EVALUATION AND EXPANSION OF OBSERVABLE DYNAMIC FROTH FLOTATION MODELS FOR CONTROL

by

Jaco-Louis Venter

Submitted in partial fulfillment of the requirements for the degree
Master of Engineering (Electronic Engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

February 2023

SUMMARY

EVALUATION AND EXPANSION OF OBSERVABLE DYNAMIC FROTH FLOTATION MODELS FOR CONTROL

by

Jaco-Louis Venter

Supervisor:	Prof. JD le Roux
Co-supervisor:	Prof. IK Craig
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Master of Engineering (Electronic Engineering)
Keywords:	Dynamic model validation, froth flotation, moving horizon estimator

This work builds on existing observable dynamic models of froth flotation circuits, aimed at on-line parameter estimation and model-based control. The models are analysed and two main limitations are identified and addressed: the lack of explicit modelling of reagent effects and the need for dynamic validation on large-scale industrial plant data.

The feasibility of expanding a froth flotation model to include reagent effects is investigated. A Sobol sensitivity analysis is used to identify the crucial parameters. The model is expanded with two different reagent effect models. Both expansions include mass balance models of the frother concentration in each cell. The first model expands an empirical parameter in the air recovery model, related to the froth height at which peak air recovery (PAR) is achieved, as a linear function of frother concentration. The second model adds a linear frother concentration term to the existing air recovery model to modify the steady-state air recovery directly. Observability analyses of the expanded models show that all states and the important time-varying model parameters are observable (and identifiable) from the available on-line measurements. Most importantly, the frother concentrations are shown to be observable without concentration measurements. Simulations of the model expansions show that the second model can

qualitatively predict the impact of increased frother dosage on air recovery, grade and recovery, while the first model can only predict the correct effect under certain conditions.

The implementation of a Moving Horizon Estimator (MHE) based on the model (excluding reagent effects) on data from an industrial rougher bank is investigated with the aim of validating the model and parameter estimation approach. The available plant data and its limitations are discussed and additional model analysis is conducted. An expanded observability analysis of the model identifies groups of parameters for which identifiability is linked. It is shown that without on-line compositional measurements only a reduced model that lumps all recovery mechanisms into a single empirical equation is observable.

The reduced model is used to develop the MHE which is implemented on data from the Mogalakwena North Concentrator (MNC) historian. The state and parameter estimates are then used to evaluate the model prediction accuracy over a shifting control horizon, as would be done in model predictive control (MPC). Estimation results show that there are substantial amounts of unmodelled dynamics and/or disturbances. Parameter estimates compensate somewhat, but the model predictions are only accurate over some sections of the data.

The lack of on-line compositional measurements as well as uncertainty regarding the validity of calculated measurements and assumptions prevented a fair evaluation of the full potential of the model, but served to highlight drawbacks and challenges that will need to be addressed in future work.

AKNOWLEDGEMENTS

I would like to thank Daniel Rabopape and MacDonald Molomo of Anglo American Platinum for making the industrial data used in this study available and their unique process knowledge. Lidia Auret and Adriaan Haasbroek from StoneThree, thank you for your continued support and interest in this project and thank you for inspiring us young engineers.

I would also like to thank the South African Council for Automation and Control (SACAC) and the IFAC Foundation for the funding that made my participation in 19th Symposium on Control, Optimization and Automation in Mining, Mineral and Metal Processing (MMM 2022) in Montreal possible.

To my supervisors, Prof. Derik le Roux and Prof. Ian Craig, thank you for your invaluable guidance and support over the course of this work and all the opportunities you have provided.

Finally, to my parents, thank you for always being there, your support, patience and love. Without you, none of this would have been possible.

LIST OF ABBREVIATIONS

EKF	extended Kalman filter
HSL	Harwell Subroutine Library
LOIC	linear observability-identifiability condition
MHE	moving horizon estimator
MPC	model predictive control
MNC	Mogalakwena North Concentrator
MV	manipulated variable
NOIC	non-linear observability-identifiability condition
PAR	peak air recovery
PF	particle filter
PUP	possibly unidentifiable parameter
PUX	possibly unidentifiable state
SCADA	supervisory control and data acquisition
SISO	single-input single-output
SVM	support vector machine
UKF	unscented Kalman filter
XRF	X-ray fluorescence

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1
1.1 BACKGROUND 1
1.2 PROBLEM STATEMENT 3
1.2.1 Context of the problem 3
1.2.2 Research gap 4
1.3 RESEARCH OBJECTIVES AND QUESTIONS 5
1.4 APPROACH 5
1.5 RESEARCH GOALS 6
1.6 RESEARCH CONTRIBUTION 6
1.7 RESEARCH OUTPUTS 7
1.8 OVERVIEW OF STUDY 7

CHAPTER 2 LITERATURE STUDY 9
2.1 CHAPTER OVERVIEW 9
2.2 FLOTATION BACKGROUND 9
2.2.1 Basic phenomena 9
2.2.2 Reagents 12
2.3 MODELLING 14
2.3.1 Model requirements 14
2.3.2 Modelling approaches 16
2.3.3 Models for control: validation and implementation 18
2.3.4 Reagent models in literature 22
2.4 STATE ESTIMATION IN FLOTATION 26
2.5 CONCLUDING REMARKS 29

CHAPTER 3 METHODS 30

3.1 CHAPTER OVERVIEW 30
3.2 SENSITIVITY ANALYSIS THEORY 30
3.3 OBSERVABILITY ANALYSIS THEORY 31
3.3.1 Non-linear observability condition 31
3.3.2 Input excitation 33
3.3.3 Practical implementation 34
3.3.4 Identifying unobservable variables 34
3.4 MOVING HORIZON ESTIMATOR 35
3.5 PERFORMANCE METRICS 36
3.6 CONCLUDING REMARKS 38

CHAPTER 4 FLOTATION MODELLING AND MODEL ANALYSIS 39
4.1 CHAPTER OVERVIEW 39
4.2 ORIGINAL MODEL 39
4.2.1 Volume balances 39
4.2.2 Water recovery 42
4.2.3 Mineral recovery 42
4.2.4 Froth model 45
4.2.5 Measurements 45
4.3 UPDATED MODEL 46
4.4 MODEL CAUSALITY 47
4.5 OBSERVABILITY 49
4.5.1 Existing observability analyses 49
4.5.2 Additional analysis and results 50
4.6 CONCLUDING REMARKS 54

CHAPTER 5 FROTHER MODEL FEASIBILITY 56
5.1 CHAPTER OVERVIEW 56
5.2 IMPACT OF REAGENTS 56
5.3 SENSITIVITY ANALYSIS RESULTS 57
5.4 PROPOSED MODEL EXPANSIONS 58
5.4.1 Model expansion 1 (ME1) 59
5.4.2 Model expansion 2 (ME2) 60
5.5 SIMULATION 61

5.5.1 Simulation set-up . . . 61
5.5.2 Model expansion 1 (ME1) results . . . 63
5.5.3 Model expansion 2 (ME2) results . . . 64
5.6 CONCLUDING REMARKS . . . 66

CHAPTER 6 MHE IMPLEMENTATION ON REAL WORLD DATA . . . 67
6.1 CHAPTER OVERVIEW . . . 67
6.2 PLANT AND MEASUREMENT DESCRIPTION . . . 68
6.2.1 Plant overview . . . 68
6.2.2 Available dynamic data . . . 69
6.2.3 Static data . . . 76
6.3 REDUCED MODEL . . . 76
6.3.1 Model summary . . . 76
6.3.2 Observability results . . . 77
6.4 PROCESS PARAMETER FITS . . . 79
6.4.1 Air recovery model parameters . . . 79
6.4.2 Bubble size model parameters . . . 81
6.4.3 Volume balance parameters . . . 82
6.5 MHE DESIGN . . . 83
6.5.1 Basic implementation . . . 83
6.5.2 Sampling time and horizon . . . 84
6.5.3 Parameters and noise . . . 85
6.5.4 Initialisation . . . 88
6.5.5 Tuning . . . 89
6.6 ESTIMATION AND PREDICTION PERFORMANCE . . . 90
6.6.1 Estimation results . . . 90
6.6.2 Prediction results . . . 95
6.7 CONCLUDING REMARKS . . . 99

