DEPARTMENT OF CHEMICAL ENGINEERING

Multiphase Flow in Pipelines: An Analysis of the Influence of Empirical Correlations on Mechanistic Models

Adrian-George Brustur

This thesis is presented for the Degree of
Master of Chemical Engineering
of
Curtin University

March 2014

Declaration

I declare that to the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature	:	
Name	:	Adrian-George Brustur
Date	:	March 2014

Preface

The subject of this study came about during an investigation into mechanistic flow models. Previous discussions with A/Prof. Martyn Ray about offshore oil and gas exploration had drawn my research interest to the phenomena associated with hydrocarbon flow in pipelines and flow assurance. Browsing through SPE papers in an attempt to find a subject for my thesis, I came across a paper discussing a unified mechanistic flow model, whose authors were part of a very active research team at Tulsa University. The paper briefly described several empirical correlations, as part of the mechanistic flow model, and referred to other previous mechanistic models. As I started following the references and reading more about mechanistic flow models, I became more familiar with the subject, but also started being intrigued by the fact that for the same parameter, different models were using different empirical correlations. In fact, the subject of empirical correlations is so vast, that certain studies are entirely dedicated to reviewing correlations for certain parameters, such as the friction factor or the void fraction. This sparked the idea of analysing whether an empirical correlation in a particular mechanistic flow model could simply be "unplugged" and a new correlation "plugged in", as you would with a mechanical component which was part of a mechanical assembly. This is not just an academic question, because it addresses the practical situation where a better empirical correlation becomes available and it needs to be "fitted" into the mechanistic model. Even though authors specify the empirical correlations to be used with their model, one has to raise the question whether there are better correlations that could be used. Although the idea of analysing the effect of empirical correlations on mechanistic flow models may seem simple, once one comes to the realisation that replacing empirical correlations may lead to significant changes to the calculation algorithm, complexity comes into play.

The previous situation described the case when we had a correlation and we needed to "plug" it into the model. Considering it in a different way, one could think of a case where we have the model and we need to find the empirical correlation that leads to the best predictions. This becomes an optimisation problem.

Later, reading more about unified mechanistic flow models, I became even more baffled by the number of empirical correlations used by certain models. There are models which use more than one empirical correlation for the estimation of the same model parameter. An example is the use of two correlations for the interfacial friction factor, according to whether the flow is stratified or annular in the Taylor bubble zone of the slug flow. This led me to believe that the interfacial friction factor is a particularly important parameter of the flow model. However, not much research is dedicated to the subject of model sensitivity analysis rela-

tive to the interfacial friction factor. In fact, little research has been done in the area of sensitivity analysis of mechanistic flow models. This set me on a path of model experimentation, by performing model numerical simulations, while keeping some of the parameters constant. My interest turned to determining whether it was important to find better correlations for some of the model parameters. If the model is fairly insensitive to variations of some of the input parameters, obviously, a better correlation would not improve the accuracy of the model much. This becomes both a sensitivity analysis and a mechanistic model simplification problem.

The title of this thesis encompasses all these problems, which will be discussed in detail in the following chapters. In my view, research is meant to generate solutions, but also and equally important to generate questions. I hope this research study has generated a mix of both.

This study progressed from initially one question, to a couple of questions, then to many questions, until it finally gained momentum, after a very slow start. Finding answers to these questions was not easy, particularly when not much research at all has been done in the area of the use of empirical correlations for mechanistic model optimisation and simplification. Fleshing out the ideas and expressing them in a concise and unambiguous manner was even more difficult. Some of the chapters had to be rewritten from scratch and others modified, as both my supervisors, A/Prof. Nicoleta Maynard and Dr. Gordon Ingram, took on the enormous task of reading chapters of muddled reasoning and confusing ideas, while I was still trying to get some clarity as to where I was heading with my research.

I would like to thank A/Prof. Martyn Ray for pointing me in the right direction in the early stages of the research, reading the content of the thesis and making valuable suggestions. I am also grateful to my supervisors A/Prof. Nicoleta Maynard and Dr. Gordon Ingram, who helped me refine the ideas and provided advice and guidance at different stages of this research.

The author alone assumes responsibility for the conclusions of this thesis and any errors it may contain.

Adrian-George Brustur
Perth, March 2014

Abstract

Mechanistic models represent process behaviour through sets of equations based upon fundamental conservation laws and scientific knowledge of the process. In order to compensate for the lack of sufficient knowledge of the physical phenomena, empirical correlations are used to estimate certain model parameters. Empirical correlations in mechanistic models make them data-sensitive, to the extent that they become data-driven. The current industry practice for model predictions is to check the data fit of a selection of mechanistic models, for the purpose of choosing the most accurate model.

This study proposes a method to use a set of alternative empirical correlations in a mechanistic model with the view of optimising or simplifying the model. This suggests that the current industry practice needs to be changed to first test the data fit of a set of empirical correlations in each model and only then select the most accurate mechanistic model.

To achieve mechanistic model optimisation a two stage modelling approach is taken. During the first stage, training, a history table of optimal correlation sets is created. For the second stage, prediction, two methods are proposed:

1. A history matching algorithm based on sensitivity analysis. This is an algorithm based on the minimisation of an objective function, with the view of identifying an optimal correlation set to be used for prediction.

2. Data mining. Sorting and grouping techniques are used to identify regions where a certain empirical correlation set is optimal. Mechanistic model simplification is useful for situations where it is difficult to find empirical correlations for certain model parameters, or the empirical correlations have a computationally-intensive component. A method based on sensitivity analysis is proposed to keep constant empirical parameters with a low influence on the model output.

Based on the observation that each empirical correlation depends on a possibly different set of input parameters, the structure of the calculation algorithm used by the model may change each time a different correlation is used. Essentially, one cannot "unplug" an empirical correlation out of a mechanistic model and replace it with an alternative correlation, without first analysing whether the calculation algorithm has to be modified. This study proposes a method to analyse the changes in the calculation algorithm, when a new empirical correlation is used.

The theoretical concepts developed in this investigation were applied to the Zhang et al. (2000) mechanistic slug flow model, widely used in the oil and gas industry for the prediction of pressure gradient and liquid holdup in two-phase pipe flow. In order to achieve model optimisation, a set of empirical correlations

pertaining to slug flow was selected and systematically used to replace the original correlations in the model. The results showed a better performance of the optimised model compared to the original Zhang et al. (2000) model.

Sensitivity analysis of the empirical parameters used in the Zhang et al. model determined:

- An easy method to improve the accuracy of the model, by identifying the empirical parameter with the greatest influence on the model output, in this case the liquid holdup fraction in the slug.
- A direction for experimental research, to find better empirical correlations for the empirical parameters that have a higher sensitivity factor, in this case liquid holdup, Taylor bubble velocity and wetted wall fraction.

This study concludes that before a mechanistic model is used on production data, the empirical correlations used by the model have to be "tuned" by selecting the appropriate correlation to ensure the best model performance.

Contents

List of figures

List of tables

Glossary

IPM Prosper

PROSPER is a well performance, design and optimisation program for modelling most types of well configurations found in the worldwide oil and gas industry today. PROSPER can assist the production or reservoir engineer to predict tubing and pipeline hydraulics and temperatures with accuracy and speed. PROSPER's sensitivity calculation features enable existing well designs to be optimised and the effects of future changes in system parameters to be assessed.

NODAL

An analytical tool used in forecasting the performance of the various elements comprising the well completion and production system. This analysis is used to optimize the completion design to suit the reservoir deliverability, identify restrictions or limits present in the production system and identify any means of improving production efficiency. NODAL^{TM} is a trademark of Schlumberger.

PipeSim

Steady-state, multiphase flow simulator software for the design and diagnostic analysis of oil and gas production systems. The software tools model multiphase flow from the reservoir to the wellhead. The software also analyses flowline and surface facility performance to generate comprehensive production system analysis. PIPESIM was created by Schlumberger.

Acronyms

ANN	Artificial Neural Networks.
CA	Calculation Algorithm.
DCP	Downstream Choke Pressure.
DFP	Downstream Flowline Pressure.
DFS	Depth First Search.
EO	Equation Oriented.
IPR	Inflow Performance Relationship.
LHS	Left Hand Side.
MARS	Multivariate Adaptive Regressive Splines.
PG	Pressure Gradient.
RHS	Right Hand Side.
SM	Sequential Modular.
THP	Tubing Head Pressure.
TUFFP	Tulsa University Fluid Flow Project.
URR	Ultimate Recoverable Reserves.

List of symbols

A_F	Area occupied by the liquid phase in the film zone of the slug unit.
A_G	Area occupied by the gas phase in the film zone of the slug unit.
A_k	Cross-sectional area occupied by phase k.
A_{CD}	Cross-sectional area embraced by the wetted wall and its chord.
A_{pipe}	Pipe cross-sectional area.
C_e	Zhang elevation coefficient used in the Zhang et al. correlation for H_{LLS}.
D	Dependency chain.
$E_{1,k}$	Average percent error for flow model k.
$E_{2,k}$	Absolute average percent error for flow model k.
$E_{3,k}$	Percent standard deviation for flow model k.
$E_{4,k}$	Average error for flow model k.
$E_{5,k}$	Absolute average error for flow model k.
$E_{6,k}$	Standard deviation for flow model k.
F	External force applied to a control volume.
F	Objective function.
F_{PR_k}	Relative performance factor for model k.
Fr	Froude number.
H_{LLS}	Liquid holdup in the slug zone of the unit.
H_{LTB}	Liquid holdup in the film zone of the slug unit.
H_{LU}	Slug unit liquid holdup.
L_F	Length of the liquid film in the slug unit.

L_S	Length of the liquid slug zone.
L_U	Length of the slug unit.
O	Optimal correlation set
P_k	External pressure applied on phase k in the control volume.
$P_{Separator}$	Separator pressure.
P_{wf}	Well flow pressure.
ΔP_{Choke}	Pressure drop across the choke.
$\Delta P_{Flowline}$	Pressure drop along the flowline.
$\Delta P_{GravelPack}$	Pressure drop across the gravel pack, if this is used as a sand control measure.
$\Delta P_{Perforations}$	Pressure drop across the casing perforations.
$\Delta P_{Reservoir}$	Pressure drawdown (Pressure drop between the reservoir and the bottomhole).
$\Delta P_{SafetyValve}$	Pressure drop across the safety valve.
$\Delta P_{TubingRestrictions}$	Pressure drop across various tubing restrictions. This includes landing nipples, internal upsets, etc.
ΔP_{Tubing}	Frictional pressure loss along the production line, from the well bottom hole to the well head.
$\Delta P_{WellHead}$	Pressure drop across the well head.
$\frac{dP}{dL}$	Pressure gradient.
$\overline{P}_{Reservoir}$	Average reservoir pressure.
R	Residual. The norm of the difference between the model prediction and the measured process output.
Re_S	Reynolds number of the gas-liquid mixture in the slug zone.
Re_{GW}	Reynolds number of the gas phase.
Re_{LW}	Reynolds number of the liquid phase.
$\hat{R}$	Predicted value of the residual during forecasting stage.
S_F	Liquid wetted wall perimeter.
S_G	Gas contacted wall perimeter.
S_S	Perimeter occupied by the gas-liquid in the slug zone.

S_i	Interfacial perimeter.
S_{CD}	Chord between the highest liquid level points in a pipe cross-section through which a mixture of gas-liquid flows.
T_F	Time interval required by the liquid film in a slug unit to pass through an imaginary pipe cross-section plane.
T_S	Time interval required by the liquid slug in a slug unit to pass through an imaginary pipe cross-section plane.
T_U	Slug period. This assumes the condition of equal slug unit length.
T_{sm}	Momentum exchange coefficient in the Zhang et al. correlation for H_{LLS}.
W	Total mass flow rate.
W_G	Gas mass flow rate.
W_L	Liquid mass flow rate.
We	Weber number.
$c_{W,k}$	Weight coefficient for model input k.
d	Pipe diameter.
e	Empirical parameter.
e_i	Absolute error at data point i.
f_S	Friction factor of the gas-liquid mixture in the slug zone.
f_i	Interfacial friction factor.
f_{GW}	Gas-wall friction factor.
f_{LW}	Liquid-wall friction factor.
f_{water}	Water cut.
g	Empirical correlation functional dependency.
g	Gravitational acceleration.
m	Mass.
p	Grid size used by the Elementary Effects method.
pe_i	Relative error at data point i.
q	Total volumetric flow rate.
q_G	Gas volumetric flow rate.

q_L	Liquid volumetric flow rate.
r	Number of random trajectories used by the Elementary Effects method.
t	Time.
v_M	Mixture velocity.
v_k	Superficial velocity of phase k.
v_{GTB}	Absolute velocity of the gas inside the liquid film zone.
v_{LTB}	Absolute velocity of the liquid film under the Taylor bubble in the liquid film zone.
v_{SG}	Superficial gas velocity.
v_{SLIP}	Slip velocity.
v_{SL}	Superficial liquid velocity.
v_{TB}	Taylor bubble velocity.
v_{drift}	Drift velocity.
x	Distance measured in the direction of the fluid flow.
x	Quality.
x_h	Data point in the history table.
$\hat{y}$	Model prediction during forecasting stage.
$\tilde{y}$	Model estimation.
y	Measured process output.
y^e	Model prediction during training stage.
y_t	True model value.
α_k	Volume fraction of phase k.
β	Pipe inclination angle.
Δ	Value of the cell size used by the Elementary Effects method.
δ_k	Projection of the distance between an input data point and a history table point on axis k of the hypercube.
Φ	Vector containing estimations of all the empirical parameters in the model for a given input data point.
Γ	Mass transfer parameter.

Ω	Set of empirical correlations.
ρ_G	Gas density.
ρ_L	Liquid density.
ρ_S	Gas-liquid density of the mixture in the slug zone.
ρ_k	Density of phase k.
σ	Surface tension.
Θ	Wetted wall fraction.
τ_S	Shear stress of the gas-liquid mixture at the wall interface.
τ_i	Shear stress at the gas-liquid interface.
τ_{GW}	Shear stress at the gas-wall interface.
τ_{LW}	Shear stress at the liquid-wall interface.
θ	Pipe inclination angle.

Chapter 1

Introduction

Multiphase flow occurs in many industrial situations and applications. In the petroleum industry, it occurs in oil and gas wells, piping and gathering systems, and equipment in refineries such as boilers, condensers, separators and distillation towers. Out of all the types of multiphase flow, gas-liquid (oil-natural gas), liquid-liquid (oil-water), gas-liquid-liquid (natural gas-oil-water) and solid-liquid-liquid-gas flows (sand-oil-water-natural gas), this study will concentrate on gas-liquid flows.

The challenges posed by transporting multiphase fluids through pipelines are significant due to complex physical and chemical phenomena. The NODAL^{TM} analysis approach commonly used in the petroleum industry, allows for well system performance analysis and surface utilities sizing. The NODAL^{TM} approach assumes the selection of a number of nodes in the well production system and the division of the system at a chosen node. The system components upstream from the chosen node constitute the inflow section and the components downstream from the node, the outflow section. In NODAL^{TM} analysis, if the node is chosen at the wellflow level, Equations 1.1 and 1.2 determine the wellflow pressure P_{wf} on the Inflow Performance Relationship (IPR) curve. The inflow equation is:

$$P_{wf} = \overline{P}_{Reservoir} - \Delta P_{Reservoir} - \Delta P_{GravelPack} - \Delta P_{Perforations} \tag{1.1}$$

The outflow equation is:

$$\begin{aligned} P_{wf} = P_{Separator} + \Delta P_{Flowline} + \Delta P_{Choke} + \Delta P_{WellHead} + \\ + \Delta P_{SafetyValve} + \Delta P_{TubingRestrictions} + \Delta P_{Tubing} \end{aligned} \tag{1.2}$$

Figure 1.1 shows the IPR curve, Tubing Head Pressure (THP), Downstream Choke Pressure (DCP), Downstream Flowline Pressure (DFP) and the separator pressure lines, plotted together. Starting from the reservoir pressure, the drawdown pressure is the first pressure drop $\Delta P_{Reservoir}$, after which each component introduces a flow rate dependent pressure drop. Figure 1.1 shows the importance of the tubing ΔP_{Tubing} and flowline $\Delta P_{Flowline}$ pressure drops in the context of petroleum production optimisation. A shift in the operating line, as shown in Figure 1.1, which may occur due to reservoir depletion or change of flow rate q, leads to changes in the value of the separator pressure. This pressure may fall below the minimum separator pressure set by the design. **This explains**

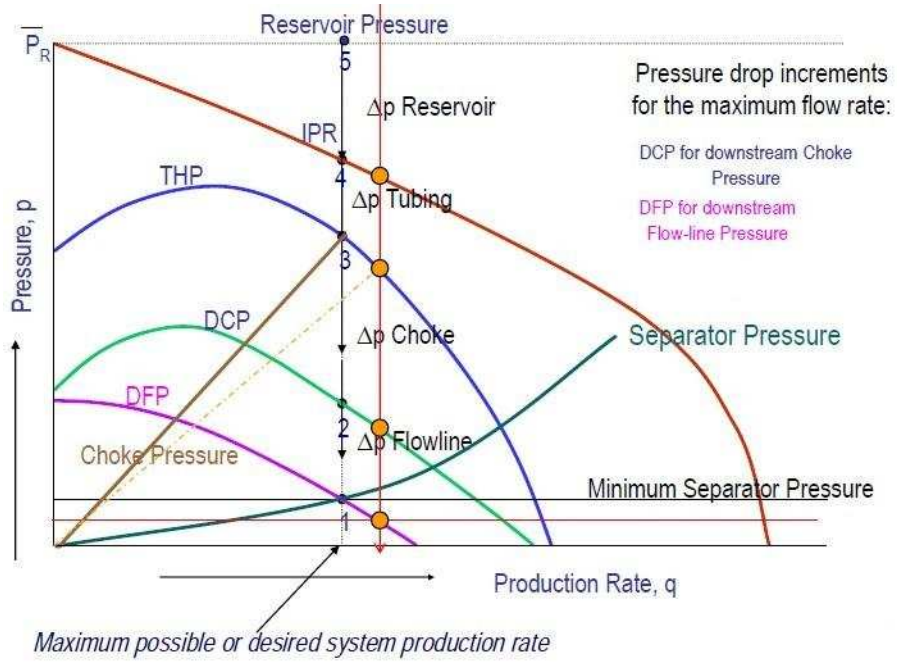


Figure 1.1: NODAL^{TM} system analysis on a well production system (Curtin 2011).

the importance of having accurate models for the calculation of multiphase flow pressure drop in pipelines of different inclination angles, in the context of petroleum production optimisation. Not knowing the pressure drops accurately, the production facilities can not be sized properly. The subject of this study is related to the flow models for the calculation of the pressure drop in the production string and transmission pipelines of different inclination angles.

In the petroleum industry, a reliable multiphase flow model is required in order to perform the following operations (Ismail 1999; Bratland 2009):

- **Selection of the appropriate tubing size**. If the tubing diameter is too large the well acts as a gas-liquid separator and the excessive gas slippage results in high bottom hole pressures. If the tube is too small, excessive frictional pressure drop will occur.
- **Prediction of the lift curve minima**, which determines when a well needs to be "kicked off", artificially lifted or recompleted.
- **Design of the gas lift system**.
- **Prediction of bottomhole pressures**. This is the case in offshore production facilities, especially when the wells are located deep below the water level. An inappropriate multiphase flow correlation may lead to inaccurate calculation of the bottomhole pressure, and subsequent errors in the facilities design.
- **Avoiding problems with gas hydrates, wax, asphaltene and scale buildup**.
- **Providing data for corrosion, erosion and corrosion-erosion calculations**.

Due to the complex nature of multiphase flow, the problem was first approached by conducting experiments and determining **empirical models**. For the development of an empirical model a large number of experiments are required and unless a dimensional analysis is carried out, the model will only apply to the particular limited set of conditions (Falcone, Hewitt, and Alimonti 2010). Empirical models lack the understanding of the fundamental physical mechanisms, but have the great advantage of being simple.

During the last two decades most multiphase flow research has been directed towards the development of **mechanistic flow models**. They use fundamental knowledge of the interactions between process variables to define the model structure and require a fundamental understanding of the physics and chemistry governing the process.

Chapter 2

Research objectives, data and system selection

2.1 Introduction

Critical thinking is a concept that will be used in this chapter to show how the subject of this research investigation was chosen. Critical thinking will be used as a tool to determine the project objectives, its boundaries and provide a sketch of the path to follow to the solutions for the stated objectives.

2.2 Critical thinking process

For a deeper understanding of the project objectives, it is important to show the process that led to their conception. The best way to describe this process is to apply critical thinking analysis to the multiphase flow knowledge domain. Figure 2.1 is a graphical depiction of such a process. Not only does such a figure allow ideas to crystallise, but it also explains how they were arrived at and it promotes logical consistency. It also shows the natural process of raising questions around the use of empirical correlations in mechanistic multiphase flow models and the attempt to find answers.

Figure 2.1 is divided into four "swimlanes": logical inference, question, solution method and outcome. The "Logical inference" lane starts with the problem to be analysed. In this case we analyse the initiative of using alternative empirical correlations in the same mechanistic model. The "Logical inference" lane contains all the inferences we can make as a result of this initiative. The nodes are connected by "leads to" oriented edges and show the logical flow of the train of thoughts. The last nodes of the inference chains are the conclusions. The next lane, "Questions", contains questions that can be asked relative to the conclusions in the previous lane. The "Solution method" lane suggests methods to solve the questions asked in the previous lane. The last lane, "Outcome", contains the outcomes of the methods applied to the questions. This diagram contains both vertical and horizontal layers. The leftmost column develops the initial idea further into a set of ideas. This is the vertical layer. Each of these ideas branch out horizontally and as they traverse the different swimlanes, questions, solutions and outcomes are suggested. A horizontal layer is thus created for each idea in

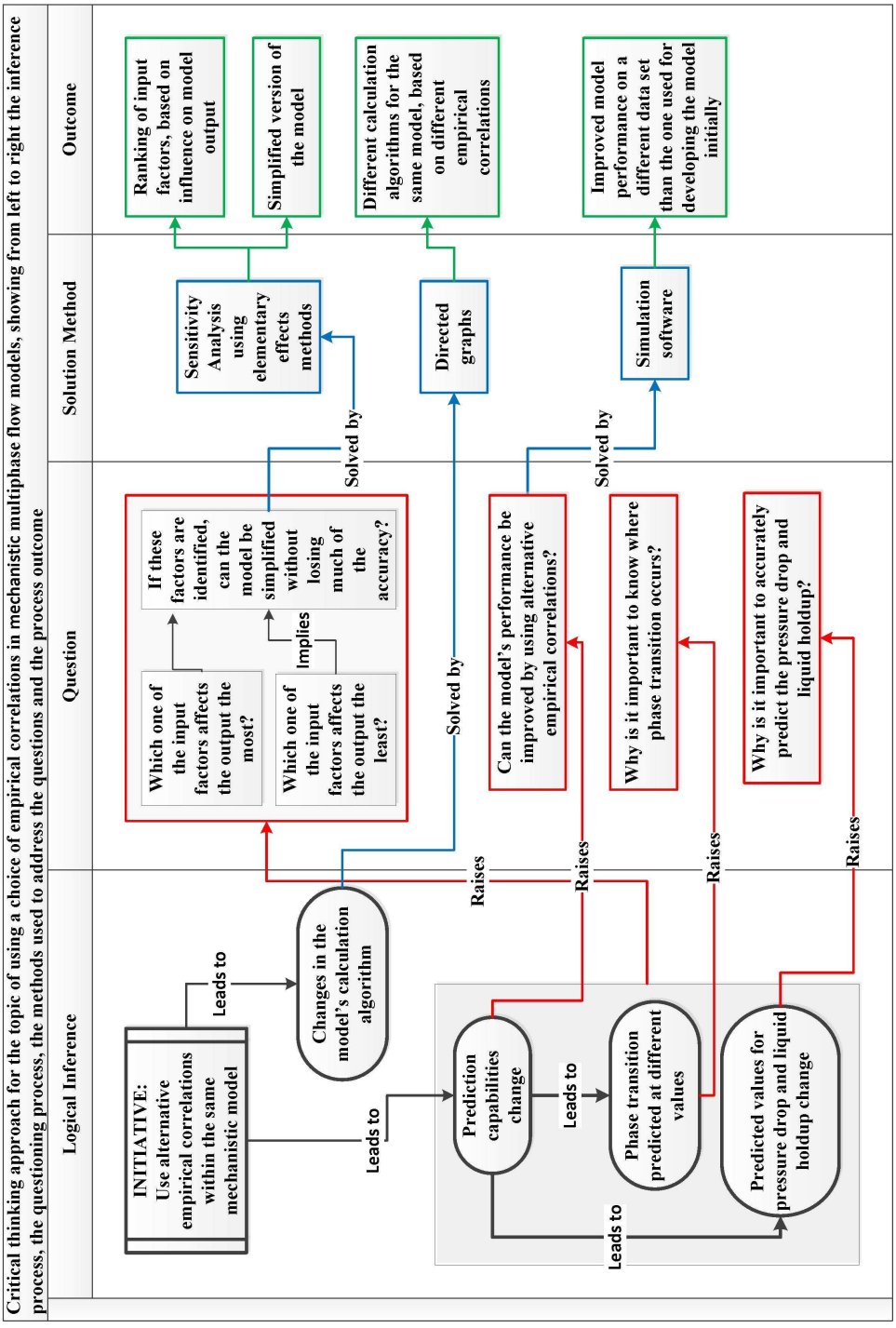


Figure 2.1: Overview of the critical thinking process applied to the multiphase flow.

the leftmost column.

The questions in Figure 2.1 presented the following possible research paths:

1. Changes of the mechanistic model's calculation algorithm.

2. Analysis of the possibility of improving the mechanistic model's performance by using a selection of empirical correlations.

3. Analysis of the possibility of simplifying the mechanistic model by maintaining constant some of the parameters, which would otherwise be calculated by using empirical correlations.

4. Analysis of the changes in the flow pattern transition prediction, associated with the process of using alternative empirical correlations.

This study will approach the first three research subjects. As far as the last research path is concerned, the concepts and methods discussed in the following chapters can be applied to determine the model predictions on each side of the boundary between adjacent flow regimes. This research path will not be explored in this study, because the discontinuity issue is already known in mechanistic flow models and a solution exists. The discontinuities vary when the empirical correlations for the two flow pattern models (such as slug-annular in vertical flow) are being replaced with alternative correlations. The agreed practice is to perform linear interpolation between the two different pressure gradient values, each pertaining to an adjacent flow regime. The linear interpolation eliminates the discontinuity. The transition boundaries can be determined with a flow pattern identification model, such as Barnea's (1987).

2.3 Project objectives

This study is an attempt to shed new light on the path that research has taken in the area of mechanistic models, by showing that **one of the areas ignored so far has been the optimisation of existing mechanistic models by finding more suitable empirical correlations**. In effect this is a tuning procedure, because there is no "silver bullet" mechanistic model to simulate the flow behaviour for all gas-liquid combinations, for all flow parameter ranges. The study will describe the steps involved in such a tuning process.

In summary, the project objectives are:

1. **To analyse the effects of using alternative empirical correlations on the mechanistic model's calculation algorithm**. The discussion in Chapter 7 will show that this is a complex process that may lead to a series of different calculation algorithms. A structured procedure to determine these different calculation algorithms is proposed in Section 7.5.

2. **To show that the use of a choice of empirical correlations leads to mechanistic model optimisation**. The effect of using different empirical correlations on the predicting capabilities of the mechanistic model are analysed in Chapter 8. It will be shown that for a data set of measured pressure gradients and liquid holdups, the iterative process of replacing the

correlations can bring the model prediction closer to, or at least as close to, the measured values as the original mechanistic model. This creates an association between the input data point and an optimal correlation set. The algorithm to achieve this will be described in Section 8.2 and constitutes the first step towards model optimisation.

The second step, a "history matching" method based on sensitivity analysis, is proposed in Chapter 9. Section 9.3.2 will show that sensitivity analysis together with history matching provide a method to optimise a mechanistic model.

3. **To show that the use of a selection of empirical correlations, in conjunction with sensitivity analysis, may lead to a simplified version of the mechanistic model**. Chapter 10 will show that ranking the effect of empirical correlations based on the contribution they have to the model's output is important, because it shows which correlations need to be replaced with better ones for more accurate predictions. Model simplification also reduces the execution time for computationally-intensive models, if the parameter with negligible contribution to the output involves iterations in its calculation procedure.

2.4 Mechanistic model selection

2.4.1 Flow pattern selection

As oil and gas reserves are becoming depleted or coming close to the end of their lifecycle, exploration is moving offshore, where the cost of constructing and operating oil and gas rigs is very high. The alternative offered by research is to utilise production systems with minimum offshore processing. Due to the latest advancements in multiphase flow modelling, driven mainly by the petroleum industry, it is now possible to transport unseparated multiphase gas-oil-water mixtures over relatively long distances to an onshore separation facility. This has generated massive savings for oil and gas developments, because multiphase trunklines have replaced topside offshore oil and gas rigs. This is the case of Snohvit and Ormen Lange fields off the Norwegian coast. Trunkline lengths of well over 100 km were used. At Tyrihans in the North Sea, where production started in 2009, the subsea tieback connected a new field to existing offshore structures and generated savings by avoiding a new topside. The flow line length in this case was 43 km (Bratland 2009). In these cases it is important to accurately predict the multiphase flow characteristics in long transportation pipelines. This not only involves the prediction of the flow pattern, but also the flow characteristics of the determined flow pattern.

Slug flow is the most common flow pattern in oil transportation pipelines (Marcano et al. 1998) and can be generated by hilly terrain, or changes in flow rate. Zhang et al. (2003b) discussed the issue of slug dissipation and generation in hilly terrain. **Slug flow is important to predict and characterise**, because of its undesirable effects:

- **Variation in oil and gas flow rates** entering the downstream processing facilities.

- **Mechanical damage to pipeline connections and support**.

However, slug flow is **desirable for artificial lift in heavy-oil reservoirs**. Its understanding is also important for designing effective gas lift and unloading strategies. The prediction of the slug-annular flow transition is important for controlling the stability of multiphase flow from the perforations, along the production string, to the wellhead. This is also an important issue in the design and operation of continuous flow in gas-lift installations used in floating production systems, where the produced fluids traverse a flexible riser to a Floating Production, Storage and Offloading (FPSO) vessel (Guerrero-Sarabia and Fairuzov 2006).

To summarise these considerations, **all the aspects of slug flow modelling are crucial for the design and operation of oil and gas transportation and processing facilities**. For this reason, a slug flow mechanistic model is selected in Section 2.4.2, in order to exemplify the theoretical aspects of this study.

2.4.2 Model selection

Numerous slug flow models can be found in the literature, but this study considered the Zhang et al. (2000) slug flow mechanistic model (Appendix E.7). The reasons the Zhang et al. (2000) model was chosen are:

1. The model was the basis for an extensive flow research project at Tulsa University (Tulsa University 2008). Improvements to the model have been made as a result of the project and it has become the cornerstone of a unified model for gas-liquid-liquid pipe flow (Tulsa University 2008).

2. The model was improved in 2003 and the new model was the first approach ever to determine flow pattern transitions based on slug flow dynamics, that is, considering all other flow patterns as particular cases of slug flow. Zhang et al. (2003c) showed that this flow pattern encompasses all the other flow patterns. Due to turbulence, gas is trapped in the liquid slug, which leads to dispersed bubble flow in this region. Each of the possible flow regimes can be obtained as a particular case of the more general case of slug flow. For example, stratified or annular flow can be obtained by taking the limit of $L_S \to 0$ and $L_F \to \infty$, where L_S is the liquid slug length and L_F is the length of the film zone in the slug unit. The opposite case for dispersed bubble flow can be obtained by taking the limits $L_S \to \infty$ and $L_F \to 0$.

3. The model is one of the advanced slug flow models and has good prediction capabilities. It was the first slug flow model to take into account the momentum exchange term, as described by Equation E.145 in Appendix E.7.

4. The model uses a large number of empirical factors, making it suitable for analysis from the point of view of this study.

Although the discussion revolves around this particular slug flow model, the concepts can be applied to any mechanistic model that has an empirical component. The Zhang et al. (2000) model was used only for the purpose of demonstrating how the theoretical concepts could be applied practically.

Table 2.1: Empirical correlations for the calculation of parameters L_S, H_{LLS}, v_{TB}, f_i and Θ used by the Zhang et al. (2000) and Zhang et al. (2003) slug flow models.

Correlation	L_S	H_{LLS}	v_{TB}	f_i	Θ
Zhang et al.(2000)	Zhang	Zhang	Bendiksen	Cohen & Hanratty	Grolman
Zhang et al.(2003)	Zhang	Zhang	Bendiksen	Andritsos & Hanratty/ Asali & Ambrosio	Grolman

Despite the Zhang et al. (2003c) model being a unified mechanistic model, which can determine the flow pattern and predict both the pressure drop and liquid holdup, the algorithm for the slug flow is very similar to the 2000 version. As the 2003 slug flow model takes into account liquid entrainment in the gas core of the slug unit, it also makes the notable change of using different empirical correlations for the friction factor f_i compared to the 2000 model (Appendix E.7.3).

Choosing the Zhang et al. (2000) model as a benchmark and not its 2003 version makes sense, because this way the theoretical framework discussed in the following chapters can analyse whether the replacement of the empirical correlation for f_i in the new model was beneficial for the accuracy of the model or not. Table 2.1 lists the empirical correlations used by the Zhang et al. (2000) and (2003) models, which can be found in Appendix F. A detailed description of the slug flow characteristics, parameters and notation is given in Appendix C.

2.5 Research data

2.5.1 Research data sources

In order to achieve the objectives mapped out earlier, data sets from multiple sources have been analysed:

- Tulsa University Fluid Flow Project (TUFFP) database - The team of researchers at TUFFP (Tulsa University 2008) conducted a four year project, with the following objectives:

  1. To investigate and understand the fundamental physical mechanisms describing the interaction between gas, oil and water.
  2. To develop a comprehensive unified mechanistic multiphase flow model.
  3. To conduct experimental studies at the Tulsa University Three Phase Flow Facility and generate new data sets.

  The TUFFP data bank contains both well and flowline data points, pertaining to all flow patterns. A 2008 report shows that the TUFFP well data bank had 2052 data points (Tulsa University 2008).

- Brill et al. (1981) - Twenty-nine multiphase flow tests were conducted in two 3-mile long pipelines in the Prudhoe Bay field, in Alaska. The pipelines used were 12" and 16", respectively. The flow line geometry was complex with sections of the pipeline following the terrain profile. In order to calculate the output pressure and liquid holdup, the line was divided into 36 segments with various inclination angles.

  Very few experiments have ever been conducted in an oil field with real formation fluids. Most of them have used flow test facilities with air and water, air and oil or other combinations, but very rarely formation fluids.

2.5.2 Research data selection

Although the Brill et al. (1981) data bank has the advantage of containing real formation fluid data, the complexity of the pipeline profile and the variation of flow patterns along the pipeline makes the identification of slug flow very difficult. This led to the selection of the TUFFP data bank, which has the advantage of having used straight flow lines. This makes the identification of slug flow data convenient.

The TUFFP database contains data referring to pipeline (0°, ±5°, ±10°) and wellbore (vertical and near vertical) inclination angles. All the flow patterns ranging from stratified smooth to annular misty flow are included. Out of all the database records, a subset of 95 flow data points for the slug flow pattern was selected for flow simulation. The selection process was random and the only criterion applied was for the data point to pertain to slug flow. The number of slug flow data points is a result of analysing TUFFP database points. As the database does not have a field for flow pattern, the FLOPATN software (Shoham 2006) was used to determine the flow pattern for each individual data point analysed.

The selected subset contains data points from Andritsos (1986), Beggs and Brill (1972) and Mukherjee and Brill (1985) flow tests. The two-phase liquid mixtures used to obtain the data are described in Appendix A, Table A.1.

All the 95 data points were used in the numerical simulations in Chapters 8 and 10. In Chapter 9, these 95 points are divided into two subsets:

- A first set of 71 data points was used as a history database.
- The remaining 24 data points were used as input data to test the optimised and simplified models.

The ratio between the number of data points in the history database and the number used for the analysis of model optimisation and simplification was chosen according to the criteria:

- A sufficient number of data points has to exist in the history database for the optimisation and simplification algorithms to function properly.
- A sufficient number of test data has to be available for the study to reach reasonable conclusions.

Based on this criteria, a ratio of:

$$\text{ratio} = \frac{\text{number of test data points}}{\text{number of history database points}} = 0.3 \tag{2.1}$$

was deemed satisfactory.

2.6 Selection of programming language

All the algorithms and methods discussed in this study have been simulated numerically by software modules written in GNU Octave.

GNU Octave is a high-level programming language primarily intended for numerical computations, which was developed by John W. Eaton in 1988 at the University of Wisconsin-Madison. It is named after Octave Levenspiel and it was initially used as a companion to the reaction engineering course at the university (Eaton 2013). The main advantage of GNU Octave over MATLAB is that it is free software, under the terms of the GNU Public License. GNU Octave is used in academia and industry, and was used for various projects on the massive parallel computer at the Pittsburgh supercomputing centre.

The following considerations were taken into account when GNU Octave was chosen for the implementation of the programs:

- License - for the future development of the programs, not having to purchase a license to run the software was considered an advantage.
- Support and documentation - GNU Octave has an active support group and the language is very well documented.
- Capabilities - GNU Octave has been built with MATLAB compatibility in mind, as an Open Source, free alternative to it.

The following software modules have been written for the numerical simulation chapters and can be found on the attached CD:

1. For the numerical simulation of the Zhang et al. slug flow model: *zhang.m* in the *Zhang Model* folder .
2. For global sensitivity analysis: *main.m, myfunction.m, randorient.m, screeningplan.m, screeningplot.m* in the *Sensitivity Analysis* folder.
3. For the optimisation of the Zhang et al. model through history matching: *zhangoptim.m and zhangcc.m* in the *Optimisation* folder.
4. For the Zhang et al. model simplification: *zhangsimple.m* in the *Simplification* folder.

The Zhang et al. (2000) slug flow model was run on the considered data bank and the prediction errors were calculated. The results are shown in Appendix A, Table A.7. The results constitute the starting point and the benchmark for all the other algorithm predictions.

2.7 Summary

This chapter analysed the thinking process that led to the choice of research subject. A multi-layered diagram is presented to show different research paths

that result from the graphical representation of the thinking process. The choice of three of these research paths leads to the project objectives. These three research objectives are:

1. Analysis of the effects of using alternative empirical correlations on the mechanistic model's calculation algorithm.

2. The investigation whether the use of a choice of empirical correlations leads to mechanistic model optimisation.

3. The verification whether the use of a selection of empirical correlations, in conjunction with sensitivity analysis, leads to a simplified version of the mechanistic model.

Although the theoretical framework presented in the following chapters applies to any mechanistic model, the Zhang et al. (2000) mechanistic slug flow model has been selected. The numerical simulation of the optimisation and simplification concepts in Chapters 9 and 10 has been applied to a subset of 95 slug flow data points. These constitute a subset of the TUFFP data bank.

The algorithms used in the simulation have been implemented as scripts in GNU Octave, a high-level language, primarily intended for numerical computations.

Chapter 3

Research motivation and significance

3.1 Introduction

Chapter 2 provided an insight into the process of selecting the subject for this research, by using critical thinking. However, this does not provide an answer to the question of the research motivation and industry significance of the subject. This chapter will provide a list of business and research drivers that created a demand for this investigation.

3.2 Project motivation and significance

As onshore oil and gas reservoirs gradually come to the end of their lifecycle, exploration has moved offshore in the search for new reservoirs. Various crude oil production forecasts currently exist. Until not long ago there was a general agreement that there would be a decline in forthcoming years, as shown in a May 2009 forecast in Figure 3.1 (Eriksen 2009). The Ultimate Recoverable Reserves (URR) shown by the red line in Figure 3.1 forecast a clear decline worldwide, mainly due to big declines in Russia, Norway, the UK, Mexico and Saudi Arabia where the production allegedly peaked in 2005.

However, new studies have shown that this is not entirely true. It seems that peak oil has not happened and it is unlikely to happen for a very long time. The Hubbert's peak theory states that fossil fuel production in a given region over time follows a bell-shaped curve with a decline in the future. Studies published by Harvard University show that a new oil boom has begun (Maugeri 2012). This is partly true because of the new investment in unconventional oil, especially shale oil. What is certain though, is that conventional vertical land well drilling has almost reached an end, and oil and gas exploration has either moved offshore or adopted the newer technologies of horizontal drilling and fracking for shale oil.

A new set of challenges come along with the move to offshore exploration. Flow models, most of them empirical, could not predict the fluid flow behaviour any longer, due to the very different flow conditions. For example, low subsea temperatures at even relatively low pressures may lead to the formation of gas hydrates (Figures 3.2 and 3.3), which is not the case for onshore production

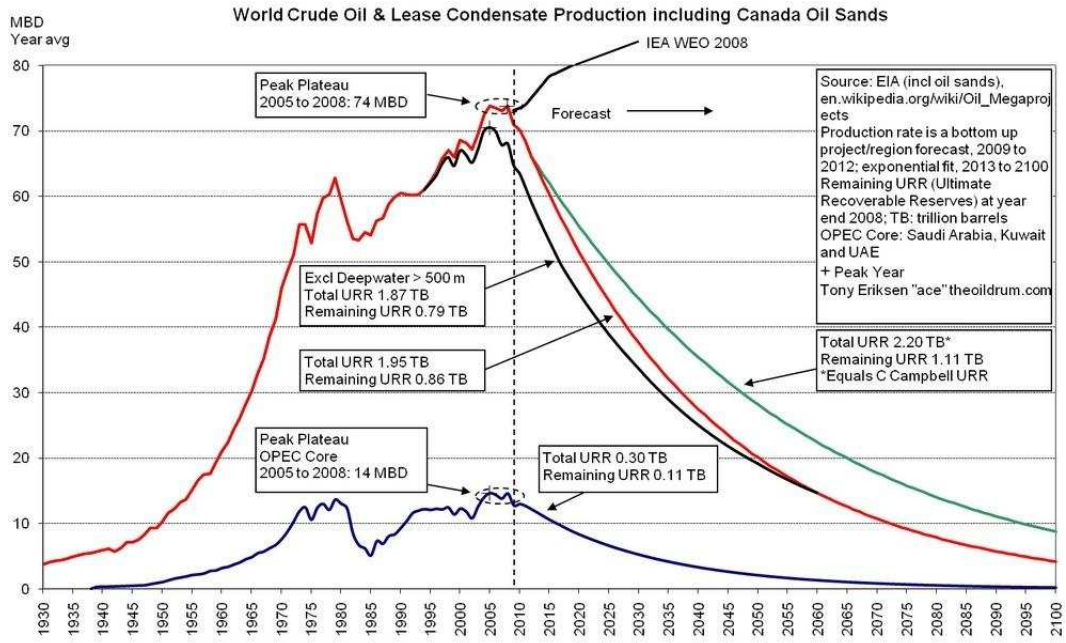


Figure 3.1: World crude oil production forecast (Eriksen 2009).

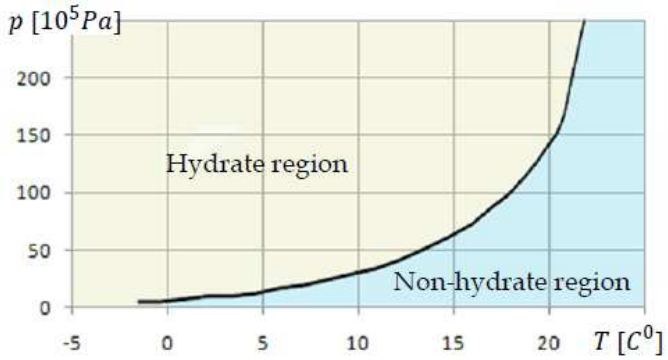


Figure 3.2: Gas hydrate formation curve as a function of pressure and temperature (Bratland 2009).

facilities.

Traditional models needed a re-assessment, hence a new approach was taken by the scientific community and mechanistic models became increasingly the preferred choice, mainly because extrapolation of the empirical flow models led to inaccurate results in the new offshore field conditions. However, mechanistic models use empirical correlations. It is in this context that the analysis of the influence of empirical correlations on the mechanistic models was considered important.

To this end, this study selected the Zhang et al. (2000) mechanistic slug flow model. This model used a set of empirical correlations, whose replacement with alternative correlations will be analysed. There are many reasons why this endeavour is worthwhile:

1. In the years since these unified models were released, new empirical correlations have been developed. A lot of the existing correlations were developed in test facilities with fluid mixtures such as air and water, or

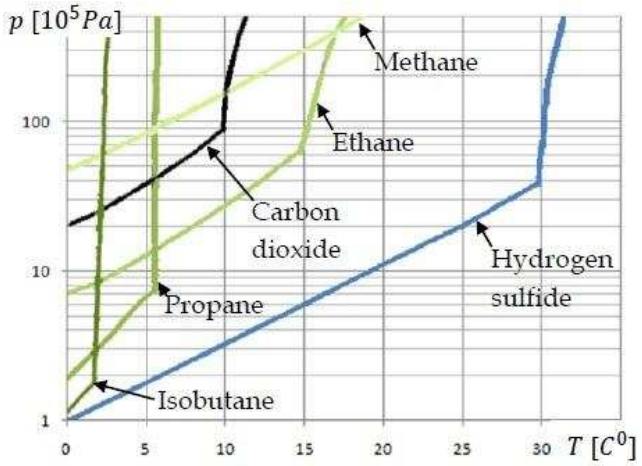


Figure 3.3: Gas hydrate formation for different pure components (Bratland 2009).

air and kerosene, and used small diameter smooth vinyl or glass pipelines. However, the offshore transportation trunklines are 20"-40" in diameter and made out of steel. The fluid mixtures are very different from the ones used in the development of most of the empirical correlations. This raises the question whether these new correlations would significantly improve the output of the mechanistic flow model. In other words, the question is whether any improvement to the accuracy of the flow model is made by continuously **updating the empirical correlations with new, more accurate ones, as they become available**.

2. The design engineer is faced with the difficult task of choosing the right empirical correlation in the process of sizing the equipment for gathering, pumping, transporting and storing the two-phase hydrocarbon mixtures. Woldesmayat and Ghajar (2007) presented a paper where 68 empirical correlations for void fraction were compared, out of hundreds of such correlations. The literature abounds in reviews of empirical correlations. These studies normally compare a large number of correlations and focus on the accuracy of these correlations on estimating a certain model parameter, such as friction factor (Shoham and Xiao 1991; Garcia et al. 2007; Aziz and Ouyang 1995; Spedding, Cole, and Donnelly 2004), or void fraction (A.Woldesmayat and Ghajar 2007). Considering that mechanistic models have several such empirical parameters, the number of empirical correlation combinations that could be used in the model increases exponentially. This could potentially mean that the engineer has to consult a very large number of empirical correlations to make the right decision. One could argue that nowadays simulation software incorporates all this complexity. However, in most cases only a very limited number of models and correlations are being considered by the software and **without prior knowledge and deep understanding of which particular empirical correlation each flow model is using, the engineer may blindly choose a model**. For example, as shown in Figure 3.4 the *Pipesim* software offers a multitude of choices for vertical flow. However, the engineer will need to know that *Duns and Ros*, *Hagedorn and Brown*, *Orkiszewski*, *Mukherjee and Brill* are

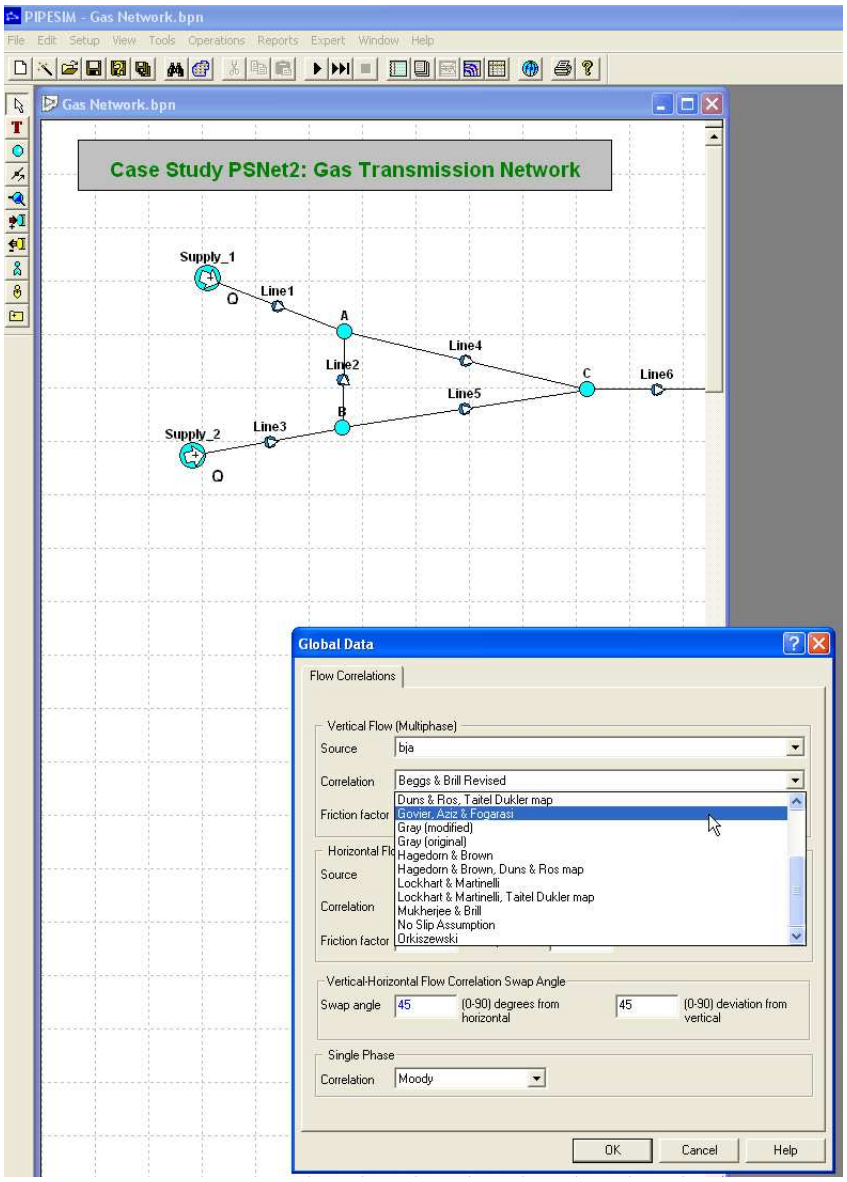


Figure 3.4: Options for choosing a flow model in the commercial software application *Pipesim*.

empirical flow models, unlike *Govier, Aziz and Fogarasi*. If one were to choose the latter to model the flow, the immediate question would be what empirical correlations does this flow model use and to what extent could they be applied to the field operations? As this is not shown by the software, not only would the engineer need to know all these details for all the flow models included in the application, but also the eventual updates to the models, such as the difference between the *Gray original* and the *Gray modified* flow models. To emphasize this complexity even further, one of the most common restrictions of empirical correlations is that they are flow regime dependent or only apply for a certain range of inclination angles (horizontal and near horizontal, vertical or upward inclined). This means that from the start, they can only be applied to a certain flow regime and this needs to be known prior to their application.

These arguments show that an off-the-shelf software application encapsulates the complexity of the flow in a non-transparent manner and the user has the ultimate responsibility of doing the extensive research to check the model's applicability to the flow conditions. **This study is an insight into the process of choosing the right correlation for the data, and shows the framework for incorporating this complexity into the software application**.

3. It is important to quantify **the extent to which empirical correlations influence the output of the flow model**. Only when this is correctly identified, do we know that **further research in the area related to the parameter estimated by the empirical correlation is worthwhile.** Model sensitivity analysis relative to a certain empirical parameter allows us to **focus where it matters on finding better and more accurate empirical correlations**. In other words, if we aimed to improve the accuracy of the mechanistic model we would look at the empirical parameter that has the greatest influence on the model accuracy, and try to find a better empirical correlation for it. The opposite would happen with the least sensitive empirical parameters, which due to their low influence would not need to have their empirical correlation replaced, because rough estimates would do as good a job as accurate ones.

4. As progress in science leads to a deeper understanding of the physical flow phenomena, mechanistic flow models and their calculation algorithms need to be re-visited, with the view of changing their use of empirical correlations. **Flow parameters that were estimated initially by empirical correlations in the original model may now be available to be calculated mechanistically**. From this point of view, the whole algorithm needs to be changed to allow for the calculation equations of these parameters to be an integral part of the model. This process should only be undertaken if the real benefit is a more accurate model, which is not detrimental to numerical stability.

5. A general consensus exists amongst the petroleum production engineers that the empirical correlations used in the field have to be selected according to the crude oil properties. Research (Hanafy et al. 1997) also emphasized that the empirical correlations should be chosen to the meet the geographical, regional properties of the production fluid. Empirical correlations used in the production fields of the North Sea cannot be used in the Gulf of Suez, Western Desert or Sinai production fields. **There is a worldwide requirement for all the production fields to replace empirical correlations in mechanistic models to match the regional, local data, corresponding to the local production fluid composition and flow conditions**.

6. Mechanistic models use empirical correlations as closure relationships. This is the case for well-known mechanistic models (Zhang et al. 2000, 2003c; Ansari et al. 1994; Petalas and Aziz 2000; Gomez et al. 2000; Bonizzi, Andreussi, and Banerjee 2009). Due to these correlations, **the models**

become data-sensitive and there is no guarantee that the model can be applied successfully to new field data.

7. The current, established software used in petroleum production (PipeSim, IPM Prosper) offers a data-matching option, which allows the user to check which one of a selected list of models performs best on a certain test data set. Given that each one of these models uses empirical correlations, the software determines the best match by determining each model's predictions and comparing them to the measured data, by using one empirical correlation set per model only, the one that comes with the model. No attempt to optimise the model by trying different correlation sets is made. **This study will show the advantage of attempting the extra step of optimising the mechanistic model by trying different empirical correlation sets on each model, before determining which mechanistic model fits best**.

3.3 Summary

This chapter has outlined industry and research challenges that this study attempts to address. These challenges are:

1. Updating an existing mechanistic model with new empirical correlations.
2. Replacing an empirical component with a mechanistic component within a mechanistic model.
3. Determining the most sensitive empirical parameter for the mechanistic model. This helps with the selection of an accurate empirical correlation for the data. In terms of future research work, this determines which parameters need better experimental estimation.
4. Finding the best empirical correlation set for the data.
5. Adjusting an existing mechanistic model to include geographically-dependent empirical correlations.

The empirical component of mechanistic models makes them data-sensitive. The main deficiency of the current approach of using mechanistic models is that they do not adjust their empirical correlations to match the data set particularities. This argument will be elaborated and demonstrated in the following chapters.

Chapter 4

Literature review

4.1 Introduction

Chapter 2 introduced the project objectives and specified that the Zhang et al. (2000) mechanistic slug flow model will be used throughout this investigation as a case study to support the discussion around using alternative empirical correlations in mechanistic models. To better understand the evolution of flow models and the latest advancements, an introduction into the field of empirical and mechanistic models is essential.

The empirical correlation replacement framework discussed in the following chapters relies on a systemic view of the mechanistic models. Therefore, this chapter will review the relevant literature to show how the optimisation process discussed in Chapter 9 relates to current systemic approaches.

4.2 Empirical correlations

Empirical models are based on direct observation through a large number of experiments, measurements and extensive data collection and analysis. The most commonly used empirical models in petroleum production are Dukler et al. (1964) and Beggs and Brill (1973), for flow in pipelines, and the Hagedorn and Brown (1964) and Duns and Ros (1961) correlations in wellbores, to name just a few of them. The predictive performances of these models are poor and errors of $\pm 30\%$ are common.

The conventional approach to pressure gradient calculation, before the proliferation of computer technology, was to use pre-prepared pressure traverse curves. These curves were prepared using a correlation such as Hagedorn and Brown (1964), for different values of the parameters: inside pipe diameter d, liquid flow rate q_L, water fraction f_W, average flowing temperature T and oil, gas and water specific gravities. The calculation is graphical and involves the selection of the diagram that corresponds to the well completion and fluid parameters. Such a diagram is presented in Figure 4.1. This approach hides the complexity of the multiphase flow in the correlation used to develop the pressure traverse curves.

The early models were not entirely empirical and still considered the mass and momentum conservation equations as their theoretical foundation. The pressure

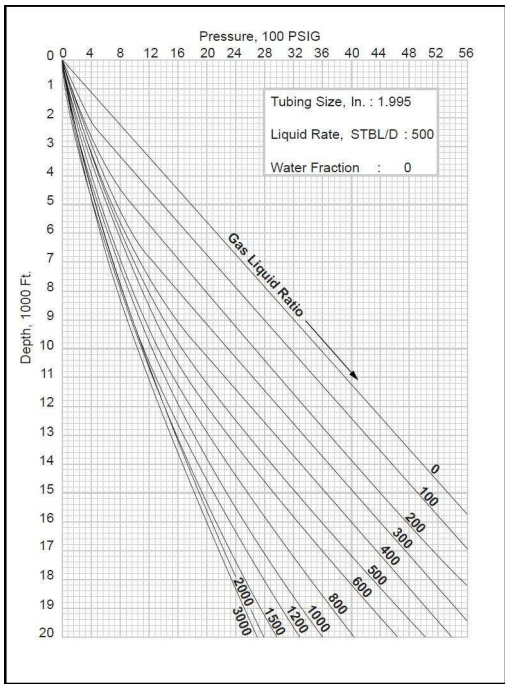


Figure 4.1: Pressure traverse curve used for the calculation of pressure drop (Heriot Watt 2009).

gradient equation derived from the continuity and momentum equations is:

$$-\alpha_k \frac{\partial P_k}{\partial x} + \left.\frac{dP}{dx}\right|_{\text{due to friction}} - \underbrace{\alpha_k \rho_k g \sin\theta}_{\text{due to elevation}} - \underbrace{\alpha_k \rho_k v_k \frac{\partial v_k}{\partial x}}_{\text{due to change in velocity}} = 0 \qquad (4.1)$$

where:

A_{pipe}	=	pipe cross-sectional area.
A_k	=	cross-sectional area occupied by phase k.
α_k	=	$\frac{A_k}{A_{pipe}}$ = volume fraction of phase k.
ρ_k	=	density of phase k.
v_k	=	superficial velocity of phase k.
g	=	gravitational acceleration.
θ	=	pipe inclination angle.
P_k	=	external pressure applied on phase k in the control volume.
$\frac{dP}{dx}$ due to friction	=	pressure gradient of phase k due to friction from other phases and pipe wall.

Equation 4.1 is called the **mechanical energy balance** and is derived in Appendix D. The values of the terms in Equation 4.1 change with the inclination angle, so that for vertical flow the elevation pressure drop prevails and for horizontal flow the frictional losses prevail. The acceleration term becomes non-negligible at high velocities. The contribution of the different pressure drop terms in the total pressure drop is illustrated in Table 4.1. The acceleration term was not included, because many of the early models did not include the term at all.

The complexity of the Equation 4.1 lies in the fact that the elevation pressure drop depends on the density of the two-phase mixture, which in turn depends

Table 4.1: Typical pressure loss values in the production string (Heriot Watt 2009).

Pressure Loss Term		Pressure loss due to elevation [%]	Pressure loss due to friction [%]
Well Orientation	(vertical) well model	85-98	2-15
	(horizontal) pipeline model	0-30	70-100

on the liquid holdup. Also, the frictional losses depend on the two-phase friction factor. However, these terms change with the flow regime. Most of the correlations for pressure drop and liquid holdup are based on these considerations. Nevertheless, they differ in their approach regarding the flow regimes, the two-phase mixture density and the gas slippage. From this point of view, a possible classification would be (Brill and Beggs 1991):

1. **Homogenous no-slip model**. This does not take into account the slip or the flow regimes. The two phase mixture is treated as a pseudo-single phase with an average mixture velocity and fluid properties. The averaging factor is the non-slip liquid holdup. The gas and liquid are assumed to be travelling at the same velocity and the density is calculated based on the input gas-liquid ratio. Equation 4.1 is thus greatly simplified and the only matter requiring further consideration is the friction factor.

2. **Slip considered, but no flow regime consideration**. Since the gas slippage is taken into account, a calculation method is provided to predict the liquid holdup and the friction factor at different locations in the pipe. These calculation methods apply without consideration of the flow regime.

3. **Slip considered, flow regime considered**. Unlike the previous categories, the models in this category have to determine the flow regime first and once this is determined, an appropriate calculation method for the pressure gradient and liquid holdup is provided.

To add to the complexity related to the use of a correlation, some correlations are only available for vertical, others for horizontal and very few for inclined flow conditions. Table 4.2 lists some of the well known early models and their classification (Brill and Beggs 1991).

Comparative research for two-phase and three-phase correlations has been done. Spedding (2006) reviewed two-phase flow correlations for horizontal flow and showed that correlations such as Lockhart-Martinelli had errors of $\pm 50\%$, while Beggs and Brill had $\pm 20\%$. The notable result was that some two-phase correlations that performed poorly on two-phase flow prediction, gave better results when applied to three-phase flow.

Most of the experimental research in the area of multiphase flow used test facilities with plastic pipes of relatively small diameter. However, offshore oil and gas production facilities use trunklines with diameters of between 20" and

Table 4.2: The vertical flow correlation classification.

Vertical Flow Correlation	Classification
Poettmann and Carpenter	1
Baxendell and Thomas	1
Fancher and Brown	1
Hagedorn and Brown	2
Duns and Ross	3
Orkiszewski	3
Aziz, Govier and Fogarasi	3
Chierici, Ciucci and Sclocehi	3
Beggs and Brill	3

40" (Tulsa University 2008). This is one of the common flaws of the empirical multiphase flow models, which for large pipe diameters exhibit "high diameter errors" (Liejin et al. 2009; Bratland 2009; Newton, Bhardwaj, and Behnia 1992). In the early 80s, explorations on the North Slope in Alaska resulted in decisions to transport oil and associated gas in large diameter flow lines, mainly for economic considerations. This required use of empirical correlations, which were mostly based on small diameter pipes. In most cases the prediction of liquid holdup was poor (Brill et al. 1981).

4.3 Mechanistic modelling

Mechanistic models are based on an understanding of the behaviour of a system's components, generally through the application of fundamental laws such as mass, momentum and energy conservation. As this study uses the Zhang et al. (2000) mechanistic slug flow model, a review of the most important mechanistic flow models is necessary.

Figure 4.2 shows schematically the structure of a mechanistic model and provides a graphical interpretation of different research paths. Most of the models currently used in the petroleum industry are one-dimensional, two-phase and do not take into account the interphase mass transfer as a result of phase behaviour. However, more complex models exist, but very few consider the entire complexity of the multiphase flow. The initial approach in the modelling of multiphase flow was to determine the pattern transition criteria and then develop calculation models for the hydrodynamics and heat transfer, for each one of the patterns.

Taitel et al. (1980) presented a mechanistic model for the flow-pattern transitions in upward two-phase flow and identified the mechanisms and transition boundaries between them. Barnea et al. (1982) extended the applicability of the model to inclined flows. A few years later, in 1987, Barnea combined all the flow pattern prediction models into one unified model for the flow pattern transition prediction (Barnea 1987). Barnea's unified model is valid for inclination angles from $-90°$ to $+90°$. The term "unified" referred to the fact that, unlike previous research, the model was now valid for a range of pipe inclinations and not just "wellbore", which had been used typically for near-vertical flow, or "pipeline", for near-horizontal flow.

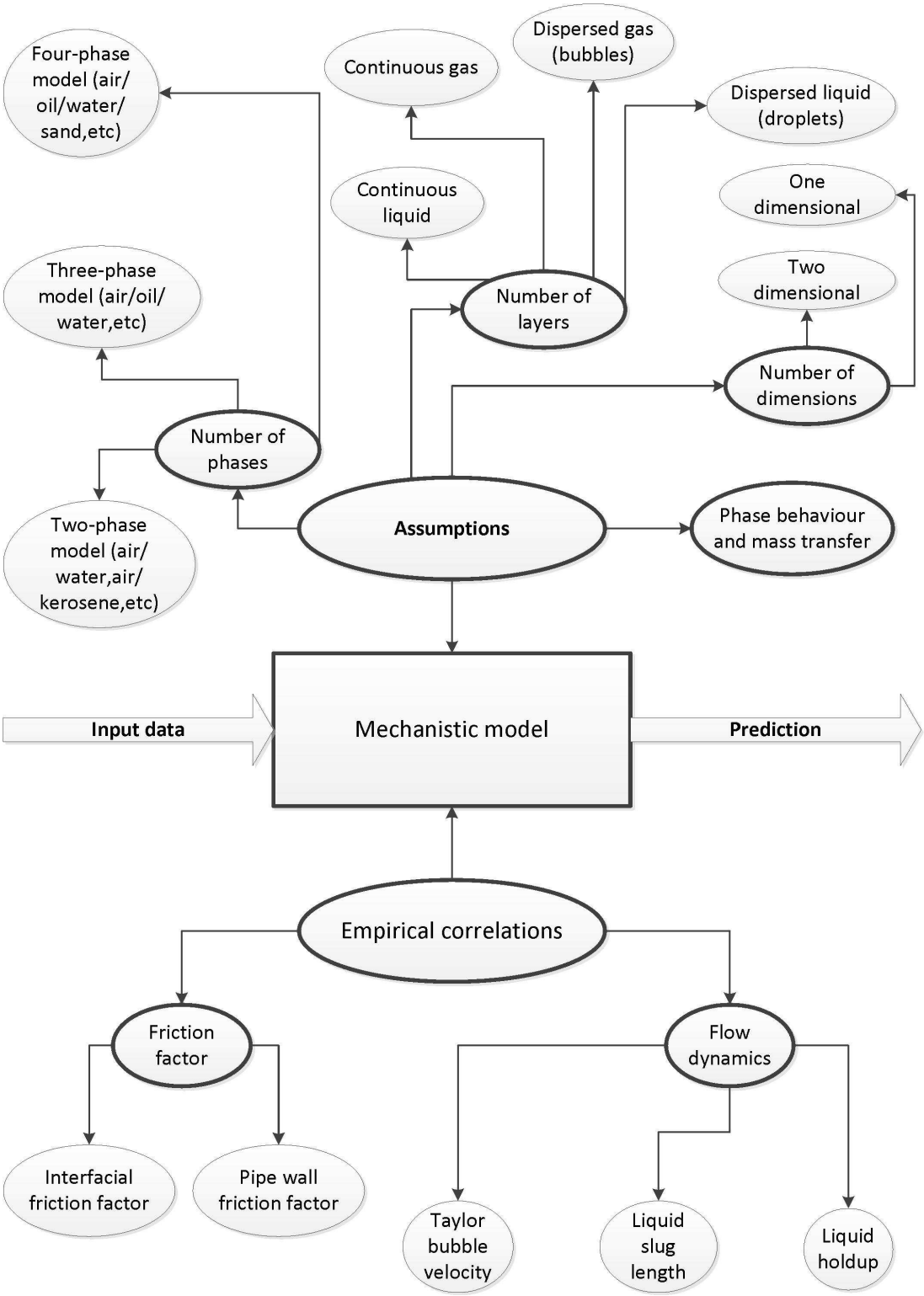


Figure 4.2: The structure of a mechanistic flow model - both the assumptions made and the empirical correlations are important, integral parts of the model.

The mechanistic models developed over the last four decades or so can be classified as wellbore, pipeline or unified.

4.3.1 Pipeline models

Mechanistic models appeared in the early 70s with the works of Agrawal et al. (1973) and Taitel and Dukler (1976). These models are applicable for horizontal and near-horizontal flow conditions, namely $\pm 10°$. The models are one-dimensional and even though two-dimensional models have been developed in the meantime by Shoham and Taitel (1984), Issa (1988) and Hall (1992), the industry is still using one-dimensional models. Due to the pioneering work of Dukler and Hubbard, and because of its durability, a description of the model is given in Appendix E.5.

4.3.2 Wellbore models

These models are applicable not only for vertical flow, which rarely occurs in the industry, but for off-vertical as well, for inclination angles between 60° and 90°. Specific models for slug flow were presented by Fernandes et al. (1983) and then Sylvester (1987) (Shoham 2006). Annular flow in wellbores was modelled by the Alves et al. (Shoham 2006). A recent model for two-phase flow in vertical annuli was proposed by Zhang et al. (2010).

4.3.3 Unified models

New unified models were developed by Felizola and Shoham (1995), who presented a unified slug flow model, valid for the entire range of angles of inclination between 0° and +90° (Shoham 2006). This model is included in Appendix E.6. Ansari et al. (1994) presented a comprehensive model for the prediction of upward two-phase flow patterns and included calculation methods for the pressure drop and liquid holdup for bubble, slug and annular flow. Petalas and Aziz (2000) developed a unified mechanistic model with a wide range of applicability, from downward, to horizontal and upward flow. Over the last decade an even more unified approach has been taken and mechanistic models have changed to incorporate a unified flow pattern prediction model and separate unified models for the different flow patterns. This was the approach taken by (Zhang et al. 2003c; Zhang and Sarica 2006). This model will be further analysed as a case study for the theoretical development in this thesis. Therefore, the base model is described in Appendix E.7. The model was extended by the researchers at Tulsa University to three-phase flow (gas-oil-water) (Tulsa University 2008).

