MULTIPHASE FLOW MEASUREMENT USING CORIOLIS FLOWMETERS INCORPORATING SOFT COMPUTING TECHNIQUES

A Thesis Submitted to the University of Kent
For the Degree of Doctor of Philosophy
In Electronic Engineering

By

Lijuan Wang BEng MSc PhD

June 2017

Abstract

This thesis describes a novel measurement methodology for two-phase or multiphase flow using Coriolis flowmeters incorporating soft computing techniques. A review of methodologies and techniques for two-phase and multiphase flow measurement is given, together with the discussions of existing problems and technical requirements in their applications. The proposed measurement system is based on established sensors and data-driven models. Detailed principle and implementation of input variable selection methods for data-driven models and associated data-driven modelling process are reported.

Three advanced input variable selection methods, including partial mutual information, genetic algorithm-artificial neural network and tree-based iterative input selection, are implemented and evaluated with experimental data. Parametric dependency between input variables and their significance and sensitivity to the desired output are discussed.

Three soft computing techniques, including artificial neural network, support vector machine and genetic programming, are applied to data-driven modelling for two-phase flow measurement. Performance comparisons between the data-driven models are carried out through experimental tests and data analysis.

Performance of Coriolis flowmeters with air-water, air-oil and gas-liquid two-phase carbon dioxide flows is presented through experimental assessment on one-inch and two-inch bore test rigs. Effects of operating pressure, temperature, installation orientation and fluid properties (density and viscosity) on the performance of Coriolis flowmeters are quantified and discussed. Experimental results suggest that the measurement system using Coriolis flowmeters together with the developed data-driven models has significantly reduced the original errors of mass flow measurement to within ±2%. The system also has the capability of predicting gas volume fraction with the relative errors less than ±10%.

Acknowledgements

The author wishes to express grateful thanks to the following:

University of Kent

Prof. Yong Yan	My first supervisor, whose advice, encouragement and contributions made it possible for me to complete this work.
Dr Xue Wang	My co-supervisor, whose support was important in this research.
Dr Tao Wang	My co-supervisor, for his industrial expertise, support and invaluable advice drove the project to progress at all times.

KROHNE LTD for providing a PhD scholarship and in-kind support by supplying the Coriolis flowmeters and a number of other useful devices.

A number of academic staff, technicians and students at the University of Kent, Krohne Ltd, Tianjin University, and North China Electric Power University, in particular, Dr Lijun Sun, for their help during experimental trials.

Special thanks to the friends I've known here who kept me going, Dr Huiling Zhu, Dr Jingyuan Sun, Dr Mengyang Wei, Jingqiong Zhang and Jinyu Liu.

This thesis is dedicated to my whole family, especially my parents and my sister, for their encouragement, help and support throughout the project.

Content

Nomenclature

A	Cross-sectional area of the pipe [m^2]
b	Bias
B_s	Sensor balance
f	Tube frequency [Hz]
$\hat{f}$	Estimated probability density function
G	Relevance of the variable in explaining the output
I_D	Drive level [A]
K	Damping
K_h	Gaussian kernel function
L	Number of hidden nodes
m	Number of variables
M	Number of different trees
$\hat{m}$	Conditional expectation
n	Number of sample observations
q_m	Mass flowrate [kg/h]
$q_{m,g}$	Mass flowrate of gas [kg/h]
$q_{m,o}$	Mass flowrate of oil [kg/h]
$q_{m,w}$	Mass flowrate of water [kg/h]
q_v	Volumetric flowrate [m^3/h]
$\boldsymbol{s}$	Selected variable set
$\bar{\mathbf{s}}$	Mean of $\boldsymbol{s}$
S_a	Sensitivity
T	Process temperature [°C]

t_d	Time shift [s]
u	Residual information in potential input variables
v	Residual information in output
V_A	Sensor-A amplitude [V]
V_B	Sensor-B amplitude [V]
v_g	Instantaneous velocity of gas [m/s]
v_o	Instantaneous velocity of oil [m/s]
v_w	Instantaneous velocity of water [m/s]
x_i	The i^{th} input vector
x_i	The i^{th} input value
X*	Training samples
Y	Output vector
y_i	The i^{th} output value
$\bar{y}$	Mean of y_i
$\hat{y}_i$	The i^{th} value of the model output
Y_1	Desired liquid mass flowrate [kg/h]
Y_2	Desired gas volume fraction [%]
$\hat{Y}_1$	Estimated liquid mass flowrate [kg/h]
$\hat{Y}_2$	Estimated gas volume fraction [%]
$\hat{\boldsymbol{z}}_i$	The i^{th} value of the model output after small changes added to the inputs
$\lvert Z \rvert$	Magnitude of impedance
α_g	Cross-sectional fraction of gas [%]
α_o	Cross-sectional fraction of oil [%]

α_w	Cross-sectional fraction of water [%]
θ	Phase factor of impedance
ρ_g	Density of gas [kg/m^3]
ρ_o	Density of oil [kg/m^3]
ρ_w	Density of water [kg/m^3]
ω	Weight
Ω	Number of non-terminal nodes in the tree

List of Abbreviations

AIC	Akaike Information Criterion
ANFIS	Adaptive Network based Fuzzy Inference System
ANN	Artificial Neural Network
AWNN	Adaptive Wavelet Neural Network
BP-ANN	Backpropagation Artificial Neural Network
CCS	Carbon Capture and Storage
DP	Differential Pressure
ELM	Extreme Learning Machine
EU-ETS	European Union-Emissions Trading Scheme
GA-ANN	Genetic Algorithm-Artificial Neural Network
GAGRNN	Genetic Algorithm and General Regression Neural Network
GP	Genetic Programming
GVF	Gas Volume Fraction
ICA	Independent Component Analysis
IIS	Tree-based Iterative Input Selection
MISO	Multiple Inputs Single Output
NRMSE	Normalised Root-Mean-Square Error
PCA	Principal Component Analysis
PCIS	Partial Correlation Input Selection
PMI	Partial Mutual Information
SISO	Single Input Single Output
SOM	Self Organizing Map
SVM	Support Vector Machine

List of Tables

List of Figures

Chapter 1

Technical Requirements for Multiphase Flow Measurement

1.1 Introduction

Multiphase flow is defined as a simultaneous flow of materials with two or more different phases (i.e. gas, liquid or solid) or unseparated components (e.g. water and oil) [1]. Multiphase flow (including two-phase flow which is a particular example of multiphase flow) is widely seen in many industries, such as food, energy, chemical and transport industries. There are four basic types of multiphase flows: gas-liquid flow (in boilers, nuclear reactors and pump cavitation etc.), gas-solid flow (in fluidized beds and pneumatically conveyed pipelines etc.), liquid-solid flow (in river bed sediments and coal-water slurry etc.) and liquid-liquid flow (in oil production and chemical emulsions etc.). In such industrial processes, accurate measurement of multiphase flow is highly desirable to realize flow quantification, operation monitoring, process optimization and product quality control. This research focuses on the solutions to the measurement problems of multiphase flow in the process industry, including oil and gas applications and the Carbon Capture and Storage (CCS) chain.

In the oil and gas industry most of gas and oil reservoirs naturally contain water. In some circumstances, additional water has to be injected into reservoirs to maintain the required pressure for production. Therefore, water occurs along with oil and gas in wellbores and hence multiphase flows in wells, transportation lines and risers. The problem of how to measure oil-gas-water mixtures has been of interest to the petroleum industry since the early 1980s [2]. Despite over three decades of intensive research and development worldwide, accurate measurement of the flowrate of oil-gas-water mixture in a pipeline still remains challenging [3]. One of the original solutions is to separate the components of the mixed flow and then measure the flowrate of each individual phase

using traditional single-phase instruments. However, separation systems in practical applications have the disadvantages of large volume, high cost, low efficiency and difficult installation and maintenance. As a result of the limitations of separation techniques, some on-line multiphase flowmeters such as those based on radiation sensor (e.g., gamma ray) and cross correlation techniques have been developed for specific applications. However, the available multiphase flowmeters still have limitations to some extent in terms of accuracy, range, reliability, applicability and cost. In addition, accurate mass flow metering is becoming more and more essential in the custody transfer of marine fuel (bunkering) in coastal ports around the world, as air entrainment is unavoidable particularly during the start and stop processes of the bunkering. For these reasons, new techniques are increasingly desirable for multiphase flow metering in the oil and gas industry and bunkering centres, especially for fiscal purpose, oil well monitoring and allocation metering.

Measurement of CO_2 flows in the CCS chain is also a challenging and pressing problem with the recent development and deployment of the CCS scheme in many countries. CCS is considered as an effective technology to reduce CO_2 emission from electrical power generation and other industrial processes. Accurate measurement of captured CO_2 is necessary not only for environmental purposes to detect CO_2 leakage, but also for verification of the CO_2 account under emissions trading schemes. To put the importance of accurate flow measurement into perspective, take the UK's largest emitting power station Drax for example, it emits approximately 22 million tonnes of CO_2 per annum. Every 1% uncertainty in flow measurement could result in a £6.6 million financial exposure in the trading scheme (based on a carbon trading price of £30/tonne, as projected between 2013 and 2015) [4]. Moreover, CO_2 flow in the CCS chain is more difficult to measure than other multiphase flows as the physical properties and flow regimes of CO_2 could change significantly through long-distance transportation, especially with the variations of environmental temperature and pressure or in the presence of inevitable impurities (N_2 and CH_4). Therefore, CO_2 flow measurement issues and challenges are arising in recent years, however, there has been no significant research undertaken in this area. Despite the traditional single-phase flowmeters, orifice plate meters and turbine meters, have been used in some CCS plants, none of the current techniques is able to meet the specified 1.5% measurement

uncertainty in the EU-ETS (European Union - Emissions Trading Scheme) under all expected flow conditions. For these reasons, new techniques for multiphase CO_2 flow measurement are highly desirable.

Coriolis flowmeters are one of the most accurate single phase mass flowmeters. The main advantage of Coriolis flowmeters is that they provide directly a measurement of mass flow rate, regardless of the variations of temperature, pressure and flow profile. Potential application of Coriolis flow metering technology to multiphase flow in the oil and gas industry and CCS systems is a recent trend in the field of flow measurement and instrumentation. With the rapid development of artificial intelligence and machine learning, soft computing techniques provide new and effective approaches to engineering problems. This research programme aims to combine Coriolis flow metering and soft computing techniques to determine the mass flowrate and volume fractions of multiphase flow. The proposed measurement methodology is to be applied and verified through metering air-water, air-oil and gas-liquid CO_2 flows.

1.2 Technical Challenges

Multiphase flow is very complex and difficult to understand, predict and model due to the simultaneous presence of different phases or components in the same stream and their interactions within the pipeline. Therefore, the measurement of multiphase flow faces a number of technical challenges. This study focuses on the measurement of gas-liquid multiphase flow in the oil and gas industry and in the CCS systems. The main technical difficulties to fulfil the requirement have been identified in the following sections, respectively.

1.2.1 Challenges in Multiphase Flow Measurement in the Oil and Gas Industry

An ideal multiphase flowmeter for the oil and gas industry should meet the measurement requirements in terms of accuracy, range, ruggedness and reliability etc.

However, there have been very few commercial flowmeters that can meet all the identified requirements.

- High accuracy for the intended application

The required accuracy depends on the specific application and local conditions. In general, more accurate flowmeters are required for the purpose of custody transfer. In some circumstances, oil can be economically produced in some sites with water fractions above 90%. In this case the water fraction has to be accurately measured in order to optimize the production of the field. Nowadays, the relative uncertainties of ±5% for liquid flowrate, ±5% for gas flowrate and an absolute uncertainty of ±2.5% for water cut are typically achievable by commercial flowmeters used in many industrial applications [3].

- Validity over the full fraction range of each phase

Being able to measure over the full fraction range of each phase is a challenge for many current commercial multiphase flowmeters, practically in some extreme cases. For example, when the gas volume fraction is greater than 95% the flow is regarded as wet gas. On the contrary, water could be the main stream with high water cut in a mixed flow. In order to deal with these situations, many manufacturers produce specialized wet gas flowmeters or high water cut flowmeters based on different techniques.

- Measurement independent of flow regime

The flow regime of multiphase flow depends mainly on the individual phase flowrates, phase properties, local operating conditions and the pipeline geometry and orientation. As the flow regime present in a pipeline is difficult to control or reliably predict, an ideal multiphase flowmeter should be able to achieve the same measurement accuracy regardless of variations in flow regime. However, many phase fraction measurement techniques, such as electrical impedance methods, are flow regime dependent which brings challenges to the development of multiphase flowmeters.

1.2.2 Challenges in CO_2 Flow Measurement in the CCS Chain

In addition to the measurement issues and challenges specific to different types of flowmeter technologies, a number of problems arise in the measurement of CO_2 flow throughout the CCS chain due to the special characteristics of CO_2 and the process and plant conditions.

- Physical properties of CO_2 flows in the CCS chain

It can be seen from the CO_2 phase diagram in Figure 1.1 that transition boundaries between CO_2 phases are very close and lie around ambient conditions. Unlike oil and water which are transported at pressures and temperatures much lower than their critical points, very small changes in temperature and pressure can lead to rapid and substantial changes in the CO_2 physical properties. This makes CO_2 more difficult to control and measure, as many industrial processes operate around ambient conditions. Moreover, pipelines are considered to be the most efficient and viable method for onshore transportation of high volume of CO_2 from capture facilities to storage sites through long distances. This also imposes significant challenges to the measurement and control of CO_2 flows in CCS pipelines. In addition, some constituents such as N_2 and CH_4 are very likely mixed in the captured stream, which makes pure CO_2 phase diagram and state equations highly unreliable and results in a high likelihood of two phase or multiphase flow conditions. This will affect the operating conditions, economics and flow metering, since the majority of flowmeters can operate only under single phase conditions. In this case, the desired flowmeter should be able to measure either single phase or multiphase flows with satisfactory accuracy. Furthermore, any substantial contamination of CO_2 with water will give rise to acid formation, leading to corrosion of flowmeters.

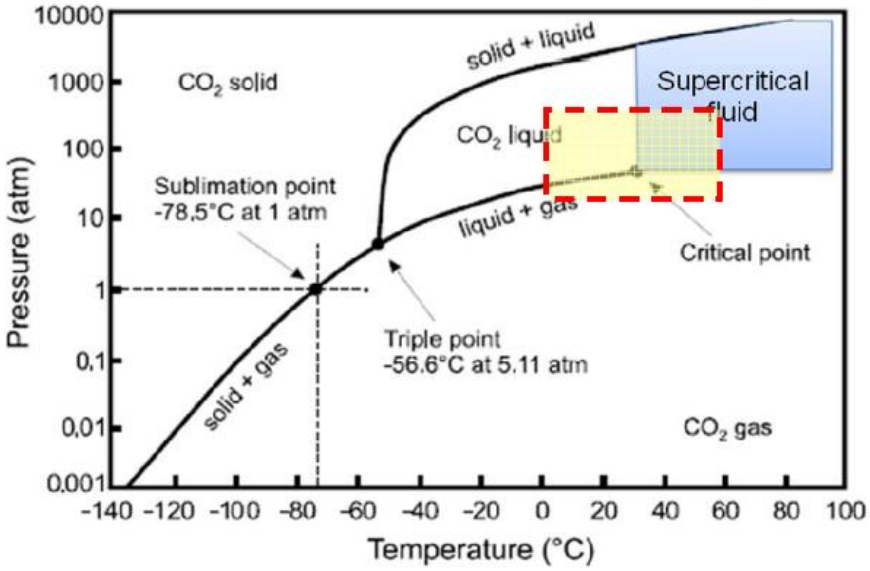


Figure 1.1 CO_2 phase diagram [4]

- Performance requirements for multiphase flowmeters

The measurement of CO_2 flow across the installation boundary, into the pipeline network, and from the pipeline head into the storage reservoir is necessary using flowmeters designed to custody transfer standard. The metering process itself is required to produce data with a measurement uncertainty within 2% [4]. To avoid the potential misallocation and measurement of fluid density, a mass flowmeter is more preferred to a volumetric flowmeter. Transportation pipelines through the CCS chain are normally in hundreds of millimetres in diameter to transport large volumes of CO_2 economically, which restricts the application of small-diameter flowmeters. The installation of the desired flowmeter should be as simple and straightforward as possible to allow retrofitting and to avoid significant pressure drop of CO_2 flows.

- Test and calibration facilities

So far there are very few CO_2 flow test and calibration facilities that can recreate CCS conditions particularly two-phase CO_2 flow in pipelines together with accurate measurement standards. Trials may be undertaken in water, oil and gas, but such deviations from realistic CCS conditions may have a detrimental effect and impact on measurement uncertainty. Therefore, development of suitable CO_2 test and calibration facilities may be necessary to test CO_2 multiphase flowmeters and support the deployment of CCS measurement schemes.

1.3 Research Objectives

This research programme aims to develop a measurement methodology for two-phase or multiphase flow using Coriolis flowmeters together with soft computing techniques. The objectives of the research programme are defined as follows:

- To review and define the state-of-the-art of multiphase flow measurement. Existing techniques for multiphase flow measurement using Coriolis flowmeters and soft computing techniques will be reviewed and gaps that require further research in the field identified.
- To identify the significant parameters from Coriolis flowmeters and additional sensors that are suitable for the measurement of multiphase flow. Advanced input variable selection methods will be employed to process the experimental data. Parametric dependency between input variables and their significance and sensitivity to the desired output will be quantified.
- To develop data-driven models for the estimation of mass flowrate and phase fractions. Soft computing techniques will be exploited to develop data-driven models. Performance comparisons between the data-driven models will be conducted in terms of accuracy, robustness, generalization ability and computing complexity.
- To evaluate the proposed measurement methodology through extensive experimental tests under a wide range of conditions. The effectiveness of Coriolis flowmeters incorporating soft computing data models for the measurement of two-phase flow will be assessed.

1.4 Thesis Outline

The contributions of this thesis to the state-of-the-art in multiphase flow metering include (1) investigations into the performance of Coriolis flowmeters under different two-phase flow conditions, (2) identification of parameter dependency, significance and sensitivity of the variables from Coriolis flowmeters and additional sensors for multiphase flow metering using input variable selection methods, (3) use of soft

computing techniques including ANN (Artificial Neural Network), SVM (Support Vector Machine) and GP (Genetic Programming) to develop data-driven models, and (4) experimental evaluation of the proposed measurement methodology with air-water, air-oil and gas-liquid CO_2 flows.

This thesis is organised in seven chapters as follows:

- Chapter 1 introduces the importance of multiphase flow measurement, covers the technical requirements in the measurement of multiphase flow in the oil and gas industry and CO_2 flow in the CCS chain, and outlines the proposed objectives of the research programme.
- Chapter 2 reviews the state-of-the-art techniques for multiphase flow measurement using Coriolis flowmeters and soft computing techniques.
- Chapter 3 provides the principle of the proposed measurement methodology, related input variable selection methods and soft computing techniques.
- Chapter 4 presents and discusses the experimental results obtained from a 1-inch air-water two-phase flow test rig.
- Chapter 5 reports and analyses the experimental results acquired from 2-inch air-water and air-oil two-phase flow test facilities.
- Chapter 6 provides and discusses the experimental results obtained from a $1/2$ inch gas-liquid two-phase CO_2 flow test rig.
- Chapter 7 draws conclusions from the work that has been presented and provides suggestions and recommendations for future work.

Chapter 2

Review of Soft Computing Techniques for Multiphase Flow Measurement

2.1 Introduction

The study of multiphase flow measurement has a long history and involves different areas from fundamental laboratory studies to industrial applications. The literature survey does not only provide necessary background knowledge for this work but also assists in demonstrating a clear contribution of this work to the state-of-the-art of multiphase flow metering. Substantial improvements have been made to develop effective and efficient techniques that may offer solutions to the measurement of multiphase flow over the past few decades, experimentally or theoretically. Different techniques have been developed to measure phase flowrate and phase fractions of multiphase flow. However, there still exists some limitations to meet all the requirements in the oil and gas industry and the CCS chain.

This chapter is organised as follows. Firstly, the chapter begins with a review of existing multiphase flow measurement systems and techniques for the measurement of phase flowrate and phase fractions. The basic principle, advantages and limitations of each technique are described. A comprehensive survey about three-phase flow measurement in the petroleum industry has been conducted by Thorn [2].

Secondly, this chapter reviews the soft computing techniques which are used for two-phase or multiphase flow measurement. Published research work and design of industrial Coriolis flowmeter have been reviewed in details by Wang and Baker [5]. This part of review focuses on the correction methods for Coriolis flowmeters under

aeration conditions. In addition, the input variable selection methods for data-driven models are summarized and reported.

Lastly, the chapter concludes with a summary on the various techniques reviewed and highlights the novelties of this research programme and why these techniques used in the current research are decided upon.

2.2 Multiphase Flow Measurement Techniques

Flowrates of the individual phases, including volumetric flowrate (q_v) or mass flowrate (q_m), are usually used to characterise a multiphase flow. So far there is no ideal flowmeter which is able to directly measure the mixed and individual mass flowrate in an industrial environment. Alternatively, an inferential measurement method is commonly used to acquire the instantaneous velocity and cross sectional fraction of each component. Take oil-gas-water three-phase flow as an example, the mass flowrate of individual component and the total mass flowrate of the mixture can be determined by:

$$q_m = q_{m,g} + q_{m,w} + q_{m,o} = A(\alpha_g v_g \rho_g + \alpha_w v_w \rho_w + \alpha_o v_o \rho_o) \tag{2-1}$$

$$\alpha_g + \alpha_w + \alpha_o = 1 \tag{2-2}$$

where q_m is the total mass flowrate of the mixture, $q_{m,g}$, $q_{m,w}$ and $q_{m,o}$ are individual mass flowrates of gas, water and oil phases, respectively. *A* is the cross-sectional area of the pipe. α_g , α_w and α_o are cross-sectional fractions of gas, water and oil components, respectively. v_g , v_w and v_o are instantaneous velocities of gas, water and oil phases, respectively. ρ_g , ρ_w and ρ_o are the densities of the phases, respectively.

Since the densities of oil, water and gas components can be made available from other production processes, at least five variables, including oil, water and gas velocities and

two of the component fractions (usually gas void fraction α_g and water fraction α_w), are to quantify in equation (2-1).

During the past several decades, significant efforts have been made to develop reliable multiphase flow measurement techniques and instruments. Thorn et al. [2],[3] and Falcone et al. [6] reviewed the developments of three-phase flowmeters, particularly for the petroleum industry. In general, multiphase flow measurement systems can be classified into three main categories: separation systems, on-line sampling systems and on-line multiphase flowmeters.

2.2.1 Separation Systems

As shown in Figure 2.1, the conventional production or well-test systems separate the oil-gas-water three-phase flow into pure gas, oil and water single-phase flows according to the differences in fluid properties [7],[8]. Then traditional single-phase flowmeters are used for each phase measurement, such as orifice plates or Coriolis flowmeters for liquid flow and ultrasonic, vertex or Coriolis flowmeters for gas flow. The separation systems are commonly applied for custody transfer purpose due to high measurement accuracy required. However, the disadvantages of this kind of system include large volume, high cost, low efficiency and difficult installation and maintenance. The commercial Accuflow 3 Phase System [9] presents an example of such three-phase separation systems.

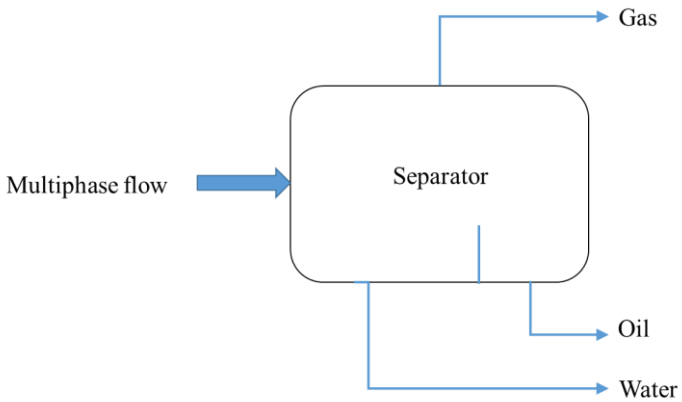


Figure 2.1 Principle of a production or well-test separation system in the oil and gas industry

Two-phase separation flowmeters normally contain a complete separation system or a partial separation system. It can be seen from Figure 2.2 that the complete separation system divides the three-phase flow into liquid (oil plus water) and pure gas streams before measurement. Once the liquid and gas are separated, each stream is measured independently using industry proven measurement devices. A conventional liquid meter (Coriolis and turbine etc.) is used to measure liquid flowrate. Water cut measurement can be obtained using traditional methods (density differential and microwave etc.). Gas flow is also measured using conventional technologies (vortex, turbine and ultrasonic etc.). After measurement, the gas and liquid streams are recombined and returned to the multiphase flow line. The commercial Accuflow JR Series Multiphase Metering system [10] and MB CCM Multiphase Meters [11] belong to the complete separation systems.

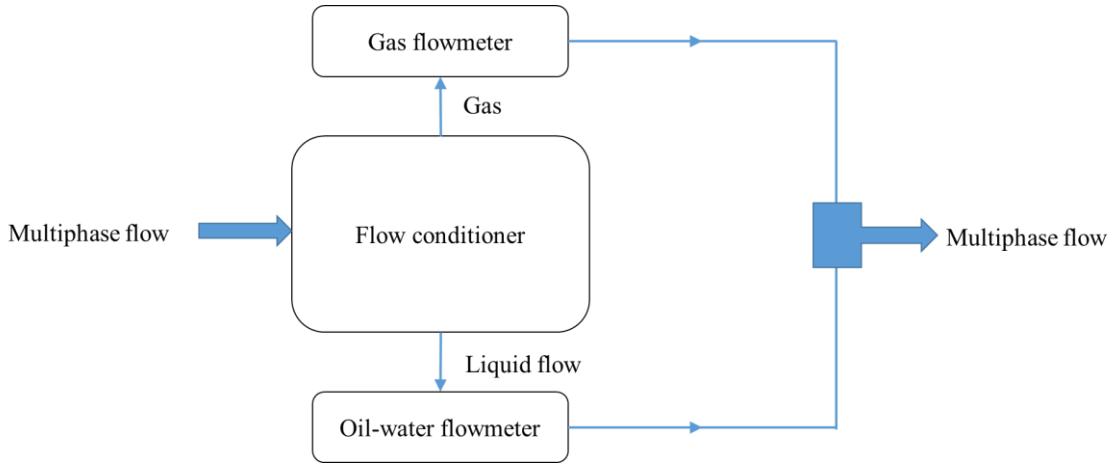


Figure 2.2 Principle of a complete separation system

Partial separation systems, such as Weatherford Alpha VSRD [12], partially separate the flow into wet gas and liquid dominant flow. The commercial multiphase flowmeter is based on an extended-throat Venturi meter, a sonar flowmeter, a Red Eye MP water-cut meter and a gamma densitometer and provides the measurement of wet-gas and liquid dominant multiphase streams.

The separation based multiphase flowmeters are capable of simplifying the multiphase metering process into single-phase or two-phase metering process and hence measurement accuracy mainly depends on the separation efficiency. Although this kind

of flowmeter is considered to be accurate and reliable, it also requires expensive and bulky separation equipment which significantly limits its usability.

2.2.2 On-line Sampling Systems

The extracting and separating method is a new generation of multiphase flowmeter. As shown in Figure 2.3, a small fraction of gas-liquid mixture which can represent the component of the total flow is extracted by a sampler and separated through a small separator. Once separated, gas flow and liquid flow are metered using corresponding flowmeters. Finally, the sampled fluid flows back to the main stream. The total gas and liquid flowrate is determined by the measurements of gas and liquid flowmeters and extraction ratio. The sampler is the most crucial device as it determines the relationship between the sampled fluid and the total gas-liquid mixture. Consequently, the related research is mostly concentrated on the structure design of the samplers [13]-[17] to make sure the extracted stream can represent the total stream.

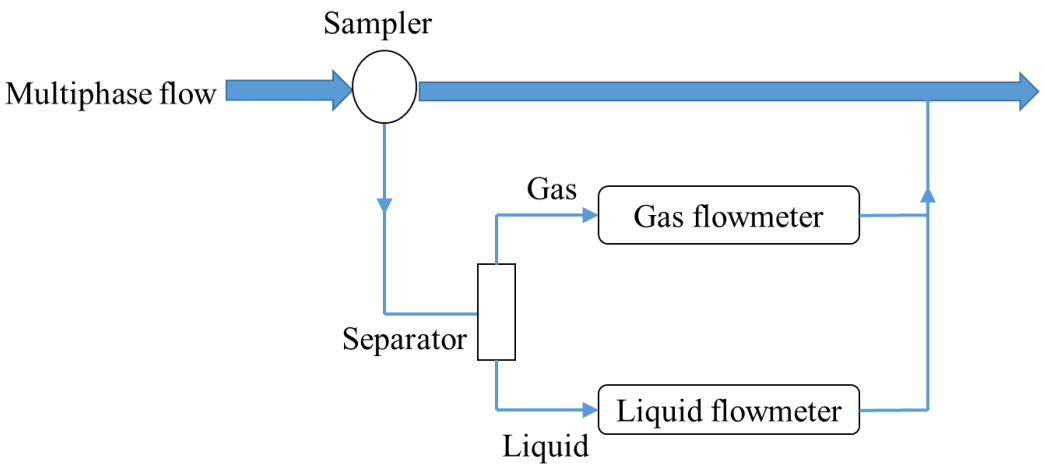


Figure 2.3 Principle of a sampling system

2.2.3 On-line Multiphase Flowmeters

On-line multiphase flowmeters are the devices to measure the mixed flow without any separators and sampling lines. They can be classified into direct and indirect measurement groups according to measurement strategies deployed.

- Direct approach

Direct measurement (Figure 2.4) aims to measure the phase velocity and phase fraction of each individual component of the mixture. Phase velocity is often determined by cross-correlation, Venturi flowmeter and Coriolis flowmeter, while phase fraction is usually measured based on gamma energy absorption, electrical impedance sensors and microwave techniques. The commercial three-phase flowmeters such as AGAR MPFM-50 [18], Multi Phase Meters AS MPM [19], Roxar Flow Measurement MPFM 2600 [20] and Subsea MPFM [21] are developed based on the combination of these techniques. In recent years, several techniques based on conductance [22]-[26], process tomography [27]-[37], ultrasonic [38]-[40] and nuclear magnetic resonance [41]-[42] have been proposed to measure phase velocity or phase fraction.

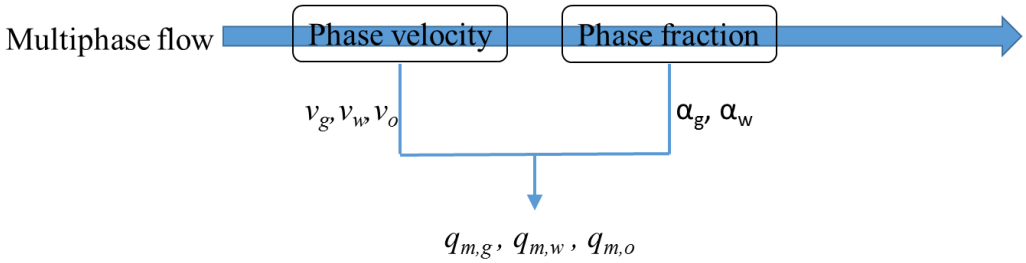


Figure 2.4 Principle of the direct approach to on-line multiphase flow measurement

- Indirect approach

Indirect measurement (Figure 2.5) usually determines the three phases from analysis of the time variant signals acquired from a set of sensors. In general, the relationship between sensor outputs and the flowrate or fraction of each phase cannot be deduced theoretically. In this case, empirical models are commonly developed from experimental data using statistical methods or soft computing techniques. The statistical methods such as dual regression, kernel partial least squares regression and multivariate

adaptive regression splines have been proposed to estimate the phase fraction or phase velocity of two-phase flow [43]-[45]. With the recent development of artificial intelligence and machine learning, soft computing techniques provide alternative approaches to traditional statistical methods and extend the capabilities of empirical models. The review of soft computing techniques for multiphase flow measurement is included in Section 2.3. The main limitation of this kind of method is the requirement for primary calibration on a flow test facility and secondary calibration using on-site field tests.

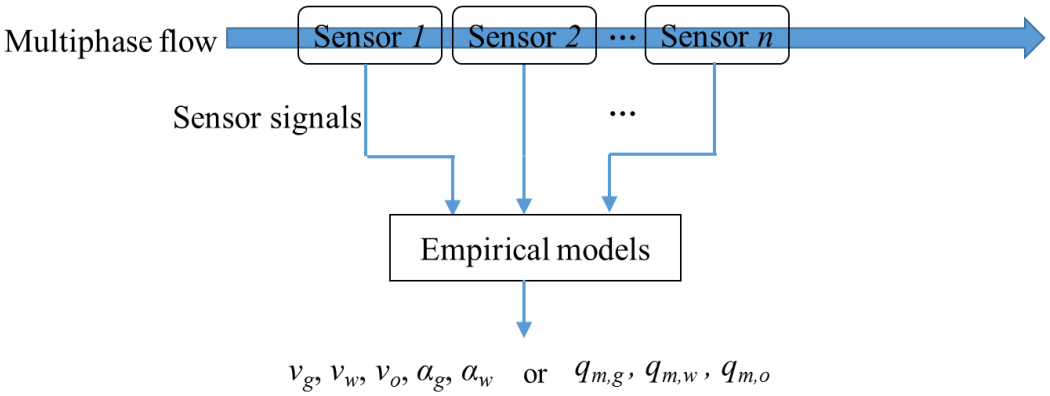


Figure 2.5 Principle of the indirect approach to on-line multiphase flow measurement

The developed multiphase flowmeters are mostly applied to oil/gas/water two-phase or three-phase flow measurement. There is no reported flowmeter specialized for CO_2 flow measurement in the CCS chain and the accuracy requirement from EU-ETS is challenging for the current flowmeters. With regard to CO_2 flow metering, Orifice plate meters and turbine meters have been employed in single-phase CO_2 measurement in Enhanced Oil Recovery (EOR) projects for many years [4]. However, the Orifice plate and differential pressure metering used for slugging two-phase mixture measurement at the well-head was reported to give errors up to 80% [46]. Coriolis flowmeters, as one of the most accurate single-phase mass flowmeters, have been tested with gas or liquid CO_2 single-phase flow [47]-[48]. For two-phase CO_2 flow measurement, commercial Coriolis mass flowmeters were field-tested with slugging two-phase CO_2 flow and the difference between the Coriolis flowmeters under test and the reference meter was 5% [46]. However, the effect of different flow patterns on the performance of Coriolis flowmeters was not reported.

2.3 Soft Computing Techniques for Multiphase Flow Measurement

Soft computing is described by L.A. Zadeh as: "*a collection of methodologies that aim to exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost*" [49]. As shown in Figure 2.6, the principal constituents of soft computing techniques are machine learning (including neural network and support vector machine, etc.), evolutionary computation (including evolutionary programming, genetic algorithm, evolution strategy and genetic programming), fuzzy logic and probabilistic reasoning (including Bayesian belief net and Dempster-Shafer theory, etc.). The first two are data-driven search and optimization approaches while the other two are based on knowledge-driven reasoning. In recent years, soft computing has become a promising tool in resolving engineering challenges due to its ability to handle highly complex, dynamic and non-linear problems and computational simplicity over analytical methods. For these reasons, they have been widely used in many fields, in particular, computer engineering, environmental engineering, material engineering and medical diagnosis [50].

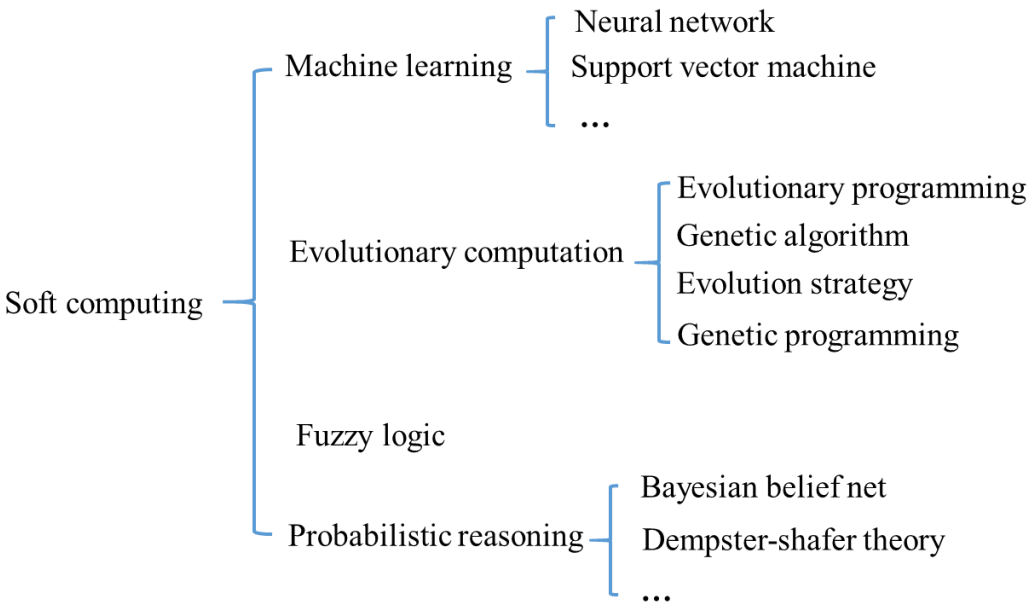


Figure 2.6 Principal constituents of soft computing techniques

The applications of soft computing techniques to multiphase flow measurement are mainly concentrated on the estimation of phase flowrates and phase fractions and identification of flow regime. The estimation of phase flowrates and fractions is equivalent to solve a problem of function approximation while the identification of flow regime is a classification problem. As the review focuses on the measurement of phase flowrates and phase fractions, the research purely on flow regime identification using soft computing techniques is out of the scope of this review. The following section reviews indirect multiphase measurement systems combining traditional sensors and soft computing techniques.

2.3.1 Ultrasonic Sensors

Figueiredo et al. [51] made use of four ultrasonic sensors incorporating an ANN to identify the flow pattern and obtain the gas volume fraction of two-phase flow, respectively. It can be seen from Figure 2.7 that the input of the ANN is comprised of energy ratios from four acoustic sensors. There are two hidden layers, including five and two hidden neurons respectively in the ANN model. The output layer generates the identified flow pattern or the estimated gas volume fraction. Experimental work with air-oil flow was conducted on 1-inch and 2-inch vertical test sections, respectively. During the experimental tests, dispersed bubbles, intermittent flow, churn flow and annular flow were observed with the liquid velocity ranging from 0.1 to 0.3 m/s and the gas volume fraction from 0 to 85%. Experimental results showed that the overall successful recognition rate of flow pattern was 98.3% while the overall variation of the estimated gas volume fraction was ±4.2.

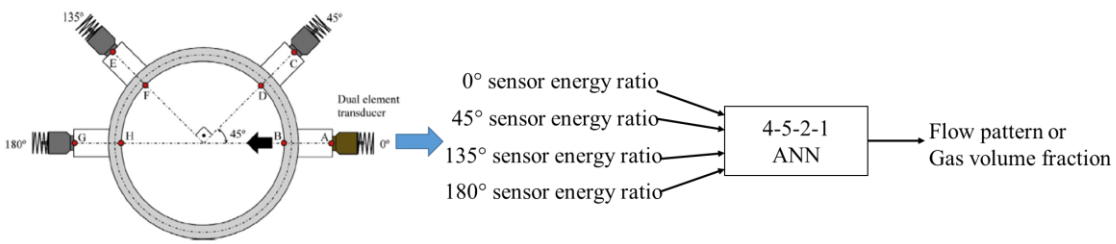


Figure 2.7 Measurement system based on ultrasonic sensors and ANN (Figueiredo et al. [51])

2.3.2 Differential Pressure Devices

Xu et al. [52] proposed a novel approach to the measurement of wet gas flow using a throat-extended Venturi meter and soft computing approximation techniques. The measurement system is shown in Figure 2.8. A backpropagation ANN (BP-ANN) model and a SVM model were developed to estimate the gas flowrate and liquid flowrate through the static and dynamic features of differential pressures. Experimental tests were carried out with natural gas and water two-phase flow on a 2-inch test rig. The gas flowrate was between 0.0139 and 0.0444 m^3/s and liquid flowrate ranging from 3.0556 $\times 10^{-4}$ to 0.0015 m^3/s. It was found that both ANN and SVM models were valid in approximation of the complex relationship between the signal features and the two-phase flowrates. With the usage of the BP-ANN, the mean prediction error and standard deviation were 3.14% and 4.22% for gas flowrate, respectively, whereas the mean and standard deviation were 4.77% and 6.33% for water flowrate, respectively. Through the SVM model, the mean and standard deviation of the relative prediction errors were 2.86% and 4.39%, respectively, for gas flowrate, whereas the mean and standard deviation were 4.25% and 6.09%, respectively, for water flowrate. In comparison with the ANN model, the SVM model was clearly of merit as the means of the relative prediction errors of the gas and water flowrates were improved by 8.9% and 10.9%, respectively.