CHAPTER 7 CONCLUSION . . . 101
7.1 GENERAL DISCUSSION . . . 101
7.2 RECOMMENDATIONS FOR FUTURE WORK . . . 103

REFERENCES . . . 104

ADDENDUM A REAGENT MODEL EXPANSION REVISIT 114

LIST OF FIGURES

1.1 Example flotation circuit. 2
1.2 Dissertation structure diagram showing basic contents of each chapter and the interrelationships between chapters. 8

2.1 General MPC layout 16

3.1 Parity plot of the prediction results for the α_2 parameter fits from Section 6.4, reproduced here for ease of reference. Each group of data-points contain the same number of observations. 37

4.1 Diagram of forced air flotation cell indicating the measured variables relevant to a cell and cell dimensions. 40
4.2 Model information flow diagram. This shows the interaction between the various parameters and variables in the Oosthuizen (2023) model. 48
4.3 Second rougher bank at Mogalakwena North Concentrator. Variables that are applicable to multiple cells are indicated with a subscript k. 51

5.1 Four-cell rougher circuit, described in Hadler et al. (2010), adapted from Oosthuizen (2023). Green: manipulated variables. Blue: measured variables. Orange: variables not measured. Purple: intermediate variables. 62
5.2 Frother concentration in all cells with changes in frother addition rate. 62
5.3 Calculated $K_{0,\alpha_{h_f},4}$ compared to the froth heights over the different simulation runs. Note that $K_{0,\alpha_{h_f},4}$ crosses the froth height in Run 2, resulting in an opposite air recovery gradient. 64
5.4 Cell 4 results for the three simulations of ME1. Results show the incorrect inversion of frother effect. 65

5.5 Cell 4 results for the three simulation runs of ME2. Results show the model responding as expected. 65

6.1 Second rougher bank at Mogalakwena North Concentrator. Variables that are applicable to multiple cells are indicated with a subscript k. 68
6.2 Available dynamic data for FT1. As the densities are not used for the MHE implementation, they were omitted. The data has been normalised for commercial reasons. Step tests of J_{g_1} were conducted during 13:00–17:00 , 25 June 2019. 70
6.3 Available dynamic data for FT2. As the densities are not used for the MHE implementation, they were omitted. The data has been normalised for commercial reasons. . . . 71
6.4 SCADA to J_{g_k} conversion options. All J_{g_k} values, are the mean J_{g_k} over the period 25–26 June 2019. The values shown here have been offset from the actual values for commercial reasons. 73
6.5 Distribution of the raw mean D_{BF_k} measurements over a normal 8 h shift (00:00–08:00, 26 June 2019), sampled every 2 min. Each box represents 240 data-points. Outliers are shown as black circles. 75
6.6 Parity plots of the α_k model predictions over 8 h periods. Note how the "Modified" parameter set captures more of the variance than the 01h–09h predictions. 81
6.7 Top of froth bubble size parameter fit predictions. Parameters to be used as initial guesses for the MHE. 82
6.8 Poles of the updated and reduced models implemented on the two available datasets. 84
6.9 Extract of L_k estimation results for the three different MHEs. Note that MHE2 and MHE3 are virtually indistinguishable here. The lines were offset from their actual values to protect sensitive information. 86
6.10 Extract of α_k estimation results for the three MHEs. 87
6.11 Estimated parameter values for MHE3 over the entire 25–26 June dataset. Note the much more gradual parameter changes after 21:00, which represents normal operation. 90
6.12 State estimates for α_k during the J_{g_k} and frother step tests. 92
6.13 State estimates for α_k during normal operation with less input variance. 93
6.14 State estimates for D_{BF_k} during the step tests. Similar results were obtained for the other sections of data. 94
6.15 FT1 state predictions initialised every 5 min in the step test time period. L_1 and J_{g_1} were offset from their actual values to protect sensitive operational information. . . . 96

6.16 FT2 state predictions initialised every 5 min in the step test time period. L_2 and J_{g_2} were offset from their actual values to protect sensitive operational information. . . . 97

6.17 L_{H_q} predictions during the step test period, initialised every 5 min and the input Q_{H_1} along with the predicted $\sum Q_{C_k}$. 99

LIST OF TABLES

3.1 Prediction errors for α_2 parameter fits. 38

4.1 Model states and intermediate variables. 41
4.2 Model inputs and exogenous variables. 42
4.3 Model parameters. 43
4.4 Number of variables. 53
4.5 Observability analysis results. 53

5.1 Sensitivity indexes (in percentage). 58
5.2 ME1 Linear observability results. 60
5.3 ME2 linear observability results. 61
5.4 Initial values of model variables for the flotation cells and hopper (Oosthuizen et al., 2021). 63
5.5 Nominal values of model variables for the flotation bank. 63

6.1 Reduced model equations. 77
6.2 Reduced number of variables. 78
6.3 Reduced model observability results. 78
6.4 Reduced model with steady-state D_{BF_k} model observability results. 79
6.5 Fitted parameters for α_k model at different times. 80
6.6 Fitted parameters for D_{BF_k} model at different times. 81
6.7 Volume balance parameter values. 83
6.8 L_k estimation performance metrics for three MHE designs. 87
6.9 α_k estimation performance metrics for three MHE designs. 88
6.10 MHE tuning parameters. 89
6.11 Estimation and prediction error metrics. 92

A.1 Number of variables. 115
A.2 Observability analysis results for ME3. 115

CHAPTER 1 INTRODUCTION

1.1 BACKGROUND

Froth flotation is the separation of particles with different surface properties (specifically hydrophobic from hydrophilic) achieved by bubbling air through a slurry (also called pulp) in a vessel (cell or column). Hydrophobic particles attach to the bubbles and float to the top forming a froth, while the hydrophilic particles remain in the pulp. This process is called *true flotation*. *Entrainment* is the other main recovery mechanism by which both desired and unwanted (gangue) minerals enter the froth. The froth is removed as concentrate, while the pulp at the bottom of the vessel is removed as tailings. Typically, the valuable mineral is recovered in the concentrate while the gangue is left in the tailings (Wills and Finch, 2015).

With the main goal of concentrating a valuable target mineral in a low grade ore body to economically viable levels, the performance of flotation is measured as the grade and mineral recovery of the concentrate stream. *Grade* is the target mineral concentration (weight percentage), while *recovery* is the fraction of a given mineral in the feed that reports to the concentrate product (Hu, 2014). A trade-off exists between grade and recovery, as conditions that promote mineral recovery typically also increase the amount of gangue reporting to the concentrate, lowering the grade (Wills and Finch, 2015). Apart from physical operating conditions such as aeration rate, agitation and cell levels, various reagents are used to improve performance by, for example, selectively modifying the surface properties of the minerals or stabilising the froth to aid recovery.

There are two main categories of flotation process units: mechanical cells and pneumatic columns (henceforth just cells and columns respectively) (Wills and Finch, 2015). Although most literature and theory is applicable to both (Shean and Cilliers, 2011), there are some important differences to take note of.

Mechanical cells have a driven impeller that agitates the pulp and introduces and disperses air into the cell. Air is either introduced by action of the impeller alone (self-aspirated cells) or additional compressed air is added (forced-air cells). Various flow, impeller and stator configurations are available (Wills and Finch, 2015).

Flotation columns are less common than cells, as they were developed slightly later (Wills and Finch, 2015). They are, however, seeing increased use, especially in the Northern hemisphere. Again, there are pulp (collection) and froth (cleaning) zones. The slurry is fed to the collection zone and meets rising air, introduced by sparging compressed air into the bottom of the column. Wash water, meant to reduce the amount of entrained gangue, is fed in or above the froth zone. Columns are capable of achieving high-grade concentrates at very low energy expenditure because of the washing process; however, it is very water-intensive.

Since a cell or column has limited efficiency and must balance grade and recovery, flotation processes rarely consist of a single unit, but rather a network (or "circuit") of cells, columns, or cells organised into cascading banks (Hu, 2014). A bank is a group of cells (or columns), typically in series, of which the outputs are combined and which are operated to cumulatively achieve a certain grade and recovery specification. Each unit of the circuit aims to achieve a different balance between grade and recovery. Figure 1.1 shows a possible flotation circuit flow diagram.