Most of the models mentioned so far are two-phase models. Three phase flow (oil, water, gas) was studied experimentally by Acikgogz et al. (1991), who gave a classification of the three-phase flow regimes. Khor et al. (1997) analysed the three-phase stratified flow and determined flow holdups. Cengizhan et al. (2007) have extended the previous work of Acikgogz and identified and classified gas-oil-water flow patterns, by using the flow facilities at Tulsa University.

An interesting approach was taken by Bonizzi, Andreussi and Banerjee (2009) who presented a one-dimensional model that does not require that the flow

regimes be specified and does not use flow regime dependent closure relationships. The model is valid for slow-transients in near-horizontal pipes. Four "fields", or "layers" are considered in the equations: continuous liquid, dispersed liquid, continuous gas and dispersed gas.

Research in the area of mechanistic flow algorithms and sensitivity analysis on the empirical parameters has been fairly scarce and very few notable papers can be mentioned (Sarica, Zhang, and Wilkens 2011). **Despite of a fairly large number of papers conducting reviews of empirical correlations of certain parameters**, such as friction factor (Shoham and Xiao 1991; Garcia et al. 2007; Aziz and Ouyang 1995; Spedding, Cole, and Donnelly 2004) and liquid volume fraction (Gregory, Nicholson, and Aziz 1978; Eissa 2009), **none of them seems to have gone the extra step of analysing what happens to the model prediction if all these correlations are applied consecutively**.

4.4 Combined mechanistic and empirical modelling

Flow modelling is just one particular case of system modelling. At this more general level, the mechanistic models represent process behaviour through a set of equations, based upon fundamental conservation laws and scientific knowledge of the process. The shortcomings of the mechanistic models are that, in order to compensate for the lack of sufficient understanding and knowledge of the relevant physical phenomena, or ability to solve complex mathematical problems, they make assumptions and simplifications. Mechanistic models allow a certain degree of extrapolation outside the region for which experimental data confirms the model's predictions with a certain level of accuracy. The mechanistic approach to modelling is called "white box" modelling.

The opposite of this approach is the so-called "black box", empirical modelling. The empirical approach extracts knowledge directly from the operating data and is fitted to the data according to a minimisation criterion. This is usually based on the minimisation of the residuals, which are the norm of the differences between prediction and observation values. Due to the method used to derive them, a large number of input-output data points are required for these models to lead to reliable and accurate predictions. Empirical models do not permit extrapolation outside the data domain they were developed for and tested against.

The approach of combining both "white box" and "black box" modelling is called "grey box" or "hybrid" modelling. Combining the two paradigms creates models that address the issues of both worlds. The mechanistic model is improved to work outside the area where existing scientific laws have been confirmed, and the empirical model is extrapolated outside its region of validity by the fundamental conservation laws.

The combination of the two approaches can be performed in serial or parallel arrangements:

1. In the serial approach, the empirical module is fed with operating data and provides estimates for the empirical parameters to the mechanistic model, as shown in Figure 4.3. This configuration was used by Psichogios and Ungar (1992) for modelling the process of fermentation. The empirical module

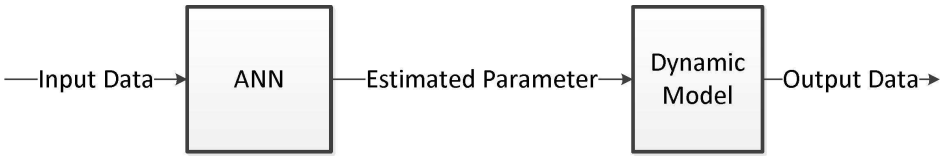


Figure 4.3: Structure of a serial hybrid model (Psichogios and Ungar 1992).

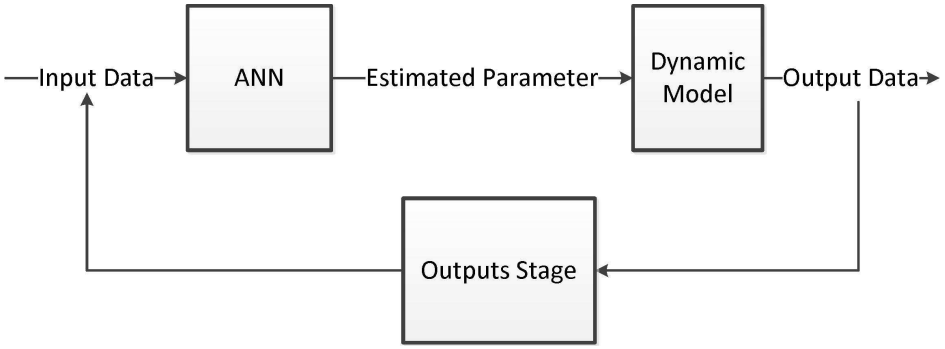


Figure 4.4: Structure of a serial with feedback hybrid model (Schubert et al. 1994).

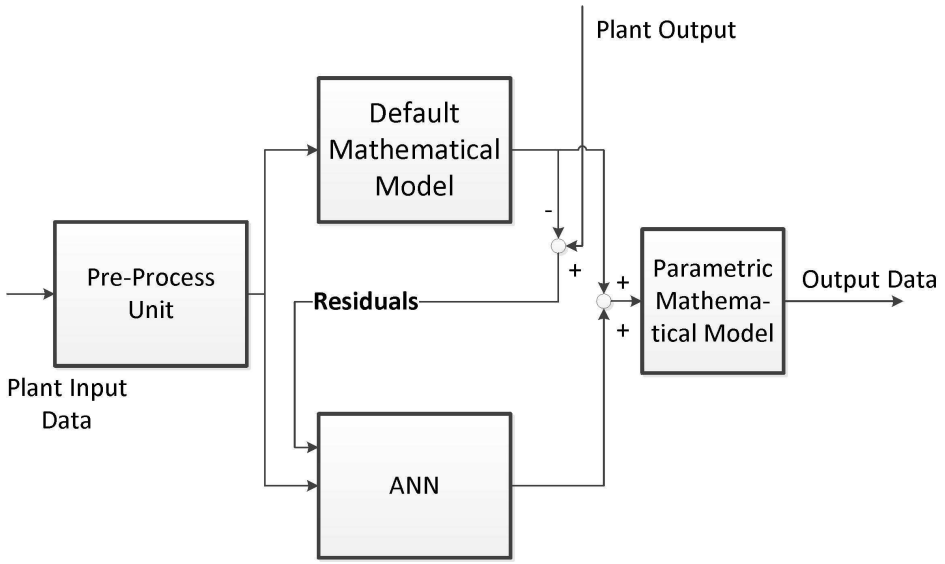


Figure 4.5: Structure of a parallel hybrid model (Thompson and Kramer 1994).

used an Artificial Neural Networks (ANN) module in this case. In order to improve the parameter estimation, the configuration can be improved with feedback to the empirical module (Figure 4.4). This approach was used by Schubert et al. (1994) for yeast cultivation process modelling.

2. In the parallel approach, both the empirical and the mechanistic models are fed the same data, as illustrated in Figure 4.5. The role of the empirical module is different in these configurations, as it aims to forecast the "residuals", the corrections that have to be added to the mechanistic model predictions to estimate the process output. Thompson and Kramer (1994) applied such a structure to model a penicillin fermentation plant.

Belmiro et al. (2004) used a hybrid model based on ANN with the structure in Figure 4.6. The model works in two phases: training and forecasting. In the

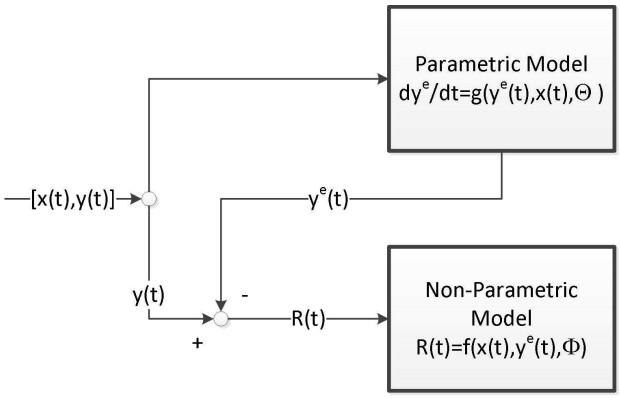


Figure 4.6: Structure of a Belmiro et al. hybrid model during the training phase (Belmiro, Saraiva, and Pantelides 2004).

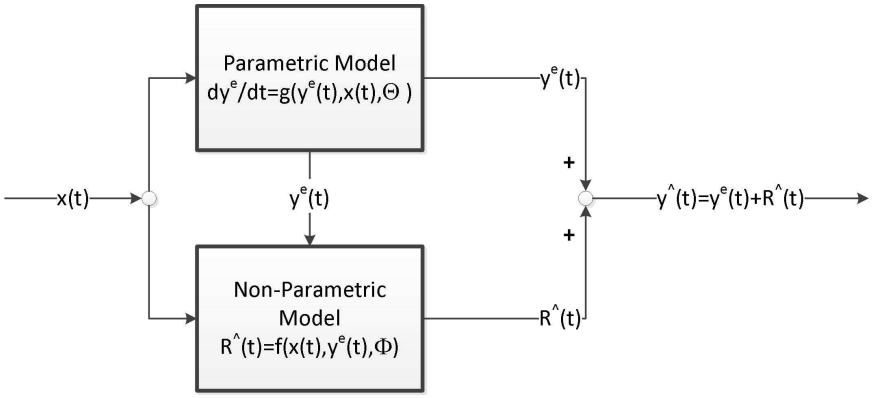


Figure 4.7: Structure of a Belmiro et al. hybrid model during the forecasting phase (Belmiro, Saraiva, and Pantelides 2004).

training phase in Figure 4.6 the inputs x are fed to the mechanistic model, which calculates the estimates $\dot{y^e} = g\left(y^e, x, \Theta\right)$, where Θ is a vector of mechanistic model parameters and y^e is the model prediction of the "true", or measured process output y. The empirical module is fed with the residuals $R = y - y^e$. Thus, the residuals have the form:

$$\hat{R} = f\left(x, y^e, \Phi\right) \tag{4.2}$$

where f is a local-continuous space of functions and Φ is a vector of parameters optimised in order to minimise the square norm of the difference $\hat{R} - R$. The space of functions f depends on the choice of empirical modelling technique, for example piecewise linear polynomials for Multivariate Adaptive Regressive Splines (MARS), sigmoidal basis functions for ANN. (Belmiro, Saraiva, and Pantelides 2004). In the forecasting phase in Figure 4.7, the mechanistic module predictions y^e and the inputs x are fed to the empirical module, which predicts the residuals R. The model prediction is then calculated by adding the mechanistic module estimate and the residual:

$$\hat{y} = y^e + \hat{R} \tag{4.3}$$

The model used in this study is also a two-phase model, like that of Belmiro et al., but instead of using ANN for the empirical module, this study uses a set of

empirical correlations. Consequently, the space of functions f used in the model, consists of piecewise functions characterising sets of empirical correlations. The proposed hybrid model has a parallel structure, similar to the one used by Belmiro et al., but with some changes due to differences in the process modelled.

4.5 Summary

This chapter presented a review of the relevant empirical and mechanistic flow models used in petroleum production optimisation, with an emphasis on mechanistic models. Such an introduction was necessary, because it outlines the Zhang et al. (2000) mechanistic model in the context of other flow models. Different relevant approaches to mechanistic modelling are also presented. They are essential for the understanding of the hybrid modelling and optimisation topics discussed in Chapter 6 and 9. The "training phase" concept used in the Belmiro and al. model will be used for hybrid modelling.

Chapter 5

Zhang et al. mechanistic slug flow model

5.1 Introduction

The Zhang et al. slug flow model was selected as a case study to validate the theoretical framework for empirical correlation replacement. Therefore, a presentation of the model calculation algorithm is necessary.

Chapters 9 and 10 will discuss model optimisation and simplification. In order to compare the performance of the modified model with that of the original model, statistical measures are used. A presentation of these measures is therefore also essential.

This chapter will present:

- A list of terms, definitions and the mathematical symbols used in slug flow modelling in Section 5.2.
- The Zhang et al. (2000) and (2003) calculation algorithms in Section 5.3. The models are derived in Appendix E.7 in the context of the fundamental flow equations.
- Flow model statistical measures in Section 5.5.

5.2 Slug flow terminology

Figure 5.1 shows the structure of a slug unit. The following notation will be adopted throughout this study:

Lengths:

d	=	pipe internal diameter.
L_U	=	length of the slug unit.
L_S	=	length of the liquid slug zone.
L_F	=	length of the liquid film zone.
L_M	=	length of the mixing zone.

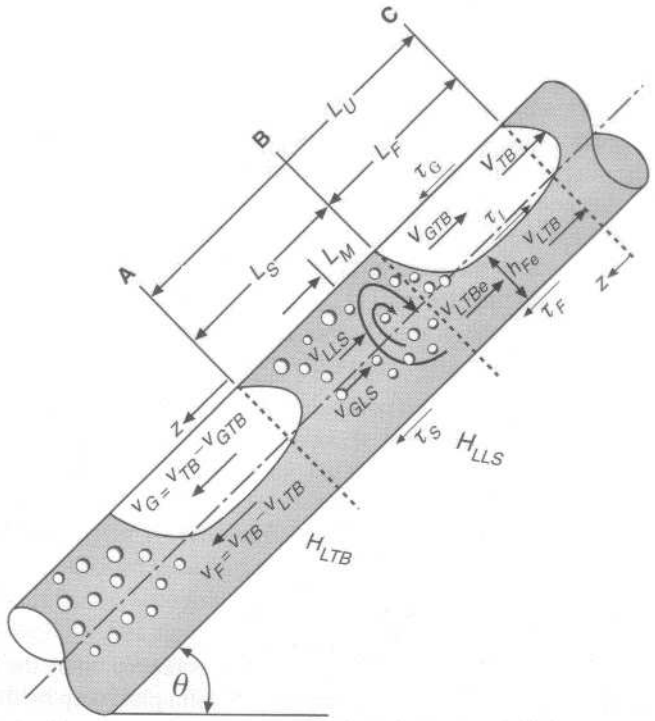


Figure 5.1: Detailed description of a slug unit (Shoham 2006).

Liquid holdups:

H_{LLS} = liquid holdup in the liquid slug zone.
H_{LTB} = liquid holdup in the liquid film zone.
H_{GLS} = gas void fraction in the liquid slug zone.
H_{GTB} = gas void fraction in the liquid film zone.

Velocities:

v_{LLS} = absolute velocity of the liquid slug.
v_{GLS} = absolute velocity of the dispersed bubbles inside the liquid slug zone.
v_M = slug mixture velocity.
v_{LTB} = absolute velocity of the liquid film under the Taylor bubble in the liquid film zone.
v_{GTB} = absolute velocity of the gas inside the liquid film zone. This is not the velocity of the Taylor bubble.
v_{TB} = absolute velocity of the Taylor bubble inside the liquid film zone.
v_F = relative velocity of the liquid film in the liquid film zone, with respect to the Taylor bubble velocity.
v_G = relative velocity of the gas in the liquid film zone, with respect to the Taylor bubble velocity.

Friction factors:

f_{LW} = liquid-wall friction factor.
f_{GW} = gas-wall friction factor.
f_i = interfacial friction factor.

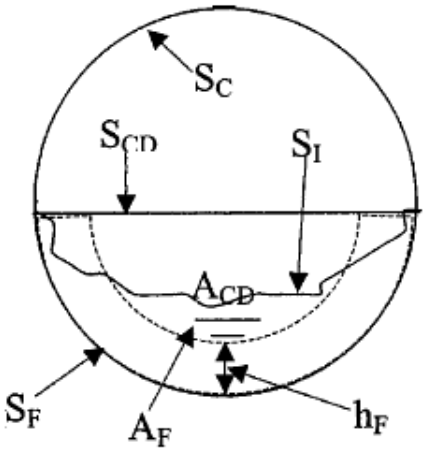


Figure 5.2: Geometrical parameters in slug flow (Zhang et al. 2003c).

Fluid properties:

ρ_L = liquid density.
ρ_G = gas density.
μ_L = liquid viscosity.
μ_G = gas viscosity.
σ = surface tension.

Geometrical parameters:

These parameters refer to Figures 5.1 and 5.2.

θ = pipe inclination angle.
A_{pipe} = pipe cross-sectional area.
Θ = wetted wall fraction. It is the ratio between the liquid wetted wall length measured along the pipe circumference, and the pipe circumference.
S_F = liquid wetted wall arc length.
S_G = gas contacted wall arc length.
S_{CD} = chord between the highest liquid level points.
S_i = interfacial perimeter.
A_F = cross-sectional liquid flow area.
A_G = cross-sectional gas flow area.
A_{CD} = cross-sectional area bounded by the wetted wall and its chord.

Shear stresses:

τ_{GW} = shear stress at the gas-wall interface.
τ_{LW} = shear stress at the liquid-wall interface.
τ_S = shear stress of the gas-liquid mixture at the wall interface.
τ_i = shear stress at the gas-liquid interface.

5.3 Zhang et al. (2000) and (2003) models

5.3.1 Assumptions

The Zhang et al. (2003) model is a unified mechanistic flow model; that is, it covers all the flow regimes, including slug flow. The slug flow component of the model assumes a fluid mixture of gas, liquid and droplets entrained in the gas core. This is an improvement on the Zhang et al. (2000) slug flow model, which only considered gas and liquid.

As this study investigates the use of various empirical correlations in slug flow modelling, and in order to compare the Zhang et al. (2000) slug flow model with the 2003 unified model, the following assumptions had to be made:

- As the study compares the two slug flow models, only the slug flow component of the 2003 unified model will be used for comparison. The other model components of the 2003 model refer to different flow regimes and will not be considered.

- As the Zhang et al. (2000) slug flow model assumes only gas and liquid fluid mixtures, we will assume that there is no droplet entrainment in the slug flow component of the 2003 unified mechanistic model.

5.3.2 Calculation algorithms

The two calculation algorithms, Zhang et al. (2000) and Zhang et al. (2003), are presented together, because the only three differences are:

1. The 2003 version includes **Zhang step 9**, whereas the 2000 version does not. This is because the 2003 version included a more accurate correlation for H_{LLS} (Zhang et al. 2003a), whereas the 2000 version only relied on the initial estimation provided by the Gregory (1978) correlation.

2. The correlations for f_i in **Zhang step 14** are different. The 2000 version used the Cohen and Hanratty correlation for f_i, whereas the 2003 version used the combination Andritsos and Hanratty (1987) and Asali and Ambrosio (1984).

3. The 2003 version calculates parameter τ_{LW} with a slightly different expression in **Zhang step 15**.

The combined presentation of the algorithms helps understand the concepts in Chapter 7, where an algorithm for the placement of empirical correlations according to their functional dependencies is developed. By showing the 2003 version of the Zhang et al. calculation algorithm, there is a clearer indication of where the Zhang et al. (2003a) correlation for H_{LLS} and other similar correlations can be positioned.

The **model input parameters** are:
d, θ, v_{SL}, v_{SG}, ρ_L, ρ_G, μ_L, μ_G and σ.

Zhang step 1: Start with an estimated value for H_{LTB}.

Zhang step 2: Calculate L_S, v_M. L_S is known as a closure relationship. The **Zhang et al.** empirical correlation is used for the calculation of L_S:

$$L_S = \left(32\cos^2\theta + 16\sin^2\theta\right) d \tag{5.1}$$

$$v_M = v_{SL} + v_{SG} \tag{5.2}$$

Zhang step 3: Calculate Taylor bubble velocity v_{TB}:

$$v_{TB} = c_0 v_M + v_{drift} \tag{5.3}$$

where the flow coefficient $c_0 = 2$ for laminar flow and $c_0 = 1.2$ for turbulent flow. The Taylor bubble drift velocity v_{drift} is calculated with the **Bendiksen** (1984) correlation:

$$v_{drift} = 0.54\sqrt{gd}\cos\theta + 0.35\sqrt{gd}\sin\theta \tag{5.4}$$

for values of $\theta \in [0°, 90°]$.

Zhang step 4: Estimate H_{LLS} using the **Gregory** correlation:

$$H_{LLS} = \frac{1}{1 + \left(\frac{v_M}{8.66}\right)^{1.39}} \tag{5.5}$$

Zhang step 5: Calculate v_{LTB} and v_{GTB} by simultaneously solving:

$$\begin{cases} H_{LLS}\left(v_{TB} - v_M\right) & = H_{LTB}\left(v_{TB} - v_{LTB}\right) \\ \left(1 - H_{LLS}\right)\left(v_{TB} - v_M\right) & = \left(1 - H_{LTB}\right)\left(v_{TB} - v_{GTB}\right) \end{cases} \tag{5.6}$$

Zhang step 6: Calculate ρ_S:

$$\rho_S = \rho_G\left(1 - H_{LLS}\right) + \rho_L H_{LLS} \tag{5.7}$$

Zhang step 7: Calculate Re_S:

$$Re_S = \frac{\rho_S v_M d}{\mu_L} \tag{5.8}$$

Zhang step 8: Calculate f_S:

$$f_S = C{Re_S}^n \tag{5.9}$$

where:

$$C = \begin{cases} 1.0 & \text{for laminar flow} \\ 0.046 & \text{for turbulent flow} \end{cases} \tag{5.10}$$

$$n = \begin{cases} 1 & \text{for laminar flow} \\ 0.2 & \text{for turbulent flow} \end{cases} \tag{5.11}$$

Zhang step 9 (Only in 2003 version): The **Zhang et al.** correlation for H_{LLS} shown in Equation 5.12 is used to improve the accuracy of the predictions:

$$H_{LLS} = \frac{1}{1 + \frac{T_{sm}}{3.16\sqrt{(\rho_L - \rho_G) g \sigma}}} \tag{5.12}$$

The parameter T_{sm} has the form:

$$T_{sm} = \frac{1}{C_e}\left[f_S \frac{\rho_S {v_M}^2}{2} + \frac{d}{4}\frac{\rho_L H_{LTB}\left(v_{TB} - v_{LTB}\right)\left(v_M - v_{LTB}\right)}{L_S}\right] \tag{5.13}$$

and C_e is:

$$C_e = \frac{2.5 - |\sin\theta|}{2} \tag{5.14}$$

Zhang step 10: Calculate L_F and L_U by simultaneously solving:

$$\begin{cases} L_U v_{SL} & = L_S v_M H_{LLS} + L_F v_{LTB} H_{LTB} \\ L_U & = L_S + L_F \end{cases}$$

Zhang step 11: Calculate the wetted wall fraction Θ, using the **Grolman** correlation:

$$\begin{aligned} \Theta &= \Theta_0 \left(\frac{\sigma_W}{\sigma}\right)^{0.15} \\ &\quad + \frac{\rho_G}{(\rho_L - \rho_G)\cos\theta}\left(\frac{\rho_L {v_{SL}}^{0.25} d}{\sigma}\right)^{0.25}\left[\frac{{v_{SG}}^2}{(1 - H_{LTB})^2 g d}\right]^{0.8} \end{aligned} \tag{5.15}$$

Zhang step 12: Calculate the geometrical parameters S_F, S_G, S_i, A_F and A_G:

$$S_S = \pi d \tag{5.16}$$
$$S_F = \pi d \Theta \tag{5.17}$$
$$A_F = H_{LTB} A_{pipe} \tag{5.18}$$
$$S_G = \pi d - A_F \tag{5.19}$$
$$A_G = (1 - H_{LTB})\, A_{pipe} \tag{5.20}$$
$$S_{CD} = d \sin(\pi\Theta) \tag{5.21}$$
$$A_{CD} = \frac{d^2}{4}\left(\pi\Theta - \frac{\sin(2\pi\Theta)}{2}\right) \tag{5.22}$$
$$S_i = \frac{S_F\left(A_{CD} - A_F\right) + S_{CD} A_F}{A_{CD}} \tag{5.23}$$

Zhang step 13: Calculate the Reynolds numbers Re_{LW} and Re_{GW}:

$$Re_{LW} = \frac{4A_F v_{LTB} \rho_L}{S_F \mu_L} \tag{5.24}$$

$$Re_{GW} = \frac{4A_G v_{GTB} \rho_G}{(S_G + S_i)\, \mu_G} \tag{5.25}$$

Zhang step 14 (2000 version): Calculate the friction factors f_{GW}, f_{LW} and f_i:

$$f_{LW} = C_1 Re_{LW}^{-n_1} \tag{5.26}$$

$$f_{GW} = C_2 Re_{GW}^{-n_2} \tag{5.27}$$

$$f_i = 0.0142 \tag{5.28}$$

where the C_i and n_i factors $(i = 1, 2)$ are:

$$C_i = \begin{cases} 1.0 & \text{for laminar flow} \\ 0.046 & \text{for turbulent flow} \end{cases} \tag{5.29}$$

$$n_i = \begin{cases} 1 & \text{for laminar flow} \\ 0.2 & \text{for turbulent flow} \end{cases} \tag{5.30}$$

The interfacial friction factor f_i, as calculated in Equation 5.28, is using the **Cohen and Hanratty** correlation.

Zhang step 14 (2003 version): Calculate the friction factors f_{GW}, f_{LW} and f_i:

$$f_{LW} = C_1 Re_{LW}^{-n_1} \tag{5.31}$$

$$f_{GW} = C_2 Re_{GW}^{-n_2} \tag{5.32}$$

where the C_i and n_i factors $(i = 1, 2)$ are:

$$C_i = \begin{cases} 1.0 & \text{for laminar flow} \\ 0.046 & \text{for turbulent flow} \end{cases} \tag{5.33}$$

$$n_i = \begin{cases} 1 & \text{for laminar flow} \\ 0.2 & \text{for turbulent flow} \end{cases} \tag{5.34}$$

If $\boldsymbol{\Theta < 0.8}$ use the **Andritsos and Hanratty** correlation to calculate f_i. Otherwise, use the **Asali and Ambrosio** correlation to calculate f_i.

These correlations are used for stratified and annular flow, respectively. They are applied to the film zone of the slug unit, where the flow can be either stratified or annular.

Andritsos and Hanratty (1987): The correlation in Equation 5.35 applies to stratified flow, under the assumption of a flat gas-liquid interface.

$$f_i = f_C \left[1 + 14.3 {H_{LTB}}^{0.5} \left(\frac{v_{SG}}{v_{SG,t}} - 1\right)\right] \tag{5.35}$$

f_C is the friction factor of the gas in the core of the Taylor bubble.

$$v_{SG,t} = 5 \left[\frac{m}{s}\right] \left(\frac{\rho_{GO}}{\rho_G}\right)^{0.5} \tag{5.36}$$

ρ_{GO} in Equation 5.36 is the gas density at atmospheric pressure.

Asali and Ambrosio (1984): The correlation applies to annular flow.

$$f_i = f_G \left[1 + 13.8 {We_G}^{0.2} {Re_G}^{-0.6} \left(h_F^+ - 200\sqrt{\frac{\rho_G}{\rho_L}}\right)\right] \tag{5.37}$$

The parameter h_F^+ is the dimensionless thickness of the liquid film and is defined in Equation 5.38:

$$h_F^+ = \frac{\rho_G h_F v_C^*}{\mu_G} \tag{5.38}$$

$$v_C^* = \sqrt{\frac{\tau_i}{\rho_G}} \tag{5.39}$$

For a smooth pipe f_G is defined as in Equation 5.40:

$$f_G = 0.046 {Re_G}^{-0.2} \tag{5.40}$$

The Weber and Reynolds numbers are calculated with Equations 5.41 and 5.42:

$$We_G = \frac{\rho_G v_{SG}^2 d}{\sigma} \tag{5.41}$$

$$Re_G = \frac{\rho_G v_{SG} d}{\mu_G} \tag{5.42}$$

Equation 5.37 shows the dependency of f_i on h_F^+, which according to Equation 5.38 depends on v_C^*. This parameter in its turn, according to Equation 5.39 depends on τ_i, the interfacial shear stress. On the other hand, τ_i is defined as:

$$\tau_i = f_i \frac{\rho_G \left(v_{SG} - v_{SL}\right) \left|v_{SG} - v_{SL}\right|}{2} \tag{5.43}$$

This shows an interdependency, which from the point of view of the calculation of f_i can only be solved by assuming an initial value for f_i and then iterating until convergence is achieved, as in the following

algorithm:
Step 1 Assume $f_i = f_G$
Step 2 Calculate τ_i according to Equation 5.43.
Step 3 Calculate v_C^* according to Equation 5.39.
Step 4 Calculate h_F^+ from Equation 5.38.
Step 5 Calculate $f_{i,calculated}$ using Equation 5.37. If $\frac{|f_{i,calculated}-f_i|}{f_{i,calculated}} < \epsilon$ the algorithm has converged and $f_{i,calculated}$ represents the value of f_i. Otherwise, repeat from **Step 2**, by assuming $f_i = f_{i,calculated}$.

Zhang step 15: Calculate the shear stresses τ_{GW}, τ_{LW}, τ_S and τ_i:

$$\begin{aligned}
\tau_{GW} &= f_{GW}\frac{\rho_G {v_{GTB}}^2}{2} \\
\tau_{LW} &= f_{LW}\frac{\rho_L {v_{LTB}}^2}{2} \\
\tau_S &= f_S\frac{\rho_S {v_M}^2}{2} \\
\tau_i &= f_i\frac{\rho_G\left(v_{GTB}-v_{LTB}\right)\left|v_{GTB}-v_{LTB}\right|}{2}
\end{aligned} \tag{5.44}$$

The (2003) version is using a slightly different expression for τ_{LW}:

$$\tau_{LW} = \frac{3\mu_L v_{LTB}}{h_F} - \frac{\tau_i}{2} \tag{5.45}$$

where the h_F is calculated with:

$$h_F = \frac{2A_{pipe}H_{LTB}}{s_F + S_i} \tag{5.46}$$

Zhang step 16: Calculate the margin of error m_{err} of Equation 5.47:

$$\begin{aligned}
m_{err} \;=\; & \frac{\rho_L\left(v_{TB}-v_{LTB}\right)\left(v_M-v_{LTB}\right)}{L_F} \\
& - \frac{\rho_G\left(v_{TB}-v_{GTB}\right)\left(v_M-v_{GTB}\right)}{L_F} \\
& + \frac{\tau_{GW}S_G}{\left(1-H_{LTB}\right)A_{pipe}} - \frac{\tau_{LW}S_F}{H_{LTB}A_{pipe}} \\
& + \tau_i S_i\left(\frac{1}{H_{LTB}A_{pipe}} + \frac{1}{\left(1-H_{LTB}\right)A_{pipe}}\right) \\
& - \left(\rho_L-\rho_G\right)g\sin\theta
\end{aligned} \tag{5.47}$$

Zhang step 17: Check if the margin of error m_{err} is less than the allowable tolerance ϵ:

$$|m_{err}| < \epsilon \tag{5.48}$$

If Condition 5.48 is satisfied, go to **Zhang step 18**. Otherwise, recalculate H_{LTB} from Equation 5.47 so that m_{err} is zero in **Zhang step 16** and go back to **Zhang step 5** to reiterate the algorithm.

Table 5.1: Empirical parameters and the empirical correlations used in the Zhang et al. (2000) slug flow model.

Empirical parameter	Empirical correlation
v_{TB}	Bendiksen
H_{LLS}	Gregory
L_S	Zhang
Θ	Grolman
f_i	Cohen and Hanratty

Table 5.2: Empirical parameters and the empirical correlations used in the Zhang et al. (2003) slug flow model.

Empirical parameter	Empirical correlation
v_{TB}	Bendiksen
H_{LLS}	Gregory, Zhang
L_S	Zhang
Θ	Grolman
f_i	Andritsos and Hanratty / Asali and Ambrosio

Zhang step 18: Calculate the slug unit pressure gradient $\frac{dP}{dL}\big|_U$ and the slug unit liquid holdup H_{LU}:

$$\left.\frac{dP}{dL}\right|_U = -\frac{L_F}{L_U}\frac{\tau_{LW}S_F + \tau_{GW}S_G}{A_{pipe}} - \frac{L_S}{L_U}\frac{\tau_S S_S}{A_{pipe}} - \frac{L_F g \sin\theta}{L_U}\left[\rho_L H_{LTB} + \rho_G\left(1 - H_{LTB}\right)\right] - \frac{L_S}{L_U}\rho_S g \sin\theta \tag{5.49}$$

$$H_{LU} = H_{LLS}\frac{L_S}{L_U} + H_{LTB}\frac{L_F}{L_U} \tag{5.50}$$

The **model outputs** are predictions for the **slug unit pressure gradient** and **slug unit liquid holdup**. This study will only analyse the pressure gradient output. All the concepts and methods presented in this investigation can be applied to the H_{LU} output in a similar way.

Tables 5.1 and 5.2 list the **empirical parameters** used in the Zhang et al. (2000) and (2003) models and the empirical correlations used for their estimation.

To compare the Zhang et al. (2000) model with other models developed here, a series of statistical measures widely applied in the literature was used. They are described next.

5.4 An equation oriented approach to solving the Zhang et al. mechanistic model

Two main process modelling solution approaches exist in chemical engineering:

1. The Sequential Modular approach.

2. The Equation Oriented approach.

The modelling approach in Section 5.3.2 is Sequential Modular (SM). The Direct Substitution method:

$$y_{k+1} = F(y_k) \tag{5.51}$$

where y is the model output vector and k is the iteration number, is used to iteratively calculate the model output.

An alternative to solve the Zhang et al. model is to use an Equation Oriented (EO) approach. This approach considers the entire model as a collection of algebraic equations, which are solved simultaneously for all the unknowns. A common strategy for solving systems of non-linear equations is to use local linearisation techniques based on the Newton-Raphson root-finding method.

Another strategy to solve the system of equations using the EO approach takes into account the fact that most engineering problems are characterised by sparse matrices. This strategy simplifies the problem considerably by reducing the structural matrix associated with the set of equations to a lower block-triangular form. The blocks identify the sets of equations that need to be solved simultaneously. The procedure of partitioning the matrix into blocks uses techniques to determine the strong components of the directed graph associated with the system of equations. The lower block-triangular form of the structural matrix offers the advantage of allowing for an easy calculation of the variables through the Forward Substitution method. Despite the reduced complexity of the final system of equations, root-finding methods are still required, as some equations are non-linear. Numerous variations of this approach exist in the literature, depending on the method used to identify the strong components (Barton 1995; Sargent and Westerberg 1964; Barkley and Motard 1972; Duff 1981b, 1981a).

The Block Decomposition method is based on the strategy of using the sparse matrix block decomposition to simplify the set of equations associated with the model. This method applied to the Zhang et al. model provides some benefits over the Direct Substitution method. The following points describe the advantages and disadvantages of the two methods:

1. Unlike the Direct Substitution method, which is highly iterative and requires a computational time of $O(noi)$ (n = number of variables, o =average time to compute an output variable, i =number of iterations), the Block Decomposition method reduces the complexity of the set of equations to solving smaller size sets of equations. Generally, EO methods offer greatly reduced computational times for high values of n or when there is no prior knowledge of the solution range of the guess variable in the Direct Substitution method.

2. The Block Decomposition method involves mathematical algorithms to reduce the sparse matrix associated with the system of equations to its lower block-triangular form. Not all sparse matrices are reducible to this form, but in the case of the Zhang et al. model the matrix is already in this format. In general, EO methods may fail during the pre-processing steps that set up the solving method and further analysis is required to ensure the success of the pre-processing operations.

3. Similar to the Direct Substitution method, the root-finding method used by the Block Decomposition method also needs to provide an initial guess for variable x_1. If the Newton-Raphson method is chosen, the success of the algorithm may be jeopardised if the Jacobian becomes numerically singular at one or more iterations, or if the initial guess is not within the convergence region. However, other numerical root-finding methods exist to mitigate this risk (e.g. Levenberg-Marquardt method).

4. The code associated with solving the Zhang et al. model using the Block Decomposition method uses library functions for boolean matrix multiplication and root-finding methods. This would have further required matrix factorisation library functions if the structural matrix had not already been in lower block-triangular form. Generally, the EO approach requires a programming language with a rich library of mathematical functions.

5. In contrast to the SM approach, EO allows for model optimisation and tuning.

5.4.1 The effect of the empirical correlation replacement on the modelling approach

An empirical correlation replacement in a mechanistic model leads to a new set of equations. A subject that warrants clarification is the effect of the change of the set of equations pertaining to the model on the chosen simulation approach.

In the EO approach, as the strategy involves a level of abstraction between the set of equations to be solved and the method to solve them, theoretically the strategy should not be affected. However, initial values to start a root-finding method such as Newton-Raphson may be affected, as the initialisation value may have to change as the empirical correlation equations are being replaced with another set of equations.

In the SM approach, the topology of the directed graph changes every time an empirical correlation is replaced in the model. Regardless of the Block Decomposition method applied, Path Tracing methods (Sargent and Westerberg 1964; Barkley and Motard 1972) or Reachability matrix based methods (Steward 1965; Himmelblau 1966), the method is not affected by the change.

Later in Chapter 7, an algorithm to replace empirical correlations in mechanistic models is described. This algorithm assumes that the structural matrix of the model has been reduced to a lower block-triangular form. The algorithm is in essence an algorithm to permute the rows of the matrix based on functional dependencies. This algorithm can be applied to both SM and EO simulation approaches.

5.5 Statistical measures used for flow model comparison

The following collection of statistical measures was first used by Ansari et al. (1994). Since then, they have been used widely in petroleum production engineering for characterising flow model fit.

Let us assume a number K of flow models, whose predictions for pressure gradient are to be compared on the same data bank of n input points. For each flow model $k \in \{1, 2, \ldots, K\}$, the following six statistical measures are calculated as functions of the output y = pressure gradient:

1. Average percent error:

$$E_{1,k} = \left(\frac{1}{n} \sum_{i=1}^{n} pe_i \right) \times 100 \tag{5.52}$$

where:

$$pe_i = \frac{y_{predicted,i} - y_{measured,i}}{y_{measured,i}} \tag{5.53}$$

is calculated for each input data point $x_i, i = 1, 2, \ldots, n$.

2. Absolute average percent error:

$$E_{2,k} = \left(\frac{1}{n} \sum_{i=1}^{n} |pe_i| \right) \times 100 \tag{5.54}$$

3. Percent standard deviation:

$$E_{3,k} = \left(\sum_{i=1}^{n} \sqrt{\frac{(pe_i - E_{1,k})^2}{(n-1)}} \right) \tag{5.55}$$

4. Average error:

$$E_{4,k} = \left(\frac{1}{n} \sum_{i=1}^{n} e_i \right) \times 100 \tag{5.56}$$

where:

$$e_i = y_{predicted,i} - y_{measured,i} \tag{5.57}$$

is calculated for each input data point $x_i, i = 1, 2, \ldots, n$.

5. Absolute average error:

$$E_{5,k} = \left(\frac{1}{n} \sum_{i=1}^{n} |e_i| \right) \times 100 \tag{5.58}$$

6. Standard deviation:

$$E_{6,k} = \left(\sum_{i=1}^{n} \sqrt{\frac{(e_i - E_{4,k})^2}{(n-1)}} \right) \tag{5.59}$$

To compare the K models, a relative performance factor is calculated for each model:

$$\begin{aligned} F_{PR_k} &= \frac{|E_{1,k}| - |E_{1_{MIN}}|}{|E_{1_{MAX}}| - |E_{1_{MIN}}|} + \frac{E_{2,k} - E_{2_{MIN}}}{E_{2_{MAX}} - E_{2_{MIN}}} + \frac{E_{3,k} - E_{3_{MIN}}}{E_{3_{MAX}} - E_{3_{MIN}}} \\ &+ \frac{|E_{4,k}| - |E_{4_{MIN}}|}{|E_{4_{MAX}}| - |E_{4_{MIN}}|} + \frac{E_{5,k} - E_{5_{MIN}}}{E_{5_{MAX}} - E_{5_{MIN}}} + \frac{E_{6,k} - E_{6_{MIN}}}{E_{6_{MAX}} - E_{6_{MIN}}} \end{aligned} \tag{5.60}$$

where:

$$E_{j_{MIN}} = \min\{E_{j,k}\} \tag{5.61}$$

$$E_{j_{MAX}} = \max\{E_{j,k}\} \tag{5.62}$$

for all $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, 6$.

The minimum value:

$$F_{PR_{k_m}} = \min\{F_{PR_k}\} \qquad k = 1, 2, \ldots, K \tag{5.63}$$

corresponds to model k_m, considered the most accurate model of all the K models compared.

5.5.1 Flow model comparison requirements

To understand the meaning of all these parameters, we need to understand the requirements for a fair model comparison. For several flow models to be compared and ranked, the following issues need to be addressed:

1. The statistical measures should reflect the performance of the models for both **high and low value ranges**. It is important to know whether a pressure variance of 20 psi, for example, was associated with a 100 psi, or a 1000 psi pressure drop. Therefore, a fair comparison takes into account the value of the variance relative to the measured value. The model with the lower variance relative to the measured value is preferred. Measures pe, E_1, E_2 and E_3 address this requirement by calculating the average and standard deviation relative to the measured values.

   On the other hand, the introduction of the measured value in the statistical measures, although necessary, does not offer the complete information to allow model comparison. To understand this, one can easily imagine the situation described in Figure 5.3, where two models of variances:

   $$V_1 = pressure_{predicted,1} - pressure_{measured}$$
   $$V_2 = pressure_{predicted,2} - pressure_{measured}$$

   are compared. Assuming that the measured values are in the range:

   $$pressure_{measured} \in [100 \text{ psi } \ldots 1000 \text{ psi}] \tag{5.64}$$

   and that the variances have the values:

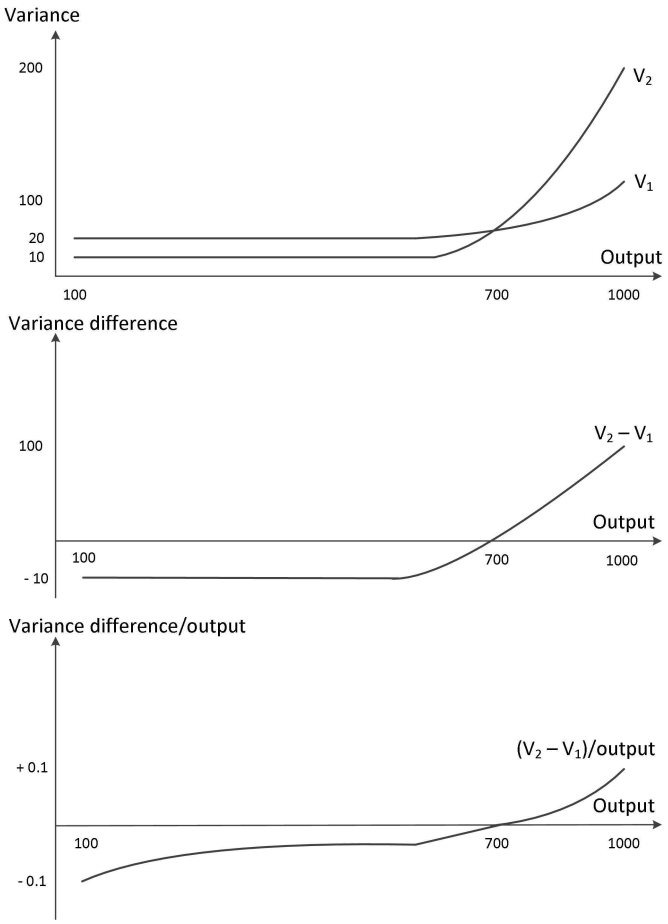


Figure 5.3: The variance difference relative to the measured output is not a good comparison parameter in the high range of output values.

(a) Variance $V_1 = 20$ psi for the range of values lower than 700 psi, while $V_2 = 10$ psi for the same range.

(b) Variance $V_1 = 100$ psi and $V_2 = 200$ psi for a measured output value of 1000 psi.

one could conclude the following:

- Model 2 performs better than model 1 for 67% of the output range.
- Model 1 performs better in the higher 33% of the output range.
- At the two ends of the output range, model 2 is better than model 1 by 10% for the minimum output value, whereas model 1 is better than model 2 by 10%, for the maximum value of the output range.

If no other statistical parameters independent of the measured value were considered, one may make the final decision that model 2 performed better.

However, this is a skewed view of the real situation, because model 1 performs only marginally worse than model 2, by 10 psi in the low to medium range. In the higher range of values, model 1 performs a lot better, having a variance of 100 psi, unlike model 2 whose variance is 200 psi. This distinction is not made if only the relative statistical measures are considered.

The difference between the pe based and e based measures is that the first are relative measures of the variance to the measured value $y_{measured}$, whereas the second are independent of the measured value. In light of this discussion, if pe, E_1, E_2 and E_3 were the only parameters used to compare the models, the comparison would not be fair in the high range of values.

To avoid this sort of situation, the statistical parameters e, E_4, E_5 and E_6 have been introduced, because they are independent of the measured values.

2. The model predictions should **follow the measured values closely**. The statistical factors to measure the "closeness" are E_1, E_2, E_4 and E_5. Some terms in the averages E_1 and E_4 might cancel out, depending on which side of the measured values the predictions fall. Low values of these measures show that the prediction follows the measured output closely and traversing this curve from opposite directions is a sign that the prediction follows the "true" output line. However, the absolute deviations on either side of the measured line have to be minimal, and this is quantified by the measures E_2 and E_5.

3. The mean of the model prediction values should follow the measured values curve as closely as possible. In other words, there should be **minimum scatter**. Parameters E_3 and E_6 respectively measure the standard deviation.

In the original paper that introduced the F_{RP} factor, Ansari et al. (1994) argue:

- "E_1 indicates the overall trend of the performance, relative to the measured pressure drop.
- E_2 indicates how large the errors are on the average.
- E_3 indicates the degree of scattering of the errors about their average value.
- E_4 indicates the overall trend, independent of the measured pressure drop.
- E_5 is also independent of the measured pressure drop and indicates the magnitude of the average error.
- E_6 indicates the scattering of the results, independent of the measured pressure drop.
- The minimum and maximum possible values for F_{RP} are 0 and 6, indicating the best and worst performances, respectively."

The parameters E_1, E_2 and E_3 depend on the magnitude of the measured value and are therefore suitable for model evaluations in the small range of values. Parameters E_4, E_5 and E_6 are appropriate for model evaluations at large values, because they are independent of the measured values (Majeed 1996).

An interesting result is obtained when only two models are compared. In this case, each one of the terms $\frac{E-E_{MIN}}{E_{MAX}-E_{MIN}}$ is either 1 or 0, respectively. This means that for the particular case of two model comparisons, the final F_{RP} score is always a positive integer.

5.6 Summary

This chapter introduced the slug flow nomenclature and presented both the 2000 and 2003 versions of the Zhang et al. calculation algorithm.

As this research study aims to establish optimised and simplified versions of the Zhang et al. model, statistical measures for flow model performance comparison were presented.

Chapter 6

Hybrid modelling of mechanistic flow models

6.1 Introduction

This chapter approaches the optimisation problem mentioned as a project objective in Chapter 2 by first taking a global view of mechanistic flow models. This view allows us to define the problem in the general context of hybrid modelling. Once defined, Chapters 7, 8 and 9 build on the theoretical foundations of this chapter and discuss the details of the optimisation process.

6.2 Overview of hybrid flow modelling

Most of the mechanistic flow models, including the Zhang et al. (2000) model selected in this study, have the structure shown in Figure 6.1. The "white box" feeds the model status parameter vector φ to the "black box", in order to provide a vector Φ containing estimates of the model parameters, which, for the lack of scientific knowledge, need to be calculated by empirical correlations. The Φ vector is fed back to the "white box" for the final calculation of the model estimate $\tilde{y}$.

For the particular case of the Zhang et al. (2000) model, whose calculation algorithm is described in Chapter 5, these vectors have the form:

$$\varphi = [H_{LTB}] \tag{6.1}$$

$$\Phi = [L_S \quad H_{LLS} \quad v_{TB} \quad f_i \quad \Theta] \tag{6.2}$$

Equations 6.1 and 6.2 are expressing the functional dependencies of L_S, H_{LLS}, v_{TB}, f_i, Θ in terms of model parameters. H_{LTB} is the only parameter used by the "black box" to provide estimates for the empirical parameters in Φ. Chapter 7 elaborates upon the functional dependency concept. The component parameters in Equations 6.1 and 6.2 were defined in Section 5.2.

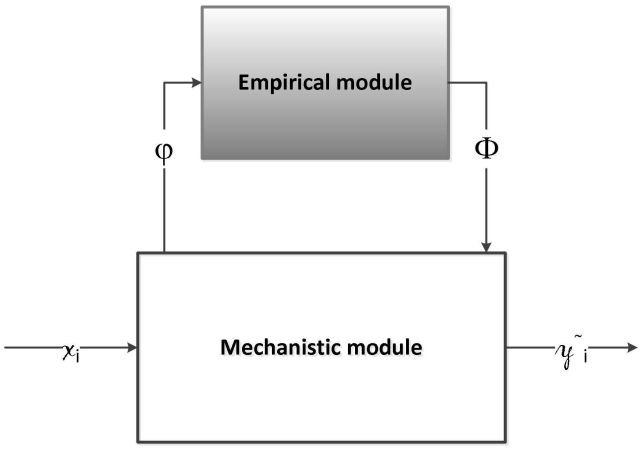


Figure 6.1: Most of the mechanistic flow models have a hybrid structure in which the empirical correlations provide estimations for the components in vector Φ, based on a status vector φ.

6.3 Definitions and notations

Suppose a set of known model input values:

$$X^{known} = \{x_1, x_2, \ldots, x_n\} \tag{6.3}$$

and the known measured values of the process:

$$Y^{true} = \{y_{t,1}, y_{t,2}, \ldots, y_{t,n}\} \tag{6.4}$$

such that $y_{t,i}$ is the "true" model value for the input x_i, $i = 1, 2, \ldots, n$. The term "true" refers to the observed, or measured values of the process. In reality there is no such thing as a "true" model, because no model can predict the real process outputs with perfect accuracy.

Definition The pairs $\langle x_i, y_{t,i} \rangle$, $i = 1, 2, \ldots, n$ are called the **training data set**.

Suppose there are r empirical parameters e_k, $k = 1, 2, \ldots, r$ in the mechanistic model. For each empirical parameter e_k, a set of alternative empirical correlations Ω_k may be used for its calculation:

$$\Omega_k = \{g_{k,1}, g_{k,2}, \ldots, g_{k,k_k}\}$$

where $g_{k,j}$ is the empirical correlation that may be used for the calculation of e_k.

For each input data point x_i, $i = 1, 2, \ldots, n$ the empirical parameters are calculated with their respective correlations $e_k(x_i) = g_{k,j_k}(x_i)$ where $j_k = 1, 2, \ldots, k_k$. For each input point x_i the empirical correlations in the sets Ω_k provide:

k_1 estimates for the empirical parameter e_1
k_2 estimates for the empirical parameter e_2
$\vdots$
k_r estimates for the empirical parameter e_r

We define the vector:

$$\Phi(x_i) = [e_1(x_i)\, e_2(x_i) \ldots e_r(x_i)] \tag{6.5}$$

containing estimations of the empirical parameters at point x_i for some combination of correlations $j_1, j_2, \ldots, j_r$.

The iterative process of selection of an empirical correlation $g_{k,j}$ from each set Ω_k leads to different values of the vector components $e_k(x_i)$ in $\Phi(x_i)$, depending on which empirical correlation was used for each empirical parameter. To denote this, we use the notation:

$$\Phi_l(x_i), \quad l = 1, 2, \ldots, L \tag{6.6}$$

to refer to:

$$\Phi_l(x_i) = [g_{1,j_1}(x_i)\, g_{2,j_2}(x_i) \ldots g_{r,j_r}(x_i)] \tag{6.7}$$

where:

$$\begin{aligned} j_1 &= 1, 2, \ldots, k_1 \\ j_2 &= 1, 2, \ldots, k_2 \\ &\vdots \\ j_r &= 1, 2, \ldots, k_r \\ l &= 1, 2, \ldots, L \end{aligned} \tag{6.8}$$

Despite Equation 6.5 suggesting $\Phi(x_i)$ is a vector, because each parameter e_k is calculated by a number of alternative correlations, $\Phi(x_i)$ is actually a set of vectors:

$$\Phi(x_i) = \{\Phi_1(x_i), \Phi_2(x_i), \ldots, \Phi_L(x_i)\} \tag{6.9}$$

where $L = k_1 \times k_2 \times \cdots \times k_r$.

The mechanistic model provides estimates of the true values. We use the notation:

$$Y_l^{estimates} = \{\tilde{y}_{1,l}, \tilde{y}_{2,l}, \ldots, \tilde{y}_{n,l}\} \tag{6.10}$$

where $\tilde{y}_{i,l}$ is an estimation of the "true" value $y_{t,i}$ for:

- The input x_i, $i = 1, 2, \ldots, n$.
- The empirical parameter vector $\Phi_l(x_i)$, $l = 1, 2, \ldots, L$.

The values $\tilde{y}_{i,l}$ are influenced by both mechanistic and empirical components in the model. We use the notation $\tilde{y}_{i,l}$ to mean:

$$\tilde{y}_{i,l} = \tilde{y}(\Phi_l(x_i), x_i) \tag{6.11}$$

where $i = 1, 2, \ldots, n$ and $l = 1, 2, \ldots, L$. Equation 6.11 shows that $\tilde{y}_{i,l}$ is calculated taking into account:

1. The empirical component $\Phi_l(x_i)$, which contains the empirical parameter values calculated with set l of empirical correlations, $l = 1, 2, \ldots, L$.
2. The mechanistic component, which depends on the input x_i.

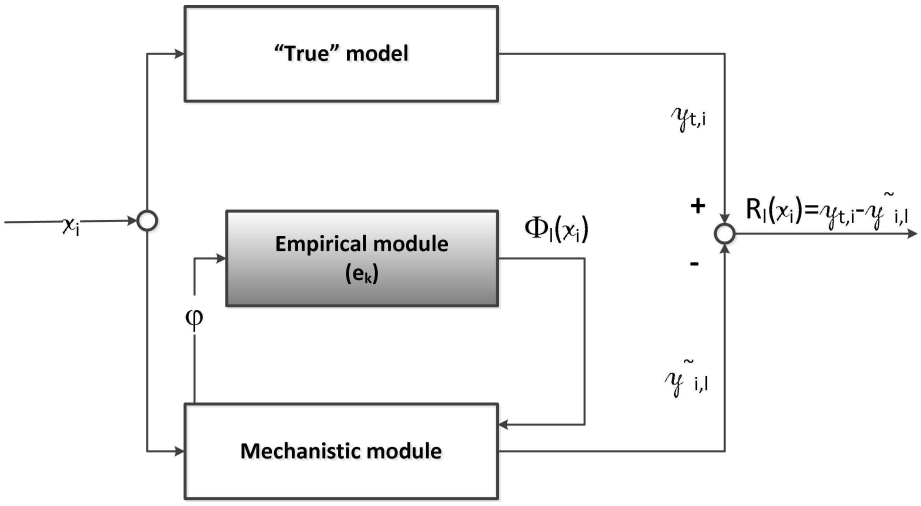


Figure 6.2: Structure of the hybrid model used during the training phase.

6.4 Optimisation method using alternative empirical correlation sets

This study proposes a **two stage** flow model optimisation:

1. The first stage is **training**. During this stage, a parallel configuration is proposed, in which the model is first "trained" on an input-output training data set, consisting of pairs $\langle x_i, y_{t,i} \rangle$, $i = 1, 2, \ldots, n$ with $x_i \in \Gamma$, where Γ is the input domain of discrete values, as shown in Figure 6.2.

   For the same discrete value x_i, one correlation from each of the empirical correlation sets $\Omega_1, \Omega_2, \ldots, \Omega_r$ in the "black box" is used to estimate the empirical parameters $e_1, e_2, \ldots, e_r$ respectively. As a different correlation is chosen from each set Ω_k each time, $L = k_1 \times k_2 \times \cdots \times k_r$ vectors $\Phi_l(x_i)$ containing empirical parameter estimations are generated. This was shown in Equation 6.7.

   The model estimation $\tilde{y}_{i,l}$ depends on the vectors $\Phi_l(x_i)$, as shown in Equation 6.11. The L values of $\tilde{y}_{i,l}$ are compared to the "true" model outputs $y_{t,i}$. The "true" model was included in Figure 6.2 for the sole purpose of outlining that the known process outputs, $y_{t,i}$ are used to calculate the residuals $R_l(x_i)$, which are the difference between the measured value of the process output $y_{t,i}$ and the model estimation $\tilde{y}_{i,l}$:

   $$R_l(x_i) = y_{t,i} - \tilde{y}_{i,l} \quad l = 1, 2, \ldots, L \tag{6.12}$$

   Let us assume that the minimum residual value for the input x_i:

   $$R_{min}(x_i) = \min\{R_l(x_i), l = 1, 2, \ldots, L\} \tag{6.13}$$

   is obtained for a certain combination of empirical correlations. We use the notation $O(x_i)$ to refer to the set of empirical correlations whose values

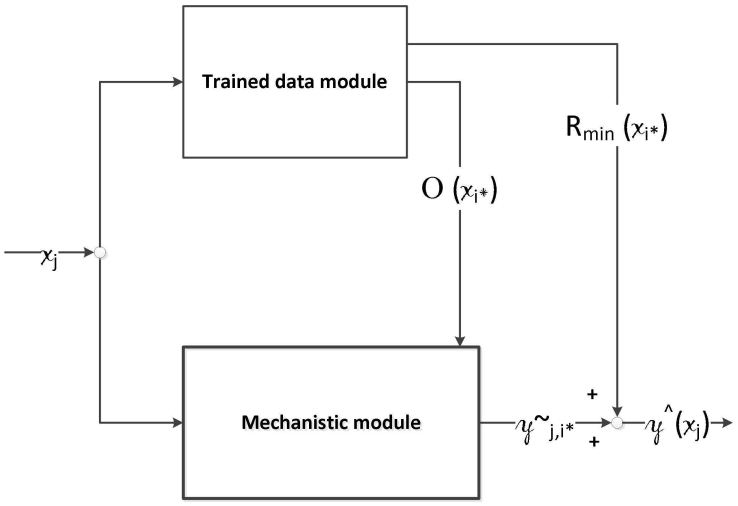


Figure 6.3: Structure of the hybrid model used during the prediction phase.

$\Phi_l(x_i)$ minimise the residual at point x_i:

$$\begin{aligned} O(x_i) &= \{g_{1,j_1^*}, g_{2,j_2^*}, \ldots, g_{r,j_r^*}\} \\ j_1^* &\in \{1, 2, \ldots, k_1\} \\ j_2^* &\in \{1, 2, \ldots, k_2\} \\ &\vdots \\ j_r^* &\in \{1, 2, \ldots, k_r\} \end{aligned} \tag{6.14}$$

The set of empirical correlations $O(x_i)$ is **optimal** for the input point x_i, because it ensures a minimum residual.

The 3-tuple $\langle x_i, O(x_i), R_{min}(x_i)\rangle$ is stored in a table during the training stage. At the end of the iteration, which goes through all the training data points $\langle x_i, y_{t,i}\rangle$ $i = 1, 2, \ldots, n$, a table is constructed. Each entry in the table associates a training input point x_i with an optimal correlation set and the minimum residual: $\langle x_i, O(x_i), R_{min}(x_i)\rangle, i = 1, 2, \ldots, n$.

2. During the second stage, **prediction**, the model uses the 3-tuple $\langle x_i, O(x_i), R_{min}(x_i)\rangle$ determined during the training stage, to predict the model output (Figure 6.3). This time, the model makes predictions on a different set of input values x_j for which the "true" output $y_{t,j}$ is not necessarily known. That is, the input data set is no longer the training data set. The model prediction $\hat{y}$ for an input point x_j is determined as follows:

$$\hat{y}(x_j) = \tilde{y}_{j,i^*} + R_{min}(x_i) \tag{6.15}$$

where:

$$\tilde{y}_{j,i^*} = \tilde{y}(O(x_{i^*}), x_j) \tag{6.16}$$

Equation 6.16 illustrates that the model estimation $\tilde{y}_{j,i^*}$ at a point x_j that is not part of the training data set takes into account:

(a) The optimal correlation set at a point x_{i^*} that is part of the training data set. An objective function $F(x_j, x_i)$, such as the distance between x_j and the input points in the training data set $x_i \in \Gamma, i = 1, 2, \ldots, n$, is defined and its values are calculated. The index i^* in Equations 6.15 and 6.16 is the index of one of the input points in the training data set, $x_{i^*} \in \Gamma$, that minimises the objective function F.

(b) The mechanistic component, reflected by the influence of x_j.

The interpretation of Equation 6.15 is that in order to make a prediction at an input point x_j:

(a) The "closest" input point x_{i^*} that is part of the training data set is determined. An objective function F is used to define the "closeness".

(b) Using the table generated in the training stage, the elements in the 3-tuple associated with x_{i^*}, the optimal correlation set $O(x_{i^*})$ and the minimum residual $R_{min}(x_{i^*})$ are determined.

(c) The model prediction is calculated at point x_j, using the optimal empirical correlation set for the "closest" point x_{i^*} and the minimum residual calculated for x_{i^*}.

Equation 6.15 ensures that the residuals are zero, if $x_j \in \Gamma$. That is, if x_j is in the training data set, the residuals are zero and $\hat{y}(x_j) = y_{t,j}$. For all the other values $x_j \notin \Gamma$, the residuals have the role of a correction factor.

If there are multiple points $x_{i_1^*}, x_{i_2^*}, \ldots, x_{i_M^*}$ that minimise the objective function F, there are M optimal correlation sets associated with these points. As only one model prediction exists, $\hat{y}(x_j)$, a further criterion has to be specified to select only one empirical correlation set from $O\left(x_{i_1^*}\right)$, $O\left(x_{i_2^*}\right)$, ..., $O\left(x_{i_M^*}\right)$. This subject will be discussed in detail in Section 9.3.4, as part of the optimisation of the Zhang et al. mechanistic model.

6.5 Summary

This chapter served both as an introduction and back reference for Chapters 7, 8 and 9, as the concepts developed in those chapters fit into the hybrid model structure and algorithm explained here.

The two stages of the optimisation process are discussed in detail in the following chapters:

- Chapter 7 shows the issues related to replacing empirical correlations with alternative correlations, as part of the **training** phase.
- Chapter 8 discusses the **training** phase of the optimisation procedure.
- Chapter 9 deals with the **prediction** phase.

Chapters 7, 8 and 9 describe the steps that need to be taken towards mechanistic model optimisation through empirical correlation set replacement, and are the theoretical building blocks of the optimisation framework proposed in this study.

Chapter 7

The effect of empirical correlations on the calculation algorithm of the mechanistic model

7.1 Introduction

To apply the concepts in Chapter 6 to model optimisation and simplification, the empirical correlations in the analysed mechanistic models have to be substituted by other correlation sets. This may involve changes to the model calculation algorithm, because each empirical correlation depends on a different set of input parameters, which need to be determined prior to the use of the correlation. Essentially, we cannot "unplug" an empirical correlation out of a mechanistic model and replace it with an alternative correlation, without first analysing whether the Calculation Algorithm (CA) has to be modified. This leads to the conclusion that different combinations of empirical correlations may lead to different CAs. Consequently, an algorithm for replacing empirical correlations in mechanistic models is required.

The algorithm has the main objective to determine the step in the CA of a mechanistic model where an empirical parameter can be placed, based on the empirical correlation used for its estimation. To reach this aim, a gradual approach was taken by:

- Defining the notation used - Section 7.2.
- Stating the problem - Section 7.3.
- Describing some of the supporting algorithms needed for Chapter 8 - Section 7.4.
- Describing the algorithm to replace empirical correlations in mechanistic models - Section 7.5.
- Demonstrating in detail the steps of the algorithm for the case study of the Zhang et al. (2000) slug flow model - Section 7.6.

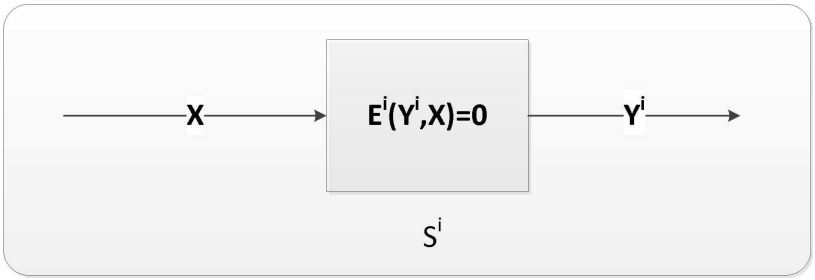


Figure 7.1: Step S^i in the CA finds the roots Y^i for the equations $E^i\left(Y^i, X\right) = 0$, $i = 1, 2, \ldots, N$.

7.2 Definitions and notations

Definition: A Calculation Algorithm (CA) is a sequence of independent calculation steps $S^i, i = 1, 2, \ldots, N$:

$$CA = \{S^1, S^2, \ldots, S^N\} \tag{7.1}$$

At each step S^i, a vector of parameters of unknown value $Y^i = (y_1^i, y_2^i, \ldots, y_n^i)$, $Y^i \in \mathbb{R}^n$ is calculated based on a vector of parameters of known value $X = (x_1, x_2, \ldots, x_m)$, $X \in \mathbb{R}^m$, according to a set of n functions $E^i = \{E_1^i, E_2^i, \ldots, E_n^i\}$, where $E_j^i : \mathbb{R}^{m+n} \to \mathbb{R}$, $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, n$. The set of equations:

$$\text{Step } S^i : \begin{cases} E_1^i\left(y_1^i, y_2^i, \ldots, y_n^i, x_1, x_2, \ldots, x_m\right) & = 0 \\ E_2^i\left(y_1^i, y_2^i, \ldots, y_n^i, x_1, x_2, \ldots, x_m\right) & = 0 \\ \quad \vdots & \\ E_n^i\left(y_1^i, y_2^i, \ldots, y_n^i, x_1, x_2, \ldots, x_m\right) & = 0 \end{cases} \tag{7.2}$$

is solved for Y^i at step S^i. The notation y_j^i refers to variable y_j calculated at step i in CA. This is illustrated in Figure 7.1.

Without any loss of generality we assumed real values for y_j^i and x_k, despite constraints that may apply to them (e.g. area is always positive, angle values have upper and lower bounds). Solving the system of Equations 7.2 allows the component variables y_j^i be expressed as a function f of the components of the input vector X, as in the following equation:

$$y_j^i = f\left(x_{j1}, x_{j2}, \ldots, x_{jp}\right) \tag{7.3}$$

where $\{x_{j1}, x_{j2}, \ldots, x_{jp}\} \subseteq \{x_1, x_2, \ldots, x_m\}$ is a subset of the input X.

Definition: The factors in the set $\{x_{j1}, x_{j2}, \ldots, x_{jp}\}$, that directly influence y_j^i are called **influencing factors** of variable y_j^i.

There are situations when the explicit form of function f is not known. However, this does not matter, because we are interested in outlining the functional dependency of variable y_j^i on $\{x_{j1}, x_{j2}, \ldots, x_{jp}\}$.

The execution of a CA is a sequence of steps $\{S^1, S^2, \ldots, S^k, \ldots, S^k, \ldots, S^N\}$, in which some of the steps may be repeated a number of times. The order of the steps is such that one step's inputs are outputs of various previous steps.

Definition: Suppose a functional dependency at step S^i in CA:

$$y_j^i = f\left(\ldots, x_{jl}, \ldots\right) \qquad l = 1, 2, \ldots, p \tag{7.4}$$

If the functional dependency in Equation 7.4 is not empty, there will always exist a variable y_k^s, calculated at step S^s, such that:

$$x_{jl} = y_k^s, \qquad s = 1, 2, \ldots, N. \tag{7.5}$$

A functional dependency is called **empty** when all its influencing factors are constants, input parameters, or a combination of these. We use the notation $f\left(\right)$ to denote an empty functional dependency.

To be part of a CA, the variables in the functional dependency in Equation 7.4 have to satisfy the condition in Equation 7.5. We will call Equation 7.5 the **step pre-requisites condition** for step S^i.

Definition: A mechanistic model is **properly ordered** if all the variables in the model are calculated at the minimum step where the pre-requisites condition is met.

Suppose the functional dependency $y_j^i = f\left(\ldots, x_{jl}, \ldots\right)$, where $x_{jl} = A_{pipe}$ and that y_j^i is the only variable in CA using A_{pipe} in its calculation. There are two possible approaches to placing the step where $A_{pipe} = \frac{\pi d^2}{4}$ is calculated:

1. At step S^{i-1} in preparation for step S^i, where y_j^i needs A_{pipe} as part of the step pre-requisites condition.
2. At any other step before S^{i-1}, such as the first step in CA.

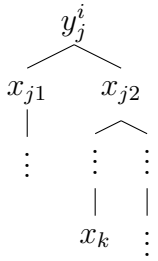
If A_{pipe} is calculated at the first step, the model is properly ordered, because variable A_{pipe} is placed at the minimum step in relation to the model input parameter d = pipe diameter.