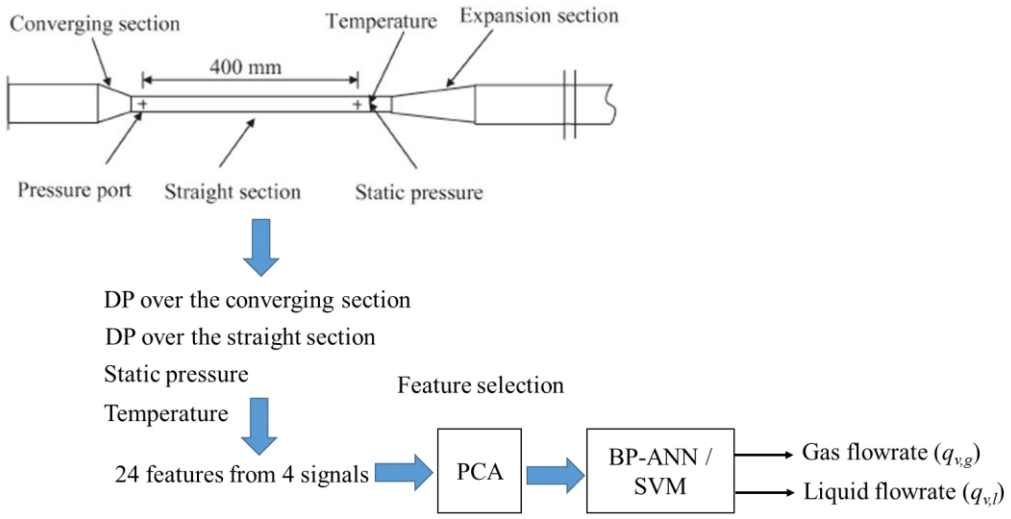


Figure 2.8 Measurement system based on a throat-extended Venturi meter with ANN/SVM (Xu et al. [52])

Shaban and Tavoularis [53] proposed a method for the measurement of gas and liquid flowrates in a vertical upward pipe using differential pressure (DP) signals. As shown in Figure 2.9, the probability density function and the power spectral density of the normalized DP signals were obtained and processed by principal component analysis (PCA) and independent component analysis (ICA). The two-phase flow regime was firstly identified through the application of the elastic maps method on the probability density function of the DP signals. Then a multi-layer BP-ANN taking the extracted features as inputs was developed for each flow regime (slug, churn and annular) to produce phase flowrates. Experimental tests were conducted with air-water in a vertical pipe of diameter 32.5 mm with air superficial velocity between 0.014 m/s and 22 m/s and liquid velocity from 0.04 m/s to 0.4 m/s. Experimental results suggested that the average relative errors of liquid flowrates for slug, churn, annular flow regimes were -0.3%, -0.1% and -0.4% respectively, and the average relative errors of gas flowrate were 5.5%, 0.5% and 0.6% respectively.

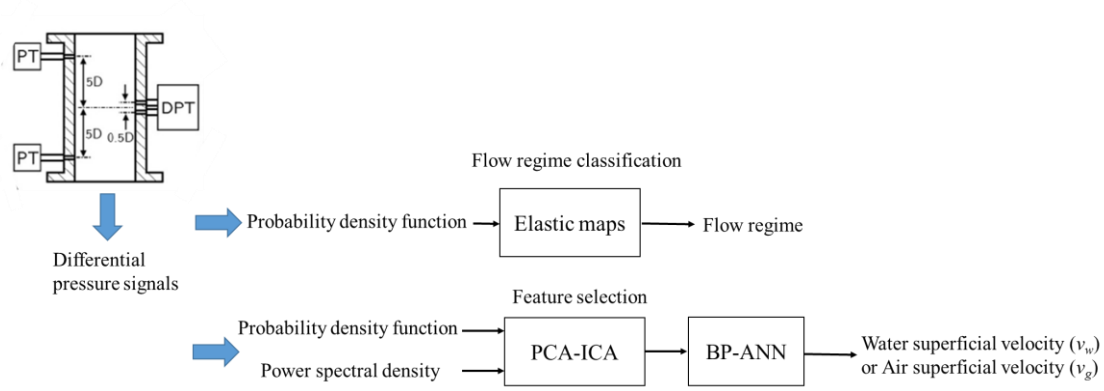


Figure 2.9 Measurement system based on a differential pressure transducer with elastic maps, PCA-ICA and ANN (Shaban and Tavoularis [53])

2.3.3 Electrical Sensors

Fan and Yan [54] presented an ANN based method to obtain gas and liquid flowrates of two-phase air-water slug flow in a 50 mm bore horizontal pipe through conductance probes. It can be seen from Figure 2.10 that five characteristic parameters of the mechanistic slug flow model, including translational velocity, slug holdup, film holdup,

slug length and film length, were extracted from the conductance signals and taken as inputs to the neural network. A feed forward neural network with ten hidden nodes was adopted to predict gas and liquid flowrates. Experimental assessment of the measurement system was conducted with air superficial velocity from 0.58 to 1.86 m/s and water superficial velocity between 0.35 to 1.62 m/s. Results suggested that the overall performance was within ±10% of full scale for the prediction of both liquid and gas flowrates.

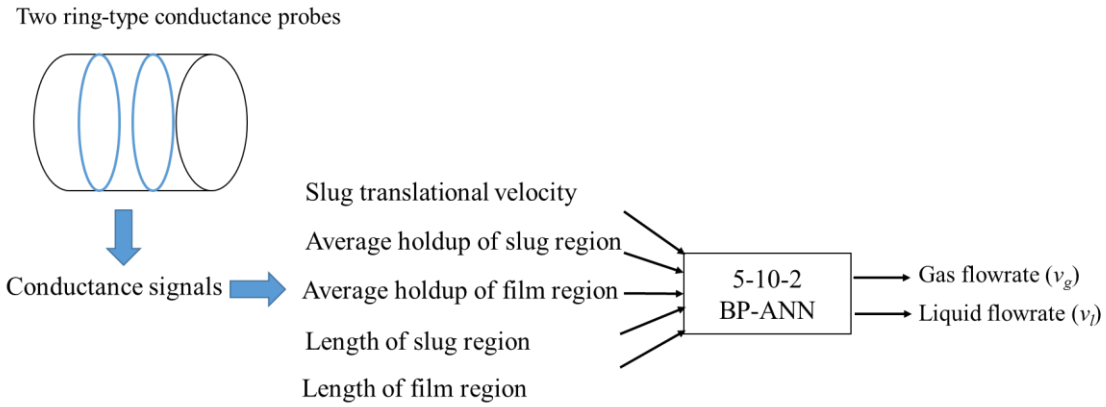


Figure 2.10 Measurement system based on two ring-type conductance probes and ANN (Fan and Yan [54])

2.3.4 Optical Sensors

Li et al. [55] applied a laser source, a 12×6 photodiode array sensor and an SVM model to quantify the void fraction of gas-liquid two-phase flow in small channels. As shown in Figure 2.11, the extracted features from the mean value and standard deviation of the 72 measured signals were taken as inputs of the SVM model. Through experimental tests with Nitrogen-water flow on a horizontal pipe with inner diameters of 4.22, 3.03, 2.16 and 1.08 mm, the flow patterns, including bubble flow, slug flow, stratified flow and annular flow, were observed. The maximum absolute error of the void fraction was found to be 7%.

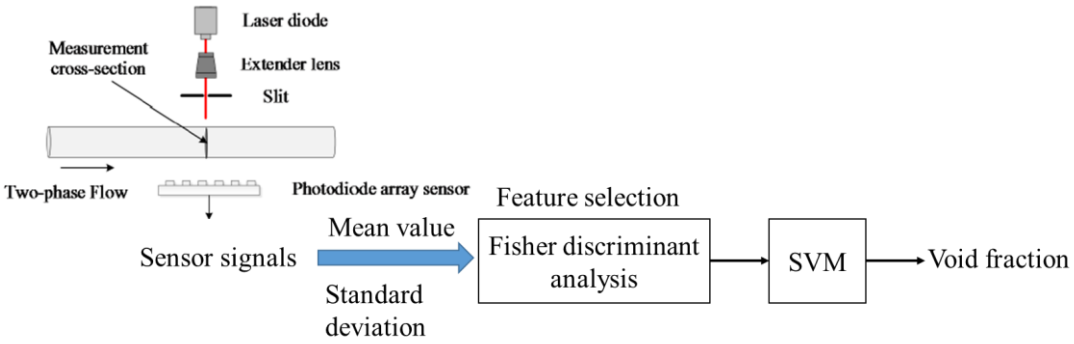


Figure 2.11 Measurement system based on a laser source, photodiode array sensor and SVM (Li et al. [55])

2.3.5 Combination of Multiple Sensors

Zheng et al. [56] proposed a measurement system using a turbine flowmeter, conductance sensors and SVM to identify the flow pattern and obtain the water cut of air-water two-phase flow in a vertical upward pipe with an inner diameter of 18 mm. As shown in Figure 2.12, the flow pattern was identified through chaotic attractor morphologic analysis of the conductance signals. The total flowrate of the mixture was obtained from the rotating speed of the turbine through a polynomial regression. To estimate the water cut of the mixed flow, a total of 10 features, extracted from fluctuant conductance signals in both time and frequency domains together with the average of rotating speed of the turbine were taken as inputs of the SVM model. The total flowrate of gas-water flow ranged from 0.1 to 2.5 m^3/h during the tests. Experimental results suggested that the success rate of the flow pattern identification was higher than 96% and the average errors of the water flowrate and gas flowrate measurements were 7.36%.

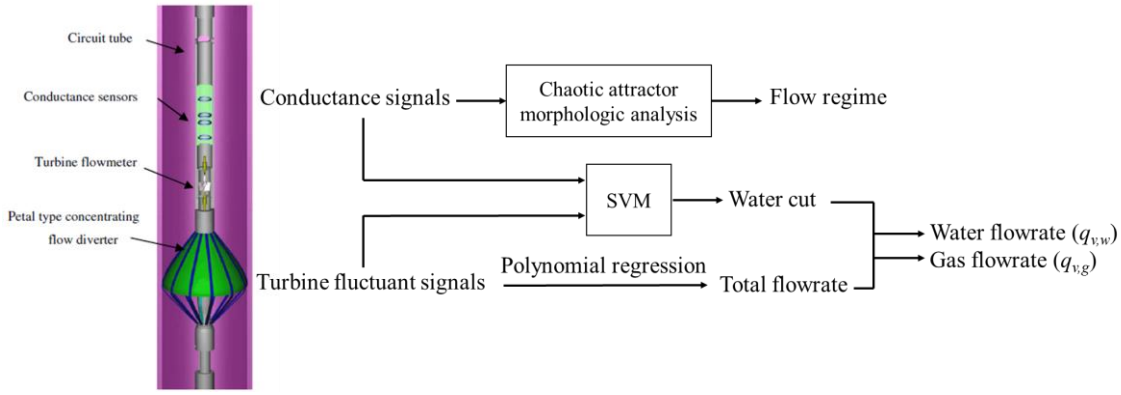
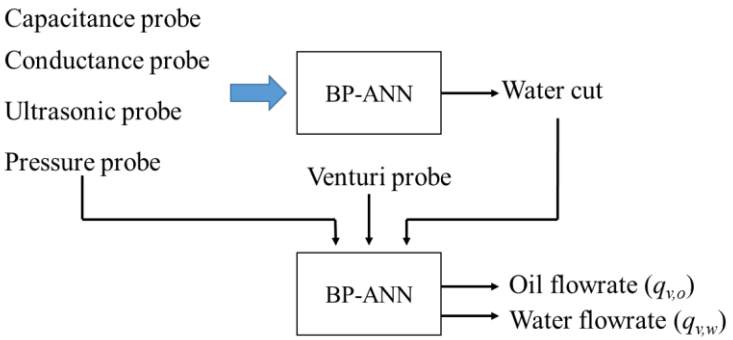
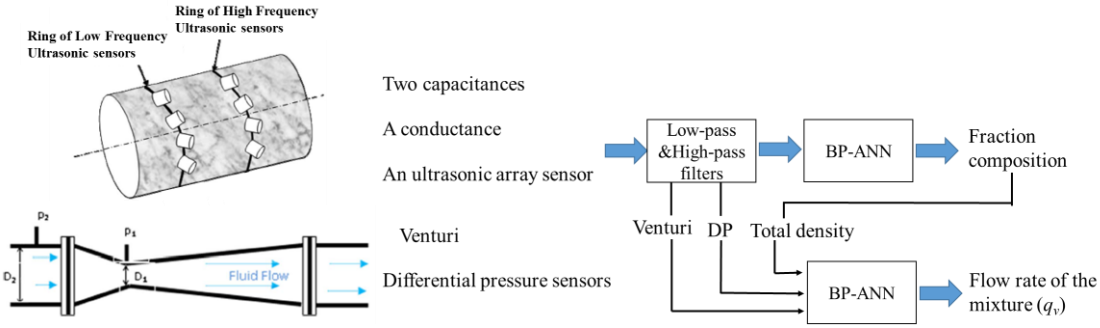


Figure 2.12 Measurement system based on conductance sensors, a turbine flowmeter and SVM (Zheng et al. [56])

Meribout et al. [57] integrated impedance measurements with ultrasonic measurements to provide the volumetric flowrate of oil-water flow with the water-cut ranging from 0 to 100%. As shown in Figure 2.13 (a), a pattern recognition algorithm based on an ANN was implemented. In the first stage, a feed forward three-layer ANN was developed with the signals extracted from the conductance, capacitance, ultrasonic and Venturi and an output of the mixture density. With the usage of the ANN, the water cut can be interpreted by the sensor signals. In the second stage, another three-layer ANN was established to obtain oil and water flowrates by combining the estimated water-cut from the first stage and the differential pressure and Venturi outputs. Experimental work was conducted on a two-inch oil-water test rig and the results showed that, for both water-cut and total flowrate determinations the relative error was less than 5% for any flow regimes. In subsequent research, they applied the same measurement method to oil-gas-water three-phase flow [58]. As shown in Figure 2.13 (b), two rings of high and low frequency ultrasonic sensors were used for low and high gas fractions, respectively. In this case, an ANN was developed in the first stage with the signals from capacitance, conductance, ultrasonic, pressure and Venturi sensors. Then the flowrate of the mixture was obtained using the estimated total density, differential pressure and Venturi outputs. The experimental results demonstrated that the average relative errors were 3.91% for water flowrate, 4.68% for gas flowrate and 6.2% for oil flowrate.



(a) Measurement system for oil-water two-phase flow measurement

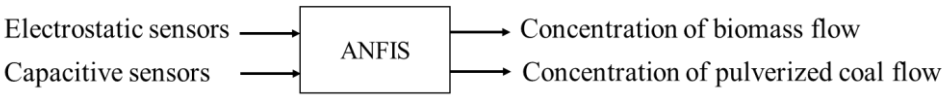


(b) Measurement system for oil-air-water three-phase flow measurement

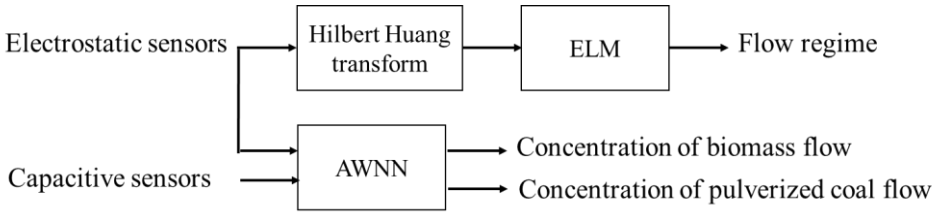
Figure 2.13 Measurement system based on capacitance, conductance, ultrasonic, DP sensors and a Venturi flowmeter (Meribout et al. [57], [58])

Wang et al. [59] proposed a data fusion method by combining capacitive and electrostatic sensors to realize online volumetric concentration measurement of pulverized coal/biomass fuel flow in co-fired power plants (See Figure 2.14). An adaptive network based fuzzy inference system (ANFIS) was developed through training with gradient descent method and hybrid method by combining the Kalman filter algorithm with a gradient descent algorithm. Experimental results on a 36 mm bore horizontal quartz glass pipe showed that the ANFIS based on the hybrid learning rule outperformed the system based on the gradient descent learning rule and the fiducial errors of biomass and pulverized coal flows were 1.2% and 0.7%, respectively. Following this research, an extreme learning machine (ELM) based on the electrostatic fluctuation signals was applied to identify the flow regime of coal/biomass/air three-phase flow and an adaptive wavelet neural network (AWNN) based on electrostatic and

capacitance sensors was created to predict the volume concentration of each phase [60]. Experimental work was carried out on a 94 mm bore horizontal quartz glass pipe. This methods yielded 2.1% fiducial error for biomass and 1.2% for pulverised coal. It was claimed that the training time, identification time and prediction time were much less than the other methods.



(a) Measurement system based on ANFIS



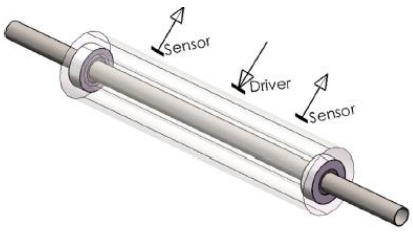
(b) Measurement system based on ELM and AWNN

Figure 2.14 Measurement system based on capacitive and electrostatic sensors (Wang et al. [59], [60])

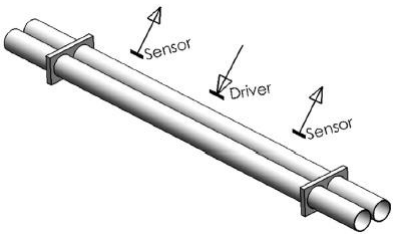
2.3.6 Coriolis Flowmeters

As shown in Figure 2.15, a Coriolis measuring system is a symmetrical design and consists of one or two measuring tubes (either straight or curved), a drive unit and two motion sensors. The driver sets the measuring tube into a uniform fundamental oscillation mode. Once liquid or gas flow moves through the measuring tube, Coriolis forces are generated in oscillating systems and result in an extremely slight distortion on the measuring tube. The distortion is picked up by motion sensors and the phase difference between the two sensor signals is proportional to the mass flowrate of the process fluid. The natural vibration frequency is directly related the density of the fluid. Since the oscillatory characteristics of the measuring tube are dependent on the fluid

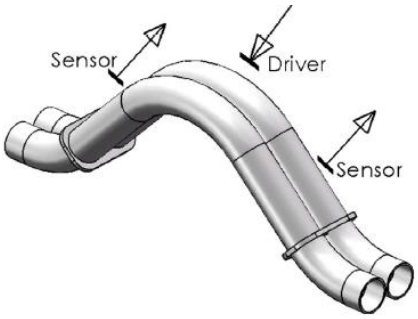
temperature, the temperature is measured continuously and the measured mass flowrate and density are compensated accordingly.



(a) A single straight-tube flow sensor



(b) A twin straight-tube flow sensor



(c) A twin bent-tube flow sensor

Figure 2.15 Typical Coriolis flow sensor designs [5]

Several theoretical analysis has been undertaken to investigate the performance of Coriolis flowmeters under two-phase flow conditions. Bubble effect, damping factor, sensor balance, compressibility and phase decoupling have been proposed to interpret and quantify the mass flow errors from Coriolis flowmeters under aeration condition [61]-[64]. Since there were some assumptions in developing the mathematical models and certain important factors such as flow regime, bubble distribution and pipe geometry were not taken into consideration, there were significant differences between the theoretical analysis and experimental results. Meanwhile, some researchers applied soft computing techniques to establish the complicated relationship between the mass flow errors and the parameters of Coriolis flowmeters. Then the erroneous mass flowrates were corrected according to the estimated mass flow errors. This sub-section mainly reviews the soft computing techniques based correction models.

Liu et al. [65] used a neural network to estimate the mass flow errors for a 1-inch Coriolis mass flowmeter on a horizontal pipe. As shown in Figure 2.16, the multi-layer

perceptron and radial basis function networks accept four inputs, including temperature, damping, density drop and flowrate, to estimate the mass flow errors. Then the mass flow reading from the Coriolis flowmeters is corrected with the estimated mass flow error. Experimental tests were conducted with the liquid mass flowrate ranging from 1.5 to 3.6 kg/s and density drop up to 35%. Although most of the mass flow errors were reduced to within ±2%, the gas entrainment was not quantified and other installation conditions were not considered.

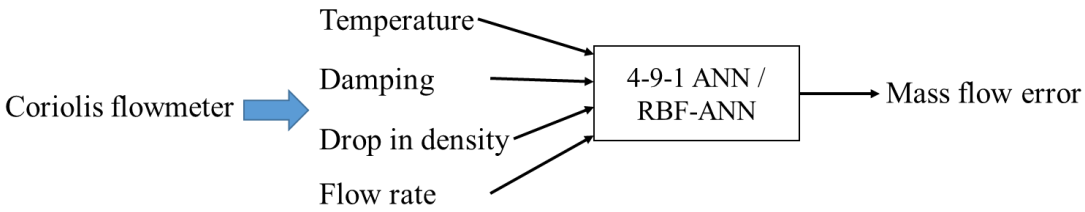


Figure 2.16 Measurement system based on a Coriolis flowmeter and ANN (Liu et al. [65])

Henry et al. [66],[67] reported a case study about two-phase flow metering of heavy oil using a Coriolis mass flowmeter. Trials were carried out on a 75 mm bore flowmeter with a mass flowrate between 1 kg/s and 10 kg/s and gas void fraction up to 80%. Experimental results demonstrated that the corrected measurements were typically within 1%-5% of the nominal mass flow and density for both steady and slugging two-phase flows.

The field tests using Coriolis flowmeters incorporating the ANN correction model were conducted to measure slugging two-phase CO_2 flow [46]. Due to the pressure losses through the pipeline network, the well-heads received the CO_2 at the pressure of 5.52 MPa -7.03 MPa and the temperature of 4°C-32°C. The significant variations in pressure and temperature resulted in gas-liquid two-phase CO_2 flow. In this case, the comparison results showed a 5% difference between the Coriolis flowmeters and the Pecos station sales meter.

Henry et al. [68] described another empirical method by combining a Coriolis mass flowmeter with a commercial water cut meter (Weaherford Red eye MP water cut meter) to achieve three-phase flow measurement. The water cut meter was used to

indicate the proportion of water in the liquid flow. Experimental tests were undertaken on a 50 mm bore test rig, with the total liquid flowrate ranging from 2.4 kg/s up to 11 kg/s, the water cut from 0% to 100% and gas volume fraction between 0% and 50%. Results demonstrated that the total liquid mass flow error was reduced within ±2.5% and the gas mass flow error within ±5%. The oil mass flow error limit was ±6% for water cut less than 70% and ±15% for water cut between 70% and 95%.

A method based on fuzzy inference was proposed to correct the mass flow error of a Coriolis mass flowmeter for two-phase flow measurement [69]. The system accepted damping, density drop and apparent mass flowrate as inputs to the fuzzy inference system to generate the corrected mass flowrate. Lari and Shabaninia [70] applied a neuro-fuzzy algorithm to correct the error of a Coriolis mass flowmeter for air-water two-phase flow measurement. However, the experimental data and results were not explained in detail in their reports.

Hou et al. [71] developed a digital Coriolis flow transmitter and tested a commercial 1-inch Coriolis flowmeter. The measurement errors under gas-liquid two-phase flow conditions were corrected using a feed-forward neural network with two inputs - apparent liquid mass flowrate and observed density drop. The online correction results showed that, when water flowrate varied from 3 to 15 kg/min with gas volume fraction up to 25%, the flowrate errors were within ±3.5% while density errors were within ±1.5%.

Ma et al. [72] used a 25 mm bore Coriolis flowmeter together with SVM algorithms to measure the overall mass flowrate of oil-water two-phase flow. The oil flowrate ranged from 0.27 to 5 m^3/h and water flowrate between 0.2 and 7 m^3/h. Experimental results showed that the relative error of the total mass flowrate was within ±1% while the individual mass flowrate had the maximum error of ±8%.

All the measurement methods mentioned in Section 2.3 using traditional sensors and soft computing techniques are summarized in Table 2.1.

Table 2.1 Indirect measurement approaches using traditional sensors and soft computing techniques

Sensors	**Soft computing method**	**Multiphase flow**	**Pipe diameter (mm)**	**Average relative error**
Ultrasonic sensor [48]	ANN	air-oil	25, 50	GVF variation: ±4.2
Throat-extended Venturi [50]	ANN	natural gas-water	50	$q_{v,g}$: 3.14%, $q_{v,l}$: 4.77%
Throat-extended Venturi [50]	SVM	natural gas-water	50	$q_{v,g}$: 2.86%, $q_{v,l}$: 4.25%
DP sensors [51]	ANN	air-water	32.5	v_g:2.4%, v_l: -0.3%
Conductance sensors [52]	ANN	air-water	50	v_g, v_l: <±10%
Laser diode, photodiode array [53]	SVM	nitrogen-water	4.22, 3.03, 2.16,1.08	GVF variation <7%
Conductance+Turbine flowmeter [54]	ANN	air-water	18	v_g, v_l: 7.36%
Capacitance+Conductance [55] +Ultrasonic+DP+Venturi	ANN	oil-water	50	v: <5%
Capacitance+Conductance [56] +Ultrasonic+DP+Venturi	ANN	oil-air-water	50	v_o: 6.2%, v_g: 4.68%, v_l:3.91%
Capacitance+Electrostatic sensor [57]	ANFIS	coal-biomass-air	36	C_b: 1.2%, C_c: 0.7%
Capacitance+Electrostatic sensor [58]	ELM	coal-biomass-air	94	C_b: 2.1%, C_c: 1.2%
Coriolis flowmeter [59]	ANN	air-water	25	$q_{m,l}$: <±2%
Coriolis flowmeter [60]	ANN	air-heavy oil	75	$q_{m,l}$: <1%~5%
Coriolis flowmeter [44]	ANN	gas-liquid CO_2	-	$q_{m,l}$: 5%
Coriolis flowmeter+Water cut meter [65]	ANN	oil-air-water	50	$q_{m,l}$: ±2.5%, $q_{m,g}$: ±5%, $q_{m,o}$: ±6%, ±15%
Coriolis flowmeter [66]	Fuzzy system	air-water	-	-
Coriolis flowmeter [67]	Neuro-Fuzzy	air-water	-	-
Coriolis flowmeter [68]	ANN	air-water	25	$q_{m,l}$: <±3.5%
Coriolis flowmeter [69]	SVM	oil-water	25	q_m:<±1%, $q_{m,o}$,$q_{m,w}$:<±8%

2.4 Input Variable Selection Methods for Data-Driven Models

Input variable selection is in fact to establish the relationships within the available data and identify suitable predictors of the model output. In the case of artificial neural networks or other similar data-driven models due to the imprecision and uncertainty nature of the soft computing approach, input variable selection is indispensable for developing a suitable model. May et al. [73] summarized the input variable selection methods for artificial neural networks and presented several key considerations in determining the most appropriate approach to input variable selection for a given application. As shown in Figure 2.17, the input variable selection techniques can be classified into model-based and model-free algorithms. In order to choose an appropriate approach, the first consideration is to decide whether the case under consideration is a linear or non-linear problem. Subsequently, the computational requirements in the model-based techniques and the selection accuracy should be balanced as well.

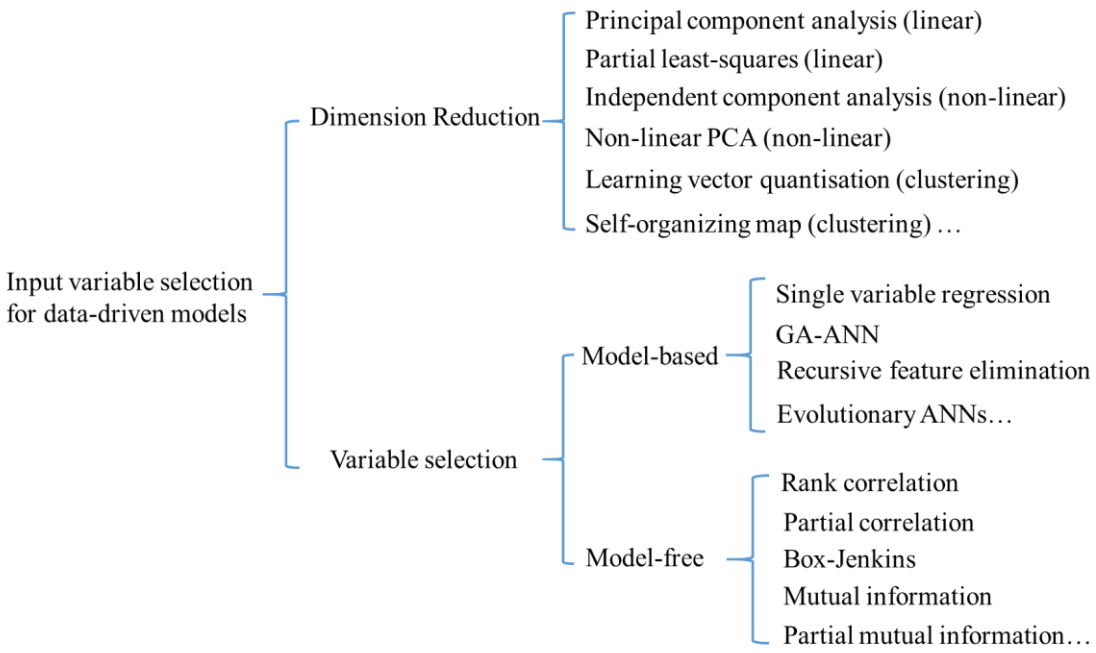


Figure 2.17 Principal input variable selection algorithms

Dimension reduction is often performed to reduce the computational effort associated with data processing, or to identify a suitable subset of variables to include in the analysis. Although the goal of dimension reduction differs from that of variable selection, it is closely related and regularly employed in data pre-processing. The commonly used dimension reduction method in the field of flow measurement is principal component analysis. Principal component analysis was applied in [52],[53],[74] to extract the features from several statistical variables before developing an ANN or SVM model. Ye et al. [75] extracted the statistical feature from pressure fluctuations through principal component analysis and established a least squares support vector machine to recognize multiphase flow pattern in a pipeline-riser system. Although principal component analysis and independent component analysis may be useful in removing noise from the data, they are not possible to distinguish the unique contributions of individual variables to the output.

So far there have been few research reports on the application of variable selection methods to multiphase flow measurement. The input variables in the proposed models in Section 2.3.6 are all selected by experience and may underestimate the performance of data-driven models. However, significant progress has been achieved in hydrology engineering as the complex environmental data are more difficult to interpret in physical meaning. There are some research conducted to compare the performance among different variable selection methods.

Bowden et al. [76],[77] tested PMI (Partial Mutual Information) and SOM-GAGRNN (Self Organizing Map - Genetic Algorithm and General Regression Neural Network) on a number of synthetic data sets and real-time forecasting simulation. It turned out that both approaches were acceptable when predictive performance was the primary aim. The variables determined using the PMI algorithm were most robust for the validation set and the PMI scores can reveal useful information about the order of importance of each significant input. May et al. [78] further researched on the partial mutual information based input variable selection for non-linear artificial neural networks. In order to resolve the problems of underlying assumption of linearity and redundancy within the available data, three novel termination criteria, i.e. Tabulated critical values, Akaike Information Criterion (AIC) and Hampel test criterion, were proposed to

improve computational efficiency and accuracy of the algorithm. Because the data distributions in real-world applications are often unknown and the assumption of Gaussian data may be inappropriate, the AIC and Hampel criteria are recommended for wider applicability. Galelli et al. [79] proposed a hybrid approach combining both model-free and model-based methods, namely, the tree-based Iterative Input Selection (IIS) algorithm. In this algorithm a tree-based ranking method is used in place of the information-theoretic measure such as PMI to estimate the information gained from the data. Results indicate that IIS is capable of selecting the most significant and non-redundant inputs in different test conditions. Subsequently, Galelli et al. [80] made an inter-comparison of four input variable selection methods, including PMI, PCIS (Partial Correlation Input Selection), IIS and GA-ANN, through testing on several datasets, and evaluated the four algorithms in terms of selection accuracy, computational efficiency and qualitative criteria. Li et al. [81] presented preliminary guidelines for the selection of most appropriate methods for obtaining the required kernel density estimates in the PMI method. The use of alternative bandwidth estimators can result in significant improvements in the PMI method for non-normally distributed data.

2.5 Summary

This chapter has reviewed the existing multiphase flow measurement systems, including separation systems, on-line sampling systems and on-line direct and indirect multiphase flowmeters. The soft computing techniques used for multiphase flow measurement and input variable selection for data-driven models have also been reported.

As can be seen from the cited work (Section 2.2), a variety of multiphase flowmeters have been developed for two-phase or multiphase flow measurement. However, there is limited research on CO_2 flow measurement in CCS, especially under two-phase or multiphase flow conditions.

As can be seen from the cited work (Section 2.3), traditional sensors incorporating soft computing techniques have been applied to measure the phase flowrate and phase fraction. The measurement systems based on Coriolis flowmeters are able to directly

provide mass flowrate of total or individual flow and have the potential to be more accurate than other measurement systems. As shown in Table 2.1, ANN is a popular approach to map the relationship between the input variables and desired outputs. However, ANN is based on empirical risk minimization and all the parameters are tuned iteratively and hence may suffer from overfitting. With the progress of soft computing techniques, more advanced algorithms such as SVM, evolutionary programming and hybrid techniques could be utilized to provide a better solution to these issues.

Moreover, the input variables for the proposed correction models of Coriolis flowmeters were selected by experience and may understate the performance of correction models. Although certain theoretical Coriolis flowmeters is available, parameter dependency, significance and sensitivity to the desired output are not quantified. In addition, there has been no research on the prediction of gas volume fraction using Coriolis flowmeters.

The literature review has demonstrated that there is a certain gap between the existing technologies and the requirements of multiphase flow measurement, including CO_2 flow in the CCS chain. Based on the literature review, Coriolis flowmeters incorporating soft computing techniques are believed to be a promising, feasible and cost-efficient solution to multiphase flow measurement.

Chapter 3

Theoretical Aspects of Data-Driven Modelling for Coriolis Flowmeters

3.1 Introduction

This chapter, firstly, presents the proposed measurement methodology using Coriolis flowmeters incorporating soft computing techniques. The measurement method is based on prior information captured from Coriolis flowmeters and additional sensors to estimate the liquid mass flowrate and gas volume fraction of two-phase flow.

Then, the chapter presents three advanced input variable selection methods for data-driven models, covering PMI (Partial Mutual Information), GA-ANN (Genetic Algorithm-Artificial Neural Network) and IIS (Tree-based Iterative Input Selection) approaches. The principle and selection procedures of the three methods are introduced specifically.

At last, the chapter gives the theory of three soft computing techniques for data-driven modelling, including ANN (Artificial Neural Network), SVM (Support Vector Machine) and GP (Genetic Programming). The structure of the data-driven models based on these techniques and modelling process are described in detail.

3.2 Measurement Methodology

Although the mass flowrate measurement from Coriolis flowmeters under two-phase or multiphase conditions has greater uncertainty than single-phase conditions, the outputs from Coriolis flowmeters and other additional sensors, such as DP (Differential Pressure) transducers and electrical impedance sensors, are related to the phase

components to some extent. Consequently, it is reasonable to assume the phase flowrate and phase fraction can be obtained if the relationship between the sensor outputs and the phase parameters is established. With the development of soft computing techniques, data-driven models are becoming more and more popular in constructing an input-output mapping by learning from history data and given examples and then performing estimation for the desired output.

As shown in Figure 3.1, the measurement methodology is based on data-driven models, which accept variables from a Coriolis flowmeter and additional sensors (a DP transducer and an electrical impedance sensor are taken into account in Chapter 4) while the output gives the liquid mass flowrate or gas volume fraction. Since the volume of data is often limited in practice, it is appropriate to design a separate model for each desired output.

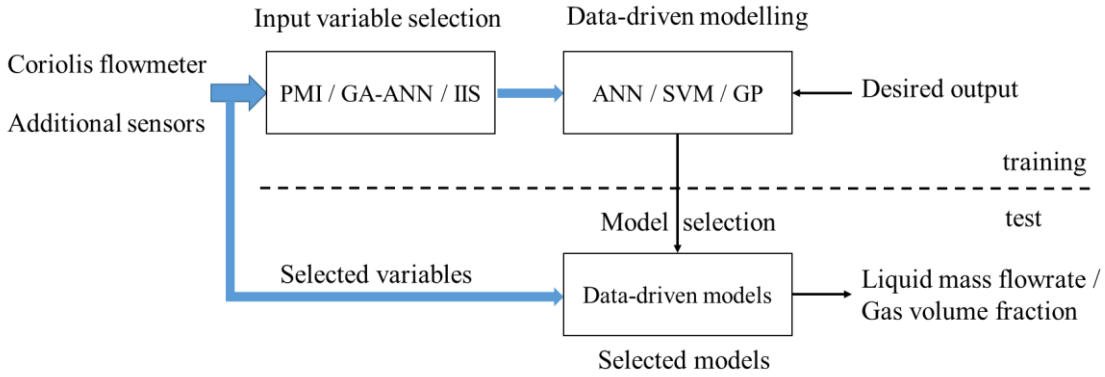


Figure 3.1 Principle of the measurement methodology

In order to develop a proper model for the desired output, there are two aspects - input variable selection and data-driven modelling. Input variable selection is to extract useful information from the available data and identify suitable input variables which are able to well explain the desired output for a data-driven model. Through input variable selection to eliminate the irrelevant or redundant variables, the complexity of the model structure is simplified and the computational efficiency is improved. Data-driven modelling is a process to find connections between the input variables and output variables through analysing the available data, even without any explicit knowledge of the physical behaviour of the system. Once the model is trained, it can be tested using

an independent data set to determine how well it can generalise to the unseen data. The approaches to input variable selection and data-driven modelling are presented in Sections 3.3 and 3.4, respectively. The evaluation of input variable selection methods and data-driven modelling techniques based on real experimental data is reported in Chapter 4.

3.3 Input Variable Selection

In this study, three input variable selection methods, including PMI, GA-ANN and IIS, are considered to identify input variables before modelling of Coriolis flowmeters for two-phase flow measurement.

3.3.1 Partial Mutual Information

The PMI approach, as an advanced model-free variable selection method, is able to minimize redundancy and maximize relevance in the available data by estimating the maximum joint mutual information. The PMI input variable selection method was first proposed by Sharma in 2000 for the identification of inputs for hydrological models [80]. PMI is a model-free variable selection method, which utilizes a measure of the partial dependence between a potential input variable and the output, conditional on any inputs that have already been selected. Earlier research on evaluating the performance of PMI on synthetic data sets and real-world data sets shows that PMI is a promising method [78],[81].

Given a dependent variable $\boldsymbol{Y}=(y_1,y_2,\ldots,y_n)^{\mathrm{T}}$, the potential input variable pool $\mathbf{X}=\{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_m}\}$ and the already selected variable set $\boldsymbol{S}$ which is a subset of $\mathbf{X}$ and can be represented as $(\mathbf{s_1}, \mathbf{s_2}, \ldots, \mathbf{s_n})^{\mathrm{T}} \in \mathbf{X}$, m is the number of potential variables and n is the number of sample observations. The PMI value of a potential input variable $\mathbf{x}=(x_1, x_2,\ldots, x_n)^{\mathrm{T}}$ can be formulated as:

$$PMI = \frac{1}{n}\sum_{i=1}^{n}\log\left[\frac{\hat{f}(u_i,v_i)}{\hat{f}(u_i)\hat{f}(v_i)}\right] \tag{3-1}$$

where, $\mathbf{u} = (u_1, u_2, \ldots, u_n)^T$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)^T$ represent the residual information in potential input variable **x** and dependent variable ***Y*** once the effect of the already selected inputs ***S*** has been taken into consideration. $\hat{f}_{(u)}$, $\hat{f}_{(v)}$ and $\hat{f}_{(u,v)}$ are the estimated marginal and joint probability density functions which are realized by kernel density estimation.

u and **v** in equation (3-1) are determined by

$$\mathbf{u} = \mathbf{x} - E[\mathbf{x} \mid \mathbf{S}] = \mathbf{x} - \hat{m}_{\mathbf{x}}(\mathbf{s}) \tag{3-2}$$

$$\mathbf{v} = \mathbf{Y} - E[\mathbf{Y} \mid \mathbf{S}] = \mathbf{Y} - \hat{m}_{\mathbf{Y}}(\mathbf{s}) \tag{3-3}$$

where $\hat{m}_{\mathbf{x}}(\mathbf{s})$ is the conditional expectation of **x** given an observed ***S***. $\hat{m}_{\mathbf{Y}}(\mathbf{s})$ is the conditional expectation of ***Y*** given an observed ***S***. The estimator $\hat{m}_{\mathbf{Y}}(\mathbf{s})$ for the regression of ***Y*** on ***S*** is written as

$$\hat{m}_{\mathbf{Y}}(\mathbf{s}) = \frac{1}{n} \frac{\sum_{i=1}^{n} y_i K_h(\|\mathbf{s} - \mathbf{s}_i\|)}{\sum_{i=1}^{n} K_h(\|\mathbf{s} - \mathbf{s}_i\|)} \tag{3-4}$$

where K_h is the Gaussian kernel function which can be described as

$$K_h = \frac{1}{\left(\sqrt{2\pi}h\right)^d \sqrt{|\Sigma|}} \exp\left(\frac{-\|\mathbf{s} - \mathbf{s}_i\|}{2h^2}\right) \tag{3-5}$$

d denotes the number of dimensions of ***s***, $\sum$ is the sample covariance matrix and $\|\mathbf{s} - \mathbf{s}_i\|$ is the squared Mahalanobis distance which is given by

$$\|\mathbf{s} - \mathbf{s}_i\| = (\mathbf{s}_i - \bar{\mathbf{s}})\Sigma^{-1}(\mathbf{s}_i - \bar{\mathbf{s}})^T \tag{3-6}$$

where $\bar{\mathbf{s}}$ is the mean of **s**. *h* denotes the kernel bandwidth which is determined by

$$h = \left(\frac{1}{d+2}\right)^{\frac{1}{d+4}} \sigma n^{\frac{-1}{d+4}} \tag{3-7}$$

where σ is the standard deviation of the data sample.