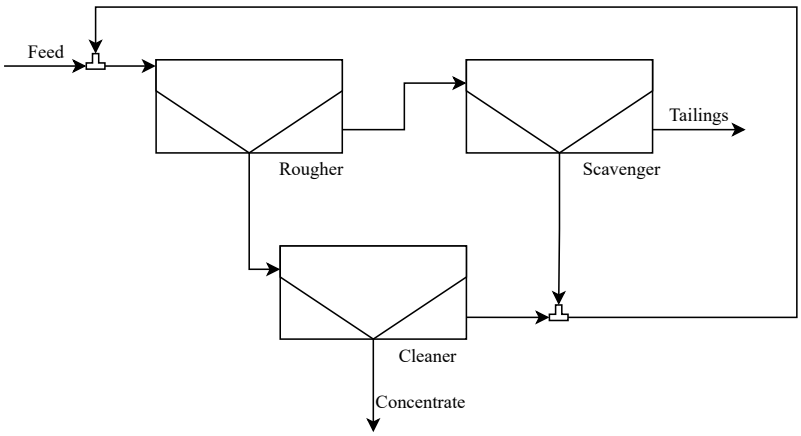

Figure 1.1. An example of a flotation circuit with multiple banks and recycles.

Incoming low-grade ore is first passed through a rougher circuit, optimised for a good balance between recovery of the easily floatable target mineral at a reasonable grade. The tailings from rougher banks still contain target mineral, albeit in much lower concentrations. A scavenger circuit treats these

rougher tailings and is optimised for maximum recovery of the target mineral in its concentrate, which is typically recycled to the rougher bank. The scavenger tailings contain very little valuable mineral and are typically discarded (Hu, 2014). Cleaner circuits are used to maximise the grade of the rougher concentrate and their concentrate is typically the ultimate product of the circuit. Cleaner tailings still contain significant amounts of the target mineral and they can also be recycled to the rougher feed.

Flotation in mineral processing is almost always preceded by a grinding circuit. This ensures that the particle size is small enough to be floatable and that the target mineral is sufficiently liberated from the gangue. It is also possible to have intermediate re-grinding steps in a flotation circuit (Hu, 2014; Wills and Finch, 2015; Gupta and Yan, 2016).

1.2 PROBLEM STATEMENT

1.2.1 Context of the problem

Froth flotation is currently one of the largest mineral separation techniques in use. Its scale means that even slight improvements in efficiency could have a large impact on the global minerals processing industry (Quintanilla, Neethling and Brito-Parada, 2021). As such, the control and optimisation of flotation circuits has received considerable attention since its development (Bascur Riquelme, 1982; Shean and Cilliers, 2011; Oosthuizen et al., 2017). Modelling has played an important role in much of this research and several models have been proposed to capture some of the fundamental mechanisms involved in the flotation process. These models are covered well in several recent literature reviews, such as Oosthuizen et al. (2017), Wang et al. (2018), Quintanilla, Neethling and Brito-Parada (2021) and several other studies on flotation (Neethling and Cilliers, 2009; Shean and Cilliers, 2011; Shean et al., 2017; Neethling et al., 2019). Some facilitate the understanding of some of the complex mechanisms involved in flotation while others allow for improved design and optimisation (Shean and Cilliers, 2011; Quintanilla, Neethling and Brito-Parada, 2021; Oosthuizen, 2023).

It is well-established that traditional PID and other single-input single-output (SISO) control methods are not sufficient to properly control flotation beyond base level stabilisation or regulatory control (Shean and Cilliers, 2011; Jovanović and Miljanović, 2015; Quintanilla, Neethling and Brito-Parada, 2021) and there is a need for more advanced control techniques, such as model predictive control (MPC), to increase the efficiency of flotation. This requires models that capture the important dynamics of the system while remaining simple enough to simulate and optimise in real time. These models

should also be easily maintainable from the limited on-line measurements, as detailed sampling campaigns are a costly and time-consuming process (Shean et al., 2018; Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021; Quintanilla, Neethling, Navia and Brito-Parada, 2021; Quintanilla, Neethling, Mesa, Navia and Brito-Parada, 2021; Steyn and Sandrock, 2021).

Of particular relevance to this study are the arguments and approach put forward by Oosthuizen et al. (2017) and Oosthuizen et al. (2021), aimed at addressing this last issue. They develop a dynamic model of a flotation circuit, combining established phenomenological models and mass balances with empirical models of the less well-understood phenomena, such as the froth dynamics. Importantly, the states and parameters of the model are shown to be observable and identifiable from the limited measurements typically available on industrial flotation plants.

Simply put, observability (or identifiability in the case of parameters) is a mathematical property of a model which means that the model states (or parameters) can be estimated from the input-output history of the system (Villaverde et al., 2019; Oosthuizen et al., 2021). This is a crucial, yet often ignored, part of almost all dynamic model-based control strategies, as full state-feedback to the controller is required as a starting point for predictions and the states are rarely all measured directly. By ensuring observability and identifiability, Oosthuizen et al. (2021) aim to estimate the model parameters in real time from the available measurements. This dissertation builds on the work of Oosthuizen et al. (2021) and Oosthuizen (2023).

The use of reagents, is another major factor in the performance of flotation. Flotation reagents (*e.g.,* xanthates) are typically very expensive, yet provide significant boosts in performance. Therefore, the optimisation and automated control of reagents, hold significant potential monetary benefit as shown by Brooks and Koorts (2017). Currently reagent addition is mostly controlled retrospectively by plant operators, based on their experience and observed plant performance (Shean and Cilliers, 2011; Xie et al., 2017), with some plants including a form of feed-forward ratio control based on mineral feed rate (Shean and Cilliers, 2011).

1.2.2 Research gap

While Oosthuizen (2023) has shown good parameter tracking and control results in simulation and fitted parameters to steady-state plant data from Hadler et al. (2010) with good correlation, the prediction accuracy of the model has yet to be validated on dynamic plant data.

Furthermore, despite the acknowledged importance and potential benefits of reagent optimisation (Brooks and Koorts, 2017), none of the recently developed dynamic models aimed at control (Shean et al., 2018; Oosthuizen et al., 2021; Quintanilla, Neethling, Navia and Brito-Parada, 2021; Steyn and Sandrock, 2021) explicitly account for varying reagent effects, the exception being Brooks and Koorts (2017). If advanced control techniques are to be developed using dynamic models, reagents effects would need to be accounted for and the lack of accepted models of reagent effects must be addressed.

1.3 RESEARCH OBJECTIVES AND QUESTIONS

Addressing the research gap discussed in Section 1.2.2 requires addressing several questions. However, the scope of this study will be limited to the questions listed below.

- Is expanding the model to include reagent effects feasible?
- Can a non-linear state estimator based on the Oosthuizen (2023) model provide realistic state and parameter estimates from real industrial data?
- Can the updated model predict future states and outputs well enough for use in model-based feedback control?
- What are the current limitations of the model?

Answering these questions will require the following objectives to be reached:

- Show that the model can be expanded to include reagent effects while maintaining observability and identifiability.
- Show that the implementation of the state and parameter estimation is theoretically possible with the information available at a particular flotation plant.
- Implement the state estimator using dynamic historical data from the plant that would be available as real-time measurements.
- Evaluate the efficacy of the estimation and k-step ahead prediction of the model.

1.4 APPROACH

A literature review is conducted to determine the current state of validated dynamic models aimed at control, how state estimation is used in flotation and how reagent effects are accounted for in flotation models.

The observability analysis of the Oosthuizen (2023) model is re-evaluated and additional analysis of the model is conducted. The observability analysis is expanded to determine the impact of missing measurements and identify possible challenges in the state estimation implementation.

The feasibility of reagent model expansion is tested on the Oosthuizen et al. (2021) version of the model, as reliable parameters based on a complete set of measurements are already available. First the parameters likely to be affected by changes in reagent conditions are identified and a Sobol sensitivity analysis is performed to determine the dominant parameters. Based on the sensitivity analysis, two possible model expansions are proposed and analysed for observability. The realism of these models is evaluated through simulations.