The first approach at step S^{i-1}, although correct from the algorithmic point of view, does not make the model properly ordered as defined here.

Definition: Suppose a variable $x_k \in \mathbb{R}$ on which variable y_j^i depends, such that:

$$y_j^i = f\left(x_{j1}, x_{j2}, \ldots, x_{jp}\right) = f\left(f\left(\ldots f\left(\ldots x_k \ldots\right) \ldots\right)\right) \tag{7.6}$$

The functional dependency of y_j^i on a parameter x_k can be represented as a tree, where the functional dependencies are represented as edges and the influencing parameters as vertices:



The **dependency chain** of variable y_j^i is the sequence of vertices in the tree obtained by a Depth First Search (DFS) walk in the tree. We use the notation $D\left(y_j^i\right)$ for the dependency chain of variable y_j^i.

$D\left(y_j^i\right)$ is the set of the dependency chains of the directly influencing factors $x_{j_1}, x_{j_2}, \ldots, x_{j_m}$:

$$D\left(y_j^i\right) = \{D\left(x_{j_1}\right), D\left(x_{j_2}\right), \ldots, D\left(x_{j_m}\right)\} \tag{7.7}$$

Definition: We classify model variables, based on the type of parameters that exist in their dependency chain. Variable y_j^i is:

1. **Entirely non-empirical**, if the dependency chain does not contain any empirical parameters.
2. **Partially non-empirical**, if there is at least one empirical parameter in the dependency chain.
3. **Empirical**, if it is estimated by an empirical correlation.

Section 7.4.2 will describe a method for determining the dependency chain of a variable.

7.3 Problem statement

Let us assume a CA $= \{S^1, S^2, \ldots, S^N\}$ for a mechanistic model. The set of empirical parameters used in the model is $\{e_1, e_2, \ldots, e_r\}$. Suppose a set of empirical correlations Ω_k is used for the estimation of each empirical parameter e_k, such that:

$$\Omega_k = \{g_{k,1}, g_{k,2}, \ldots, g_{k,k_k}\}$$

where $g_{k,j}$ is the empirical correlation that may be used for the calculation of e_k, $k = 1, 2, \ldots, r$.

Problem 1: For a given CA $= \{S^1, S^2, \ldots, S^N\}$, find the permutation of steps:

$$\text{CA}^{new} = Permutation\left(S^1, S^2, \ldots, S^N\right) \tag{7.8}$$

associated with the replacement of the empirical correlation $g_{k,i}$ with another correlation $g_{k,j}$ from the same set Ω_k, $i \neq j$, $i, j = 1, 2, \ldots k_k$, so that the step pre-requisites conditions are still met. Both correlations, $g_{k,i}$ and $g_{k,j}$ calculate the same empirical parameter e_k.

The equivalent statement is to find the step in CA^{new} where an empirical parameter $e_j, j = 1, 2, \ldots, r$ can be placed, when the empirical correlation for $e_k, k \neq j$ is replaced by another empirical correlation in the same set Ω_k, $k = 1, 2, \ldots, r$.

Problem 2: Determine the position of the empirical parameters e_j, $j = 1, 2, \ldots, r$ to minimise the number of different CA^{new} generated by substituting one correlation at a time, using all the correlations in all the sets at least once.

7.4 Supporting algorithms for graph generation and functional dependency analysis

Suppose we intend to replace the empirical correlation for parameter e_k in CA. The method proposed in this study is based on converting the empirical correlation for the calculation of e_k from Equation 7.9:

$$e_k = g\left(x_1, x_2, \ldots, x_m\right) \tag{7.9}$$

to the functional dependency in Equation 7.10:

$$\begin{aligned} &e_k = f\left(n_1^{n_1}, \ldots, n_n^{n_n}, p_1^{p_1}, \ldots, p_p^{p_p}, e_1, \ldots, e_e\right) \\ &e_k \notin \{e_1, e_2, \ldots, e_e\} \end{aligned} \tag{7.10}$$

where:

g = empirical correlation.
$\{x_1, x_2, \ldots, x_m\}$ = set of influencing parameters.
$n_i^{n_i}$ = non-empirical parameters calculated at step S^{n_i} of CA.
$p_j^{p_j}$ = partially empirical parameters calculated at step S^{p_j} of CA.
e_l = empirical parameters.
$i = 1, 2, \ldots, n_n$, $j = 1, 2, \ldots, p_p$ and $l = 1, 2, \ldots, e_e$.
S^i is the $i-$th step in the first iteration of the CA.

Function f is not explicitly determined and is just a placeholder for expressing that parameter e_k depends on the factors on the Right Hand Side (RHS) of Equation 7.10.
Different notations are used for the steps in CA and CA^{new}:

- $S^i \in$ CA refers to a calculation step in CA.
- $Step^i \in$ CA^{new} refers to a calculation step in CA^{new}.

Equation 7.10 takes into account the three possible types of influencing parameters that are used in a functional dependency:

1. Entirely non-empirical.
2. Partially non-empirical.
3. Empirical.

Equation 7.10 allows us to determine the step in CA^{new} where the empirical parameter e_k can be placed, by considering the type of the influencing parameter:

1. Entirely non-empirical parameters $n_i^{n_i}$ can be placed after the maximum of all the steps of the entirely non-empirical parameters:

   $$Step^{e_k} \geq \max\left(S^{n_1}, S^{n_2}, \ldots, S^{n_n}\right) + 1 \tag{7.11}$$

   The value on the RHS of Condition 7.11 is constant, because $Step^{n_1} = S^{n_1}$, $Step^{n_2} = S^{n_2}$, $\ldots$, $Step^{n_1} = S^{n_1}$. This is because the position of the entirely non-empirical parameters does not change when a new empirical correlation is used in the algorithm.

2. Partially non-empirical parameters $p_i^{p_i}$ can be placed after all the steps where empirical parameters along the dependency chain are calculated. Suppose the tree representation of $D(e_k)$ and that $e_j, j \neq k$ is the first empirical parameter encountered along the dependency chain:

$$D(e_k) = D\left(p_j^{p_j}\right) = \ldots = D(e_j) \tag{7.12}$$

As correlations for e_j are alternated with others from the same set, the step of e_j in CA^{new} may change. Because none of the parameters between e_k and e_j are empirical, the position of e_k relative to e_j will always be the same:

$$Step^{e_k} - Step^{e_j} = \text{constant} \tag{7.13}$$

We can state that:

$$\begin{aligned} &Step^{e_k} \geq \max\left(Step^{e_j} + \Delta(e_k, e_j)\right), \\ &\quad j = \text{ index of all empirical parameters in } D(e_k) \end{aligned} \tag{7.14}$$

where:

$$\Delta(e_k, e_j) = S^{e_k} - S^{e_j} \tag{7.15}$$

is the relative distance in steps between the step where e_k is calculated and the step where e_j is calculated. As explained in the lead up to Equation 7.13, Δ is constant.

3. Empirical parameters e_i can be placed after the steps where $e_1, e_2, \ldots, e_e$ have been calculated. These steps may change position as the empirical correlations used for their calculation are being alternated. $Step^{e_1}, Step^{e_2}, \ldots, Step^{e_e}$ are not known yet, because Equation 7.10 has not been obtained for all the empirical parameters in the model yet. However we can state that:

$$Step^{e_k} \geq \max\left(Step^{e_1}, Step^{e_2}, \ldots, Step^{e_e}\right) + 1 \tag{7.16}$$

Remark: Empirical parameters are a particular case of partially non-empirical parameters that are obtained for $\Delta = 1$. Equation 7.16 can be obtained from Equation 7.14 for $\Delta = 1$. Therefore, when we calculate $Step^{e_k}$, we only need to classify the influencing factors in two categories:

1. **Entirely non-empirical**.
2. **Having an empirical component**, which includes empirical and partially non-empirical parameters.

To summarise these results, $Step^{e_k}$ can be expressed as the maximum between a constant and a variable component:

$$Step^{e_k} \geq \max \begin{cases} \textit{constant} \text{ determined by entirely non-empirical factors} \\ \textit{variable} \text{ determined by factors having an empirical component} \end{cases} \tag{7.17}$$

Condition 7.17 implies that in order to determine $Step^{e_k}$ we need to classify the influencing factors of e_k into the two categories and then apply Equations 7.11 or 7.14, accordingly. Consequently, we need to determine:

- The dependency chain $D(e_k)$.
- The influencing factor classification.

Although Condition 7.17 contains a certain level of uncertainty relative to the position of e_k, once Equation 7.10 is obtained for all the empirical parameters in the model, we can centralise the information and make a decision of where each empirical parameter can be placed.

As the dependency chain is a tree map of the variables and functional dependencies, we need to extend the tree concept to all the variables in the mechanistic model. Therefore, for the purpose of this analysis we will use a **graph representation** of CA. The **vertices** in the graph represent the **variables** and the **directed edges** show both a **functional dependency** and the **direction** of it.

Sections 7.4.1 and 7.4.2 illustrate the processes of:

- Building the graph.
- Determining the dependency chain of a variable.
- Determining the variable classification, necessary for Condition 7.17.

7.4.1 Building a graph representation of a mechanistic model

In order to build a graph representation of a mechanistic model, two operations are required:

- **Mapping model variables to graph vertices**. The result is a disconnected graph.
- **Connecting the vertices in the graph**. This links the vertices in the graph, based on the functional dependencies of the variables they represent.

1. **Mapping model variables to graph vertices**

   The calculation algorithm CA is mapped to a graph G by defining a map $M : CA \rightarrow G$ that associates a step $S^i \in CA$ to a set of n vertices:

   $$V^i = \{v_1^i, v_2^i, \ldots, v_n^i\} \tag{7.18}$$

   such that the name of each one of the Y^i components is assigned to a label of a vertex:

   $$M\left(S^i\right) = \{v_1^i, v_2^i, \ldots, v_n^i | Label\left(v_k^i\right) = Name\left(y_k^i\right), k = 1, 2, \ldots, n\} \tag{7.19}$$

   The $Name$ values used in Equation 7.19 refers to the text string values, rather than numeric value of y_j^i.

   The map is illustrated in Figure 7.2.

   The calculation steps in CA are parsed and sequentially mapped to graph vertices. The pseudocode is shown in Algorithm 7.1. At the end of the procedure, a graph $G = V^1 \cup V^2 \cup \ldots \cup V^N$ is obtained, as shown in Figure 7.3. The vertices of the graph are disconnected.

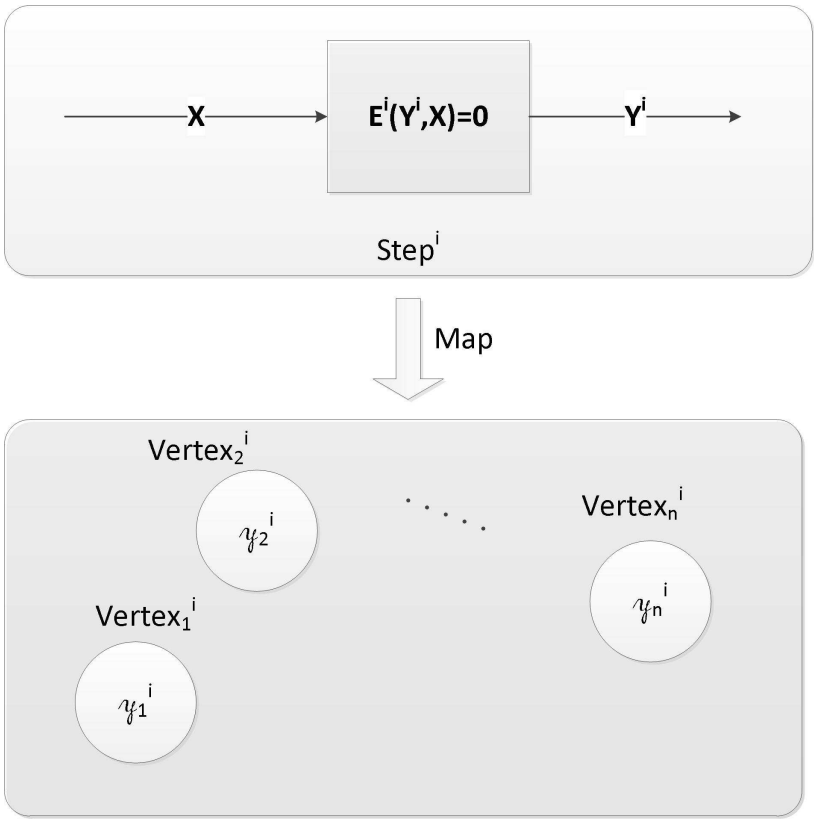


Figure 7.2: Mapping step S^i to n graph vertices.

Algorithm 7.1 Build disconnected graph

for all steps S^i in CA **do**
$V^i \leftarrow M\left(S^i\right)$ such that $Label\left(v_j^i\right) = Name\left(y_j^i\right), j = 1, 2, \ldots, n$
end for

2. **Connecting the vertices in the graph**

   Two entities are required for this operation:

   (a) The graph **vertices,** whose labels are named after the y_j^i variables.

   (b) The **CA steps** S^i, which include the functional dependency statement for y_j^i:

   $$y_j^i = f\left(x_{j1}, x_{j2}, \ldots, x_{jp}\right) \tag{7.20}$$

   where $\{x_{j1}, x_{j2}, \ldots, x_{jp}\} \subseteq \{x_1, x_2, \ldots, x_m\}$ is a subset of the input X, $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, n$.

Assuming the current vertex is v_j^i, there are **four stages** involved in the process of linking v_j^i to its adjacent graph vertices:

(a) From the label of the current vertex **determine the calculation step** that mapped to it. The inverse map permits the convenient identification of the calculation step S^i from the superscript index of the label of the vertex:

$$M^{-1}\left(v_j^i\right) = S^i \tag{7.21}$$

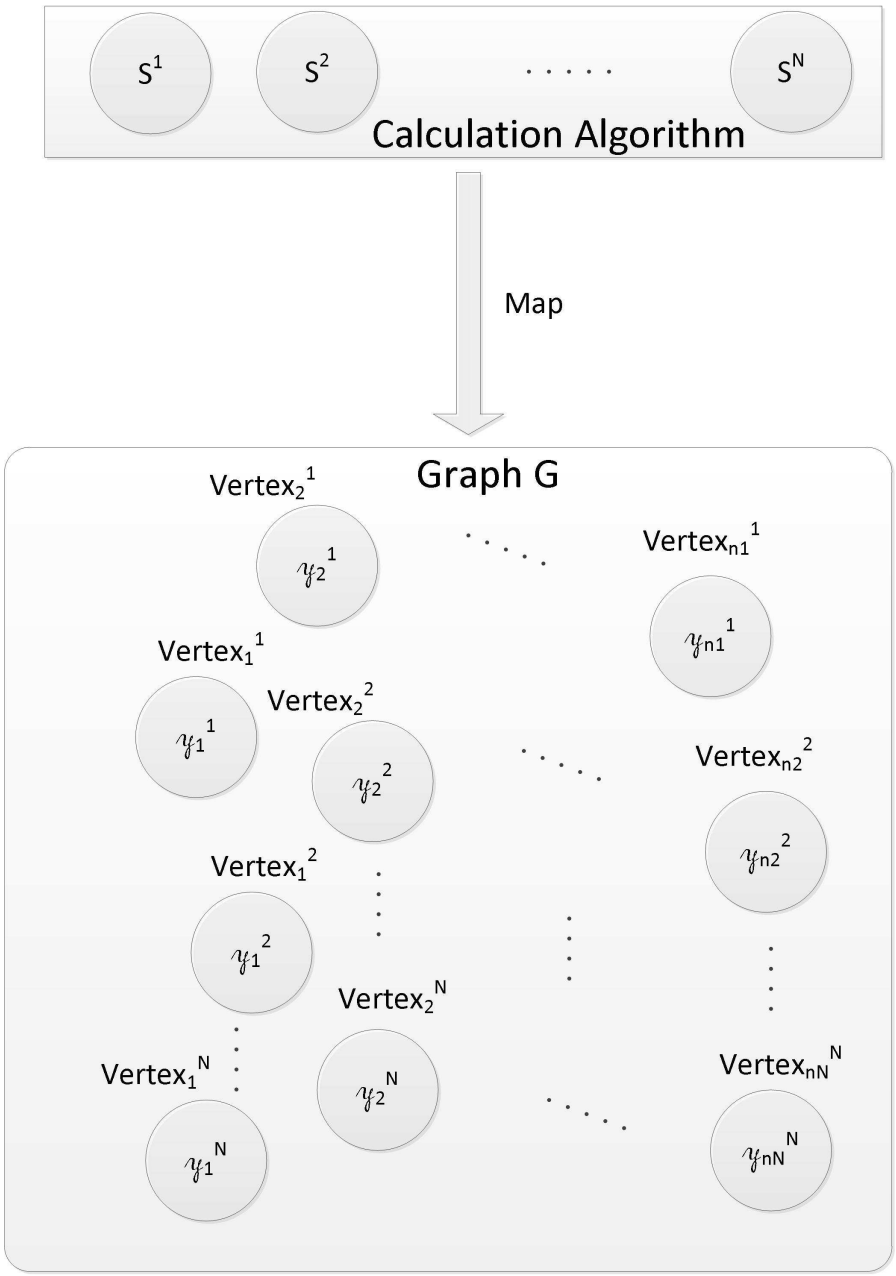


Figure 7.3: Mapping of the CA to a disconnected graph. There are ni variables calculated at step $S^i, i = 1, 2, \ldots, N$.

(b) From the functional dependency **determine the influencing factors**. Equation 7.21 allowed the identification of step S^i, where the factor y_j^i had already been determined to have the functional dependency:

$$y_j^i = f\left(x_{j_1}, x_{j_2}, \ldots, x_{j_p}\right) \tag{7.22}$$

The factors that y_j^i depends on are $\{x_{j_1}, x_{j_2}, \ldots, x_{j_p}\}$.

(c) **Search graph** G for the p vertices whose labels are $\{x_{j_1}, x_{j_2}, \ldots, x_{j_p}\}$. The influencing factors $\{x_{j_1}, x_{j_2}, \ldots, x_{j_p}\}$ of y_j^i have been previously calculated in other steps and there exists a set $\{y_1^{k_1}, y_2^{k_2}, \ldots, y_p^{k_p}\}$ of

parameters calculated at steps $\{S^{k_1}, S^{k_2}, \ldots, S^{k_p}\}$ such that:

$$\begin{cases} x_{j_1} & = y_1^{k_1} \\ x_{j_2} & = y_2^{k_2} \\ \quad \vdots & \\ x_{j_p} & = y_p^{k_p} \end{cases} \tag{7.23}$$

because the calculation steps are connected such that the output of one step is the input of another. This is the step pre-requisites condition discussed in Section 7.2.

As the graph G is a representation of the model parameters, there exists a set of p vertices such that:

$$\begin{cases} Label\left(v_1^{k_1}\right) & = Name\left(y_1^{k_1}\right) \\ Label\left(v_2^{k_2}\right) & = Name\left(y_2^{k_2}\right) \\ \quad \vdots & \\ Label\left(v_p^{k_p}\right) & = Name\left(y_p^{k_p}\right) \end{cases} \tag{7.24}$$

This proves that the graph searches for vertices of labels $x_{j_1}, x_{j_2}, \ldots, x_{j_p}$ will always be successful.

(d) **Connect** the p vertices found, $\{v_1^{k_1}, v_2^{k_2}, \ldots, v_p^{k_p}\}$, to vertex v_j^i by using directed edges that originate in the p vertices.

The four-stage process of connecting a vertex to its adjacent vertices using functional dependency is illustrated in Figure 7.4. The visual representation of the graph is complex even for a small size model. Therefore, we will use the adjacency matrix, defined as:

$$A_{R,R} = \begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,R} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ a_{R,1} & a_{R,2} & \ldots & a_{R,R} \end{bmatrix}$$

where:

R = number of model parameters.
$a_{i,j} = 1$ denotes a functional dependency of the variable in the vertex v_i on the variable in the vertex v_j and $a_{i,j} = 0$ otherwise.
$i, j = 1, 2, \ldots, R$.

The pseudocode of the recursive procedure to connect the vertices in graph G based on the four stages discussed earlier is shown in Algorithm 7.2. The procedure starts from vertices V^N and populates the edges in the graph until the stop condition $y = f()$ is reached. At the end of the procedure, the adjacency matrix A contains the graph representation of the model parameters and their functional dependencies.

As mentioned at the start of this chapter, this procedure will be demonstrated in Section 7.6 using a case study: the Zhang et al. slug flow model.

Algorithm 7.2 The pseudocode of the recursive procedure to connect the vertices in graph G based on the steps in CA. The procedure generates the adjacency matrix A.

procedure CONNECT(G, v)
INPUT:
v=current vertex.
G=directed graph.
OUTPUT:
A=adjacency matrix for graph G.

Build a zero square matrix of size R:
$a_{i,j} = 0$ where $i, j = 1, 2, \ldots, R$.
Mark v as explored:
$v\,(Explored) =$ **TRUE**
1. Determine the calculation step where variable y_j^i, for which $Label\,(v) = Name\left(y_j^i\right)$, is calculated.
The inverse map:
$M^{-1}\,(v) = S^i$ is used for this operation.
2. Determine the input factors from:
$y_j^i = f\,(x_{j1}, x_{j2}, \ldots, x_{jp})$
if $f = f\,()$ **then**
return
else
3. Search graph G for vertices $\{w_1, w_2, \ldots, w_p\}$ such that:
$Label\,(w_1) = x_{j1}^i$ and
$Label\,(w_2) = x_{j2}^i$ and
$\vdots$
$Label\,(w_p) = x_{jp}^i$
This determines $G.adjacentVertices\,(v) = \{w_1, w_2, \ldots, w_p\}$
4. Update the elements of the adjacency matrix:
$a_{j,j1} = 1, a_{j,j2} = 1, \ldots, a_{j,jp} = 1$
for all w in $G.adjacentVertices\,(v)$ **do**
if $w\,(Explored) =$ **FALSE then**
CALL Connect(G, w)
else
return
end if
end for
end if
end procedure

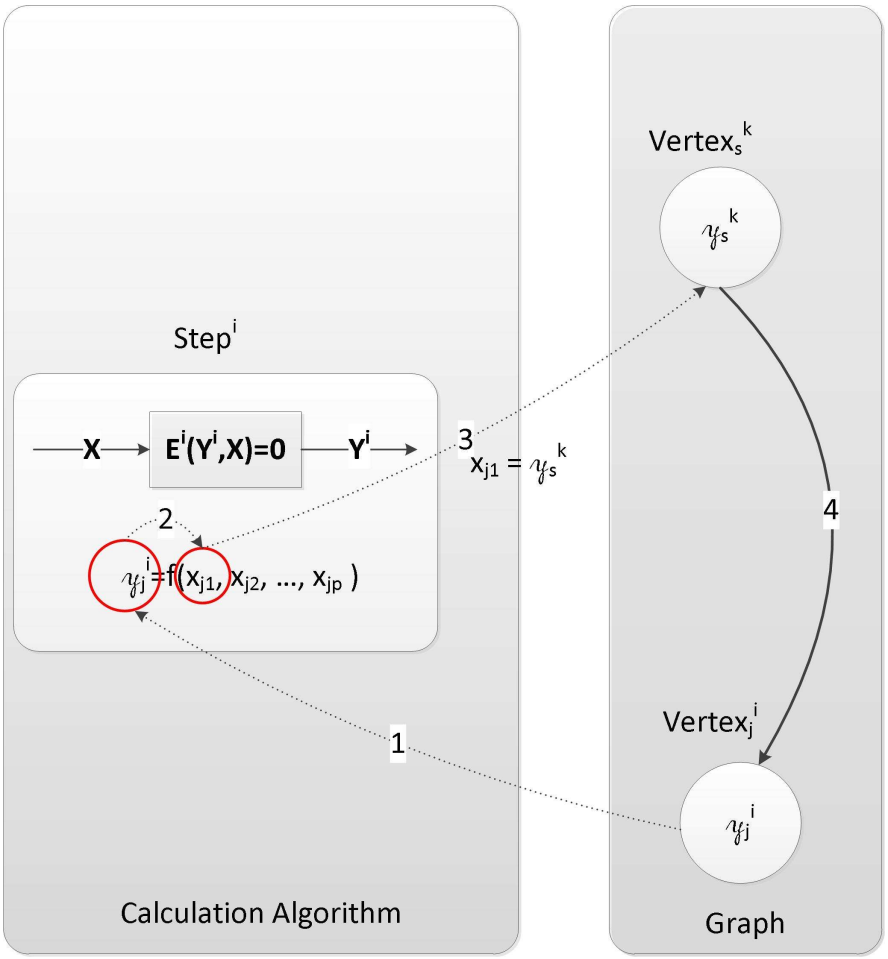


Figure 7.4: The four-stage process of determining the adjacent vertices for a vertex in graph G.

7.4.2 Dependency chain identification

Let us assume a mechanistic model mapped to a directed graph G, as described in Section 7.4.1, and an empirical parameter:

$$e_k = g\left(x_1, x_2, \ldots, x_m\right) \tag{7.26}$$

calculated with the empirical correlation g.
This section aims to:

- Determine the dependency chain $D\left(e_k\right)$.
- Determine the category of the influencing factors $\{x_1, x_2, \ldots, x_m\}$.
- Determine the position of $Step^{e_k}$.

The weight of the edges in graph G are all equal to 1, as reflected by the elements of the adjacency matrix. The distance between two vertices in G is defined as the sum of the weights of the connecting edges.

Definition: The **depth** of vertex x_j relative to vertex e_k is defined as:

$$Depth\left(x_j, e_k\right) = distance\left(x_j, e_k\right) \tag{7.27}$$

where vertex x_j is an ancestor of vertex e_k in the directed graph.

The dependency chain is determined by using a Depth First Search (DFS) in the graph, starting sequentially from the vertices of the influencing factors $x_1, x_2, \ldots, x_m$. To determine the category of each one of these factors, the DFS stop condition was modified to stop the graph traversal if:

1. An empirical parameter is encountered.
2. An empty functional dependency is detected.

This permits classifying the influencing parameters as:

1. Having an empirical component, if stop condition (1) is encountered.
2. Entirely non-empirical, if stop condition (2) is encountered.

The pseudocode of the recursive procedure that determines the dependency chain and establishes the calculation step for an empirical parameter is described in Algorithm 7.3. The procedure is initialised and called according to Algorithm 7.4.

Two global variables are used in Algorithms 7.3 and 7.4 for the case when the current influencing factor, x_j, is entirely non-empirical:

- $ChainMaxStep$. This variable records the maximum step of all the entirely non-empirical parameters in $D(x_j)$.
- $DependencyChainMaxStep$. This variable records the maximum of all $ChainMaxStep$ values. This is the equivalent of calculating the RHS of Condition 7.11.

The two variables require global memory allocation, because they need to retain their value between successive calls of the $IdentifyDependencyChain$ procedure.

If an empirical parameter is encountered while traversing $D(x_j)$, the procedure in Algorithm 7.3 reaches the stop condition and the name of the empirical parameter and the depth are saved to the stack. This is equivalent to calculating the RHS of Condition 7.14.

Before the start of Algorithm 7.3 we assumed e_k had the influencing factors x_1 and x_2, as shown in Figure 7.5. The x_1 and x_2 factors are represented in the superscript notation showing the step in CA where they are calculated. The dependency of e_k on x_1 and x_2 is outlined by the two 1 values on the row pertaining to e_k and two columns pertaining to x_1 and x_2 respectively. The graphical tracing of the execution of Algorithm 7.3 is shown in Figure 7.6. The procedure starts from x_1 as an input to correlation g. Due to the step pre-requisites condition, there exists a step in CA that calculates x_1 and, hence, for which x_1 is the output. This is the red line 1 in Figure 7.6. Next, the influencing factors of x_1 are determined. These are the red lines labelled 2 in Figure 7.6. The process continues in the same fashion until one of the following stop conditions is met:

1. The row pertaining to a factor does not contain any values of 1. This means that the factor can not be expanded any more. This means that x_1 is entirely non-empirical.

Algorithm 7.3 Recursive procedure to determine $D(e_k)$. Empirical parameters along the dependency chain and their depth are pushed to the stack. e_k has a functional dependency $e = f(X)$, where $X = \{x_1, x_2, \ldots, x_m\}$ are the influencing factors.

procedure IDENTIFYDEPENDENCYCHAIN($CurrentNode, X, A, Depth$)

INPUT:

$CurrentNode$=current vertex.
X=set of influencing factors.
A=adjacency matrix for graph G.
$Depth$=current depth value.

OUTPUT:

1. Stack contains empirical parameters in the dependency chain and their depth.
2. $DependencyChainMaxStep$ has the maximum in the step of all entirely non-empirical parameters dependency chain.

---

for all x in X **do**
  $ChainMaxStep = 0$
  if $isEmpirical(x) =$ **TRUE then**
    Push x to Stack
    Push $Depth$ to Stack
    Continue
  else
    $Step = getStep(x)$
    if $Step > ChainMaxStep$ **then**
      $ChainMaxStep = Step$
    end if
    $X_1 = identifyAdjacencyList(x)$
    if $isEmpty(X_1)$ **then**
      if $ChainMaxStep > DependencyChainMaxStep$ **then**
        $DependencyChainMaxStep = ChainMaxStep$
      end if
      Continue
    else
      $CurrentNode = x$
      $Depth = Depth + 1$
      CALL IdentifyDependencyChain($CurrentNode, X_1, A, Depth$)
    end if
  end if
end for
end procedure

Algorithm 7.4 Calculate the step where e_k can be placed.

procedure CALCULATESTEPS(e_k, g, A)
INPUT:
e_k=empirical parameter.
g=empirical correlation.
A=adjacency matrix for directed graph G.
OUTPUT:
Stack contains (empirical factor, Depth) pairs.
DependencyChainMaxStep is the maximum of all the steps of the entirely non-empirical factors.
The step where e_k can be placed is:
max(*DependencyChainMaxStep*+1, step of empirical factor+Depth from the stack)

Global: $DependencyChainMaxStep$
Global: $ChainMaxStep$
$CurrentNode = e_k$
$X = identifyAdjacencyList\,(e_k)$
$Depth = 1$
CALL IdentifyDependencyChain($CurrentNode, X, A, Depth$)
end procedure

		Inputs								
		$y_1^{j_1}$	$y_2^{j_2}$	$\cdots$	$x_1^{i_1}$	$\cdots$	$x_2^{i_2}$	$\cdots$	e_k	$y_R^{j_R}$
Outputs	$y_1^{j_1}$	0		$\cdots$		$\cdots$		$\cdots$		
	$y_2^{j_2}$		0	$\cdots$		$\cdots$		$\cdots$		
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_1^{i_1}$			1	0	$\cdots$		$\cdots$		
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$x_2^{i_2}$			$\cdots$		$\cdots$	0	$\cdots$		
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	e_k			$\cdots$	1	$\cdots$	1	$\cdots$	0	
	$y_R^{j_R}$			$\cdots$		$\cdots$				0

Figure 7.5: Adjacency matrix showing the functional dependency of parameter $e_k = g\left(x_1^{i_1}, x_2^{i_2}, \ldots, x_m^{i_m}\right)$.

2. An empirical factor is encountered. This means that x_1 is a factor having an empirical component.

Once either of the two stop conditions is encountered while traversing the dependency chain of x_1, the procedure continues in the same way with x_2.

As noted in Section 7.4.1, the Zhang et al. slug flow model will be used in Chapter 8 to demonstrate the application of these supporting algorithms.

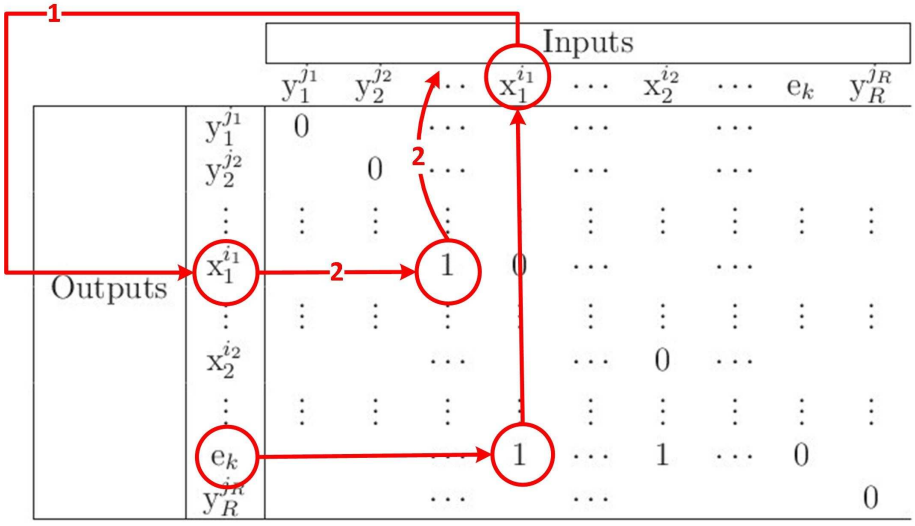


Figure 7.6: Empirical parameter e_k depends on $x_1^{i_1}$ and $x_2^{i_2}$. DFS first explores the dependency chain of $x_1^{i_1}$. The red lines show the path taken to determine the sequence of vertices in $D\left(x_1^{i_1}\right)$. After $D\left(x_1^{i_1}\right)$ is determined, the dependency chain of $x_2^{i_2}$ is explored.

7.5 Algorithm to replace empirical correlations in mechanistic models

Section 7.4 presented a set of procedures, each solving a particular part of the problems stated in Section 7.3. This section merges all the procedures into one algorithm that solves both Problems 1 and 2. The algorithm determines:

Objective 1: One CA^{new} that does not change when the empirical correlations are alternated.

Objective 2: The step in CA^{new} of each empirical parameter in the model.

The **empirical correlation replacement algorithm** is as follows:

Step 1: Convert the model to a properly ordered model.

Most of the time, this can be achieved by introducing an S^0 in CA and calculating all the model parameters that are based entirely on model input factors in this step.

Step 2: For all steps $S^i \in \{S^1, S^2, \ldots, S^N\}$ of CA do:

2.1 Express the variables $y_j^i, j = 1, 2, \ldots, n$ calculated in S^i as a functional dependency:

$$y_j^i = f\left(x_1, x_2, \ldots, x_m\right) \tag{7.28}$$

2.2 Map step S^i to a set of n vertices, one for each variable y_j^i, $j = 1, 2, \ldots, n$. Label the vertex with the variable name.

At the end of Step 2 the disconnected graph G is the graph representation of the model.

The procedure performing these tasks was shown in Algorithm 7.1.

Step 3: For all vertices $v \in G$, starting with the vertices in:

$$V^N = \{v_1^N, v_2^N, \ldots, v_{n_N}^N\}$$

do:

3.1 Determine the adjacent vertices of v.
Use the inverse map and the functional dependency in Equation 7.28 to determine the adjacent vertices, as was explained in Section 7.4.1.

3.2 Update the adjacency matrix to reflect the adjacency.
$a_{i,j} = 1$, if vertex v_i is adjacent with v_j; $a_{i,j} = 0$ otherwise.

At the end of Step 3, the graph representation is complete and the functional dependencies are represented as links in the graph.

The procedure that builds the adjacency matrix representation of the graph was shown in Algorithm 7.2.

Step 4: Group the empirical correlations in sets Ω_k by the empirical parameter e_k they estimate.

Step 5: For each group Ω_k and for all empirical correlations $g_{k,i}$ in each group:

5.1 Write the empirical parameter e_k estimated by the correlation as a functional dependency:

$$e_k = g_{k,i}\left(x_1, x_2, \ldots, x_m\right) \tag{7.29}$$

If the functional dependency is empty, the empirical parameter e_k can be placed anywhere in CA and steps **5.2** and **5.3** are skipped.

5.2 Identify the influencing factors x_j as elements $x_j^{i_j}$ in the *Outputs* column of the adjacency table (Figure 7.5).

5.3 For each influencing factor $x_j^{i_j}, j = 1, 2, \ldots, m$ determine:

- The dependency chain $D\left(x_j^{i_j}\right)$.
- The maximum step of the entirely non-empirical parameters in $D\left(x_j^{i_j}\right)$.
- Record the empirical parameters and their depth in $D\left(x_j^{i_j}\right)$.

While traversing the graph, the stop conditions are:

1. x_j is an empirical parameter.
2. An empty functional dependency $D\,()$ is detected.

5.4 Express the position of $Step^{e_k}$ as:

$$\begin{aligned} Step^{e_k} \geq \{&\text{step of entirely non-empirical parameters in CA,} \\ &\text{step of empirical parameters in CA}^{new} + \Delta\} \end{aligned} \tag{7.30}$$

where Δ is the depth.
The pseudocode of the procedures implementing this step were shown in Algorithms 7.3 and 7.4

Table 7.1: Slug flow empirical parameters and empirical correlations selected in this study.

Parameter	Correlations used for replacement	Number of correlations per parameter
L_S	Zhang, Scott, Felizola and Shoham	3
H_{LLS}	Eissa Al-Safran, Woldesmayat and Ghajar, Toshiba, Gomez, Zhang, Spedding and Spence, Minami and Brill, Barnea and Brauner, Gregory, Nicklin	10
v_{TB}	Bendiksen, Hassan and Kabir	2
f_i	Cohen and Hanratty, Andritsos and Hanratty, Ouyang and Aziz, Bendiksen, Vlachos, Asali and Ambrosio	6
Θ	Grolman, Biberg, Fan, Hart	4

Step 6: Minimise the number of calculation algorithms.

6.1 Identify the empirical parameters that have associated more than two empirical correlation groups. A group was formed as a set of empirical correlations having the same position conditions in the CA for the given empirical parameter.

6.2 For each one of these parameters, reduce the number of groups to one, by making all the correlations part of the group with the highest step number.

7.6 Case study — Applying the empirical correlation replacement algorithm to the Zhang et al. (2000) slug flow mechanistic model

We will now follow all the steps in the algorithm in Section 7.5 and use the Zhang et al. (2000) model as a case study. A set of empirical correlations was selected for the estimation of each one of the empirical parameters: L_S, v_{TB}, H_{LLS}, Θ and f_i. These correlations are shown in Table 7.1. While reading this section, it may help to refer to Section 5.3, which presented the Zhang model.

Remark: Step S^i, which was the notation for a step in a CA, translates into **Zhang step** i for the case of the Zhang et al. (2000) model, from Section 5.3.

Step 1: Convert the model to a properly ordered model.

This can be achieved by introducing step S^0 and calculating the parameters:

- A_{pipe}=cross-sectional area
- S_S=pipe circumference
- $v_M = v_{SL} + v_{SG}$

as part of S^0. This involves bringing:

- S_S from S^{12} to S^0.
- v_M from S^2 to S^0.

Step 2: For all steps $S^i \in \{S^1, S^2, \ldots, S^{18}\}$ of CA do:

2.1 Express the variables $y_j^i, j = 1, 2, \ldots, n$ calculated in S^i as a functional dependency:

$$y_j^i = f\left(x_1, x_2, \ldots, x_m\right) \tag{7.31}$$

S^0 was not included because all the functional dependencies of the parameters in S^0 are empty. As S^0 is not mapped to the graph, when we express the functional dependencies we will not take into account S^0 parameters. In this case, they are A_{pipe}, S_S and v_M.
Starting from **Zhang step 1**, which in this notation is S^1, we express the variables in each step as in Equation 7.31:

S^1 :

$$H_{LTB}^1 = f\left(\right)$$

S^2 :

$$L_S^2 = f\left(\right)$$

S^3 :

$$v_{TB}^3 = f\left(\right)$$

S^4 :

$$H_{LLS}^4 = f\left(\right)$$

S^5 :

$$v_{LTB}^5 = f\left(H_{LTB}^1, v_{TB}^3, H_{LLS}^4\right)$$
$$v_{GTB}^5 = f\left(H_{LTB}^1, v_{TB}^3, H_{LLS}^4\right)$$

S^6 :

$$\rho_S^6 = f\left(H_{LLS}^4\right)$$

S^7 :

$$Re_S^7 = f\left(\rho_S^6\right)$$

S^8 :

$$f_S^8 = f\left(Re_S^7\right)$$

S^9 : This step does not occur in the 2000 version of the Zhang model considered in this example.

S^{10} :

$$L_U^{10} = f\left(H_{LTB}^1, L_S^2, v_{GTB}^3, H_{LLS}^4\right)$$
$$L_F^{10} = f\left(H_{LTB}^1, L_S^2, v_{GTB}^3, H_{LLS}^4\right)$$

S^{11} :

$$\Theta^{11} = f\left(H_{LTB}^1\right)$$

S^{12} :

$$S_F^{12} = f\left(\Theta^{11}\right)$$
$$S_G^{12} = f\left(\Theta^{11}\right)$$
$$S_{CD}^{12} = f\left(\Theta^{11}\right)$$
$$S_i^{12} = f\left(H_{LTB}^1, \Theta^{11}\right)$$
$$A_F^{12} = f\left(H_{LTB}^1\right)$$
$$A_G^{12} = f\left(H_{LTB}^1\right)$$
$$A_{CD}^{12} = f\left(\Theta^{11}\right)$$

S^{13} :

$$Re_{LW}^{13} = f\left(v_{LTB}^5, A_F^{12}, S_F^{12}\right)$$
$$Re_{GW}^{13} = f\left(v_{GTB}^5, A_G^{12}, S_G^{12}, S_i^{12}\right)$$

S^{14} :

$$f_{LW}^{14} = f\left(Re_{LW}^{13}\right)$$
$$f_{GW}^{14} = f\left(Re_{GW}^{13}\right)$$
$$f_i^{14} = f\left(\right)$$

S^{15} :

$$\tau_{LW}^{15} = f\left(v_{LTB}^5, f_{LW}^{14}\right)$$
$$\tau_{GW}^{15} = f\left(v_{GTB}^5, f_{GW}^{14}\right)$$
$$\tau_i^{15} = f\left(v_{LTB}^5, v_{GTB}^5, f_i^{14}\right)$$
$$\tau_S^{15} = f\left(\rho_S^6, f_S^8\right)$$

S^{16} and S^{17} : These steps are not included here, because they do not calculate model parameters. They represent the conditional reiteration of some of the CA steps.

S^{18} :

$$\frac{dP}{dL}^{18} = f\left(H_{LTB}^1, L_S^2, L_U^{10}, L_F^{10}, S_F^{12}, S_G^{12}, \tau_{LW}^{15}, \tau_{GW}^{15}, \tau_S^{15}\right)$$

Table 7.2: The empty adjacency matrix is a preliminary step in generating the graph representation of the Zhang et al. (2000) model. It has 26 rows and 26 columns.

	H^1_{LTB}	L^2_S	v^3_{TB}	$\dots$	τ^{15}_{LW}	τ^{15}_{GW}	τ^{15}_i	τ^{15}_S	$\frac{dP}{dL}^{18}$
H^1_{LTB}									
L^2_S									
v^3_{TB}									
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
τ^{15}_{LW}									
τ^{15}_{GW}									
τ^{15}_i									
τ^{15}_S									
$\frac{dP}{dL}^{18}$									

2.2 Map step S^i to a set of n vertices, one for each variable y^i_j, $j = 1, 2, \dots, n$. Assign the variable name to the vertex label. This is equivalent to mapping to the empty adjacency matrix in Table 7.2.

Step 3: For all vertices $v \in G$, starting with the vertices in $V^{18} = \{v^{18}_1\}$ do:

3.1 Determine the adjacent vertices of v.

3.2 Update the adjacency matrix to reflect the new adjacency.

All the graph vertices are found in the *Outputs* column of the adjacency table. Therefore, to traverse all the vertices in the graph, we start with the last element in the column and progress towards the first.

The last element in the *Outputs* column has the $Label = Name\left(\frac{dP}{dL}^{18}\right)$. The inverse map points to S^{18}, where the functional dependency for $\frac{dP}{dL}$ determined in Step 2 is:

$$\frac{dP}{dL}^{18} = f\left(H^1_{LTB}, L^2_S, L^{10}_U, L^{10}_F, S^{12}_F, S^{12}_G, \tau^{15}_{LW}, \tau^{15}_{GW}, \tau^{15}_S\right)$$

The adjacent vertices are: $\{H^1_{LTB}, L^2_S, L^{10}_U, L^{10}_F, S^{12}_F, S^{12}_G, \tau^{15}_{LW}, \tau^{15}_{GW}, \tau^{15}_S\}$.

Factors S_F, S_G and S_{CD} have been merged into one variable S_*, because as S^{12} shows, they have the same functional dependency. Similarly A_F and A_G are merged into $A_{F,G}$. These simplifications help unclutter the adjacency matrix without affecting the results. Based on these considerations, the adjacent vertices are: $\{H^1_{LTB}, L^2_S, L^{10}_U, L^{10}_F, S^{12}_{F,G}, \tau^{15}_{LW}, \tau^{15}_{GW}, \tau^{15}_S\}$.

The adjacency relationship is reflected in the following matrix elements:

$a_{26,1} = 1 \quad a_{26,2} = 1 \quad a_{26,10} = 1 \quad a_{26,11} = 1$
$a_{26,13} = 1 \quad a_{26,22} = 1 \quad a_{26,23} = 1 \quad a_{26,25} = 1$

The next vertex has the label τ^{15}_S. Inverse mapping to S^{15}, we find the functional relationship for τ^{15}_S:

$$\tau^{15}_S = f\left(\rho^6_S, f^8_S\right)$$

The following matrix elements are updated:

$$a_{25,7} = 1 \quad a_{25,9} = 1$$

The next vertex has the label τ_i^{15}, for which the inverse map points to S^{15}. The functional relationship:

$$\tau_i^{15} = f\left(v_{LTB}^5, v_{GTB}^5, f_i^{14}\right)$$

is reflected in the adjacency matrix:

$$a_{24,5} = 1 \quad a_{24,6} = 1 \quad a_{24,21} = 1$$

The procedure continues in a similar fashion until all the vertices are traversed and the adjacent vertices determined. At the end, the adjacency matrix has the form shown in Table 7.3. This table is the graph representation of the Zhang et al. (2000) model.

Step 4: Group the empirical correlations in sets Ω by the empirical parameter they estimate. Table 7.1 lists the empirical correlations that have been selected for each empirical parameter in the Zhang et al. (2000) slug flow model.

$\Omega_1 = \{g_{1,1}, g_{1,2}, g_{1,3}\}$ for the estimation of $e_1 = L_S$
$\Omega_2 = \{g_{2,1}, g_{2,2}, g_{2,3}, g_{2,4}, g_{2,5}, g_{2,6}, g_{2,7}, g_{2,8}, g_{2,9}, g_{2,10}\}$
for the estimation of $e_2 = H_{LSS}$
$\Omega_3 = \{g_{3,1}, g_{3,2}\}$ for the estimation of $e_3 = v_{TB}$
$\Omega_4 = \{g_{4,1}, g_{4,2}, g_{4,3}, g_{4,4}, g_{4,5}, g_{4,6}\}$ for the estimation of $e_4 = f_i$
$\Omega_5 = \{g_{5,1}, g_{5,2}, g_{5,3}, g_{5,4}\}$ for the estimation of $e_5 = \Theta$

$g_{1,1} =$ Zhang correlation
$g_{1,2} =$ Scott correlation
$g_{1,3} =$ Felizola and Shoham correlation

$g_{2,1} =$ Eissa Al-Safran correlation
$g_{2,2} =$ Woldesmayat and Ghajar correlation
$g_{2,3} =$ Toshiba correlation
$g_{2,4} =$ Gomez correlation
$g_{2,5} =$ Zhang correlation
$g_{2,6} =$ Spedding and Spence correlation
$g_{2,7} =$ Minami and Brill correlation
$g_{2,8} =$ Barnea and Brauner correlation
$g_{2,9} =$ Gregory correlation
$g_{2,10} =$ Nicklin correlation

	H^1_{LTB}	L^2_S	v^3_{TB}	H^4_{LLS}	v^5_{LTB}	v^5_{GTB}	ρ^6_S	Re^7_S	f^8_S	L^{10}_U	L^{10}_F	Θ^{11}	S^{12}_*	S^{12}_i	$A^{12}_{F,G}$	A^{12}_{CD}	Re^{13}_{LW}	Re^{13}_{GW}	f^{14}_{LW}	f^{14}_{GW}	f^{14}_i	τ^{15}_{LW}	τ^{15}_{GW}	τ^{15}_i	τ^{15}_S	$\frac{dP}{dL}^{18}$
H^1_{LTB}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L^2_S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v^3_{TB}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H^4_{LLS}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v^5_{LTB}	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
v^5_{GTB}	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ρ^6_S	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Re^7_S	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f^8_S	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L^{10}_U	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L^{10}_F	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Θ^{11}	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S^{12}_*	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S^{12}_i	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$A^{12}_{F,G}$	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A^{12}_{CD}	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Re^{13}_{LW}	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
Re^{13}_{GW}	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
f^{14}_{LW}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
f^{14}_{GW}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
f^{14}_i	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
τ^{15}_{LW}	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
τ^{15}_{GW}	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
τ^{15}_i	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
τ^{15}_S	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\frac{dP}{dL}^{18}$	1	1	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	0

Table 7.3: Adjacency matrix for the Zhang et al. (2000) model. Note that $S_* = \{S_F, S_G, S_{CD}\}$ and $A_{F,G} = \{A_F, A_G\}$

$g_{3,1}$ = Bendiksen correlation
$g_{3,2}$ = Hassan and Kabir correlation

$g_{4,1}$ = Cohen and Hanratty correlation
$g_{4,2}$ = Andritsos and Hanratty correlation
$g_{4,3}$ = Ouyang and Aziz correlation
$g_{4,4}$ = Bendiksen correlation
$g_{4,5}$ = Vlachos correlation
$g_{4,6}$ = Asali and Ambrosio correlation

$g_{5,1}$ = Grolman correlation
$g_{5,2}$ = Biberg correlation
$g_{5,3}$ = Fan correlation
$g_{5,4}$ = Hart correlation

Step 5: The details of Step 5 were discussed in Section 7.5 and will be demonstrated for each one of the empirical parameters e_k (H_{LLS}, L_S, v_{TB}, f_i and Θ) and for each empirical correlation $g_{k,i}$ used in this study, as shown in Table 7.1 and Step 4.

1. **Group - Parameter slug liquid holdup, H_{LLS}**

   (a) **Eissa Al-Safran (2009):** The Eissa Al-Safran correlation for H_{LLS} is:

$$H_{LLS} = 1.05 - \frac{0.0417}{mr - 0.123}$$
$$mr = \frac{H_{LTB}\left(v_{TB} - v_{LTB}\right)\left(v_M - v_{LTB}\right)}{v_M^2}$$

   5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\left(H_{LTB}, v_{TB}, v_{LTB}, v_M\right)$$

   5.2 Determine the influencing factors in the superscript notation.

$$\{H_{LTB}^1, v_{TB}^3, v_{LTB}^5\}$$

   5.3 For each influencing factor determine the dependency chain. Stop the dependency chain when an empty dependency or an empirical parameter is encountered.

   The following is a tracing of the execution of the pseudocode in Algorithm 7.3. $MaxStep$ is the variable $DependencyChainMaxStep$ in the pseudocode.

Depth: $+1$, $MaxStep = 0$

$D\left(H^1_{LTB}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} > MaxStep$. Assign $MaxStep = 1$. Stop.
$D\left(v^3_{TB}\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB}+1\}$. Stop.
$D\left(v^5_{LTB}\right) = \{D\left(H^1_{LTB}\right), D\left(v^3_{TB},\right), D\left(H^4_{LLS}\right)\}$. Explore further.

Depth: $+2$

$D\left(H^1_{LTB}\right) = \left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\left(v^3_{TB}\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB}+2, v_{TB}+1\}$. Stop.
$D\left(H^4_{LLS}\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS}+2, v_{TB}+2, v_{TB}+1\}$. Stop.
Return to Depth $+1$.

Depth: $+1$

Finished exploring all influencing factors.

5.4 H_{LLS} placement restrictions by influencing parameter type:
Entirely non-empirical: $\{H^1_{LTB}\}$.

$$Step^{H_{LLS}} \geq MaxStep + 1 = 2$$

Having an empirical component: $\{v^5_{LTB}, v^3_{TB}\}$.

$$\begin{aligned} Step^{H_{LLS}} &\geq \max\{Step^{H_{LLS}}+2, Step^{v_{TB}}+2, Step^{v_{TB}}+1\} \\ &= \max\{Step^{H_{LLS}}+2, Step^{v_{TB}}+2\} \end{aligned}$$

H_{LLS} final placement restrictions:
Combining the previous results, we conclude that:

$$Step^{H_{LLS}} \geq \max\{2, Step^{H_{LLS}}+2, Step^{v_{TB}}+2\} \quad (7.32)$$

(b) **Woldesmayat and Ghajar (2006):** The Woldesmayat and Ghajar correlation for H_{LLS} is:

$$\begin{aligned} H_{LLS} &= 1 - \frac{v_{SG}}{\tilde{v}} \\ \tilde{v} &= v_{SG}\left[1 + \left(\frac{v_{SL}}{v_{SG}}\right)^{\left(\frac{\rho_G}{\rho_L}\right)^{0.1}}\right] \\ &\quad + 2.9\left[\frac{gd\sigma\left(1+\cos\theta\right)\left(\rho_L-\rho_G\right)}{{\rho_L}^2}\right]^{0.25} \\ &\quad \left(1.22 + 1.22\sin\theta\right)^{\frac{P_{atmospheric}}{P_{system}}} \end{aligned}$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\,()$$

5.4 Consequently, the empirical parameter can be placed anywhere in CA^{new}.

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.33}$$

(c) **Toshiba (2005):** The Toshiba correlation for H_{LLS} is:

$$H_{LLS} = 1 - \frac{v_{SG}}{1.08 v_M + 0.45}$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\,()$$

Similarly:
5.4 H_{LLS} final placement restrictions:

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.34}$$

(d) **Gomez (2000):** The Gomez correlation for H_{LLS} is:

$$H_{LLS} = 1.0 \times e^{-\left(7.8510^{-3}\theta + 2.4810^{-6} Re_{LS}\right)}$$
$$Re_{LS} = \frac{\rho_L v_M d}{\mu_L}$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\,()$$

5.4 This has the same functional dependency like the Toshiba correlation, analysed earlier. Therefore:

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.35}$$

(e) **Zhang (2000):** The Zhang correlation for H_{LLS} is:

$$\begin{aligned} H_{LLS} &= \frac{1}{1 + \frac{T_{sm}}{3.16\sqrt{(\rho_L - \rho_G) g \sigma}}} \\ T_{sm} &= f_S \frac{\rho_S {v_M}^2}{2 C_e} \\ &\quad + \frac{d}{4} \frac{\rho_L H_{LTB} \left(v_{TB} - v_{LTB}\right) \left(v_M - v_{LTB}\right)}{L_S C_e} \\ C_e &= \frac{2.5 - |\sin\theta|}{2} \end{aligned}$$

5.1 Write the empirical parameter as a functional dependency.

$H_{LLS} = g\left(\rho_S, f_S, H_{LTB}, v_{TB}, v_{LTB}, L_S\right)$

5.2 Determine the influencing factors in the superscript notation.

$\{\rho_S^6, f_S^8, H_{LTB}^1, v_{TB}^3, v_{LTB}^5, L_S^2\}$

5.3 For each influencing factor determine the dependency chain.

Depth: $+1$, $MaxStep = 0$

$D\left(\rho_S^6\right) = D\left(H_{LLS}\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 1\}$. Stop.
$D\left(f_S^8\right) = D\left(Re_S^7\right) = D\left(\rho_S^6\right)$. Explore further.

Depth: $+2$

$D\left(\rho_S^6\right) = D\left(H_{LLS}^4\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 2, H_{LLS} + 1\}$. Stop.
Return to Depth $+1$.

Depth: $+1$

$D\left(H_{LTB}^1\right) = ()$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} > MaxStep$. Assign $MaxStep = 1$. Stop.
$D\left(v_{TB}^3\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB} + 1, H_{LLS} + 2, H_{LLS} + 1\}$. Stop.
$D\left(v_{LTB}^5\right) = \{D\left(H_{LTB}^1\right), D\left(v_{TB}^3, D\left(H_{LLS}^4\right)\right.\right.$. Explore further.

Depth: $+2$

$D\left(H_{LTB}^1\right) = D()$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\left(v_{TB}^3\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB} + 2, v_{TB} + 1, H_{LLS} + 2, H_{LLS} + 1\}$. Stop.
$D\left(H_{LLS}^4\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 2, v_{TB} + 2, v_{TB} + 1, H_{LLS} + 2, H_{LLS} + 1\}$. Stop.
Return to Depth $+1$.

Depth: $+1$

$D\left(L_S^2\right)$: Empirical : Push L_S and depth to stack : $\{L_S + 1, H_{LLS} + 2, v_{TB} + 2, v_{TB} + 1, H_{LLS} + 2, H_{LLS} + 1\}$. Stop.
Finished H_{LLS} dependency chain.

5.4 H_{LLS} placement restrictions by influencing parameter type:
Entirely non-empirical: $\{H_{LTB}^1\}$.

$$Step^{H_{LLS}} \geq MaxStep + 1 = 2$$

Having an empirical component: $\{L_S^2, v_{TB}^3, \rho_S^6, f_S^8, v_{LTB}^5\}$.

$$Step^{H_{LLS}} \geq Step^{L_S} + 1, Step^{H_{LLS}} + 2, Step^{v_{TB}} + 2,$$
$$Step^{v_{TB}} + 1, Step^{H_{LLS}} + 2, Step^{H_{LLS}} + 1$$

Combining the results, we conclude that:

$$Step^{H_{LLS}} \geq \max\{2, Step^{L_S} + 1, Step^{v_{TB}} + 2, Step^{H_{LLS}} + 2\} \tag{7.36}$$

(f) **Spedding and Spence (1989):** The Spedding and Spence correlation for H_{LLS} is:

$$\frac{\epsilon}{1-\epsilon} = \left[0.45 + 0.08e^{-100\left(0.25 - v_{SL}^2\right)}\right]\left(\frac{v_{SG}}{v_{SL}}\right)^{0.65}$$
$$H_{LLS} = 1 - \epsilon$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\,()$$

5.4 Consequently, the empirical parameter can be placed anywhere in CA^{new}.

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.37}$$

(g) **Minami and Brill (1987):** The Minami and Brill correlation for H_{LLS} is:

$$\begin{aligned} H_{LLS} &= 1 - e^{-\left[\frac{\ln Z_1 + 9.21}{8.7115}\right]^{4.3374}} \\ Z_1 &= \frac{1.84{v_{SL}}^{0.575}}{v_{SG}d^{0.0277}}\left(\frac{{\rho_L}^{0.5804}}{g^{0.3696}\sigma^{0.1804}}\right)^{-0.25} \\ &\quad \left(\frac{P_{system}}{101325}\right)^{0.05}{\mu_L}^{0.1} \end{aligned}$$

5.1 - 5.4 Similarly, all the inputs in the Minami and Brill correlation are mechanistic model inputs. Therefore:

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.38}$$

(h) **Barnea and Brauner (1985):** The Barnea and Brauner correlation for H_{LLS} is:

$$\begin{aligned} H_{LLS} &= 1 - 0.058h^2 \\ h &= 2\sqrt{\frac{0.4\sigma}{(\rho_L - \rho_G)\,g}}\left(\frac{\rho_L}{\sigma}\right)^{0.666}\left(\frac{f_S{v_M}^3}{2d}\right)^{0.4} - 0.725 \end{aligned}$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\left(f_S\right)$$

5.2 Determine the influencing factors in the superscript notation. Using the adjacency matrix, we determine that the influencing factors are:

$$\{f_S^8\}$$

5.3 For each influencing factor determine the dependency chain.

Depth: $+1$, $MaxStep = 0$

$D\left(f_S^8\right) = D\left(Re_S^7\right) = D\left(\rho_S^6\right)$. Explore further.

Depth: $+2$

$D\left(\rho_S^6\right) = D\left(H_{LLS}^4\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 2\}$. Stop.
Return to Depth $+1$.

Depth: $+1$

Finished exploring the influencing factors of H_{LLS}.

5.4 H_{LLS} final placement restrictions:
Combining the results, we conclude that:

$$Step^{H_{LLS}} \geq \max\{Step^{H_{LLS}} + 2\} \tag{7.39}$$

This result may seem incorrect, but the interpretation is that the step where H_{LLS} calculated with the Barnea and Brauner correlation can be placed is at least two steps after the initialisation of the parameter with the Gregory correlation.

(i) **Gregory (1978):** The Gregory correlation for H_{LLS} is:

$$H_{LLS} = \frac{1}{1 + \left(\frac{v_M}{8.66}\right)^{1.39}}$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\left(\right)$$

5.4 H_{LLS} placement restrictions:

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.40}$$

(j) **Nicklin (1962):** The Nicklin correlation for H_{LLS} is:

$$H_{LLS} = 1 - \frac{v_{SG}}{1.2v_M + 0.35\sqrt{gd}}$$

5.1 Write the empirical parameter as a functional dependency.

$$H_{LLS} = g\,() \tag{7.41}$$

5.4 This is similar to the previous case, therefore the position is:

$$Step^{H_{LLS}} = \text{anywhere} \tag{7.42}$$

Based on their functional dependencies, the empirical correlations for parameter H_{LLS} form the following **four** groups, as shown in Table 7.4:

Group 1: Woldesmayat and Ghajar, Toshiba, Gomez, Spedding and Spence, Minami and Brill, Gregory, Nicklin. H_{LLS} estimated with these correlations can be placed anywhere in the calculation algorithm.

Group 2: Barnea and Brauner. H_{LLS} calculated with this correlation can be placed at a step that is two steps after the correlation that initially estimates H_{LLS}, such as the Gregory correlation or any other Group 1 correlation.

Group 3: Eissa Al-Safran. H_{LLS} calculated with this correlation can be placed at a step that is a maximum of the following: 3, two steps after v_{TB} and two steps after the correlation that initialises H_{LLS}.

Group 4: Zhang. H_{LLS} calculated with this correlation can be placed at a step that is a maximum of the following: 2, one step after L_S, two steps after v_{TB} and two steps after the correlation that initially estimates H_{LLS}.

2. **Group - Parameter liquid slug length, L_S**

(a) **Zhang at al (2000):**
The correlations is:

$$L_S = \left(32\cos^2\theta + 16\sin^2\theta\right) d \tag{7.43}$$

All the influencing parameters are model input factors. Therefore:

$$Step^{L_S} = \text{anywhere} \tag{7.44}$$

(b) **Scott et al. (1989):**
The Scott correlation is:

$$L_S = \max\left\{30d, e^{-26.8+28.5\left[\ln\left(\frac{d}{0.0254}\right)\right]^{0.1}}\right\} \tag{7.45}$$

Similarly:

Table 7.4: There are four groups of empirical correlations for H_{LLS}. The grouping criteria are the step placement conditions in the algorithm.

Group	Empirical correlations	Position conditions
1	Woldesmayat and Ghajar, Toshiba, Gomez, Spedding and Spence, Minami and Brill, Gregory, Nicklin	$Step^{H_{LLS}} = \text{anywhere}$
2	Barnea and Brauner	$Step^{H_{LLS}} \geq \max\{Step^{H_{LLS}} + 2\}$
3	Eissa Al-Safran	$Step^{H_{LLS}} \geq \max \begin{cases} 2 \\ Step^{v_{TB}} + 2 \\ Step^{H_{LLS}} + 2 \end{cases}$
4	Zhang	$Step^{H_{LLS}} \geq \max \begin{cases} 2 \\ Step^{L_S} + 1, \\ Step^{v_{TB}} + 2, \\ Step^{H_{LLS}} + 2 \end{cases}$

Table 7.5: There is only one group of correlations for the calculation of L_S.

Group	Empirical correlations	Position condition
1	Zhang, Scott, Felizola and Shoham	$Step^{L_S} = \text{anywhere}$

$$Step^{L_S} = \text{anywhere} \tag{7.46}$$

(c) **Felizola and Shoham (1995):**
The Felizola and Shoham correlation is:

$$L_S = 20d\left(1 + 1.5\frac{\theta}{90}\right) \tag{7.47}$$

$$Step^{L_S} = \text{anywhere} \tag{7.48}$$

These three correlations can be placed anywhere in the calculation algorithm. Therefore, all these correlations can be made part of **one** group, as shown in Table 7.5.

3. **Group - Parameter Taylor bubble translational velocity,** v_{TB}

(a) **Bendiksen (1984):** The Bendiksen correlation is:

$$v_{TB} = C_S v_M + v_{drift} \tag{7.49}$$

Table 7.6: There is only one group of correlations for the calculation of v_{TB}.

Group	Empirical correlations	Position condition
1	Bendiksen, Hassan and Kabir	$Step^{v_{TB}}$ = anywhere

$$v_{drift} = 0.54\sqrt{gd}\cos\theta + 0.35\sqrt{gd}\sin\theta \tag{7.50}$$

where C_S is 1.2 for turbulent flow and 2 for laminar flow.
5.1 Write the empirical parameter as a functional dependency.

$$v_{TB} = g\,() \tag{7.51}$$

5.4 v_{TB} placement restrictions:

$$Step^{v_{TB}} = \text{anywhere} \tag{7.52}$$

(b) **Hassan and Kabir (1992):** The Hassan and Kabir correlation for v_{TB} is:

$$v_{drift} = 0.35\sqrt{\frac{g\sigma_L\left(\rho_L - \rho_G\right)}{{\rho_L}^2}}\sqrt{\sin\theta}(1+\cos\theta)^2 \tag{7.53}$$

Similarly:

$$Step^{v_{TB}} = \text{anywhere} \tag{7.54}$$

There is **one** group of correlations for parameter v_{TB}, as shown in Table 7.6.

4. **Group - Parameter interfacial friction factor,** f_i

(a) **Cohen and Hanratty (1965):** The Cohen and Hanratty correlation for f_i is:

$$f_i = 0.0142 \tag{7.55}$$

$$Step^{f_i} = \text{anywhere} \tag{7.56}$$

(b) **Andritsos and Hanratty (1987):**
The Andritsos and Hanratty correlation for f_i is:

$$f_i = f_{GW}\left(1 + 14.3{H_{LTB}}^{0.5}\left(\frac{v_{SG}}{v_{SG,t}} - 1\right)\right) \tag{7.57}$$

$$v_{SG,t} = 5\left[\frac{m}{s}\right]\left(\frac{\rho_{GO}}{\rho_G}\right)^{0.5} \tag{7.58}$$

5.1 Write the empirical parameter as a functional dependency.

$$f_i = g\left(H_{LTB}, f_{GW}\right)$$

5.2 Determine the influencing factors in the superscript notation. Using the adjacency matrix, we determine that the influencing factors are:

$$\{f_{GW}^{14}, H_{LTB}^{1}\}$$

5.3 For each influencing factor determine the dependency chain.

Depth: $+1$, $MaxStep = 0$

$D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} > MaxStep$. Assign $MaxStep = 1$. Stop.
$D\left(f_{GW}^{14}\right) = D\left(Re_{GW}^{13}\right) = D\left(v_{GTB}^{5}, S_{G}^{12}, S_{i}^{12}, A_{G}^{12}\right)$. Explore further.

Depth: $+2$

$D\left(v_{GTB}^{5}\right) = D\left(H_{LTB}^{1}, v_{TB}^{3}, H_{LLS}^{4}\right)$. Explore further.

