values could include imputation (replacing missing values with estimated values) or removing rows or columns with missing data.
- b) Outliers are the data points which is different from the rest of the data should be identified. The outlier needs to be removed, transform, or keep them as valid based on the context of the problem and the model being used.
- c) Categorical variables must be converted into numerical format so that they can be used by machine learning algorithms. Common techniques include one-hot encoding and label encoding.
- d) Scaling and the distribution of numerical features is key factor based on the range it need to be scaled and the decision to apply feature scaling,

such as standardization or normalization, is made to ensure that all features have a similar scale.

- e) The relevance and importance of each feature to the target variable is must to be analysed. Irrelevant or redundant features need to be discarded to reduce dimensionality and improve model performance.
- f) The class imbalances is essential to check, and techniques like oversampling, under sampling, or using different evaluation metrics must be applied to handle the imbalance.
- g) The raw data transformation is need in few cases, such as logarithmic or exponential transformation, to make it more suitable for the chosen machine learning model.
- h) The dataset must be divided into training, validation, and test sets to train the model, tune hyperparameters, and evaluate its performance, respectively.
- i) Time Series Data the temporal order must be preserved during the analysis, and appropriate features must be extracted.
- j) Correlations between features must be examined to identify multicollinearity, which can affect some models like linear regression. Multicollinearity must be addressed by removing or combining highly correlated features.
- k) Noisy data points that may negatively impact the model's performance must be identified and handled.
- l) The data must be visualized to gain insights, identify patterns, and understand the relationships between features and the target variable.
- m) Skewed data handling is very keen in this process, the data transformations or specialized models can be skewed distributions effectively must be considered.
- n) Large datasets, data partitioning techniques like stratified sampling is must to consider ensuring representative subsets for analysis and model training.

5.4 Hardware and code environment

Python code is used for data processing in anaconda navigator platform as well as Jupiter notebook and Jupiter lab environment. The entire program executed on a laptop powered 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz 2.30 GHz 32GB RAM.

5.5 Serpentine pipe data Pre-processing

The experimental results from the serpentine pipe are divided into two data sets, one for analysing the void fraction percentage and the other for primarily examining the liquid layer thickness. For both data sets 2700 data points are available for different inputs.

5.5.1 Dataset – 01: Void fraction data

The void data experimental results separated into the flowing inputs and outputs:

- liquid velocities which vary from 0.07, 0.1, 0.2, 0.38, 0.48, 0.7, 1.0, 1.2 and 1.5 m/s.
- the gas flow rates varies from 6, 10, 20, 30, 40, 50, 80, 100, 150, 200, 300, 400, 600, 800 and 1000 Sm^3/h .
- the standard flow rate is modified to gas flow rates for all the flow conditions.
- The static and absolute pressures is used as inputs.
- Temperatures at the inlet and outlet is considered.

The output pressures and the gas void fraction at six locations are monitored as the output.

5.5.2 Dataset – 02: Liquid film thickness data

The liquid film thickness data experimental results separated into following inputs and outputs.

- liquid velocities which vary from 0.1, 0.2, 0.3, 0.48, 0.7, and 1.0 m/s
- the gas flow rates varies from 50, 100, 150, 200, 300, 400, 600, 800 and 1000 Sm^3/h .

- the standard flow rate is modified to gas flow rates for all the flow conditions.
- The static and absolute pressures is used as inputs.
- Temperatures at the inlet and outlet is considered.

The void fraction, output pressure and liquid film thickness are considered as output.

5.5.3 Step by step code used for data pre-processing:

Python's core libraries for machine learning and data analysis are loaded by this code as shown in Figure 5-2. 'pd,' a powerful data manipulation library, is used by pandas. NumPy is then imported as 'np,' offering support for multi-dimensional arrays and mathematical operations. 'plt,' a well-known data visualization library, is used to import Matplotlib. For feature scaling, a preprocessing step to standardize or normalize data, the Scikit-learn classes 'StandardScaler' and 'MinMaxScaler' are imported. Then, the 'LinearRegression' class from Scikit-learn, a straightforward machine learning algorithm, is loaded. Last but not least, Seaborn is imported as 'sns,' a Matplotlib-based library for data visualization. These libraries are crucial for many data-driven applications because they make data processing, analysis, visualization, and modelling chores easier.

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.svm import SVR
from sklearn.preprocessing import PolynomialFeatures, MinMaxScaler
from sklearn.pipeline import make_pipeline
from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error
from sklearn.linear_model import Ridge
```

Figure 5-2: Libraries used in the code.

5.5.4 Data pre-processing: Min Max Scalar function

As shown in Figure 5-3 next stage in the pre-processing is reading the data, the data are stored in `void_fraction_data.csv` and the `liquid_film_thickness.csv` `pd.read_csv('*****.csv')`: This line imports data into a pandas DataFrame called `data` by reading it from a CSV file with the name `"*****.csv"` (Burstyn et al., 1983).

Data set is divided into inputs and outputs, in both data sets 8 columns is input and remaining is maintained as output using this simple code in `Data.iloc[:,8].values = X` The input variables `X` are chosen as the first 8 columns (0 to 7) of the DataFrame `data` in this line. To pick rows and columns by integer position, it makes use of the `iloc` indexer. The selected DataFrame subset is transformed into a NumPy array by the `values` attribute. `data.iloc[:, 8:] = y.values`: This line chooses the last column of the DataFrame `data` (index 12), starting at column 9, as the output variable `y`.

```
# Load the data from the CSV file
data = pd.read_csv('*****.csv')

# Split the data into input and output variables
X = data.iloc[:, :8].values
y = data.iloc[:, 8:].values

# Create a scaler for the input variables
scalerX = MinMaxScaler()
X_scaled = scalerX.fit_transform(X)

# Create a scaler for the output variables
scalerY = MinMaxScaler()
y_scaled = scalerY.fit_transform(y)
```

Figure 5-3: Code for reading and scaling data

Model's inputs and outputs contain data with different ranges, which could potentially result in mistakes when the model is run. Therefore, standardizing the data using the provided scaling code is essential to ensuring a seamless and accurate model performance. It addresses the issue of differences in their scales by scaling the input and output data to bring all features to a common range. By

removing any potential biases brought on by different scales and allowing the model to interact effectively with the data, this preprocessing phase improves the model's overall accuracy and stability. The ScalerX is created by calling `scalerX = MinMaxScaler()` and is a Min-Max Scaler object. The input variables X are transformed to a particular range (often between 0 and 1) using the Min-Max Scaler.

This line adjusts the input variables X using the scalerX that was previously built, `X_scaled = scalerX.fit_transform(X)`. The `fit_transform()` method transforms the data by scaling it to the appropriate range after fitting the scaler to the data (calculating the minimum and maximum values for scaling). X_scaled contains the input data that has been scaled. By calling `scalerY = MinMaxScaler()`, a new scalerY object of the Min-Max Scaler is created. This scaler's function is to change the output variables y to a particular range. This line adjusts the output variables y by utilizing the scalerY object, `y_scaled = scalerY.fit_transform(y)`. In a manner similar to the prior phase.

5.5.4.1 Min Max Scalar function: Input values data set

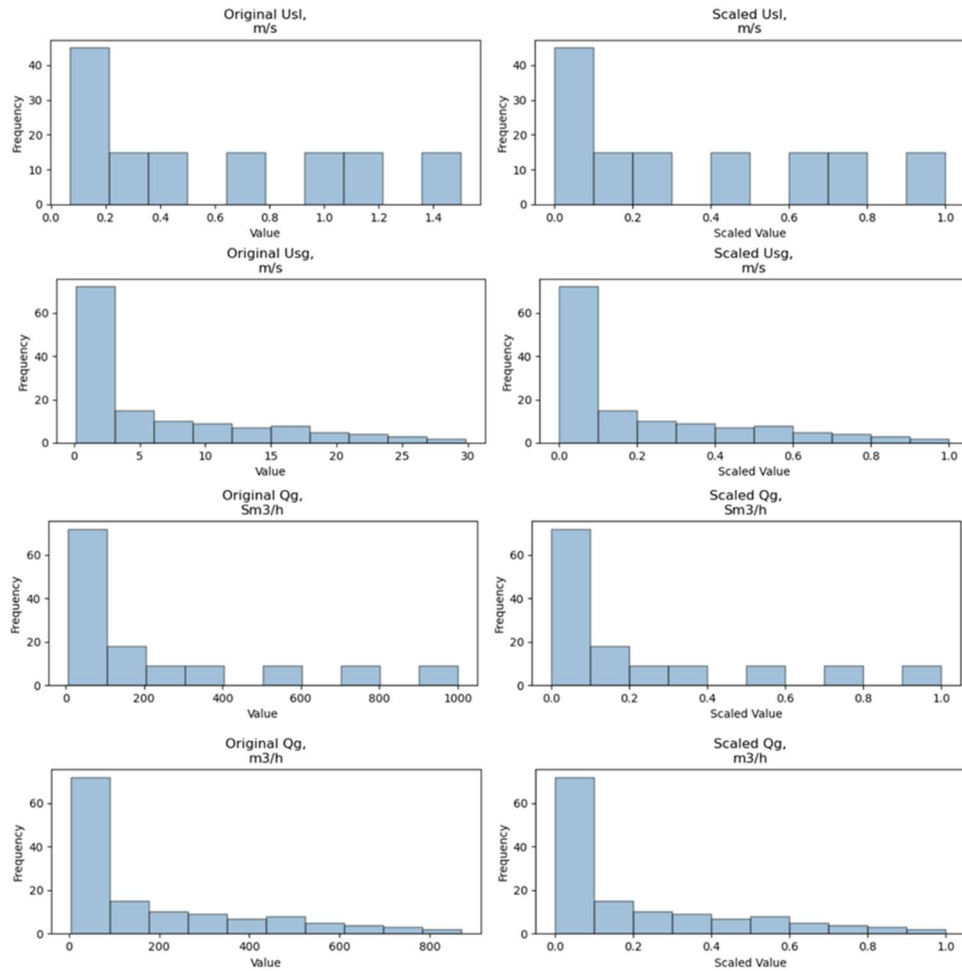


Figure 5-4: Scaled value vs original value: superficial velocities and flow rates

Figure 5-4 shows the Min-Max scaling comparison between the dataset's original values and the corresponding scaled values. Scaling data helps normalize the characteristics to a common range, which is essential preprocessing step in machine learning. This standardization is necessary because machine learning algorithms frequently operate more effectively when the magnitude of the input features is identical.

With Min-Max scaling, the basic goal is to change the data so that all of the values lie between $[0, 1]$. The minimum value of a feature is subtracted from each data point, and the result is divided by the feature's range, which is the difference between its maximum and minimum values, to achieve scaling.

The formula for Min-Max scaling is as follows for a single data point "x":

$$scaled_value = (x - min) / (max - min) \quad \text{Equation 5-3}$$

Where "min" is the minimum value of the feature, "max" is the maximum value of the feature, and "scaled_value" is the resulting scaled value of "x."

The features that have drastically differing magnitudes, for flow rates, gas and liquid superficial velocities and flow, the use of Min-Max scaling is quite advantageous. These velocities can vary greatly in terms of flow rates and most likely have unique units of measurement. It effectively bring both types of velocities to a common scale by scaling the data to a range of [0, 1], which removes the influence of their original magnitudes throughout the machine learning process.

Because of its huge magnitude, this scaling process makes sure that no single feature dominates the learning process and prevents biased predictions. It also aids in improving the machine learning algorithm's overall performance and convergence rate. The Figure 5-5 shows the input data scaled values for temperature and pressure.

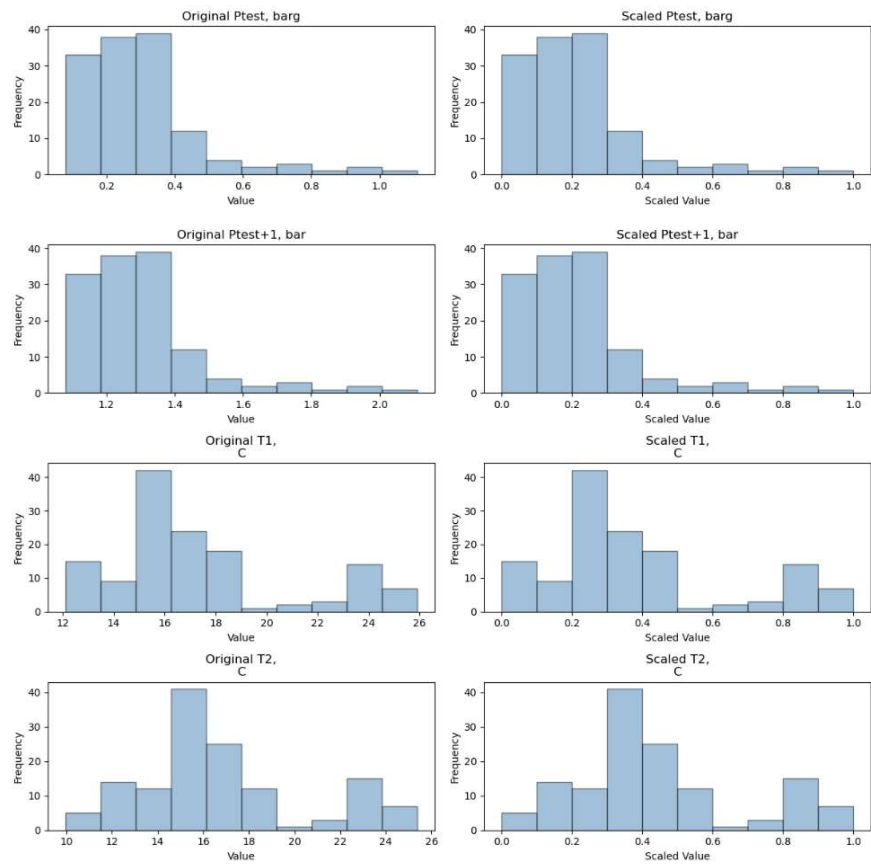


Figure 5-5: Scaled value vs original value: input pressure and temperatures

5.5.4.2 Min Max Scalar function: Data set 1 output values

The Figure 5-6 shows the input data scaled values of pressure at each station and Figure 5-7 shows the scaled values of the void fraction at each stations. Those scaled values used to train the data set for validation and predictions.

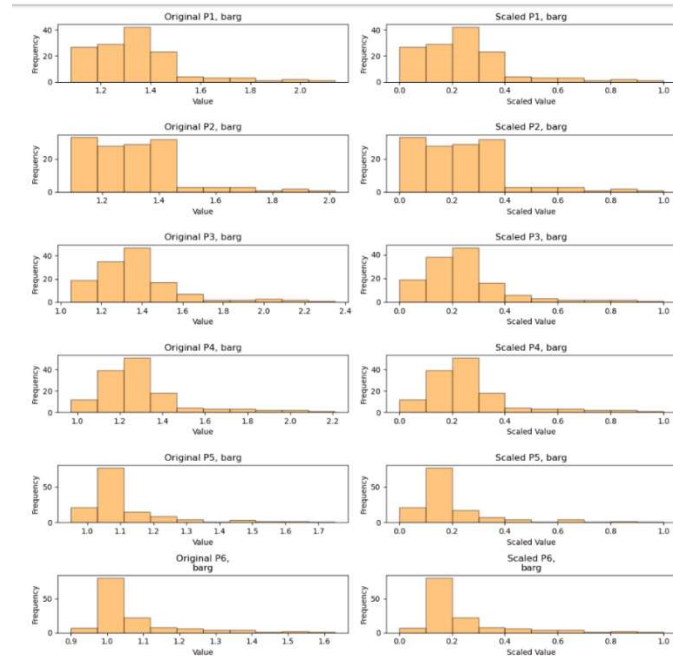


Figure 5-6: Scaled value vs original value: output pressure at stations

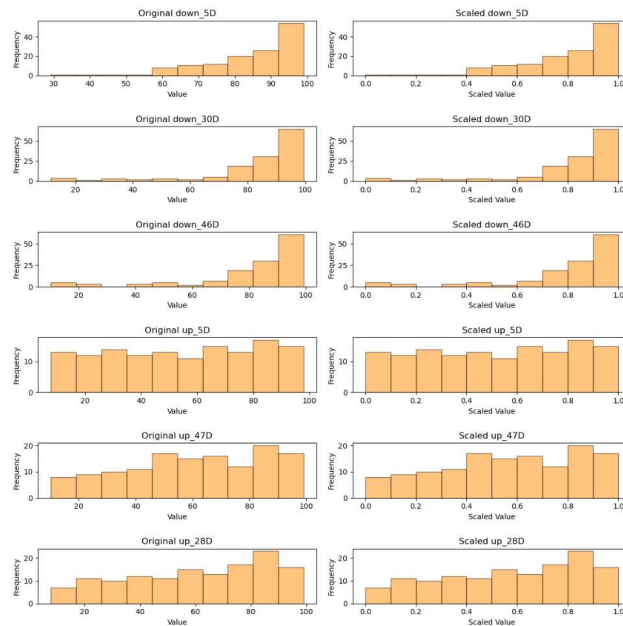


Figure 5-7: Scaled value vs original value: void fraction at each station

5.5.4.3 Min Max Scalar function: Data set 2 output values

The Figure 5-8, Figure 5-9, Figure 5-10, Figure 5-11, Figure 5-12 and Figure 5-13 shows the data set film thickness output scaled values at each stations.

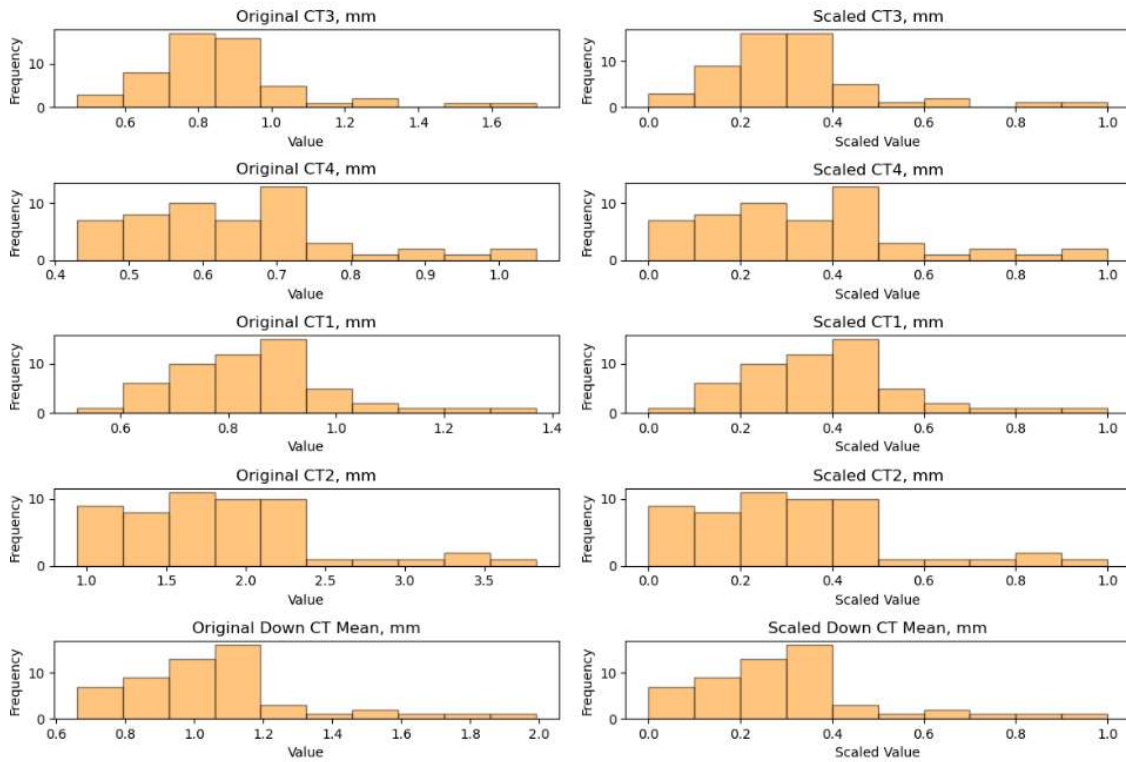


Figure 5-8: Scaled value vs original value: film thickness at down CT(a)

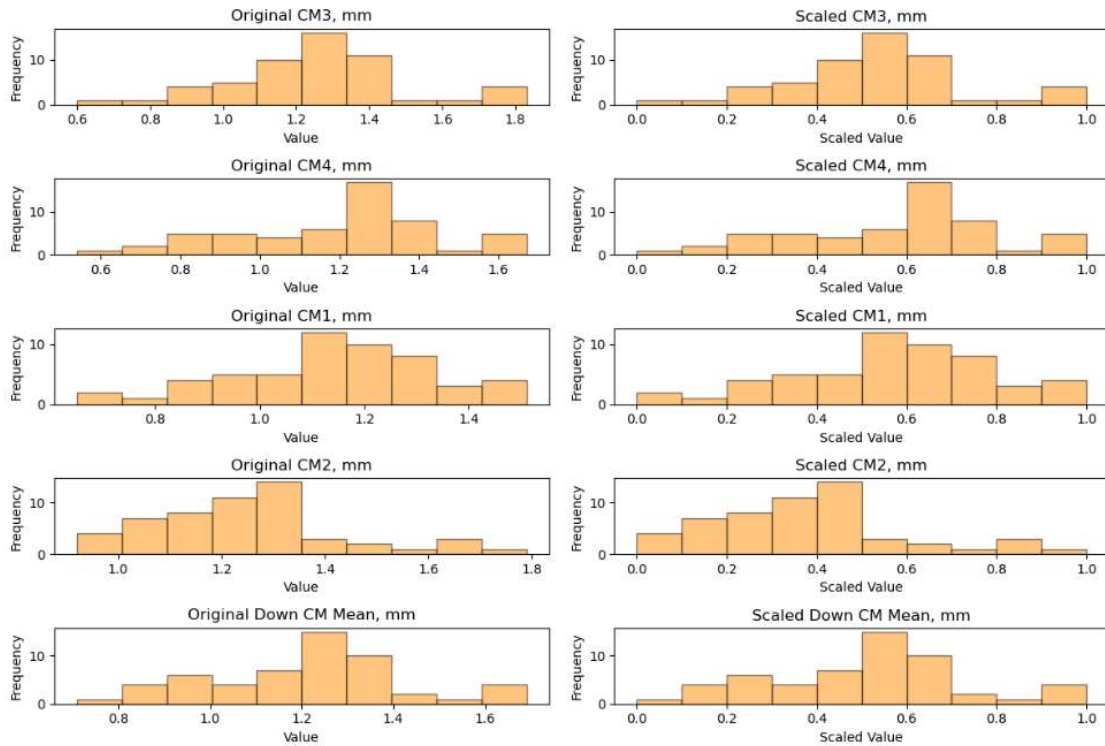


Figure 5-9: Scaled value vs original value: film thickness at down CM(a)

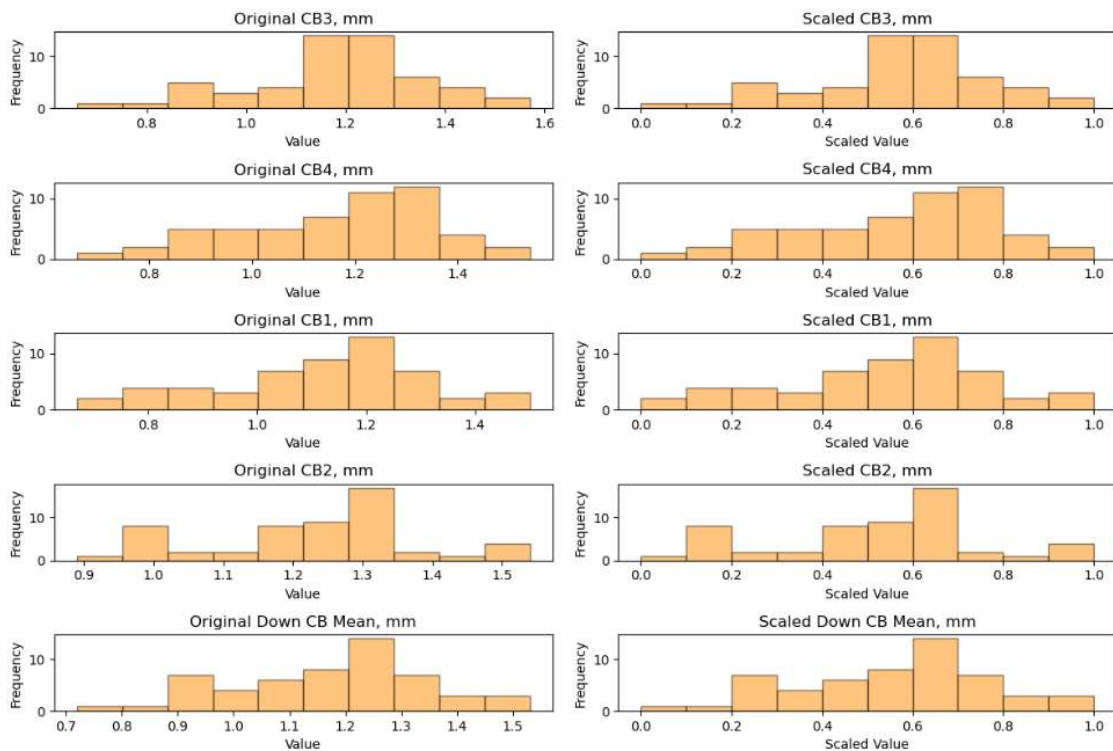


Figure 5-10: Scaled value vs original value: film thickness at down CB(a)

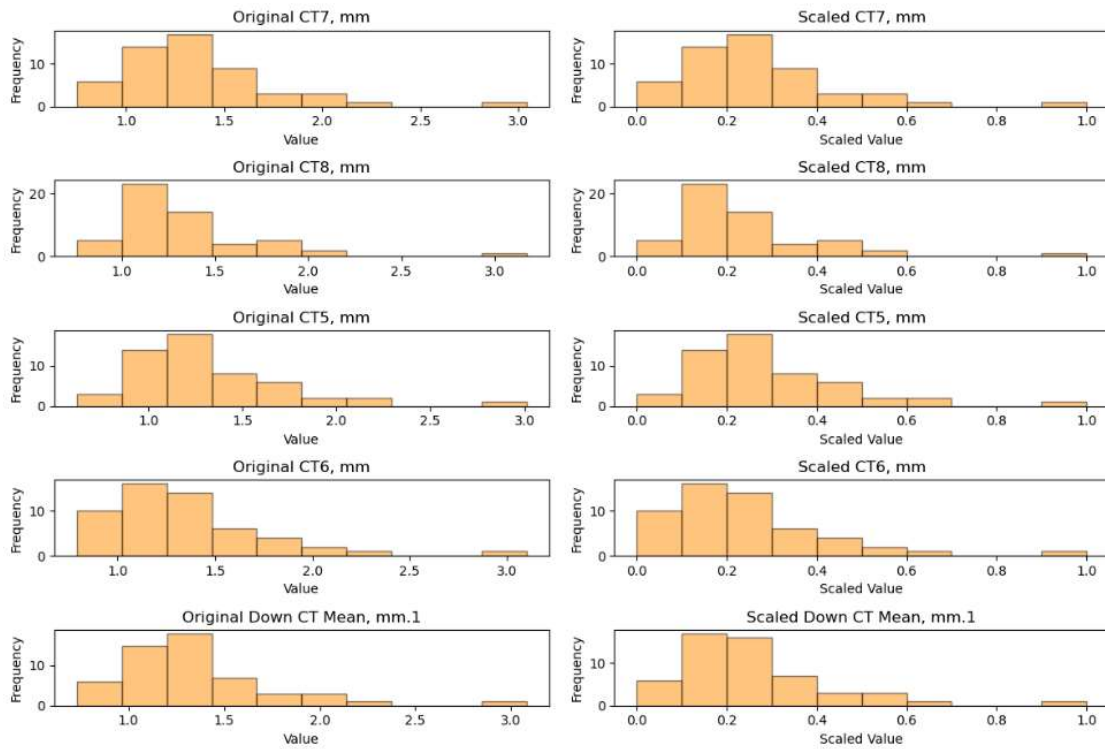


Figure 5-11: Scaled value vs original value: film thickness at down CT(b)

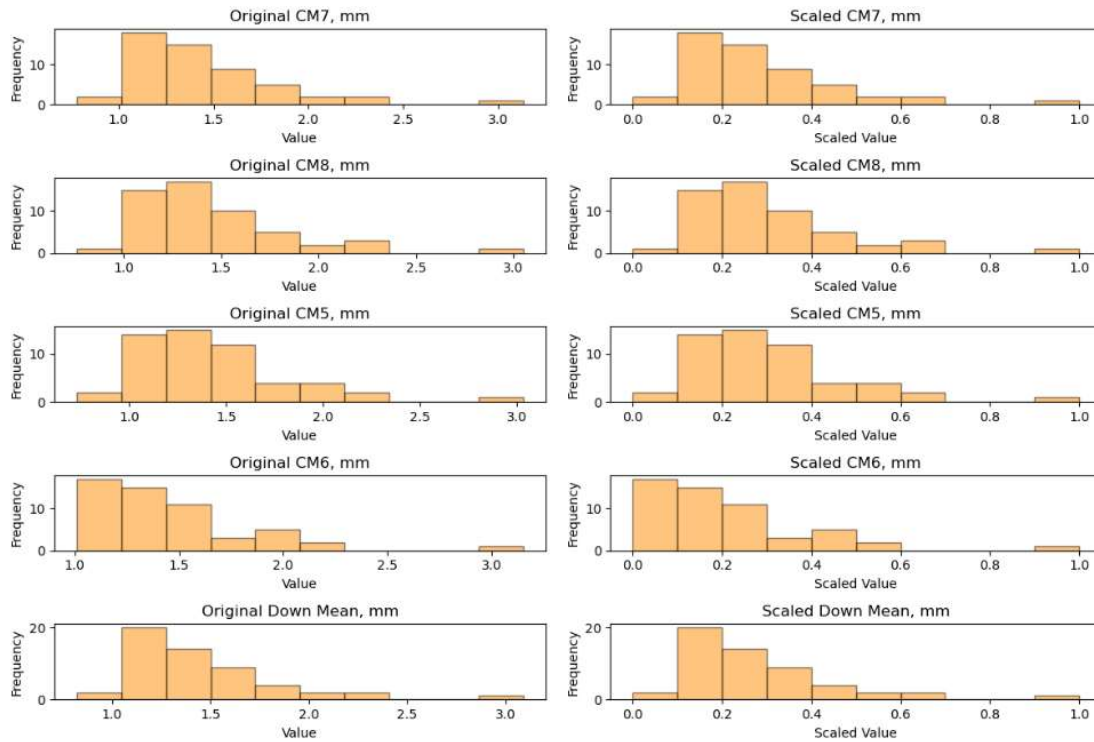


Figure 5-12: Scaled value vs original value: film thickness at down CM(b)

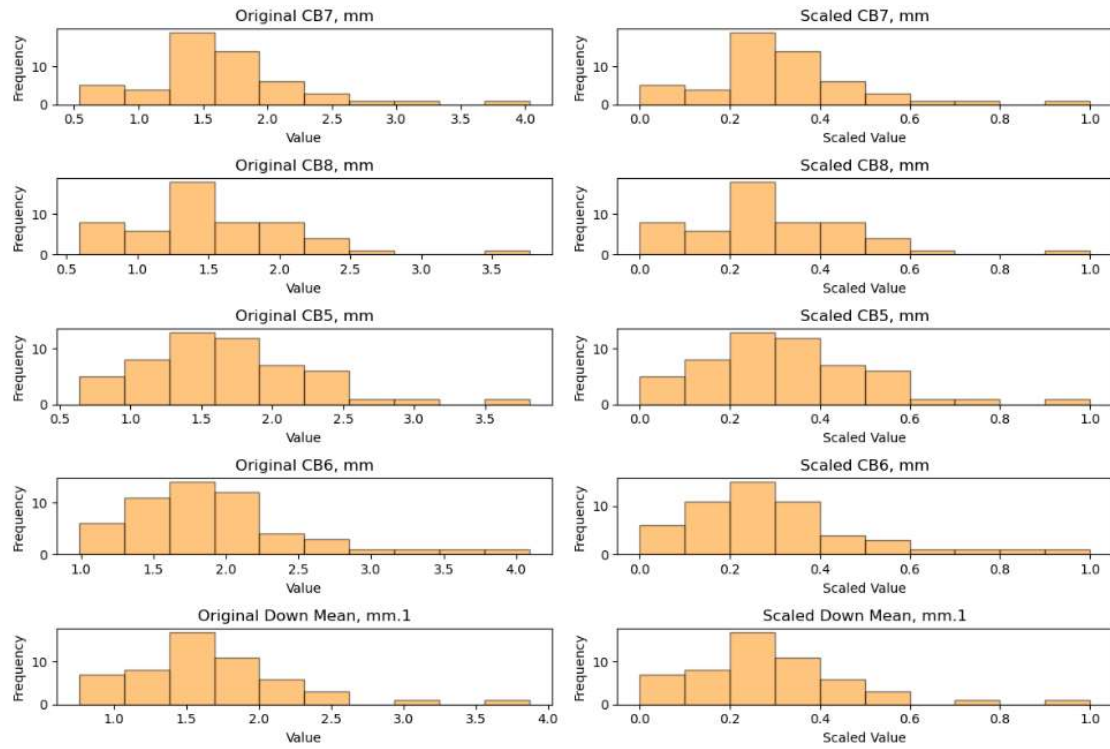


Figure 5-13: Scaled value vs original value: film thickness at down CB (b)

5.5.5 Data pre-processing: Null error identification

The code focuses on finding and dealing with missing values, which are frequent in datasets from the real world. But, certain data points may be missing and represented in the dataset as "null" values due to a variety of factors, such as measurement errors, sensor malfunctions, or problems with data recording. This issue will be addressed by the code, which is made to find any missing values in the dataset kept in the data DataFrame. The code is explained in full here. `print(data.isnull().sum())`. The output from the code is shown in Figure 5-14

Data set 01		Data set 02			
Us1, \nm/s	0	Us1, \nm/s	0	CB3, mm	0
Usg, \nm/s	0	Usg, \nm/s	0	CB4, mm	0
Qg, \nSm3/h	0	Qg, \nSm3/h	0	CB1, mm	0
Qg, \nm3/h	0	Qg, \nm3/h	0	CB2, mm	0
Ptest, barg	0	Ptest, barg	0	Down CB Mean, mm	0
Ptest+1, bar	0	Ptest+1, bar	0	CT7, mm	0
T1, \nC	0	T1, \nC	0	CT8, mm	0
T2, \nC	0	T2, \nC	0	CT5, mm	0
P1, barg	0	CT3, mm	0	CT6, mm	0
P2, barg	0	CT4, mm	0	Down CT Mean, mm.1	0
P3, barg	0	CT1, mm	0	CM7, mm	0
P4, barg	0	CT2, mm	0	CM8, mm	0
P5, barg	0	Down CT Mean, mm	0	CM5, mm	0
P6, \nbarg	0	CM3, mm	0	CM6, mm	0
down_5D	0	CM4, mm	0	Down Mean, mm	0
down_30D	0	CM1, mm	0	CB7, mm	0
down_46D	0	CM2, mm	0	CB8, mm	0
up_5D	0	Down CM Mean, mm	0	CB5, mm	0
up_47D	0			CB6, mm	0
up_28D	0			Down Mean, mm.1	0
dtype: int64				dtype: int64	

Figure 5-14: Null error results

5.5.6 Data pre-processing: Standard deviation

This code in the Figure 5-15 aims to analyse the input and output data of a serpentine flow model using the concept of standard deviation. Standard deviation is a statistical measure that indicates the degree of dispersion or variability of a set of values (Obaid, Dheyab and Sabry, 2019).