The estimator $\hat{f}(u)$ in equation (3-1) is given by

$$\hat{f}(u) = \frac{1}{n}\sum_{i=1}^{n} K_h(u-u_i) = \frac{1}{n(\sqrt{2\pi}h)^d\sqrt{|\Sigma|}}\sum_{i=1}^{n}\exp\left(\frac{-\|u-u_i\|}{2h^2}\right) \tag{3-8}$$

The estimators $\hat{m}_{\mathbf{x}}(\mathbf{s})$, $\hat{f}_{(\mathrm{v})}$ and $\hat{f}_{(\mathrm{u,v})}$ can be similarly constructed.

Akaike information criterion (AIC) is considered as the termination criterion of the PMI algorithm, as it can provide a general measure of the trade-off between information gain and the complexity introduced to the model by the addition of input variables. It is based on the analysis of the output variable residual **u**. Once there is no further reduction in the information contained in **u**, the optimal input variable set is reached and the selection is terminated.

$$AIC = n\log_e\left(\frac{1}{n}\sum_{i=1}^{n}u_i^2\right) + 2p \tag{3-9}$$

where p is the number of model parameters. **u** is the residual of the desired and estimated outputs and **u** can be calculated from equation (3-2).

The search strategy is a forward selection procedure which is described as follows:

(1) Initialize the selected input set ***S*** with null.

(2) Calculate the PMI value (equation (3-1)) for each potential variable which is beyond ***S***.

(3) Find out the variable **x** from {**X**-***S***} with the highest PMI value.

(4) Calculate the AIC score (equation (3-9)), assuming **x** is included in ***S***.

(5) If AIC score decreases, variable **x** can be added to S and go to step (2).

Otherwise variable **x** is rejected and the selection is terminated.

3.3.2 Genetic Algorithm-Artificial Neural Network

GA-ANN is a kind of model-based input variable selection method. It comprises of a simple 1-hidden node multilayer perceptron neural network as a regression model and

genetic algorithm as the heuristic search strategy. The potential variable subsets are generated by the genetic operations: selection, crossover and mutation. The suitability of the input variables is determined by assessing the performance of the neural network which is established based on the corresponding variables. The out-of-sample AIC is considered to quantify the accuracy of the neural network after k-fold cross validation. The termination criterion of the genetic algorithm is either the maximum number of generations or convergence of the fitness function is achieved [82].

A genetic algorithm usually includes five steps: population initialization, fitness function, selection, crossover and mutation. The selection procedure is as follows:
(1) Initialise a population with a random population of chromosomes.
(2) Each chromosome in the population is decoded into a solution. Calculate its fitness using an objective function.
(3) The chromosome with best fitness is selected to generate a second generation.
(4) Partially exchange genetic information between two parent chromosomes during crossover.
(5) With some low probability, a portion of the new individuals have some of their chromosomes flipped during mutation, which is used to keep the population diverse and prevent the algorithm from prematurely converging onto a local minimum.
(6) If the maximum number of generations or convergence of the fitness function is reached, the selection procedure is over. Otherwise, go to (2).

3.3.3 Tree-Based Iterative Input Selection

The IIS method is a combination of model-free and model-based methods. It utilizes extra-trees to estimate the relative contribution of each candidate input. The ranking-based evaluation does not require any assumption on the statistical properties of the input data set (e.g. Gaussian distribution) and can be applied to any sort of samples. Moreover, this approach does not rely on computationally intensive methods (e.g. bootstrapping) to estimate the information content in the data and thus is generally faster and more efficient [83].

Extra-trees is a nonparametric tree-based regression method. The process of building nodes and branches of a tree involves partitioning the input space into mutually exclusive regions according to a splitting criterion, progressively narrowing down the size of the regions. The splitting criterion is defined by the identification of the best input to split a node and the corresponding splitting value. Eventually, when the number of instances in a region becomes smaller than a specific user-defined value, the partitioning of that region stops the tree, a specific path is followed according to the splitting rules adopted in the tree-building procedure, and the predicted output is then obtained according to the values stored in the leaf.

The particular structure of extra-trees can be exploited to rank the importance of the input variables in explaining the output behaviour. It is based on the idea of scoring each input variable by estimating the variance reduction it can be associated with by propagating the training data set D over the M different trees composing the ensemble. The relevance G of the variable $\mathbf{x}_i$ in explaining the output $\boldsymbol{Y}$ can be evaluated by:

$$G = \frac{\sum_{t=1}^{M}\sum_{j=1}^{\Omega} k \cdot \Delta_{\mathrm{var}}(w_{tj})|D|}{\sum_{t=1}^{M}\sum_{j=1}^{\Omega} \Delta_{\mathrm{var}}(w_{tj})|D|} \tag{3-10}$$

where M is the number of different trees and Ω is the number of non-terminal nodes in the tree. w_{tj} is the j^{th} non-terminal node in the t^{th} tree. k is equal to 1 if the variable $\mathbf{x}_i$ is used to split the node w_{tj}, otherwise k is 0. $\Delta_{\mathrm{var}}(w_{tj})$ is the variance reduction associated to node w_{tj}. $|D|$ is the number of observations in training dataset D.

The input variables are sorted by decreasing the relevance of potential variables, and thus the first variable should be most significant. In order to reduce miss-selection and minimize the redundancy, the first q variables in the ranking are further evaluated by extra-trees based SISO (Single Input Single Output) and MISO (Multiple Inputs Single Output) models. The termination criterion is defined as either the best variable obtained in the current iteration is already in the selected variable set $\boldsymbol{S}$, or the improvement of the model performance reaches to tolerance ε.

The search strategy is forward selection and the selection procedure is as follows:
(1) Calculate the input ranking of the potential input variables according to the explained variance and pick out the q most relevant input variables.
(2) Build SISO models for each of the first q-ranked variables. Compute the distance metric between $\boldsymbol{Y}$ and the model output, and then select the most relevant variable $\mathbf{x}_i$ and add it to the selected variable set $\boldsymbol{S}$.
(3) Build a MISO model with selected variable set S and compute the distance metric between $\boldsymbol{Y}$ and the model output $M(\boldsymbol{S})$.
(4) The procedure is repeated using the residuals $[\boldsymbol{Y}\text{-}M(\boldsymbol{S})]$ as the new output variable in steps (1) and (2). These operations iterated until either the best variable returned is already in the set $\boldsymbol{S}$ or the performance of the underlying model does not significantly improve.

The evaluation of the distance metric follows a k-fold cross-validation approach. The estimated prediction accuracy is then the average value of the distance metric over the k validations.

3.4 Soft Computing Techniques

Data-driven modelling is a process to develop models using soft computing techniques, particularly machine learning. This section describes three soft computing techniques, including ANN, SVM and GP, for modelling a nonlinear system with multiple inputs and outputs [84]-[88]. These techniques are to be applied in the following study.

3.4.1 Artificial Neural Network

Neural network models are developed by training the network to represent the relationships and processes that are inherent within the data. Being essentially non-linear regression models, they perform an input-output mapping using a set of interconnected simple processing nodes or neurons. Each neuron takes in inputs either externally or from other neurons and passes it through an activation or transfer function

such as a logistic or sigmoid curve. As shown in Figure 3.2, the ANN consists of an input layer, one or more hidden layers and an output layer. The inputs can be any combination of variables that are thought to be important for predicting the output. Therefore, some knowledge of the process and input variable selection are very important. The hidden layer is the essential component that allows the neural network to learn the relationships in the data. The estimated values are generated from the output layer. In general, a neural network with a single-hidden layer of sufficient neurons is able to represent any nonlinear problem. In consideration of the simplicity of the ANN structure, a single-hidden layer is chosen and investigated in this study. The ANN trained with a backpropagation learning algorithm is defined as BP-ANN.

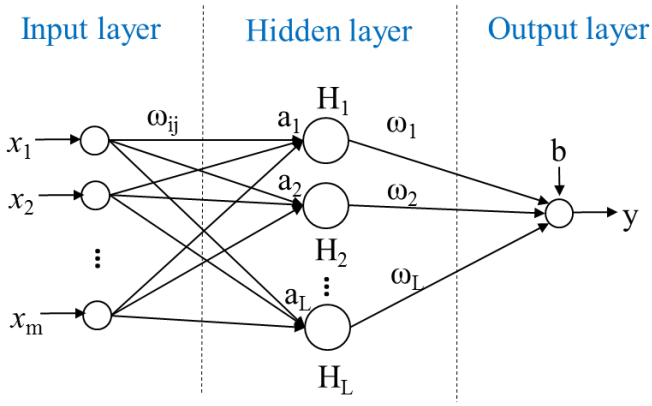


Figure 3.2 Structure of a BP-ANN

As shown in Figure 3.2, $\mathbf{x}=(x_1,x_2, \ldots, x_m)^T$ is an input sample and y is the desired output. Assume y is the linear output of the hidden neurons and a transfer function $f(x)$ is used on the neurons, the ANN is modelled as:

$$y_{BP} = \sum_{j=1}^{L} \omega_j H_j + b = \sum_{j=1}^{L} \omega_j f\left(\sum_{i=1}^{m} \omega_{ij} x_i + a_j\right) + b \tag{3-11}$$

where m and L are the numbers of input variables and hidden nodes, respectively. ω_j is the weight connecting the j^{th} hidden node and the output node, ω_{ij} is the weight connecting the i^{th} input node to the j^{th} hidden node. a_j and b are the biases on the j^{th} hidden node and the output node, respectivley. In this study, the hyperbolic tangent sigmoid function is used as a transfer function on hidden neurons and presented by

$$f(x) = \frac{2}{1+e^{-2x}} - 1 \tag{3-12}$$

The weights (ω_{ij} and ω_j) and biases (a_j and b) are obtained through the training process. Backpropagation algorithm is a variation of a gradient descent optimisation algorithm that minimises the error between the predicted and actual output values. The weighted connects between neurons are adjusted after each training cycle until the error in the validation data set begins to rise. The BP algorithm has been widely applied to solve practical problems. However, the BP algorithm has the disadvantage of slow convergence and long training time. Additionally, the success of the BP algorithm depends on the user-dependent parameters, such as initialization and structure of the ANN.

3.4.2 Support Vector Machine

3.4.2.1 SVM

SVM was developed by Vapnik in 1995 to solve the classification problem based on the statistic learning theory and structural risk minimization [89]. Since then, this method has been extended to the domain of regression and prediction problems [90]. The estimation of liquid mass flowrate and gas volume fraction is a regression problem. As shown in Figure 3.3, the input vector $\mathbf{x}=(x_1,x_2, \ldots, x_m)^T$ is first mapped into an L-dimensional feature space using transfer functions $\varphi(\mathbf{x})$ and then a linear model is constructed in this feature space. The nonlinear function f between input $\mathbf{x}$ and output y_{svm} is given by

$$y_{SVM} = f(\mathbf{x}) = \langle \boldsymbol{\omega}, \varphi(\mathbf{x}) \rangle + b \tag{3-13}$$

where $\boldsymbol{\omega}$ is the weight vector and b is the bias term.

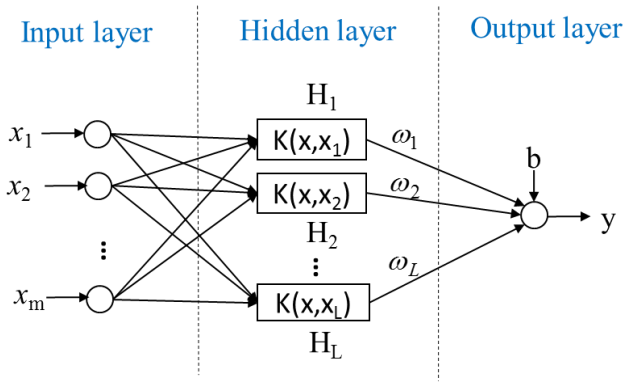


Figure 3.3 Structure of an SVM

SVM regression performs a linear regression in the high dimensional feature space using ε-insensitive loss and tends to reduce the model complexity by minimizing $\|\omega\|^2$. This can be described by introducing slack variables ξ_i and ξ_i' (i=1,2,…,n) to measure the deviation of training samples (X*, D) outside ε-insensitive zone. X*=($\mathbf{x_1}$, $\mathbf{x_2}$,…, $\mathbf{x_n}$) represents n input vectors of training samples and D=(y_1,y_2, …, y_n) is the corresponding desired output. Regression estimates can be obtained by minimizing the empirical risk on the training data and thus the optimization problem can be formulated as

$$\min \frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

$$\text{s.t.} \quad \begin{cases} y_i - \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle - b \le \varepsilon + \xi_i \\ \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{3-14}$$

where n is the number of training samples. ε is the approximation accuracy that can be violated by means of the slack variables ξ_i, ξ_i^* for the non-feasible case. C is a positive constant as a regularization parameter that allows tuning the trade-off between the flatness of the function $f(\mathbf{x})$ and the tolerance of deviations larger than ε.

The Lagrangian function is given by

$$\begin{aligned} L_{SVM} = \frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) - \sum_{i=1}^{n}\alpha_i(\varepsilon + \xi_i - y_i + \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle + b) \\ - \sum_{i=1}^{n}\alpha_i^*(\varepsilon + \xi_i^* + y_i - \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle - b) - \sum_{i=1}^{n}(\eta_i\xi_i + \eta_i^*\xi_i^*) \end{aligned} \tag{3-15}$$

where α_i , α_i^* , η_i , $\eta_i^* \geq 0$ are Lagrange multipliers. The conditions for optimality are given by

$$\begin{cases} \dfrac{\partial L_{SVM}}{\partial \boldsymbol{\omega}} = 0 \rightarrow \boldsymbol{\omega} = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\varphi(\mathbf{x}_i) \\ \dfrac{\partial L_{SVM}}{\partial b} = 0 \rightarrow \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0 \\ \dfrac{\partial L_{SVM}}{\partial \xi_i} = 0 \rightarrow C - \alpha_i - \eta_i = 0 \\ \dfrac{\partial L_{SVM}}{\partial \xi_i^*} = 0 \rightarrow C - \alpha_i^* - \eta_i^* = 0 \end{cases} \tag{3-16}$$

The extreme value problem is transferred into a dual problem by substituting (3-16) to (3-15):

$$\max_{\alpha,\alpha^*} Q = -\frac{1}{2}\sum_{i,k=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_k - \alpha_k^*)\langle \varphi(\mathbf{x}_i), \varphi(\mathbf{x}_k)\rangle - \varepsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*)$$

$$\text{s.t. } \begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases} \tag{3-17}$$

The parameter *b* can be computed by exploiting the Karush-Kuhn-Tucker (KKT) conditions and given by

$$\begin{cases} b = y_i - \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle - \varepsilon & for \quad \alpha_i \in (0, C) \\ b = y_i - \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle + \varepsilon & for \quad \alpha_i^* \in (0, C) \end{cases} \tag{3-18}$$

The resulting SVM for nonlinear function estimation takes the form:

$$f(\mathbf{x}) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\langle \varphi(\mathbf{x}), \varphi(\mathbf{x}_i)\rangle + b \tag{3-19}$$

According to Mercer's condition, the inner product $\langle \varphi(\mathbf{x}), \varphi(\mathbf{x}_i) \rangle$ can be defined through a kernel $K(\mathbf{x}, \mathbf{x}_i)$, so the final product of a training process in the SVM method can be presented by:

$$f(\mathbf{x}) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(\mathbf{x}, \mathbf{x}_i) + b \tag{3-20}$$

There are some optional kernel functions for SVM, such as:

Linear kernel: $K(\mathbf{x}, \mathbf{x}_i) = \langle \mathbf{x}, \mathbf{x}_i \rangle$ (3-21)

Polynominal kernel: $K(\mathbf{x}, \mathbf{x}_i) = (\langle \mathbf{x}, \mathbf{x}_i \rangle + p)^d \quad d \in m, p > 0$ (3-22)

Gaussian radial basis function kernel: $K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\sigma^2} \right)$ (3-23)

Sigmoid kernel: $K(\mathbf{x}, \mathbf{x}_i) = \tanh(\phi \langle \mathbf{x}, \mathbf{x}_i \rangle + \theta) \quad \phi > 0, \theta > 0$ (3-24)

3.4.2.2 LSSVM

LSSVM is a least squares version of support vector machine and proposed by Auykens and Vandewale [91]. Unlike the inequality constraint (equation (3-10)) used in SVM, LSSVM applies equality constraints and thus finds the optimal solution through solving a set of linear equations instead of a convex quadratic programming problem in classical SVM. As a result, computational complexity is significantly reduced. Meanwhile, LSSVM has been proven to have excellent generalization performance and low computational cost in many applications [92],[93].

Similar to equation (3-14), the optimization problem in LSSVM is formulated

$$\min \frac{1}{2} \|\boldsymbol{\omega}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^{n} {e_i}^2 \tag{3-25}$$

$$\text{s.t. } y_i = \langle \boldsymbol{\omega}, \varphi(\mathbf{x}_i) \rangle + b + e_i, \; i = 1, \ldots, n \tag{3-26}$$

where γ is a penalty parameter that balances model complexity and approximation accuracy, e_i is the i^{th} error variable. With the usage of the Lagrangian method, the optimization problem can be converted into a group of linear equations:

$$L_{LSSVM} = \frac{1}{2}\|\boldsymbol{\omega}\|^2 + \frac{1}{2}\gamma\sum_{i=1}^{n} e_i^2 - \sum_{i=1}^{n}\alpha_i\left[\langle\boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle + b + e_i - y_i\right] \tag{3-27}$$

where α_i (i=1,…,n) are the Lagrange multipliers. The conditions for optimality are given by:

$$\begin{cases} \dfrac{\partial L_{LSSVM}}{\partial \boldsymbol{\omega}} = 0 \rightarrow \boldsymbol{\omega} = \sum_{i=1}^{n}\alpha_i\varphi(\mathbf{x}_i) \\ \dfrac{\partial L_{LSSVM}}{\partial b} = 0 \rightarrow \sum_{i=1}^{n}\alpha_i = 0 \\ \dfrac{\partial L_{LSSVM}}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \\ \dfrac{\partial L_{LSSVM}}{\partial \alpha_i} = 0 \rightarrow \langle\boldsymbol{\omega}, \varphi(\mathbf{x}_i)\rangle + b + e_i - y_i = 0 \end{cases} \tag{3-28}$$

The solutions can be given in a group of linear equations by eliminating the variables $\boldsymbol{\omega}$ and e:

$$\begin{bmatrix} 0 & \mathbf{1}_v^T \\ \mathbf{1}_v & \boldsymbol{\Omega} + \mathbf{I}/\gamma \end{bmatrix}\begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \tag{3-29}$$

where $\mathbf{1}_v = [1,...,1]^T$, $\mathbf{y} = [y_1,...,y_n]^T$, $\boldsymbol{\alpha} = [\alpha_1,...,\alpha_n]^T$, $\mathbf{I}$ is an n×n identity matrix. The Mercer condition is applied:

$$\Omega_{ik} = \langle\varphi(\mathbf{x}_i), \varphi(\mathbf{x}_k)\rangle = K(\mathbf{x}_i, \mathbf{x}_k) \;\; i,k=1,\ldots,n \tag{3-30}$$

The parameter $\boldsymbol{\alpha}$ and b can be calculated as follows:

$$\begin{cases} b = \dfrac{\mathbf{1}_v^T\mathbf{H}^{-1}\mathbf{y}}{\mathbf{1}_v^T\mathbf{H}^{-1}\mathbf{1}_v} \\ \boldsymbol{\alpha} = \mathbf{H}^{-1}(\mathbf{y} - \mathbf{1}_v b) \end{cases} \tag{3-31}$$

where **H** is a positive definite matrix and described as

$$\mathbf{H}=\mathbf{\Omega}+\frac{1}{\gamma}\mathbf{I}=\begin{bmatrix} K(\mathbf{x}_1,\mathbf{x}_1)+\frac{1}{\gamma} & K(\mathbf{x}_1,\mathbf{x}_2) & \cdots & K(\mathbf{x}_1,\mathbf{x}_n) \\ K(\mathbf{x}_2,\mathbf{x}_1) & K(\mathbf{x}_2,\mathbf{x}_2)+\frac{1}{\gamma} & \cdots & K(\mathbf{x}_2,\mathbf{x}_n) \\ \vdots & \vdots & \ddots & \vdots \\ K(\mathbf{x}_n,\mathbf{x}_1) & K(\mathbf{x}_n,\mathbf{x}_2) & \cdots & K(\mathbf{x}_n,\mathbf{x}_n)+\frac{1}{\gamma} \end{bmatrix} \tag{3-32}$$

The choices of the kernel function $K(\mathbf{x}_i,\mathbf{x}_k)$ in LSSVM are the same as those in SVM. Finally, the LSSVM regression model can be obtained:

$$f(\mathbf{x})=\langle \boldsymbol{\omega},\varphi(\mathbf{x})\rangle+b=\sum_{i=1}^{n}\alpha_i K(\mathbf{x},\mathbf{x}_i)+b \tag{3-33}$$

3.4.3 Genetic Programming

Genetic programming is a method for evolving equations by taking various mathematical building blocks such as functions, constants and arithmetic operations and combining them into a single expression and was originally developed by Koza in 1992. GP as an evolutionary computation technique is an extension of genetic algorithms and is widely applied to symbolic data mining (symbolic regression, classification and optimization) [94]-[96]. Unlike traditional regression analysis, GP based symbolic regression automatically evolves both the structure and parameters of the mathematical model from the available data. Meanwhile, it is superior to other machine learning techniques due to the ability to generate an empirical mathematical equation without assuming prior form of the existing relationships. In this study, multigene symbolic regression is applied to establish a model for two-phase flow measurement. The structure of a multigene symbolic regression model is shown in Figure 3.4.

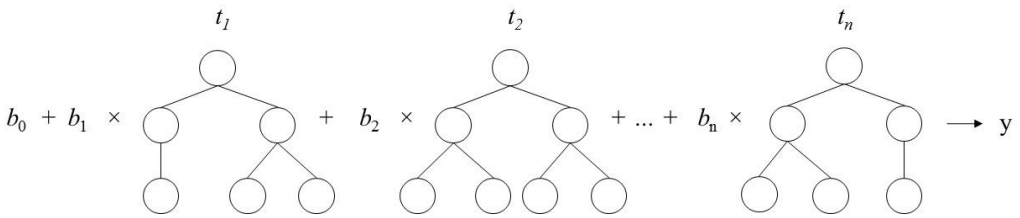


Figure 3.4 Structure of a GP model

The GP model can be regarded as a linear combination of lower-order nonlinear transformations of the input variables. The output y_{GP} is defined as a vector output of n trees modified by the bias term b_0 and scaling parameters $b_1,\ldots, b_n$:

$$y_{GP} = b_0 + b_1 t_1 + \ldots + b_n t_n \tag{3-34}$$

where t_i (i=1,…, n) is the (m×1) vector of outputs from the i^{th} tree comprising a multigene individual.

The evolutionary process starts with initial population by creating individuals containing GP trees with different genes generated randomly. The evolutionary process continues with an evaluation of the fitness of the new population, two-point high-level crossover to acquire and delete genes and low-level crossover on sub-trees. Then the created trees replace the parent trees or the unaltered individual in the next generation through mutation operators. The best program that appeared in any generation, the best-so-far solution, defines the output of the GP algorithm [95].

3.5 Summary

This chapter has firstly described the measurement methodology using Coriolis flowmeters incorporating soft computing techniques. Data-driven models accept the useful information extracted from sensor signals and estimate the liquid mass flowrate and gas volume fraction of two-phase flow. Before developing the data-driven models, input variable selection is an essential step to identify the significance of the variables to the desired output and determine a suitable subset of variables as inputs of the model.

Three different approaches to input variable selection have been presented, including PMI, GA-ANN and IIS. The principle and selection procedures have been described in detail. In addition, the chapter has also presented the principles of soft computing techniques for data-driven modelling, including ANN, SVM and GP. The structure of the model based on each technique and modelling process have been reported specifically. The performances of the input variable selection methods and data-driven

models based on different soft computing techniques are demonstrated and compared through the real experimental data obtained from 1-inch air-water two-phase flow test rig (Chapter 4). The comparison of generalization ability of SVM and LSSVM is reported in Chapter 5.

Chapter 4

Experimental Tests with Air-Water Two-Phase Flow

4.1 Introduction

This chapter presents the results using soft computing techniques on 1-inch Coriolis flowmeters (KROHNE OPTIMASS 6400 S25) with air-water two-phase flow. Experimental work was conducted on the air-water two-phase flow test rig at KROHNE Ltd. The objectives of the tests were defined as follows:

1) To assess the performance of 1-inch Coriolis flowmeters with air-water two-phase flow on horizontal and vertical installation orientations.
2) To evaluate the effectiveness of the input variable selection methods (i.e. PMI, GA-ANN, IIS) and soft computing techniques (ANN, SVM, GP) presented in Chapter 3.
3) To investigate the performance of the measurement methodology for two-phase flow using Coriolis flowmeters incorporating soft computing techniques.

A range of experimental tests were carried out with liquid flowrate from 700 kg/h to 14500 kg/h and gas volume fraction between 0 and 30%. Previous studies suggested that the parameters from Coriolis flowmeters, DP transducers and electrical impedance sensors are varying with the liquid flowrate and entrained gas. The impacts of sensor outputs on the measurement of liquid mass flowrate and gas volume fraction are reported in this chapter. The performance comparisons of input variable selection method and soft computing techniques are presented in detail.

4.2 Experimental Conditions

Experimental tests of 1-inch Coriolis flowmeters (KROHNE OPTIMASS 6400 S25) with air-water two-phase flow were conducted at KROHNE Ltd in July 2015. As shown in Figure 4.1, the gas flow is set to enter to the liquid flow through a by-pass on the pipe. The liquid mass flowrate is controlled by adjusting the pump frequency from 15 to 80%. The gas flowrate is varied by adjusting the opening of the valve in a gas flow controller. Two independent Coriolis flowmeters (KROHNE OPTIMASS 6400 S25 and Bronkhorst mini CORI-FLOW M15) were installed before the mixer to provide references for the individual mass flow rates of the liquid and gas phases respectively. Both reference meters' measurement uncertainties under single-phase conditions were verified according to the manufacturer's technical specification. Due to the effect of gravity and buoyancy on the gas-liquid two-phase flow, bubbles distributed within the Coriolis measuring tubes are different between horizontal and vertical installations and hence Coriolis flowmeters perform differently. For this reason, two additional Coriolis flowmeters (see Figure 4.2) of the same type were installed in the horizontal and vertical test sections, respectively, to measure the liquid mass flowrate. The Coriolis flowmeters under test are in twin bent-tube design, as shown in Figure 4.3. Two DP transducers were also used to record the DP value across each flowmeter under test. In addition, an electrical impedance sensor was installed in the upstream of the Coriolis flowmeter on the horizontal test section.

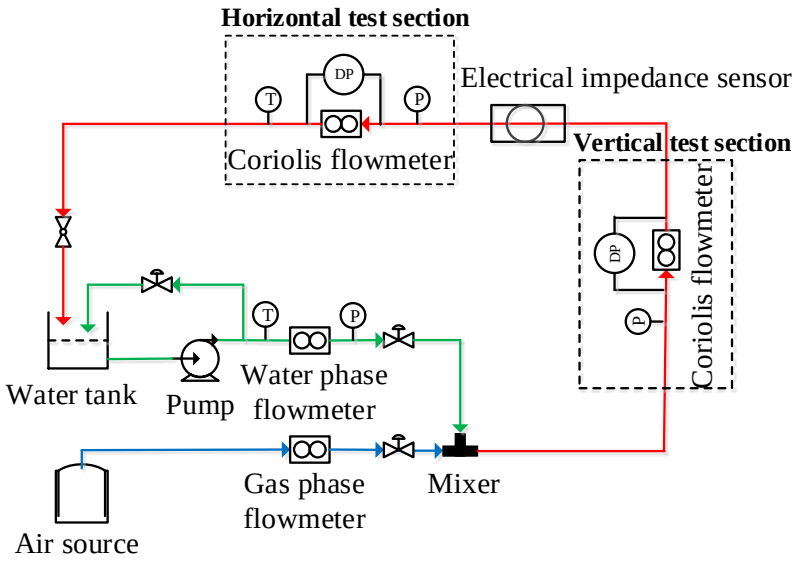


Figure 4.1 Schematic of the two-phase flow test rig

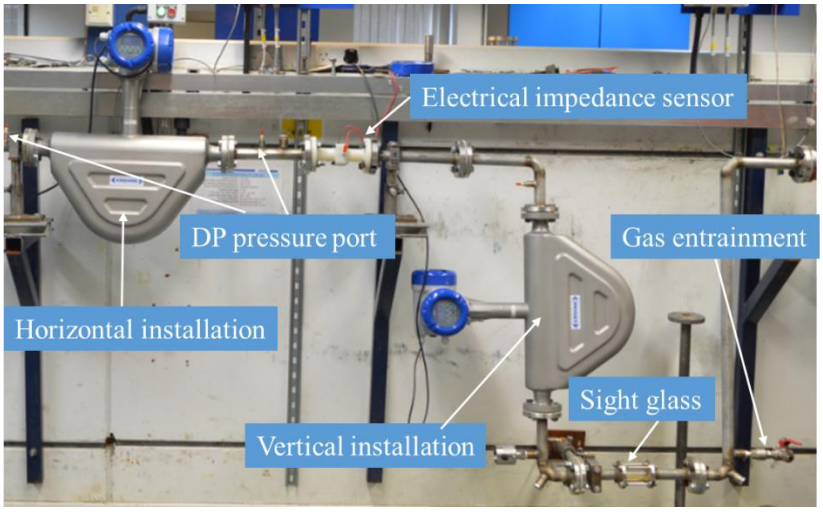


Figure 4.2 Photo of the Coriolis flowmeters under test

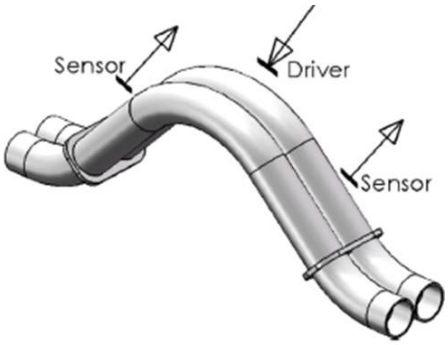


Figure 4.3 Typical design of twin bent Coriolis measuring tubes [5]

A series of experimental tests were conducted with the liquid mass flow rate ranging from 700 kg/h to 14500 kg/h and gas volume fraction from 0 to 30%. The test points with variations in liquid mass flowrate and gas volume fraction are summarized in Figure 4.4. For the purpose of model training, a dataset of 237 records (circular makers in Figure 4.4) was collected while the dataset of 24 records (triangular markers in Figure 4.4) for testing the model. The test dataset includes some experimental data which were collected at different flowrates from those in the training dataset. The new conditions as in the test dataset which were conducted on a different day and obtained under different flowrate from the training dataset are useful to assess the generalization capability of data-driven models.

Each dataset represents the average of all recorded values within an approximate window of 100 s. The fluid temperature during the tests was between 18°C and 24°C, which varied with the ambient temperature in the laboratory. It was observed during the

experiments that the flow pattern in the horizontal pipe was mostly slug or dispersed bubbly flow whilst the flow was of bubbly and dispersed bubbly nature in the vertical pipe.

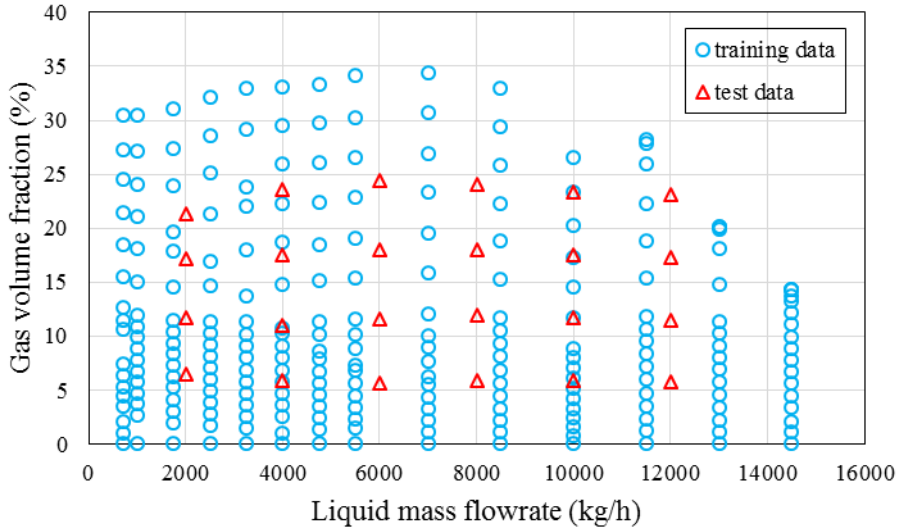
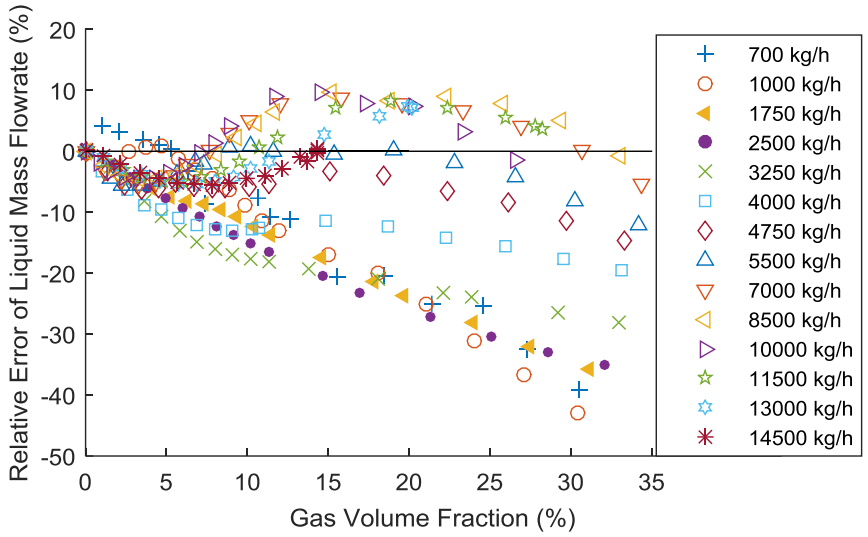


Figure 4.4 Experimental test points of the 1-inch air-water flow

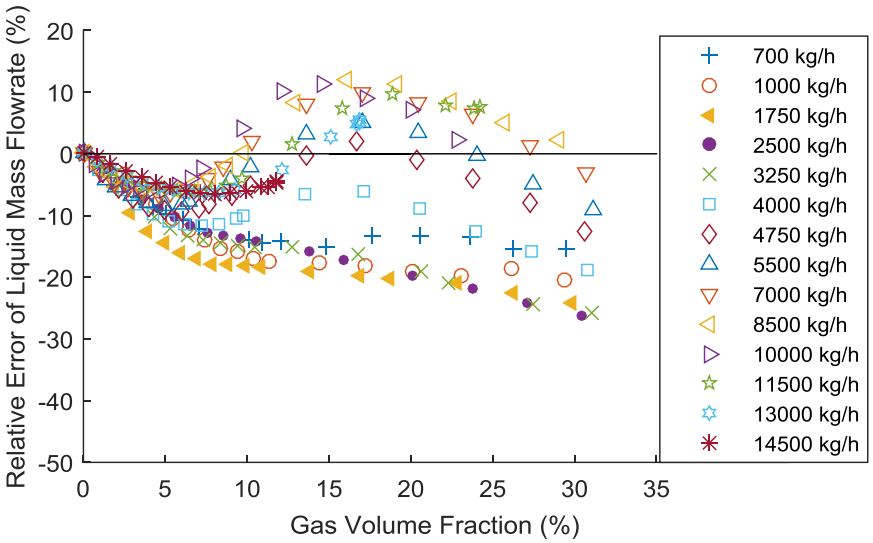
4.3 Analysis of Original Errors

The typical original mass flow errors of the Coriolis flowmeters in training data are plotted in Figure 4.5. The Coriolis flowmeter on the vertical section gives negative errors at flowrates below 4000 kg/h. At a higher flowrate (>5500 kg/h), the mass flow errors become positive and crossing the zero line and then return to negative errors again along with increasing entrained gas. This is believed to be due to the flow regime effects on the fluid-tube coupling system at different flowrates. At a lower flowrate (< 2000 kg/h), the flow was nearly slug flow as observed during the test while the flow regime became gradually dispersed bubbly flow as the flowrate and entrained gas increase. For the Coriolis flowmeter on the horizontal pipeline, the range of mass flow errors is different from that on the vertical pipeline most likely due to the effects of gravity and buoyancy on the flow regime. Positive errors occur at mass flowrates of 700 kg/h and 1000 kg/h when the gas volume fraction is below 6%. Gas volume fraction is calculated according to the temperature and pressure in the upstream of the test meter. By comparing the mass flow errors at the same flowrate in Figure 4.5 and Figure 4.6,

the errors are generally reproducible for the same installation and thanks to the new-generation flow transmitter [97].

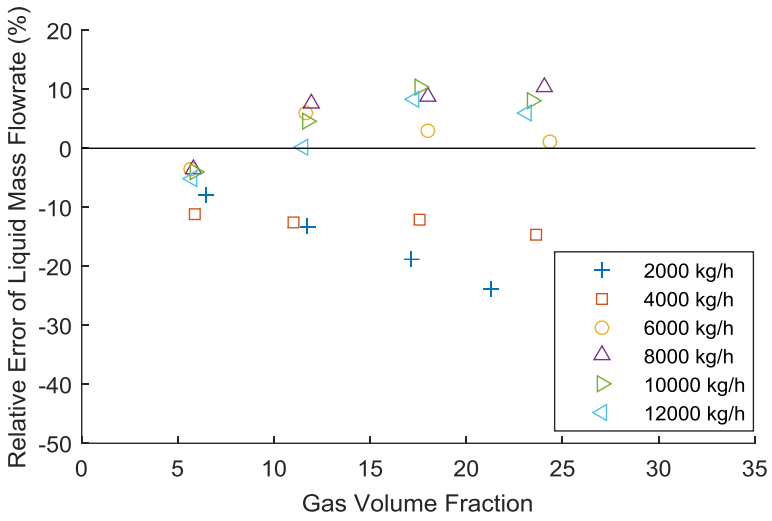


(a) Horizontal pipeline

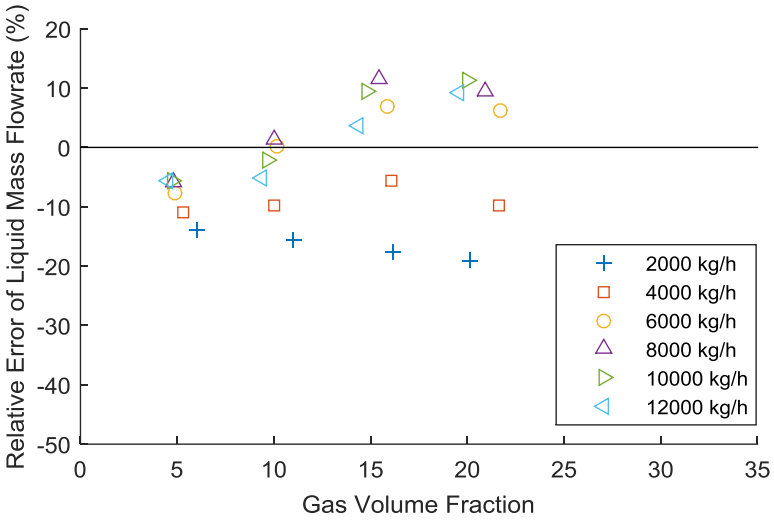


(b) Vertical pipeline

Figure 4.5 Original errors of the liquid mass flowrate from training data



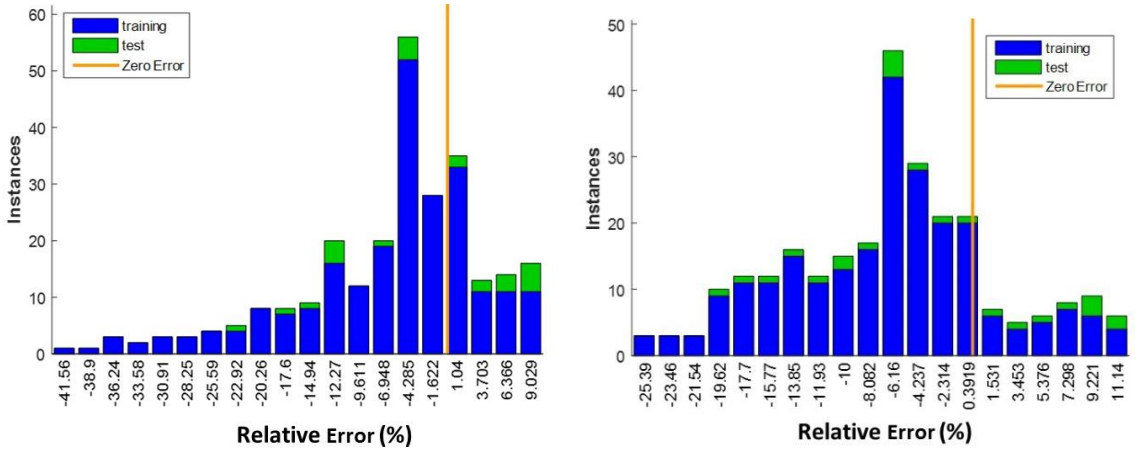
(a) Horizontal pipeline



(b) Vertical pipeline

Figure 4.6 Original errors of the liquid mass flowrate from test data

Figure 4.7 depicts the distribution of the relative errors of the measured liquid mass flowrate on both horizontal and vertical pipelines. Each color (blue or green) in the figure represents training or test datasets respectively. The Coriolis flowmeter on the horizontal pipeline yields the liquid mass flowrate with a relative error between -41% and 9% whilst the meter on the vertical pipeline gives an error from -25% to 11%. The difference in errors between the vertical and horizontal installations is due to the fact that the bubbles in a vertical flow are distributed evenly in the pipe cross section due to the effect of gravity, resulting in less interruptions on the tube vibration inside the Coriolis flowmeter and hence different errors.



