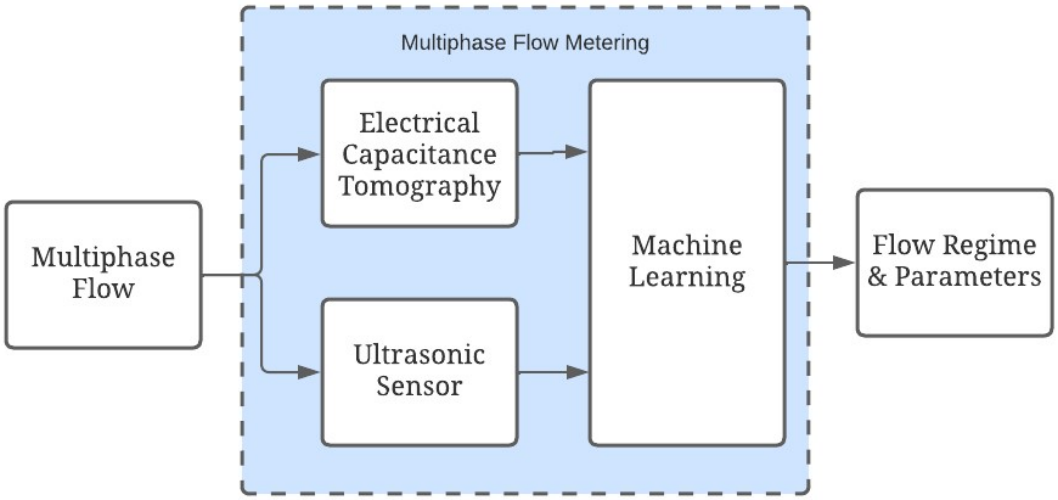
University of South-Eastern Norway

www.usn.no

FMH606 Master's Thesis 2023

Master of Science, Industrial IT & Automation

Multiphase Flow Metering with Multimodal Sensor Suite for Identification of Flow Regimes, and Estimation of Phase Fractions and Velocities



Noorain Syed Kazmi

Faculty of Technology, Natural sciences and Maritime Sciences

Campus Porsgrunn

University of South-Eastern Norway

www.usn.no

Course: FMH606 Master's Thesis, 2023

Title: Multiphase flow metering with multimodal sensor suite for identification of flow regimes, and estimation of phase fractions and velocities

Number of pages: 122

Keywords: Multiphase flow metering, Flow Regimes, Flow Parameters, Classification, Regression, Exploratory Data Analysis, Machine Learning, Cross-Correlation

Student: Noorain Syed Kazmi

Supervisors: Ru Yan, Saba Mylvaganam, Håkon Viumdal

External partner: **None**

Summary:

In this Master Thesis, data-driven multiphase flow metering models are developed to classify flow regimes and to estimate phase fractions and velocities of two-phase air/water flow after collecting data from the horizontal flow rig situated at University of South-Eastern Norway (USN).

Two types of multimodal sensors are used in this thesis namely Electrical Capacitance Tomography (ECT) and Ultrasonic Sensors (US). Experiments are performed on the flow rig at USN with ECT to collect capacitance data of the flow. Conventional measurements of pressure and flow rate are also collected during the experiments. Ultrasonic transit time data was already available through historical experiments.

Exploratory data analysis is performed on ECT and US data for feature engineering. ECT and US features are used to train and test flow regime classification models. Machine learning algorithms including Decision Tree, K-Nearest Neighbors, Artificial Neural Networks and Support Vector Machines are mainly employed to train flow classification and phase fractions estimation models. For flow velocity estimation, cross correlation technique is employed on 2-planes ECT data. Lastly, comparison of flow visualization with ECT and US data is performed. The images produced from ECT and US data are also compared for interface detection in multiphase air/water flow.

The flow regime classification models using ECT achieved an accuracy of more than 96%. Sensor fusion models of flow regime classification achieved accuracy of more than 97%. Flow velocity was accurately estimated using cross-correlation for slug regime. The phase fraction estimation Neural Network model achieved an R value of more than 0.95.

Preface

This thesis is submitted for the degree of Master of Science in Industrial IT & Automation at University of South-Eastern Norway.

I am deeply grateful to my thesis supervisors, Ru Yan, Saba Mylvaganam and Håkon Viumdal, for their firm support, guidance and supervision of this thesis work.

My sincere thanks go to Fredrik Hansen for his role in training me on conducting experiments for this thesis and for assisting me during the experiments.

I wish to extend my appreciation to Tonni F. Johansen for guiding me on the right path to understanding ultrasonic technology.

Porsgrunn, 15 May 2023

Noorain Syed Kazmi

Contents

Nomenclature

List of Acronyms

AE	Acoustic Emission
ANN	Artificial Neural Network
CSV	Comma-Separated Values
DAQ	Data Acquisition
DNN	Deep Neural Network
DPM	Differential Pressure Meters
DT	Decision Tree
ECT	Electrical Capacitance Tomography
EDA	Exploratory Data Analysis
EIT	Electrical Impedance Tomography
FFT	Fast Fourier Transform
FPS	Frames per second
FT	Flow Transmitter
GRT	Gamma Ray Tomography
KNN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
LR	Logistic Regression
MPF	Multiphase Flow
MPFM	Multiphase Flow Metering
MRI	Magnetic Resonance Imaging
NN	Neural Network
P&ID	Piping & Instrumentation Diagram
PCA	Principal Component Analysis
PDT	Pressure Differential Transmitter
PFM	Physical Flow Metering
PID	Proportional-Integral-Differential Controller
PSD	Power Spectral Density
PT	Pressure Transmitter
ReLU	Rectified Linear Unit
SVM	Support Vector Machine

TT	Temperature Transmitter
US	Ultrasonic
USN	University of South-Eastern Norway
VFM	Virtual Flow Metering
VR	Volume Ratio
VT	Velocity Transmitter

List of Symbols

A_L	Cross-sectional area of a pipe occupied by liquid phase
A_p	Cross-sectional area of a pipe
b_n	Neural Network bias for neuron *n*
C_m	Inter-electrode raw capacitance
C_L	Capacitance when the pipe is full with lower permittivity material
C_h	Capacitance when the pipe is full with higher permittivity material
C_n	Normalized capacitance
CXY	Normalized capacitance between ECT electrode X and Y
D_g	Distance from the top of B-scan image to the interface line
D_w	Distance from the bottom of B-scan image to the interface line
e	Total number of ECT measurement electrodes
$f(.)$	Neural Network activation function
N_n	Neural Network neuron *n*
VR	Volume Ratio
$w_{n,m}$	Neural Network weight between neuron *n* and input *m*
x_m	Neural Network input *m*
y_n	Neural Network output *n*

1 Introduction

Multiphase flow in a pipe is a complex phenomenon which is common in petrochemical industry. Fluids flowing at high speeds in a pipe produce a variety of flow patterns some of which can be detrimental to the overall system in which the pipe is installed. A lot of research has been performed on the real-time monitoring of this phenomenon due to its implication on the performance, reliability and availability of a flow system. Detection of flow regime instantaneously in a multiphase flow pipe is a challenge in petrochemical industry because some of the flow regimes such as slugs cause harmful vibration in the piping system and sometimes cause offshore production plant shutdowns and platform trips [1]. If flow regimes can be detected in real-time through indirect non-intrusive methods, the flow can be actively controlled to prevent damage to the piping system increasing reliability of offshore production facilities. Machine learning / artificial intelligence techniques can remove the requirement of a human supervisor for flow regime identification in flow pipelines.

In this thesis, detailed investigation and research are performed on the application of non-intrusive Electrical Capacitance Tomography (ECT) and Ultrasonic (US) sensors for data-driven multiphase flow metering with main focus on classification of two-phase air/water flow regimes and estimation of two-phase air/water flow parameters in a horizontal flow rig. The multiphase flow rig installed at the Process Hall of University of South-Eastern Norway (USN), Porsgrunn, Norway is employed to collect experimental flow data. Various algorithms are developed to classify flow regime and estimate flow-based parameters by using features extracted from the collected experimental data.

1.1 Background

Multiphase flow rig in USN, Campus Porsgrunn, has been used in a multitude of experiments involving Bachelor, Master, and PhD candidates and researchers (for both Norwegian and International), resulting in many student projects and research publications. Measurements using acoustic emission (AE) sensors, tomographic measurements using electrical impedance tomography (EIT) equipment, along with conventional measurements, such as acceleration, temperature, and pressure, have been used in these multiphase flow measurements. (Appendix – A)

This thesis deals with the usage of Ultrasonic and ECT sensors together with conventional measurements for measuring flow types, flow velocities and void fractions. This work is closely coupled to an ongoing project SAM (SAM: Self Adapting Model-based system for Process Autonomy - SINTEF), where Equinor and SINTEF are central partners. (Appendix – A)

1.2 Objective

The objective of this thesis is to develop data-driven multiphase flow metering models using multimodal sensor suite as depicted in Figure 1.1. Here multiphase flow metering refers to classification of flow regimes and estimation of flow velocities and phase fractions. The multimodal sensor suite includes ECT and US sensors. The suitability and performance of ECT and US sensors in combination with machine learning / artificial intelligence algorithms in the

field of multiphase flow metering is explored. A comparison between ECT and US images is presented to detect two-phase air/water interface from these images.

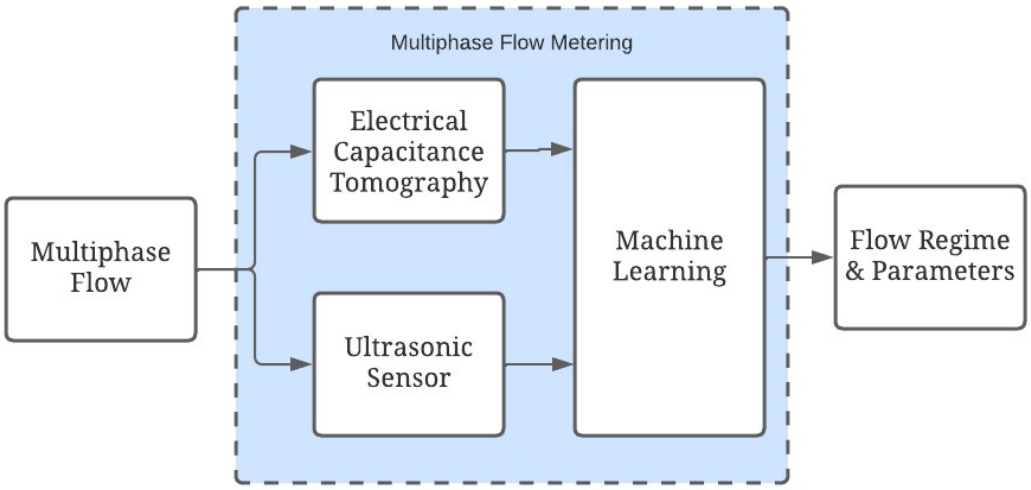


Figure 1.1: Main objective of this thesis showing main sensing technologies of ECT and US to classify flow regimes and estimate flow parameters

1.3 Methods

In this thesis, novel data-driven multiphase flow metering methods are developed by using ECT and US data to classify flow regimes and estimate flow parameters. Fresh experiments are conducted to collect new ECT data while US data was already available through past experiments.

Exploratory data analysis (EDA) is implemented on these data to extract features and labels for various classification and regression algorithms. In EDA, the methods implemented are spectrograms, correlation and box plots.

For classification of flow regimes, supervised classification algorithms of K-Nearest Neighbors, Decision Tree, Logistic Regression, Linear Discriminant and Support Vector Machine are mainly implemented in Classification Learner App in MATLAB. Both ECT and US data is tested for flow regime classification problem. A supervised artificial neural network is also developed in the Neural Net Pattern Recognition App of MATLAB for flow regime classification.

For estimation of flow velocity, cross-correlation algorithm is developed in MATLAB. Data from the two planes of ECT is cross-correlated for velocity estimation.

For estimation of phase fractions, supervised regression algorithm is implemented by creating an Artificial Neural Network in the Neural Net Fitting App of MATLAB.

MATLAB, Python, Microsoft Notepad and Microsoft Excel are the applications used in this thesis for data processing, modeling and analysis.

1.4 Scope of Thesis

The scope of this thesis covers literature survey, experimental work, exploratory data analysis and, classification and regression modelling.

A brief survey of multiphase flow, multiphase flow metering as well as usage of ECT and US modalities for multiphase flow metering is covered in this thesis. An overview of the flow rig at USN, its specification, the instruments installed on it and data acquisition methods from the rig are concisely described in this thesis.

Experiments on the flow rig to gather data from sensors by creating a new test matrix are in the scope of this thesis.

Exploratory data analysis for feature engineering and developing classification and regression models including machine learning models based on the collected data from the flow rig is included in this thesis. Classification of flow regimes and estimation of flow parameters are in the scope of this thesis. Lastly, a comparison of ECT and US data for flow visualization is briefly touched upon in this thesis.

The scope of work is shown as Figure 1.2.

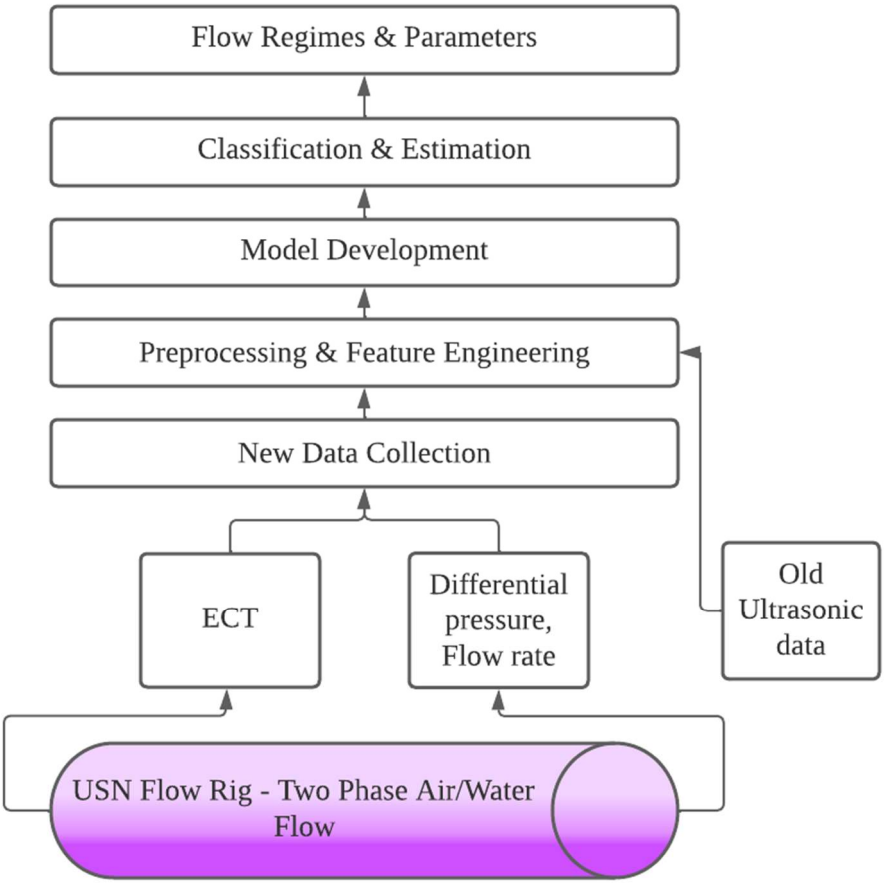


Figure 1.2: Flow chart of Data-driven virtual MPFM used in this thesis work

1.5 Report Structure

This report is divided into eight major sections.

Section 1 covers the introduction, background, objective, methods and scope of this thesis.

In Section 2, a brief survey of multiphase flow and multiphase flow metering is reported.

Section 3 includes literature survey of ECT and US sensor technologies. It also includes virtual flow metering techniques.

In Section 4, the multiphase flow rig in USN is described in detail along with its sensors, instrumentation and DAQ modules.

In Section 5, the experiments conducted on the flow rig at USN are discussed in detail including creation of test matrix, experimental setup and details about collected data.

In Section 6, the methods and results are presented. It includes exploratory data analysis, feature engineering, development of flow classification and regression models, and the results obtained from implementing these models. Moreover, a comparison of ECT and US sensors for flow visualization is also included in this section.

Sections 7 and 8 form discussion and conclusion sections respectively.

2 Brief Survey of Multiphase Flow and Multiphase Flow Metering

This section covers a brief description of the multiphase flow and multiphase flow metering methods.

2.1 Multiphase Flow

Fluid mechanics involves the study of multiphase flow, which refers to the movement of two or more phases of matter through a pipe [2]. Common multiphase flow combinations are provided in Figure 2.1. Gas/Solid is prevalent in pneumatic conveyors, dust collectors, fluidized beds, heterogeneous reactors and metallized propellant rockets. Gas/liquid flow can be seen in atomizers, scrubbers, dryers and combustors. Liquid/liquid droplet flow is observed in extraction, homogenizing and emulsifying. Liquid/solid is present in flotation and sedimentation [3].

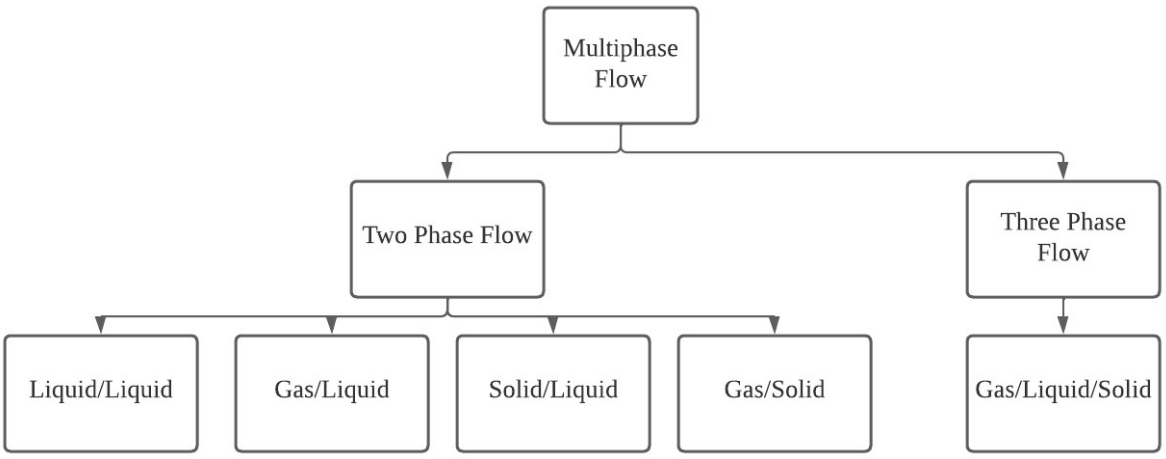


Figure 2.1: Some types of multiphase flows, including two-phase and three-phase flows [4]

2.1.1 Flow Regime / Pattern

Multiphase flow is a complex phenomenon. The geometric distribution of constituent phases within a multiphase flow is known as the flow regime or pattern [5]. There are various types of flow regimes observed in multiphase flow including slug, plug, stratified, annular, wavy, and bubble, among others [6]. The specific flow regime that develops is highly dependent on factors such as the orientation of the pipe, the direction of flow, the density and viscosity of the phases, and the mass flow rates of each phase [7]. The operating conditions of the system can also have a significant impact on the creation and stability of different flow regimes. Factors such as operating pressure, temperature, valves and bends have a direct effect on the flow regimes [8].

The petrochemical industry involves the flow of both gases and liquids within pipelines, leading to a variety of flow regimes. Two-phase system refers to the simultaneous flow of gas

and liquid in a pipeline. Identifying and classifying the different flow regimes in a two-phase flow pipeline is a significant challenge in the field of flow analysis [9]. Flow regimes can be classified subjectively through graphics or by employing the probability density function of pressure or void fractions signals from sensors [10], [11]. Accurate classification of flow regimes is critical for measuring parameters such as flow velocities and phase fractions [11].

Flow regime is difficult to predict accurately. To generate specific flow regimes in a pipe in a research laboratory, flow regime maps are used such as the one given in Figure 2.2 [12]. This flow regime map is based on an air-water system with superficial velocities given in ft/s.

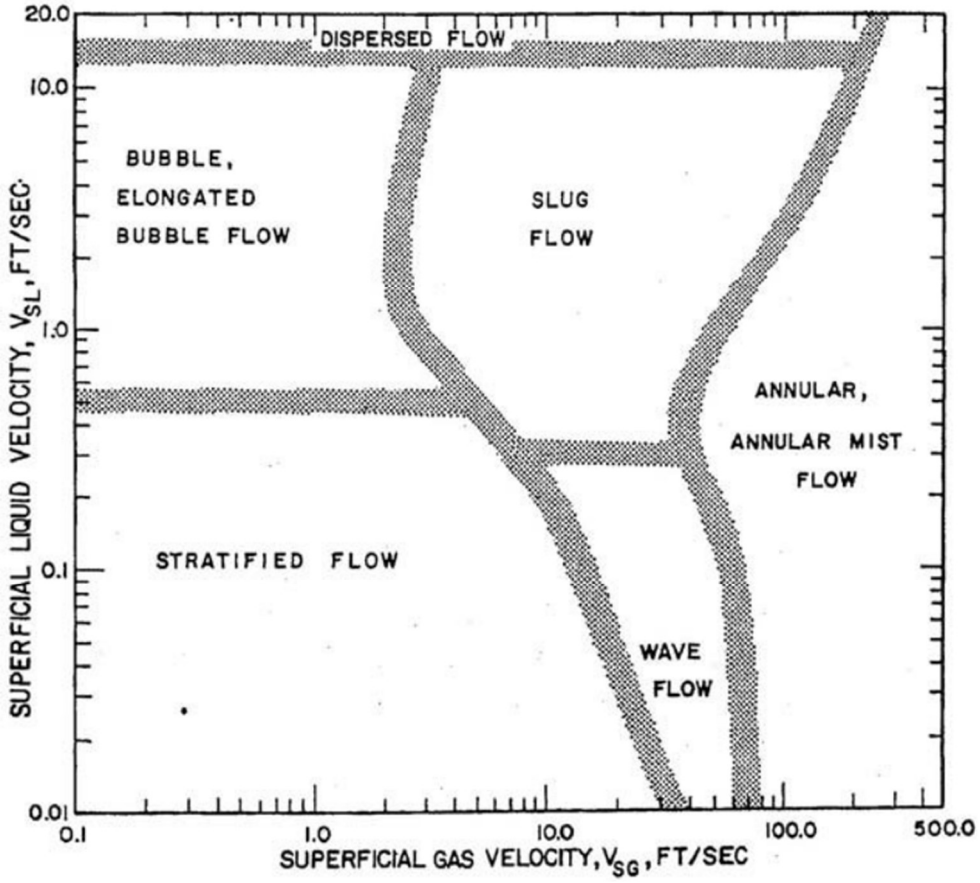


Figure 2.2: A two-phase flow regime map [12]

Figure 2.3 displays various flow regimes that are commonly observed in horizontal gas/liquid multiphase flows [6]. These flow regimes are described below:

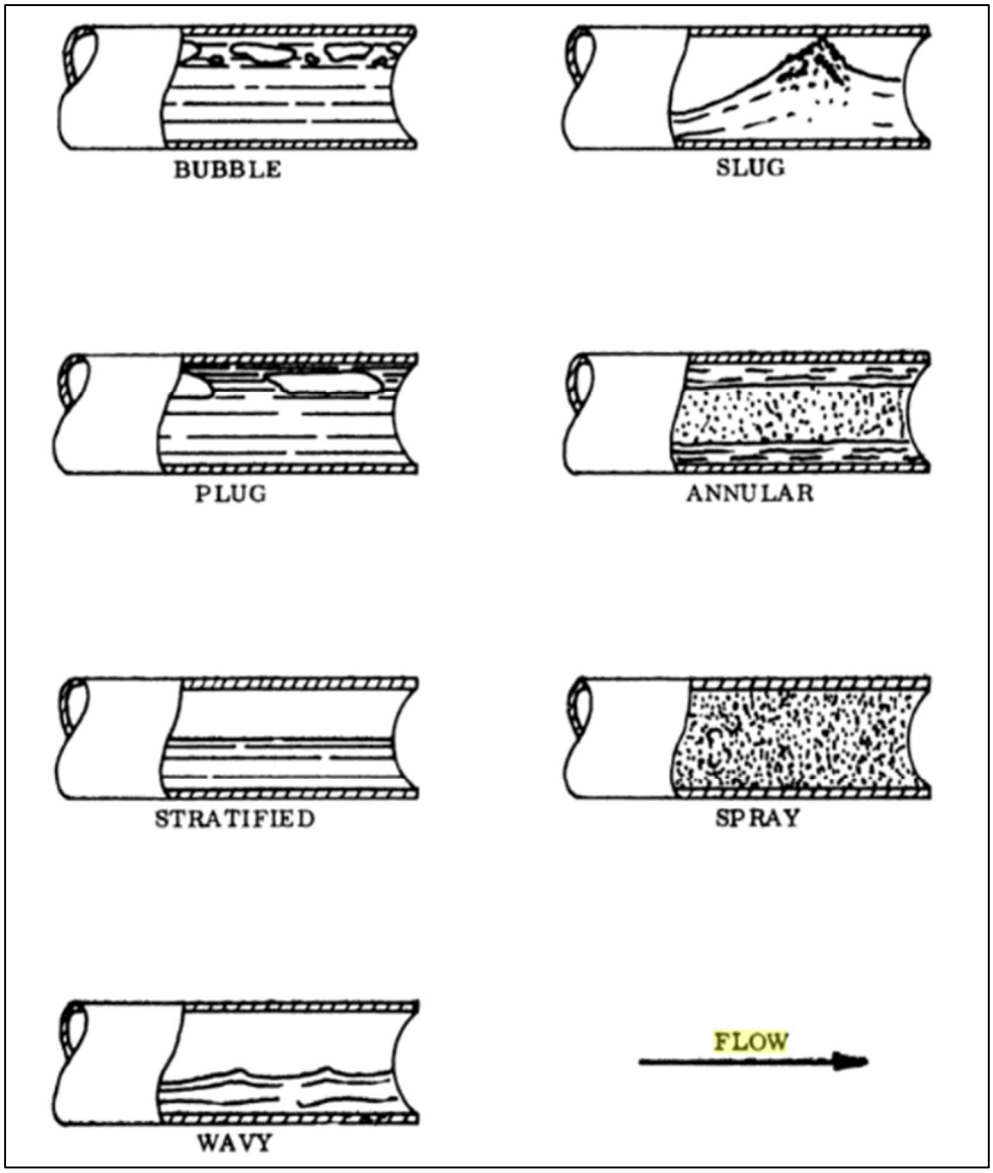


Figure 2.3: Gas/liquid Flow Regimes in horizontal pipe [6]

2.1.1.1 Stratified

When gas and liquid flow rates are low, stratified flow is observed. It is applicable in horizontal flow direction. There is no mixing of the two phases and the liquid phase remains as a film at the lower portion of the pipe [13].

2.1.1.2 Wavy

At higher gas flow rates, the stratified flow converts to wavy flow in which ripples or waves are observed on the top of the liquid layer. It appears like waves in a sea. [14]

2.1.1.3 Annular

At higher gas velocity, wavy flow converts to annular flow in which liquid flows at the periphery of the pipe while gas flows at the center of the pipe. [15]

2.1.1.4 Plug / Elongated Bubble

Plug is a kind of flow pattern in which large bubbles of gas float on the top surface of the liquid phase spanning a large area in the pipe. The gas phase is dispersed in the liquid phase. [16]

2.1.1.5 Slug

Slug flow is intermittent flow in which slugs of liquid with dispersed bubbles flow along with large gas pockets. The flow is always unsteady. The bubble is in the shape of a bullet. This flow alternates between high liquid fraction and high gas fraction. [17]

2.1.1.6 Bubble

Bubble flow is similar to plug flow and is characterized by small gas bubbles dispersed in the liquid phase. [18]

2.1.1.7 Spray / Annular Mist

In spray flow, droplets of gas are dispersed in the liquid phase. [19]

2.1.2 Phase Velocity

Phase velocity, as defined in [20], refers to the velocity of a specific phase (such as oil, water, or gas) within a pipe. It is measured in m/s.

2.1.3 Phase Fraction

For multiphase flow in a pipe, phase fraction is the fraction of one phase with respect to the whole fluid in a cross-section of pipe [5]. Mathematically, it is calculated as the ratio of the cross-sectional area of the phase of interest to the entire cross-sectional area of the pipe, as shown in Equation (2.1) below. Phase fraction is typically expressed as a percentage. In Equation (2.1), A_L is the cross-sectional area occupied by liquid phase and A_p is the cross-sectional area of a pipe.

$$\text{Phase Fraction of liquid} = \frac{A_L}{A_p} \tag{2.1}$$

Phase fraction can be defined in terms of Volume Ratio in gas/liquid flow which is the percentage of oil or water in the cross-section of a pipe [21].

2.2 Multiphase Flow Metering

Multiphase flow (MPF) is a highly complex phenomenon. Measurement of multiphase flow patterns and parameters is challenging, requiring specialized meters and sensors. Knowledge of multiphase flow dynamics is required to build physics-based multiphase flow meters [22]. Some measuring techniques used in multiphase flow metering (MPFM) are process tomography and imaging, ultrasound, X-ray, Magnetic Resonance Imaging (MRI), differential pressure and turbine meters [22]. The applied technique is highly dependent on the application [22].

Measuring multiphase flow in the oil and gas industry can be challenging, particularly for subsurface and offshore installations where pipelines may contain oil, water, and gas [8]. Subsurface meters and instrumentation are often expensive [8]. However, multiphase flow metering (MPFM) can help control slug patterns in the pipeline by providing more accurate measurements [8]. MPF measurements are better than pressure measurements for flow regime control [8].

MPFM is the measurement of constituent phase parameters of a MPF pipeline without separating the phases [8].

2.2.1 MPFM Technologies

The MPFM technologies can be grouped under the following major headings [8]:

- Tomography
- Gamma Densitometry
- Differential Pressure Meters
- Wet Gas
- Ultrasonic Sensing
- Coriolis

2.2.1.1 Tomography

Tomography is a non-intrusive image processing technology for measuring MPF. It is non-radioactive. The tomography sensors are located around the periphery of pipe without touching the flow [8].

2.2.1.2 Gamma Densitometry

Gamma densitometry employs a radioactive source to measure MPF. The concept of its operation is that the radiated gamma rays from the radioactive source are attenuated based on the flow pattern to detect phase fractions [8].

2.2.1.3 Differential Pressure Meters (DPM)

DPM is also known as virtual flow meter (VFM). Difference in pressure at two points in the pipeline can be measured by DPM to calculate the flow rate based on Bernoulli's principle [8].

2.2.1.4 Wet Gas

Wet gas method is applicable when water and oil fractions are less than 5% in the pipe and gas fraction is more than 95%. Wet gas flow then can be obtained by Lockhart-Martinelli parameter and, gas and liquid densimetric Froude number [8].

2.2.1.5 Ultrasonic Sensing

Ultrasonic methods of MPFM are non-intrusive and radiation-free. They do not affect the flow [23]. They employ time-of-flight, cross-correlation and Doppler techniques to measure flow parameters [23].

2.2.1.6 Coriolis Flow Meters

Coriolis flow meters are widely used in MPFM. They can measure mass flow rates accurately. They work on the principle of the Coriolis effect. [24]

2.2.2 Methods of Multiphase Flow Metering

2.2.2.1 Physical Flow Metering (PFM)

In physical flow metering, flow parameters are directly measured by MPFM mounted on the pipeline. The examples of these meters are Coriolis mass flow meters and Gamma densitometry. [25]

2.2.2.2 Virtual Flow Metering (VFM)

VFM use available field data to generate estimations of flow parameters [25]. There are two types of VFM, namely first principles VFM and data-driven VFM.

In first principles VFM, mechanistic model of process is used to estimate flow parameters from the available pressure and temperature field data [25].

In data-driven MPFM data is collected from the flow and employed in mathematical or machine learning models to generate estimations about the flow [25]. This type of MPFM requires no detailed knowledge of the physical model of the process [25].

2.2.2.3 Comparison of VFM and PFM

Table 2.1 describes comparison between VFM and PFM. Both methods are suitable for real-time monitoring of flow. The cost of physical flow metering is higher than virtual flow metering. PFM requires periodic calibration as compared to VFM. Experienced personnel are required to operate PFM in contrast with VFM. Exploratory data analysis and feature engineering is an essential requirement in VFM but not in PFM. PFM requires high maintenance cost as compared to VFM. The computational power requirement for VFM is very high with respect to PFM since VFM are data-driven.

Table 2.1: Comparison of PFM and VFM [25]

	PFM	**VFM**
Real-time monitoring	Yes	Yes
Cost	High	Low
Periodic Calibration	Yes	No
Operational Experience	High	Limited
Feature Engineering	No	Yes
Maintenance	High	Low

Computational Power	Low	High

3 Brief Survey on Sensing and Data-Driven Metering Techniques

As discussed in Section 2, there are a number of MPFM technologies to measure MPF parameters. In this section, a literature survey of contemporary publications in MPFM is presented with emphasis on ECT and US modalities.

In [11], a technique to classify 2-phase gas/liquid flow regime in an S-shaped pipeline using Doppler ultrasonic sensors is presented. Deep Neural Networks (DNNs), AdaBoost, Bagging, Extra trees and Decision tree machine learning methods were employed in this paper along with Fast Fourier Transform (FFT) of the ultrasonic signals. Twin-window features were input to the DNN while flow regimes were the output of the classification DNN. Four hidden layers were used in the DNN with rectified linear unit (ReLU) as an activation function in all of them.

The work presented in [26] focused on measuring the flow velocity and phase concentrations of horizontal two-phase flow using a dual-plane ECT with gamma-ray tomography (GRT). It achieved an estimation accuracy of ±10% for the volumetric flow rates of the phases.

In [27], flow regime identification using ECT time-series statistical data analysis is discussed.

In [28], two-phase flow regime identification using ECT is presented. In this paper, Support Vector Machine (SVM) is used to classify flow regimes with good accuracy. Principal Component Analysis (PCA) was used to reduce the number of inputs to the SVM.

[29] discusses velocity measurement of two-phase flow through ECT using cross-correlation techniques.

In [30], two-phase flow regime identification is presented using the ultrasonic power spectral density (PSD) technique.

3.1 Brief Survey on Electrical Capacitance Tomography (ECT)

ECT is a flow sensing technique in which the spatial distribution of the mixture of dielectric materials in a flow inside a pipe is measured from electrical capacitance measurement between electrodes placed on the periphery of the pipe [21], [31]. The capacitances can also be converted into images for visual depiction of the flow. It is a non-invasive technique and non-radioactive. Figure 3.1 provides the cross-sectional illustration of an ECT system mounted on a pipeline [31]. It has 12 electrodes to measure capacitances. With 12 electrodes, 66 pair of capacitances can be measured.

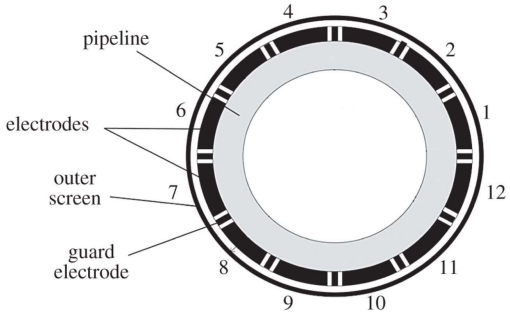


Figure 3.1: Cross-sectional view of a pipeline mounted with a 12-electrode ECT system [31]

ECT is generally used in two-phase liquid/gas flow to measure permittivity distribution inside a pipe [21]. ECT measurements at two cross-sections of a pipe can measure the velocity profile of flow [21]. Volume ratio or phase fraction can be obtained by ECT. ECT works best when the fluid has low electric conductivity and a variable permittivity [21]. The working principle of ECT is shown in Figure 3.2. An alternating voltage is applied between the source and the ground. The currents at the detector electrodes are measured with respect to the source electrode which are directly proportional to the capacitance between the source and detectors [21]. The source is changed to the next electrode and the currents are measured at all the other electrodes during one frame of measurement [21]. When only one electrode is excited then for N electrodes, there are M = N(N-1)/2 unique capacitances making up one frame of measurement [21]. The output of ECT is a capacitance file with M capacitances data per frame of measurement [21].

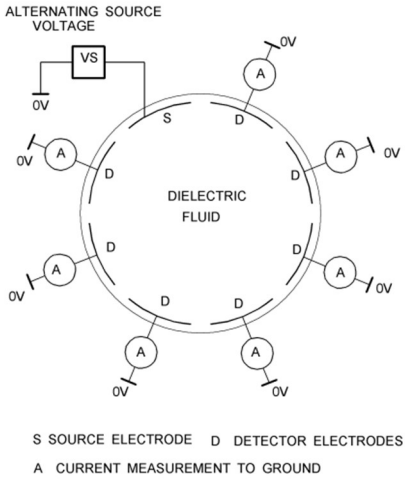


Figure 3.2: Principle of ECT with 8 electrodes [21]

The measured capacitances can be normalized by using Equation (3.1) as given in [21]

$$C_n = \frac{C_m - C_L}{C_\mathrm{h} - C_L} \tag{3.1}$$

In Equation (3.1), C_m is the inter-electrode raw capacitance. C_L is the capacitance when the pipe is full with lower permittivity material such as air. C_h is the capacitance when the pipe is full with higher permittivity material such as water. C_n is the normalized capacitance. C_n is

dimensionless and normalized, making it suitable as input for mathematical operations and algorithms. [21]

3.2 Brief Survey on Ultrasonic Sensing (US)

The range of audio spectrum is from 20Hz to 20kHz. Humans can perceive audio frequencies in this range. The audio frequency band above 20kHz is known as ultrasonic band. Ultrasonic (US) band continues up to 1GHz. Most of the industrial applications of ultrasonic ranges from 1MHz to 100MHz. [32]

The speed of sound varies according to the medium through which it travels. In dry air the speed of sound is 300m/s whereas in water, the speed of sound is about 1500m/s. [32]

There are advantages to use US technology for industrial applications. The speed of US waves is lower than electromagnetic waves. They can be easily visualized on a screen due to their low velocity. They can travel through solid and opaque matter. Light cannot travel through opaque matter [32]. They are least affected by color type of target surfaces and materials. They can detect small anomalies over long distances. [33]

There are two types of ultrasonic experimental techniques that can be used in flow pipelines to detect flow characteristics: Pulse-Echo and Pitch-Catch techniques.

3.2.1 Pulse Echo Technique

The basic principle of pulse-echo technique is shown in Figure 3.3. The transmitter converts electrical signal to an ultrasonic wave. The wave travels through the material and is reflected by an anomaly or the back wall. The reflected wave is converted to an electrical signal by the receiver. Generally, the transmitter and receiver are placed in a common housing. [34]

Pulse-echo technique can be used for detection of flaws and defects in metals. [35]

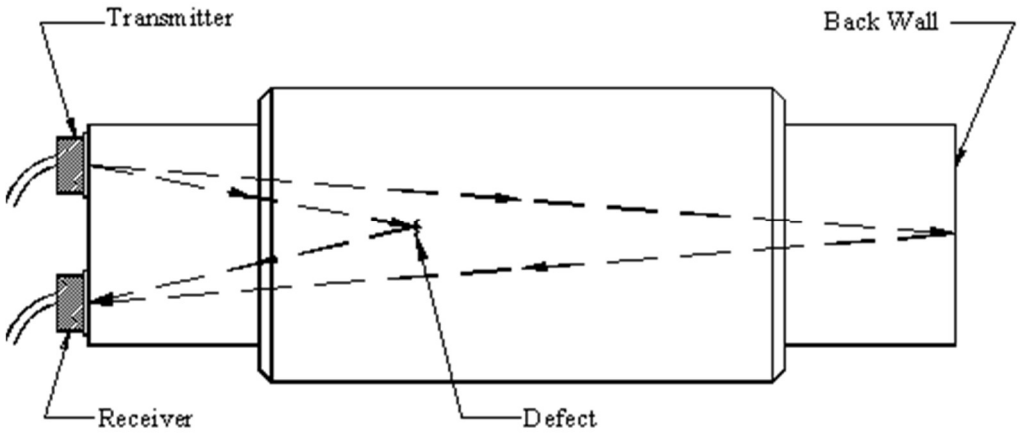


Figure 3.3: Basic Principle of Pulse-Echo Technique [34]

3.2.2 Pitch Catch Technique

The basic principle of pitch-catch technique is shown in Figure 3.4. The transmitter converts electrical signal to an ultrasonic wave. The wave travels through the material and is reflected

by an anomaly or the boundary. The reflected wave is converted to an electrical signal by the receiver. The transmitter and receiver are placed at a distance apart. [34]

Pitch-Catch technique can be used to determine the location of defect in metals along with the depth of the defect. [35]

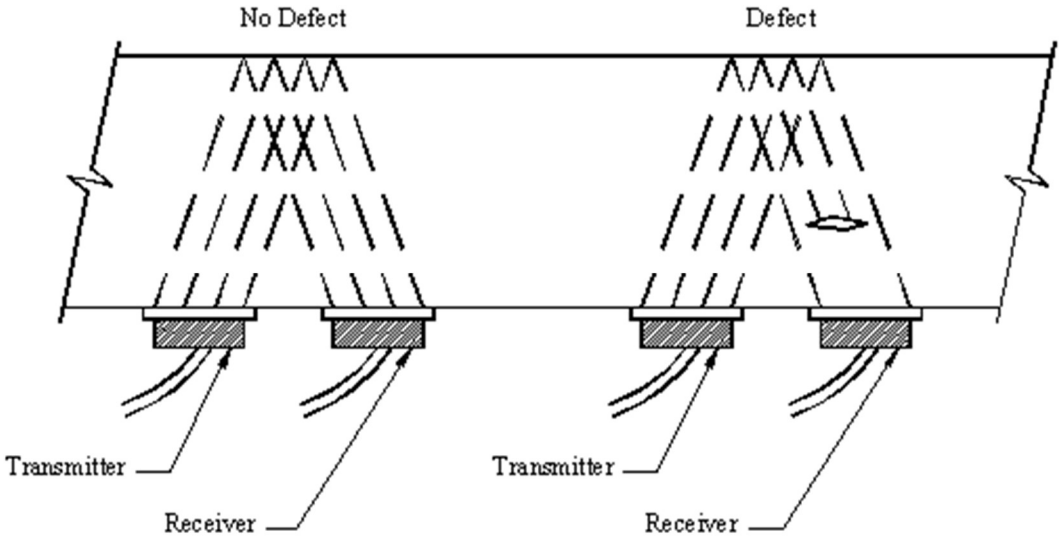


Figure 3.4: Basic Principle of Pitch-Catch Technique [34]

3.2.3 Generated Output from Ultrasonic Techniques

Figure 3.5 shows a typical ultrasonic A-scan (Amplitude Scan) from pulse-echo and pitch-catch techniques. A-scan is the amplitude trace of ultrasound echoes with respect to time in one-dimension [32]. The first peak is due to the material boundary at the transmitter end. The second peak represents the reflected echo at the anomaly inside the test material. The third peak is the reflected echo from the back wall of the test material. [34]

Based on this scan, the location and presence of anomaly can be detected. A large number of similar A-scans can be stacked together to generate a two-dimensional image of the test material which is known as B-scan.