Depth: $+3$

$D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\left(v_{TB}^{3}\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB} + 3\}$
$D\left(H_{LLS}^{4}\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 3, v_{TB} + 3\}$

Depth: $+2$

$D\left(S_{G}^{12}\right) = D\left(\Theta^{11}\right)$: Empirical : Push Θ and depth to stack : $\{\Theta + 2, H_{LLS} + 3, v_{TB} + 3\}$
$D\left(S_{i}^{12}\right) = D\left(\Theta 1\right)$: Empirical : Push Θ and depth to stack : $\{\Theta + 2, H_{LLS} + 3, v_{TB} + 3\}$
$D\left(A_{G}^{12}\right) = D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop

5.4 Combining the results, we conclude that:

$$Step^{f_i} \geq \max\{2, Step^{\Theta}+2, Step^{H_{LLS}}+3, Step^{v_{TB}}+3\} \quad (7.59)$$

(c) **Asali and Ambrosio (1984):**
The correlation was used in Zhang step 14, in Section 5.3.2 and involves the following calculation algorithm:

Asali step 1 Assume $f_i = f_G$. For a smooth pipe f_G is defined as in Equation 7.60:

$$f_G = 0.046 {Re_G}^{-0.2} \tag{7.60}$$

The Weber and Reynolds numbers are calculated with:

$$We_G = \frac{\rho_G v_{SG}^2 d}{\sigma} \tag{7.61}$$

$$Re_G = \frac{\rho_G v_{SG} d}{\mu_G} \tag{7.62}$$

Asali step 2 Calculate τ_i from Equation 7.63:

$$\tau_i = f_i \frac{\rho_G \left(v_{SG} - v_{SL}\right) \left|v_{SG} - v_{SL}\right|}{2} \tag{7.63}$$

Asali step 3 Calculate v_C^* according to Equation 7.64:

$$v_C^* = \sqrt{\frac{\tau_i}{\rho_G}} \tag{7.64}$$

Asali step 4 Calculate h_F^+ from Equation 7.65:

$$h_F^+ = \frac{\rho_G h_F v_C^*}{\mu_G} \tag{7.65}$$

Asali step 5 Calculate $f_{i,calculated}$ using Equation 7.66:

$$f_i = f_G \left(1 + 13.8 {We_G}^{0.2} {Re_G}^{-0.6} \left(h_F^+ - 200\sqrt{\frac{\rho_G}{\rho_L}}\right)\right) \tag{7.66}$$

If $\frac{\left|f_{i,calculated} - f_i\right|}{f_{i,calculated}} < \epsilon$ the algorithm has converged and $f_{i,calculated}$ represents the value of f_i. Otherwise, repeat from **Asali step 2**, by assuming $f_i = f_{i,calculated}$.

$$Step^{f_i} = \text{anywhere} \tag{7.67}$$

(d) **Ouyang and Aziz (1996):** The Ouyang and Aziz correlation for f_i is:

$$f_i = 10^{-8.0942 + 4.2893 H_{LTB}^{\sin\theta}} \frac{f_{LW}^{0.8732} N_{v_L}^{0.3072} N_D^{1.0365}}{N_G^{1.914} \left(\frac{v_{GTB}}{v_{LTB}}\right)^{0.9783}} \tag{7.68}$$

where f_{LW} is the Fanning wall friction factor for the liquid phase:

$$R_{volumetric} = \frac{v_{SG}}{v_{SL}} \tag{7.69}$$

$$Re_{liquid} = \frac{\rho_L v_{LTB} H_{LTB} d}{\mu_L} \tag{7.70}$$

$$f_{LW} = 1.6291 \left({Re_{liquid}}^{-0.5161}\right) \left({R_{volumetric}}^{0.0926}\right) \tag{7.71}$$

The following dimensionless groups are defined:
Liquid velocity number:

$$N_{v_L} = v_{SL}\sqrt[4]{\frac{\rho_L}{g\sigma}} \tag{7.72}$$

Gas viscosity number:

$$N_G = \mu_G\sqrt[4]{\frac{g}{\rho_L\sigma^3}} \tag{7.73}$$

Pipe diameter number:

$$N_D = d\sqrt{\frac{\rho_L g}{\sigma}} \tag{7.74}$$

5.1 Write the empirical parameter as a functional dependency.

$$f_i = g\left(f_{LW}, v_{GTB}, v_{LTB}, H_{LTB},\ \right) \tag{7.75}$$

5.2 Determine the influencing factors in the superscript notation.

$$\{f_{LW}^{14}, v_{GTB}^{5}, v_{LTB}^{5}, H_{LTB}^{1}\}$$

5.3 For each influencing factor determine the dependency chain.

Depth: $+1$, $MaxStep = 0$

$D\left(f_{LW}^{14}\right) = D\left(Re_{LW}^{13}\right) = D\left(v_{LTB}^{5}, S_F^{12}, A_F^{12}\right)$. Explore further.

Depth: $+2$

$D\left(v_{LTB}^{5}\right) = D\left(H_{LTB}^{1}, v_{TB}^{3}, H_{LLS}^{4}\right)$. Explore further.

Depth: $+3$

$D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} > MaxStep$. Assign $MaxStep = 1$. Stop.
$D\left(v_{TB}^{3}\right)$: Empirical : Push v_{TB} and depth to stack :
$\{v_{TB} + 3\}$
$D\left(H_{LLS}^{4}\right)$: Empirical : Push H_{LLS} and depth to stack :
$\{H_{LLS} + 3, v_{TB} + 3\}$

Depth: $+2$

$D\left(S_F^{12}\right) = D\left(\Theta\right)$: Empirical : Push
Θ and depth to stack : $\{\Theta + 2, H_{LLS} + 3, v_{TB} + 3\}$
$D\left(A_F^1 2\right) = D\left(H_{LTB}^{1}\right) = D\left(\right)$. Stop.

Depth: $+1$

$D\left(v_{GTB}^{5}\right) = D\left(H_{LTB}^{1}, v_{TB}^{3}, H_{LLS}^{4}\right)$. Explore further.

Depth: +2

$D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\left(v_{TB}^{3}\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB}+2, \Theta+2, H_{LLS}+3, v_{TB}+3\}$
$D\left(H_{LLS}^{4}\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS}+2, v_{TB}+2, \Theta+2, H_{LLS}+3, v_{TB}+3\}$

Depth: +1

$D\left(v_{LTB}^{5}\right) = D\left(H_{LTB}^{1}, v_{TB}^{3}, H_{LLS}^{4}\right)$. Explore further.

Depth: +2

$D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\left(v_{TB}^{3}\right)$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB}+2, H_{LLS}+2, v_{TB}+2, \Theta+2, H_{LLS}+3, v_{TB}+3\}$
$D\left(H_{LLS}^{4}\right)$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS}+2, v_{TB}+2, H_{LLS}+2, v_{TB}+2, \Theta+2, H_{LLS}+3, v_{TB}+3\}$

Depth: +1

$D\left(H_{LTB}^{1}\right) = D\left(\right)$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.

5.4 f_i final placement restrictions:

$$Step^{f_i} \geq \max\{2, Step^{\Theta}+2, Step^{H_{LLS}}+3, Step^{v_{TB}}+3\} \quad (7.76)$$

(e) **Bendiksen et al. (1989):** The Bendiksen correlation for f_i calculates a Δh_{Wave} factor first.

$$\Delta h_{Wave} = \left[\frac{\rho_G (v_{SG}-v_{SL})^2}{4\left(\rho_L-\rho_G\right) g\cos\theta}\right] + \sqrt{\left[\frac{\rho_G (v_{SG}-v_{SL})^2}{4\left(\rho_L-\rho_G\right) g\cos\theta}\right]^2 - \frac{\sigma}{\left(\rho_L-\rho_G\right) g\cos\theta}} \quad (7.77)$$

$$f_i = 0.001375\left[1+\left(2\times 10^4 \frac{\Delta h_{Wave}}{d}+\frac{10^6}{Re_G}\right)0.333\right] \quad (7.78)$$

$$Re_G = \rho_G \frac{|v_{GTB} - v_{LTB}|\,(1 - H_{LTB})\,d}{\mu_G} \tag{7.79}$$

5.1 Write the empirical parameter as a functional dependency.

$$f_i = g\,(H_{LTB}, v_{GTB}, v_{LTB}\ ) \tag{7.80}$$

5.2 Determine the influencing factors in the superscript notation.

$$\{H^1_{LTB}, v^5_{GTB}, v^5_{LTB},\}$$

5.3 For each influencing factor determine the dependency chain.

Depth: $+1$, $MaxStep = 0$

$D\,(H^1_{LTB}) = D\,()$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} > MaxStep$. Assign $MaxStep = 1$. Stop.
$D\,(v^5_{GTB}) = D\,(H^1_{LTB}, v^3_{TB}, H^4_{LLS})$. Explore further.

Depth: $+2$

$D\,(H^1_{LTB}) = D\,()$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\,(v^3_{TB})$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB} + 2\}$
$D\,(H^4_{LLS})$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 2, v_{TB} + 2\}$

Depth: $+1$

$D\,(v^5_{LTB}) = D\,(H^1_{LTB}, v^3_{TB}, H^4_{LLS})$. Explore further.

Depth: $+2$

$D\,(H^1_{LTB}) = D\,()$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} = MaxStep$. Keep $MaxStep = 1$. Stop.
$D\,(v^3_{TB})$: Empirical : Push v_{TB} and depth to stack : $\{v_{TB} + 2, H_{LLS} + 2, v_{TB} + 2\}$
$D\,(H^4_{LLS})$: Empirical : Push H_{LLS} and depth to stack : $\{H_{LLS} + 2, v_{TB} + 2, H_{LLS} + 2, v_{TB} + 2\}$

5.4 f_i final placement restrictions are:

$$Step^{f_i} \geq \max\{2, Step^{H_{LLS}} + 2, Step^{v_{TB}} + 2\} \tag{7.81}$$

(f) **Vlachos (1997):** The Vlachos correlation is:

$$f_i = 0.024 {H_{LTB}}^{0.35} {Re_{SL}}^{0.18} \tag{7.82}$$

Table 7.7: Groups of correlations for the calculation of f_i.

Grp	Empirical correlations	Position conditions
1	Cohen and Hanratty, Asali and Ambrosio	$Step^{f_i}$ = anywhere
2	Andritsos and Hanratty, Ouyang and Aziz	$Step^{f_i} \geq \max\{2, Step^{\Theta}+2, Step^{H_{LLS}}+3, Step^{v_{TB}}+3$
3	Vlachos	$Step^{f_i} \geq 2$
4	Bendiksen	$Step^{f_i} \geq \max\{2, Step^{H_{LLS}}+2, Step^{v_{TB}}+2\}$

where the Reynolds number is:

$$Re_{SL} = \frac{\rho_L v_{SL} d}{\mu_L} \tag{7.83}$$

f_i final placement restrictions:

$$Step^{f_i} \geq 2 \tag{7.84}$$

There are **four** different groups of correlations for f_i, as shown in Table 7.7.

5. **Group - Parameter wetted wall fraction, Θ**

   (a) **Grolman (1994):** The Grolman correlation is:

$$\Theta = \Theta_0 \left(\frac{\sigma_{water}}{\sigma}\right)^{0.15} + \frac{\rho_G}{(\rho_L - \rho_G)} \frac{1}{\cos\theta} \left(\frac{\rho_L {v_{SL}}^{0.25} d}{\sigma}\right)^{0.25} \left[\frac{{v_{SG}}^2}{(1-H_{LTB})^2 g d}\right]^{0.8} \tag{7.85}$$

$$\Theta_0 = 0.624 (H_{LTB})^{0.374} \tag{7.86}$$

5.1 Write the empirical parameter as a functional dependency.

$$\Theta = g\,(H_{LTB}) \tag{7.87}$$

5.2 Determine the influencing factors in the superscript notation.

$\{H^1_{LTB}\}$

5.3 For each influencing factor determine the dependency chain.

Depth: $+1$, $MaxStep = 0$

$D\,(H^1_{LTB}) = D\,()$. $S^{H_{LTB}} = 1$. $S^{H_{LTB}} > MaxStep$. Assign $MaxStep = 1$. Stop.

5.4 Θ final placement restrictions:

$$Step^{\Theta} \geq 2 \tag{7.88}$$

(b) **Biberg (1999):** The Biberg correlation is:

$$\beta = 2\pi - 2\left\{\pi H_{LTB} + \left(\frac{3\pi}{2}\right)^{\frac{1}{3}}\left[1 - 2H_{LTB} + {H_{LTB}}^{\frac{1}{3}} - (1 - H_{LTB})^{\frac{1}{3}}\right]\right\} \tag{7.89}$$

5.1 Write the empirical parameter as a functional dependency.

$$\Theta = g\left(H_{LTB}\right) \tag{7.90}$$

5.4 Θ final placement restrictions:

$$Step^{\Theta} \geq 2 \tag{7.91}$$

(c) **Fan (2005):** The Fan correlation is:

$$\Theta = \left[0.57{H_{LTB}}^{0.345} + 0.0637{Fr_L}^{0.68}\left(\frac{v_{SG}}{v_{SG,C}}\right)^{0.68}\right]\left(\frac{\sigma_W}{\sigma}\right)^{0.15} \tag{7.92}$$

for values of $\Theta \in [0, 0.5]$.

$$\Theta = \left[0.57{H_{LTB}}^{0.345} + 0.0637{Fr_L}^{0.68}\left(\frac{v_{SG}}{v_{SG,C}}\right)^{0.55}\right]\left(\frac{\sigma_{water}}{\sigma}\right)^{0.15} \tag{7.93}$$

for values of $\Theta \in (0.5, 1]$.
The critical value of the superficial gas velocity, $v_{SG,C}$ is:

$$v_{SG,C} = 5\left[\frac{m}{s}\right]\left[\frac{1.24}{\rho_G\left[\frac{\text{kg}}{m^3}\right]}\right]^{0.5} \tag{7.94}$$

$$Fr_L = \frac{\rho_L {v_{SL}}^2}{(\rho_L - \rho_G)\, gd\cos\theta} \tag{7.95}$$

5.1 Write the empirical parameter as a functional dependency.

$$\Theta = g\left(H_{LTB}\right) \tag{7.96}$$

5.4 Θ final placement restrictions:

Table 7.8: There is only one group of correlations for the calculation of Θ.

Group	Correlations	Position condition
1	Grolman, Biberg, Fan, Hart	$Step^{\Theta} \geq 2$

$$Step^{\Theta} \geq 2 \tag{7.97}$$

(d) **Hart et al. (1989):** The Hart correlation is:

$$\Theta = \min\left\{1, \Theta_0 + CFr^{0.58}\right\} \tag{7.98}$$

$$\Theta_0 = 0.624(H_{LTB})^{0.374} \tag{7.99}$$

where the constant $C = 0.26$ and the Froude number is:

$$Fr = \frac{\rho_L {v_{SL}}^2}{(\rho_L - \rho_G)\, gd} \tag{7.100}$$

5.1 Write the empirical parameter as a functional dependency.

$$\Theta = g\,(H_{LTB}\,) \tag{7.101}$$

5.4 Θ final placement restrictions:

$$Step^{\Theta} \geq 2 \tag{7.102}$$

There is **one** group of empirical correlations for parameter Θ, as shown in Table 7.8.

Step 6: Minimise the number of calculation algorithms. The correlations for each empirical parameter were grouped based on their position restrictions in the algorithm. Table 7.9 lists the number of groups of equivalent correlations for each empirical parameter in the Zhang et al. (2000) and (2003) models.

Based on the number of groups, the total number of distinct flow calculation algorithms may be as high as:

$$\begin{aligned}\text{Number calculation algorithms} &= \text{Groups } H_{LLS} \times \text{Groups } L_S \\ &\quad \times \text{Groups } v_{TB} \times \text{Groups } f_i \\ &\quad \times \text{Groups } \Theta \\ &= 4 \times 1 \times 1 \times 4 \times 1 = 16\end{aligned} \tag{7.103}$$

This is a significant reduction from the initial number:

$$\text{Number calculation algorithms} = 3 \times 10 \times 2 \times 6 \times 4 = 1440 \tag{7.104}$$

Table 7.9: Number of groups of equivalent correlations for each empirical parameter in the Zhang et al. (2000) model.

Empirical parameter	Number of groups
H_{LLS}	4
L_S	1
v_{TB}	1
f_i	4
Θ	1

that might be expected from a preliminary consideration of Table 7.1.

To further reduce the number of algorithms, we need to:

6.1 Identify the empirical parameters that have a number of groups ≥ 2:

These parameters are H_{LLS} and f_i. They lead to a high number of calculation algorithms.

6.2 For each one of these parameters, reduce the number of groups to one, by making all the correlations part of the group with the highest step number.

1. $\boldsymbol{H_{LLS}}$ **:** Four groups of correlations for parameter H_{LLS} were determined, as shown in Table 7.4. The solution to reduce the number of groups for parameter H_{LLS} to one, is to make all the groups part of the same group, by placing them after the maximum step required by any correlation in these four groups.

   Table 7.4 shows a dependency of H_{LLS} on H_{LLS}. As mentioned earlier, the second H_{LLS} refers to the initial estimation of H_{LLS}, used to start the calculations. The first H_{LLS} is used as a more "accurate" correlation, to recalculate H_{LLS}. Consequently, we need to split the correlations for H_{LLS} into two groups:

   Group 1: Woldesmayat and Ghajar, Toshiba, Gomez, Spedding and Spence, Minami and Brill, Gregory, Nicklin. These are the correlations in group 1 in Table 7.4 and do not have placement restrictions. We choose to keep Zhang step 4 (Section 5.3.2) for the initial estimation of H_{LLS}.

   Group 2: Zhang, Barnea and Brauner and Eissa Al-Safran. According to the conditions in Table 7.4, these correlations require they be placed at a step that needs to be determined based on $Step^{L_S}$, $Step^{v_{TB}}$ and $Step^{H_{LLS}}$. As H_{LLS} for the initial estimation was decided to be placed at Zhang step 4, we now need to determine the positions of $Step^{L_S}$ and $Step^{v_{TB}}$.

   - According to the result in Table 7.5, L_S correlations do not have any requirements in terms of where L_S has to be placed.
   - Similarly, Table 7.6 shows that v_{TB} correlations do not have placement restrictions.

Therefore, we place:

- L_S at Zhang step 2, which is the same position as in the original Zhang et al. (2000) algorithm.
- v_{TB} at Zhang step 3, keeping the same position.
- H_{LLS} initial estimation at the same Zhang step 4, using the correlations of Woldesmayat and Ghajar, Toshiba, Gomez, Spedding and Spence, Minami and Brill, Gregory, Nicklin.

To make all H_{LLS} correlations part of the same group we need to find the maximum step required by the correlations:

$$\begin{cases} Step^{BarneaandBrauner} \geq \{Step^{H_{LLS}} + 2\} \\ Step^{EissaAl-Safran} \geq \max\{2, Step^{v_{TB}} + 2, Step^{H_{LLS}} + 2\} \\ Step^{Zhang} \geq \max\{2, Step^{L_S} + 1, Step^{v_{TB}} + 2, Step^{H_{LLS}} + 2\} \end{cases} \tag{7.105}$$

which, given the assignments for L_S, v_{TB} and initialisation of H_{LLS}, leads to:

$$\begin{cases} Step^{BarneaandBrauner} \geq \{6\} \\ Step^{EissaAl-Safran} \geq \max\{2, 5, 6\} \\ Step^{Zhang} \geq \max\{2, 3, 5, 6\} \end{cases} \tag{7.106}$$

Therefore, if we place H_{LLS} at a step greater than 6, we could make all the H_{LLS} correlations part of the same group.

These considerations lead to the following placement of the correlations for H_{LLS} and L_S:

Zhang step 2: L_S calculated by all the selected correlations.

Zhang step 3: v_{TB} calculated by all the selected correlations.

Zhang step 4: H_{LLS} calculated with any of the correlations:

- Woldesmayat and Ghajar
- Toshiba
- Gomez
- Spedding and Spence
- Minami and Brill
- Gregory
- Nicklin

Zhang step 6 onwards: H_{LLS} calculated with any of the correlations, so long as it is not the same correlation used for the initial estimation. That is, the condition is that the correlations in step 4 and 6 be different.

2. $\boldsymbol{f_i}$: Four groups of correlations for f_i were identified, as shown in Table 7.7. In order to make all the correlations part of the same group, we need to determine $Step^{\Theta}$.

Table 7.10: Empirical correlation groups and the position of the empirical parameter they estimate in CA^{new}.

Parameter	Final position	Empirical correlations
H_{LLS}	**Zhang step 4**	Woldesmayat and Ghajar, Toshiba, Gomez, Spedding and Spence, Minami and Brill, Gregory, Nicklin
H_{LLS}	**Zhang step 9**	All the correlations in Table 7.4, on the condition that the correlation in Zhang step 9 is not the same as in Zhang step 4
L_S	**Zhang step 2**	All the correlations in Table 7.5
v_{TB}	**Zhang step 3**	All the correlations in Table 7.6
f_i	**Zhang step 14**	All the correlations in Table 7.7
Θ	**Zhang step 11**	All the correlations in Table 7.8

We have already established that $Step^{H_{LLS}} = 4$ and $Step^{v_{TB}} = 3$. According to Table 7.8, all the correlations for Θ have the requirement $Step^{\Theta} \geq 2$, therefore we keep the existing position for the calculation of Θ, at Zhang step 11. With these selections, the conditions in Table 7.7 can be stated as:

$$Step^{f_i} \geq \max \begin{cases} \{2, (11+2), (4+3), (2+3)\} \\ \{2, (4+2), (3+3)\} \end{cases} \tag{7.107}$$

which results in:

$$Step^{f_i} \geq 13 \tag{7.108}$$

Therefore, keeping the existing position Zhang step 14 for the calculation of f_i makes all correlations part of the same group.

To summarise, we have determined the positions in the calculation algorithm for the selection of empirical correlations in Table 7.1. These positions are shown in Table 7.10.

7.7 Summary

This chapter presented an **empirical correlation replacement algorithm** that:

- For a given mechanistic model with a Calculation Algorithm (CA), and
- For an empirical correlation set chosen for each empirical parameter in the model,

determines a new CA such that the empirical correlations can be swapped without modifying the new CA. The algorithm is based on a graph representation of the CA.

The empirical correlation replacement algorithm was designed to address the problem of replacing an empirical correlation used in the mechanistic model, with

an alternative empirical correlation. The algorithm determines the step in the calculation algorithm where the empirical parameter calculated by the alternative correlation can be placed. The algorithm can be used in all the situations where it is necessary to replace an empirical correlation in a model:

- Using an alternative correlation that matches better the local production data, instead of a more generic empirical correlation used by the model.
- Replacing the empirical correlation with a set of mechanistic equations, when the phenomenon described initially by the correlation becomes better understood.

However, in this study the empirical correlation replacement algorithm is used as a supporting algorithm for the optimisation method described in Section 6.4.

The algorithm was demonstrated on the Zhang et al. (2000) and (2003) slug flow models. Five sets of empirical correlations for parameters L_S, H_{LLS}, v_{TB}, f_i and Θ were selected. The number of correlations in each set was shown in Table 7.1. According to Equation 7.104, the initial assessment of the number of possible calculation algorithms was 1440. The empirical correlation replacement algorithm presented here and demonstrated using the Zhang et al. (2000) and (2003) models, reduced this number to just 1.

Chapter 8

Mechanistic model training using alternative empirical correlations

8.1 Introduction

Section 6.4 stated that the first step in the model optimisation using alternative empirical correlation sets was the **training phase**. The goal of this chapter is to determine the **best empirical correlation set** for a **training data set**, which consists of input data points and their associated output measured values. A **curve fitting method using alternative correlations** is proposed for this purpose. The method is applied to the Zhang et al. (2000) slug flow model and the best empirical correlation set is determined for the data set selected in Section 2.5, which is used as a training data set.

8.2 Curve fitting method using alternative empirical correlations

There are many curve fitting methods used for interpolation in numerical analysis, such as polynomial, spline, rational and Gaussian. This chapter will describe a procedure to determine the **best empirical correlation set for each input training data point** x_i, with the view of using it for interpolation in the vicinity of that point. In mathematical terms, the method will generate piecewise functions, such that in the vicinity of the input data point x_i, the same empirical functions will be used for empirical parameter estimations. As the input moves towards x_{i+1}, a possibly different set of functions corresponding to another set of empirical correlations will be used by the model. Figures 8.1 and 8.2 illustrate the concept of selecting one empirical correlation out of a pool of correlations, for each empirical model parameter and the use of one correlation set in the vicinity of each input data point.

Unlike other numerical interpolation methods, which are not based on modelling the physical phenomena and only aim to determine a continuous function to run through the data points, the method presented here offers the advantage of still being based on a mechanistic model, which means that if the input points are too far apart, it can still predict the output with a certain level of accuracy.

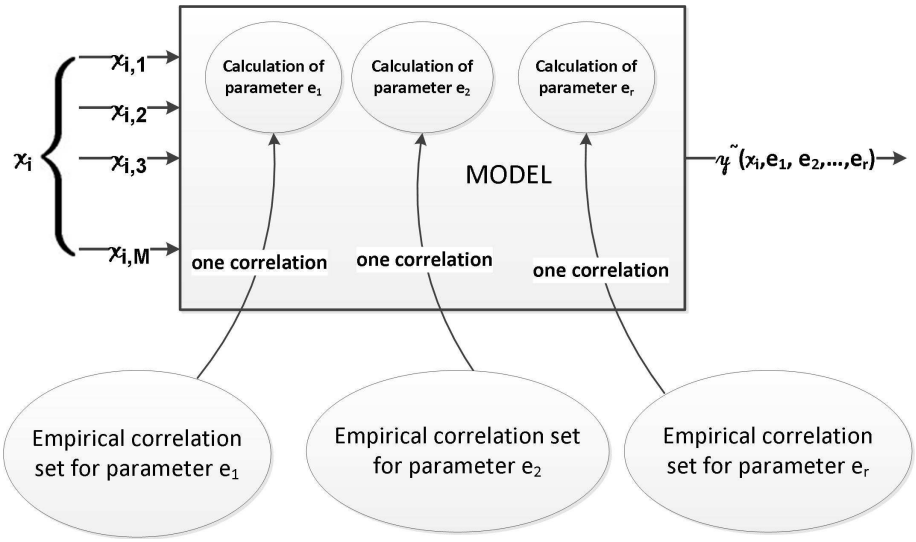


Figure 8.1: A mechanistic model showing the dependency of the model output on empirical correlations. This study will analyse the effect on the model predictions of selecting one empirical correlation for each empirical parameter e_j, $j = 1, 2, \ldots, r$, from a predefined pool of correlations.

Let us consider an input data set $(x_1, x_2, \ldots, x_n)$, a "true" model whose measured outputs are $(y_{t,1}, y_{t,2}, \ldots, y_{t,n})$, and a mechanistic model whose estimations are $(\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n)$. Each input data point is a vector with M components: $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,M}), i = 1, 2, \ldots, n$. The problem discussed in this chapter is to find the best empirical correlations for the parameters $(e_1, e_2, \ldots, e_r)$ in the "black" boxes of the mechanistic model, for each individual input data point x_i, such that the residuals $R_i = R(x_i)$ are minimal.

The purpose of the method is to determine the **best empirical correlation set for each input data point** x_i and save this information into a **history table**. To establish which empirical correlation set fits best the curve of the "true" values in Figure 8.2, the model estimation $\tilde{y}_i$ is compared to the measured data $y_{t,i}$ and the one that generates a minimum variance is associated with the data point x_i. Expressed in terms of residuals, the empirical correlation set that minimises the residual R_i, for an input data point x_i, is considered the best. The notation used for the best correlation set at point x_i is $O(x_i)$.

The curve fitting method replaces one empirical correlation at a time. The model estimations and residuals are calculated for each input data point x_i. The 3-tuple:

$$\langle x_i, O(x_i), R_{min}(x_i) \rangle$$

where:

$$\begin{aligned} O(x_i) &= \text{ optimal correlation set} \\ R_{min}(x_i) &= \text{ the minimum residual at } x_i \end{aligned}$$

is stored with the view of being used during the **prediction** stage. The collection of all these records constitutes the **history table**. The curve fitting method is the implementation of the **training** phase in the optimisation process described in Section 6.4.

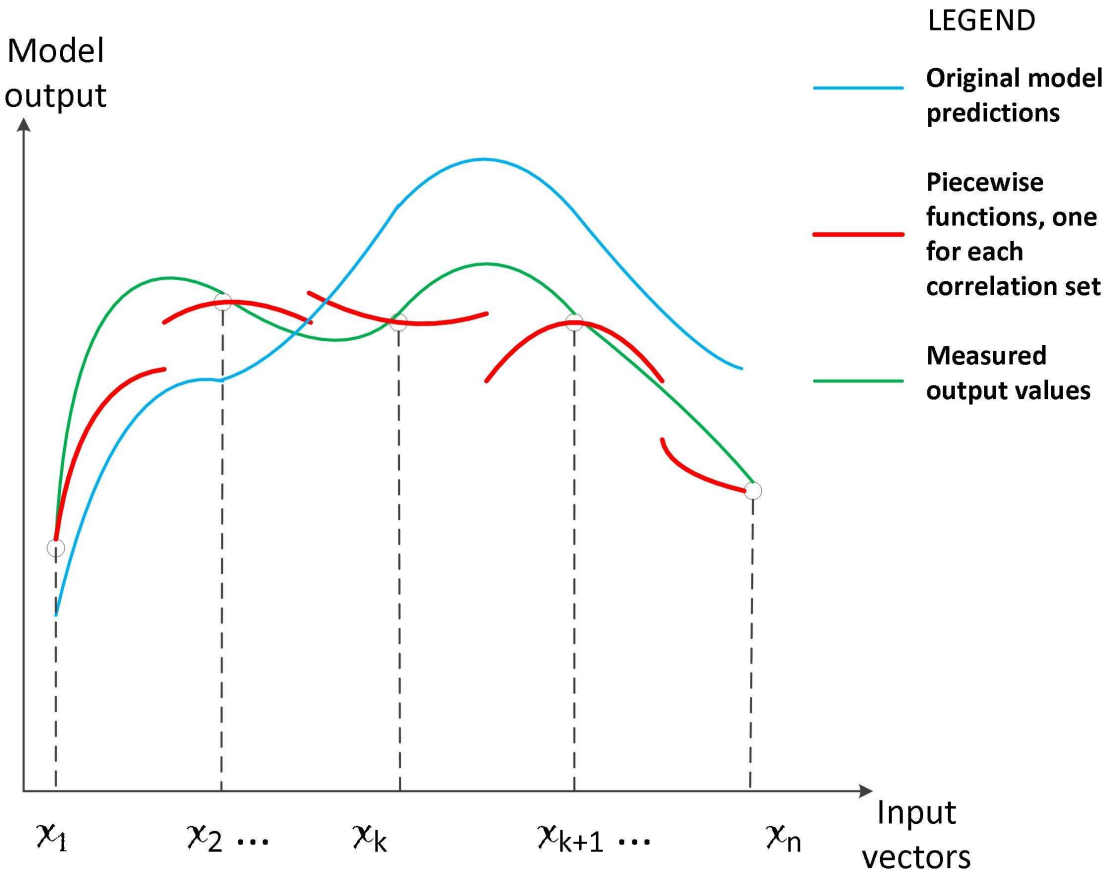


Figure 8.2: Piecewise functions are curve-fitting the "true model". Each piecewise function represents the model prediction for the best empirical correlation at input $x_k = (x_{k,1}, x_{k,2}, \ldots, x_{k,M})$ and around its vicinity.

As the curve fitting method forms a new correlation set every time a correlation has been replaced in the calculation algorithm, the empirical parameters calculated with the new combination of correlations, have to be placed at different steps in the calculation algorithm. This is where the results of Chapter 7 are used. The **correlation replacement algorithm** described in Chapter 7 determined the position in the calculation algorithm where the empirical parameters could be placed, depending on the empirical correlation used for their estimation. Positioning the empirical parameters is a pre-requisite for the use of the curve fitting method. Figure 8.3 summarises the building blocks involved in constructing the **history table**.

8.3 Case study - Building a history database for the Zhang et al. slug flow model

The **curve fitting method using alternative correlations** was applied to the Zhang et al. (2000) model to the end of determining the best empirical correlation set for each input data point. The 95 slug flow input data points from the TUFFP data bank used for the simulation are vectors with the following components:

$$x_i = (x_{i,1}, x_{i,2}, x_{i,3}, x_{i,4}, x_{i,5}, x_{i,6}, x_{i,7}, x_{i,8}, x_{i,9}) \qquad i = 1, 2, \ldots, n \qquad M = 9$$

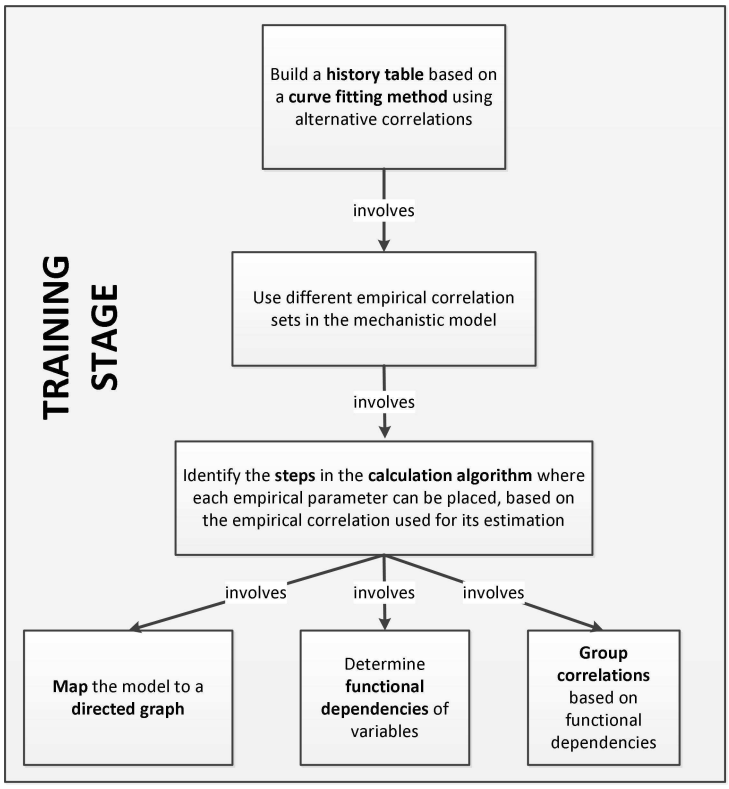


Figure 8.3: Top down structure showing the dependency relationships between the building blocks involved in optimising a mechanistic model.

$$\begin{array}{lll} x_{i,1} = d & x_{i,2} = \beta & x_{i,3} = v_{SL} \\ x_{i,4} = v_{SG} & x_{i,5} = \rho_L & x_{i,6} = \rho_G \\ x_{i,7} = \mu_L & x_{i,8} = \mu_G & x_{i,9} = \sigma \end{array}$$

The empirical correlation vector $\Phi = (g_1, g_2, \ldots, g_6)$, $r = 6$ used for the simulation was:

g_1 is one correlation for L_S
from the set: Zhang, Scott, Felizola and Shoham.

g_2 is one correlation for v_{TB}
from the set: Bendiksen, Hassan and Kabir.

g_3 is one correlation for H_{LLS}
from the group 1 set: Gregory, Gomez, Woldesmayat and Ghajar, Toshiba and Leung, Spedding and Spence, Minami and Brill, Nicklin.

g_4 is one correlation for H_{LLS}
from the group 2 set: Gregory, Gomez, Woldesmayat and Ghajar, Toshiba and Leung, Spedding and Spence, Minami and Brill, Nicklin, Eissa Al-Safran, Barnea and Brauner, Zhang

provided that g_4 is not the same correlation as g_3.

g_5 is one correlation for Θ
from the set: Grolman, Biberg, Fan, Hart.

g_6	is one correlation for f_i from the set:	Cohen and Hanratty, Andritsos and Hanratty, Vlachos, Ouyang and Aziz, Bendiksen, Asali and Ambrosio.

The position in the calculation algorithm of the empirical parameters, which is dependent on the empirical correlation, has to be determined before applying the curve fitting method to the Zhang et al. (2000) model. This was established in Chapter 7, and Table 7.10 shows the steps in the calculation algorithm where the parameters can be placed. Based on these results, the curve fitting method has the form illustrated by the pseudocode in Algorithm 8.1.

The GNU Octave program in the *Model Training* folder on the attached CD implements the pseudocode. The output of running the simulation program on the data set is the history table and is shown in Table A.9, Appendix A. The numerical codes associated with the empirical correlations are given in Tables A.2-A.6, Appendix A.

8.4 Data mining techniques for result interpretation

The **curve fitting method using alternative correlations** determined the best empirical correlation set for each **individual** input data point in the **training data** set. These correlations are shown in the **history table** in Table A.9, Appendix A. However, rather than knowing the best correlation at the granular level of the individual point, we are also interested to know whether there are any **patterns** where the same correlation set performs best.

Data mining is an effective process in the attempt to find such patterns in data sets. This is useful for our purpose of determining **regions** where a certain empirical correlation set performs best. This undertaking aims to show that organising the input data in certain ways, such as grouping and sorting by certain keys may reveal regions, or patterns of regions, where only one empirical correlation set performs best and is therefore, **optimal**.

The input data used in the case study was sorted in ascending order of v_{SL}. Using the threshold $v_{SL} = 0.1$ m/s, the data were then separated into two groups corresponding to values smaller and greater than 0.1 m/s respectively. Both groups were then further partitioned according to whether the inclination angle β was zero or not. This created four data sets that were analysed for regions where one correlation set prevailed. This sorting and partitioning procedure is illustrated in Figure 8.4, which emphasizes that the correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$, $\Theta^c = 1$ is the dominant optimal correlation set in the series of values $v_{SL} < 0.1$ m/s and $\beta = 0$. The notation $parameter^c$ refers to the empirical correlation number in Tables A.2-A.6 used for the calculation of the *parameter*.

Table A.10 in Appendix A shows the data organised by ascending values of v_{SL} and then grouped by pipe inclination angles β. Figure 8.4 is a simplification of the data in Table A.10.

This procedure to organise the data is not random. Sensitivity analysis calculations in Section 9.3.1 show that v_{SL} is the most influential parameter on the

Algorithm 8.1 The pseudocode of the curve fitting method applied to the Zhang et al. calculation algorithm. The purpose of the method is to build a history table.

procedure TRAINING($X, \Omega_1, \Omega_2, \Omega_3, \Omega_4, \Omega_5, \Omega_6$)
    REQUIRES: The step in the calculation algorithm where the empirical parameters are calculated, depending on the correlation used for their calculation.
    INPUT: $X = \{x_1, x_2, \ldots, x_n\}$
    The sets of correlations Ω_1,Ω_2,Ω_3,Ω_4,Ω_5,Ω_6
    OUTPUT: History database. Each record is $\langle x_i, O(x_i), R_{min}(x_i)\rangle$
    ---
    for all $x_i \in X$ **do**
        for all $g_{LS} \in \Omega_1$ **do**
            for all $g1_{H_{LLS}} \in \Omega_2$ **do**
                for all $g_{v_{TB}} \in \Omega_3$ **do**
                    for all $g_{f_i} \in \Omega_4$ **do**
                        for all $g_\Theta \in \Omega_5$ **do**
                            for all $g2_{H_{LLS}} \in \Omega_2$ **do**
                                Zhang step 1: Guess H_{LTB}
                                Zhang step 2: Calculate $L_S = g_{L_S}(x_i)$
                                Zhang step 3: Calculate $v_{TB} = g_{v_{TB}}(x_i)$
                                Zhang step 4: Calculate $H_{LLS} = g1_{H_{LLS}}(x_i)$
                                Zhang step 5: Calculate v_{TB} and v_{GTB}
                                Zhang step 6: Calculate ρ_S
                                Zhang step 7: Calculate Re_S
                                Zhang step 8: Calculate f_S
                                if $g2_{H_{LLS}} \neq g1_{H_{LLS}}$ **then**
                                    Zhang step 9: Calculate $H_{LLS} = g2_{H_{LLS}}(x_i)$
                                else
                                    Continue
                                end if
                                Zhang step 10: Calculate L_F and L_U
                                Zhang step 11: Calculate $\Theta = g_\Theta(x_i)$
                                Zhang step 12: Calculate S_F, A_F, S_G, A_G, S_{CD}, A_{CD} and S_i
                                Zhang step 13: Calculate Re_{LW} and Re_{GW}
                                Zhang step 14: Calculate f_{LW}, f_{GW} and $f_i = g_{f_i}(x_i)$
                                Zhang step 15: Calculate τ_{GW}, τ_{LW} and τ_i
                                Zhang step 16: Calculate m_{err} = margin of error

Zhang step 17:
if $m_{err} \geq \epsilon$ **then**
Go To **Zhang step 5**
else
Calculate residual:
$R(x_i) = |dP/dL - (dP/dL)_{measured}|$
if $R(x_i) < R_{min}(x_i)$ **then**
$R_{min}(x_i) = R(x_i)$
Create a record in the history table
Save: x_i,$g_{LS}, g1_{H_{LLS}}, g_{v_{TB}}, g_{f_i}, g_{\Theta}, g2_{H_{LLS}}$
Save: $R_{min}(x_i)$
end if
end if
end for
end for
end for
end for
end for
end for
end procedure

model predictions. The second most influential parameter is the inclination angle β. This explains why the data were organised by taking into account these two factors, in the order of their influence.

Figures 8.6-8.8 represent the occurrence fraction of each of the optimal correlations for the parameters L_S, H_{LLS}, v_{TB}, f_i and Θ, in the four series of data: ($v_{SL} < 0.1$ m/s and $\beta = 0$), ($v_{SL} \geq 0.1$ m/s and $\beta = 0$), ($v_{SL} < 0.1$ m/s and $\beta \neq 0$) and ($v_{SL} \geq 0.1$ m/s and $\beta \neq 0$).

Figure 8.5 shows the distribution of the correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ over the interval $v_{SL} < 0.1$ m/s and $\beta = 0$. This correlation set offers **better prediction capabilities in this domain** than all the other correlation sets and is therefore **optimal**.

Other patterns of correlations were investigated, but as Figures 8.6-8.8 show, no obvious regions where one pattern dominated could be determined.

8.4.1 Comparison between the optimal empirical correlation set and the set in the Zhang et al. (2000) slug flow model

The Zhang et al. (2000) used the empirical correlation set in Table 8.1. This is the optimal correlation set for $v_{SL} < 0.1$ m/s and $\beta = 0$. Consequently, we can ascertain that the Zhang et al. (2000) model **is already optimised in the domain of horizontal flow at low superficial liquid velocities**.

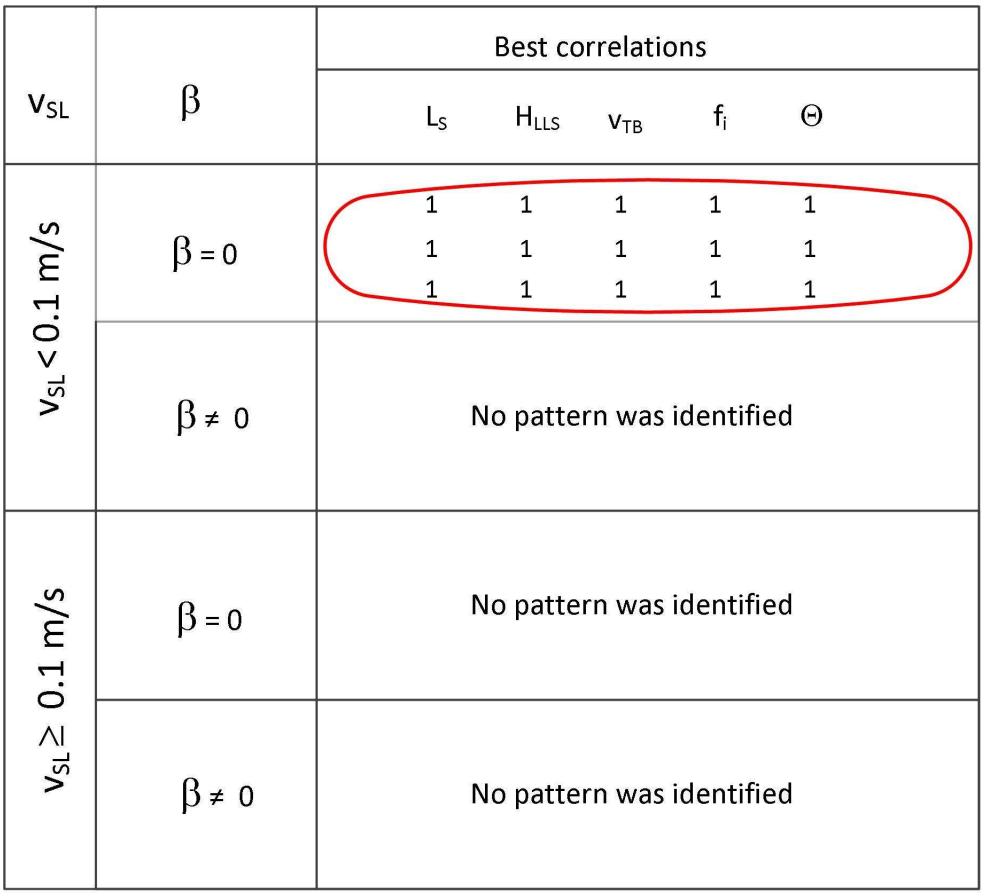


Figure 8.4: Sorting and partitioning the input data revealed the existence of the prevailing optimal set of empirical correlations $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$, $\Theta^c = 1$. This is a simplification of the data in Table A.10, for the purpose of revealing a region with an optimal correlation set.

Table 8.1: Empirical parameters and the empirical correlations used in the Zhang et al. (2000) slug flow model.

Empirical parameter	Empirical correlation	Correlation code
v_{TB}	Bendiksen	1
H_{LLS}	Gregory	1
L_S	Zhang	1
Θ	Grolman	1
f_i	Cohen and Hanratty	1

8.4.2 Comparison between the optimal correlation set and the set in the Zhang et al. (2003) slug flow model

The Zhang et al. (2003) used the empirical correlation set in Table 8.2. The two sets are different in the correlations for H_{LLS} and f_i.

This finding is surprising. The Zhang et al. (2003a) correlation for slug liquid holdup H_{LLS} was developed in 2003 and was based on the balance between the turbulent kinetic energy of the liquid phase and the surface energy of the dispersed

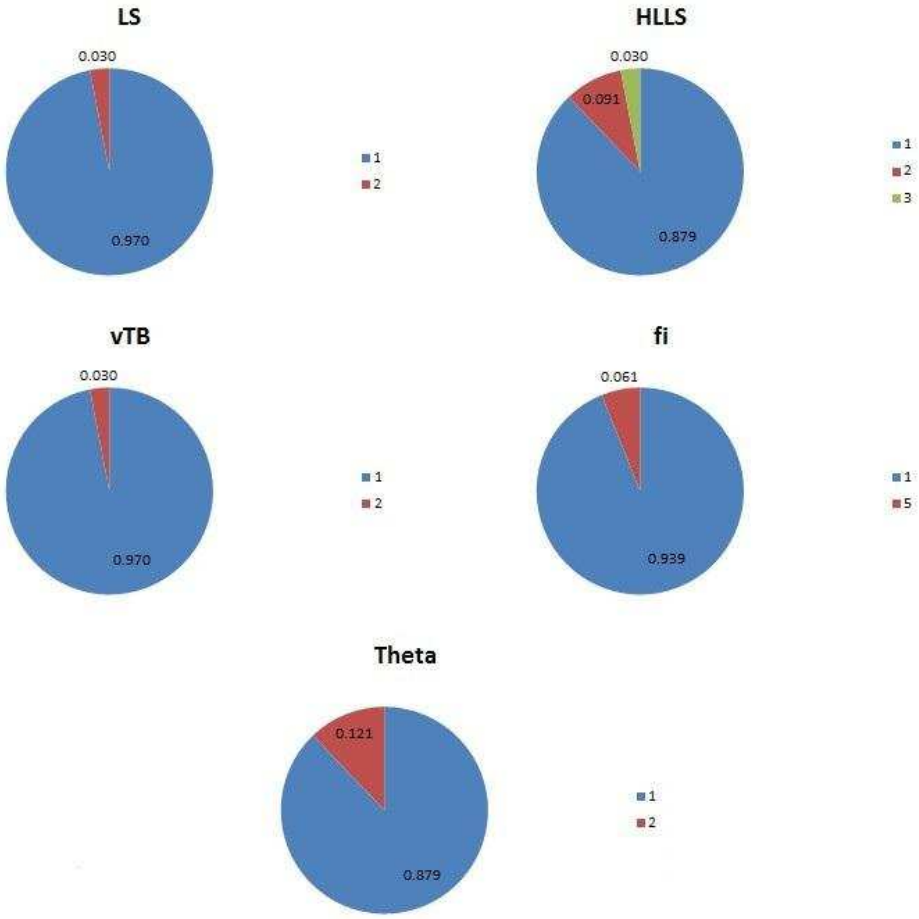


Figure 8.5: The correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ is optimal for horizontal flow ($\beta = 0$) in the region of values of $v_{SL} < 0.1$ m/s.

Table 8.2: Empirical parameters and the empirical correlations used in the Zhang et al. (2003) slug flow model.

Empirical parameter	Empirical correlation	Correlation code
v_{TB}	Bendiksen	1
H_{LLS}	Gregory, Zhang	1,10
L_S	Zhang	1
Θ	Grolman	1
f_i	Andritsos and Hanratty, Asali and Ambrosio	2,6

spherical gas bubbles, unlike the Gregory correlation which was developed in 1977 and had a simpler approach to calculating H_{LLS}. In fact, the Zhang et al. (2003) flow model uses the Gregory (1987) correlation as a "less accurate", initial estimate for H_{LLS} and refines the value of the liquid holdup by using its own Zhang et al. (2003a) H_{LLS} correlation, during the iterations of the calculation algorithm. The Gregory correlation holds for horizontal slug flow, while the Zhang et al. H_{LLS} correlation was developed for both horizontal and inclined flow. Hence, **both correlations were in their validity domain**. Both correlations are

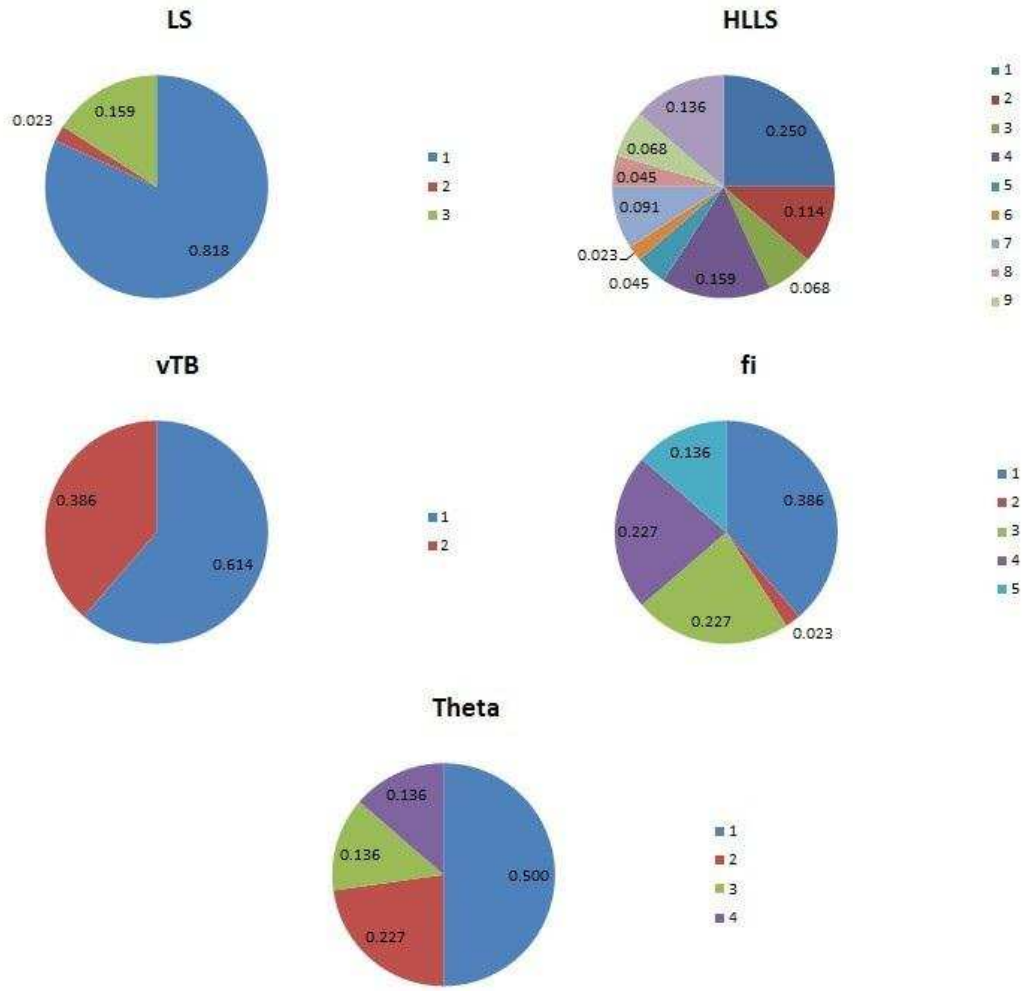


Figure 8.6: The correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ is not optimal for horizontal flow ($\beta = 0$) in the region of values of $v_{SL} \geq 0.1$ m/s.

described in Section F.3. These results prove that for certain flow conditions of horizontal flow at low superficial liquid velocity, the Gregory correlation for liquid holdup is more accurate than the Zhang et al. correlation. More importantly, this emphasizes the **data sensitivity** of mechanistic models and that **the embedded empirical correlation should be chosen according to the data fit on the input domain**.

The correlations for f_i used in the Zhang et al. (2003) model depend on the flow regime in the liquid film area of the slug unit. The model used the following correlations for f_i:

- Andritsos and Hanratty (1987) in the case of stratified flow.
- Asali and Ambrosio (1984) in the case of annular flow.

The Cohen and Hanratty (1968) correlation for the interfacial friction factor f_i was developed for stratified flow, for a fully developed rough gas-liquid interface, corresponding to small liquid waves. Based on empirical experiments, the correlation simply suggests that the best estimation for the interfacial friction factor is $f_i = 0.0142$. On the other hand, the Andritsos and Hanratty (1987) correlation

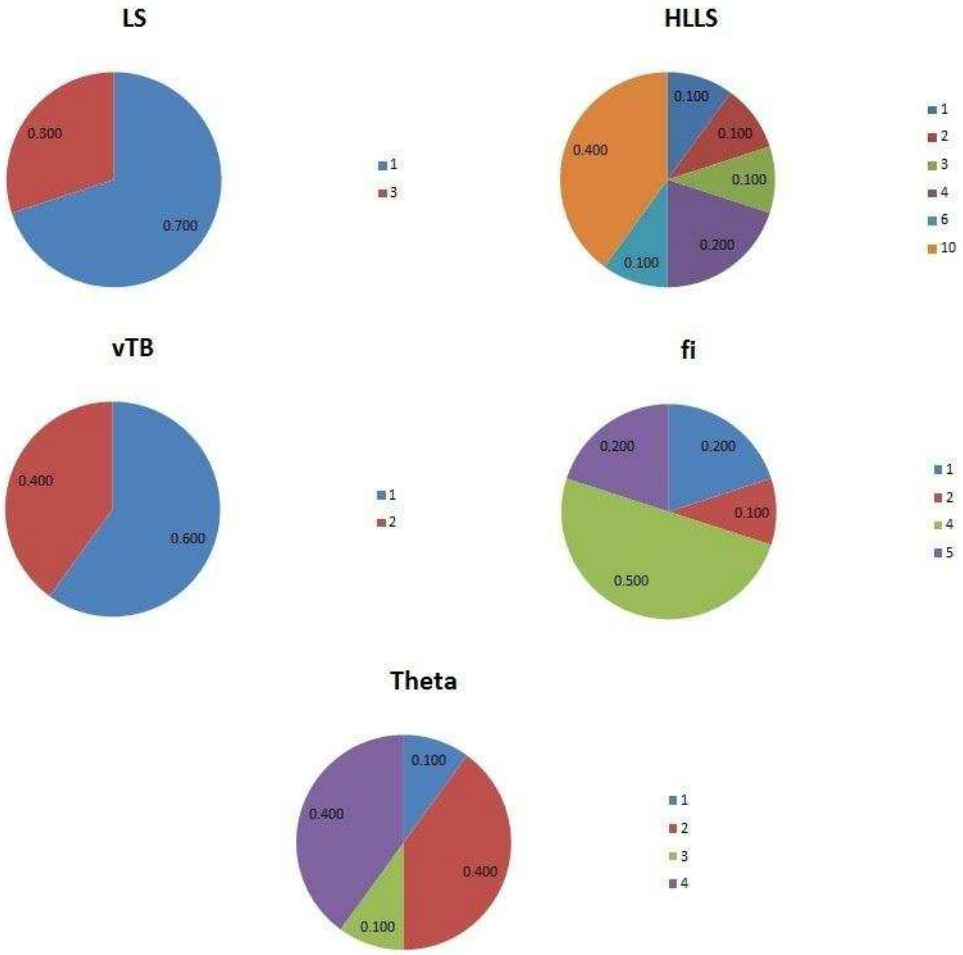


Figure 8.7: The correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ is not optimal for non-horizontal flow ($\beta \neq 0$) in the region of values of $v_{SL} < 0.1$ m/s.

described in Section F.6, Equation F.32, applies to stratified flow in cases when the gas-liquid interface is flat. The better prediction capabilities offered by the former correlation rather than the latter are puzzling, given the simplicity of the first equation.

It is important to point out that the "improvement" brought to the Zhang et al. (2000) model by its 2003 version was not beneficial for horizontal flow at low superficial liquid velocities. **Zhang et al. (2000) used the optimal correlation set, while Zhang et al. (2003) did not**. This emphasizes again the idea that upgrading a model with "more accurate" correlations is not justified, if the correlation is not data-fitted first. That is, we can not apply a mechanistic model to some data without first testing the fitting of its internal empirical correlations. **This is a change to the current industry practice**, which only checks the data fit of a selection of mechanistic models, **without checking the fit of their internal empirical correlations to the data set**. This argument is discussed in more detail in Section 9.5.

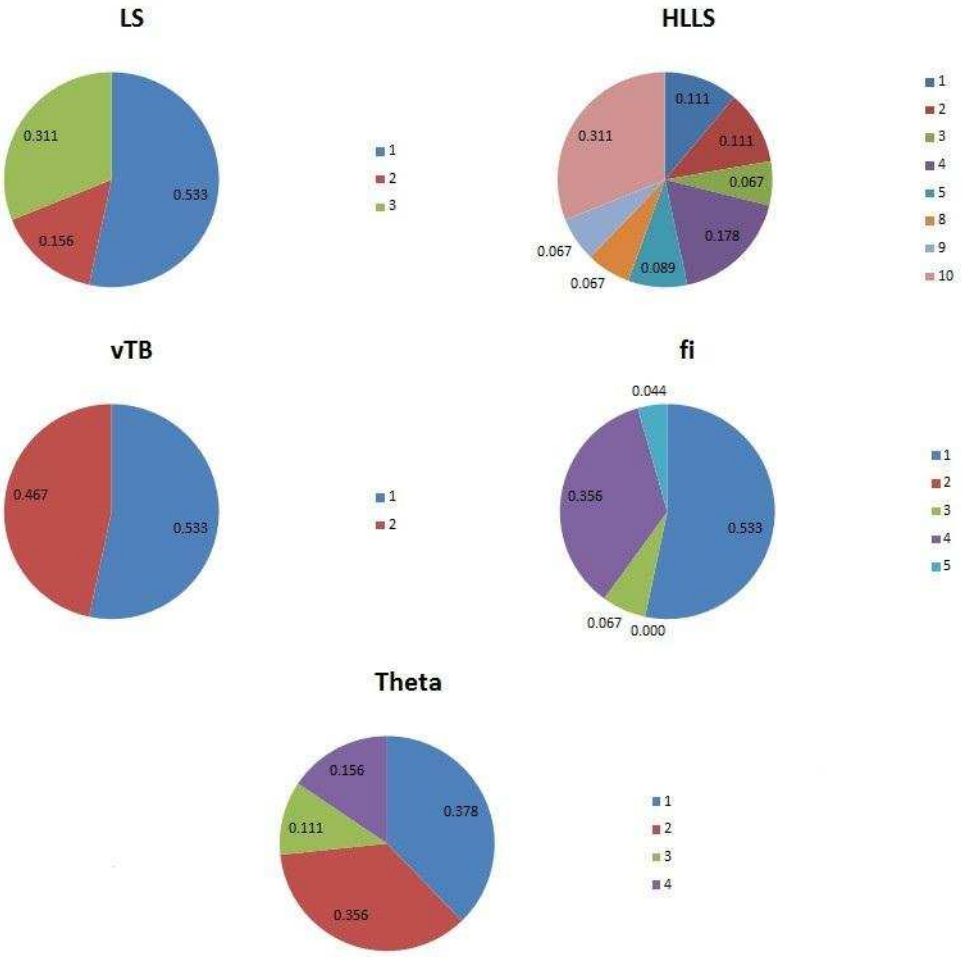


Figure 8.8: The correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ is not optimal for non-horizontal flow ($\beta \neq 0$) in the region of values of $v_{SL} \geq 0.1$ m/s.

8.5 Summary

This chapter discussed the **curve fitting method using alternative empirical correlations**, as part of the **training** stage, the first step towards mechanistic model optimisation using correlation replacement. The method determined the **best empirical correlation set** for each **individual** input data point. This information is stored in a database that will be used for model optimisation in the next chapter.

Data mining techniques applied to the simulation results identified the region ($v_{SL} < 0.1$ m/s, $\beta = 0$), where the empirical correlation set $L_S = 1$, $H_{LLS} = 1$, $v_{TB} = 1$, $f_i = 1$ and $\Theta = 1$ performs best. Consequently, this correlation set is **optimal** for the identified domain.

The comparison between the optimal correlation set for the domain of horizontal flow at low superficial velocities, the Zhang et al. (2000) and the Zhang et al. (2003) models showed:

- The Zhang et al. (2000) model used the optimal correlation set for the selected input domain.
- The correlations for H_{LLS} and f_i in the Zhang et al. (2003) model, despite

being "more accurate", according to their authors (Zhang et al. 2003a, 2003c) are **not optimal**. This makes the newer Zhang et al. (2003) model **not optimal for the input domain of horizontal flow at low superficial liquid velocities**.

The following two conclusions can be drawn:

1. Mechanistic models are **data sensitive** and **one can not apply a mechanistic model to a data set without first testing the data fit of its internal empirical correlations**.

2. The **current industry practice needs to be changed** from checking the data fit of a selection of mechanistic models, to **first checking the data fit of their internal empirical correlation set**.

Chapter 9

Model optimisation

9.1 Introduction

The outcome of the **training** stage was a **history table** of optimal empirical correlation sets for the **training data set**. This chapter will discuss the **prediction** stage of the algorithm in Section 6.4 and will use the information in the history table to generate the optimal prediction.

9.2 Model optimisation using correlation extrapolation and data interpolation

As petroleum exploration moves to new offshore fields with more extreme physical conditions, where fluid properties are outside the validity range of known empirical correlations, hydrocarbon pipeline flow modelling becomes more difficult. As new exploratory wells are drilled, flow data become available before the start of production. The general approach taken in the literature (Kahrs and Marquardt 2007) is to consider the entire flow model invalid outside the validity domain of the empirical correlations and to use neural networks on the available data, in order to adjust the model. A lot of research is focussed on establishing the criteria for determining the model validity domain, based on the empirical correlations used by the model (Kahrs and Marquardt 2007; Fiedler and Schuppert 2008).

This study takes a fundamentally different approach. Instead of completely discarding an empirical correlation outside the data domain it was tested against, this investigation recognises that the empirical correlation may be valid outside it. In light of this view, this study tries to extrapolate the empirical correlation outside its validity domain. From this point of view, this is an **extrapolation** problem.

Chapter 8 identified the best empirical correlation sets for the training data set, on an individual point basis. The challenge is to make a prediction for an input data point when this is not part of the training data set, but is in its proximity. This is an **interpolation** problem.

The selection of the best empirical correlation set, for each individual input data point, in order to minimise the distance between the model's prediction and the measured output, constitutes an **optimisation** problem.

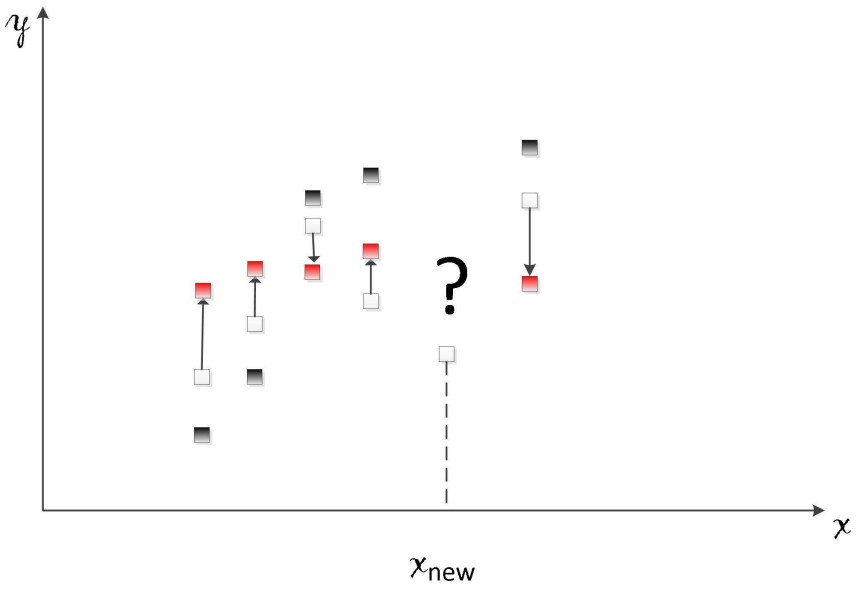


Figure 9.1: The black rectangular markers show the predicted output of the original mechanistic model. The white rectangular markers show the estimations of the curve fitting model, which rely on prior knowledge of the measured, or "true" output values, shown here as red markers. The arrows illustrate the residuals. For a new x_{new} input data point, the lack of knowledge about the "true" output raises the question about predicting the output.

To summarise these concepts, this study deals with **extrapolation and interpolation in order to optimise a mechanistic model**. This is illustrated in Figure 9.1, where the optimised model has to provide a good estimation of the real value, while taking as input a new data point x_{new} that is not part of the training set.

Two different methods to optimise a mechanistic model are proposed and presented.

1. The first optimisation method uses **history matching** to determine the closest input data point in the data set for which an optimal correlation set is known. This involves the minimisation of an objective function. This method is presented in Section 9.3.

2. The second optimisation method uses **data mining** to detect prevailing patterns of optimal correlations. In this case, the optimisation is not performed over the entire input domain, but over a subset for which such a pattern exists. The method is presented in Section 9.4.

Both methods assume that a table of best correlations has been built for a training data set, as described in the **curve fitting** algorithm in Chapter 8. This **history table** has the schema in Table 9.1 and was built in Section 8.3. The records in the history table were shown in Table A.9.

Table 9.1: The schema of the history table, where n = test number, t = test name, d = diameter, β = inclination angle, v_S = superficial velocity, ρ = density, μ = viscosity and σ = interfacial surface tension, L_S^c = optimal correlation number for the calculation of L_S, H_L^c = optimal correlation number for the calculation of H_{LLS}, v_{TB}^c = optimal correlation number for the calculation of v_{TB}, f_i^c = optimal correlation number for the calculation of f_i, Θ^c = optimal correlation number for the calculation of Θ, R = minimum residual and subscripts L and G refer to liquid and gas, respectively.

n	t	d	β	v_{SL}	v_{SG}	ρ_L	ρ_G	μ_L	μ_G	σ	L_S^c	H_L^c	v_{TB}^c	f_i^c	Θ^c	R

9.3 Model optimisation using a history matching algorithm based on sensitivity analysis

This optimisation method uses the **history table** schema shown in Table 9.1 in conjunction with a **history matching** algorithm, discussed later in Section 9.3.2. The new input data point with its flow dynamics (v_{SL}, v_{SG}), fluid properties (ρ_L, ρ_G, μ_L, μ_G, σ), certain pipe diameter and inclination (d, β) is compared to the training input data set, for which the optimal correlation sets have already been identified. If such a match is found, that is, if $x_{new} = x_{history}$, the best correlation set to fit the new input data point is determined from the **history table** as the correlation set associated with the matched entry $x_{history}$. However, this poses a new question. It is very unlikely that the new input data point will match value-for-value each component of the existing data point vectors ($d_{new} = d_{history}, \beta_{new} = \beta_{history}, v_{SL_{new}} = v_{SL_{history}}, \ldots, \sigma_{new} = \sigma_{history}$). If we were to plot two data points in a nine-dimensional hypercube (Figure 9.2), the new data point and one of the points in the history table, the problem could be stated as to find the data point in the history table to minimise the distance in the nine-dimensional hypercube. Generalising the three-dimensional distance between two data points in the input space, the objective function would be:

$$F = \sqrt{\sum_{k=1}^{9} \left(x_{new,k} - x_{history,k}\right)^2} \tag{9.1}$$

However, this approach is simplistic and does not take into account an important aspect. Each element of the input data point vector:

$$x = [d, \beta, v_{SL}, v_{SG}, \rho_L, \rho_G, \mu_L, \mu_G, \sigma] \tag{9.2}$$

has a different effect on the output of the model. Equation 9.1 assumes an equal weight of the input factors in the model's output. This is an unfounded assumption. One approach would be to consider the following objective function:

$$F_h = \sqrt{\sum_{k=1}^{9} {c_{W,k}}^2 \left(x_{new,k} - x_{h,k}\right)^2} \tag{9.3}$$

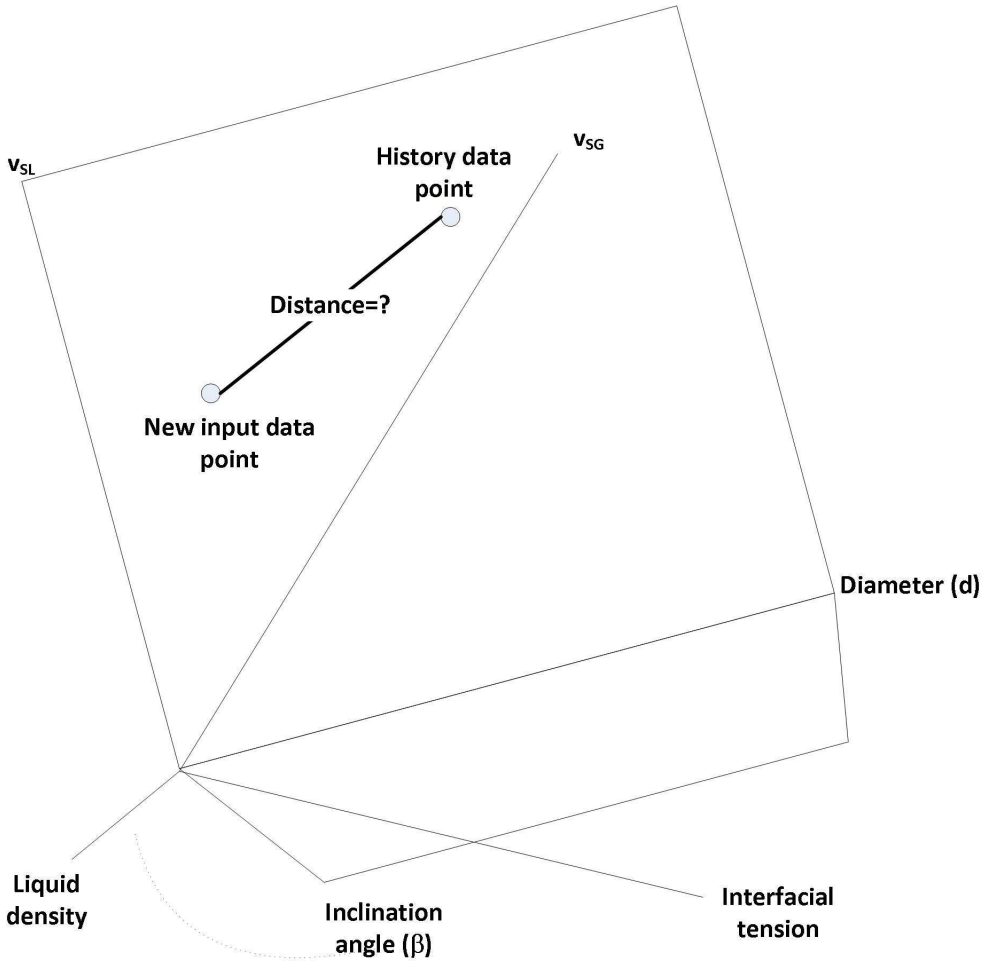


Figure 9.2: Distance between the new input data point and the matched data point in the hypercube.

where each weight factor $c_{W,k}$ is calculated by determining the sensitivity of the input component k. The index h refers to a data point in the history table. Function F_h defined in Equation 9.3 takes into account the different contributions of the input parameters to the model output. Hence, when the function is minimised across all the h points in the history table in the search for the closest history point, the minimisation is performed in such a way that the highest contributing factor is most strongly taken into account.