(a) Horizontal pipeline (b) Vertical pipeline

Figure 4.7 Relative error histogram of the liquid mass flowrate

4.4 Evaluation of Input Variable Selection Methods

4.4.1 Definition of Variables

The variables acquired from sensors and their corresponding physical definitions are outlined in Table 4.1. All input variables (x_1-x_{14}) are obtained through measuring or transforming the internal parameters of the two Coriolis flowmeters, DP transducers and the electrical impedance sensor. Variables x_1-x_3 are direct measurement outputs from Coriolis flowmeters while x_4-x_{11} are their internal parameters. The measurement principle of mass flowrate (q_m) of a single-phase flow using Coriolis flowmeters is represented as [98],[99]

$$q_m = K_R[1+(K_T+e)\Delta T + C_1\Delta\sigma + C_2\Delta\tau + C_3\Delta T + C_4\Delta T^2]t_d \tag{4-1}$$

where K_R is a calibration factor for the measurement of the mass flowrate of a single-phase flow at a reference temperature, K_T is a material temperature dependence factor, e is a device-specific temperature dependence factor, and ΔT is a relevant temperature difference. $\Delta\sigma$ and $\Delta\tau$ indicate the relevant differences on circumferential stress and axial stress of a measuring tube. C_1, C_2, C_3 and C_4 are corresponding calibration coefficients. t_d is the time shift between the signals from motion sensors A and B.

The fluid density (ρ) from a Coriolis flowmeter is determined by

$$\rho = A_R(1+K_E\Delta T)\frac{1}{f^2} - B_R \tag{4-2}$$

where A_R and B_R are calibration factors to measure the density of single phase flow, K_E is a temperature dependence factor, and f is the frequency of a measuring tube.

Coriolis flowmeters are capable of providing accurate measurements of mass flowrate and density of single-phase flow. However, the mass flowrate and density readings are erroneous under the condition of two-phase or multiphase flows due to the effect of additional phases on the vibration coupling of the measuring tube. In this case, the erroneous mass flowrate and density are defined as apparent mass flowrate and observed density (i.e. x_1 and x_2 in Table 4.1).

x_{12} (DP) from the DP transducer represents the pressure difference across a Coriolis flowmeter. x_{13} and x_{14} are from the electrical impedance sensor and represent the magnitude and phase angle of the impedance of the flow between two electrodes. Under two phase flow conditions, it is impossible to measure the liquid mass flowrate and gas volume fraction directly using each of the instruments alone. However, the variations of the variables in Table 4.1 reflect the changing of the desired outputs Y_1 (desired liquid mass flowrate) and Y_2 (desired gas volume fraction) to some extent. Due to the reproducibility of the experimental tests, it is possible to establish data-driven models using the available experimental data to estimate liquid mass flowrate ($\hat{Y}_1$) and gas volume fraction ($\hat{Y}_2$). Because some of the variables from different sensors are independent and some are related to each other in physical sense, it is difficult to empirically determine which variable has more contribution to explain the desired outputs. It is necessary to formulate the relationship between the potential input variables and the desired outputs and then identify significant variables which are able to explain the desired outputs completely. Meanwhile, the repeated information should be eliminated to reduce the complexity of the model structure and improve the performance of the model.

Table 4.1 Variables and their corresponding physical definition

ID	Variable name and symbol	Physical definition	Source
x_1	Apparent mass flowrate (q_m)	The mass flowrate reading from Coriolis flowmeters based on calibration characteristics for single-phase flows.	Coriolis flowmeter
x_2	Observed density (ρ)	The density reading from Coriolis flowmeters based on calibration characteristics for single-phase flows.	Coriolis flowmeter
x_3	Process temperature (T)	The temperature reading from Coriolis flowmeters	Coriolis flowmeter
x_4	Sensor-A amplitude (V_A)	The voltage amplitude of signals from motion sensor-A.	Coriolis flowmeter
x_5	Sensor-B amplitude (V_B)	The voltage amplitude of signals from motion sensor-B.	Coriolis flowmeter
x_6	Drive level (I_D)	The current amplitude of the driver output.	Coriolis flowmeter
x_7	Time shift (t_d)	The time delay between the signals from the two motion sensors induced by flow.	Coriolis flowmeter
x_8	Tube frequency (f)	The oscillation frequency of the Coriolis measuring tube(s) inside Coriolis flowmeters.	Coriolis flowmeter
x_9	Sensor balance (B_s)	The ratio of the voltage amplitude of sensor-A signal to sensor-B signal.	Coriolis flowmeter
x_{10}	Damping (K)	Damping factor of Coriolis measuring tube(s)	Coriolis flowmeter
x_{11}	Two phase indicator	An indicator for the detection of a two-phase or multiphase flow [100]	Coriolis flowmeter
x_{12}	Differential pressure (DP)	DP is the differential pressure across the Coriolis flowmeter.	DP transducer
x_{13}	Magnitude of impedance ($\lvert Z \rvert$)	The ratio of the voltage difference amplitude to the current amplitude.	Electrical impedance sensor
x_{14}	Phase factor of impedance (θ)	The phase difference between voltage and current.	Electrical impedance sensor
Y_1	Desired liquid mass flowrate	Desired liquid mass flowrate on the test section is obtained from liquid reference Coriolis flowmeter	Liquid reference Coriolis flowmeter
Y_2	Desired gas volume fraction	Desired gas volume fraction on the test section is obtained by calculation according to liquid and gas reference Coriolis flowmeters	Liquid and gas reference flowmeters
$\hat{Y}_1$	Estimated liquid mass flowrate	Output of data-driven models for the measurement of liquid mass flowrate under two-phase flow conditions.	Data-driven model
$\hat{Y}_2$	Estimated gas volume fraction	Output of data-driven models for the measurement of gas volume fraction under two-phase flow conditions.	Data-driven model

In consideration of different performances of the Coriolis flowmeter on horizontal and vertical installations, two separate models are to be built for each position. Model-H-L and Model-H-G present the models built for the Coriolis flowmeter in the horizontal section to estimate liquid mass flowrate and gas volume fraction, respectively. Similarly, Model-V-L and Model-V-G indicate the models built in the vertical section to estimate liquid mass flowrate and gas volume fraction, separately. Since apparent mass flowrate (x_1) and observed density (x_2) are derived directly from time shift (x_7) and tube frequency (x_8), respectively, with temperature compensation, the possibility of replacing x_7 and x_8 with x_1 and x_2 is considered, respectively. For this reason, x_7 and x_8 are excluded in Table 4.2. Regarding each model, there are 12 potential input variables and one output. The input variables and outputs from the four models are outlined in Table 4.2.

Table 4.2 Symbols of the input variables and corresponding outputs for the four models

Installation	**Model**	**Potential input variables**	**Model output**
Horizontal pipe	H-L	x_1-x_6, x_9-x_{14}	$\hat{Y}_1$
	H-G	x_1-x_6, x_9-x_{14}	$\hat{Y}_2$
Vertical pipe	V-L	x_1-x_6, x_9-x_{14}	$\hat{Y}_1'$
	V-G	x_1-x_6, x_9-x_{14}	$\hat{Y}_2'$

4.4.2 Implementation of Input Variable Selection

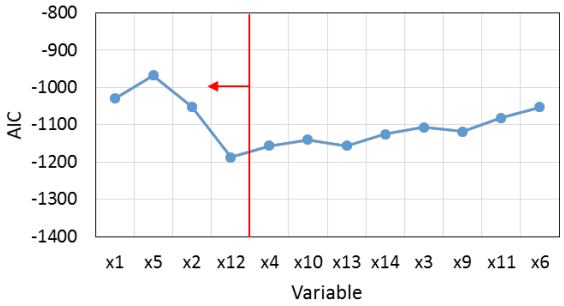
4.4.2.1 Implementation of PMI

The PMI algorithm was implemented on the experimental dataset for each model. The variable with the highest PMI value is selected as the candidate input variable set for each iteration. As defined in equation (3-9), AIC is a measure of the trade-off between accuracy of the regression and the size of the input sets. The behaviour of AIC will initially be dominated by a reduction in the magnitude of the residual terms and decreases with increasing the number of inputs before reaching a minimum value.

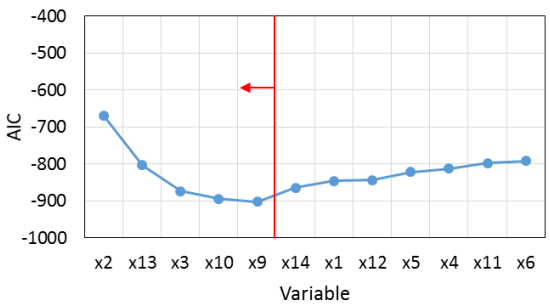
Afterwards, AIC increases due to the term of 2*p* in equation (3-9), which penalises the selection of addition variables. Hence, the optimal size of inputs corresponds to the minimum AIC. Figure 4.8 shows the change of AIC score when a new variable is added. As Figure 4.8 (a)-(d) shown, the AIC score has a decreasing trend until a new variable is selected so that the termination criterion is reached. The variables before the red line in Figure 4.8 are the final selected variables for the four models.

Model-H-L and Model-V-L are applied to the Coriolis flowmeters on horizontal and vertical pipelines, respectively, for the measurement of liquid mass flowrate. The results of PMI selection (Figure 4.8 (a) and (c)) show that variables x_1 (apparent mass flowrate), x_2 (observed density) and x_{12} (DP) are main factors for the two models to estimate the desired liquid mass flowrate. This means, although the Coriolis flowmeters on different orientations have different performances, the main factors affecting liquid mass flowrate are the same. Variable x_5 (sensor-B amplitude) provides additional information for Model-H-L while x_{10} (damping) is helpful for Model-V-L. The combined effect of the variables is more significant than that of an individual variable on the output. The selection results demonstrate x_1 has more contribution than the other variables to the measurement of liquid mass flowrate as one would expect from a purely physical point of view. The difference in selection sequence between the models for horizontal and vertical pipes is due to the effect of installation on the performance of Coriolis flowmeters.

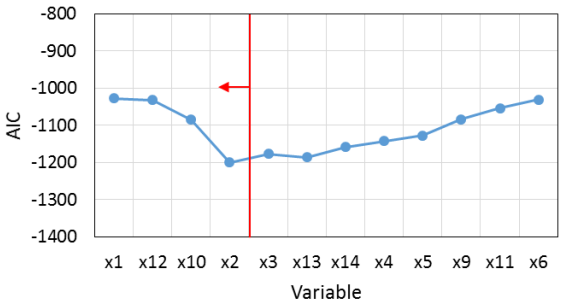
Model-H-G and Model-V-G are applied for gas volume fraction prediction of Coriolis flowmeters on horizontal and vertical pipelines, separately. Figure 4.8 (b) and (d) show that variables x_2 (observed density), x_{13} (magnitude of impedance) and x_3 (process temperature) are obtained from the PMI selection procedures for both Model-H-G and Model-V-G. Apart from the three variables, x_{10} and x_9 are also beneficial to Model-H-G for fitting the desired output. As expected, variable x_2 was first selected for Model-H-G and Model-V-G since the observed density changes significantly in relation to the amount of gas entrained in the liquid flow.



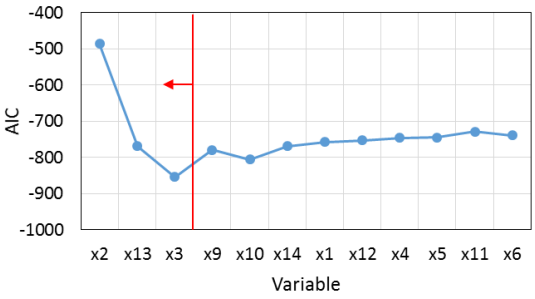
(a) PMI for Model-H-L



(b) PMI for Model-H-G



(c) PMI for Model-V-L



(d) PMI for Model-V-G

Figure 4.8 Procedures of PMI input variable selection for Models H-L, H-G, V-L and V-G

4.4.2.2 Implementation of GA-ANN

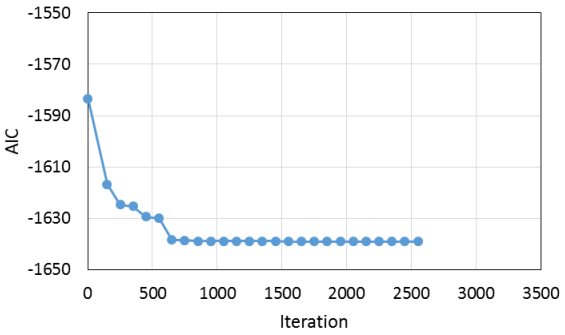
GA-ANN input variable selection was implemented on the dataset of Models H-L, H-G, V-L and V-G. The GA algorithm takes thousands of iterations to conduct selection, crossover and mutation before the AIC value is converged. The iteration process for each model is shown in Figure 4.9. The out-of-sample AIC for the four models are individually 1639.06, 958.65, 1555.46 and 911.36.

The selected input variables are x_1, x_5, x_6, x_{10} and x_{12} for Model-H-L, while x_1, x_5, x_6, x_9, x_{10} and x_{12} for Model-V-L. It is obvious that the common variables x_1 (apparent mass flowrate), x_5 (sensor-B amplitude), x_6 (drive level), x_{10} (damping) and x_{12} (DP) are important for the pre-defined ANN models to fit the desired liquid mass flowrate.

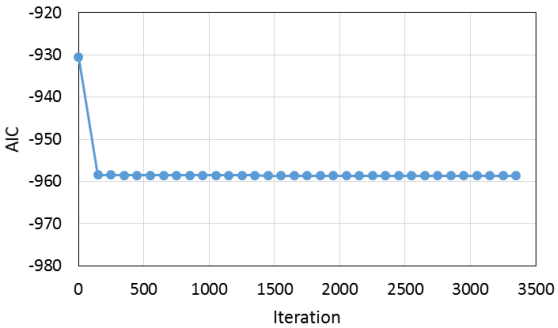
The selected inputs variables are x_2, x_3, x_4, x_5, x_{11}, x_{12}, x_{13}, x_{14} for Model-H-G, while x_2, x_3, x_4, x_9, x_{11}, x_{12}, x_{13}, x_{14} for Model-V-G. As the GA-ANN selection results suggest,

the variables x_2 (observed density), x_3 (process temperature), x_4 (sensor-A amplitude), x_{11} (two phase indicator), x_{12} (DP), x_{13} (magnitude of impedance) and x_{14} (phase factor of impedance) are common important variables for the pre-defined models to predict gas volume fraction.

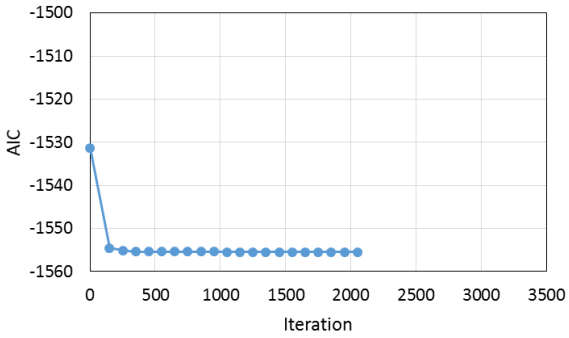
The outcomes of GA-ANN include more variables than PMI. The selected subset is restricted by the structure of ANN and the problem of redundancy in the selected variables is more serious than PMI. Moreover, it cannot provide the information of significance level of the selected input variables.



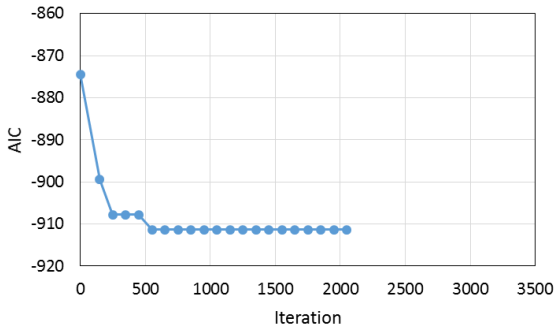
(a) GA-ANN for Model-H-L



(b) GA-ANN for Model-H-G



(c) GA-ANN for Model-V-L



(d) GA-ANN for Model-V-G

Figure 4.9 Iteration of GA algorithm for Models H-L, H-G, V-L and V-G

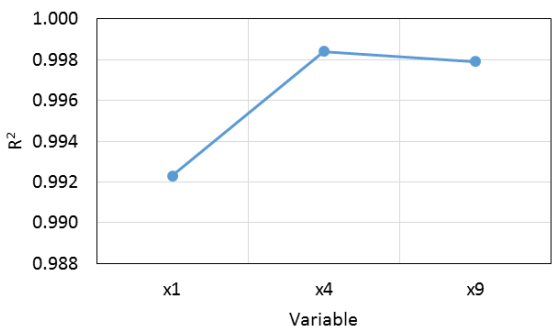
4.4.2.3 Implementation of IIS

IIS algorithm was executed on the dataset of Models H-L, H-G, V-L and V-G, respectively. The candidate subset was incrementally built through ranking, SISO (Single Input Single Output) and MISO (Multiple Inputs Single Output) evaluation. The selection process and the performance of MISO are shown in Figure 4.10 for each

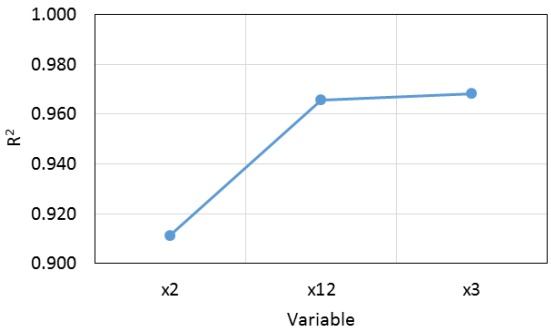
model. For Model-H-L, the tolerance ε is reached after x_1, x_4 and x_9 are all included. For Model-H-G, the selection process is terminated because variable x_{12} is selected twice. For Model-V-L, the tolerance ε is reached after x_1, x_4, x_2 and x_9 are selected. For Model-V-G, input variable x_{12} is also selected twice resulting in the selection procedure terminated.

From the results of Model-H-L, the variables x_1 (apparent mass flow), x_4 (sensor-A amplitude), x_9 (sensor balance) are selected while variables x_1, x_4, x_2 (observed density), x_9 are selected for Model-V-L. This presents the information of apparent mass flow and sensor balance are two most useful variables to improve the performance of MISO models.

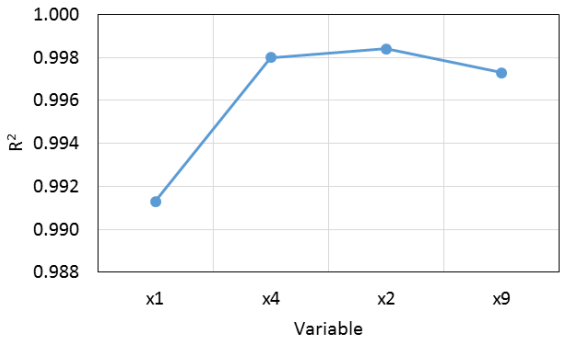
From the results of Model-H-G, the variables x_2, x_{12} (DP), x_3 (process temperature) were selected while variables x_2, x_1, x_{12} were selected for Model-V-G. The fitting results of MISO models with the inputs of the information of observed density and mass flowrate information (DP or apparent mass flowrate) are significantly improved.



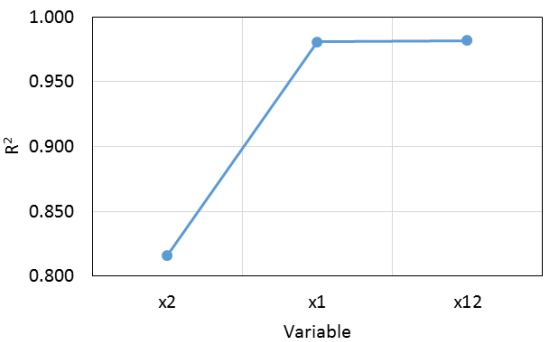
(a) IIS for Model-H-L



(b) IIS for Model-H-G



(c) IIS for Model- V-L



(d) IIS for Model-V-G

Figure 4.10 Results of IIS for Models H-L, H-G, V-L and V-G

4.4.3 Comparison of Input Variable Selection Methods

The selected variables by PMI, GA-ANN and IIS methods and corresponding running time (the laptop processor is Intel(R) Core (TM) i5-4200U CPU @ 1.60GHz) are summarized in Table 4.3. It is clear that GA-ANN algorithm takes longer time to implement the heuristic search than the other two methods. The running time of GA-ANN is about 10 times of IIS running time and 100 times of PMI running time. A large number of calibration and validation processes in GA-ANN result in high computational requirement while the model-free approaches are directly based on the information obtained from the available dataset. In view of time efficiency, PMI algorithm is the most effective approach and IIS is in the second place.

In terms of the number of selected input variables, GA-ANN produces more candidate variables than PMI and IIS. Moreover, there is much more redundant and unnecessary information in the selected subset. IIS generates smallest subset without such redundancy. Comparing the outcomes from IIS and PMI, the important information of the candidate variables from IIS are mostly included in the selected variables from PMI.

Table 4.3 Variable selection outcomes from PMI, GA-ANN and IIS

Model	PMI		GA-ANN		IIS	
	Variables	**Time (s)**	**Variables**	**Time (s)**	**Variables**	**Time (s)**
H-L	x_1, x_5, x_2, x_{12}	11.81	x_1, x_5, x_6, x_{10}, x_{12}	5281.93	x_1, x_4, x_9	217.33
V-L	x_1, x_{12}, x_{10}, x_2	13.14	x_1, x_5, x_6, x_9, x_{10}, x_{12}	4152.62	x_1, x_4, x_2, x_9	279.63
H-G	x_2, x_{13}, x_3, x_{10}, x_9	13.10	x_2, x_3, x_4, x_5, x_{11}, x_{12}, x_{13}, x_{14}	2689.13	x_2, x_{12}, x_3	242.92
V-G	x_2, x_{13}, x_3	12.63	x_2, x_3, x_4, x_9, x_{11}, x_{12}, x_{13}, x_{14}	721.12	x_2, x_1, x_{12}	247.57

4.4.3.1 Validation

In order to compare the performance of the three variable selection methods, a typical data-driven model is established for each case with SVM. Since an SVM model is based on the statistic learning theory and structural risk minimization, the output is not affected by the initial parameters and pre-defined structure of the model. Due to the relatively constant structure of the SVM model, it is applied to establish the relationship between the input variables and the output. In this study, the parameters for building the SVM models are optimized through 10-fold cross validation. The performance of the SVM model is quantified by Normalised Root-Mean-Square Error (NRMSE), which is defined as

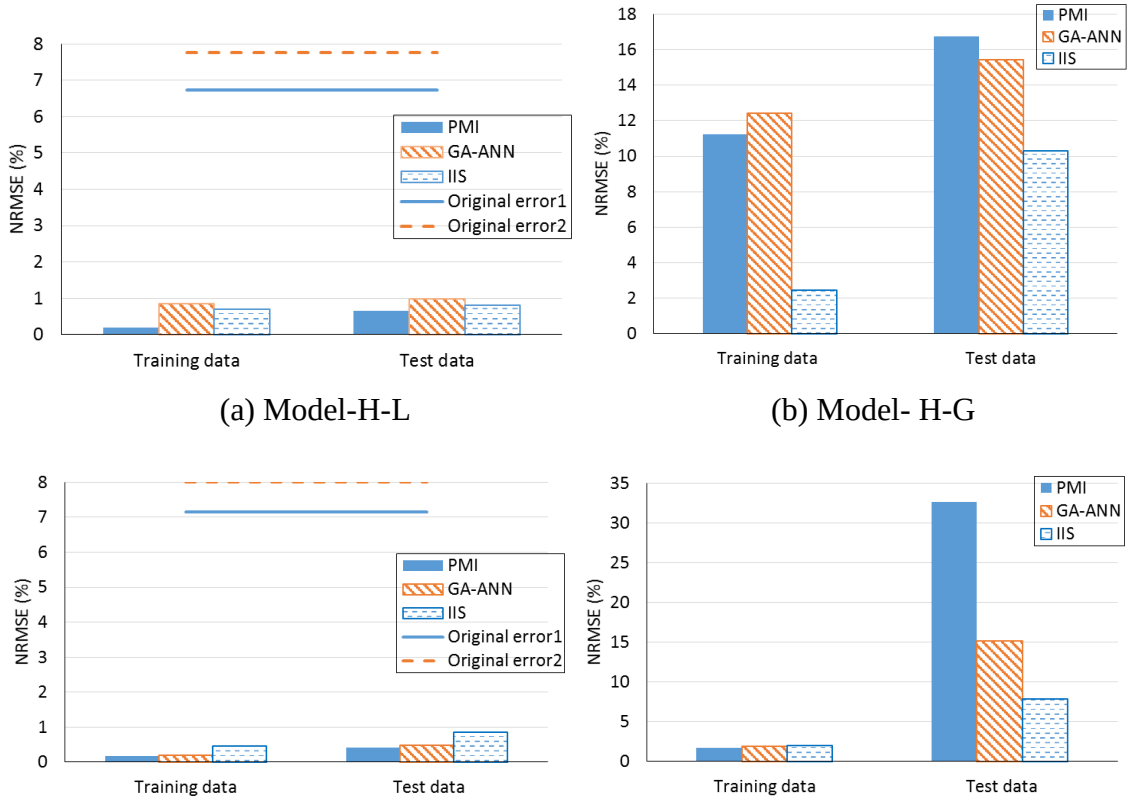
$$NRMSE = \frac{1}{\bar{y}} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \hat{y}_i \right)^2} \tag{4-3}$$

where y_i is the reference value of the model output, $\bar{y}$ is the mean of y_i , $\hat{y}_i$ is the model output, and n is the number of observations.

For each model, three SVM-based models were established based on the selected variables from PMI, GA-ANN and IIS, respectively. Each SVM-based model was tested on the training data and test data to evaluate the performance of the models which was mainly affected by the input variables. The solid line (original error1) and dash line (original error2) in Figure 4.11 (a) and (c) represent the original errors of mass flowrate on training data and test data. After correction by SVM-based models, the error of liquid mass flowrate from Coriolis flowmeters on horizontal and vertical pipes are both dramatically reduced. Due to the inherent limitation of generalization ability of data-driven models, the errors on test data are larger than those on training data. Through comparing the NRMSE values of the three SVM-based models, the model with the variables selected by PMI has the lowest error than the other two. This means the selected variables by PMI include better and more completed information to explain the liquid mass flowrate.

As shown in Figure 4.11 (b) and (d), the SVM-based models with PMI input variables does not perform well to predict gas volume fraction. Alternatively, IIS input variables

make the model predict gas volume fraction with relatively lower error. This illustrates IIS provides the fundamental variables for the prediction of gas volume fraction. The performance of SVM-based models with GA-ANN input variables is not as good as the models based on variables from PMI and IIS.



(a) Model-H-L (b) Model- H-G

(c) Model-V-L (d) Model-V-G

Figure 4.11 Performance of SVM with input variables selected for PMI, GA-ANN and IIS

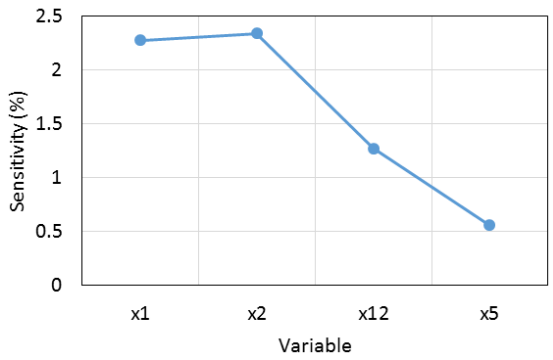
4.4.3.2 Sensitivity Analysis

Sensitivity analysis is used to evaluate how sensitive the model output is to the changes in the value of input variables and also identify which input variables are important in contributing to the prediction of the output variables [101],[102]. In performing the sensitivity analysis, each variable of the model's inputs is increased by 5% in turn. The aim is to assess the effect of small changes in each input on the model output. The percentage change in the output as a result of the increase in each of the inputs is the sensitivity, which is defined as:

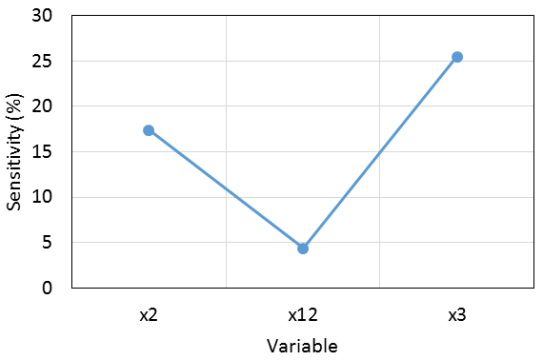
$$S_a = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{z}_i - \hat{y}_i}{\hat{y}_i}\right)^2} \times 100\% \tag{4-4}$$

where, $\hat{z}_i$ is the model output after small changes added to the input variables, $\hat{y}_i$ is the model output with no changed input variables and n is the number of observations.

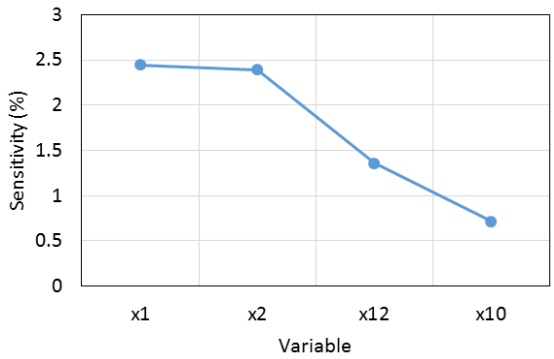
Sensitivity analysis was conducted on the SVM-based models with PMI input variables for liquid mass flowrate measurement, and the SVM-based models based on IIS input variables for gas volume fraction prediction. The sensitivity of each variable is plotted in Figure 4.12. From Figure 4.12 (a) and (c), it is clear that variables x_1 (apparent mass flowrate) and x_2 (observed density) have the higher sensitivity to the model output than the other variables. They have more relative influence on the measurement of liquid mass flowrate. Figure 4.12 (b) and (d) show that small changes in variable x_2 (observed density) can result in more significant variations on the model output. Moreover, temperature fluctuation has more effect on the measurement of gas volume fraction than liquid mass flowrate.



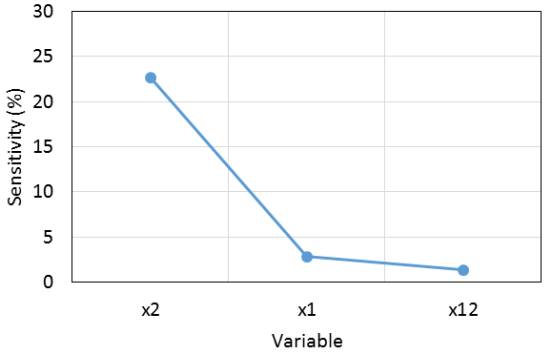
(a) Sensitivity for Model-H-L



(b) Sensitivity for Model- H-G



(c) Sensitivity for Model-V-L



(d) Sensitivity for Model-V-G

Figure 4.12 Sensitivity of input variables for the four models

The validation and analysis results suggest that the input variables selected from the PMI algorithm provide more effective information to estimate the liquid mass flowrate while the IIS algorithm provides a fewer but more effective variables to predict the gas volume fraction. With regard to the experimental data obtained in this study, the most important variable set for the measurement of liquid mass flowrate includes observed density, apparent mass flowrate, DP and damping while those for the prediction of gas volume fraction include observed density, apparent mass flowrate and DP. Although variable selection approaches can provide some valuable information to determine the input variables of a data-driven model, the accuracy of the methods also depends on the observational dataset, such as data size and their distributions. A dataset with less data or low-quality may result in underestimation or overestimation of the candidate variables for a data-driven model. Consequently, in order to ensure the selection accuracy with limited size of a dataset, it is necessary to determine the input subset using variable selection methods incorporating physical interpretation of the variables.

4.5 Evaluation of Data-Driven Models

4.5.1 Data-Driven Modelling

4.5.1.1 BP-ANN Models

The BP-ANN model is developed through a training process. For each installation condition a separate model is established for the estimation of the measured liquid mass flowrate and gas volume fraction. The inputs of the BP-ANN for liquid mass flowrate correction include four variables, i.e. observed density, apparent mass flowrate, damping and DP which are concluded from Section 4.4 while the inputs of the BP-ANN for gas volume fraction prediction include observed density, apparent mass flowrate and DP. The number of neurons (L) in the hidden layer is determined using the equations below, as proposed in [77]:

$$L \leq 2n+1 \tag{4-5}$$

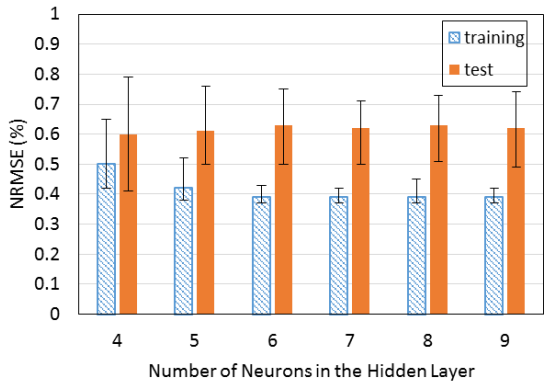
$$L \leq \frac{m}{n+1} \tag{4-6}$$

where n and m are the numbers of input variables and training samples, respectively. However, equations (4-5) and (4-6) give only the range of L for BP-ANN models. The exact L for a model can be selected by a trial-and-error method to compromise between minimizing errors and achieving good generalization capability. The output layer has one neuron for each model since there is only one output variable.

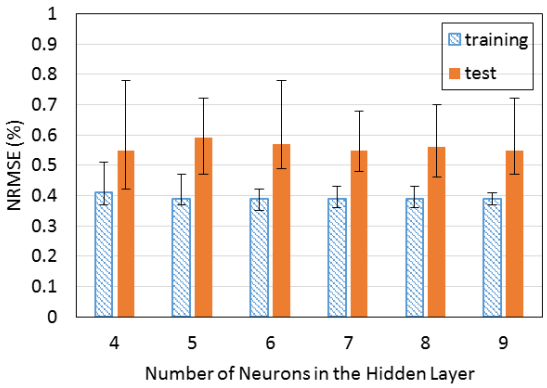
The BP-ANN transfer function between the input and hidden layers is hyperbolic tangent sigmoid transfer function. The pure linear function is taken as the transfer function connecting the hidden layer to the output layer. The training function is Bayesian regularization whilst the learning function is gradient descent with momentum weight and bias learning function. Training stops when the maximum number of epochs is reached or the performance is minimized to the goal. In this study, NRMSE is used to assess the performance of a data driven model, which is defined in equation (4-3).

As the weights and biases between the neurons are initialized randomly, a different BP-ANN is obtained for each training, resulting in different performance. A preliminary study of averaging NRMSE of more than 200 BP-ANNs did not show any noticeable difference. Therefore, in order to minimize the effect of random initialization of an ANN, the average NRMSE of 200 BP-ANNs with the same structure is calculated to assess the effect of the hidden neurons on the performance of the ANN.

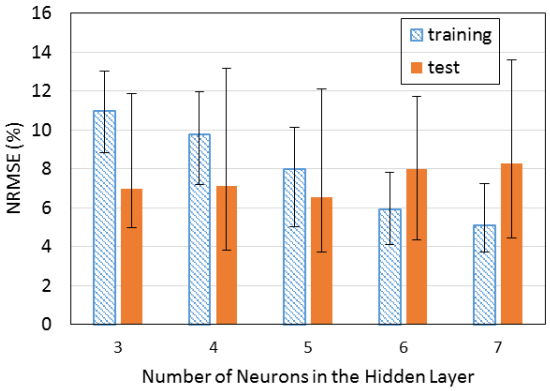
For the models for liquid mass flowrate estimation, the number of neurons in the hidden layer is set from 4 to 9 as per equations (4-5) and (4-6). The NRMSE values of the BP-ANNs are summarized in Figure 4.13. The error bars indicate the maximum and minimum errors of 200 BP-ANNs for the same structure. In view of the errors on both training and test datasets, the BP-ANN with 7 neurons in the hidden layer performs better than other structures under both horizontal and vertical conditions. The BP-ANN used for gas volume fraction prediction has lower NRMSE when the number of the hidden neurons is 6.



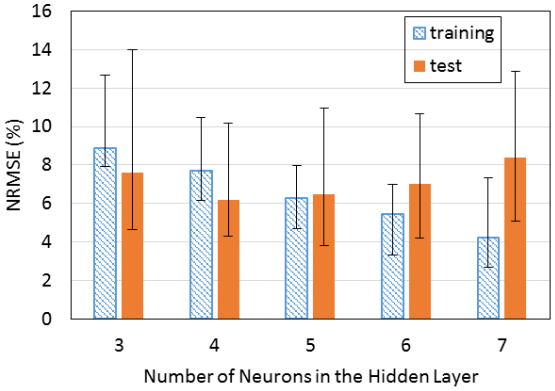
(a) BP-ANN: H-L



(b) BP-ANN: V-L



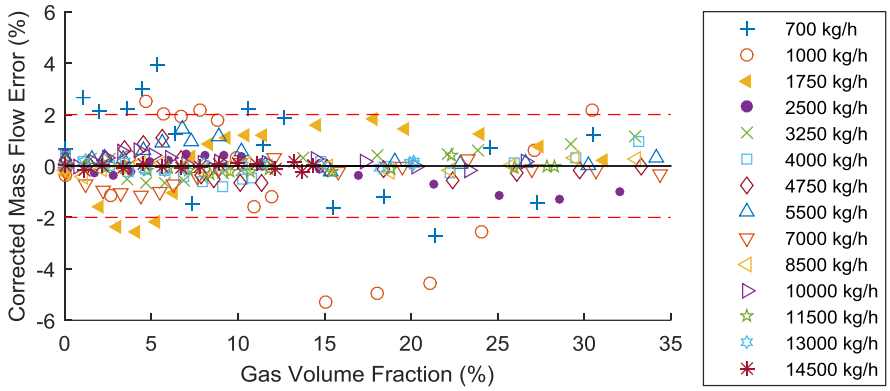
(c) BP-ANN: H-G



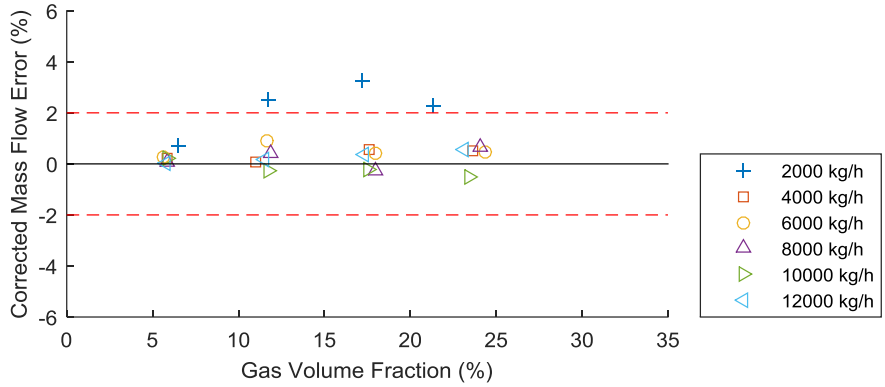
(d) BP-ANN: V-G

Figure 4.13 Performance of BP-ANNs with different numbers of neurons in the hidden layer

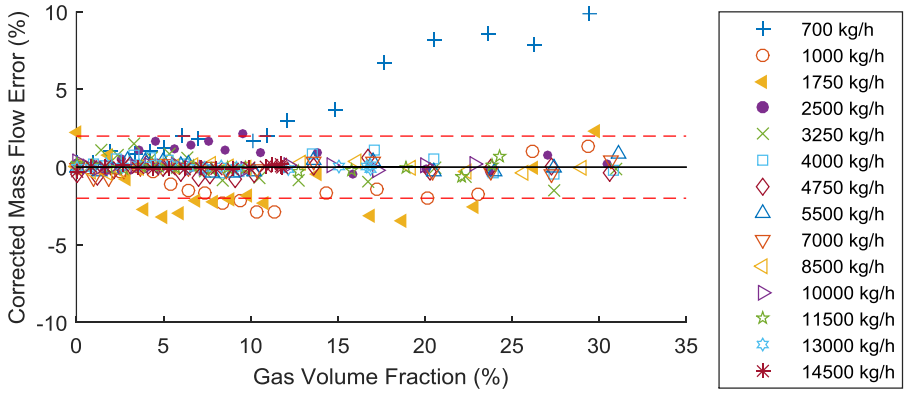
Once the structure of a BP-ANN is determined, the trained neural network which has the minimum error with the test dataset is selected. Figure 4.14 shows the errors of the corrected liquid mass flowrate from the BP-ANNs. For the horizontal and vertical pipelines, the relative errors are mostly less than ±2% (the red dash lines in Figure 4.14) with the training dataset except some larger errors at low flowrates of 700 kg/h and 1000 kg/h. This is very likely due to larger bubbles or slugs appearing in the flow tubes under low flowrate which affects the Coriolis flowmeter behaving differently from smaller bubbles. The trained BP-ANN has relatively larger errors at low flowrates and hence results in unsatisfactory performance with the test dataset under the same experimental conditions.