```
# Compute standard deviation of inputs and outputs
input_std = np.std(X_scaled, axis=0)
output_std = np.std(y_scaled, axis=0)

# Create a bar chart of the standard deviations for inputs
fig, ax = plt.subplots()
ax.bar(range(len(input_std)), input_std)
ax.set_xticks(range(len(input_std)))
ax.set_xticklabels(['Input {}'.format(i+1) for i in range(len(input_std))])
ax.set_xlabel('Variable')
ax.set_ylabel('Standard Deviation')
ax.set_title('Standard Deviation of Inputs')
plt.show()

# Set the output color variable to a dark shade of orange with outline
outputcolor = np.array([255, 140, 0]) / 255 # Dark orange color
outlinecolor = 'black' # Black outline color

# Set the size of the plot (adjust the width and height as desired)
plt.figure(figsize=(12, 8)) # Example: width=10 inches, height=6 inches

# Create a bar chart of the standard deviations for outputs
fig, ax = plt.subplots()
ax.bar(range(len(output_std)), output_std)
ax.set_xticks(range(len(output_std)))
ax.set_xticklabels(['Output {}'.format(i+1) for i in range(len(output_std))])
ax.set_xlabel('Variable', fontsize=14)
ax.set_ylabel('Standard Deviation')
ax.set_title('Standard Deviation of Outputs')

plt.show()
```

Figure 5-15: Standard deviation program

In the context of this thesis, the input data includes various parameters such as flow, temperatures, and velocity, while the output data set 1 consists of pressure and void fraction at different locations and data set 2 hold the liquid film thickness within the serpentine channel. The code starts by calculating the standard deviation for both the input and output data, which is done by computing the standard deviation along each column (axis 0) of the scaled input and output datasets (*X_scaled* and *y_scaled*, respectively). This code is useful for gaining insights into the variability of the input data, which can help in understanding the potential impact of different input variables on the liquid film thickness, a critical parameter in serpentine flow analysis. The bar chart provides a clear visual representation of the standard deviations, allowing researchers to identify which input variables contribute most to the variability of the output, thereby aiding in the interpretation and improvement of the serpentine flow model.

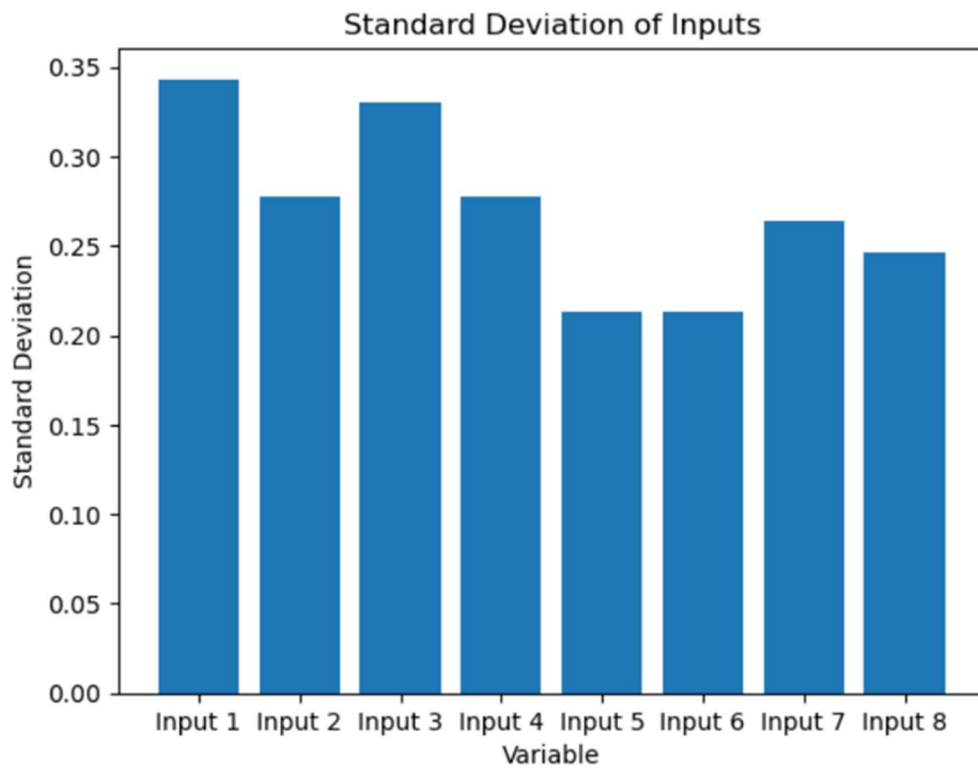


Figure 5-16: Input data standard deviation plot

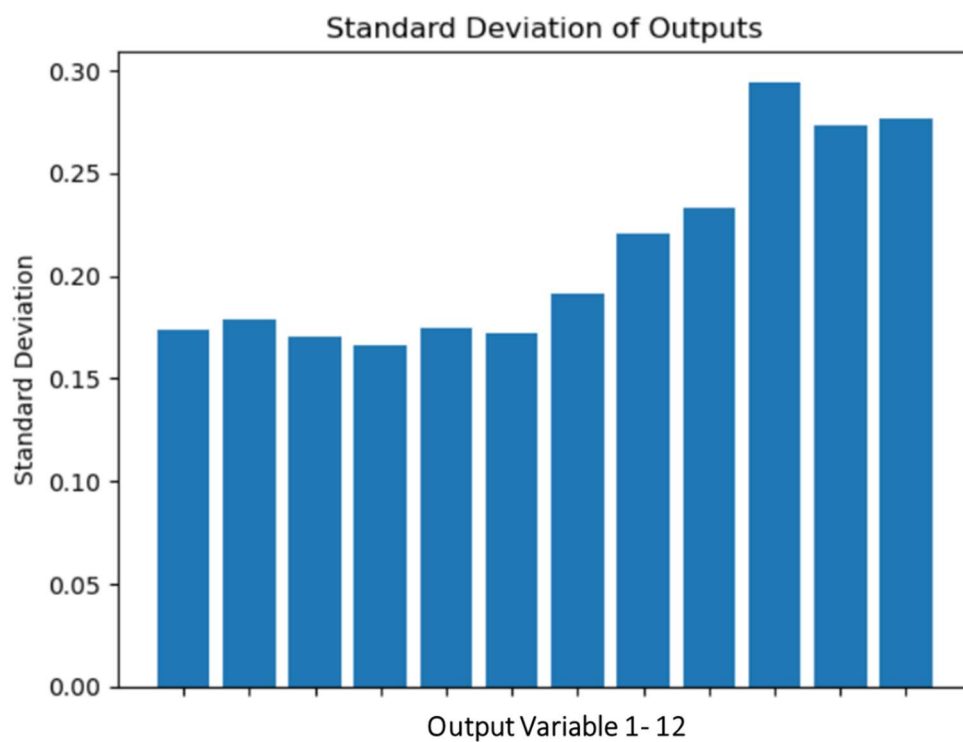


Figure 5-17: Data set 1 output standard deviation plot

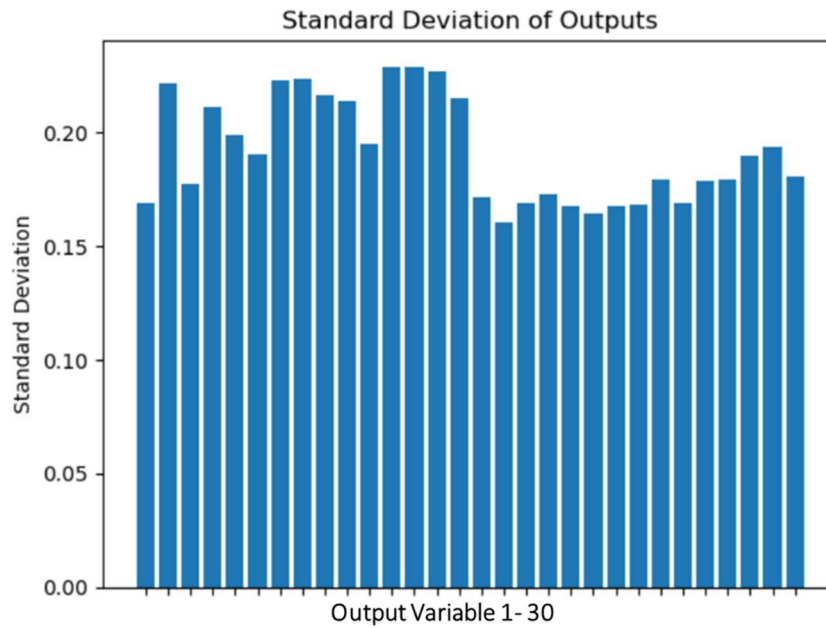


Figure 5-18: Data set 2 output standard deviation plot

Figure 5-17 indicates the standard deviation plot for the input, the variation is between 0.22 -0.34. Figure 5-17 first dataset (data set-1) falls within the range of 0.22 to 0.35, while the standard deviation of the second dataset Figure 5-18 (data set-02) lies between 0.17 and 0.29. This information is crucial for understanding the variability or spread of the data in each dataset. A smaller standard deviation implies that the data points are closer to the mean, indicating less variability or more consistency within the dataset. On the other hand, a larger standard deviation indicates that the data points are more spread out from the mean, signifying greater variability. The deviation in the outputs and outputs are not high, which is a good sign for proceeding with the training and predictions.

5.5.7 Data pre-processing: Histogram plots

In data analysis histogram plots serve as invaluable tools for visualizing the distribution of data within a given dataset. These plots provide a graphical representation of the frequency or count of data points falling within specified ranges or "bins." In the context of this thesis, the focus is on the input variables, which are critical factors influencing the observed outcomes (Obaid, Dheyab and Sabry, 2019).

The provided code Figure 5-19 segment is responsible for generating histogram plots for each of the input variables in the dataset, denoted as 'df' in the code.

The loop iterates over the columns corresponding to the input variables, and for each input, a histogram is constructed. The 'sns.displot' function from the Seaborn library is employed to create these histograms. This function automatically organizes the data into bins (here, 50 bins) and represents the count or density of data points within each bin. The x-axis of the histogram corresponds to the range of the input variable, while the y-axis represents the density or frequency of data points within each bin.

The visual outcome of this code is a series of histograms, with each plot showcasing the distribution pattern of a specific input variable. By examining these histograms, researchers gain insights into the central tendencies, spread, and potential outliers within the input data. This visual exploration is crucial for understanding the characteristics of the input variables, which, in turn, aids in subsequent data analysis, modelling, and drawing meaningful conclusions about the problem being investigated in the thesis.

```
# Plot histograms for each of the inputs
for column in data.columns[:8]:
    sns.displot(data[column], bins=50)
    plt.xlabel(column)
    plt.ylabel('Density')
    plt.title(f'Histogram of {column}')
    plt.show()
```

Figure 5-19: Histogram code

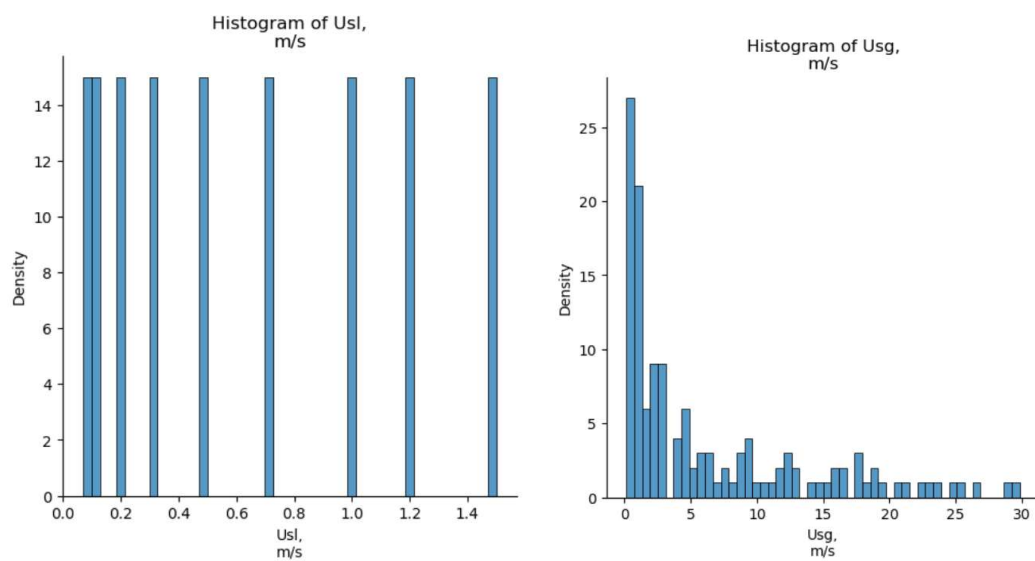


Figure 5-20: Input histogram plot-superficial liquid (left) and gas (right) velocity

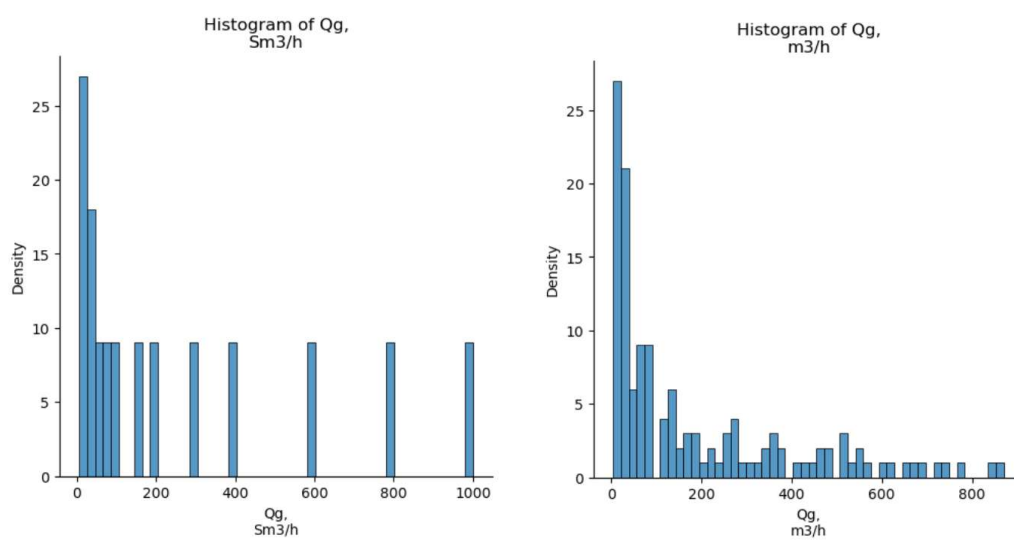


Figure 5-21: Input histogram plot-standard (left) and original (right) flow rates

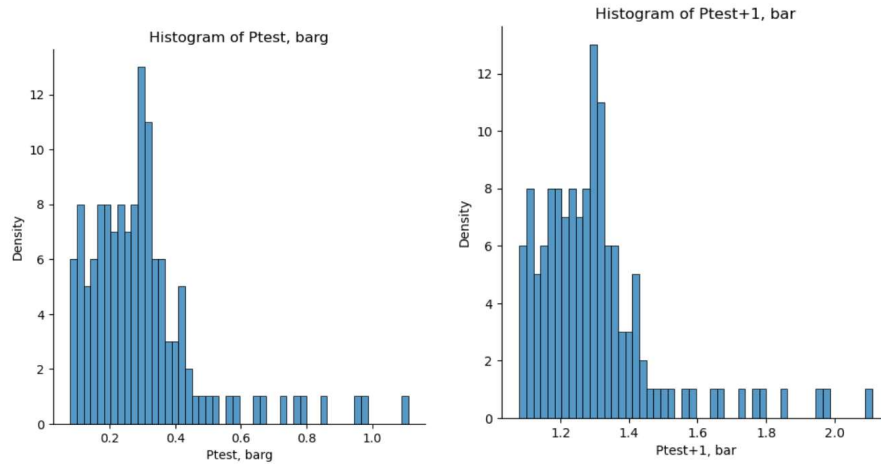


Figure 5-22: Input histogram plot gauge (left) and absolute pressure (right)

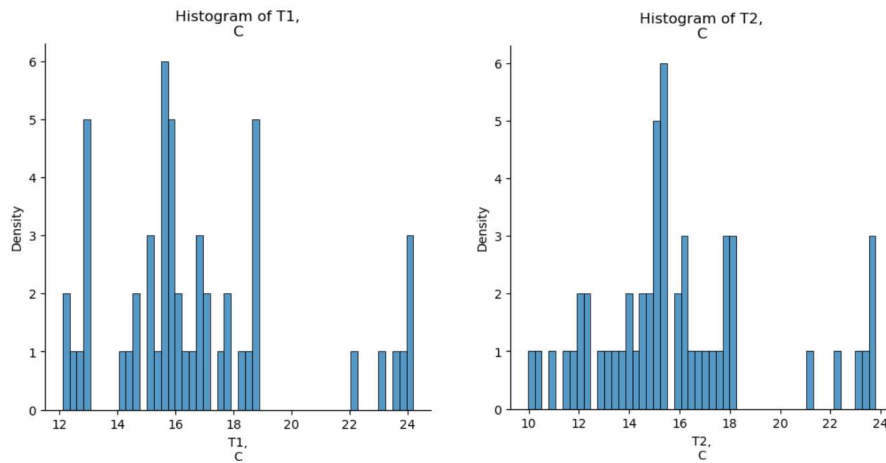


Figure 5-23: Input histogram plot-inlet (left) and outlet (right) temperature.

The histogram results for the input data is shown in the Figure 5-20, Figure 5-21 Figure 5-22 and Figure 5-23.

5.5.8 Data pre-processing: pair plot

In the pair plot: The scatterplots in the upper triangle of the matrix show the relationship between pairs of input variables (Cox, 2007) and each point represents a data point in the dataset, and the scatterplot visually represents the correlation or lack thereof between the two variables.

The histograms in the diagonal plots show the distribution of each individual input variable. These histograms provide insights into the range and frequency of values for each variable.

The pair plot helps to quickly assess whether certain variables are correlated, whether there are any outliers, and what kind of distributions the variables have. Patterns, clusters, or trends between variables can also be observed, which can be helpful for feature selection or further analysis.

```
# Create a pair plot of the input variables
sns.pairplot(data.iloc[:, :8], diag_kind='hist')
plt.show()
```

Figure 5-24: Pair plot code

The code is used shown in the Figure 5-24 and Figure 5-25 shows the plot for the input data.

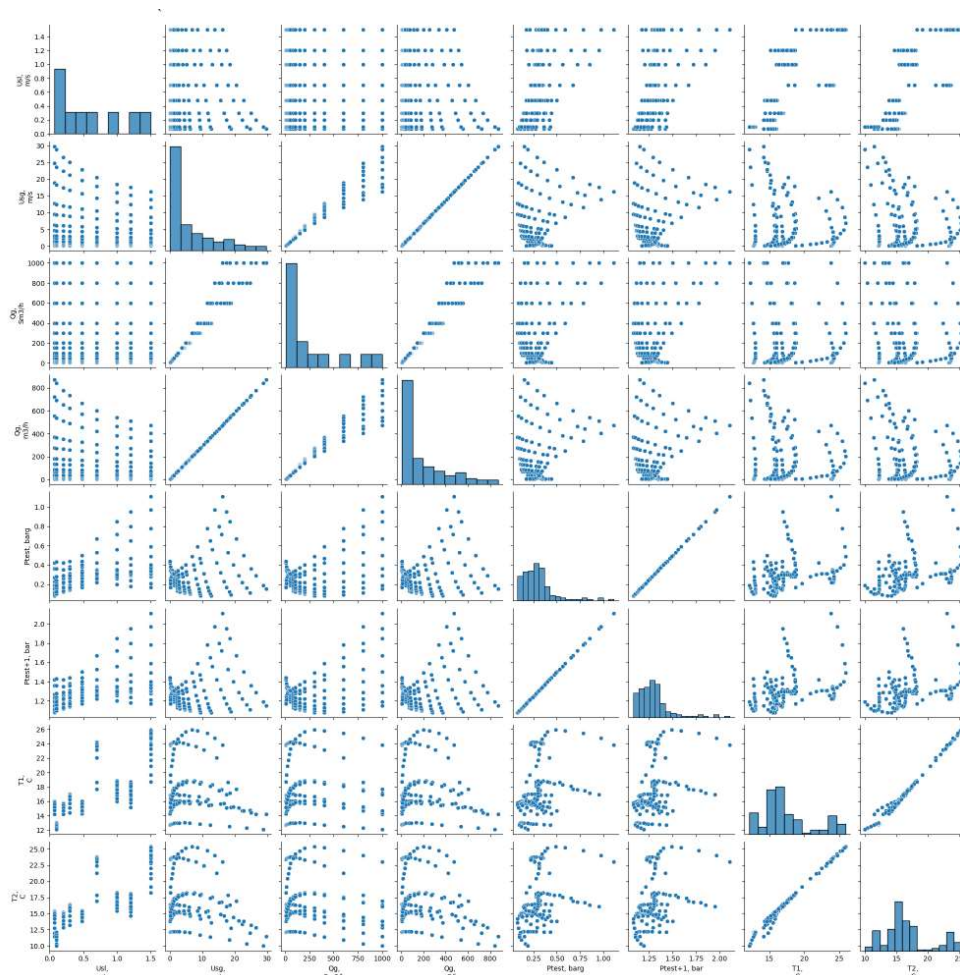


Figure 5-25: Input data pair plot

5.5.9 Data pre-processing: coefficient of the model

The coefficients are crucial in understanding the relationship between the input features and the target variable in a linear regression model. They provide insights into the magnitude and direction of influence that each feature has on the predicted outcome (Kuhn Max and Johnson, 2013). Positive coefficients suggest a positive impact on the target variable, while negative coefficients suggest a negative impact.

By analysing the coefficients, the features have the most significant influence on the prediction can be identified. Features with larger absolute coefficients tend to contribute more to the model's predictions. Additionally, coefficients can help to understand whether an increase or decrease in a feature leads to an increase or decrease in the predicted outcome. The code to use to find the coefficient is for the data sets is shown in Figure 5-26 and the output of the coefficient's are shown in Figure 5-27 and Figure 5-28.

```
# Split the data into input and output variables
X = df.iloc[:, :8]
y = df.iloc[:, 8:]

# Fit a linear regression model
model = LinearRegression()
model.fit(X, y)

# Print the coefficients of the model
print("Coefficients:", model.coef_)
|
```

Figure 5-26: Code for coefficient calculation

```

Coefficients: [[ 5.97986437e-02 -1.21590317e+00  1.90414587e-04  4.13432419e-02
 4.48668504e-01  4.48668504e-01  1.62007313e-02 -1.89200086e-02]
 [ 8.72687758e-02 -1.18005495e+00  8.82635019e-05  4.02655260e-02
 3.89118216e-01  3.89118216e-01  1.41465411e-02 -1.65303054e-02]
 [-8.07826753e-02  1.18922331e+00 -3.35774897e-04 -4.02641710e-02
 7.02182348e-01  7.02182348e-01 -3.54169653e-02  3.76989740e-02]
 [-9.16072896e-02  1.52872299e+00 -1.06139840e-04 -5.21866761e-02
 6.16170074e-01  6.16170074e-01 -2.02842019e-02  2.30518817e-02]
 [ 1.33540707e-02  4.16071245e-02  7.64132142e-04 -2.16633776e-03
 1.51592020e-01  1.51592020e-01  1.01267631e-02 -7.06944030e-03]
 [-1.24115994e-02  5.77597946e-01  8.56655698e-04 -2.06370256e-02
 8.64917923e-02  8.64917923e-02  9.90375011e-03 -6.10400680e-03]
 [-3.08640500e+01  1.64748744e+02 -1.17447180e-02 -5.61532356e+00
 1.30980848e+01  1.30980848e+01 -1.72125152e+00  2.43172359e+00]
 [-4.07525885e+01  4.47379375e+02 -7.20962044e-02 -1.52193146e+01
 3.01089064e+01  3.01089064e+01 -5.46726470e+00  6.39199908e+00]
 [-4.40387103e+01  4.79741890e+02 -6.57404405e-02 -1.63347964e+01
 3.00079899e+01  3.00079899e+01 -5.11883547e+00  6.18206399e+00]
 [-7.92213918e+00  2.26492262e+02  1.93467193e-01 -7.89811852e+00
 -4.69593117e+01 -4.69593117e+01 -1.95065201e+00  3.21002643e+00]
 [-1.09537476e+01  3.08036712e+02  1.40107160e-01 -1.06357379e+01
 -3.43607575e+01 -3.43607575e+01 -3.26476748e+00  4.53776024e+00]
 [-1.57400395e+01  5.43015435e+01  1.95274143e-01 -2.01101003e+00
 -4.68959801e+01 -4.68959801e+01 -1.75439322e-01  2.25435675e+00]]

```

Figure 5-27: Data set -01 coefficient

```

Coefficients: [[ 8.44801683e-01  4.71502333e+00 -1.43563498e-03 -1.59550447e-01
-1.70376126e-01 -1.70376126e-01 -3.43919686e-01  3.16250824e-01]
[ 6.17860229e-01  1.72346477e+00 -1.01188471e-03 -5.74722795e-02
-4.55789462e-02 -4.55789462e-02 -1.30694594e-01  1.13798935e-01]
[ 4.99468878e-01  2.31219597e+00 -1.37062306e-03 -7.71941642e-02
-1.17719657e-01 -1.17719657e-01 -3.29705343e-01  3.17799470e-01]
[ 3.08385502e+00  6.11512202e+00  1.72933759e-03 -2.11361978e-01
-2.00229371e+00 -2.00229371e+00 -2.72557831e-01  2.20009508e-01]
[ 1.25680539e+00  2.44551023e+00 -3.21730097e-04 -8.30863047e-02
-6.33671042e-01 -6.33671042e-01 -2.63898694e-01  2.37295512e-01]
[ 8.48893482e-01 -5.03498340e-01 -5.36697415e-04  1.84534200e-02
-3.48038182e-01 -3.48038182e-01 -5.94033247e-01  5.68722927e-01]
[ 8.07254084e-01  5.80131466e-01  6.43108728e-04 -2.02148038e-02
-5.68938119e-01 -5.68938119e-01 -5.57023508e-01  5.45370220e-01]
[ 6.02269346e-01 -2.35489980e-01  5.92286685e-05  8.39266448e-03
-4.14026504e-01 -4.14026504e-01 -4.54770281e-01  4.48024428e-01]
[ 9.25483650e-01  1.47257588e+00  7.38363803e-05 -5.00651687e-02
-5.58157080e-01 -5.58157080e-01 -2.55917853e-01  2.40745391e-01]
[ 7.93195952e-01  4.18349765e-01  5.40046653e-05 -1.39357343e-02
-4.68963524e-01 -4.68963524e-01 -4.64482762e-01  4.49459680e-01]
[ 5.59294758e-01 -3.77863503e-02 -7.53993824e-04  3.06993849e-03
-2.51738298e-01 -2.51738298e-01 -6.95688879e-01  6.76157176e-01]
[ 6.32704369e-01  2.72099985e+00 -4.56737997e-04 -9.19909854e-02
-2.63107347e-01 -2.63107347e-01 -6.52795567e-01  6.33492059e-01]
[ 6.49696683e-01  2.60320255e+00  7.83847367e-05 -8.88091768e-02
-4.27286647e-01 -4.27286647e-01 -4.74031452e-01  4.64588044e-01]
[ 5.09089283e-01  2.75205840e+00 -1.18116916e-03 -9.20840360e-02
-1.53739601e-01 -1.53739601e-01 -5.27603327e-01  5.14686593e-01]
[ 5.88732629e-01  2.21792236e+00 -5.77602039e-04 -7.45936506e-02
-2.72638044e-01 -2.72638044e-01 -5.89990343e-01  5.74217431e-01]
[ 4.89856561e-02  1.37543824e+01 -5.85918392e-03 -4.65461942e-01
 1.36918306e+00  1.36918306e+00 -2.00987883e-01  1.68777156e-01]
[ 1.96129112e-01  1.60671091e+01 -5.45487195e-03 -5.45168546e-01
 1.21497490e+00  1.21497490e+00 -1.64370541e-01  1.27916924e-01]
[ 1.87725026e-01  8.41925974e+00 -4.53630339e-03 -2.84450455e-01
 1.07867658e+00  1.07867658e+00 -1.19892270e-01  8.78245635e-02]
[ 1.34782841e-01  1.47118518e+01 -5.35478686e-03 -4.98916956e-01
 1.18120959e+00  1.18120959e+00 -1.66023367e-01  1.35394740e-01]
[ 1.38653687e-01  1.31947556e+01 -5.32978965e-03 -4.46976888e-01
 1.21817532e+00  1.21817532e+00 -1.66173908e-01  1.33211214e-01]
[ 2.51859882e-01  1.45869142e+01 -4.89205731e-03 -4.95115250e-01
 1.15624585e+00  1.15624585e+00 -2.40315183e-01  2.08268374e-01]
[ 2.62242678e-01  1.54212634e+01 -5.30861071e-03 -5.23110870e-01
 1.20319953e+00  1.20319953e+00 -3.21360521e-01  2.79729940e-01]
[ 3.41256673e-01  1.06282913e+01 -3.93356302e-03 -3.60826661e-01
 9.34245135e-01  9.34245135e-01 -6.25899619e-02  3.43139860e-02]
[ 2.95824937e-01  1.71182211e+01 -4.95440546e-03 -5.81962633e-01
 1.19965228e+00  1.19965228e+00 -6.46173859e-02  3.41124788e-02]
[ 2.84961143e-01  1.45583686e+01 -4.79099970e-03 -4.94332165e-01
 1.13179059e+00  1.13179059e+00 -1.74275171e-01  1.40998861e-01]
[ 1.32591301e+00  1.89778454e+01 -6.42749064e-03 -6.44444786e-01
 9.04401125e-01  9.04401125e-01  1.15280075e-01 -1.53760429e-01]
[ 7.73940307e-01  1.72796607e+01 -7.96055877e-03 -5.83779657e-01
 1.16757423e+00  1.16757423e+00 -3.97783558e-01  3.37436766e-01]
[ 1.26420056e+00  2.17939743e+01 -8.41345276e-03 -7.37911467e-01
 1.29692122e+00  1.29692122e+00 -5.52740646e-01  4.78144074e-01]
[ 1.85218826e+00  2.90674006e+01 -9.72092228e-03 -9.84535359e-01
 1.00130630e+00  1.00130630e+00 -5.49821912e-01  4.88212667e-01]
[ 1.06011790e+00  2.22990981e+01 -8.16741950e-03 -7.55381204e-01
 1.16361971e+00  1.16361971e+00 -3.78577333e-01  3.29022592e-01]]

```

Figure 5-28: Data set -02 coefficient

5.5.10 Data pre-processing: Correlation matrix heat map

This code is of significant importance in the context of analyzing flow data correlations. It's designed to generate a heatmap representation of a correlation matrix, providing a visually intuitive way to grasp the relationships between different variables in a dataset (Zhang et al., 2015). Here's why this code is crucial as shown in Figure 5-29 :

Visualization of Relationships: The heatmap produced by this code offers a clear visual representation of the correlations between various flow data variables. It allows researchers and analysts to quickly identify which variables are positively or negatively correlated and to what degree.

Identifying Patterns: Correlation heatmaps help in uncovering patterns or trends within the dataset. Variables that move in tandem or exhibit inverse behaviors become readily apparent, aiding in the understanding of complex interdependencies within the flow data.

Variable Prioritization: By highlighting strong correlations, the heatmap aids in prioritizing variables that exhibit significant relationships. This can guide further investigations or analysis, allowing researchers to focus on the most influential factors.

Data-Driven Insights: Correlation analysis is essential for data-driven decision-making. By identifying which flow data variables are highly correlated, stakeholders can make informed choices about potential interventions or optimizations.

Data Quality Check: The heatmap can also help in spotting inconsistencies or anomalies in the data. Sudden deviations from expected correlations may indicate data quality issues or outliers that require further investigation.

Effective Communication: Visualizing correlations through heatmaps simplifies the communication of complex relationships to stakeholders who may not be well-versed in the technical details of the data. It translates intricate data insights into easily understandable visuals.

Parameter Tuning: For flow simulations or experiments, understanding how different parameters correlate can guide researchers in setting up experiments or simulations. Adjusting parameters based on their correlations can help capture accurate representations of real-world scenarios.

Model Enhancement: When coupled with machine learning, correlation analysis can guide feature selection and model building. Variables with high correlations might be strong predictors, enhancing the predictive power of models.

```
# Compute correlation matrix
corr_matrix = data.corr()

# Assuming you have defined the correlation matrix 'corr_matrix'

# Set the output color variable to a dark shade of orange with outline
outputcolor = np.array([255, 140, 0]) / 255 # Dark orange color
outlinecolor = 'black' # Black outline color

# Increase the size of the heatmap
plt.figure(figsize=(16, 8)) # Adjust the width and height as desired

# Create heatmap of correlation matrix
sns.heatmap(corr_matrix, cmap='coolwarm', xticklabels=corr_matrix.columns, yticklabels=corr_matrix.columns,
            annot=True, fmt='.2f', annot_kws={"size": 10}, linewidths=0.5) # Increase annot_kws font size to 10 or any other desired value
plt.xticks(rotation=45, ha='right', fontsize=10) # Increase xticklabels font size to 10 or any other desired value
plt.yticks(fontsize=10) # Increase yticklabels font size to 10 or any other desired value
plt.title('Correlation Matrix Heatmap', fontsize=14) # Increase the title font size to 14 or any other desired value
plt.show()
```

Figure 5-29: Code for correlation matrix heat map

In summary, this code serves as a powerful tool for extracting meaningful insights from flow data by visually depicting correlations. It empowers researchers, engineers, and analysts to make informed decisions, identify hidden patterns, and enhance their understanding of the underlying dynamics within the dataset.

Figure 5-30 shows the data set 1 matrix and Figure 5-31 shows the data set 2 matrix, which will be used to fine tune the model in later stages.

Positive Correlation (Values between 0 and 1): A positive correlation value between 0 and 1 indicates that as one variable increases, the other variable tends to increase as well. The closer the value is to 1, the stronger the positive correlation. This implies that the two variables move in the same direction, suggesting a potential linear relationship where both tend to rise or fall together.

Negative Correlation (Values between -1 and 0): A negative correlation value between -1 and 0 signifies that as one variable increases, the other variable tends to decrease. The closer the value is to -1, the stronger the negative correlation.

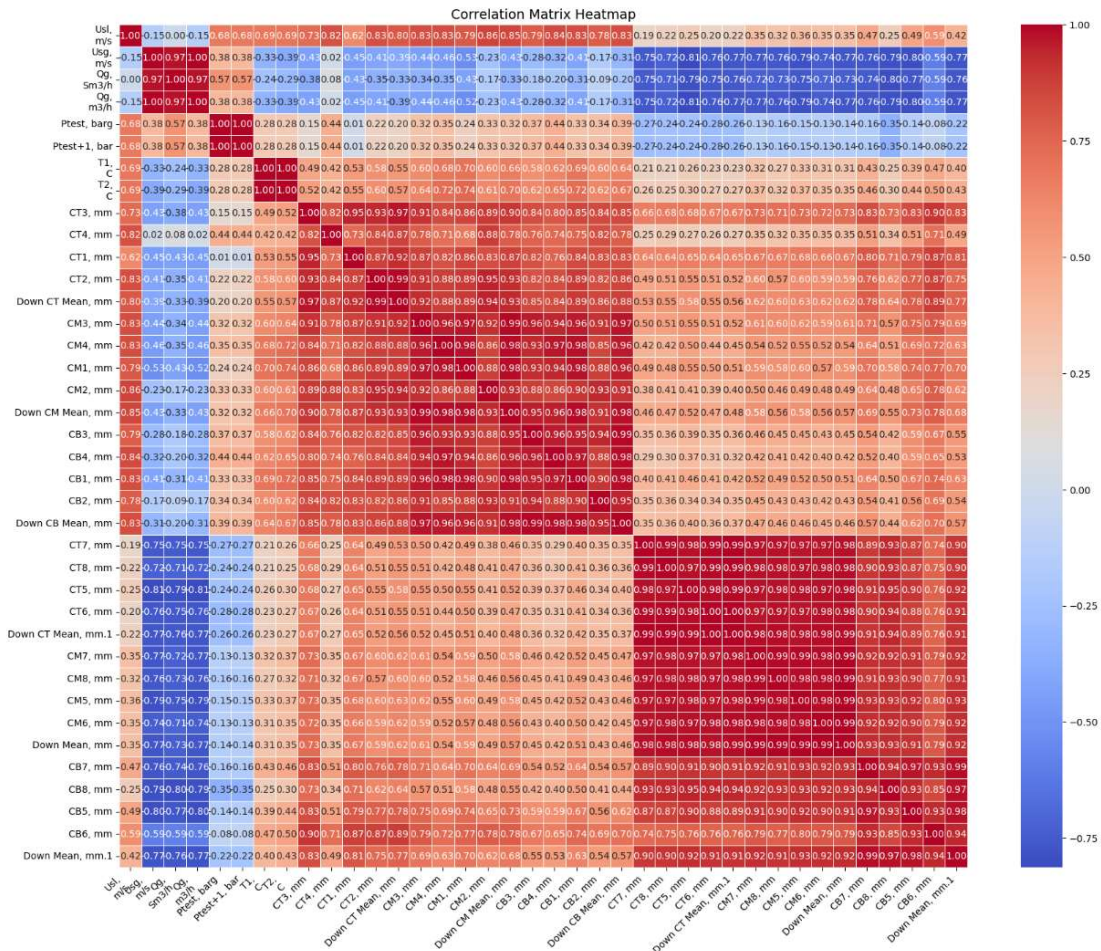


Figure 5-31: Data set 02 correlation matrix heatmap

5.6 Regression Models Implementation in this Study

Certainly, the provided code serves as a pivotal component in the realm of machine learning applied to flow physics. It plays a crucial role in preprocessing data, constructing and training predictive models, and subsequently making predictions based on input data (Majchrzak, Marciniak-Lukasiak and Lukasiak, 2023). Here's a comprehensive breakdown of the code's functionality and its interplay with both machine learning and flow physics concepts:

Importing Required Libraries: The code shown in Figure 5-32 begins by importing essential libraries such as pandas, numpy, and various modules from the sklearn library for data preprocessing, model training, and evaluation.

```

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.svm import SVR
from sklearn.preprocessing import PolynomialFeatures, MinMaxScaler
from sklearn.pipeline import make_pipeline
from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error
from sklearn.linear_model import Ridge

```

Figure 5-32: Import libraries used in the model.

Loading and Preprocessing Data: The code loads data from a CSV file (path specified as ".csv") into a pandas DataFrame.

It then splits the data into input features (denoted as X) and output variables (denoted as Y).

Data Scaling: Min-Max scaling is applied to both input (X) and output (Y) data to normalize their ranges between 0 and 1. This step is vital for effective training and prediction in machine learning models.

Train-Test Split: The data is split into training and testing sets using the `train_test_split` function from `sklearn.model_selection`.

The loading and preprocessing code used as shown Figure 5-33.