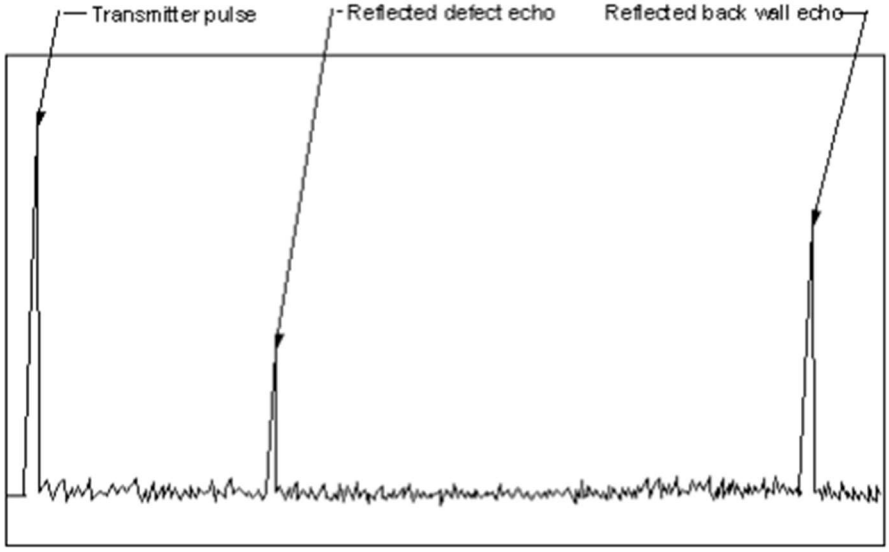


Figure 3.5: A typical ultrasonic A-scan [34]

3.3 Data-Driven Virtual Flow Metering Techniques

This section provides a brief overview of various methods used for VFM, with a focus on the specific techniques applied in this thesis for flow regime identification and flow parameters estimation.

3.3.1 Machine Learning Algorithms

In this thesis, only supervised machine learning algorithms are used for classification and regression problems. In supervised machine learning, the data available for analysis always includes features that are associated with a corresponding label or output. [36]

3.3.1.1 Decision Tree

The decision tree algorithm is used for classification and regression tasks. It is a non-parametric algorithm. In this algorithm, classification and regression decisions are made based on a sequence of decisions resembling a tree-like structure. The input dataset is divided into small subsets that act as tree nodes. [37]

Consider an ECT dataset comprising of four samples, each consisting of three features and a label as shown in Table 3.1. The dataset has 4 samples. The three features, namely C12, C13, and C14, represent the normalized capacitance values between electrodes 1 and 2, 1 and 3, and 1 and 4, respectively. The label, referred to as the Regime, is the flow pattern observed corresponding to the measured capacitances.

Table 3.1: An example dataset of ECT

C12	**C13**	**C14**	**Regime**
0.1	0.2	0.3	Wavy

0.2	0.4	0.6	Annular
0.15	0.21	0.33	Wavy
0.23	0.43	0.66	Annular

A classifier decision tree model developed using the dataset given in Table 3.1 is illustrated in Figure 3.6. Decision tree is constructed by starting with C12 as the root node. In this node, a decision is taken based on the value of C12. Depending on the decision on C12, the tree is traversed to the next node where the value of C13 is compared to a decision point. After C13, C14 is finally checked to find a final decision on regime classification.

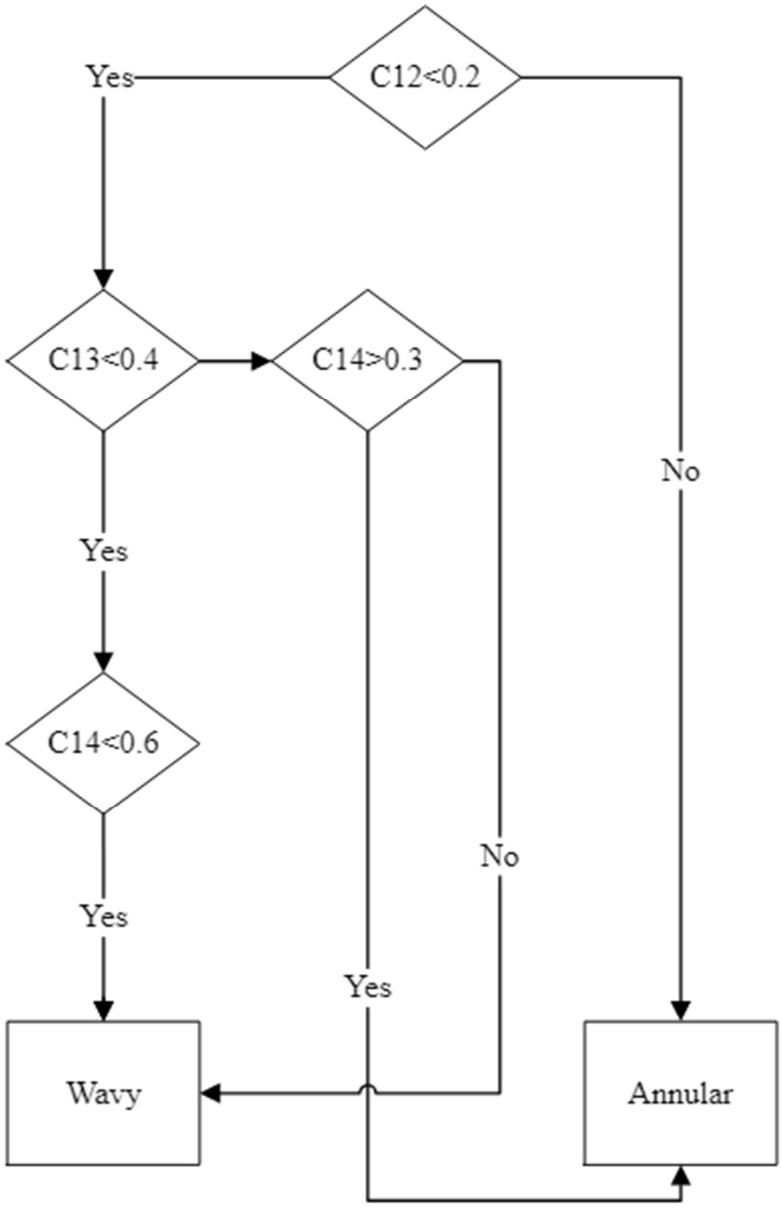


Figure 3.6: An example of decision tree algorithm

3.3.1.2 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is a non-parametric and supervised learning algorithm. It is a grouping algorithm. It divides the data into groups and then new data is classified into one of the groups based on its proximity to the individual groups. It is used for classification and regression tasks. [38]

Decision boundaries are formed by using Voronoi diagrams. To calculate proximity between points, distance measurements such as Euclidean distance can be used. [38]

In Figure 3.7, assume input-1 as C12, the normalized capacitance between electrode 1 and 2 of ECT, input-2 as C13, the normalized capacitance between electrode 1 and 3 of ECT, output-1 as Wavy, output-2 as Annular and output-3 as Stratified. There are clearly demarcated areas of flow regimes in the diagram based on the grouping of data points. The Euclidean distance of a new data point will be calculated from each of the existing data points. The new data point will be classified based on its closeness from the other already classified points.

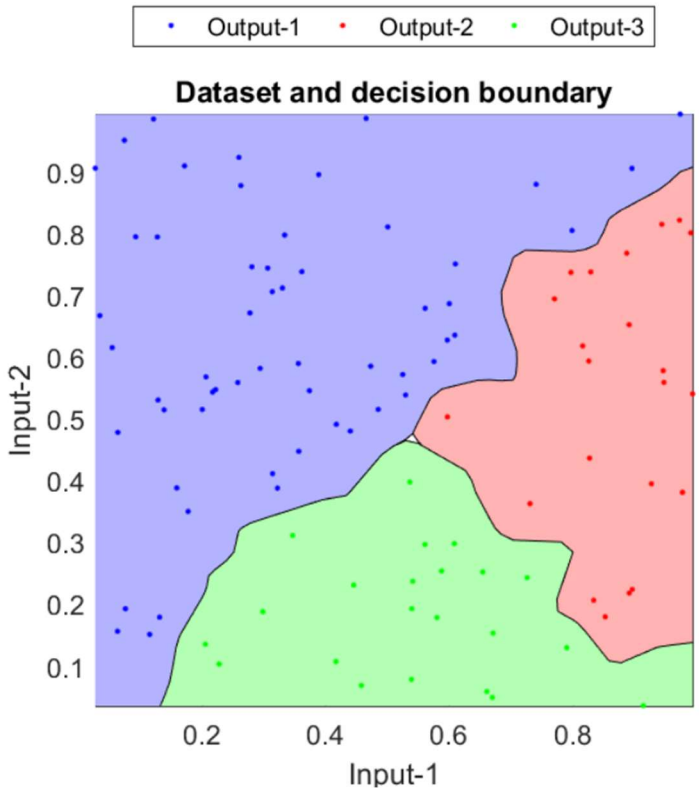


Figure 3.7: An example of Voronoi diagram taken from [39]

3.3.1.3 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a supervised classification algorithm which reduces dimensions of the datasets. It is used to find linear combination of features separating classes. It creates a lower-dimensional discriminant plane maximizing the ratio of the parameters 'between-class variances' to 'within-class variances'. [40]

3.3.1.4 Logistic Regression

Logistic Regression (LR) is a supervised classification model mainly used to classify two classes. It estimates the probability on the classes based on training dataset, a value between 0 and 1. Based on the probability, the data can be classified into two classes. For a binary classification problem, if the probability is more than 50% then the data is classified in class 2 and if the probability is less than 50% then the data is classified in class 1. [41]

3.3.1.5 Support Vector Machines

Support Vector Machines (SVM) are mainly supervised machine learning algorithms. They can be used for classification as well as regression tasks. In a classification problem, a new data point is to be classified in a particular class. SVM creates hyperplanes in a higher dimension. The hyperplane with the highest distance from the nearest training datasets on each side is selected as the classifying hyperplane as shown in Figure 3.8. [42]

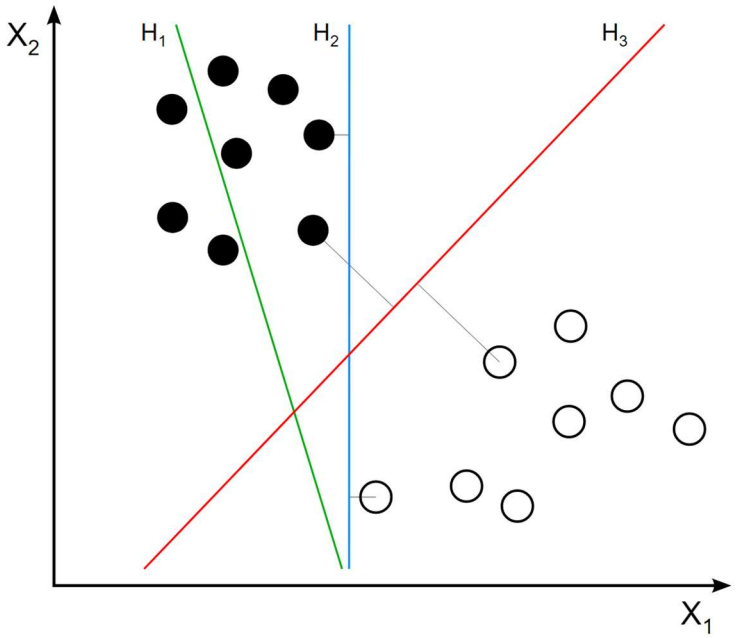


Figure 3.8: A 2-dimensional plot showing SVM selecting a hyperplane in a classification problem [42]. H_3 plane will be selected in this case because of its distance from datasets is the largest [42]

3.3.1.6 Artificial Neural Network

Artificial Neural Networks (ANN) are machine learning algorithms developed on the concept of neural networks in animal brains. These algorithms can be applied for supervised as well as unsupervised classification and regression tasks. Figure 3.9 shows an example of feedforward neural network with m inputs (x_1 to x_m) and n outputs (y_1 to y_n). There is one hidden layer of n neurons, N_1 to N_n. $w_{n,m}$ is the weight between neuron n and input m. b_n is the bias for neuron n. $f(.)$ is the activation function. The output Y is given by Equation (3.2) where W, x and b are matrices. [43]

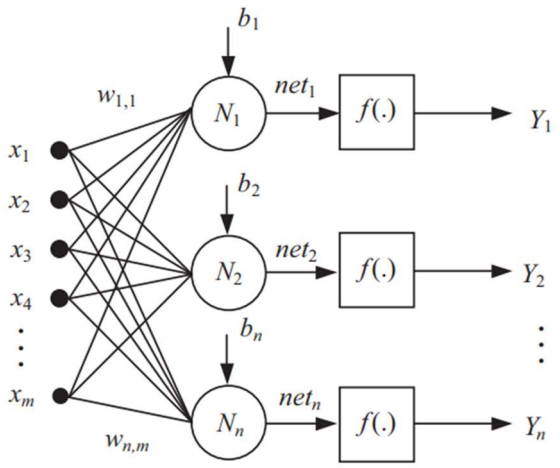


Figure 3.9: A feedforward neural network with one hidden layer [43]

$$Y = f(W.x+b) \tag{3.2}$$

3.3.1.7 Ensemble Learning

Ensemble methods combine a lot of learners and provide a solution to the same problem as shown in Figure 3.10. The base learners can be any of the machine learning algorithms such as Decision Tree, KNN or ANN. When the learning power of algorithms is weak, ensemble learning is used to boost the learning algorithms by combining learners. There are two types of Ensemble Learning methods, boosting and bagging. Boosting involves training the learners sequentially and then combining the result. Bagging involves training the learners parallelly and then using a voting system for classification problem and averaging for regression problem. [44]

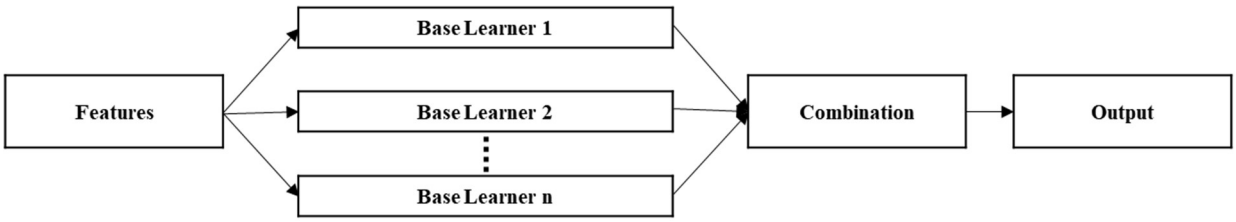


Figure 3.10: Concept of Ensemble Learning Algorithm [44]

3.3.2 Cross-Correlation Technique

Cross-correlation is a technique in which correlation is computed between two signals with sampling times as lags. It computes the similarity between two signals when one signal moves with respect to the other. [45]

4 Overview of USN Multiphase Rig

This section provides an overview of USN multiphase flow rig present in the Process Hall of Campus Porsgrunn with its accompanying sensors and instrumentation.

4.1 USN Multiphase Rig

The USN multiphase flow rig is located in the Process Hall at USN, Porsgrunn campus. The rig has the capability to generate different types of flows by passing water, air and mineral oil through a horizontal pipe. The pipe can be tilted by ±10º with respect to the horizontal surface [27]. Mass flow rates in kg/min of air, water and oil are monitored and controlled by Coriolis flow meters and control valves [27]. The operational limit of mass flow rate for air is 5 kg/min whereas for liquid is 150 kg/min [27]. The rig can simulate different flow regimes in the pipe by injecting combinations of air, water, and oil mass flow rates. These regimes, including stratified, slug, plug, annular, and wavy, can be visually observed through a transparent Plexiglass section, as shown in Figure 4.1. [27].

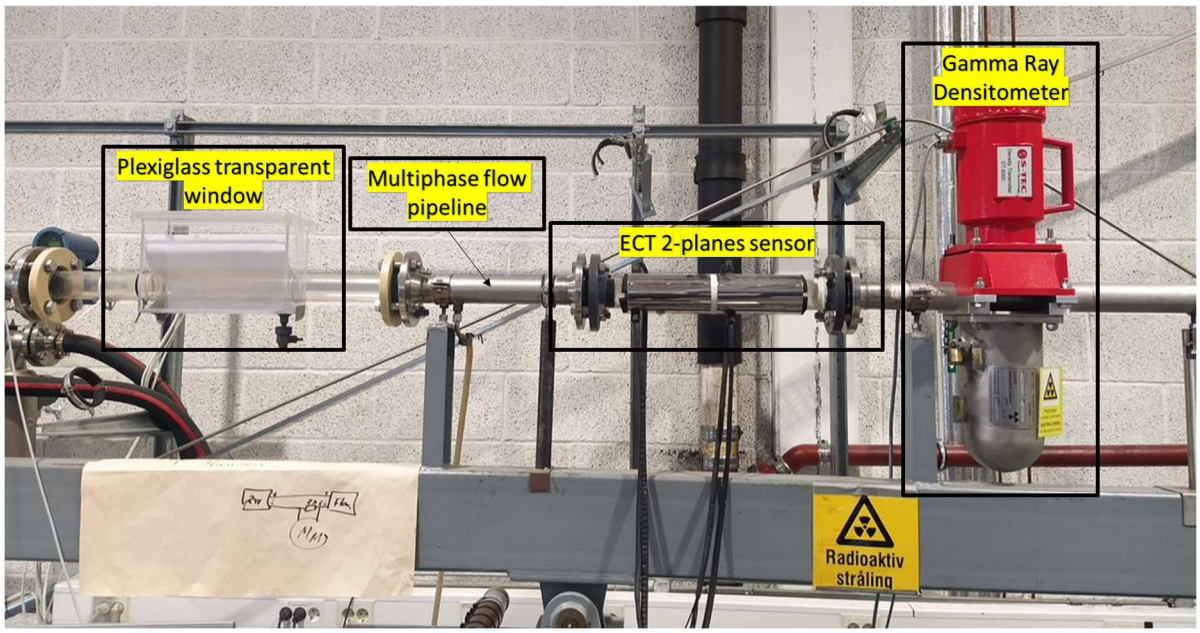


Figure 4.1: USN multiphase setup showing plexiglass window, ECT and Gamma ray densitometer

There is a TOMOFLOW TFLR5000 ECT system manufactured by Process Tomography Limited mounted on the rig as shown in Figure 4.1. There is also a gamma ray densitometer DT9300 manufactured by S-TEC [46] using Cs-137 isotope as seen in Figure 4.1.

Figure 4.2 depicts the operating control panel of the rig which is developed in LabVIEW through which the flow rig is operated. It also shows various transmitters installed on the rig to control and monitor the flow conditions during its operation. PID control is used to control the flow rates of gas and liquids. The mass flow rates of oil, water and air are controlled through this panel. TT is temperature transmitter in ºC. VT is velocity (volume) transmitter in l/min. FC is flow controller. DT is density transmitter in kg/m^3. FT is flow transmitter in kg/min. PDT121 is short distance differential pressure transmitter in mbar. PDT120 is long distance differential pressure transmitter in mbar. PT131 is air inlet pressure transmitter in mbar. P101

is a large water pump used for flow rates more than 12 kg/min. P102 is a small water pump used for flow rates up to 12 kg/min. When the big pump is running liquid flow rate is measured by FT114A whereas when using the small pump, the flow rate is measured by FT114B. FT131 measures the air flow rate. P100 is an oil pump. R100 is a separator tank to separate oil and water at the pipe outlet. T100 is oil tank. T101 is water tank. Six Coriolis meters are shown in the P&ID that measure mass flow rate, temperature, density, and velocity (volume). A simplified P&ID is shown in Figure 4.3. Important parameters of the multiphase flow rig are given in Table 4.1.

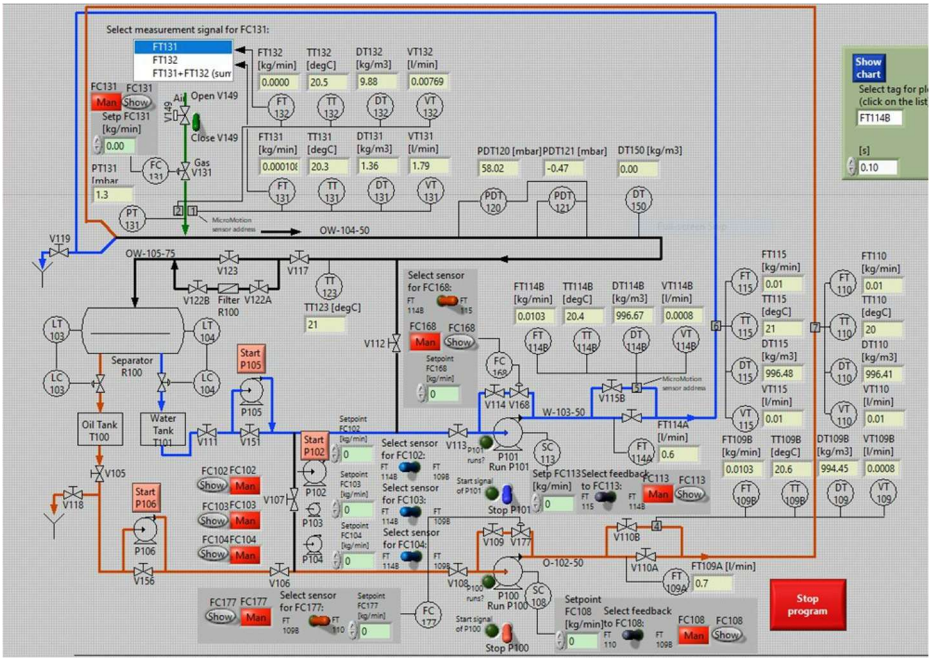


Figure 4.2: P&ID of USN flow rig in LabVIEW (Developed by USN)

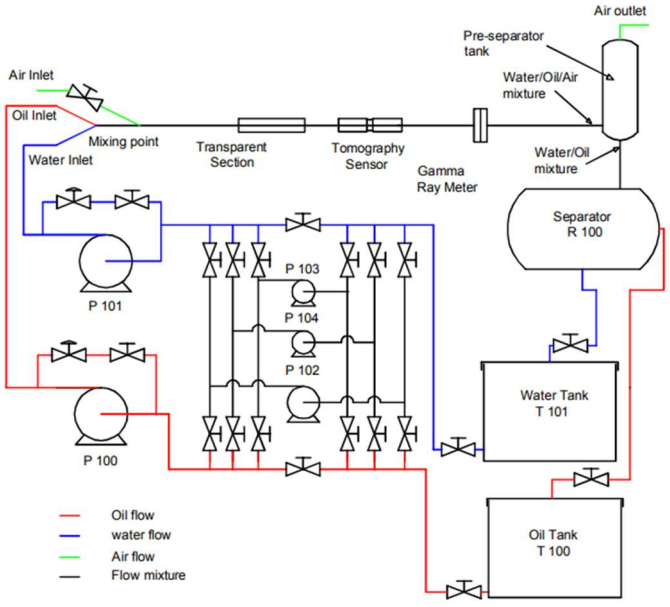


Figure 4.3: P&ID of USN flow rig [47]

Table 4.1: Rig parameters for multiphase flow [47]

Parameters	**Value**
Inside pipe diameter	56 mm
Outside pipe diameter	60 mm
Length of pipe	15 m
Water density	996 kg/m^3 at 27ºC and 1bar
Air density	1 kg/ m^3
Oil density	790 kg/m^3
Water viscosity	0.00102 Pa-s (20ºC)
Oil viscosity	0.00164 Pa-s (25ºC)

4.2 Sensors, Instrumentation and DAQ Modules

This section describes major instruments and sensors installed with the USN multiphase rig.

4.2.1 ECT System: TOMOFLOW TFLR5000

As seen from Figure 4.1, there is an ECT system present on the USN multiphase rig pipe. It is developed by Process Tomography Limited [21]. Its model is TOMOFLOW TFLR5000 (Figure 4.4) multiphase flow measurement system and is suitable for research laboratories [21]. It can measure flow parameters of an uneven two-phase flow when the constituents have dielectric properties [21]. In the USN rig, the permittivity of air is 1, oil is 2.7 and of water is 80 [27]. Therefore, ECT is suitable to detect permittivity distribution in the USN flow pipe with a mixture of two-phase flow. Currently, the application of this device is research in laboratories for two-phase flow measurements and analysis. [21]

Figure 4.4: TFLR5000 ECT system control unit showing the communication and power ports

4.2.2 Pressure Meters Including Differential Pressure Meters

PDT120 and PDT121 are the differential pressure meters (DPM) mounted on the rig whereas PT131 is the air inlet pressure transmitter as seen from Figure 4.2. Their details are provided in Table 4.2. PDT120 and PDT121 are shown in Figure 4.5. From Figure 4.6, it can be seen that PDT120 is measuring differential pressure between two points 10.22m apart in the pipe and PDT121 is measuring differential pressure between two points 5.38m apart in the pipe.

Table 4.2: Differential pressure meters installed on the rig [47]

Transducer	**Model**	**Range(mbar)**
PDT120	Rosemount 3051CD	0-100
PDT121	Rosemount 3051CD	0-50
PT131	Vika S-10	0-100

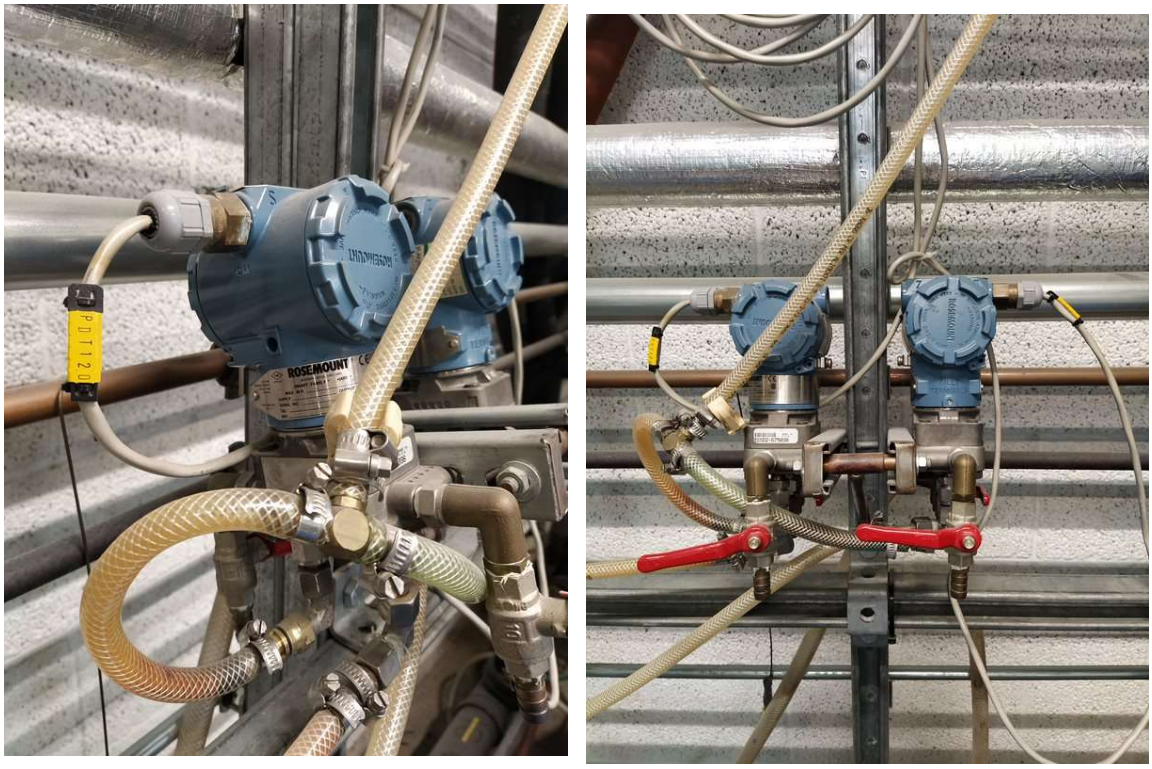


Figure 4.5: Differential Pressure Meters at USN Rig, PDT120 and PDT121

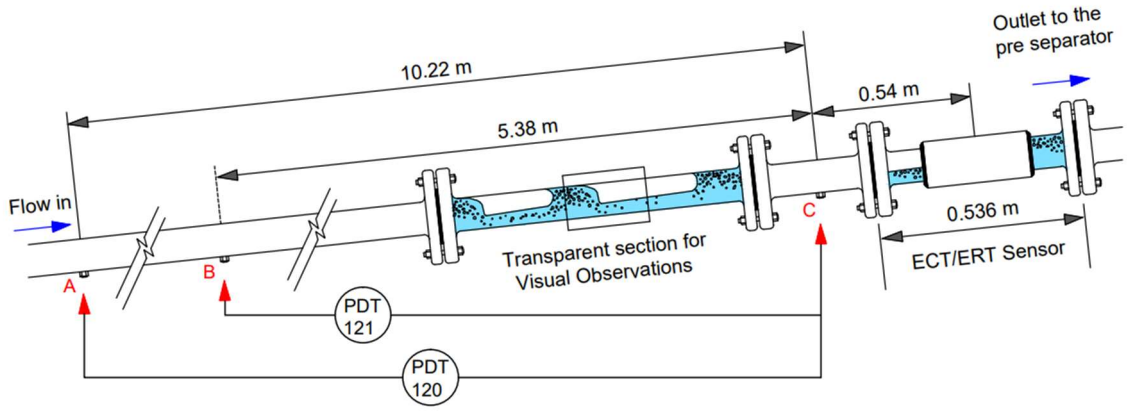


Figure 4.6: DPM measurement regions: PDT120-10.22m & PDT121-5.38m [47]

4.2.3 Ultrasonic Device: US-Key from Lecoeur Electronique

Figure 4.7 displays ultrasonic transmitter and receiver Pitch-Catch setup mounted on the USN flow pipeline. A device named US-Key manufactured by Lecoeur Electronique is used to transmit and receive US waves. It is shown in Figure 4.8. This device is used in industrial applications such as non-destructive testing. [48]

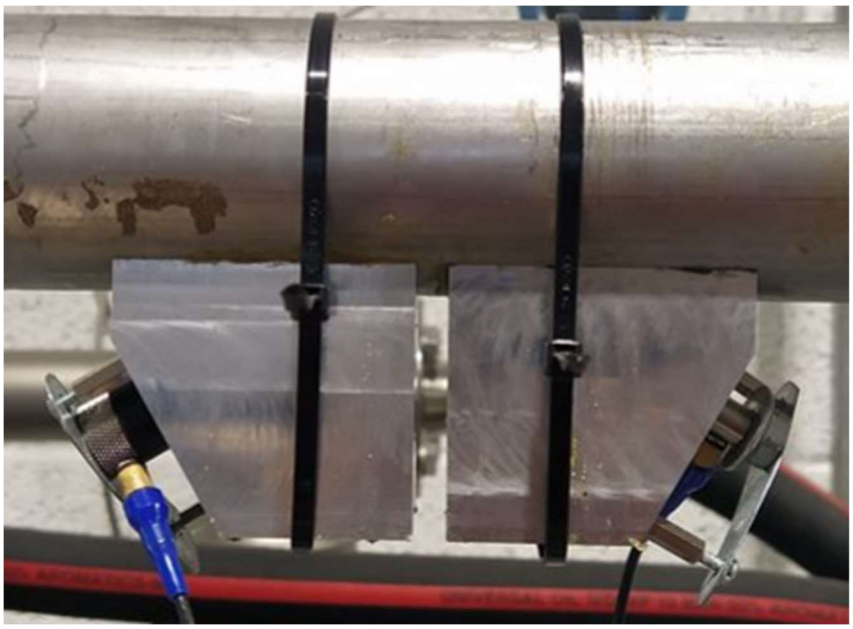


Figure 4.7: Ultrasonic transducers in pitch-catch configuration on the USN pipeline

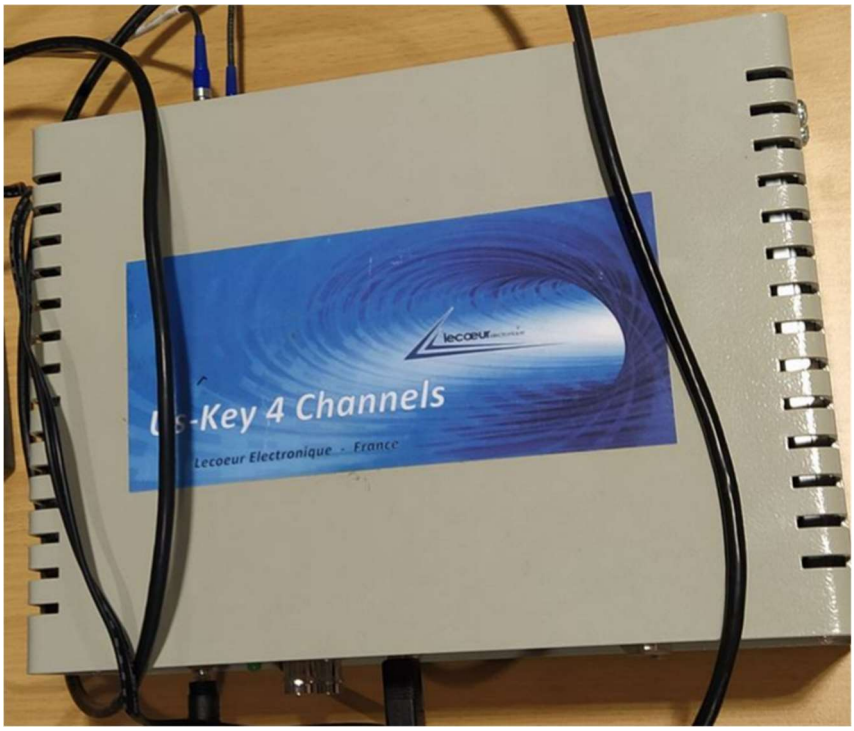


Figure 4.8: US-Key device from Lecoeur Electronique

4.2.4 Coriolis Flow Meters from Emerson

FT109B, FT110, FT114B, FT115, FT132, FT131 are Coriolis flow meters that measure the mass flow rate in kg/min accurately with an uncertainty of ± 0.01kg/min [47]. They are Emerson's Micro Motion make [49]. They can also measure temperature, density, and velocity (volume). A Coriolis flow meter FT115 installed on the flow rig for water phase is shown in Figure 4.9.

Figure 4.9: Coriolis meter FT115 for water installed at USN Rig

5 Experiments

To implement data-driven multiphase metering models, the first requirement is the availability of data to use as inputs to the models. If data is available from previous experiments then it can be used for development of models. Collection of new data by conducting fresh sets of experiments is always encouraged. The purpose of experiments in this thesis is to collect sensor data corresponding to five flow regimes of slug, plug, stratified, annular and wavy generated in the flow rig by using different combinations of air and water mass flow rates. For the case of this thesis, the aim is to collect data mainly from ECT, US and pressure sensors. In this thesis, new fresh sets of experiments are conducted to collect data from ECT and pressure sensors while already available data from US sensors is considered for development of classification and regression models.

It was decided that two-phase experiments using air and water would be conducted for this thesis. After this decision, the first step was to develop a test matrix for air/water experiments so that it remains clear that the experiments are focused on a specific pre-decided plan. Planning experiments beforehand by creating a test matrix saves a lot of time during actual experiments.

The creation of test matrix was the first challenge in this thesis. To create a test plan, the flow regime map of USN flow rig is used as a reference which has been adapted from the version of Figure 2.2. The map is shown in Figure 5.1. It has mass flow rate in kg/min as axes rather than flow velocities of air and water. There are five types of flow regimes that can be generated in the USN multiphase flow rig namely plug, slug, annular, stratified, and wavy.

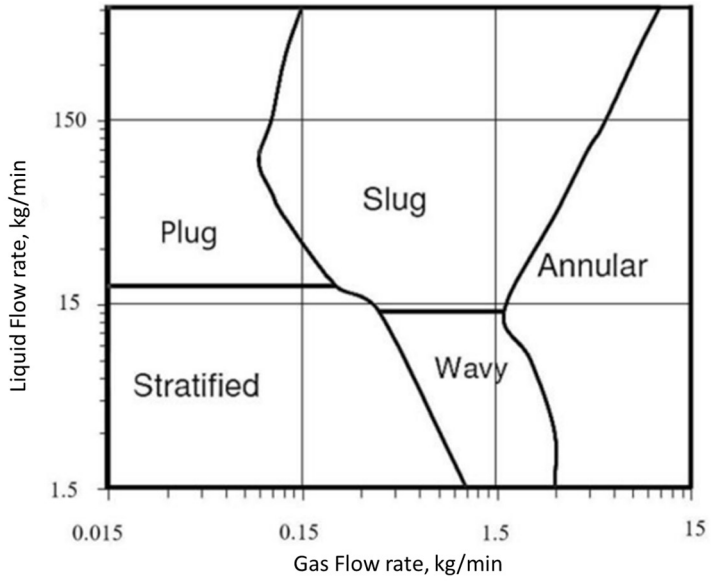


Figure 5.1: USN's rig flow regime map based on [12] using mass flow rate developed by USN

A new test matrix developed for use in this thesis based on the flow regime map is provided in Figure 5.2. The matrix has 120 experiments divided into five regimes, 30 for stratified, 30 for plug, 20 for wavy, 20 for slug and 20 for annular. By using this matrix, the flow rates of air and water can be set in various combinations to create a two-phase flow in the pipeline generating specific flow regimes. For instance, a combination of flow rate of 5 kg/min for air and 5 kg/min for water is supposed to generate a two-phase flow depicting Annular regime as

seen from Figure 5.2. Based on the mass flow rates of air and water, flow regimes can be observed through visual inspection by looking at the Plexiglas transparent section on the pipe to confirm the accuracy of the test matrix [12].

2-phase Air & Water		Water (kg/min)									
		2	3	4	5	6	75	90	105	120	135
Air (kg/min)	0.02	Stratified	Stratified	Stratified	Stratified	Stratified	Plug	Plug	Plug	Plug	Plug
	0.03	Stratified	Stratified	Stratified	Stratified	Stratified	Plug	Plug	Plug	Plug	Plug
	0.04	Stratified	Stratified	Stratified	Stratified	Stratified	Plug	Plug	Plug	Plug	Plug
	0.05	Stratified	Stratified	Stratified	Stratified	Stratified	Plug	Plug	Plug	Plug	Plug
	0.06	Stratified	Stratified	Stratified	Stratified	Stratified	Plug	Plug	Plug	Plug	Plug
	0.07	Stratified	Stratified	Stratified	Stratified	Stratified	Plug	Plug	Plug	Plug	Plug
	1.5	Wavy	Wavy	Wavy	Wavy	Wavy	Slug	Slug	Slug	Slug	Slug
	1.7	Wavy	Wavy	Wavy	Wavy	Wavy	Slug	Slug	Slug	Slug	Slug
	1.9	Wavy	Wavy	Wavy	Wavy	Wavy	Slug	Slug	Slug	Slug	Slug
	2	Wavy	Wavy	Wavy	Wavy	Wavy	Slug	Slug	Slug	Slug	Slug
	3.5	Annular	Annular	Annular	Annular	Annular					
	4	Annular	Annular	Annular	Annular	Annular					
	4.5	Annular	Annular	Annular	Annular	Annular					
	5	Annular	Annular	Annular	Annular	Annular					

Figure 5.2: Initial test matrix for 2-phase flow of Air & Water developed for this thesis

The test matrix of Figure 5.2 is completely based on theoretical knowledge of the flow regime map of the USN flow rig. At the time of actual experiments, some of the experiments were impossible to conduct on the current configuration of the USN flow rig. When the flow rate of air was tried to be set above 5 kg/min, dangerous vibrations were observed in the flow rig. These vibrations can disintegrate the entire flow rig setup. Also, similar vibrations can be observed when the flow rate of water is tried beyond 80 kg/min. Due to these operational constraints the test matrix of Figure 5.2 was modified to a matrix shown in Figure 5.3 and Figure 5.4. The air flow rate was limited to 5 kg/min and the water flow rate was limited to 80 kg/min due to unstable vibrations in the rig at higher flow rates of air and water. In this test matrix, the minimum flow rate of air is 0.07 kg/min and the maximum is 5 kg/min. The minimum flow rate of water is 2 kg/min and maximum flow rate is 77 kg/min. This is the final test matrix which was used to collect new data from ECT sensor. It has 45 experiments for various flow regime, 9 for stratified, 6 for wavy, 6 for annular, 15 for plug and 9 for slug.