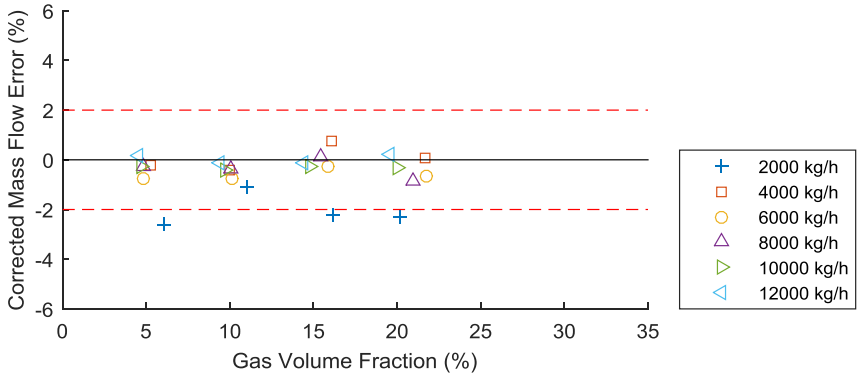
(a) BP-ANN for Model-H-L with training data



(b) BP-ANN for Model-H-L with test data



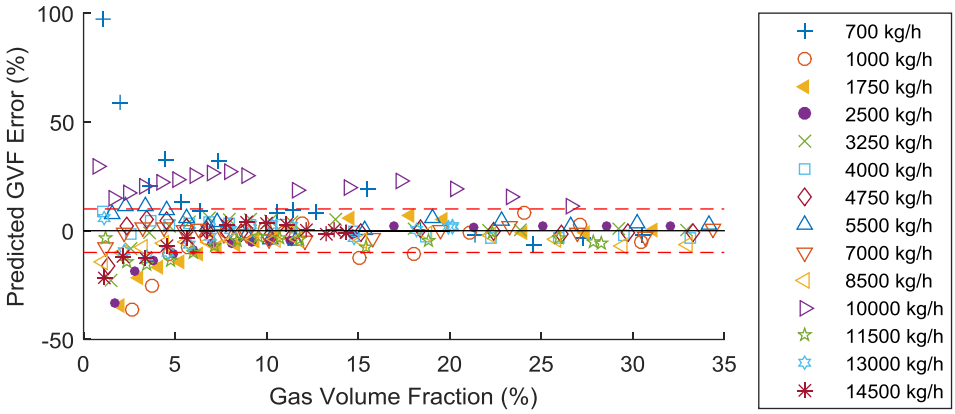
(c) BP-ANN for Model-V-L with training data



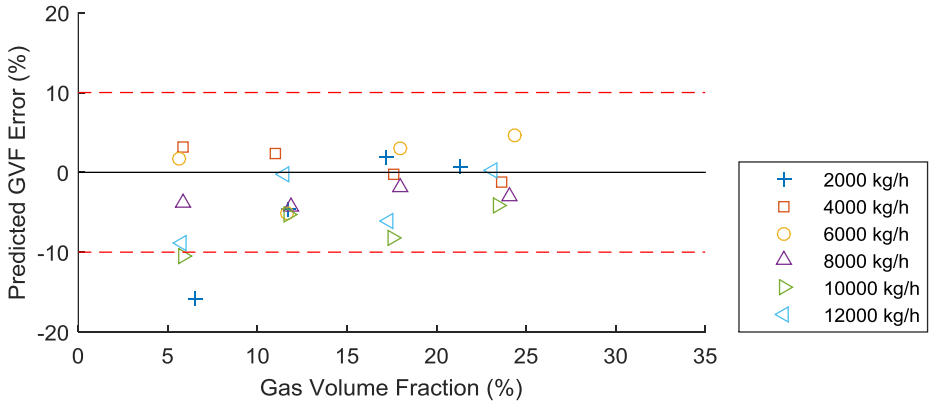
(d) BP-ANN for Model-V-L with test data

Figure 4.14 Errors of the corrected liquid mass flowrate from the trained BP-ANNs

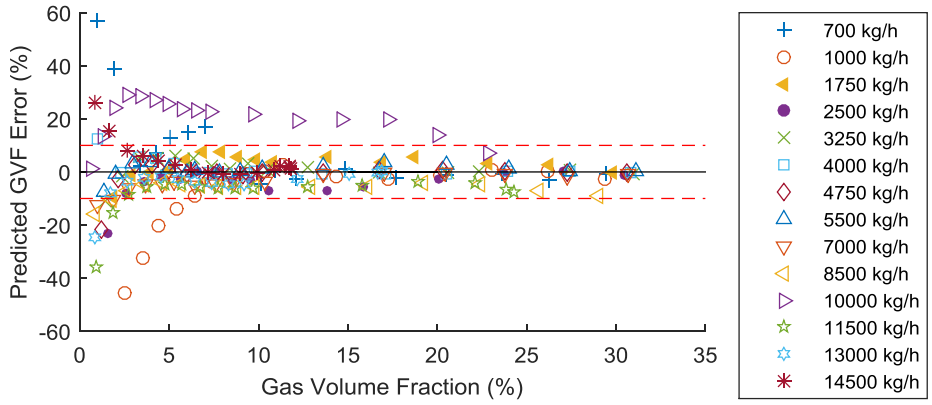
Since the gas volume fraction under the experimental conditions ranges from 0 to 30% and the intrinsic complexity of two-phase flow, the relative errors of the predicted gas volume fraction from the BP-ANNs are quite large when the gas volume fraction is below 5%. As the entrained gas increases, the errors from the training dataset are mostly within ±10% (the red dash lines in Figure 4.15). For the test dataset, however, all the errors are less than ±10% on the vertical pipeline, even under the low flowrate conditions.



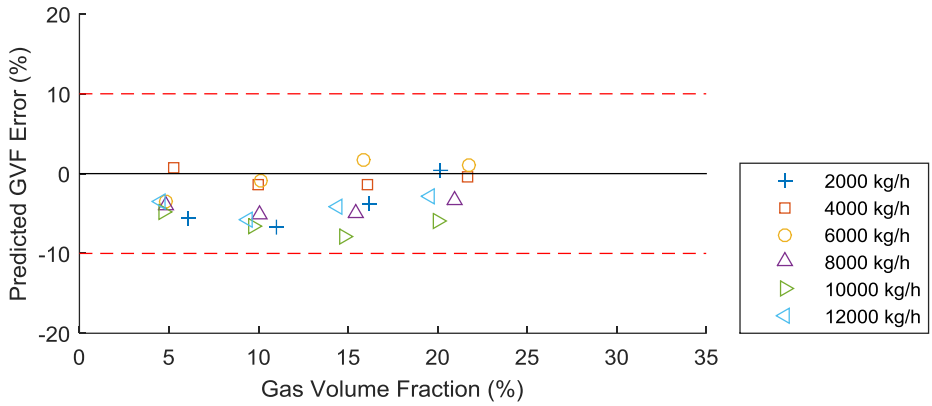
(a) BP-ANN for Model-H-G with training data



(b) BP-ANN for Model-H-G with test data



(c) BP-ANN for Model-V-G with training data



(d) BP-ANN for Model-V-G with test data

Figure 4.15 Errors of the predicted gas volume fraction from the trained BP-ANNs

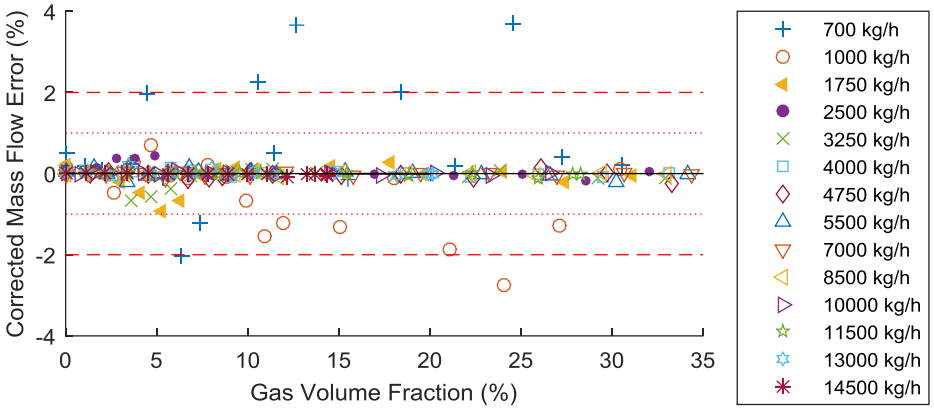
4.5.1.2 SVM Models

SVM models are also established for both installation conditions. An important difference between the SVM and ANN models is that the SVM leads to a unique deterministic model for each dataset while ANNs depend on a random initial choice of synaptic weights and cannot produce the fixed results. Through a direct comparison of the performances of SVM between the four kinds of kernel function (Table 4.4), we know that the SVM with radial basis function (RBF) generates the smallest NRMSE among the four models.

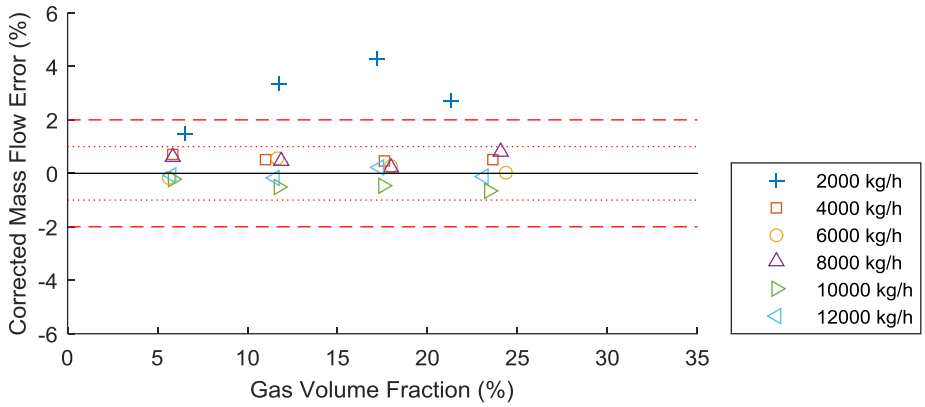
Table 4.4 NRMSE of the SVM models with different kernel functions

Model	**Training data**				**Test data**			
	Linear	**Polynomial**	**RBF**	**Sigmoid**	**Linear**	**Polynomial**	**RBF**	**Sigmoid**
H-L	5.62%	11.12%	0.11%	889.50	7.44%	10.97%	0.58%	738.32
V-L	6.32%	10.33%	0.10%	911.90	9.39%	11.42%	0.57%	777.32
H-G	21.37%	28.37%	3.44%	606.58	2.6%	5.68%	3.29%	138.03
V-G	27.27%	34.08%	2.16%	683.13	3.71%	6.78%	3.2%	171.56

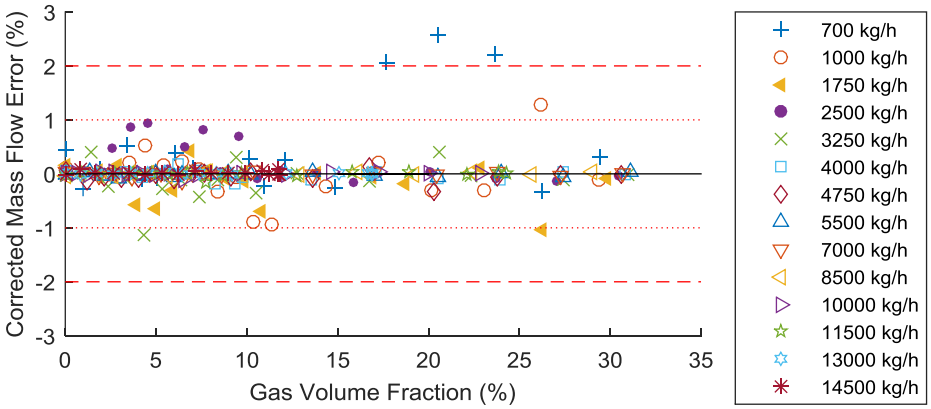
From Figure 4.16 (a) and (c), the SVM model performs well to fit with training data and limit the relative errors on horizontal and vertical pipelines to ±1% or less, except some points at 700 kg/h and 1000 kg/h, which is a common problem for the ANN and SVM models. The generalization ability of the SVM model is proven as shown in Figure 4.16 (b) and (d). Most errors from the SVM models with the test data are reduced to ±1%.



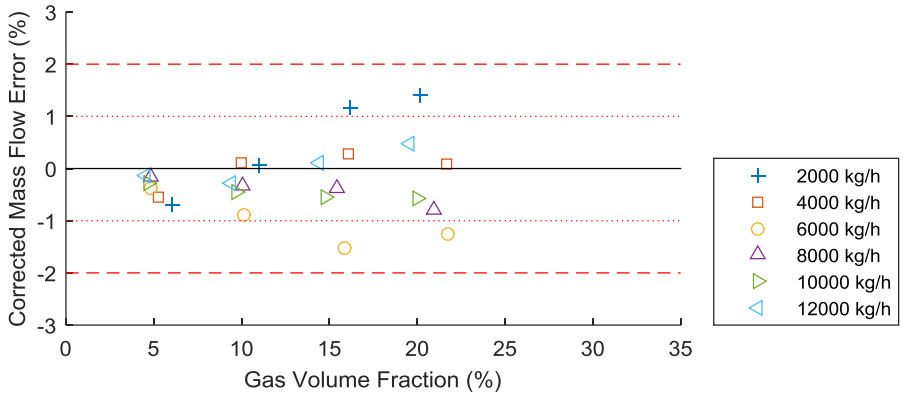
(a) SVM for Model-H-L with training data



(b) SVM for Model-H-L with test data



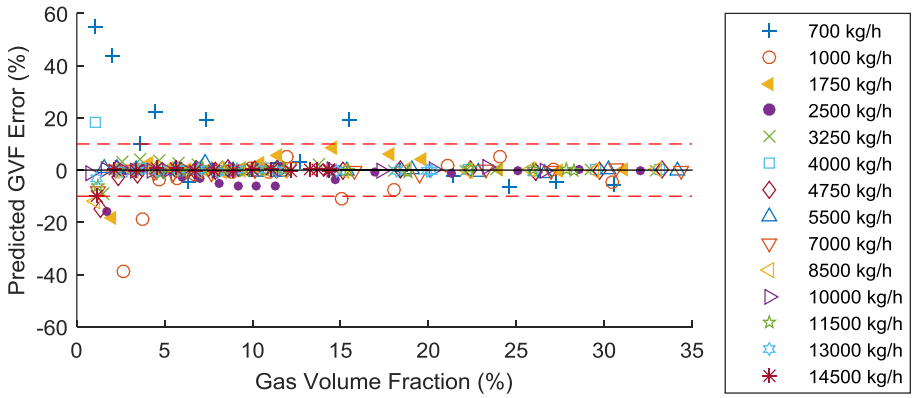
(c) SVM for Model-V-L with training data



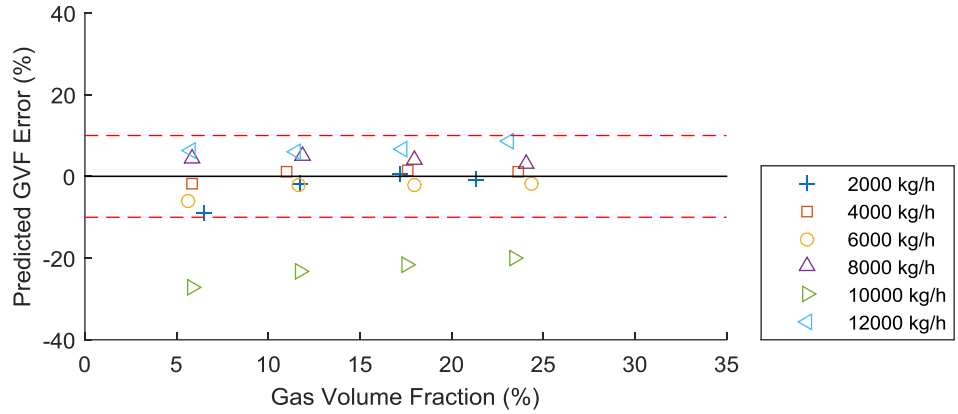
(d) SVM for Model-V-L with test data

Figure 4.16 Errors of the corrected liquid mass flowrate from the SVMs

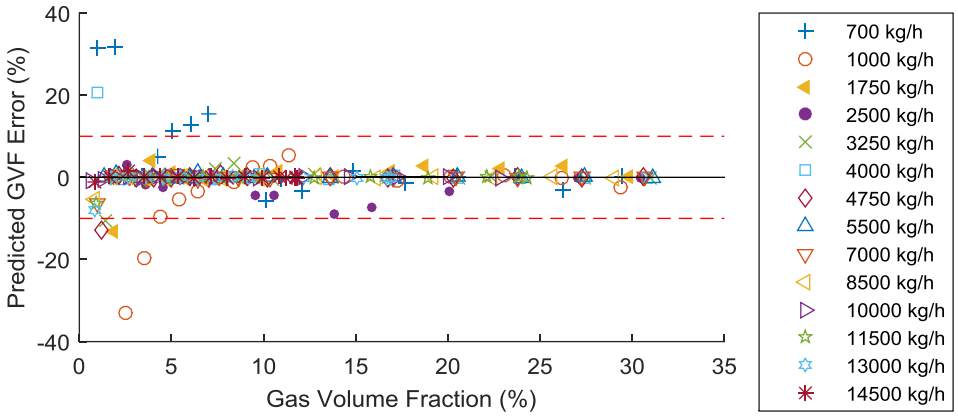
Figure 4.17 indicates that, for gas volume fraction prediction, a less number of points from the SVM models have an error beyond ±10% with the training dataset. This is probably due to the fact that the samples at 1000 kg/h flowrate are far away from the support vectors in the network. The relative errors in the predicted gas volume fraction with the test dataset at the flowrate of 1000 kg/h is larger than other test data.



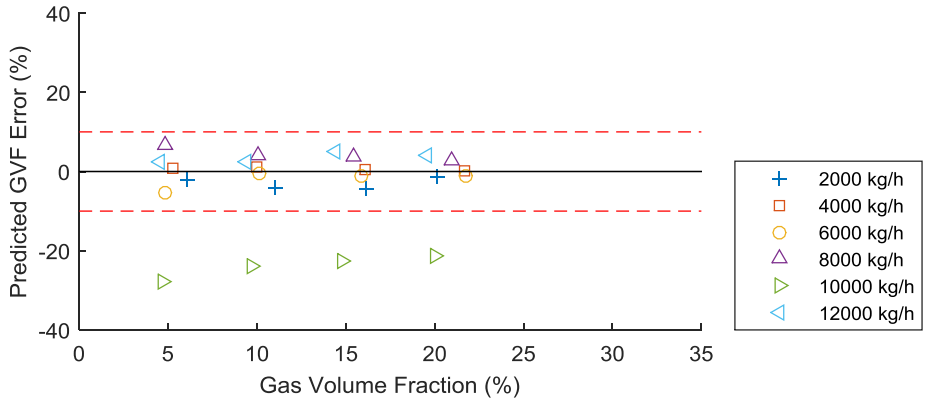
(a) SVM for Model-H-G with training data



(b) SVM for Model-H-G with test data



(c) SVM for Model-V-G with training data



(d) SVM for Model-V-G with test data

Figure 4.17 Errors of the predicted gas volume fraction from the SVMs

4.5.1.3 GP Models

Four GP models are established in this study for correcting the liquid mass flowrate and predicting the gas volume fraction, respectively, for horizontal and vertical installations of Coriolis flowmeters. The parameters that were set in the GP algorithms include: population size 250, tournament size 25, elitism 0.7, maximum number of genes allowed in an individual 6, function set {×, -, +, tanh, mult3, add3}, terminal sets {x_1, x_2, x_3, x_4} for models H-L and V-L and {x_1, x_2, x_3} for models H-G and V-G.

The GP-based formulations for the four models are given below:

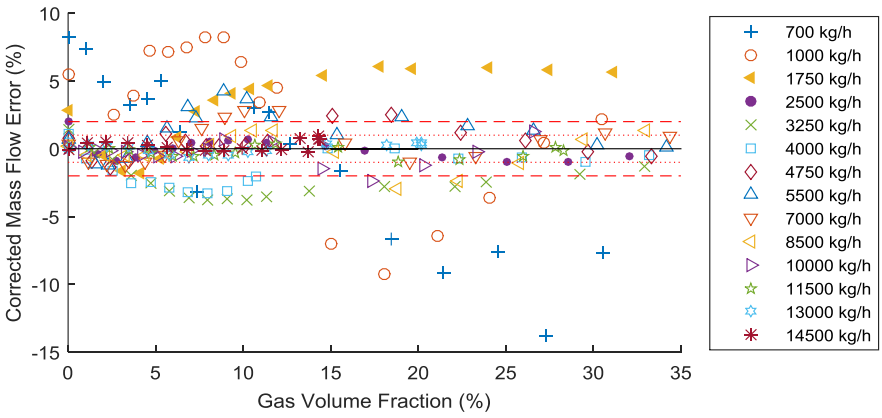
$$y_{H-L} = 0.994x_2 - 2633x_1 + 4300\tanh(x_1)\tanh(x_3) + 13.2x_1x_4 + 0.00571x_2x_3 - 0.0995x_2x_3\tanh(x_1) + 62.4 \quad (4\text{-}7)$$

$$y_{V-L} = x_2 + 57.6x_3 - 0.161x_4 + 29.8x_1x_4 + 871\tanh(x_3^2x_4)\tanh(x_1) - 0.00913x_4(x_3 + x_1x_4) - 0.122x_1x_2x_3 + 32.5 \quad (4\text{-}8)$$

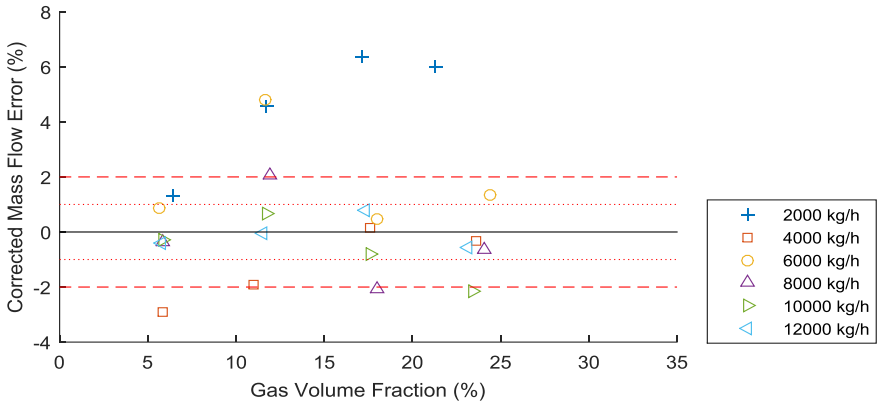
$$y_{H-G} = 0.783x_1 + 1.6e^{-6}x_2 + 0.00278x_3 - 0.114x_1x_3 + 0.159x_1^2x_3 + 6.82e^{-5}x_3^3 - 0.0182 \quad (4\text{-}9)$$

$$y_{V-G} = 1.01x_1 - 5.49e^{-7}x_2 - 0.0217x_3 - 2.74e^{-7}\tanh(x_1) - 1.05e^{-4}x_1x_2 + 2.74e^{-6}x_2x_3 + 0.00253x_1x_3^2 - 1.05e^{-4}x_1^2x_3 - 2.74e^{-7}x_1x_2x_3 + 0.00587 \quad (4\text{-}10)$$

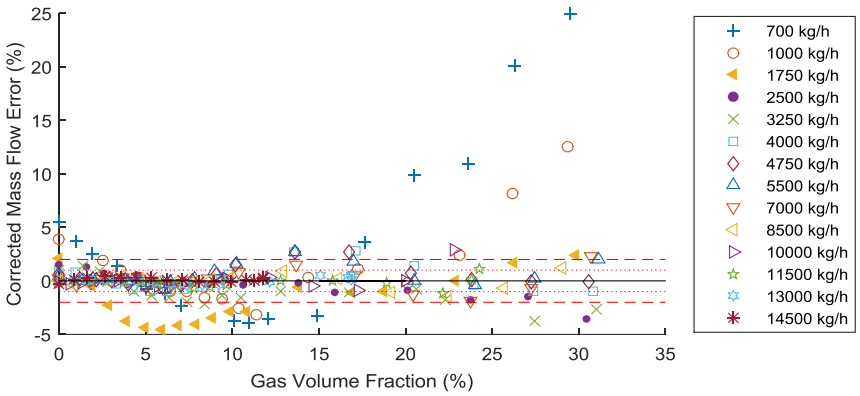
The errors of the corrected mass flowrate on the training dataset using GP are higher by -15% and 25%, respectively, under horizontal and vertical installations (Figure 4.18 (a) and (c)), which results in larger errors on the test dataset (Figure 4.18 (b) and (d)). As can be seen that larger errors normally occur at low flowrates, which indicates that the GP models are unable to approximate all the data.



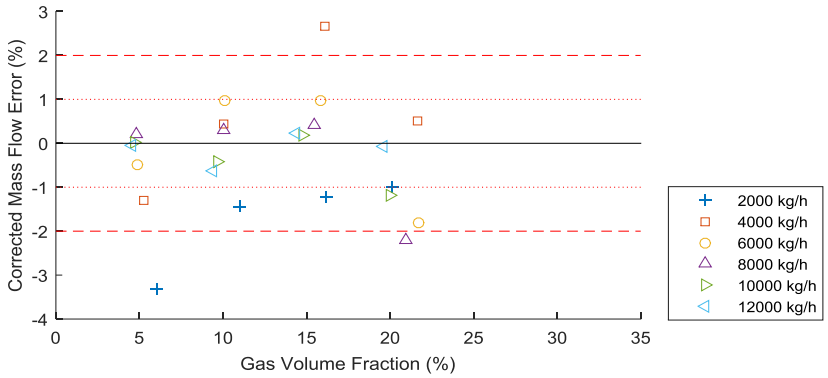
(a) GP for Model-H-L with training data



(b) GP for Model-H-L with test data



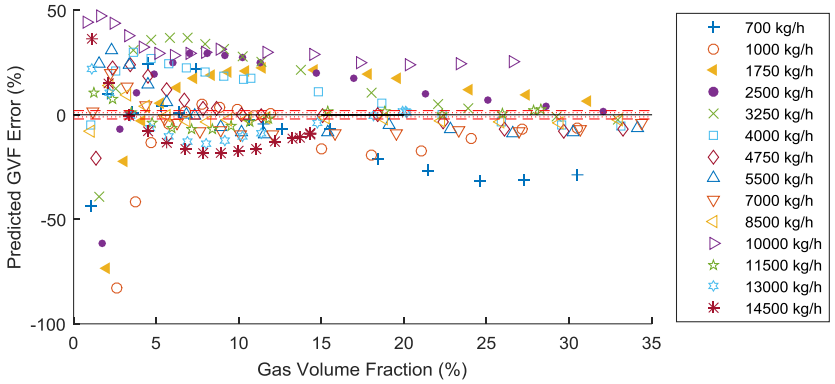
(c) GP for Model-V-L with training data



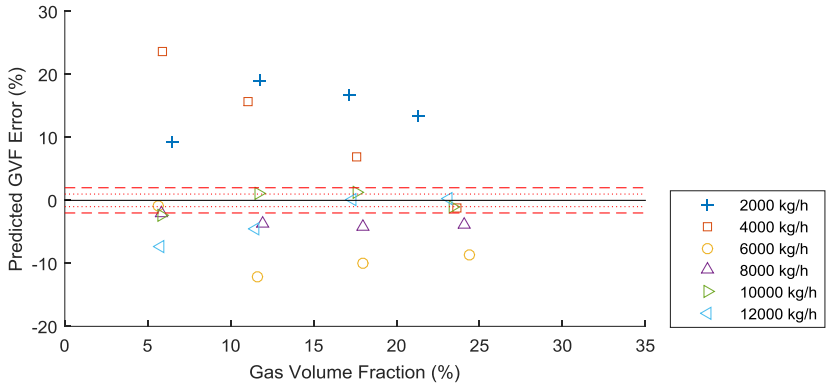
(d) GP for Model-V-L with test data

Figure 4.18 Errors of the corrected liquid mass flowrate from the GPs

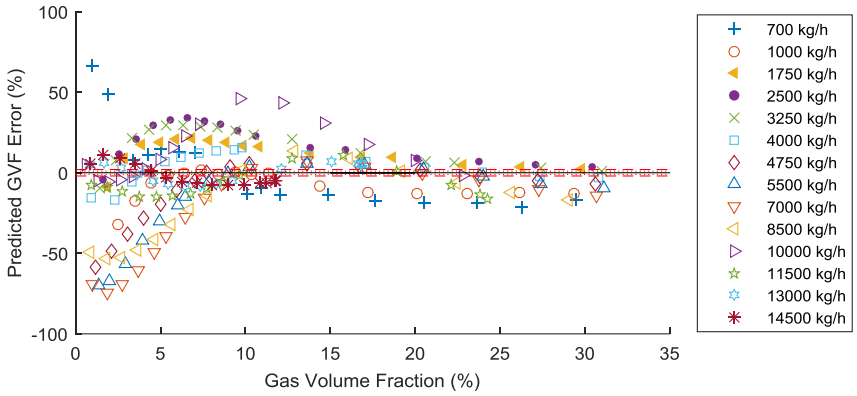
As shown in Figure 4.19, for the prediction of gas volume fraction, the outputs of GP models have large errors for low gas entrainment and low flowrates. The relative errors with test data reach 25% and -50% on horizontal and vertical pipes, respectively.



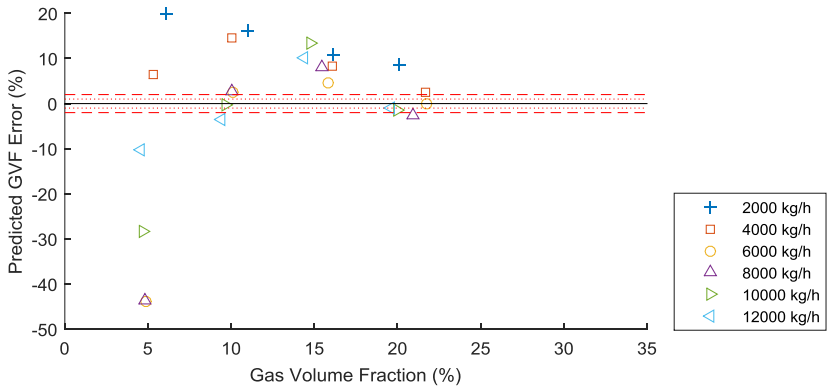
(a) GP for Model-H-G with training data



(b) GP for Model-H-G with test data



(c) GP for Model-V-G with training data



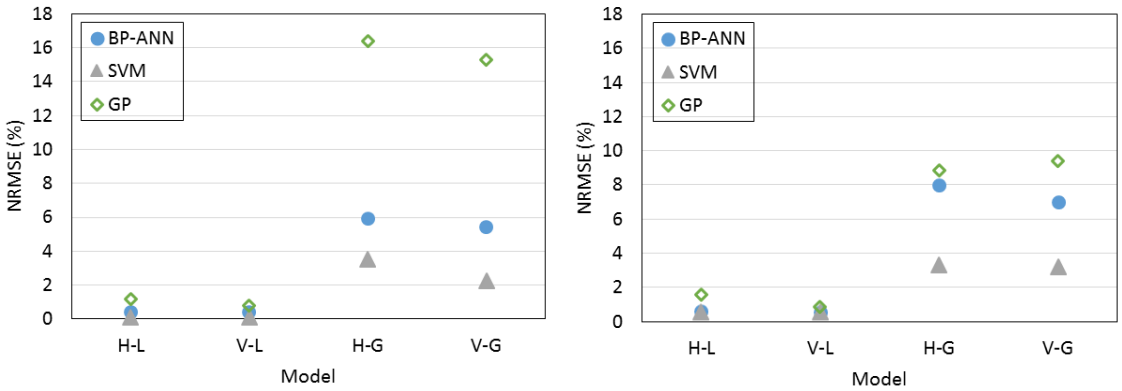
(d) GP for Model-V-G with test data

Figure 4.19 Errors of the predicted gas volume fraction from the GPs

4.5.2 Performance Comparison between Data-Driven Models

4.5.2.1 Robustness

In order to assess the robustness of the developed models for two-phase flow measurement, the averaged NRMSE values are shown in Figure 4.20. Among the models for liquid mass flowrate correction and gas volume fraction prediction, GP produces larger errors than the other three techniques, while the SVM models yield less errors. With the test dataset, BP-ANN and SVM methods perform similarly on Model-H-L and Model-V-L. However the SVM models are significantly better than the BP-ANN and GP models for the prediction of gas volume fraction. Moreover, BP-ANN has uncertain parameters to optimize which could result in differences in performance. However, due to their fixed structure, the SVM models produce repeatable results all the time. This outcome suggests that the SVM models are superior to both ANN and GP models in terms of robustness.



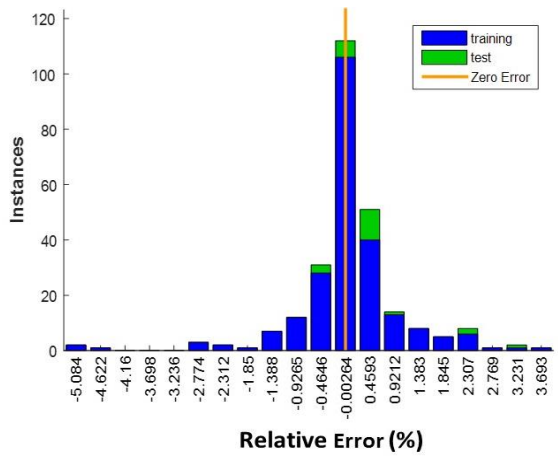
(a) Performance with training dataset (b) Performance with test dataset

Figure 4.20 Performance comparison between ANNs, SVMs and GPs

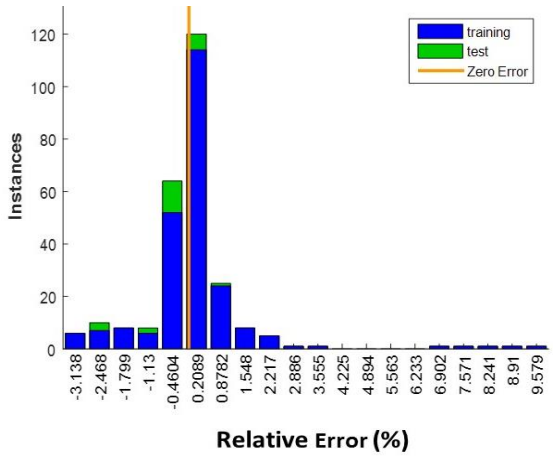
4.5.2.2 Accuracy

Figure 4.21 depicts the relative error histograms of the ANNs, SVMs and GPs for corrected liquid mass flowrate. It is clear that the error distributions of the GP and ANN models are much wider and dispersive than the SVM models. Through comparing the

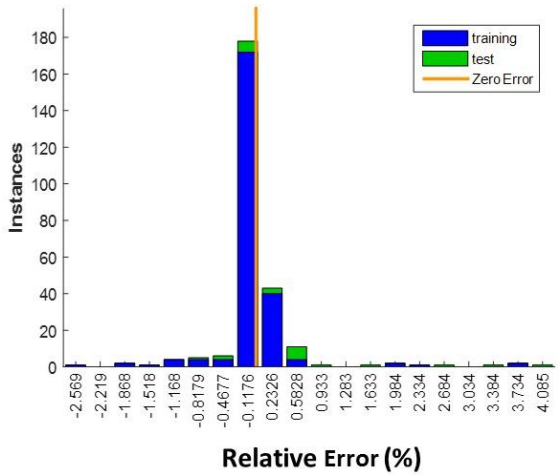
mean value and standard deviation of the errors between the eight error distributions (Table 4.5), we can see that the SVM models with the lowest mean value and standard deviation outperform the BP-ANN and GP models for liquid mass flowrate measurement on both horizontal and vertical pipelines. Moreover, the data driven models (mean value 0.0008% and standard deviation 0.40%) on the vertical pipeline perform better than those on the horizontal pipeline (mean value 0.0585% and standard deviation 0.66%).



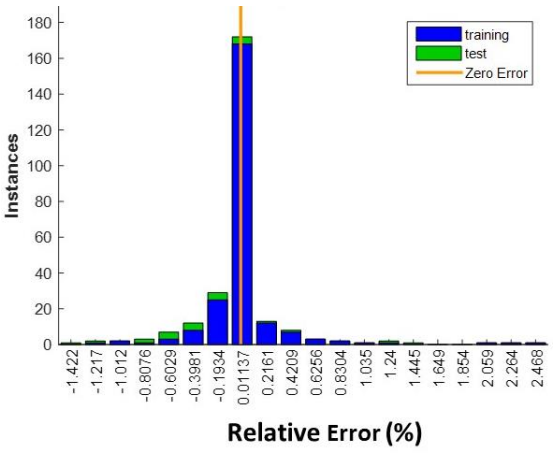
(1) BP-ANN: H-L



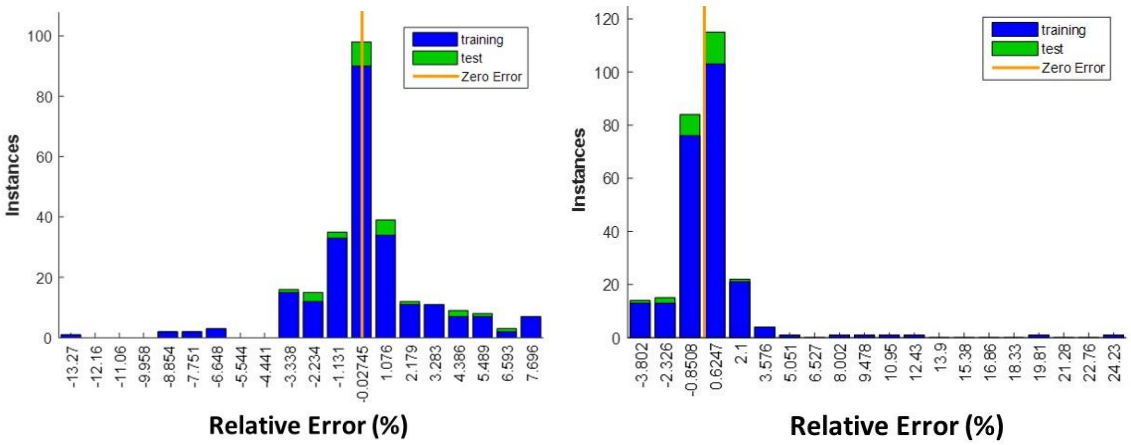
(2) BP-ANN: V-L



(3) SVM: H-L



(4) SVM: V-L



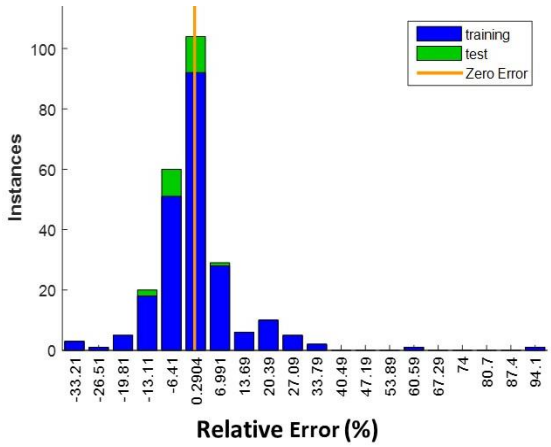
(5) GP: H-L (6) GP: V-L

Figure 4.21 Relative error histogram of ANNs, SVMs and GPs for corrected liquid mass flowrate

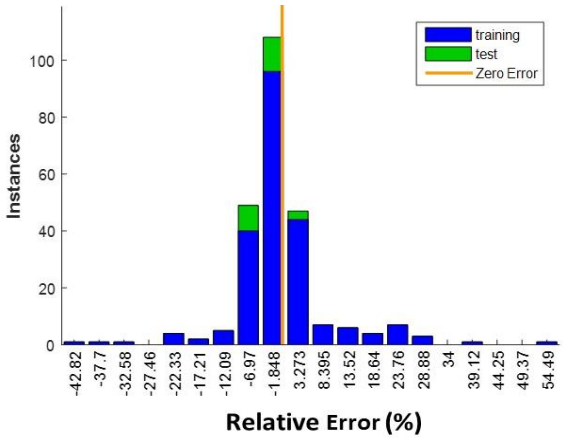
Table 4.5 Mean and standard deviation of the relative error distribution for liquid mass flowrate correction

Model		BP-ANN	SVM	GP
Model H-L	Mean (%)	0.0823	0.0585	0.2405
	Standard deviation (%)	1.03	0.66	2.83
Model V-L	Mean (%)	0.0548	0.0008	0.1660
	Standard deviation (%)	1.50	0.40	2.77

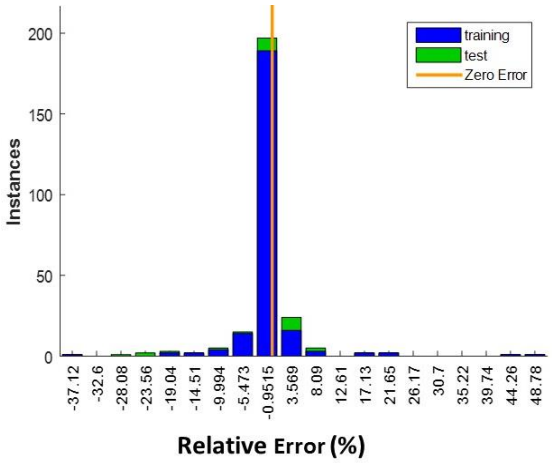
Figure 4.22 shows the relative error histograms of the three kinds of models for gas volume fraction prediction. GP models have a larger range of errors than all other models. The error distribution of the SVM model is much narrower than the ANN models for the measurement of gas volume fraction. It can be seen that most errors of the SVM models are concentrated around zero line.