```

# Load data
data = pd.read_csv(".csv")

# Split data into input features and output variable
X = data.iloc[:, 0:8] # input features
Y = data.iloc[:, 8:20] # output variable

# Create a scaler for the input variables
scalerX = MinMaxScaler()
X_scaled = scalerX.fit_transform(X)

# Create a scaler for the output variables
scalerY = MinMaxScaler()
y_scaled = scalerY.fit_transform(Y)

# Split data into training and testing sets
X_train, X_test, Y_train, Y_test = train_test_split(X_scaled, y_scaled, test_size=0.2, random_state=0)

# Build linear regression model
lin_reg = LinearRegression()
lin_reg.fit(X_train, Y_train)
lin_reg_pred = lin_reg.predict(X_test)

```

Figure 5-33: Loading and preprocessing section

Model Building and Training:

The code as shown in Figure 5-34 builds and trains various regression models using different algorithms:

- Linear Regression (LinearRegression)
- Polynomial Regression (PolynomialFeatures + Ridge)
- Decision Tree Regression (DecisionTreeRegressor)
- Random Forest Regression (RandomForestRegressor)
- Support Vector Regression (SVR)

```
# Build polynomial regression model
degree = 2 # degree of polynomial
poly_reg = make_pipeline(PolynomialFeatures(degree), Ridge(alpha=0.1))
poly_reg.fit(X_train, Y_train)
poly_reg_pred = poly_reg.predict(X_test)
lin_reg_mae = mean_absolute_error(Y_test, lin_reg_pred)
lin_reg_mse = mean_squared_error(Y_test, lin_reg_pred)
lin_reg_rmse = mean_squared_error(Y_test, lin_reg_pred, squared=False)

# Build decision tree regression model
dt_reg = DecisionTreeRegressor(random_state=0)
dt_reg.fit(X_train, Y_train)
dt_reg_pred = dt_reg.predict(X_test)
poly_reg_mae = mean_absolute_error(Y_test, poly_reg_pred)
poly_reg_mse = mean_squared_error(Y_test, poly_reg_pred)
poly_reg_rmse = mean_squared_error(Y_test, poly_reg_pred, squared=False)

# Build random forest regression model
rf_reg = RandomForestRegressor(n_estimators=10, random_state=0)
rf_reg.fit(X_train, Y_train)
rf_reg_pred = rf_reg.predict(X_test)