Dynamic data is obtained from an industrial flotation circuit. A reduced version of the model is developed based on limitations in the available data and used to develop a moving horizon estimator (MHE). The MHE is implemented on historical plant data. The state and parameter estimates are evaluated by testing the prediction accuracy of the model over typical control horizons over 24 h of operational data.

1.5 RESEARCH GOALS

The aim of this study is twofold:

- to determine the feasibility of expanding the existing model with observable reagent effect models
- to determine the potential efficacy and feasibility of implementing the Oosthuizen (2023) model on a plant.

1.6 RESEARCH CONTRIBUTION

It is demonstrated that it is possible to expand the older Oosthuizen et al. (2021) model to include reagent effects such that observability and parameter identifiability are maintained. Importantly, the model expansion shows that frother concentrations based on mass balance differential equations are observable without the need for any measurements not already available on modern flotation circuits. While practical observability is not investigated, this could aid in the development of soft sensors for reagent concentrations.

This study expands on the work of Oosthuizen (2023). The expanded observability analysis demonstrates how to identify possibly unidentifiable parameter groups and link them to measurements. The

analysis also highlights the source of possible challenges in the state and parameter estimation. Shortcomings in the available measurements, the model and the estimation algorithm are identified. This may serve as guidance for future investigations into this modelling and estimation approach.

1.7 RESEARCH OUTPUTS

The feasibility study of expanding the existing model to include reagent effect while remaining observable in Chapter 5 has been published in:

- Venter, J., Le Roux, J.D., Craig, I.K., 2022. Observable dynamic models of reagent effects for model-based froth flotation control. *IFAC-PapersOnLine* **55**(21): 102–107.

1.8 OVERVIEW OF STUDY

A literature review covering the basics of froth flotation, the current state of modelling for control and how reagents are incorporated (or not) in flotation models is given in Chapter 2. Chapter 3 contains the theoretical background for the analysis and state estimation methods used throughout the rest of the dissertation. The Oosthuizen et al. (2021) and Oosthuizen (2023) models used in this work are described and analysed in Chapter 4. Chapter 5 describes a feasibility study of expanding the model to include reagent effects while remaining observable. In Chapter 6 the model and state estimation performance are evaluated using dynamic data from an industrial flotation circuit. A summary of the study, concluding remarks and recommendations for future work are given in Chapter 7. Figure 1.2 gives an overview of the different chapters and how their topics relate to other chapters.

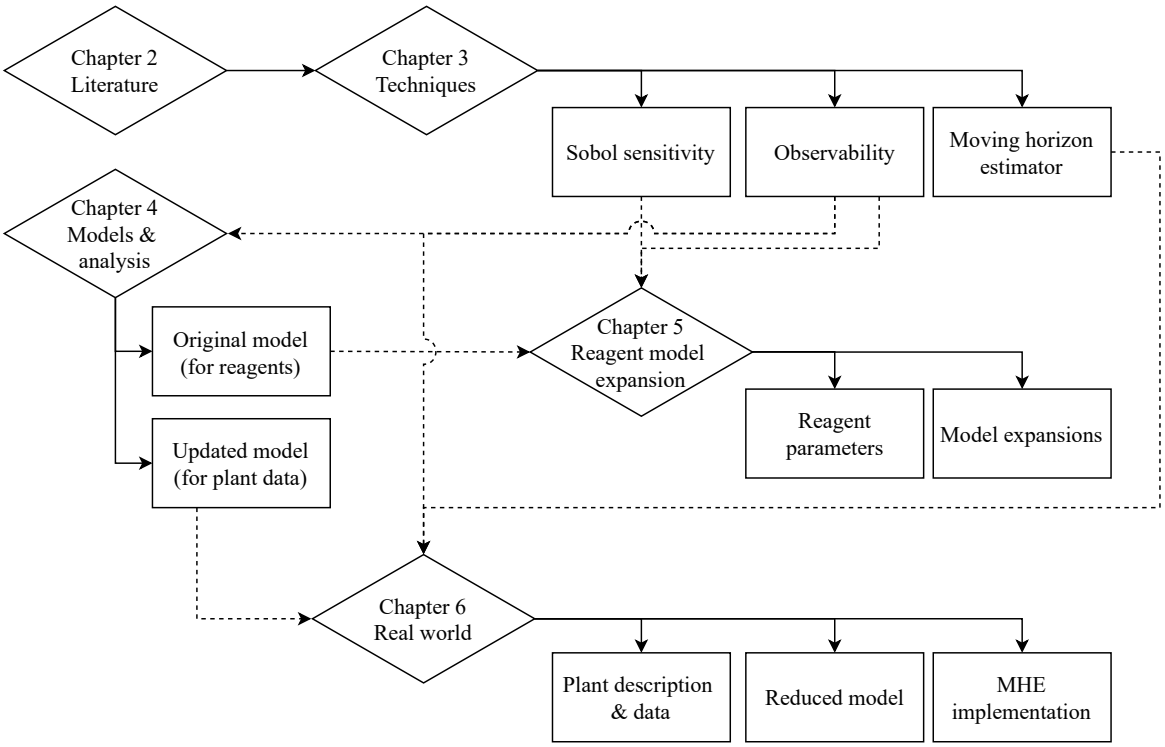

Figure 1.2. Dissertation structure diagram showing basic contents of each chapter and the interrelationships between chapters.

CHAPTER 2 LITERATURE STUDY

2.1 CHAPTER OVERVIEW

As mentioned in Chapter 1, this study aims to evaluate the possibility of expanding the Oosthuizen (2023) model with reagent models as well as investigate the validity of the model and proposed joint state-parameter estimation using dynamic industrial data. This requires some background in the fundamentals of flotation and flotation reagents (Section 2.2), as well as the typical modelling approaches in flotation (Section 2.3). Focus will be placed specifically on the validation of models aimed at control (Section 2.3.3) and the modelling of reagent effects (Section 2.3.4). Finally state estimation in flotation is discussed in Section 2.4.

2.2 FLOTATION BACKGROUND

2.2.1 Basic phenomena

With more than 100 variables affecting a typical flotation process (Jovanović and Miljanović, 2015), a detailed discussion of all possible mechanisms and parameters is not practical for the current work. However, there appears to be consensus in literature regarding which phenomena are most important to consider and several authors (Shean and Cilliers, 2011; Oosthuizen et al., 2017; Quintanilla, Neethling and Brito-Parada, 2021; Oosthuizen, 2023) have identified the following variables as crucial, specifically for control applications:

- Flow-rates, densities and concentrations (grades) of the feed, concentrate and tailings
- Reagent concentrations and addition rates
- Air flow-rate and recovery
- Froth properties (loading, bubble size distribution, water content *etc.*)
- Froth depth
- Pulp levels

- Wash water flow-rate (columns)
- Particle properties (size, shape, degree of liberation *etc.*)
- Electrochemical potentials

Based on the modelling work by various authors (Neethling and Cilliers, 2003; Wang et al., 2015; Shean et al., 2018; Wang et al., 2018; Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021) and the list of key variables above, the following important phenomena were identified:

- True flotation (particle-bubble attachment in pulp)
- Entrainment
- Froth stability (bursting, coalescence)

This corresponds to the phenomena identified by Oosthuizen et al. (2017). While there are other phenomena that influence the flotation process, the same three fundamental phenomena were selected as they appear to have the greatest impact on grade and recovery and receive the most research attention in studies that consider the overall performance of flotation circuits.

2.2.1.1 True flotation

The attachment of the hydrophobic particles to the air bubbles and their subsequent transportation to the froth, also known as true flotation, is selective toward the surface properties of the different minerals and is the primary mechanism by which the target mineral is recovered in the froth (Wills and Finch, 2015). However, not only the target mineral is recovered in this manner. Any particle that meets the requirements for floatability can potentially be recovered. For this reason collectors and depressants are used to modify the surface properties of the different minerals to increase the selectivity of the process.