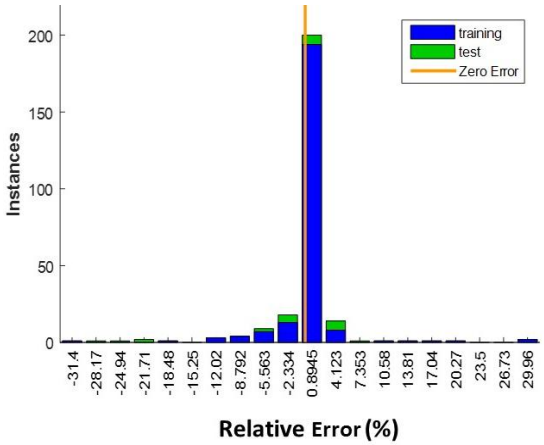
(1) BP-ANN: H-G



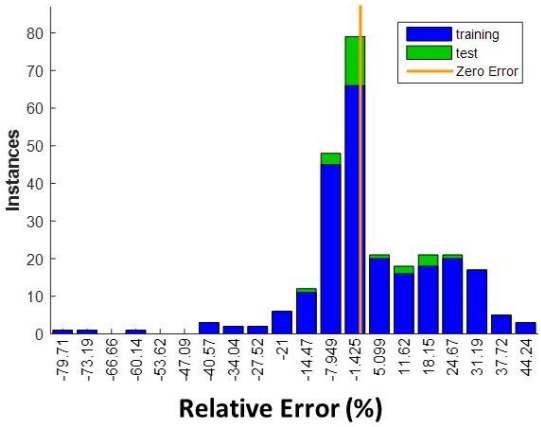
(2) BP-ANN: V-G



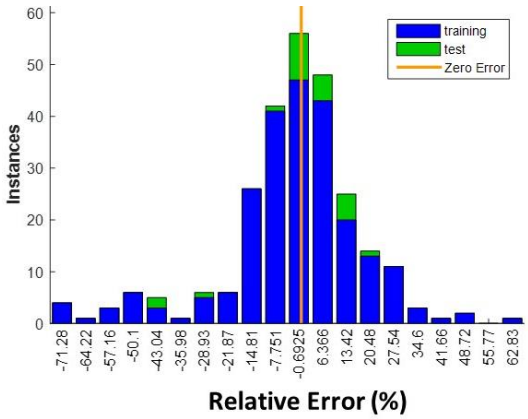
(3) SVM: H-G



(4) SVM: V-G



(5) GP: H-G



(6) GP: V-G

Figure 4.22 Relative error histogram of ANNs, SVMs and GPs for gas volume fraction prediction

Table 4.6 shows that the standard deviations of the SVM models are smaller than those of the BP-ANN and GP models on both horizontal and vertical pipelines.

Table 4.6 Mean and standard deviation of the relative error distribution for gas volume fraction prediction

Model		**BP-ANN**	**SVM**	**GP**
Model H-G	Mean (%)	0.17	-0.25	3.15
	Standard deviation (%)	11.88	6.95	17.70
Model V-G	Mean (%)	-0.18	-0.38	-1.99
	Standard deviation (%)	9.70	5.57	20.62

In order to assess the accuracy of the ANN, SVM and GP models, the percentage of experimental data for each model which can achieve the accuracy of ±2% and ±1%, respectively, for liquid mass flowrate measurement and ±10% for gas volume fraction prediction is calculated and summarized in Table 4.7. For liquid mass flowrate measurement with the SVM models, 93.49% of the experimental data yield a relative error less than ±1% on the horizontal pipeline whilst 96.17% of the results are within ±1% on the vertical installation. The SVM models predict the gas volume fraction with a relative error less than 10% for 93.10% and 94.25% of the test conditions on horizontal and vertical installations, respectively. Therefore, the SVM models perform significantly better than the BP-ANN and GP models for two phase flow measurement in terms of robustness and accuracy.

Table 4.7 Accuracy comparisons of ANN, SVM and GP models

Model	**H-L**		**V-L**		**H-G**	**V-G**
Error Limit	**≤ ±2%**	**≤ ±1%**	**≤ ±2%**	**≤ ±1%**	**≤ ±10%**	**≤ ±10%**
BP-ANN	91.95%	80.08%	89.66%	79.69%	79.31%	86.21%
SVM	96.93%	93.49%	98.85%	96.17%	93.10%	94.25%
GP	68.20%	54.41%	83.14%	67.05%	55.56%	54.79%

4.6 Summary

A wide range of experimental tests of 1-inch Coriolis flowmeters have been conducted with air-water two-phase flow. The test results have demonstrated the effectiveness of the proposed two-phase flow measurement strategy using Coriolis flowmeters incorporating soft computing techniques.

The evaluation results of PMI, GA-ANN and GP demonstrated that 1) In terms of time efficiency the PMI and IIS algorithms have outperformed the GA-ANN algorithm; 2) In terms of selection accuracy PMI provides more effective variables for the measurement of liquid mass flowrate than the other two approaches. For the prediction of gas volume fraction, the IIS has selected the most important information and thus generate a better performing model; 3) With regard to the experimental data obtained in this study, the most important variable set for the measurement of liquid mass flowrate includes observed density, apparent mass flowrate, DP and damping while those for the prediction of gas volume fraction include observed density, apparent mass flowrate and DP.

The comparison results between BP-ANN, SVM and GP suggested that the SVM models are superior to the ANN and GP models for two-phase flow measurement in terms of robustness and accuracy. For liquid mass flowrate measurement with the SVM models, 93.49% of the experimental data yield a relative error less than ±1% on the horizontal pipeline, while 96.17% of the results are within ±1% on the vertical installation. The SVM models predict the gas volume fraction with a relative error less than ±10% for 93.10% and 94.25% of the test conditions on the horizontal and vertical installations, respectively.

Chapter 5

Experimental Tests with Air-Water and Air-Oil Flows

5.1 Introduction

This chapter presents results using soft computing techniques on 2-inch Coriolis flowmeters (KROHNE OPTIMASS 6400 S50) with air-water and air-oil two-phase flows. Experimental work was conducted on an air-water two-phase flow test rig at KROHNE Ltd and an air-water-oil three-phase flow test rig at Tianjin University, respectively. Through analysing the original errors of liquid mass flowrate from the Coriolis flowmeter under test, effects of operating pressure, temperature, installation orientation of the Coriolis flowmeter and fluid properties (i.e. density and viscosity) on the performance of Coriolis flowmeters are discussed.

It is concluded in Chapter 4 that SVM models outperform BP-ANN and GP models in terms of accuracy and robustness. Consequently, SVM models are applied to estimate liquid mass flowrate and gas volume fraction of air-water and air-oil flows in this Chapter. In addition, Least Squares Support Vector Machine (LSSVM) models are developed for each case. Performance comparison between the SVM and LSSVM models is carried out and discussed in terms of generalization ability and computational complexity.

5.2 Experimental Conditions

Experimental tests of 2-inch Coriolis flowmeters (OPTIMASS 6400 S50) with air-water flow were conducted at KROHNE Ltd in July 2015. As shown in Figure 5.1, a Coriolis flowmeter as a reference meter was installed in the upstream and the Coriolis flowmeter

under test was installed in the downstream of the horizontal section. A gas flow controller between the two meters was used to control the amount of gas entrainment. The same Coriolis flowmeter was tested on air/water/oil three-phase flow test rig at Tianjin University in August 2015. As shown in Figure 5.2, the different positions including horizontal and vertical installations of the flowmeter under test were taken in account. Additionally, the Coriolis flowmeter was tested with air-water and air-oil flows respectively. A turbine flowmeter, a magnetic flowmeter and an Alicat flowmeter, with uncertainty of 1%, 1% and 0.35% respectively, were installed on single oil, water and air flow sections as reference meters, respectively.

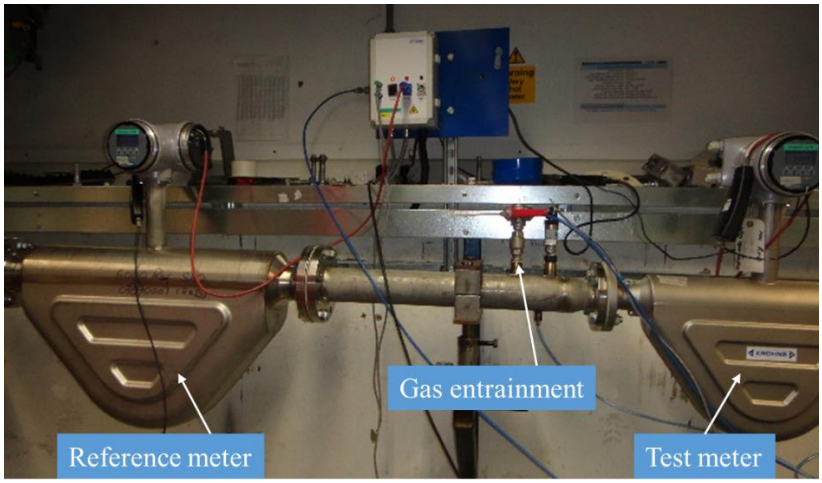


Figure 5.1 Installation of the Coriolis flowmeter under test at KROHNE Ltd

(a) Horizontal section

(b) Vertical section

Figure 5.2 Installation of the Coriolis flowmeter under test at Tianjin University

Experimental conditions are summarized in Table 5.1. Five experimental cases are named according to the test rig, meter size, test flow and installation orientation. All the tests were conducted with the liquid flowrate ranging from 2000 kg/h to 12000 kg/h and the gas volume fraction between 0 and 60%. In the case of KROHNE tests, the fluid temperature was around 20°C and the inlet pressure of the Coriolis flowmeter under test was about 0.5 bar. However, there was no cooling system on the Tianjin test rig and hence the fluid temperature was varying between 30°C and 36°C. Meanwhile, the inlet pressure was between 0.58 bar and 2.96 bar.

Table 5.1 Experimental conditions

Test code	**Test rig**	**Meter size**	**Flow**	**Position**	**Pressure (bar)**	**Temperature (°C)**
K2AWH	KROHNE	2 inch	air-water	Horizontal	0.47~0.54	18.1~21.9
T2AWH	Tianjin	2 inch	air-water	Horizontal	1.00~2.61	32.6~35.9
T2AWV	Tianjin	2 inch	air-water	Vertical	0.63~1.48	31.0~33.8
T2AOH	Tianjin	2 inch	air-oil	Horizontal	0.82~2.96	30.6~34.1
T2AOV	Tianjin	2 inch	air-oil	Vertical	0.58~1.45	29.6~33.1

It is observed from the sight window on the horizontal test section that the flow pattern is intermittent flow, including slug flow and plug flow for the five cases. For each test, 90 experimental data were acquired as training data while 24 data as test data. The circular markers in Figure 5.3 indicate the training data and triangular markers present the test data. It can be seen from the test matrix that some test data are within the range of training data and the other points at the flowrates of 2000 kg/h and 12000 kg/h are out of range. The data out of range are regarded as unknown data and are used to assess the generalization ability of the data-driven models.

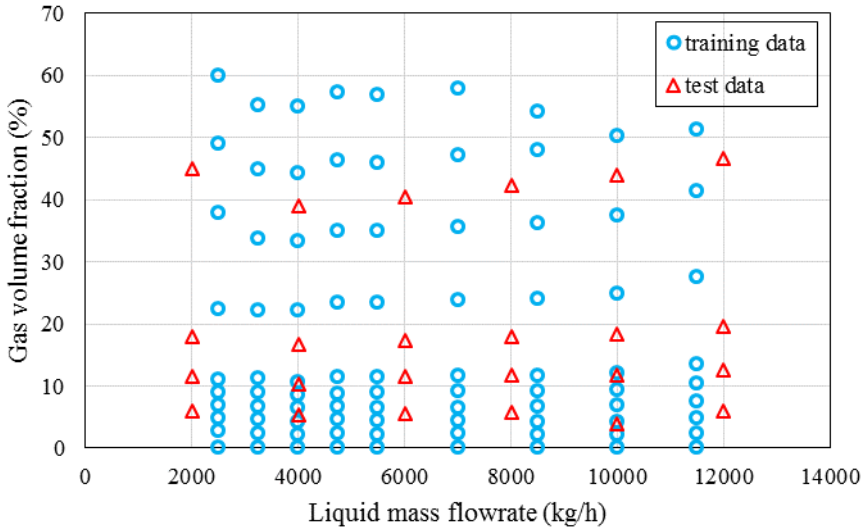


Figure 5.3 Experimental test points of the 2-inch oil/air/water flows

5.3 Analysis of Original Errors

The original mass flow errors of the 2-inch Coriolis flowmeter in the case of K2AWH are plotted in Figure 5.4. The results from Tianjin tests with air-water and air-oil flows are shown in Figures 5.5-5.8, respectively. The effects of several factors such as pressure, temperature, installation orientation and fluid properties are discussed in the following sections.

5.3.1 Effect of Pressure and Temperature on Mass Flow Errors

It is well known that pressure and temperature are two important factors affecting the physical properties of the fluid such as density, viscosity and compressibility. Although the measurement of single phase flow using Coriolis flowmeters is not affected by variations of density and viscosity, the performance of Coriolis flowmeters is different under two-phase flow conditions due to relative motion between liquid and gas phases. As shown in Figures 5.4 and 5.5, the original errors of the Coriolis flowmeter have a decreasing trend with different gradients in the cases of K2AWH and T2AWH. Most errors are negative except that one point at the flowrate of 2500 kg/h is positive. Due to

the different environmental conditions between the two experimental cases, the higher temperature and pressure of the fluid in the Tianjin tests result in smaller errors than the KROHNE tests.

5.3.2 Effect of Installation Orientation on Mass Flow Errors

As discussed in Section 4.2.2, different installation positions of the 1-inch Coriolis flowmeter under test result in different performances with air-water two-phase flow. Original errors of the 2-inch Coriolis flowmeter with air-water flow on horizontal and vertical positions are shown in Figures 5.5 and 5.6, respectively. Similarly, the results with air-oil flow are plotted in Figures 5.7 and 5.8. As bubbles in the vertical pipe are more evenly distributed than those in the horizontal pipe, they have less influence on the vibration of Coriolis measuring tubes. As a result, the mass flow errors of Coriolis flowmeters on the vertical position are smaller than those on the horizontal position.

5.3.3 Effect of Fluid Properties on Mass Flow Errors

Different liquids have different density and viscosity and thus the relative motion between the liquid phase and the gas phase is different. The density of water and oil is 995.68 kg/m3 and 876.14 kg/m3, respectively, at the temperature of 30 °C. They both decrease with gas entrained. The kinematic viscosity of water and oil are 0.00101 Pa·s and 0.0147 Pa·s, respectively. Through a comparison between the results from T2AWH and T2AOH, the errors from air-oil flow on horizontal position are more linear than those from air-water flow. Similarly, the results from T2AWV and T2AOV present that the errors from air-oil flow are larger than those from air-water flow.

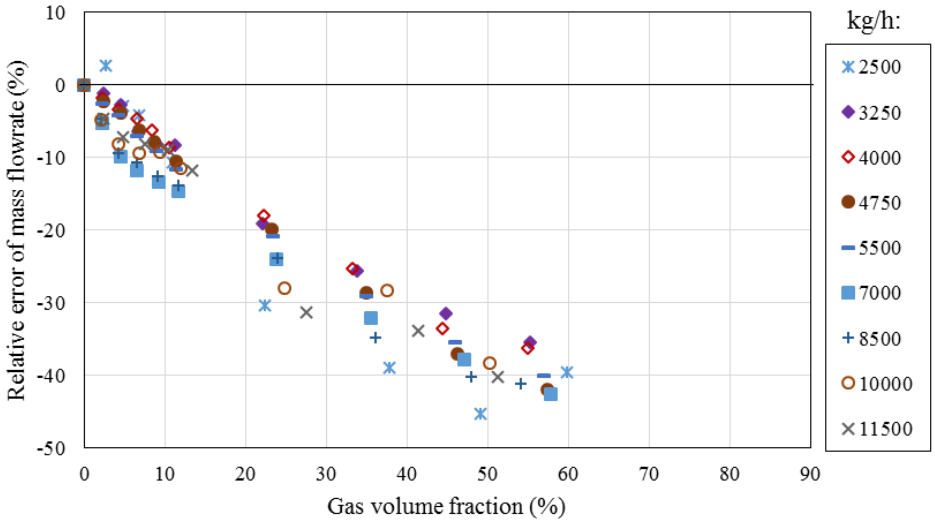


Figure 5.4 Original errors of the liquid mass flowrate for K2AWH

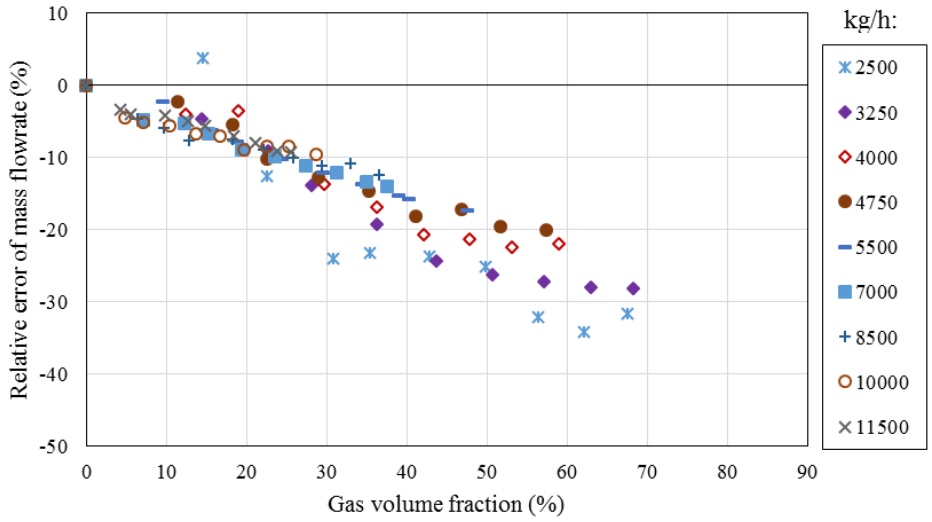


Figure 5.5 Original errors of the liquid mass flowrate for T2AWH

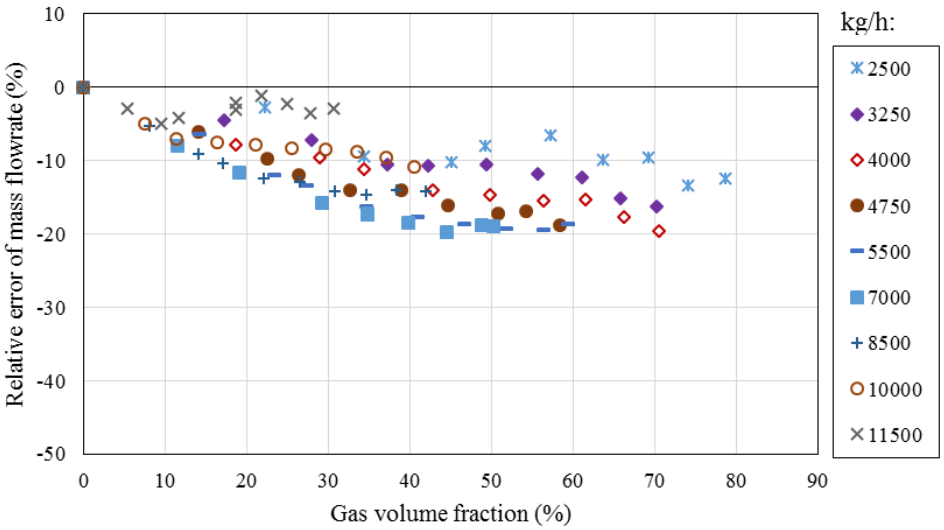


Figure 5.6 Original errors of the liquid mass flowrate for T2AWV

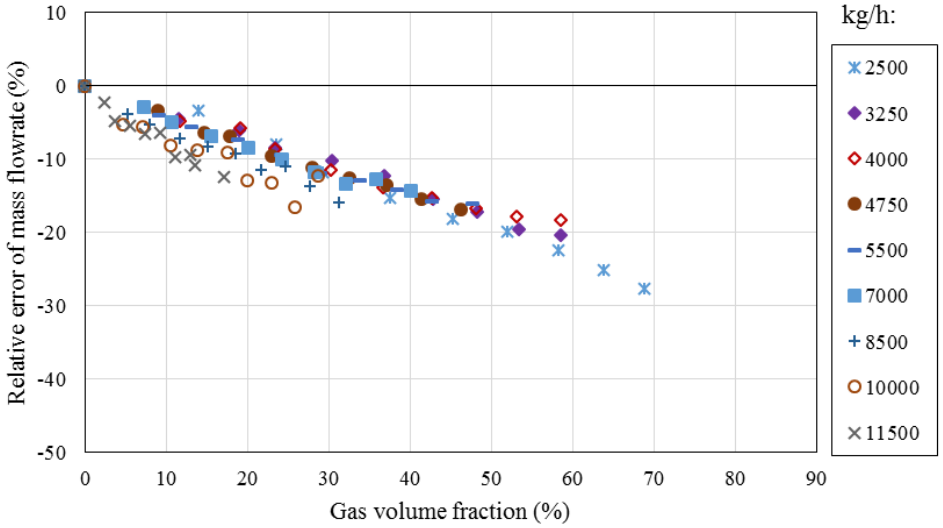


Figure 5.7 Original errors of the liquid mass flowrate for T2AOH

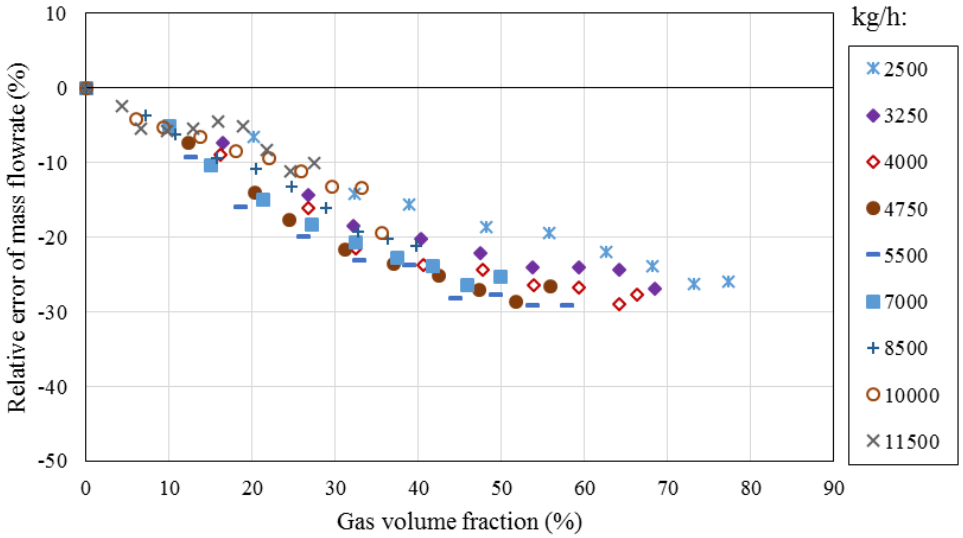


Figure 5.8 Original errors of the liquid mass flowrate for T2AOV

5.4 Comparison between SVM and LSSVM Models

As SVM models outperform BP-ANN and GP models in terms of accuracy and robustness through comparative investigations in Chapter 4, SVM is adopted in this Chapter to build data-driven models for the five cases. In practice, data-driven models are usually restricted by their generalization abilities. A model that not only performs well with the training data but also deals well with the test data which are not part of the training process is regarded as a good model. In this section, SVM and LSSVM models are established for the five cases to estimate liquid mass flowrate and gas volume

fraction, respectively. Performance comparison, especially the generalization ability, between SVM and LSSVM models is discussed through analysing the relative error of model output with test data. The test data (the triangular markers in Figure 5.3) are divided into two sets: the data points within the range of training data are defined as test I while the data out of range constitute test II.

5.4.1 SVM and LSSVM Models for the Correction of Liquid Mass Flowrate

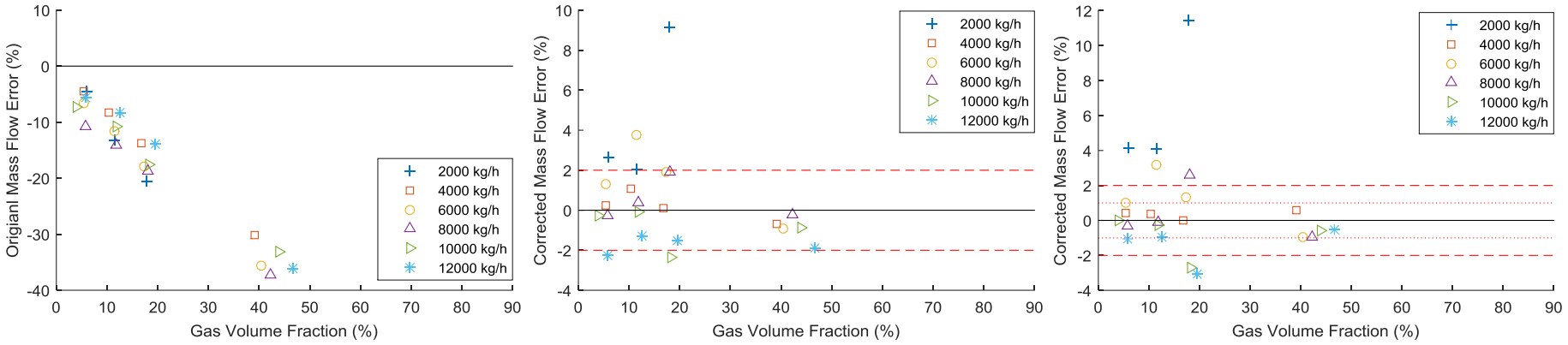
For each case, there are 14 variables available as outlined in Table 4.1. The PMI based input variable selection is applied to the five cases respectively and thus four variables including apparent mass flowrate, observed density, damping and DP are obtained and determined as the inputs of the data-driven models.

In developing the SVM and LSSVM models, the cost parameter C and kernel parameter γ should be determined appropriately. To obtain the most suitable values of C and γ for data-driven models, a cross-validation approach is used. Specifically, the training set is divided into a number of subsets (in the present study, five subsets) with an equal size. Sequentially one subset is tested using the model trained on the remaining subsets. Thus, each instance of the whole training set is predicted once. The C and γ with which the best validation accuracy is obtained are taken as the kernel parameters to train the whole training set. The cross-validation procedure can ameliorate or prevent the over-fitting problem. The kernel function is determined through comparing the performance of the model with different functions including linear, polynomial, radial basis function and sigmoid function. It is turned out that the SVM and LSSVM models with radial basis function outperform the others in the cases of K2AWH, T2AWH, T2AWV and T2AOV, while the linear kernel function is more suitable for the SVM and LSSVM models in the case of T2AOH.

The performance of SVM and LSSVM models on test data is shown in Figures 5.9-5.13. For each case, original mass flow errors from the Coriolis flowmeter under two-phase flow condition and corrected errors of the liquid mass flowrate from SVM and LSSVM

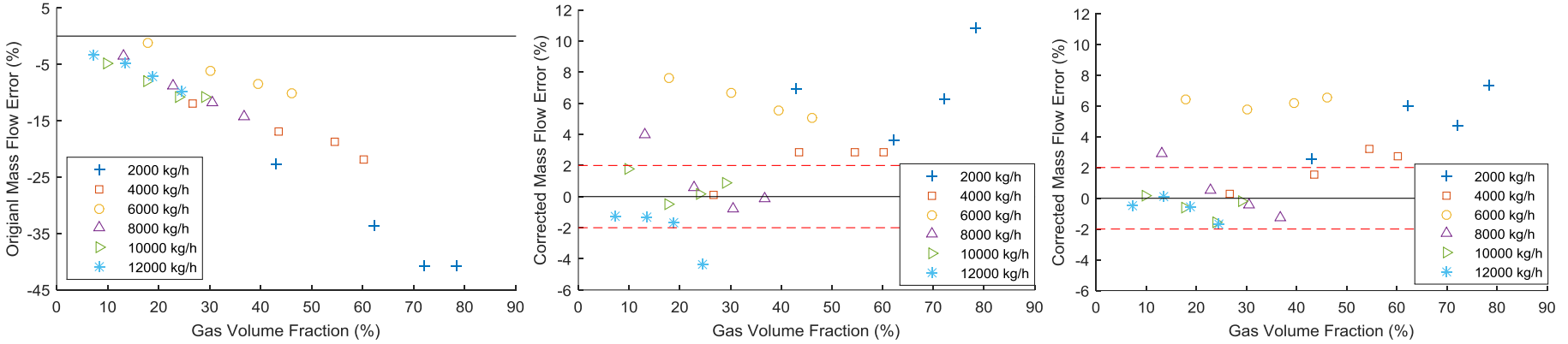
are presented. In the case of K2AWH, the original errors are all negative and drop dramatically with the gas entrainment. The maximum error is up to -40%. After correction by SVM and LSSVM, most errors are significantly reduced to ±2% (the red dashed lines) except for some points at the lowest flowrate of 2000 kg/h. It is observed that SVM and LSSVM obtain similar performance for K2AWH. In the other four cases, the corrected errors beyond ±2% are observed at the lowest (2000 kg/h) and highest flowrates (12000 kg/h) as these test points are out of range of the training data. The errors at 2000 kg/h seem to be larger than the errors at 12000 kg/h. This is due to the unstable slug flow at the low flowrate which is difficult to extrapolate from the available data while the flow at high flowrate is more stable. It is clear that the errors at highest and lowest flowrate from LSSVM models are smaller than the errors from SVM models through comparing (b) and (c) in Figures 5.10-5.13.

The NRMSE of original errors and corrected errors from SVM and LSSVM models are outlined in Table 5.2. It particularly highlights the performance comparison between SVM and LSSVM with test data. It can be seen from the NRMSE values that the original NRMSEs are drastically reduced after correction with SVM and LSSVM models. However, there exists overcorrection in SVM models with test II and hence larger errors after correction in the case of T2AWV. As observed from KROHNE test, the SVM and LSSVM models obtain similar generalization ability with the test data. However, the LSSVM models achieve smaller NRMSEs than the SVM models and result in much better generalization performance for the Tianjin tests.



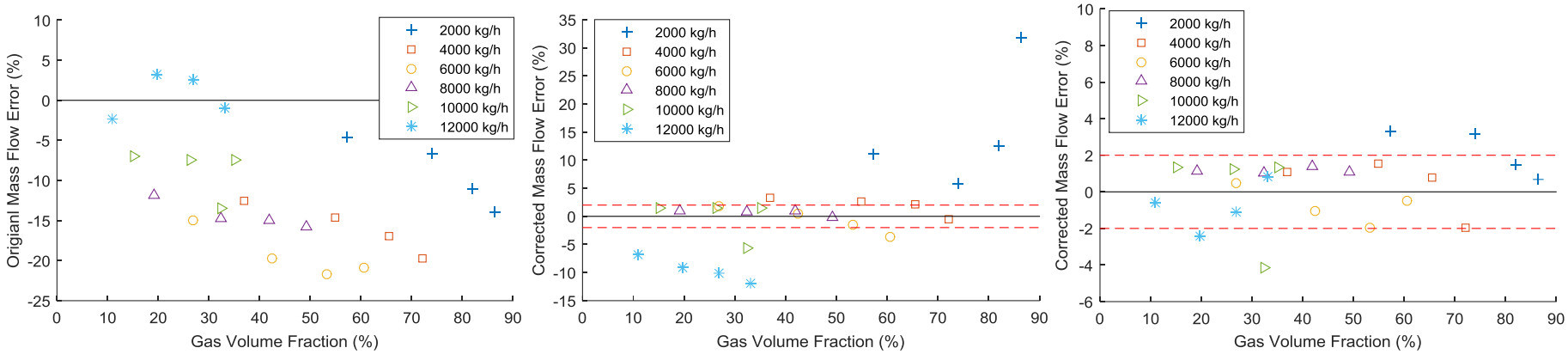
(a) Original errors for K2AWH (b) Corrected errors from SVM for K2AWH (c) Corrected errors from LSSVM for K2AWH

Figure 5.9 Relative errors of the liquid mass flowrate for K2AWH



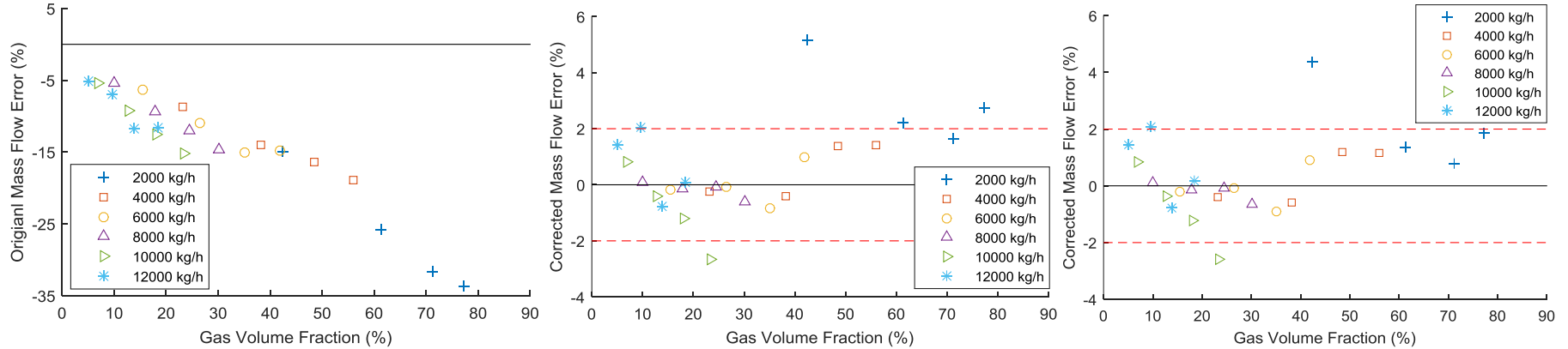
(a) Original errors for T2AWH (b) Corrected errors from SVM for T2AWH (c) Corrected errors from LSSVM for T2AWH

Figure 5.10 Relative errors of the liquid mass flowrate for T2AWH



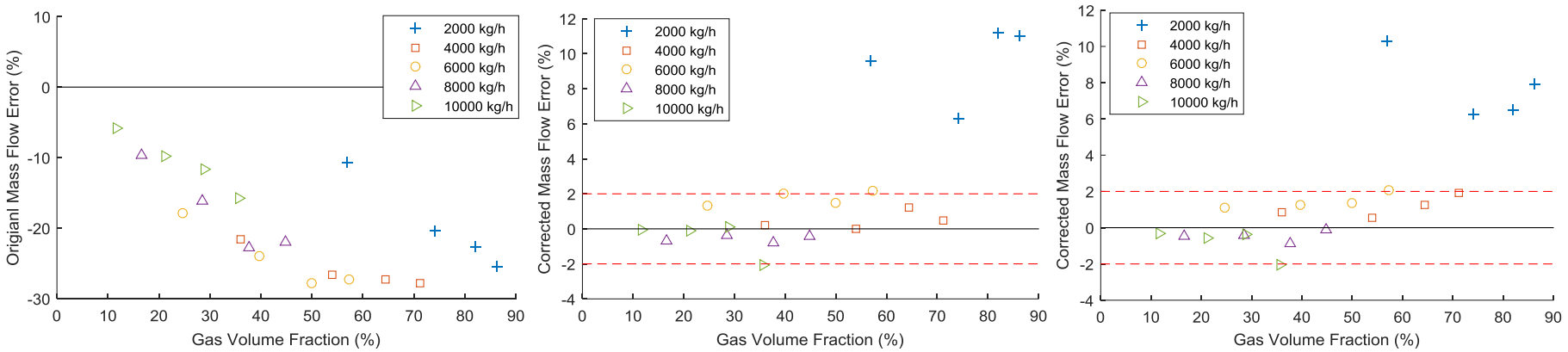
(a) Original errors for T2AWV (b) Corrected errors from SVM for T2AWV (c) Corrected errors from LSSVM for T2AWV

Figure 5.11 Relative errors of the liquid mass flowrate for T2AWV



(a) Original errors for T2AOH (b) Corrected errors from SVM for T2AOH (c) Corrected errors from LSSVM for T2AOH

Figure 5.12 Relative errors of the liquid mass flowrate for T2AOH



(a) Original errors for T2AOV (b) Corrected errors from SVM for T2AOV (c) Corrected errors from LSSVM for T2AOV

Figure 5.13 Relative errors of the liquid mass flowrate for T2AOV

Table 5.2 NRMSE comparison of the SVM and LSSVM models for Model-L (%)

Model-L	**K2AWH**		**T2AWH**		**T2AWV**		**T2AOH**		**T2AOV**	
	Test I	**Test II**	**Test I**	**Test II**	**Test I**	**Test II**	**Test I**	**Test II**	**Test I**	**Test II**
Original	19.82	19.79	10.53	10.93	14.30	3.52	12.13	12.53	18.33	20.70
SVM	1.44	2.33	3.15	3.37	2.64	12.29	1.16	1.73	1.12	9.75
LSSVM	1.47	2.40	3.01	1.58	1.97	1.79	1.14	1.71	1.10	7.87

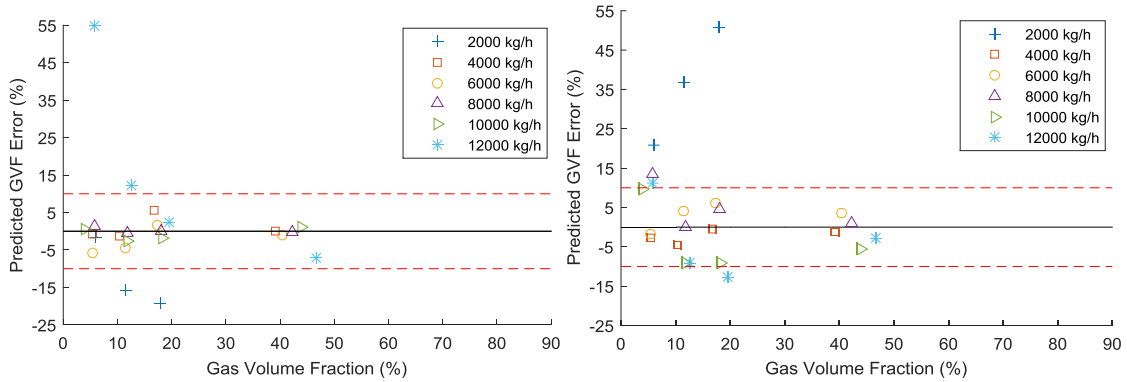
Table 5.3 summarises the training time of the SVM and LSSVM models for liquid mass flowrate correction. It can be seen that training time by SVM is much more than LSSVM. For LSSVM, the main computational cost comes from calculating the Lagrange multipliers α based on equation (3-29) while SVM need more time to solve a dual problem in equation (3-17).