This concept would work well for input data whose values were fairly close to those in the history table. If, for example, the model were using a history table of laboratory experimental data and the model were to be tested on field data that differed considerably from those obtained in the laboratory, the objective function could minimise at a point considerably far from the input data point, simply because there was no point in the history table close to the input point. For such cases, it is expected that the optimised model will not perform any better than the Zhang et al. (2000) original model and a decision is made to use the latter model for these situations. In other words, the optimisation will only work well if:

- The input data fed into the model have a similar range to the data in the history table. Generally this is the case, because well test data are obtained

from the field before the production start.

- Sufficient data exist in the history table.

- A good objective function is used, capable of discovering the closest point in the history table to the input point.

Assuming that the objective function minimises at the history data point with the index $h^* \in \{1, 2, \ldots, n\}$:

$$x_{h^*} = (x_{h^*,1}, x_{h^*,2}, \ldots, x_{h^*,9}) \tag{9.4}$$

the following test criterion was applied to revert back to the Zhang model when the objective function minimises too far from the input data:

$$\delta_k = c_{W,k}|x_{new,k} - x_{h^*,k}| \qquad k \in [1..9] \tag{9.5}$$

$$x_{new} = \begin{cases} x_{h^*} & \text{if } \delta_k \leq 10^{-3}|x_k^{MAX} - x_k^{MIN}| \qquad \forall k \in [1 \ldots 9] \\ x_{new} & \text{otherwise} \end{cases} \tag{9.6}$$

where:

$$\begin{aligned} x_k^{MAX} &= \max\{x_{j,k}\} \text{ maximum of component } k \text{ across all input data points} \\ x_k^{MIN} &= \min\{x_{j,k}\} \text{ minimum of component } k \text{ across all input data points} \\ j &= 1, 2, \ldots, n \quad \text{and} \quad k = 1, 2, \ldots, M \end{aligned}$$

are the upper and lower limit of the range of variation of component k across all the input data points $\{x_{1,k}, x_{2,k}, \cdots, x_{n,k}\}$.

The meaning of Condition 9.6 is that if there exists a projection of the distance between x_{new} and x_{h^*} on one of the axes of the hypercube, such that the norm of the projection is greater than 10^{-3} of the maximum variation range of that component, the Zhang et al. (2000) original model will be used for prediction. The value 10^{-3} was chosen as data resolution, meaning that a change in the 3^{rd} digit was observed to trigger a correlation set change.

If all the input vector components satisfy the inequality $\delta_k \leq 10^{-3}|x_k^{MAX} - x_k^{MIN}|$, then the Zhang et al. optimised model will be used for prediction.

Other test criteria can be applied, and many other objective functions can be constructed. The purpose of this study is not to focus on objective functions, but to show that even relatively simple objective functions can lead to improved results, better than those given by the original mechanistic model. Nonetheless, this does not mean that the form of the objective function is not important. A good objective function is crucial for the functioning of the history matching algorithm.

Table 9.2: The minimum and maximum values of the input factor for the considered data sample.

Input factor	Minimum	Maximum	Unit [SI]
d	0.025	0.174	m
β	-0.174	0.174	rad
v_{SL}	0.027	3.983	m/s
v_{SG}	0.091	3.419	m/s
ρ_L	790.673	1000.035	kg/m^3
ρ_G	1.156	30.435	kg/m^3
μ_L	0.000810	0.048170	Pa s
μ_G	0.000012	0.000019	Pa s
σ	0.023	0.075	N/m

9.3.1 Calculation of the weight coefficients $c_{w,k}$ using sensitivity analysis of the input factors

A screening method, such as the Elementary Effects method (Morris 1991), described in detail in Section B.1, Appendix B, can be applied to rank the input factors in terms of their influence on the model's output. This ranking is necessary, because it will be used to determine the values of the weight coefficients in the objective function in Equation 9.3. A summary of the sensitivity analysis calculations is shown here. Please refer to Section B.1, Appendix B for full details of the application of the Morris Elementary Effects method.

Sensitivity analysis of the pipe geometry, flow dynamics and fluid properties input parameters

The following input factors were considered for the analysis: diameter (d), inclination angle (β), superficial liquid velocity (v_{SL}), superficial gas velocity (v_{SG}), liquid and gas densities (ρ_L, ρ_G), liquid and gas viscosities (μ_L, μ_G) and surface tension (σ). The minimum and maximum values for each one of the factors in the 95 point data sample discussed in Section 2.5 were determined and shown in Table 9.2.

The distribution of these parameters in the minimum-maximum range is uniform, because, for the purpose of this analysis, any value in the range is assumed with an equal probability. Therefore, the Elementary Effects method does not need to be modified as it would for a normal distribution.

Selection of the cell, Δ and parameter r = number of random trajectories

The method ensures the transformation of the input domain values from the range $[min \dots max]$ to the interval $[0 \dots 1]$. A grid value of $p = 100$ was considered sufficient for the division of the interval $[0 \cdots 1]$ into equal cells. The method's Δ value was chosen according to the recommendations, $\Delta = \frac{p}{2(p-1)}$ (Saltelli et al. 2008). The parameter r = number of random trajectories in the input domain needs to be chosen in such a way that the values of the Morris normalised means do not change between dif-

Table 9.3: The value of the input factor normalised Morris means for different values of r = number of trajectories.

Input factor	$r = 5$ run 1 case 1	$r = 5$ run 2 case 2	$r = 10$ run 1 case 3	$r = 10$ run 2 case 4	$r = 100$ run 1 case 5	$r = 1000$ run 1 case 6	$r = 10000$ run 1 case 7
d	0.1328	0.4774	0.1553	0.0000	0.1776	0.1413	0.1568
β	0.2396	0.3839	0.4217	0.3673	0.2477	0.2646	0.2671
v_{SL}	0.2966	0.0323	0.3412	0.2035	0.2454	0.2939	0.2780
v_{SG}	0.2800	0.0000	0.0082	0.3532	0.1551	0.2063	0.1821
ρ_L	0.0000	0.0923	0.0185	0.0434	0.0184	0.0124	0.0189
ρ_G	0.0006	0.0015	0.0000	0.0010	0.0001	0.0006	0.0018
μ_L	0.0467	0.0084	0.0441	0.0167	0.0716	0.0421	0.0582
μ_G	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
σ	0.0033	0.0039	0.0107	0.0145	0.0837	0.0385	0.0366

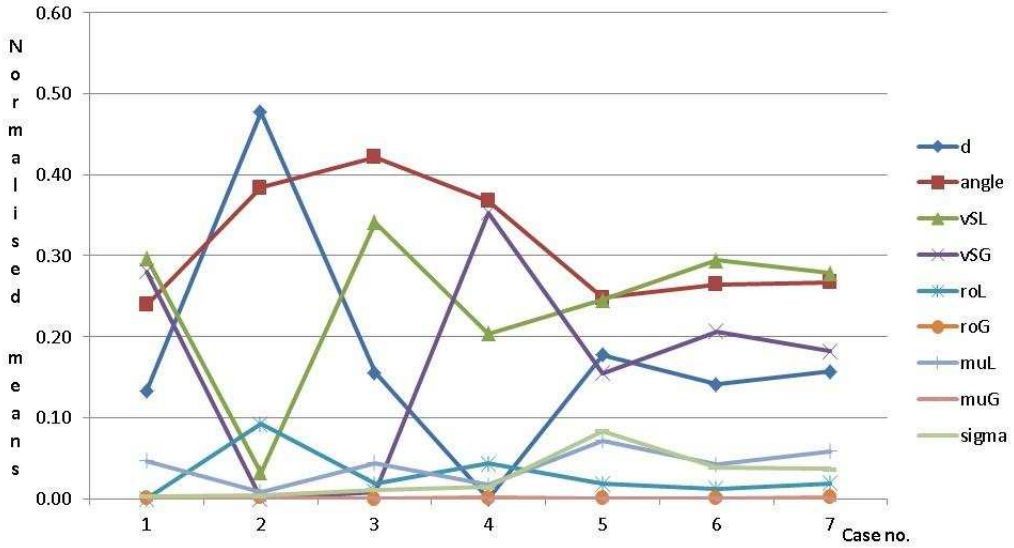


Figure 9.3: Variation of the normalised Morris means for different cases and values of the r parameter. The values have stabilised for case 7 ($r = 10000$).

ferent runs of the sensitivity analysis method. A number of trajectories between 4 and 10 is normally sufficient for most sensitivity analysis cases (Saltelli et al. 2008), because the ranking of the input factors usually does not change anymore for $r \geq 10$. However, because we use the normalised Morris means as weight coefficients, we also need to make sure that the values of the means stabilise too. In this case it was noticed (Figure 9.3) that the normalised means did not stabilise for $r = 10$ between different runs of the sensitivity method. The value of r was increased gradually from $r = 10$ to 100, then 1000 and then 10000 while the values of the Morris normalised means were examined. The values are shown in Table 9.3. The purpose of these calculations was to choose the value of r for which these values have stabilised.

As the values of the means converged to a stable value at case 7 ($r = 10000$), a value of $r = 10000$ was chosen in the calculation. The normalised

Table 9.4: The Morris means μ_i and the normalised Morris means ($\frac{\mu_i}{\Sigma\mu_i}$) of the input factors, obtained by applying the Elementary Effects method for $p = 100$ and $r = 10000$.

Input	Morris mean μ_i	Normalised Morris mean $\frac{\mu_i}{\Sigma\mu_i}$
d	μ_d =7368.643	0.156810
β	μ_β =12552.692	0.267131
v_{SL}	$\mu_{v_{SL}}$ =13067.850	0.278094
v_{SG}	$\mu_{v_{SG}}$ =8561.372	0.182192
ρ_L	μ_{ρ_L} =892.657	0.018996
ρ_G	μ_{ρ_G} =87.783	0.001868
μ_L	μ_{μ_L} =2737.998	0.058267
μ_G	μ_{ρ_G} =0.039	0.000001
σ	μ_σ =1721.808	0.036641

Table 9.5: Ranking of the input factors from the most influential (high value) to the least influential (low value).

Input	Morris mean μ_i	Normalised Morris mean $\frac{\mu_i}{\Sigma\mu_i}$
v_{SL}	$\mu_{v_{SL}}$ =13067.850	0.278094
β	μ_β =12552.692	0.267131
v_{SG}	$\mu_{v_{SG}}$ =8561.372	0.182192
d	μ_d =7368.643	0.156100
μ_L	μ_{μ_L} =2737.998	0.058267
σ	μ_σ =1721.808	0.036641
ρ_L	μ_{ρ_L} =892.657	0.018996
ρ_G	μ_{ρ_G} =87.783	0.001868
μ_G	μ_{ρ_G} =0.039	0.000001

means obtained for this value of r are shown in Table 9.4.

It is important to mention here that the figures in Table 9.4 are obtained by averaging the derivatives of the model's output as a function of each input factor, for a random selection of input domain values. Sorting these values in descending order leads to a ranking of the input factors in the order from the most influential to the least influential factor, as shown in Table 9.5 and Figure 9.4. This ranking applies to the entire input domain.

Coefficient calculation

Sorting the means calculated by the Elementary Effects method in descending order determines the influence of the input factors, from the highest to the lowest, as shown in Table 9.5. By assigning the weight factors the values of the normalised means, the coefficients have the role of minimising preferentially in the direction of the more influential input components.

The ranking in Table 9.5, where the inclination angle has a significant effect on the output, confirms previous research, which has shown that even the slightest change in pipe inclination, such as 0.1°, can have a significant effect on the output and can even lead to a flow pattern transition (Shoham

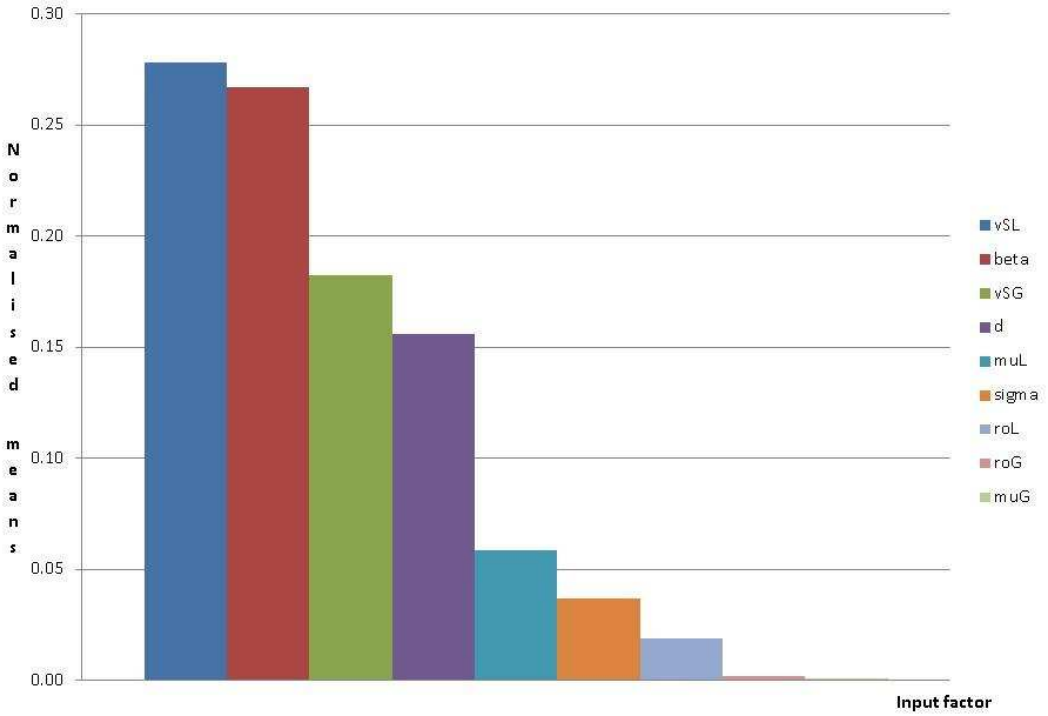


Figure 9.4: Ranking of the input factors from the highest value as the most influential (high normalised mean) to the least influential (lowest normalised mean).

2006). Figures 9.5 and 9.6 illustrate the significant change in the flow pattern map for a small inclination of 0.25°. The maps were obtained by applying the Barnea flow pattern transition criteria (Barnea 1987) to a wide range of superficial gas and liquid velocities, v_{SL} and $v_{SG} \in [0.05 \ldots 50]$ m/s. Nevertheless, despite the fact that results like this can be generalised to all input data ranges, others are data-dependent and only apply for the input range applied to the model. For example the fact that the liquid velocity v_{SL} has a slightly higher influence than the inclination angle β cannot be generalised for all data sets.

9.3.2 History matching algorithm

The curve fitting algorithm in Chapter 8 built a **history table**, which was an association between a set of input data points and the best empirical correlations to be used for each point. The **history matching** algorithm developed here, identifies the closest **history table** point, by minimising the distance objective function. Sensitivity analysis is used to calculate the weight coefficients in the objective function. An illustration of how these concepts come together is shown in Figure 9.7.

The algorithm to optimise a mechanistic model using history matching based on sensitivity analysis is:

Step 1: Apply a sensitivity analysis method to the input factors.

For the selected data set (Section 2.5) and the Zhang et al. (2000) mechanistic model, Section 9.3.1 applied the Morris Elementary Effects method

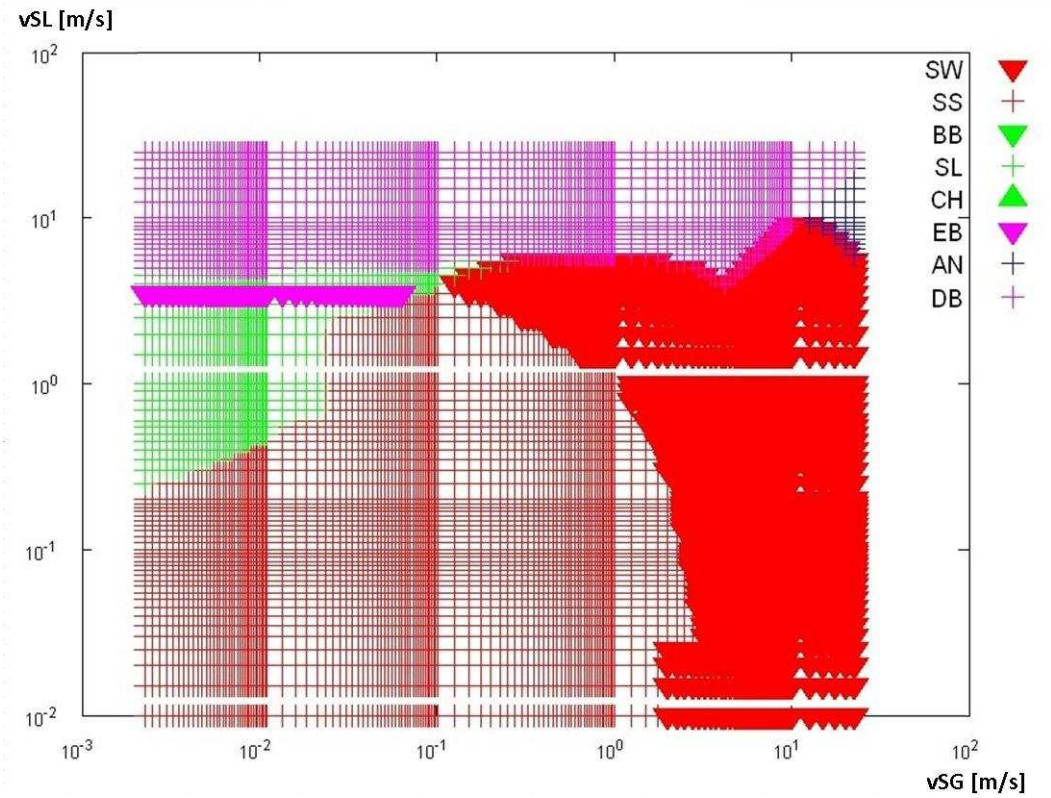


Figure 9.5: Flow regime map for a 21" pipe diameter and 0° inclination. SW = stratified wavy, SS = stratified smooth, BB = bubble, SL = slug, CH = churn, EB = elongated bubble, AN = annular, DB = dispersed bubble.

to the input factors d, β, v_{SL}, v_{SG}, ρ_L, ρ_G, μ_L, μ_G, σ and calculated their sensitivities.

Step 2: Normalise the results and rank the input factors. Assign the normalised values to the weight coefficients $c_{w,k}$.

For the example of Zhang et al. (2000) model, the weight coefficients are: $c_{w,v_{SL}} = 0.278094$, $c_{w,\beta} = 0.267131$, $c_{w,v_{SG}} = 0.182192$, $c_{w,d} = 0.156810$, $c_{w,\mu_L} = 0.058267$, $c_{w,\sigma} = 0.036641$, $c_{w,\rho_L} = 0.018996$, $c_{w,\rho_G} = 0.001868$, $c_{w,\mu_G} = 0.000001$.

Step 3: In order to calculate the prediction, for each input data point build the objective function defined in Equation 9.3. Calculate and store the value of the objective function for all the points h in the history table $(h = 1, 2, \ldots, H)$.

The objective function is the distance in the hypercube between the new input data point and a point in the history table. Both the new input data point and the history point are vectors with the following components:

$$\begin{aligned} x_{new} &= (x_{new,d}, x_{new,\beta}, x_{new,v_{SL}}, \ldots, x_{new,\mu_L}, x_{new,\mu_G}, x_{new,\sigma}) \\ x_h &= (x_{h,d}, x_{h,\beta}, x_{h,v_{SL}}, \ldots, x_{h,\mu_L}, x_{h,\mu_G}, x_{h,\sigma}) \\ h &= 1, 2, \ldots, H \end{aligned} \tag{9.7}$$

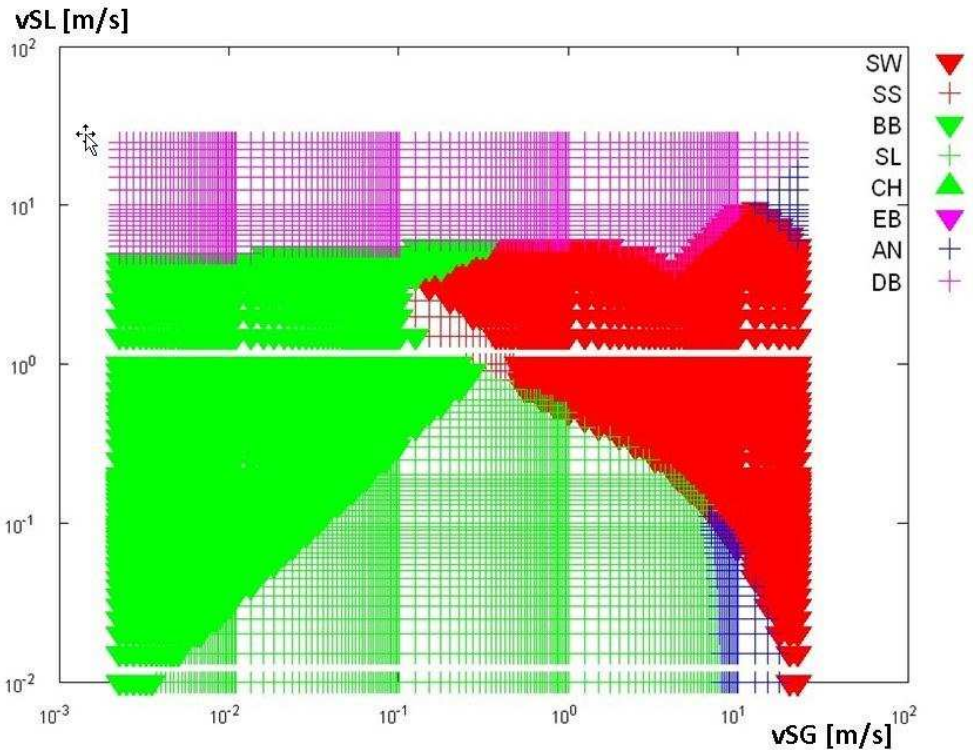


Figure 9.6: Flow regime map for a 21" pipe diameter and 0.25° inclination. A significant change in flow pattern distribution is observed from the previous case for a very small inclination increment.

where H is the size of the history table.

Taking into account the weight coefficients calculated in Step 2 and the x vectors in Equations 9.7, the distance between the new input data point and a history data point becomes:

$$\begin{aligned} F_h\left(x_{new}, x_h\right) = \{ & {c_{W,d}}^2 (x_{new,d} - x_{h,d})^2 + {c_{W,\beta}}^2 (x_{new,\beta} - x_{h,\beta})^2 \\ & + {c_{W,v_{SL}}}^2 (x_{new,v_{SL}} - x_{h,v_{SL}})^2 + {c_{W,v_{SG}}}^2 (x_{new,v_{SG}} - x_{h,v_{SG}})^2 \\ & + {c_{W,\rho_L}}^2 (x_{new,\rho_L} - x_{h,\rho_L})^2 + {c_{W,\rho_G}}^2 (x_{new,\rho_G} - x_{h,\rho_G})^2 \\ & + {c_{W,\mu_L}}^2 (x_{new,\mu_L} - x_{h,\mu_L})^2 + {c_{W,\mu_G}}^2 (x_{new,\mu_G} - x_{h,\mu_G})^2 \\ & + {c_{W,\sigma}}^2 (x_{new,\sigma} - x_{h,\sigma})^2 \}^{0.5} \end{aligned} \quad (9.8)$$

Calculate and store the value of the objective function in Equation 9.8 for all the points h in the history table ($h = 1, 2, \ldots, H$).

Step 4: Iterate through all the values of the objective function determined in Step 3 and identify the minimum value and the index h^* for which this is reached.

The history data point x_{h^*} is the closest entry in the history table to the new input data point x_{new}, and has an optimal correlation set associated with it. The position of this optimal correlation set in the history table entry is shown in the schema in Table 9.1.

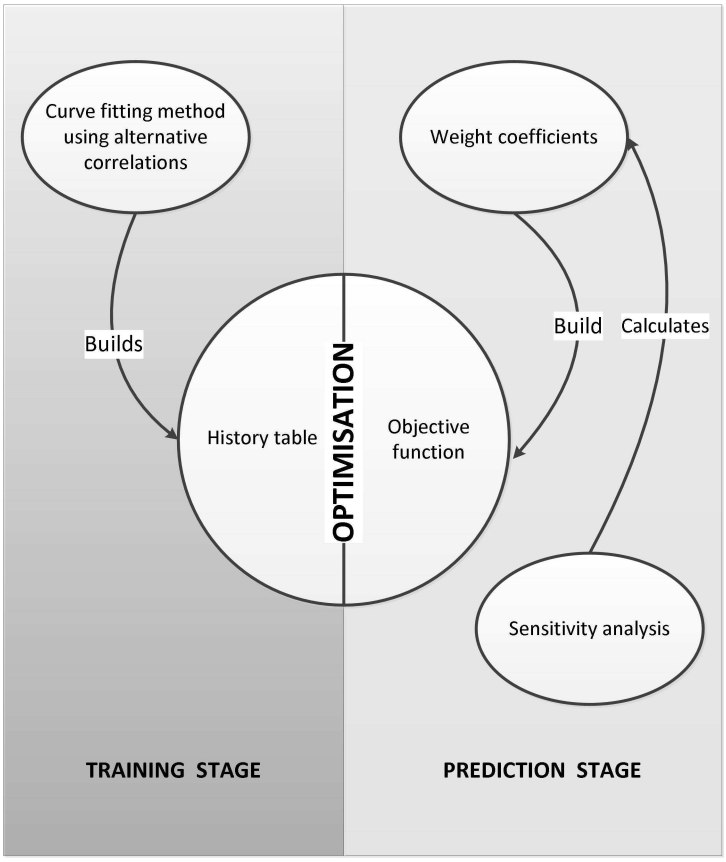


Figure 9.7: This diagram describes the two main components of this optimisation method: a history table and an objective function. The former relies on the curve fitting method described in Chapter 8 to build a history table and the latter relies on a sensitivity analysis method to calculate the objective function's weight coefficients. The first is part of the "training" stage, whereas the second is part of the "prediction" stage.

Step 5: Verify that all the projections of the line between x_{new} and x_{h^*} on all the axes of the hypercube satisfy the condition:

$$\delta_k \leq 10^{-3}|x_k^{MAX} - x_k^{MIN}| \qquad k = 1, 2, \ldots, 9 \tag{9.9}$$

This criterion was described in Equation 9.6 and means that the input point is in the proximity of the history point x_{h^*}.

Otherwise, the objective function minimises at a history point that is too far from the input point and the optimal correlation set known for this history point cannot be used. Hence, in these situations, the optimal correlation set is the one used by the original mechanistic model and Step 6 is skipped.

For the Zhang et al. (2000) model and the selected data bank, if any of the inequalities in Step 5 is not met, the original empirical correlation set used in the Zhang et al. (2000) model will be used: $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$, $\Theta^c = 1$ (the decoding correlation number-correlation name is shown in Tables A.2-A.6).

Step 6: Read the entry for x_{h^*} in the history table and extract the optimal correlation set and the minimum residual $R_{h^*,min}$ associated with it.

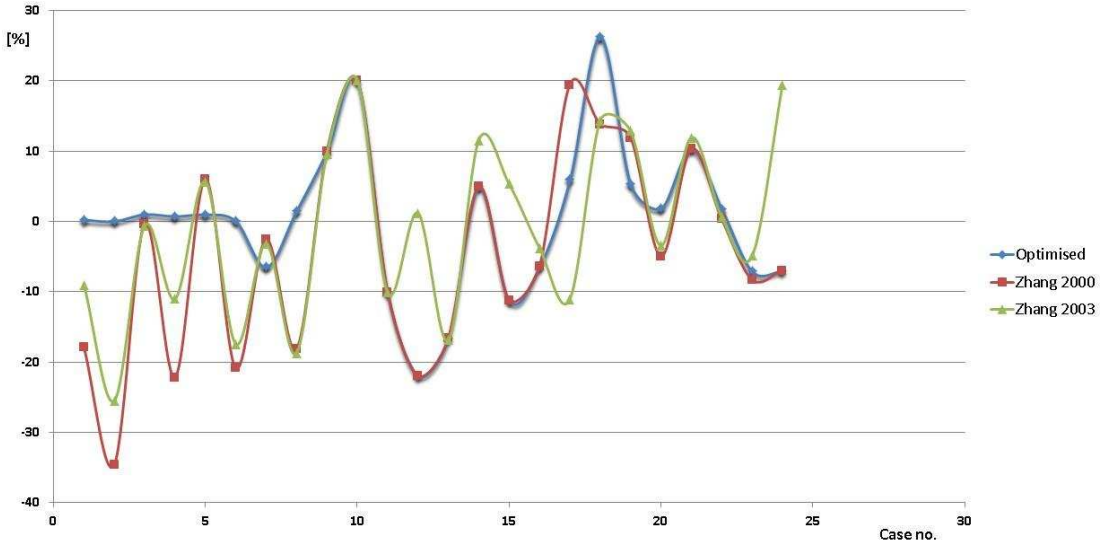


Figure 9.8: Comparison of pressure gradient prediction errors: the Optimised model vs. the Zhang et al. (2000) model vs. Zhang et al. (2003) model. The Optimised model outperforms the other models.

Step 7: Use the optimal correlation set and the input point x_{new} to calculate the model output. Add the residual to the output to construct the prediction, according to Equation 6.15 in Chapter 6.

The pseudocode of this algorithm is shown in Algorithm 9.1.

9.3.3 Case study - Zhang et al. slug flow model optimisation using the history matching algorithm based on sensitivity analysis

The best correlation sets for the 95 slug flow data points were calculated as part of the simulation in Section 8.3. These data points and the associated optimal correlations constituted the history table. Based on the considerations in Section 2.5.2, in order to test the performance of the optimised model, 24 data points were taken randomly out of the history table and considered as new input data. Thus, the optimised model was run against a history table of 71 points.

9.3.4 Simulation results

Algorithm 9.1 was implemented as a GNU Octave script in the *Optimisation* folder on the attached CD. The combined results of the predictions of the Optimised, Zhang et al. (2000) and (2003) models are shown in Figures 9.8 and 9.9, which are based on the results in Table A.11, in Appendix A. The Optimised model performs best with an overall score of $F_{PR} = 0$, followed by the Zhang et al. (2003) model with a score of $F_{PR} = 3.12$ and Zhang et al. (2000) with a score of $F_{PR} = 6$.

The E statistical measures for the comparison are shown in Table 9.6. The terms $Term_1 - Term_6$ in Table 9.7 are the six components of the sum in Equation 5.60.

Algorithm 9.1 The pseudocode of the optimal prediction algorithm.

procedure OPTIMAL(x_i)
  REQUIRES: Rank of the input factors is known.
  INPUT: x_i,
  $Sensitivities = \{Sens_1, Sens_2, \ldots, Sens_M\}$
  Maximum values for each input factor vector component:
  $MAX = \{x_1^{MAX}, x_2^{MAX}, \ldots, x_M^{MAX}\}$
  Minimum values for each input factor vector component:
  $MIN = \{x_1^{MIN}, x_2^{MIN}, \ldots, x_M^{MIN}\}$
  OUTPUT: $\hat{y}$

  Calculate thresholds on each axis:
  $C_{W,1} = Sens_1, C_{W,2} = Sens_2, \ldots, C_{W,M} = Sens_M$
  $Threshold_k = 10^{-3} \left| x_k^{MAX} - x_k^{MIN} \right|$
  for all $x_h \in H$ **do**
    Calculate objective function F_h
    if $F_h < F_{min}$ **then**
      $F_{min} = F_h$
      $x_{h^*} = x_h$
    end if
  end for
  for all $x_{i,k}, k \in \{1, 2, \ldots, M\}$ **do**
    if $|x_{i,k} - x_{h^*,k}| < Threshold_k$ **then**
      $CorrelationFlag = Optimal$
    else
      $CorrelationFlag = OriginalModel$
    end if
  end for
  if $CorrelationFlag = Optimal$ **then**
    Search history table for record containing x_{h^*}
    Read the empirical correlation set $O(x_{h^*})$
    Set $CorrelationSet = O(x_{h^*})$
    Read minimum residual associated with x_{h^*}: $R_{min}(x_{h^*})$
    Set residual: $Residual = R_{min}(x_{h^*})$
  else
    $CorrelationSet = original$
    Set residual: $Residual = 0$
  end if
  Calculate the model estimation $\tilde{y}$:
  $\tilde{y}$ =PredictOutput($x_i, CorrelationSet$)
  Calculate the prediction $\hat{y}$:
  $\hat{y} = \tilde{y} + Residual$
end procedure

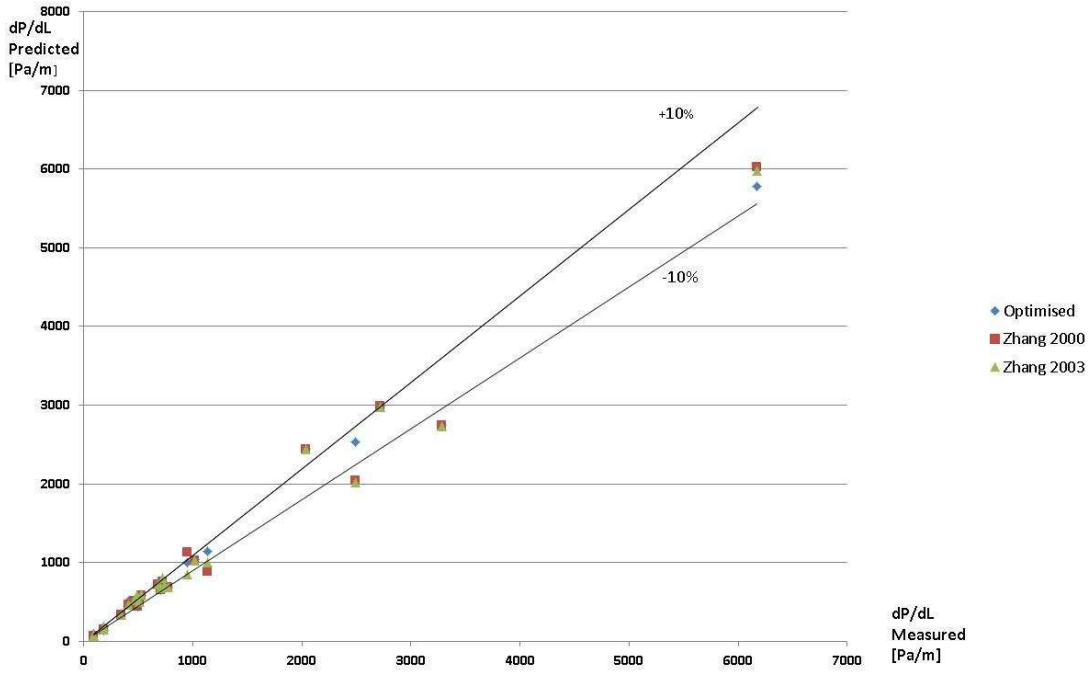


Figure 9.9: Comparison of pressure gradient prediction errors: the Optimised, the Zhang et al. (2000) and (2003) models. Most of the model predictions fall within the range of measured value $\pm 10\%$.

Table 9.6: E parameter comparison between the Optimised model and the Zhang (2000) and (2003) models.

Model	E_1	E_2	E_3	E_4	E_5	E_6
Optimised	0.20	7.42	0.04	-833.07	9544.63	4127.28
Zhang (2003)	-0.96	10.33	0.20	-2375.12	11552.08	11767.07
Zhang (2000)	-4.40	12.48	0.94	-2973.03	12226.83	14729.33

Table 9.7: The F_{PR} score shows that the Optimised model outperforms the Zhang (2000) and (2003) models. The lower value of F_{PR} corresponds to the more accurate model.

Model	$Term_1$	$Term_2$	$Term_3$	$Term_4$	$Term_5$	$Term_6$	F_{PR}
Optimised	0	0	0	0	0	0	0
Zhang (2003)	0.18	0.57	0.18	0.72	0.74	0.72	3.12
Zhang (2000)	1	1	1	1	1	1	6

Despite the history data containing the best correlation set for every training data point, the success of the optimisation method described here lies in the effectiveness of the objective function to rediscover the optimal correlation set. The objective function used in this study was fairly simple, but other more effective functions may give even better results. Figure 9.8 shows that for some of the data points the Optimised model performed worse. This is because the objective function minimised too far from the input point and the threshold criteria did not work for that particular point. It may be argued that the "too far" criteria may need to be changed, but there is also another more important reason for

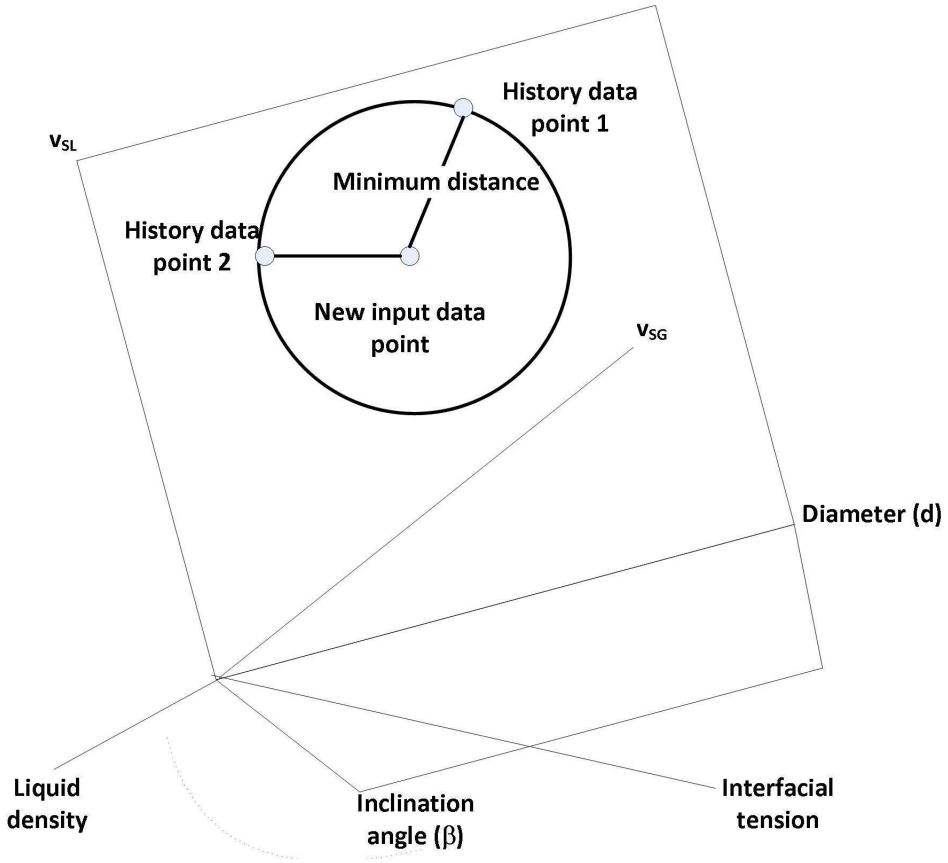


Figure 9.10: History points that lead to the minimisation of the objective function are situated on a spheroid of radius F_{min}.

which this could happen.

It is important to know whether the minimisation of the objective function F_h, defined in Equation 9.3 has a unique solution x_{h^*} in the history data set.

For a minimum value F_{min}, a given input point of components $x_{new,k}$ and a known set of weight coefficients $c_{W,k}$, $k = 1, 2, \ldots, 9$, the locus of the points x in the 9-dimensional hyperplane satisfying Equation 9.10:

$$F_{min} = \sqrt{\sum_{k=1}^{9} {c_{W,k}}^2 (x_{new,k} - x)^2} \tag{9.10}$$

is a hyperspheroid of radius F_{min} (Figure 9.10). This means that all the history points situated on the surface of the hyperspheroid, lead to a minimum value of the objective function. This may create an ambiguous situation, because each history point on the surface of the hyperspheroid is associated with a possibly different optimal correlation set. The existence of a multitude of history points leading to a minimum objective function raises the question of which one of them to choose, and consequently which optimal correlation set to associate with the new input data point x_{new}.

The solution to choosing the right history table point relies on the ranking of the input factors. Let us assume two history data points $x_{1,history}$ and $x_{2,history}$, situated on the same hyperspheroid of radius F_{min}. As the sensitivity analysis determined that the order of influence of the input factors on the output is

$(v_{SL}, \beta, v_{SG}, d, \mu_L, \sigma, \rho_L, \rho_G, \mu_G)$, the solution of choosing one history data point over the other involves the calculation of $|\delta_k| = |c_{W,k}\,(x_{k,new} - x_{k,h})|$ values, for both points $h = 1, 2$, where $k = v_{SL}, \beta, v_{SG}, d, \mu_L, \sigma, \rho_L, \rho_G, \mu_G$.

If $\left|\delta_{1_{v_{SL}}}\right| < \left|\delta_{2_{v_{SL}}}\right|$, history point 1 is chosen, and the associated optimal correlation set is used for the prediction of the input x_{new}.

If $\left|\delta_{1_{v_{SL}}}\right| = \left|\delta_{2_{v_{SL}}}\right|$, then if $\left|\delta_{1_\beta}\right| < \left|\delta_{2_\beta}\right|$, point 1 is chosen. Otherwise, point 2 is chosen.

The idea is to minimise the components in the order dictated by the sensitivity analysis. This procedure allows the selection of a single history data point and therefore the selection of the optimal correlation set for the new input data point.

On the other hand, the Elementary Effects method is a global sensitivity method and determines the input factor ranking based on the averages of the derivatives of the output as a function of input factors. This leads to one ranking over the entire input domain. This "blanket" ranking used in the objective function does not lead to a good objective function to discover the closest point in the history table, especially when the input factor ranking may differ locally, from point to point. A different approach for the sensitivity analysis needs to be taken for a more effective model optimisation.

If a local sensitivity method is applied, the influence of each input factor on the model output varies across the range of values of the factor. Assuming the same F_h objective function, this change means that the weight coefficients become functions of the history data points. The objective function changes accordingly:

$$F_h = \sqrt{\sum_{k=1}^{9} \left[c\,(x_{k,h})_{W,k}\right]^2 (x_{k_{new}} - x_{k,h})^2} \tag{9.11}$$

There are problems with defining the objective function in this way. Figure 9.11 shows an input data point and two history data points situated on the hyperspheroid of radius F_{min}. Let us assume the following:

- In the area around history point 1 we assume v_{SL} as the most influential factor.
- In the area around history point 2 we assume d as the most influential factor.
- For an input data point such that $d_{input} = d_{history2}$ and $v_{SLinput} = v_{SLhistory1}$, we need to associate one of the history data points with the new input data point, in order to determine the optimal correlation set to be used with the input.

The selection of one history point for this case is impossible. Unlike the previous situation when we could use δ_k and the ranking of the global sensitivity values to determine a single history point, this is not possible anymore because the local sensitivities differ locally, from point to point. The solution to this problem is beyond the scope of this study, but would lead to an algorithm with better capabilities to find the optimal correlation set for each input point. Nevertheless, even when the optimisation procedure uses global sensitivity analysis the method ensured better model predictions than the original mechanistic model.

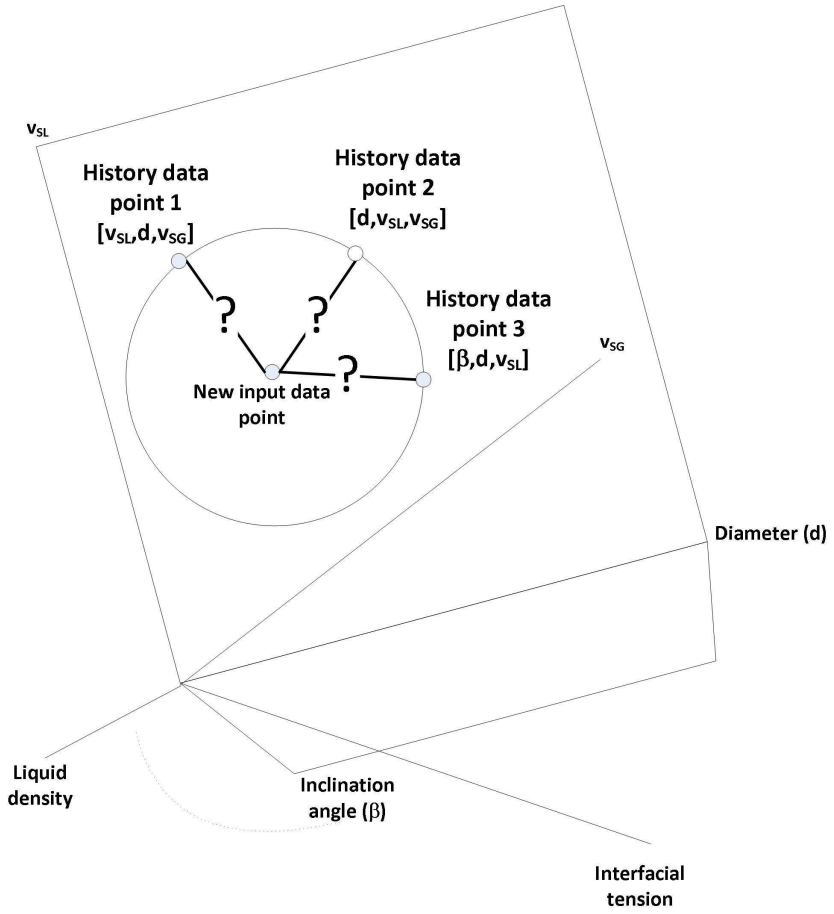


Figure 9.11: The history data points have different input factor rankings. For example, for the history data point 1 v_{SL} has the highest influence on the model output, followed by d and v_{SG}.

9.4 Model optimisation using data mining

Data mining for regions of prevailing optimal correlation sets was discussed in Section 8.4. If such a region is identified over a range of input values, then the model using that prevailing correlation set is optimised over that input range.

Based on the input parameter ranking results in Section 9.3.1, which showed that parameters v_{SL} and β had the greatest influence on the model output, the input data were sorted in ascending order of v_{SL}. The sensitivity analysis performed in Section 9.3.1 explains now why the data were sorted by values of v_{SL} first and then by values of β, which at the time of the discussion in Section 8.4 had not been explained.

Section 8.4 showed that by sorting the data pertaining to horizontal flow ($\beta = 0$) by values of v_{SL}, a fairly consistent pattern of correlations was observed for the lower range of v_{SL} ($v_{SL} < 0.1$ m/s). The combination $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ prevailed as the optimal correlation set in the range of values $v_{SL} < 0.1$ m/s and $v_{SG} < 5.9$ m/s, as shown in Table A.10 and Figure 8.5. A suggestive depiction of the sorting and partitioning procedure was shown in Figure 8.4.

We can conclude that despite the Zhang et al. (2003) model obtaining a better

overall score ($F_{PR} = 3.12$) than the Zhang et al. (2000) model ($F_{PR} = 6$) for the 24 data points, there are data sets where the correlation set $L_S^c = 1$, $H_{LLS}^c = 10$, $v_{TB}^c = 1$, $f_i^c = 2$ for stratified flow or $f_i^c = 3$ for annular flow, $\Theta = 1$, used by the Zhang et al. (2003) model is not optimal. These data correspond to low superficial liquid velocities, where the optimal correlation set is $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$, $\Theta^c = 1$. This way, using data mining techniques, the original Zhang et al. (2003) model can be optimised by adjusting the empirical correlation set over the input range of low superficial liquid velocities, $v_{SL} < 0.1$ m/s.

9.4.1 Optimisation method comparison

Although both the **data mining** and the **history matching** methods rely on the existence of a **history table**, the fundamental difference between them is that for a given input point, the first method **does not need an objective function** to identify the optimal correlation set. This is because the prevailing optimal correlation set, if it exists, is valid for an entire range of input values. If the input point is part of this range, then the optimal correlation set to be associated with the input point, is the prevailing correlation set.

It is important to mention that unlike the optimisation method using **history matching** and sensitivity analysis, the **data mining** method may lead to **regions of optimality**. For example, the optimal correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$ and $\Theta^c = 1$ is not optimal over the entire input range, but only for horizontal flow ($\beta = 0$) at low superficial liquid velocities ($v_{SL} < 0.1$ m/s).

9.5 Recommendations on the use of empirical correlations in mechanistic models

Numerous papers and books (Shoham 2006; Zhang et al. 2003c; Zhang and Sarica 2006; Fan et al. 2007; Gomez et al. 2000; Ansari et al. 1994; Petalas and Aziz 2000; Zhang and Sarica 2006) describe mechanistic models, along with the empirical correlations used as closure relationships. The models are then tested against data sets and the performance is analysed according to a set of statistical measures. In light of this discussion, this study raises the following criticisms to this widely used approach:

1. The selection of the empirical correlations in these models is **independent** of the **particular input data**. The empirical correlations are first chosen and then the model is tested on the data, but the feedback on the model's performance is **not used for the selection** of the correlation. This ignores the **data dependency** aspect of mechanistic models. It may be argued that the empirical correlations are chosen according to their validity domain, so they are not data independent. However, this is a **weak dependency**, because it does not take into account the **particularities** of the data set. Sections 9.4 and 8.4 showed that the correlation for f_i used in the Zhang et al. (2000) model performed better than the correlations for f_i in the Zhang et al. (2003) model in the low superficial velocity domain, despite both correlations being within their validity region. As discussed here, the

right approach is to use the **optimal correlation** for predictions. As the optimality criterion is data-dependent, optimisation using alternative correlations will always outperform the classic approach, where the empirical correlations are **not tuned to the data**, in spite of them being within their validity domain.

2. The practice of data fitting mechanistic models without replacing the embedded empirical correlations is an **incomplete solution** to the problem. The performance of the mechanistic models varies on different input data sets, because of the empirical component. Trying to data fit the entire mechanistic model, when the root cause is the empirical component is not the right approach.

9.6 Summary

This chapter approached the optimisation problem both as an interpolation and extrapolation problem. The **new paradigm** is:

mechanistic model optimisation = correlation extrapolation + data interpolation

Given a set of input data points for which the optimal empirical correlation set is known, the prediction of the model for a new input point in their vicinity needs to be calculated. This is an **interpolation** problem. The method suggested in this study recognises that there is a likelihood that an empirical correlation may be valid outside its validity range. This use of empirical correlations, constitutes an **extrapolation** problem. The data interpolation combined with the empirical correlation extrapolation are used for mechanistic **model optimisation**.

Two optimisation methods were discussed:

1. **Optimisation using a history matching algorithm based on sensitivity analysis**. This method relies on two elements:
    - A **history table**, which is an association between an input data point and both an optimal correlation set and a residual. The table was built as part of the **training** stage, as discussed in Chapter 6.
    - An **objective function** whose minimisation determines the closest point in the history table to the input for which the output is to be predicted. The objective function was built as a distance function in a hyperspace, and the weight coefficient calculated based on sensitivity analysis. The minimisation, which leads to finding the optimal correlation set for the input point, is part of the **prediction** phase, as discussed in Chapter 6. A thorough discussion of the selection of the history data point which leads to both the minimisation of the objective function and the selection of the optimal correlation set for the input data point was given in Section 9.3.4.

    The numerical predictions performed with the Optimised, the Zhang et al. (2000) and (2003) models using a set of 24 data points showed the following

ranking, in the order of decreasing performance:

(a) Optimised model.

(b) Zhang et al. (2003) model.

(c) Zhang et al. (2000) model.

2. **Optimisation using data mining**. This method analysed the **history table** by using sorting and grouping techniques. For the data set analysed and the Zhang et al. (2000) model, it was noticed that the correlation set $L_S^c = 1$, $H_{LLS}^c = 1$, $v_{TB}^c = 1$, $f_i^c = 1$, $\Theta^c = 1$ prevailed and was therefore **optimal**. It was concluded that in order to be optimal in the region of low superficial velocities:

   - The empirical correlation set in the Zhang et al. (2000) model should not be replaced.
   - The empirical correlation set in the Zhang et al. (2003) model should be replaced with the optimal correlation set.

Unlike the **history matching** method, the optimisation using **data mining** may provide only **regional optimisation**.

Chapter 10

Model simplification

10.1 Introduction

Chapter 9 showed the use of sensitivity analysis in model optimisation. Sensitivity analysis was applied to model input factors to rank their effect on the output of the mechanistic model.

The approach taken in this chapter is to use sensitivity analysis to determine the effect of the empirical parameters on the model's prediction. The difference between the two approaches is illustrated in Figure 10.1. This undertaking is important, because it establishes whether an empirical parameter requires an accurate empirical correlation for its estimation. Sensitivity analysis applied to empirical parameters has the following benefits:

1. It determines an easy method to improve the accuracy of the mechanistic

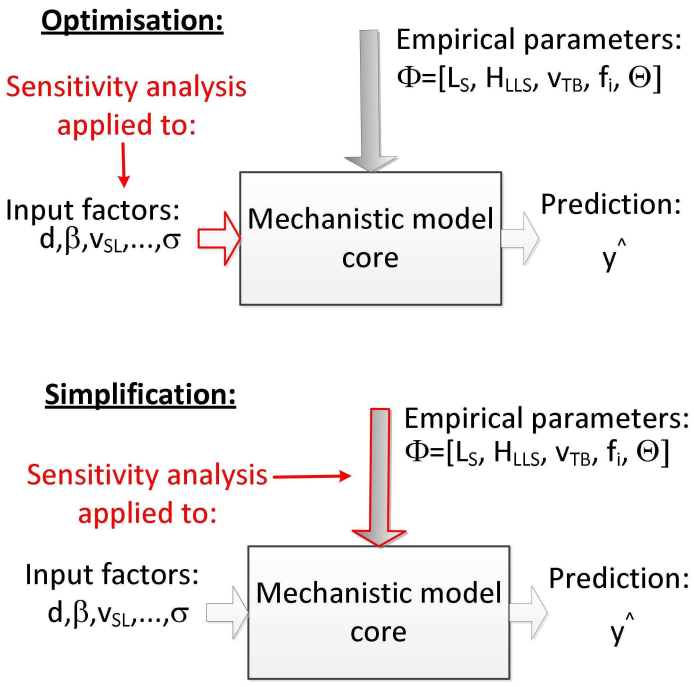


Figure 10.1: Mechanistic model simplification using sensitivity analysis applied to the empirical correlation vector Φ. This is different to the optimisation case in which sensitivity analysis was applied to the model input factors.

model, by identifying the parameters with the greatest influence on the model output. This allows the production engineer to easily establish a method to improve model prediction by replacing the empirical correlation that estimates the parameter with the greatest influence on the output, with a more accurate one.

2. It sets a direction for experimental research to find better empirical correlations for the empirical parameters that have greater influence on the model output.

3. It determines the model parameters that have a minimal influence on the model output and hence, only need to be assigned educated guess values or be calculated with less accurate empirical correlations. Therefore, the model calculation algorithm could be simplified by maintaining constant these empirical parameters. This is beneficial for cases when an empirical correlation is difficult to find, or be developed, or has a highly computational component.

The investigation was triggered by the fact that numerous papers discuss friction factor correlations and some focus on the interfacial friction factor (Liu et al. 2008; Shoham and Xiao 1991; Ouyang and Aziz 1996; Kowalski 1987; Cohen and Hanratty 1968). On the other hand, there are mechanistic models such as the Shoham slug flow model, described in Section E.6 in Appendix E, that in spite of not taking into account the interfacial friction factor, still offer reasonable prediction capabilities (Shoham 2006). This raised the question of the effect of the interfacial friction factor f_i on the prediction of slug flow mechanistic models.

10.2 Case study - Empirical parameter sensitivity of the Zhang et al. mechanistic model

10.2.1 Sensitivity analysis

We will now determine the effect of the empirical parameters L_S, v_{TB}, H_{LLS}, f_i and Θ on the Zhang et al. model prediction and review the three benefits mentioned previously in Section 10.1. The Morris Elementary Effects method (Section B.1, Appendix B) is applied to the Zhang model core. The mechanistic model core is defined as the model without its empirical components.

Ranking of the empirical parameters $\mathbf{L_S}$, $\mathbf{H_{LLS}}$, $\mathbf{v_{TB}}$, $\mathbf{f_i}$ and $\mathbf{\Theta}$

The minimum and maximum values of the empirical parameters were determined for the selected data sample of 95 slug flow data points (Section 2.5) and are shown in Table 10.1. The experimental measurements of Roumazeilles (1994) and Yang (1996) show that the slug liquid holdup can become as low as 0.24. Therefore, the minimum value of H_{LLS} was set to 0.24.

For the purpose of model simplification analysis, these empirical parameters have a uniform distribution, so the elementary effects method can be

Table 10.1: The minimum and maximum values of the input factor for the considered data sample.

Input factor	Minimum	Maximum	Unit [SI]
L_S	0.423	45.439	m
H_{LLS}	0.240	1.000	-
v_{TB}	0.178	8.777	m/s
f_i	0.000	0.172	-
Θ	0.269	1.000	-

Table 10.2: The value of the input factor normalised Morris means for different values of r = number of trajectories.

Input factor	$r = 5$ run 1 case 1	$r = 5$ run 2 case 2	$r = 10$ run 1 case 3	$r = 10$ run 2 case 4	$r = 100$ run 1 case 5	$r = 1000$ run 1 case 6
L_S	0.000001	0.000001	0.000000	0.000000	0.000001	0.000000
H_{LLS}	0.794618	0.419211	0.347085	0.260221	0.451215	0.438789
v_{TB}	0.165120	0.470886	0.347472	0.421766	0.399322	0.366599
f_i	0.000000	0.000002	0.000000	0.000000	0.000003	0.000011
Θ	0.040261	0.109900	0.305443	0.318013	0.149459	0.194601

Table 10.3: Ranking of the empirical parameters in the order from the most influential (highest value) to the least influential (lowest value).

Input Factor	Normalised Morris mean
Liquid holdup in the slug (H_{LLS})	0.438789
Velocity of Taylor bubble (v_{TB})	0.366599
Wetted wall fraction (Θ)	0.194601
Interfacial friction factor (f_i)	0.000011
Liquid slug length (L_S)	0.000000

applied without any prior processing.

Selection of the cell, Δ and parameter r = number of random trajectories
The number of trajectories in the Morris Elementary Effects method was determined as the minimum value for which the values of Morris normalised means have stabilised, as shown in both Table 10.2 and Figure 10.2. For a number of cells $p = 100$, this number of trajectories is $r = 1000$. The last column in Table 10.2 is transferred to Table 10.3, which has the final stabilised values of the normalised means. The sensitivities are assigned these values.

Result interpretation The suite of GNU Octave programs in the *Sensitivity Analysis* folder on the attached CD were written to implement the Morris method.

The means calculated by the Morris method, given in Table 10.3 and Fig-

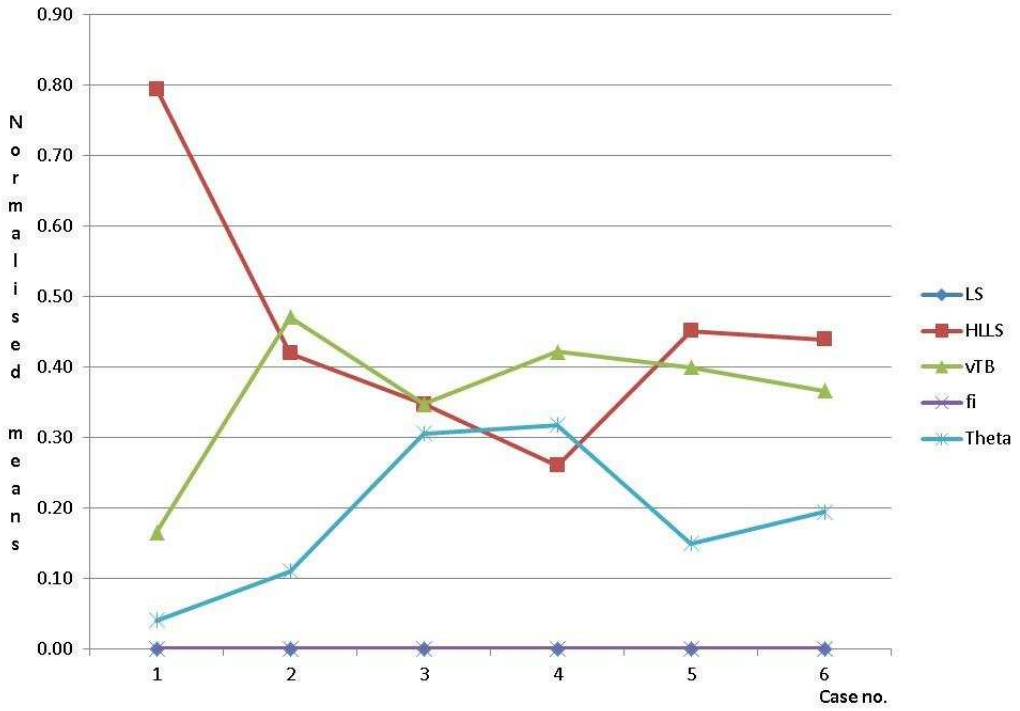


Figure 10.2: Selection of the r = number of trajectories parameter, based on stabilisation of the normalised Morris means.

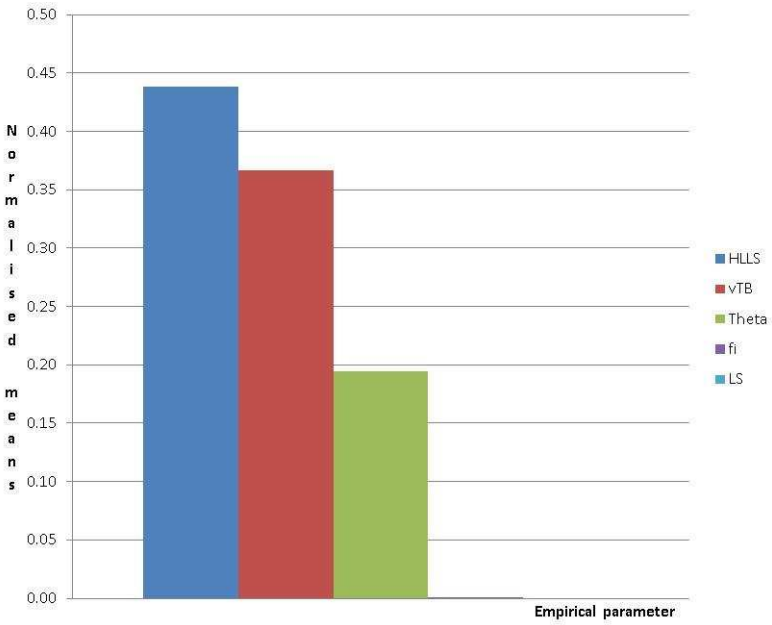


Figure 10.3: Ranking of the empirical parameters from the highest value as the most influential (highest normalised mean) to the least influential (lowest normalised mean).

ure 10.3 show the ranking of the empirical parameters, in the order of the most influential to the least influential. Similarly, Table 10.4 lists the standard deviations, which show the interaction effects between parameters. Low values of both Morris mean (μ) and standard deviation (σ), as shown in both Tables 10.3 and 10.4 for the entries of L_S and f_i, correspond to

Table 10.4: Degrees of correlation between the empirical parameters, with the highest correlation value first.

Parameter	Normalised Standard Deviations
H_{LLS}	0.393804
v_{TB}	0.378810
Θ	0.227305
f_i	0.000080
L_S	0.000001

non-influential input. Due to the difference in the order of magnitude between the last two terms (f_i and L_S) and the first three (H_{LLS}, v_{TB} and Θ) in Table 10.3, we can conclude the following:

1. A method to quickly improve the accuracy of the model, is to find **better correlations for $\mathbf{H_{LLS}}$, $\mathbf{v_{TB}}$ and $\mathbf{\Theta}$** for the input domain where the model will be used.
2. Research in the areas of **liquid holdup, liquid wetted wall fraction and Taylor bubble translational velocity** will have a positive influence on the model's prediction capability.
3. The model could be **simplified** by keeping **$\mathbf{f_i}$ and $\mathbf{L_S}$ constant**. In fact, Cohen and Hanratty (1968) suggested a constant value of $f_i = 0.0142$ for stratified wavy flow in the liquid film area of the slug unit.

10.2.2 A simplified mechanistic model obtained from the Zhang et al. slug flow model

To analyse the performance of the simplified models obtained by keeping constant L_S and f_i, we need to use a mechanistic model whose empirical correlations do not assign constant values to these parameters. Therefore, the Zhang et al. (2000) model is not suitable, because it uses the Cohen and Hanratty correlation, which simply assigns $f_i = 0.0142$. A suitable candidate is the Zhang et al. (2003) model, because it does not use the Cohen and Hanratty correlation for f_i. Instead, the model uses:

- The Andritsos and Hanratty correlation for f_i when the flow in the film zone is stratified.
- The Asali and Ambrosio correlation for f_i, when the flow in the film zone is annular (Zhang et al. 2003c).

Therefore, the numerical simulation compares the following **four** models:

1. The Zhang et al. (2003) mechanistic model that keeps **both $\mathbf{L_S}$ and $\mathbf{f_i}$ constant**.
2. The Zhang et al. (2003) model that only keeps **$\mathbf{L_S}$ constant**.
3. The Zhang et al. (2003) model that only keeps **$\mathbf{f_i}$ constant**.

4. The Zhang et al. (2003) model, with its **original empirical correlations** for L_S and f_i.

on the same data set of 95 slug flow data points, discussed in Section 2.5.

The following constant values are assigned to L_S and f_i in the first three models:

- $\mathbf{f_i = 0.0142}$ following Cohen and Hanratty (1968), who suggested this value for stratified wavy flow. The data sample pertains to near-horizontal flow. This means that stratified flow is very likely to occur. If the inclination angle had been outside the range $[0° \ldots + 10°]$, stratified flow would not have occurred and this value would have been incorrect.

- $\mathbf{L_S}$ kept constant at a value equal to its **average value**. The maximum and minimum values were obtained by running the Optimised Zhang model on the history data. A minimum value of 0.423 m and a maximum value of 45.439 m were determined for L_S. Therefore, the average value of $L_S =$ 22.931 m was chosen.

10.3 Analysis of the performance of the simplified Zhang models

The program that implements the simplified models is found on the attached CD in the *Simplification* folder.

The predictions of the simplified models are compared to the predictions of the Zhang et al. (2003) model. The simulation results are shown in Tables A.12-A.15. The performance comparison results are presented in Table 10.5 and show that for the input data considered:

1. The first two simplified models (L_S and f_i constant, and L_S constant) have a very similar performance factor: $F_{PR}\,(L_S = ct, f_i = ct) = 0.854$, $F_{PR}\,(L_S = ct) = 0.861$.

2. The last two models (f_i constant and the original Zhang et al. (2003) model) have almost the same performance factor: $F_{PR}\,(f_i = ct) = 2.996$, $F_{PR}\,(Zhang2003) = 3.002$.