Flotation requires bubble-particle collisions with enough contact time (induction time) for the particle to attach to the bubble, forming a bubble-particle aggregate. For hydrophobic particles the required time is much shorter than for hydrophilic particles (Wills and Finch, 2015). The bubble surface area flux in the pulp, S_b, which in turn depends on bubble size, D_{BP}, and superficial gas velocity, J_g, as well as the gas hold-up (fraction of the total pulp volume that is gas), determines the probability of bubble-particle collisions and therefore also plays an important role in the flotation rate (Maldonado, Desbiens, Villar, Poulin and Riquelme, 2010). A further requirement for the bubble-particle aggregate to float is that the bulk density of the aggregate must be lower than the density of the surrounding fluid (Wills and Finch, 2015). This limits the floatability of large particles and also means that bubbles have a limited carrying capacity.

2.2.1.2 Entrainment and entrapment

Wills and Finch (2015) mention two other mechanisms that can cause particles to end up in the froth, both of which are not selective towards either valuable mineral or gangue: entrainment of particles in the upward moving water between the bubbles and entrapment (also called aggregation) between particles attached to the froth bubbles. However, from the literature studied for the current work it appears that these two mechanisms are often lumped together under the term "entrainment" (Neethling and Cilliers, 2009; Alves Dos Santos et al., 2014; McFadzean et al., 2016; Oosthuizen et al., 2021) or the effect of entrapment is acknowledged and accounted for in the entrainment factor model (Yianatos and Contreras, 2010; Wang et al., 2015). This lumping is due to the unlikelihood of being able to distinguish by which mechanism a given particle was recovered (Neethling and Cilliers, 2009). For the remainder of this report, the more general definition of entrainment will be used, *i.e.,* the mechanism by which any particles in the concentrate not recovered through true flotation, were recovered.

A thorough review on the factors affecting entrainment as well as the measurement and modelling of entrainment is given in Wang et al. (2015). Of these parameters the most notable are particle size and concentration, air flow-rate and froth stability. The latter is particularly important; for entrained particles to be recovered in the concentrate, they have to be entrained into the froth and then remain in the froth long enough to reach the overflow. Froth stability and structure determine how much of the entrained particles and water between the bubbles drain back into the pulp, reducing the recovery due to entrainment. In this regard, frothers play a role in the impact that entrainment can have on flotation performance (see Section 2.2.2).

2.2.1.3 Froth stability

The froth that forms on top of the pulp phase plays an important role in the overall flotation performance. By preferentially retaining hydrophobic particles in the bubble lamella, while letting entrained solids and water drain, it serves to increase the selectivity of the process as well as concentrate the target mineral into a dryer slurry (Wills and Finch, 2015). The stability of the froth is therefore a crucial parameter to consider. For example, if bubbles burst or coalesce, some of the floated mineral is dumped back into the pulp, reducing the recovery of the valuable material and if the structure does not allow for adequate drainage, the grade of the concentrate will suffer due to the increased impact of entrainment.

There are several measures of froth stability including bubble size, water recovery and air recovery.

Several studies (Hadler and Cilliers, 2009; Hadler et al., 2010, 2012) have shown that there is a peak in air recovery (PAR) with increasing aeration rate in mechanical flotation cells and that this corresponds to the highest mineral recovery (Hadler and Cilliers, 2009) in the cell bank. It was also shown that the PAR air rate increased with increasing froth depth and that deeper froths resulted in higher grades and mineral recoveries despite decreasing mass pull (Hadler et al., 2010, 2012). This phenomenon was found to be intimately linked with the stability and structure of the froth.

The particle size, surface chemistry and loading in the froth are important parameters for determining froth stability. Wiese et al. (2011) found that depressants could have a destabilising effect when used in too high concentration in platinum ore flotation. This was due to the depression of naturally floatable gangue which had previously stabilised the froth. The use of frothers typically stabilise the froth (Harris et al., 2013; McFadzean et al., 2016) and was found to counteract the effect observed in Wiese et al. (2011).

2.2.2 Reagents

As already mentioned, chemicals are commonly used to enhance the performance of the flotation process. Wills and Finch (2015) and Bulatovic (2007) categorise these chemicals into three groups by their function: collectors increase the hydrophobicity of the target mineral, frothers stabilise the froth phase and regulators are added to achieve a variety of effects.

The field of flotation reagents is extremely large with many different reagent suites (or “recipes”) available for a single ore type. As this work is more focussed on the modelling of flotation processes, specifically the effect of reagent addition, only a quick overview of reagents and their purpose is given. For a detailed discussion on the available reagents, their fundamental mechanisms and application, the reader is referred to works such as Bulatovic (2007), Wills and Finch (2015) and McFadzean et al. (2016). In this regard the Centre for Minerals Research at the University of Cape Town is involved in research into the effect of different reagents (Wiese et al., 2011; McFadzean et al., 2016; Taguta et al., 2017). There are also large contributions to flotation reagent research from researchers in China (Zhu et al., 2014; Xie et al., n.d., 2021).

2.2.2.1 Collectors

Since few minerals are naturally hydrophobic enough for effective flotation, chemicals that can selectively increase the hydrophobicity of the target mineral, called collectors, are considered the most important reagents in flotation (Wills and Finch, 2015). These reagents are added to the pulp and given

time to adsorb to the particles before being fed to the cell. This is known as conditioning. They are used in relatively small amounts to prevent increased flotation of unwanted minerals (Wills and Finch, 2015; Xie et al., 2021). However, this means seemingly small differences in concentration can lead to quite significant differences in performance.

Typical collectors are ionising organic molecules with a strong polar functional group that attaches to the mineral surface and long non-polar hydrocarbon chains that provide the hydrophobicity (Wills and Finch, 2015; Taguta et al., 2017). This increases the overall hydrophobic nature of the particle surface and reduces induction time, making the particle more likely to float. Some of the most commonly used functional groups for sulphide minerals are xanthates, dithiophosphates and dithiocarbamates (Taguta et al., 2017), with others covered in Urbina (2003) and Wills and Finch (2015). For non-sulphide minerals, fatty acids are the most prevalent collectors followed by amines (Wills and Finch, 2015). Adsorption, either chemical or physical, is a crucial mechanism for collectors and has been studied in some detail (Bicak et al., 2007; Taguta et al., 2017).

2.2.2.2 Frothers

The main purpose of frothers is to reduce the bubble size in the pulp by inhibiting coalescence and stabilise the froth by lowering the bubble burst rate (Harris et al., 2013; Wills and Finch, 2015). Smaller bubbles have larger surface area and slower rise time which increases the probability of bubble-particle collisions and improves the flotation kinetics. Typical frothers have a hydrophilic (polar) head with hydrophobic hydrocarbon tail. They act to stabilise the air-water interface of bubbles. Alcohols and poly-glycols are the most commonly used frothers (Wills and Finch, 2015). Frothers are usually used as blends and McFadzean et al. (2016) has shown that the frother mixtures investigated outperform their individual components.

There is a limit to the bubble size reduction that can be achieved by frother addition, known as the critical coalescence concentration (CCC) (Harris et al., 2013; Wills and Finch, 2015), beyond which the bubble size is determined by the air dispersion mechanism. While further frother addition does not reduce the bubble size, it can still increase froth stability. Ideally, enough frother is added to the process to surpass the CCC, as this ensures the highest possible bubble surface area for true flotation (Harris et al., 2013). It is however, undesirable to add frother in much higher concentrations than the CCC as this is not only more costly, but can lead to increased water recovery in the froth which in turn increases entrainment (Wills and Finch, 2015).

2.2.2.3 Regulators

Wills and Finch (2015) classify reagents that do not fall under either collectors or frothers as regulators (or "modifiers" in the case of Bulatovic (2007)). Some of the most common regulators include depressants which increase the inherent hydrophilicity of gangue and activators which increase the effectiveness of collectors. Dispersants prevent aggregation of particles which among other effects, reduces entrainment, improving the concentrate grade. Many depressants also act as dispersants, which further enhance their effectiveness (Wills and Finch, 2015). Examples of studies that consider the effects of regulators, specifically depressants, on flotation performance are Wiese et al. (2005), Bicak et al. (2007), Wiese et al. (2011) and Harris et al. (2013).

2.3 MODELLING

The available models of froth flotation are extremely varied in scope, goal and approach. A basic discussion of the requirements and general approaches is given. Several recent reviews (Oosthuizen et al., 2017; Wang et al., 2018; Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021) cover the models in detail. Rather than repeating the model reviews already available, the following sections focus on identifying the model validation methods in these studies (Section 2.3.3) and aim to illustrate the gap in reagent effect models (Section 2.3.4).