Table 5.3 Training time of the SVM and LSSVM models for Model-L

Model-L	K2AWH (s)	T2AWH (s)	T2AWV (s)	T2AOH (s)	T2AOV (s)
SVM	3.353	22.525	2.351	25.144	4.205
LSSVM	0.668	0.454	0.457	0.534	0.503

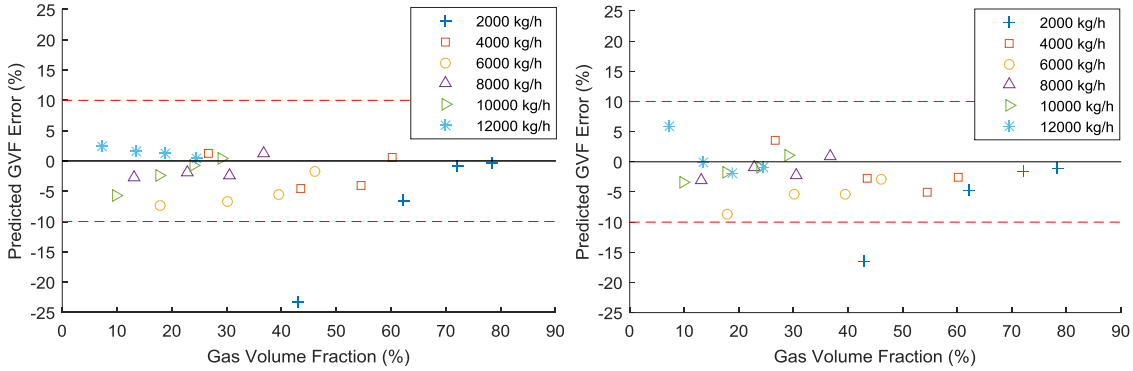
5.4.2 SVM and LSSVM Models for the Prediction of Gas Volume Fraction

Three variables (apparent mass flowrate, observed density and DP) are determined as inputs of the data-driven models through input variable selection. SVM and LSSVM models are established respectively for each case to predict gas volume fraction. The kernel function used in these models is radial basis function. The optimal combination of C and γ is determined by 5-fold cross validation. As shown in Figures 5.14-5.18, the relative errors of the predicted gas volume fraction are mostly within ±10% (the red dashed lines). It is notable that large errors occur at the flowrate of 2000 kg/h, practically in the cases of K2AWH, T2AWH and T2AWV, due to the unstable flow. For the air-oil flow, all the predicted errors with the test data are less than ±10%.



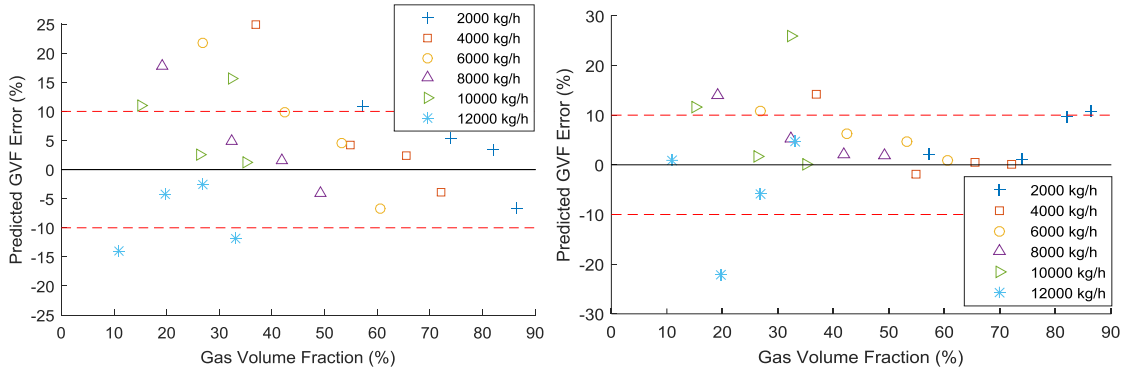
(a) The SVM model for K2AWH (b) The LSSVM model for K2AWH

Figure 5.14 Relative errors of gas volume fraction for K2AWH



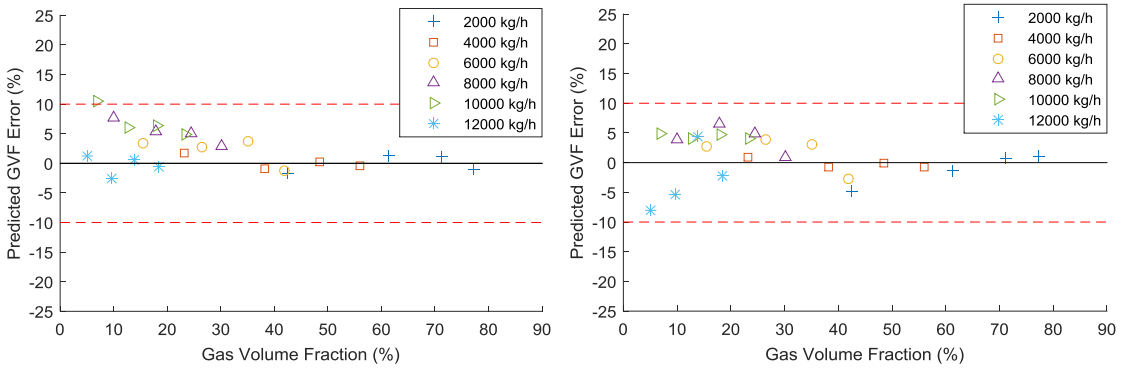
(a) The SVM model for T2AWH (b) The LSSVM model for T2AWH

Figure 5.15 Relative errors of gas volume fraction for T2AWH



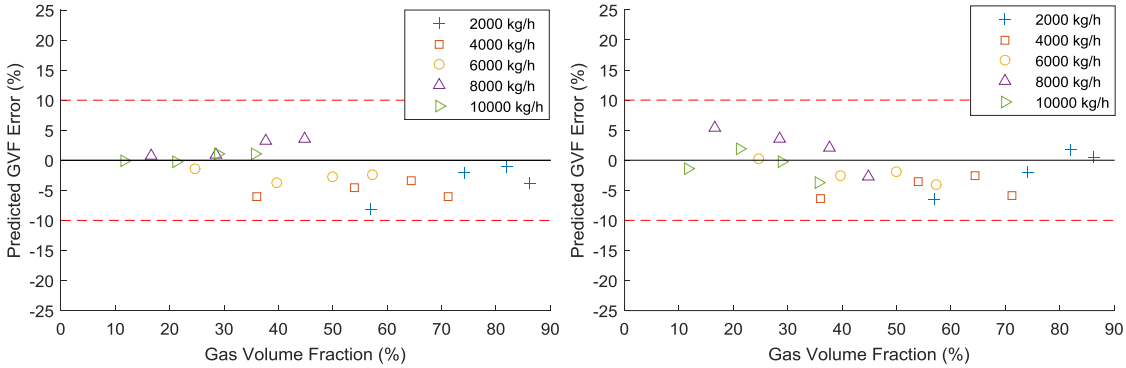
(a) The SVM model for T2AWV (b) The LSSVM model for T2AWV

Figure 5.16 Relative errors of gas volume fraction for T2AWV



(a) The SVM model for T2AOH (b) The LSSVM model for T2AOH

Figure 5.17 Relative errors of gas volume fraction for T2AOH



(a) The SVM model for T2AOV (b) The LSSVM model for T2AOV

Figure 5.18 Relative errors of gas volume fraction for T2AOV

The NRMSE of predicted gas volume fraction from the SVM and LSSVM models are summarized in Table 5.4. In the case of K2AWH, the SVM model outperforms LSSVM with test I and test II. In the other cases, the SVM and LSSVM models present similar generalization ability with test data. With regard to the training time shown in Table 5.5, the SVM models require much more time than LSSVM to complete the training process. Therefore, LSSVM is more efficient than SVM in practical application.

Table 5.4 NRMSE of the SVM and LSSVM models for Model-G (%)

Model-G	**K2AWH**		**T2AWH**		**T2AWV**		**T2AOH**		**T2AOV**	
	Test I	**Test II**	**Test I**	**Test II**	**Test I**	**Test II**	**Test I**	**Test II**	**Test I**	**Test II**
SVM	1.87	13.84	3.73	9.53	8.98	7.79	3.04	1.52	4.31	4.03
LSSVM	5.28	23.33	3.96	6.95	7.04	9.67	2.84	2.50	4.22	2.89

Table 5.5 Training time of the SVM and LSSVM models for Model-G (s)

Model-G	K2AWH	T2AWH	T2AWV	T2AOH	T2AOV
SVM	21.648	20.140	19.715	24.396	20.223
LSSVM	0.486	0.460	0.518	0.461	0.462

5.5 Summary

Experimental tests of 2-inch Coriolis flowmeters with air-water and air-oil flows have been conducted on KROHNE and Tianjin test rigs, respectively. The original errors of Coriolis flowmeters with different two-phase flows were presented. The effects of pressure, temperature, installation orientation and physical properties of the fluid on the performance of Coriolis flowmeters have been quantified and discussed.

SVM and LSSVM models have been established for each case to correct liquid mass flowrate and predict gas volume fraction of two-phase flows. The performance of the models has been assessed in terms of accuracy, generalization ability and computational complexity. The results have demonstrated that the developed SVM and LSSVM models are capable of reducing the mass flow errors to within ±2% and gas volume fraction less than ±10% in most cases. Even though test I and test II are different from the training data, the data-driven models have yielded better results in test I than in test II due to the fact that test I data are within the range of the training process. Large errors are often observed with the unstable slug flow at 2000 kg/h, however, the LSSVM models outperform the SVM models in most cases. As for the training time, the LSSVM is faster than the SVM model to compute the model parameters in the training process. Consequently, the LSSVM models have similar and even better generalization ability when compared to the SVM models. Meanwhile, the LSSVM is more computationally efficient than the SVM models.

Chapter 6

Experimental Tests with Gas-Liquid Two-Phase CO_2 Flow

6.1 Introduction

This chapter assesses the performance of Coriolis flowmeters (KROHNE OPTIMASS 6400 S15) with gas-liquid two-phase CO_2 flow. Experimental work was conducted on a purpose-built two-phase CO_2 flow test rig. The original errors of Coriolis flowmeters in horizontal and vertical installation orientations are presented and interpreted.

SVM and LSSVM models are established for the Coriolis flowmeters in horizontal and vertical installation positions respectively. A performance comparison between the models is discussed. As three typical flow patterns were observed during the experimental tests with gas-liquid two-phase CO_2, the modelling work is more complicated than that in Chapters 4 and 5. Consequently, flow pattern recognition is necessary before developing the individual data-driven models for each flow pattern. Performances of pattern recognition and two-phase CO_2 flow measurement using Coriolis flowmeters and LSSVM models are reported.

6.2 Experimental Conditions

Experimental tests of Coriolis flowmeters (KROHNE OPTIMASS 6400 S15) with gas-liquid two-phase CO_2 flow were conducted at NCEPU (North China Electric Power University) in August 2016. As shown in Figure 6.1, the CO_2 test facility at NCEPU is capable of providing single-phase (liquid or gas) or two-phase (liquid/gas) CO_2 flows in horizontal and vertical pipelines with pressure up to 72 bar. Two independent Coriolis flowmeters were installed before the mixer to provide references for the individual mass

flow rates of the liquid and gas CO_2 phases. The reference Coriolis flowmeters equipped on the facility offer uncertainties of 0.16% for CO_2 liquid flow and 0.3% for CO_2 gas flow [103]. In the downstream, two additional Coriolis flowmeters (Figure 6.2) of the same type were installed in the horizontal and vertical test sections, respectively. These are the meters under test to assess their performance of the developed data-driven models. In view of the effects of gravity and buoyancy on two-phase fluid, both horizontal and vertical installations of the meters are considered. Temperature, pressure and DP transducers were used to record the flow conditions in the pipelines.

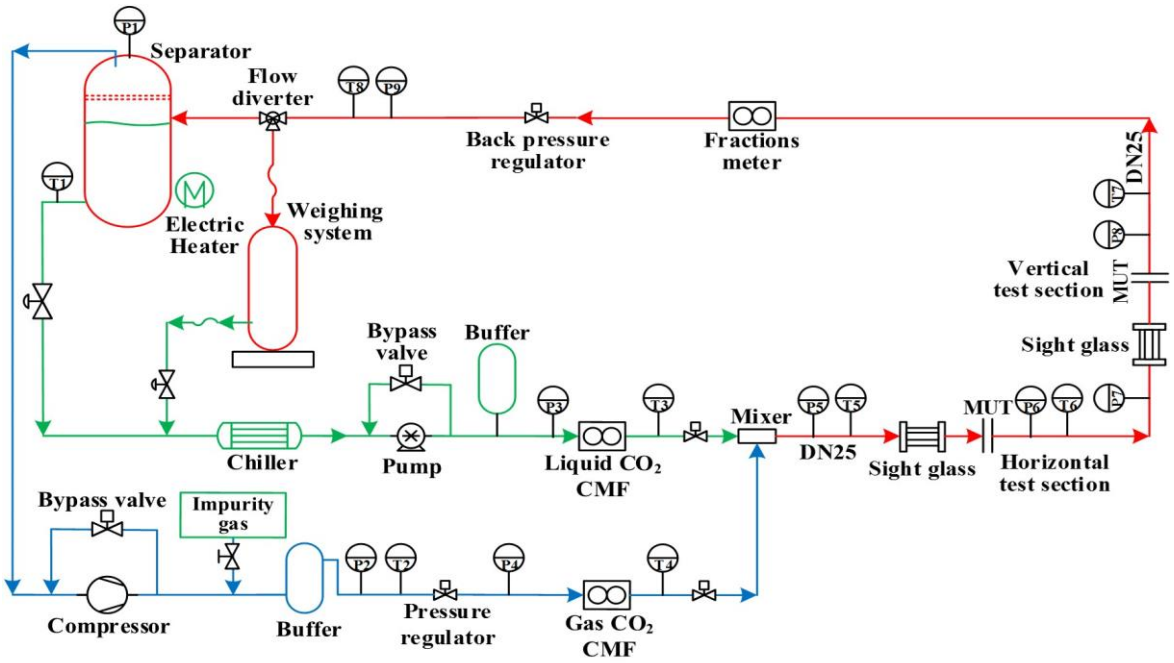


Figure 6.1 Schematic of the gas-liquid two-phase CO_2 flow test rig

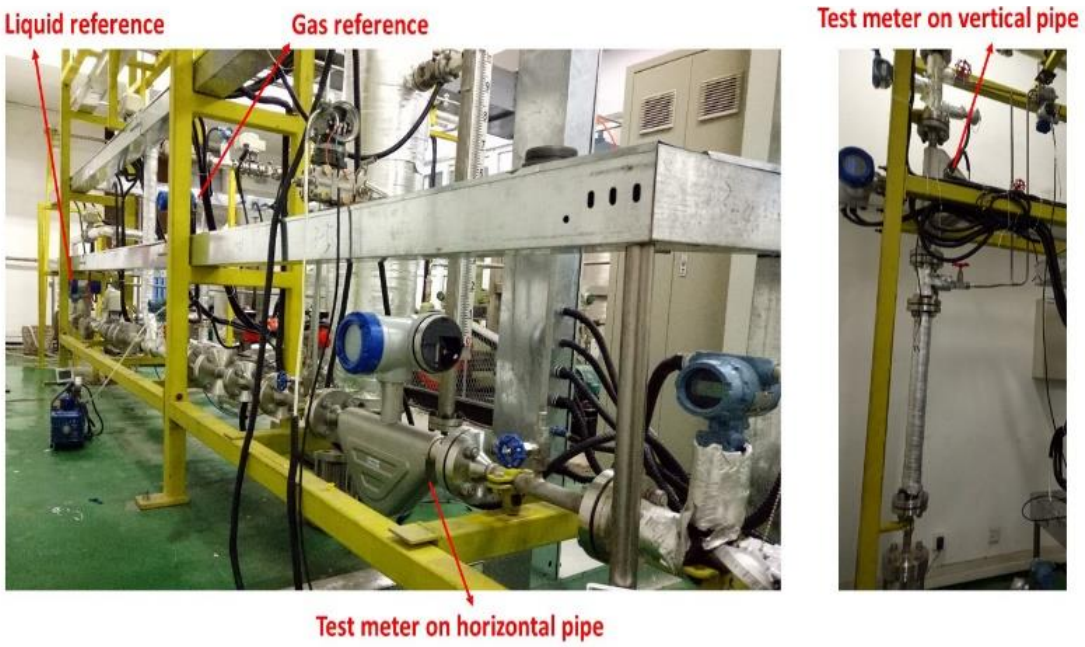


Figure 6.2 Meters under test and reference meters on the test rig

In order to achieve liquid CO_2 at single phase section and stable gas-liquid mixture at test section, the fluid temperature was controlled between 19°C and 21°C through a cooling system and pressure was from 54 bar to 58 bar. The liquid and gas CO_2 mass flowrates were varied from 250 kg/h to 3200 kg/h and from 0 to 330 kg/h, respectively. The test points in terms of liquid CO_2 mass flowrate and gas CO_2 mass flowrate are plotted in Figure 6.3. A total of 232 data sets (circular markers in Figure 6.3, representing liquid flowrates of 250, 400, 700, 800, 1050, 1300, 1800, 2300, 2800 and 3200 kg/h) were collected for the purpose of training the data-driven models whilst 89 data sets (triangular markers in Figure 6.3, presenting liquid flowrates of 300, 550, 900, 1550, 2050, 2550 and 3050 kg/h) for testing the models. Each data set represents the average of all recorded values within an approximate window of 100 seconds. On the horizontal test section three typical flow regimes, including stratified flow, intermittent flow and dispersed flow, were observed. The typical flow patterns on the vertical test section include bubbly flow, intermittent flow and dispersed flow. As shown in Figure 6.3, the test matrix is divided into three sections according to the flow patterns.

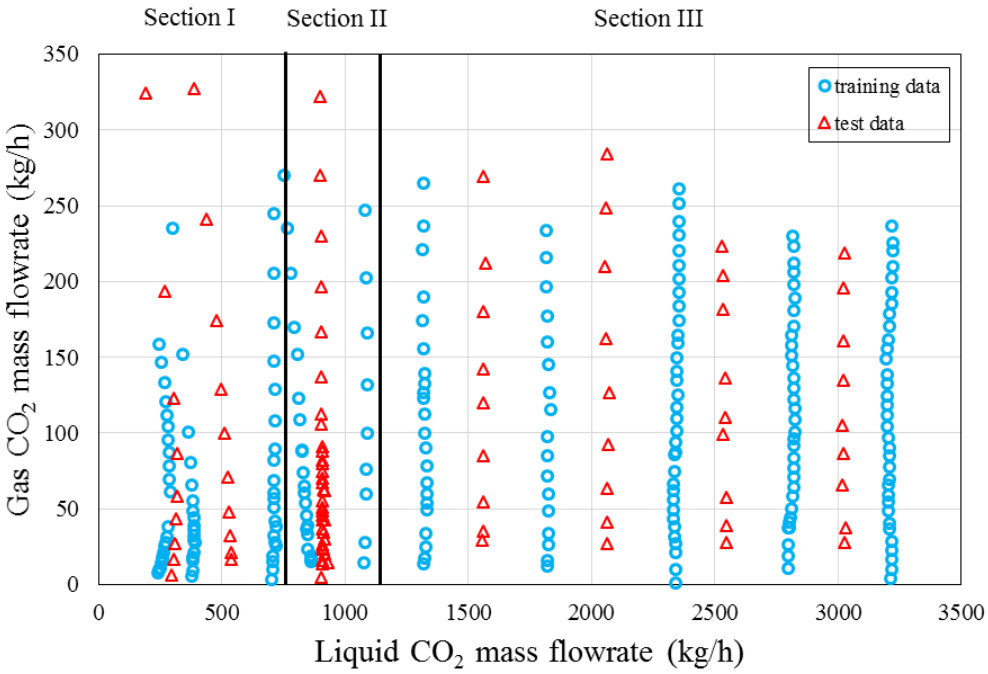
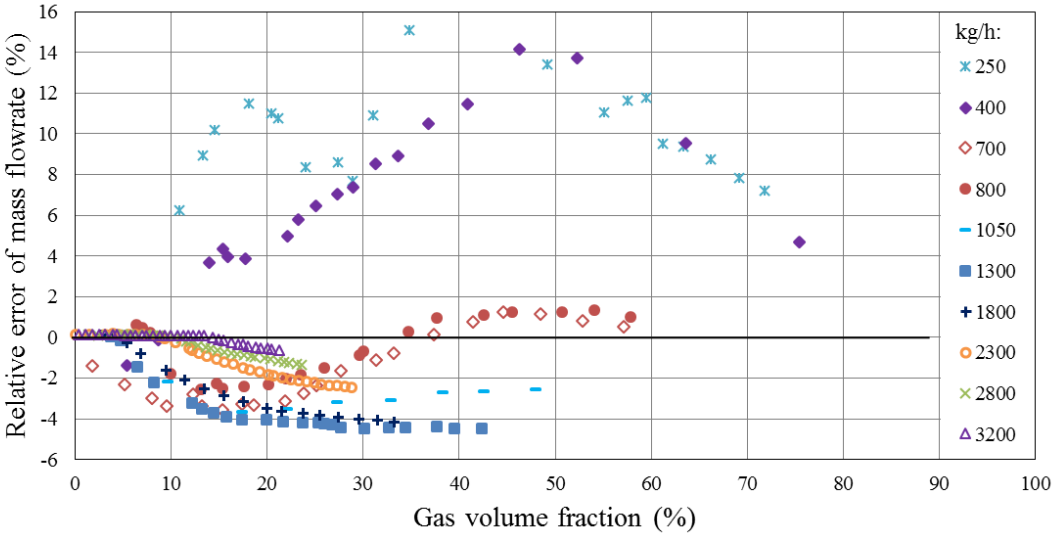


Figure 6.3 Experimental test points of the gas-liquid CO_2 two-phase flow

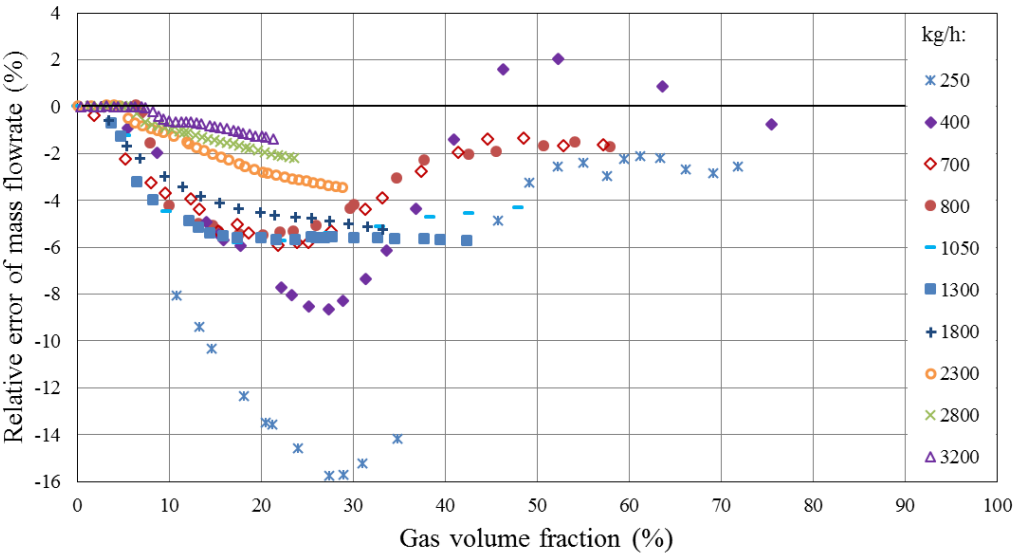
6.3 Analysis of Original Errors

The typical original, uncorrected mass flow errors of the Coriolis flowmeters on horizontal and vertical test sections under two-phase CO_2 flow conditions are plotted in Figure 6.4. As gas mass flowrate cannot be ignored at higher pressure, the error presented in this Chapter is against the total mass flowrate of liquid and gas CO_2. When the liquid mass flowrate is lower than 800 kg/h, gas and liquid phases are completely separated and stratified flow is observed in the horizontal pipe. The high volume of gas in the liquid introduces large positive errors to the Coriolis flowmeter on the horizontal position. The flowmeter on the vertical position yields smaller errors as bubbles go upwards in the liquid. From 800 kg/h to 1000 kg/h, the flow is observed as intermittent flow. As gas CO_2 increases, the two Coriolis flowmeters both generate negative errors for the dispersed flow. Different flow patterns of the two-phase flow present different trends on the error curves due to the chaotic nature of gas phase distributions within the liquid phase. The meter orientation also affects the phase distribution in the Coriolis measuring tubes. In the horizontal installation, the Coriolis measuring tubes are in downward position and bubbles may get trapped on the inlet side at low rates due to the

buoyancy effect. Consequently, the mass flow errors of the Coriolis flowmeters under the test conditions are either positive or negative and have different trends from horizontal and vertical installations.



(a) Coriolis flowmeter in horizontal installation



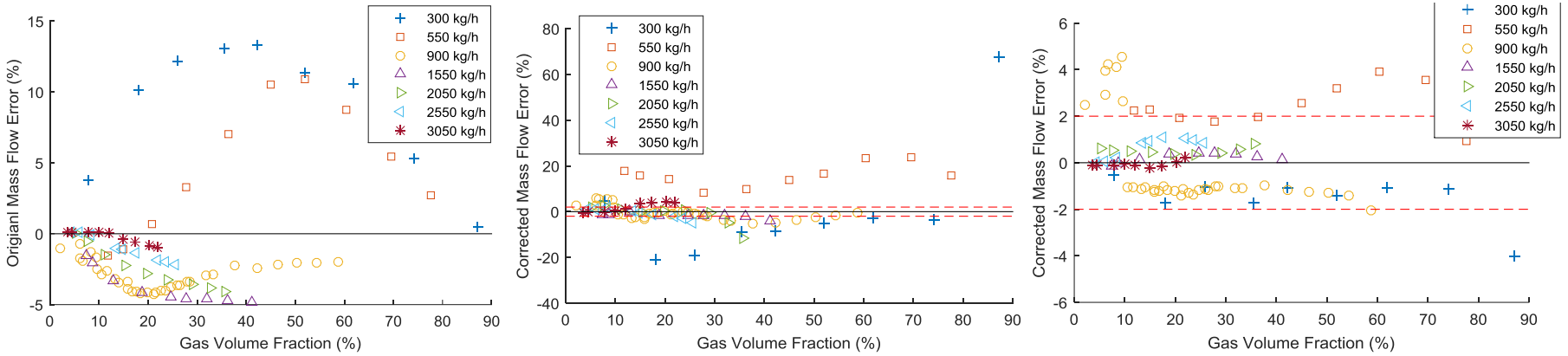
(b) Coriolis flowmeter in vertical installation

Figure 6.4 Original errors of mass flowrate of two-phase CO_2

6.4 Measurement of Gas-Liquid Two-Phase CO_2 Flow

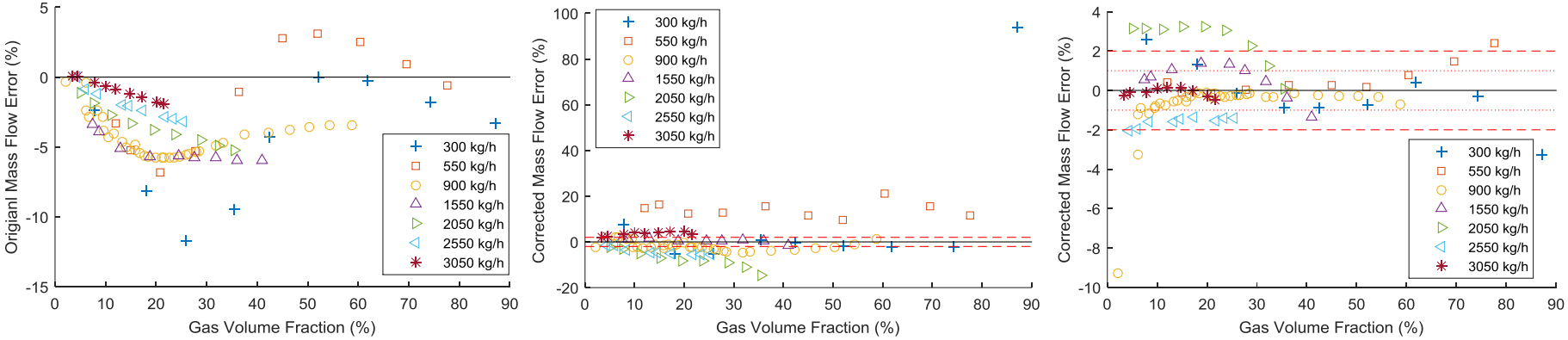
6.4.1 SVM and LSSVM Models

Input variable selection is conducted based on the experimental data from gas-liquid two-phase CO_2 flow to determine the inputs of correction and prediction models. Similar to the results of input variable selection in Chapter 4, four variables (apparent mass flowrate, observed density, damping and DP) are selected out by PMI for the correction models of mass flowrate. A total of three variables including apparent mass flowrate, observed density and DP are taken as inputs for the prediction models of gas volume fraction. SVM and LSSVM models are established to correct the mass flowrate of two-phase CO_2 flow, respectively. The relative errors of CO_2 mass flowrate from horizontal and vertical Coriolis flowmeters are shown in Figures 6.5 and 6.6. It can be seen from the original errors that the errors at the flowrate of 300 kg/h and 550 kg/h are very different from the rest due to the difference in flow patterns. The CO_2 mass flow from SVM has relative large errors at the flowrate of 300 kg/h and 550 kg/h on horizontal position and 550 kg/h and 2050 kg/h on vertical position. Compared with original errors, it is notable that these errors are overcorrected by SVM models. Even though the penalty parameters C and γ have been optimized through five-fold cross validation and a proper kernel function (radial basis function) is utilized, the optimal interface is difficult to satisfy all the test points and results in some unexpected errors.



(a) Original errors for H-installation (b) Corrected errors from the SVM model (c) Corrected errors from the LSSVM model

Figure 6.5 Relative errors of CO_2 mass flowrate for horizontal installation

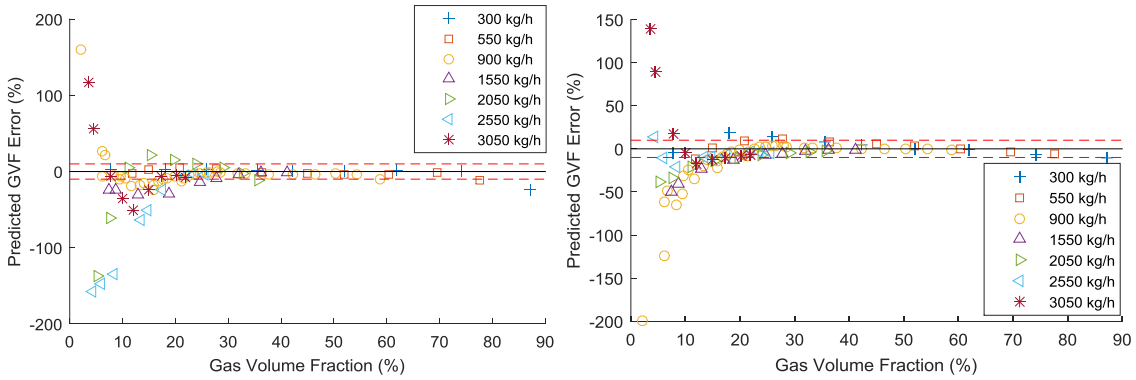


(a) Original errors for V-installation (b) Corrected errors from the SVM model (c) Corrected errors from the LSSVM model

Figure 6.6 Relative errors of CO_2 mass flowrate for vertical installation

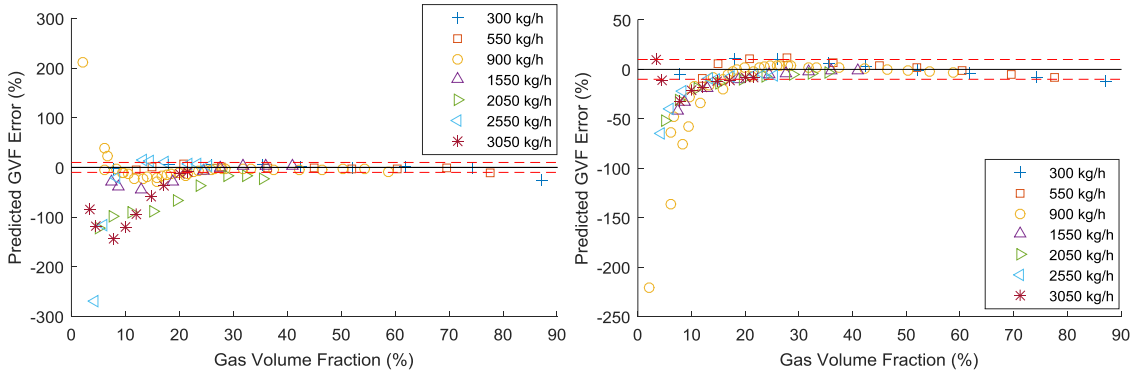
When LSSVM is applied, all the errors on the horizontal position are reduced to ±5% and most of them are less than ±2%. The test points at 900 kg/h on both installations are over-corrected at low gas entrainment, but the results are still much better than those from the SVM models.

As for the prediction of gas volume fraction, as shown in Figures 6.7 and 6.8, some of the relative errors from SVM and LSSVM models are over 100% when the gas volume fraction is lower than 0.1. As gas CO_2 increases, the predicted errors are all within 10%. Therefore, it is evident that the performances of SVM and LSSVM are affected by the flow pattern, especially at lower flowrate and less gas entrainment.



(a) Predicted errors from the SVM model (b) Predicted errors from the LSSVM model

Figure 6.7 Relative errors of gas volume fraction for horizontal installation



(a) Predicted errors from the SVM model (b) Predicted errors from the LSSVM model

Figure 6.8 Relative errors of gas volume fraction for vertical installation

6.4.2 Flow Pattern Recognition

In order to further reduce the corrected errors and avoid overcorrection at lower flowrate, individual models are established for a specific flow pattern. As shown in Figure 6.9, the flow pattern based measurement methodology in this case consists of a classifier model for flow pattern recognition, a correction model and a prediction model for each flow pattern. The flow pattern based models are developed for horizontal and vertical installations, respectively. Since flow pattern is related to the flowrate of liquid and gas phases, the classifier model can be established based on the available data from Coriolis flowmeters and DPs.

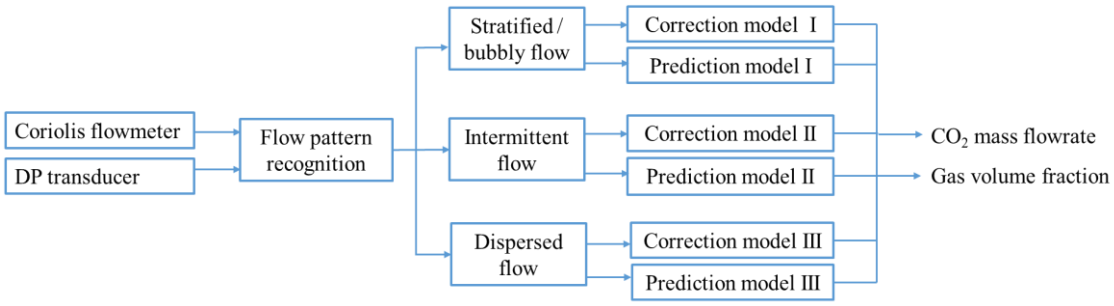


Figure 6.9 Principle of the flow pattern based measurement methodology for horizontal and vertical installations

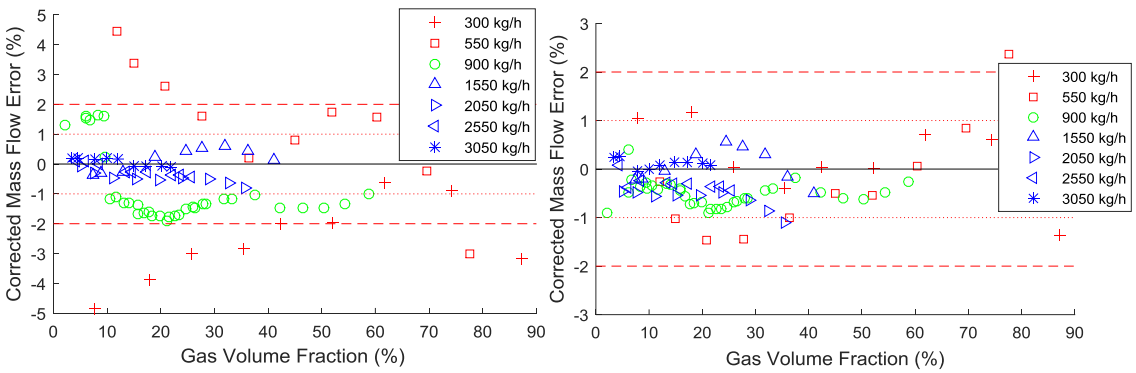
The variable selection results in Chapters 4 and 5 show that the variables, including apparent mass flowrate, observed density, damping and DP, have more significance to estimate the liquid mass flowrate and gas volume fraction which are closely related to the flow pattern. Consequently, these four variables are selected as inputs for the flow pattern recognition model. Due to the fact that LSSVM models have similar performance to SVM models and more computationally efficient, LSSVM is thus applied to train the flow pattern classifier. For the test purpose, there are 19 data for stratified flow on the horizontal section and bubbly flow on the vertical section, respectively, 34 data from intermittent flow and 36 data from dispersed flow on both horizontal and vertical sections, respectively. The results of flow pattern recognition are summarised in Table 6.1. Due to the high performance of LSSVM for classification, all

the test points can be correctly classified into the corresponding flow patterns and result in 100% successful recognition rate.

Table 6.1 Results of flow pattern recognition

Horizontal flow patterns	**Success rate**	**Vertical flow patterns**	**Success rate**
Stratified flow (19)	19	Bubbly flow (19)	19
Intermittent flow (34)	34	Intermittent flow (34)	34
Dispersed flow (36)	36	Dispersed flow (36)	36
Overall success rate	100% (89/89)	Overall success rate	100% (89/89)

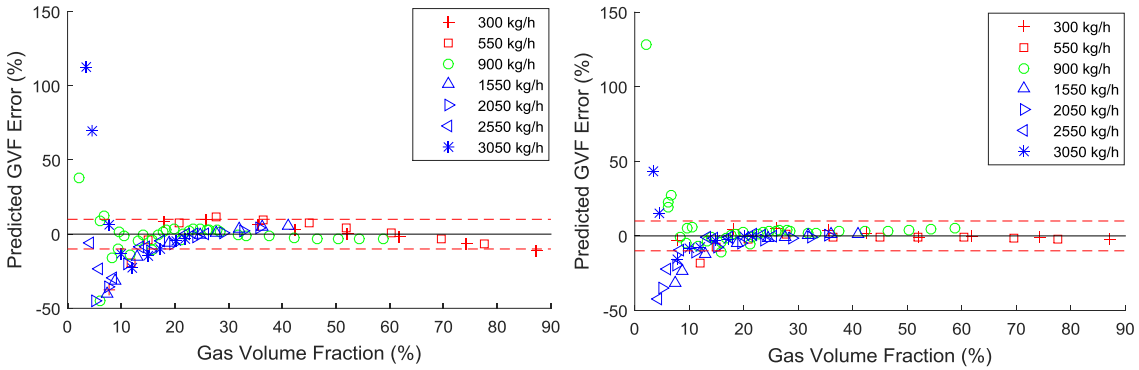
Once the test point is classified into a certain flow pattern, the corresponding correction model and prediction model are determined. The relative errors of the corrected CO_2 mass flowrate on the horizontal and vertical sections are plotted in Figure 6.10. The red, green and blue markers represent the test points from Sections I, II and III, respectively. It can be seen that the test points at Sections II and III are all within ±2% for the horizontal installation and ±1% for the vertical installation. Meanwhile, the errors of mass flowrate at Section I are largely reduced in comparison with the results from the SVM and LSSVM models. The maximum errors at low flowrate are all within ±5%. The correction model on the vertical installation outperforms the horizontal one at the flowrates of 300 kg/h and 550 kg/h and results in relative errors within ±2%.



(a) Corrected errors for H-installation (b) Corrected errors for V-installation

Figure 6.10 Relative errors of CO_2 mass flowrate from the flow pattern based LSSVM

The relative errors of gas volume fraction from the flow pattern based LSSVM model are plotted in Figure 6.11. The large errors occur at the intermittent and dispersed flows with gas volume fraction less than 5% while the predicted errors of other test points are all within ±10%.



(a) Predicted errors for H-installation (b) Predicted errors for V-installation

Figure 6.11 Relative errors of gas volume fraction from the flow pattern based LSSVM

NRMSE results from the SVM, LSSVM and the flow pattern based LSSVM (FP_LSSVM) models for mass flowrate correction and gas volume fraction prediction are summarized in Table 6.2. For the mass flow correction models (CO_2_M_H and CO_2_M_V), the NRMSE values increase after correction by SVM due to overcorrection occurred at some test points. The results from the FP_LSSVM are much smaller than the LSSVM models and they both improve the performance of Coriolis flowmeters for the measurement of two-phase CO_2 mass flow.

As for gas volume fraction prediction, the SVM and LSSVM models have similar performance on both horizontal and vertical installations while FP_LSSVM models outperform the SVM and LSSVM models significantly.