# Build support vector regression model for each output variable
svr_regs = []
for i in range(Y_train.shape[1]):
    svr_reg = SVR(kernel='rbf', C=0.1)
    svr_reg.fit(X_train, Y_train[:, i])
    svr_regs.append(svr_reg)
```

Figure 5-34: Regression models code

The code's overarching purpose is to demonstrate how machine learning techniques can be applied to predict flow-related outcomes. It addresses the complexities of multiphase flow dynamics and attempts to model the behavior of flow based on input features. The data preprocessing steps, feature scaling, and

model training mirror the real-world intricacies of capturing and simulating fluid flow behavior.

By combining machine learning methodologies and the nuances of flow physics, this code represents a potent tool to predict flow data accurately and analyze deviations between simulated and experimental results. It highlights the iterative process of refining models, understanding the relationship between input features and output variables, and ultimately advancing the understanding of complex flow phenomena.

6 Computational Fluid Dynamics and Machine learning

6.1 Introduction

Computational Fluid Dynamics (CFD) is one of the accomplished methods to study the dynamics of different fluid dynamics in the multiphase flow. A multiphase flow system can be complex and influenced by various factors, so there is no generalised mathematical technique to capture all systems, and each system is unique. It is important to develop a strategical simulation methodology specific to the system with the help of literature when needed. This knowledge helps to choose the right geometry creation, meshing techniques, turbulence model selection and precisely analysing the results.

The interface transferring momentum, geometry, defining the right boundary conditions specific to the problem in the hand, mesh independence study and outcome parameters expected need to be defined clearly, specifically for the simulation objectives. It is also crucial to select the right computational resources when solving a specific problem.

Experiments are usually expensive and a time-consuming process for multiphase flows when it comes to setting up the test rig. CFD involves computational resources which will be expensive and will require time to solve multiphase flow problems as many cells and small time-steps are usually needed. A Machine Learning (ML) approach leverages this through integrating CFD and experimental data using computational algorithms, to understand the flow patterns and make predictions. The large data set required from the experimental and CFD. Experimental has a limitation is populating the for different geometries, where CFD can help in generating data where experimental has its limitations(Rizzo et al., 2020).

By feeding proper inputs like superficial flow velocities along with other flow properties, the output data such as phase variations can learn by the ML algorithm to predict the results for the unseen cases. Regression models and deep learning techniques can be used to tackle multiphase flow problems(Rizzo et al., 2020). These techniques can be used for tasks such as flow regime

classification, phase distribution prediction, bubble, or droplet tracking, or even optimization of flow parameters. Integrating ML concepts along with experimental test rig data and CFD simulation data will enable large scale studies, with less computational costs and will provide more insights for complex flow phenomena. The successful implementation of ML technique in multiphase flow depends on the data available, the cleaning of this data and the selection of the right algorithm. Overall, the incorporation of machine learning techniques in CFD multiphase flow holds great potential for advancing our understanding and predictive capabilities in studying complex fluid systems, opening new avenues for research and engineering applications.

6.2 Simulation requirement and approach

A requirement is to use CFD simulation data, empirical data, and machine learning to accurately represent the characteristics of serpentine pipes with a diameter range of 50 mm to 150 mm for the given flow conditions. There are several limitations when using CFD modelling to examine serpentine pipes with multiphase flow in very turbulent environments. There has not been much research done on the effects of three bends in the multiphase flow domain, especially when it comes to bigger diameter pipes. The experimental method takes time, and adding this impact to CFD simulations makes things more complicated. Because of this, there is a chance for research to use machine learning approaches to investigate the impact of bent sections in various pipe sizes. Figure 6-1 shows the flow diagram of the CFD simulation process.

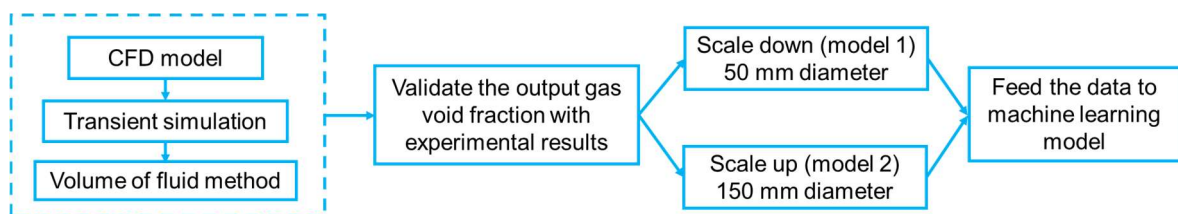


Figure 6-1: Simulation overview flow chart

6.3 Geometries used in the study.

In this section, the geometries used for the numerical simulations is detailed, which is crucial since capturing the real physical system will help us to get accurate results for the different flow conditions using validated model. The geometries are created with the SOLIDWORKS 2021 CAD tool. Figure 6-2 shows the schematic section of 101.56 mm serpentine pipes. There are six sections marked in the geometries which will be used for the study purpose, those six sections are marked in the straight downward and upward sections, located in the top, middle and bottom regions of each vertical pipe section.

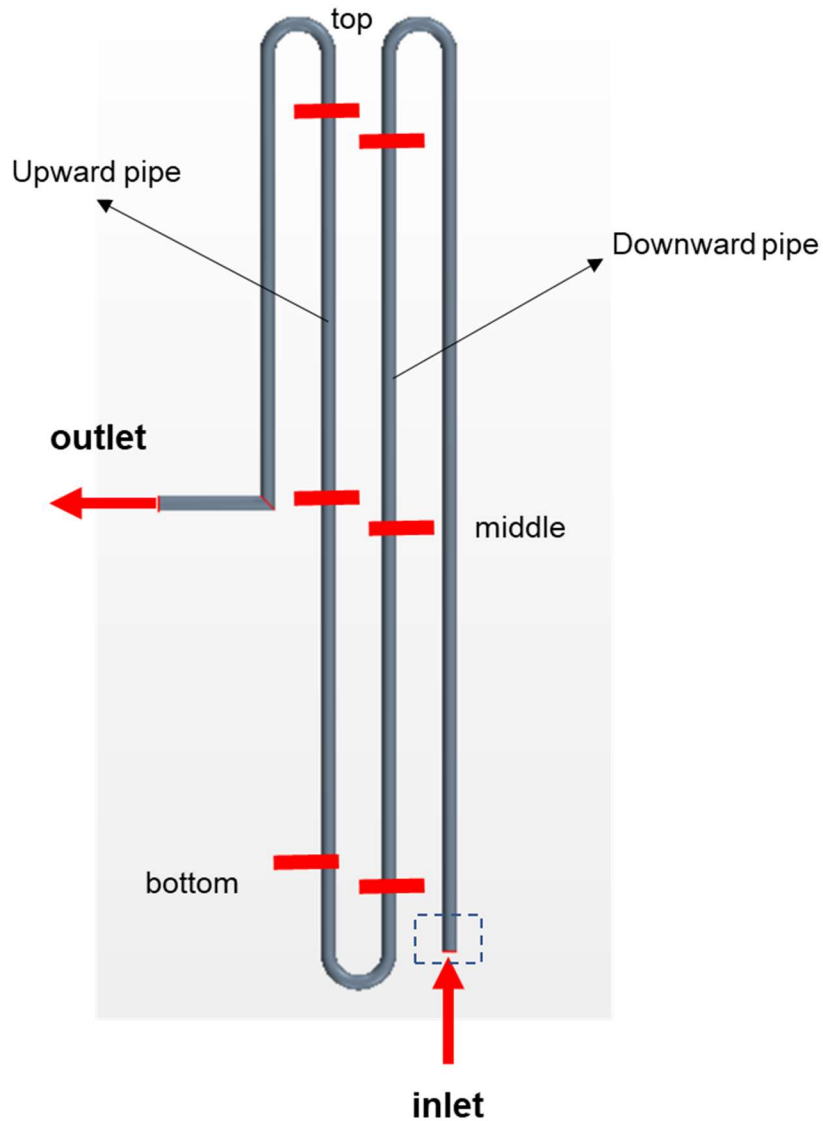


Figure 6-2: CAD model representing the experimental pipe setup

6.3.1 The main piping section.

The pipe diameter remains constant at 101.6 mm throughout the section, comprising three curved bends and four straight sections. The pipeline consists of an inlet section, a downward pipe, an upward pipe, and an outlet section. For the study of flows, particular attention is given to the upward and downward sections.

6.3.2 Model 1 and Model 2 dimensions.

Two CAD models were created, one with a 50 mm diameter and another with a 150 mm diameter, both designed to closely resemble the base model in representation.

6.3.3 Inlet and outlet sections

The inlet section is modified using an iterative approach to closely match the experimental setup. In actual setup, liquid is fed from the bottom of the pipe, while gas is introduced through a 90-degree pump placed in between the straight pipe sections. For simulation purposes, the gas feed pipe is excluded. Both gas and liquid enter the pipe through the same inlet. The diameter of the inlet is reduced to accommodate both gas and liquid flows, as show in Figure 6-3. The outlet section is maintained like the experimental one.

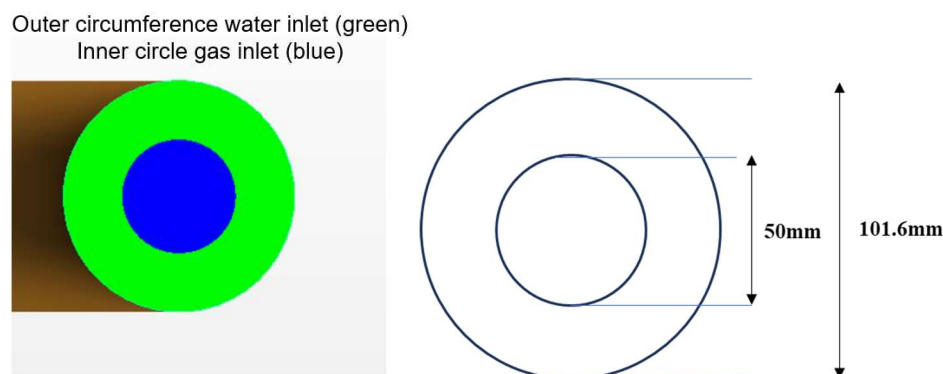


Figure 6-3: Inlet section for liquid and gas

The area of the inlets and outlet are tabulated in the Table 6-1, the inlet 8082.02 mm² is divided into two section water area is being 6124.62 mm² and the air is 1957.39mm²

Table 6-1: Inlet air and water diameter area

Location	Area (mm²)
Inlet (water)	6124.62
Inlet (air/gas)	1957.39
Outlet	8082.02

6.4 Boundary Conditions

The simulation involved the utilization of both air and water. Initially, the computational domain was fully initialized with water, followed by the introduction of air and water through their respective inlets. The properties of air and water are presented in Table 6-2. The velocity boundary conditions utilized in the experiment data were applied. These input conditions were employed across all optimization scenarios to facilitate a numerical solution that could be compared with experimental measurements. At the inlet, a mass flow inlet boundary type was implemented, specifying the mass flow rates of water and gas, along with the liquid volume fraction. A pressure outlet boundary condition was utilized at the outlet.

Table 6-2: Fluid properties

Material properties	
Water	
Density	997.561 kg/m ³
Dynamic Viscosity	8.8871E-4 Pa-s
Air	
Density	1.18415 kg/ m ³
Dynamic Viscosity	1.85508E-5 Pa-s

6.5 Mathematical modelling and CFD

The computational fluid dynamics (CFD) research employs the volume of fluid (VOF) method in Simcenter STAR-CCM+ 2206.0001 (17.04.008-R8). The VOF method is selected over other techniques due to its ability to accurately handle multiphase flows, making it particularly suitable for simulating the complex behaviour of serpentine flows. Additionally, the adoption of the k-omega turbulence model enhances the accuracy of this study's simulations. The k-omega model is well-suited for capturing turbulent behaviours in intricate flow scenarios like those found in serpentine configurations (Iaccarino and Elkins, 2006). Its proficiency in handling turbulent flows, even in scenarios with high curvature, recirculation zones, and intricate flow geometries, makes it a favorable choice for capturing the intricate fluid dynamics inherent to the serpentine flow problem.

6.5.1 Assumptions

The study encompasses a three-dimensional fluid flow scenario characterized by turbulence and incompressibility. The analysis assumes no involvement of chemical reactions within the fluid, and the fluid properties are maintained as constants throughout the investigation. This approach facilitates a comprehensive understanding of the flow dynamics and behaviours.

6.5.2 VOF theory

The VOF model was employed within Starccm+ to capture the phase distributions in the study. This model hinges on conservation equations for the engaged phases, incorporated into a single set for VOF modelling (Katopodes, 2019). The consideration of conservation equations was restricted to an isothermal system, excluding mass transfer and phase change dynamics. The equation governing mass conservation is expressed as follows:

$$\frac{\partial(\rho)}{\partial t} + \nabla \cdot (\rho u) = 0 \quad \text{Equation 6-1}$$

where ρ represents the density, t is the time, and u denotes the velocity.

The momentum equation, encompassing external factors like static pressure (p), external body force (F), and gravitational body force (ρg), is given by:

$$\frac{\partial(\rho u)}{\partial t} + \nabla \cdot (\rho u \cdot u) = -\nabla p + \nabla \cdot [\mu(\nabla u + \nabla u^T)] + \rho g + F \quad \text{Equation 6-2}$$

with μ the dynamic viscosity, ∇ the gradient operator, and g is the acceleration due to gravity.

In the context of dispersed flow interactions, parameters p , F , and ρg are indicative of static pressure, external body force, and gravitational body force, respectively. The interface between the primary phase (water) and the secondary phase (air) was tracked simultaneously, involving conservation equations for the secondary phase's void fraction (α_o):

$$\partial \frac{(\rho_g \alpha_g)}{\partial t} + \nabla \cdot (\rho_g \alpha_g u) = 0 \quad \text{Equation 6-3}$$

It is worth noting that the volume fraction equation was not solved for the primary phase (water). Instead, the volume fraction of the primary phase was determined through the constraint:

$$\alpha_g + \alpha_w = 1 \quad \text{Equation 6-4}$$

This constraint signifies that the total volume fraction of either gas or water within the control volume ranges between 0 and 1, representing the complete filling of the control volume by either phase.

The properties of matter, such as density (ρ) and viscosity (μ), within the transport equations were calculated within each control volume, accounting for the presence of component phases. In the context of two-phase flow (gas-water), the density and viscosity in each cell were evaluated as:

$$\rho = \alpha_g \rho_g + \alpha_w \rho_w \quad \text{Equation 6-5}$$

$$\mu = \alpha_g \mu_g + \alpha_w \mu_w \quad \text{Equation 6-6}$$

In these equations, subscripts g and w refer to gas and water, respectively, and α_o and α_w represent their respective volume fractions. These formulations underpin the comprehensive analysis of the fluid dynamics in the studied system.

6.5.3 k-omega turbulence model equations

Transport Equation for Turbulence Kinetic Energy (k):

The time derivative of the density multiplied by the turbulence kinetic energy (k) plus its convection term along the x direction is balanced by the diffusion term in the y direction (Wilcox, 2012). This equation represents the evolution of turbulence kinetic energy in the flow.

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho k u)}{\partial x} = \frac{\partial}{\partial x} \left(\left(\mu + \frac{\mu_t}{\sigma_k} \right) \frac{\partial k}{\partial y} \right) + G_k - \rho \varepsilon - Y_M + S_k \quad \text{Equation 6-7}$$

where:

ρ is the density of the fluid.

k represents the turbulence kinetic energy.

u denotes the velocity in the x-direction.

μ stands for dynamic viscosity.

μ_t is the turbulent viscosity.

σ_k is the turbulent Prandtl number for kinetic energy.

G_k represents the generation of turbulence kinetic energy.

S_k is a source term related to the turbulence model.

Y_M represents the contribution of fluctuating dilatation in compressible turbulence.

ε represents the dissipation rate of turbulence kinetic energy.

The time derivative of the density multiplied by the turbulence dissipation rate (ε) plus its convection term along the x direction is balanced by the diffusion term in the y direction. This equation describes the evolution of turbulence dissipation in the flow.

$$\frac{\partial(\rho\varepsilon)}{\partial t} + \frac{\partial(\rho\varepsilon u)}{\partial x} = \frac{\partial}{\partial x} \left(\left(\mu + \frac{\mu_t}{\sigma_\varepsilon} \right) \frac{\partial \varepsilon}{\partial y} \right) + C_{1\varepsilon} \frac{k}{\varepsilon} (G_k + C_3 G_b) - C_{2\varepsilon} \rho \frac{\varepsilon^2}{k} + S_\varepsilon \quad \text{Equation 6-8}$$

where:

ε represents the turbulence dissipation rate.

$C_{1\varepsilon}$, $C_{2\varepsilon}$, and $C_{3\varepsilon}$ are model constants.

G_b represents the buoyancy production of turbulence kinetic energy.

S_ε is a source term related to the turbulence model.

Other variables have the same meanings as in the previous equation.

6.5.4 CFD model physics settings:

Table 6-3: Physics setting solver

Setting Description	Value/Setting
Turbulence Model	k-Omega
Turbulence Intensity	0.01
Turbulent Velocity Scale	1.0 m/s
Turbulent Viscosity Ratio	100
Linear Constitutive Option (a1)	0.31
Compressibility Correction	Enabled
Normal Stress Term	Enabled
Low Re Damping Modification	Disabled
Material Properties (Water)	Density: 997.561 kg/m ³
	Dynamic Viscosity: 8.8871E-4 Pa-s
Material Properties (Air)	Density: 1.18415 kg/m ³
	Dynamic Viscosity: 1.85508E-5 Pa-s
Surface Tension	0.072 N/m
Artificial Viscosity	0.1
Segregated Flow Solver Settings	Positivity Rate Limit: 0.2
	AMG Solver Cycles: 30
	Convergence Tolerance: 0.1
VOF Modeling (HRIC Scheme)	Sharpening Factor: 0.0
	Angle Factor: 0.05
	CFL_l: 0.5, CFL_u: 1.0
Eulerian Phases	Water (Primary Phase, Index 1)
	Air (Secondary Phase, Index 2)
Simulation Duration	30 seconds (149,999 time steps)

In this section, the essential physics settings employed in the CFD simulation are detailed. The choice of the k- Ω turbulence model is based on its capability to capture turbulence dynamics effectively. This model accounts for generation, dissipation, and transport of turbulence kinetic energy and dissipation rate, enabling accurate simulation of turbulent flows in fluid mechanics. The simulation employs the VOF method to investigate dynamic interactions between water and air. This method is chosen for its steadiness in prediction and computational cost. Furthermore, specific physics settings are adjusted to optimize quality and convergence, offering a comprehensive setup to explore multiphase flow interactions as shown in Table 6-3.

6.5.5 CFD model Solver settings:

Table 6-4: CFD solver settings

Setting Description	Value/Setting
Time-Step	0.001 seconds
Temporal Discretization Scheme	2nd-order
Simulation Duration	30 seconds (150,000 iterations)
Partitioning Method	"Per-Continuum"
Wall Distance Calculation	Dynamic and Non-Frozen
Turbulence Intensity	0.01
Turbulent Velocity Scale	1.0 m/s
Turbulent Viscosity Ratio	100
Segregated Flow Solver Settings	Positivity Rate Limit: 0.2
	AMG Solver Cycles: 30
	Convergence Tolerance: 0.1
VOF Modeling (HRIC Scheme)	Sharpening Factor: 0.0
	Angle Factor: 0.05
	CFL _I : 0.5, CFL _u : 1.0
Turbulent Viscosity Solver	Under-Relaxation Factor: 0.8
	AMG Solver (similar settings)
Stopping Criteria	Inner Iterations: 5
	Steps: 1000
	Physical Time: 30 seconds
Solution Storage	Efficient Utilization
Single-Step Under-Relaxation	Factor: 0.9
Solution History Monitoring	At Every Iteration
Current Solution's Iteration	Number: 150,000
Time Step	29,999 (Solution Time: 30 seconds)

In this section, the solver settings utilized in the CFD simulation are outlined as shown in Table 6-4 providing insights into their rationale and significance. The implicit unsteady approach with a small time-step of 0.001 seconds is adopted to accurately capture fluid behaviour over time. Partitioning through the "Per-Continuum" method ensures efficient resource utilization. Dynamic and non-frozen wall distance calculation enhances boundary layer modelling. The k-Omega method is chosen for turbulence modelling due to its robustness. The segregated flow solver employs various techniques, including velocity and pressure under-relaxation factors, to achieve numerical stability and convergence. The VOF modelling, using the HRIC scheme, captures interface behaviour precisely. By detailing various settings such as convergence criteria and solution history monitoring, the simulation setup ensures the generation of reliable and accurate data for comprehensive analysis.

6.6 Mesh Selection Study

Grid generation is an essential step in any analysis, and the choice of mesh size depends significantly on the Reynolds number. This selection process typically involves categorizing the mesh sizes into three types: fine, medium, and coarse. Estimation of y^+ is used to choose the correct mesh for validation the CFD model. The average y^+ in the domain and average void fraction is measured at the top of the downward pipe.

6.6.1 Flow input data

As previously mentioned, the experimental data from the slug flow zone will be utilized to conduct a study on mesh independence and to validate the CFD model. Table 6-5 presents the three sets of experimental data on void fraction used in this investigation.

Table 6-5 : Experimental data used for the study.

positions	down5D (top)	down30D (middle)	down46D (bottom)	up5D (top)	up28D (middle)	up47D (bottom)
water – 1 m/s, air 0.2 m/s	58.55	36.72	21.71	12.6	17.7	17.64
water – 1 m/s, air 1 m/s	71.27	74.51	68.81	30.51	38.82	38.16
water – 1 m/s, air 2 m/s	73.52	78.5	77.76	44.38	54.93	51.59

6.6.2 Boundary conditions

For this mesh study and model validation, water of superficial velocity of 1 m/s and air superficial velocity of 1 m/s case are employed. The mass flow rate was calculated based on these velocities and subsequently utilized as the inlet boundary conditions. The flow rate values are tabulated in Table 6-6

Table 6-6: Boundary conditions – inlet flow rates

Water velocity (m/s)	Gas velocity (m/s)	Water inlet area (m ²)	Gas inlet area (m ²)	Mass flow rate water (kg/s)	Mass flow rate gas (kg/s)
1	1	0.008011847		8.01184666	0.01020709

6.6.3 Y+ criteria

Capturing the sub-viscous layer is important to capture the wall effect in the simulation, so the y^+ values should be less than 5 particularly for wall-bounded flows (Abuhatira, Salim and Vorstius, 2021). The choice of keeping y^+ below 5 is based on the desire to accurately resolve the near-wall region and capture the viscous sublayer. It is also due to the choice of turbulence model, here k- Ω .

Reason for Using $y^+ < 5$:

In CFD simulations, the y^+ value plays a crucial role in accurately modelling the behaviour of fluid near the wall boundary. The y^+ value represents the ratio of the distance from the wall (y) to the molecular viscous length scale (μ/ρ). Keeping the y^+ values below 5 is a widely accepted practice for several reasons:

- **Resolving viscous sublayer:** The region near the wall, known as the viscous sublayer, experiences significant velocity gradients and viscous effects. Resolving this region accurately is essential for capturing the correct boundary layer behaviour and turbulence characteristics. y^+ values less than 5 ensure that the mesh provides sufficient resolution to accurately simulate these near-wall effects.
- **Accurate turbulence modeling:** Most turbulence models used in CFD simulations are designed to work well in the logarithmic region of the velocity profile. By keeping y^+ values low, to ensure that the model is applied in a region where it is more accurate and reliable, improving the overall accuracy of simulation results.
- **Minimizing sensitivity to mesh:** y^+ values less than 5 help reduce the sensitivity of simulation results to the mesh density near the wall. Higher y^+ values may lead to inaccurate predictions if the mesh is not fine enough to resolve the near-wall gradients adequately.
- **Avoiding wall-interference:** Larger y^+ values can lead to mesh-induced interference with the near-wall behavior, affecting the predicted flow patterns and boundary layer characteristics. Keeping y^+ below 5 helps mitigate these interference effects and provides a more accurate representation of the flow.

The y^+ definition is:

$$y^+ = \frac{\rho U y}{\mu} \quad \text{Equation 6-9}$$

where:

U is the velocity of the fluid,

y is the distance from the wall, and

μ is the dynamic viscosity of the fluid.

6.6.4 Mesh and grid

During the formulation of precise meshing strategies for the CFD model development, three notable techniques come to the forefront, and it is used in this work: the surface remesher, trimmed cell, and prism layer. The surface remesher serves as a refining mechanism for mesh elements along the geometry's surfaces, ensuring enhanced resolution in regions characterized by intricate shapes, curvature, or complex flow interactions. Complementing this, the trimmed cell technique facilitates the representation of intricate geometries featuring cutouts or intricate shapes with precision, yielding a mesh that closely conforms to the geometry's surface boundaries as shown in Figure 6-4. Additionally, the incorporation of a prism layer within the mesh strategy proves invaluable near solid boundaries. This layer, composed of stacked prismatic elements oriented perpendicular to the surface, impeccably captures the intricate flow characteristics within the viscous sublayer.

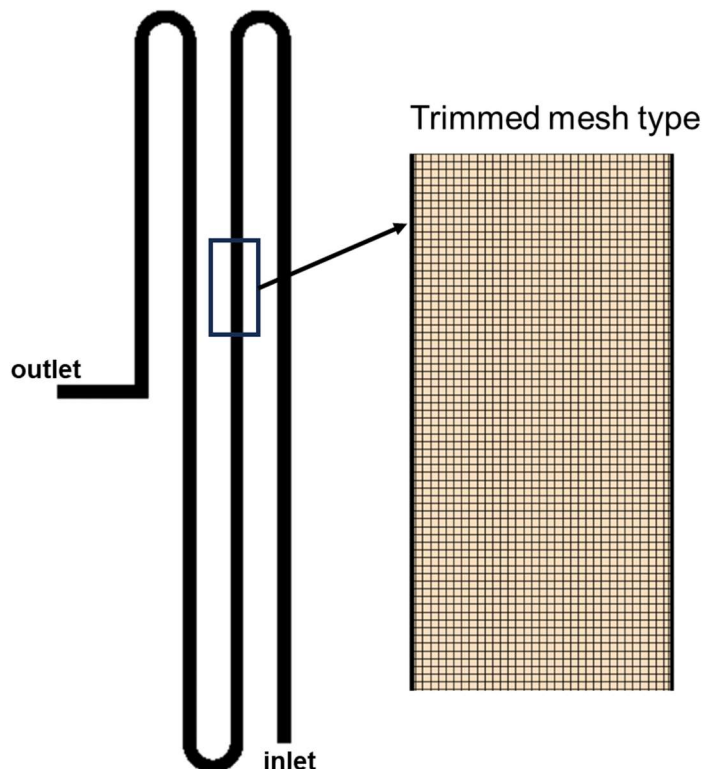


Figure 6-4: Trimmed mesh

By amalgamating these techniques, a comprehensive mesh emerges that adeptly captures both geometric intricacies and near-wall phenomena. In this investigation, three different mesh sizes were examined: coarse (8 mm), medium (6 mm), and fine (3 mm). The prism layer was systematically varied between 2, 4, 6, and 8, employing a stretching factor of 1.5. This adjustment was made to ensure that the wall y^+ remained below 5, enabling the accurate capture of sub-viscous layers. These adjustments were made in response to the experimental results, where the void fraction was carefully measured to validate the model's performance. The mesh dependence study results are tabulated in the Table 6-7.

Table 6-7: Mesh dependence study results

Mesh	Mesh size (mm)	Prism layer	Prism layer stretching	Mesh count	Y+	void fraction downward pipe (top) (%)	
						Experimental	CFD simulation
Coarse	8	2	1.5	603281	15	71.27	28
		4	1.5	824721	8		26
		6	1.5	1046161	6		29
		8	1.5	1267601	3		34
Medium	6	2	1.5	1298105	15		32
		4	1.5	1700009	8		34
		6	1.5	2101913	6		37
		8	1.5	2503817	3		41
Finer	3	2	1.5	10470653	15		58
		4	1.5	11560914	8		58
		6	1.5	12651175	6		60
		8	1.5	13741437	3		61

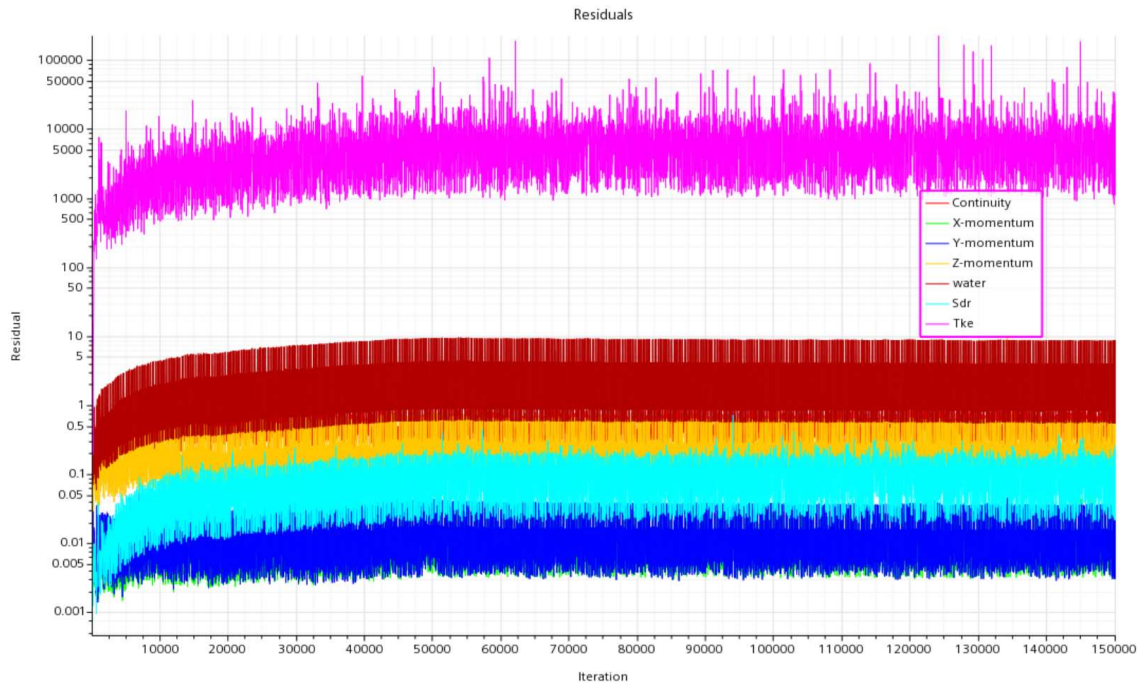


Figure 6-5: Residuals plot (finer mesh, 8 prism layer)