2-phase Air & Water		Air (kg/min)													
		0.07	0.09	0.1	0.11	0.13	0.15	0.3	0.4	0.5	1	2.5	4	5	
Water (kg/min)	2	Minimum air flow rate is 0.065 kg/min for the rig		Stratified			Stratified			Stratified	Wavy	Wavy	Annular	Annular	Small Pump <12kg/min water
	3			Stratified			Stratified			Stratified	Wavy	Wavy	Annular	Annular	
	4			Stratified			Stratified			Stratified	Wavy	Wavy	Annular	Annular	
	75	Plug	Plug		Plug	Plug	Plug	Slug	Slug	Slug	Vibrations observed in the rig at this region of air flow rate of more than 0.5 kg/min				Large Pump >12kg/min water
	76	Plug	Plug		Plug	Plug	Plug	Slug	Slug	Slug					
	77	Plug	Plug		Plug	Plug	Plug	Slug	Slug	Slug					

Figure 5.3: Final test matrix based on operational constraints for this thesis

There are two electric pumps that inject water into the pipe. There is a small water pump that operates for water flow rate less than 12kg/min and there is a large water pump that operates for water flow rates equal to and above 12kg/min. Only one pump can be functioning at a time. So, first all the experiments for large pump were conducted. After that the pump was changed to the smaller one to conduct remaining experiments. The flow rates of air and water are controlled by the control program of LabVIEW from Figure 4.2. Lets see an example of generating a flow regime in the pipe with flow rate of 0.13kg/min for air and flow rate of 77kg/min for water. First the large pump is brought into operation. Then the water flow rate is taken to 77kg/min by increasing the flow rate of water in increments of 0.5kg/min. When the

water flow rate is fixed at 77kg/min, the flow rate of air is increased carefully and slowly in incremements of 0.01kg/min to avoid vibrations in the pipe to 0.13kg/min. When both the flow rates are reached, the flow regime is observed throught the transparent section of the pipe and noted down whether the observed flow regime conforms to the flow regime depicted in the test matrix. Vibrations were observed in the flow rig for flow rate of air above 0.5kg/min when the flow rate of water was 75 to 77 kg/min. So, for plug and slug regimes, the flow rate of air was limited to 0.5kg/min.

In all the 45 experiments, the flow regime was validated by observing them in the transparent section. All the generated flow regimes in the experiments conformed to the estimated flow regimes in the test matrix.

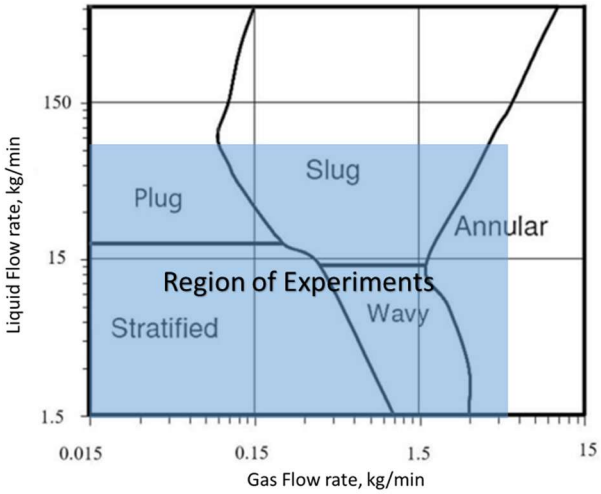


Figure 5.4: Region of experiments on USN flow rig for this thesis

The 45 experiments of Figure 5.3 are plotted on a flow regime map as shown in Figure 5.5. This plot matches the regions of flow regimes given in Figure 5.4.

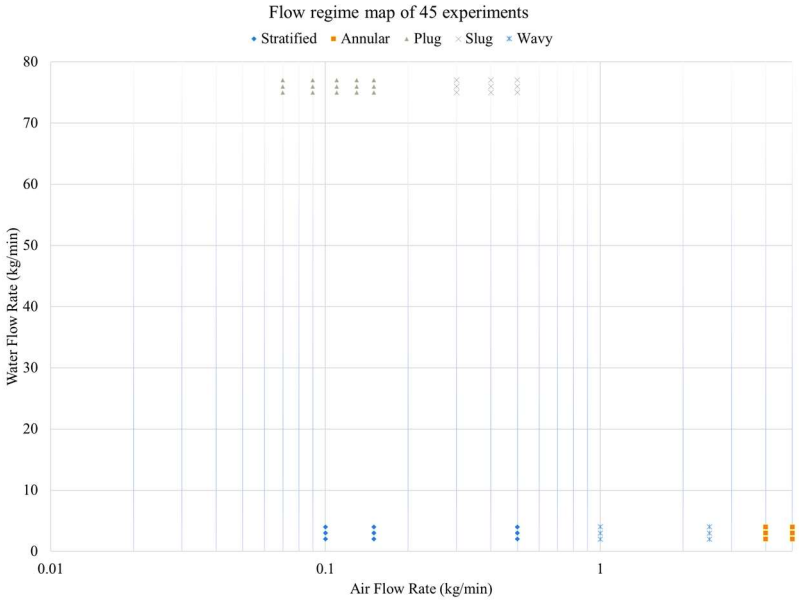


Figure 5.5: Flow Regime Map based on final test matrix

5.1 Data Collection

5.1.1 ECT Data Collection

For this study, 45 air/water 2-phase experiments based on the test matrix of Figure 5.3 were conducted in the Process Hall using the TOMOFLOW TFLR5000 ECT system in 8-electrode 2-planes configuration. The two planes are separated by a distance of 18.7cm. To begin with the ECT system was calibrated first by filling the pipe completely with air first and then water. After calibration, the ECT system was ready for experiments.

Figure 5.6 shows the 8-electrode configuration of the ECT system during the experiments. The ECT electrodes mounted on the flow rig has 12-electrodes but TFLR5000 is an 8-electrode system in dual plane configuration. To mitigate this challenge, only 8 electrodes from the 12 available electrodes were connected to the TFLR5000 by using the configuration of Figure 5.6. The same configuration implemented on TFLR5000 is depicted in Figure 5.7. From the 12 electrodes, electrodes 3,6,9 and 12 are left disconnected. This same configuration is repeated in both the ECT planes. The entire periphery of the pipe is not covered by operational electrodes as seen from Figure 5.6. The electrodes are numbered as per Figure 5.6 for the experiments. The flow direction is from left to right.

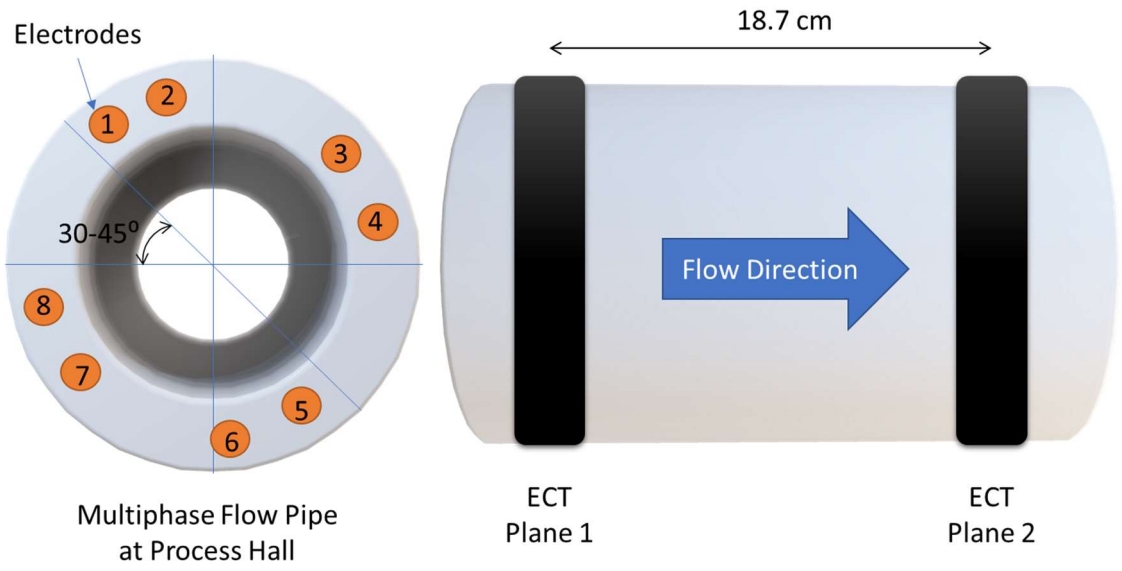


Figure 5.6: Setup of Experiment showing 2 planes of ECT and distribution of 8 electrodes around the pipe

Figure 5.7: Connection of electrodes on 3 Feb 2023 with ECT 2-planes sensor

Table 5.1 provides the parameters set during the experiments. Based on these parameters, 100 frames of permittivity distribution in the pipe are captured by the system every second providing a sampling time of 10ms. Since the interval of logging for each flow regime experiment is 30 seconds, a total of 3000 permittivity frames are collected for each flow regime experiment. The capacitance model used in these experiments is parallel capacitance model. Each frame is composed of 28 capacitances from 8-electrode configuration.

Table 5.1: Parameters set during the experiments in TFLR5000 for this thesis

Parameter	**Value**
Frames per second per plane	100
Sampling interval per plane (ms)	10
Logging duration per experiment (s)	30
Data points (frames) captured per experiment per plane	3000
Number of electrodes per plane	8
Number of planes	2
Number of capacitances per measured frame	28
Capacitance Model	Parallel [21]

Figure 5.8 shows ongoing experiments using ECT system with stratified flow regime. Through the laptop, parameters of ECT system can be modified and flow regimes can be monitored in real-time through cross-sectional images of the pipe in terms of permittivity distribution.

Figure 5.8: Ongoing ECT experiments on 3 Feb 2023 with 2-planes ECT sensor

Figure 5.9 provides the distribution of experimental data in terms of flow regimes. 45 experiments are divided into plug, slug, stratified, annular and wavy with percentages of 34%, 20%, 20%, 13% and 13% respectively. A total of 135000 frames of ECT capacitances were collected during the experiments. The distribution in terms of number of frames is shown in Table 5.2.

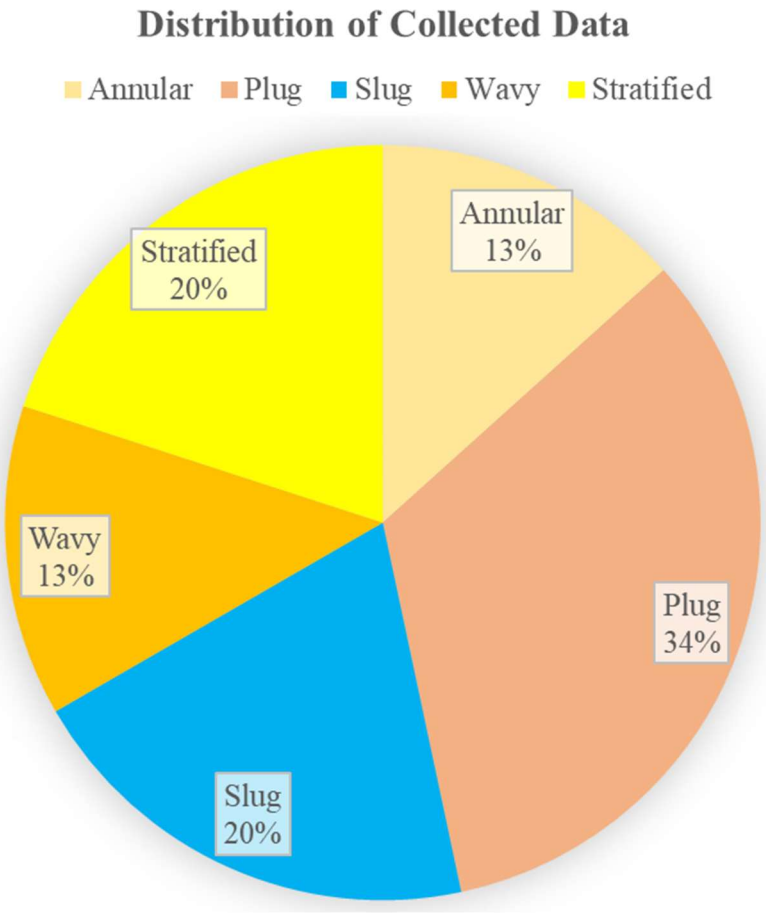


Figure 5.9: Distribution of collected data on 3 Feb 2023 with ECT 2-planes sensor

Table 5.2: Distribution of collected data in terms of frames collected per plane of ECT on 3 Feb 2023 with ECT 2-planes sensor

Regime	**Number of frames**
Annular	18000
Plug	45000
Slug	27000
Wavy	18000
Stratified	27000
Total	**135000**

5.1.1.1 Additional Data Collection

On 27 February 2023, 18 additional experiments were carried out to collect more data from the ECT as outlined in Table 5.3. The basis for these additional experiments was to match the flow rates of already available Ultrasonic sensor data for air/water two-phase flow so that ECT and Ultrasonic data can be compared at similar test conditions for flow visualization. The Ultrasonic data included flow rates of water at 80, 40, 25, 15, 10 and 5 kg/min with their combination with flow rates of air at 0.1, 0.3, 0.4, 0.5, 0.6, 1, 2, 3, 4 kg/min. The test matrix available from already conducted Ultrasonic experiments generating two-phase air/water flow is shown in Table 5.3. The same test matrix was used to collect additional ECT data. The parameters during these 18 experiments were kept same as provided in Table 5.1. Data from 2 planes of ECT was collected in this case.

During these 18 experiments, slow-motion videos were captured at 240fps through Mi A4 smartphone to observe and verify the flow regimes. The images captured from these videos of various regimes are provided in the next subsections.

Table 5.3: Test matrix of additional tests on ECT on 27 Feb 2023 to match Ultrasonic data in terms of air and water flow rates

Water (kg/min)	Air (kg/min)	Observed Regime
80	0.1	Plug
80	0.3	Slug
40	0.1	Plug
40	0.6	Slug
25	0.1	Wavy
25	0.5	Slug
25	2	Annular
15	0.1	Stratified
15	0.4	Wavy
15	1	Wavy
15	3	Annular
10	0.1	Stratified
10	0.4	Stratified
10	1	Wavy
10	4	Annular
5	0.1	Stratified
5	1	Wavy
5	4	Annular

5.1.1.1.1 Visuals of Slug Flow

Slug regime is displayed in Figure 5.10. Six frames are provided showing the entire slug flow regime sequence. Water flows at the bottom and air at the top of the pipe volume intermittently with a bullet of water with dispersed air bubbles passing at very high speed in the pipe filling it completely. This kind of flow regime creates vibrations in the pipe that can hamper the functioning of the system in which the pipe is installed. As seen personally during experiments, the bullet of water, frame 3 and 4 of Figure 5.10, travels very fast through the pipe creating vibrations in the pipe and sound of high speed water flowing in the pipe. The bullet has a lot of bubbles and it appears intermittently with a quiet phase as seen in frames 1 and 2. This flow regime is dangerous to mechanical pipes due to the force with which the water flows in the pipe and it needs to be detected quickly for mitigation of the risk of breakdowns in pipelines.

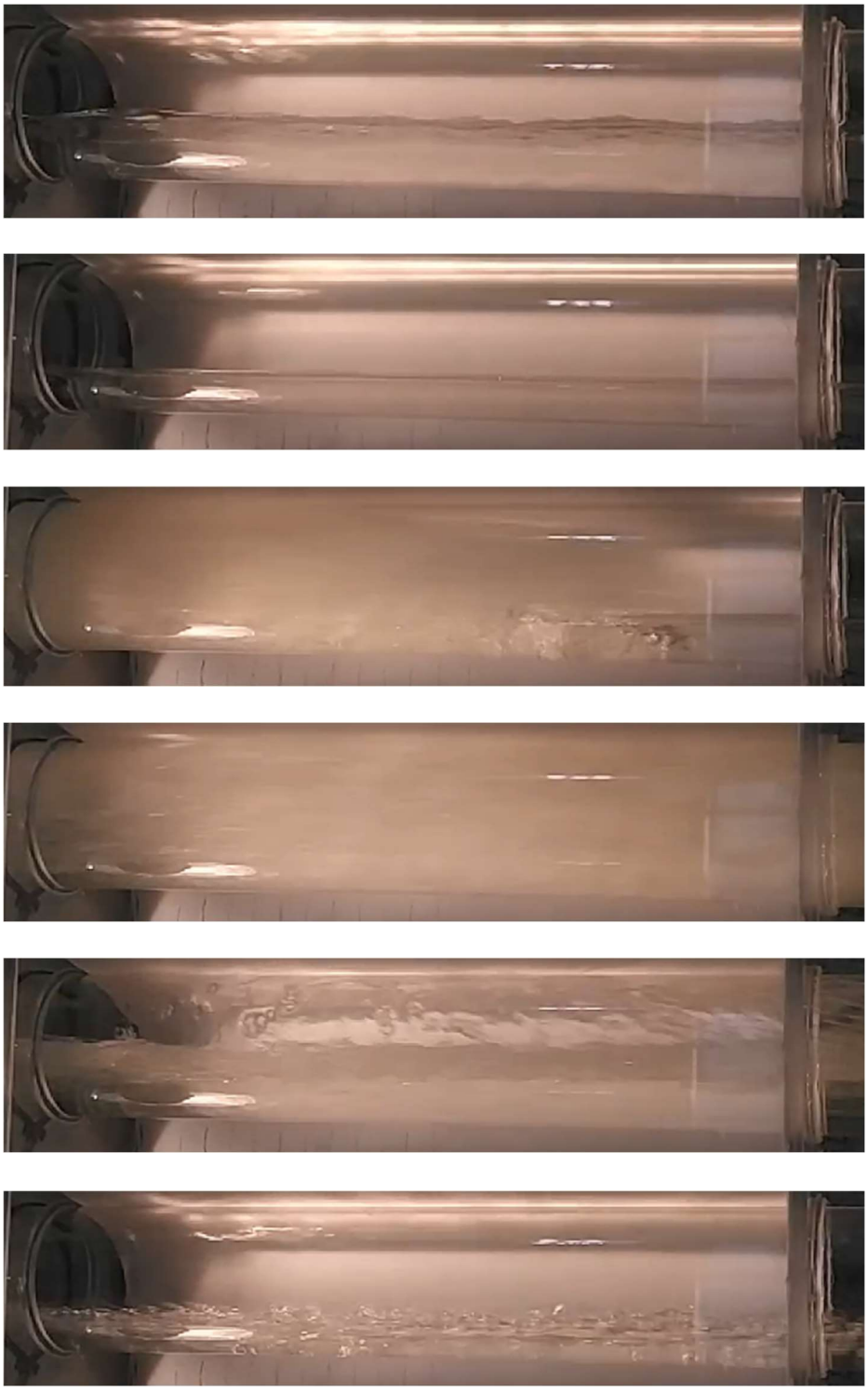


Figure 5.10: Sequence of frames (top to bottom) showing slug regime captured during experiments

5.1.1.1.2 Visuals of Plug Flow

Plug regime is displayed in Figure 5.11. Five frames are provided showing the entire plug flow regime sequence. Comparing it with Figure 2.3, large bubbles of air on top of the water phase

are observed intermittently with the pipe filled completely with water. Plug flows at a slower speed than slug as seen from visual inspections during experiments.

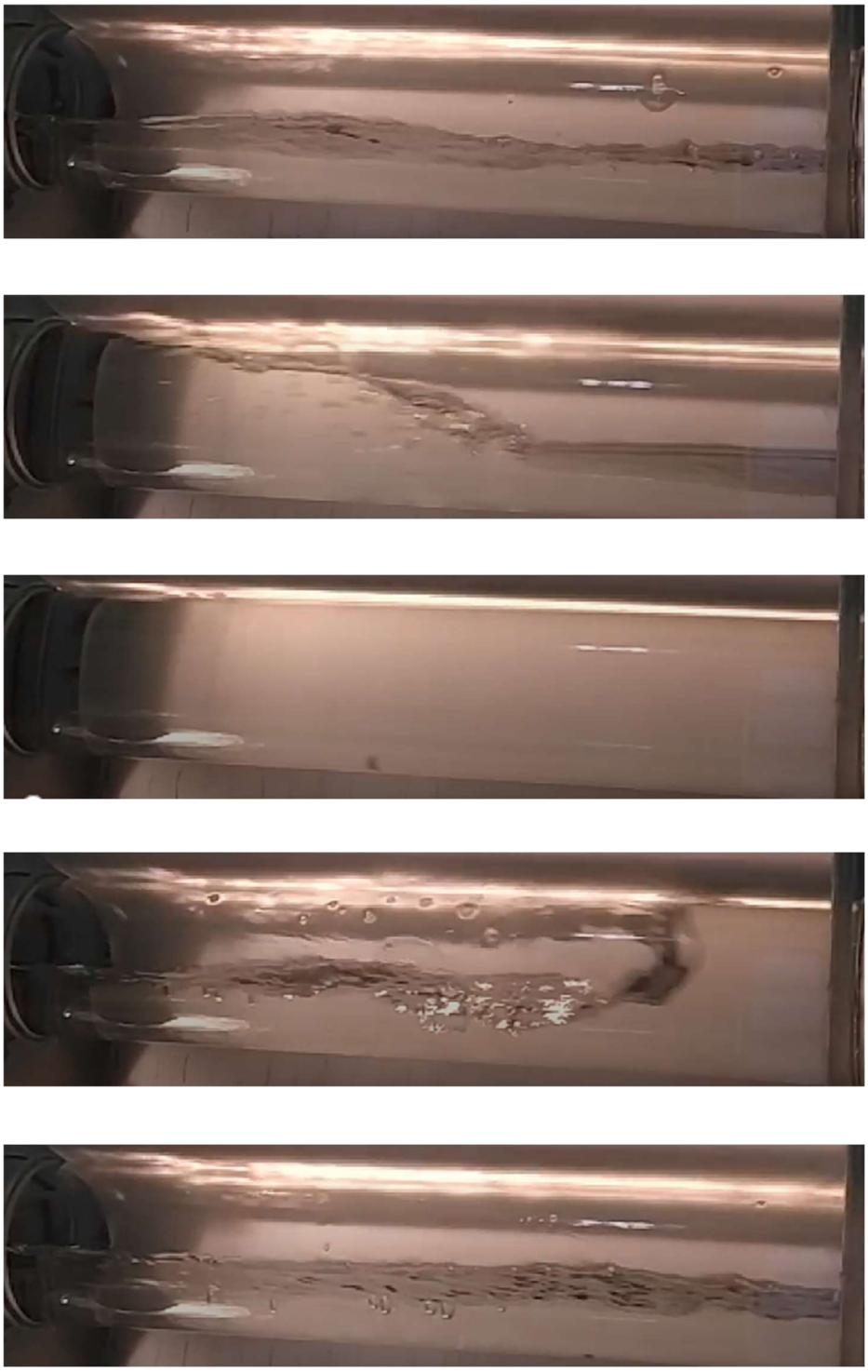


Figure 5.11: Sequence of frames (top to bottom) showing plug regime captured during experiments

5.1.1.1.3 Visual of Annular Flow

Annular regime is displayed in Figure 5.12. Comparing it with Figure 2.3, a film of water is formed at the periphery of the pipe while air flows at the center of the pipe.

Figure 5.12: Annular regime as seen in experiments

5.1.1.1.4 Visual of Wavy Flow

Wavy regime is displayed in Figure 5.13. Comparing it with Figure 2.3, waves on water are formed at the bottom of the pipe while air flows at the top of the pipe.

Figure 5.13: Wavy regime as seen in experiments

5.1.1.1.5 Visual of Stratified Flow

Stratified regime is displayed in Figure 5.14. Comparing it with Figure 2.3, a smooth film of water is formed at the bottom of the pipe while air flows at the top of the pipe.

Figure 5.14: Stratified regime as seen in experiments

5.1.1.2 Format of Collected data

The collected data from each experiment using the TFLR5000 ECT system is in the form of binary capacitance files with the extension ‘BCP’. This file can be converted to a normalized

capacitance file (‘ANC’), an image file (‘AIM’) or a volume ratio file (‘AVR’) depending on usage through the software named ‘ECT32’ provided by Process Tomography Limited.

5.1.2 Ultrasonic Data Collection

Experiments using ultrasonic pitch-catch technique were previously conducted at USN. These experiments included the air/water two flow experiments according to the test matrix of Table 5.3. The data collected from these air/water flow experiments are used in this thesis for further analysis and comparison with new data from ECT.

5.1.2.1 Format of Collected data

The collected data from the experiments using ultrasonic sensor is in the form of a binary file for each set of experiment. To convert these binary files into B-scan images, a MATLAB script developed by Tonni F. Johansen was utilized which is attached as Appendix-J. Features from this file can be extracted through an extended MATLAB script of Appendix-J provided in Appendix-K.

6 Methods & Results

In this section, methods used for data-driven multiphase flow metering are described in detail starting from exploratory data analysis to modelling metering algorithms in MATLAB. The results from the various metering algorithms are also presented here.

6.1 Exploratory Data Analysis (EDA)

6.1.1 ECT Data Analysis

The format of normalized capacitances files is given in Figure 6.1. This type of file is obtained from binary capacitance file by using the ECT32 software in Windows. The extension of the file is 'ANC'. Here each frame in the ANC file is the capacitance matrix consisting of 28 inter-electrode normalized capacitances from 8 electrodes. There are 3000 frames in each of these files. These files are flattened to 28 capacitances feature file using a Python script given in Appendix-B. Flattening of the frames means that the 28 values of capacitances in one frame which were in a non-linear form as shown in Figure 6.1 are processed in Python to get a linear form of 28 capacitances in one row of a comma-separated values (CSV) file. This step is very important since classification and regression algorithm mainly require inputs in the form of tables with all the samples in one row of a table. A snippet of the flattened form of capacitances is shown in Figure 6.2. This form is saved as a CSV file. The first 28 columns in the format corresponds to 28 normalized capacitances. The type of observed flow regime is included in the column after column 28.

The algorithm used in the Python script is straightforward. Since the 28 capacitances are in 7 rows, these 7 rows of capacitances are first stored in a pandas dataframe and then concatenated as 1 row in another dataframe. This algorithm is repeated 3000 times for one ANC file of each experiment. The last column of the flattened file has the flow regime name corresponding to the experiment which is directly inserted in the Python script from the file name of the ANC file. This script operates when all ANC files are stored in one folder. It will convert all the ANC files to a flattened version while Appendix-C saves a combined flattened file in CSV format.

In this thesis, C12 is referred to as the normalized capacitance between electrode 1 and 2, C13 is referred to as the normalized capacitance between electrode 1 and 3, and so on. To generalize, the normalized capacitance between electrode X and Y is denoted by CXY.

```
## Normalised capacitance data file
## Capacitance model: Parallel
## Created by: ECT32v3 for Win32 3.xx SVN 659 +mods [2009/11/30 19:54:36]
## Source: w2a0_1.bcp
##   Description: Contents of capture mode buffer
##   Date: 11:26:57 GMT Standard Time, 03Feb10
## Electrodes = 8, Measurements = 28
## Data for Plane 1


## frame 1 (0.0 msec)
-0.019 -0.011 -0.007 -0.011 0.078 0.117 0.104
-0.006 -0.004 -0.008 0.009 0.012 0.008
0.002 -0.009 -0.003 0.005 0.008
-0.007 -0.008 0.003 0.007
-0.428 0.041 0.057
1.793 1.658
3.939
```

C12	C13		...	C1e
C23	C24	...	C2e	
C34	C35	...	C3e	
.	.	.		
C(e-2)(e-1)	C(e-2)e			
C(e-1)e				

Figure 6.1: Normalized capacitance ANC file contents in general in comparison with the format of measurement representation where e is the total number of measurement electrodes [21]

	C12	C13	C14	C15	C16	C17	C18	C23	C24	C25	...	C46	C47	C48	C56	C57	C58	C67	C68	C78	Regime
0	-0.032	-0.014	-0.008	-0.010	0.123	0.155	0.141	-0.007	-0.004	-0.009	...	-0.008	0.005	0.009	-0.596	0.056	0.073	1.954	1.853	3.673	Stratified
1	-0.032	-0.014	-0.008	-0.010	0.123	0.155	0.141	-0.007	-0.004	-0.009	...	-0.008	0.005	0.009	-0.596	0.056	0.073	1.954	1.853	3.672	Stratified
2	-0.032	-0.014	-0.008	-0.010	0.123	0.155	0.141	-0.007	-0.004	-0.009	...	-0.008	0.005	0.009	-0.597	0.056	0.073	1.955	1.853	3.672	Stratified
3	-0.032	-0.014	-0.009	-0.010	0.123	0.155	0.141	-0.007	-0.004	-0.009	...	-0.008	0.005	0.009	-0.596	0.056	0.073	1.955	1.853	3.672	Stratified
4	-0.032	-0.014	-0.008	-0.010	0.123	0.155	0.141	-0.007	-0.004	-0.009	...	-0.008	0.005	0.009	-0.597	0.056	0.073	1.955	1.853	3.672	Stratified
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
134994	0.842	-0.009	-0.006	-0.021	0.105	0.244	0.298	-0.019	-0.003	-0.007	...	-0.006	0.005	0.009	1.537	0.061	0.054	1.858	1.686	3.864	Slug
134995	0.838	-0.009	-0.006	-0.020	0.103	0.241	0.294	-0.019	-0.003	-0.007	...	-0.006	0.005	0.009	1.547	0.057	0.051	1.849	1.679	3.877	Slug
134996	0.833	-0.009	-0.006	-0.020	0.102	0.238	0.288	-0.019	-0.003	-0.007	...	-0.006	0.005	0.008	1.521	0.054	0.049	1.849	1.678	3.884	Slug

Figure 6.2: Flattened normalized capacitance form with flow regimes

6.1.1.1 Distribution of Normalized Capacitances

After converting the binary capacitances files to normalized capacitances files through ECT32 software and flattening the ANC files, the normalized capacitances corresponding to the 8 electrodes can be analyzed to explore their distribution for the five flow regimes considered in this thesis. There are many methods to analyze the distribution of capacitances. In this thesis, box plots, correlation and spectrograms are the primary methods used to explore the distribution of capacitances, examine their relationships, and to find outliers in the data. The nature of information that can be extracted from raw normalized capacitance data from ECT will be investigated in this section. Based on this analysis, features can be selected for further use as inputs to classification and regression algorithms.

Box plots are old ways to investigate distribution of data and to find outliers in the data. They give a lot of information such as median, range, minimum, maximum, first quartile and third quartile in a dataset. By looking at a box plot, the distribution of data can be easily seen and compared with other box plots on the same axes. [50]

In this section, the distribution of normalized capacitances for various regimes is compared and explored by using box plots. From 135000 frames collected through experiments, box plots are created in MATLAB by using the script provided in Appendix-F. The horizontal axis has flow regimes while the vertical axis has normalized capacitances. So, the box plots are created comparing the normalized capacitances with respect to the five flow regimes of stratified, wavy, annular, plug and slug. Each box plot belongs to the normalized capacitances of a

particular pair of electrodes such as C12, C13, etc. The box plots for the 28 inter-electrode capacitances from C12 to C78 are given in Appendix-M.

The box plot of C12 is shown as Figure 6.3. The range of distribution of C12 is low for stratified and wavy regimes as compared to annular, plug and slug regimes. From this box plot, plug, slug, annular can be distinguished easily from stratified and wavy. An overlapping distribution is observed for stratified and wavy.

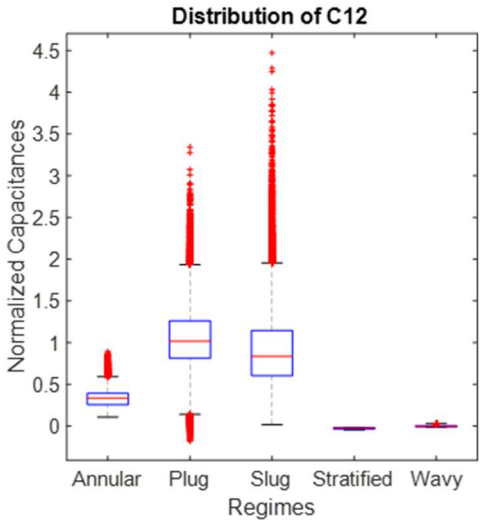


Figure 6.3: Box Plot of C12

Similarly other box plots can be investigated from Appendix-M. None of them gives a clear demarcation or clusters of regimes. Therefore, regime classification through box plots is difficult to observe. But similarities between inter-electrode capacitances can be observed easily for such a large amount of data.

By observing the box plots of Figure 6.4 carefully, it can be inferred that C16, C17 and C18 are having the same distribution profile.

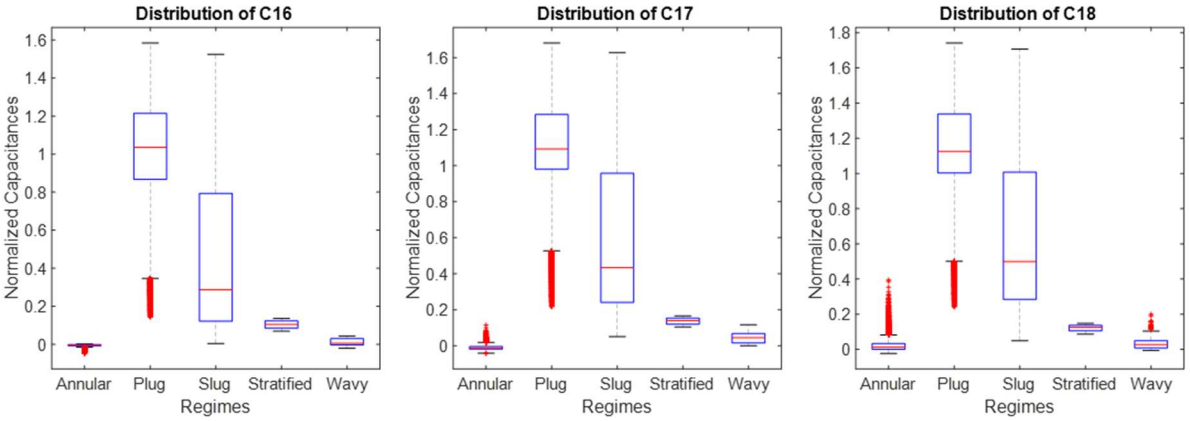


Figure 6.4: Box Plots of C16, C17 and C18

By observing the box plots of Figure 6.5 carefully, it can be inferred that C25, C26, C27 and C28 are having the same distribution profile.

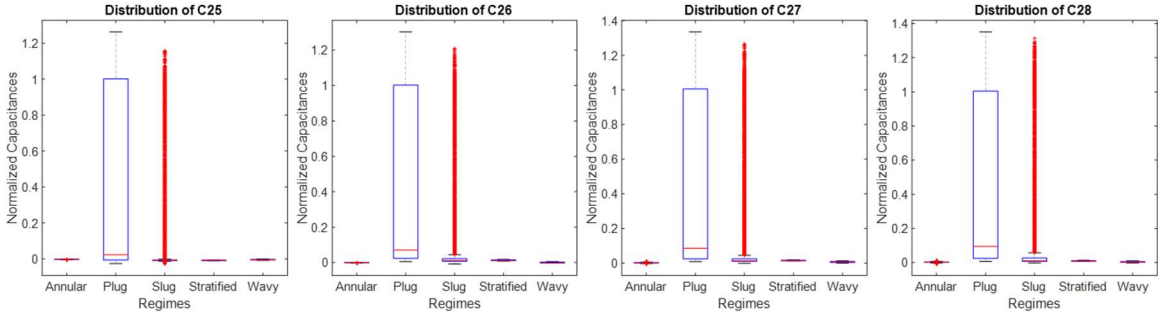


Figure 6.5: Box Plots of C25, C26, C27 and C28

By observing the box plots of Figure 6.6 carefully, it can be inferred that C35, C36, C37 and C38 are having the same distribution profile.

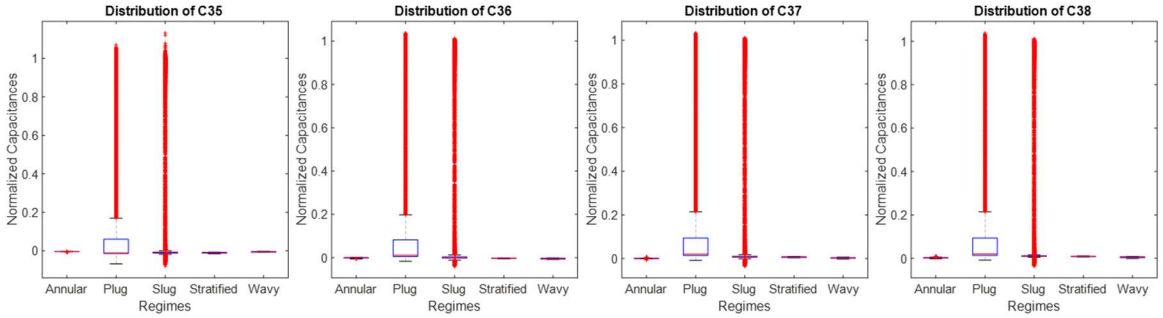


Figure 6.6: Box Plots of C35, C36, C37 and C38

By observing the box plots of Figure 6.7 carefully, it can be inferred that C45, C46, C47 and C48 are having the same distribution profile.

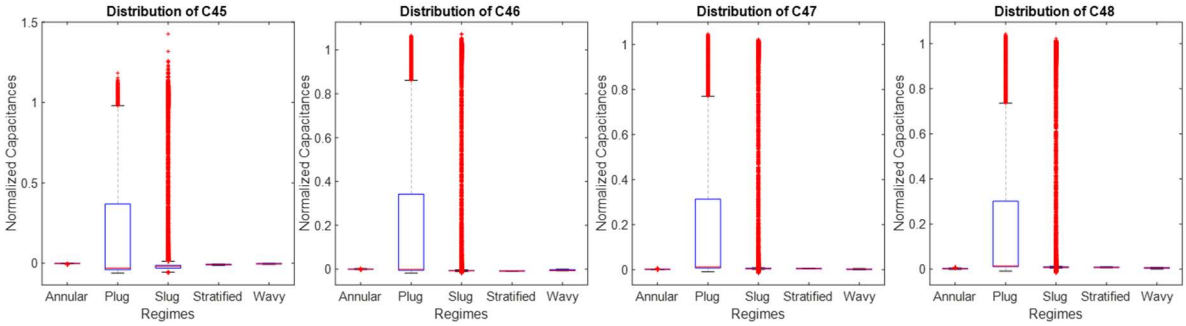


Figure 6.7: Box Plots of C45, C46, C47 and C48

By observing the box plots of Figure 6.8 carefully, it can be inferred that C57 and C58 are having the same distribution profile.

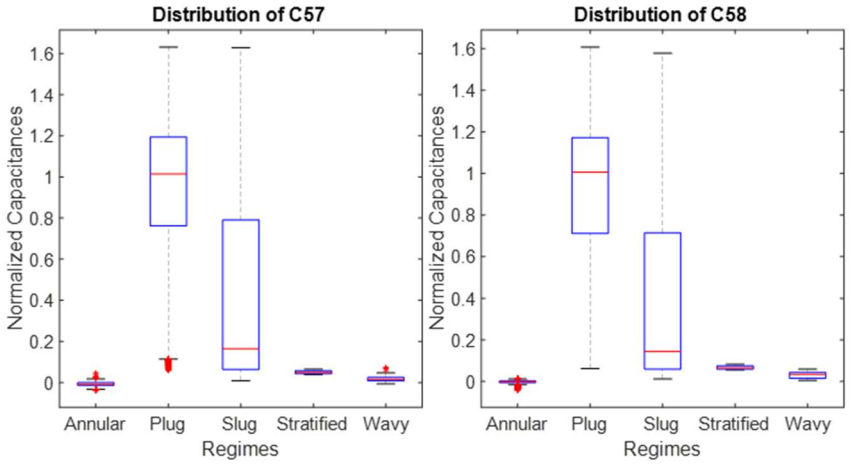


Figure 6.8: Box Plots of C57 and C58

By observing the box plots of Figure 6.9 carefully, it can be inferred that C67 and C68 are having the same distribution profile.

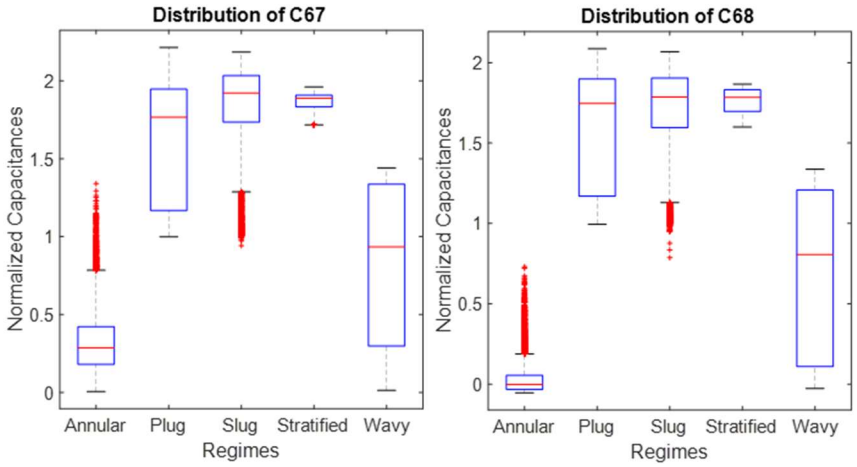


Figure 6.9: Box Plots of C67 and C68

The last sets of capacitances corresponding to each electrode are found to have a similar distribution. This similarity in the distribution of capacitances can be used to decrease the number of features when dealing with classification and regression algorithms saving a lot of computation power and storage.