3. The first two models perform better than the last two models.

10.3.1 Discussion of the simulation results

A spike is noticed in the output of the first two simplified models, as shown in Figure 10.4, for data point (test 934 CRE37, Tables A.12-A.13). This occurs for the model that keeps constant both L_S and f_i, and for the model that keeps constant L_S. The sensitivity analysis in Table 10.3 concluded that f_i had a greater influence on the output than L_S. Therefore, relative to the spike in Figure 10.4, it would be expected that a higher contribution to the spike would come from f_i. This means that the model that keeps f_i constant should exhibit a similar spike for data point (test 934 CRE37). However, the line for $f_i =$ constant in Figure 10.4 shows exactly the opposite. Once f_i is maintained constant and

Table 10.5: E parameter comparison between the cases when ($L_S = ct$ and $f_i = ct$), ($L_S = ct$), ($f_i = ct$) and the Zhang et al. model. The lower values of F_{PR} show the first two models perform better.

Model	E_1	E_2	E_3	E_4	E_5	E_6	F_{PR}
$L_S = ct$ and $f_i = ct$	-2.350	16.630	22.800	573.620	10750.190	5659.374	0.854
$L_S = ct$	-2.343	16.628	22.737	575.599	10750.420	5677.752	0.861
$f_i = ct$	0.159	11.111	1.714	1267.311	10855.920	12293.570	2.996
Zhang (2003)	0.162	11.110	1.735	1268.115	10856.040	12301.370	3.002

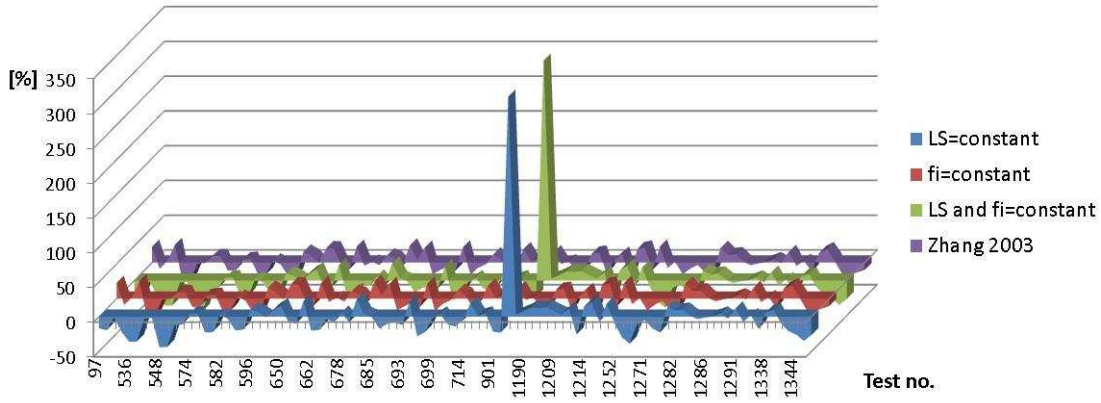


Figure 10.4: The simplified models (blue, red and green lines) follow closely the Zhang et al. (2003) model predictions (purple line).

L_S estimated by an empirical correlation, the spike disappears. This apparently contradicts the statement that f_i has a greater influence than L_S on the model output.

The explanation lies in the fact that the statement we made was incomplete. It is true that the influence of f_i is greater, but on the average, not on a point by point basis. If we examine the method we used in the sensitivity analysis section to rank the empirical parameters (Section B.1, Appendix B), we see that the Elementary Effects method is a global sensitivity method and it calculates the mean of the derivatives for each input parameter. The ranking of the influence of the input parameters on the output of the model is based on sorting these means. In other words, the level of influence is based on averages of the input domain and consequently does not apply to an individual point, but on the average of the values in the input domain.

As this study was approaching completion a paper was published (Sarica, Zhang, and Wilkens 2011) by some of the authors of the Zhang et al. slug flow model, in which they perform a sensitivity analysis study on the influence of the slug length L_S on the slug flow models of Ansari (1994) and Zhang et al. (2000, 2003c). They conclude that for all three models the liquid slug length L_S input factor showed negligible influence on the model output. Their study showed that the Zhang et al. model manifested an increased L_S sensitivity at high gas flow

rates.

These results are in complete agreement with this study. Nevertheless, this study also shows a low influence of the interfacial friction factor f_i and that the important correlations that affect the prediction performance of the Zhang et al. mechanistic model are: the liquid holdup in the slug, H_{LLS} and the Taylor bubble velocity v_{TB}. This means that in order to improve the prediction capabilities of the Zhang et al. model, the empirical correlations for these two factors need to be updated.

10.3.2 The performance of the model with f_i = constant

Over the years many research papers have been dedicated to correlations for the prediction of the interfacial friction factor f_i. The global sensitivity result in Table 10.3, which shows that for the Zhang et al. model the interfacial friction factor has a very low contribution to the model output, is puzzling. Further investigation of the reasons why the research community allocated such a high importance to f_i was required. Most of the research papers review or determine new correlations for f_i in stratified flow in near-horizontal pipelines (Liu et al. 2008; Shoham and Xiao 1991; Ouyang and Aziz 1996; Kowalski 1987; Cohen and Hanratty 1968). It is widely accepted that the interfacial friction factor in stratified flow is associated with the onset of roll waves and slug flow (Liu et al. 2008; Shoham and Xiao 1991). Since slug flow derives from different interfacial wave patterns, Zhang et al. (2008) analyse the mechanism for the transition to slug flow. Classical Kelvin-Helmholtz stability was usually applied for inviscid fluids to explain the transition. Zhang et al. (2008) argue that the instability of a stratified flow is only a necessary condition for slugs to appear. The sufficient condition is slug stability, without which slugs form but gradually may disappear in time. Liu et al. (2008) show that the fraction $\frac{f_i}{f_{GW}}$ has a direct effect on the stability of the slug flow. They suggest that if there is no fluctuation between the gas and liquid phases, f_i could be approximated by f_{GW}, the gas-wall friction factor. However, as soon as the interfacial waves appear, f_i becomes larger than f_{GW}. This is partially in agreement with the results of Cohen and Hanratty (1968) who suggested $f_i = f_{GW}$ for $f_{GW} > 0.0142$ and $f_i = 0.0142$ for stratified wavy flow.

This explains the result obtained in this study, which considered a near-horizontal pipeline data sample of stable slug flow, with mostly stratified flow in the film zone area. According to Cohen and Hanratty, f_i should be set constant, which explains the low sensitivity shown in Table 10.3. In other words, the low sensitivity of the model towards f_i applies to the chosen data sample, but if another data sample had been chosen (which included annular flow in the film zone area of the slug unit), the sensitivity may have been higher.

10.4 Summary

The following can be concluded:

- The slug length L_S was found to have little influence on the output of the Zhang et al. model. This is in agreement with the current research (Sarica,

Zhang, and Wilkens 2011), which showed low influence of L_S on several slug flow models.

- The interfacial friction factor was found to have a low influence on the output. This is in agreement with the research (Cohen and Hanratty 1968), which predicted low sensitivity for stratified flow, so we conclude that this may be due to the nature of the data sample. For the particular input data sample used in this study, the model that assumes f_i = constant performs slightly better than the Zhang et al. (2003) model.

- The empirical correlations for H_{LLS}, v_{TB}, Θ have a high contribution to the model predictions and better empirical correlations for these model parameters would definitely improve the model's performance.

- For the selected data bank, the simplified models that keep L_S and f_i constant, and L_S constant, offered better predictions than the model using empirical correlations for the estimation of these parameters.

Chapter 11

Conclusions and recommendations for future work

As offshore production moves into deeper waters farther offshore, the formation of undesirable substances, such as wax and hydrates, starts occurring. The majority of empirical correlations, developed in flow test facilities for onshore conditions, do not produce accurate predictions for offshore conditions. Consequently, the mechanistic models using these empirical correlations may become unusable offshore. As flow data from offshore production facilities become available, either new correlations developed for the new conditions need to replace existing correlations in the model, or, as this study suggests, other correlation sets, not necessarily specifically developed for the new offshore conditions could be used. In either of these situations, the empirical correlations in the mechanistic flow model need to be replaced by other correlations.

This study addresses this practical problem and proposes a framework for replacing empirical correlations in mechanistic models. **This approach to empirical correlations in mechanistic models has not been taken before by any other research studies and the framework proposed here is the first in this field**. This theoretical framework can be applied in a very similar fashion to any mechanistic model, **irrespective of the process being modelled**.

The conclusions of this study are:

1. The process of using alternative empirical correlations for the same model parameter may involve **changes in the structure of the calculation algorithm** of the mechanistic model. This means that one cannot "unplug" an existing empirical correlation and replace it by "plugging in" another correlation. To this end, a method was proposed in Chapter 7 to analyse the structural changes of the algorithm and place the empirical parameter at the appropriate step. The method, which uses functional dependency analysis and directed graphs, was shown in Section 7.5.

2. Two main applications of the framework for using alternative empirical correlation sets have been identified:

   (a) **Mechanistic model optimisation**. Chapter 6 outlined a hybrid modelling approach to model optimisation. The **training** phase discussed in Chapter 8, proposes a **curve-fitting method** based on us-

ing multiple empirical correlation sets, for the purpose of building a **history table**. This ensures that the model is "trained" first on a given input/output data set, to know the optimal correlation sets for each such data 2-tuple. The table is the fundamental element for **two prediction methods** proposed by this study:

i. **Optimal prediction using a history matching algorithm based on sensitivity analysis**. Along with the history table, this prediction method is based on the minimisation of an objective function, whose weight coefficients are calculated using sensitivity analysis. The minimisation determines the optimal correlation set to be used for output prediction, depending on where the input is located in the input domain.
This was exemplified in Section 9.3.1, as a case study of the Zhang et al. (2000) slug flow model, where an objective function was built based on the elementary effects global sensitivity analysis method. It was shown that the optimised model had a performance superior to that of the original model. A general procedure to optimise a mechanistic model was described in Section 9.3.2.

ii. **Optimal prediction using data mining**. This method is based on data mining techniques, such as sorting and partitioning of the history table, in order to determine regions where one optimal correlation set prevails. Unlike the previous prediction method, this may only lead to regional optimisations.
For the TUFFP data set analysed, it was noticed that the correlation set $L_S = 1$, $H_{LLS} = 1$, $v_{TB} = 1$, $f_i = 1$, $\Theta = 1$ prevailed in the region of low superficial liquid velocities, $v_{SL} < 0.1$ m/s. This means that:
   - The empirical correlations in the Zhang et al. (2000) model are optimal for the selected data bank.
   - In order to be optimal in the region of low superficial velocities, the Zhang et al. (2003) model's empirical correlation set has to be replaced with the set: $L_S = 1$, $H_{LLS} = 1$, $v_{TB} = 1$, $f_i = 1$, $\Theta = 1$.

(b) **Mechanistic model simplification**. In cases where empirical correlations are difficult to find, or the computational intensity of the correlations is high, sensitivity analysis, as shown in Chapter 10 may offer the answer to simplifying the model. The case study of the Zhang et al. model described in Section 10.2 showed that for the particular data sample used, the simplified models obtained by keeping the interfacial friction factor f_i and the length of the slug L_S constant, performed better than the Zhang et al. (2003) model, which used empirical correlations for the calculation of these empirical parameters.

3. In order to **identify which model parameter brings about the highest improvement in model prediction**, the study proposes a sensitivity analysis method be applied to the model, for the input domain used in the application. This will determine which empirical parameter requires an investigation for more accurate empirical correlations.

In the case of Zhang et al. model, the sensitivity analysis performed in Section 10.2 and shown in Table 10.3 revealed that correlations for H_{LLS} have the highest influence on the model predictions. Therefore, for the Zhang et al. model, investigations into finding better correlations for H_{LLS} can bring significant improvements to the model.

Empirical correlations in mechanistic models make them **data-sensitive**, to the extent that they become **data-driven**. This study suggests that before applying a mechanistic model that contains empirical correlations to production data, a set of alternative empirical correlations be applied first, in order to optimise the model. Section 9.4 gives the example of the Zhang et al. mechanistic flow model, where the authors replaced empirical correlations between successive versions of the model in order to improve its performance. However, this study has shown that for the selected data range, the "improved" model performed worse. This emphasizes the idea that **a framework for empirical correlation replacement should be applied to the model first, before it is used on any production data.**

11.1 Recommendations for further research

1. The study showed that in the case of the Zhang et al. (2000) and (2003) mechanistic models, the parameter L_S had a negligible influence on the output. This is a confirmation of recent research (Sarica, Zhang, and Wilkens 2011), which analysed multiple slug flow models and reached the same conclusion. The sensitivity analysis performed in Section 10.2 showed that parameters H_{LLS}, v_{TB} and Θ have a high influence on the model output. A recommendation is made to assess the influence of these factors on other well-known mechanistic models. Should the results be consistent with the results in this study, which only analysed the Zhang et al. model, it could be concluded that research in the area of improving empirical correlations for **liquid holdup, Taylor bubble velocity and wetted wall fraction** is of the greatest importance for improving the accuracy of mechanistic slug flow models.

2. The previous recommendation can be generalised by making a research proposition to investigate multiple mechanistic models for a certain process, and perform sensitivity analysis on their empirical parameters. This investigation would determine the most influential empirical parameters for each one of them and identify a common set of influential parameters. Further on, these parameters would determine the class of empirical correlations that needs to be further researched to bring the best accuracy improvement in the entire group of mechanistic models considered.

3. The development of a mechanistic model, from a collection of first principle equations and empirical correlations into a calculation algorithm is a complex process. The mapping is sometimes not unique, and multiple calculation algorithms may exist for the same series of equations. The functional dependency graph discussed in Chapter 7 could be used to this end. A research initiative is suggested here to develop an algorithm to convert the

series of model equations into a calculation algorithm. Not only would this help with the development of a calculation algorithm, but also with the identification of equivalent calculation algorithms and the selection of the best performing one.

Appendix A

Tables

A.1 Data tables

Table A.1: Experimental conditions of the TUFFP pipeline databank. A=air, W=water, K=kerosene, LO=lube oil. The data sources are: Andritsos (1986), Beggs (1972), Cheremisinoff (1977), Minami (1983), Mukherjee (1985).

Data source	Fluid used	d [in]	v_{SL} [ft/s]	v_{SG} [ft/s]	ρ_L [lb/ft^3]	μ_L [cp]	σ [dynes/cm]	P [psi]	β [°]
And.	A/W, A/G	0.99, 3.75	0.0009 - 1.099	2.62 - 535.6	62.4 - 76.2	1 - 80	66 - 73	14.3 - 28.4	0
Beg.	A/W	1.0, 1.5	0.074 - 17.07	0.98 - 83.8	62.4	0.78 - 1.4	70.6 - 75.6	16 - 98.9	10, 5, 0, -5, -10
Che.	A/W	2.5	0.054 - 0.228	8.47 - 82.8	62.4	0.93 - 1.15	70	14.7	0
Min.	A/W, A/K	3.07	0.017 - 3.12	1.56 - 54.4	49.4 - 62.4	0.58 - 1.99	26 - 72.1	16 - 96.7	0
Muk.	A/K, A/LO	1.5	0.05 - 13.07	0.3 - 93.5	49.1 - 53.9	0.92 - 61.9	22.6 - 37.0	16 - 91.9	5, 0, -5

A.2 Empirical correlation codification

Tables throughout this research study make reference to empirical correlations by their code number. The codification correlation number-correlation name is specified in Tables A.2- A.6.

Table A.2: Empirical correlations for the calculation of parameter L_S and their codification.

Code	Empirical correlation for L_S
1	Zhang
2	Scott
3	Felizola and Shoham

Table A.3: Empirical correlations for the calculation of parameter H_{LLS} and their codification.

Code	Empirical correlation for H_{LLS}
1	Gregory
2	Gomez
3	Barnea and Brauner
4	Eissa Al-Safran
5	Woldesmayat and Ghajar
6	Toshiba and Leung
7	Spedding and Spencer
8	Minami and Brill
9	Nicklin
10	Zhang

Table A.4: Empirical correlations for the calculation of parameter v_{TB} and their codification.

Code	Empirical correlation for v_{TB}
1	Bendiksen
2	Dukler and Hubbard

Table A.5: Empirical correlations for the calculation of parameter f_i and their codification.

Code	Empirical correlation for f_i
1	Cohen and Hanratty
2	Andritsos
3	Ouyang and Aziz
4	Bendiksen
5	Vlachos

Table A.6: Empirical correlations for the calculation of parameter Θ and their codification.

Code	Empirical correlation for Θ
1	Grolman
2	Biberg
3	Fan
4	Hart

A.3 Model comparison tables

Table A.7: Zhang et al. (2000) PG predicted values against the measured values for each test in the selected data bank. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no	Predicted PG [Pa/m]	Measured PG [Pa/m]	Predicted PG Err [%]
97	AND97	32.75	26.69	-14.29
113	AND113	80.47	88.44	-17.87
129	AND129	196.42	192.95	0.55
343	AND343	64.78	51.57	-22.08
536	BEG1	68.79	92.74	-34.81
537	BEG2	69.08	92.74	-34.53
538	BEG3	333.74	339.30	-1.39
539	BEG4	337.24	339.30	-0.33
548	BEG13	59.38	54.28	-44.51
549	BEG14	59.37	54.28	-44.62
550	BEG15	998.12	1131.02	-22.23
551	BEG16	1006.31	1131.02	-22.16
574	BEG39	709.96	678.61	4.99
575	BEG40	716.94	678.61	6.00
580	BEG45	148.74	180.96	-21.22
581	BEG46	149.42	180.96	-20.85
582	BEG47	5483.87	5315.82	3.90
583	BEG48	5977.64	6175.40	-2.50
588	BEG53	1996.46	2488.25	-19.16
589	BEG54	2022.40	2488.25	-18.09
596	BEG61	1205.91	1040.54	15.98
617	BEG82	2975.82	2714.46	10.00
619	BEG84	6617.60	6673.05	2.35
629	BEG94	689.95	570.03	20.62
650	BEG115	2443.98	2035.84	20.05
659	BEG124	324.29	339.30	-9.15
660	BEG125	691.40	769.09	-10.06
661	BEG126	2391.87	1945.36	23.97
662	BEG127	269.09	282.75	-20.48
665	BEG130	183.09	180.96	-21.92
666	BEG131	1699.97	1764.40	-10.67
673	BEG138	399.79	361.92	-7.06
678	BEG143	1132.10	1063.16	6.59
681	BEG146	405.62	429.78	-8.26
682	BEG147	4233.17	3393.07	25.55
684	BEG149	7532.78	7238.56	4.41
685	BEG150	385.82	316.68	-0.56
689	BEG154	2725.35	3279.97	-16.54
690	BEG155	870.42	972.68	-10.71

Continued on next page

Table A.7 – continued from previous page

No	Test no	Predicted PG [Pa/m]	Measured PG [Pa/m]	Predicted PG Err [%]
692	BEG157	587.71	622.06	-9.74
693	BEG158	1058.41	1198.88	-12.01
694	BEG159	2895.03	2375.15	22.48
695	BEG160	409.57	497.65	-29.60
698	BEG163	313.89	339.30	-24.25
699	BEG164	2399.22	2465.63	-7.79
705	BEG170	807.34	723.85	4.96
706	BEG171	524.17	497.65	-11.21
711	BEG176	1166.20	1357.23	-14.16
714	BEG179	674.79	701.23	-6.41
715	BEG180	4778.89	4003.83	19.88
722	BEG187	3441.57	3393.07	1.66
723	BEG188	1517.20	1447.71	4.54
901	CRE4	23.20	26.69	-21.24
906	CRE9	113.87	99.07	-23.07
934	CRE37	15.30	15.60	36.92
1188	MUK1	3992.32	3923.30	5.56
1190	MUK3	860.87	1016.11	13.82
1191	MUK4	844.36	949.38	19.46
1202	MUK15	401.91	356.04	12.57
1203	MUK16	466.35	407.84	13.83
1209	MUK22	869.17	1030.81	13.40
1210	MUK23	422.14	407.84	2.98
1211	MUK24	518.92	459.87	11.90
1213	MUK26	166.61	215.12	-25.24
1214	MUK27	521.86	437.48	12.55
1227	MUK40	819.57	674.99	21.15
1245	MUK58	3854.73	4323.91	-1.27
1250	MUK63	1223.83	993.71	15.76
1252	MUK65	607.48	660.06	-9.09
1256	MUK69	164.54	163.09	-34.54
1258	MUK71	426.23	511.67	-47.61
1270	MUK83	473.34	519.14	-10.66
1271	MUK84	494.19	511.67	-4.94
1272	MUK85	444.97	519.14	-16.63
1274	MUK87	570.75	667.53	-26.08
1281	MUK94	296.85	244.75	13.78
1282	MUK95	468.85	422.77	8.58
1283	MUK96	589.40	526.60	10.32
1284	MUK97	623.85	615.50	1.13
1285	MUK98	699.78	726.79	-3.70
1286	MUK99	1005.07	1016.11	-1.57
1287	MUK100	1022.26	1016.11	0.36
1288	MUK101	1065.56	1016.11	4.37

Continued on next page

Table A.7 – continued from previous page

No	Test no	Predicted PG [Pa/m]	Measured PG [Pa/m]	Predicted PG Err [%]
1289	MUK102	609.65	615.50	-2.92
1291	MUK104	857.06	741.72	9.87
1292	MUK105	705.25	741.72	-8.26
1298	MUK111	941.71	919.75	2.31
1337	MUK150	297.36	341.11	-34.10
1338	MUK151	524.42	459.87	-9.07
1339	MUK152	584.24	489.50	-6.98
1340	MUK153	260.31	252.21	-28.87
1343	MUK156	156.36	185.48	-35.87
1344	MUK157	73.20	88.89	-40.75
1345	MUK158	216.16	252.21	-60.22
1354	MUK167	119.24	133.46	-51.26

Table A.8: Zhang et al. (2003) pressure gradient (PG) predicted values against the measured values for each test in the selected data bank. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no	Predicted PG [Pa/m]	Measured PG [Pa/m]	Predicted PG Err [%]
97	AND97	32.75	26.69	22.71
113	AND113	80.47	88.44	-9.01
129	AND129	196.42	192.95	1.79
343	AND343	64.78	51.57	25.60
536	BEG1	68.79	92.74	-25.82
537	BEG2	69.08	92.74	-25.51
538	BEG3	333.74	339.30	-1.63
539	BEG4	337.24	339.30	-0.60
548	BEG13	59.38	54.28	9.38
549	BEG14	59.37	54.28	9.37
550	BEG15	998.12	1131.02	-11.75
551	BEG16	1006.31	1131.02	-11.02
574	BEG39	709.96	678.61	4.61
575	BEG40	716.94	678.61	5.64
580	BEG45	148.74	180.96	-17.80
581	BEG46	149.42	180.96	-17.42
582	BEG47	5483.87	5315.82	3.16
583	BEG48	5977.64	6175.40	-3.20
588	BEG53	1996.46	2488.25	-19.76
589	BEG54	2022.40	2488.25	-18.72
596	BEG61	1205.91	1040.54	15.89
617	BEG82	2975.82	2714.46	9.62
619	BEG84	6617.60	6673.05	-0.83
629	BEG94	689.95	570.03	21.03
650	BEG115	2443.98	2035.84	20.04
659	BEG124	324.29	339.30	-4.42
660	BEG125	691.40	769.09	-10.10
661	BEG126	2391.87	1945.36	22.95
662	BEG127	269.09	282.75	-4.83
665	BEG130	183.09	180.96	1.17
666	BEG131	1699.97	1764.40	-3.65
673	BEG138	399.79	361.92	10.46
678	BEG143	1132.10	1063.16	6.48
681	BEG146	405.62	429.78	-5.62
682	BEG147	4233.17	3393.07	24.75
684	BEG149	7532.78	7238.56	4.06
685	BEG150	385.82	316.68	21.83
689	BEG154	2725.35	3279.97	-16.90
690	BEG155	870.42	972.68	-10.51
692	BEG157	587.71	622.06	-5.52
693	BEG158	1058.41	1198.88	-11.711

Continued on next page

Table A.8 – continued from previous page

No	Test no	Predicted PG [Pa/m]	Measured PG [Pa/m]	Predicted PG Err [%]
694	BEG159	2895.03	2375.15	21.88
695	BEG160	409.57	497.65	-17.69
698	BEG163	313.89	339.30	-7.48
699	BEG164	2399.22	2465.63	-2.69
705	BEG170	807.34	723.85	11.53
706	BEG171	524.17	497.65	5.32
711	BEG176	1166.20	1357.23	-14.07
714	BEG179	674.79	701.23	-3.77
715	BEG180	4778.89	4003.83	19.35
722	BEG187	3441.57	3393.07	1.42
723	BEG188	1517.20	1447.71	4.79
901	CRE4	23.20	26.69	-13.05
906	CRE9	113.87	99.07	14.93
934	CRE37	15.30	15.60	-1.94
1188	MUK1	3992.32	3923.30	1.75
1190	MUK3	860.87	1016.11	-15.27
1191	MUK4	844.36	949.38	-11.06
1202	MUK15	401.91	356.04	12.88
1203	MUK16	466.35	407.84	14.34
1209	MUK22	869.17	1030.81	-15.68
1210	MUK23	422.14	407.84	3.50
1211	MUK24	518.92	459.87	12.84
1213	MUK26	166.61	215.12	-22.54
1214	MUK27	521.86	437.48	19.28
1227	MUK40	819.57	674.99	21.41
1245	MUK58	3854.73	4323.91	-10.85
1250	MUK63	1223.83	993.71	23.15
1252	MUK65	607.48	660.06	-7.96
1256	MUK69	164.54	163.09	0.88
1258	MUK71	426.23	511.67	-16.69
1270	MUK83	473.34	519.14	-8.82
1271	MUK84	494.19	511.67	-3.41
1272	MUK85	444.97	519.14	-14.28
1274	MUK87	570.75	667.53	-14.49
1281	MUK94	296.85	244.75	21.28
1282	MUK95	468.85	422.77	10.89
1283	MUK96	589.40	526.60	11.92
1284	MUK97	623.85	615.50	1.35
1285	MUK98	699.78	726.79	-3.71
1286	MUK99	1005.07	1016.11	-1.08
1287	MUK100	1022.26	1016.11	0.60
1288	MUK101	1065.56	1016.11	4.86
1289	MUK102	609.65	615.50	-0.95
1291	MUK104	857.06	741.72	15.54

Continued on next page

Table A.8 – continued from previous page

No	Test no	Predicted PG [Pa/m]	Measured PG [Pa/m]	Predicted PG Err [%]
1292	MUK105	705.25	741.72	-4.91
1298	MUK111	941.71	919.75	2.38
1337	MUK150	297.36	341.11	-12.82
1338	MUK151	524.42	459.87	14.03
1339	MUK152	584.24	489.50	19.35
1340	MUK153	260.31	252.21	3.20
1343	MUK156	156.36	185.48	-15.69
1344	MUK157	73.20	88.89	-17.643
1345	MUK158	216.16	252.21	-14.29
1354	MUK167	119.24	133.46	-10.65

Table A.9: The history table generated by the curve-fitting method during the training stage.

No	Test no.	L_S corr. no.	H_{LLS} corr. no.	v_{TB} corr. no.	f_i corr. no.	Θ corr. no.	$\tilde{y}$ [Pa/m]	y_t [Pa/m]	R_{min} [Pa/m]
97	AND97	1	4	2	5	2	27.43	26.69	0.74
113	AND113	3	10	1	4	3	88.56	88.45	0.11
129	AND129	1	1	1	4	1	194.01	192.95	1.06
343	AND343	1	4	1	4	4	51.66	51.57	0.09
536	BEG1	3	3	2	3	3	92.34	92.74	0.40
537	BEG2	3	10	2	2	1	92.75	92.74	0.01
538	BEG3	3	10	1	1	1	338.68	339.31	0.63
539	BEG4	1	1	1	1	1	338.16	339.31	1.15
548	BEG13	1	4	1	1	2	55.22	54.29	0.93
549	BEG14	1	4	1	1	2	55.15	54.29	0.86
550	BEG15	1	2	1	4	4	1027.94	1131.03	103.09
551	BEG16	1	2	1	4	4	1037.06	1131.03	93.97
574	BEG39	1	7	1	3	1	678.49	678.62	0.13
575	BEG40	1	4	2	4	2	682.54	678.62	3.92
580	BEG45	1	2	2	1	2	180.56	180.96	0.40
581	BEG46	1	2	2	1	2	180.76	180.96	0.20
582	BEG47	3	9	2	1	1	5318.42	5315.82	2.60
583	BEG48	1	10	2	3	1	6193.62	6175.40	18.22
588	BEG53	1	7	1	3	2	2415.46	2488.26	72.80
589	BEG54	1	7	1	3	2	2453.93	2488.26	34.33
596	BEG61	3	10	1	3	2	1035.18	1040.54	5.36
617	BEG82	3	2	2	4	1	2724.01	2714.46	9.55
619	BEG84	1	9	1	4	1	6672.28	6673.05	0.77
629	BEG94	1	10	2	1	1	572.12	570.04	2.08
650	BEG115	1	1	2	3	2	2020.63	2035.85	15.22
659	BEG124	2	4	1	5	4	336.67	339.31	2.64
660	BEG125	2	10	2	1	1	767.66	769.10	1.44
661	BEG126	2	10	2	5	2	1941.90	1945.36	3.46
662	BEG127	1	2	2	1	1	277.24	282.76	5.52
665	BEG130	1	4	1	2	2	180.78	180.96	0.18
666	BEG131	1	2	1	4	4	1721.87	1764.40	42.53
673	BEG138	3	4	2	1	1	351.90	361.93	10.03
678	BEG143	1	3	1	1	3	1058.92	1063.16	4.24
681	BEG146	1	8	1	1	3	428.58	429.79	1.21
682	BEG147	1	4	1	1	2	3497.53	3393.08	104.45
684	BEG149	2	4	1	4	1	7296.15	7238.57	57.58
685	BEG150	1	1	2	4	2	317.60	316.69	0.91
689	BEG154	3	5	2	1	2	2773.46	3279.97	506.51
690	BEG155	2	4	2	4	4	1009.56	972.68	36.88
692	BEG157	3	10	1	4	4	627.04	622.06	4.98
693	BEG158	1	10	2	4	1	1211.75	1198.89	12.86

Continued on next page

Table A.9 – continued from previous page

No	Test no.	L_S corr. no.	H_{LLS} corr. no.	v_{TB} corr. no.	f_i corr. no.	Θ corr. no.	$\tilde{y}$ [Pa/m]	y_t [Pa/m]	R_{min} [Pa/m]
694	BEG159	1	3	2	1	2	2373.40	2375.15	1.75
695	BEG160	3	10	1	4	4	498.61	497.65	0.96
698	BEG163	3	10	1	4	4	341.30	339.31	1.99
699	BEG164	2	2	1	4	4	2407.96	2465.64	57.68
705	BEG170	1	4	1	4	1	728.15	723.86	4.29
706	BEG171	1	10	1	4	3	506.79	497.65	9.14
711	BEG176	3	10	2	1	1	1316.61	1357.23	40.62
714	BEG179	3	10	1	1	1	684.92	701.24	16.32
715	BEG180	1	4	1	1	2	4041.62	4003.83	37.39
722	BEG187	1	5	2	1	2	3404.72	3393.08	11.64
723	BEG188	1	8	2	1	1	1447.58	1447.71	0.13
901	CRE4	1	10	1	1	2	26.72	26.69	0.03
906	CRE9	3	2	2	4	4	99.32	99.08	0.24
934	CRE37	1	10	1	1	1	15.31	15.61	0.30
1188	MUK1	1	8	2	3	1	3967.79	3923.30	44.49
1190	MUK3	1	5	1	5	1	1019.75	1016.11	3.64
1191	MUK4	1	4	2	4	3	949.59	949.38	0.21
1202	MUK15	1	5	1	3	1	355.45	356.05	0.60
1203	MUK16	1	2	1	1	2	405.83	407.85	2.02
1209	MUK22	1	9	2	4	3	1033.61	1030.82	2.79
1210	MUK23	1	3	2	1	1	406.71	407.85	1.14
1211	MUK24	2	1	2	4	2	460.41	459.88	0.53
1213	MUK26	1	7	1	3	3	216.18	215.12	1.06
1214	MUK27	1	3	2	4	2	456.30	437.48	18.82
1227	MUK40	1	4	1	3	3	672.23	675.00	2.77
1245	MUK58	3	1	2	3	2	4320.49	4323.91	3.42
1250	MUK63	2	1	2	4	2	995.57	993.72	1.85
1252	MUK65	1	1	1	1	2	659.83	660.07	0.24
1256	MUK69	1	10	1	4	4	163.72	163.09	0.63
1258	MUK71	3	10	1	4	3	514.76	511.68	3.08
1270	MUK83	3	10	1	4	2	520.73	519.14	1.59
1271	MUK84	1	3	1	1	2	510.98	511.68	0.70
1272	MUK85	3	10	2	1	2	510.33	519.14	8.81
1274	MUK87	1	9	2	4	3	669.01	667.53	1.48
1281	MUK94	1	10	2	5	2	245.40	244.75	0.65
1282	MUK95	1	3	1	5	4	452.94	422.78	30.16
1283	MUK96	1	6	1	1	2	487.31	526.61	39.30
1284	MUK97	1	2	1	1	2	615.25	615.50	0.25
1285	MUK98	3	4	2	1	4	726.51	726.80	0.29
1286	MUK99	3	4	1	1	1	1015.23	1016.11	0.88
1287	MUK100	1	9	1	4	1	1017.72	1016.11	1.61
1288	MUK101	1	8	1	1	2	1016.19	1016.11	0.08
1289	MUK102	3	10	1	1	1	614.34	615.50	1.16

Continued on next page

Table A.9 – continued from previous page

No	Test no.	L_S corr. no.	H_{LLS} corr. no.	v_{TB} corr. no.	f_i corr. no.	Θ corr. no.	$\tilde{y}$ [Pa/m]	y_t [Pa/m]	R_{min} [Pa/m]
1291	MUK104	1	1	2	1	1	705.33	741.73	36.40
1292	MUK105	1	5	2	4	3	741.73	741.73	0.00
1298	MUK111	1	5	2	1	1	917.09	919.75	2.66
1337	MUK150	3	4	2	5	4	341.78	341.12	0.66
1338	MUK151	1	4	1	1	1	458.94	459.88	0.94
1339	MUK152	1	4	1	5	4	493.88	489.51	4.37
1340	MUK153	1	9	1	5	1	253.05	252.22	0.83
1343	MUK156	3	8	2	5	2	185.80	185.49	0.31
1344	MUK157	1	6	2	3	1	88.81	88.90	0.09
1345	MUK158	1	10	2	5	4	258.73	252.22	6.51
1354	MUK167	2	10	1	5	2	133.55	133.46	0.09

Table A.10: Best correlation sets for horizontal flow, with data sorted by increasing values of v_{SL}. In the low superficial velocity range, $v_{SL} < 0.1$ [m/s], the dominant correlation set is $L_S = 1$, $H_{LLS} = 1$, $v_{TB} = 1$, $f_i = 1$ and $\Theta = 1$.

v_{SL} [m/s]	Angle [°]	v_{SG} [m/s]	d [m]	L_S corr. no.	H_{LLS} corr. no.	v_{TB} corr. no.	f_i corr. no.	Θ corr. no.
0.005791	0	0.643128	0.077978	1	1	1	1	1
0.005791	0	1.011936	0.077978	1	1	1	1	1
0.005791	0	1.743456	0.077978	1	1	1	1	1
0.005791	0	3.166872	0.077978	1	1	1	1	1
0.005791	0	5.574792	0.077978	1	1	1	1	1
0.009754	0	0.600456	0.077978	1	1	1	1	1
0.009754	0	1.075944	0.077978	1	1	1	1	1
0.009754	0	1.847088	0.077978	1	1	1	1	1
0.009754	0	3.118104	0.077978	1	1	1	1	1
0.009754	0	5.879592	0.077978	1	1	1	1	1
0.014326	0	0.85344	0.077978	1	1	1	1	1
0.017069	0	1.072896	0.077978	1	1	1	1	1
0.017069	0	1.837944	0.077978	1	1	1	1	1
0.017069	0	3.060192	0.077978	1	1	1	1	1
0.017069	0	5.903976	0.077978	1	1	1	1	1
0.017374	0	0.542544	0.077978	1	1	1	1	1
0.03048	0	1.033272	0.077978	1	1	1	1	1
0.03048	0	1.780032	0.077978	1	1	1	1	1
0.03048	0	3.096768	0.077978	1	1	1	1	1
0.03048	0	5.635752	0.077978	1	1	1	1	1
0.03109	0	0.551688	0.077978	1	1	1	1	1
0.042062	0	2.310384	0.0381	1	4	1	1	2
0.042062	0	2.313432	0.0381	1	4	1	1	2
0.05334	0	3.13944	0.077978	1	1	1	1	1
0.055169	0	5.733288	0.077978	1	1	1	1	1
0.055778	0	0.981456	0.077978	1	1	1	1	1
0.056083	0	1.795272	0.077978	1	1	1	1	1
0.056388	0	0.612648	0.077978	1	1	1	1	1
0.063002	0	1.161288	0.025146	1	4	2	5	2
0.082296	0	0.192024	0.0381	2	10	1	5	2
0.098146	0	1.834896	0.077978	1	1	1	1	1
0.098755	0	1.014984	0.077978	1	1	1	1	1
0.098755	0	5.715	0.077978	1	1	1	1	1
0.101194	0	3.084576	0.077978	1	1	1	1	1
0.109728	0	0.09144	0.0381	1	6	2	3	1
0.109728	0	0.371856	0.0381	1	10	2	5	4
0.139995	0	1.161288	0.025146	3	10	1	4	3
0.148986	0	3.160776	0.09525	1	4	1	4	4
0.165842	0	1.182624	0.0381	3	3	2	3	3
0.165842	0	1.185672	0.0381	3	10	2	2	1

Continued on next page

Table A.10 – continued from previous page

v_{SL} [m/s]	Angle [°]	v_{SG} [m/s]	d [m]	L_S corr. no.	H_{LLS} corr. no.	v_{TB} corr. no.	f_i corr. no.	Θ corr. no.
0.173431	0	1.70688	0.077978	1	1	1	1	1
0.17465	0	5.586984	0.077978	1	1	1	1	1
0.17587	0	3.392424	0.077978	1	1	1	1	1
0.176479	0	2.584704	0.077978	1	1	1	1	1
0.201686	0	1.109472	0.0254	1	2	2	1	2
0.201686	0	1.11252	0.0254	1	2	2	1	2
0.277368	0	0.103632	0.0381	3	8	2	5	2
0.277368	0	0.286512	0.0381	1	9	1	5	1
0.291998	0	1.170432	0.025146	1	1	1	4	1
0.291998	0	2.901696	0.077978	1	1	1	1	1
0.32065	0	5.202936	0.077978	1	1	1	1	1
0.44196	0	0.155448	0.0381	3	4	2	5	4
0.4572	0	0.6096	0.17145	1	10	1	1	2
0.530352	0	0.688848	0.0381	1	7	1	3	3
0.552115	0	1.185672	0.0381	3	10	1	1	1
0.552115	0	1.197864	0.0381	1	1	1	1	1
0.562356	0	4.681728	0.077978	1	1	1	1	1
0.566928	0	1.749552	0.0381	1	3	2	4	2
0.582168	0	0.435864	0.0381	1	4	1	5	4
0.603504	0	0.35052	0.0381	1	4	1	1	1
0.622676	0	1.100328	0.0254	1	7	1	3	1
0.622676	0	1.11252	0.0254	1	4	2	4	2
1.136904	0	0.676656	0.0381	2	1	2	4	2
1.149096	0	0.3048	0.0381	1	5	1	3	1
1.155192	0	0.454152	0.0381	1	3	2	1	1
1.161288	0	0.4572	0.0381	1	2	1	1	2
1.603431	0	0.362712	0.0381	1	2	1	4	4
1.603431	0	0.371856	0.0381	1	2	1	4	4
1.697736	0	0.402336	0.0381	1	4	1	3	3
1.85547	0	0.301752	0.0254	1	7	1	3	2
1.85547	0	0.316992	0.0254	1	7	1	3	2
2.231136	0	0.316992	0.0381	1	4	2	4	3
2.231136	0	0.332232	0.0381	1	5	1	5	1
2.231136	0	0.39624	0.0381	1	9	2	4	3
2.61619	0	1.027176	0.0254	3	9	2	1	1
2.61619	0	1.194816	0.0254	1	10	2	3	1
3.983736	0	1.956816	0.0381	1	8	2	3	1

Table A.12: Performance of the simplified model for the case L_S=ct and f_i=ct vs measured pressure gradient values. PG = pressure gradient. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no.	Simplified L_S=ct f_i=ct PG [Pa/m]	Measured PG [Pa/m]	Simplified L_S=ct f_i=ct PG Err [%]
97	AND97	22.36	26.69	-16.22
113	AND113	71.15	88.44	-19.54
129	AND129	190.69	192.95	-1.17
343	AND343	40.11	51.57	-22.21
536	BEG1	59.29	92.74	-36.06
537	BEG2	59.47	92.74	-35.86
538	BEG3	329.77	339.30	-2.80
539	BEG4	333.18	339.30	-1.80
548	BEG13	30.51	54.28	-43.79
549	BEG14	30.56	54.28	-43.69
550	BEG15	997.38	1131.02	-11.81
551	BEG16	1005.54	1131.02	-11.09
574	BEG39	705.95	678.61	4.02
575	BEG40	712.84	678.61	5.04
580	BEG45	140.21	180.96	-22.51
581	BEG46	140.66	180.96	-22.26
582	BEG47	5483.26	5315.82	3.14
583	BEG48	5976.54	6175.40	-3.22
588	BEG53	1996.45	2488.25	-19.76
589	BEG54	2022.31	2488.25	-18.72
596	BEG61	1145.86	1040.54	10.12
617	BEG82	2903.57	2714.46	6.96
619	BEG84	6627.14	6673.05	-0.68
629	BEG94	623.37	570.03	9.35
650	BEG115	2367.42	2035.84	16.28
659	BEG124	311.42	339.30	-8.21
660	BEG125	691.88	769.09	-10.03
661	BEG126	2386.54	1945.36	22.67
662	BEG127	227.05	282.75	-19.70
665	BEG130	145.75	180.96	-19.45
666	BEG131	1699.27	1764.40	-3.69
673	BEG138	334.39	361.92	-7.60
678	BEG143	1128.98	1063.16	6.19
681	BEG146	395.35	429.78	-8.01
682	BEG147	4229.46	3393.07	24.64
684	BEG149	7533.45	7238.56	4.07
685	BEG150	313.39	316.68	-1.03
689	BEG154	2725.36	3279.97	-16.90
690	BEG155	871.28	972.68	-10.42
692	BEG157	569.11	622.06	-8.51
693	BEG158	1059.89	1198.88	-11.59

Continued on next page

Table A.12 – continued from previous page

No	Test no.	Simplified L_S=ct f_i=ct PG [Pa/m]	Measured PG [Pa/m]	Simplified L_S=ct f_i=ct PG Err [%]
694	BEG159	2889.92	2375.15	21.67
695	BEG160	356.79	497.65	-28.30
698	BEG163	264.76	339.30	-21.96
699	BEG164	2398.62	2465.63	-2.71
705	BEG170	757.33	723.85	4.62
706	BEG171	442.88	497.65	-11.00
711	BEG176	1167.81	1357.23	-13.95
714	BEG179	660.58	701.23	-5.79
715	BEG180	4775.00	4003.83	19.26
722	BEG187	3441.58	3393.07	1.42
723	BEG188	1519.61	1447.71	4.96
901	CRE4	20.95	26.69	-21.50
906	CRE9	76.28	99.07	-23.00
934	CRE37	64.99	15.60	316.41
1188	MUK1	4047.59	3923.30	3.16
1190	MUK3	1094.39	1016.11	7.70
1191	MUK4	1074.78	949.38	13.20
1202	MUK15	400.74	356.04	12.55
1203	MUK16	464.25	407.84	13.83
1209	MUK22	1110.44	1030.81	7.72
1210	MUK23	419.96	407.84	2.97
1211	MUK24	515.44	459.87	12.08
1213	MUK26	160.54	215.12	-25.36
1214	MUK27	499.55	437.48	14.18
1227	MUK40	819.13	674.99	21.35
1245	MUK58	4025.24	4323.91	-6.90
1250	MUK63	1182.82	993.71	19.03
1252	MUK65	600.15	660.06	-9.07
1256	MUK69	114.41	163.09	-29.84
1258	MUK71	314.16	511.67	-38.60
1270	MUK83	464.38	519.14	-10.54
1271	MUK84	487.07	511.67	-4.80
1272	MUK85	433.11	519.14	-16.57
1274	MUK87	516.39	667.53	-22.64
1281	MUK94	278.79	244.75	13.90
1282	MUK95	459.53	422.77	8.69
1283	MUK96	581.23	526.60	10.37
1284	MUK97	623.67	615.50	1.32
1285	MUK98	700.75	726.79	-3.58
1286	MUK99	1004.37	1016.11	-1.15
1287	MUK100	1021.19	1016.11	0.49
1288	MUK101	1063.26	1016.11	4.63
1289	MUK102	598.23	615.50	-2.80
1291	MUK104	829.82	741.72	11.87

Continued on next page

Table A.12 – continued from previous page

No	Test no.	Simplified L_S=ct f_i=ct PG [Pa/m]	Measured PG [Pa/m]	Simplified L_S=ct f_i=ct PG Err [%]
1292	MUK105	686.66	741.72	-7.42
1298	MUK111	941.02	919.75	2.31
1337	MUK150	281.55	341.11	-17.46
1338	MUK151	488.63	459.87	6.25
1339	MUK152	536.82	489.50	9.66
1340	MUK153	227.45	252.21	-9.81
1343	MUK156	146.45	185.48	-21.04
1344	MUK157	66.02	88.89	-25.72
1345	MUK158	163.66	252.21	-35.11
1354	MUK167	98.33	133.46	-26.31

Table A.11: Comparison of the performance of the Optimised, Zhang et al. (2000) and (2003) models. PG = pressure gradient. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no.	Measured output PG [Pa/m]	Optimised prediction PG [Pa/m]	Zhang 2003 prediction PG [Pa/m]	Zhang 2000 prediction PG [Pa/m]	Optimised prediction PG Err. [%]	Zhang 2003 prediction PG Err. [%]	Zhang 2000 prediction PG Err. [%]
113	AND113	88.44	88.67	80.47	72.63	0.25	-9.01	-17.87
537	BEG2	92.74	92.84	69.08	60.71	0.10	-25.51	-34.53
539	BEG4	339.30	342.97	337.24	338.16	1.08	-0.60	-0.33
551	BEG16	1131.02	1140.14	1006.31	880.38	0.80	-11.02	-22.16
575	BEG40	678.61	685.77	716.94	719.39	1.05	5.64	6.00
581	BEG46	180.96	181.16	149.42	143.21	0.11	-17.42	-20.85
583	BEG48	6175.40	5977.64	6020.49		-6.44	-3.20	-2.50
589	BEG54	2488.25	2526.72	2022.40	2038.05	1.54	-18.72	-18.09
617	BEG82	2714.46	2975.82	2975.82	2985.96	9.62	9.62	10.00
650	BEG115	2035.84	2443.98	2443.98	2444.13	20.04	20.04	20.05
660	BEG125	769.09	699.80	691.40	691.67	-9.00	-10.10	-10.06
665	BEG130	180.96	250.78	183.09	141.28	38.58	1.17	-21.92
689	BEG154	3279.97	2785.10	2725.35	2737.16	-15.08	-16.90	-16.54
705	BEG170	723.85	807.34	807.34	759.81	11.53	11.53	4.96
706	BEG171	497.65	585.34	524.17	441.83	17.62	5.32	-11.21
714	BEG179	701.23	604.28	674.79	656.28	-13.82	-3.77	-6.41
1191	MUK4	949.38	1005.70	844.36	1134.14	5.93	-11.06	19.46
1203	MUK16	407.84	515.08	466.35	464.27	26.29	14.34	13.83
1211	MUK24	459.87	518.92	518.92	514.60	12.84	12.84	11.90
1271	MUK84	511.67	521.39	494.19	486.36	1.90	-3.41	-4.94
1283	MUK96	526.60	589.40	589.40	580.96	11.92	11.92	10.32
1287	MUK100	1016.11	1034.79	1022.26	1019.77	1.83	0.60	0.36
1292	MUK105	741.72	705.25	705.25	680.41	-4.91	-4.91	-8.26
1339	MUK152	489.50	456.81	584.24	455.32	-6.67	19.35	-6.98

Table A.13: Performance of the simplified model for the case L_S=ct vs measured pressure gradient values. PG = pressure gradient. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no.	Simplified L_S=ct PG [Pa/m]	Measured PG [Pa/m]	Simplified L_S=ct PG Err [%]
97	AND97	22.36	26.69	-16.22
113	AND113	71.15	88.44	-19.54
129	AND129	190.69	192.95	-1.17
343	AND343	40.11	51.57	-22.21
536	BEG1	59.29	92.74	-36.06
537	BEG2	59.47	92.74	-35.86
538	BEG3	329.77	339.30	-2.80
539	BEG4	333.18	339.30	-1.80
548	BEG13	30.51	54.28	-43.79
549	BEG14	30.56	54.28	-43.69
550	BEG15	997.38	1131.02	-11.81
551	BEG16	1005.54	1131.02	-11.09
574	BEG39	705.95	678.61	4.02
575	BEG40	712.84	678.61	5.04
580	BEG45	140.21	180.96	-22.51
581	BEG46	140.66	180.96	-22.26
582	BEG47	5483.26	5315.82	3.14
583	BEG48	5976.54	6175.40	-3.22
588	BEG53	1996.45	2488.25	-19.76
589	BEG54	2022.31	2488.25	-18.72
596	BEG61	1145.86	1040.54	10.12
617	BEG82	2903.57	2714.46	6.96
619	BEG84	6627.14	6673.05	-0.68
629	BEG94	623.37	570.03	9.35
650	BEG115	2367.42	2035.84	16.28
659	BEG124	311.42	339.30	-8.21
660	BEG125	691.88	769.09	-10.03
661	BEG126	2386.54	1945.36	22.67
662	BEG127	227.05	282.75	-19.70
665	BEG130	145.75	180.96	-19.45
666	BEG131	1699.27	1764.40	-3.69
673	BEG138	334.39	361.92	-7.60
678	BEG143	1128.98	1063.16	6.19
681	BEG146	395.35	429.78	-8.01
682	BEG147	4229.46	3393.07	24.64
684	BEG149	7533.45	7238.56	4.07
685	BEG150	313.39	316.68	-1.03
689	BEG154	2725.36	3279.97	-16.90
690	BEG155	871.28	972.68	-10.42
692	BEG157	569.11	622.06	-8.51
693	BEG158	1059.89	1198.88	-11.59

Continued on next page

Table A.13 – continued from previous page

No	Test no.	Simplified L_S=ct PG [Pa/m]	Measured PG [Pa/m]	Simplified L_S=ct PG Err [%]
694	BEG159	2889.92	2375.15	21.67
695	BEG160	356.79	497.65	-28.30
698	BEG163	264.76	339.30	-21.96
699	BEG164	2398.62	2465.63	-2.71
705	BEG170	757.33	723.85	4.62
706	BEG171	442.88	497.65	-11.00
711	BEG176	1167.81	1357.23	-13.95
714	BEG179	660.58	701.23	-5.79
715	BEG180	4775.00	4003.83	19.26
722	BEG187	3441.58	3393.07	1.42
723	BEG188	1519.61	1447.71	4.96
901	CRE4	20.95	26.69	-21.50
906	CRE9	76.28	99.07	-23.00
934	CRE37	64.99	15.60	316.40
1188	MUK1	4047.59	3923.30	3.16
1190	MUK3	1094.39	1016.11	7.71
1191	MUK4	1074.78	949.38	13.21
1202	MUK15	400.74	356.04	12.55
1203	MUK16	464.25	407.84	13.83
1209	MUK22	1110.44	1030.81	7.72
1210	MUK23	419.96	407.84	2.97
1211	MUK24	515.44	459.87	12.08
1213	MUK26	160.54	215.12	-25.36
1214	MUK27	499.55	437.48	14.18
1227	MUK40	819.13	674.99	21.35
1245	MUK58	4025.24	4323.91	-6.90
1250	MUK63	1182.82	993.71	19.03
1252	MUK65	600.15	660.06	-9.07
1256	MUK69	114.41	163.09	-29.84
1258	MUK71	314.16	511.67	-38.60
1270	MUK83	464.38	519.14	-10.54
1271	MUK84	487.07	511.67	-4.80
1272	MUK85	433.11	519.14	-16.57
1274	MUK87	516.39	667.53	-22.64
1281	MUK94	278.79	244.75	13.95
1282	MUK95	459.53	422.77	8.69
1283	MUK96	581.23	526.60	10.36
1284	MUK97	623.67	615.50	1.32
1285	MUK98	700.75	726.79	-3.58
1286	MUK99	1004.37	1016.11	-1.15
1287	MUK100	1021.19	1016.11	0.49
1288	MUK101	1063.26	1016.11	4.63
1289	MUK102	598.23	615.50	-2.80
1291	MUK104	829.82	741.72	11.87

Continued on next page

Table A.13 – continued from previous page

No	Test no.	Simplified L_S=ct PG [Pa/m]	Measured PG [Pa/m]	Simplified L_S=ct PG Err [%]
1292	MUK105	686.66	741.72	-7.42
1298	MUK111	941.02	919.75	2.31
1337	MUK150	281.55	341.11	-17.44
1338	MUK151	488.63	459.87	6.31
1339	MUK152	536.82	489.50	9.74
1340	MUK153	227.45	252.21	-9.73
1343	MUK156	146.45	185.48	-21.01
1344	MUK157	66.02	88.89	-25.68
1345	MUK158	163.66	252.21	-34.98
1354	MUK167	98.33	133.46	-26.22

Table A.14: Performance of the simplified model for the case f_i=ct vs measured pressure gradient values. PG = pressure gradient. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no.	Simplified f_i =ct PG [Pa/m]	Measured PG [Pa/m]	Simplified f_i =ct PG Err [%]
97	AND97	32.75	26.69	22.71
113	AND113	80.47	88.44	-9.01
129	AND129	196.42	192.95	1.79
343	AND343	64.78	51.57	25.60
536	BEG1	68.79	92.74	-25.82
537	BEG2	69.08	92.74	-25.51
538	BEG3	333.74	339.30	-1.63
539	BEG4	337.24	339.30	-0.60
548	BEG13	59.38	54.28	9.38
549	BEG14	59.37	54.28	9.37
550	BEG15	998.12	1131.02	-11.75
551	BEG16	1006.31	1131.02	-11.02
574	BEG39	709.96	678.61	4.61
575	BEG40	716.94	678.61	5.64
580	BEG45	148.74	180.96	-17.80
581	BEG46	149.42	180.96	-17.42
582	BEG47	5483.87	5315.82	3.16
583	BEG48	5977.64	6175.40	-3.20
588	BEG53	1996.46	2488.25	-19.76
589	BEG54	2022.40	2488.25	-18.72
596	BEG61	1205.91	1040.54	15.89
617	BEG82	2975.82	2714.46	9.62
619	BEG84	6617.60	6673.05	-0.83
629	BEG94	689.95	570.03	21.03
650	BEG115	2443.98	2035.84	20.04
659	BEG124	324.29	339.30	-4.42
660	BEG125	691.40	769.09	-10.10
661	BEG126	2391.87	1945.36	22.95
662	BEG127	269.09	282.75	-4.83
665	BEG130	183.09	180.96	1.17
666	BEG131	1699.97	1764.40	-3.65
673	BEG138	399.79	361.92	10.46
678	BEG143	1132.10	1063.16	6.48
681	BEG146	405.62	429.78	-5.62
682	BEG147	4233.17	3393.07	24.75
684	BEG149	7532.78	7238.56	4.06
685	BEG150	385.82	316.68	21.83
689	BEG154	2725.35	3279.97	-16.90
690	BEG155	870.42	972.68	-10.51
692	BEG157	587.71	622.06	-5.52
693	BEG158	1058.41	1198.88	-11.71

Continued on next page

Table A.14 – continued from previous page

No	Test no.	Simplified f_i =ct PG [Pa/m]	Measured PG [Pa/m]	Simplified f_i =ct PG Err [%]
694	BEG159	2895.03	2375.15	21.88
695	BEG160	409.57	497.65	-17.69
698	BEG163	313.89	339.30	-7.48
699	BEG164	2399.22	2465.63	-2.69
705	BEG170	807.345058	723.85	11.53
706	BEG171	524.170067	497.65	5.32
711	BEG176	1166.20	1357.23	-14.07
714	BEG179	674.79	701.23	-3.77
715	BEG180	4778.89	4003.83	19.35
722	BEG187	3441.57	3393.07	1.42
723	BEG188	1517.20	1447.71	4.80
901	CRE4	23.20	26.69	-13.05
906	CRE9	113.87	99.07	14.93
934	CRE37	15.30	15.60	-1.92
1188	MUK1	3992.32	3923.30	1.75
1190	MUK3	860.87	1016.11	-15.27
1191	MUK4	844.36	949.38	-11.06
1202	MUK15	401.91	356.04	12.88
1203	MUK16	466.35	407.84	14.34
1209	MUK22	869.17	1030.81	-15.68
1210	MUK23	422.14	407.84	3.50
1211	MUK24	518.92	459.87	12.84
1213	MUK26	166.61	215.12	-22.54
1214	MUK27	521.86	437.48	19.28
1227	MUK40	819.57	674.99	21.41
1245	MUK58	3854.73	4323.91	-10.85
1250	MUK63	1223.83	993.71	23.15
1252	MUK65	607.48	660.06	-7.96
1256	MUK69	164.54	163.09	0.88
1258	MUK71	426.23	511.67	-16.69
1270	MUK83	473.34	519.14	-8.82
1271	MUK84	494.19	511.67	-3.41
1272	MUK85	444.97	519.14	-14.28
1274	MUK87	570.75	667.53	-14.49
1281	MUK94	296.76	244.75	21.25
1282	MUK95	468.84	422.77	10.89
1283	MUK96	589.45	526.60	11.93
1284	MUK97	623.85	615.50	1.35
1285	MUK98	699.78	726.79	-3.71
1286	MUK99	1005.07	1016.11	-1.08
1287	MUK100	1022.26	1016.11	0.60
1288	MUK101	1065.56	1016.11	4.86
1289	MUK102	609.65	615.50	-0.95
1291	MUK104	857.06	741.72	15.54

Continued on next page

Table A.14 – continued from previous page

No	Test no.	Simplified f_i =ct PG [Pa/m]	Measured PG [Pa/m]	Simplified f_i =ct PG Err [%]
1292	MUK105	705.25	741.72	-4.91
1298	MUK111	941.71	919.75	2.38
1337	MUK150	297.33	341.11	-12.83
1338	MUK151	524.30	459.87	14.01
1339	MUK152	584.10	489.50	19.32
1340	MUK153	260.19	252.21	3.16
1343	MUK156	156.34	185.48	-15.70
1344	MUK157	73.19	88.89	-17.66
1345	MUK158	215.98	252.21	-14.36
1354	MUK167	119.15	133.46	-10.71

Table A.15: Prediction errors of the simplified models (L_S = ct and f_i = ct, L_S = ct, f_i = ct) vs Zhang et al. (2003) prediction errors. PG Err[%] = (PG-measured PG)/measured PG × 100.

No	Test no.	L_S =ct f_i =ct PG Err [%]	L_S =ct PG Err [%]	f_i =ct PG Err [%]	Zhang (2003) PG Err [%]
97	AND97	-16.22	-16.22	22.71	22.71
113	AND113	-19.54	-19.54	-9.01	-9.01
129	AND129	-1.17	-1.17	1.79	1.79
343	AND343	-22.21	-22.21	25.60	25.60
536	BEG1	-36.06	-36.06	-25.82	-25.82
537	BEG2	-35.86	-35.86	-25.51	-25.51
538	BEG3	-2.80	-2.80	-1.63	-1.63
539	BEG4	-1.80	-1.80	-0.60	-0.60
548	BEG13	-43.79	-43.79	9.38	9.38
549	BEG14	-43.69	-43.69	9.37	9.37
550	BEG15	-11.81	-11.81	-11.75	-11.75
551	BEG16	-11.09	-11.09	-11.02	-11.02
574	BEG39	4.02	4.02	4.61	4.61
575	BEG40	5.04	5.04	5.64	5.64
580	BEG45	-22.51	-22.51	-17.80	-17.80
581	BEG46	-22.26	-22.26	-17.42	-17.42
582	BEG47	3.14	3.14	3.16	3.16
583	BEG48	-3.22	-3.22	-3.20	-3.20
588	BEG53	-19.76	-19.76	-19.76	-19.76
589	BEG54	-18.72	-18.72	-18.72	-18.72
596	BEG61	10.12	10.12	15.89	15.89
617	BEG82	6.96	6.96	9.62	9.62
619	BEG84	-0.68	-0.68	-0.83	-0.83
629	BEG94	9.35	9.35	21.03	21.03
650	BEG115	16.28	16.28	20.04	20.04
659	BEG124	-8.21	-8.21	-4.42	-4.42
660	BEG125	-10.03	-10.03	-10.10	-10.10
661	BEG126	22.67	22.67	22.95	22.95
662	BEG127	-19.70	-19.70	-4.83	-4.83
665	BEG130	-19.45	-19.45	1.17	1.17
666	BEG131	-3.69	-3.69	-3.65	-3.65
673	BEG138	-7.60	-7.60	10.46	10.46
678	BEG143	6.19	6.19	6.48	6.48
681	BEG146	-8.01	-8.01	-5.62	-5.62
682	BEG147	24.64	24.64	24.75	24.75
684	BEG149	4.07	4.07	4.06	4.06
685	BEG150	-1.03	-1.03	21.83	21.83
689	BEG154	-16.90	-16.90	-16.90	-16.90
690	BEG155	-10.42	-10.42	-10.51	-10.51
692	BEG157	-8.51	-8.51	-5.52	-5.52
693	BEG158	-11.59	-11.59	-11.71	-11.71

Continued on next page

Table A.15 – continued from previous page

No	Test no.	L_S =ct f_i =ct PG Err [%]	L_S =ct PG Err [%]	f_i =ct PG Err [%]	Zhang (2003) PG Err [%]
694	BEG159	21.67	21.67	21.88	21.88
695	BEG160	-28.30	-28.30	-17.69	-17.69
698	BEG163	-21.96	-21.96	-7.48	-7.48
699	BEG164	-2.71	-2.71	-2.69	-2.69
705	BEG170	4.62	4.62	11.53	11.53
706	BEG171	-11.00	-11.00	5.32	5.32
711	BEG176	-13.95	-13.95	-14.07	-14.07
714	BEG179	-5.79	-5.79	-3.77	-3.77
715	BEG180	19.26	19.26	19.35	19.35
722	BEG187	1.42	1.42	1.42	1.42
723	BEG188	4.96	4.96	4.80	4.79
901	CRE4	-21.50	-21.50	-13.05	-13.05
906	CRE9	-23.00	-23.00	14.93	14.93
934	CRE37	316.41	316.40	-1.92	-1.94
1188	MUK1	3.16	3.16	1.75	1.75
1190	MUK3	7.70	7.71	-15.27	-15.27
1191	MUK4	13.20	13.21	-11.06	-11.06
1202	MUK15	12.55	12.55	12.88	12.88
1203	MUK16	13.83	13.83	14.34	14.34
1209	MUK22	7.72	7.72	-15.68	-15.68
1210	MUK23	2.97	2.97	3.50	3.50
1211	MUK24	12.08	12.08	12.84	12.84
1213	MUK26	-25.36	-25.36	-22.54	-22.54
1214	MUK27	14.18	14.18	19.28	19.28
1227	MUK40	21.35	21.35	21.41	21.41
1245	MUK58	-6.90	-6.90	-10.85	-10.85
1250	MUK63	19.03	19.03	23.15	23.15
1252	MUK65	-9.07	-9.07	-7.96	-7.96
1256	MUK69	-29.84	-29.84	0.88	0.88
1258	MUK71	-38.60	-38.60	-16.69	-16.69
1270	MUK83	-10.54	-10.54	-8.82	-8.82
1271	MUK84	-4.80	-4.80	-3.41	-3.41
1272	MUK85	-16.57	-16.57	-14.28	-14.28
1274	MUK87	-22.64	-22.64	-14.49	-14.49
1281	MUK94	13.90	13.95	21.25	21.28
1282	MUK95	8.69	8.69	10.89	10.89
1283	MUK96	10.37	10.36	11.93	11.92
1284	MUK97	1.32	1.32	1.35	1.35
1285	MUK98	-3.58	-3.58	-3.71	-3.71
1286	MUK99	-1.15	-1.15	-1.08	-1.08
1287	MUK100	0.49	0.49	0.60	0.60
1288	MUK101	4.63	4.63	4.86	4.86
1289	MUK102	-2.80	-2.80	-0.95	-0.95
1291	MUK104	11.87	11.87	15.54	15.54

Continued on next page

Table A.15 – continued from previous page

No	Test no.	L_S =ct f_i =ct PG Err [%]	L_S =ct PG Err [%]	f_i =ct PG Err [%]	Zhang (2003) PG Err [%]
1292	MUK105	-7.42	-7.42	-4.91	-4.91
1298	MUK111	2.31	2.31	2.38	2.38
1337	MUK150	-17.46	-17.44	-12.83	-12.82
1338	MUK151	6.25	6.31	14.01	14.03
1339	MUK152	9.66	9.74	19.32	19.35
1340	MUK153	-9.81	-9.73	3.16	3.20
1343	MUK156	-21.04	-21.01	-15.70	-15.69
1344	MUK157	-25.72	-25.68	-17.66	-17.64
1345	MUK158	-35.11	-34.98	-14.36	-14.29
1354	MUK167	-26.31	-26.22	-10.71	-10.65

Appendix B

Sensitivity analysis calculations

Conventional sensitivity analysis methods can be broadly classified into:

- Sensitivity testing - this involves studying the response of the model for a set of changes in the model formulation. This approach is often used to evaluate the robustness of a model by testing whether the response of the model changes significantly in relation to changes in model parameters and structural formulation of the model.
- Analytical methods - involve either the differentiation of the model equations, or the reformulation of the original model using stochastic algebraic equations.
- Sampling based methods - do not require the differentiation of the model equations. These methods involve running the code of the model on a set of sample points. Some of the widely used methods are: Monte Carlo and Latin Hypercube Sampling, Fourier Amplitude Sensitivity Test (FAST), reliability based methods and response surface methods.

The Morris Elementary Effects method (Morris 1991) is a sampling screening method. The input domain is randomly sampled through a number of trajectories. A trajectory is obtained by varying one vector component at a time by a small Δ value. This method was selected for this study because it is a generalisation of the conventional local derivative method and it offers a method to rank the effect of the input factors on the output of the model.

B.1 The Morris Elementary Effects method

B.1.1 Elementary Effect definition

Let us consider a model with k independent inputs $x_i, i = 1, 2, \ldots, k$. These inputs can take values across p selected levels in the k-dimensional unit cube ($0 \leq x_i \leq 1$). This way the input space is discretised into a p-level grid Ω. For a given value of an input vector $X = (x_1, x_2, \ldots, x_k)$, the Elementary Effect of the i-th input factor is defined as:

$$EE_i = \frac{Y\left(x_1, x_2, \ldots, x_{i-1}, x_i + \Delta, x_{i+1}, \ldots, x_k\right) - Y\left(x_1, x_2, \ldots, x_{i-1}, x_i, \ldots, x_k\right)}{\Delta}$$

(B.1)

where Y is the model output for vector X, p is the number of levels and Δ is an element of the set $\{\frac{1}{(p-1)}, \ldots, 1 - \frac{1}{(p-1)}\}$. p is chosen to be an even number and $\Delta = \frac{p}{2(p-1)}$.

$X \in \Omega$ and the transformation $(X + e_i\Delta)$ is such that the new vector is in Ω $((X + e_i\Delta) \in \Omega)$. The vector e_i is defined as a vector of zeroes, but with a unit as its i-th component: $e_i = (0, 0, \ldots, 1, \ldots, 0)$.

The distribution of Elementary Effects F_i associated with the i-th input factor is obtained by randomly sampling different input vectors X from the grid Ω. Morris proposed that the mean μ and the standard deviation σ of the distribution F_i be used as sensitivity measures. The mean μ determines the overall influence of the input factor on the output and the standard deviation σ determines the correlation level with other input factors.

Campolongo et al. (2007) (Saltelli et al. 2008) proposed the value μ^* be used instead of μ, where μ^* is the mean of the distribution G_i of the absolute values of the Elementary Effects, $|EE_i|$.