Table 6.2 NRMSE comparison of SVM, LSSVM and FP_LSSVM models

NRMSE (%)	**Original**	**SVM**	**LSSVM**	**FP_LSSVM**
CO_2_M_H	2.82	4.97	1.05	0.80
CO_2_M_V	3.69	6.93	1.77	0.51
CO_2_G_H	-	9.65	10.24	8.11
CO_2_G_V	-	12.81	10.49	4.32

6.5 Summary

The performance of Coriolis flowmeters (KROHNE OPTIMASS 6400 S15) has been assessed with gas-liquid two-phase CO_2 flow under horizontal and vertical orientations. The liquid flowrate in this experimental test is as low as 250 kg/h and hence stratified flow on horizontal section, which is different and more complex than the tests reported in Chapters 4 and 5. SVM and LSSVM models have been established to estimate the mass flowrate and gas volume fraction of the two-phase flow. The performance of SVM and LSSVM models were evaluated with test data for different flow patterns. Test results have demonstrated that some points for flowrate measurement were overcorrected by the SVM models especially at the lower flowrate (300 kg/h and 550 kg/h) and large predicted error for gas volume fraction at less gas entrainment. In order to reduce the measurement errors and avoid overcorrection, flow pattern recognition was incorporated in the measurement strategy. The LSSVM based flow pattern classifier is capable of identifying the flow pattern with a success rate of 100%. After correction with the flow pattern based LSSVM model, the mass flowrate is mostly with ±2% on the horizontal section and ±1.5% on the vertical section. The predicted errors of gas volume fraction is largely reduced, compared with the SVM and LSSVM models.

Chapter 7

Conclusions and Recommendations for Future Work

7.1 Introduction

The research work presented in this thesis is concerned with the measurement of multiphase flow using Coriolis flowmeters incorporating soft computing techniques. Experimental tests of different sized Coriolis flowmeters (KROHNE OPTIMASS 6400 S25, S50 and S15) with air-water, air-oil and gas-liquid two-phase CO_2 flows have been conducted. The performance of the Coriolis flowmeters under different two-phase flow conditions has been presented and analysed. Data-driven models based on ANN, GP, SVM and LSSVM were established with the experimental data to estimate mass flowrate and gas volume fraction. Performance of these models in terms of accuracy, robustness, generalization ability and computational efficiency was evaluated with test data. Considering the effect of flow pattern on the model performance, individual LSSVM models for the estimation of mass flowrate and gas volume fraction of gas-liquid two-phase CO_2 flow have been developed for each flow pattern.

This chapters presents the conclusions that have been drawn from the research programme conducted and makes recommendations for further work in the field.

7.2 Conclusions

7.2.1 Input Variable Selection and Model Selection

Input variable selection and model selection have been taken into consideration based on the experimental data from 1-inch Coriolis flowmeters (KROHNE OPTIMASS 6400

S25) with air-water flow. The performance of Coriolis flowmeters under air-water two-phase flow condition has been reported in Chapter 4.

Three input variable selection approaches including PMI, GA-ANN and IIS have been applied to determine the most suitable variable subset for the correction models of liquid mass flowrate and prediction models of gas volume fraction, respectively. The validity of the selected variables has been assessed on SVM-based models through a comparison of the NRMSE and sensitivity analysis. It can be concluded from the results that the PMI and IIS algorithms have outperformed the GA-ANN algorithm in terms of time efficiency and selection validity. With regard to the experimental data obtained, the most important variables for the correction of liquid mass flowrate include observed density, apparent mass flowrate, DP and damping while those for the prediction of gas volume fraction include observed density, apparent mass flowrate and DP.

Experimental and analytical investigations have been carried out to assess the performance of Coriolis flowmeters together with BP-ANN, SVM and GP models for gas-liquid two-phase flow measurement. Results presented have suggested that the SVM models are superior to the ANN and GP models in terms of robustness and accuracy. Once the SVM models are applied to analyse the test data, the NRMSE of the liquid mass flowrate is reduced from 8% to 0.58% for both horizontal and vertical installation orientations while the NRMSE of the predicted gas volume fraction is about 3.20%.

7.2.2 Tests with Air-Water and Air-Oil Flows

Experimental tests of 2-inch Coriolis flowmeters (KROHNE OPTIMASS 6400 S50) have been conducted with air-water and air-oil two-phase flows. Through the analysis of the original errors of liquid mass flowrate from the Coriolis flowmeters under test, the effects of pressure, temperature, installation orientation and physical properties of the fluid on the performance of Coriolis flowmeters have been discussed in Chapter 5.

SVM and LSSVM models have been developed for each case to correct the liquid mass flowrate and predict the gas volume fraction of two-phase flows. The performance of the models has been assessed in terms of accuracy, generalization ability and computational complexity. Experimental results have demonstrated that SVM and LSSVM models are able to reduce the relative errors of mass flowrate to within ±2% and predict the gas volume fraction with relative errors less than ±10% in most cases. LSSVM models have similar or even better generalization ability compared with SVM models. Meanwhile, LSSVM is more computational efficiency than SVM. Once the LSSVM models are applied to process the test data from the five experimental cases, the NRMSEs of the liquid mass flowrate are reduced from more than 10% to less than 2% while the NRMSEs of the predicted gas volume fraction are less than 7%.

7.2.3 Tests with Gas-Liquid Two-Phase CO_2 Flow

Experimental tests of Coriolis flowmeters (KROHNE OPTIMASS 6400 S15) have been conducted with gas-liquid two-phase CO_2 flow. The original errors of Coriolis flowmeters on horizontal and vertical installations have been presented and interpreted in Chapter 6.

SVM and LSSVM models have been established to estimate the mass flowrate and gas volume fraction of gas-liquid two-phase CO_2 flow. The performance of SVM and LSSVM models was evaluated with test data for different flow patterns. Test results have demonstrated that some points were overcorrected by SVM models especially at lower flowrates (300 kg/h and 550 kg/h) and hence large predicted error of gas volume fraction at less gas entrainment (<10%). In order to reduce the measurement errors and avoid overcorrection, the flow pattern based measurement methodology with a flow pattern classifier was proposed. The LSSVM based flow pattern classifier is capable of achieving a successful recognition rate of 100%. After correction with the flow pattern based LSSVM model, the relative errors of mass flowrate are mostly with ±2% on the horizontal section and ±1.5% on the vertical section. Meanwhile, the NRMSE of the mass flowrate is reduced from 2.82% to 0.80% on the horizontal position and from

3.69% to 0.51% on the vertical positions while the NRMSEs of the predicted gas volume fraction are less than 8%.

The work reported in this thesis offers very useful measurement strategies to develop data-driven models of Coriolis flowmeters for applications to two-phase or multiphase flow measurement. The proposed measurement method has been validated through a series of experimental tests. The work reported in this thesis has demonstrated the potential of the Coriolis flowmeters incorporating soft computing techniques in multiphase flow measurement. It is envisaged that the implementation of such technique will lead to an effective method for multiphase flow measurement in industrial processes.

7.3 Recommendations for Future Research

The work presented in this thesis has demonstrated the usefulness and potential of the outcomes of the research programme. However, the soft computing techniques for multiphase flow measurement are still in their development stage. There are a number of areas that require further research and improvements in the near future and have been identified as follows:

1) Experimental work in this study was conducted on laboratory test rigs with relatively stable temperature and pressure. However, the temperature and pressure in industrial processes are quite different in practical applications. Moreover, the installation orientation of Coriolis flowmeters may be not limited to horizontal or vertical position in an industrial site. Consequently, different fluid temperature, pressure and installation orientation of Coriolis flowmeters will be taken into consideration in the modelling work. The modelling method will be extended from air-water, air-oil and gas-liquid CO_2 two-phase flows to oil-gas-water multiphase flow. Further field trials in the oil and gas industry and bunkering centres are also required to evaluate the measurement system under a wider range of conditions.

2) Gas-liquid two-phase CO_2 flow measurement was carried out in this study. In real CCS chain, there could exist more states of CO_2 in the mixture due to the diverse climate and variant temperature and pressure during the long transport pipelines. In addition, some constituents such as N_2 and CH_4 are very likely mixed in the captured stream, which makes pure CO_2 phase diagram and state equations highly unreliable and results in two-phase or multiphase flow conditions. The present study will be further extended by considering the solid state of CO_2 or impurities of N_2 and CH_4 involved in the multiphase CO_2 flow.

3) With the rapid development of soft-computing techniques, some more advanced, improved and hybrid algorithms are being developed to achieve high-performance data-driven models. Further investigations into the techniques used for data-driven modelling should be conducted.

4) The data-driven models based on soft computing techniques provide an effective and cost-efficient approach for Coriolis flowmeters in the application of two-phase or multiphase flow measurement. Mass flowrate and gas volume fraction can be estimated through data-driven models without any modifications on the Coriolis flowmeters. However, this approach is a kind of ‘black box’ and cannot interpret the mass flow errors of Coriolis flowmeters due to gas entrainment. Hence, further research on the measurement principle and analysis on low-level signals is a necessary research direction to explore the essence of air entrainment on Coriolis flowmeters.

References

[1] R. Baker, Flow measurement handbook-industrial designs, operating principles, performance, and applications (2nd Edition), Cambridge University Press, 2016.

[2] R. Thorn, G. Johansen and E. Hammer, “Recent developments in three-phase flow measurement,” Measurement Science and Technology, vol. 8, pp. 691-701, 1997.

[3] R. Thorn, G. Johansen and B. Hjertaker, “Three-phase flow measurement in the peroleum industry,” Measurement Science and Technology, vol. 24, pp. 1-17, 2013.

[4] L. Hunter and G. Leslie, “National physical laboratory (NPL): A study of measurement issues for carbon capture and storage (CCS),” TUV NEL Ltd, Glasgow, U.K., Tech. Rep. 2009/54, 2009.

[5] T. Wang and R. Baker, “Coriolis flowmeters: a review of developments over the past 20 years, and an assessment of the state of the art and likely future directions,” Flow Measurement and Instrumentation, vol. 40, pp. 99-123, 2014.

[6] G. Falcone, G. Hewitt, C. Alimonti and B. Harrison, “Multiphase flow metering: current trends and future developments,” Journal of Petroleum Technology, vol. 54, no. 4, pp. 77-84, 2002.

[7] [On-line] An introduction to multiphase flow measurement, http://www.tuvnel.com/_x90lbm/An_Introduction_to_Multiphase_Flow_Measurement.pdf , Accessed 08/03/2017.

[8] S. Neumann et al., Handbook of multiphase flow metering, Norwegian Society for Oil and Gas Measurement, 2005.

[9] [On-line] Accuflow 3-Phase Separation Metering System, http://www.accuflow.com/wp-content/uploads/2012/05/Brochure3P.pdf, Accessed 09/03/2017.

[10] [On-line] Accuflow JR Series Multiphase Metering system, http://www.accuflow.com/wp-content/uploads/2012/05/Brochure_JR.pdf, Accessed 09/03/2017.

[11] [On-line] MB CCM Multiphase Meters, http://www.phasedynamics.com/index.php/products/ccmmultiphasemeter, Accessed 09/03/2017.

[12] [On-line] Multiphase flow measurement, http://www.weatherford.com/en/products-services/production/flow-measurement/multiphase-flow-measurement, Accessed 08/03/2017.

[13] D. Wang and L Hu, "Gas-liquid two-phase flow measurement using ESM," Experimental Thermal and Fluid Science, vol. 26, pp. 827-832, 2002.

[14] D. Wang, F. Liang, Z. Peng, Y. Wang and Z. Hu, "Gas-liquid two-phase flow measurements by full steam batch sampling," International Journal of Multiphase Flow, vol. 40, pp. 113-125, 2012.

[15] F. Liang, D. Wang, J. Chen and G. Yang, "Gas-liquid two-phase flow sampling measurement using a swirl sampler," Flow Measurement and Instrumentation, vol. 33, pp. 145-152, 2013.

[16] D. Wang, J. Tan, X. Zhang and Y. Lin, “High pressure steam-water two-phase flow measurements by flow division and separation method,” Experimental Thermal and Fluid Science, vol. 44, pp. 468-474, 2013.

[17] F. Liang, Y. Sun, Z. Fang and S. Sun, “Application of multi-slot sampling method for gas-liquid two-phase flow rate measurement,” Experimental Thermal and Fluid Science, vol. 79, pp. 213-221, 2016.

[18] [On-line] AGAR MPFM-50, http://www.agarcorp.com/technology/mpfm50.html, Accessed 08/03/2017.

[19] [On-line] Multi Phase Meters AS MPM, http://www.fmctechnologies.com/en/Multiphase-meters/Benefits.aspx, Accessed 08/03/2017.

[20] [On-line] Roxar Flow Measurement MPFM 2600, http://www.emerson.com/catalog/en-us/roxar-2600-multiphase , Accessed 08/03/2017.

[21] [On-line] Roxar Flow Measurement Subsea MPFM, http://www.emerson.com/catalog/en-us/roxar-subsea-multiphase, Accessed 08/03/2017.

[22] J. Kim, Y. Ahn and M. Kim, “Measurement of voild fraction and bubble speed of slug flow with three-ring conductance probes,” Flow Measurement and Instrumentation, vol. 20, pp. 103-109, 2009.

[23] C. Tan, W. Dai, H. Wu and F. Dong, “A conductance ring coupled cone meter for oil-water two-phase flow measurement,” IEEE Sensors Journal, vol. 14, no. 4, 2014.

[24] L. Xu, W. Xu, Z Cao, X Liu and J Hu, “Multiple parameters’ estimation in horizontal well logging using a conductance-probe array,” Flow Measurement and Instrumentation, vol. 40, pp. 192-198, 2014.

[25] L. Zhai, P. Bian, Y. Han, Z. Gao and N. Jin, “The measurement of gas-liquid two-phase flows in a small diameter pipe using a dual-sensor multi-electrode conductance probe,” Measurement Science and Technology, vol. 27, pp. 1-19, 2016.

[26] H. Shaban and S. Tavoularis, “The wire-mesh sensor as a two-phase flow meter,” Measurement Science and Technology, vol. 26, pp. 1-16, 2015.

[27] W. Yang, “Design of electrical capacitance tomography sensors,” Measurement Science and Technology, vol. 21, pp. 1-13, 2010.

[28] M. Mao, J. Ye, H. Wang and W. Yang, “Evaluation of extraction strategies for 3D ECT in gas-solids flow measurement,” IEEE Sensors Journal, vol. 16, no. 23, pp. 8523-8530, 2016.

[29] J. Ye, H. Wang, Y. Li and W. Yang, “Coupling of fluid field and electrostatic field for electrical capacitance tomography,” IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 12, pp. 3334-3353, 2015.

[30] F. Dong, Y. Xu, L. Xu, L. Hua and X. Qiao, “Application of dual-plane ERT system and cross-correlation technique to measure gas-liquid flows in vertical

upward pipe," Flow Measurement and Instrumentation, vol. 16, pp. 191-197, 2005.

[31] F. Dong, C. Xu, Z. Zhang and S. Ren, "Design of parallel electrical resistance tomography system for measuring multiphase flow," Chinese Journal of Chemical Engineering, vol. 20, no. 2, pp. 368-379, 2012.

[32] B. Wang, Y. Hu, H. Ji, Z. Huang and H. Li, "A novel electrical resistance tomography system based on C4D technique," IEEE Transactions on Instrumentation and Measurement, vol. 62, no. 5, pp. 1017-1024, 2013.

[33] Z. Liu, G. Yang, N. He and X. Tan, "Landweber iterative algorithm based on regularization in electromagnetic tomography for multiphase flow measurement," Flow Measurement and Instrumentation, vol. 27, pp. 53-58, 2012.

[34] T. Leeungculsatien and G. Lucas, "Measurement of velocity profiles in multiphase flow using a multi-electrode electromagnetic flow meter," Flow Measurement and Instrumentation, vol. 31, pp. 86-95, 2013.

[35] Y. Faraj, M. Wang, J. Jia, Q. Wang, C. Xie, G. Oddie, K. Primrose and C. Qiu, "Measurement of vertical oil-in-water two-phase flow using dual-modality ERT-EMT system," Flow Measurement and Instrumentation, vol. 46, pp. 255-261, 2015.

[36] J. Jia, M. Wang and Y. Faraj, "Evaluation of EIT systems and algorithms for handling full void fraction range in two-phase flow measurement," Measurement Science and Technology, vol. 26, pp. 1-7, 2015.

[37] M. Mallach, P. Gebhardt and T. Musch, “2D microwave tomography system for imaging of multiphase flows in metal pipes,” Flow Measurement and Instrumentation, vol. 53, pp. 80-88, 2017.

[38] C. Tan, Y. Yuan, X. Dong and F Dong, “Oil-water two-phase flow measurement with combined ultrasonic transducer and electrical sensors,” Measurement Science and Technology, vol. 27, pp. 1-11, 2016.

[39] V. Drobkov, V. Melniko and S. Labutin, “Ultrasonic flowmeter and velocity meter for the components of a multiphase stream,” Measurement Techniques, vol. 45, no. 12, pp. 1254-1255, 2002.

[40] L. Xing, Y. Geng, C. Hua, H. Zhu, A. Rieder, W. Drahm and M. Bezdek, “A combination method for metering gas-liquid two-phase flows of low liquid loading applying ultrasonic and Coriolis flowmeters,” Flow Measurement and Instrumentation, vol. 37, pp. 135-143, 2014.

[41] A. Bilgic, J. Kunze, V. Stegemann, J. Hogendoorn and M. Zoeteweij, “Multiphase flow metering with nuclear magnetic resonance spectroscopy,” TM-Technisches Messen, vol. 82, no. 11, pp. 539-548, 2015.

[42] W. Zhu, J. Yang and J. Lao, “A novel signal identifying method of oil-gas-water phase fractions in NMR Measurement,” in the Proceedings of 2015 International Conference on Information and Intelligent Control, Sanya, China, 2015, pp. 551-558.

[43] N. Wang, C. Tan and F. Dong, "Water holdup measurement of oil-water two-phase flow based on KPLS regression," in the Proceedings of 2015 Chinese Automation Congress, Wuhan, China, 27-29 Nov. 2015, pp. 1896-1900.

[44] J. Zhang, H. Hu, J. Dong and Y. Yan, "Concentration measurement of biomass/coal/air three-phase flow by integrating electrostatic and capacitive sensors," Flow Measurement and Instrumentation, vol. 24, pp. 43-49, 2012.

[45] T. Vendruscolo, R. Fischer, C. Martelli, P. Rodrigues, R. Morales and M. Silva, "Multiphase flow parameter estimation based on laser scattering," Measurement Science and Technology, vol. 26, pp. 1-8, 2015.

[46] T. Green, M. Reese and M. Henry, "Two-phase CO_2 measurement and control in the Yates oil field," Measurement and Control, vol. 41, no. 7, pp. 205-207, 2008.

[47] K. Adefila, Y. Yan, L. Sun and T. Wang, "Calibration of an averaging pitot tube for gaseous CO_2 flow metering," IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 5, pp. 1240-1249, May 2015.

[48] C. Lin, A. Bhattacharji, G. Spicer and M. Maroto-Valer, "Coriolis Metering Technology for CO_2 Transportation for Carbon Capture and Storage," Energy Procedia, vol. 63, pp. 2723-2726, 2014.

[49] L.A. Zadeh, "Soft computing and fuzzy logic," IEEE Software, vol. 11, no. 5, pp. 48-56, 1994.

[50] S. Das, A. Kumar, B. Das and A. Burnwal, "On soft computing techniques in various areas," in Proceedings of National Conference on Advancement of Computing in Engineering Research, 2013, pp. 59-68.

[51] M. Figueiredo, J. Goncalves, A. Nakashima, A. Fileti and R. Carvalho, “The use of an ultrasonic technique and neural networks for identification of the flow pattern and measurement of the gas volume fraction in multiphase flows,” Experimental Thermal and Fluid Science, vol. 70, pp. 29-50, 2016.

[52] L. Xu, W. Zhou, X. Li and S. Tang, “Wet gas metering using a revised Venturi meter and soft-computing approximation techniques,” IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 3, pp. 947-956, 2011.

[53] H. Shaban and S. Tavoularis, “Measurement of gas and liquid flow rates in two-phase pipe flows by the application of machine learning techniques to differential pressure signals,” International Journal of Multiphase Flow, vol. 67, pp. 106-117, 2014.

[54] S. Fan and T. Yan, “Two-phase air-water slug flow measurement in horizontal pipe using conductance probes and neural network,” IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 2, pp. 456-466, 2014.

[55] H. Li, H. Ji, Z. Huang, B. Wang, H. Li and G. Wu, “A new void fraction measurement method for gas-liquid two-phase flow in small channels,” Sensors, vol. 16, no. 2, pp. 1-13, 2016.

[56] G. Zheng, N. Jin, X. Jia, P. Lv and X. Liu, “Gas-liquid two phase flow measurement method based on combination instrument of turbine flowmeter and conductance sensor,” International Journal of Multiphase Flow, vol. 34, pp. 1031-1047, 2008.

[57] M. Meribout, N. Al-Rawahi, A. Al-Naamany, A. Al-Bimani, K. Al-Busaidi and A. Meribout, "Integration of impedance measurements with acoustic measurements for accurate two phase flow metering in case of high water-cut," Flow Measurement and Instrumentation, vol. 21, pp. 8-19, 2010.

[58] M. Meribout, N. Al-Rawahi, A. Al-Naamany, A. Al-Bimani, K. Al-Busaidi and A. Meribout, "A multisensor intelligent device for real-time multiphase flow metering in oil fields," IEEE Transactions on Instrumentation and Measurement, vol. 59, no. 6, pp. 1507-1519, 2010.

[59] X. Wang, H. Hu and A. Zhang, "Concentration measurement of three-phase flow based on multi-sensor data fusion using adaptive fuzzy inference system," Flow Measurement and Instrumentation, vol. 39, pp. 1-8, 2014.

[60] X. Wang, H. Hu and X. Liu, "Multiphase data fusion techniques with ELM for pulverized-fuel flow concentration measurement in cofired power plant," IEEE Transactions on Instrumentation and Measurement, vol. 64, no. 10, pp. 2769-2780, 2015.

[61] D. Gysling and T. Banach, "Accurate liquid phase density measurement of aerated liquids using speed of sound augmented Coriolis meters," ISA EXPO, Houston, USA, 5-7 October 2004.

[62] A. Rieder, W. Drahm and H. Zhu, "Coriolis mass flowmeter: on measurement errors in two-phase conditions," in the Proceedings of the 13th International Flow Measurement Conference, 2016.

[63] J. Hemp and J. Kutin, “Theory of errors in Coriolis flowmeter readings due to compressibility of the fluid being metered,” Flow Measurement and Instrumentation, vol. 17, pp. 359-369, 2006.

[64] N. Basse, “A review of the theory of Coriolis flowmeter measurement errors due to entrained particles,” Flow Measurement and Instrumentation, vol. 37, pp. 107-118, 2014.

[65] R. Liu, M. Fuent, M. Henry and M. Duta, “A neural network to correct mass flow errors caused by two-phase flow in a digital Coriolis mass flowmeter,” Flow Measurement and Instrumentation, vol. 12, no. 1, pp. 53–63, 2001.

[66] W. Mattar, M. Henry, M. Duta and M. Tombs, Multiphase Coriolis flowmeter, Patent US20060161366A1, 2006.

[67] M. Henry, M. Tombs, M. Duta, F. Zhou, R. Mercado, F. Kenyery, J. Shen, M. Morles, C. Garcia and R. Langansan, “Two-phase flow metering of heavy oil using a Coriolis mass flow meter: A case study,” Flow Measurement and Instrumentation, vol. 17, pp. 399-413, 2006.

[68] M. Henry, M. Tombs, M. Zamora and F. Zhou, “Coriolis mass flow metering for three-phase flow: a case study,” Flow Measurement and Instrumentation, vol. 30, pp. 112-122, 2013.

[69] B. Safarinejadian, M. Tajeddini, and L. Mahmoodi, “A new fuzzy based method for error correction of Coriolis mass flow meter in presence of two-phase fluid,” in Proceedings of International Conference on Artificial Intelligence and Image Processing, Dubai, UAE, 2012, pp. 192-196.

[70] V. Lari and F. Shabaninia, “Error correction of a Coriolis mass flow meter in two-phase flow measurement using neuro-fuzzy,” in Proceedings of 16th CSI International Symposium on Artificial Intelligence and Signal Processing, Shiraz, Iran, 2012, pp. 611–616.

[71] Q. Hou, K. Xu, M. Fang, Y. Shi, B. Tao and R. Jiang, “Gas-liquid two-phase flow correction method for digital CMF,” IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 10, pp. 2396-2404, 2014.

[72] L. Ma, H. Zhang, H. Zhou, and Q. He, “Mass flow measurement of oil water two-phase flow based on Coriolis flow meter and SVM,” Journal of Chemical Engineering of Chinese Universities, vol. 21, no. 2, pp. 201–205, 2007.

[73] R. May, G. Dandy and H. Maier, “Review of input variable selection methods for artificial neural networks,” Chapter 2 in book of Artificial Neural Networks - Methodological Advances and Biomedical Applications, InTech, ISBN: 978-953-307-243-2, 2011.

[74] L. Xu, J. Chen, Z. Cao and W. Zhang, “Identification of oil-water flow patterns in a vertical well using a dual-ring conductance probe array,” IEEE Transactions on Instrumentation and Measurement, vol. 65, no 5, pp. 1249-1258, 2016.

[75] J. Ye and L. Guo, “Multiphase flow pattern recognition in pipeline-riser system by statistical feature clustering of pressure fluctuations,” Chemical Engineering Science, vol. 102, pp. 486-501, 2013.

[76] G. Bowden, G. Dandy and H. Maier, “Input determination for neural network models in water resources applications. Part 1 - Background and methodology,” Journal of Hydrology, vol. 301, pp. 75-92, 2005.

[77] G. Bowden, H. Maier and G. Dandy, “Input determination for neural network models in water resources applications. Part 2 - Case study: forecasting salinity in a river,” Journal of Hydrology, vol. 301, pp. 93-107, 2005.

[78] R. May, H. Maier, G. Dandy and T. Fernando, “Non-linear variable selection for artificial neural networks using partial mutual information,” Environmental Modelling & Software, vol. 23, pp. 1312-1326, 2008.

[79] S. Galelli and A. Castelletti, “Tree-based iterative input variable selection for hydrological modelling,” Water Resources Research, vol. 49, pp. 4295-4310, 2013.

[80] S. Galelli, G. Humphrey, H. Maier, A. Castelletti, G. Dandy and M. Gibbs, “An evaluation framework for input variable selection algorithms for environmental data-driven models,” Environmental Modelling & Software, vol. 62, pp. 33-51, 2014.

[81] X. Li, H. Marier and A. Zecchin, “Improvement PMI-based input variable selection approach for artificial neural network and other data driven environmental and water resource models,” Environmental Modelling & Software, vol. 65, pp. 15-29, 2015.

[82] A. Sharma, “Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: part 1—A strategy for system predictor identification,” Journal of Hydrology, vol. 239, no. 1-4, pp. 232-239, 2000.

[83] S. Galelli and A. Castelletti, “Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling,” Hydrology & Earth System Sciences, vol. 17, no. 7, pp. 2669-2684, 2013.

[84] H. Farfani, F. Behnamfar and A. Fathollahi, “Dynamic analysis of soil-structure interaction using the neural networks and the support vector machines,” Expert Systems with Applications, vol. 42, no. 22, pp. 8971–8981, 2015.

[85] S. Belaid and A. Mellit, “Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate,” Energy Conversion and Management, vol. 118, pp. 105–118, 2016.

[86] M. Ramli, S. Twaha and Y. Al-Turki, “Investigating the performance of support vector machine and artificial neural networks in predicting solar radiation on a tilted surface: Saudi Arabia case study,” Energy Conversion and Management, vol. 105, pp. 442–452, 2015.

[87] F. Chen, H. Li, Z. Xu, S. Hou and D. Yang, “User-friendly optimization approach of fed-batch fermentation conditions for the production of iturin a using artificial neural networks and support vector machine,” Electronic Journal of Biotechnology, vol. 18, no. 4, pp. 273–280, 2015.

[88] A. Gandomi and A. Alavi, “A new multi-gene genetic programming approach to nonlinear system modeling. Part I, materials and structural engineering

problems," Neural Computing and Applications, vol. 21, no. 1, pp. 171–187, 2012.

[89] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[90] H. Drucker, C. Burges, L. Kaufman, A. Smola and V. Vapnik, "Support vector regression machines," Neural Information Processing System, vol. 9, pp. 155–161, 1997.

[91] J. Suykens and J. Vandewalle. Least squares support vector machine classifiers, Neural Processing Letters, vol. 9, no. 3, pp. 293-300, 1999.

[92] G. Huang, H. Zhou, X. Ding and R. Zhang. Extreme learning machine for regression and multiclass classification, IEEE Transactions on System, Man, and Cybernetics-Part B: Cybernetics, vol. 42, no. 2, pp. 513-529, 2012.

[93] X. Liu, C. Gao and P. Li. A comparative analysis of support vector machines and extreme learning machines, Neural Networks, vol. 33, pp. 58-66, 2012.

[94] J. Koza, Genetic Programming: On the Programming of Computers by Means of Natural Evolution. Cambridge, MA, USA: MIT Press, 1992.

[95] J. Madár, J. Abonyi and F. Szeifert, "Genetic programming for the identification of nonlinear input-output models," Industrial Engineering Chemistry Research, vol. 44, no. 9, pp. 3178–3186, 2005.

[96] D. Searson, “GPTIPS 2: An open-source software platform for symbolic data mining,” in Handbook of Genetic Programming Applications. New York, NY, USA: Springer, 2015, ch. 22.

[97] J. Kunze, R. Storm and T. Wang, “Coriolis mass flow measurement with entrained gas,” in Proceedings of Sensors and Measuring Systems, Nuremberg, German, 2014, pp. 1–6.

[98] T. Wang and Y. Hussain, Method for installing and operating a mass flowmeter and mass flowmeter, US Patent 20100326204 A1, 2010.

[99] T. Wang and Y. Hussain, “Coriolis mass flow measurement at cryogenic temperatures,” Flow Measurement and Instrumentation, vol. 20, no. 3, pp. 110-115, 2009.

[100] R. Storm, Process for operating a Coriolis mass flow rate measurement device, Patent US20080011101 A1, 2008.

[101] M. Gevrey, I. Dimopoulos and S. Lek, “Review and comparison of methods to study the contribution of variables in artificial neural network models,” Ecological Modelling, vol. 160, no. 3, pp. 249-264, 2003.

[102] M. Gevrey, I. Dimopoulos and S. Lek, “Two-way interaction of input variables in the sensitivity analysis of neural network models,” Ecological Modelling, vol. 195, no. 1-2, pp. 43-50, 2006.

[103] [On-line] CO_2 flow metering through multi-modal sensing and statistical data fusion , Autumn 2016 Biannual Meeting, Edinburgh, 14-15, Sep. 2016.

https://ukccsrc.ac.uk/sites/default/files/documents/event/edinburghBASep16/edinburgh_biannual_sep16_proceedings_final.pdf Accessd 12 April 2017.

Appendix 1 Program for BP-ANN

```
% ********************************************************************%
% File name:   BPANN.m
% Train data: 1_apparent mass flowrate; 2_observed density; 3_process temperature;
% 4_sensorA; 5_sensor B; 6_drive level; 7_time shift; 8_tuve frequency;
% 9_sensor balance; 10_damping; 11_two phase indicator; 12_DP; 13_Z; 14_θ
% Desired output: 1_liquid mass flowrate; 2_gas volume fraction
% Toolbox: Matlab Neural Network Toolbox
 % ********************************************************************%

%% load data
      load T2gw_h;
% set the number of variables
      p_train=ptrain(:,[1,2,10,12])';
      t_train=ttrain(:,1)';
      p_test=ptest(:,[1,2,10,12])';
      t_test=ttest(:,1)';

      hiddennode=9; % the number of hidden neurons

% normalise the training data
      [inputn,inputps]=mapminmax(p_train,-1,1);
      [outputn,outputps]=mapminmax(t_train,-1,1);

%% create BP net
% set parameters for training a BP-ANN
      net=newff(inputn,outputn,hiddennode,{'tansig', 'purelin'},'trainbr','learngdm');
      net.trainParam.epochs=1500;
      net.trainParam.lr=0.1;
      net.trainParam.goal=0.00001;

      net.divideParam.trainRatio = 80/100;
```

```
net.divideParam.valRatio = 10/100;
net.divideParam.testRatio =10/100;

%% BP net training
net=train(net,inputn,outputn);
% normailise the perdict data
inputn_test=mapminmax('apply',p_test,inputps);

%% BP net on training data
an=sim(net,inputn);
% converse normalization of the output
bp_prediction_result1=mapminmax('reverse',an,outputps);
% calculate NRMSD after correction
absolute_error1=bp_prediction_result1-t_train;

% calculate relative error after correction
corrected_error1=bp_prediction_result1./t_train-1;
corrected_error1=corrected_error1';
L=length(absolute_error1);
NRMSD1=sqrt(sum(absolute_error1.^2)/L)/mean(t_train);
disp(['NRMSD on training data after correction']);
NRMSD1

% calculate orginal NRMSD of mass flow before correction
absolute_error_original1=ptrain(:,2)-ttrain(:,1);
% calculate relative of mass flow errors before correction
original_error1=ptrain(:,2)./ttrain(:,1)-1;
L=length(absolute_error_original1);
NRMSD_original1=sqrt(sum(absolute_error_original1.^2)/L)/mean(t_train);
disp(['NRMSD on training data before correction']);
NRMSD_original1

%% BP net on test data
```

```
an2=sim(net,inputn_test);
% converse normalization of the output
bp_prediction_result2=mapminmax('reverse',an2,outputps);
% calculate NRMSD after correction
absolute_error2=bp_prediction_result2-t_test;
% calculate relative error after correction
corrected_error2=bp_prediction_result2./t_test-1;
corrected_error2=corrected_error2';
L=length(absolute_error2);
NRMSD2=sqrt(sum(absolute_error2.^2)/L)/mean(t_test);
disp(['NRMSD on test data after correction']);
NRMSD2

% calculate orginal NRMSD of mass flow before correction
absolute_error_original2=ptest(:,2)-ttest(:,1);
% calculate relative of mass flow errors before correction
original_error2=ptest(:,2)./ttest(:,1)-1;
L=length(absolute_error_original2);
NRMSD_original2=sqrt(sum(absolute_error_original2.^2)/L)/mean(t_test);
disp(['NRMSD on test data before correction']);
NRMSD_original2
```

Appendix 2 Program for SVM

```
% ********************************************************************%
% File name:   SVM.m
% Train data: 1_apparent mass flowrate; 2_observed density; 3_process temperature;
% 4_sensorA; 5_sensor B; 6_drive level; 7_time shift; 8_tuve frequency;
% 9_sensor balance; 10_damping; 11_two phase indicator; 12_DP; 13_Z; 14_θ
% Desired output: 1_liquid mass flowrate; 2_gas volume fraction
% Toolbox: Libsvm
 % ********************************************************************%

%% load data
      load T2gw_h;
% set the number of variables
      p_train=ptrain(:,[1,2,10,12])';
      t_train=ttrain(:,1)';
      p_test=ptest(:,[1,2,10,12])';
      t_test=ttest(:,1)';
% data preprocessing
      [inputn,inputps]=mapminmax(p_train,-1,1);
      inputn=inputn';
      [outputn,outputps]=mapminmax(t_train,-1,1);
      outputn = outputn';

% optimal parameters selection
       [bestmse,bestc,bestg] = SVMcgForRegress(outputn,inputn,-8,8,-8,8,3,1,1,0.1);
% print the selected result
        disp('print the reslut of fine slection');
        str = sprintf( 'Best Cross Validation MSE = %g Best c = %g Best g =
%g',bestmse,bestc,bestg);
        disp(str);

%% Train SVM network with the optimal parameters
```

```
% -s (svm type)0-C-SVC; 1-nu-SVC; 2-one-class SVM; 3-epsilon-SVR; 4-nu-SVR
% -t (kernel_type): 0-linear; 1-polynomial; 2-radial basis function; 3-sigmoid
        cmd = ['-c ', num2str(bestc), ' -g ', num2str(bestg) , ' -s 3 -t 2 -p 0.01'];
        model = svmtrain(outputn,inputn,cmd);
        save best_model model

%% SVM network regression prediction
% test SVM on training data
        [predict,mse,dicision_values] = svmpredict(a,inputn,model,'-b 0');
        predict_SVM1 = mapminmax('reverse',predict,outputps);
        predict_SVM1 = predict_SVM1';

% test SVM on test data
        inputn_test=mapminmax('apply',p_test,inputps);
        outputn_test=mapminmax('apply',t_test,outputps);
        inputn_test=inputn_test';
        outputn_test=outputn_test';
        [predict,mse,dicision_values] = svmpredict(a,inputn_test,model,'-b 0');
        predict_SVM2 = mapminmax('reverse',predict,outputps);
        predict_SVM2 = predict_SVM2';
```

Appendix 3 Program for LSSVM

```
% ********************************************************************%
% File name:   LSSVM.m
% Train data: 1_apparent mass flowrate; 2_observed density; 3_process temperature;
% 4_sensorA; 5_sensor B; 6_drive level; 7_time shift; 8_tuve frequency;
% 9_sensor balance; 10_damping; 11_two phase indicator; 12_DP; 13_Z; 14_θ
% Desired output: 1_liquid mass flowrate; 2_gas volume fraction
% Toolbox: LS-SVMLab
 % ********************************************************************%

% load data
   load T2go_v;
%  set the number of variables
   p_train=ptrain(:,[1,2,10,12]);
   t_train=ttrain(:,1);
   p_test=ptest(:,[1,2,10,12]);
   t_test=ttest(:,1);

% lin_(Linear); poly_(Polynominal); RBF_ (Radial Basis Function); MLP_(Multiplayer
% Perceptron)
   model=initlssvm(p_train,t_train,'f',[],[],'RBF_kernel','o');
   model=tunelssvm(model,'simplex','crossvalidatelssvm',{5,'mse'});
   model=trainlssvm(model);

   save best_model model

%  test LSSVM on training data
   predict_SVM1=simlssvm(model,p_train);

%  test LSSVM on test data
   predict_SVM2=simlssvm(model,p_test);
```

Appendix 4 Program for GP

```
% ************************************************************************%
% File name:   GP_config.m
% Train data: 1_apparent mass flowrate; 2_observed density; 3_process temperature;
% 4_sensorA; 5_sensor B; 6_drive level; 7_time shift; 8_tuve frequency;
% 9_sensor balance; 10_damping; 11_two phase indicator; 12_DP; 13_Z; 14_θ
% Desired output: 1_liquid mass flowrate; 2_gas volume fraction
% Toolbox: GPTIPS
 % ************************************************************************%

% run control parameters
    gp.runcontrol.pop_size = 250;
    gp.runcontrol.timeout = 10;
    gp.runcontrol.runs = 3;

% selection
    gp.selection.tournament.size = 25;
    gp.selection.tournament.p_pareto = 0.7;
    gp.selection.elite_fraction = 0.7;
    gp.nodes.const.p_int= 0.5;

% fitness
    gp.fitness.terminate = true;
    gp.fitness.terminate_value = 0.2;

% load data
    load K1gw_17_V
    p_train=ptrain(:,[1,2,10,12]);
    t_train=ttrain(:,2);
    p_test=ptest(:,[1,2,10,12]);
    t_test=ttest(:,2);
```

```
    gp.userdata.xtest = p_test; %testing set (inputs)
    gp.userdata.ytest = t_test; %testing set (output)
    gp.userdata.xtrain = p_train; %training set (inputs)
    gp.userdata.ytrain = t_train; %training set (output)
    gp.userdata.name = 'Mass flow';

%genes
    gp.genes.max_genes = 6;

%define building block function nodes
    gp.nodes.functions.name = {'times','minus','plus','tanh','mult3','add3'};
```

Publications and Dissemination

The following publications were produced during the course of the work leading to the preparation of this thesis.

1. **L. Wang**, Y. Yan, X. Wang and T. Wang, "Application of soft computing techniques to multiphase flow measurement: A review," *International Journal of Multiphase Flow*, 2017. (under review)

2. **L. Wang**, Y. Yan, X. Wang and T. Wang, "Mass flow measurement of gas-liquid two-phase CO_2 in CCS transportation pipelines using Coriolis flowmeters," *International Journal of Greenhouse Gas Control*, 2017. (under review)

3. **L. Wang**, Y. Yan, X. Wang and T. Wang, "Input variable selection for data-driven models of Coriolis flowmeters for two-phase flow measurement," *Measurement Science and Technology*, vol. 28, no.3, pp. 1-12, 2017.

4. **L. Wang**, J. Liu, Y. Yan, X. Wang and T. Wang, "Gas-liquid two-phase flow measurement using Coriolis flowmeters incorporating artificial neural network, support vector machine and genetic programming algorithms," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 5, pp. 852-868, 2017.

5. **L. Wang**, J. Liu, Y. Yan, X. Wang and T. Wang, "Mass flow measurement of two-phase carbon dioxide using Coriolis flowmeters," in *Proceedings of IEEE International Instrumentation and Measurement Technology Conference*, pp. 1299-1303, Torino, Italy, May 22-25, 2017.

6. **L. Wang**, Y. Yan, J. Liu, X. Wang and T. Wang, "Gas-liquid two-phase flow measurement using Coriolis flowmeters incorporating neural networks," in *Proceedings of IEEE International Instrumentation and Measurement Technology Conference*, pp.747-751, Taipei, Taiwan, May 23-26, 2016.

7. **L. Wang,** J. Liu, Y. Yan, X. Wang and T. Wang, "Gas-liquid Two-phase Flow Measurement using Coriolis Flowmeters incorporating Neural Networks," *The 9th International Symposium on Measurement Techniques for Multiphase Flows*, September 23-25, Sapporo, Hokkaido, Japan 2015. (Best Presentation Award)