The convergence analysis of the different meshes in relation to the void fraction was conducted by closely monitoring the void fraction values at the upper section of the downward pipe. Figure 6-5 shows the residuals plot, focusing on the finer mesh case with a mesh size of 3 mm and 8 prism layers.

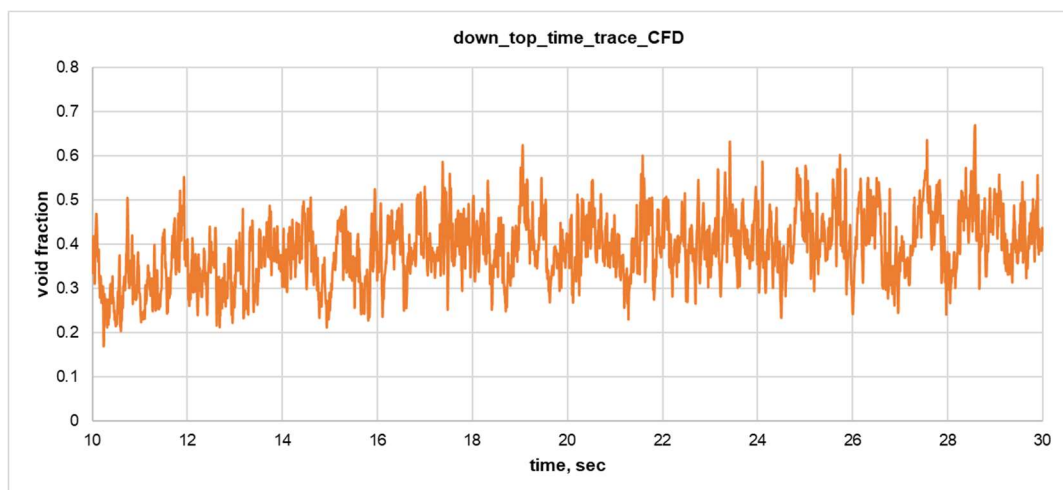


Figure 6-6: Void fraction convergence plot (finer mesh, 8 prism layer)

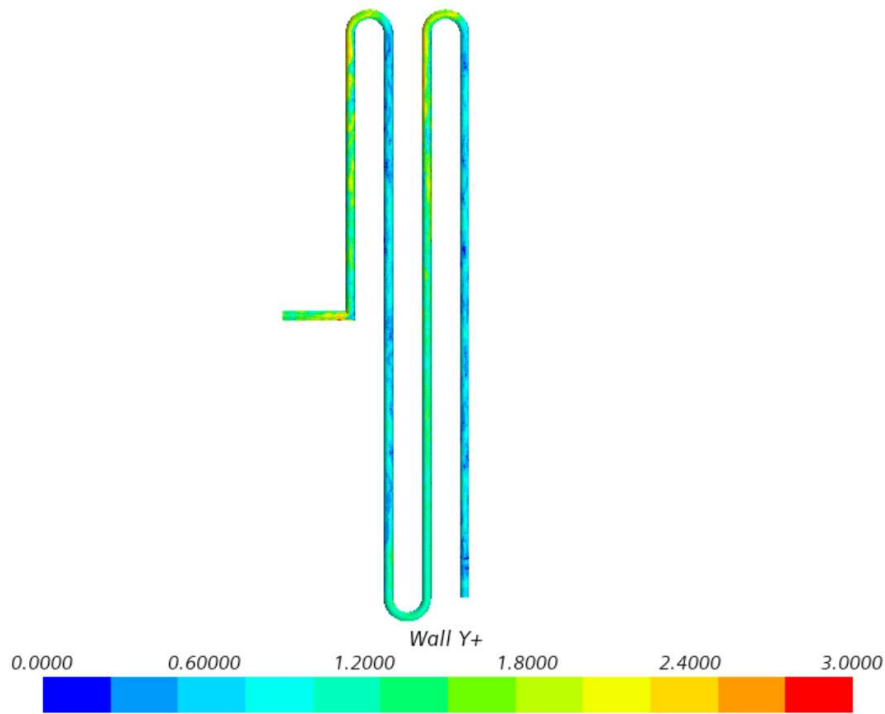


Figure 6-7: Wall y^+ plot (finer mesh, 8 prism layer)

To obtain the numerical results, transient simulations were performed over a duration of 30 seconds. The void fraction variation plot in Figure 6-6 shows that the flow is developed after about 30 seconds. Additionally, the wall y^+ figure in Figure 6-7 shows that the entire domain maintains a wall y^+ value below 5.

The Figure 6-8 illustrates the relationship between mesh count and void fraction. Notably, the finer mesh configuration exhibits a more favourable convergence trend. Within this context, the minimum wall y^+ condition was selected to ensure proximity to experimental void fraction values. Consequently, the mesh configuration with a finer mesh size of 3 mm and 8 prism layers emerged as the optimal choice for further investigation and study.

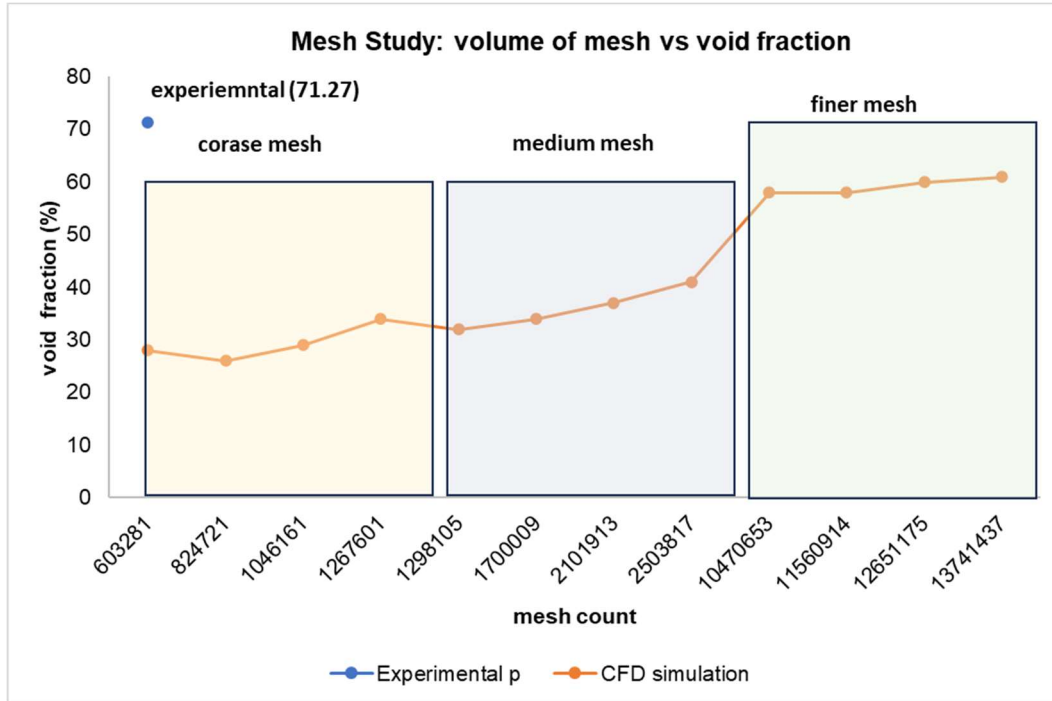


Figure 6-8: Mesh convergence plot

In summary, through a comprehensive convergence analysis utilizing the finer mesh setting, coupled with transient simulations and maintaining wall y^+ values, the study successfully identified the mesh configuration. This optimized configuration, consisting of a finer mesh size of 3 mm and 8 prism layers, will serve as the foundation for advancing and refining the ongoing study.

6.6.5 Multi region mesh refinement study

The final mesh configuration selected for the model's results holds significant promise for further enhancing the model's accuracy and performance (Jimack et al., 2002). The volume fraction plot illustrates intriguing findings, shedding light on complex flow dynamics. Notably, after the bend, both the upper and lower positions exhibit turbulent flow characteristics, indicative of high-energy dissipation and effective mixing of the two fluid phases. Within the bend itself, a distinct pattern emerges: water tends to flow closer to the curvature, while air occupies the upper portion.

Particularly noteworthy is the behavior within the bottom part of the serpentine pipe bend. Here, a compelling phenomenon unfolds, as air predominantly

occupies the inner side of the bend, while water flows along the outer side. Given the focal points of the upward and downward sections of the pipe, the logical step is to further refine the mesh utilizing a multi-block approach.

The multi-region mesh strategy is a sophisticated technique in CFD that subdivides the computational domain into distinct regions and applies mesh refinement strategies tailored to each region's specific complexities. This is especially advantageous for flows with notable variations in length scales, gradients, or turbulence profiles across distinct domain segments. In this thesis, the multi-region mesh approach is used to capture the intricate flow complexities under investigation. By deploying varying mesh resolutions in regions where intricate or turbulent flow behaviors manifest, a finer level of precision can be achieved without unwarranted computational overhead in regions of lesser complexity.

For example, finer mesh elements can be strategically employed to resolve intricate flow phenomena such as separation and vortices in regions where they occur. Conversely, regions characterized by smoother flow patterns can be assigned coarser mesh elements, optimizing computational efficiency while still encapsulating overarching flow trends.

Adapting mesh resolutions in response to inherent flow complexities offers a compelling avenue for attaining a faithful representation of fluid dynamics. By leveraging this approach, simulations gain heightened fidelity, empowering deeper insights into underlying flow physics and facilitating more informed engineering judgments. This deliberate mesh refinement strategy amplifies the value of CFD simulations, fostering a more profound comprehension of the intricate interplay within the fluid flow.

Figure 6-9 shows the volume fraction of air plot, it is observed the more turbulent flow after the bend, it is better to capture the flow effectively to match the simulation closely to the real time flow, for further studies the mesh is refined for different regions with maintaining the same prism layers by using multilevel meshing approach.

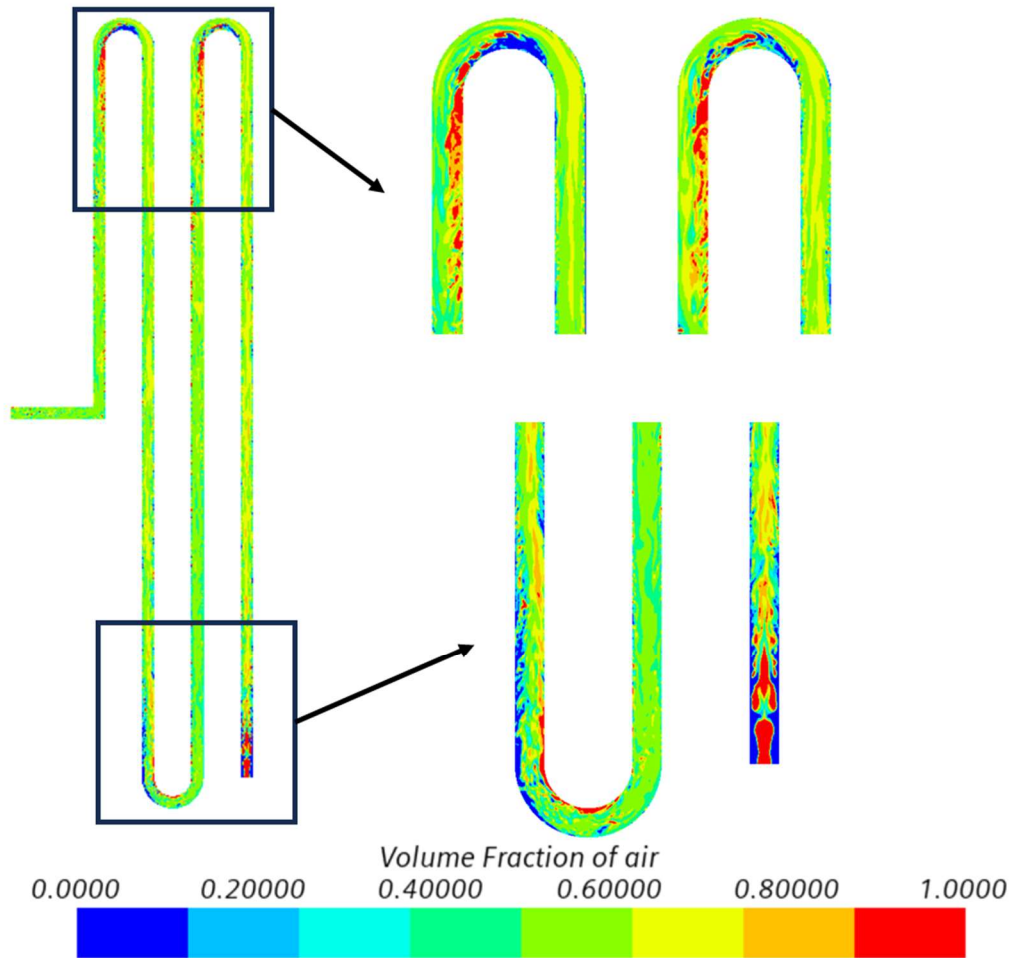


Figure 6-9: Volume fraction air plot (finer mesh, 8 prism layer)

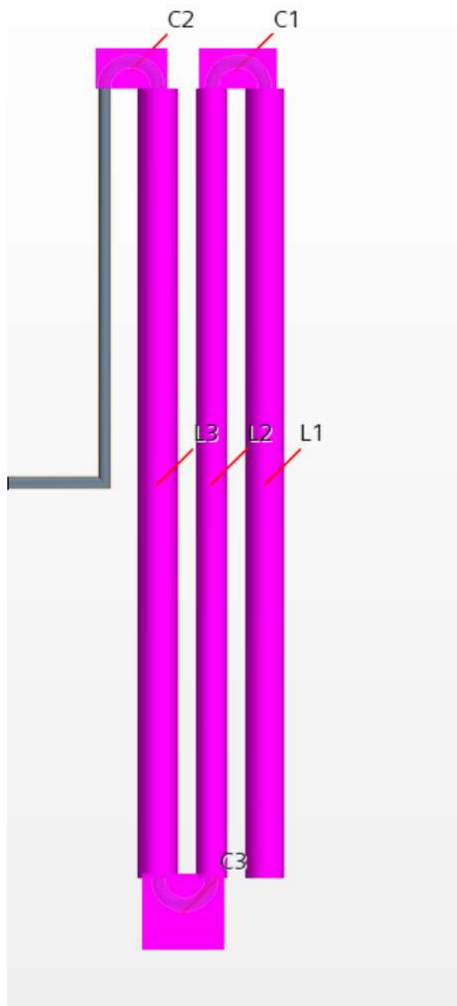


Figure 6-10: Multi region mesh

As shown in Figure 6-10, the mesh is separated into different regions and the mesh size is increased further in specific regions, as per Table 6-8..

Table 6-8: Multi-region mesh size

Region	Mesh size (mm)
L1	3
L2	2.5
L3	2.5
C1	2
C2	2
C3	2

6.7 CFD Model Validation

The validation of the CFD model involves a refined mesh configuration, and this simulation is meticulously compared against experimental results. This validation procedure encompasses three distinct flow conditions:

- **Air Flow at Different Velocities:** The study considers air flow at velocities of 0.15 m/s for Case 1, 1 m/s for Case 2, and 9 m/s for Case 3, while maintaining a constant water velocity of 1 m/s. Flow rates are precisely calculated and adopted as flow boundary conditions.

To ensure a robust validation process, three key parameters are evaluated for direct comparison with experimental data. The validation is conducted exclusively on the downward pipe, focusing on the following parameters:

- a) **Time-Average Void Fraction:** This parameter offers insights into the overall distribution of voids within the multiphase flow, serving as a crucial metric for comparison.
- b) **Probability Density Function (PDF):** The PDF characterizes the likelihood of void fraction occurrences at different levels. Its evaluation in both the CFD and experimental data facilitates a comprehensive comparison.
- c) **Time Trace Void Fraction Variation:** Tracking the temporal evolution of void fraction provides valuable dynamic information. The analysis of this parameter across CFD simulations and experimental observations offers a multi-dimensional comparison.

To optimize the simulation time while maintaining accuracy, the CFD simulations were conducted over a 30-second timeframe, with the resultant data employed for the analysis. This approach effectively balances computational efficiency with the capture of pertinent flow behaviours.

6.7.1 Case – 1 air 0.15 m/s & water 1m/s:

The simulation for Case 1 has been executed utilizing the specified boundary conditions. To illustrate this, Figure 6-11 showcases the void fraction distribution plotted across distinct sections: bottom, middle, and top.

Analysing the plot reveals distinctive trends:

- In the top section, air dominance is notably diminished.
- In the middle section, the prevalence of air becomes more pronounced.
- Within the bottom section, a relatively consistent void fraction distribution is observed.

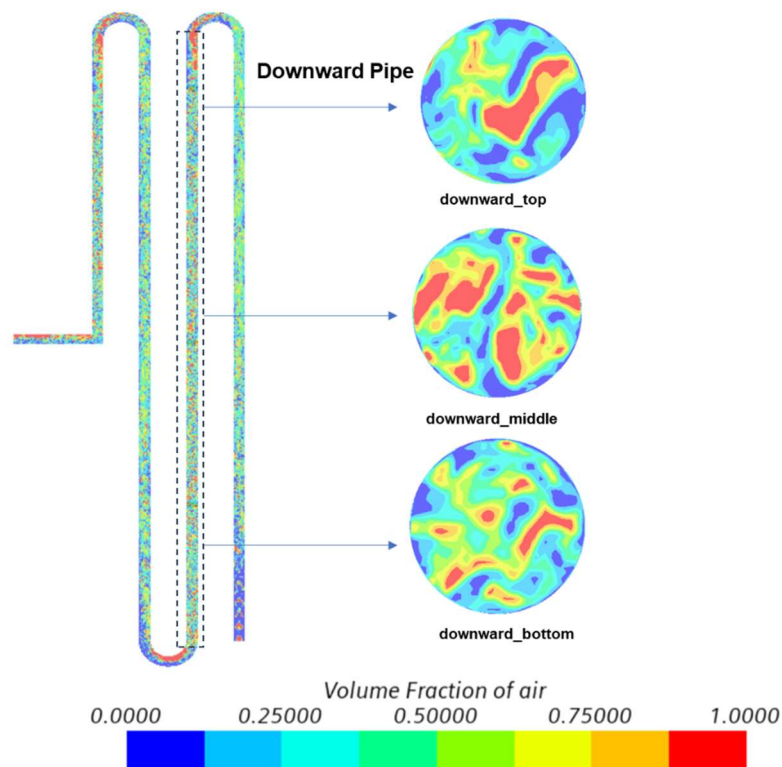


Figure 6-11: Case 1- Void fraction plot downward pipe X section

Figure 6-12 presents a juxtaposition of time-averaged data derived from both experimental and CFD sources, all under the specified boundary conditions. This comparison serves to underscore the divergence and similarities between simulation outcomes and real-world observations.

Analysing the plot reveals intriguing dynamics:

- In the experimental data, the upper section predominantly features air, accounting for approximately 50% of the composition.
- In contrast, the CFD simulation registers a lower air percentage of around 39% in the same top section.
- Within the middle and bottom sections, both the CFD and experimental data converge to showcase air percentages ranging from 10% to 20%.

This comparison, while revealing variations, also highlights certain areas of agreement between the two datasets. The discrepancies in air distribution, particularly in the top section, draw attention to the nuances of the simulation approach and its alignment with the boundary conditions and fluid dynamics being modelled.

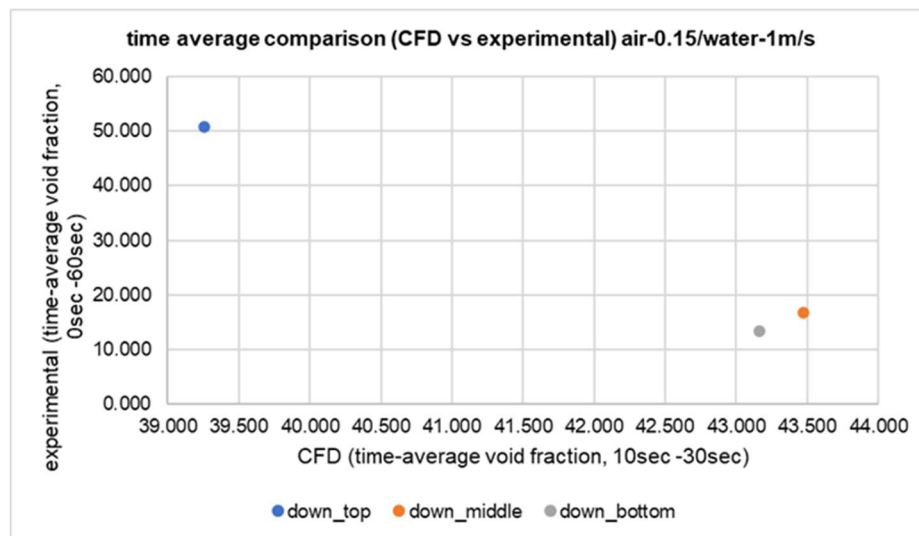


Figure 6-12: Case 1 - Time average void fraction comparison

Figure 6-13, Figure 6-14 and Figure 6-15 shows the variation of void fraction in the three stations, top middle and bottom respectively across a time period of 30 seconds. At the top section the variation is between 0%-80% as sinusoidal, but in CFD it is varying between 20% - 50%, which is not showing good matching with the experimental data. Similarly the

middle and top sections are also not matching either the experimental results,

Given attention to Figure 6-13, Figure 6-14 and Figure 6-15, it is delve into the temporal evolution of void fraction across three distinct stations: top, middle, and bottom, each spanning a 30-second duration. This now discussed.

- In the upper section, the void fraction showcases a sinusoidal fluctuation pattern, oscillating between 0% and 80%. Notably, when juxtaposed with CFD results, this pattern contrasts. In the simulation, the variation assumes a range of 20% to 50%, thus diverging from experimental findings.
- Similarly, the middle and bottom sections also exhibit incongruities when compared to experimental outcomes. This pronounced discrepancy underscores the complexities inherent in capturing real-world multiphase flow behaviors within computational simulations.

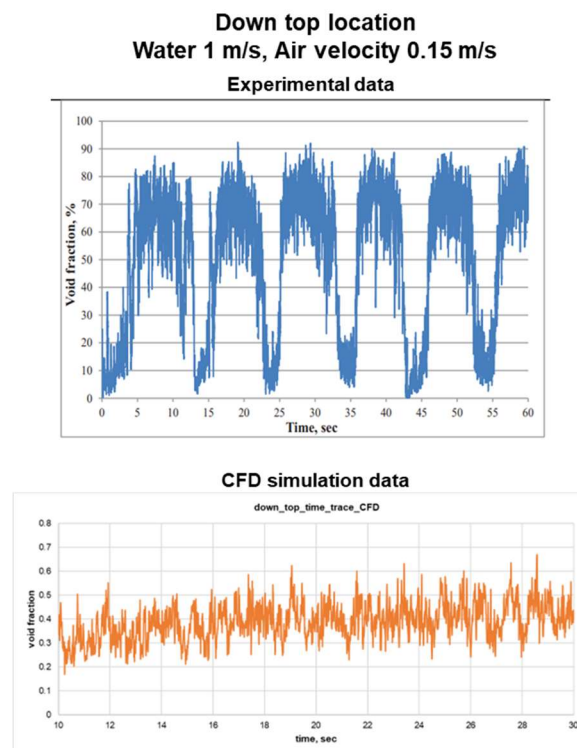


Figure 6-13: Case 1 – Down top time trace CFD vs experiment comparison

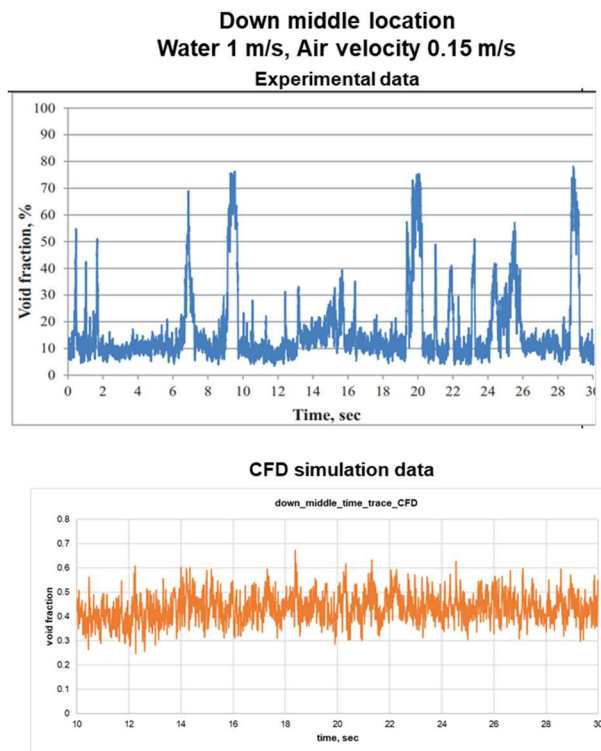


Figure 6-14: Case 1 – Down middle time trace CFD vs experiment comparison

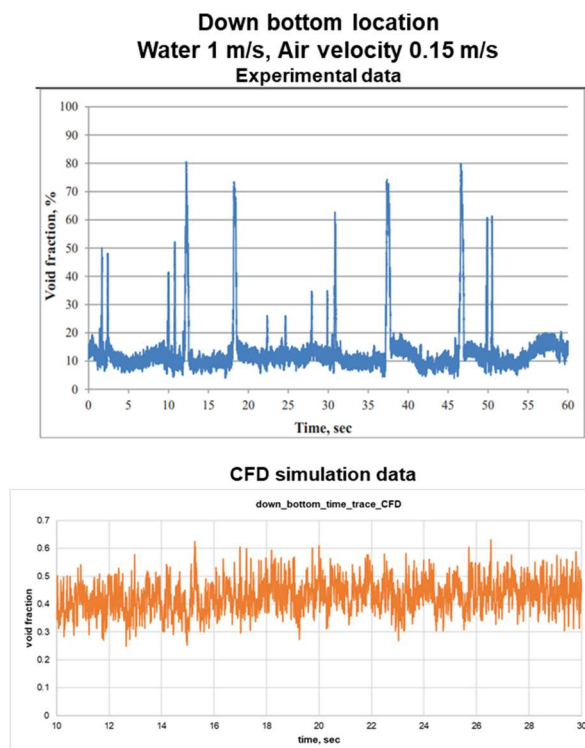


Figure 6-15: Case 1 – Down bottom time trace CFD vs experiment comparison

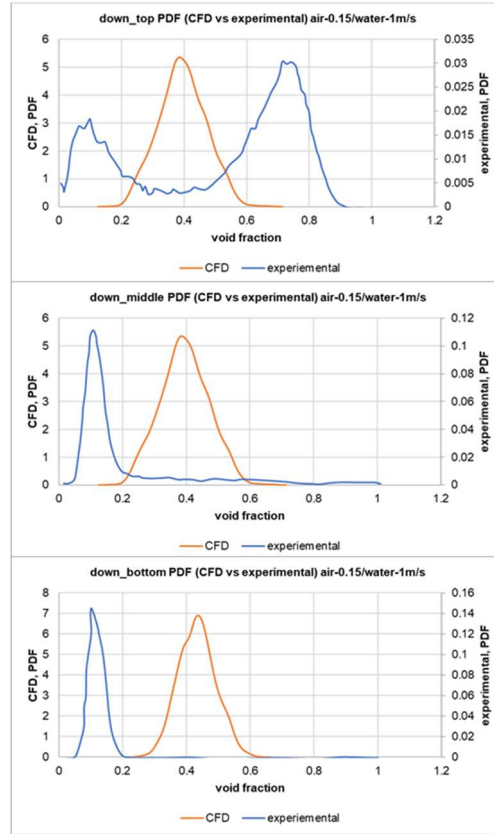


Figure 6-16: Case1 - Probability density function comparison

Building upon the insights gleaned from previous data, now by examine Figure 6-16, which presents the probability density function (PDF) for void fraction at three distinct locations. Our approach entails aligning this data against the backdrop of experimental results, thereby offering a robust comparative analysis.

Observing the plot, a salient distinction emerges:

- Within the CFD results, a conspicuous trend unfolds—a consistency of higher air occupancy spanning the entirety of the downward pipe.
- Conversely, the experimental findings tell a divergent tale, showcasing reduced air occupation at the bottom of the pipe.

The CFD outcomes, while portraying a cohesive pattern, diverge from the experimental data, indicating that subtle intricacies inherent in the multiphase flow dynamics are yet to be fully replicated within the simulation framework. This disparity calls for a nuanced evaluation of simulation parameters, boundary

conditions, and underlying flow physics to bridge the gap between theoretical and real-world outcomes.

Similarly, to Case 1 Case 02 and Case 03 three results are now described to check the accuracy of the model.

6.7.2 Case – 02 air 1 m/s & water 1m/s:

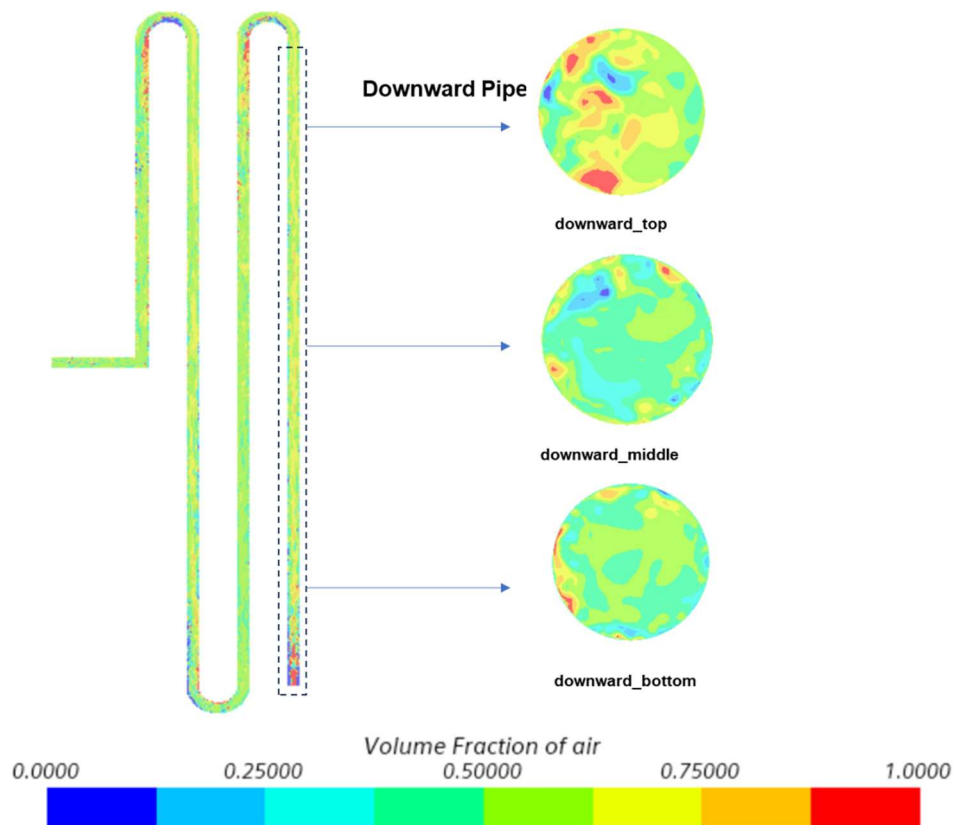


Figure 6-17: Case 2 Void fraction plot downward pipe X section

A detailed assessment of the void fraction distribution for Case 2 with air velocity of 1 m/s and water velocity of 1 m/s, is carried out. This meticulous analysis unveils intricate dynamics, providing a more comprehensive understanding of the multiphase flow behaviour under these specific conditions.

- In the upper section shown in Figure 6-17 and Figure 6-18, an intriguing trend surfaces. The void fraction plot showcases a distinctive pattern where a higher proportion of air is confined.

- This observation remains consistent when compared to the time-average plot, further enhancing its significance. Specifically, the experimental data show that 72% of the composition constitutes trapped air, whereas the CFD results yield a value of 61%. This alignment between CFD and experimental findings is notably closer compared to the results obtained for Case 1.
- Intriguingly, as transition to the middle section, the CFD plot reveals a reduction in air presence when contrasted with the upper section. This trend persists in the bottom section as well, aligning with the conclusions drawn from the time-average plot. The middle section's time-average void fraction in the CFD stands at 52%, while the experimental counterpart registers a higher value of 73%, signifying a slight increase from the upper portion.

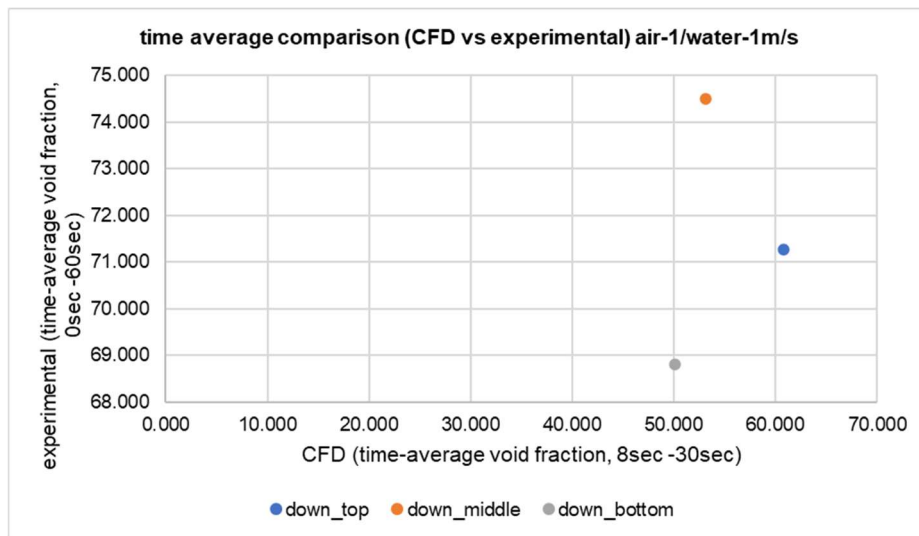


Figure 6-18: Case 2 - Time average void fraction comparison

- The bottom section's time-average plot indicates a 65% void fraction in the experimental data, while the CFD results display a lower value of 50%. Despite this slight divergence from experimental outcomes, it's worth noting that the model's results exhibit an improvement compared to Case 1. This positive trend shows the model's enhanced performance in capturing the intricate complexities of energy dissipation and fluid mixing within the flow.

This granular examination highlights the model's adaptability to air velocity 1 m/s and water velocity 1 m/s. The interplay of parameters, evident in the void fraction distribution, showcases the intricate relationship between simulation dynamics, boundary conditions, and the complex fluid behaviour that governs the flow. These findings contribute to a deeper appreciation of the model's capabilities and its potential for further refinement in future investigations.

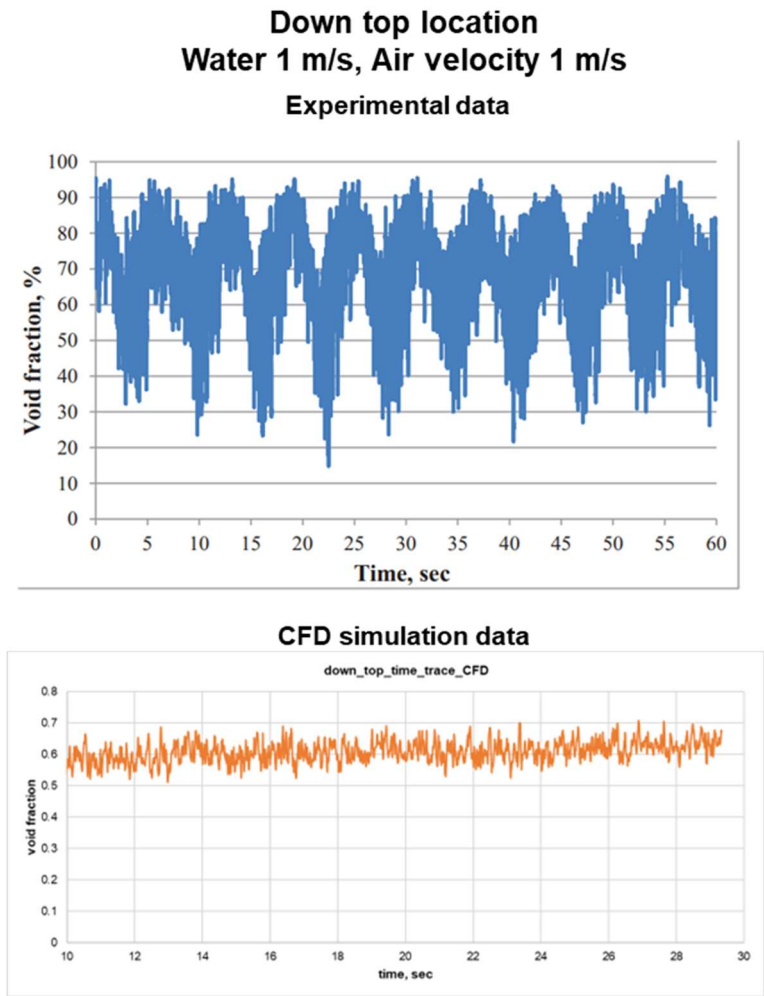
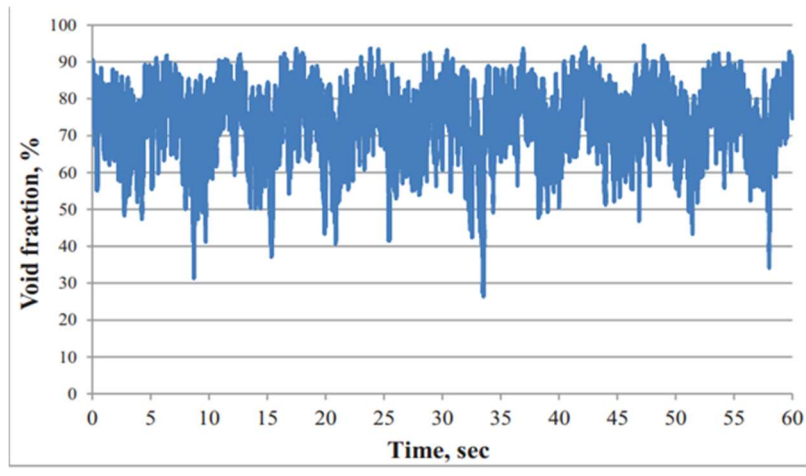


Figure 6-19: Case 2 – Down top time trace CFD vs experiment comparison

Down middle location
Water 1 m/s, Air velocity 1 m/s
Experimental data



CFD simulation data

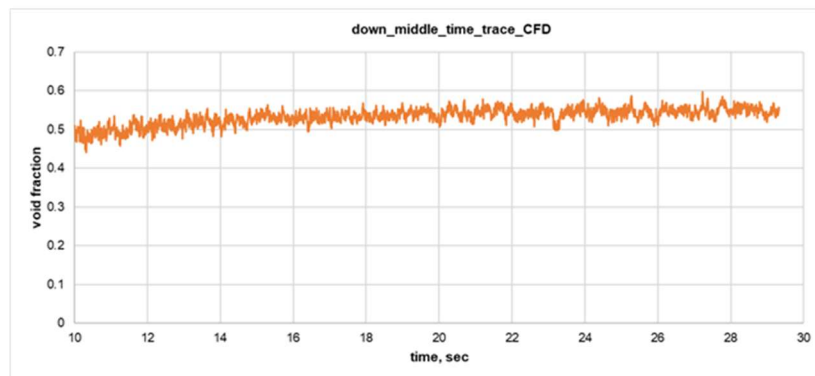


Figure 6-20: Case 2 – Down middle time trace CFD vs experiment comparison

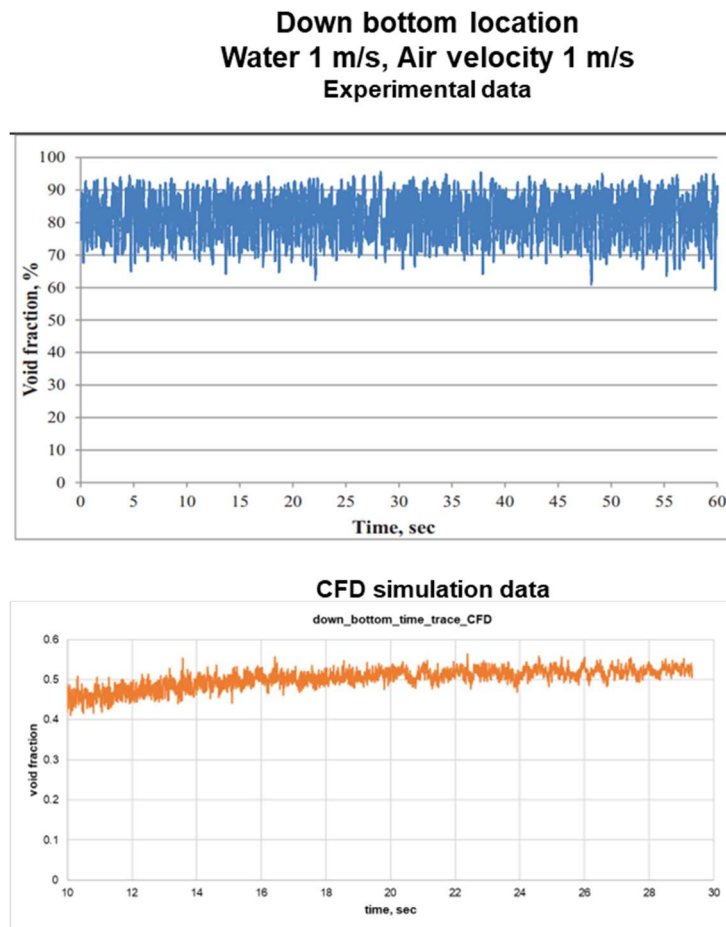


Figure 6-21: Case 2 – Down bottom time trace CFD vs experiment comparison

Case 2 - Time Trace Analysis:

Within the confines of Case 2, a study of the time trace behaviour in the top section of the downward pipe reveals some insights. Figure 6-19, Figure 6-20 and Figure 6-21 experimental observations unveil a sinusoidal variation in air distribution, oscillating between 90% and 30% at the top section. However, the CFD counterpart diverges from this pattern, exhibiting a range between 50% and 70%. This noteworthy contrast highlights the necessity for CFD model refinement to faithfully capture the intricate flow dynamics inherent in this scenario. In the middle region, the sinusoidal pattern exits, as shown in Figure 5-20, but the range is reduced. In the bottom section, the variation reduces further, as shown in Figure 6-21.

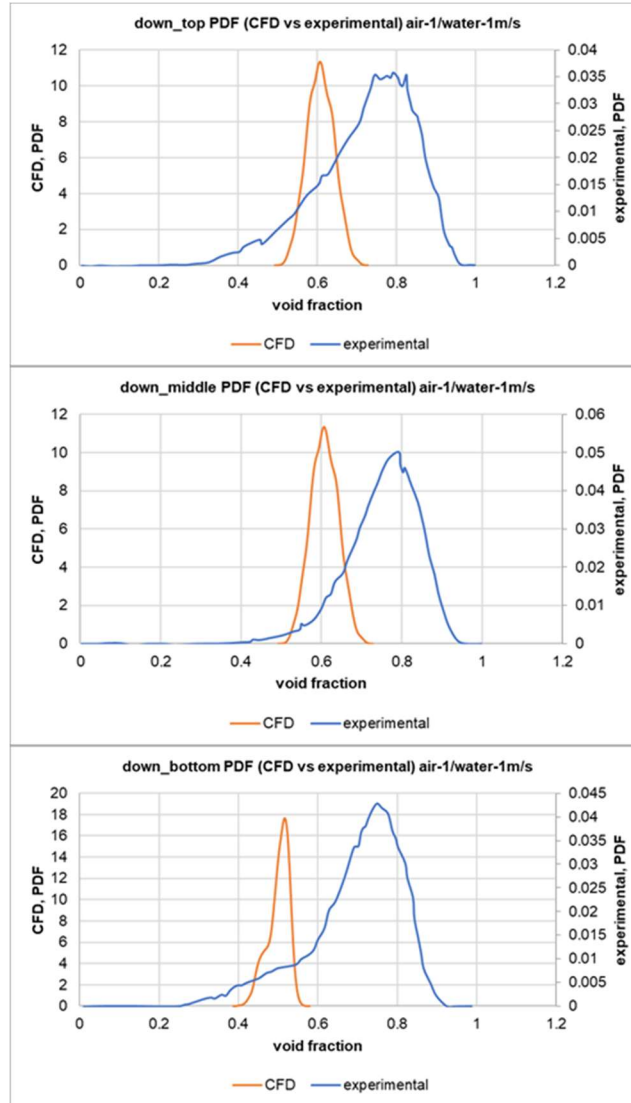


Figure 6-22: Case 2 - Probability density function comparison

This pattern extends to the probability density function (PDF) plot, as shown in Figure 6-22 exemplified in Case 2. Here, the clear variance between CFD and experimental data becomes evident, affirming the challenges in faithfully reproducing the observed dynamics.

6.7.3 Case – 03 air 9 m/s & water 1m/s:

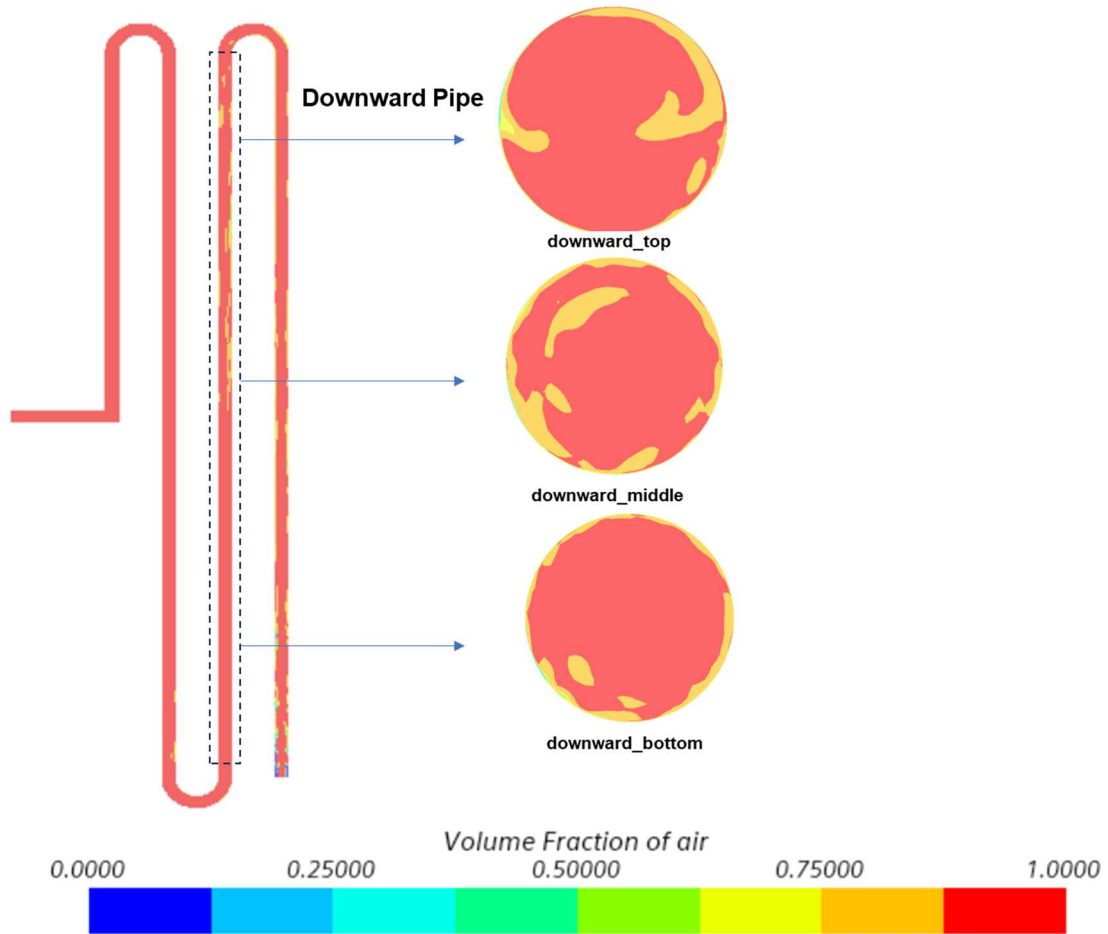


Figure 6-23: Case 3-Time average void fraction comparison

For Case 3 air velocity 9 m/s and water velocity 1 m/s, the simulation outcomes are plotted in the Figure 6-23 and Figure 6-24 .

The void fraction distribution plot showcases a pronounced trend within the CFD results:

- Across all sections—top, middle, and bottom—the void fraction reveals a dominant capture of air, exceeding 90%. This uniformity in air trapping is a noteworthy facet of the CFD simulation under the specified conditions.
- Comparing these findings with experimental data, a subtle divergence comes to light. In the experimental realm, the air trapped within the void fraction stands at 81% for the top section, 80% for the middle, and 78% for the bottom. These values, while marginally contrasting with the CFD

outcomes, offer a comprehensive perspective on the multiphase flow behaviour.

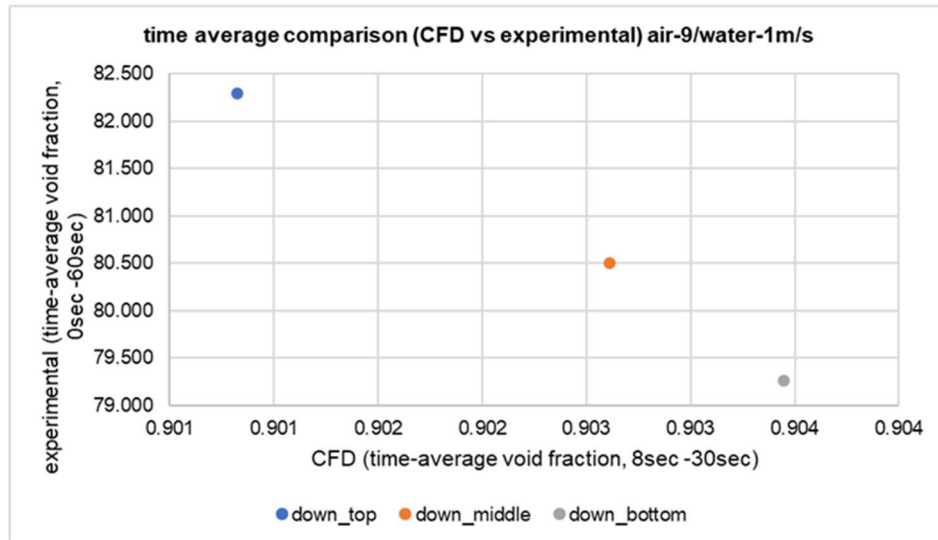
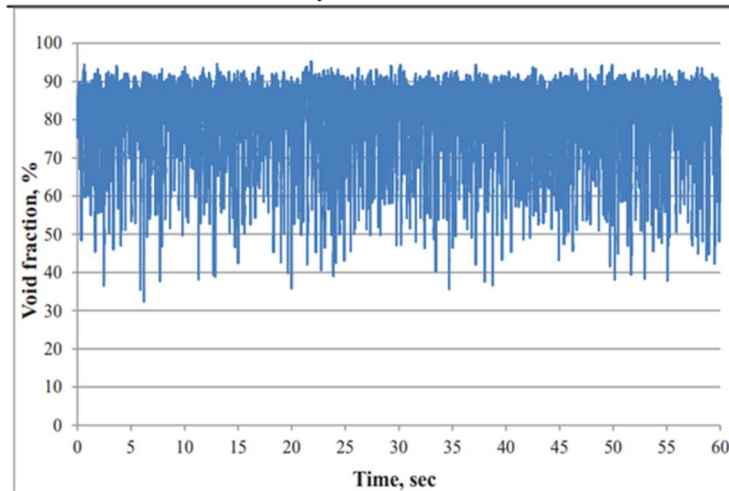


Figure 6-24: Case 3 – Down top time trace CFD vs experiment comparison

While the disparities between CFD and experimental results exist, it's pivotal to acknowledge the inherent significance within the CFD findings. The meaningful values derived from the simulation offer valuable insights into the fluid dynamics at play. This intricate interplay—characterized by air trapping and distribution—serves as a testament to the model's ability to capture complex phenomena even in conditions characterized by high velocities. The observations from Case 3 underline the potential of the CFD simulation to depict intricate flow behaviours, providing a foundation for enhanced understanding and potential refinements in future modelling endeavours.

Down top location
Water 1 m/s, Air velocity 9 m/s

Experimental data



CFD simulation data

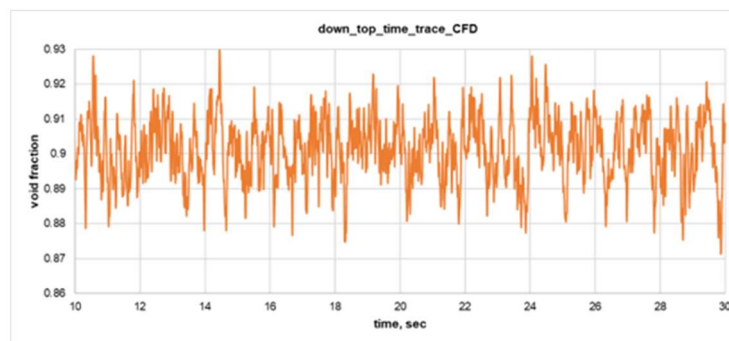
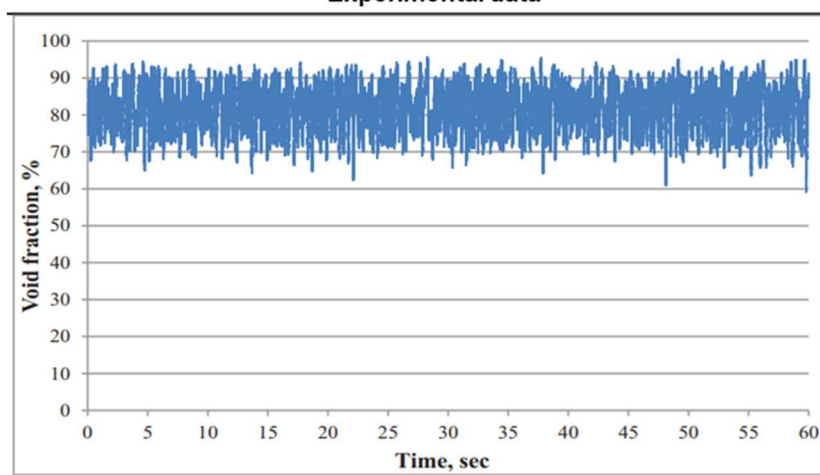


Figure 6-25: Case 3 – Down top time trace CFD vs experiment comparison

Down middle location
Water 1 m/s, Air velocity 9 m/s
Experimental data



CFD simulation data

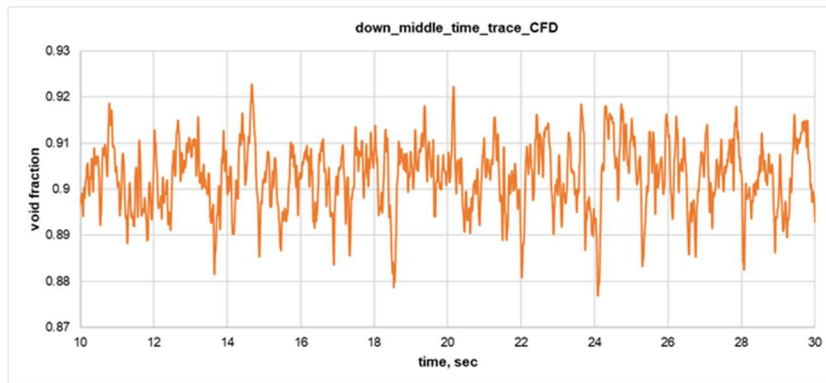


Figure 6-26: Case 3 – Down middle time trace CFD vs experiment comparison

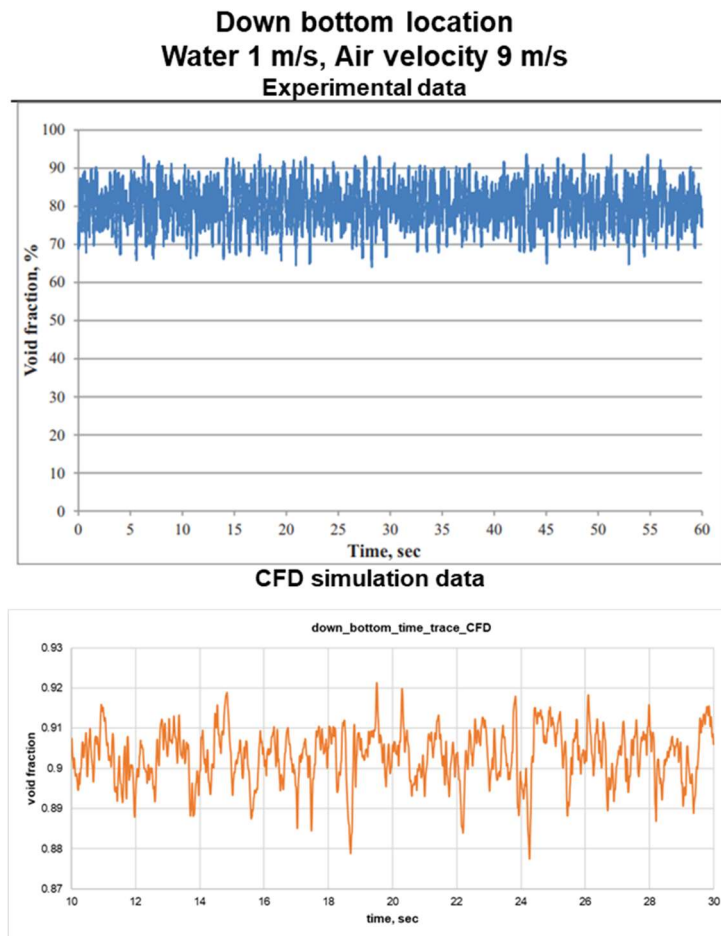


Figure 6-27: Case 3 – Down bottom time trace CFD vs experiment comparison

For Case 3, the variation in air distribution within the top section spans a range of 88% to 92% as shown in Figure 6-25. Interestingly, the sinusoidal oscillations observed in the experimental data are notably absent in the CFD simulation results. This observation points out the need for refining the CFD model to capture the dynamic behaviours that seem to elude the current framework. Due to higher velocity, the sinusoidal variation dies as shown in middle and bottom as shown in Figure 6-26 and Figure 6-27. Similarly to Case 2, CFD results for Case 3 show that the void fraction is almost constant.

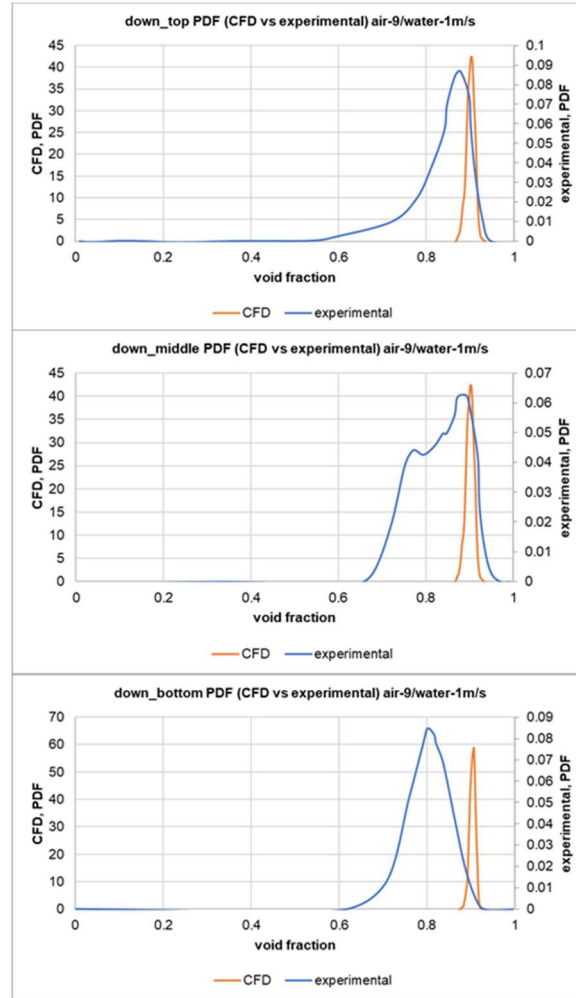


Figure 6-28: Case 3 - Probability density function comparison

Figure 6-28 shows the trend, the probability density function (PDF) plot for Case 3 exhibits a distinct difference between CFD and experimental data, underlining the complexity of the flow that the CFD model is yet to entirely capture.

6.8 Summary of Validation Findings and Future Steps:

The CFD validation across three distinct cases, namely Case 1, Case 2, and Case 3, has provided valuable insights into the performance and potential applications of the numerical model. This summary encapsulates the conclusions drawn from the validation phase, and outlines the next steps for optimizing the geometry through a fusion of CFD and machine learning methodologies.

Case 1:

The validation process for Case 1 has revealed intriguing trends in void fraction distribution and time-averaged results. While certain observations align closely with experimental data, notable disparities also emerged, particularly in air distribution. This indicates the need for a more comprehensive analysis to delve into the intricacies of the simulation approach under varying conditions.

Case 2:

The validation of Case 2 yielded encouraging outcomes. The CFD model's results exhibited fair agreement with the experimental data across the void fraction distribution, time-averaged findings, and probability density function. This consistency highlights the model's aptitude to effectively capture the multiphase flow behavior under those specified conditions.

Case 3:

For Case 3, where air velocity was 9 m/s and water velocity was 1 m/s, the CFD simulation showcased substantial air trapping across the sections. While subtle deviations from experimental data were noted, the CFD findings still hold valuable insights, shedding light on intricate fluid dynamics.

As we pivot towards the utilization of the CFD model's outcomes for machine learning, it is evident that the model's performance varies across cases. Case 2 serves as a promising foundation, while Cases 1 and 3 warrant further analyses. The culmination of CFD and machine learning techniques is poised to unlock novel avenues for geometry optimization.

The research's amalgamation of CFD and machine learning to optimize geometry is indicative of a powerful approach. As we transition to harnessing this synergy for 150 mm and 50 mm pipe diameters, the established process demonstrates its efficacy. This process, rooted in rigorous validation and exploration, equips us with the means to extract meaningful simulation data. This data, in turn, will serve as a pivotal resource in refining geometries and ultimately enhancing the overall understanding of fluid dynamics within complex systems.

7 Results and Discussion

7.1 Flow characterisation using ML models.

In the previous chapters, the intricate dynamics of void fraction, pressure, and liquid film thickness within the serpentine pipes were explored. This chapter deals with predictions and validation. The focus is on comprehensively understanding the flow characteristics by leveraging the trained data and employing regression models. The goal is to validate the trained data and subsequently discuss the predictions obtained through these models. Each station within both the upward and downward pipes will be meticulously explored, unravelling the nuances of flow behaviour, and shedding light on the predictive power of the regression models. Through this analysis, this section aims to offer a robust understanding of how these key variables interact within the serpentine pipes, thereby contributing to a more profound comprehension of the fluid dynamics at play. For validation purpose Data -1 Table 7-1 is used the trained models tested at three conditions at their results compared for different regression models developed and the predictions are performed for intermediate slug flow conditions.

Table 7-1: Inputs used for validation.

U_{sl} , m/s	U_{sg} , m/s	Q_g , Sm ³ /h	Q_g , m ³ /h	Ptest, barg	Ptest+1, bar	T1, °C	T2, °C
0.10	1.44	50	41.87	0.19	1.19	12.96	12.23
0.48	1.36	50	39.67	0.26	1.26	15.52	0.48
1.00	1.32	50	38.46	0.30	1.30	17.84	17.32

Table 7-2: Prediction input values

U_{sl} , m/s	U_{sg} , m/s	Q_g , Sm ³ /h	Q_g , m ³ /h	Ptest, barg	Ptest+1, bar	T1, °C	T2, °C
1.00	1.995	75	58.245	0.29	1.29	18.13	17.585
1.00	6.455	250	188.4	0.325	1.325	18.82	18.215
1.00	8.635	350	251.96	0.385	1.385	18.735	18.09
1.00	14.565	700	424.97	0.640	1.640	17.93	17.17

For predictions the slug will be studied in detail, with water superficial velocity 1 m/s and air velocity between 0.1 m/s – 20 m/s.

U_{sl},	U_{sg},	Q_g,	Q_g,	P_{test}, barg	P_{test+1}, bar	T1,	T2,
m/s	m/s	Sm³/h	m³/h			°C	°C
0.10	1.44	50	41.87	0.19	1.19	12.96	12.23
0.48	1.36	50	39.67	0.26	1.26	15.52	0.48
1.00	1.32	50	38.46	0.30	1.30	17.84	17.32

Table 7-2 undergo a meticulous selection process. The range of chosen input values encompasses the slug flow zone for gas void fraction and annular flow for liquid film thickness. To validate the models, a handful of cases were specifically selected within this range, primarily due to the absence of available empirical data for those scenarios.

7.2 Gas void data and Liquid Film Thickness validation and prediction.

As outlined in Chapter 5, the data that has undergone preprocessing is utilized for this phase of analysis. Three crucial parameters are considered: variation in pressure, void fraction, and thickness of the liquid film. Within this section, the model's accuracy is assessed by applying it to the trained data, validating it against experimental data, and subsequently making predictions and engaging in discussions concerning new values.