6.1.1.2 Correlation Analysis of ECT Inter-Electrode Capacitances

After performing a distribution analysis of the inter-electrode capacitances through box plots and finding similar patterns of distribution, another method of similarity analysis is used to confirm these observations. The method is called correlation analysis. Microsoft Excel, Python and MATLAB are used to implement this correlation analysis.

After performing correlation analysis of all the 28 normalized capacitances, the correlation coefficients are generated in Python and Excel as shown in Figure 6.10 and Figure 6.11 respectively. The coefficients greater than 0.99 are highlighted in red in Figure 6.11. These coefficients show a high correlation and similarity between normalized capacitances. This analysis confirms the observations from box plots that there is a similarity in their distribution.

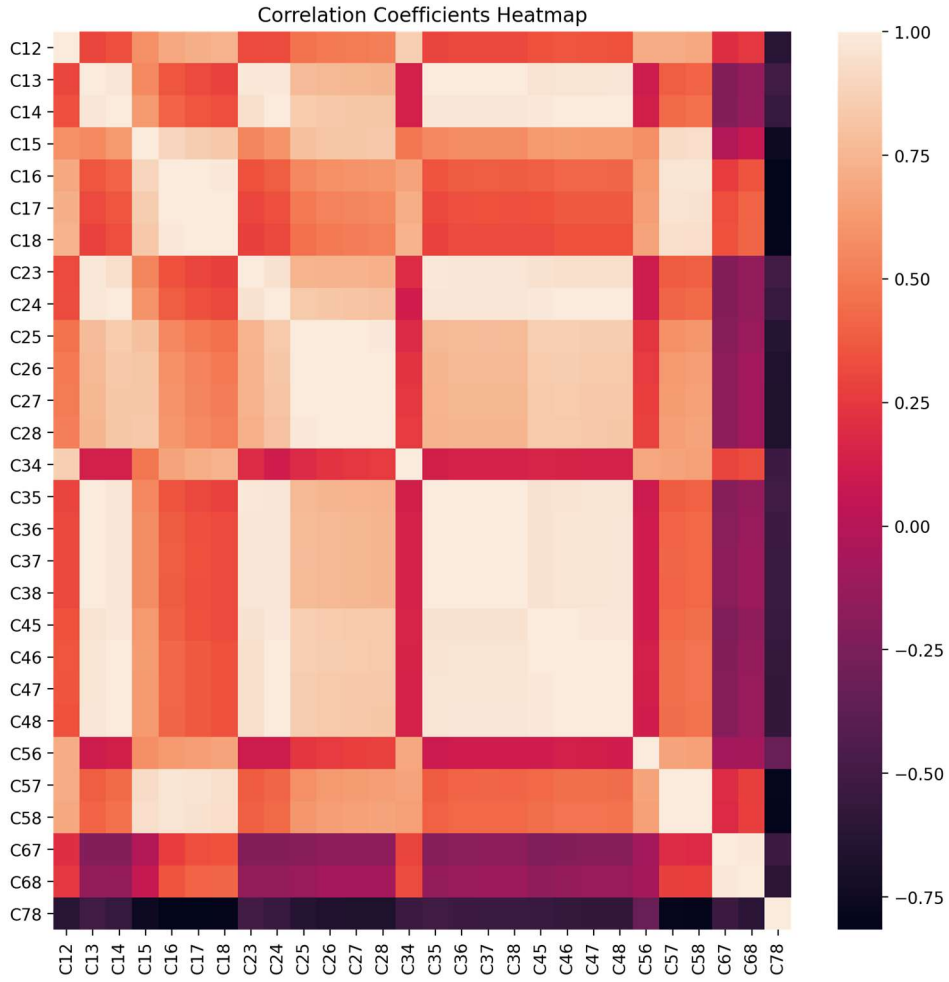


Figure 6.10: Correlation coefficients heatmap of normalized capacitances created in Python (Appendix-N)

	C12	C13	C14	C15	C16	C17	C18	C23	C24	C25	C26	C27	C28	C34	C35	C36	C37	C38	C45	C46	C47	C48	C56	C57	C58	C67	C68	C78
C12	1.00	0.30	0.34	0.58	0.69	0.72	0.74	0.31	0.32	0.47	0.49	0.50	0.51	0.86	0.29	0.31	0.31	0.31	0.35	0.36	0.35	0.34	0.71	0.71	0.69	0.20	0.24	-0.62
C13	0.30	1.00	0.98	0.56	0.36	0.31	0.29	0.98	0.98	0.77	0.76	0.75	0.75	0.13	1.00	1.00	1.00	1.00	0.96	0.97	0.97	0.97	0.10	0.39	0.41	-0.21	-0.15	-0.51
C14	0.34	0.98	1.00	0.62	0.41	0.36	0.33	0.95	1.00	0.85	0.83	0.83	0.82	0.13	0.97	0.98	0.98	0.98	0.99	1.00	1.00	1.00	0.12	0.44	0.46	-0.21	-0.14	-0.56
C15	0.58	0.56	0.62	1.00	0.90	0.86	0.83	0.54	0.60	0.80	0.82	0.83	0.84	0.49	0.55	0.57	0.57	0.57	0.62	0.64	0.63	0.63	0.58	0.92	0.94	-0.02	0.08	-0.76
C16	0.69	0.36	0.41	0.90	1.00	0.99	0.99	0.34	0.39	0.55	0.59	0.60	0.61	0.67	0.35	0.38	0.39	0.38	0.39	0.42	0.42	0.42	0.63	0.97	0.97	0.27	0.35	-0.82
C17	0.72	0.31	0.36	0.86	0.99	1.00	1.00	0.30	0.34	0.49	0.53	0.54	0.55	0.71	0.31	0.34	0.34	0.34	0.34	0.37	0.37	0.37	0.64	0.96	0.95	0.33	0.41	-0.82
C18	0.74	0.29	0.33	0.83	0.99	1.00	1.00	0.28	0.31	0.46	0.50	0.51	0.52	0.74	0.28	0.31	0.32	0.31	0.31	0.34	0.34	0.34	0.67	0.95	0.94	0.34	0.41	-0.80
C23	0.31	0.98	0.95	0.54	0.34	0.30	0.28	1.00	0.95	0.74	0.74	0.73	0.73	0.19	0.98	0.98	0.98	0.98	0.95	0.95	0.95	0.95	0.10	0.38	0.40	-0.21	-0.15	-0.49
C24	0.32	0.98	1.00	0.60	0.39	0.34	0.31	0.95	1.00	0.84	0.82	0.81	0.81	0.12	0.98	0.98	0.98	0.98	0.99	1.00	1.00	1.00	0.10	0.41	0.44	-0.21	-0.14	-0.54
C25	0.47	0.77	0.85	0.80	0.55	0.49	0.46	0.74	0.84	1.00	1.00	1.00	0.99	0.19	0.76	0.77	0.77	0.77	0.86	0.86	0.86	0.85	0.24	0.59	0.61	-0.20	-0.12	-0.64
C26	0.49	0.76	0.83	0.82	0.59	0.53	0.50	0.74	0.82	1.00	1.00	1.00	1.00	0.23	0.75	0.76	0.76	0.76	0.85	0.85	0.84	0.84	0.26	0.62	0.65	-0.17	-0.08	-0.67
C27	0.50	0.75	0.83	0.83	0.60	0.54	0.51	0.73	0.81	1.00	1.00	1.00	1.00	0.24	0.74	0.75	0.75	0.75	0.84	0.84	0.84	0.83	0.27	0.63	0.66	-0.17	-0.08	-0.67
C28	0.51	0.75	0.82	0.84	0.61	0.55	0.52	0.73	0.81	0.99	1.00	1.00	1.00	0.26	0.74	0.75	0.75	0.75	0.84	0.84	0.83	0.83	0.28	0.64	0.67	-0.16	-0.08	-0.67
C34	0.86	0.13	0.13	0.49	0.67	0.71	0.74	0.19	0.12	0.19	0.23	0.24	0.26	1.00	0.13	0.15	0.15	0.14	0.16	0.15	0.15	0.14	0.69	0.67	0.65	0.29	0.33	-0.53
C35	0.29	1.00	0.97	0.55	0.35	0.31	0.28	0.98	0.98	0.76	0.75	0.74	0.74	0.13	1.00	1.00	1.00	1.00	0.96	0.96	0.97	0.97	0.09	0.38	0.40	-0.20	-0.14	-0.50
C36	0.31	1.00	0.98	0.57	0.38	0.34	0.31	0.98	0.98	0.77	0.76	0.75	0.75	0.15	1.00	1.00	1.00	1.00	0.96	0.97	0.97	0.97	0.11	0.41	0.43	-0.19	-0.12	-0.53
C37	0.31	1.00	0.98	0.57	0.39	0.34	0.32	0.98	0.98	0.77	0.76	0.75	0.75	0.15	1.00	1.00	1.00	1.00	0.95	0.97	0.97	0.97	0.11	0.41	0.43	-0.18	-0.11	-0.54
C38	0.31	1.00	0.98	0.57	0.38	0.34	0.31	0.98	0.98	0.77	0.76	0.75	0.75	0.14	1.00	1.00	1.00	1.00	0.95	0.97	0.97	0.97	0.11	0.41	0.43	-0.18	-0.11	-0.53
C45	0.35	0.96	0.99	0.62	0.39	0.34	0.31	0.95	0.99	0.86	0.85	0.84	0.84	0.16	0.96	0.96	0.95	0.95	1.00	0.99	0.99	0.99	0.11	0.43	0.45	-0.23	-0.16	-0.54
C46	0.36	0.97	1.00	0.64	0.42	0.37	0.34	0.95	1.00	0.86	0.85	0.84	0.84	0.15	0.96	0.97	0.97	0.97	0.99	1.00	1.00	1.00	0.13	0.45	0.47	-0.21	-0.14	-0.57
C47	0.35	0.97	1.00	0.63	0.42	0.37	0.34	0.95	1.00	0.86	0.84	0.84	0.83	0.15	0.97	0.97	0.97	0.97	0.99	1.00	1.00	1.00	0.12	0.45	0.47	-0.20	-0.13	-0.57
C48	0.34	0.97	1.00	0.63	0.42	0.37	0.34	0.95	1.00	0.85	0.84	0.83	0.83	0.14	0.97	0.97	0.97	0.97	0.99	1.00	1.00	1.00	0.12	0.45	0.47	-0.19	-0.12	-0.57
C56	0.71	0.10	0.12	0.58	0.63	0.64	0.67	0.10	0.10	0.24	0.26	0.27	0.28	0.69	0.09	0.11	0.11	0.11	0.11	0.13	0.12	0.12	1.00	0.67	0.66	-0.08	-0.06	-0.33
C57	0.71	0.39	0.44	0.92	0.97	0.96	0.95	0.38	0.41	0.59	0.62	0.63	0.64	0.67	0.38	0.41	0.41	0.41	0.43	0.45	0.45	0.45	0.67	1.00	1.00	0.20	0.28	-0.79
C58	0.69	0.41	0.46	0.94	0.97	0.95	0.94	0.40	0.44	0.61	0.65	0.66	0.67	0.65	0.40	0.43	0.43	0.43	0.45	0.47	0.47	0.47	0.66	1.00	1.00	0.18	0.27	-0.80
C67	0.20	-0.21	-0.21	-0.02	0.27	0.33	0.34	-0.21	-0.21	-0.20	-0.17	-0.17	-0.16	0.29	-0.20	-0.19	-0.18	-0.18	-0.23	-0.21	-0.20	-0.19	-0.08	0.20	0.18	1.00	0.99	-0.53
C68	0.24	-0.15	-0.14	0.08	0.35	0.41	0.41	-0.15	-0.14	-0.12	-0.08	-0.08	-0.08	0.33	-0.14	-0.12	-0.11	-0.11	-0.16	-0.14	-0.13	-0.12	-0.06	0.28	0.27	0.99	1.00	-0.60
C78	-0.62	-0.51	-0.56	-0.76	-0.82	-0.82	-0.80	-0.49	-0.54	-0.64	-0.67	-0.67	-0.67	-0.53	-0.50	-0.53	-0.54	-0.53	-0.54	-0.57	-0.57	-0.57	-0.33	-0.79	-0.80	-0.53	-0.60	1.00

Figure 6.11: Correlation coefficients of normalized capacitances. Red color indicates correlation coefficients with values more than 0.99

From Figure 6.12 to Figure 6.17, correlation matrices are provided to observe a visual display of correlations between the normalized capacitances. These matrices are developed in MATLAB through the script provided in Appendix-G.

From Figure 6.12 it is clearly evident that capacitances C16, C17 and C18 have a high correlation coefficient, which is more than 0.99. So out of these three only one can be used as a feature for machine learning model. This enables reduction of features by two.

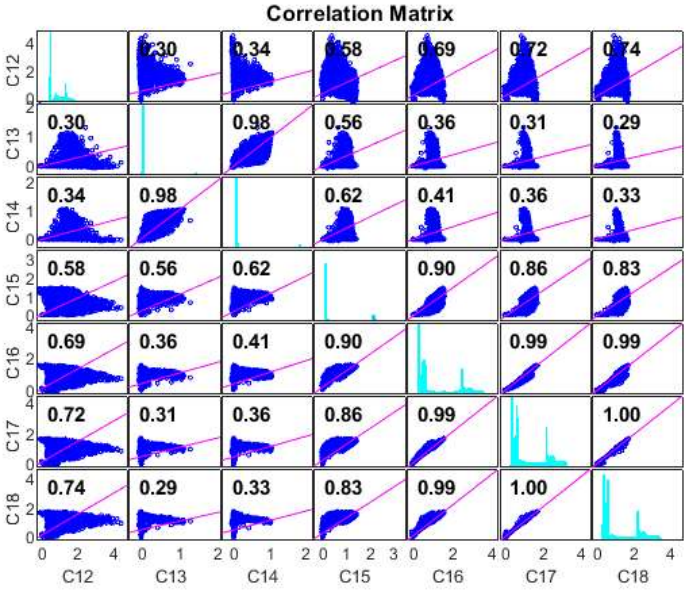


Figure 6.12: Correlation matrix C12 to C18

As seen from Figure 6.13, capacitances C25, C26, C27 and C28 have a high correlation coefficient between them. The correlation coefficient is more than 0.99 between them. This enables the use of only one of these capacitances as feature of machine learning model. This makes it possible the reduction of feature by three.

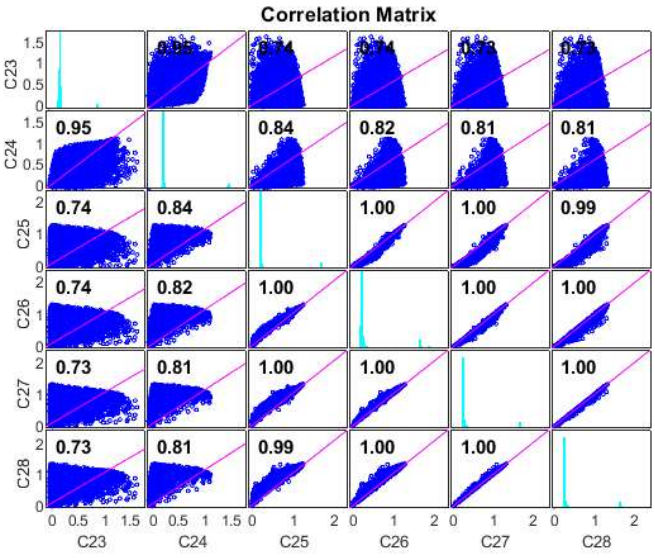


Figure 6.13: Correlation matrix C23 to C28

Figure 6.14 is a correlation matrix having C34, C35, C36, C37 and C38 capacitances. C35, C36, C37 and C38 have a correlation coefficient of 1 between them. Therefore, only one out of these four can be used as a feature for machine learning model. Four features have been reduced to one through this analysis.

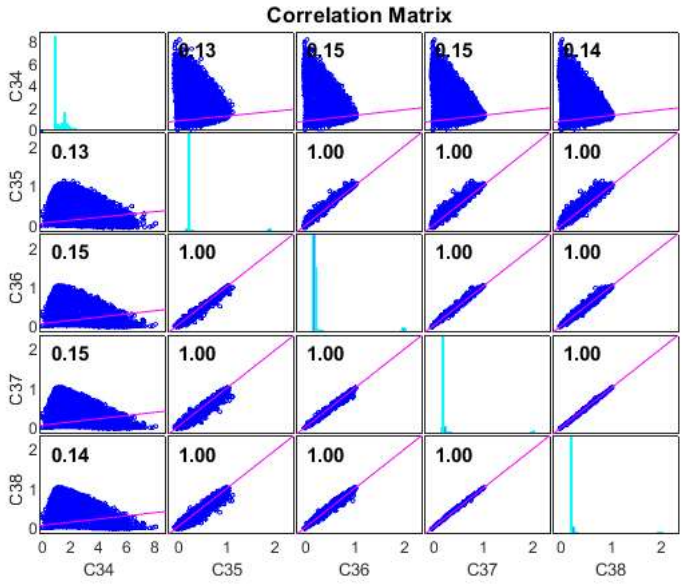


Figure 6.14: Correlation matrix C34 to C38

As evident from Figure 6.15, C46, C47, C48 and C45 show a high level of correlation with that correlation coefficient more than 0.99. Out of these four capacitances only one can be used as a feature for machine learning model. This enabled the reduction of features by three.

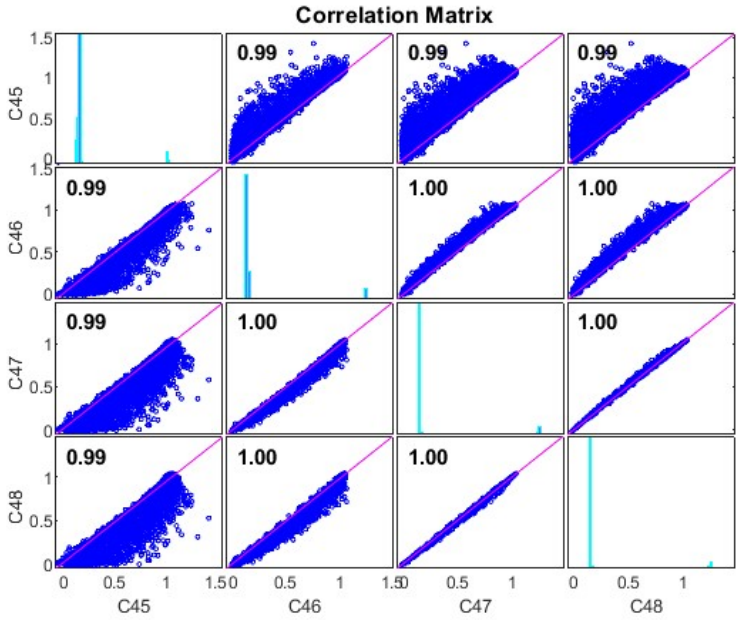


Figure 6.15: Correlation matrix C45 to C48

Looking at Figure 6.16, C57 and C58 are having a perfect correlation coefficient of 1. Only one of them can be used as a feature for machine learning model. This made it possible to reduce features by 1.

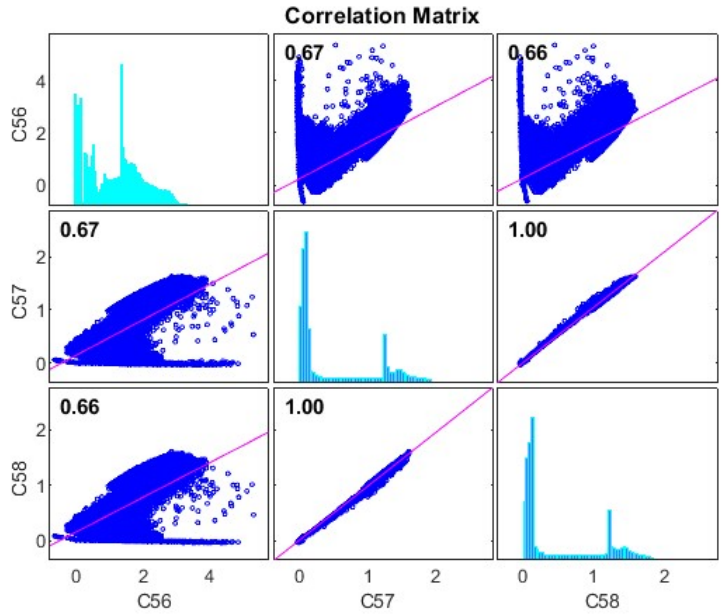


Figure 6.16: Correlation matrix C56 to C58

As seen in Figure 6.17, normalized capacitance C67 is showing high correlation with C68 with a correlation coefficient of 0.99. Only one of them can be used as a feature for machine learning models based on the correlation coefficient. This reduced features by one.

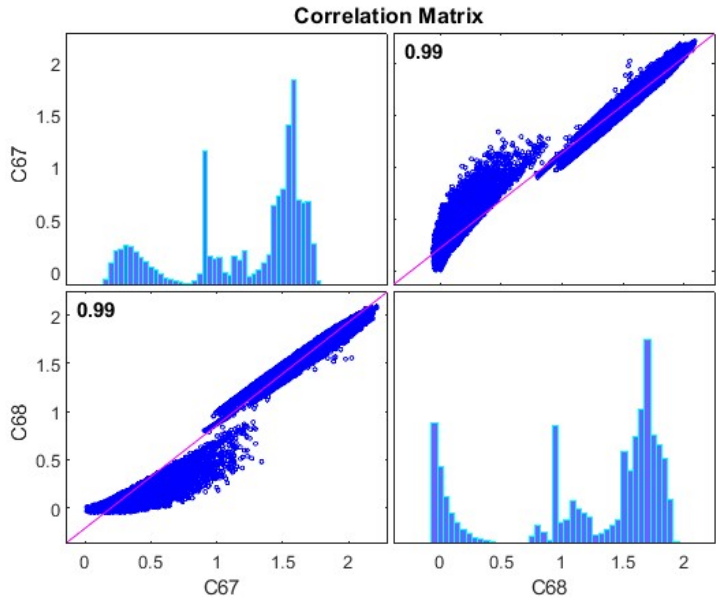


Figure 6.17: Correlation matrix C67 to C68

Therefore, after correlation analysis of inter-electrode normalized capacitances, 28 normalized capacitances can be reduced to 15 normalized capacitances and used as features for flow regime classification models.

6.1.1.3 Spectrograms of Flow Regimes from ECT Normalized Capacitances

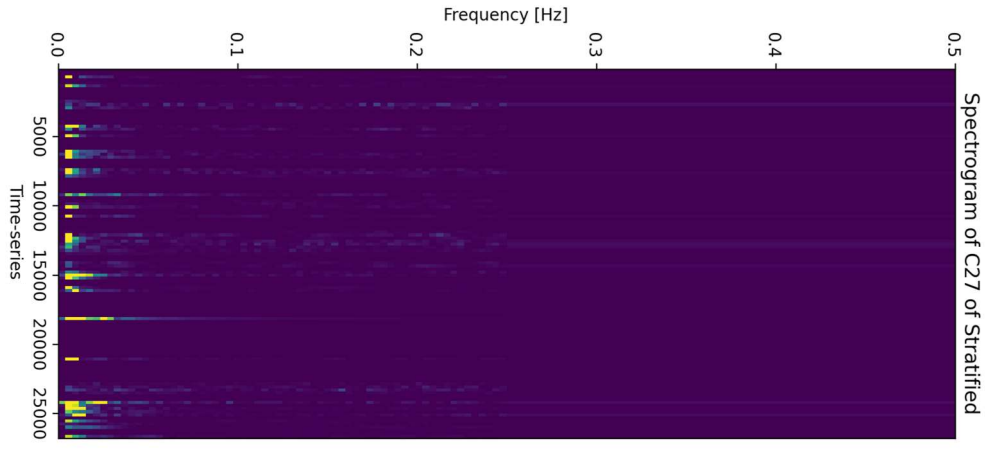
There is another way to see the overview of ECT data by using spectrograms. Appendix – H provides a Python script to generate spectrograms of normalized capacitances of specific flow regimes [51]. In this particular script, 'signal' library from 'SciPy' library is imported to use its signal processing functions. The main function, to generate spectrograms in this script, is 'signal.spectrogram' that takes inputs of normalized capacitances for a particular flow regime.

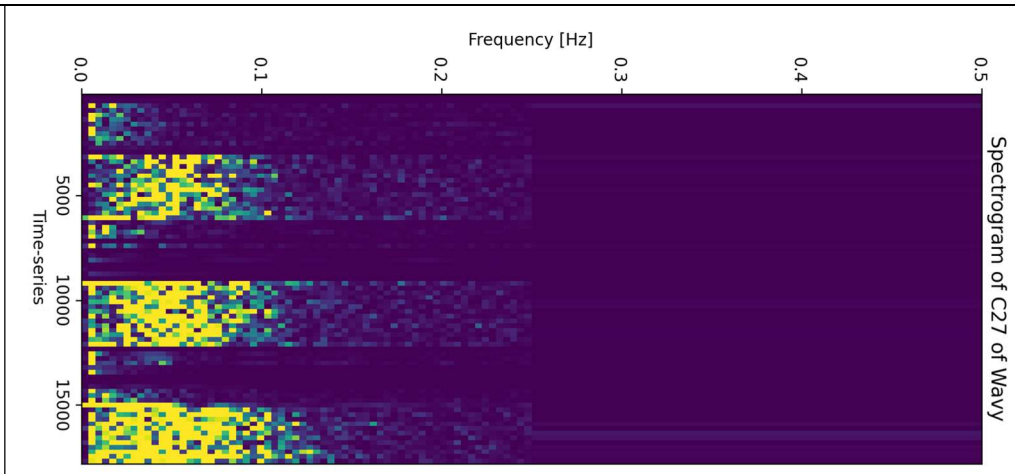
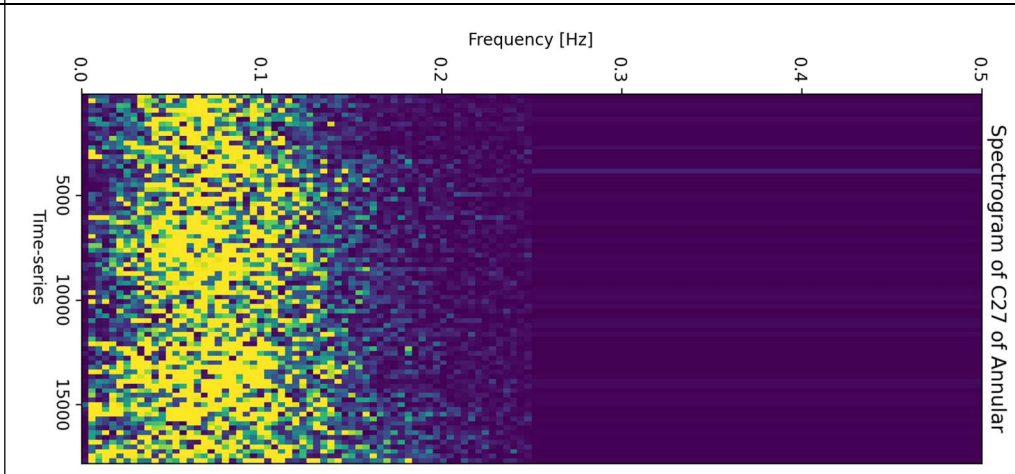
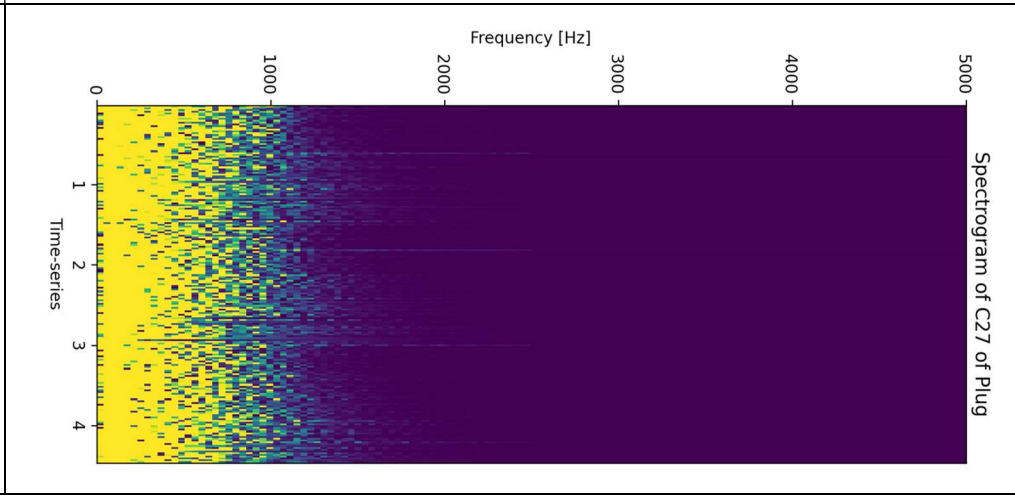
A comparison of spectrograms of capacitance C27 for five flow regimes is given in Table 6.1. The images are rotated in the table. The horizontal axis is the total number of experiments taken in a series of time for a regime. The vertical axis is the frequency in Hz. The density of color in the spectrogram determines the capacitances region of frequency.

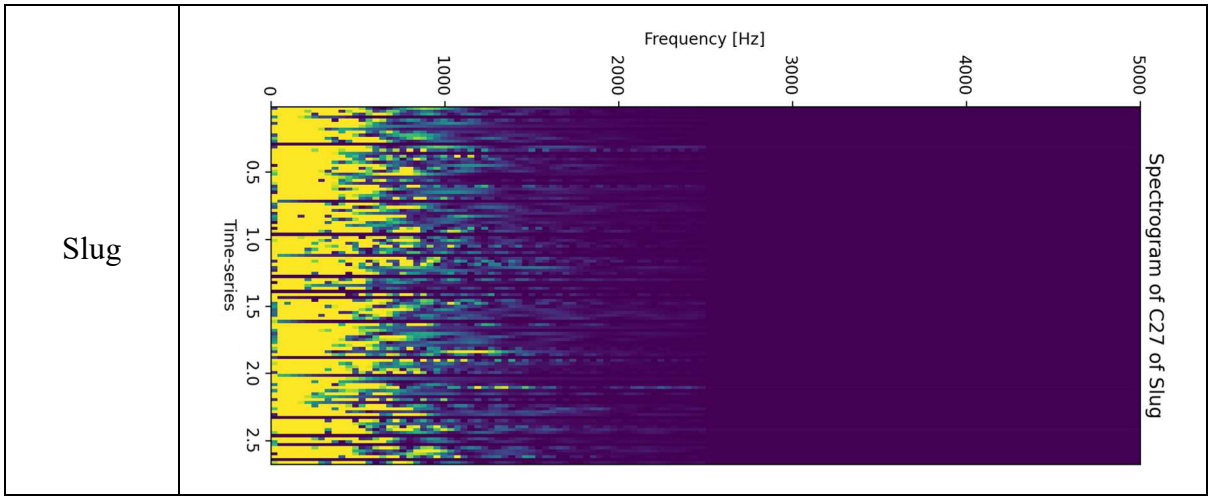
It is readily observed that for C27, the spectrograms provide clear distinction between the five flow regimes. Stratified regime lies mainly in the region of 0.01 to 0.05Hz. The frequency response varies with time. Wavy flow provides a wavy spectrogram with frequencies reaching up to 0.25Hz. Annular flow follows the same pattern of annular flow in the pipe with frequencies lying consistently from 0 to 0.25Hz. Plug is denser than annular in the spectrograms and in the spectrogram of slug, there are absence of frequencies intermittently during the time-series.

These images can be used as inputs to flow regimes classification neutral network to create classification models. Similar spectrograms can be developed for the other 27 capacitances. Spectrograms are a better option than box plots to observe normalized capacitance data in terms of distinction of flow regimes.

Table 6.1: Comparison of spectrograms of normalized capacitance C27 of flow regimes

Stratified	

Wavy	
Annular	
Plug	



6.1.1.4 Feature Engineering

6.1.1.4.1 Flow Regime Classification

After EDA, the task is to decide on the features to be used for flow regime classification algorithms. Since there is availability of 28 capacitances per frame of ECT data and EDA was performed on these 28 capacitances, these can be used as features. So, after performing EDA on ECT data, it is decided to use 28 normalized capacitances as inputs to machine learning algorithms for flow regime classification. Also, based on box plots and correlation analysis, a case of using only 15 capacitances as features is also considered reducing the inputs by 53%. Both of these cases are considered in the flow classification models to compare their effect on model accuracy. The basic way of feature extraction from ECT data is shown in Figure 6.18. The Python script of Appendix-B is used to achieve this transformation.

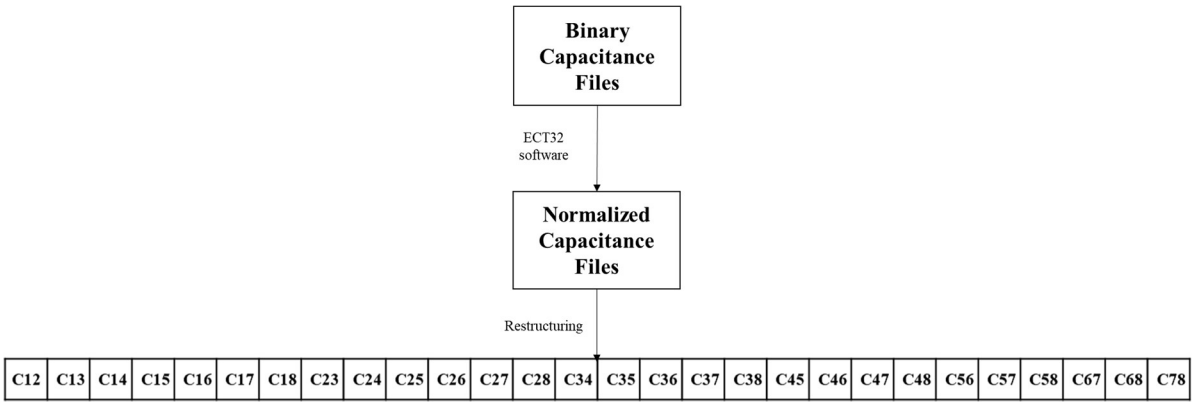


Figure 6.18: Feature Engineering for Flow Regime Classification by using normalized capacitances of ECT

The method to develop the flow regime classification model is illustrated in Figure 6.19. In this model, 28 normalized capacitances from ECT are used as inputs to the machine learning algorithms while flow regimes are used as labels.

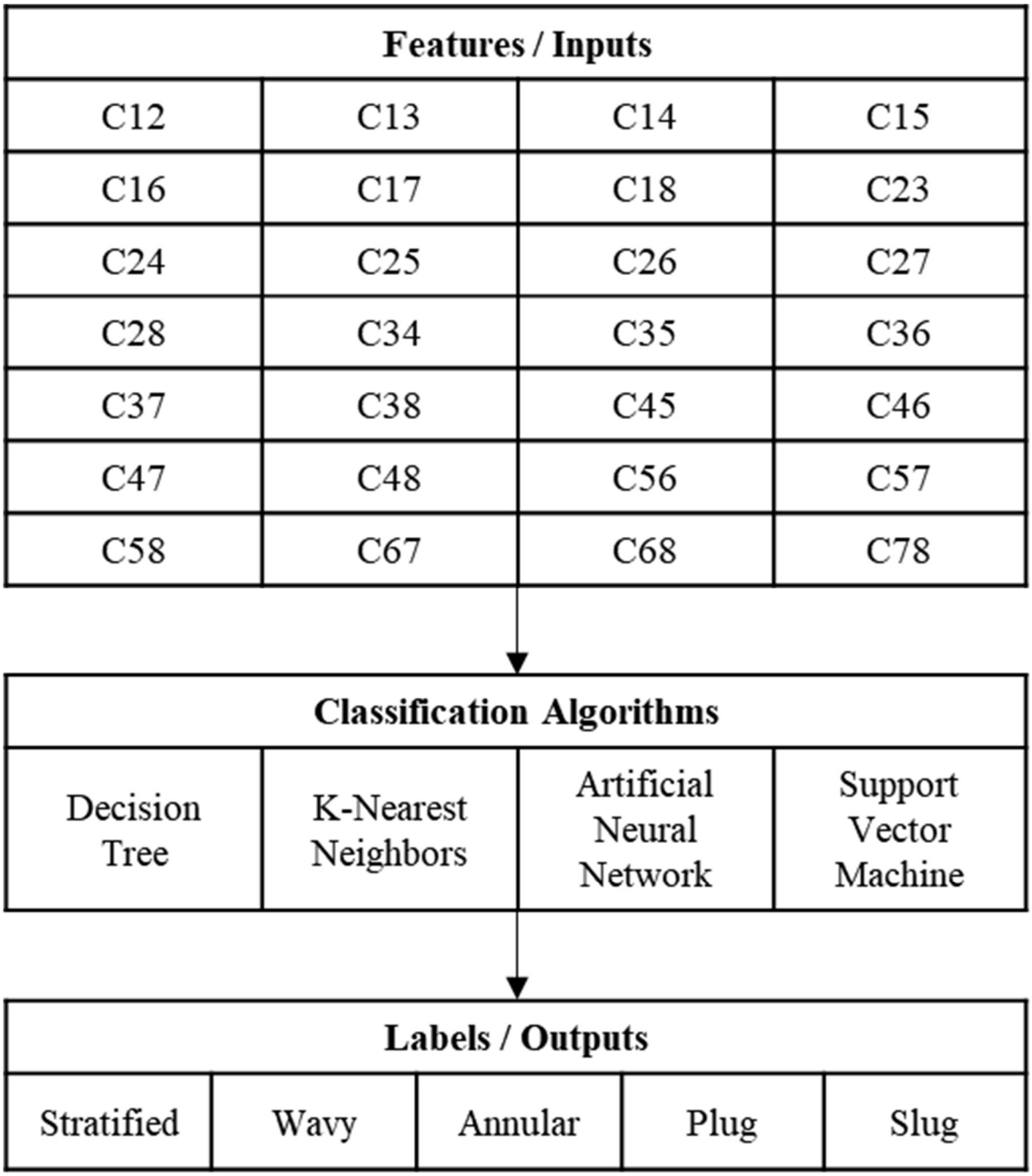


Figure 6.19: Model for flow regime classification using 28 normalized capacitances as features

Another method to classify flow regime is by fusing different types of sensors. This method is generally known as Sensor Fusion. The model for Sensor Fusion concept is shown as Figure 6.20. In this method, 28 normalized capacitances along with pressure sensor signals in mbar are used as inputs to the machine learning algorithms while flow regimes are used as labels. This transformation of features is achieved through a Python script provided in Appendix – D. The pressure and flow sensors data were sampled at 1s while ECT sampling was at 10ms. Therefore, every 100^{th} frame from ECT normalized capacitance file was extracted and clubbed with the pressure and flow sensors data.