2.3.1 Model requirements

The ultimate goal of any model should inform the approach followed. This includes the choice of mathematical structure, simplifying assumptions, considered variables and the required level of accuracy. As in most other industrial process fields, the field of flotation modelling has two broad research angles that require the use of models (Quintanilla, Neethling and Brito-Parada, 2021): simulation and control implementation.

2.3.1.1 Simulation

Models aimed at accurate process simulation tend to be very detailed and focus on specific mechanisms. These models are mostly used to improve the fundamental understanding of the system in question (Wang et al., 2015, 2018), during process design (Hu et al., 2013), or to simulate existing systems to investigate and optimise operating conditions and practices (Steyn and Sandrock, 2021). They tend to have many parameters that have to be fitted to experimental data or, in the case of design, chosen by the user and additional complexity is often accepted for gains in accuracy.

2.3.1.2 Control

Detailed simulation models still have an important role to play in control research. The development of control strategies, operating practices and optimal operating conditions are aided with detailed process simulations. This is especially important for new systems where control should play an integral role in informing design decisions (Seborg et al., 2011).

The models for model-based control have a specific set of requirements (Shean and Cilliers, 2011; Jovanović and Miljanović, 2015; Quintanilla, Neethling and Brito-Parada, 2021). The development of these advanced control strategies require models that capture the important dynamics and interactions of the system with reasonable accuracy, while remaining robust for long periods of implementation (Quintanilla, Neethling and Brito-Parada, 2021). This last point is emphasised in recent modelling work and reviews (Shean et al., 2018; Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021; Steyn and Sandrock, 2021), which point out that most existing models require detailed, dedicated sampling campaigns to fit and update the parameters – a costly and time-consuming process. This means that with the usually limited on-line measurements, long term continuous implementations using these models will fail due to model-plant mismatch introduced over time or significantly different operating conditions (Oosthuizen et al., 2021; Steyn and Sandrock, 2021).

With model-based control there is also the need for state-feedback (Oosthuizen et al., 2021). Figure 2.1 shows a typical MPC feedback loop. At each time step, the current system states (x) are used as initial conditions for the plant model (here shown as a state-space model also dependent on a set of parameters (p)). An optimisation algorithm varies the system inputs (u) to minimise some objective function, which typically involves comparing the predicted outputs (y) to a desired trajectory (y_{SP}). However, in most processes (and certainly in flotation), the model states are rarely *all* measured directly ($x \not\subset y$) (Shean and Cilliers, 2011; Oosthuizen et al., 2017; Oosthuizen, 2023). The unmeasured states must therefore be estimated if MPC is to be used. To obtain a state estimate ($\tilde{x}$), the model states must be observable given the available measurements.

Observability is discussed in much more detail in later chapters. For now, let it suffice to say that almost all implementations of model-based control will require the model to be observable. Granted, many models will be observable by default. However, while it may be considered by researchers, it is rarely explicitly addressed in flotation literature. Some exceptions are discussed in Section 2.4.

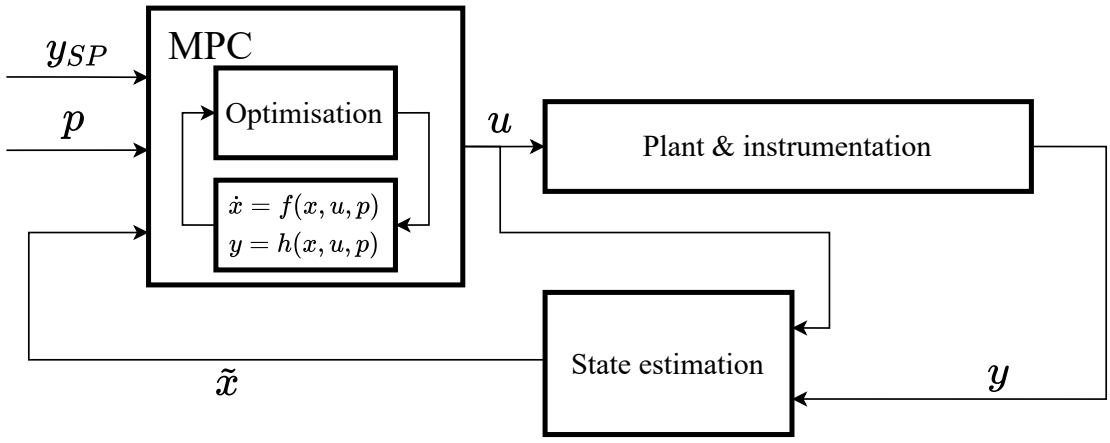

Figure 2.1. Typical MPC flow diagram with state-space model and state estimation.

2.3.2 Modelling approaches

Several modelling approaches can be used to achieve the different goals in simulation and control discussed in Section 2.3.1. Beyond classification based on the purpose of the model, Quintanilla, Neethling and Brito-Parada (2021) divides flotation modelling into two overarching approaches: empirical (or data-driven) models and phenomenological models. They also suggest "hybrid" or "switching" models as a third category, referring to the combined use of continuous and discrete variables, with the latter being logic rules switching between different model modes. For the purpose of this work, this category as a whole will not be considered in detail, as it is mostly used to switch between models that fall under the other two categories. Putz and Cipriano (2015) is a good example of the use of hybrid models in flotation control strategies.

As noted by Quintanilla, Neethling and Brito-Parada (2021), the categorisation of a particular model is somewhat arbitrary as almost all phenomenologically based models include simplifications using fitted parameters and some empirical models are influenced in structure by mechanistic considerations. The following categorisation should, therefore, only serve to outline the types of approaches that are observed in existing models, rather than representing a set of mutually exclusive categories into which models can be divided. Other model properties that can be used for classification are (Quintanilla, Neethling and Brito-Parada, 2021): steady-state or dynamic, deterministic or stochastic (random variables and probability distributions), linear or non-linear, mathematical equations or fuzzy logic based and micro (single phenomena) or macro-scale (linking micro-scale models to model larger parts of a system and their interaction).

2.3.2.1 Empirical models

Empirical models choose a set of variables and some model structure relating these variables with adjustable parameters that modify the relationships. This includes the use of transfer function approximations for the system when the transfer function is not derived from fundamental differential equations. The parameters are fitted (or "trained") based on data gathered from the system under consideration, typically using algorithms such as least square minimisation. Models using neural networks and machine-vision are also classified under empirical models.

Empirical models are easy to build if data are available and can work well once fitted, but typically fail to predict performance if some part of the process conditions change that is not included in the set of selected variables or falls outside a certain calibration range of these variables (Oosthuizen et al., 2017; Quintanilla, Neethling and Brito-Parada, 2021). This makes them quite attractive for pilot scale projects (Oosthuizen et al., 2017), but less suitable for more general process models when several conditions vary over large ranges. Another major drawback is that empirical models have to be re-calibrated for each new plant or operating range (Quintanilla, Neethling and Brito-Parada, 2021).

One field in flotation where data-driven models, such as convolutional neural networks (CNN), has found wide-spread implementation is image-based instrumentation or machine vision (Aldrich et al., 2010; Shean and Cilliers, 2011). Commercial froth cameras and software that extract bubble size and froth velocity are readily available (Shean and Cilliers, 2011). Another very recent development in machine-vision instrumentation is the StoneThree pulp sensor (Horn et al., 2022). Compared to the manual, offline sampling of the Anglo Platinum Bubble Sizer (APBS) (Harris et al., 2013), the Stone Three On-line Pulp Sensor (OPS) provides on-line measurements of pulp bubble size, superficial gas velocity (J_g), gas hold-up and bubble surface area flux (S_b) (Horn et al., 2022).

2.3.2.2 Phenomenological models

Phenomenological models try to replicate system behaviour based on the fundamental mechanisms involved in the process, such as mass and energy balances, surface chemistry, physics, *etc.* (Quintanilla, Neethling and Brito-Parada, 2021). Again, it must be emphasised that, due to practicality, most of these models still make simplifying assumptions that lump certain effects and include some parameters that must still be calibrated using experimental or process data. Their biggest advantage is that they can predict performance reasonably well for a much wider range of conditions and are applicable to

more than one plant (Oosthuizen et al., 2017).