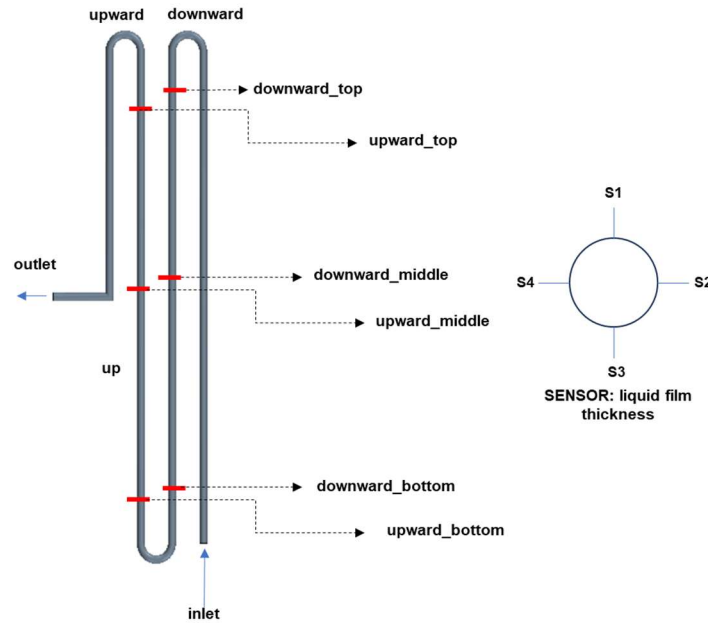


Figure 7-1: Sensor and reading location

The behaviour of both the upward and downward pipes at each station is comprehensively examined. The liquid film thickness values are derived from employed sensors, with their mean value being a focal point of this study. The precise locations of the sensors and the highlighted areas depicting void fraction are indicated in the Figure 7-1.

7.2.1 Pressure, void fraction and liquid film thickness

Pressure: The pressure behaviour within a serpentine configuration of upward and downward pipes, characterized by two-phase flow involving air velocities between 1 and 20 m/s and water velocities between 0.1 and 1.5 m/s, is intricate and dynamic. The interplay between air and water velocities across this spectrum significantly impacts the pressure distribution along the serpentine path.

- In regions where air velocity is low and water velocity is high, pressure decreases due to resistance from rapidly moving water. Conversely, higher air velocity and lower water velocity lead to pressure fluctuations from phase interactions affecting pipe walls.
- Factors like shifts in flow regime (e.g., bubbly to slug flow) and varying void fraction, representing mixture volume proportions, contribute to pressure

changes. Elevated void fractions cause pressure reduction due to reduced mixture density and increased compressibility.

The serpentine geometry introduces complexities; flow direction alterations between upward and downward pipes impact pressure distribution due to centrifugal forces and gravity. Pressure fluctuations within the described serpentine configuration, involving diverse air and water velocities, result from phase interactions, flow regime changes, void fraction variations, and serpentine geometry.

Gravity is pivotal in this context. With air velocities from 1 to 20 m/s and water velocities from 0.1 to 1.5 m/s, gravity's influence on the two-phase mixture becomes evident. In upward pipes with lower water and higher air velocities, gravity separates phases, leading to water settling and increased pressure due to hydrostatic pressure. Conversely, in downward pipes with higher water and lower air velocities, gravity sustains phase mixing, influencing pressure dynamics. Data points illustrate this:

Air Velocity: 1 m/s, Water Velocity: 0.1 m/s

- Upward Pipe: Gravity-induced settling increases pressure
- Downward Pipe: Gravity maintains mixing, impacting pressure
- Air Velocity: 10 m/s, Water Velocity: 0.5 m/s
- Upward Pipe: Gravity-induced phase separation affects pressure
- Downward Pipe: Gravitational effects on phase distribution influence pressure

Air Velocity: 20 m/s, Water Velocity: 1.5 m/s

- Upward Pipe: Hydrostatic pressure and phase separation impact pressure
- Downward Pipe: Momentum exchange and gravity-driven phase behavior influence pressure

These data points underscore the intricate connection between velocities, gravity, and phase interactions, emphasizing the complex nature of pressure variations in the serpentine configuration. Furthermore, the pressure in such a system is

influenced by factors such as velocities, densities, pipe dimensions, bends, and flow patterns. The pressure decreases as velocities and pipe diameter increase, while bends and flow patterns also contribute to pressure variations.

For a specific scenario with upward-downward serpentine pipes and varying air and water velocities, the pressure drop will be notable, dependent on these factors. Well-established correlations, like Lockhart-Martinelli (Lun, Calay and Holdo, 1996), Chisholm (Cheah, Kevrekidis and Benziger, 2013), and Chen (Jiang et al., 2020), can estimate pressure drop with reasonable accuracy based on experimental data.

Void Fraction: The void fraction in the serpentine configuration, comprising upward and downward pipes with two-phase flow involving air velocities ranging from 1 to 20 m/s and water velocities from 0.1 to 1.5 m/s, introduces dynamic complexity. The interplay between these varied velocities significantly influences the distribution of void fraction along the serpentine path.

- In areas where air velocity surpasses water velocity, void fraction could increase due to air displacing water. Conversely, with lower air velocity and higher water velocity, a decrease in void fraction might occur due to the prominence of the denser water phase.
- Flow regime shifts (e.g., bubbly to slug flow) and alterations in pressure and temperature contribute to void fraction fluctuations. In regions with elevated void fraction, the mixture's reduced density affects compressibility, leading to variations.
- The serpentine geometry adds complexity; its bends and direction changes impact void fraction distribution.

Liquid Film Thickness:

The liquid film thickness within this serpentine arrangement, encompassing upward and downward pipes with two-phase flow and diverse air velocities from 1 to 20 m/s and water velocities from 0.1 to 1.5 m/s, represents a crucial factor in the system's dynamics. The interplay between these velocities significantly affects liquid film thickness distribution along the serpentine path.

- In areas where air velocity is high and water velocity is low, the liquid film thickness might decrease due to air displacing water. Conversely, lower air velocity and higher water velocity may lead to an increase in liquid film thickness due to water's prominence.
- Flow pattern changes, surface characteristics, and fluid properties contribute to liquid film thickness variations. Rough surfaces or changes in viscosity can influence the film's thickness.

The serpentine geometry introduces complexity; bends and direction shifts can result in liquid film thickness changes due to centrifugal forces and gravity.

To anticipate these complexities, various regression models are employed and extensively examined for each station. The objective is to identify the most suitable model that can accurately predict these intricacies while minimizing the time required for analysis.

7.2.1.1 Downward top section

The flow is first disturbed at the top section of the downward pipe, immediately after bend 1. Figure 7-2 and Figure 7-3 shows the pressure data points validation and predictions respectively. The validation shows that the support vector model holds good for higher gas velocities and at low velocities it is over predicting. Similarly, the prediction values shows the same trend.

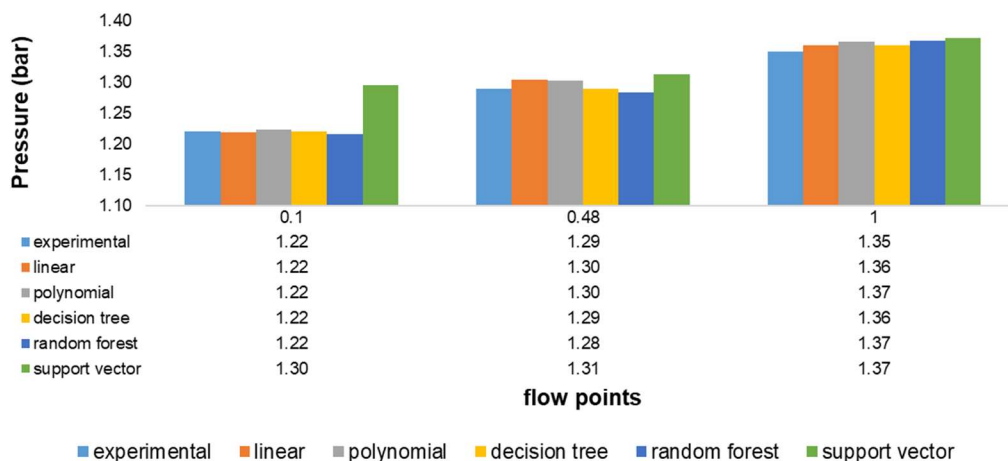


Figure 7-2: Validation – pressure at downward pipe -top

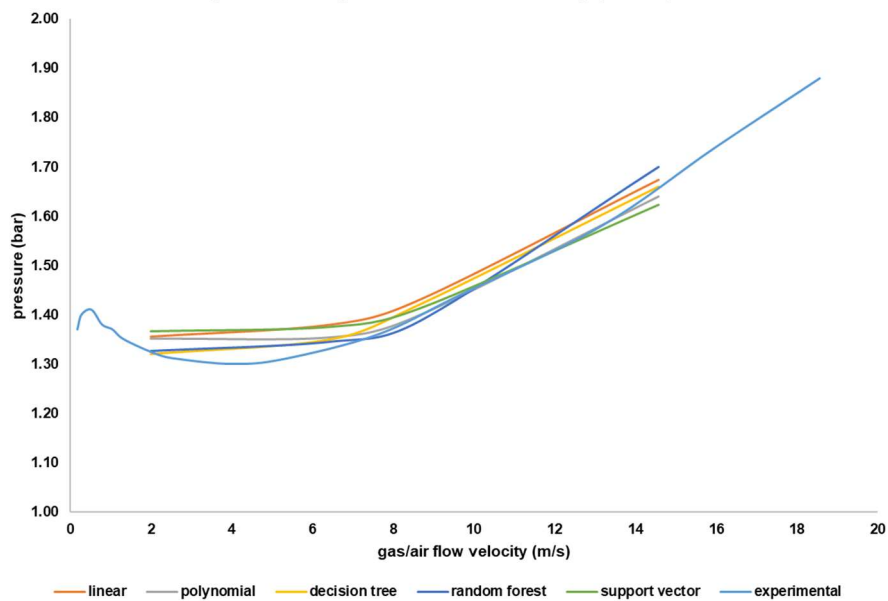


Figure 7-3: Prediction – pressure at downward pipe -top

Figure 7-4 and Figure 7-5 plots show the validation and prediction plot for the void fraction at the downward top data point respectively. Support vector model under prediction marginally with other models, in the prediction point of view polynomial and random forest hold good for this study.

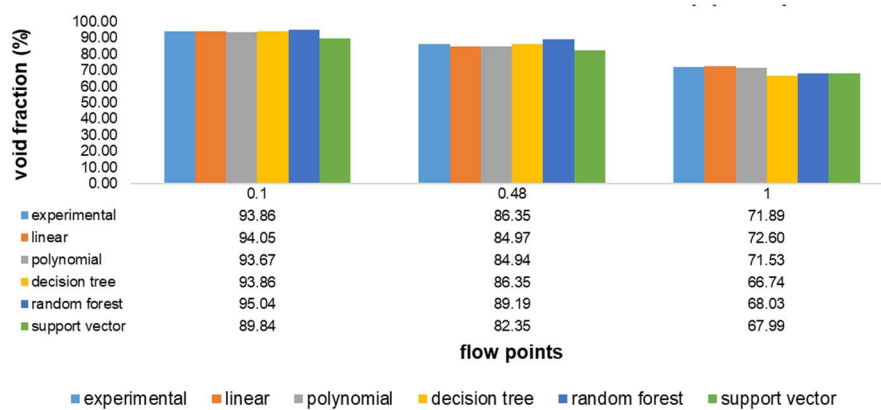


Figure 7-4: Validation – void fraction at downward pipe - top

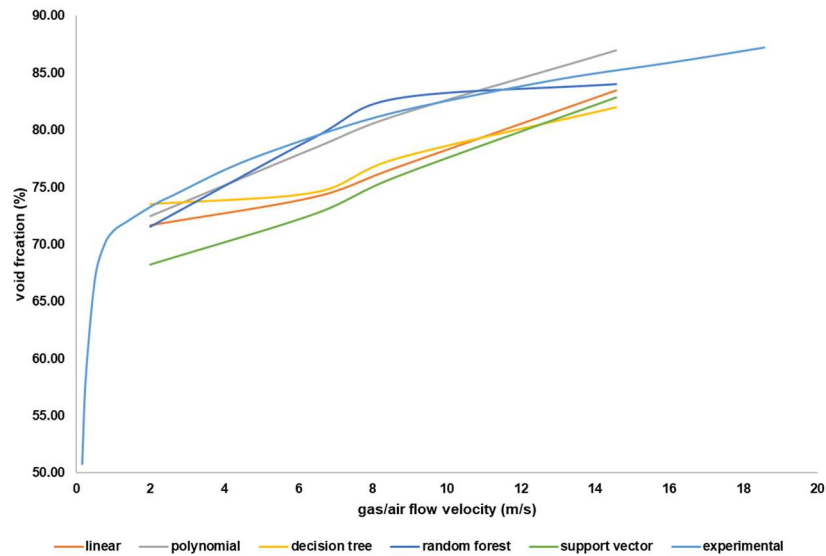


Figure 7-5: Prediction – void fraction at downward pipe – top

Figure 7-6 and Figure 7-7 plot shows the validation and prediction plot for the liquid film thickness at downward top data point respectively. The film thickness prediction is very challenging because of the gas and water velocity changes. The decision tree predicts very well the thickness behaviour for different velocity changes.

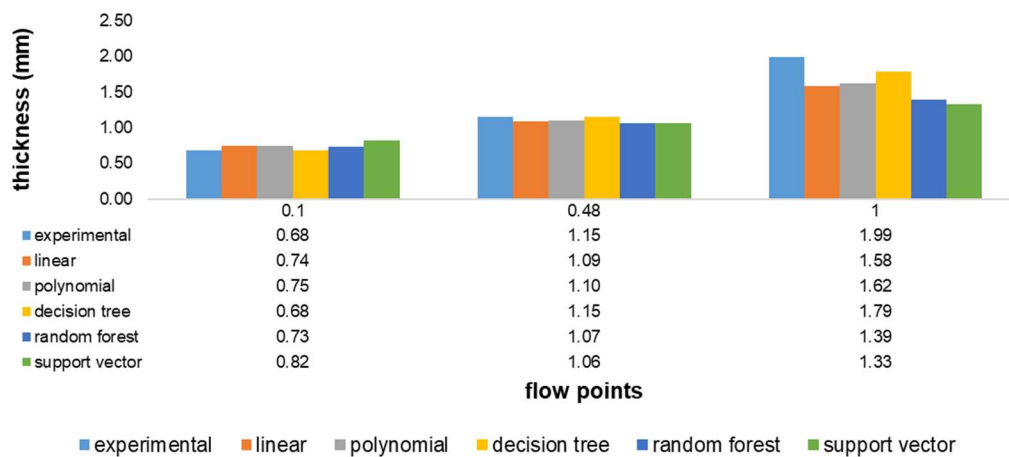


Figure 7-6: Validation – liquid film thickness at downward pipe - top

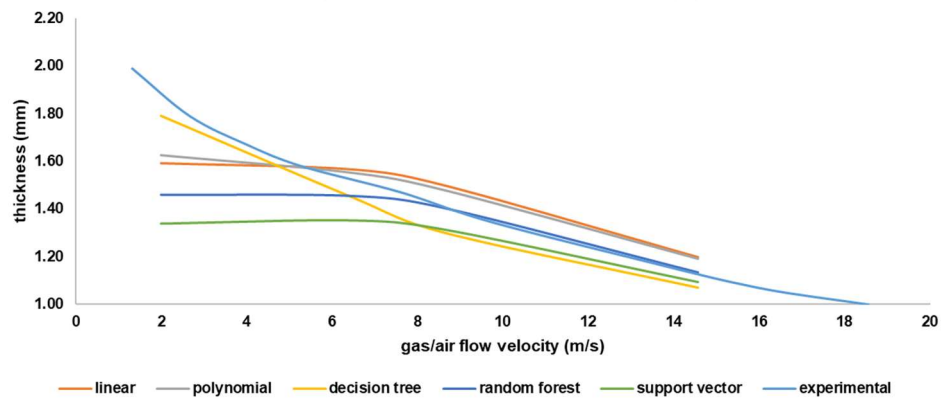


Figure 7-7: Prediction – liquid film thickness at downward pipe – top

7.2.1.2 Downward middle section

In the downward middle section, pressure data points are also validated and predicted. The fluctuations are illustrated for distinct regression models in Figure 7-8 and Figure 7-9. As seen in the top section, the support and rest other models predicting the pressure variation effectively.

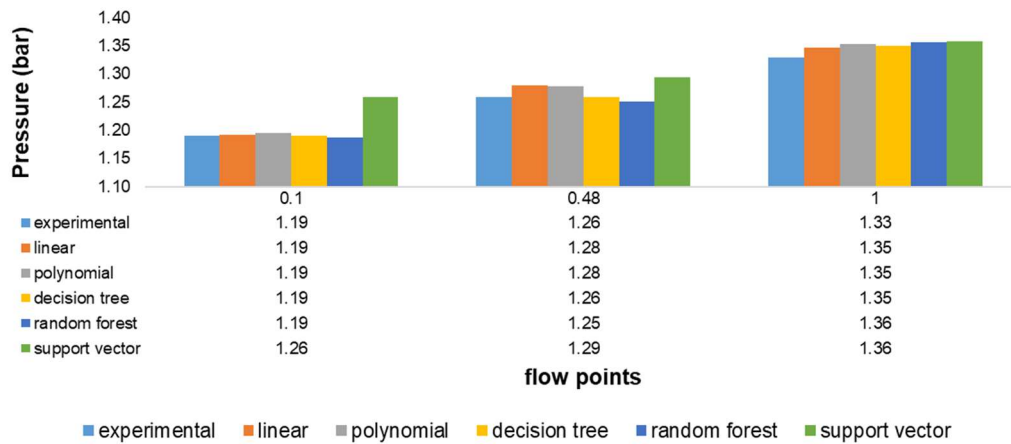


Figure 7-8: Validation – pressure at downward pipe - middle

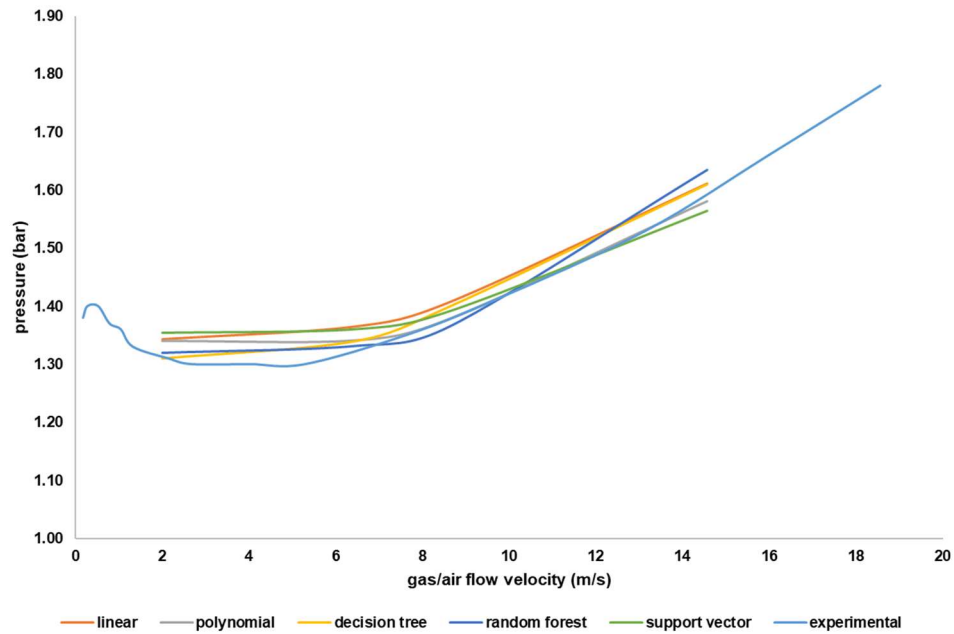


Figure 7-9: Prediction – pressure at downward pipe – middle

Figure 7-10 and Figure 7-11 show the validations and predictions for the void fraction at downward middle data point respectively. In the prediction point of view random forest hold good.

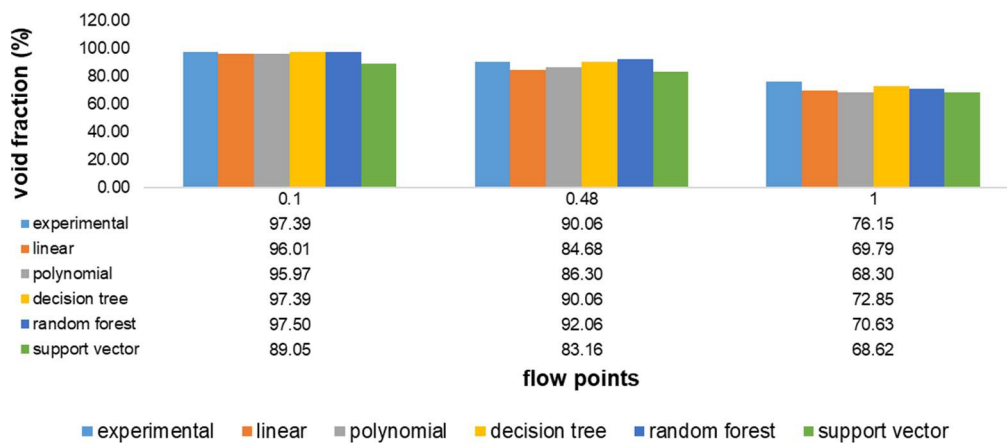


Figure 7-10: Validation – void fraction at downward pipe - middle

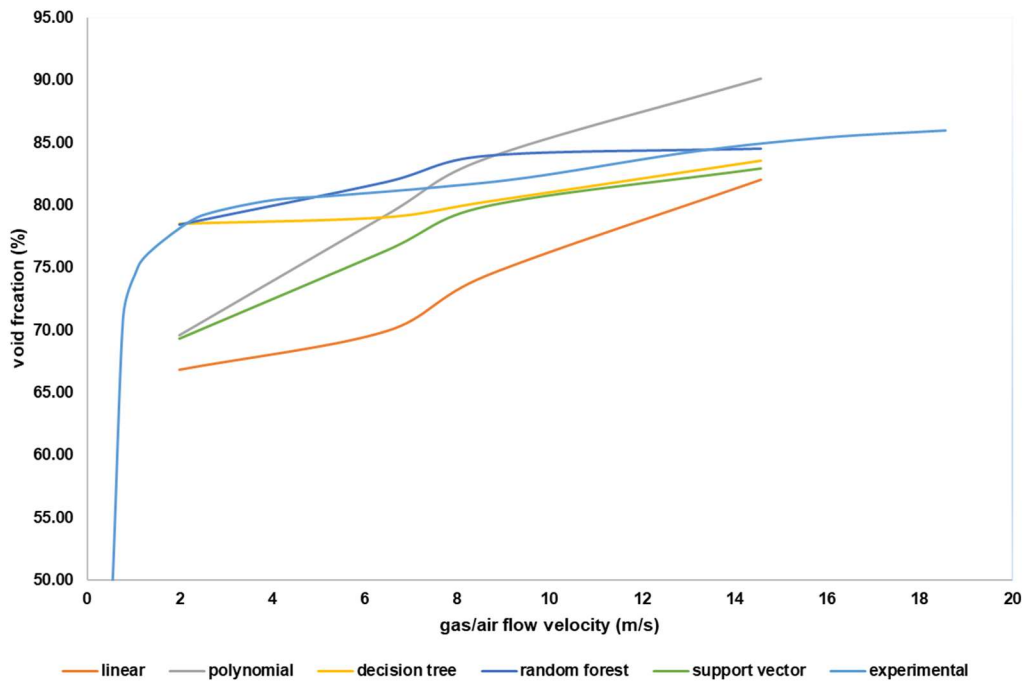


Figure 7-11: Prediction – void fraction at downward pipe – middle

Figure 7-12 and Figure 7-13 show the validations and prediction plot for the liquid film thickness at the downward middle data point respectively. The decision tree predicts the flow thickness effectively, the other models predict the trend better, but the decision tree holds at this location as well.

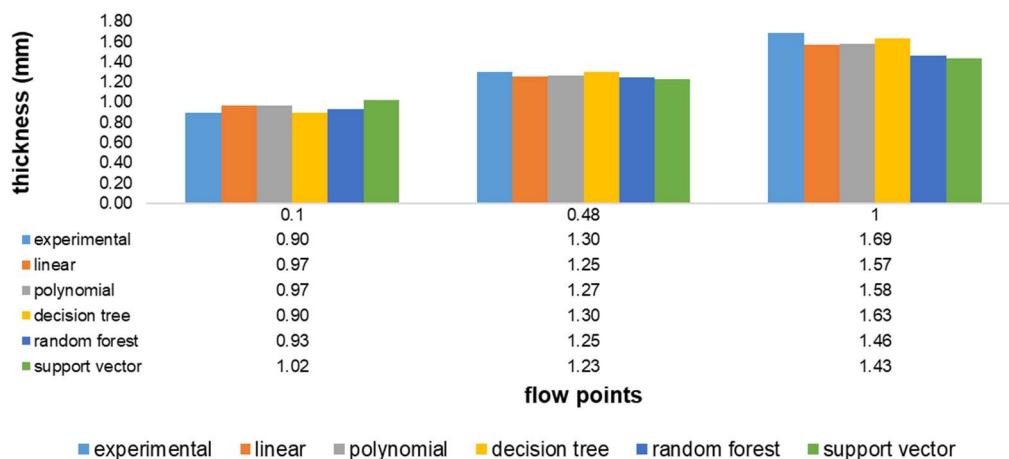


Figure 7-12: Validation – liquid film thickness at downward pipe - middle

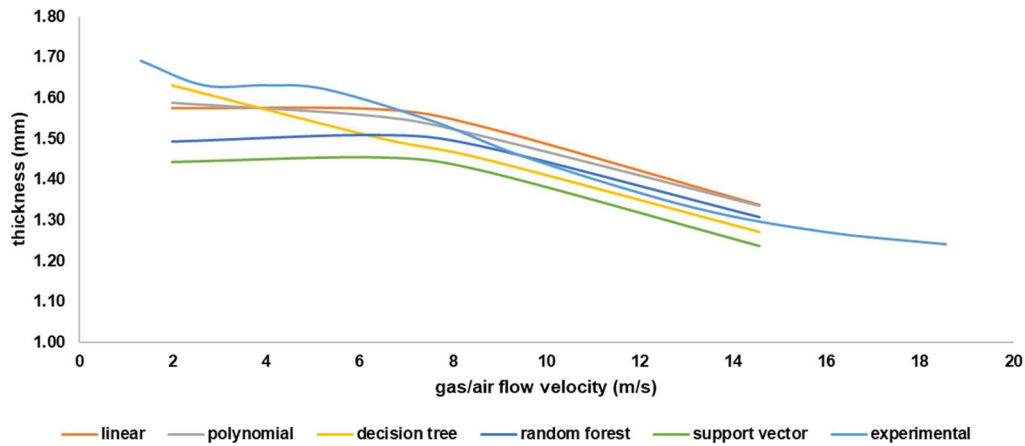


Figure 7-13: Prediction – liquid film thickness at downward pipe – middle

7.2.1.3 Downward bottom section

In the downward bottom section, pressure data points are validated and predicted. The fluctuations are illustrated for distinct regression models in Figure 7-14 and Figure 7-15. As seen in the top and middle section, the support vector model is not suitable at the location.

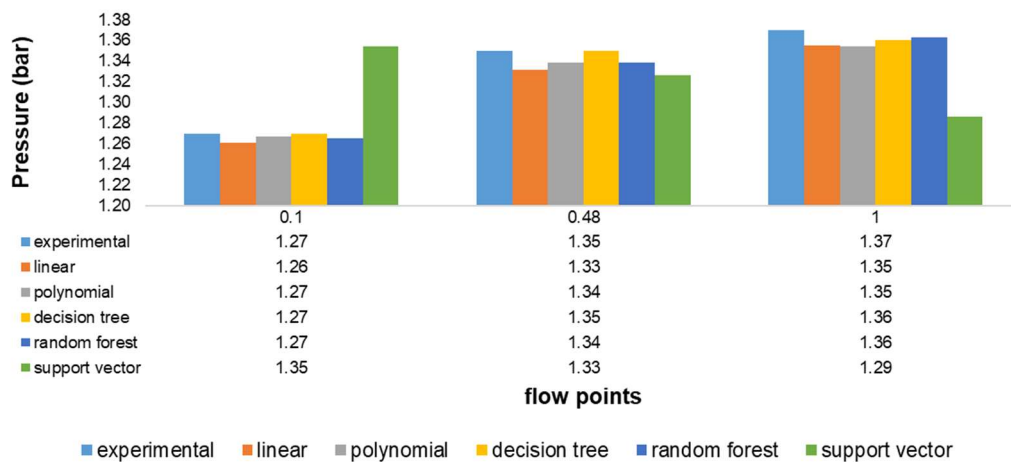


Figure 7-14: Validation – pressure at downward pipe - bottom

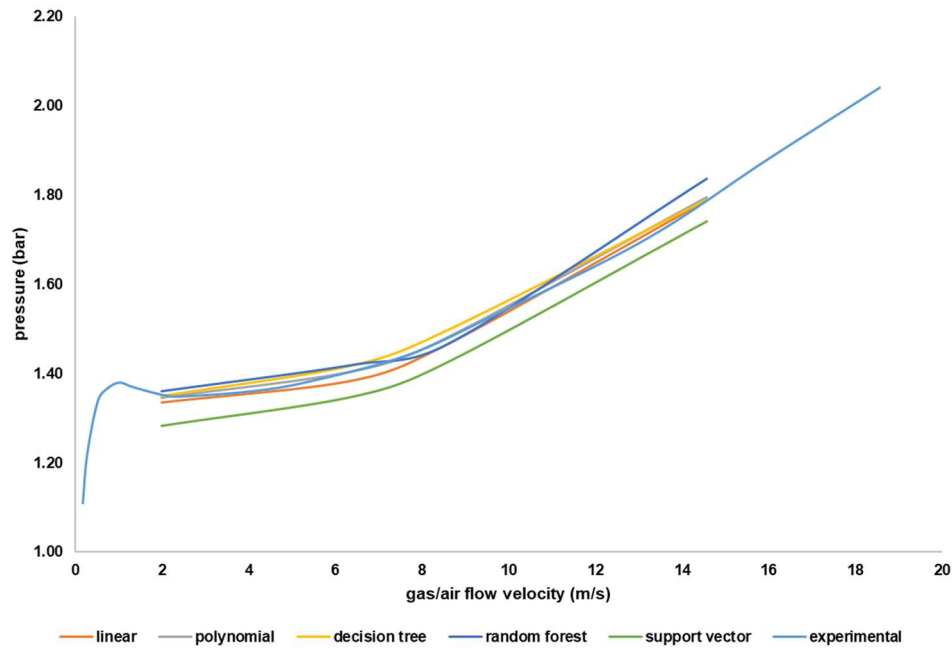


Figure 7-15: Prediction – pressure at downward pipe – bottom

Figure 7-16 and Figure 7-17 show the validation and prediction plot for the void fraction at the downward middle data point, respectively. The validation shows that all models are suitable for this study. However, the random forest is the best model for the void fraction prediction.

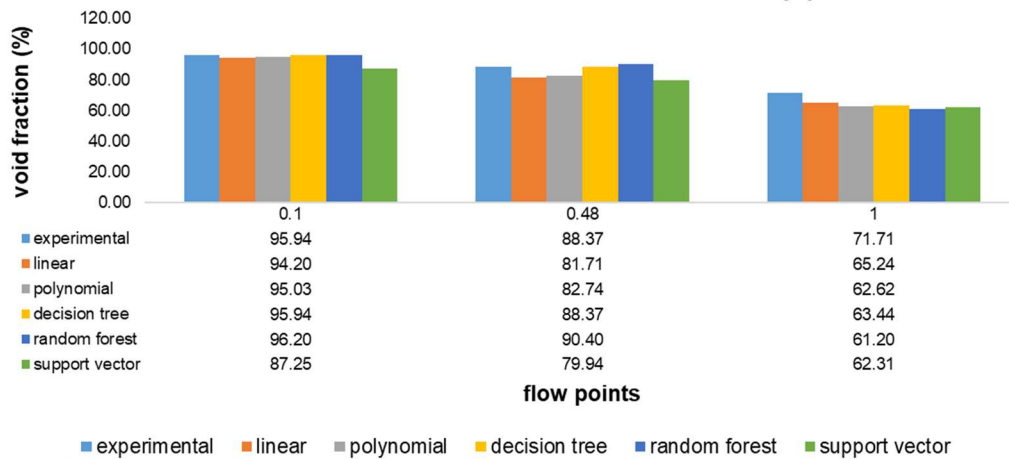


Figure 7-16: Validation – void fraction at downward pipe - bottom

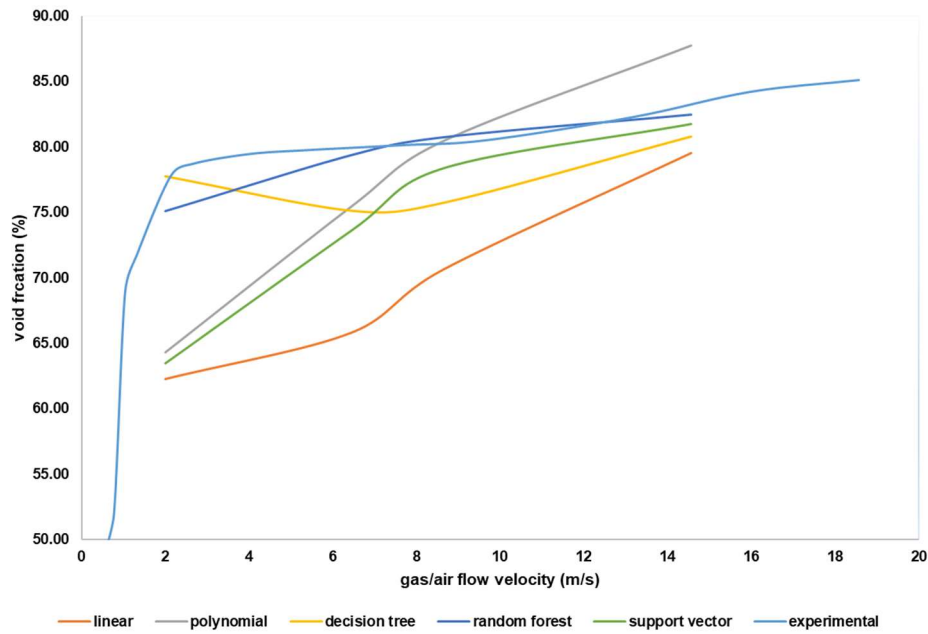


Figure 7-17: Prediction – void fraction at downward pipe – bottom

Figure 7-18 and Figure 7-19 show the validation and prediction plot for the liquid film thickness at the downward middle data point, respectively. All models show poor prediction at low gas velocity. At higher velocity, all models show the same trend, but the random forest appears closer to the experimental value.

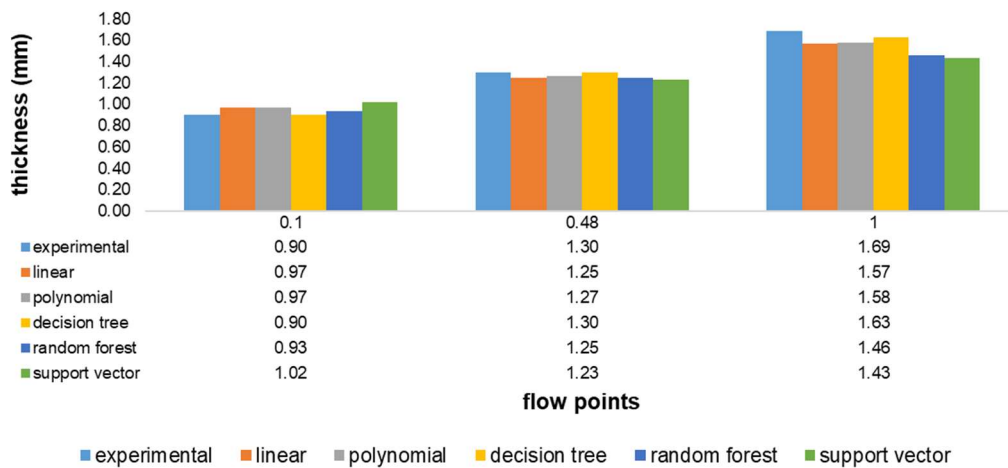


Figure 7-18: Validation – liquid film thickness at downward pipe - bottom

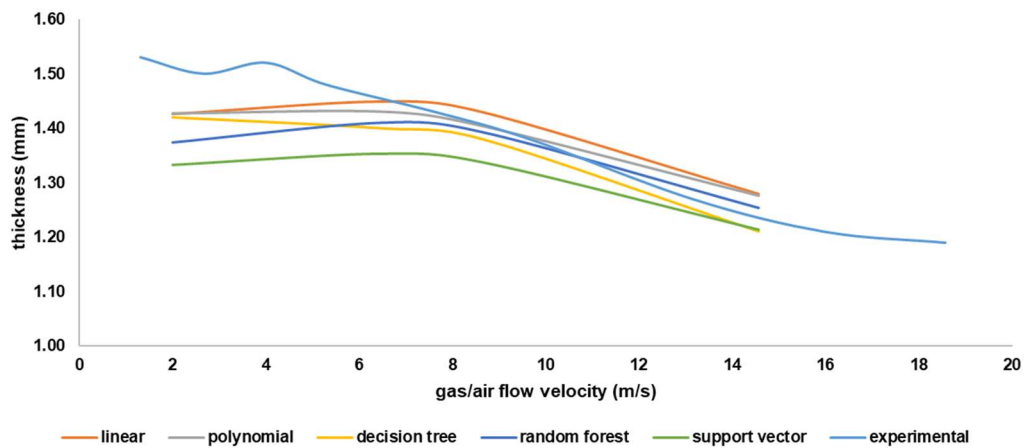


Figure 7-19: Prediction – liquid film thickness at downward pipe – bottom

7.2.1.4 Upward bottom section

In the upward bottom section, pressure data points are validated and predicted. The fluctuations are illustrated for distinct regression models in Figure 7-20 and Figure 7-21. Similarly to the downward section location, the support vector model does not appear suitable at the upward bottom section either.

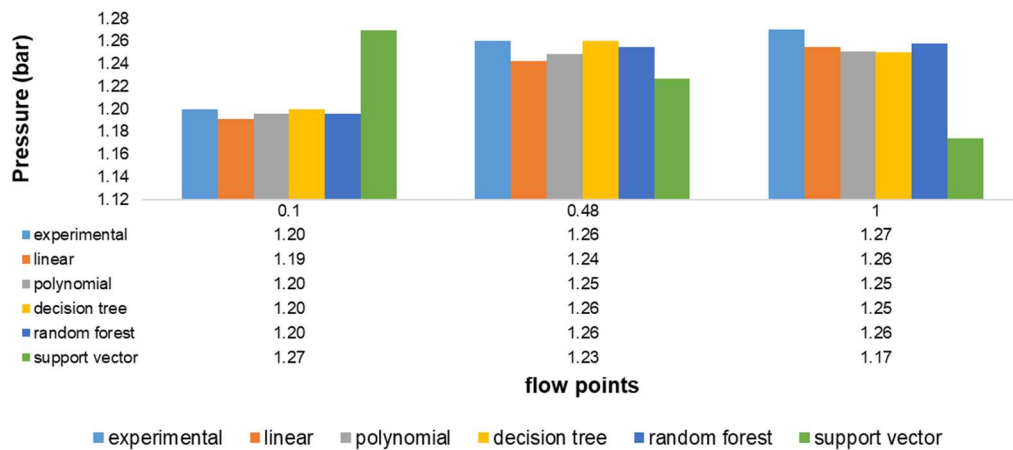


Figure 7-20: Validation – pressure at upward pipe – bottom

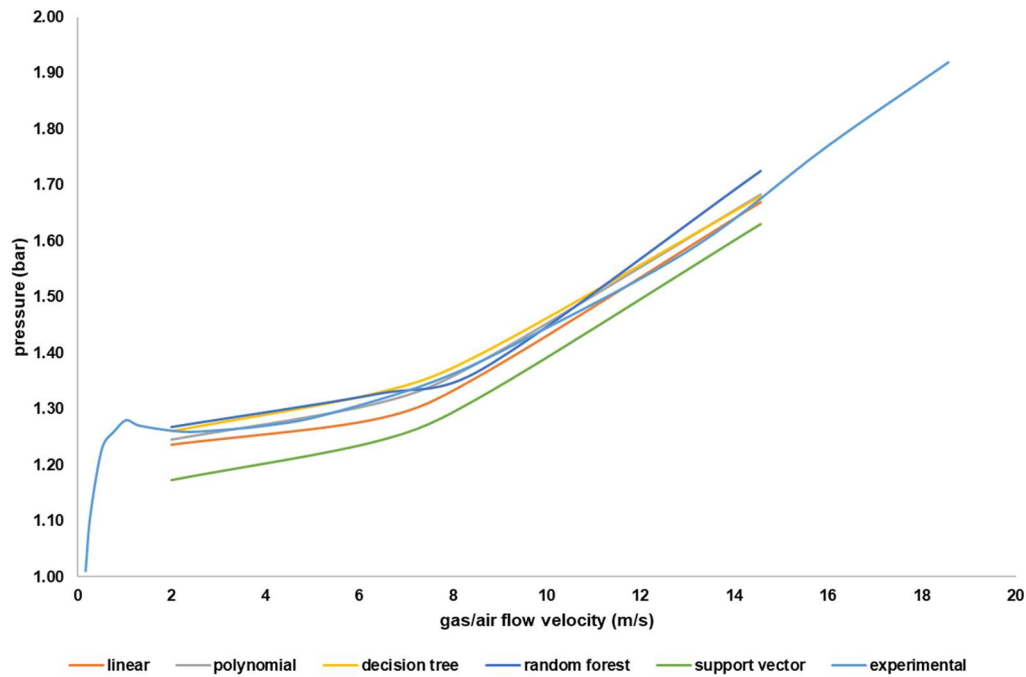


Figure 7-21: Prediction – pressure at upward pipe – bottom

Figure 7-22 and Figure 7-23 show the validation and prediction plot for the void fraction at the upward bottom data point, respectively. Apart from the linear model, all other models predict values close to the experimental results.

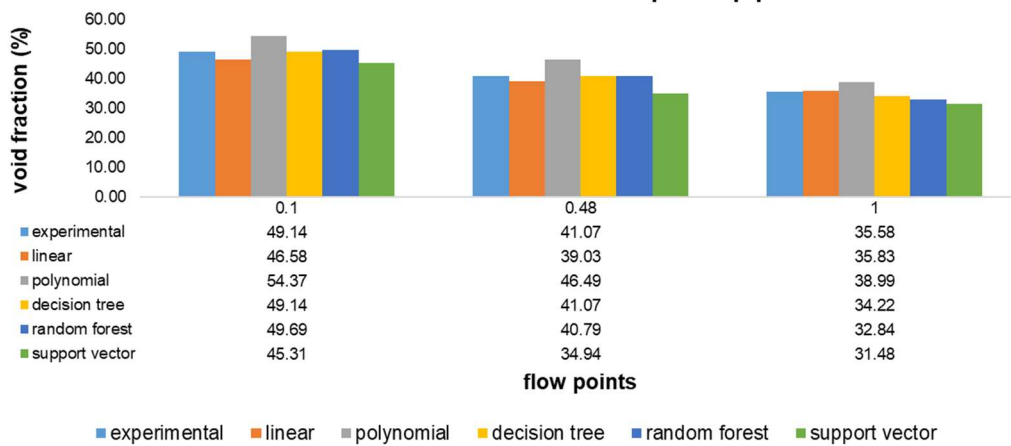


Figure 7-22: Validation – void fraction at upward pipe - bottom

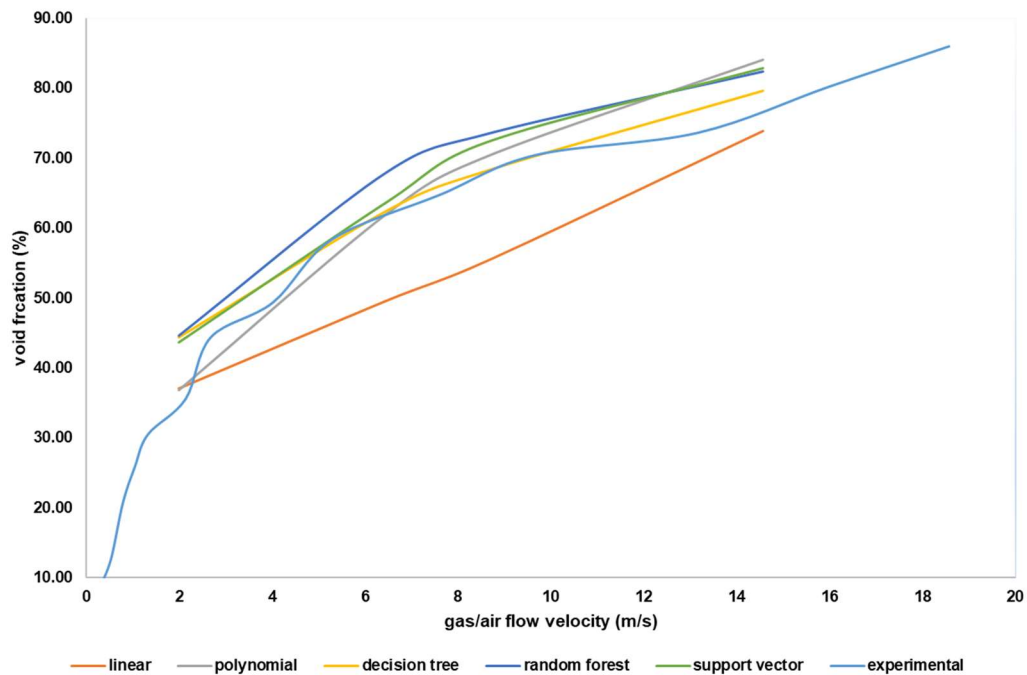


Figure 7-23: Prediction – void fraction at upward pipe – bottom

Figure 7-24 and Figure 7-25 show the validation and prediction plot for the liquid film thickness at the downward bottom data point, respectively. As can be seen, all models are very effective to predict the liquid film thickness at gas velocities higher than 5 m/s.

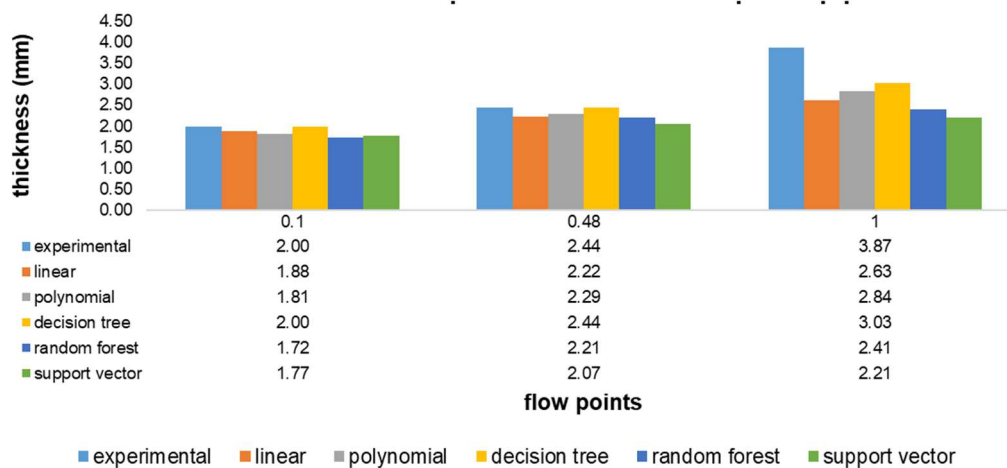


Figure 7-24: Validation – liquid film thickness at upward pipe - bottom

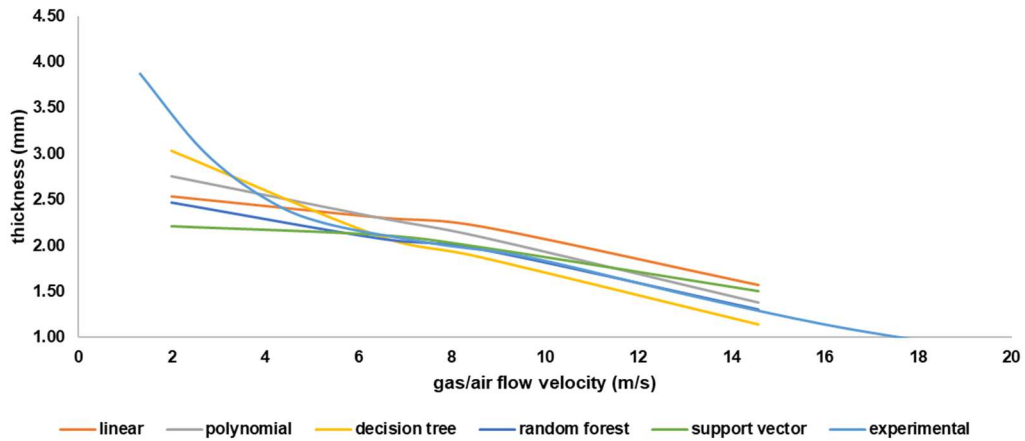


Figure 7-25: Prediction – liquid film thickness at upward pipe – bottom

7.2.1.5 Upward middle section

In the upward middle section, the pressure data points are also validated and predicted. The fluctuations are illustrated for the regression models in Figure 7-26 and Figure 7-27. Similarly to the downward bottom section, the support vector model is not suitable either at this location.

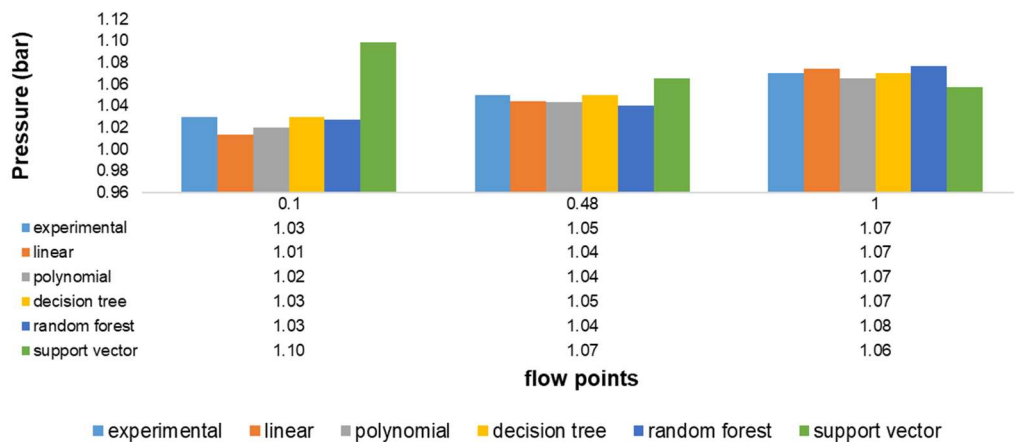


Figure 7-26: Validation – pressure at upward pipe - middle

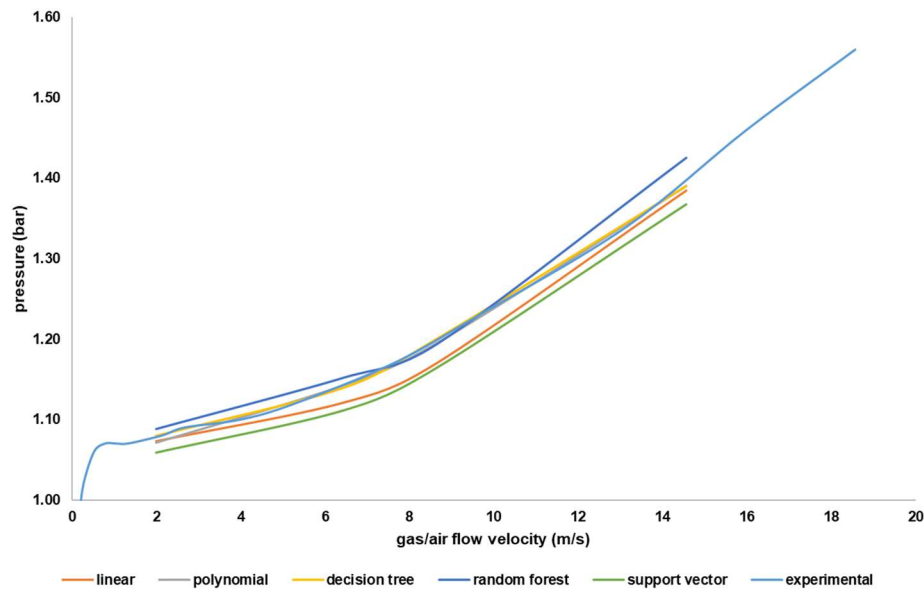


Figure 7-27: Prediction – pressure at upward pipe – middle

Figure 7-28 and Figure 7-29 show the validation and prediction for the void fraction at the upward middle data point, respectively. The random forest matches closely the experimental results. The other models however show poor performance.

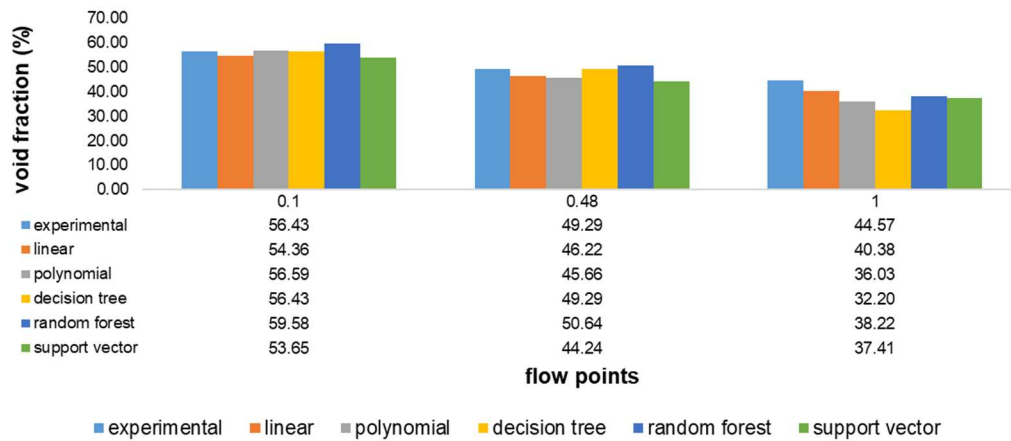


Figure 7-28: Validation – void fraction at upward pipe - middle

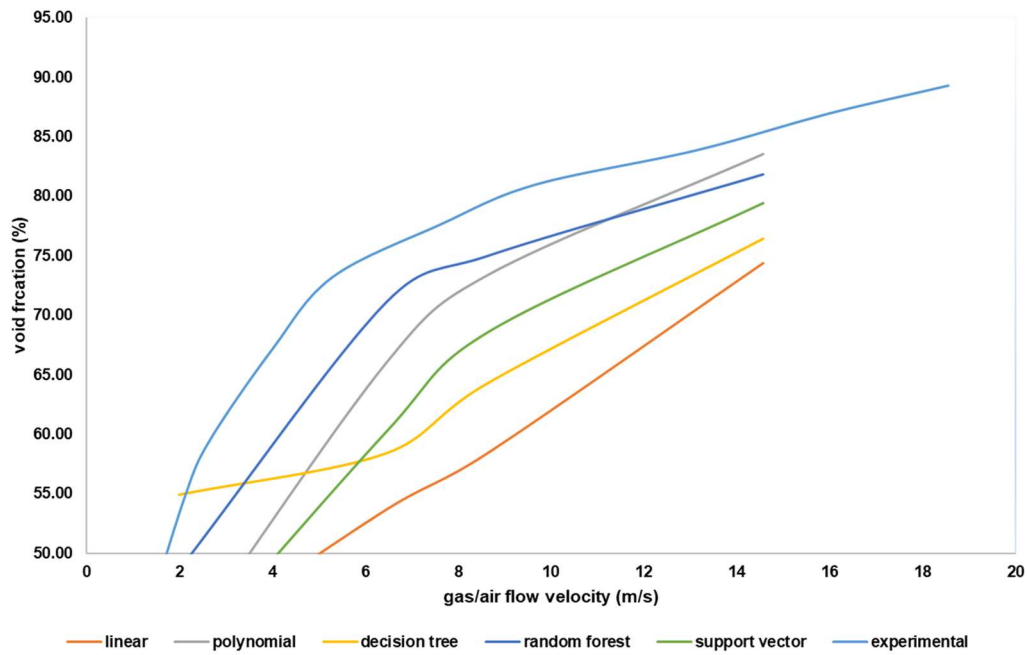


Figure 7-29: Prediction – void fraction at upward pipe – middle

Figure 7-30 and Figure 7-31 plot shows the validation and prediction plot for the liquid film thickness at downward middle data's point respectively. All models, predicting very effectively the liquid film thickness.

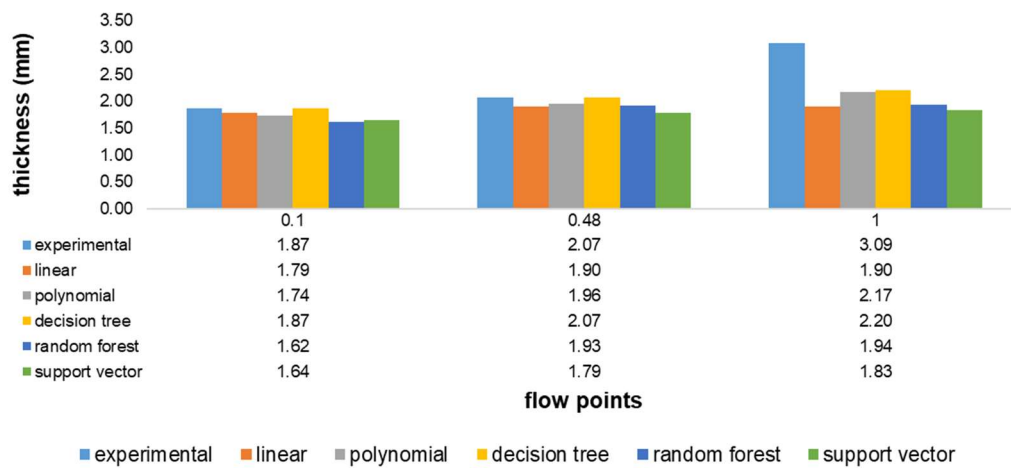


Figure 7-30: Validation – liquid film thickness at upward pipe - middle

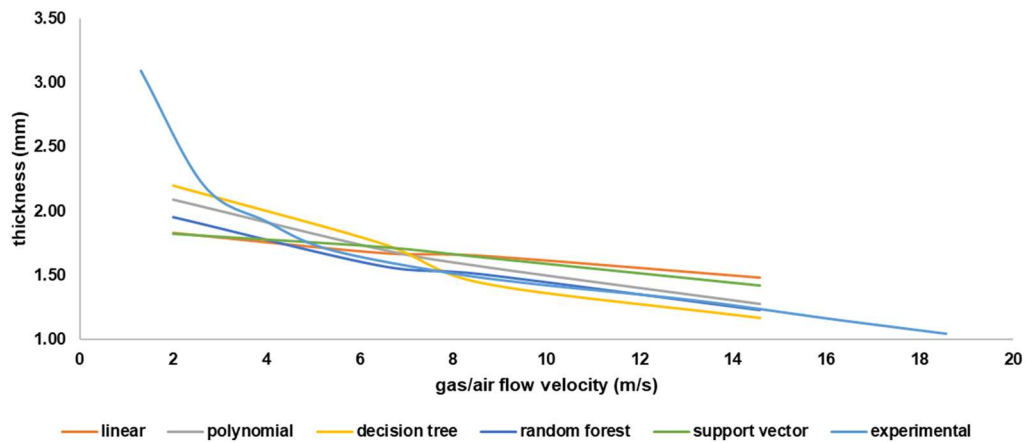


Figure 7-31: Prediction – liquid film thickness at upward pipe – middle

7.2.1.6 Upward top section

In the upward middle section, pressure data points are validated and predicted. The fluctuations are illustrated for distinct regression models in Figure 7-32 and Figure 7-33. As seen, the support vector model will not suitable for pressure predictions.

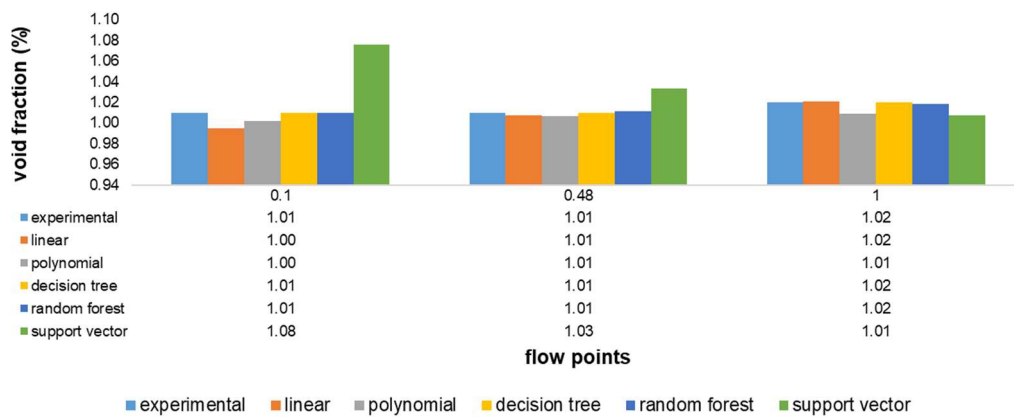


Figure 7-32: Validation – pressure at upward pipe - top

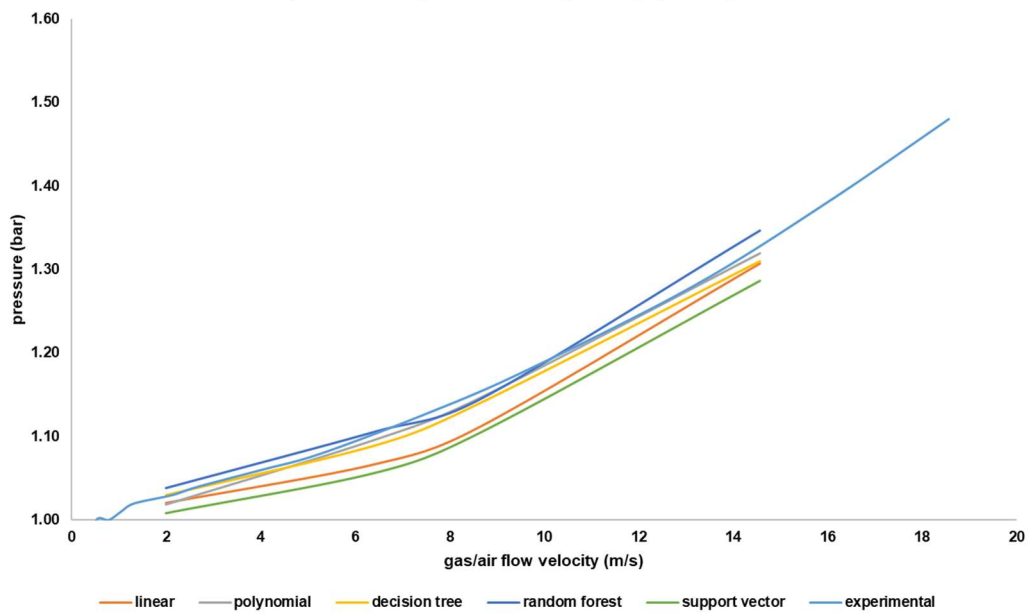


Figure 7-33: Prediction – pressure at upward pipe – top

Figure 7-34 and Figure 7-35 show the validation and prediction of the void fraction at the upward top data point, respectively. The Random forest model shows good performance for the void fraction prediction.

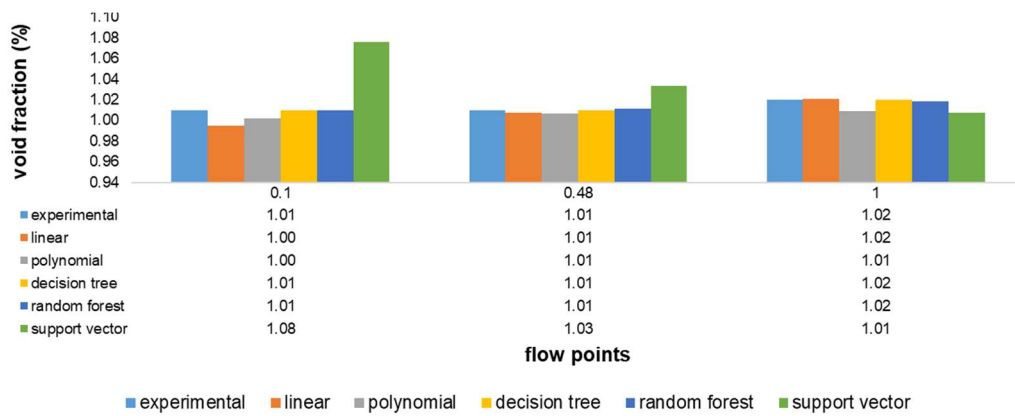


Figure 7-34: Validation – void fraction at upward pipe - top

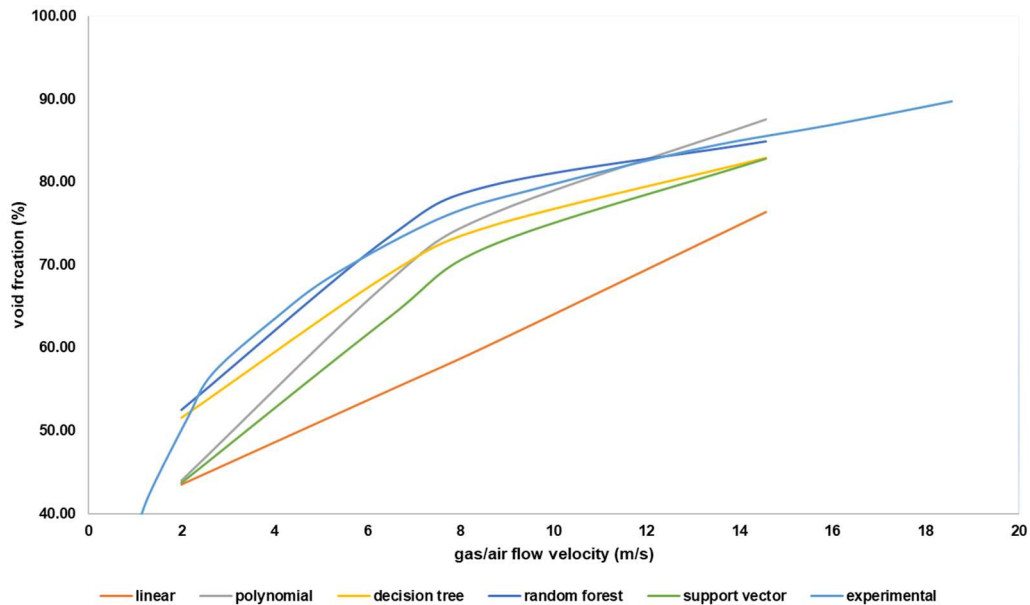


Figure 7-35: Prediction – void fraction at upward pipe – top

Figure 7-36 and Figure 7-37 show the validation and predictions for the liquid film thickness at the downward middle data point, respectively. As can be seen, all models predict well the liquid film thickness. The decision tree's curve seems the most accurate at low gas velocity, but a large difference is seen around 6-7 m/s, before the curves re-establishes itself as the air velocity increases.

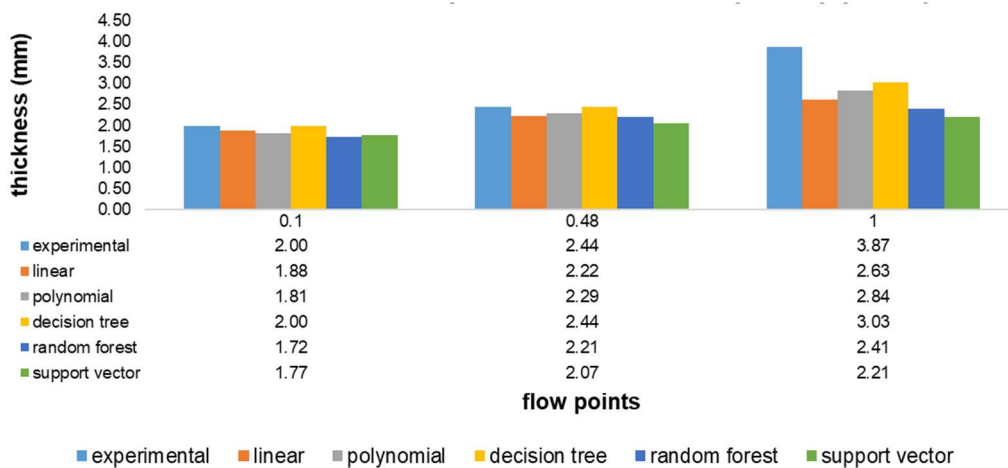


Figure 7-36: Validation – liquid film thickness at upward pipe – top

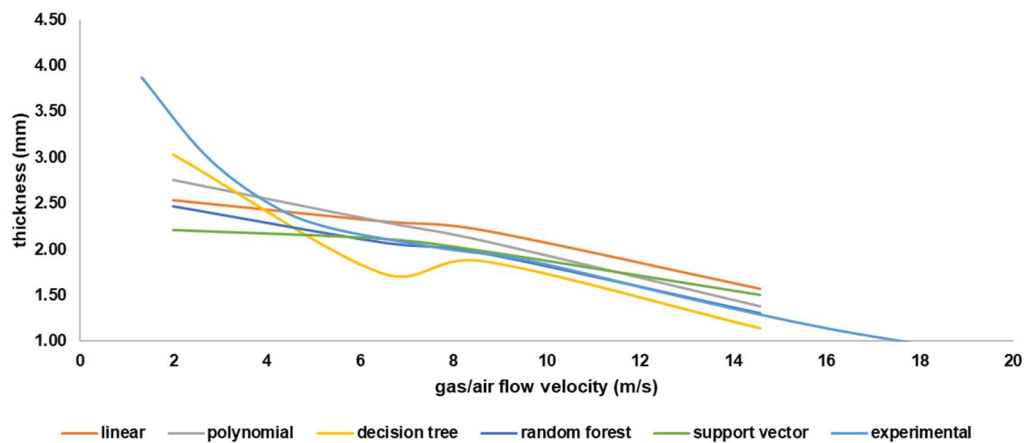


Figure 7-37: Prediction – liquid film thickness at upward pipe – top

7.2.2 ML's Research Findings Summary

7.2.2.1 Pressure

Table 7-3 summarised the trends observed in the results, which indicates the performance of the regression models in capturing pressure variations at various locations within the serpentine pipe.

Table 7-3: Pressure data ML models performance matrix

pressure					
upward top	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Good	Good	Good	Good	Bad
upward middle	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Good	Good	Good	Good	Bad
upward bottom	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Good	Good	Good	Good	Bad

downward bottom	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Good	Good	Good	Good	Bad
downward middle	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Good	Good	Good	Good	Bad
downward top	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Good	Good	Good	Good	Bad

Support Vector Regression Limitation: The Support Vector Regression exhibited consistent challenges in accurately capturing the pressure trends across various locations of the serpentine pipe. This deficiency is marked by the "Bad" rating in both validation and prediction tasks for all segments, including downward top, downward middle, downward bottom, upward bottom, upward middle, and upward top.

Consistent Positive Performance Trend: Across Linear, Polynomial, Decision Tree, and Random Forest: In contrast, Linear, Polynomial, Decision Tree, and Random Forest regression models displayed a positive trend. They consistently demonstrated good performance in capturing pressure variations at all locations.

Promising Predictive Capability: These four models – Linear, Polynomial, Decision Tree, and Random Forest – exhibited a consistent "Good" designation for both validation and prediction tasks across all sections. This indicates their ability to capture and predict pressure behaviours effectively.

Interpreting the Trend: The observed trend can be attributed to the nature of the data and the complexities inherent in the serpentine pipe's flow dynamics. The Support Vector Regression model's limitations in capturing the intricate pressure variations might stem from its sensitivity to outliers or its complexity not aligning well with the specific pressure trends in this dataset.

On the other hand, Linear, Polynomial, Decision Tree, and Random Forest models are likely more adaptable to the pressure behaviors encountered within

the serpentine pipe. Their flexibility allows to capture non-linear relationships, adapt to varying patterns, and handle complexities that can arise in fluid dynamics scenarios.

This trend underscores the importance of choosing models that align well with the data's characteristics and the underlying physics. It is also a reminder that a single model might not universally excel across all scenarios, especially in situations as complex as two-phase flow within a serpentine pipe. A careful consideration of model suitability and interpretation of the results are crucial in deriving meaningful insights from such analyses.

7.2.2.2 Void fraction:

The trends observed in the results of the void fraction are summarised in Table 7-4, indicates the performance of different regression models in capturing the void fraction at various locations within the serpentine pipe. The following summary points can be extracted.

Regression Model Performance: Random Forest model consistently demonstrates robust performance in effectively capturing void fraction variations during both validation and prediction stages across different sections of the serpentine pipe. Polynomial and Decision Tree models exhibit proficient performance, successfully capturing void fraction dynamics in both validation and prediction tasks for specific sections.

Limitation of Support Vector Regression, Linear, and Polynomial Models: The Support Vector Regression, Linear, and Polynomial models consistently struggle to accurately capture void fraction variations at various positions within the serpentine pipe, as indicated by "Bad" designations in both validation and prediction tasks.

Table 7-4: Void Fraction data ML models performance matrix

downward top	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Bad	Good	Bad	Good	Bad
downward middle	linear	polynomial	decision tree	random forest	support vector

validation	Good	Good	Good	Good	Bad
prediction	Bad	Bad	Bad	Good	Bad
downward bottom	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Bad	Bad	Bad	Good	Bad
upward bottom	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Bad	Good	Good	Good	Good
upward middle	linear	polynomial	decision tree	random forest	support vector
validation	Bad	Bad	Bad	Good	Bad
prediction	Bad	bad	bad	Good	Bad
upward top	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Bad
prediction	Bad	Bad	Bad	Good	Bad

Interpreting the Trend: The observed trend underscores the superior adaptability of the Random Forest model in capturing intricate void fraction dynamics. Polynomial and Decision Tree models also demonstrate promise in certain scenarios.

Model Adaptability and Complexity: The flexibility of the Random Forest model enables it to effectively capture both linear and non-linear relationships, adapting to diverse patterns and complexities inherent in fluid dynamics scenarios.

Context-Dependent Model Suitability: The findings emphasize the necessity of selecting regression models that align well with data characteristics and the underlying physics of the system.

The study serves as a reminder that model suitability should be context-dependent, considering the complexities of two-phase flow in a serpentine pipe.

Practical Implications: The research findings highlight the significance of a meticulous model selection in the context of intricate fluid dynamics scenarios.

7.2.2.3 Liquid film thickness:

The findings derived from the data presented in Table 7-5 offer valuable insights into the performance of the regression models in capturing the liquid film thickness variations at distinct locations within the serpentine pipe. Analyzing the results reveals the following noteworthy trends:

Table 7-5: Liquid Film Thickness data ML models performance matrix

liquid film thickness					
downward top	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Bad	Bad	Good	Bad	Bad
downward middle	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Bad	BA	Good	Bad	Bad
downward bottom	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Bad	Bad	Good	Bad	Bad
upward bottom	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Good	Good	Good	Good	Good
upward middle	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Good	Good	Good	Good	Good
upward top	linear	polynomial	decision tree	random forest	support vector
validation	Good	Good	Good	Good	Good
prediction	Good	Good	Good	Good	Good

Downward Top Section:

- Linear, Polynomial, Decision Tree, Random Forest, and Support Vector models achieved a "Good" validation performance.
- However, during prediction, the Support Vector model exhibited limitations with a "Bad" designation, whereas Linear, Polynomial, Decision Tree, and Random Forest models displayed less accurate results.

Downward Middle Section:

- All models consistently demonstrated good performance in validation tasks.
- For prediction, Linear, Polynomial, Decision Tree, and Random Forest models maintained their effectiveness, while the Support Vector model encountered challenges.

Downward Bottom Section:

- Similar to the previous sections, all models exhibited "Good" validation results.
- Yet again, during prediction, Linear, Polynomial, Decision Tree, and Random Forest models remained reliable, while the Support Vector model's accuracy declined.

Upward Bottom Section:

- All models achieved "Good" validation and prediction results, indicating their robustness in capturing liquid film thickness dynamics.
- Upward Middle Section:
- Both Linear and Polynomial models displayed "Good" validation and prediction performance.
- The Decision Tree, Random Forest, and Support Vector models demonstrated "Bad" validation results, with prediction accuracy declining for the Support Vector model.

Upward Top Section:

- All models showcased "Good" validation results.
- In terms of prediction, Linear, Polynomial, and Decision Tree models maintained their accuracy, while Random Forest and Support Vector models faced limitations.

Implications and Interpretation:

The trends observed in the results underscore the complexity of capturing liquid film thickness dynamics. The diverse behavior of different models across various sections of the serpentine pipe emphasizes the significance of model suitability