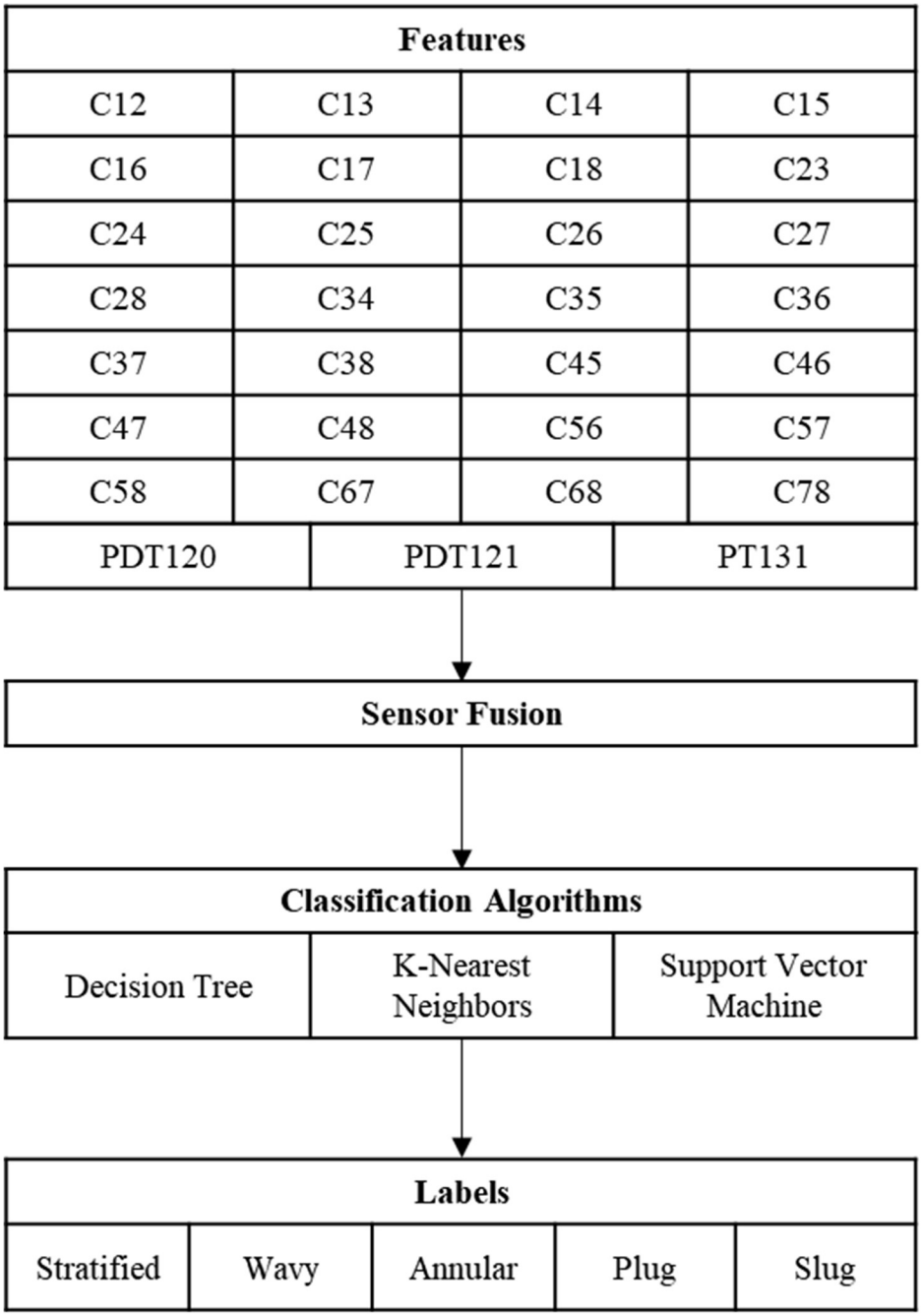


Figure 6.20: Model for flow regime classification using 28 normalized capacitances and pressure signals as features

6.1.1.4.2 Flow Velocity Estimation

Flow Velocity is a dynamic parameter. Prediction of flow velocity is not a classification problem. The applicability of two planes of ECT comes into picture in this case of flow velocity estimation. Since flow velocity involves time, the normalized capacitances from the two planes of ECT can be cross-correlated to find similarities between them. So, if the highest correlation coefficient is found after 4 lags, then the time taken by the flow from one plane to another is corresponding to 4 lags. 1 lag is equal to 10ms, the frame sampling time. So the time taken in 4 lags is 40ms. When the distance between the planes is divided by this time, the approximate

value of flow velocity can be calculated in m/s. The entire concept is shown in Figure 6.21. 28 normalized capacitances are cross-correlated. For instance, C12 of Plane 1 is cross-correlated with C12 of Plane 2 to find the lag at the highest correlation coefficient. The same cross-correlation is repeated with the other 27 capacitances. MATLAB script of Appendix – I is used to analyse cross-correlation of dual plane ECT data.

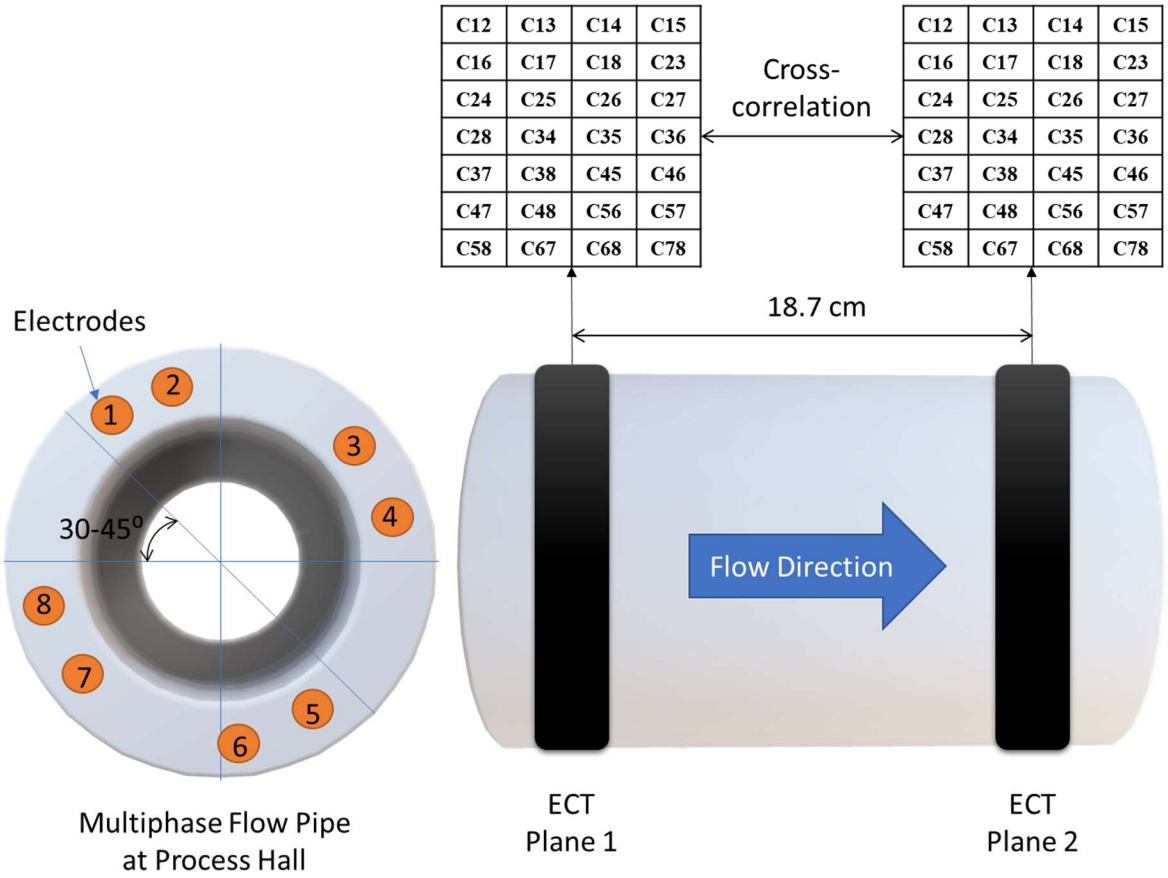


Figure 6.21: Concept of Flow velocity estimation by using cross correlation of dual-plane ECT data

6.1.1.4.3 Volume Ratio Estimation

Volume ratio in this thesis refers to the ratio of cross-sectional area of liquid phase to the total cross-sectional inside area of the pipe. This volume ratio can be estimated by processing images generated from the normalized capacitances of ECT for every frame.

Regression algorithms can be employed to estimate this volume ratio by selecting the features properly. In this thesis, 3 features namely differential pressure from differential pressure sensors PDT120 and PDT121, and input air flow rate from flow meter FT131 are used as inputs to the regression algorithms. The labels are the inverse of volume ratio in percentages extracted from the Volume Ratio 'AVR' files obtained from the normalized capacitances of ECT as shown in Figure 6.22. Python script of Appendix-E is created to bring about this feature transformation. The inverse of volume ratio is used as label of these algorithms as shown in Figure 6.22.

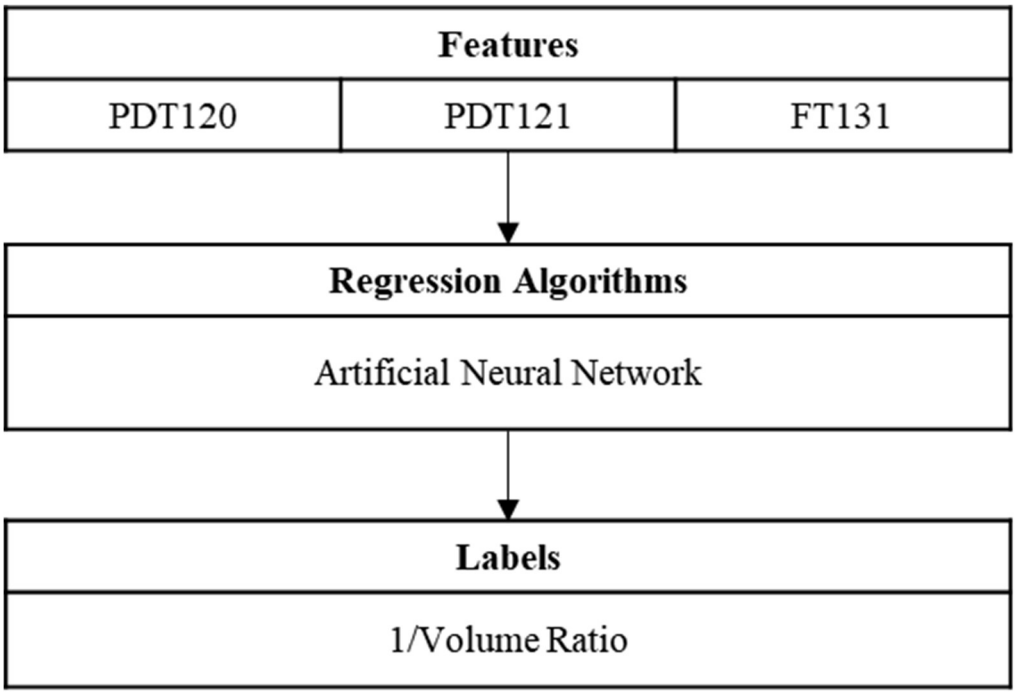


Figure 6.22: Model of volume ratio estimation by using differential pressure meters and a flow meter

6.1.2 Ultrasonic Data Analysis

6.1.2.1 Feature Engineering

The ultrasonic data collected during experiments was in the form of binary files containing 863 A-scans for each experiment of a specific flow regime. Each A-scan consisted of 9596 ultrasonic signals received by the receiver in the pitch-catch technique. Since the quantity of data was huge to be used as input to machine learning algorithm, statistical features were extracted from each A-scan as shown in Figure 6.23. The features of mean, maximum, standard deviation, normalized standard deviation, sum of all the points in the envelope, square of sum of all the points in the envelope were extracted by using the MATLAB script of Appendix-K. By using this technique, the total number of input data points to the machine learning algorithms was reduced by 99.94% for each experiment.

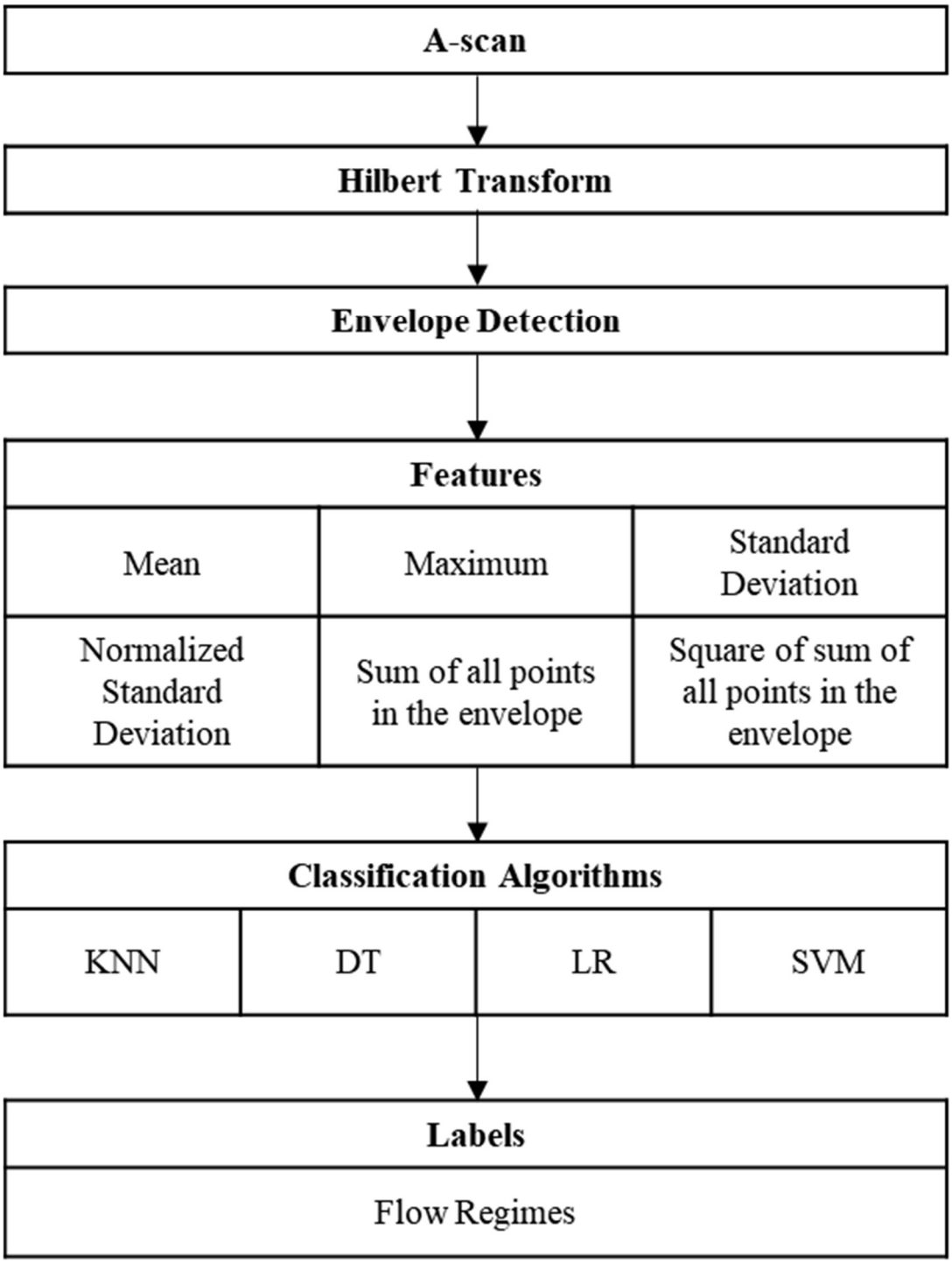


Figure 6.23: Feature Engineering on ultrasonic data for regime classification

6.2 Data-Driven Multiphase Flow Metering Models

6.2.1 ECT Classification and Estimation Models

In this section, models to identify flow regimes and to estimate flow velocity and volume ratio are developed with collected data from ECT by using the features extracted from the data.

6.2.1.1 Flow Regime Identification – Model 1

Flow regime classification is a multi-class classification problem. In this problem, the classification model should be able to identify correct flow regime by looking at the input data provided to it. A flow regime classification model is shown in Figure 6.24 which is based on the feature engineering described in Figure 6.19 in the previous section. The architecture is shown in Table 6.2 which has 28 inputs and 1 output. The idea behind this model is that each frame of 28 normalized capacitances is used as inputs to the model and after training, the model is expected to classify these 28 inputs to a specific flow regime. To train the model 135000 samples containing features and labels are used.

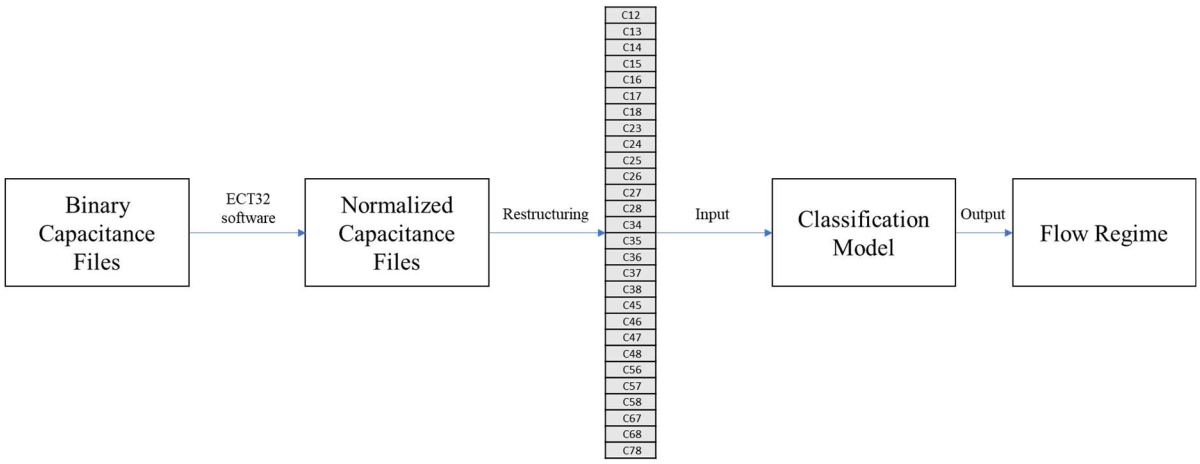


Figure 6.24: Model 1 for flow regime classification using 28 normalized capacitances as inputs and flow regime as output

Table 6.2: Architecture for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

Architecture No.	**1**	
Samples	135000 rows	29 columns
Inputs / Features	28	C12, C13, C14, C15, C16, C17, C18, C23, C24, C25, C26, C27, C28, C34, C35, C36, C37, C38, C45, C46, C47, C48, C56, C57, C58, C67, C68, C78

Outputs / Labels	1	Flow Regime (1 - Stratified, 2 - Wavy, 3 - Annular, 4 - Slug, 5 - Plug)

6.2.1.1.1 Neural Network

The first model developed for this multi-class classification problem is an Artificial Neural Network in MATLAB. MATLAB's Neural Net Pattern Recognition App is used to develop a feed forward neural network (NN) with one hidden layer containing 10 neurons. The details of the neural network are provided in Table 6.3. The architecture of the NN from MATLAB is shown in Figure 6.25.

The type of Neural Network used is Feedforward Neural Network with 28 inputs and 5 outputs. There are five outputs because MATLAB's Neural Net Pattern Recognition App takes nominal data type as inputs to the App. There is one hidden layer with 10 neurons and an output layer that provides a nominal output in the form of 1s and 0s. The 135000 samples are divided into training, validation and testing data in the ratio 70:15:15. The activation function used in the hidden layer is *tansig*.

Table 6.3: Details of Neural Network for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

Type	Feedforward
Number of inputs	28
Input data type	Float
Number of outputs	5
Output data type	Nominal
Number of hidden layers	1
Number of neurons in hidden layer	10
Training: Validation: Testing Ratio	70:15:15
Activation Functions in Hidden Layer	tansig

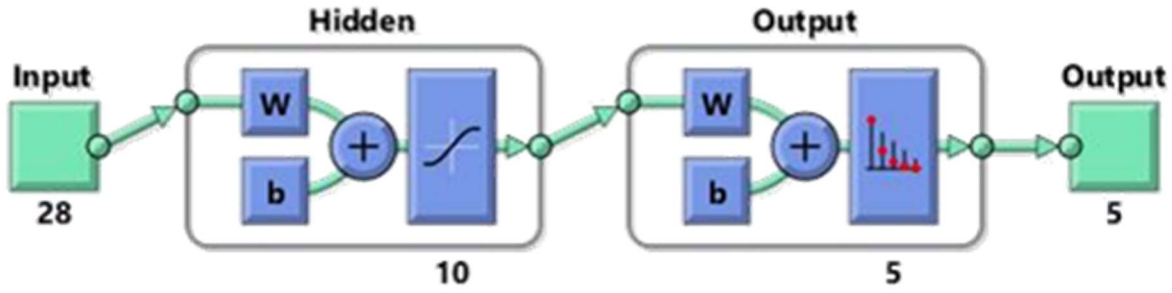


Figure 6.25: Neural Network Architecture in MATLAB for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

After setting up the parameters of the neural network, the neural network is trained. The algorithms used in training and the training progress is shown in Figure 6.26. It was trained in 336 iterations.

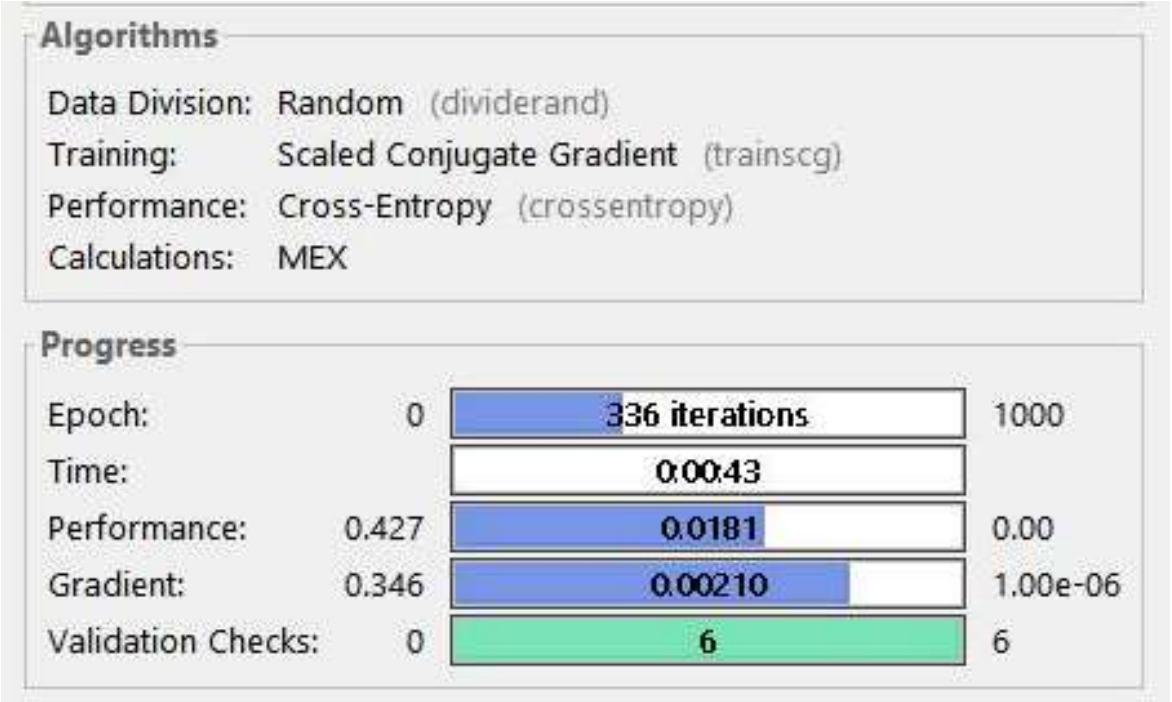


Figure 6.26: Training of NN for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

The confusion matrices for training, validation and testing are shown in Figure 6.27. It can be observed that Stratified, Annular and Wavy regimes have a classification accuracy of more than 99.6% for test data set. The overall testing accuracy is 96.6%. The training performance is shown in Figure 6.28. The accuracies obtained from the NN are summarized in Table 6.4.

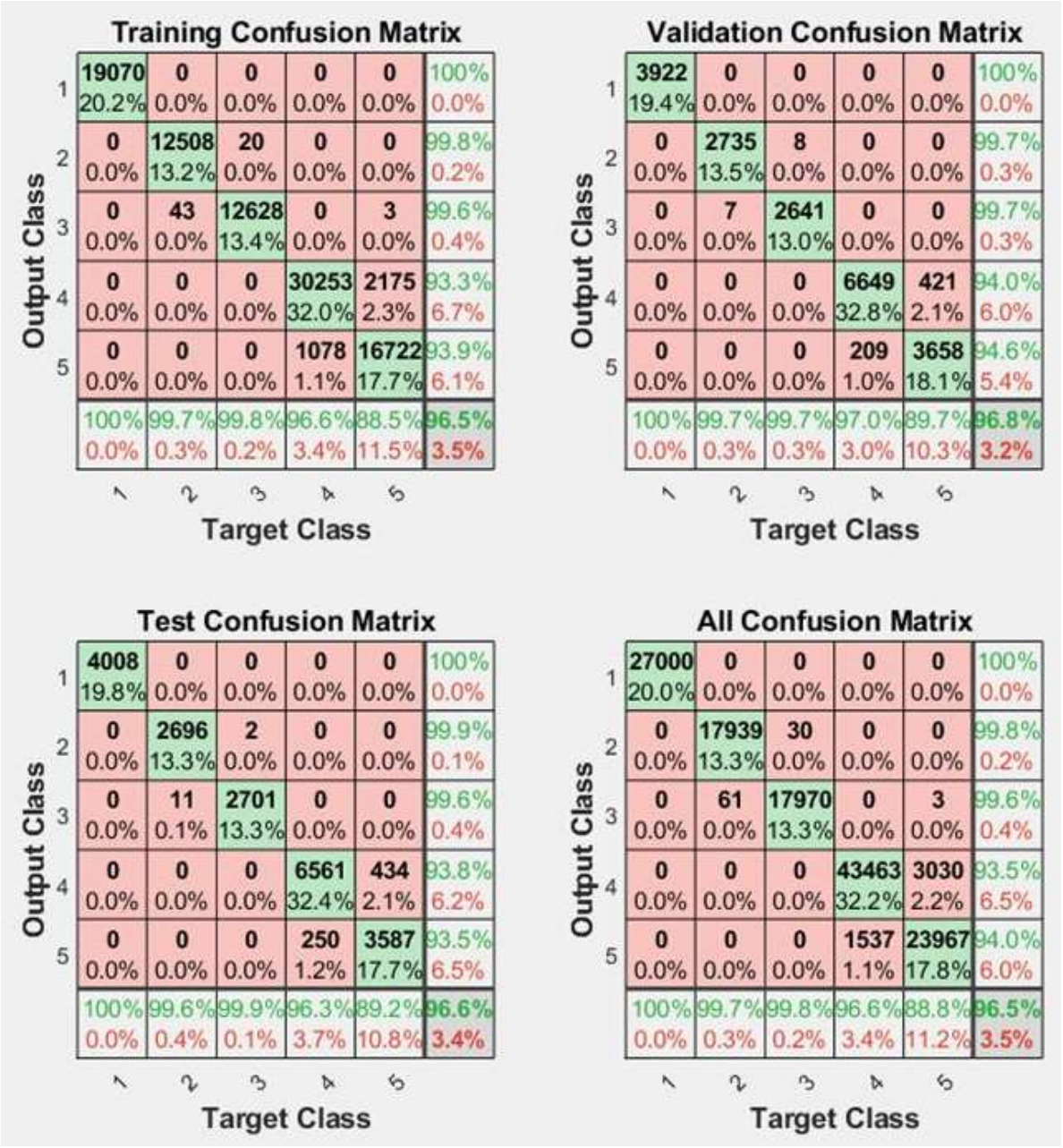


Figure 6.27: Confusion Matrices of NN for flow regime identification with 28 normalized capacitances as inputs and flow regime as output (1-Stratified, 2-Wavy, 3-Annular, 4-Plug, 5-Slug)

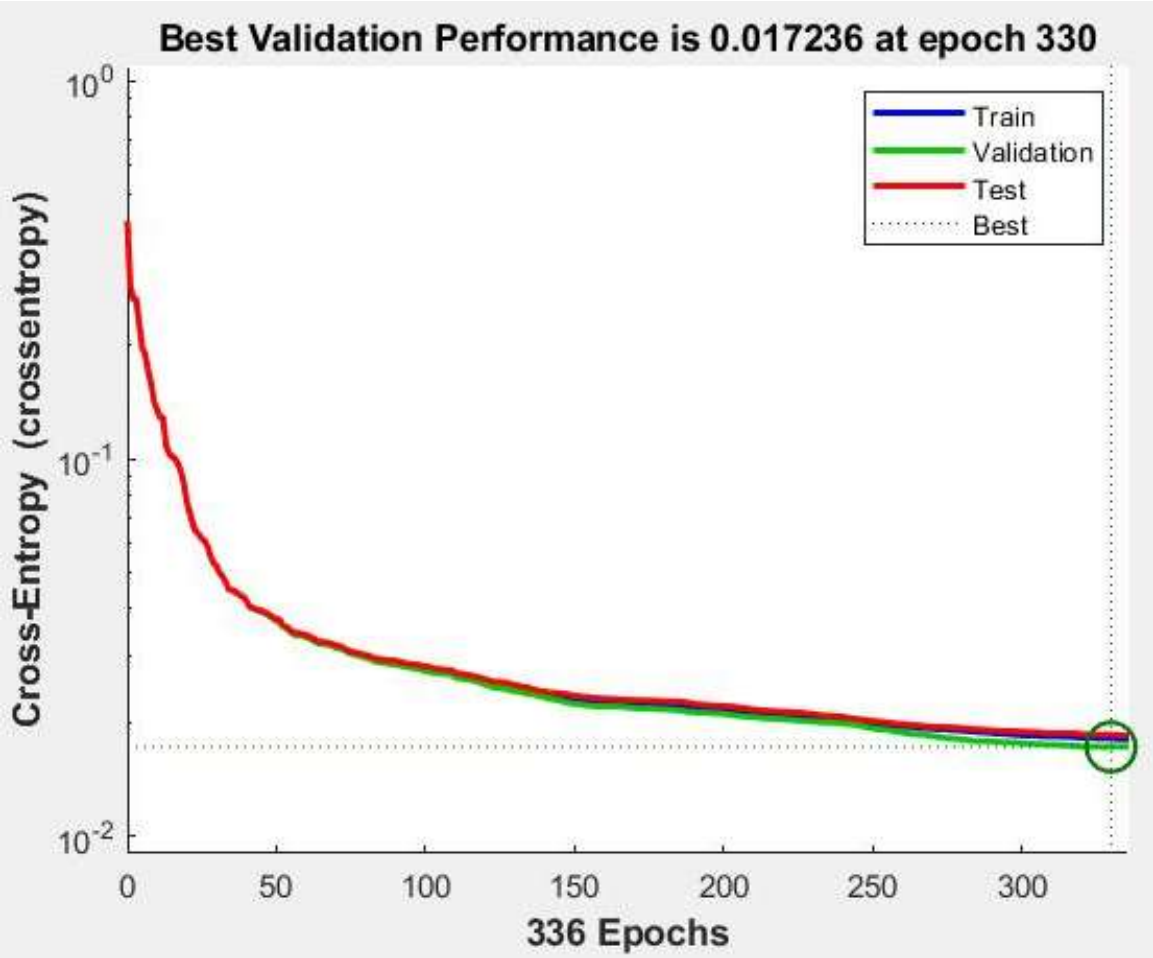


Figure 6.28: Training performance of the NN for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

Table 6.4: Results of the NN model for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

Results	**Accuracy**
Training Accuracy	96.5%
Validation Accuracy	96.8%
Testing Accuracy	96.6%
Overall Accuracy	96.5%

6.2.1.1.2 Results from Other Machine Learning Models

In the previous section, a flow regime classification model using Neural Network was trained. Other than Neural Network, the Classification Learner App of MATLAB is employed to develop a plethora of machine learning models for flow regime classification. KNN, SVM,

Decision Tree (DT) and Linear Discriminant (LD) algorithms are used to train machine learning models for flow regime classification. 25% of the data was kept aside for validation and 75% was used for training. The overall validation accuracy from these algorithms to classify flow regime is given in Figure 6.29. KNN gives the highest accuracy while LD has the lowest accuracy.

1 ☆ Tree	Accuracy: 96.6%	
Last change: Disabled PCA	28/28 features	
2 ☆ KNN	Accuracy: **98.8%**	
Last change: Fine KNN	28/28 features	
3 ☆ Linear Discriminant	Accuracy: 92.7%	
Last change: Linear Discriminant	28/28 features	
4 ☆ SVM	Accuracy: 94.8%	
Last change: Linear SVM	28/28 features	

Figure 6.29: Comparison of accuracy from various algorithms for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

The validation confusion matrices of KNN model is provided in Figure 6.30. It is able to classify Annular, Stratified and Wavy regimes perfectly. Slug and Plug regimes have a low percentage of misclassification.

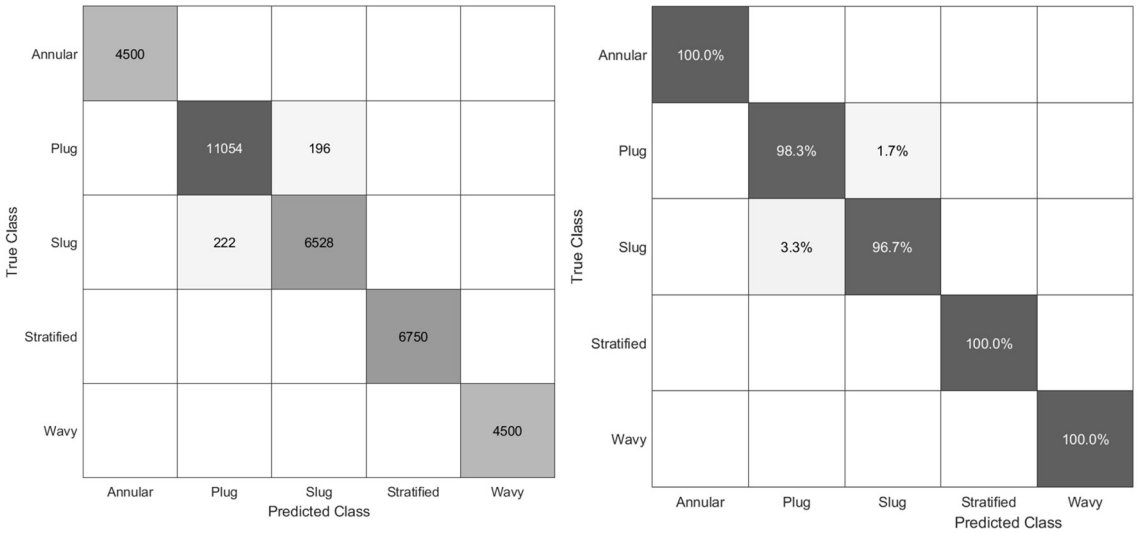


Figure 6.30: Confusion Matrices of KNN (Validation) for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

Table 6.5 provides details of parameters of the machine learning algorithms to classify flow regimes. LD and DT are the fastest algorithms with good accuracy for flow regime classification.

Table 6.5: Comparison of parameters of machine learning models for flow regime identification with 28 normalized capacitances as inputs and flow regime as output

Algorithm	**Accuracy (%)**	**Total Misclassification cost**	**Prediction speed (observations/s)**	**Training time (s)**	**Other details**
KNN	98.8%	1084	1400	131.38	Fine KNN, 1 neighbor
DT	96.6%	2322	1500000	1.66	Fine Tree, Maximum number of splits: 100
LD	92.7%	4812	570000	1.3	Full covariance
SVM	94.8%	3523	180000	119.4	Linear Kernel, one-vs-one multiclass method

6.2.1.2 Flow Regime Identification – Model 2

After using only ECT data to classify flow regimes, another model is created based on Sensor Fusion concept in this section. The data from pressure sensors including differential pressure sensors are clubbed with the normalized capacitances from ECT to develop a flow regime classification model. This model is tested to observe any performance improvements in flow regime model of the previous section. The model's framework is shown in Figure 6.31. The detailed architecture is provided in Table 6.6. The architecture has 31 inputs and 1 output.

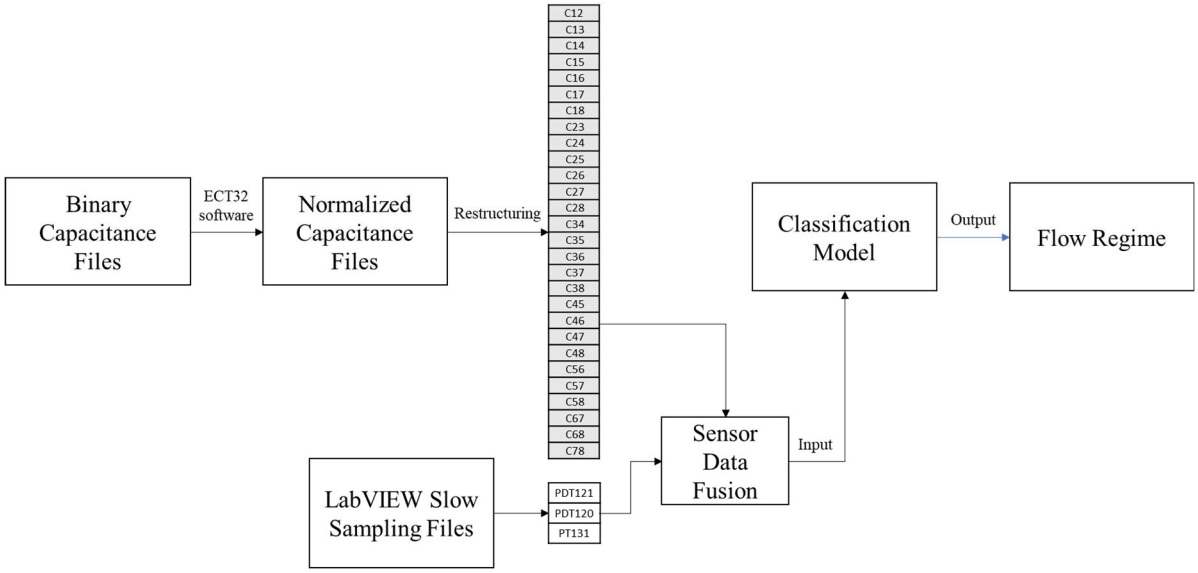


Figure 6.31: Sensor Fusion Model for flow regime identification for flow regime identification with 28 normalized capacitances and pressure sensors as inputs and flow regime as output

Table 6.6: Architecture for flow regime identification for flow regime identification with 28 normalized capacitances, PDT120, PDT121 and PT131 as inputs and flow regime as output

Architecture No.	**2**	
Samples	1350 rows	32 columns
Inputs	31	C12,C13,C14,C15,C16,C17,C18,C23,C24,C25,C26,C27,C28,C34,C35,C36,C37,C38,C45,C46,C47,C48,C56,C57,C58,C67,C68,C78,PDT120,PDT121,PT131
Outputs	1	Flow Regime (1 - Stratified, 2 - Wavy, 3 - Annular, 4 - Slug, 5 - Plug)

MATLAB's Classification Learner App is used to train this model. There are 1350 samples for inputs. The performance of the all the algorithms improves drastically when all 31 features are considered as shown in Figure 6.32. In this case, SVM performs better than other algorithms with a 5-fold cross-validation accuracy of 99.1%. The confusion matrix for SVM is given in Figure 6.33.

1 ☆ Tree	Accuracy: 98.5%
Last change: Fine Tree	31/31 features
2 ☆ SVM	Accuracy: **99.1%**
Last change: Linear SVM	31/31 features
3 ☆ KNN	Accuracy: 98.4%
Last change: Fine KNN	31/31 features
4 ☆ Linear Discriminant	Accuracy: 97.7%
Last change: Linear Discriminant	31/31 features

Figure 6.32: Performance of algorithms for flow regime identification with 28 normalized capacitances, PDT120, PDT121 and PT131 as inputs and flow regime as output

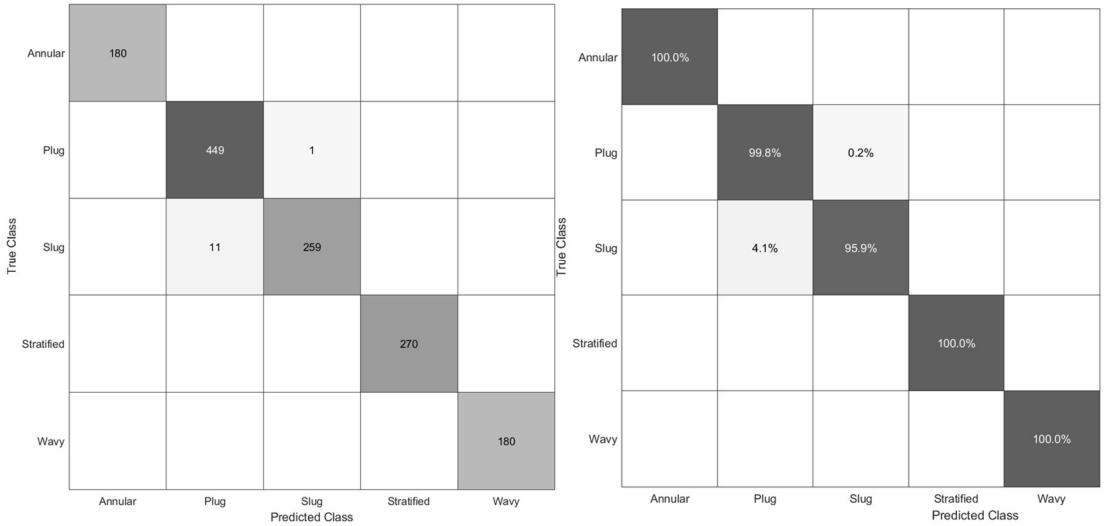


Figure 6.33: Confusion matrices of SVM for flow regime identification with 28 normalized capacitances, PDT120, PDT121 and PT131 as inputs and flow regime as output

To test the impact of only pressure sensors on the flow regime classification, a model with only 3 inputs from pressure sensors is developed. By using PT131, PDT121 and PDT120 as inputs, 96.2% accuracy was observed in KNN algorithm as shown in Figure 6.34 and its confusion matrices are shown in Figure 6.35. It is observed here that the accuracy is less than the previous models with ECT and with sensor fusion. Therefore, it can be said that sensor fusion improves flow regime classification accuracy.