B.1.2 The sampling strategy

The Elementary Effects method builds r trajectories of $(k + 1)$ points in the input space, each trajectory providing k Elementary Effects values, one for each input factor, with a total of $r\,(k + 1)$ sampling points. The trajectories are generated starting from a base value X^*, which is not part of the trajectory, but is used to generate the trajectory points by increasing one or more of its k components by Δ. This way, the first trajectory point $X^{(1)}$ is obtained from X^* by adding Δ to one or more components of X^*, such that $X^{(1)} \in \Omega$. The second trajectory point $X^{(2)}$ is obtained from $X^{(1)}$ by adding $+\Delta$ or $-\Delta$ to its i-th component: $X^{(2)} = X^{(1)} + e_i\Delta$. The index i is randomly selected from the set $\{1, 2, \ldots, k\}$. The third sampling point is generated from $X^{(2)}$ by randomly selecting an index j such that $j \neq i$ and $X^{(3)} = X^{(2)} + e_i\Delta$. The procedure continues until $X^{(k+1)}$, which is the end point of the trajectory. This design generates a trajectory of $(k + 1)$ sampling points $\{X^{(1)}, X^{(2)}, \ldots, X^{(k+1)}\}$, which has the fundamental property that two consecutive points differ by Δ in one component only, and all the k indices have been used. The procedure is illustrated in Figure B.1. The following equation could generate the r trajectories with the above properties:

$$B^* = J_{k+1,k}X^* + \Delta B \tag{B.2}$$

where B is a $(k + 1) \times k$ strictly lower triangular matrix of 1s:

$$B = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{bmatrix} \tag{B.3}$$

and $J_{k+1,k}$ is a $(k + 1) \times k$ matrix of 1s, and X^* is the randomly chosen base, required for building the trajectory. Equation B.2 builds a matrix of trajectories,

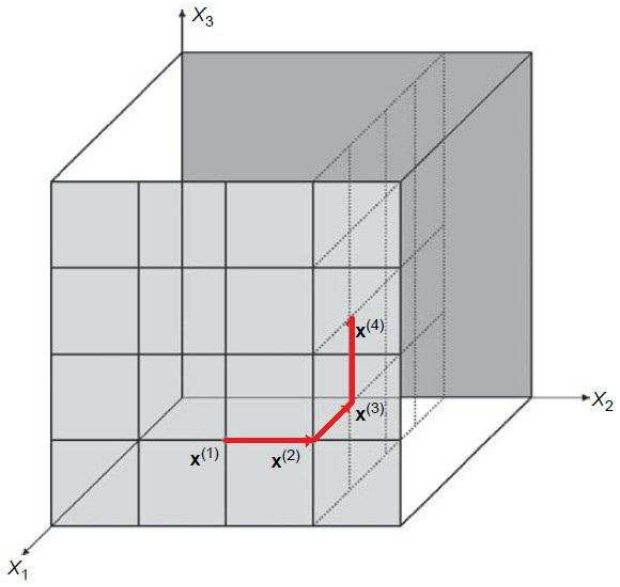


Figure B.1: An example of a trajectory in the input space for $k = 3$. The red line shows the transition between consecutive points (Saltelli et al. 2008).

whose rows are sampling points $X^{(1)}, X^{(2)}, \ldots, X^{(k+1)}$, but has the major drawback that the sampling points are not chosen randomly, given the form of matrix B.

A randomised version of the sampling matrix proposed by Morris is given in Equation B.4:

$$B^* = \{J_{k+1,k}X^* + \frac{\Delta}{2}\left[(2B - J_{k+1,k})\, D^* + J_{k+1,k}\right]\}P^* \tag{B.4}$$

where D^* is a k-dimensional diagonal matrix in which each diagonal element is either $+1$ or -1 with equal probability:

$$D^* = \begin{bmatrix} +1 & 0 & 0 & \cdots & 0 \\ 0 & +1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix} \tag{B.5}$$

and P^* is a $k \times k$ matrix in which each row contains one element equal to 1 while all others are 0, and no two columns have 1s in the same position. An example of such matrix is given in Equation B.6. This is called the random permutation matrix.

$$P^* = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & \cdots & 0 \end{bmatrix} \tag{B.6}$$

Equation B.4 calculates a matrix B^* which provides one Elementary Effect per input, which is randomly selected.

B.1.3 The computation of the sensitivity measure

Let us consider trajectory j in the input space and point $X^{(L)}$ on this trajectory. The Elementary Effect associated with point $X^{(L)}$ and its i-th component is:

$$EE_i^j\left(X^{(L)}\right) = \frac{Y\left(X^{(L+1)}\right) - Y\left(X^{(L)}\right)}{\Delta} \tag{B.7}$$

$$EE_i^j\left(X^{(L)}\right) = \frac{Y\left(x_{L_1}^{(j)}, x_{L_2}^{(j)}, \ldots, x_{L_i}^{(j)} + \Delta, \ldots, x_{L_k}^{(j)}\right) - Y\left(x_{L_1}^{(j)}, x_{L_2}^{(j)}, \ldots, x_{L_i}^{(j)}, \ldots, x_{L_k}^{(j)}\right)}{\Delta} \tag{B.8}$$

The mean μ, mean of absolute Elementary Effect values μ^* and standard deviation σ^2 are calculated with Equations B.9, B.10 and B.11:

$$\mu_i = \frac{1}{r}\sum_{j=1}^{r} EE_i^j \qquad i = 1, 2, \ldots, k \tag{B.9}$$

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r} \left|EE_i^j\right| \qquad i = 1, 2, \ldots, k \tag{B.10}$$

$$\sigma_i^2 = \frac{1}{r-1}\sum_{j=1}^{r} \left(EE_i^j - \mu_i\right)^2 \qquad i = 1, 2, \ldots, k \tag{B.11}$$

The mean μ in Equation B.9 assesses the global influence of the input factor x_i on the output. Campologno et al. (2007) showed that μ^* is a better measure of the overall effect of the input x_i on the output. If both μ and μ^* are high, the input factor has a strong effect on the output and that the sign of the effect is always the same. If μ is low while μ^* is high, the input factor has a strong effect, but with variable direction. The standard deviation σ^2 is a measure of input factor correlation, with high values indicating a high degree of correlation.

Appendix C

Multiphase flow terminology and definitions

C.1 Note

This appendix contains materials found in the literature. As the definition of the multiphase flow terminology is important for the presentation of the concepts described in Chapters 1-10, a list of multiphase flow terms is presented here.

C.2 Notations

Mass Flow Rate [kg/s]

W_L = liquid mass flow rate.
W_G = gas mass flow rate.
W = total mass flow rate.
$W = W_L + W_G$

Volumetric Flow Rate [m^3/s]

q_L = liquid volumetric flow rate.
q_G = gas volumetric flow rate.
q = total volumetric flow rate.
$q = q_L + q_G$

C.3 Definitions

C.3.1 Phase

For the purpose of this research, a phase is a physically and chemically homogenous sub-system, separated by definite boundaries, that allows mass transfer to other phases.

C.3.2 Component

Components are pure substances that make up the system. For this purpose, they will be considered as molecular species of fixed elemental composition and molecular weight.

Gas	Composition	Range
Methane	CH_4	70-90%
Ethane	C_2H_6	0-20%
Propane	C_3H_8	0-20%
Butane	C_4H_{10}	0-20%
Pentane and higher	C_5H_{12}	0-10%
Carbon dioxide	CO_2	0-8%
Oxygen	O_2	0-0.2%
Nitrogen	N_2	0-5%
Hydrogen sulfide, carbonyl sulfide	H_2S, COS	0-5%
Argon, Helium, Neon, Xenon	A, He, Ne, Xe	traceable

Figure C.1: Typical composition of natural gas (Bratland 2009).

Most of the time, the literature does not specifically mention *multiphase, multicomponent* when it refers to a mixture having different phases and it adopts the more concise term of *multiphase*. This is the convention adopted throughout this study too. The term *multiphase*will refer to a mixture of components with different phases. The typical example is a mixture of natural gas, oil and water. Natural gas in itself is a mixture of more components, as shown in Figure C.1. Some of the liquid droplets may be entrained in the gas flow, for example in an annular mist flow, when the liquid forms a thin film around the pipe circumference. According to the previous definition of the phase concept, droplets constitute a phase. Droplets are torn away from the liquid film through entrainment and also deposited back into the liquid film through deposition. This way, there is a continuous mass transfer between the two phases. Obviously, the droplet phase will have the same composition as the liquid phase it comes from.

C.3.3 Liquid holdup H_L and gas void fraction α

The liquid holdup H_L is the fraction of a volume element occupied by the liquid phase. As the liquid level in the pipe fluctuates, the notion of instantaneous liquid holdup is a more accurate measurement of the liquid fraction in the unit volume. The instantaneous liquid holdup $H_L\,(x,t)$ is the liquid holdup inside a differential volume at a spatial and temporal point (x,t). For practical reasons, the spatial and temporal average of the instantaneous holdup is a more useful definition:

$$\overline{H_L} = \frac{\iint H_L\,(x,t)\,dxdt}{\int dx \int dt} \tag{C.1}$$

For simplicity, the average liquid holdup for a designated volume will be denoted by H_L, knowing that $\overline{H_L}$ is actually meant.
The gas void fraction is the fraction of the volume element occupied by the gas phase. Hence, it can be calculated from the liquid holdup:

$$\alpha = 1 - H_L \tag{C.2}$$

Both the liquid holdup and the gas void fraction have values in the range $[0,1]$.

C.3.4 Superficial velocities, v_{SG} and v_{SL}

The superficial velocity of a phase is the velocity at which the phase would travel through the pipe if this were the only phase inside the pipe. This can be expressed mathematically as:

$$v_{SL} = \frac{q_L}{A_{pipe}} \quad v_{SG} = \frac{q_G}{A_{pipe}} \tag{C.3}$$

The superficial velocities are not actual velocities, because each phase only occupies a fraction of the pipe cross-section.

C.3.5 Mixture velocity v_M

The mixture velocity v_M is the total volumetric flow rate $q = q_L + q_G$ per unit area.

$$v_M = \frac{q_L + q_G}{A_{pipe}} = v_{SL} + v_{SG} \tag{C.4}$$

C.3.6 Actual velocities, v_G and v_L

The actual velocities, as opposed to the superficial velocities, are the real velocities of the gas and liquid phases, as measured by a measuring device. They can be expressed with respect to the superficial velocities:

$$v_L = \frac{q_L}{A_L} = \frac{q_L}{A_{pipe}} \times \frac{A_{pipe}}{A_L} = \frac{v_{SL}}{H_L} \tag{C.5}$$

Similarly:

$$v_G = \frac{v_{SG}}{\alpha} = \frac{v_{SG}}{1 - H_L} \tag{C.6}$$

C.3.7 Slip velocity v_{SLIP}

The difference between the actual phase velocities is called the slip velocity.

$$v_{SLIP} = v_G - v_L \tag{C.7}$$

C.3.8 Slippage and holdup

The multiphase flow equations are generally complex and some of the earlier flow models assumed that both the gas and liquid phase travelled at the same velocity, $v_G = v_L$. This assumption is called the *non-slip condition*. For this condition:

$$v_{SLIP} = v_G - v_L = 0 \tag{C.8}$$

$$v_G = \frac{v_{SG}}{\alpha} \tag{C.9}$$

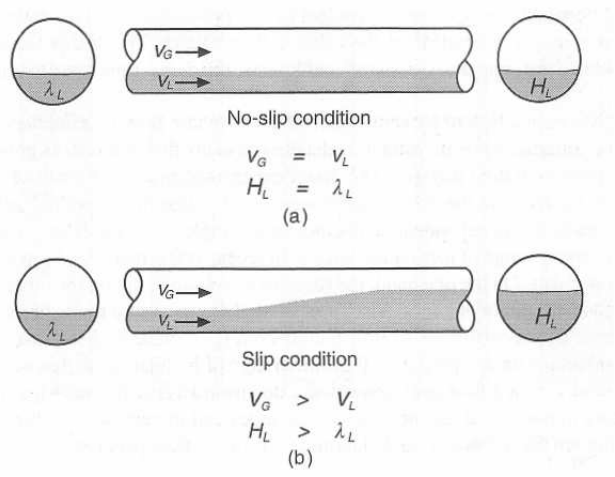


Figure C.2: The effect of slippage on liquid holdup (Shoham 2006).

$$\alpha = 1 - H_L \tag{C.10}$$

$$v_L = \frac{v_{SL}}{H_L} \tag{C.11}$$

$$0 = \frac{v_{SG}}{1 - H_L} - \frac{v_{SL}}{H_L} \tag{C.12}$$

$$H_L = \frac{v_{SL}}{v_{SL} + v_{SG}} \tag{C.13}$$

For the non-slip condition, the liquid level is the same throughout the pipe and at any pipe cross-section:

$$\lambda_L = \frac{A_G}{A_{pipe}} = \frac{\frac{q_L}{v_G}}{\frac{q_L}{v_{SL}}} = \frac{v_{SL}}{v_{SL} + v_{SG}} \tag{C.14}$$

Equations C.13 and C.14 show that the non-slip liquid holdup is equal to the average liquid holdup:

$$\lambda_L = H_{L-non-slip} \tag{C.15}$$

If $v_G > v_L$ the non-slip condition is not satisfied any longer and there is a slippage between the gas and the liquid phase. This is reflected by:

$$H_L > \lambda_L \tag{C.16}$$

Figure C.2 depicts graphically these considerations.

C.3.9 Drift velocity v_{drift}

The drift velocity is the velocity of a phase relative to the mixture velocity.

$$v_{L-drift} = v_L - v_M \tag{C.17}$$

$$v_{G-drift} = v_G - v_M \tag{C.18}$$

C.3.10 Quality x

Quality is a measure of the gas mass flow rate with respect to the total mass flow rate.

$$x = \frac{W_G}{W_G + W_L} = \frac{W_G}{W} \tag{C.19}$$

C.4 Average fluid properties

The average density and viscosity of the multiphase mixture are:

$$\rho_M = \rho_L H_L + \rho_G (1 - H_L) \tag{C.20}$$

$$\mu_M = \mu_L H_L + \mu_G (1 - H_L) \tag{C.21}$$

The liquid phase is a combination of oil and water in most cases. The liquid phase average density, viscosity and interfacial tension are:

$$\rho_L = \rho_{water} f_{water} + \rho_{oil} (1 - f_{water}) \tag{C.22}$$

$$\mu_L = \mu_{water} f_{water} + \mu_{oil} (1 - f_{water}) \tag{C.23}$$

$$\sigma_{LG} = \sigma_{water-G} f_{water} + \sigma_{oil-G} (1 - f_{water}) \tag{C.24}$$

where the water cut f_{water} is defined as the volumetric water flow rate with respect to the total liquid flow rate:

$$f_{water} = \frac{q_{water}}{q_{water} + q_{oil}} \tag{C.25}$$

In reality, because oil and water flows are extremely complex, correlations have been developed for the calculations of the average properties, instead of Equations C.22 and C.23.

Appendix D

Derivation of fundamental equations of two-phase fluid flow

D.1 Note

Most of the slug flow mechanistic models rely on the mass and momentum conservation laws. Although the derivation of the continuity and momentum conservation equations is widely found in the literature, the presentation of the derivation of these equations in this appendix is the foundation upon which Appendix E was built.

D.2 Notation

m = mass.
v = velocity.
x = distance measured in the direction of the fluid flow.
A = cross-sectional area.
F = external force applied to the control volume.
P = pressure.
α = phase holdup.
ρ = density.
θ = pipe inclination angle.

Subscript k applied to these elements means that the element refers to phase k.

D.3 Continuity equation

Figure D.1 shows the control volumes used in the derivations of the fundamental equations of two-phase flow, used as basis by the slug flow models described in Appendix E.

The mass conservation law for phase k can be expressed as:

$$\underbrace{\text{Net mass flow of phase } k \text{ into control volume from phase } k}_{\text{IN-OUT}} + \underbrace{\text{Net mass flow of phase } k \text{ from other sources or phases}}_{\text{IN-OUT}} = \text{Mass of phase } k \text{ accumulated} \quad \text{(D.1)}$$

The first term in Equation D.1 is:

$$\rho_k v_k A_k dt - \left(\rho_k + \frac{\partial \rho_k}{\partial x}dx\right)\left(v_k + \frac{\partial v_k}{\partial x}dx\right)\left(A_k + \frac{\partial A_k}{\partial x}dx\right)dt = \rho_k v_k A_k dt - \left(\rho_k v_k A_k + v_k A_k \frac{\partial \rho_k}{\partial x}dx + \rho_k A_k \frac{\partial v_k}{\partial x}dx + \rho_k v_k \frac{\partial A_k}{\partial x}dx\right) \quad \text{(D.2)}$$

Since $A_k = \alpha_k A$ and ignoring the negligible terms containing more the one partial derivative, the first term in Equation D.1 becomes:

$$\underbrace{\text{Net mass flow of phase } k \text{ into control volume from phase } k}_{\text{IN-OUT}} = -A\frac{\partial\left(\alpha_k \rho_k v_k\right)}{\partial x}dx \quad \text{(D.3)}$$

Substituting this term in the initial mass balance, Equation D.1 becomes:

$$-A\frac{\partial\left(\alpha_k \rho_k v_k\right)}{\partial x}dx + \sum_{p=1,p\neq k}^{n} \dot{m}_{p\to k} - \sum_{p=1,p\neq k}^{n} \dot{m}_{k\to p} = A\frac{\partial\left(\alpha_k \rho_k\right)}{\partial t}dx \quad \text{(D.4)}$$

$$\frac{\partial \alpha_k \rho_k}{\partial t} + \frac{\partial\left(\alpha_k \rho_k v_k\right)}{\partial x} = \frac{\sum_{p=1,p\neq k}^{n} \dot{m}_{p\to k} - \sum_{p=1,p\neq k}^{n} \dot{m}_{k\to p}}{Adx} \quad \text{(D.5)}$$

To simplify this form of the equation, we can define two mass transfer parameters:

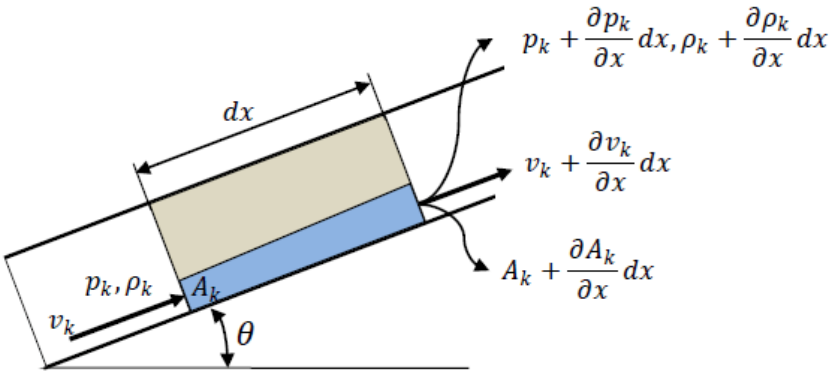


Figure D.1: Two phase (gas-liquid) flow control volume (Bratland 2009).

$$\Gamma_k^{IN} = \frac{\sum_{p=1,p\neq k}^{n} \dot{m}_{p\to k}}{Adx}$$
$$\Gamma_k^{OUT} = \frac{\sum_{p=1,p\neq k}^{n} \dot{m}_{k\to p}}{Adx} \tag{D.6}$$

Parameters Γ_k^{IN} and Γ_k^{OUT} defined in Equations D.6 represent the mass transfer rate per unit volume of pipe, from other phases to phase k, and from phase k to other phases, respectively. Γ is measured in [kg/m^3 s].

The mass balance can be re-written in terms of Γ parameters:

$$\frac{\partial\left(\alpha_k\rho_k\right)}{\partial t} + \frac{\partial\left(\alpha_k\rho_k v_k\right)}{\partial x} = \Gamma_k^{IN} - \Gamma_k^{OUT} \tag{D.7}$$

$$\sum_{k=1}^{n} \Gamma_k^{IN} - \sum_{k=1}^{n} \Gamma_k^{OUT} = 0 \tag{D.8}$$

Equation D.8 is zero because the mass gain of one phase is the mass loss of another phase.

$$\sum_{k=1}^{n} \alpha_k = 1 \tag{D.9}$$

Equation D.9 is called the *saturation constraint*. Equations D.7-D.9 represent the **mass conservation equations**.

For steady-state conditions, Equation D.7 becomes:

$$\frac{\partial\left(\alpha_k\rho_k v_k\right)}{\partial x} = \Gamma_k^{IN} - \Gamma_k^{OUT} \tag{D.10}$$

D.4 Momentum equation

The momentum equation for phase k is:

$$\rho_k\alpha_k A dx\frac{dv_k}{dt} = \sum_{i=1}^{m} F \tag{D.11}$$

A Taylor series is used to expand the term $\frac{dv_k}{dt}$ in the Equation D.11. The second order and higher derivative terms are neglected:

$$dv_k = \frac{\partial v_k}{\partial t}dt + \frac{\partial v_k}{\partial x}dx \tag{D.12}$$

Now the term $\frac{dv_k}{dt}$ can be expanded into:

$$\frac{dv_k}{dt} = \frac{\partial v_k}{\partial t} + \frac{\partial v_k}{\partial x}\frac{dx}{dt} \tag{D.13}$$

Equation D.11 becomes:

$$\alpha_k A \rho_k \left(\frac{\partial v_k}{\partial t} + v_k \frac{\partial v_k}{\partial x} \right) dx = \sum_{i=1}^{m} F \tag{D.14}$$

The term $\frac{\partial v_k}{\partial t}$ in Equation D.14 will be expanded:

$$\frac{\partial \left[(\alpha_k \rho_k) v_k \right]}{\partial t} = \alpha_k \rho_k \frac{\partial v_k}{\partial t} + v_k \frac{\partial (\alpha_k \rho_k)}{\partial t} \tag{D.15}$$

From this equation we calculate:

$$\alpha_k \rho_k \frac{\partial v_k}{\partial t} = \frac{\partial \left[(\alpha_k \rho_k) v_k \right]}{\partial t} - v_k \frac{\partial (\alpha_k \rho_k)}{\partial t} \tag{D.16}$$

This term can now be substituted in Equation D.14, which becomes:

$$\frac{\partial (\alpha_k \rho_k v_k)}{\partial t} - v_k \frac{\partial (\alpha_k \rho_k)}{\partial t} + \alpha_k \rho_k v_k \frac{\partial v_k}{\partial x} = \frac{\sum_{i=1}^{m} F}{A dx} \tag{D.17}$$

The term $\alpha_k \rho_k v_k \frac{\partial v_k}{\partial x}$ will be expanded and for this we calculate:

$$\frac{\partial (\alpha_k \rho_k {v_k}^2)}{\partial x} = \alpha_k \rho_k v_k \frac{\partial v_k}{\partial x} + v_k \frac{\partial (\alpha_k \rho_k v_k)}{\partial x} \tag{D.18}$$

The first term of the Left Hand Side (LHS) of this equation can be substituted into Equation D.17:

$$\frac{\partial (\alpha_k \rho_k v_k)}{\partial t} - v_k \frac{\partial (\alpha_k \rho_k)}{\partial t} + \frac{\partial (\alpha_k \rho_k {v_k}^2)}{\partial x} - v_k \frac{\partial (\alpha_k \rho_k v_k)}{\partial x} = \frac{\sum_{i=1}^{m}}{A dx} \tag{D.19}$$

$$\frac{\partial (\alpha_k \rho_k v_k)}{\partial t} + \frac{\partial (\alpha_k \rho_k {v_k}^2)}{\partial x} = v_k \underbrace{\left[\frac{\partial (\alpha_k \rho_k)}{\partial t} + \frac{\partial (\alpha_k \rho_k v_k)}{\partial x} \right]}_{\text{mass conservation}} + \frac{\sum_{i=1}^{m} F}{A dx} \tag{D.20}$$

$$\frac{\partial (\alpha_k \rho_k v_k)}{\partial t} + \frac{\partial (\alpha_k \rho_k {v_k}^2)}{\partial x} = v_k \left(\Gamma_k^{IN} - \Gamma_k^{OUT} \right) + \frac{\sum_{i=1}^{m} F}{A dx} \tag{D.21}$$

The term $v_k \left(\Gamma_k^{IN} - \Gamma_k^{OUT} \right)$ in Equation D.21 shows that fluids going from one phase to another take momentum with them.
Equation D.21 can be re-written as:

$$\begin{aligned}\frac{\sum_{i=1}^{m} F}{Adx} &= \left(\left.\frac{dP}{dx}\right|_{external-pressure}\right) + \left(\left.\frac{dP}{dx}\right|_{gravity}\right) + \\ &+ \left(\left.\frac{dP}{dx}\right|_{friction-from-other-phases}\right) + \\ &+ \left(\left.\frac{dP}{dx}\right|_{friction-from-wall}\right) + \\ &+ \left(\left.\frac{dP}{dx}\right|_{surface-tension-from-other-phases}\right) + \\ &+ \left(\left.\frac{dP}{dx}\right|_{surface-tension-from-wall}\right) \end{aligned} \tag{D.22}$$

The external pressure term in Equation D.22 can be expressed in terms of the pressures applied at the control volume ends:

$$\left(\left.\frac{dP}{dx}\right|_{external-pressure}\right) = \frac{p_k \alpha_k A + F_{k-internal} - A\left(\alpha_k + \frac{\partial \alpha_k}{\partial x} dx\right)\left(p_k + \frac{\partial p_k}{\partial x} dx\right)}{Adx} \tag{D.23}$$

$F_{k-internal}$ is the internal force on phase k, due to the pressure on the surface of phase k inside the control volume.
Assuming that both the pressure on phase k and α_k vary linearly:

$$F_{k-internal} = \left(p_k + \frac{1}{2}\frac{\partial p_k}{\partial x} dx\right)\left(A\frac{\partial \alpha_k}{\partial x} dx\right) \tag{D.24}$$

$$\begin{aligned} Adx \left.\frac{dP}{dx}\right|_{external-pressure} &= A\alpha_k p_k + p_k A\frac{\partial \alpha_k}{\partial x} dx + \frac{A}{2}\frac{\partial p_k}{\partial x}\frac{\partial \alpha_k}{\partial x}(dx)^2 - A\alpha_k p_k \\ &\quad - A\alpha_k \frac{\partial p_k}{\partial x} dx - A p_k \frac{\partial \alpha_k}{\partial x} dx - A\frac{\partial \alpha_k}{\partial x}\frac{\partial p_k}{\partial x}(dx)^2 \\ &= -A\alpha_k \frac{\partial p_k}{\partial x} dx + \frac{A}{2}\frac{\partial p_k}{\partial x}\frac{\partial \alpha_k}{\partial x}(dx)^2 \end{aligned} \tag{D.25}$$

If we ignore the higher order terms, the pressure gradient due to external forces becomes:

$$\left.\frac{dP}{dx}\right|_{external-pressure} = -\alpha_k \frac{\partial p_k}{\partial x} \tag{D.26}$$

The pressure gradient due to gravity is:

$$\left.\frac{dP}{dx}\right|_{gravity} = -\alpha_k \rho_k g \sin\theta \tag{D.27}$$

The pressure gradients due to external forces and gravity from Equations D.26 and D.27 can be substituted in Equation D.22:

$$\begin{aligned}\frac{\partial\left(\alpha_k\rho_k v_k\right)}{\partial t}+\frac{\partial\left(\alpha_k\rho_k {v_k}^2\right)}{\partial x} &= \underbrace{-\alpha_k\frac{\partial p_k}{\partial x}}_{\text{due to external forces}} \\ &+\left(\left.\frac{dP}{dx}\right|_{due-to-friction-from-other-phases}\right) \\ &+\left(\left.\frac{dP}{dx}\right|_{due-to-friction-from-wall}\right) \\ &+\left(\left.\frac{dP}{dx}\right|_{due-to-surface-tension-from-other-phases}\right) \\ &+\left(\left.\frac{dP}{dx}\right|_{due-to-surface-tension-from-wall}\right) \\ &-\underbrace{\alpha_k\rho_k g\sin\theta}_{\text{due to gravitation}} \\ &+\underbrace{v_k\Gamma_k^{IN}-v_k\Gamma_k^{OUT}}_{\text{due to mass transfer}} \end{aligned} \tag{D.28}$$

As any force acting on phase k must have an opposite counterforce, the sum of all forces between different phases must be zero. In this model there are three such forces:

1. Friction forces.
2. Surface tension forces.
3. Forces due to momentum exchange.

This observation can be expressed mathematically as:

$$\begin{aligned}&\sum_{k=1}^{n}\left(\left.\frac{dP}{dx}\right|_{due-to-friction-from-other-phases}\right) \\ &+\sum_{k=1}^{n}\left(\left.\frac{dP}{dx}\right|_{due-to-surface-tension-from-other-phases}\right) \\ &+\sum_{k=1}^{n}v_k\left(\Gamma_k^{IN}-\Gamma_k^{OUT}\right)=0\end{aligned} \tag{D.29}$$

Most of the models described in this study will use the following assumptions:

1. Steady-state conditions.
2. No mass transfer occurs between phases.
3. The effect of the surface tension forces due to other phases on momentum is ignored.

On these assumptions, Equation D.28 becomes:

$$\frac{\partial\left(\alpha_k \rho_k {v_k}^2\right)}{\partial x} = -\alpha_k \frac{\partial p_k}{\partial x} + \left.\frac{dP}{dx}\right|_{due-to-friction} - \alpha_k \rho_k g \sin\theta \tag{D.30}$$

$$\frac{\partial\left(\alpha_k \rho_k {v_k}^2\right)}{\partial x} = \frac{\partial\left[\left(v_k\right)\left(\alpha_k \rho_k v_k\right)\right]}{\partial x} = \alpha_k \rho_k v_k \frac{\partial v_k}{\partial x} + v_k \frac{\partial\left(\alpha_k \rho_k v_k\right)}{\partial x} = \alpha_k \rho_k v_k \frac{\partial v_k}{\partial x} \tag{D.31}$$

From Equation D.10, the assumption of no mass transfer between phases leads to:

$$\frac{\partial\left(\alpha_k \rho_k v_k\right)}{\partial x} = 0 \tag{D.32}$$

Equation D.31 becomes:

$$\alpha_k \rho_k v_k \frac{\partial v_k}{\partial x} = -\alpha_k \frac{\partial p_k}{\partial x} + \left.\frac{dP}{dx}\right|_{due-to-friction} - \alpha_k \rho_k g \sin\theta \tag{D.33}$$

$$-\alpha_k \frac{\partial p_k}{\partial x} + \left.\frac{dP}{dx}\right|_{due-to-friction} - \underbrace{\alpha_k \rho_k g \sin\theta}_{\text{due to elevation}} - \underbrace{\alpha_k \rho_k v_k \frac{\partial v_k}{\partial x}}_{\text{due to change in velocity}} = 0 \tag{D.34}$$

Equation D.34 is called the **mechanical energy balance** equation. The equation shows that the pressure drop in multiphase flow under steady-state conditions is due to:

1. Friction - the friction sources are either friction forces between the current phase and other phases, or between the current phase and the pipe wall.
2. Gravitation - the term called **due to elevation** is the term showing the gravitational effect on pressure drop. This term is normally referred to as **elevation pressure gradient**.
3. Change in velocity - the term called **due to change in velocity** is normally referred to as the **acceleration pressure gradient**.

It is important to reinforce the assumptions under which Equation D.34 is valid:

1. Steady-state conditions.
2. No mass transfer between the phases occurs.
3. Surface tension forces are ignored in so far as the momentum is concerned.

Appendix E

Slug flow modelling

E.1 Note

Although this appendix contains materials readily found in the literature, it was felt to be important to present a clear and coherent coverage of the foundation of the slug flow mechanistic models, which are at the heart of this research study.

E.2 Structure of a slug unit

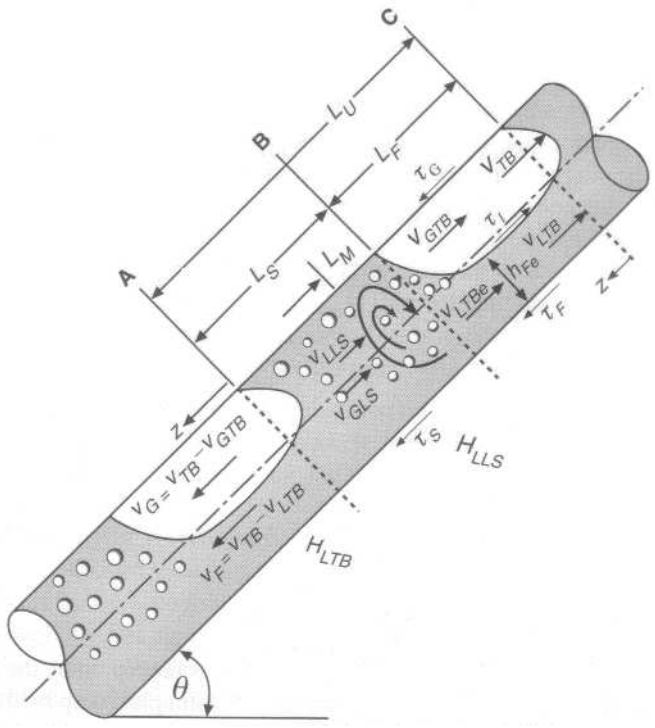


Figure E.1: Detailed description of a slug unit (Shoham 2006).

Figure E.1 shows the structure of a slug unit. The notation and definition of the parameters used in this chapter is given in Section 5.2.

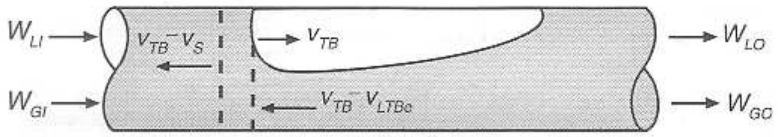


Figure E.2: Mass balance equation applied on a control volume inside the slug body (Shoham 2006).

E.3 Introduction

The slug flow pattern takes place in a wide range of pipe diameters and superficial gas and liquid velocities. The gas pockets called Taylor bubbles in the literature, characterize this flow pattern. Taylor bubbles occupy most of the pipe cross section. If two parallel planes sectioned the pipe across the Taylor bubble region, the flow inside the volume could be described as either stratified or annular, depending on the wetting degree of the pipe circumference. Fully developed slug flow is characterized by an alternation of gas pocket and liquid slugs, the period at which the slug units in the train alternate being quasi-equal. In vertical flow, the Taylor bubbles have a round nose and an almost flat stern. The rise velocity of the gas pockets does not depend on their length, as shown by Nicklin (1962).

E.4 Slug velocity vs. Taylor bubble velocity

As the slug body travels at an absolute velocity v_S, there are two simultaneous processes occurring at the ends of the slug body:

1. Scooping - because the liquid slug body moves at a higher velocity v_S than the liquid film in front of it, v_{LTB}, the liquid particles of the liquid film at the border with the slug body are scooped and engulfed by the slug body. A turbulent mixing zone is created inside the liquid slug as a result.

2. Shedding - after being accelerated to the slug body velocity and after traveling for a while with the slug body, the liquid particles are finally shed back into the liquid film of the following slug unit and decelerated back to the liquid film velocity v_{LTB}.

It is important to point out that the slug body does not travel at the same velocity as the Taylor bubble and $v_S \neq v_{TB}$.
The two processes can be described mathematically by applying the mass conservation equations.

E.4.1 The continuity equation on a control volume inside the slug body

Assuming a control volume bounded by the front of the slug body and a parallel plane inside the slug body, as shown in Figure E.2, a coordinate system moving at the velocity of the Taylor bubble is set. With respect to this coordinate system,

the interface of the slug body is stationary. The liquid in the slug body flows backward with a relative velocity of $v_{TB} - v_{LLS}$. The liquid in the film zone ahead of the slug body moves backward at a relative velocity of $v_{TB} - v_{LTBe}$. The liquid film particles are sped up from v_{LTB} to a higher velocity v_{LTBe} in the entry zone adjacent to the slug body.

$$\frac{dm_{scooped}}{dt} = \rho_L \times \underbrace{\underbrace{A_{pipe} H_{LTBe}}_{\text{liquid area in film in front of slug body}} \times (v_{TB} - v_{LTBe})}_{\text{scooping liquid volumetric flow rate}} \tag{E.1}$$

$$\frac{dm_{shed}}{dt} = \rho_L \times \underbrace{\underbrace{A_{pipe} H_{LLS}}_{\text{liquid area inside the slug body}} \times (v_{TB} - v_{LLS})}_{\text{shedding liquid volumetric flow rate}} \tag{E.2}$$

As the scooping flow rate is equal to the shedding flow rate, Equations E.1 and E.2 lead to:

$$H_{LTBe} (v_{TB} - v_{LTBe}) = H_{LLS} (v_{TB} - v_{LLS}) \tag{E.3}$$

Similarly, applying the mass balance on the gas phase yields to:

$$(1 - H_{LTBe}) (v_{TB} - v_{GTBe}) = (1 - H_{LLS}) (v_{TB} - v_{GLS}) \tag{E.4}$$

Some slug flow models make the assumption that there is no slip between the gas and liquid phases inside the slug body. In these cases:

$$v_S = v_{LLS} = v_{GLS} \tag{E.5}$$

Even though Figure E.2 shows this particular case, the mass balance Equations E.3 and E.4 can still be described by the same figure.

E.4.2 The continuity equation on a control volume in both the slug body and stratified region

Maintaining the same coordinate system moving at the translational velocity, a control volume bounded by a plane in the slug body (at point 1) and a plane in the stratified region (point 2) can be set, as depicted in Figure E.3. A mass balance can be carried out on this control volume:

$$(v_{TB} - v_{LLS}) (1 - H_{LLS}) = (v_{TB} - v_{GTB}) (1 - H_{LTB}) \tag{E.6}$$

E.4.3 The continuity equation on a control volume over the entire slug unit

A mass balance can be written for both liquid and gas phases, considering a control volume the size of an entire slug unit.
For the liquid phase, this yields:

$$\rho_L v_{SL} T_U A_{pipe} = \underbrace{\rho_L v_{LLS} T_S A_{pipe} H_{LLS}}_{\text{liquid mass average in the slug body region}} + \underbrace{\rho_L v_{LTB} T_F A_{pipe} H_{LTB}}_{\text{liquid mass average in the liquid film region}} \tag{E.7}$$

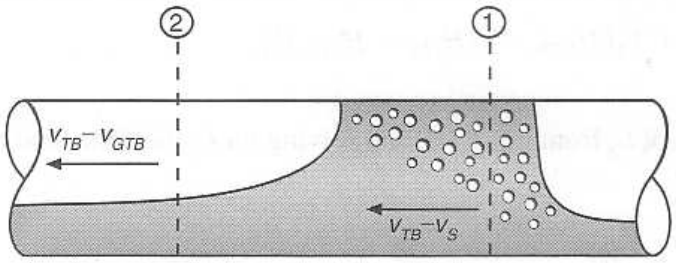


Figure E.3: Control volume between two cross-sections, one in the slug body, the other in the stratified region, for the mass conservation equation in the two regions (Shoham 2006).

Defining a parameter $\beta = \frac{L_F}{L_U} = \frac{v_{TB}T_F}{v_{TB}T_U} = \frac{T_F}{T_U}$, the previous equation becomes:

$$v_{SL} = v_{LLS}H_{LLS}\left(1-\beta\right) + v_{LTB}H_{LTB}\beta \tag{E.8}$$

Similarly, for the gas phase:

$$\begin{aligned}\rho_G v_{SG} T_U A_{pipe} \;=\; & \underbrace{\rho_G v_{GLS} T_S A_{pipe} H_{GLS}}_{\text{gas mass average in the slug body region}} \\ & + \underbrace{\rho_G v_{GTB} T_F A_{pipe} H_{GTB}}_{\text{gas mass average in the liquid film region}}\end{aligned} \tag{E.9}$$

$$v_{SG} = v_{GTB}\left(1-H_{LTB}\right)\frac{T_F}{T_U} + v_{GLS}\left(1-H_{LLS}\right)\frac{T_S}{T_U} \tag{E.10}$$

$$T_U = T_S + T_F \tag{E.11}$$

$$v_{SG} = v_{GLS}\left(1-H_{LLS}\right)\left(1-\beta\right) + v_{GTB}\left(1-H_{LTB}\right)\beta \tag{E.12}$$

E.4.4 Mass of liquid inside the slug unit

There are two ways to measure the amount of liquid inside a slug unit:

1. Assuming a pipe cross-section, the mass of liquid which flows through the cross-section in an amount of time equal to the slug unit period would be the mass of liquid in a slug unit.

2. The second method assumes "freezing" the slug unit, by assuming a coordinate system traveling at the same velocity at the slug unit and measuring the inputs, outputs and the mass accumulated.

Determination of the mass of liquid in a slug unit by using the cross-section method

The mass of liquid in a slug unit can be determined by calculating the amount of liquid passing through a cross-section of the pipe during a time interval equal to the slug unit period, T_U:

$$W_L = \frac{\underbrace{\rho_L A_{pipe} H_{LLS} v_S T_S}_{\text{mass of liquid in slug body zone}} + \underbrace{\int_0^{T_F} \rho_L A_{pipe} H_{LTB}(t)\, v_{LTB} dt}_{\text{mass of liquid in film zone}}}{T_U} \tag{E.13}$$

Equation E.13 can be transformed from the time to the space domain, by using the conversion terms:

$$T_S = \frac{L_S}{v_{TB}} \quad T_F = \frac{L_F}{v_{TB}} \quad dt = \frac{dL}{v_{TB}} \tag{E.14}$$

$$W_L = \rho_L A_{pipe} H_{LLS} \frac{v_{LLS}}{v_{TB}} \frac{L_S}{T_U} + \int_0^{L_F} \rho_L A_{pipe} H_{LTB} \frac{v_{LTB}}{T_U v_{TB}} dL \tag{E.15}$$

$$L_U = v_{TB} T_U \tag{E.16}$$

$$W_L = \frac{\rho_L A_{pipe} H_{LLS} v_{LLS} L_S + \int_0^{L_F} \rho_L A_{pipe} H_{LTB} v_{LTB} dL}{L_U} \tag{E.17}$$

Determination of the mass of liquid in a slug unit by using the freezing method

Assuming a coordinate system moving with the same translational velocity as the slug, the following mass balance can be written:

$$W_L = \frac{\underbrace{\rho_L A_{pipe} H_{LLS} L_S}_{\text{mass of liquid in the slug body zone}} + \underbrace{\int_0^{L_F} \rho_L A_{pipe} H_{LTB} dL}_{\text{mass of liquid in the film zone}}}{T_U} - x \tag{E.18}$$

where, x is the mass rate for the scooping or shedding process:

$$x = \rho_L A_{pipe} H_{LLS} (v_{TB} - v_{LLS}) = \rho_L A_{pipe} H_{LTB} (v_{TB} - v_{LTB}) \tag{E.19}$$

The liquid superficial velocity, v_{SL} is:

$$\begin{aligned}
v_{SL} &= \frac{q_L}{A_{pipe}} \\
&= \frac{W_L}{\rho_L A_{pipe}} = \frac{H_{LLS}L_S + \int_0^{L_F} H_{LTB}dL}{\frac{L_U}{v_{TB}}} - x \\
&= v_{TB}H_{LLS}\frac{L_S}{L_U} + \frac{v_{TB}}{L_U}\int_0^{L_F} H_{LTB}dL - \left(v_{TB} - v_{LLS}\right)H_{LLS} \\
&= v_{TB}H_{LLS}\left(\frac{L_S}{L_U} - 1\right) + v_{LLS}H_{LLS} + \frac{v_{TB}}{L_U}\int_0^{L_F} H_{LTB}dL \\
&= v_{LLS}H_{LLS} - v_{TB}H_{LLS}\frac{L_F}{L_U} + \frac{v_{TB}}{L_U}\int_0^{L_F} H_{LTB}dL \\
&= \underbrace{v_{LLS}H_{LLS} + v_{TB}\frac{L_F}{L_U} - v_{TB}H_{LLS}\frac{L_F}{L_U} + \frac{v_{TB}}{L_U}\int_0^{L_F} H_{LTB}dL - v_{TB}\frac{L_F}{L_U}}_{v_{TB}\ L_F/L_U \text{ was added and subtracted}} \\
&= v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right)\frac{L_F}{L_U} + \frac{v_{TB}}{L_U}\int_0^{L_F} H_{LTB}dL - \frac{v_{TB}}{L_U}\int_0^{L_F} dL \\
&= v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right)\frac{L_F}{L_U} - \frac{v_{TB}}{L_U}\int_0^{L_F}\left(1 - H_{LTB}\right)dL \qquad \text{(E.20)}
\end{aligned}$$

This gives the expression of the liquid superficial velocity:

$$v_{SL} = v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right)\frac{L_F}{L_U} - \frac{v_{TB}}{L_U}\int_0^{L_F}\left(1 - H_{LTB}\right)dL \qquad \text{(E.21)}$$

The average gas void fraction in the slug unit, H_{GSU} is defined as:

$$\begin{aligned}
H_{GSU} = & \underbrace{H_{GLS}\frac{L_S}{L_U}}_{\text{average gas void fraction in the liquid slug region}} \\
& + \underbrace{\frac{\int_0^{L_F} H_{GTB}dL}{L_U}}_{\text{average gas void fraction in the liquid film region}} \qquad \text{(E.22)}
\end{aligned}$$

This takes into account the gas void fraction values in both the liquid slug and liquid film regions.
The average slug unit liquid holdup is the opposite of the average slug unit gas void fraction:

$$H_{LSU} = 1 - H_{GSU} = 1 - \frac{\left(1 - H_{LLS}\right)L_S + \int_0^{L_F}\left(1 - H_{LTB}\right)dL}{L_U} \qquad \text{(E.23)}$$

Equation E.23 can be combined with Equation E.21:

$$\begin{aligned} v_{SL} &= v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right)\left(1 - \frac{L_S}{L_U}\right) - \frac{v_{TB}}{L_U}\int_0^{L_F}\left(1 - H_{LTB}\right)dL \\ &= v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right) - v_{TB}\left(1 - H_{LLS}\right)\frac{L_S}{L_U} \\ &\quad - \frac{v_{TB}}{L_U}\int_0^{L_F}\left(1 - H_{LTB}\right)dL \\ &= v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right) \\ &\quad - v_{TB}\left\{\frac{\left(1 - H_{LLS}\right)L_S + \int_0^{L_F}\left(1 - H_{LTB}\right)dL}{L_U}\right\} \\ &= v_{LLS}H_{LLS} + v_{TB}\left(1 - H_{LLS}\right) - v_{TB}\left(1 - H_{LSU}\right) \\ &= v_{LLS}H_{LLS} + v_{TB} - v_{TB}H_{LLS} - v_{TB} + v_{TB}H_{LSU} \end{aligned} \tag{E.24}$$

This equation allows the calculation of H_{LSU}:

$$H_{LSU} = \frac{v_{TB}H_{LLS} - v_{LLS}H_{LLS} + v_{SL}}{v_{TB}} \tag{E.25}$$

Equation E.25 allows the calculation of H_{LSU} in terms of v_{SG}:

$$v_M = v_{SL} + v_{SG} = v_{LLS}H_{LLS} + v_{GLS}\left(1 - H_{LLS}\right) \tag{E.26}$$

This equation holds because the mass balance equation for both liquid and gas phases results in a constant volumetric flow rate through any cross section of the slug unit, in this case the slug body.

$$v_{SL} = v_{GLS}\left(1 - H_{LLS}\right) + v_{LLS}H_{LLS} - v_{SG} \tag{E.27}$$

This result can be used in Equation E.25:

$$H_{LSU} = \frac{v_{TB}H_{LLS} + v_{GLS}\left(1 - H_{LLS}\right) - v_{SG}}{v_{TB}} \tag{E.28}$$

Slug unit average holdup concluding remarks

Equations E.25 and E.28 show an interesting result. The average liquid holdup in a slug unit H_{LSU} does not depend on the slug structure, namely it does not depend on the length of the slug body zone L_S, or the liquid film zone, L_F. It only depends on:

- Liquid and gas superficial flow rates, v_{SL} and v_{SG}.
- Liquid and gas particle/bubble velocities in the slug body, v_{LLS} and v_{GLS}.
- Taylor bubble translational velocity, v_{TB}.
- Liquid holdup within the slug body, H_{LLS}.

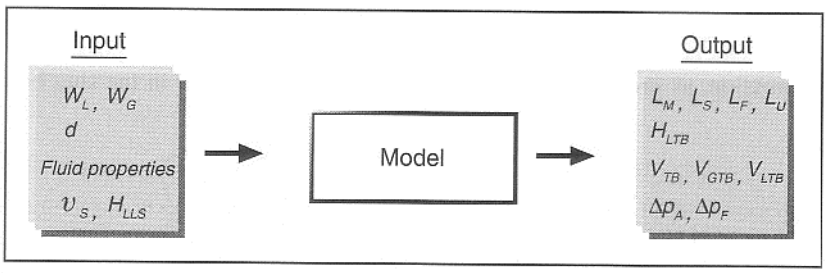


Figure E.4: Dukler and Hubbard model (1975) (Shoham 2006).

E.5 Dukler and Hubbard (1976)

E.5.1 Assumptions

The following assumptions were made in the Dukler and Hubbard model:

1. The model considers fully developed, steady-state, two-dimensional flow. A consequence of the steady-state assumptions is that liquid does not accumulate inside the slug and, as explained further in the next section, the volumetric scooping rate at the front of the slug body is equal to the volumetric shedding rate at the end of it. The immediate consequence of this is that the slug train consists of slug units of same length L_U.

2. There is no slip between gas and liquid in the liquid slug body. This assumption translates into the condition:

$$v_{LLS} = v_{GLS} = v_S \tag{E.29}$$

3. Both the gas and liquid phases are incompressible.

4. The flow is horizontal.

The model is not *complete*, because two additional variables need to be specified as input parameters: $freq_S$, the slug frequency, and H_{LLS}, the liquid holdup in the slug body.

E.5.2 Mass conservation

From Equation E.2, the translational velocity of the Taylor bubble can be expressed as:

$$v_{TB} = v_S + \underbrace{\frac{x}{\rho_L A_{pipe} H_{LLS}}}_{\text{additional velocity gained by the scooping process}} \tag{E.30}$$

Defining $c = \frac{x}{\rho_L A_{pipe} H_{LLS}}$, Equation E.30 can be re-written as:

$$v_{TB} = v_S + cv_S = v_S\,(1 + c) = c_0 v_S \tag{E.31}$$

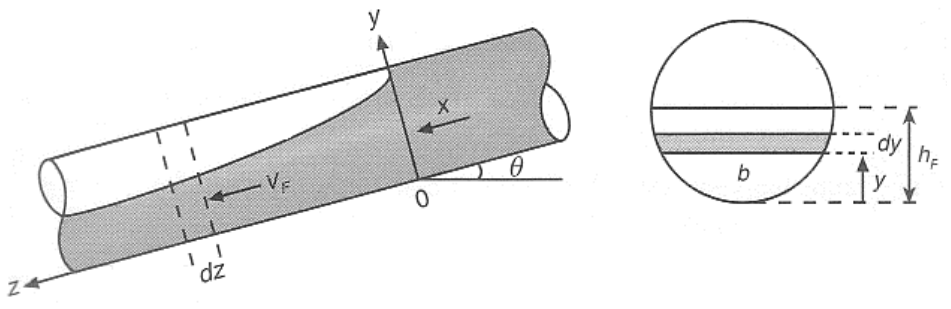


Figure E.5: Control volume for the study of the liquid film hydrodynamics (Shoham 2006).

where $c_0 = 1 + c$. c is a function of Reynolds number and the following expression holds:

$$c = 0.021 \ln Re_S + 0.022 \tag{E.32}$$

where:

$$Re_S = \frac{\rho_L v_M d_{pipe}}{\mu_L} \tag{E.33}$$

$$v_M = \frac{q_L + q_G}{A_{pipe}} = v_S \tag{E.34}$$

Equation E.32 was obtained based on considerations related to the shedding process, for a wide range of Reynolds numbers.

The values of c were also confirmed by experiments (Fabre 1994). It was found that for turbulent flow $c = 0.2$ and for laminar flow $c = 1$. Consequently, c_0 has the values:

$$c_0 = \begin{cases} 1.2 & \text{for turbulent flow} \\ 2 & \text{for laminar flow} \end{cases} \tag{E.35}$$

The meaning of the c_0 coefficient is that of a flow distribution coefficient. It shows the contribution of the mixture velocity to the translational velocity.

E.5.3 Hydrodynamics of the liquid film

This section aims to determine:

1. The velocity profile $v_F(z)$ relative to the coordinate system moving with the slug front velocity of v_{TB}.
2. The liquid holdup distribution along the liquid film $H_{LTB}(z)$, or alternatively, $h_F(z)$, the liquid film height at location z. As shown in Figure E.5, the axis z has its origin at the end of a slug body and the positive values are opposite to the direction of flow.

The velocity of the liquid film in this coordinate system is:

$$v_F = v_{TB} - v_{LTB} = \frac{x}{\rho_L A_{pipe} H_{LTB}} \tag{E.36}$$

Momentum Balance

A momentum balance can be applied to the control volume of width dz in Figure E.5:

$$\frac{\frac{d\left(m_{liquid}v_{liquid}\right)}{dt}}{dz} = \frac{\sum F}{dz} \tag{E.37}$$

where:

$$\begin{aligned}\sum F &= \text{Wall Friction} - \text{Taylor Bubble Pushing Force} \\ &\quad + \text{Gravitational Component along } z \text{ axis}\end{aligned} \tag{E.38}$$

Equations E.37 and E.38 can be further expanded into:

$$\begin{aligned}\frac{d}{dz}\left(\underbrace{\underbrace{\underbrace{\rho_L \quad \underbrace{A_{pipe}H_{LTB}}_{\text{liquid flow area}} v_F}_{\text{liquid volumetric flow rate}} \quad v_F}_{\text{liquid mass flow rate}}}_{\text{momentum rate}}\right) &= \tau_F S_F - \frac{d}{dz}\left(\overline{P}A_{pipe}H_{LTB}\right) \\ &\quad + \underbrace{\left(\underbrace{\underbrace{\rho_L \; \underbrace{A_{pipe}H_{LTB}}_{\text{liquid flow area}} \, dz \, g \sin\theta}_{\text{volume of liquid}}}_{\text{mass of liquid}}\right)}_{\text{gravitational component along z axis}} \frac{1}{dz}\end{aligned} \tag{E.39}$$

The Taylor bubble pushing force was expressed in terms of $\overline{P}$, the average hydrostatic gas pressure acting on the control volume.

$$\begin{aligned}\frac{d}{dz}\left(\overline{P}A_{pipe}H_{LTB}\right) &= \underbrace{\left(\rho_L \quad g \quad dh_F \quad \cos\theta\right)}_{\text{hydrostatic weight of the liquid column of height } dh_F \text{ and length } dz \text{ on } z \text{ axis}} \\ &\quad \times \frac{A_{pipe}H_{LTB}}{dz}\end{aligned} \tag{E.40}$$

Substituting the value of v_F from E.36 into Equation E.39 yields:

$$\begin{aligned}\frac{d}{dz}\left(\rho_L A_{pipe}H_{LTB}v_F^2\right) &= \frac{d}{dz}\left[\left(\rho_L A_{pipe}H_{LTB}\right)\left(\frac{x^2}{\rho_L A_{pipe}H_{LTB}}\right)^2\right] \\ &= \frac{d}{dz}\left(\frac{x^2}{\rho_L A_{pipe}H_{LTB}\left(h_F\right)}\right) \\ &= \frac{x^2}{\rho_L A_{pipe}} - \frac{\frac{dH_{LTB}}{dh_F}}{H_{LTB}^2}\frac{dh_F}{dz}\end{aligned} \tag{E.41}$$

$$-\frac{x^2}{\rho_L A_{pipe}} \frac{\frac{H_{LTB}}{dh_F}}{H_{LTB}^2} \frac{dh_F}{dz} = \tau_F S_F - \rho_L A_{pipe} H_{LTB} g \cos\theta \frac{dh_F}{dz} + \rho_L A_{pipe} H_{LTB} g \sin\theta \tag{E.42}$$

$$\frac{dh_F}{dz} = -\frac{\tau_F S_F + \rho_L A_{pipe} H_{LTB} g \sin\theta}{\frac{x^2}{\rho_L A_{pipe} H_{LTB}^2} \frac{dH_{LTB}}{dh_F} - \rho_L A_{pipe} H_{LTB} g \cos\theta} \tag{E.43}$$

The boundary conditions for E.43 are:

$$h_F\left(z=0\right) = H_{LLS} d \tag{E.44}$$

corresponding to:

$$v_F\left(z=0\right) = v_{TB} - v_S \tag{E.45}$$

The following conclusions can be drawn from E.43:

1. The equilibrium level $h_{F-equilibrium}$ is reached when $\frac{dh_F}{dz} = 0$. This yields to:

   $$\tau_F S_F + \rho_L A_{pipe} H_{LTB} g \sin\theta = 0 \tag{E.46}$$

   where τ_F is the shear stress caused by the friction between the liquid film and the pipe wall. This must be expressed using the actual velocity of the liquid film v_{LTB}, not the relative velocity $v_F = v_{TB} - v_{LTB}$, because force is invariant in non-accelerating coordinate systems.

   $$\tau_F = f_F \frac{\rho_L v_{LTB}^2}{2} \tag{E.47}$$

2. The critical level $h_{F-critical}$ is the value of h_F for which the denominator is zero:

   $$\frac{x^2}{\rho_L A_{pipe} H_{LTB}^2} \frac{dH_{LTB}}{dh_F} - \rho_L A_{pipe} H_{LTB} g \cos\theta = 0 \tag{E.48}$$

   If the boundary condition of Equation E.43, $h_F\left(z=0\right) > h_{F-critical}$, which is the solution of E.48, $h_F\left(z=0\right)$ needs to be adjusted and made equal to $h_{F-critical}$.
   In other words, the boundary condition of Equation E.43 becomes:

   $$h_F\left(z=0\right) \begin{cases} H_{LLS} d & \text{if } H_{LLS} d \leq h_{F-critical} \\ h_{F-critical} & \text{if } H_{LLS} d > h_{F-critical} \end{cases} \tag{E.49}$$

E.5.4 Slug length

The same Equation E.13 that provided the value of the mass of liquid in a slug unit could be used to determine the liquid slug length. The method measured the mass of liquid passing through a cross-section of the pipe during an interval equal to the slug unit period T_U:

$$W_L = \frac{\underbrace{\rho_L A_{pipe} H_{LLS} v_S T_S}_{\text{mass of liquid in slug body zone}} + \underbrace{\int_0^{T_F} \rho_L A_{pipe} H_{LTB}(t)\, v_{LTB} dt}_{\text{mass of liquid in film zone}}}{T_U} \tag{E.50}$$

The term H_{LTB} is not constant in the interval T_F. Therefore, the purpose of the following calculations is to determine the product $H_{LTB} v_{LTB}$.
As previously shown in Equation E.3, the following equation holds:

$$H_{LTB}(v_{TB} - v_{LTB}) = H_{LLS}(v_{TB} - v_S) \tag{E.51}$$

Also, Equation E.31 shows that:

$$v_{TB} = v_S + c v_S = v_S(1 + c) \tag{E.52}$$

$$v_{TB} - v_S = c v_S \tag{E.53}$$

Substituting this result in E.51 yields to:

$$H_{LTB}(v_{TB} - v_{LTB}) = H_{LLS} c v_S \tag{E.54}$$

$$v_{TB} H_{LTB} - c v_S H_{LLS} = v_{LTB} H_{LTB} \tag{E.55}$$

$$(1 + c) v_S H_{LTB} - c v_S H_{LLS} = v_{LTB} H_{LTB} \tag{E.56}$$

$$v_S [H_{LTB} - c(H_{LLS} - H_{LTB})] = v_{LTB} H_{LTB} \tag{E.57}$$

This result can now be introduced in Equation E.50:

$$\begin{aligned} W_L &= \{\rho_L A_{pipe} H_{LLS} v_S \frac{L_S}{v_{TB}} \\ &\quad + \int_0^{L_F} \rho_L A_{pipe} v_S [H_{LTB} - c(H_{LLS} - H_{LTB})] \frac{dL}{v_{TB}} \} freq_U \end{aligned} \tag{E.58}$$

The expression was converted from the time to the space domain, using:

$$T_S = \frac{L_S}{v_{TB}} \quad T_F = \frac{L_F}{v_{TB}} \quad dt = \frac{dL}{v_{TB}} \tag{E.59}$$

$$\frac{W_L}{\rho_L A_{pipe} v_S} = \frac{freq_U}{v_{TB}} \left\{ H_{LLS} L_S + \int_0^{L_F} \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right] dL \right\} \quad \text{(E.60)}$$

Assuming that the liquid film is at equilibrium in the stratified region of the slug unit:

$$H_{LTB} = H_{LTBe} = constant \quad \text{(E.61)}$$

$$\frac{W_L}{\rho_L A_{pipe} v_S} = \frac{freq_U}{v_{TB}} \left\{ H_{LLS} L_S + L_F \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right] \right\} \quad \text{(E.62)}$$

$$L_F = L_U - L_S = \frac{v_{TB}}{freq_U} - L_S \quad \text{(E.63)}$$

This result can be introduced in Equation E.62:

$$\begin{aligned} \frac{W_L}{\rho_L A_{pipe} v_S} &= \frac{freq_U}{v_{TB}} H_{LLS} L_S \\ &\quad + \frac{freq_U}{v_{TB}} \left(\frac{v_{TB}}{freq_U} - L_S \right) \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right] \\ &= \frac{freq_U}{v_{TB}} H_{LLS} L_S + \left(1 - L_S \frac{freq_U}{v_{TB}} \right) \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right] \\ &= \frac{freq_U}{v_{TB}} H_{LLS} L_S + \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right] \\ &\quad - L_S \frac{freq_U}{v_{TB}} \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right] \end{aligned} \quad \text{(E.64)}$$

This allows us to solve for L_S:

$$\begin{aligned} L_S &= \frac{\frac{W_L}{\rho_L A_{pipe} v_S} - \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right]}{\frac{freq_U}{v_{TB}} \left(H_{LLS} - H_{LTB} + c H_{LLS} - c H_{LTB} \right)} \\ &= \frac{\frac{W_L}{\rho_L A_{pipe} v_S} - \left[H_{LTB} - c \left(H_{LLS} - H_{LTB} \right) \right]}{\frac{freq_U}{(1+c) v_S} \left(1 + c \right) \left(H_{LLS} - H_{LTB} \right)} \end{aligned} \quad \text{(E.65)}$$

This gives the solution L_S:

$$L_S = \frac{v_S}{freq_U \left(H_{LLS} - H_{LTB} \right)} \left[\frac{W_L}{\rho_L A_{pipe} v_S} - H_{LTB} + c \left(H_{LLS} - H_{LTB} \right) \right] \quad \text{(E.66)}$$

E.5.5 Velocity of the gas pocket

As shown in Equation E.6:

$$\left(v_{TB} - v_S \right) \left(1 - H_{LLS} \right) = \left(v_{TB} - v_{GTB} \right) \left(1 - H_{LTB} \right) \quad \text{(E.67)}$$

$$v_{GTB} = v_{TB} - \left(v_{TB} - v_S \right) \left(\frac{1 - H_{LLS}}{1 - H_{LTB}} \right) \quad \text{(E.68)}$$

As $v_{TB} = \left(1 + c \right) v_S$, the previous equation can be written as:

$$v_{GTB} = v_{TB} - c v_S \left(\frac{1 - H_{LLS}}{1 - H_{LTB}} \right) \quad \text{(E.69)}$$

E.5.6 Length of the mixing zone

According to Dukler and Hubbard (1976), based on a correlation for the velocity head, the length of the mixing zone is:

$$L_{mixing} = 0.3 v_{head} = 0.3 \frac{(v_S - v_{LTBe})^2}{2g} \tag{E.70}$$

E.5.7 Algorithm for the calculation of pressure gradient and liquid holdup in horizontal slug flow

Input parameters:

- W_L and W_G = liquid and gas mass rates.
- d_{pipe} = pipe diameter.

Fluid properties:

- ρ_L and ρ_G = liquid and gas densities.
- μ_L and μ_G = liquid and gas viscosities.
- σ = surface tension.

Additional parameters:

- H_{LLS} = liquid holdup in the slug body.
- $freq_U$ = slug unit frequency.

Step 1 Calculate the slug velocity:

$$v_S = \frac{\frac{W_L}{\rho_L} + \frac{W_G}{\rho_G}}{A_{pipe}} \tag{E.71}$$

Step 2 Calculate the Taylor bubble velocity, v_{TB} and the scooping mass rate, x:

$$\rho_S = \rho_L H_{LLS} + \rho_G (1 - H_{LLS}) \tag{E.72}$$

$$\mu_S = \mu_L H_{LLS} + \mu_G (1 - H_{LLS}) \tag{E.73}$$

$$Re_S = \frac{\rho_S v_S d_{pipe}}{\mu_S} \tag{E.74}$$

$$c = 0.021 \ln (Re_S) + 0.22 \tag{E.75}$$

$$v_{TB} = (1 + c)\, v_S \tag{E.76}$$

$$x = c\,(\rho_L A_{pipe} H_{LLS} v_S) \tag{E.77}$$

Step 3 Assume a value for L_S:

$$L_S = L_{S-estimate} \tag{E.78}$$

Step 4 Calculate the length of the liquid film region, L_F:

$$L_F = \frac{v_{TB}}{freq_U} - L_S \tag{E.79}$$

Step 5 Integrate numerically the equation:

$$\frac{dh_F}{dz} = -\frac{\tau_F S_F + \rho_L A_{pipe} H_{LTB} g \sin\theta}{\frac{x^2}{\rho_L A_{pipe} H_{LTB}^2} \frac{dH_{LTB}}{dh_F} - \rho_L A_{pipe} H_{LTB} g \cos\theta} \tag{E.80}$$

from $z = 0$ to $z = L_F$. The boundary conditions are $h_F\,(z = 0) = H_{LLS} d$. The shear stress in the liquid film region is calculated with:

$$\tau_F = f_F \frac{\rho_L (v_{TB} - v_F)^2}{2} \tag{E.81}$$

This step determines $H_{LTS}\,(z)$ and $v_{LTB}\,(z)$.

Step 6 Calculate the length of the liquid slug, L_S from:

$$L_S = \frac{v_S}{freq_U\,(H_{LLS} - H_{LTB})} \left[\frac{W_L}{\rho_L A_{pipe} v_S} - H_{LTB} + c\,(H_{LLS} - H_{LTB}) \right] \tag{E.82}$$

Step 7 Compare the value of the assumed $L_{S-estimate}$ with the value of the L_S calculated from Equation E.82:

$$|L_S - L_{S-estimate}| < \epsilon \tag{E.83}$$

If the absolute value of the difference between the two values is less the acceptable tolerance ϵ, convergence is reached and continue with **Step 8**.
Otherwise, go back to **Step 4**.

Step 8 Calculate the pressure drops:

$$-\Delta P_{acceleration} = \frac{x}{A_{pipe}} (v_S - v_{TBe}) \tag{E.84}$$

$$-\Delta P_{friction} = f_S \frac{\rho_S v_S^2}{2} \frac{L_S}{d_{pipe}} \tag{E.85}$$

$$-\Delta P_U = -\Delta P_{acceleration} - \Delta P_{friction} \tag{E.86}$$

Step 9 The total pressure gradient is:

$$-\frac{dP}{dL} = -\frac{\Delta P_U}{L_U} \tag{E.87}$$

E.6 Felizola and Shoham mechanistic model for upward slug flow (1995)

This model builds up on the previous slug flow models by Fernandes et al. (1983), Sylvester (1987) and Vo and Shoham (1989). Many of the equations used by this model, Equations E.3, E.4, E.8 and E.12 have already been discussed in the previous sections and were derived based on the mass balance equation applied to different control volumes. The model considers that the liquid film in the stratified region of the slug flows backwards, so the term v_{LTB} is replaced by $-v_{LTB}$ in the aforementioned equations:

$$H_{LTB}\left[v_{TB}-(-v_{LTB})\right]=H_{LLS}\left(v_{TB}-v_{LLS}\right) \tag{E.88}$$

$$\left(1-H_{LTB}\right)\left(v_{TB}-v_{GTB}\right)=\left(1-H_{LLS}\right)\left(v_{TB}-v_{GLS}\right) \tag{E.89}$$

$$v_{SL}=v_{LLS}H_{LLS}\left(1-\beta\right)+v_{LTB}H_{LTB}\beta \tag{E.90}$$

$$v_{SG}=v_{GLS}\left(1-H_{LLS}\right)\left(1-\beta\right)+v_{GTB}\left(1-H_{LTB}\right)\beta \tag{E.91}$$

where $\beta=\frac{L_F}{L_U}=\frac{T_F}{T_U}$.
A force balance over the liquid film zone yields:

$$\underbrace{\underbrace{\underbrace{\rho_L \underbrace{H_{LTB}A_{pipe}}_{\text{area occupied by liquid in liquid film zone}}}_{\text{volume occupied by liquid in liquid film zone}}}_{\text{mass of liquid in the film zone}}\, dL\, g\sin\theta}_{\text{gravitational component parallel to the pipe axis}} = \tau_{LW}\frac{S_L}{A_{pipe}}+\tau_i\frac{S_L}{A_{pipe}}$$

$$= f_L\frac{\rho_L {v_{LTB}}^2}{2}\frac{S_i}{A_{pipe}} + f_i\frac{\rho_G\left[v_{GTB}-(-v_{LTB})\right]}{2}\frac{S_L}{A_{pipe}} \tag{E.92}$$

The geometrical parameters S_L and S_i are defined in Figure E.6.