However, it is often very difficult to obtain the required information and parameters of these models due to limitations in instrumentation. While this is not such a drawback when simulation or design is considered, as some parameters can be specified by the user, it severely limits the application of such models to control (Oosthuizen et al., 2017; Quintanilla, Neethling and Brito-Parada, 2021). Quintanilla, Neethling and Brito-Parada (2021) further divide phenomenological models into kinetic (Shean et al., 2018; Oosthuizen et al., 2021; Steyn and Sandrock, 2021), population balance (Savassi, 2005; Bascur and Kuyvenhoven, 2010; Shean and Cilliers, 2011; Alves Dos Santos et al., 2014), probabilistic (Zhu et al., 2014) and hydrodynamic (Wang et al., 2018) models.

2.3.3 Models for control: validation and implementation

The typical model development process is an iterative cycle of formulation and validation. After establishing the goal of the modelling work and understanding the system, be it through phenomenology or data, model equations are formulated. The candidate model must then be validated against the original goals and adjusted as necessary. In the case of process control or simulation, the goal is typically to capture input-output relationships, be it dynamic or steady-state. In the case of feedback control and especially MPC, capturing the dominant process dynamics in addition to steady-state characteristics is very important. However, correctly predicting directionality is often sufficient for successful feedback control.

The importance of modelling for control development in froth flotation was already emphasised in very early modelling work such as Bascur Riquelme (1982), Pérez-Correa et al. (1998), Hodouin et al. (2000) and Casali et al. (2002). These early studies focused more on developing simulator models that could be used to develop and test control strategies. All of the above-mentioned studies (Bascur Riquelme, 1982; Pérez-Correa et al., 1998; Hodouin et al., 2000; Casali et al., 2002) report the use of some form of dedicated sampling of pilot or industrial scale flotation cells to fit and validate the models. The model fitting and validation in Bascur Riquelme (1982) and Casali et al. (2002) are discussed in much more detail than in Pérez-Correa et al. (1998) and Hodouin et al. (2000).

Bascur Riquelme (1982) shows excellent correlation between the steady-state model predictions and experimental data obtained from a pilot scale flotation circuit. His model can predict dynamic changes over 25 min and 30 min experiments. However, the same data used for the fitting of the model was

used in the comparison. Ideally, a model would be trained on one set of data and validated on another (Steyn and Sandrock, 2021). In practice this can be challenging to achieve as experiments involve significant investment of time, effort and funds. To account for measurement error and unmeasured data, Bascur Riquelme (1982) calculate unmeasured variables using material balances and adjusted the "best estimate" of the steady-state values (both measured and unmeasured) using a weighted least-squares minimisation approach which takes into account the reliability of the various measurements.

The cell used in Bascur Riquelme (1982) was very well-instrumented, with flow and density meters on almost all process streams and flow-rate measurement of the reagent, air and water streams. The froth and pulp levels were measured using a custom conductivity probe and light attenuation device. An on-line X-ray fluorescence (XRF) compositional analyser was also used to gather compositional data from the cell tailings.

Casali et al. (2002) used 20 h of on-line data from an industrial 9 cell copper rougher bank. The slurry levels, feed flow-rate, solids content, copper and iron grade and the rich concentrate copper grade were measured on-line. The available measurements were not sufficient to fit all model parameters. Instead, viable parameters were fitted to steady-state data and some of the parameters related to unmeasured dynamics were fixed at the best estimates. Casali et al. (2002) point out that fitting the remaining phenomenological parameters on the available plant data is difficult, as many of the inputs such as feed rate and grade could not be varied at will. Instead, they fitted auto-regressive moving average with exogenous inputs (ARMAX) models to the available plant data by minimising the root mean square error (RMSE) of the concentrate grade prediction. The fitted ARMAX model was then used to generate simulated step and pulse responses of the plant which were used to calibrate the remaining model parameters. This study is a clear example of the inherent difficulties of obtaining the necessary parameters for phenomenological models from industrial data.

In the interim years there have been continued modelling efforts, with a clearly-stated preference for phenomenological models. Some examples of models for relating the froth dynamics to water recovery, entrainment and grade are Ventura-Medina and Cilliers (2002), Neethling and Cilliers (2003) and Neethling and Cilliers (2009). The very detailed nature of these models also means that there are a number of parameters that are very difficult to obtain experimentally and certainly not purely from typical on-line measurements. Rather than trying to fit the parameters to experimental data, these studies assumed realistic values for the parameters and compared the predicted trends to trends

observed in industrial data, experimental work and dedicated sampling campaigns. While this served to significantly improve the understanding of the phenomena involved, these models are not directly applicable to control in their unsimplified form due to their complexity (Quintanilla, Neethling and Brito-Parada, 2021; Oosthuizen, 2023).

More recently, there has been an increased focus on developing models suitable for MPC and plant optimisation (Brooks and Koorts, 2017; Oosthuizen et al., 2017; Shean et al., 2018; Brooks et al., 2019; Quintanilla, Neethling and Brito-Parada, 2021; Quintanilla, Neethling, Navia and Brito-Parada, 2021; Quintanilla, Neethling, Mesa, Navia and Brito-Parada, 2021; Steyn and Sandrock, 2021; Oosthuizen, 2023). Another trend in the modelling of flotation circuits is the increased emphasis on relying on on-line measurements, rather than extensive sampling campaigns. This is the main motivation for the modelling approach followed by Oosthuizen (2023) and is also discussed in Oosthuizen et al. (2017) and Steyn and Sandrock (2021).

Brooks and Koorts (2017) and Brooks et al. (2019) are somewhat different to the other recent studies in the field of MPC on flotation circuits as they make use of linear models (finite step response models), include reagent dosage rate as a manipulated variable (MV) in the model and report on the successful industrial implementation of MPC in a flotation circuit (Brooks and Koorts, 2017). They make use of on-line XRF compositional analysers to measure the grade and recovery of the various process streams and manipulate the air flows, pulp levels and reagent dosage rates to maximise recovery while maintaining grades within a specified range.

The linear models were fitted to response data from several cells obtained using automated step test software. As the cells in a bank are typically similar, there are models that are expected to be quite similar for different cells, *e.g.,* the relationship between the air rate of a cell and the concentrate rate. The best model for one of the MVs to the CV in question was chosen and all other similar models were made to be identical. Brooks and Koorts (2017) argue that this prevents the controller from finding a solution with markedly different aeration rates down the bank that would be detrimental to performance. The MPC resulted in a clear reduction in reagent usage and even an improvement in the tails grade distribution mean and variability, with estimated annual financial gains of $1M AUD.

Steyn and Sandrock (2021) developed a steady-state model of flotation aimed at offline control and optimisation studies and trained it on predominantly on-line measurement data of a month from the

rougher bank at Mogalakwena North Concentrator (MNC). The available data was divided into training and testing sets. The resulting model performed well enough on the training data set that it was deemed suitable for further optimisation studies (Steyn and Sandrock, 2021). Interestingly, the model uses a single froth stability parameter for the entire bank to determine individual froth recoveries.

Two of the most recent dynamic models of flotation circuits aimed at MPC were proposed by Oosthuizen et al. (2021) and Quintanilla, Neethling, Navia and Brito-Parada (2021). Quintanilla, Neethling, Navia and Brito-Parada (2021) developed a primarily phenomenological model that has a strong focus on including the froth physics and interaction between pulp and froth. Compared to Quintanilla, Neethling, Navia and Brito-Parada (2021), Oosthuizen et al. (2021) make use of very similar models for the pulp and flotation rate, however much simpler empirical models for the froth properties (air recovery and bubble size). Instead, Oosthuizen et al. (2021) specifically developed the model such that all states are observable and all critical parameters (including the froth model parameters) are identifiable using only measurements that are available on modern flotation circuits (see Section 3.3 for more on observability). They propose that the parameters be updated in real time from the process measurements using state estimation, achieving long-term model prediction performance. The model was developed further in Oosthuizen (2023).