5 ☆ SVM	Accuracy: 90.7%	
Last change: Disabled PCA	3/31 features	
6 ☆ Tree	Accuracy: 94.7%	
Last change: Fine Tree	3/31 features	
7 ☆ Linear Discriminant	Accuracy: 75.9%	
Last change: Linear Discriminant	3/31 features	
8 ☆ KNN	Accuracy: 96.2%	
Last change: Fine KNN	3/31 features	

Figure 6.34: Performance of algorithms for flow regime identification with 3 pressure sensors as inputs and flow regime as output

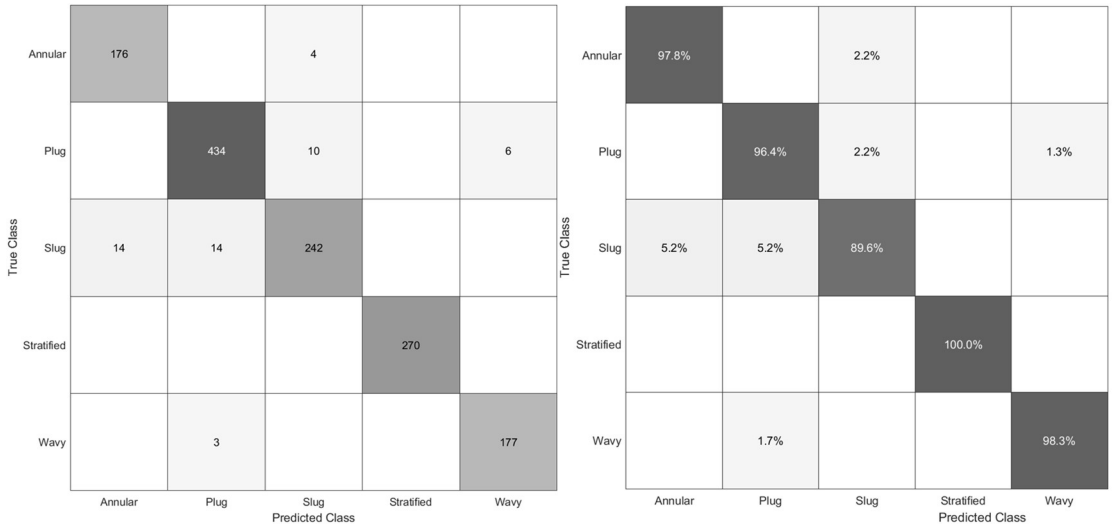


Figure 6.35: Confusion Matrices for KNN for flow regime identification with pressure sensors as inputs and flow regime as output

6.2.1.3 Flow Regime Identification – Model 3

In the previous section, samples as huge as 135000x29 in number were used to train machine learning models. This huge number of samples is not suitable for all the computing machines such as microcontrollers and wherever possible a lesser number of samples is preferred. During EDA, reduction of samples was discussed based on box plots and correlation analysis. Based on the reduction strategy, a model is proposed as shown in Figure 6.36 where input features are reduced to 15 normalized capacitances based on the similarity of distribution observed in normalized capacitances. The architecture is given in Table 6.7. The architecture has 15 inputs and 1 output.

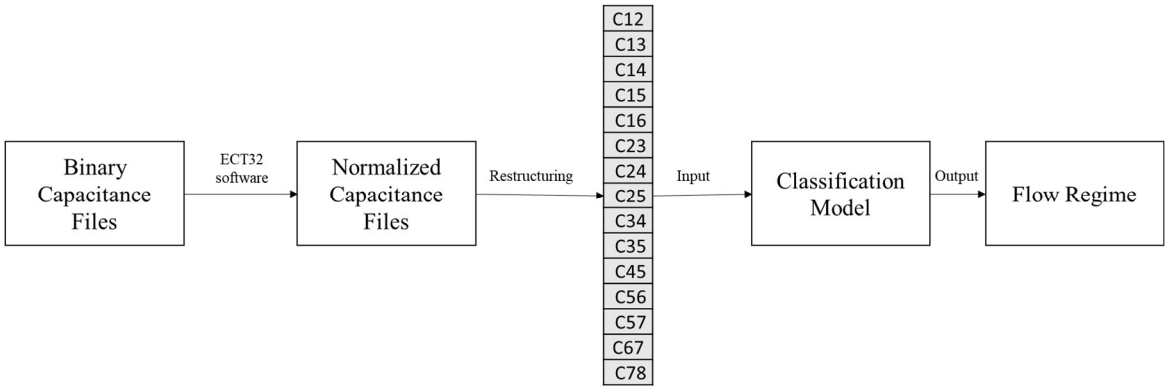


Figure 6.36: Model with reduced features through exploratory data analysis

Table 6.7: Architecture for for flow regime identification with 15 normalized capacitances as inputs and flow regime as output

Architecture No.	**3**	
Samples	135000 rows	16 columns
Inputs	15	C12,C13,C14,C15,C16,C23,C24,C25,C34,C35,C45,C56,C57,C67,C78
Outputs	1	Flow Regime (1 - Stratified, 2 - Wavy, 3 - Annular, 4 - Slug, 5 - Plug)

In this model, the performance of machine learning algorithms is compared by considering only 15 inputs based on the similarity of normalized capacitances seen in box plots and correlation analysis. In this case, KNN performs better than other algorithms with a 25% validation accuracy of 98.5% as seen in Figure 6.37. The confusion matrices for KNN are given in Figure 6.38. Good accuracy is still achieved after reduction of features.

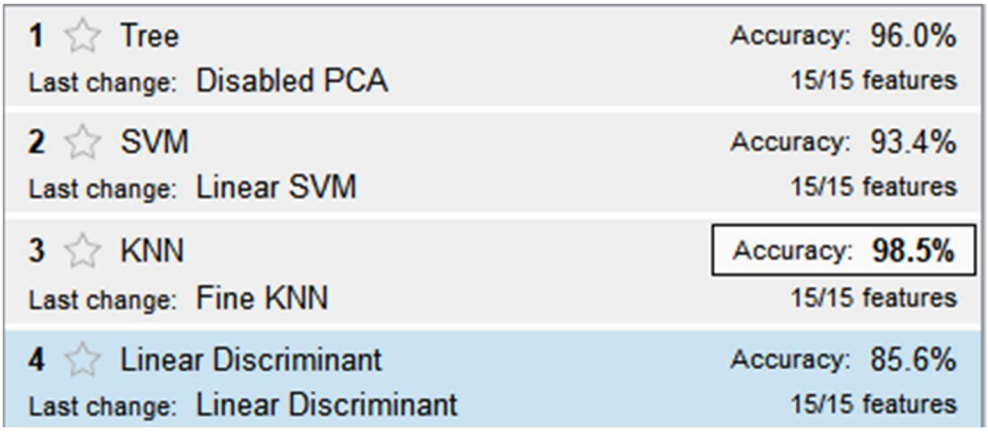


Figure 6.37: Performance of machine learning algorithms for flow regime identification with 15 normalized capacitances as inputs and flow regime as output

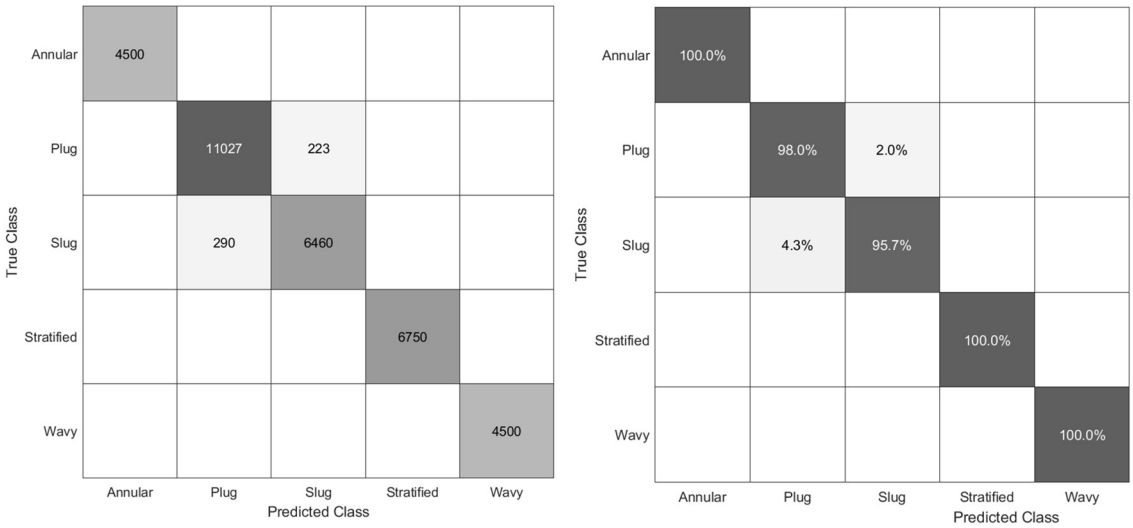


Figure 6.38: Confusion matrices of KNN algorithm for flow regime identification with 15 normalized capacitances as inputs and flow regime as output

6.2.1.4 Flow Velocity Estimation

In this section, a model based on Figure 6.21 to estimate flow velocity through cross correlation techniques is presented. The normalized capacitances at the two planes of ECT are cross-correlated to find the peak correlation between them. For instance, the series of capacitances between electrode 1 and 2 at plane 1 is correlated with the series of capacitances between electrode 1 and 2 at plane 2 of ECT. The peak corresponds to the degree of similarity between the two capacitances. Only Annular, Plug and Slug flow regimes are considered for this model because of their dynamic nature of flow. The distance between the two planes is 0.187 m. The flow velocity is calculated by dividing the distance between the planes by the lag time. The time per lag is 10ms.

6.2.1.4.1 Annular

6.2.1.4.1.1 Case 1: Water - 2 kg/min, Air – 4 kg/min

From Table 6.8, it is observed that the cross correlation lag for 8 capacitances is approximately 120 to 130ms. Therefore, the flow velocity is estimated to be 0.187/0.13 = 1.43 m/s for this experiment. Higher lags are disregarded as they likely arise from random fluctuations.

Table 6.8: Cross correlation for Water - 2 kg/min, Air – 4 kg/min

Capacitance	Peak Correlation Lag	Capacitance	Peak Correlation Lag
C12	0	C35	0
C13	0	C36	-12
C14	0	C37	-1664
C15	0	C38	-2992

C16	-13	C45	0
C17	-13	C46	-13
C18	-13	C47	-2992
C23	-175	C48	-2992
C24	0	C56	-35
C25	0	C57	-2787
C26	-12	C58	2997
C27	-2993	C67	-13
C28	-2963	C68	-13
C34	0	C78	0

6.2.1.4.1.2 Case 2: Water - 2 kg/min, Air – 5 kg/min

From Table 6.9, it is observed that the cross correlation lag for 10 capacitances is approximately 90 to 110ms. Therefore, the flow velocity is estimated to be 0.187/0.10 = 1.87 m/s for this experiment. Higher lags are disregarded as they likely arise from random fluctuations.

Table 6.9: Cross correlation for Water - 2 kg/min, Air – 5 kg/min

Capacitance	**Peak Correlation Lag**	**Capacitance**	**Peak Correlation Lag**
C12	0	C35	0
C13	0	C36	-9
C14	0	C37	-2994
C15	-6	C38	-2997
C16	-10	C45	-1
C17	-10	C46	2997
C18	-10	C47	-2997
C23	-2997	C48	-2996
C24	0	C56	-9
C25	0	C57	-10
C26	2860	C58	-9
C27	-2994	C67	-10
C28	2997	C68	-10
C34	0	C78	-11

6.2.1.4.2 Plug

6.2.1.4.2.1 Case 1: Water - 76 kg/min, Air – 0.11 kg/min

From Table 6.10, it is observed that the cross correlation lag for 28 capacitances is approximately 100 to 130ms. Therefore, the flow velocity is estimated to be 0.187/0.13 = 1.43 m/s for this experiment.

Table 6.10: Cross correlation for Water - 76 kg/min, Air – 0.11 kg/min

Capacitance	**Peak Correlation Lag**	**Capacitance**	**Peak Correlation Lag**
C12	-13	C35	-12
C13	-12	C36	-12
C14	-10	C37	-12
C15	-12	C38	-12
C16	-11	C45	-10
C17	-11	C46	-10
C18	-11	C47	-10
C23	-11	C48	-10
C24	-10	C56	-12
C25	-12	C57	-12
C26	-12	C58	-12
C27	-12	C67	-10
C28	-12	C68	-10
C34	-11	C78	-11

6.2.1.4.2.2 Case 2: Water - 77 kg/min, Air – 0.07 kg/min

From Table 6.11, it is observed that the cross correlation lag for 23 capacitances is approximately 100 to 150ms. Therefore, the flow velocity is estimated to be 0.187/0.15 = 1.24 m/s for this experiment. Lower lags are disregarded as they likely arise from random fluctuations.

Table 6.11: Cross correlation for Water - 77 kg/min, Air – 0.07 kg/min

Capacitance	**Peak Correlation Lag**	**Capacitance**	**Peak Correlation Lag**
C12	-15	C35	-15

C13	-15	C36	-15
C14	-12	C37	-15
C15	-15	C38	-15
C16	-1	C45	-13
C17	-2	C46	-13
C18	-3	C47	-12
C23	-14	C48	-12
C24	-12	C56	-14
C25	-14	C57	-2
C26	-14	C58	0
C27	-14	C67	-10
C28	-14	C68	-10
C34	-13	C78	-11

6.2.1.4.3 Slug

6.2.1.4.3.1 Case 1: Water - 75 kg/min, Air – 0.3 kg/min

From Table 6.12, it is observed that the cross correlation lag for 28 capacitances is approximately 50 to 70ms. Therefore, the flow velocity is estimated to be 0.187/0.07 = 2.67 m/s for this experiment.

Table 6.12: Cross correlation for Water - 75 kg/min, Air – 0.3 kg/min

Capacitance	**Peak Correlation Lag**	**Capacitance**	**Peak Correlation Lag**
C12	-6	C35	-6
C13	-6	C36	-6
C14	-6	C37	-6
C15	-6	C38	-6
C16	-6	C45	-6
C17	-6	C46	-6
C18	-6	C47	-6
C23	-5	C48	-6
C24	-6	C56	-7
C25	-6	C57	-7
C26	-6	C58	-7

C27	-6	C67	-5
C28	-6	C68	-5
C34	-6	C78	-6

6.2.1.4.3.2 Case 2: Water - 77 kg/min, Air – 0.5 kg/min

From Table 6.13, it is observed that the cross correlation lag for 28 capacitances is approximately 30 to 50ms. Therefore, the flow velocity is estimated to be 0.187/0.05 = 3.74 m/s for this experiment.

Table 6.13: Cross correlation for Water - 77 kg/min, Air – 0.5 kg/min

Capacitance	**Peak Correlation Lag**	**Capacitance**	**Peak Correlation Lag**
C12	-5	C35	-4
C13	-4	C36	-4
C14	-4	C37	-4
C15	-4	C38	-4
C16	-4	C45	-4
C17	-4	C46	-5
C18	-4	C47	-4
C23	-3	C48	-4
C24	-4	C56	-4
C25	-4	C57	-4
C26	-5	C58	-4
C27	-5	C67	-4
C28	-5	C68	-4
C34	-5	C78	-4

6.2.1.5 Volume Ratio Estimation

To estimate volume ratio (VR), differential pressure meters and a flow meter are considered as inputs. The volume ratio data obtained from ECT are considered as target variables. The architecture is shown in Table 6.14. PDT120, PDT121 and FT131 are taken as inputs to the NN. The visual display of the architecture is shown in Figure 6.39. The details of the NN are given in Table 6.15.

Table 6.14: NN architecture for Volume Ratio Estimation with PDT120, PDT121 and FT131 as inputs and 1/Volume Ratio as output

Architecture No.	**4**	
Samples	1350 rows	4 columns
Inputs	3	PDT120, PDT121, FT131
Outputs	1	Total Volume/Volume of liquid

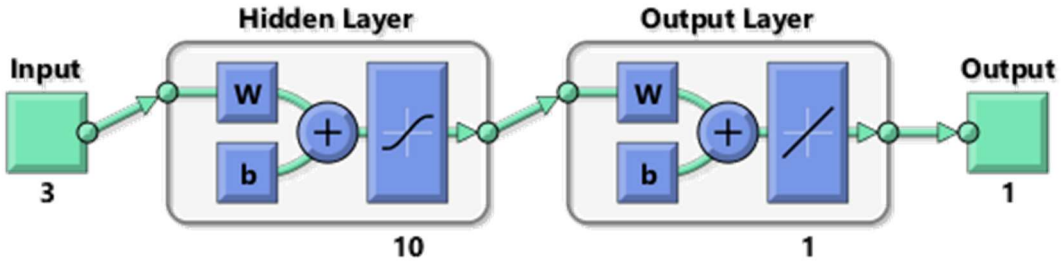


Figure 6.39: Architecture of NN for for Volume Ratio Estimation with PDT120, PDT121 and FT131 as inputs and 1/Volume Ratio as output

Table 6.15: Details of NN for Volume Ratio Estimation with PDT120, PDT121 and FT131 as inputs and 1/Volume Ratio as output

Type	Feedforward
Number of inputs	3
Input data type	Float
Number of outputs	1
Output data type	Float
Number of hidden layers	1
Number of neurons in hidden layer	10
Training: Validation: Testing Ratio	70:15:15
Activation Functions	tansig

The results of the regression NN is shown in Figure 6.40. The model obtained an R value of 0.95 during testing and an overall R value of 0.95 depicting a good fit. The performance curve is given in Figure 6.41.

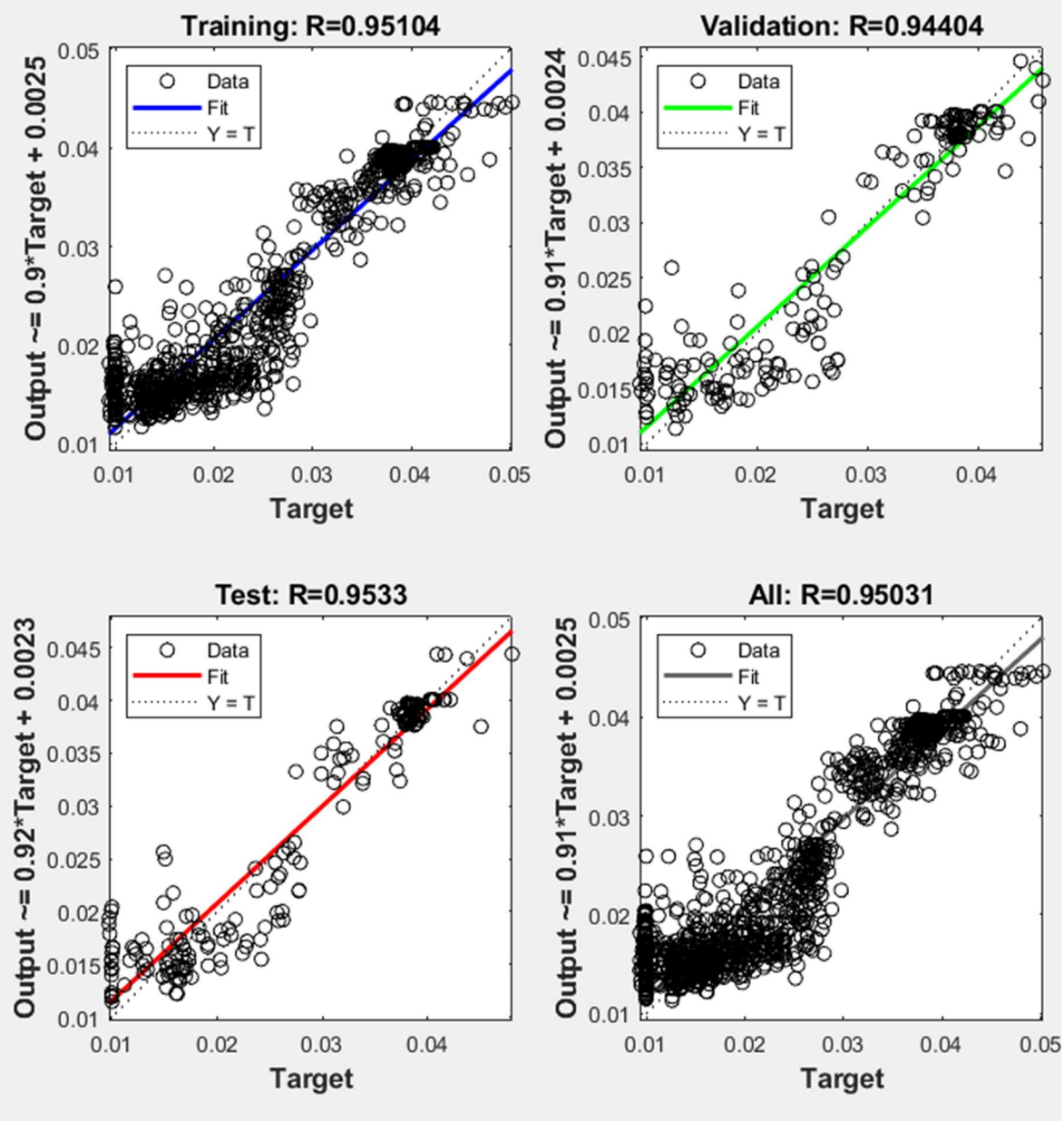


Figure 6.40: Performance of NN for Volume Ratio Estimation with PDT120, PDT121 and FT131 as inputs and 1/Volume Ratio as output

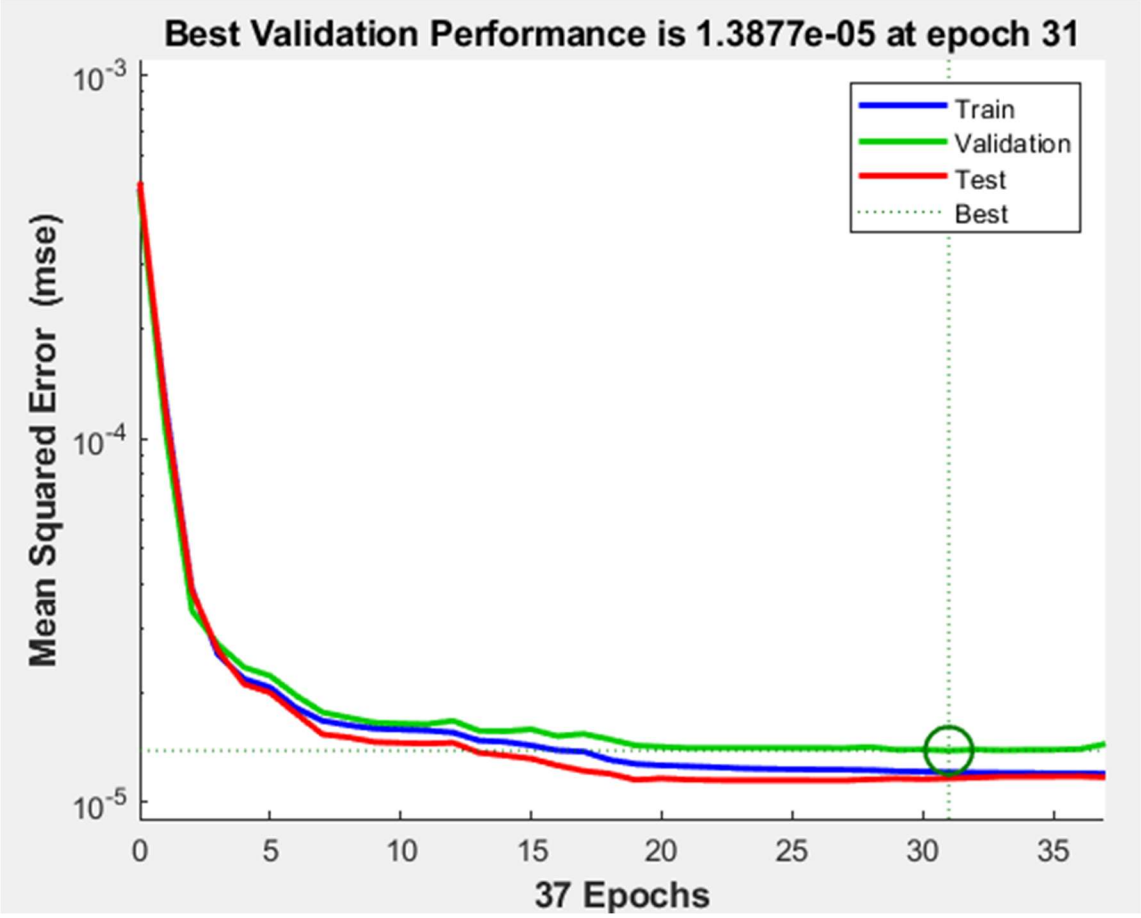


Figure 6.41: Performance of NN for Volume Ratio Estimation with PDT120, PDT121 and FT131 as inputs and 1/Volume Ratio as output

6.2.2 Ultrasonic Classification Models

In this section, models to identify flow regimes are developed with available data of the US sensor. Figure 6.23 provides the sequence to extract features from the available US data. The six extracted features are used to build flow regime classification models in this section.

6.2.2.1 Flow Regime Identification – Model 1

A binary flow regime classification model to classify flow regimes into stratified or annular regime is trained in this section. Its architecture is provided in Table 6.16. The architecture has 6 inputs and 1 output.

Table 6.16: Machine Learning architecture for flow regime classification with US features as inputs and binary flow regime as output

Architecture No.	**1**	
Samples	6904 rows	7 columns
Inputs	6	Mean of envelope, Maximum of envelope, Standard Deviation of envelope, Normalized Standard

		Deviation of envelope, Sum of all the points in the envelope, Square of sum of all the points in the envelope
Outputs	1	Flow Regime – Stratified or Annular

The model is developed in MATLAB Classification Learner App. The total number of samples used in the model is 6904. 15% of the samples were held out for validation. These samples are divided equally into stratified and annular flow regime. The model is trained using DT, LD, LR, SVM and KNN.

For this binary classification problem, the performance of the models is provided in Figure 6.42. Logistic Regression provides the best classification accuracy of 98.6% (validation).

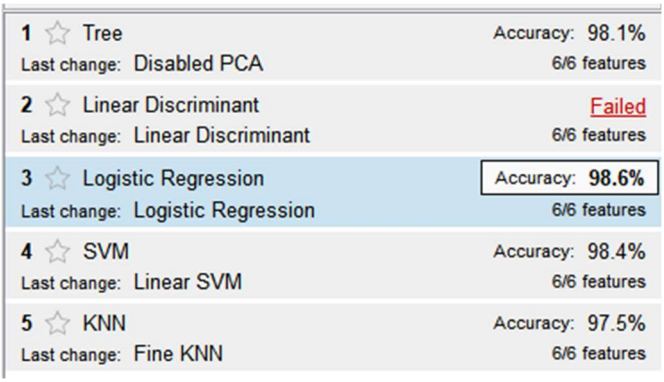


Figure 6.42: Performance of Machine Learning algorithms for flow regime classification with US features as inputs and binary flow regime as output

The confusion matrices for Logistic Regression algorithm are shown in Figure 6.43. This model is accurately able to distinguish between stratified and annular regimes.

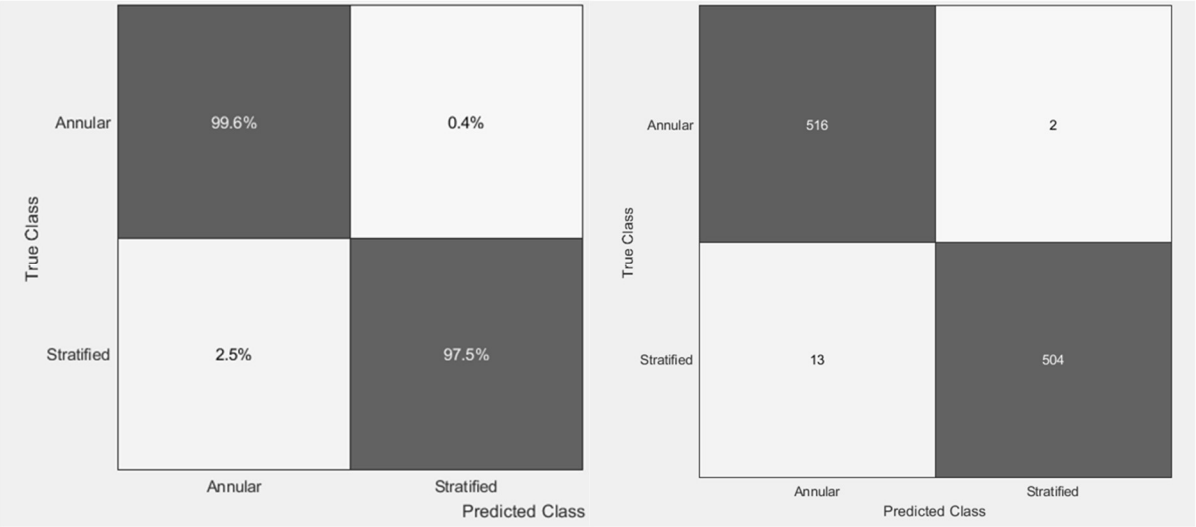


Figure 6.43: Confusion matrices of Logistic Regression algorithm for flow regime classification with US features as inputs and binary flow regime as output

6.2.2.2 Flow Regime Identification – Model 2

A multi-class classification model to classify five flow regimes is developed in this section. Its architecture is shown in Table 6.17. The architecture has 6 inputs and 1 output.

Table 6.17: Machine Learning architecture for flow regime classification with US features as inputs and 1 of 5 flow regimes as output

Architecture No.	**2**	
Samples	4315 rows	7 columns
Inputs	6	Mean of envelope, Maximum of envelope, Standard Deviation of envelope, Normalized Standard Deviation of envelope, Sum of all the points in the envelope, Square of sum of all the points in the envelope
Outputs	1	Flow Regime (1 - Stratified, 2 - Wavy, 3 - Annular, 4 - Slug, 5 - Plug)

MATLAB's Classification Learner App is again used here to develop the multi-class classification problem. The model has 4315 samples. 15% of the samples are held out for validation. The algorithms of DT, SVM and Ensemble Learning are used to develop this model. The performance of the models is shown in Figure 6.44. The accuracy for this model was found to be lesser than the binary classification model of the previous section. Ensemble Bagged Trees provided the best performance with 70.6% accuracy (validation).

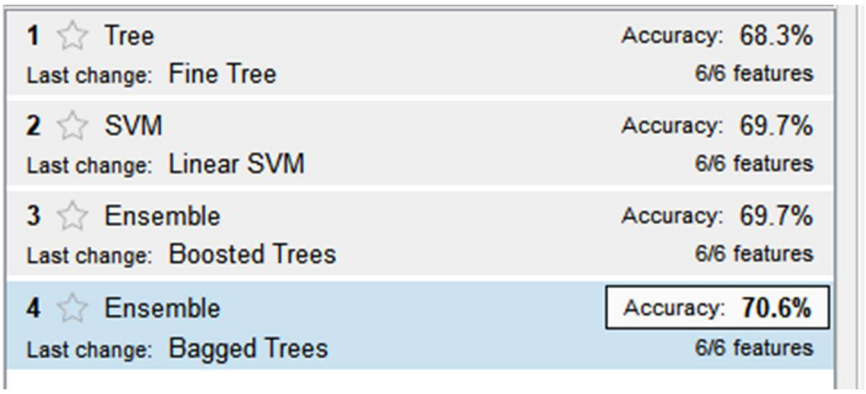


Figure 6.44: Performance of Machine Learning models for flow regime classification with US features as inputs and flow regimes as output

The confusion matrices of Ensemble Bagged Trees are given in Figure 6.45. The accuracy for stratified, annular, wavy, plug and slug are 93%, 89.1%, 71.5%, 51.5% and 48.1% respectively. As in binary classification model, this model is able to identify stratified and annular fairly well. When the interface is moving as in wavy, plug and slug regimes, the model is not able to classify the regimes accurately.

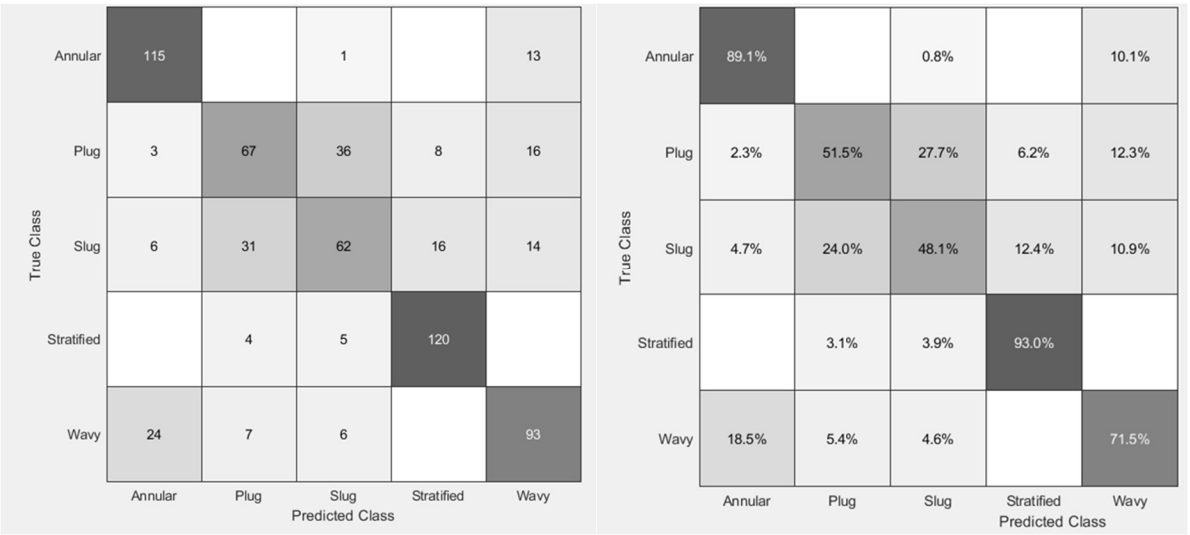


Figure 6.45: Confusion matrices of Ensemble learning algorithm for flow regime classification with US features as inputs and flow regimes as output

6.3 Flow Visualization with Ultrasonic Sensing and ECT

6.3.1 Image Processing from Ultrasonic Data

Ultrasonic data was collected as A-scan amplitude data with 9596 echoes. This collection of A-scan was repeated 863 times to generate a B-scan image of a particular flow regime. The interface between air and water can be investigated in the B-scan image.

For ultrasonic imaging, ultrasound transmitter moves around a specimen to generate a B-scan of that specimen. Since flow is moving in nature, the ultrasonic sensor can be placed at one point on a pipe to get a B-scan image of the flow dynamics as shown in Figure 6.46.

A sample of A-scan in the existing data contains 9596 received ultrasonic signals. For a B-scan, 863 samples are collected for each experiment. The transmitter sends ultrasound waves into the pipe. The ultrasound waves are reflected at the interface of air and water, and at the top of the pipe. The reflected ultrasound waves are logged by the receiver in a binary file. The time of flight is the time taken by the ultrasound waves to travel Path A and B. The transmitter sends the ultrasound wave. The echoes are captured by the receiver. This one line of data of reflected echoes is known as A-scan. A collection of A-scan data is known as B-scan. [32]

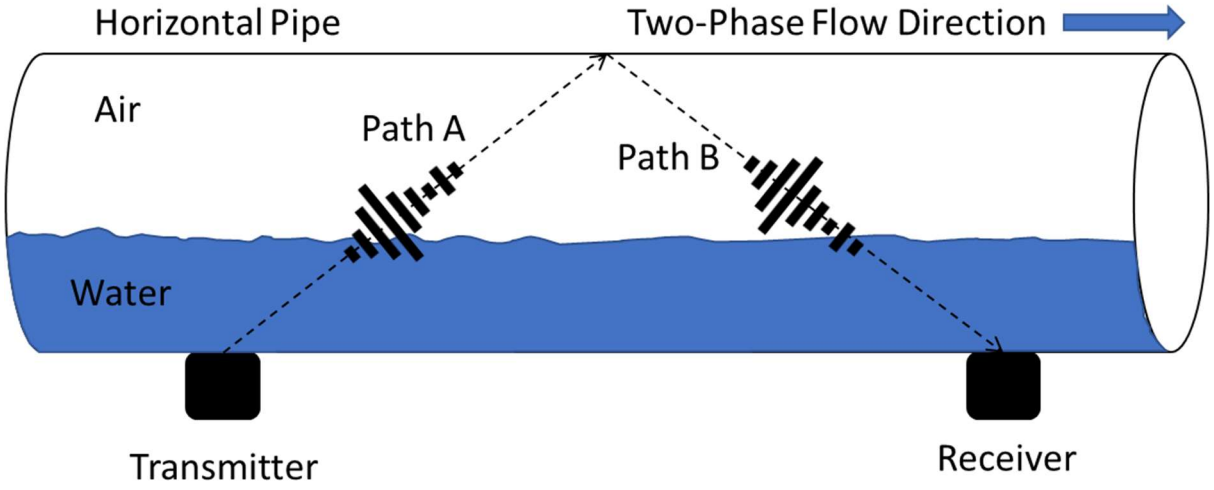


Figure 6.46: Ultrasonic Pitch-Catch Technique to obtain B-scan of two-phase flow by collecting 863 A-scans for each experiment of air and water combination

The A-scans are combined and processed as a B-scan image by using a MATLAB program given in Appendix-J written by Tonni F. Johansen. The images for stratified, wavy, annular, slug and plug flow regimes generated through this program are provided in Figure 6.47, Figure 6.48, Figure 6.49, Figure 6.50 and Figure 6.51 respectively.

The interface of air and water in a two-phase stratified flow is easily observed in the B-scan image since in stratified flow, the flow is not turbulent. When looking at the B-scan image, the strong reflections or echoes are observed at the interface of water and air in the pipe. For Stratified flow, this reflection is observed at t=60µs as a horizontal line as seen in Figure 6.47. It is the strongest, brightest and dominant reflection compared to all the other reflections. Other reflections are due to the higher order scattering between water and lower pipe wall. These reflections are called ghost reflections. [52]

B-scan of Wavy flow (Figure 6.48) shows similar reflection at the air-water interface but with a wave shape formed by the movement of water phase in the shape of waves. Here strong reflections are seen at t=60µs and 100µs. Further investigation is required to find the interface. It is easy to distinguish stratified flow and wavy flow from the B-scan images.

A B-scan of Annular flow is given in Figure 6.49. It consists of a lot of reflections because the air phase is at the center of the pipe while the water phase is at the periphery of the pipe. The reflections are also due to water falling at the bottom of the pipe due to gravity and scattering effects. Two strong reflections are observed in the B-scan at t=30µs and t=90µs.

In the B-scans of slug and plug flows (Figure 6.50 and Figure 6.51), the interface is seen to be moving vertically with time with a small gap of no reflection where the pipe is completely full with water phase.

Based on the visual inspection of the B-scan images, the five flow regimes can be distinguished.