and adaptability. While Linear and Polynomial models consistently demonstrated effective performance, the behavior of Decision Tree, Random Forest, and Support Vector models was more context-dependent.

These findings emphasize that the choice of a regression model should be guided by the specific characteristics of the data and the inherent complexities of fluid dynamics. The insights gained from this analysis serve as a valuable reminder that model selection should align with the intricacies of the system under investigation, ensuring accurate interpretation and meaningful insights.

7.3 CFD Simulation and ML-Based Geometry Optimization.

Within this section, an in-depth analysis of the CFD-based results is carried out, for serpentine pipes of diameters varying from 50 to 150 mm. The simulation endeavour encompasses three distinct boundary conditions, mirroring those previously established for the CFD validation in Section 6.7. While acknowledging the divergence between the CFD model and the empirical reality, this research trailblazes a novel methodology that seamlessly fuses CFD insights with machine learning techniques, paving a transformative path toward geometry optimization, all while harnessing minimal, yet pivotal physics-based data.

The study introduces a pioneering approach that artfully integrates CFD simulations and machine learning methodologies. This symbiotic matchup becomes the cornerstone for an innovative process, unfolding the potential for geometry optimization. In this context, the research strategically leverages the wealth of CFD results derived from pipes with diameters of 50 mm, 101 mm, and 150 mm. The focus is on the extraction of void fraction data exclusively for the downward pipe, particularly at three distinct stations. This process bypasses the limitations of sparse data points and adroitly interpolates intermediate void fraction values within these specific flow conditions. This approach not only deepens our understanding but also charts a course toward advanced insights, transcending the boundaries of data scarcity.

Acknowledging the constraints imposed by limited data availability, the research transitions into an optimization study, harnessing the power of ML models —

linear, polynomial, decision tree, random forest, and support vector models. While the ML models do not aim for perfect alignment with the CFD model, they serve as a guiding compass for the optimization process. The spotlight remains steadfast on the downward pipe's top, bottom, and middle sections, each undergoing meticulous scrutiny of the ML models.

7.3.1 Simulation results

The CFD simulations encompassed a range of conditions as shown in Table 7-7. The water velocity maintained as 0.8 m/s, 1 m/s and 2m/s throughout the simulations. A total of 135 simulation results is used for the analysis.

Table 7-6: CFD simulation boundary conditions

u_{sl} (m/s)			u_{sg} (m/s)	diameter (mm)		
0.8	1	1.2	0.15	50	101	150
0.8	1	1.2	0.27	50	101	150
0.8	1	1.2	0.52	50	101	150
0.8	1	1.2	0.78	50	101	150
0.8	1	1.2	1.04	50	101	150
0.8	1	1.2	1.32	50	101	150
0.8	1	1.2	2.13	50	101	150
0.8	1	1.2	2.67	50	101	150
0.8	1	1.2	4.01	50	101	150
0.8	1	1.2	5.29	50	101	150
0.8	1	1.2	7.62	50	101	150
0.8	1	1.2	9.65	50	101	150
0.8	1	1.2	13.15	50	101	150
0.8	1	1.2	15.98	50	101	150
0.8	1	1.2	18.56	50	101	150

Table 7-7 shows the void fraction values extracted from the top, bottom, and middle average values of void fraction for 9 cases results, noting that all results are provided in the Appendix A.1. These values are expressed in percentage terms and offer insights into the distribution and behaviour of gas and liquid phases within the fluid medium under different flow conditions. The scalar plots in Figure 7-38, Figure 7-39 and Figure 7-40 represent the proportion of void (gas)

volume within the total volume of the mixture (gas + liquid) at specific sections of the downward pipe.

Table 7-7: Void fraction simulation results

u_{sl} (m/s)	u_{sg} (m/s)	diameter (mm)	void fraction (%)		
			downward_top	downward_middle	downward_bottom
1	0.15	50	14.9	14.09	14.07
1	1	50	50.09	50.07	51.2
1	9	50	90.05	90.09	90.08
1	0.15	101	39.26	43.47	43.16
1	1	101	60.76	53.17	50.06
1	9	101	90.08	90.26	90.34
1	0.15	150	38.8	34.32	30.3
1	1	150	73.9	68.2	69.02
1	9	150	90.05	91.05	91.05

Downward Top Section:

This section refers to the uppermost region of the pipe. It provides insights into the distribution of gas and liquid phases changes near the upper surface.

- As the superficial gas velocity (u_{sg}) increases while keeping other parameters constant, the void fraction tends to increase as well. This is evident from the increasing void fraction values for the same diameter and increasing u_{sg} . For example, in the row where u_{sg} changes from 0.15 m/s to 1 m/s (while other parameters remain the same), the void fraction at the top cross-section increases significantly. This indicates that higher gas velocities result in a larger volume occupied by gas near the pipe's upper surface.
- Similarly, when comparing different diameters while keeping the gas and liquid velocities constant, it observe that larger diameters tend to have higher void fractions. For instance, when comparing the void fractions in rows with the same u_{sg} and u_{sl} but different diameters (e.g., 50 mm and 101 mm), the void fraction at the top cross-section is notably higher for the larger diameter. This suggests that in larger pipes, the gas phase occupies a greater volume near the top section.

Downward Middle Section:

This section represents the core region of the pipe's cross-section. It offers insights into the void fraction distribution within the central flow. Here's how the void fraction varies:

- The influence of superficial gas velocity (u_{sg}) remains consistent in this section. As u_{sg} increases, the void fraction generally increases as well. This trend is observed across rows with varying diameters but the same u_{sl} and varying u_{sg} . For example, higher u_{sg} values lead to higher void fractions in the middle cross-section, indicating that increased gas velocities result in a larger volume of gas within the core flow.
- When comparing different diameters while keeping gas and liquid velocities constant, similar trends as in the top section can be observed. Larger diameters tend to exhibit higher void fractions in the middle cross-section. Rows with different diameters but the same u_{sg} and u_{sl} values reveal this pattern. This indicates that larger pipes allow for more gas to occupy the core region of the flow.

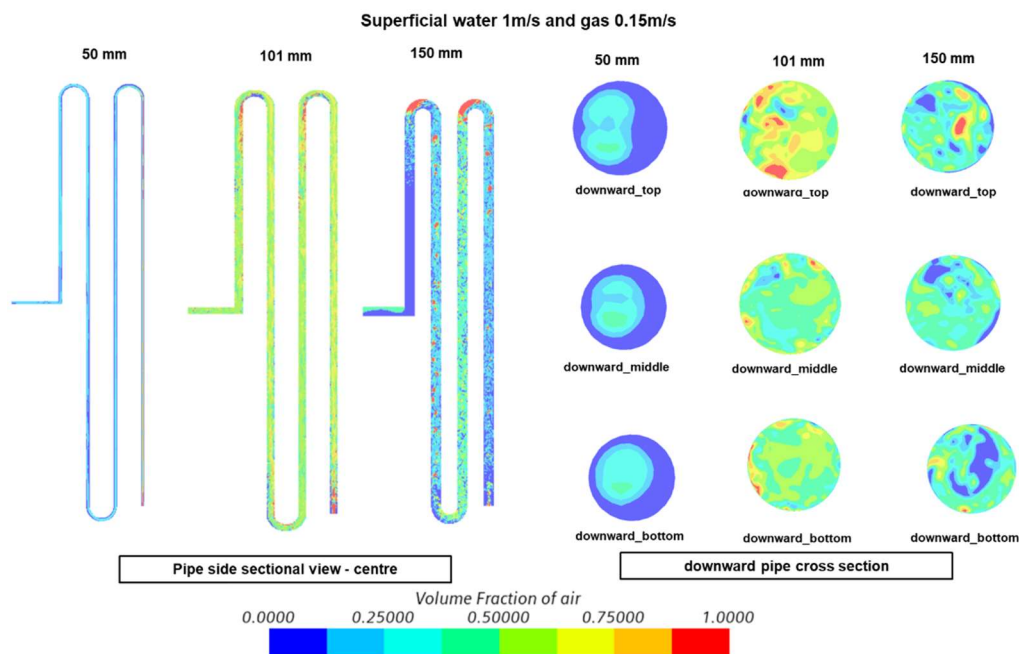


Figure 7-38: Gas velocity 0.15m/s void fraction scalar plot for different pipe diameters

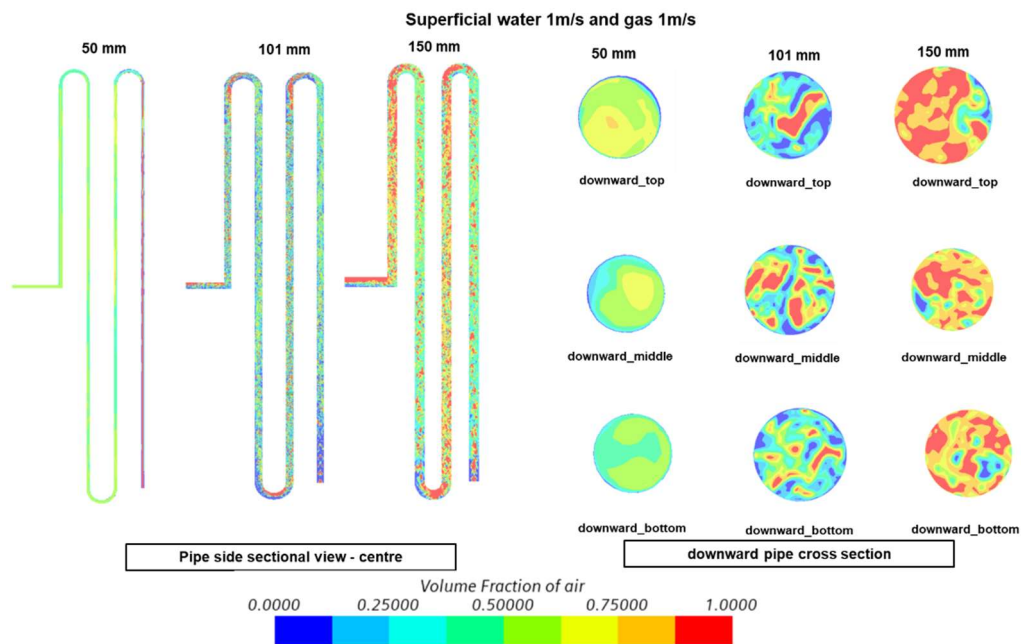


Figure 7-39: Gas velocity 1 m/s void fraction scalar plot for different pipe diameters

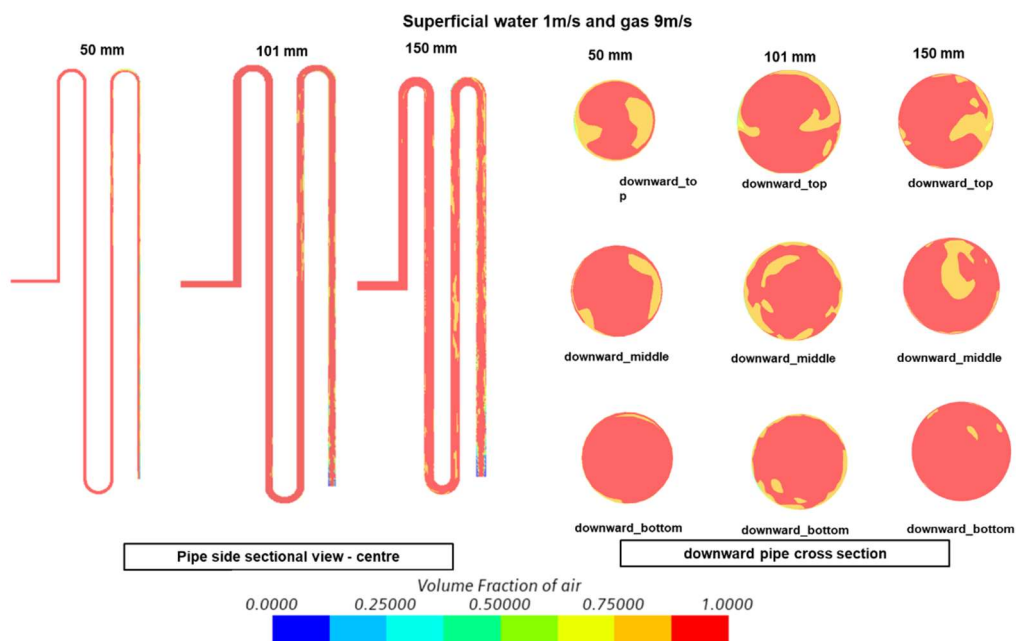


Figure 7-40: Gas velocity 9 m/s void fraction scalar plot for different pipe diameters

Downward Bottom Section:

The lower part of the pipe, near its bottom surface is discussed. This section provides insights into the dynamics of the gas-liquid distribution close to the pipe's base. Here's how the void fraction varies:

- As with the top and middle sections, an increase in superficial gas velocity (u_{sg}) results in higher void fractions in the bottom cross-section. This trend is consistent across rows with different diameters but the same u_{sl} and varying u_{sg} . An example is seen in the increasing void fractions for larger u_{sg} values in the bottom cross-section.
- Larger diameters also lead to higher void fractions in the bottom section. Rows with varying diameters but the same u_{sg} and u_{sl} values demonstrate this pattern. This indicates that, similar to the other sections, larger pipes allow for increased gas occupancy near the pipe's lower surface.

7.3.2 ML Results

The outcomes obtained through CFD simulations serve as the foundation for a comprehensive ML analysis. The primary objective of this analysis is to delve into the intricate interplay between crucial process parameters and the potential for predicting void fraction percentages for specific flow rates and diameters using advanced regression modelling techniques. While this approach holds high potential, it necessitates an increased volume of CFD data to achieve robust outcomes. Given our time constraints, expanding the dataset through additional simulations poses a challenge. The results of this study are not perfect, but they are still valuable. They can help us to understand how to predict void fraction in a simpler way. This could lead to new advances in this field in the future.

In the context of the analysis, a suite of regression modeling techniques was used to establish the relationships between input parameters (u_{sl} , u_{sg} , diameter) and output variables (top, middle, bottom). This code detailed in Appendix A.2, encompassed five distinct regression models. Commencing with Linear Regression, the study delved into linear dependencies, revealing initial trends but acknowledging limitations in capturing data complexities. Transitioning to

Polynomial Regression, non-linear associations were unearthed through a second-degree polynomial, enhancing predictive capabilities by embracing intricate non-linear patterns. For complex interactions, Decision Tree and Random Forest Regression models were employed, highlighting varying parameter importance and predicting output variables with diverse precision levels. This comprehensive modelling spanning collectively illuminates the interplay between inputs and outputs within the dataset.

7.3.2.1 Validation – Mean Absolute Error

During the regression analysis, a comprehensive assessment of various regression models was undertaken to ascertain their effectiveness in predicting the target variables. The Mean Absolute Error (MAE) metric, a pivotal measure quantifying the absolute disparities between predicted and actual target values, was used to gauge the models' predictive performance.

Table 7-8: Mean Absolute Error values.

Linear Regression MAE	14.62
Polynomial Regression MAE	7.50
Decision Tree Regression MAE	4.40
Random Forest Regression MAE	4.15

The computed MAE values were as follows: 14.62 for linear regression, 7.50 for polynomial regression, 4.40 for decision tree regression, and 4.15 for random forest regression as shown in Table 7-8. These MAE values encapsulate the average deviation between predicted and actual values.

The obtained results offer valuable insights into the models' performance. The higher MAE of 14.62 for linear regression indicates that its predictions deviated by approximately 14.62 units, on average, from actual values. Polynomial regression exhibited a lower MAE of 7.50, suggesting better alignment between predictions and actual values due to its ability to capture intricate data relationships.

Furthermore, decision tree regression achieved a more favourable MAE of 4.40, signifying improved predictive accuracy attributed to its capacity to accommodate

non-linearities. Remarkably, the random forest regression outperformed other models with the lowest MAE of 4.15. The ensemble nature of the random forest, blending multiple decision trees, contributed to its exceptional predictive capability and minimized prediction errors.

To summarise, the models' comparative performance reveals that the random forest regression, with its lowest MAE, emerges as the optimal choice for predictive analysis on similar datasets. Thus, the MAE-based evaluation underscores the efficacy of the random forest regression model as the preferred tool for accurate predictions in this context.

7.3.3 Validation – test data

To code validation, the regression models were evaluated using three distinct test datasets, and their outcomes were compared against corresponding datasets from CFD simulations. The validation procedure aimed to ascertain the predictive capability of the models in replicating void fraction values.

The findings of the validation process are presented in Table 7-9 and illustrating the outcomes for Validation-01 is shown Figure 7-41. This validation case involves a set of input conditions, including gas and liquid velocities, as well as pipe diameter. The void fraction results are tabulated for three distinct regions: downward_top, downward_middle, and downward_bottom.

Table 7-9: Validation-01 CFD and ML void fraction results

Validation -01	u_{sl} (m/s)	u_{sg} (m/s)	diameter (mm)	void fraction (%)		
				downward_top	downward_middle	downward_bottom
CFD	1	0.15	150	39.26	43.47	43.16
ML Predictions	Linear			69.1	68.7	67.3
	Polynomial			59.6	58.3	56.4
	Decision tree			38.8	34.3	30.3
	Random Forest			41.3	39.0	36.1
	Support vector			85.3	86.3	85.6

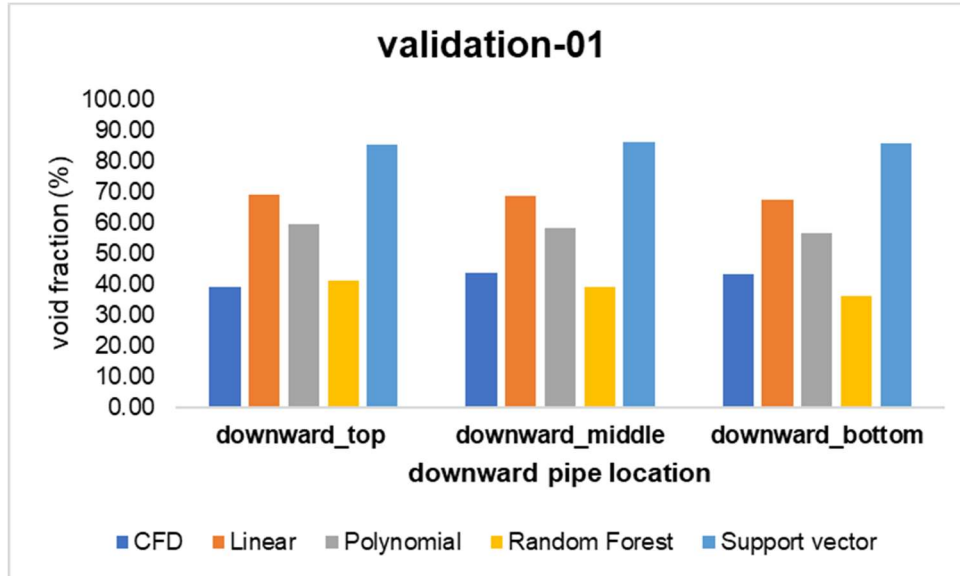


Figure 7-41: Validation 01 CFD vs ML comparison

Similarly, the outcomes of Validation-02 are presented in Table 7-10, accompanied by Figure 7-42 for a comprehensive visual representation. Finally, Validation-03 results are presented in Table 7-11 and displayed in Figure 7-43.

Table 7-10: Validation-02 CFD and ML void fraction results

Validation -02	u_{sl} (m/s)	u_{sg} (m/s)	diameter (mm)	void fraction (%)		
				downward _top	downward _middle	downward _bottom
CFD	1	2.13	101	87.30	87.31	87.29
ML Predictions	Linear			64.9	64.9	64.2
	Polynomial			74.2	74.9	74.6
	Decision tree			87.3	87.3	87.3
	Random Forest			86.7	87.2	87.1
	Support vector			84.9	86.0	85.3

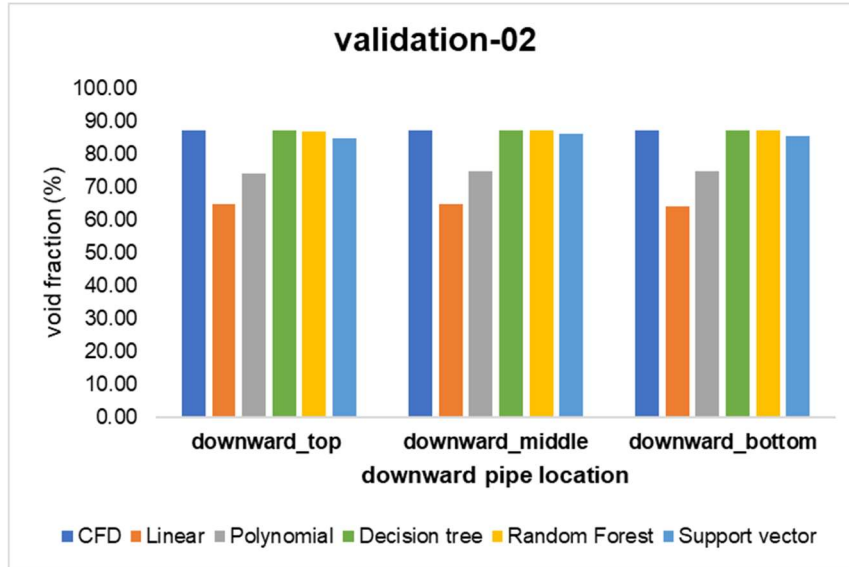


Figure 7-42: Validation-02 CFD vs ML comparison

Table 7-11: validation-0 3CFD and ML void fraction results

Validation -03	u_{sl} (m/s)	u_{sg} (m/s)	diameter (mm)	void fraction (%)		
				downward _top	downward _middle	downward _bottom
CFD	1	0.52	50	23.80	23.81	23.79
ML Predictions	Linear			50.42	50.80	50.69
	Polynomial			34.22	34.43	34.51
	Decision tree			23.80	23.81	23.79
	Random Forest			24.55	24.56	24.54
	Support vector			84.16	85.25	84.64

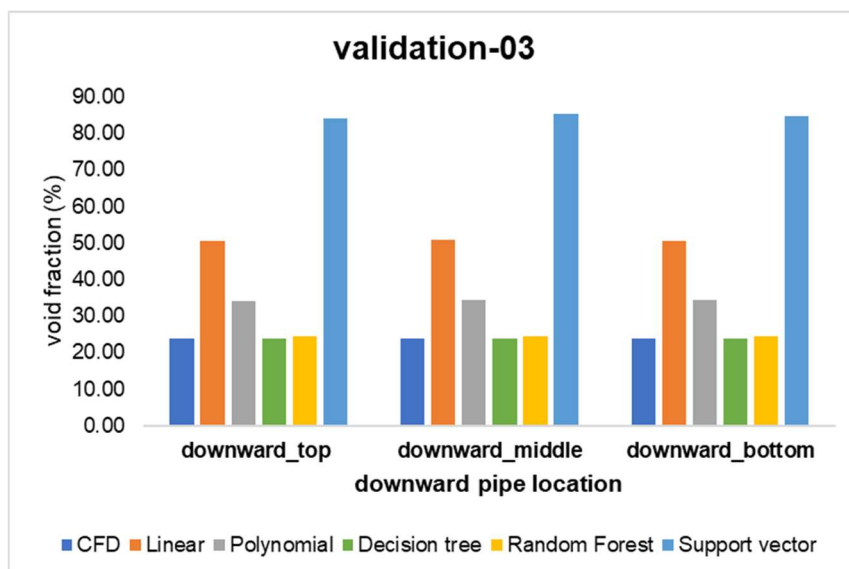


Figure 7-43: Validation-03 CFD vs ML comparison

Upon analysis of the results, it is apparent that there exists a discernible disparity between the void fraction predictions from various regression models and the CFD-derived data. In terms of Mean Absolute Error (MAE), the Decision Tree and Random Forest models emerge as noteworthy contenders, with their respective predictions closely aligning with the CFD data. This alignment is particularly evident for the downward_top, downward_middle and downward bottom regions.

The validation process underscores the efficacy of different regression models in capturing the intricate relationships governing the void fraction within the system. The Decision Tree and Random Forest models stand out for their capacity to approximate the void fraction values with commendable accuracy, as evidenced by their minimal MAE in comparison to the CFD data. While the Support Vector model also demonstrates competitive performance, the Decision Tree and Random Forest models exhibit a desirable balance between accuracy and computational feasibility. The findings thus shed light on the models' aptitude for void fraction prediction and offer insights into their applicability in scenarios characterized by similar input conditions.

7.3.4 `Predictions

Utilizing the ML methodology for geometry optimization involves the objective of ascertaining the void fraction at the upper, middle, and lower segments of the downward pipe. This endeavour targets an optimal diameter for the serpentine pipe, specifically within the scope of 50 to 150 mm. The optimization process hinges on the flow velocities outlined within the dataset for serpentine pipes. Within this segment, void fraction prediction is conducted for two scenarios: one at 75 mm and the other at 125 mm diameters, both with gas 2.1 m/s and water velocities set at 1 m/s, under the assumption of slug flow. Discussion and presentation exclusively center around the outcomes derived from the decision tree and random forest models.

The predictions generated by the code reflect the void fraction values as shown in Table 7-12. Notably, the void fraction values obtained from the random forest model fall within the range of CFD values. This observation underscores the

promising potential of the ML model, especially when combined with CFD data, in enhancing the accuracy of void fraction predictions for serpentine pipes.

Table 7-12: ML predictions vs CFD data.

ML Predictions	u_{sl} (m/s)	u_{sg} (m/s)	diameter (mm)	void fraction (%)		
				downward _top	downward _middle	downward _bottom
CFD	1.0	2.1	50.0	51.4	51.4	51.4
ML - decision tree	1.0	2.1	75.0	51.4	51.4	51.4
ML - random forest	1.0	2.1	75.0	55.5	55.1	55.1
CFD	1.0	2.1	101.0	87.3	87.3	87.3
ML - decision tree	1.0	2.1	125.0	87.3	87.3	87.3
ML - random forest	1.0	2.1	125.0	86.1	87.0	86.9
CFD	1.0	2.1	150.0	80.2	78.4	75.5

However, it is important to note that the accuracy of the predictions is not currently at an optimal level. To enhance the simulation results and achieve more reliable predictions, there is a clear need for further improvements. Increasing the quality and quantity of CFD data could significantly contribute to refining the dataset used for training and validation. This would likely lead to more robust and accurate predictions, as the model would have a broader and more representative set of inputs to learn from. By expanding the CFD dataset and focusing on improving simulation techniques, can work towards achieving more accurate and dependable results in predicting void fraction within serpentine pipes.

8 Conclusions

Conclusions can be drawn from the work carried out in this project and summarised in this thesis.

Fluid Dynamics Context

The serpentine flow pipes comprise upward and downward configurations, each marked by distinct flow behaviors. The downward pipe exhibits bubbly, intermittent, and annular flows, while the upward pipe shows variations. Dynamics in the upward pipe's flow are notably influenced by curvatures, particularly responsive to air and water superficial velocities, contributing to intricate dynamics.

Complexity in Pressure Distribution

The distribution of pressure values along the pipe is intricately nonlinear due to gravitational effects. This interplay of forces adds complexity to the underlying physics. Employing data processing techniques is pivotal in comprehending this intricate dataset, forming the foundation for robust regression models.

Complete System Understanding:

Comprehensive analysis of input and output parameters sheds light on system behaviour. The multifaceted flow behaviours in serpentine pipes, shaped by air and water superficial velocities, are further compounded by non-linear pressure variations due to gravity. Scrutinizing both input parameters (superficial velocities) and output data (pressure values) systematically paves the way for accurate regression model development.

Pressure Analysis:

- Linear, Polynomial, Decision Tree, and Random Forest models consistently capture pressure variations, demonstrating adaptability to complexity.
- Support Vector Regression faces challenges, reflected in "Bad" ratings across all sections, highlighting limitations.

- Adaptability issues may arise from Support Vector Regression's sensitivity to outliers and misalignment with complexity. In contrast, other models adeptly navigate non-linear relationships.
- Model suitability reveals nuances, underscoring the need for context-driven model selection for nuanced insights.

Void Fraction Analysis:

- Random Forest excels, proficiently capturing void fraction dynamics in validation and prediction.
- Support Vector Regression, Linear, and Polynomial models lag consistently, struggling with intricacies.
- Adaptability remains pivotal, with Random Forest showcasing superior capability, while Polynomial and Decision Tree models show promise in specific contexts.

Liquid Film Thickness Analysis:

- Liquid film thickness trends vary across sections. Linear, Polynomial, and Decision Tree models consistently maintain effectiveness.
- The persistent theme of context-driven suitability resonates, with Decision Tree, Random Forest, and Support Vector models displaying mixed outcomes.

Practical Implications:

Intricacies of model selection are paramount, necessitating alignment with data intricacies and fluid dynamics. Future research avenues should delve into adaptable model potentials for navigating complex dynamics. The significance of balanced model selection cannot be overstated, empowering accurate interpretation and fostering meaningful insights.

Revealing Insights:

Our research unravels the intricate interplay between model adaptability and fluid dynamics complexity. It underscores the importance of tailored model choices, guiding the way to uncover hidden insights. This work deepens our understanding

of fluid dynamics while offering a roadmap for effectively navigating intricacies in various systems.

The pursuit of geometry optimization for serpentine pipes has traditionally demanded a significant investment of time and resources, often requiring complex experimental procedures and detailed CFD simulations. This study explores a new way to use CFD and machine learning ML together. This new approach could revolutionize the field of CFD, which is a complex and challenging field.

A promising avenue is unveiled for predicting complex flow behavior. The integration of these approaches addresses the challenges posed by time-consuming processes, offering a streamlined methodology with the potential to deliver accurate results efficiently.

The investigation reveals that, for the specific geometry of serpentine pipes considered in this study, the fusion of CFD and machine learning holds tremendous potential. In particular, the Decision Tree and Random Forest models have emerged as the most adequate tools for generating accurate predictions. These models exhibit a noteworthy capacity to navigate the complexities of the system, resulting in predictions that closely align with actual outcomes.

However, it is essential to acknowledge the prevailing limitations. The dearth of data remains a significant hurdle in achieving more refined and dependable results. Even within the confines of a limited timeframe, the quantity and quality of data are pivotal in enhancing the accuracy of predictions. Despite this constraint, the preliminary results underscore the promising trajectory of machine learning in this realm.

Without a robust dataset, the machine learning model's capacity to comprehend and predict outcomes is compromised.

Future endeavours in this domain should prioritize the development of strategies to accumulate more data and execute comprehensive CFD simulations. This

concerted approach will not only refine the effectiveness of machine learning applications for CFD simulations but will extend its relevance to other fields.

In conclusion, this study showcases the feasibility of geometry optimization for serpentine pipes through the innovative fusion of machine learning, CFD, and empirical data. Progressing towards superior predictions and insights necessitates an earnest commitment to bolstering datasets and refining simulation techniques. As this path unfolds, the latent potential of machine learning in shaping the trajectory of fluid dynamics and optimization becomes increasingly apparent.

9 Future Work

The future of this research holds several exciting avenues for exploration and advancement. Building upon the insights gained from the current study, several key directions can be pursued to further enhance our understanding and contribute to the field of serpentine flow pipe analysis:

Advanced Numerical Simulations: Given the complexity of the flow behaviours observed in the serpentine pipes, further refinement of the CFD simulations is paramount. Employing more sophisticated turbulence models and adaptive meshing strategies can aid in capturing subtle flow dynamics and boundary effects that may have been missed in the current simulations.