E.6.1 Closure relationships

As the number of unknowns exceeds the number of equations, a set of closure relationships needs to be used. This is a set of empirical correlations.

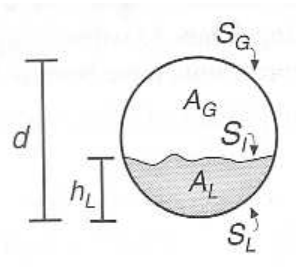


Figure E.6: Geometric parameters for the stratified region of slug flow (Shoham 2006).

Taylor bubble drift velocity v_{drift}

The Taylor bubble translational velocity v_{TB} can be expressed as:

$$v_{TB} = c_0 v_M + v_{drift} \tag{E.93}$$

where the flow coefficient c_0 is:

$$c_0 = \begin{cases} 1.2 & \text{for turbulent flow} \\ 2 & \text{for laminar flow} \end{cases} \tag{E.94}$$

Alves et al. (1993) suggested the following values for the flow coefficient c_0:

$$c_0 = \begin{cases} 1.0 & \text{for turbulent flow and } 0^\circ < \theta \leq 50^\circ \\ 1.15 & \text{for turbulent flow and } 50^\circ < \theta \leq 60^\circ \\ 1.25 & \text{for turbulent flow and } 60^\circ < \theta \leq 90^\circ \\ 2 & \text{for laminar flow} \end{cases} \tag{E.95}$$

A widely used correlation for the Taylor bubble drift velocity is Bendiksen (1984):

$$v_{drift} = 0.54\sqrt{g d_{pipe}}\cos\theta + 0.35\sqrt{g d_{pipe}}\sin\theta \tag{E.96}$$

for values of $\theta \in [0^\circ \ldots 90^\circ]$.

Gas bubble velocity in the slug body v_{GLS}

The gas bubble velocity in the slug body is the sum of the medium velocity and the bubble rise velocity. The Harmathy (1962) correlation can be used for the latter:

$$v_{GLS} = c_0 v_M + 1.53\sqrt[4]{\frac{(\rho_L - \rho_G)\, g\sigma}{{\rho_L}^2}}\, {H_{LLS}}^{0.5} \sin\theta \tag{E.97}$$

where the distribution coefficient, c_0 is:

$$c_0 = \begin{cases} 1.0 \ldots 1.5 & \text{for vertical flow} \\ 1.0 & \text{for horizontal and near horizontal flow} \end{cases} \tag{E.98}$$

The value used for inclined flow applications is $c_0 = 1.2$.

Liquid holdup in the slug body H_{LLS}

There are several correlations for H_{LLS}, but the one used by this model is Gomez et al. (2000):

$$H_{LLS} = e^{-\left(7.85\times10^{-3}\theta + 2.48\times10^{-6}Re_{LS}\right)} \tag{E.99}$$

where the Reynolds number Re_{LS} is:

$$Re_{LS} = \frac{\rho_L v_M d_{pipe}}{\mu_L} \tag{E.100}$$

Slug body length L_S

Based on experimental results, Shoham (2006) suggested that the length of the slug body in slug flow has values in the range $[20d_{pipe}, \cdots, 30d_{pipe}]$, with the lower end of the scale reached for vertical flow and the higher end for horizontal flow. Shoham improved the model by using a correlation for L_S based on a linear interpolation between the two ends of the scale.

E.6.2 Numerical solution

Two parameters, K_1 and K_2 are defined:

$$K_1 = v_{TB}H_{LLS} + v_{GLS}\left(1 - H_{LLS}\right) - v_M \tag{E.101}$$

$$K_2 = K_1 + v_M - v_{TB} \tag{E.102}$$

The eight Equations E.88-E.93, E.97 and E.99 can be reduced to just one equation with one unknown, h_F:

$$f_G\rho_G\left[\frac{K_1}{H_{LTB}} + \frac{K_2}{(1-H_{LTB})}\right]^2 S_i + f_L\rho_L\left(\frac{K_1}{H_{LTB}} - v_{TB}\right)^2 S_L - 2\rho_L g A_{pipe} H_{LTB}\sin\theta = 0 \tag{E.103}$$

because S_i, S_L and H_{LTB} can be expressed as functions of h_F:

$$S_i = d_{pipe}\sqrt{1 - \left(2\frac{h_F}{d_{pipe}}\right) - 1^2} \tag{E.104}$$

$$S_L = d_{pipe}\left[\pi - \arccos\left(2\frac{h_F}{d_{pipe}} - 1\right)\right] \tag{E.105}$$

$$H_{LTB} = \frac{\pi - \arccos\left(2\frac{h_F}{d_{pipe}} - 1\right) + \left(2\frac{h_F}{d_{pipe}} - 1\right)\sqrt{1 - \left(2\frac{h_F}{d_{pipe}} - 1\right)^2}}{\pi} \tag{E.106}$$

The solution $h_{F-solution}$ of Equation E.103 allows us to determine all other unknowns in the entire set of eight equations mentioned before.
To find the solution, the Newton-Raphson method could be used:

$$h_F|_{j+1} = h_F|_j - \frac{F}{F'} \tag{E.107}$$

where the function F is the LHS of Equation E.103.

$$\begin{aligned} F' &= \frac{dF}{dh_F} \\ &= 2\{f_G\rho_G S_i \left(\frac{K_1}{H_{LTB}} + \frac{K_2}{1-H_{LTB}}\right)\left[\frac{K_2}{(1-H_{LTB})^2} - \frac{K_1}{{H_{LTB}}^2}\right] \\ &\quad - f_L\rho_L S_L \left(\frac{K_1}{H_{LTB}} - v_{TB}\right)\frac{K_1}{{H_{LTB}}^2} - \rho_L g A_{pipe}\sin\theta\}\frac{dH_{LTB}}{dh_F} \\ &\quad + f_G\rho_G\left(\frac{K_1}{H_{LTB}} + \frac{K_2}{1-H_{LTB}}\right)^2\frac{dS_i}{dh_F} + f_L\rho_L\left(\frac{K_1}{H_{LTB}} - v_{TB}^2\right)^2\frac{dS_L}{dh_F} \end{aligned} \tag{E.108}$$

where:

$$\frac{dS_i}{dh_F} = \frac{2\left(1-2\frac{h_F}{d_{pipe}}\right)}{\sqrt{1-\left(2\frac{h_F}{d_{pipe}}\right)^2}} \tag{E.109}$$

$$\frac{dS_L}{dh_F} = \frac{2}{\sqrt{1-\left(2\frac{h_F}{d_{pipe}}\right)^2}} \tag{E.110}$$

$$\frac{dH_{LTB}}{dh_F} = \frac{4}{\pi d_{pipe}}\sqrt{1-\left(2\frac{h_F}{d_{pipe}}\right)^2} \tag{E.111}$$

There are two solutions to Equation E.103 and only the lower value is the correct one. For any iterative numeric algorithm, using an initial guess of $h_{F-initial} = 0.1d_{pipe}$ will allow the algorithm to always converge to the correct root.

E.6.3 Felizola Shoham algorithm for upward slug flow in inclined pipes

Step 1 Calculate v_{TB}, v_{GLS} and H_{LLS} from the closure relationships, Equations E.93, E.97 and E.99:

$$v_{TB} = c_0 v_M + v_{drift} \qquad v_M = v_{SL} + v_{SG} \tag{E.112}$$

where the flow coefficient c_0, with Alvez et al. (1993) suggested changes is:

$$c_0 = \begin{cases} 1.0 & \text{for turbulent flow and } 0^\circ < \theta \leq 50^\circ \\ 1.15 & \text{for turbulent flow and } 50^\circ < \theta \leq 60^\circ \\ 1.25 & \text{for turbulent flow and } 60^\circ < \theta \leq 90^\circ \\ 2 & \text{for laminar flow} \end{cases} \tag{E.113}$$

$$v_{GLS} = c_0 v_M + 1.53 \sqrt[4]{\frac{(\rho_L - \rho_G)\, g\sigma}{{\rho_L}^2}} {H_{LLS}}^{0.5} \sin\theta \tag{E.114}$$

$$H_{LLS} = e^{-\left(7.85\times10^{-3}\theta + 2.48\times10^{-6} Re_{LS}\right)} \tag{E.115}$$

Equations E.97 and E.103 are valid for $0° \leq \theta \leq 90°$.
Step 2 Calculate K_1 and K_2 parameters, as defined in E.101 and E.102:

$$K_1 = v_{TB} H_{LLS} + v_{GLS}\,(1 - H_{LLS}) - v_M \tag{E.116}$$

$$K_2 = K_1 + v_M - v_{TB} \tag{E.117}$$

with $v_M = v_{SL} + v_{SG}$.
Step 3 Determine h_F and H_{LTB} from Equation E.103:

$$f_G \rho_G \left[\frac{K_1}{H_{LTB}} + \frac{K_2}{(1 - H_{LTB})}\right]^2 S_i + f_L \rho_L \left(\frac{K_1}{H_{LTB}} - v_{TB}\right)^2 S_L - 2\rho_L g A_{pipe} H_{LTB} \sin\theta = 0 \tag{E.118}$$

The Newton-Raphson method could be used to determine the appropriate root $h_{F-solution}$. The remarks in Section E.6.2 relative to the use of the method are useful to this end.
Step 4 Calculate v_{LTB} and v_{GTB} from:

$$v_{LTB} = \frac{K_1}{H_{LTB}} - v_{TB} \tag{E.119}$$

$$v_{GTB} = \frac{K_2}{1 - H_{LTB}} + v_{TB} \tag{E.120}$$

Step 5 Solve for v_{LLS} Equation E.6:

$$(v_{TB} - v_{LLS})\,(1 - H_{LLS}) = (v_{TB} - v_{GTB})\,(1 - H_{LTB}) \tag{E.121}$$

Step 6 Solve for $\beta = \frac{L_F}{L_U}$ equation:

$$\beta = \frac{v_{LLS} H_{LLS} - v_{SL}}{v_{LLS} H_{LLS} + v_{LTB} H_{LTB}} \tag{E.122}$$

This is just another formulation of Equation E.8:

$$v_{SL} = v_{LLS} H_{LLS}\,(1 - \beta) + v_{LTB} H_{LTB} \beta \tag{E.123}$$

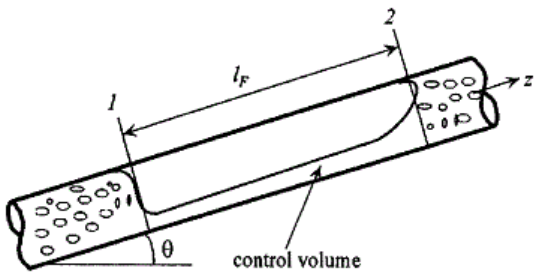


Figure E.7: Control volume used by the Zhang et al. (2000) model.

Step 7 Determine the length of the liquid film region, L_F and the length of the slug unit L_U from:

$$L_F = L_S \frac{\beta}{1-\beta} \tag{E.124}$$

$$L_U = L_S + L_F \tag{E.125}$$

Step 8 Determine the pressure drop in the slug unit and the pressure gradient:

$$-\Delta P_U = \rho_S g L_S \sin\theta + \tau_S \left(\pi d_{pipe}\right) \frac{L_S}{A_{pipe}} + \rho_F g L_F \sin\theta + \tau_F L_F \frac{S_F}{A_{pipe}} + \tau_G L_F \frac{S_G}{A_{pipe}} \tag{E.126}$$

and $-\frac{dP}{dL} = -\frac{\Delta P_U}{L_U}$.

E.7 Zhang et al. model for slug flow (2000)

Unlike the previous slug flow models by Dukler and Hubbard (1976) and Taitel and Barnea (1990), the Zhang et al. (2000) model considers the entire gas pocket-liquid film as the control volume. The model allows the calculation of:

- The pressure gradient in the slug unit.
- The slug unit frequency.
- The liquid holdup in the film region.

E.7.1 Equations for the Developed Slug Flow

Assumptions

1. No slip between the gas and liquid phases in the slug body. This is equivalent to:

$$v_{LLS} = v_{GLS} = v_S \tag{E.127}$$

2. Slug units are equally spaced, which is equivalent to L_U = constant.
3. Gas and liquid phases are incompressible.

Continuity equations

Assuming a control volume the size of the slug unit, the following mass balance can be written:

$$\rho_L A_{pipe} H_{LLS} \left(v_S - v_{TB}\right) = \rho_L A_{pipe} H_{LTB} \left(v_{LTB} - v_{TB}\right) \tag{E.128}$$

As shown earlier in Equations E.3 and E.4:

$$H_{LTB} \left(v_{TB} - v_{LTB}\right) = H_{LLS} \left(v_{TB} - v_{LLS}\right) \tag{E.129}$$

$$\left(1 - H_{LTB}\right) \left(v_{TB} - v_{GTB}\right) = \left(1 - H_{LLS}\right) \left(v_{TB} - v_{GLS}\right) \tag{E.130}$$

Summing the two Equations E.129 and E.130 the velocity v_S is obtained:

$$v_S = v_{GTB} \left(1 - H_{LTB}\right) + v_{LTB} H_{LTB} \tag{E.131}$$

Mass balances for the liquid and gas phases, as shown in Equations E.8 and E.12, lead to:

$$v_{SL} = v_{LLS} H_{LLS} \left(1 - \beta\right) + v_{LTB} H_{LTB} \beta \tag{E.132}$$

$$v_{SG} = v_{GLS} \left(1 - H_{LLS}\right) \left(1 - \beta\right) + v_{GTB} \left(1 - H_{LTB}\right) \beta \tag{E.133}$$

where $\beta = \frac{L_F}{L_U}$. With the non-slip assumption, $v_{LLS} = v_{GLS} = v_S$, the equations become:

$$v_{SL} L_U = v_S H_{LLS} L_S + v_{LTB} H_{LTB} L_F \tag{E.134}$$

$$v_{SG} L_U = v_S \left(1 - H_{LLS}\right) L_S + v_{GTB} \left(1 - H_{LTB}\right) L_F \tag{E.135}$$

Momentum equations

The momentum at section 1 in Figure E.7 is:

$$\rho_L A_{pipe} H_{LTB} \left(v_{LTB} - v_{TB}\right) v_{LTB} \tag{E.136}$$

and at section 2 is:

$$\rho_L A_{pipe} H_{LLS} \left(v_S - v_{TB}\right) v_S \tag{E.137}$$

Due to gas static and hydrostatic liquid pressures, the forces acting at section 1 are:

$$P_1 A_{pipe} H_{LTB} + \rho_L g \cos\theta \int_0^{h_{fe}} \left(h_{fe} - y\right) b\left(y\right) dy \tag{E.138}$$

and at section 2:

$$P_2 A_{pipe} H_{LLS} + \rho_L g \cos\theta \int_0^{h_{fi}} (h_{fi} - y)\, b\,(y)\, dy \tag{E.139}$$

The third type of forces acting on the curved surface of the liquid film in the direction z are due to the gas pressure P_2 pushing on the liquid slug body:

$$P_2 A_{pipe}\,(H_{LLS} - H_{LTB}) \tag{E.140}$$

Based on these facts, the momentum equation is:

$$\begin{aligned} P_1 A_{pipe} H_{LTB} &+ \rho_L g \cos\theta \int_0^{h_{fe}} (h_{fe} - y)\, b\,(y)\, dy \\ &- \left[P_2 A_{pipe} H_{LLS} + \rho_L g \cos\theta \int_0^{h_{fi}} (h_{fi} - y)\, b\,(y)\, dy \right] \\ &+ \tau_i S_i L_F - \tau_{LW} S_{LS} L_F - \rho_L g \sin\theta A_{pipe} H_{LTB} L_F \\ &= \rho_L A_{pipe} H_{LLS}\,(v_S - v_{TB}) - \rho_L A_{pipe} H_{LTB}\,(v_{LTB} - v_{TB})\, v_{LTB} \end{aligned} \tag{E.141}$$

The RHS of Equation E.141 can be further expanded based on:

$$H_{LLS}\,(v_S - v_{TB}) = H_{LTB}\,(v_{LTB} - v_{TB})$$

as shown in Equation E.129. The RHS of Equation E.141 becomes:

$$\rho_L A_{pipe} H_{LTB}\,(v_{LTB} - v_{TB})\,(v_S - v_{LTB}) \tag{E.142}$$

Replacing the RHS of Equation E.141 with E.142, the momentum Equation E.141 becomes:

$$\begin{aligned} \frac{P_2 - P_1}{L_F} &= \frac{\rho_L\,(v_{TB} - v_{LTB})\,(v_S - v_{LTB})}{L_F} \\ &+ \frac{\tau_i S_i - \tau_{LW} S_W}{H_{LTB} A_{pipe}} - \rho_L g \sin\theta \\ &- \frac{\rho_L g \cos\theta}{H_{LTB} A_{pipe} L_F} \left[\int_0^{hfi} (h_{fi} - y)\, b_1\,(y)\, dy \right] \\ &- \frac{\rho_L g \cos\theta}{H_{LTB} A_{pipe} L_F} \left[\int_0^{hfe} (h_{fe} - y)\, b_2\,(y)\, dy \right] \end{aligned} \tag{E.143}$$

Similarly, for the gas phase:

$$\frac{P_2 - P_1}{L_F} = \frac{\rho_G\,(v_{TB} - v_{GTB})\,(v_S - v_{GTB})}{L_F} + \frac{-\tau_i S_i - \tau_{GW} S_W}{(1 - H_{LTB})\, A_{pipe}} - \rho_G g \sin\theta \tag{E.144}$$

Equating the pressure drop from the previous two Equations E.143 and E.144, yields:

$$\begin{aligned} 0 \;=\; & \frac{\rho_L \left(v_{TB} - v_{LTB}\right)\left(v_S - v_{LTB}\right) - \rho_G \left(v_{TB} - v_{GTB}\right)\left(v_S - v_{GTB}\right)}{L_F} \\ & - \frac{\rho_L g \cos\theta}{L_F H_{LTB} A_{pipe}} \left[\int_0^{hfi} \left(h_{fi} - y\right) b_1\left(y\right) dy - \int_0^{hfe} \left(h_{fe} - y\right) b_2\left(y\right) dy\right] \\ & + \frac{\tau_{GW} S_{GW}}{\left(1 - H_{LTB}\right) A_{pipe}} - \frac{\tau_{LW} S_{LW}}{H_{LTB} A_{pipe}} \\ & + \tau_i S_i \left(\frac{1}{H_{LTB} A_{pipe}} + \frac{1}{\left(1 - H_{LTB}\right) A_{pipe}}\right) \\ & - \left(\rho_L - \rho_G\right) g \sin\theta \end{aligned} \tag{E.145}$$

$$L_U = L_S + L_F \tag{E.146}$$

The first term in Equation E.145 represents the momentum exchange between the slug body and the liquid film zone. Zhang et al.'s model was the first model to take into account the momentum exchange term, as described in Equation E.145.

E.7.2 Zhang et al. (2000) model calculation algorithm

The Equations used by the solution algorithm are E.129, E.130, E.131, E.134, E.145 and E.146.

Step 1 Start with an estimated value for H_{LTB}.

Step 2 L_S is known as a closure relationship. The Zhang et al. (2000) correlation found in Section F.1 is used:

$$L_S = \left(32 \quad \cos^2\theta + 16 \quad \sin^2\theta\right) d \tag{E.147}$$

Step 3 Calculate v_{TB} using the equation:

$$v_{TB} = c_0 v_S + v_{drift} \qquad v_S = v_{SL} + v_{SG} \tag{E.148}$$

where the flow coefficient $c_0 = 1.2$ for turbulent flow and $c_0 = 2$ for laminar flow. These were the values considered in the Zhang et al (2000) slug flow calculation algorithm.

Alvez et al. (1993) improved the values for the c_0 coefficient to:

$$c_0 = \begin{cases} 1.0 & \text{for turbulent flow and } 0° < \theta \leq 50° \\ 1.15 & \text{for turbulent flow and } 50° < \theta \leq 60° \\ 1.25 & \text{for turbulent flow and } 60° < \theta \leq 90° \\ 2 & \text{for laminar flow} \end{cases} \tag{E.149}$$

The Taylor bubble drift velocity is calculated with the Bendiksen (1984) correlation:

$$v_{drift} = 0.54\sqrt{g d_{pipe}} \cos\theta + 0.35\sqrt{g d_{pipe}} \sin\theta \tag{E.150}$$

for values of $\theta \in [0° \ldots 90°]$.
Step 4 Use the Gregory correlation described in Section F.3 to calculate H_{LLS}:

$$H_{LLS} = \frac{1}{1 + \left(\frac{v_S}{8.66}\right)^{1.39}} \tag{E.151}$$

Step 5 Calculate v_{LTB} and v_{GTB} using Equations E.129 and E.130:

$$\begin{cases} H_{LLS}\left(v_{TB} - v_S\right) & = H_{LTB}\left(v_{TB} - v_{LTB}\right) \\ \left(1 - H_{LLS}\right)\left(v_{TB} - v_S\right) & = \left(1 - H_{LTB}\right)\left(v_{TB} - v_{GTB}\right) \end{cases}$$

Step 6 Calculate L_F and L_U from:

$$\begin{cases} L_U v_{SL} & = L_S v_S H_{LLS} + L_F v_{LTB} H_{LTB} \\ L_U & = L_S + L_F \end{cases}$$

Step 7 Calculate the wetted wall fraction factor Θ, using the Grolman correlation:

$$\Theta = \Theta_0 \left(\frac{\sigma_W}{\sigma}\right)^{0.15} + \frac{\rho_G}{\left(\rho_L - \rho_G\right)} \frac{1}{\cos\theta} \left(\frac{{\rho_L v_{SL}}^{0.25} d}{\sigma}\right)^{0.25} \left[\frac{{v_{SG}}^2}{\left(1 - H_{LTB}\right)^2 g d}\right]^{0.8} \tag{E.152}$$

Step 8 Calculate the geometrical parameters S_F, S_G, S_i, A_F and A_G:

$$\begin{aligned} S_F &= \pi d \Theta \\ A_F &= H_{LTB} A_{pipe} \\ S_G &= \pi d - A_F \\ A_G &= \left(1 - H_{LTB}\right) A_{pipe} \\ S_{CD} &= d \sin\left(\pi\Theta\right) \\ A_{CD} &= \frac{d^2}{4}\left(\pi\Theta - \frac{\sin\left(2\pi\Theta\right)}{2}\right) \\ S_i &= \frac{S_F\left(A_{CD} - A_F\right) + S_{CD} A_F}{A_{CD}} \end{aligned} \tag{E.153}$$

These parameters are shown graphically in Figure E.8.
Step 9 Calculate the friction factors f_{GW}, f_{LW} and f_i:

$$\begin{aligned} f_{GW} &= C Re_{LW}^{-n} \\ f_{LW} &= C Re_{GW}^{-n} \\ f_i &= 0.0142 \end{aligned} \tag{E.154}$$

where the C and n factors are:

$$C = \begin{cases} 1.0 & \text{for laminar flow} \\ 0.046 & \text{for turbulent flow} \end{cases} \tag{E.155}$$

$$n = \begin{cases} 1 & \text{for laminar flow} \\ 0.2 & \text{for turbulent flow} \end{cases} \tag{E.156}$$

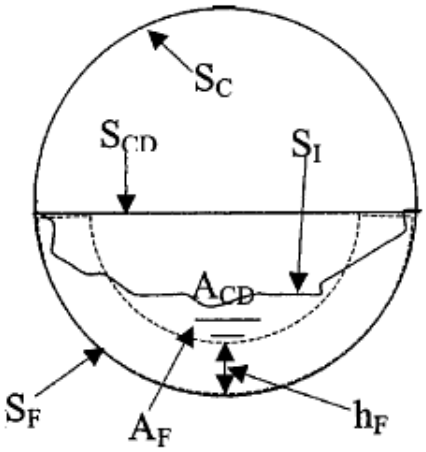


Figure E.8: Geometrical parameters in slug flow (Zhang et al. 2000).

The Reynolds numbers are calculated with:

$$Re_{LW} = \frac{4\ A_F\ v_{LTB}\ \rho_L}{S_F\ \mu_L}$$
$$Re_{GW} = \frac{4\ A_G\ v_{GTB}\ \rho_G}{(S_G + S_i)\,\mu_L} \tag{E.157}$$

The interfacial friction factor f_i in Equation E.154 is using the Cohen-Hanratty correlation shown in Section F.6.

Step 10 Calculate the shear stresses τ_{GW}, τ_{LW} and τ_i:

$$\tau_{GW} = f_{GW}\frac{\rho_G\ (v_{GTB})^2}{2}$$
$$\tau_{LW} = f_{LW}\frac{\rho_L\ (v_{LTB})^2}{2}$$
$$\tau_i = f_i\frac{\rho_G\,(v_{GTB} - v_{LTB})\,|v_{GTB} - v_{LTB}|}{2} \tag{E.158}$$

Step 11 After all the terms in Equation E.145 have been calculated, determine the margin of error m_{err}, which is the value of the RHS of the equation:

$$\begin{aligned} m_{err} \;=\; & \frac{\rho_L\,(v_{TB} - v_{LTB})\,(v_S - v_{LTB}) - \rho_G\,(v_{TB} - v_{GTB})\,(v_S - v_{GTB})}{L_F} \\ & - \frac{\rho_L g\cos\theta}{L_F H_{LTB} A_{pipe}} \\ & \times \left[\int_0^{hfi} (h_{fi} - y)\,b_1\,(y)\,dy - \int_0^{hfe} (h_{fe} - y)\,b_2\,(y)\,dy\right] \\ & + \frac{\tau_{GW} S_{GW}}{(1 - H_{LTB})\,A_{pipe}} - \frac{\tau_{LW} S_{LW}}{H_{LTB} A_{pipe}} \\ & + \tau_i S_i \left(\frac{1}{H_{LTB} A_{pipe}} + \frac{1}{(1 - H_{LTB})\,A_{pipe}}\right) \\ & - (\rho_L - \rho_G)\,g\sin\theta \end{aligned} \tag{E.159}$$

The integral terms are only about 1% of the value of the mixing term, which is the first term in the RHS of the Equation E.159, according to (Zhang et al. 2000).

Hence, they can be safely ignored.
Step 12 Check if the margin of error m_{err} is less than the admitted tolerance ϵ:

$$|m_{err}| < \epsilon \tag{E.160}$$

If Condition E.160 is satisfied, continue with **Step 13**. Otherwise, calculate H_{LTB} from m_{err} and go back to **Step 5** to reiterate the algorithm.
Step 13 Calculate the pressure gradient and the liquid holdup for the slug unit with Equations E.161 and E.162:

$$\left.\frac{dP}{dL}\right|_U = -\frac{L_F}{L_U}\frac{\tau_{LW}S_{LW} + \tau_{GW}S_{GW}}{A_{pipe}} - \frac{L_S}{L_U}\frac{\tau_S S_S}{A_{pipe}} - \frac{L_F g \sin\theta}{L_U}\left[\rho_L H_{LTB} + \rho_G\left(1 - H_{LTB}\right)\right] - \frac{L_S}{L_U}\rho_S g \sin\theta \tag{E.161}$$

$$H_{LU} = H_{LLS}\frac{L_S}{L_U} + H_{LTB}\frac{L_F}{L_U} \tag{E.162}$$

E.7.3 Zhang et al. (2003) model calculation algorithm

A unified hydrodynamic mechanistic model for gas-liquid pipe flow was developed at Tulsa University in 2003 (Zhang et al. 2003c). Unlike previous mechanistic models, this model makes predictions for both flow-pattern transition and flow behaviour using slug dynamics. The equations of slug flow are used not only to calculate the slug characteristics, but also to predict the boundary transitions from slug flow to other flow patterns. Apart from being a unified model, as far as the slug flow model is concerned, some improvements were made to the 2000 slug model:

- The model improves the previous assumptions and considers the liquid entrainment in the gas core of the slug unit.
- The Cohen and Hanratty (Equation F.32) interfacial friction factor correlation is replaced by the Andritsos and Hanratty (Equation F.33) correlation for cases when the flow is stratified, or Asali and Ambrosio correlation (Equation F.35) when the flow is annular.
- The shear stress at the pipe wall for laminar flow of the film zone of the Taylor bubble is now calculated with the equation:

$$\tau_{LW} = \frac{3\ \mu_L\ v_{LTB}}{h_F} - \frac{\tau_i}{2} \tag{E.163}$$

where the h_F is the average of liquid film height in the film zone and is calculated with:

$$h_F = \frac{2A_{pipe}H_{LTB}}{S_F + S_i} \tag{E.164}$$

This research study analyses the performance of the Zhang et al. model as it transitioned from the (2000) version to the (2003) version. The analysis focuses

on the replacement of the correlations for H_{LLS} and f_i in the new version. For a fair model comparison between the two model versions, the number of phases considered has to be the same: gas and liquid, without liquid entrainment.

With the assumption of no liquid entrainment, the calculation algorithm of the 2003 version is:

Step 1 Start with an estimated value for L_F.

Step 2 Use the Zhang et al. (2000) correlation for L_S (Section F.1):

$$L_S = \left(32\,\cos^2\theta + 16\,\sin^2\theta\right) d \tag{E.165}$$

Step 3 Calculate v_{TB} using the Bendiksen correlation:

$$v_{TB} = c_0 v_S + v_{drift} \qquad v_S = v_{SL} + v_{SG} \tag{E.166}$$

where the flow coefficient $c_0 = 1.2$ for turbulent flow and $c_0 = 2$ for laminar flow.

$$v_{drift} = 0.54\sqrt{g d_{pipe}}\cos\theta + 0.35\sqrt{g d_{pipe}}\sin\theta \tag{E.167}$$

for values of $\theta \in [0^\circ \ldots 90^\circ]$.

Step 4 Use the Gregory correlation to calculate an initial estimation of H_{LLS}:

$$H_{LLS} = \frac{1}{1 + \left(\frac{v_S}{8.66}\right)^{1.39}} \tag{E.168}$$

Step 5 Calculate v_{LTB} and v_{GTB} using Equations E.129 and E.130:

$$\begin{cases} H_{LLS}\left(v_{TB} - v_S\right) & = H_{LTB}\left(v_{TB} - v_{LTB}\right) \\ \left(1 - H_{LLS}\right)\left(v_{TB} - v_S\right) & = \left(1 - H_{LTB}\right)\left(v_{TB} - v_{GTB}\right. \end{cases}$$

Step 6 Recalculate H_{LLS} using the more "accurate" Zhang correlation described in Section F.3:

$$H_{LLS} = \frac{1}{1 + \frac{T_{sm}}{3.16\sqrt{(\rho_L - \rho_G) g \sigma}}} \tag{E.169}$$

The parameter T_{sm} has the form:

$$T_{sm} = \frac{1}{C_e}\left[f_S \frac{\rho_S {v_S}^2}{2} + \frac{d}{4}\frac{\rho_L H_{LTB}\left(v_{TB} - v_{LTB}\right)\left(v_M - v_{LTB}\right)}{L_S}\right] \tag{E.170}$$

and C_e is:

$$C_e = \frac{2.5 - |\sin\theta|}{2} \tag{E.171}$$

Step 7 Calculate L_U from:

$$L_U = L_S + L_F \tag{E.172}$$

Step 8 Calculate the wetted wall fraction factor Θ using the Grolman correlation:

$$\Theta = \Theta_0\left(\frac{\sigma_W}{\sigma}\right)^{0.15} + \frac{\rho_G}{\left(\rho_L - \rho_G\right)}\frac{1}{\cos\theta}\left(\frac{\rho_L {v_{SL}}^{0.25} d}{\sigma}\right)^{0.25}\left[\frac{{v_{SG}}^2}{\left(1 - H_{LTB}\right)^2 g d}\right]^{0.8} \tag{E.173}$$

Step 9 Calculate the geometrical parameters S_F, S_G, S_i, A_F and A_G:

$$\begin{aligned} S_F &= \pi d \Theta \\ A_F &= H_{LTB} A_{pipe} \\ S_G &= \pi d - A_F \\ A_G &= (1 - H_{LTB})\, A_{pipe} \\ S_{CD} &= d \sin(\pi \Theta) \\ A_{CD} &= \frac{d^2}{4}\left(\pi\Theta - \frac{\sin(2\pi\Theta)}{2}\right) \\ S_i &= \frac{S_F\,(A_{CD} - A_F) + S_{CD} A_F}{A_{CD}} \end{aligned} \tag{E.174}$$

Step 10 Calculate the friction factors f_{GW}, f_{LW} and f_i:

$$\begin{aligned} f_{GW} &= C Re_{LW}^{-n} \\ f_{LW} &= C Re_{GW}^{-n} \end{aligned} \tag{E.175}$$

where the C and n factors are:

$$C = \begin{cases} 1.0 & \text{for laminar flow} \\ 0.046 & \text{for turbulent flow} \end{cases} \tag{E.176}$$

$$n = \begin{cases} 1 & \text{for laminar flow} \\ 0.2 & \text{for turbulent flow} \end{cases} \tag{E.177}$$

The Reynolds numbers are calculated with:

$$\begin{aligned} Re_{LW} &= \frac{4\ A_F\ v_{LTB}\ \rho_L}{S_F\ \mu_L} \\ Re_{GW} &= \frac{4\ A_G\ v_{GTB}\ \rho_G}{(S_G + S_i)\,\mu_L} \end{aligned} \tag{E.178}$$

Depending on whether the flow in the film zone is annular or stratified, Asali and Ambrosio or Andritsos and Hanratty correlations for f_i are used. The decision about the flow type is made by testing the value of the wetted wall fraction parameter, Θ:

$$\Theta \begin{cases} \leq 0.8 & \text{stratified flow} \\ > 0.8 & \text{annular flow} \end{cases} \tag{E.179}$$

If the flow is annular, the Asali and Ambrosio correlation is used. Otherwise, the Andritsos and Hanratty correlation for f_i is used. These correlations are described in Section F.6.

Step 11 Calculate the shear stresses τ_{GW}, τ_{LW} and τ_i:

$$\begin{aligned}
\tau_{GW} &= f_{GW}\frac{\rho_G\ {v_{GTB}}^2}{2} \\
\tau_{LW} &= \frac{3\ \mu_L\ v_{LTB}}{h_F} - \frac{\tau_i}{2} \\
\tau_i &= f_i\frac{\rho_G\left(v_{GTB}-v_{LTB}\right)\left|v_{GTB}-v_{LTB}\right|}{2}
\end{aligned} \tag{E.180}$$

where the h_F is the average of the film zone and is calculated with:

$$h_F = \frac{2A_{pipe}H_{LTB}}{S_F + S_i} \tag{E.181}$$

Step 12 Calculate the margin of error m_{err}:

$$\begin{aligned}
m_{err} \;=\; & \frac{\rho_L\left(v_{TB}-v_{LTB}\right)\left(v_S-v_{LTB}\right)-\rho_G\left(v_{TB}-v_{GTB}\right)\left(v_S-v_{GTB}\right)}{L_F} \\
& - \frac{\rho_L g\cos\theta}{L_F H_{LTB} A_{pipe}} \\
& \times \left[\int_0^{hfi}\left(h_{fi}-y\right)b_1\left(y\right)dy - \int_0^{hfe}\left(h_{fe}-y\right)b_2\left(y\right)dy\right] \\
& + \frac{\tau_{GW}S_{GW}}{\left(1-H_{LTB}\right)A_{pipe}} - \frac{\tau_{LW}S_{LW}}{H_{LTB}A_{pipe}} \\
& + \tau_i S_i\left(\frac{1}{H_{LTB}A_{pipe}} + \frac{1}{\left(1-H_{LTB}\right)A_{pipe}}\right) \\
& - \left(\rho_L-\rho_G\right)g\sin\theta
\end{aligned} \tag{E.182}$$

Step 13 Check if the margin of error m_{err} is less than the admitted tolerance ϵ:

$$\left|m_{err}\right| < \epsilon \tag{E.183}$$

If Condition E.183 is satisfied, continue with **Step 14**. Otherwise, calculate L_F from m_{err} and go back to **Step 5** to re-iterate the algorithm.
Step 14 Calculate the pressure gradient and the liquid holdup of the slug unit:

$$\begin{aligned}
\left.\frac{dP}{dL}\right|_U = & -\frac{L_F}{L_U}\frac{\tau_{LW}S_{LW}+\tau_{GW}S_{GW}}{A_{pipe}} - \frac{L_S}{L_U}\frac{\tau_S S_S}{A_{pipe}} \\
& - \frac{L_F g\sin\theta}{L_U}\left[\rho_L H_{LTB} + \rho_G\left(1-H_{LTB}\right)\right] - \frac{L_S}{L_U}\rho_S g\sin\theta
\end{aligned} \tag{E.184}$$

$$H_{LU} = H_{LLS}\frac{L_S}{L_U} + H_{LTB}\frac{L_F}{L_U} \tag{E.185}$$

Appendix F

Empirical correlations used by slug flow models

This study will consider the following empirical correlations, grouped by the parameter they calculate:

F.1 Empirical correlations for the calculation of parameter Liquid Slug Length, L_S

1. **Scott et al. (1989)**:

$$L_S = \max\left\{30d, e^{-26.8+28.5\left[\ln\left(\frac{d}{0.0254}\right)\right]^{0.1}}\right\} \tag{F.1}$$

   This correlation is based on data from the Prudhoe Bay field (Bratland 2009).

2. **Zhang et al. (2003c)**:

$$L_S = \left(32\cos^2\theta + 16\sin^2\theta\right) d \tag{F.2}$$

3. **Felizola and Shoham (2006)**:

   This correlation was discussed in the Felizola and Shoham model, in the closure relationships subsection E.6.1.

F.2 Empirical correlations for the calculation of parameter Slug Frequency, $freq_{slug}$

1. **Zabaras (2000)**: The SI form of the correlation is:

$$freq_{slug} = 0.0226\left(\frac{v_{SL}}{gd}\right)^{1.2}\left[\frac{64.8}{v_M} + 3.281 v_M\right]^{1.2}\left[0.836 + 2.75\left(\sin\theta\right)\right]^{0.25} \tag{F.3}$$

This correlation is valid for inclination angles between $0° \leq \theta \leq 11°$ and pipe diameters 0.0254 m $\leq d \leq$ 0.2 m. It is worth mentioning that under steady-state conditions, $v_{TB} = v_{slug}$ and:

$$L_U = \frac{v_{slug}}{freq_{slug}} \tag{F.4}$$

F.3 Empirical correlations for the calculation of parameter Slug Liquid Holdup, H_{LLS}

Based on studies by Dukler et al. (1964), Marcano (1973), Palmer (1975), Mandhane et al. (1975), Spedding et al. (1990), Spedding (1997) and Woldesmayat and Ghajar (2006), which compared different empirical correlations for void fraction in pipes at different inclination angles, the following empirical correlations were selected. According to these studies these correlations have an accuracy range of ± 15 %:

1. **Eissa Al-Safran (2009)**:

$$H_{LLS} = 1.05 - \frac{0.0417}{mr^{\curvearrowleft} - 0.123} \tag{F.5}$$

The parameter mr is the momentum transfer rate to accelerate the mass of the liquid from v_{LTB} to the slug mixture v_M. Its mathematical expression is:

$$mr = \rho_L A_{pipe} H_{LTB} \left(v_{TB} - v_{LTB}\right) \left(v_M - v_{LTB}\right) \tag{F.6}$$

The non-dimensional form of the momentum transfer rate, $mr^{\curvearrowleft}$, is given by:

$$mr^{\curvearrowleft} = \frac{mr}{\rho_L A_{pipe} v_M^2} \tag{F.7}$$

2. **Woldesmayat and Ghajar (2006)**:

$$\begin{aligned} H_{LLS} &= 1 - \frac{v_{SG}}{\tilde{v}} \\ \tilde{v} &= v_{SG}\left[1 + \left(\frac{v_{SL}}{v_{SG}}\right)^{\left(\frac{\rho_G}{\rho_L}\right)^{0.1}}\right] \\ &\quad + 2.9\left[\frac{gd\sigma\left(1+\cos\theta\right)\left(\rho_L - \rho_G\right)}{{\rho_L}^2}\right]^{0.25} \\ &\quad \left(1.22 + 1.22\sin\theta\right)^{\frac{P_{atmospheric}}{P_{system}}} \end{aligned}$$

This correlation is a modified version of the Coddington and Macian (2002) correlation and is valid for all flow regimes.

3. **Toshiba (Leung 2005)**:

$$H_{LLS} = 1 - \frac{v_{SG}}{1.08 v_M + 0.45} \tag{F.8}$$

This correlation is valid for inclined pipes and all flow regimes.

4. **Zhang (2003a)** :

$$H_{LLS} = \frac{1}{1 + \frac{T_{sm}}{3.16\sqrt{(\rho_L - \rho_G) g \sigma}}} \tag{F.9}$$

The parameter T_{sm} has the form:

$$T_{sm} = \frac{1}{C_e}\left[f_S \frac{\rho_S {v_M}^2}{2} + \frac{d}{4}\frac{\rho_L H_{LTB}\left(v_{TB} - v_{LTB}\right)\left(v_M - v_{LTB}\right)}{L_S}\right] \tag{F.10}$$

and C_e is:

$$C_e = \frac{2.5 - |\sin\theta|}{2} \tag{F.11}$$

5. **Gomez (2000)**:
This correlation is valid between inclination angles $0 \leq \theta \leq \pi/2$.

$$H_{LLS} = 1.0 \times e^{-\left(7.85\times10^{-3}\theta + 2.48\times10^{-6} Re_{LS}\right)} \tag{F.12}$$

where the Reynolds number is:

$$Re_{LS} = \frac{\rho_L v_M d}{\mu_L} \tag{F.13}$$

6. **Spedding and Spence (1989)**:

$$\frac{\epsilon}{1-\epsilon} = \left[0.45 + 0.08 e^{-100\left(0.25 - v_{SL}^2\right)}\right]\left(\frac{v_{SG}}{v_{SL}}\right)^{0.65} \tag{F.14}$$

where:

$$H_{LLS} = 1 - \epsilon \tag{F.15}$$

7. **Minami and Brill (1987)**:

$$H_{LLS} = 1 - e^{-\left[\frac{\ln Z_1 + 9.21}{8.7115}\right]^{4.3374}} \tag{F.16}$$

where:

$$Z_1 = \frac{1.84 {v_{SL}}^{0.575}}{v_{SG} d^{0.0277}}\left(\frac{{\rho_L}^{0.5804}}{g^{0.3696}\sigma^{0.1804}}\right)^{-0.25}\left(\frac{P_{system}}{101325}\right)^{0.05} {\mu_L}^{0.1} \tag{F.17}$$

This correlation applies to all flow regimes and pipe inclinations.

8. **Barnea and Brauner (1985)**:
   This correlation is independent of the inclination angle.

$$H_{LLS} = 1 - 0.058\left[2\sqrt{\frac{0.4\sigma}{(\rho_L - \rho_G)\, g}}\frac{\rho_L}{\sigma}^{0.666}\frac{f_S {v_M}^3}{2d}^{0.4} - 0.725\right]^2 \quad \text{(F.18)}$$

9. **Gregory (1978)**:
   This correlation was designed and tested for horizontal flow:

$$H_{LLS} = \frac{1}{1 + \left(\frac{v_M}{8.66}\right)^{1.39}} \quad \text{(F.19)}$$

   Despite its limited validity, various flow models use it as a low accuracy estimate for an initial value for H_{LLS}, while the more accurate values are calculated by other correlations at later steps in the calculation algorithm.

10. **Nicklin et al. (1962)**:

$$H_{LLS} = 1 - \frac{v_{SG}}{1.2 v_M + 0.35\sqrt{gd}} \quad \text{(F.20)}$$

F.4 Empirical correlations for the calculation of parameter Taylor Bubble Translational Velocity, v_{TB}

1. **Bendiksen**:

$$v_{TB} = C_S v_M + v_{drift} \quad \text{(F.21)}$$

$$v_{drift} = 0.54\sqrt{gd}\cos\theta + 0.35\sqrt{gd}\sin\theta \quad \text{(F.22)}$$

   where C_S is 1.2 for turbulent flow and 2 for laminar flow.

2. **Hassan and Kabir**:

$$v_{drift} = 0.35\sqrt{\frac{g\sigma_L\left(\rho_L - \rho_G\right)}{{\rho_L}^2}}\sqrt{\sin\theta}(1 + \cos\theta)^2 \quad \text{(F.23)}$$

   where $v_M = v_{SL} + v_{SG}$.

F.5 Empirical correlations for the calculation of parameter Wetted Wall Fraction, Θ

1. **Fan (2005)**:

$$\Theta = \left[0.57 {H_{LTB}}^{0.345} + 0.0637 {Fr_L}^{0.68}\left(\frac{v_{SG}}{v_{SG,C}}\right)^{0.68}\right]\left(\frac{\sigma_W}{\sigma}\right)^{0.15} \quad \text{(F.24)}$$

for values of $\Theta \in [0, 0.5]$.

$$\Theta = \left[0.57{H_{LTB}}^{0.345} + 0.0637{Fr_L}^{0.68}\left(\frac{v_{SG}}{v_{SG,C}}\right)^{0.55}\right]\left(\frac{\sigma_W}{\sigma}\right)^{0.15} \tag{F.25}$$

for values of $\Theta \in [0.5, 1]$.
The critical value of the superficial gas velocity, $v_{SG,C}$ is:

$$v_{SG,C} = 5\left[\frac{m}{s}\right]\left[\frac{1.24}{\rho_G\left[\frac{kg}{m^3}\right]}\right]^{0.5} \tag{F.26}$$

$$Fr_L = \frac{\rho_L {v_L}^2}{(\rho_L - \rho_G)\, gd\cos\theta} \tag{F.27}$$

2. **Biberg (1999)**:

$$\beta = 2\pi - 2\left\{\pi H_{LTB} + \left(\frac{3\pi}{2}\right)^{\frac{1}{3}}\left[1 - 2H_{LTB} + {H_{LTB}}^{\frac{1}{3}} - (1 - H_{LTB})^{\frac{1}{3}}\right]\right\} \tag{F.28}$$

For stratified flow, this correlation is accurate within ± 0.002 radians, according to Bratland (2009).

3. **Grolman (1994)**:

$$\Theta = \Theta_0\left(\frac{\sigma_W}{\sigma}\right)^{0.15} + \frac{\rho_G}{(\rho_L - \rho_G)}\frac{1}{\cos\theta}\left(\frac{\rho_L {v_{SL}}^{0.25} d}{\sigma}\right)^{0.25}\left[\frac{{v_{SG}}^2}{(1 - H_{LTB})^2 gd}\right]^{0.8} \tag{F.29}$$

4. **Hart et al. (1989)**:

$$\Theta = \min\left\{1, \Theta_0 + CFr^{0.58}\right\} \tag{F.30}$$

where the constant $C = 0.26$ and the Froude number is:

$$Fr = \frac{\rho_L {v_L}^2}{(\rho_L - \rho_G)\, gd} \tag{F.31}$$

F.6 Empirical correlations for the calculation of the interfacial friction factor f_i

1. **Cohen and Hanratty (1968)**: Cohen and Hanratty proposed a constant value of 0.0142 for the interfacial friction factor in stratified flow, corresponding to a fully developed rough interface, due to the presence of small liquid waves.

$$f_i = 0.0142 \tag{F.32}$$

2. **Andritsos and Hanratty (1987)**: The correlation in Equation F.33 applies to stratified flow, under the assumption of a flat gas-liquid interface.

$$f_i = f_C \left[1 + 14.3\ {H_{LTB}}^{0.5} \left(\frac{v_{SG}}{v_{SG,t}} - 1\right)\right] \tag{F.33}$$

f_C is the friction factor of the gas in the core of the Taylor bubble.

$$v_{SG,t} = 5 \left[\frac{m}{s}\right] \left(\frac{\rho_{GO}}{\rho_G}\right)^{0.5} \tag{F.34}$$

ρ_{GO} in Equation F.34 is the gas density at atmospheric pressure.

3. **Ambrosio and Asali (1984)**: The correlation applies to annular flow.

$$f_i = f_G \left[1 + 13.8\ {We_G}^{0.2} {Re_G}^{-0.6} \left(h_F^+ - 200\sqrt{\frac{\rho_G}{\rho_L}}\right)\right] \tag{F.35}$$

The parameter h_F^+ is the dimensionless thickness of the liquid film and is defined in Equation F.36.

$$h_F^+ = \frac{\rho_G h_F v_C^*}{\mu_G} \tag{F.36}$$

$$v_C^* = \sqrt{\frac{\tau_i}{\rho_G}} \tag{F.37}$$

For a smooth pipe f_G is defined as:

$$f_G = 0.046\ {Re_G}^{-0.2} \tag{F.38}$$

The Weber and Reynolds numbers are calculated with:

$$We_G = \frac{\rho_G v_G^2 d}{\sigma} \tag{F.39}$$

$$Re_G = \frac{\rho_G v_G d}{\mu_G} \tag{F.40}$$

Equation F.35 shows the dependency of f_i on h_F^+. However, according to Equation F.36 h_F^+ depends on v_C^*. Equation F.37 shows the dependency of v_C^* on τ_i, the interfacial shear stress, which is defined as:

$$\tau_i = f_i \frac{\rho_G \left(v_G - v_L\right) \left|v_G - v_L\right|}{2} \tag{F.41}$$

A method for solving Equations F.35- F.41 is given in Zhang step 14 (2003 version) in Section 5.3.2.

4. **Vlachos et al. (1997)**: For stratified flow, the interfacial friction factor can be expressed by the empirical correlation:

$$f_i = 0.024\ {H_L}^{0.35} {Re_{SL}}^{0.18} \tag{F.42}$$

where the Reynolds number is:

$$Re_{SL} = \frac{\rho_L v_{SL} d}{\mu_L} \tag{F.43}$$

5. **Hamersma and Hart (1987)**:
For small liquid holdup, $H_L \leq 0.04$:

$$f_i = \frac{0.0625}{\left[\log\left(\frac{15}{Re_G} + \frac{\epsilon}{3.715d}\right)\right]^2} \tag{F.44}$$

where ϵ = estimated interfacial roughness $= 2.3\delta_L$ and δ_L = uniform liquid thickness.
For annular flow, the uniform liquid thickness can be obtained as a function of the liquid holdup H_L, as follows:

$$H_{LTB} = \frac{A_L}{A_{pipe}} = \frac{\frac{\pi}{4}\left[d^2 - (d-\delta_L)^2\right]}{\frac{\pi}{4}d^2} = 2\left(\frac{\delta_L}{d}\right) - \left(\frac{\delta_L}{d}\right)^2 \tag{F.45}$$

With the notation $x = \frac{\delta_L}{d}$, the following equation is obtained:

$$-x^2 + 2x - H_{LTB} = 0 \tag{F.46}$$

with the solutions:

$$x_{1,2} = 1 \pm \sqrt{1 - H_{LTB}} \tag{F.47}$$

and because $x \leq 1$, the only acceptable solution is:

$$\begin{aligned} x &= 1 - \sqrt{1 - H_{LTB}} \\ \delta_L &= d\left(1 - \sqrt{1 - H_{LTB}}\right) \end{aligned} \tag{F.48}$$

6. **Kowalski (1987)**:
This correlation was developed using data from experiments conducted under high pressure. Kowalski correlated the interfacial friction factor with the liquid holdup and the gas and liquid Reynolds numbers (Kowalski 1987):

$$f_i = 7.5 \times 10^{-5} H_L^{-0.25} Re_G^{-0.3} Re_L^{0.83} \tag{F.49}$$

where the Reynolds numbers are defined as:

$$\begin{aligned} Re_G &= \frac{\rho_G v_G d_{pipe}}{\mu_G} \\ Re_L &= \frac{\rho_L v_L d_{pipe}}{\mu_L} \end{aligned} \tag{F.50}$$

7. **Bendiksen et al. (1989)**:
Bendiksen et al. derived an expression for the average wave height of the interfacial waves in stratified flow, based on a force balance equation. The expression of Δh_{Wave} is:

$$\Delta h_{Wave} = \left[\frac{\rho_G (v_G - v_L)^2}{4\,(\rho_L - \rho_G)\,g\cos\theta}\right] + \sqrt{\left[\frac{\rho_G (v_G - v_L)^2}{4\,(\rho_L - \rho_G)\,g\cos\theta}\right]^2 - \frac{\sigma}{(\rho_L - \rho_G)\,g\cos\theta}} \tag{F.51}$$

The height of the wave in this equation can be used as the absolute roughness parameter to calculate the interfacial friction factor, either using the Colebrook and White, or the Hall Equation F.53.

8. **Sinai (1986)**:
Sinai derived an empirical expression for the roughness of the gas-liquid phase interface, based on the Charnock (1955) correlation for the interfacial friction factor. The relationship expresses the roughness of the interface in terms of the gas velocity:

$$\epsilon_i = 89.9\left(\frac{\rho_G}{\rho_L - \rho_G}\right)\left(\frac{S_I}{S_G + S_I}\right)\left(\frac{v_G^2}{g}\right) f_i \tag{F.52}$$

This relationship can be used in conjunction with the Colebrook and White equation (Brill and Beggs 1986), Hall correlation (1957) or Moody's equation. Moody produced the following approximate equation for the friction factor:

$$f_{Fanning} = 0.001375\left[1 + \left(2\times 10^4 \frac{\epsilon}{d_{pipe}} + \frac{10^6}{Re}\right) 0.333\right] \tag{F.53}$$

This equation gives values within $\pm 5\%$ for Reynolds numbers between 4000 and 10^7 and values of ϵ/d_{pipe} up to 0.01.

9. **Ouyang and Aziz (1996)**:
Using a regression approach, Ouyang and Aziz derived the following expression for the interfacial friction factor:

$$f_i = 10^{-8.0942 + 4.2893 H_L^{\sin\theta}} \frac{f_{WL}^{0.8732} N_{v_L}^{0.3072} N_D^{1.0365}}{N_G^{1.914} H_r^{0.9783}} \tag{F.54}$$

where H_r is the liquid holdup ratio, H_L is the liquid holdup, θ is the inclination angle and f_{LW} is the Fanning wall friction factor for the liquid phase. H_r is defined as follows:

$$H_r = \frac{v_{GTB}}{v_{LTB}} \tag{F.55}$$

The following parameters are defined:

$$R_{volumetric} = \frac{v_{SG}}{v_{SL}} \tag{F.56}$$

$$Re_{liquid} = \frac{\rho_L v_{LTB} H_{LTB} d}{\mu_L} \tag{F.57}$$

$$f_{LW} = 1.6291\left({Re_{liquid}}^{-0.5161}\right)\left({R_{volumetric}}^{0.0926}\right) \tag{F.58}$$

These parameters are used in the definition of the following dimensionless groups:

Liquid velocity number:

$$N_{v_L} = v_{SL}\sqrt[4]{\frac{\rho_L}{g\sigma}} \tag{F.59}$$

Gas viscosity number:

$$N_G = \mu_G\sqrt[4]{\frac{g}{\rho_L\sigma^3}} \tag{F.60}$$

Pipe diameter number:

$$N_D = d_{pipe}\sqrt{\frac{\rho_L g}{\sigma}} \tag{F.61}$$

Index

Bibliography

Acikgoz, M., F. Franca, and R.T. Lahey. 1991. "An experimental study of three-phase flow regimes." *International Journal of Multiphase Flow* 18 (3): 327–336.

Agrawal, S.S., G.A. Gregory, and G.W.. Govier. 1973. "An analysis of horizontal stratified two-phase flow in pipes." *The Canadian Journal of Chemical Engineering* 31 (3): 280–286.

Ansari, A., N. Sylvester, C. Sarica, O. Shoham, and J. Brill. 1994. "A Comprehensive Mechanistic Model for Upward Two-Phase Flow in Wellbores." *SPE Production and Facilities* 9 (2): 143–152.

A.Woldesmayat and A.J. Ghajar. 2007. "Comparison of void fraction correlations for different flow patterns in horizontal and upward inclined pipes." *International Journal of Multiphase Flow* 33 (4): 347–370.

Aziz, K., and L. Ouyang. 1995. "Steady-state gas flow in pipes." *Journal of Petroleum Science and Engineering* 14 (3): 137–158.

Barkley, R.W., and R.L. Motard. 1972. "Decomposition of nets." *The Chemical Engineering Journal,* no. 3:265–275.

Barnea, D. 1987. "A Unified Model for Predicting Flow-Pattern Transitions for the whole Range of Pipe Inclinations." *International Journal of Multiphase Flow* 13 (1): 1–12.

Barnea, D., Ovadia Shoham, and Y. Taitel. 1982. "Flow Pattern Transition for Vertical Downward Two-Phase Flow." *Chemical Engineering Science* 37 (5): 735–740.

Barton, P. 1995. "Structural analysis of systems of equations." *Chemical Engineering Science* 21 (7): 425–438.

Beggs, H.D., and J.P. Brill. 1973. "A Study of Two-Phase Flow in Inclined Pipes." *Journal of Petroleum Technology* 25 (5): 607–617.

Belmiro, D., P.M. Saraiva, and C.C. Pantelides. 2004. "Combined Mechanistic and Empirical Modelling." *International Journal of Chemical Reactor Engineering* 2 (1): 1542–6580.

Bonizzi, M., P. Andreussi, and S. Banerjee. 2009. "Flow regime independent, high resolution multi-field modelling of near-horizontal gas-liquid flows in pipelines." *International Journal of Multiphase Flow* 35 (1): 34–46.

Bratland, Ove. 2009. *Pipe Flow 2: Multiphase Flow Assurance.* Website. `http://www.drbratland.com`.

Brill, J., and D. Beggs. 1991. *Two-Phase Flow in Pipes.* 6th ed. 4–37. Forfatterne.

Brill, J., W. Coberly, J. Herring, and D. Moore. 1981. "Analysis of Two-Phase tests in Large-Diameter Flow Lines in Prudhoe Bay Field." *Society of Petroleum Engineering* 21 (3): 363–378.

Cengizhan, K., H-Q. Zhang, and C. Sarica. 2007. "Identification and Classification of New Three-Phase Gas/Oil/Water Flow Patterns." In *SPE Annual Technical Conference and Exhibition,* 1–13. Anaheim, California, U.S.A. doi:`110221`.

Cohen, S.L., and T.J. Hanratty. 1968. "Effects of waves at a gas-liquid interface on a turbulent air flow." *Journal of Fluid Mechanics* 31 (3): 467–469.

Curtin, University. 2011. *Petroleum Production Technology.* Lecture Notes.

Duff, I.S. 1981a. "Algorithm 575. Permutations for a zero-free diagonal." *ACM Transactions on Mathematical Software* 7:387–390.

Duff, I.S. 1981b. "On algorithms for obtaining a maximum transversal." *ACM Transactions on Mathematical Software* 7:315–330.

Dukler, A.E., M. Wickes, and R.G. Cleveland. 1964. "Frictional Pressure Drop in Two-Phase Flow: An Approach Through Similarity Analysis." *AIChE Journal* 10 (1): 44–51.

Eaton, J.W. 2013. *GNU Octave.* Website. `http://www.gnu.org/software/octave`.

Eissa, A-S. 2009. "Prediction of Slug Liquid Holdup in Horizontal Pipes." *Journal of Energy Resources Technology* 131 (2): 47–55.

Eriksen, T. 2009. *The Oil Drum - Discussions about Energy and Future.* Website. `http://www.theoildrum.com/node/5395`.

Falcone, G., G. Hewitt, and C. Alimonti. 2010. *Multiphase Flow Metering.* Developments in Petroleum Science. Elsevier.

Fan, Y., H-Q. Zhang, C. Sarica, and T. Danielson. 2007. "A model to predict liquid holdup and pressure gradient of near-horizontal wet-gas pipelines." *SPE Projects, Facilities Construction* 2 (2): 1–8.

Fiedler, B., and A.A. Schuppert. 2008. "Local identification of scalar hybrid models with tree structure." *IMA Journal of Applied Mathematics* 73 (3): 449–476.

Garcia, F., M. Garcia, R. Garcia, and D. Joseph. 2007. "Friction factor improved correlations for laminar and turbulent gas-liquid flow in horizontal pipelines." *International Journal of Multiphase Flow* 33 (2007): 1320–1336.

Gomez, L., O. Shoham, Z. Schmidt, R. Chokshi, and T. Northug. 2000. "Unified Mechanistic Model for Steady-State Two-Phase Flow: Horizontal to Vertical Upward Flow." *Society of Petroleum Engineers* 5 (3): 339–350.

Gregory, G.A., K.M. Nicholson, and K. Aziz. 1978. "Correlation of the liquid volume fraction in the slug for horizontal gas-liquid slug flow." *International Journal of Multiphase Flow* 4 (1): 33–39.

Guerrero-Sarabia, I., and Y.V. Fairuzov. 2006. "Stability Analysis of Gas-Lift Wells Used for Deep-Water Oil Production." In *First International Oil Conference and Exhibition in Mexico.* Cancun, Mexico. doi:10.2118/104037-MS.

Hall, A.R.W. 1992. "Multiphase flow of oil, water and gas in horizontal pipe." PhD diss., Imperial College, University of London, U.K.

Hanafy, H.H., S.M. Macary, Y.M. ElNady, A.A. Bayomi, and M.H. El Batanony. 1997. "Empirical PVT Correlations Applied to Egyptian Crude Oils Exemplify Significance of Using Regional Correlations." In *International Symposium on Oilfield Chemistry,* 733–737. Houston, Texas, U.S.A. doi:10.2118/37295-MS.

Heriot Watt, University. 2009. *Production Technology - Lecture Notes.* Heriot Watt University.

Himmelblau, D. 1966. "Decomposition of large scale systems - I Systems composed of lumped parameter elements." *Chemical Engineering Science* 21 (10): 425–438.

Ismail, I. 1999. *The Importance of Having an Accurate Multiphase Flow Correlation Program.* Technical report. Skudai, Johor, Malaysia: Department of Petroleum Engineering, University of Technology Malaysia.

Issa, R.I. 1988. "Prediction of turbulent, stratified, two-phase flow in inclined pipes and channels." *International Journal of Multiphase Flow* 14 (2): 141–154.

Kahrs, O., and W. Marquardt. 2007. "The validity domain of hybrid models and its application in process optimization." *Journal of Chemical Engineering and Processing* 46 (11): 1054–1066.

Khor, M., A.M. Mendes-Tatsis, and G.F. Hewitt. 1997. "One-dimensional modelling of phase holdups in three-phase stratified flow." *International Journal of Multiphase Flow* 23 (5): 885–897.

Kowalski, J. 1987. "Wall and Interfacial Shear Stress in Stratified Flow in Horizontal Pipe." *AIChE Journal* 33 (2): 274–281.

Liejin, G., B. Bofeng, W. Xin, and G. Hanyang. 2009. "Online recognition of the multiphase flow regime and study of slug flow in pipeline - The 6th International Symposium on Measurement Techniques for Multiphase Flow." In *The 6th International Symposium on Measurement Techniques for Multiphase Flows.* Tokyo, Japan. doi:10.1088/1742-6596/147/1/012047.

Liu, Y-P., H. Zhang, S-H. Wang, and J. Wang. 2008. "Gas-liquid interfacial friction factor for the transition from stratified to slug Flow." *Microgravity Science Technology* 20 (3–4): 299–305.

Majeed, G. 1996. "Liquid holdup in horizontal two-phase gas-liquid flow." *Journal od Petroleum Science and Engineering* 15 (2–4): 271–280.

Marcano, R., X.T. Chen, C. Sarica, and J.P. Bril. 1998. "A Study of Slug Characteristics for Two-Phase Horizontal Flow." In *International Petroleum Conference and Exhibition of Mexico.* Villahermosa, Mexico. doi:10.2118/39856-MS.

Maugeri, L. 2012. *Oil:The Next Revolution-The unprecedented upsurge of oil production capacity and what it means for the world.* Discussion Paper. Belfer Center for Science and International Affairs, Harvard Kennedy School.

Morris, Max D. 1991. "Factorial Sampling Plans for Preliminary Computational Experiments." *Technometrics* 33 (2): 161–173.

Newton, C., M. Bhardwaj, and M. Behnia. 1992. "Effects of pipe diameter on pressure drop calculations in horizontal two-phase flow." In *11th Australasian Fluid Mechanics Conference,* 627–630. University of Tasmania, Hobart, Australia.

Ouyang, L., and K. Aziz. 1996. "Development of New Wall Friction Factor and Interfacial Friction Factor Correlations for Gas-Liquid Stratified Flow in Wells and Pipelines." In *SPE Western Regional Meeting,* 285–295. Anchorage, Alaska. doi:10.2118/35679-MS.

Petalas, N., and K. Aziz. 2000. "A Mechanistic Model for Multiphase Flow in Pipes." *Journal of Canadian Petroleum Technology* 39 (6): 43–55.

Psichogios, D.C., and L.H. Ungar. 1992. "A Hybrid Neural Network-First Principles Approach to Process Modelling." *AIChE Journal* 38 (10): 1499–1511.

Roumazeilles, P. 1994. "An experimental study of downward slug flow in inclined pipes." Master's thesis, University of Tulsa, U.S.A.

Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. 2008. *Global Sensitivity Analysis. The Primer.* Wiley.

Sargent, R.W.H., and A.W. Westerberg. 1964. "Speed-Up in chemical engineering design." *Transactions of the Institution of Chemical Engineers,* no. 42:190–197.

Sarica, C., H-Q Zhang, and R. Wilkens. 2011. "Sensitivity of slug flow mechanistic models on slug length." *Journal of Energy Resources Technology* 133 (4): 043001–1–043001–6.

Schubert, J., R. Simutis, M. Dors, I. Havlik, and A. Lubbert. 1994. "Hybrid Modelling of Yeast Production Processes - a combination of a Priori Knowledge on Different Levels of Sophistication." *Journal of Chemical Engineering Technology* 17 (1): 10–20.

Shoham, O., and Y. Taitel. 1984. "Stratified turbulent-turbulent gas-liquid flow in horizontal and inclined pipes." *AIChE Journal* 30 (3): 377–385.

Shoham, O., and J.J. Xiao. 1991. "Evaluation of Interfacial Friction Factor Prediction Methods for Gas-Liquid Stratified Flow." In *SPE Annual Technical Conference and Exhibition,* 53–64. Dallas, Texas, U.S.A. doi:10.2118/22765-MS.

Shoham, Ovadia. 2006. *Mechanistic Modeling of Gas-Liquid Two-Phase Flow in Pipes.* Society of Petroleum Engineers.

Spedding, P.L., E. Benard, and G.F. Donnelly. 2006. "Prediction of Multiphase Horizontal Pipe Flow: A Reassessment, Parts I and II." *Developments in Chemical Engineering and Mineral Processing* 14 (3/4): 567–584.

Spedding, P.L., G.S. Cole, and G.F. Donnelly. 2004. "Friction factors in two phase horizontal pipe flow." *Developments in Chemical Engineering and Mineral Processing* 12 (1/2): 179–188.

Steward, D. 1965. "Partitioning and tearing systems of equations." *SIAM Journal on Numerical Analysis* 2 (2): 345–365.

Taitel, Y., D. Barnea, and A.E. Dukler. 1980. "Modelling Flow Pattern Transitions for Steady Upward Gas-Liquid Flow in Vertical Tubes." *AIChE Journal* 26 (3): 345–354.

Taitel, Y., and A.E. Dukler. 1976. "A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow." *AIChE Journal* 22 (1): 47–55.

Thompson, M.L., and M.A. Kramer. 1994. "Modeling chemical processes using prior knowledge and neural networks." *AIChE Journal* 40 (8): 1328–1340.

Tulsa University, Fluid Flow Projects. 2008. *Final Report - Development of Next Generation Multiphase Pipe Flow Prediction Tools.* Technical report. Office of Fossil Energy.

Yang, J-D. 1996. "A study of intermittent flow in downward inclined pipes." PhD diss., University of Tulsa.

Zhang, H-Q., and C. Sarica. 2006. "Unified modeling of gas/oil/water pipe flow-Basic approaches and preliminary validation." *SPE Projects, Facilities Construction* 1 (2): 1–7.

Zhang, H-Q., M.X. Li, C. Sarica, and T. Yu. 2010. "A mechanistic model for gas/liquid flow in upward vertical annuli." *SPE Production Operations* 25 (3): 285–295.

Zhang, H-Q., Q. Wang, C. Sarica, and C. Brill. 2003a. "A unified mechanistic model for slug liquid holdup and transition between slug and dispersed bubble flows." *International Journal of Multiphase Flow* 29 (1): 97–107.

Zhang, H-Q., A-S. Eissa, S. Jayawardena, C.L. Redus, and J.P. Bril. 2003b. "Modeling of slug dissipation and generation in a hilly terrain pipeline." *Journal of Energy Resources Technology* 125 (3): 161–168.

Zhang, H-Q., S. Jayawardena, C. Reddus, and J. Brill. 2000. "Slug dynamics in gas-liquid pipe flow." *Transactions of the ASME, Journal of Energy Resources Technology* 122 (1): 14–21.

Zhang, H-Q., Q. Wang, C. Sarica, and J. Brill. 2003c. "Unified model for gas-liquid pipe flow via slug dynamics-Part 1: Model development." *Transactions of the ASME, Journal of Energy Resources Technology* 125 (4): 266–273.

Appendix G

Afterword

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.