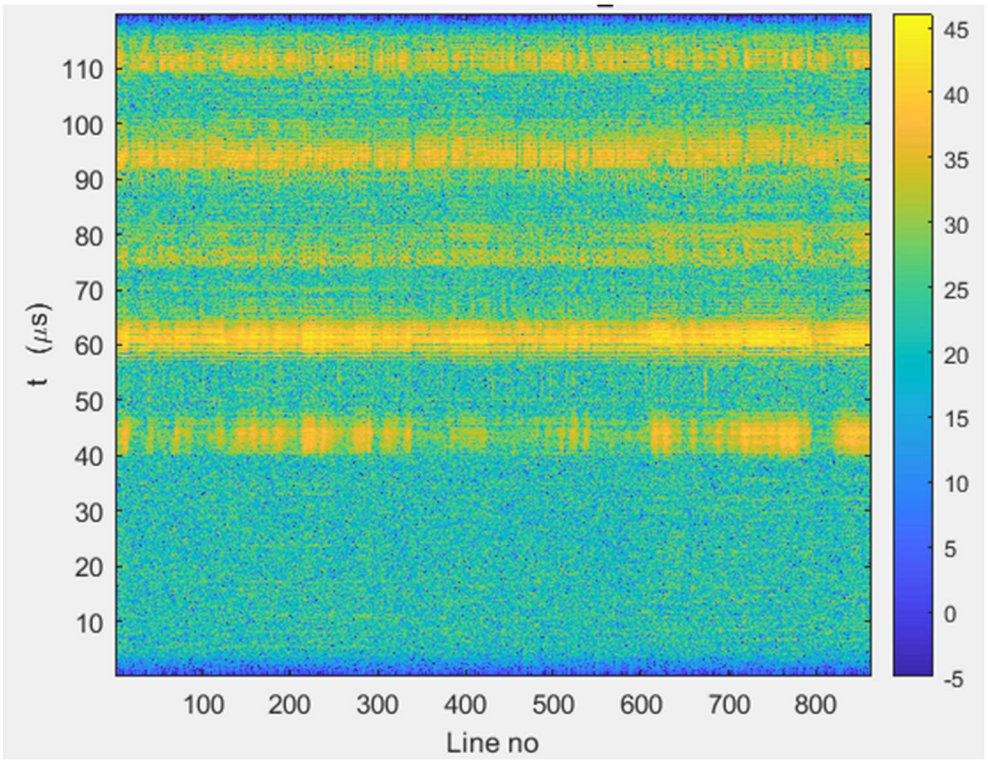


Figure 6.47: B-scan image for stratified flow

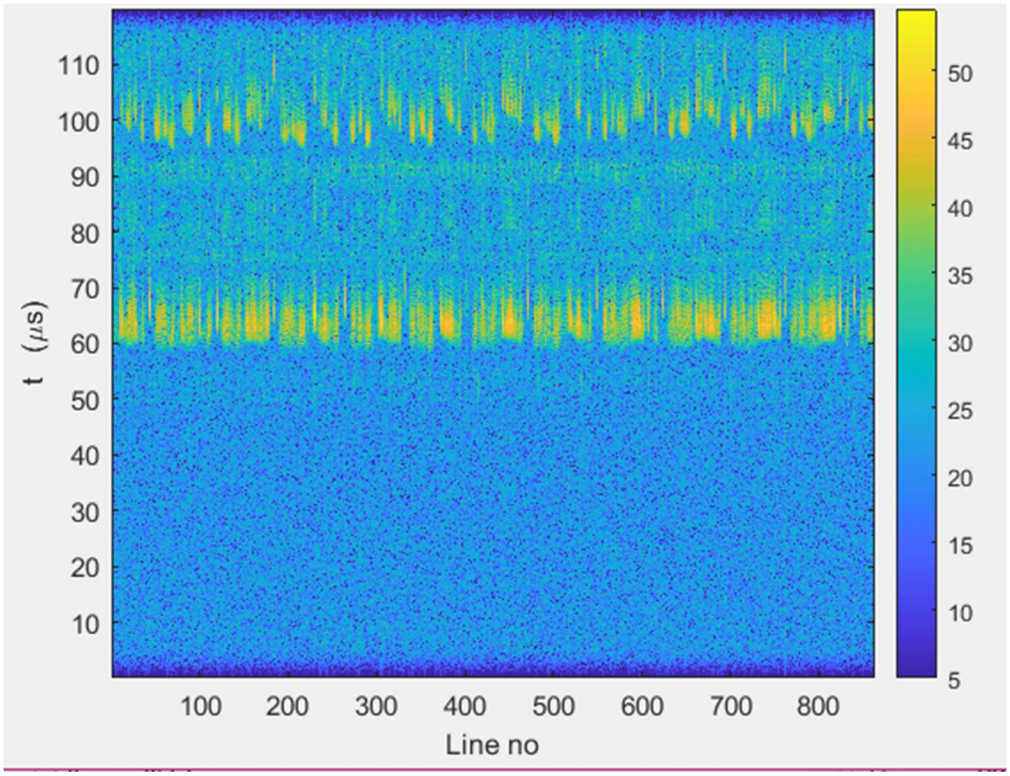


Figure 6.48: B-scan image for wavy flow

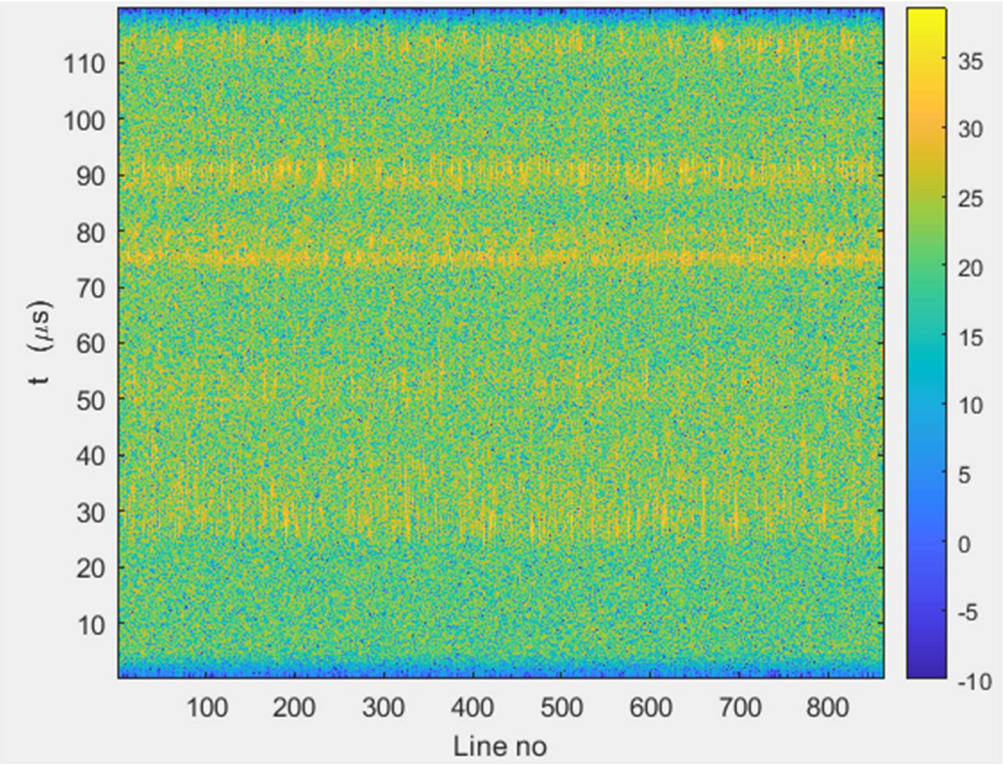


Figure 6.49: B-scan image for annular flow

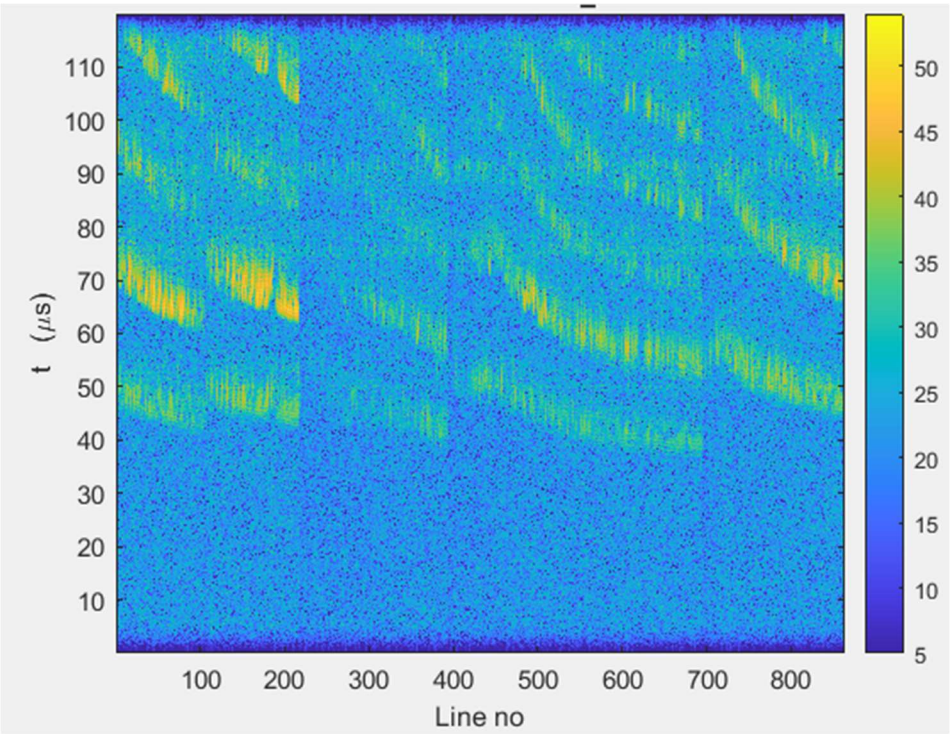


Figure 6.50: B-scan image for slug flow

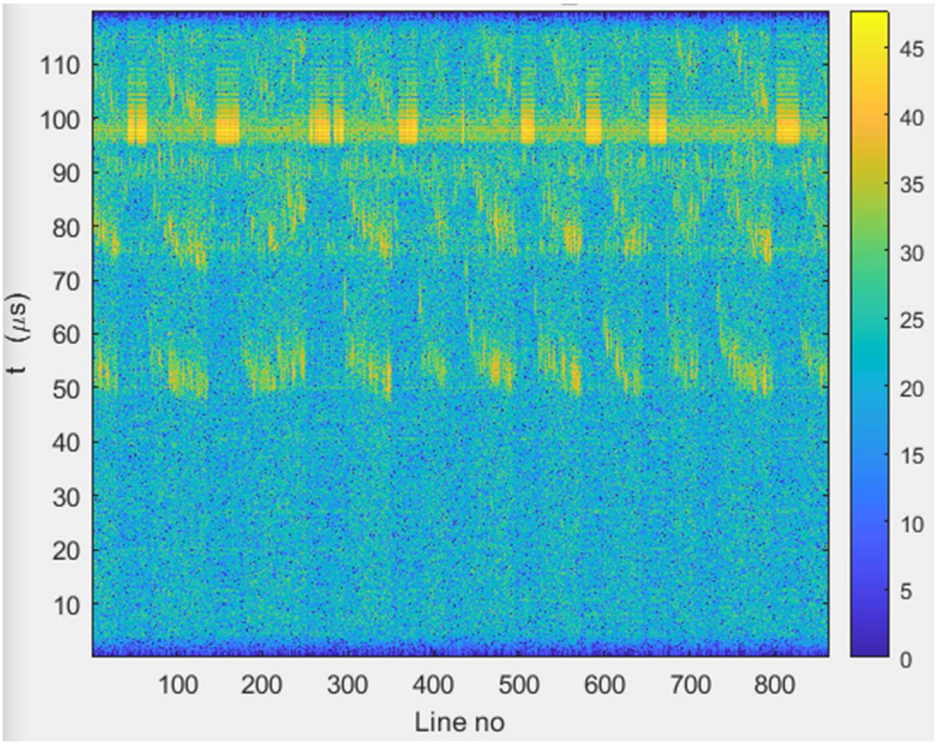


Figure 6.51: B-scan image for plug flow

Void fraction can also be calculated based on B-scan images of stratified flow regime. [52]

The distance (D_w) from the bottom of the image to the interface can be measured from the image. The distance (D_g) from the top of the image to the interface can be measured from the image. An estimate of void fraction can be calculated from Equation (6.1) [52] as:

$$\text{Void Fraction} = \frac{D_g}{D_w} \tag{6.1}$$

6.3.2 Image Processing from ECT Data

To generate an image from ECT data, the raw binary capacitance files are converted to image files with the extension 'AIM' using the ECT32 application. The image files contain 3000 frames of 32 by 32 pixels showing the cross-sectional image of the pipe. A Python program to generate an image from ECT image frame is given in Appendix-L. The generated image given in Figure 6.52 shows a cross-sectional image of the pipe with distinguishing colors for air and water in stratified flow depending on the pixel values in the 32 by 32 matrix of the ECT frame. When all the collected 3000 frames per experiment are converted to an image by the script of Appendix-L and joined in a video file where all 3000 frames are displayed successively using an appropriate frame rate, the dynamic flow in the pipe can be observed as a cross-sectional view.

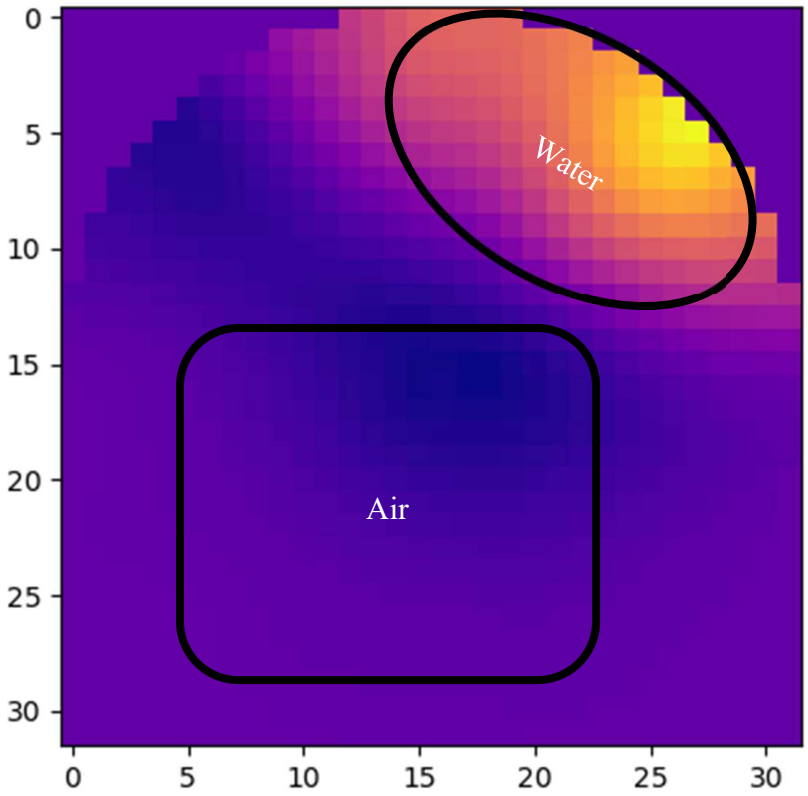


Figure 6.52: Image of one ECT frame showing air and water phase in Stratified flow

A rough estimate of void fraction from ECT image of a frame can be calculated by dividing the number of pixels with color depicting air and the number of pixels with color depicting water.

6.3.3 Comparison of ECT and Ultrasonic Tomography

As seen in the previous sections, ECT and ultrasonic can both be used for online real-time flow visualization. There are important differences between ECT and ultrasonic tomography which should be considered depending on the user requirements. The differences are provided in Table 6.18.

Table 6.18: Differences between ECT and ultrasonic tomography for flow visualization in a pipe

Sr. No.	ECT	Ultrasonic Tomography
1	The image generated from ECT provides a cross-sectional view of flow in the pipe	The image generated from ultrasonic provides a time-series side view of flow in the pipe

2	Only one cross-sectional frame can be seen at a time	Historical flow interfaces can be seen depending on the sampling time
3	Interfaces are hard to distinguish from the cross-sectional view of the flow	A clear interface between water and air can be seen in the B-scan image
4	A low resolution 32 by 32 image is generated with ECT	A high-resolution side view of the flow inside the pipe is generated with ultrasonic technology
5	Complex image processing technique is required to show side view from the cross-sectional ECT images	It is easier to generate B-scan image from A-scans.

6.3.4 Comparison of ECT Volume Ratio and US Images to Detect Interface

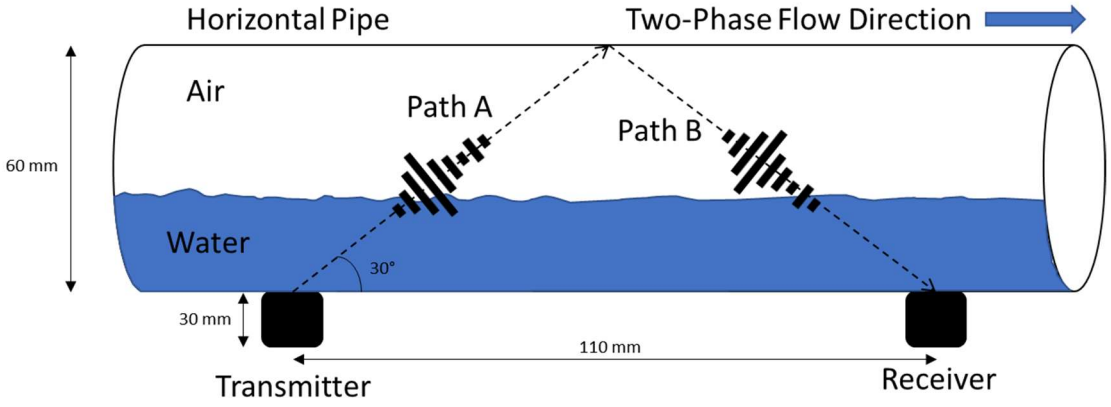


Figure 6.53: The distances and angles at which ultrasonic waves travel in the pitch-catch technique during the experiments performed at USN [53]

As seen from Figure 6.53, the ultrasonic waves cover 90 mm distance covering the wedge in which the ultrasonic transducer is placed and the pipe with outer diameter of 60 mm. The transmitter is at an angle of 30° from the horizontal. Based on this knowledge, the total path covered by the ultrasonic wave in one direction is calculated through Pythagoras theorem. The distance is found to be 105.5 mm. The distance travelled in the wedge is approximately 30 mm. Therefore, the distance covered by ultrasonic waves in the pipe is equal to 85.5 min in one direction. Based on this knowledge, the B-scan images from ultrasonic sensor can be interpreted.

The volume ratio from normalized capacitances of ECT can be used to verify the interface echo in the B-scan image. A time-series plot of volume ratio depicts the same information as the B-scan image for air/water two-phase flow. A volume ratio plot for stratified flow regime is shown in Figure 6.54 with its corresponding B-scan image. This experiment is performed with the flow rate of water at 5 kg/min and the flow rate of air at 0.1 kg/min to generate stratified flow regime. As per ECT volume ratio data, the volume ratio remains close to 30%. In the B-scan, there are 5 reflections that are observed with the brightest one at 60µs. From the volume ratio data, the interface should lie around 30% of the pipe diameter. The height of the B-scan is approximately 105.5 mm including 30 mm of wedge. If this 30 mm is deducted from the height of the B-scan, the interface should lie around 30% of 85.5 mm. This comes to 25.65 mm. Looking at the B-scan, the second line of echo is the interface line which is near to the height of 25.65 mm based on the volume ratio data from ECT.

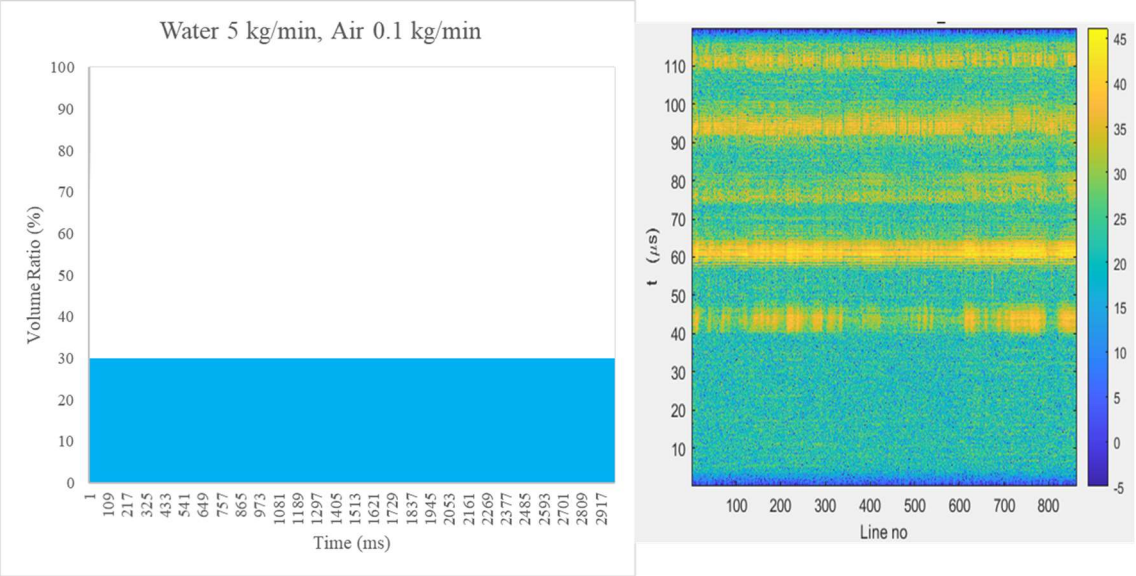


Figure 6.54: Comparison of ECT volume ratio data and B-scan images to detect air/water interface in the stratified regime

Similarly, Figure 6.55 to Figure 6.58 provides the comparison of ECT volume ratio and B-scan images of slug, wavy, plug and annular regimes respectively. The interfaces for these regimes can be verified to be at around 70, 100, 100 and 70 µs respectively from the discussion above.

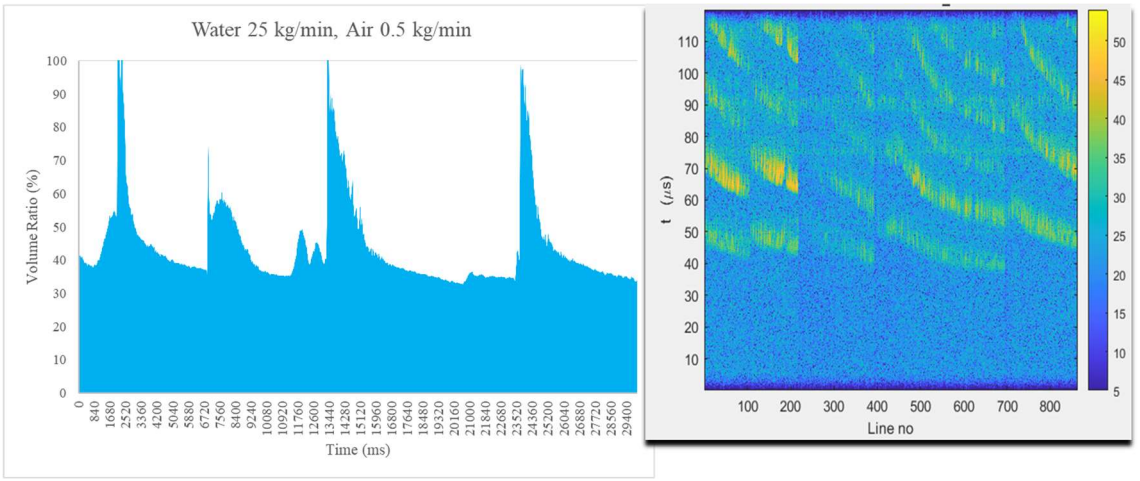


Figure 6.55: Comparison of ECT volume ratio data and B-scan images to detect air/water interface in the slug regime

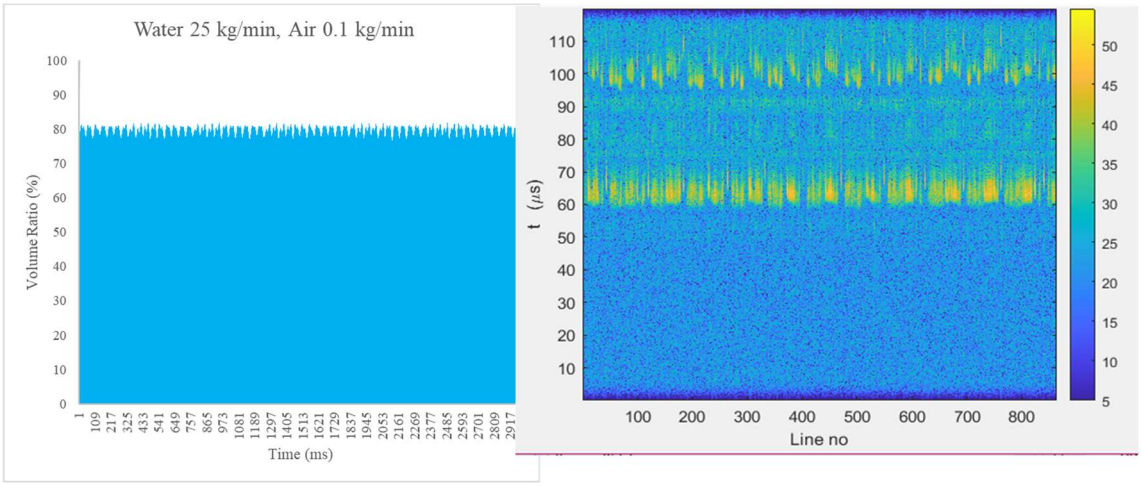


Figure 6.56: Comparison of ECT volume ratio data and B-scan images to detect air/water interface in the wavy regime

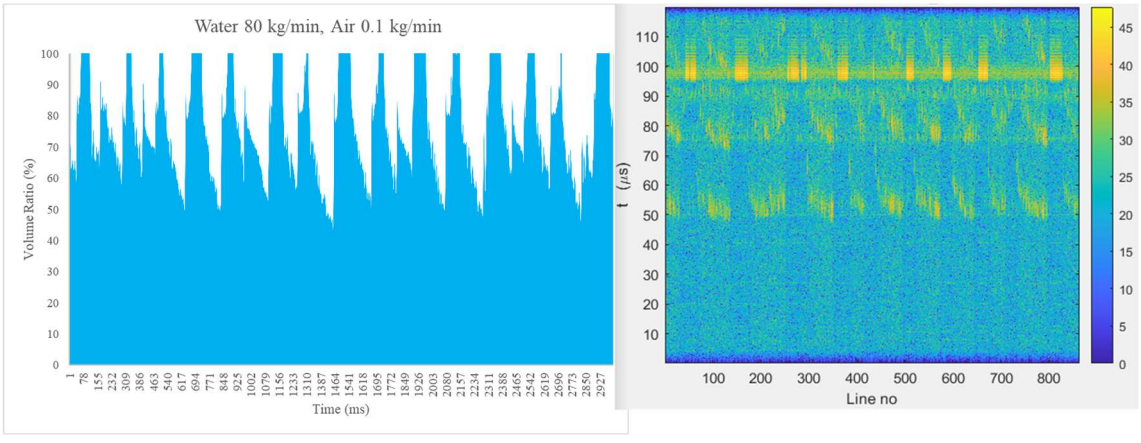


Figure 6.57: Comparison of ECT volume ratio data and B-scan images to detect air/water interface in the plug regime

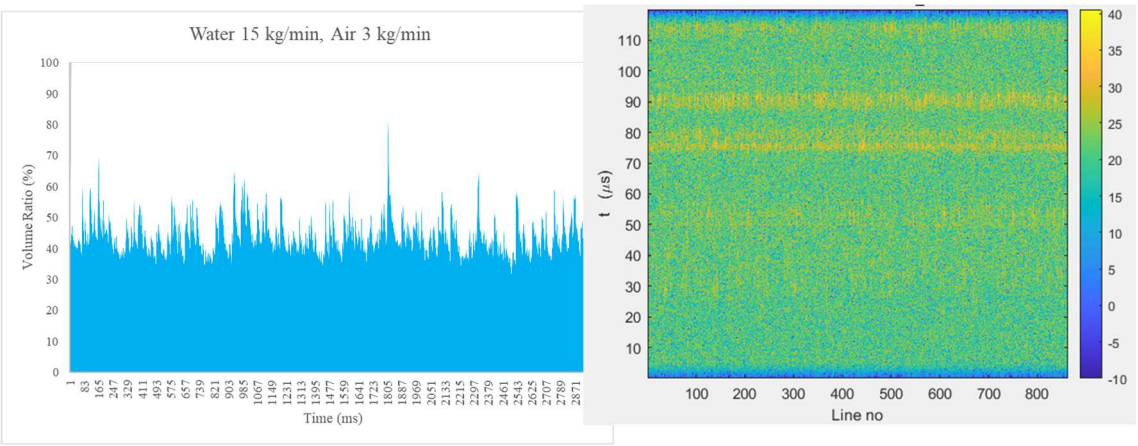


Figure 6.58: Comparison of ECT volume ratio data and B-scan images to detect air/water interface in the annular regime

Another method is explored to find the correct time-of-flight in μs from the B-scan image at which there is a presence of interface. Equation (6.2) provides the formula for this method where *VR* is the volume ratio from ECT data. Here it is assumed that the speed of sound in water is 1500 m/s. From transmitter to receiver, the ultrasound waves travel twice amount of the distance which is assumed to be equal in both directions. [54]

$$\text{Interface time in B-scan (μs)} = \frac{[2 \times (VR \times 56 + 30 + 2)]}{1.5} \qquad (6.2)$$

From Figure 6.54 of stratified regime, *VR* is equal to 30%. Putting this value in Equation (6.2), the interface time comes out to be 65.06μs.

From Figure 6.56 of wavy regime, *VR* is equal to 80%. Putting this value in Equation (6.2), the interface time comes out to be 102.4μs.

This confirms that fusing ECT volume ratio data and US B-scan images can improve the method for interface detection from B-scan images.

7 Discussion

In this thesis, non-intrusive and non-radioactive ECT and US sensors have been investigated for flow regime classification and flow parameters estimation. Raw data from ECT and ultrasonic sensors are used to create flow regime classification and flow parameters estimation models. Sensor fusion concept is also studied in this thesis to investigate the effect of collaboration of different types of sensors on flow regime classification performance. While similar work has been done before in previous theses at USN, this is the first time ECT and ultrasonic sensors are compared together for their performance in MPFM. Also, none of the previous works have studied ECT normalized capacitances in the form of spectrograms and box plots. A set of 45 dedicated experiments were performed with the necessary multi-sensor data logging for analysis in this thesis. This set of 8-electrode ECT experiments is completely different from previous experimental campaigns in scope and structure. The combination of flow rates for these experiments are different from previous theses. Volume ratio Neural Network model and flow velocity cross-correlation model are developed for the first time in this thesis as compared to other theses at USN. In this thesis, very high accuracy is achieved for flow regime classification machine learning models by using normalized capacitances only as features of classification algorithms. This reduces complex feature engineering stages such as eigenvalue extraction from the machine learning pipeline making the classification problem simple and computationally light. While 135000 samples were used to develop the classification model, lesser number of samples can be used for the same problem without loss of much accuracy as seen in the classification model after sensor fusion with 1350 samples. Comparison between US and ECT data for visualization of MPF is done for the first time in this thesis.

This thesis shows the importance of sensor fusion in the field of MPFM. In the flow regime classification models, sensor fusion has higher accuracy than using only one type of sensor. Additionally, combining ECT and US sensors revealed more information about the flow interface than by using only US sensor. These results reveal the benefits of sensor fusion in MPFM.

While training machine learning models, Decision Tree algorithm is very fast during training while providing good accuracy. This removes the need of complex ANN for MPFM. Decision Tree algorithms are computationally light and can be implemented in most of the open-source programming languages.

In flow velocity estimation, the flow velocity is calculated in m/s. Depending on the location of cross-correlation region and the type of flow regime, the phase velocity can be judged. For instance, the phase velocity of water present around C78 for plug flow from Table 6.10 is 1.7m/s.

This thesis shows the applicability of ECT and US sensors in the field of MPFM. They provide a great value to this field based on their performance in this thesis. In terms of accuracy of flow regime classification, ECT has the potential to replace visual inspection of flow regimes in industries.

In [26], cross-correlation and cross-spectrum techniques were used on ECT and GRT sensors data to estimate flow velocities without considering complex flow regimes such as plug and

slug. This thesis uses cross-correlation technique on ECT data only to find flow velocities in complex flow regimes without using radioactive sensors such as GRT.

In [27], ECT was used to develop flow regime classification models through statistical parameters and filters. In comparison, this thesis employs machine learning algorithms only for flow regime classification and flow parameters estimation.

In [28], two-phase flow regime identification using ECT is presented using SVM only. In this thesis, machine learning algorithms of DT, KNN, LR, LD, NN and SVM are used to train two-phase flow regime classification models.

In [29] velocity estimation of two-phase flow through ECT using cross-correlation technique is studied. However, complex flow regimes are not considered in the study. In this thesis, cross-correlation technique is used on ECT data to find flow velocities of complex flow regimes such as plug and slug.

In [30], two-phase flow regime identification model is developed through ultrasonic power spectral density (PSD) technique. In this thesis, machine learning algorithms are used to develop flow regime classification models on ultrasonic A-scan data.

In [11], a technique to classify 2-phase gas/liquid flow regime in an S-shaped pipeline using Doppler ultrasonic sensors is provided using machine learning techniques. In this thesis, machine learning models to classify 2-phase air/water flow regime in a horizontal pipeline using ultrasonic Pitch-Catch technique are presented.

The following are some suggestions for future work based on this thesis:

Doppler technique using ultrasonic sensor for multiphase flow metering [11] can be used to compare its results for flow classification with the transit time technique used in this thesis. For flow velocity estimation, cross-correlation technique on ultrasonic sensors, which is similar to dual plane ECT technique, can be implemented and its results compared with the flow velocity estimation technique with ECT. The accuracy of machine learning models may be further increased by tuning the hyperparameters of the models. The machine learning models can be implemented using Python, which gives a greater control and range over the hyperparameters to the user and is applicable for commercial applications. The US and ECT features can be combined to test the performance of flow regime classification models. US sensors can be set in another configuration such as pulse-echo to test improvements in the performance of flow regime classification models. PCA analysis can be performed to reduce features of ECT.

8 Conclusion

In this thesis, data-driven multiphase flow metering with multimodal sensor suite for identification of flow regimes, and estimation of phase fractions and velocities in a two-phase flow with air and water has been achieved with astounding results. The multimodal sensor suite consisted of Electrical Capacitance Tomography and Ultrasonic sensors along with conventional sensors.

A total of 63 ECT experiments were performed with a combination of air and water flow rates in the horizontal flow rig at USN.

A total of 5 ECT and 2 US data-driven MPFM models are developed in this thesis that show promising results.

Using ECT normalized capacitances measurements, machine learning flow classification models are created. In the first model, 28 normalized capacitances are used as inputs and flow regime is used as output of classification machine learning models. The results of this model are shown in Table 8.1 with various machine learning algorithms. With such a high overall accuracy, it can be concluded that normalized capacitances are good indicators of flow regimes in the horizontal flow pipe.

Table 8.1: Overall accuracy of flow classification machine learning models with 28 normalized capacitances as inputs and flow regime as output

Machine Learning Algorithm	Accuracy
Neural Network	96.8%
KNN	98.8%
Decision Tree	96.6%
Support Vector Machine	94.8%
Linear Discriminant	92.7%

Through EDA, the inputs of machine learning models are reduced to 15 normalized capacitances. The results of this model are shown in Table 8.2 with various machine learning algorithms. A similar level of overall accuracy is achieved in this case compared with usage of 28 normalized capacitances as inputs. This achieved a 53% reduction of inputs and saving of computational power without much loss of accuracy.

Table 8.2: Overall accuracy of flow classification machine learning models with 15 normalized capacitances as inputs and flow regime as output

Machine Learning Algorithm	Accuracy
KNN	98.5%
Decision Tree	96.0%

Support Vector Machine	93.4%
Linear Discriminant	85.6%

To achieve the benefits of sensor fusion, the 28 normalized capacitances are combined with 3 pressure signals to test the improvement of flow classification accuracy. The results of this model are shown in Table 8.3 with various machine learning algorithms. As compared with Table 8.1 there is an increase in overall accuracy by 4% for SVM and 2% for Decision Tree. It can be concluded that fusion of ECT normalized capacitances with conventional pressure measurements improves flow classification accuracy.

Table 8.3: Overall accuracy of flow classification machine learning models with 28 normalized capacitances and 3 pressure signals as inputs and flow regime as output

Machine Learning Algorithm	**Accuracy**
KNN	98.4%
Decision Tree	98.5%
Support Vector Machine	99.1%
Linear Discriminant	97.7%

To estimate flow velocity with ECT technology, cross-correlation technique is used on the data from two planes of ECT. The average flow velocity for various regimes is shown in Table 8.4. The highest velocity is found for slug flow at 3.74m/s.

Table 8.4: Average flow velocity for various flow regimes by using cross-correlation on dual plane ECT

Flow Regime	**Average Flow Velocity (m/s)**
Annular	1.43 to 1.87
Plug	1.24 to 1.43
Slug	2.67 to 3.74

Estimation of volume ratio is performed by using differential pressure signals and air flow rate as inputs to a Neural Network algorithm. The output is the inverse of volume ratio obtained from ECT volume ratio files. This model achieved a very high R value of 0.95.

Flow classification is also performed by using ultrasonic data. Six features are extracted from the A-scan of ultrasonic signals and used as inputs to the machine learning algorithms. The results of this model are shown in Table 8.5 for various machine learning algorithms. It is noteworthy that this model is found to be highly accurate for stratified and annular flow classification.

Table 8.5: Overall accuracy of flow classification machine learning models with 6 ultrasonic A-scan features as inputs and flow regime as output

Machine Learning Algorithm	Accuracy
SVM	69.7%
Decision Tree	68.3%
Ensemble Bagged Trees	70.6%
Ensemble Boosted Trees	69.7%

Additionally, flow visualization with both ECT and US sensors is compared. Ultrasonic B-scan image is found to be a better option for flow visualization as compared to an ECT image in terms of the angle of view and monitoring of flow regimes. It has also been discovered that by analyzing ECT volume ratio data with US B-scan image, air/water interface can be detected in a B-scan image.

To conclude, it has been shown in this thesis that multiphase flow metering is possible by using data from ECT and ultrasonic sensors. The data obtained from these sensors can be fed directly to machine learning pipelines to obtain knowledge of the flow in the pipe and, to monitor and control flow in the pipe. Use of machine learning algorithms on time-series data obtained from these sensors removes the need of complicated mathematical time-series and image processing methods. Finally, an overall comparison between ECT and US Pitch-Catch techniques for MPFM is provided in Table 8.6. Based on the comparison, ECT is a better option for MPFM in terms of its versatility of use.

Table 8.6: Comparison of ECT with US Pitch-Catch techniques for multiphase flow metering

Parameter	ECT	US Pitch-Catch
Flow Regime Classification	High Accuracy for plug and slug	Low Accuracy for plug and slug
Volume Ratio Estimation	Possible with Sensor Fusion techniques	Possible through analyzing B-scan images
Flow Velocity Estimation	Possible by use of dual-plane ECT	Not possible with the current Pitch-Catch setup
Interface Detection	Detected through volume ratio data	Detected through B-scan image in combination with ECT volume ratio data
Image Processing	Cross-sectional image with low resolution	High resolution side view of the multiphase flow

References

[1] K. Havre and M. Dalsmo, “Active Feedback Control as the Solution to Severe Slugging,” in *All Days*, New Orleans, Louisiana: SPE, Sep. 2001, p. SPE-71540-MS. doi: 10.2118/71540-MS.

[2] “Multiphase flow,” *Wikipedia*. Jan. 2023. Accessed: Feb. 18, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Multiphase_flow&oldid=1136311145

[3] S. L. Soo, *Multiphase fluid dynamics*. Beijing : Aldershot ; Brookfield, USA: Science Press ; Gower Technical, 1990.

[4] G. F. Hewitt, *Multiphase Flow*. Begel House Inc., 2011. doi: 10.1615/AtoZ.m.multiphase_flow.

[5] C. Tan and F. Dong, “Sensor Instrumentation for Flow Measurement,” in *Encyclopedia of Sensors and Biosensors (First Edition)*, R. Narayan, Ed., Oxford: Elsevier, 2023, pp. 536–554. doi: 10.1016/B978-0-12-822548-6.00074-1.

[6] J. H. Vohr, “FLOW PATTERNS OF TWO-PHASE FLOW–A SURVEY OF LITERATURE,” Columbia Univ., New York. Engineering Research Labs., TID-11514; CU-2-60-AEC-187-Ch., Dec. 1960. Accessed: Feb. 19, 2023. [Online]. Available: https://www.osti.gov/biblio/4017048

[7] M. Alssayh, A. Addali, David, and T. Dao, “IDENTIFICATION OF TWO PHASE FLOW REGIME USING ACOUSTIC EMISSION TECHNOLOGY,” 2013. Accessed: Feb. 20, 2023. [Online]. Available: https://www.semanticscholar.org/paper/IDENTIFICATION-OF-TWO-PHASE-FLOW-REGIME-USING-Alssayh-Addali/295d25a414d008aeaa42a3820b1727c89822b0f2

[8] L. S. Hansen, S. Pedersen, and P. Durdevic, “Multi-Phase Flow Metering in Offshore Oil and Gas Transportation Pipelines: Trends and Perspectives,” *Sensors*, vol. 19, no. 9, p. 2184, May 2019, doi: 10.3390/s19092184.

[9] E. Pereyra, C. Torres, R. Mohan, L. Gomez, G. Kouba, and O. Shoham, “A methodology and database to quantify the confidence level of methods for gas–liquid two-phase flow pattern prediction,” *Chem. Eng. Res. Des.*, vol. 90, no. 4, pp. 507–513, Apr. 2012, doi: 10.1016/j.cherd.2011.08.009.

[10] N. Almalki and W. H. Ahmed, “Prediction of two-phase flow patterns using machine learning algorithms,” Nov. 2020. doi: 10.11159/ffhmt20.182.

[11] S. Godfrey Nnabuife, B. Kuang, J. F. Whidborne, and Z. Rana, “Non-intrusive classification of gas-liquid flow regimes in an S-shaped pipeline riser using a Doppler ultrasonic sensor and deep neural networks,” *Chem. Eng. J.*, vol. 403, p. 126401, Jan. 2021, doi: 10.1016/j.cej.2020.126401.

[12] J. M. Mandhane, G. A. Gregory, and K. Aziz, “A flow pattern map for gas—liquid flow in horizontal pipes,” *Int. J. Multiph. Flow*, vol. 1, no. 4, pp. 537–553, Oct. 1974, doi: 10.1016/0301-9322(74)90006-8.

[13] A. Liné and J. Fabre, *Stratified Gas-Liquid Flow*. Begel House Inc., 2011. doi: 10.1615/AtoZ.s.stratified_gas-liquid_flow.

[14] S. Jayanti, *Wavy Flow*. Begel House Inc., 2011. doi: 10.1615/AtoZ.w.wavy_flow.

[15] Y. A. Zeigarnik, *Annular flow*. Begel House Inc., 2011. doi: 10.1615/AtoZ.a.annular_flow.

[16] "plug_flow." Accessed: Feb. 25, 2023. [Online]. Available: https://glossary.slb.com/en/terms/p/plug_flow

[17] "slug_flow." Accessed: Feb. 25, 2023. [Online]. Available: https://glossary.slb.com/en/terms/s/slug_flow

[18] "bubble_flow." Accessed: Feb. 25, 2023. [Online]. Available: https://glossary.slb.com/en/terms/b/bubble_flow

[19] "spray flow," *TheFreeDictionary.com*. Accessed: Feb. 25, 2023. [Online]. Available: https://encyclopedia2.thefreedictionary.com/spray+flow

[20] "phase_velocity_log." Accessed: Feb. 19, 2023. [Online]. Available: https://glossary.slb.com/en/terms/p/phase_velocity_log

[21] "Process Tomography Limited." Accessed: Feb. 22, 2023. [Online]. Available: http://www.tomography.com/

[22] R. Hytowitz, "Multiphase Flow Measurement: Techniques and Applications," *IEEE Access*. Nov. 2016. Accessed: Feb. 21, 2023. [Online]. Available: https://ieeeaccess.ieee.org/closed-special-sections/multiphase-flow-measurement-techniques-applications/

[23] C. Tan, Y. Murai, W. Liu, Y. Tasaka, F. Dong, and Y. Takeda, "Ultrasonic Doppler Technique for Application to Multiphase Flows: A Review," *Int. J. Multiph. Flow*, vol. 144, p. 103811, Nov. 2021, doi: 10.1016/j.ijmultiphaseflow.2021.103811.

[24] "What is a Coriolis Flow Meter and How Does it Work?," *https://www.omega.com/en-us/*. Accessed: Feb. 26, 2023. [Online]. Available: https://www.omega.com/en-us/resources/what-is-a-coriolis-flow-meter

[25] T. Bikmukhametov and J. Jäschke, "First Principles and Machine Learning Virtual Flow Metering: A Literature Review," *J. Pet. Sci. Eng.*, vol. 184, p. 106487, Jan. 2020, doi: 10.1016/j.petrol.2019.106487.