Multiphase Flow Modelling: The inclusion of additional phases, such as solid particles or droplets, could significantly impact the flow behaviours within the serpentine pipes. Investigating the interactions and transport phenomena of these additional phases can provide a more comprehensive understanding of the flow dynamics and their implications for practical applications.

Machine Learning Integration: Integrating machine learning techniques into the analysis can offer predictive capabilities beyond the capabilities of traditional models. By training machine learning algorithms on the extensive dataset generated through simulations and experiments, it is possible to predict flow behaviours and pressure distributions with greater accuracy.

Optimization of Geometry: With the aid of machine learning, the optimization of the serpentine pipe geometry can be explored. By establishing a direct link between geometry variations and desired flow outcomes, it becomes possible to design pipes that achieve specific performance objectives, such as enhanced mixing efficiency or reduced pressure drop.

Experimental Validation: Further experimental investigations can serve to validate and refine the insights gained from computational simulations. Advanced measurement techniques, such as high-speed imaging and particle image velocimetry, can offer detailed flow visualization and data validation, helping to corroborate simulation results.

Real-world Applications: Applying the findings of this research to practical applications within industries such as energy, manufacturing, and environmental engineering holds immense potential. Understanding the intricate flow behaviours within complex geometries can lead to the design of more efficient and optimized systems.

Sensitivity Analysis: Conducting a comprehensive sensitivity analysis to understand the influence of individual parameters on flow behaviours can help identify key drivers that dictate the system's performance. This can guide future experiments and simulations to focus on critical areas.

In conclusion, more research can be done on the understanding of the flow in serpentine pipes. By delving into more sophisticated simulations, integrating machine learning, and exploring real-world applications, the potential to revolutionize the design and operation of fluid systems is within reach. This research has laid a solid foundation for future endeavours that can contribute to advancements in fluid dynamics, numerical modelling, and engineering solutions.

REFERENCES

- Abuhatira, A.A., Salim, S.M. and Vorstius, J.B. (2021) 'Numerical Simulation of Turbulent Pipe Flow With 90-Degree Elbow Using Wall $y +$ Approach', *Volume 10: Fluids Engineering*. American Society of Mechanical Engineers.
- AL-Dogail, A.S. and Gajbhiye, R.N. (2021) 'Effects of density, viscosity and surface tension on flow regimes and pressure drop of two-phase flow in horizontal pipes', *Journal of Petroleum Science and Engineering*, 205 Elsevier B.V.
- AL-Dogail, A.S.G.R.N.A.-N.M.A.A.A.A.H.Y. and A.A.A. (2003) 'Machine Learning Approach to Predict Pressure Drop of Multi-Phase Flow in Horizontal Pipe and Influence of Fluid Properties', *Paper presented at the Gas & Oil Technology Showcase and Conference, Dubai, UAE, March 2023*
- Andersson, B.A.B., Andersson, R., Håkansson, L., Mortensen, M., Sudiyo, R. and van Wachem, B.G.M. (2011) '*Computational Fluid Dynamics for Engineers*',
- Arteaga-Arteaga, H.B., Mora-Rubio, A., Florez, F., Murcia-Orjuela, N., Diaz-Ortega, C.E., Orozco-Arias, S., delaPava, M., Bravo-Ortíz, M.A., Robinson, M., Guillen-Rondon, P. and Tabares-Soto, R. (2021) 'Machine learning applications to predict two-phase flow patterns', *PeerJ Computer Science*, 7 PeerJ Inc.
- Bennett, K.P. and Blue, J.A. (1998) 'A support vector machine approach to decision trees', *1998 IEEE International Joint Conference on Neural Networks Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98CH36227)*., Vol.3, pp. 2396–2401 vol.3.
- Berthelsen, P.A. and Ytrehus, T. (2005) 'Calculations of stratified wavy two-phase flow in pipes', *International Journal of Multiphase Flow*, 31(5) Pergamon, pp. 571–592. Available at: 10.1016/J.IJMULTIPHASEFLOW.2005.02.001 (Accessed: 22 August 2023).
- Brunton, S.L. (2021) Applying machine learning to study fluid mechanics *Acta Mechanica Sinica/Lixue Xuebao*. Springer Verlag,

Brunton, S.L., Noack, B.R. and Koumoutsakos, P. (2019) 'Machine Learning for Fluid Mechanics', *Annu. Rev. Fluid Mech.* 2020, 52, pp. 477–508.

Burstyn, H.C., Sengers, J. V, Bhattacharjee, J.K. and Ferrell, R.A. (1983) 'Dynamic scaling function for critical fluctuations in classical fluids', *Physical Review A*, 28(3) American Physical Society, pp. 1567–1578.

Calay, R.K. and Holdø, A.E. (2003) '*CFD Modelling of Multiphase Flows: An Overview*',

Cazemier, W., Verstappen, R.W.C.P. and Veldman, A.E.P. (1998) 'Proper orthogonal decomposition and low-dimensional models for driven cavity flows', *Physics of Fluids*, 10(7), pp. 1685–1699.

Cheah, M.J., Kevrekidis, I.G. and Benziger, J.B. (2013) 'Water slug formation and motion in gas flow channels: The effects of geometry, surface wettability, and gravity', *Langmuir*, 29(31), pp. 9918–9934.

Chicco Davide AND Oneto, L.A.N.D.T.E. (2022) 'Eleven quick tips for data cleaning and feature engineering', *PLOS Computational Biology*, 18(12) Public Library of Science, pp. 1–21.

Ciofalo, M. and Di Liberto, M. (2015) 'Fully developed laminar flow and heat transfer in serpentine pipes', *International Journal of Thermal Sciences*, 96 Elsevier Masson SAS, pp. 248–266.

Cotton J. S. Ph.D., Shoukri M. Ph.D. and Chang J. S. Ph.D. (2001) 'OSCILLATORY ENTRAINED DROPLET EHD TWO-PHASE FLOW', *Journal of Heat Transfer*, 123(4), p. 622.

Cox, N.J. (2007) 'PAIRPLOT: Stata module for plots of paired observations', *Statistical Software Components*

Dukler, A.E. and Taitel, Y. (1976)

van Eckeveld, A.C., Gotfredsen, E., Westerweel, J. and Poelma, C. (2018) 'Annular two-phase flow in vertical smooth and corrugated pipes', *International*

Journal of Multiphase Flow, 109 Pergamon, pp. 150–163. Available at: 10.1016/J.IJMULTIPHASEFLOW.2018.07.004 (Accessed: 22 August 2023).

Eyo, E.N., Pilario, K.E.S., Lao, L. and Falcone, G. (2021) 'Development of a Real-Time Objective Gas-Liquid Flow Regime Identifier Using Kernel Methods', *IEEE Transactions on Cybernetics*, 51(5) Institute of Electrical and Electronics Engineers Inc., pp. 2688–2698.

Faghri, A. (2012) 'Review and Advances in Heat Pipe Science and Technology', *Journal of Heat Transfer*, 134(12)

Falcone, G., Hewitt, G., Alimonti, C. and Harrison, B. (2001) *Multiphase Flow Metering: Current Trends and Future Developments*.

Frenken, K. and Nuvolari, A. (2004) 'The early development of the steam engine: an evolutionary interpretation using complexity theory', *Industrial and Corporate Change*, 13(2), pp. 419–450.

Hastie Trevor and Friedman, J. and T.R. (2001) 'Linear Methods for Regression', in *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer New York, pp. 41–78.

Hou, J., Li, H., Hu, J., Zhao, C., Guo, Y., Li, S. and Pan, Q. (2017) 'A review of the applications and hotspots of reinforcement learning', *2017 IEEE International Conference on Unmanned Systems (ICUS)*., pp. 506–511.

Hussein, H.M.A., Ahmed, S.T. and Habeeb, L.J. (2022) 'Numerical and experimental investigation of two phase flow geometrical characteristics', *Curved and Layered Structures*, 9(1) De Gruyter Open Ltd, pp. 212–226.

Iaccarino, G. and Elkins, C.J. (2006) 'Towards Rapid Analysis of Turbulent Flows in Complex Internal Passages', *Flow, Turbulence and Combustion*, 77(1), pp. 27–39.

Islam Fetoui (2015.) *Multiphase flow regimes.*, <https://production-technology.org/multiphase-flow-regimes/>

Jiang, G., Gao, J., Shi, X., Li, F. and Xu, L. (2020) 'Flow and Heat Transfer Characteristics of the Mist/Steam Two-Phase Flow Cooling the Rectangular Channel with Column-Row-Ribs', *International Journal of Heat and Mass Transfer*, 156 Pergamon, p. 119737. Available at: 10.1016/J.IJHEATMASSTRANSFER.2020.119737 (Accessed: 22 August 2023).

Jiang, Y. and Rezkallah, K.S. (1993) 'A STUDY ON VOID FRACTION IN VERTICAL CO-CURRENT UPWARD AND DOWNWARD TWO-PHASE GAS-LIQUID FLOW-II: CORRELATIONS', *Chemical Engineering Communications*, 126, pp. 245–259.

Jimack, P.K., Mahmood, R., Walkley, M.A. and Berzins, M. (2002) 'A multilevel approach for obtaining locally optimal finite element meshes', *Advances in Engineering Software*, 33(7–10) Elsevier, pp. 403–415. Available at: 10.1016/S0965-9978(02)00077-7 (Accessed: 29 August 2023).

Kaboodvand, H., Majidi, S. and Jannesari, H. (2021) 'Investigating the effect of inserting longitudinal fins on the overall thermal performance of spirally coiled tubes', *International Communications in Heat and Mass Transfer*, 123 Pergamon, p. 105232. Available at: 10.1016/J.ICHEATMASSTRANSFER.2021.105232 (Accessed: 30 August 2023).

Kang, S., Kim, H. and Kim, D. (2021) 'Mitigation of pressure fluctuation in two-phase slug flow by rib geometry', *International Journal of Multiphase Flow*, 143 Pergamon, p. 103753. Available at: 10.1016/J.IJMULTIPHASEFLOW.2021.103753 (Accessed: 22 August 2023).

Katopodes, N.D. (2019) 'Volume of Fluid Method', *Free-Surface Flow*, Butterworth-Heinemann, pp. 766–802. Available at: 10.1016/B978-0-12-815485-4.00018-8 (Accessed: 29 August 2023).

Kendoush, A.A. and Ai-Khatib, S.A.W. (1994) *Experiments on Flow Characterization in Vertical Downward Two-Phase Flow*.

Kuhn Max and Johnson, K. (2013) 'Data Pre-processing', in *Applied Predictive Modeling*. New York, NY: Springer New York, pp. 27–59.

Kulkarni, V.Y. and Sinha, P.K. (2014) '*Effective Learning and Classification using Random Forest Algorithm*',

Li, W., Liu, D., Desbrun, M., Huang, J. and Liu, X. (2021) 'Kinetic-Based Multiphase Flow Simulation', *IEEE Transactions on Visualization and Computer Graphics*, 27(7) IEEE Computer Society, pp. 3318–3334.

Lini~t, A. and Lopez, D. (1997) *TWO-FLUID MODEL OF WAVY SEPARATED TWO-PHASE FLOW*.

Liu, X., Chen, Y. and Shi, M. (2012) 'Influence of gravity on gas–liquid two-phase flow in horizontal pipes', *International Journal of Multiphase Flow*, 41 Pergamon, pp. 23–35. Available at: 10.1016/J.IJMULTIPHASEFLOW.2011.12.008 (Accessed: 22 August 2023).

Lote, D.A., Vinod, V. and Patwardhan, A.W. (2018) 'Computational Fluid Dynamics Simulations of the Air–Water Two-Phase Vertically Upward Bubbly Flow in Pipes', *Industrial & Engineering Chemistry Research*, 57(31) American Chemical Society, pp. 10609–10627.

Lun, I., Calay, R.K. and Holdo, A.E. (1996) 'Modelling two-phase flows using CFD', *Applied Energy*, 53(3) Elsevier, pp. 299–314. Available at: 10.1016/0306-2619(95)00024-0 (Accessed: 22 August 2023).

Majchrzak, M., Marciniak-Lukasiak, K. and Lukasiak, P. (2023) 'A Survey on the Application of Machine Learning in Turbulent Flow Simulations', *Energies*, 16(4)

Mays, L.W., Koutsoyiannis, D. and Angelakis, A.N. (2007) 'A brief history of urban water supply in antiquity', *Water Supply*, 7(1), pp. 1–12.

Mazumder, Q.H. (2012) 'CFD Analysis of Single and Multiphase Flow Characteristics in Elbow', *Engineering*, 04(04) Scientific Research Publishing, Inc, pp. 210–214.

Naphon, P. and Wongwises, S. (2006a) A review of flow and heat transfer characteristics in curved tubes *Renewable and Sustainable Energy Reviews*. Elsevier Ltd,

Naphon, P. and Wongwises, S. (2006b) 'A review of flow and heat transfer characteristics in curved tubes', *Renewable and Sustainable Energy Reviews*, 10(5) Pergamon, pp. 463–490. Available at: 10.1016/J.RSER.2004.09.014 (Accessed: 30 August 2023).

Nguyen, V.T., Song, C.H., Bae, B.U. and Euh, D.J. (2013) 'Modeling of bubble coalescence and break-up considering turbulent suppression phenomena in bubbly two-phase flow', *International Journal of Multiphase Flow*, 54 Pergamon, pp. 31–42. Available at: 10.1016/J.IJMULTIPHASEFLOW.2013.03.001 (Accessed: 22 August 2023).

Obaid, H.S., Dheyab, S.A. and Sabry, S.S. (2019) 'The Impact of Data Pre-Processing Techniques and Dimensionality Reduction on the Accuracy of Machine Learning', *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*., pp. 279–283.

Ohn, J. and Ustice, E. (1911.) *Experiments on Stream-line Motion in Curved Pipes*.

Ostertagová, E. (2012) 'Modelling using Polynomial Regression', *Procedia Engineering*, 48 No longer published by Elsevier, pp. 500–506. Available at: 10.1016/J.PROENG.2012.09.545 (Accessed: 29 August 2023).

Ramprasad, S., Mallikarjuna, B. and Bhatta, S.H.C.V.S. (2021) 'Numerical and linear regression analysis on MHD two-phase flow in an asymmetric nonuniform channel', *Heat Transfer*, 50(7), pp. 7100–7119.

Rizzo, F., Zazzini, P., Pasculli, A. and Di Crescenzo, A. (2020) 'Statistical approach to compute a surrogate input for building physics CFD simulations through experimental measurements', *Journal of Computational Methods in Sciences and Engineering*, 20 IOS Press, pp. 909–936.

Roul, M.K. and Dash, S.K. (2012) 'Single-phase and two-phase flow through thin and thick orifices in horizontal pipes', *Journal of Fluids Engineering, Transactions of the ASME*, 134(9)

Seo, J. and Lee, J.Y. (2016.) *CFD Analysis and Visualization of the Two Phase Flow in a Thermosyphon for a Passive Heat Removal System of a Nuclear Power Plant*.

Sharma, P., Chung, W.T., Akoush, B. and Ihme, M. (2023) A Review of Physics-Informed Machine Learning in Fluid Mechanics *Energies*. MDPI,

Shawkat, M.E., Ching, C.Y. and Shoukri, M. (2008) 'Bubble and liquid turbulence characteristics of bubbly flow in a large diameter vertical pipe', *International Journal of Multiphase Flow*, 34(8) Pergamon, pp. 767–785. Available at: 10.1016/J.IJMULTIPHASEFLOW.2008.01.007 (Accessed: 22 August 2023).

Srinivasan, V. and Christensen, R.N. (1992) 'Experimental investigation of heat transfer and pressure drop characteristics of flow through spirally fluted tubes', *Experimental Thermal and Fluid Science*, 5, pp. 820–827.

Steven L. Brunton, J.N.K. (2022) *Data-Driven Science and Engineering Machine Learning, Dynamical Systems, and Control*. 2nd edn. Cambridge University Press.

Vasiliev, L.L. (2005) 'Heat pipes in modern heat exchangers', *Applied Thermal Engineering*, 25(1), pp. 1–19.

Vaze, M.J. and Banerjee, J. (2011) 'Experimental visualization of two-phase flow patterns and transition from stratified to slug flow', *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 225(2), pp. 382–389.

Vinuesa, R. and Brunton, S. (2022) '*Emerging trends in machine learning for computational fluid dynamics*'

Vinuesa, R., Brunton, S.L. and McKeon, B.J. (2023) '*The transformative potential of machine learning for experiments in fluid mechanics*'

Wilcox, D. (2012.) 'Formulation of the k-omega Turbulence Model Revisited', in *45th AIAA Aerospace Sciences Meeting and Exhibit*.

Xin, R.C., Dong, Z.F. and Ebadian (1996) '*Three-dimensional numerical modeling of turbulent single-phase and two-phase flow in curved pipes*',

Yang, R. and Sung Fa Chang (1993) 'A numerical study of fully developed laminar flow and heat transfer in a curved pipe with arbitrary curvature ratio', *International Journal of Heat and Fluid Flow*, 14(2) Elsevier, pp. 138–145. Available at: 10.1016/0142-727X(93)90021-E (Accessed: 30 August 2023).

Yeung, H. (2009) *Almabrok Abushanaf Almabrok Gas-Liquid Two-Phase Flow in Up and Down Vertical Pipes*.

Zhang, P., Roberts, R. and Benard, A. (2012) 'Computational guidelines and an empirical model for particle deposition in curved pipes using an Eulerian-Lagrangian approach', *Journal of Aerosol Science*, 53, pp. 1–20.

Zhang, T., Cao, B., Fan, Y., Gonthier, Y., Luo, L. and Wang, S. (2011) 'Gas–liquid flow in circular microchannel. Part I: Influence of liquid physical properties and channel diameter on flow patterns', *Chemical Engineering Science*, 66(23) Pergamon, pp. 5791–5803. Available at: 10.1016/J.CES.2011.07.035 (Accessed: 30 August 2023).

Zhang, Z., McDonnell, K.T., Zadok, E. and Mueller, K. (2015) 'Visual Correlation Analysis of Numerical and Categorical Data on the Correlation Map', *IEEE Transactions on Visualization and Computer Graphics*, 21(2), pp. 289–303.

Zhou, W., Shao, Z., Yu, J. and Lin, J. (2021) 'Advances and Trends in Forming Curved Extrusion Profiles', *Materials*, 14(7) Multidisciplinary Digital Publishing Institute, p. 1603. Available at: 10.3390/ma14071603 (Accessed: 22 August 2023).

APPENDICES

CFD Results and Machine learning Code

A.1 CFD simulation void fraction average at downward, top bottom and middle.

u_{sl} (m/s)	u_{sg} (m/s)	diameter (mm)	Top (%)	Middle(%)	Bottom(%)
1	0.15	50	14.9	14.09	14.07
1	0.27	50	18.2	18.18	18.19
1	0.52	50	23.8	23.81	23.79
1	0.78	50	29.3	29.28	29.3
1	1.04	50	34.6	34.62	34.6
1	1.32	50	40.9	40.89	40.92
1	2.13	50	51.4	51.39	51.42
1	2.67	50	57.7	57.72	57.69
1	4.01	50	69.3	69.29	69.32
1	5.29	50	79.8	79.82	79.79
1	7.62	50	86.9	86.92	86.91
1	9.65	50	90.1	90.12	90.11
1	13.15	50	92.5	92.51	92.49
1	15.98	50	94.7	94.72	94.71
1	18.56	50	96	96	96
1	0.15	101	39.26	43.47	43.16
1	0.27	101	46.59	50.79	50.35
1	0.52	101	55.17	59.53	59.11
1	0.78	101	63.84	68.35	67.92
1	1.04	101	72.5	77.16	76.7
1	1.32	101	81.16	85.95	85.44
1	2.13	101	87.3	87.31	87.29
1	2.67	101	91.5	91.52	91.51
1	4.01	101	94	94.01	93.99
1	5.29	101	95.8	95.82	95.79
1	7.62	101	96.8	96.81	96.79
1	9.65	101	96	96	96
1	13.15	101	96	96	96
1	15.98	101	96	96	96
1	18.56	101	96	96	96
1	0.15	150	38.8	34.32	30.3
1	0.27	150	45.1	42.27	39.16
1	0.52	150	53.6	51.75	49.12
1	0.78	150	62.1	60.29	57.54
1	1.04	150	70.6	68.81	66.01
1	1.32	150	80.2	78.36	75.46

1	2.13	150	86.3	86.32	86.31
1	2.67	150	90.5	90.52	90.51
1	4.01	150	92.9	92.91	92.89
1	5.29	150	94.7	94.72	94.71
1	7.62	150	96	96	96
1	9.65	150	97.08	97.58	96.875
1	13.15	150	98.88	98.88	98.888
1	15.98	150	98.56	98	98
1	18.56	150	96	98.5	98.8
0.8	0.15	50	23.63	23.68	23.61
0.8	0.27	50	26.65	26.69	26.63
0.8	0.52	50	31.31	31.35	31.29
0.8	0.78	50	35.99	36.03	35.96
0.8	1.04	50	40.25	40.29	40.23
0.8	1.32	50	45.44	45.48	45.41
0.8	2.13	50	56.32	56.36	56.28
0.8	2.67	50	61.08	61.12	61.04
0.8	4.01	50	67.63	67.67	67.59
0.8	5.29	50	73.56	73.6	73.52
0.8	7.62	50	79.48	79.52	79.44
0.8	9.65	50	83.04	83.08	83
0.8	13.15	50	88.36	88.4	88.32
0.8	15.98	50	90.75	90.79	90.71
0.8	18.56	50	92.52	92.56	92.48
0.8	0.15	101	58.68	41.36	41.91
0.8	0.27	101	63.05	53.63	54.24
0.8	0.52	101	69	65.83	65.78
0.8	0.78	101	74.76	74.76	74.76
0.8	1.04	101	80.26	80.26	80.26
0.8	1.32	101	85.18	85.18	85.18
0.8	2.13	101	87.3	87.31	87.29
0.8	2.67	101	91.5	91.52	91.51
0.8	4.01	101	94	94.01	93.99
0.8	5.29	101	95.8	95.82	95.79
0.8	7.62	101	96.8	96.81	96.79
0.8	9.65	101	96	96	96
0.8	13.15	101	96	96	96
0.8	15.98	101	96	96	96
0.8	18.56	101	96	96	96
0.8	0.15	150	44.05	38.62	35.73
0.8	0.27	150	47.1	42.77	39.81
0.8	0.52	150	51.67	48.17	45.18
0.8	0.78	150	56.37	53.91	50.89
0.8	1.04	150	61.1	59.59	56.54
0.8	1.32	150	66	65.44	62.35
0.8	2.13	150	74.98	74.34	70.54

0.8	2.67	150	78.85	78.21	74.36
0.8	4.01	150	85.72	84.17	79.55
0.8	5.29	150	91.48	89.87	84.59
0.8	7.62	150	96.58	94.07	89.46
0.8	9.65	150	98.89	98.89	99.29
0.8	13.15	150	98.89	98.89	99.29
0.8	15.98	150	98.89	98.89	99.29
0.8	18.56	150	98.89	98.89	99.29
1.2	0.15	50	15.97	15.14	15.16
1.2	0.27	50	19.59	19.57	19.58
1.2	0.52	50	25.47	25.49	25.46
1.2	0.78	50	31.27	31.29	31.26
1.2	1.04	50	36.69	36.71	36.68
1.2	1.32	50	43.06	43.05	43.08
1.2	2.13	50	54.22	54.23	54.2
1.2	2.67	50	60.51	60.53	60.49
1.2	4.01	50	73.27	73.26	73.29
1.2	5.29	50	83.86	83.88	83.85
1.2	7.62	50	91.31	91.33	91.3
1.2	9.65	50	94.71	94.73	94.7
1.2	13.15	50	97.04	97.06	97.03
1.2	15.98	50	98.84	98.86	98.83
1.2	18.56	50	100.11	100.13	100.1
1.2	0.15	101	41.72	45.96	45.51
1.2	0.27	101	49.04	53.25	52.79
1.2	0.52	101	57.61	61.87	61.4
1.2	0.78	101	66.29	70.6	70.12
1.2	1.04	101	75.03	79.37	78.87
1.2	1.32	101	83.79	88.17	87.65
1.2	2.13	101	89.94	94.37	93.83
1.2	2.67	101	94.14	97.97	97.71
1.2	4.01	101	96.69	97.85	97.62
1.2	5.29	101	97.67	98.16	98
1.2	7.62	101	98.13	98.52	98.41
1.2	9.65	101	98.28	98.57	98.46
1.2	13.15	101	98.28	98.57	98.46
1.2	15.98	101	98.28	98.57	98.46
1.2	18.56	101	98.28	98.57	98.46
1.2	0.15	150	49.06	49.11	49.04
1.2	0.27	150	53.06	53.11	53.04
1.2	0.52	150	59.07	59.12	59.05
1.2	0.78	150	64.07	64.12	64.05
1.2	1.04	150	69.07	69.12	69.05
1.2	1.32	150	74.07	74.12	74.05
1.2	2.13	150	83.57	83.62	83.55
1.2	2.67	150	88.81	88.86	88.79

1.2	4.01	150	95.88	95.93	95.86
1.2	5.29	150	98.42	98.47	98.4
1.2	7.62	150	98.42	98.47	98.4
1.2	9.65	150	98.42	98.47	98.4
1.2	13.15	150	98.42	98.47	98.4
1.2	15.98	150	98.42	98.47	98.4
1.2	18.56	150	98.42	98.47	98.4

A.2 CFD simulation void fraction average at downward, top bottom and middle.

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures
from sklearn.pipeline import make_pipeline
from sklearn.linear_model import Ridge
from sklearn.tree import DecisionTreeRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.svm import SVR
from sklearn.metrics import mean_absolute_error, mean_squared_error

# Load data from CSV
data = pd.read_csv("opt_1.csv") # Update the file path accordingly

# Separate input (X) and output (Y)
X = data.iloc[:, :3] # First three columns are input features
Y = data.iloc[:, 3:] # Next three columns are output values

# Split data into training and testing sets
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=42)

# Linear regression
lin_reg = LinearRegression()
lin_reg.fit(X_train, Y_train)
lin_reg_pred = lin_reg.predict(X_test)
lin_reg_mae = mean_absolute_error(Y_test, lin_reg_pred)
lin_reg_mse = mean_squared_error(Y_test, lin_reg_pred)
lin_reg_rmse = mean_squared_error(Y_test, lin_reg_pred, squared=False)

# Polynomial regression
degree = 2 # degree of polynomial
poly_reg = make_pipeline(PolynomialFeatures(degree), Ridge(alpha=0.1))
poly_reg.fit(X_train, Y_train)
poly_reg_pred = poly_reg.predict(X_test)
poly_reg_mae = mean_absolute_error(Y_test, poly_reg_pred)
poly_reg_mse = mean_squared_error(Y_test, poly_reg_pred)
poly_reg_rmse = mean_squared_error(Y_test, poly_reg_pred, squared=False)

# Decision tree regression
dt_reg = DecisionTreeRegressor(random_state=0)
dt_reg.fit(X_train, Y_train)
dt_reg_pred = dt_reg.predict(X_test)
dt_reg_mae = mean_absolute_error(Y_test, dt_reg_pred)
dt_reg_mse = mean_squared_error(Y_test, dt_reg_pred)
dt_reg_rmse = mean_squared_error(Y_test, dt_reg_pred, squared=False)

# Random forest regression
rf_reg = RandomForestRegressor(n_estimators=10, random_state=0)
rf_reg.fit(X_train, Y_train)
rf_reg_pred = rf_reg.predict(X_test)
rf_reg_mae = mean_absolute_error(Y_test, rf_reg_pred)
rf_reg_mse = mean_squared_error(Y_test, rf_reg_pred)
rf_reg_rmse = mean_squared_error(Y_test, rf_reg_pred, squared=False)

# Support vector regression
svr_regs = []
for i in range(Y_train.shape[1]):
    svr_reg = SVR(kernel='rbf', C=0.1)
    svr_reg.fit(X_train, Y_train.iloc[:, i])
    svr_regs.append(svr_reg)
```

```

# Print results
print("Linear Regression MAE:", lin_reg_mae)
print("Polynomial Regression MAE:", poly_reg_mae)
print("Decision Tree Regression MAE:", dt_reg_mae)
print("Random Forest Regression MAE:", rf_reg_mae)

# Load new input data
new_input_data = pd.DataFrame({
    "usl": [1],
    "usg": [2.1],
    "diameter": [75]
})

# Create X_input with the same columns as X_train
X_input = new_input_data.reindex(columns=X_train.columns, fill_value=0)

# Predict using trained models
lin_reg_pred = lin_reg.predict(X_input)
poly_reg_pred = poly_reg.predict(X_input)
dt_reg_pred = dt_reg.predict(X_input)
rf_reg_pred = rf_reg.predict(X_input)
svr_reg_preds = [svr_reg.predict(X_input) for svr_reg in svr_regs]

# Print predicted values
print("Linear Regression Predictions:", lin_reg_pred)
print("Polynomial Regression Predictions:", poly_reg_pred)
print("Decision Tree Regression Predictions:", dt_reg_pred)
print("Random Forest Regression Predictions:", rf_reg_pred)
print("Support Vector Regression Predictions:", svr_reg_preds)

```