[26] S. H. Stavland, Y. Arellano, A. Hunt, R. Maad, and B. T. Hjertaker, "Multimodal Two-Phase Flow Measurement Using Dual Plane ECT and GRT," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, doi: 10.1109/TIM.2020.3034615.

[27] A. Dupré, G. Ricciardi, S. Bourennane, and S. Mylvaganam, "Electrical Capacitance-Based Flow Regimes Identification—Multiphase Experiments and Sensor Modeling," *IEEE Sens. J.*, vol. 17, no. 24, pp. 8117–8128, Dec. 2017, doi: 10.1109/JSEN.2017.2707659.

[28] H. X. Wang and L. F. Zhang, "Identification of two-phase flow regimes based on support vector machine and electrical capacitance tomography," *Meas. Sci. Technol.*, vol. 20, no. 11, p. 114007, Nov. 2009, doi: 10.1088/0957-0233/20/11/114007.

[29] H. L. M. Ameran *et al.*, “VELOCITY MEASUREMENT SIMULATIVE STUDY OF TWIN PLANE ECT USING ADVANCED CROSS CORRELATION TECHNIQUE,” *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 19, Oct. 2015, doi: October 2015.

[30] S. G. Nnabuifea, K. E. S. Pilario, L. Lao, Y. Cao, and M. Shafiee, “Identification of Gas-Liquid Flow Regimes Using a Non-intrusive Doppler Ultrasonic Sensor and Virtual Flow Regime Maps,” *Flow Meas. Instrum.*, vol. 68, Aug. 2019, Accessed: Feb. 23, 2023. [Online]. Available: https://doi.org/10.1016/j.flowmeasinst.2019.05.002

[31] I. Saied and M. Meribout, “Electronic hardware design of electrical capacitance tomography systems,” *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, vol. 374, no. 2070, p. 20150331, Jun. 2016, doi: 10.1098/rsta.2015.0331.

[32] J. D. N. Cheeke and J. D. N. Cheeke, *Fundamentals and Applications of Ultrasonic Waves*, 0 ed. CRC Press, 2010. doi: 10.1201/9781420042139.

[33] V. K. Solanki, V. G. Díaz, and J. P. Davim, Eds., *Handbook of IoT and Big Data*. Boca Raton: CRC Press, 2019. doi: 10.1201/9780429053290.

[34] “Guidelines for Ultrasonic Inspection of Hanger Pins - FHWA-HRT-04-042.” Accessed: Apr. 19, 2023. [Online]. Available: https://www.fhwa.dot.gov/publications/research/infrastructure/structures/04042/02.cfm

[35] “ULTRASONIC TESTING OF AEROSPACE MATERIALS.” Accessed: Apr. 27, 2023. [Online]. Available: http://www.klabs.org/DEI/References/design_guidelines/test_series/1422msfc.pdf

[36] “Supervised learning,” *Wikipedia*. Mar. 2023. Accessed: Apr. 04, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Supervised_learning&oldid=1144919032

[37] “Machine Learning Fundamentals,” *Packt*. Accessed: Apr. 04, 2023. [Online]. Available: https://www.packtpub.com/product/machine-learning-fundamentals/9781789803556

[38] “What is the k-nearest neighbors algorithm? \textbar IBM.” Accessed: Apr. 04, 2023. [Online]. Available: https://www.ibm.com/in-en/topics/knn

[39] “5-2 K-nearest-neighbor Classifiers.” Accessed: Apr. 04, 2023. [Online]. Available: http://mirlab.org/jang/books/dcpr/prKnnc.asp?title=5-2%20K-nearest-neighbor%20Classifiers

[40] A. Subasi, “Chapter 3 - Machine learning techniques,” in *Practical Machine Learning for Data Analysis Using Python*, A. Subasi, Ed., Academic Press, 2020, pp. 91–202. doi: 10.1016/B978-0-12-821379-7.00003-5.

[41] “Logistic Regression \textbar Machine Learning,” *Google Developers*. Accessed: Apr. 04, 2023. [Online]. Available: https://developers.google.com/machine-learning/crash-course/logistic-regression/video-lecture

[42] “Support vector machine,” *Wikipedia*. Mar. 2023. Accessed: Apr. 19, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Support_vector_machine&oldid=1144271534

[43] “Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing - Siddique, Nazmul; Adeli, Hojjat: 9781118337844 - AbeBooks.” Accessed: Apr. 04, 2023. [Online]. Available: https://www.abebooks.com/9781118337844/Computational-Intelligence-Synergies-Fuzzy-Logic-1118337840/plp

[44] Zhou, Z.-H, *Ensemble Methods: Foundations and Algorithms*. CRC Press, 2012.

[45] “Cross-correlation,” *Wikipedia*. Mar. 2023. Accessed: Apr. 19, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Cross-correlation&oldid=1145253137

[46] “Density and concentration meters,” *Sensor Technology AS*. Accessed: Feb. 22, 2023. [Online]. Available: https://s-tec.no/en/produkter-og-tjenester/densitet-og-konsentrasjonsmalere/

[47] C. Pradeep, “Tomographic Approach to Automatic and Non-Invasive Flow Regime Identification,” Aug. 2015. Accessed: Feb. 20, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Tomographic-Approach-to-Automatic-and-Non-Invasive-Pradeep/191d13b3eccda2d16428327e138307edbcc68cbb

[48] “ULTRASOUND PRODUCTS - LE CŒUR ELECTRONIQUE - MADE IN FRANCE.” Accessed: Apr. 19, 2023. [Online]. Available: https://www.lecoeur-electronique.net/

[49] “Micro Motion Flow Meters \textbar Emerson US.” Accessed: Mar. 01, 2023. [Online]. Available: https://www.emerson.com/en-us/automation/micro-motion

[50] W. L. Martinez, A. R. Martinez, and J. Solka, *Exploratory data analysis with MATLAB*, Third edition. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2017.

[51] “scipy.signal.spectrogram.” https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.spectrogram.html

[52] J. M. Blackledge, “B-scan imaging of two-phase flows,” in *IEE Colloquium on Ultrasound in the Process Industry*, Sep. 1993, p. 5/1-517.

[53] Ashim Khadka, “Multiphase flow estimation using accelerometers, ultrasonic sensors, and machine learning,” University of South-Eastern Norway, Porsgrunn, 2022. Accessed: Apr. 23, 2023. [Online]. Available: https://openarchive.usn.no/usn-xmlui/handle/11250/3005166?show=full

[54] N. M. Tole and H. Ostensen, *Basic physics of ultrasonographic imaging*. Geneva: World Health Organization, 2005.

Appendices

Appendix A – Task Description

Appendix B - Python script to extract normalized capacitances features from ECT .ANC files to .CSV files

Appendix C - Windows Shell code to combine CSV files

Appendix D - Python script to combine ECT data with pressure sensor data

Appendix E - Python script to combine ECT Volume Ratio files with pressure and flow data

Appendix F – MATLAB script to generate box plots

Appendix G – MATLAB script to generate Correlation plots of ECT

Appendix H – Python script to generate spectrograms

Appendix I - MATLAB script to detect cross correlation of dual planes of ECT normalized capacitances

Appendix J – MATLAB script to read ultrasonic binary files

Appendix K – MATLAB script to extract features from ultrasonic data

Appendix L – Python script to display image from ECT pixels

Appendix M – Box Plots

Appendix N – Python script to generate Correlation Coefficients

Appendix A – Task Description

University of South-Eastern Norway

Faculty of Technology, Natural Sciences and Maritime Sciences, Campus Porsgrunn

FMH606 Master's Thesis

Title: Multiphase flow metering with multimodal sensor suite for identification of flow regimes, and estimation of phase fractions and velocities
USN supervisor: Ru Yan, Saba Mylvaganam, Håkon Viumdal

External partner:

Task background:
Multiphase flow rigs in USN, Campus Porsgrunn, have been used in a multitude of experiments involving bachelor, master, and PhD candidates and researchers (for both Norwegian and International), resulting in many student projects and research publications. Measurements using acoustic emission (AE) sensors, tomographic measurements using electrical capacitance tomography (ECT) equipment, along with conventional measurements, such as acceleration, temperature, and pressure, have been used in these multiphase flow measurements.

A measurement campaign should be planned for multiphase flow using the existing instrumentation for ultrasonic and tomographic measurements. The master students will get practical experience by running the multiphase flow rig in collaboration with the team at USN.

This project deals with the usage of an ultrasonic flowmeter and ECT sensors together with conventional measurements for measuring flow types, flow velocities and void fractions. The focus is on the existing and collected data from the experiments for estimating flow parameters and exploiting physics-aided machine learning (ML) / artificial intelligence (AI) techniques

Task description:
The tentative list of tasks for this thesis work is as follows:
1) Brief survey of multiphase flow and multiphase flow metering.
2) Coupled to task (1), a literature survey of the contemporary publications related to the same topic, especially emphasizing the usage of ultrasonic and ECT modalities.
3) Giving an overview of the USN-multiphase setup of the rig, the instruments installed, and how the data acquisition from different sensors are taken, transferred, and stored on-premises and in the cloud.
4) Making a test plan based on existing plans, running experiments on the flow rig according to the test plan and implementing data collection.
5) Performing exploratory data analysis (EDA) on the existing data from previous tests and the collected data from the experiment.
6) Developing models including ML/AI techniques in the estimation of flow types and flow parameters using the data from ECT and ultrasonic flowmeters together with other process data from the USN rig.
7) Compare the data from ultrasonic and ECT modalities with focus on image processing
8) Submitting a Master Thesis following the guidelines of USN with necessary programs and including a well-documented and complete set of all experimental data from the measurements

Student category: IIA students

The task is suitable for online students (not present at the campus): No

Practical arrangements:
The student will be trained to run the multiphase flow rig at campus Porsgrunn, and will have access to the rig throughout the project period. This work is closely coupled to an ongoing project SAM (SAM: Self Adapting Model-based system for Process Autonomy - SINTEF), where Equinor and SINTEF are central partners and may involve partners in the project Digipro as well.

Supervision:
As a general rule, the student is entitled to 15-20 hours of supervision. This includes necessary time for the supervisor to prepare for supervision meetings (reading material to be discussed, etc).

Signatures:

Supervisor (date and signature): 01.02.2023

Student (write clearly in all capitalized letters): NOORAIN SYED KAZMI

Student (date and signature): 1/2/2023

238758

Appendix B – Python script to extract normalized capacitances features from ECT .ANC files to .CSV files

The following script works when all the files are stored in one folder and the names of the files are in the format: ‘w2a0_1_Stratified.anc’, ‘w2a1_0_Wavy.anc’, etc. These files can be accessed from USN cloud.

```
import os
import pandas as pd
import numpy as np
for k in range(len(os.listdir(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ANC\ancFiles'))):
    fileName = os.listdir(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ANC\ancFiles')[k]
    df = pd.read_csv(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ANC\ancFiles\%s'%fileName, sep=" ", header=None)
    df1 = pd.DataFrame
    i = 0;
    j=0;
    k = np.arange(0,21000,7);

    df1 = df.take(k)
    df2 = df.take(k+1)
    df3 = df.take(k+2)
    df4 = df.take(k+3)
    df5 = df.take(k+4)
    df6 = df.take(k+5)
    df7 = df.take(k+6)

    df1 = df1.reset_index()
    df2 = df2.reset_index()
    df3 = df3.reset_index()
    df4 = df4.reset_index()
    df5 = df5.reset_index()
    df6 = df6.reset_index()
    df7 = df7.reset_index()

    result = pd.concat([df1, df2, df3, df4, df5, df6, df7], axis=1)

    result = result.dropna(axis='columns')
    result = result.drop(['index'],axis=1)
    result['Regime'] = fileName.split('.')[0].split('_')[2]
    result.to_csv(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ANC\csvFiles\%s.csv'%fileName.split('.')[0],header=False,  index=False)
```

Appendix C – Windows Shell code to combine CSV files

The following command can be used in Windows Shell to combine the contents of CSV files in one TXT file.

```
C:\Users\Luna\Documents\USN\Thesis-ML\23-02-03\ECT csv Files>copy *.csv yes.txt
w2a0_1.csv
w2a0_15.csv
w2a0_5.csv
```

Appendix D – Python script to combine ECT data with pressure sensor data

The following script combines ECT and conventional sensors data in one CSV file.

```
import os
import pandas as pd
import numpy as np
import glob

path = r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-03\23-02-03\LabVIEW\230202\slow"

filenames = glob.glob(path + "\*.txt")
df = pd.DataFrame()

for file in filenames:
   # reading csv files
   print("\nReading file = ",file)
   df = df.append(pd.read_csv(file,header=None,skiprows=10,sep='\t'))

for k in range(31):
   df=df.drop(index=[k+31])

df=df.drop(index=[30])
df=df.reset_index()
df.to_csv(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-03\MMSS\mmss2.csv',header=False,
index=False)
df = pd.read_csv(r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ANC\csvFiles\combined.txt",header=None)
df = df.take(np.arange(0,134998,100))
df1 = pd.read_csv(r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\MMSS\mmss2.csv",header=None)
df=df.reset_index()
result = pd.concat([df, df1], axis=1)
result.to_csv(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\MMSS\Velocity.csv',header=False, index=False)
```

Appendix E – Python script to combine ECT Volume Ratio files with pressure and flow data

The following script combines ECT Volume Ratio and conventional sensors data in one CSV file.

```
import os
import pandas as pd
import numpy as np
import glob

df = pd.read_csv(r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ECTVR\VR.txt",header=None)



df = df.take(np.arange(0,134998,100))

df1 = pd.read_csv(r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\MMSS\mmss2.csv",header=None)

df=df.reset_index()

result = pd.concat([df, df1], axis=1)



result.to_csv(r'C:\Users\Luna\Documents\USN\Thesis-ML\23-02-03\MMSS\VR-
PDT.csv',header=False, index=False)
```

Appendix F – MATLAB script to generate box plots

The following script generates box plots for one inter-electrode capacitance at a time.

```
load P1withRegime1
boxplot(C78,Regime)
title('Distribution of C78')
xlabel('Regimes')
ylabel('Normalized Capacitances')
```

Appendix G – MATLAB script to generate Correlation plots of ECT

The following script generates correlation plots of inter-electrode capacitances.

```
corrplot(P1withRegime2(1:135000,["C12","C13","C14","C15","C16","C17","C18"]))
corrplot(P1withRegime2(1:135000,["C23","C24","C25","C26","C27","C28"]))
corrplot(P1withRegime2(1:135000,["C34","C35","C36","C37","C38"]))
corrplot(P1withRegime2(1:135000,["C45","C46","C47","C48"]))
corrplot(P1withRegime2(1:135000,["C56","C57","C58"]))
corrplot(P1withRegime2(1:135000,["C67","C68"]))
```

Appendix H – Python script to generate spectrograms

The following script generates spectrograms from one inter-electrode normalized capacitances and one flow regime.

```
import pandas as pd
import scipy
from scipy import signal
import matplotlib.pyplot as plt
from scipy.fft import fftshift
df = pd.read_csv(r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-
03\ANC\csvFiles\combined.txt")
#df = pd.read_csv('combined.txt')
df.columns =['C12','C13','C14',      'C15','C16','C17',    'C18', 'C23', 'C24',
       'C25', 'C26', 'C27'  ,'C28' ,'C34','C35'  ,'C36','C37', 'C38', 'C45',
       'C46', 'C47', 'C48', 'C56'  ,'C57' ,'C58' ,'C67' ,'C68','C78','Regime']
plt.figure(1,figsize=(4, 10),dpi=200)
plt.title('Spectrogram of C27 of Slug')
f, t, Sxx = signal.spectrogram(df['C27'][df['Regime']=='Slug'],10e3)
plt.pcolormesh(t, fftshift(f), fftshift(Sxx, axes=0), cmap='viridis',vmax =
0.00001)
plt.ylabel('Frequency [Hz]')
plt.xlabel('Time-series')
plt.show()
```

Appendix I – MATLAB script to detect cross correlation of dual planes of ECT normalized capacitances

The following script generates cross cross-correlation plots and tables of 28 dual planes inter-electrode normalized capacitances for a specific flow regime.

```
for v = 1:28
   [c,lags] = xcorr(P1(1:2998,v),P2(1:2998,v));
   %[c,lags] = xcorr(P1(1:1000,v),P2(1:1000,v));
   %figure(v);
   %stem(lags,c);
   [M,I] = max(c);
   C(v) = lags(I)
end
```

Appendix J – MATLAB script to read ultrasonic binary files

The following script can read Ultrasonic binary files to generate B-scan images. This is developed by Tonni F. Johansen.

```
[file_pc,path_pc] = uigetfile('*.bin','Pitch catch file');
fid = fopen([path_pc,file_pc]);
data_pc = fread(fid,2e7,'uint16','b');
fclose(fid);
data_info = importdata([file_pc(1:end-3),'txt'],' ',0);
dt = data_info(2)*1e-9;
fs = 1/dt;
N_samp = data_info(3);
N_line = length(data_pc)/N_samp;
rfpc = reshape(data_pc,N_samp,N_line);
t = (1:N_samp)*dt;
[bpc,apc]=butter(2,[1e5 1e7]*dt/2);
rfpc_f = rfpc-mean(rfpc,2)*ones(1,N_line,1);
rfpc_f = filtfilt(bpc,apc,rfpc_f.*tukeywin(N_samp,0.1));
rfpc_fh = hilbert(rfpc_f);
figure
imagesc(1:N_line,t*1e6,20*log10(abs(rfpc_fh)))
colorbar
ylabel('t   (\mus)')
xlabel('Line no')
title(['Pitch catch data ',strrep(file_pc,'_','\_')])
set(gca,'YDir',"normal")
[cmi,cma]=caxis;
caxis([round((cma-50)/5)*5, cma])
[maxval,indpc_max] = max(sum(abs(rfpc_fh)));
[minval,indpc_min] = min(sum(abs(rfpc_fh)));
```

Appendix K – MATLAB script to extract features from ultrasonic data

The following script can read Ultrasonic binary files to detect envelope and extract features.

```
[file_pc,path_pc] = uigetfile('*.bin','Pitch catch file');
fid = fopen([path_pc,file_pc]);
data_pc = fread(fid,2e7,'uint16','b');
fclose(fid);
data_info = importdata([file_pc(1:end-3),'txt'],' ',0);

dt = data_info(2)*1e-9;
fs = 1/dt;
N_samp = data_info(3);
N_line = length(data_pc)/N_samp;
rfpc = reshape(data_pc,N_samp,N_line);
t = (1:N_samp)*dt;
[bpc,apc]=butter(2,[1e5 1e7]*dt/2);


rfpc_f = rfpc-mean(rfpc,2)*ones(1,N_line,1);
rfpc_f = filtfilt(bpc,apc,rfpc_f.*tukeywin(N_samp,0.1));

rfpc_fh = hilbert(rfpc_f);

for i = 1:863
    [up,lo] = envelope(rfpc_f(1:9500,i));
    %[up,lo] = envelope(abs(rfpc_fh(1:9500,i)));
    meanR(i) = mean(up);
    maxU(i) = max(up);
    stdU(i) = std(up);
    normStd(i) = std(up)/max(up);
    sumA(i) = sum(up,'all');
    sumAs(i) = (sum(up,'all'))^2;

end
exp06 = [meanR; maxU;  stdU; normStd;sumA;sumAs];
exp06 = exp06';
```

Appendix L – Python script to display image from ECT pixels

The following script can read ECT 'AIM' files and generates cross-sectional images of multiphase flow in the pipe. It uses Python Imaging Library library.

```
import numpy as np
import matplotlib.pyplot as plt
import os
import pandas as pd
import glob
from PIL import Image
fr = []
for k in range(5):
    data = np.loadtxt(r"C:\Users\Luna\Documents\USN\Thesis-ML\23-02-03\image\imageFiles\w2a0_15.aim",skiprows=2+34*k,max_rows=32)
    im_mx = data.reshape((32, 32))
    plt.imshow(im_mx,cmap='plasma')
    plt.savefig('%s'%k)
    fr.append(im_mx)
```

Appendix M – Box Plots

The following box plots from Figure M.1 to M.28 provide a visual display of the distribution of normalized capacitances for inter-electrode capacitances C12 to C78. These are divided into five classes of flow regimes – Annular, Plug, Slug, Stratified and Wavy.

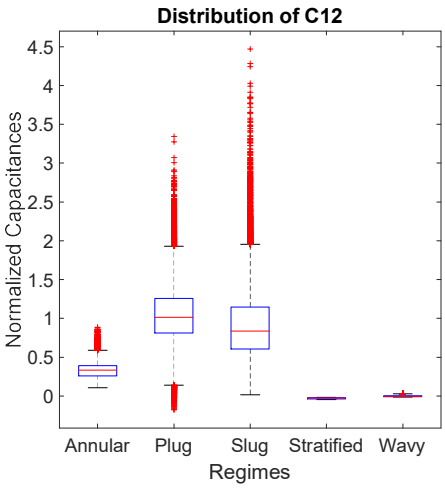


Figure M.1: Box Plot of C12

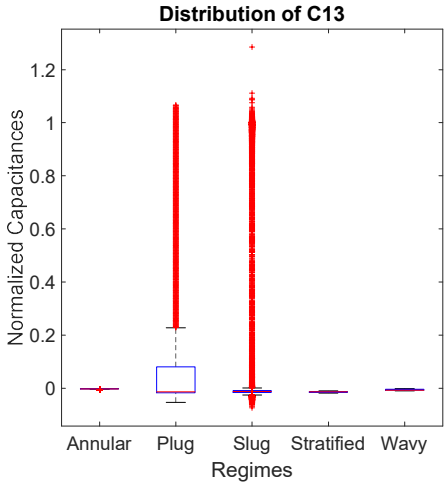


Figure M.2: Box Plot of C13

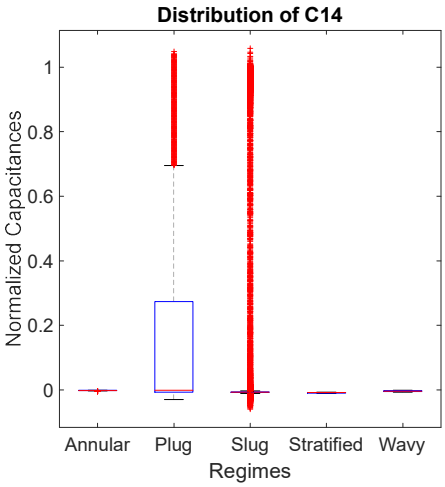


Figure M.3: Box Plot of C14

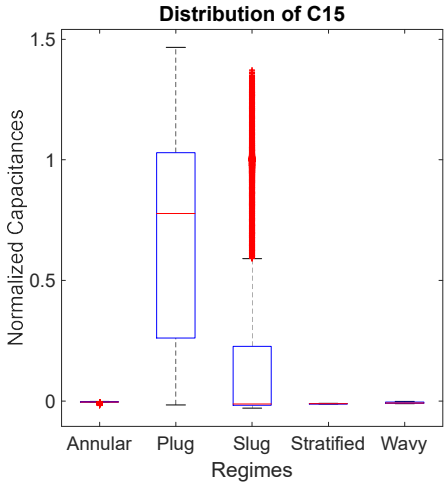


Figure M.4: Box Plot of C15

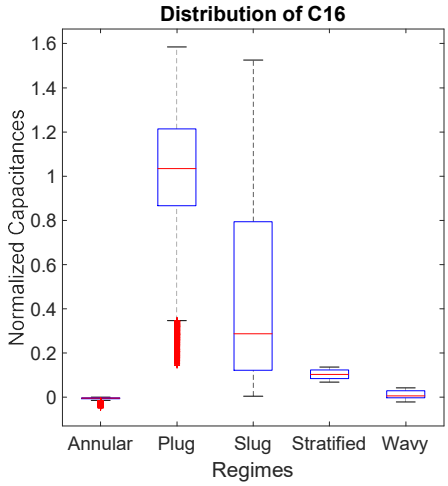


Figure M.5: Box Plot of C16

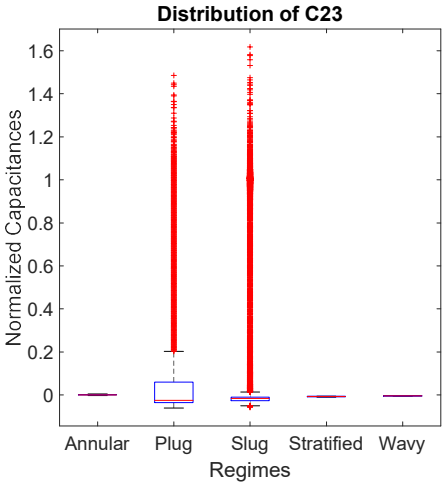


Figure M.6: Box Plot of C23

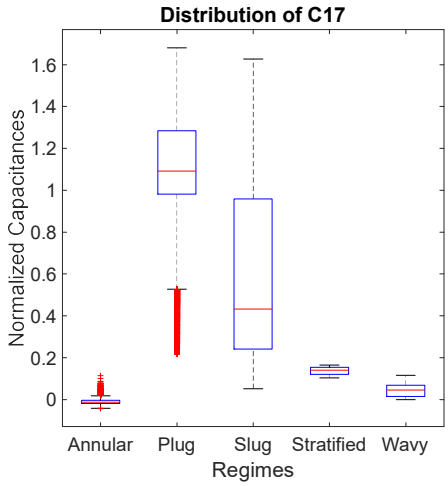


Figure M.7: Box Plot of C17

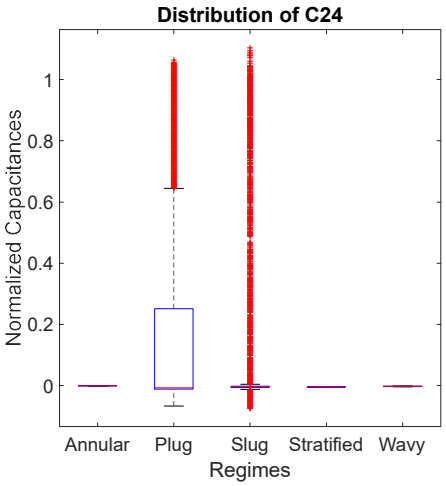


Figure M.8: Box Plot of C24

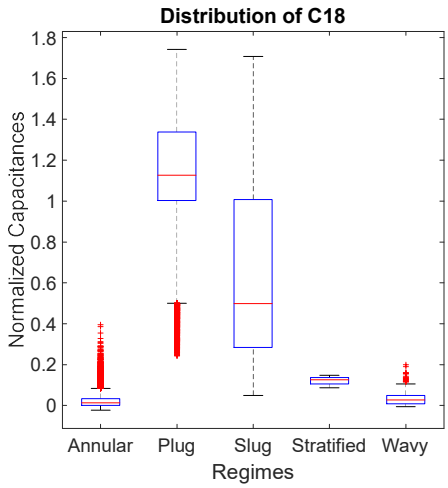


Figure M.9: Box Plot of C18

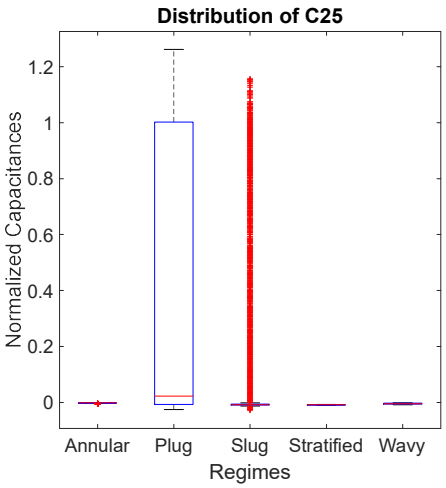


Figure M.10: Box Plot of C25

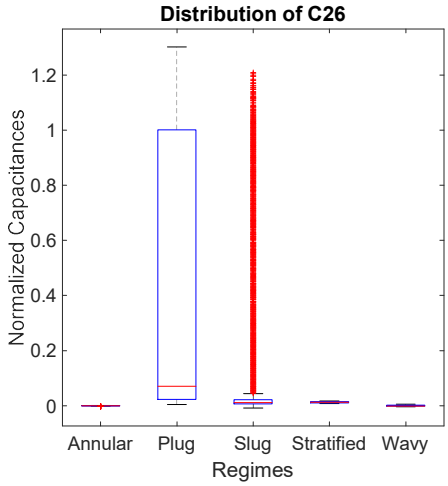


Figure M.11: Box Plot of C26

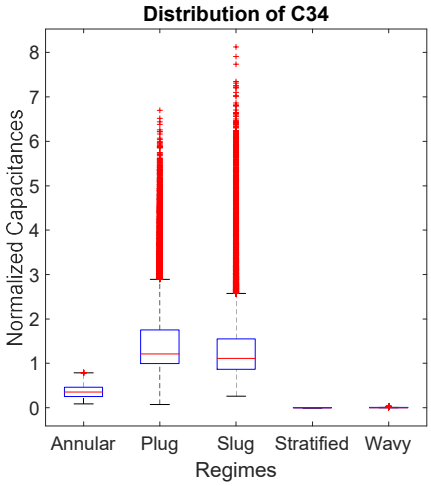


Figure M.12: Box Plot of C34

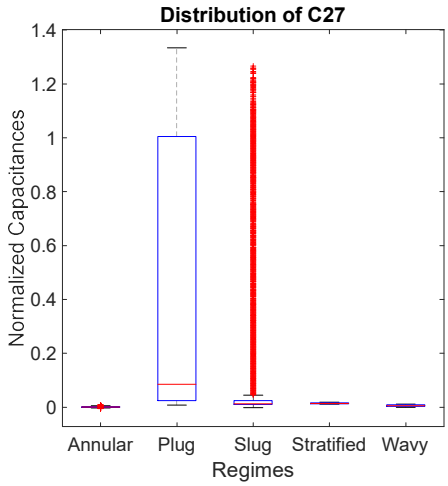


Figure M.13: Box Plot of C27

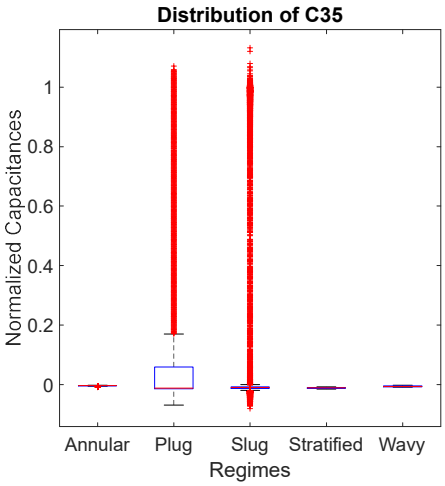


Figure M.14: Box Plot of C35

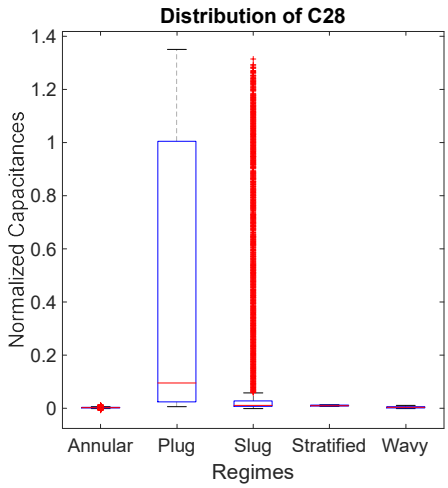


Figure M.15: Box Plot of C28

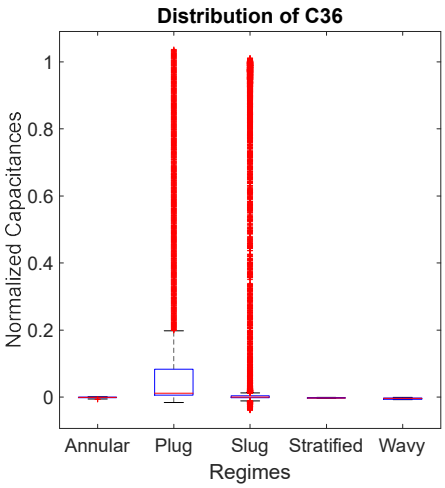


Figure M.16: Box Plot of C36

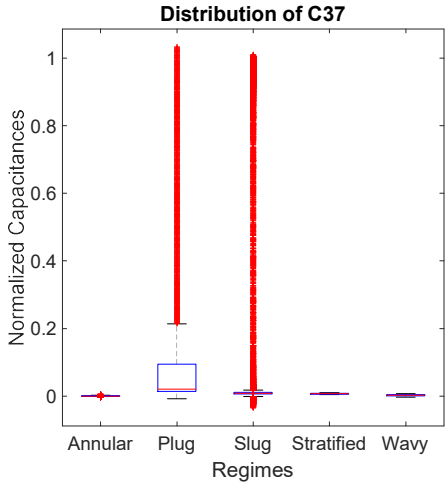


Figure M.17: Box Plot of C37

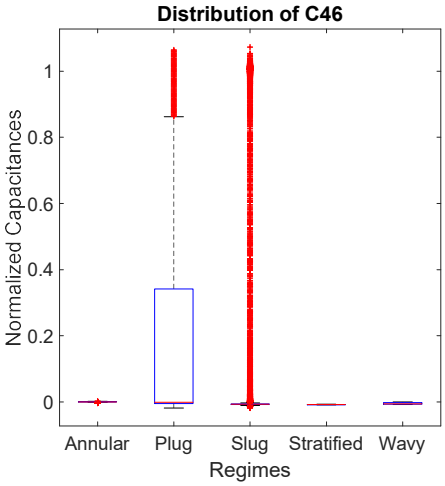


Figure M.18: Box Plot of C46

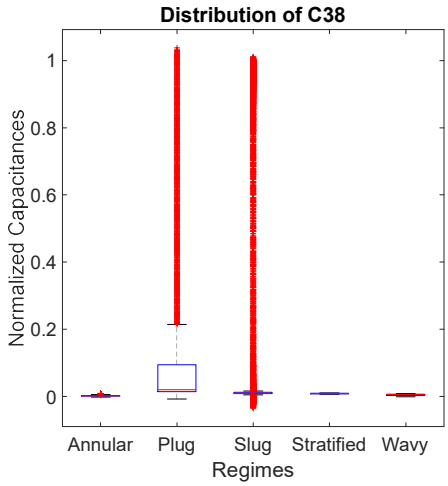


Figure M.19: Box Plot of C38

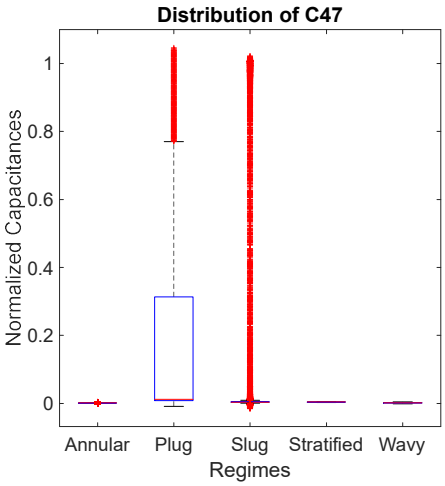


Figure M.20: Box Plot of C47

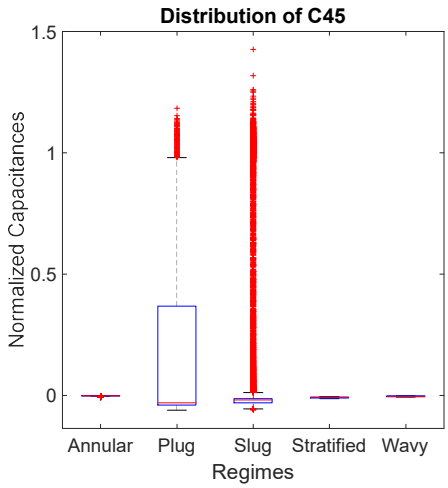


Figure M.21: Box Plot of C45

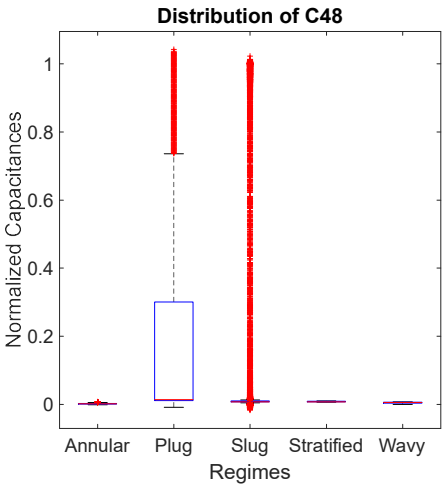


Figure M.22: Box Plot of C48

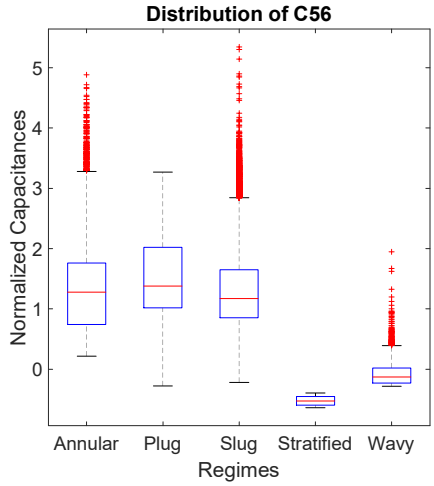


Figure M.23: Box Plot of C56

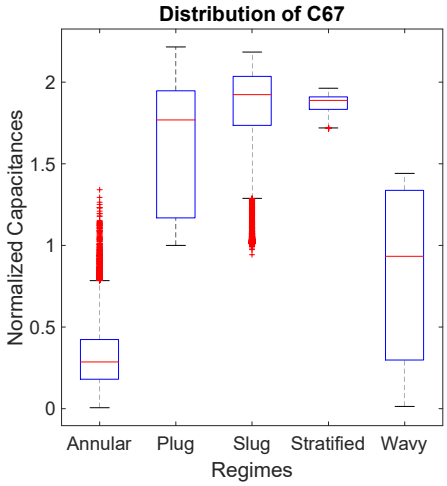


Figure M.24: Box Plot of C67

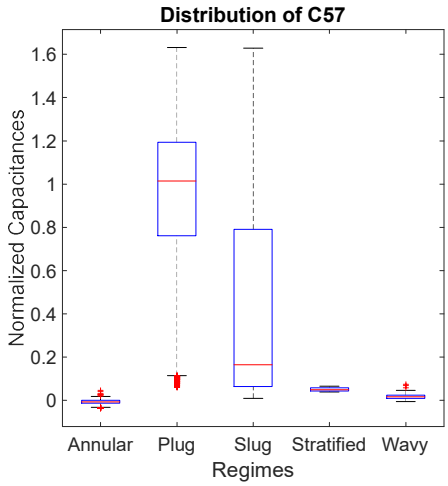


Figure M.25: Box Plot of C57

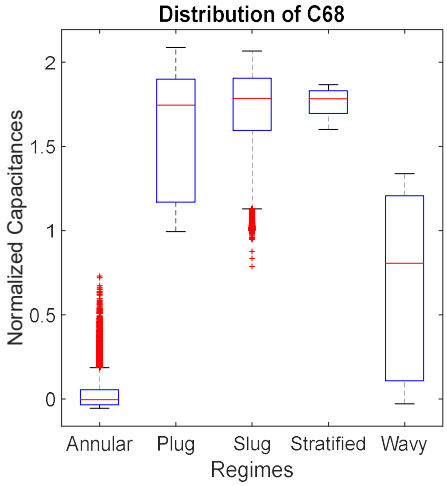


Figure M.26: Box Plot of C68

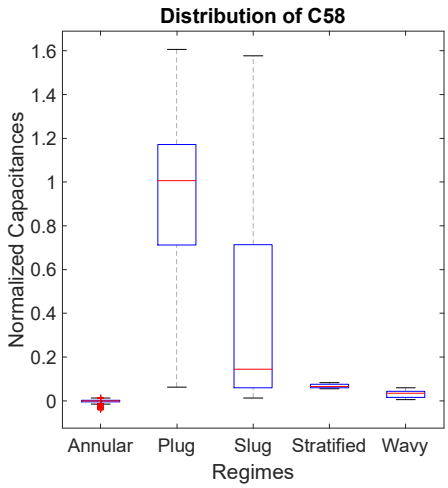


Figure M.27: Box Plot of C58

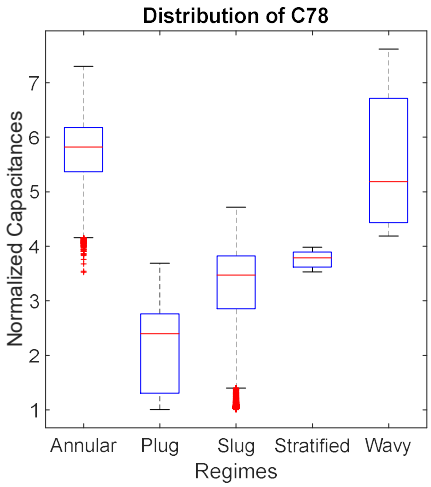


Figure M.28: Box Plot of C78

Appendix N – Python script to generate Correlation Coefficients

The following Python script generates correlation coefficients of 28 ECT normalized capacitances with respect to one other.

```
import seaborn as sns

import matplotlib.pyplot as plt

import pandas as pd

df = pd.read_csv(r"C:\Users\Luna\Documents\USN\Thesis-ML\boxPlots\P1withRegimes.txt",sep="\t")

df.corr()

corrMatrix = df.corr(method='pearson')

plt.figure(1,figsize=(10, 10),dpi=200)

plt.title('Correlation Coefficients Heatmap')

sns.heatmap(corrMatrix, annot=False)

plt.show()
```