Similar to Brooks and Koorts (2017) and Steyn and Sandrock (2021), Oosthuizen et al. (2021) used the same parameter values for the froth models for all cells down the bank. This approach simplifies the model considerably and ensures that the parameters are all identifiable from the available measurements. In Oosthuizen (2023) parameters that determine the optimal J_g value and maximum achievable air recovery in each cell were allowed to vary between cells, as this is more akin to what is seen in real flotation circuits.

Quintanilla, Neethling, Mesa, Navia and Brito-Parada (2021) validated the model in Quintanilla, Neethling, Navia and Brito-Parada (2021) using extensive experiments on a single laboratory scale flotation cell. The cell was well-instrumented and carefully controlled with known feed properties and varying aeration rate and tailings valve fraction. The instrumentation used during the experiments are all relatively typical of modern industrial plants. The model showed very good correlation with the experimental data over periods of roughly 10 min. However, the feed properties and reagent dosages were kept constant and parameter values were assumed to remain constant. The calibrated parameters primarily affected the bubble size and bursting rate (which plays an important role in the air recovery

model). The floatability parameter value was taken from Hu et al. (2013) and also assumed to remain constant. All of these variables are closely tied to the reagent (frother and collector) conditions and the amount of solids in the froth also has a significant impact on the froth stability. Without tests where these conditions where varied, it is not known how well the model would extend to such operating conditions and how much the parameters would have to change.

Oosthuizen (2023) fitted the model parameter values to steady-state data from Hadler et al. (2010), reporting good correlation. Simulations have shown that the MHE can track parameter changes and that the model behaviour matches observed industrial plants (Oosthuizen, 2023). However, the model and parameter estimation has yet to be validated against dynamic industrial data.

2.3.4 Reagent models in literature

Although there are plenty of models of various approaches in literature, few include the effect of varying reagent addition rate or concentration. Examples of how the reagent effects are accounted for are considered here.

As an example of early empirical attempts, the model in Hodouin et al. (2000) uses transfer functions to model the system around a specified reference point. This includes a discrete second-order plus dead-time transfer function with a single zero

$$G = z^{-\theta_d} \frac{K(1+\tau_a z^{-1})}{1+\tau_1 z^{-1}+\tau_1 z^{-2}}, \tag{2.1}$$

to model the effect of collector addition on both the concentrate and tails grades. As pointed out by the authors Hodouin et al. (2000), the models are limited to operation around the reference point where the parameters were fitted and disregard any non-linearity in the system. This severely limits the use of these kinds of models for large scale circuit optimisation that is investigated in more recent literature (Oosthuizen et al., 2017, 2021). A poignant counter-example to the arguments against such simple models is the successful implementation of an MPC using step response reagent models by Brooks and Koorts (2017).

Recent examples of data-driven models include the use of neural networks and machine vision to model relationships in flotation circuits (Aldrich et al., 2010; Gupta and Yan, 2016; Quintanilla, Neethling and Brito-Parada, 2021). For example, Fu and Aldrich (2019) showed that grade could be accurately determined using machine vision and convolution neural networks. While not directly related to modelling of reagent effects, this could still contribute to reagent control strategies. A very

promising trend is the direct modelling of flotation with neural networks for control implementations. Many studies in this field, such as Pu et al. (2020), claim to show great promise in predicting flotation performance, but neglect to train the model on data that capture variation in the reagent addition rates. This is quite possibly due to a lack of long term process data for training. One example of a neural network approach that includes reagent variation is Zhang et al. (2021), which developed a reagent control strategy based on a "neural network autoregressive with exogenous input". This strategy produced promising results during industrial implementation.

The main advantage of using neural network (or similar) models, is that they do not require exact knowledge of all the process mechanisms. Unfortunately, this type of modelling requires large amounts of training data that is often not available (Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021; Steyn and Sandrock, 2021). Another major drawback is that after training the model on the dataset, which is usually the result of an expensive sampling campaign, it is only applicable to that specific plant. If modelling is to make a large impact on the flotation process industry as a whole, the approach followed should preferably be easily implementable on many different processes. However, if the correct balance between data requirement and general applicability of neural network-based models can be achieved, they could be very useful tools in flotation control and optimisation.

The effects of true flotation are modelled much better through the use of phenomenological models (Shean and Cilliers, 2011; Oosthuizen et al., 2017). One of the most common phenomenological models used for this purpose, is the first-order flotation rate kinetics model (Shean and Cilliers, 2011; Oosthuizen et al., 2017; Quintanilla, Neethling and Brito-Parada, 2021)

$$\frac{\mathrm{d}}{\mathrm{d}t}N_p = -KN_p, \tag{2.2}$$

where N_p represents some particle concentration in the pulp and K is a flotation rate constant. This is derived by treating the true flotation process in the pulp as a chemical reactor with the more general kinetic rate equation (Gupta and Yan, 2016)

$$\frac{\mathrm{d}}{\mathrm{d}t}N_p = -k\prod N_i^{m_i}, \tag{2.3}$$

where k is an overall rate constant, N_i is the concentration of any species i that possibly has an effect on the flotation rate (particles, bubbles, reagents *etc.*) and m_i is its corresponding order. By reducing this to first-order kinetics the effects of all these other variables are lumped into the rate constant.

The rate constant in (2.2) can be fitted to experimental data for a set of operating conditions, but to

achieve more realistic models the rate constant is often modelled empirically or the kinetic model is expanded to include some additional factors. While several authors expand this basic model to include the effects of variables such air rate and bubble size (Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021) and particle size and class (Bergh and Yianatos, 2013; Oosthuizen et al., 2021), only a few examples of models could be found where the effect of a reagent is modelled explicitly.

In Casali et al. (2002) a dynamic simulator for a rougher circuit is developed. While the simple first-order kinetic form is used in the mass balances, the experimentally determined flotation rate constants, K_i are adjusted using correction factors, $f_{COL,j}$, $f_{GRA,j}$ and $f_{CPA,j}$ to account for the collector dosage, particle granulometry and solids concentration respectively. The constant $K_{i,j}$ for a mineralogical class i in cell j had the form

$$K_{i,j} = K_i f_{GRA,j} f_{COL,j} f_{CPA,j}. \tag{2.4}$$

The correction factors are linear functions of the related variable and must be fitted to experimental data. Due to limited plant data, some of the model parameters had to be tuned manually (Casali et al., 2002), again highlighting the drawbacks of a large parameter set. The empirical nature of this model also limits its suitability for control applications as already discussed.

Expanding on the model first suggested by Pérez-Correa et al. (1998) to account for different particle sizes, i, and granulometry classes, k, in different cells, j, Putz and Cipriano (2015) modelled both the flotation rate constant $K_{col}^{i,j}$ and a froth drainage constant, $K_e^{i,j}$ as polynomial functions of collector, Q_{col}, and frother, Q_f, addition rates

$$K_{col}^{ijk} = a_1^{ijk} Q_{col}^3 + a_2^{ijk} Q_{col}^2 + a_3^{ijk} Q_{col} + \kappa_{col,0}^{ijk}, \tag{2.5}$$

$$K_e^{ijk} = b_1^{ijk} Q_{col} + b_2^{ijk} Q_f + \kappa_{e0}^{ijk}. \tag{2.6}$$

The empirical constants a_1, a_2, a_3, b_1, b_2, κ_{e0} and $\kappa_{col,0}$ are fitted from plant data. Their model was used to develop and simulate a hybrid MPC implementation. While their results were quite promising, they did not consider variation in the reagent feed rates and did not report the values of the fitted parameters. Therefore, no conclusion as to the applicability of this model of reagent effects could be made. While the model may prove to be applicable, concerns regarding the robustness of a control strategy based on this model would need to be addressed. Without any phenomenological basis for this model structure and parameters that will not be updated on a regular basis, these empirical relations could likely show deteriorating accuracy when operating conditions vary or the strategy is implemented

for any extended period of time (Oosthuizen et al., 2021). Furthermore, Quintanilla, Neethling and Brito-Parada (2021) argues that while first-order rate equations model the pulp behaviour well, they are inadequate for use in the froth phase. This would mean that the use of a drainage rate constant is not entirely appropriate.

With regard to robustness over time, the dynamic model developed by Oosthuizen et al. (2021) provides an alternative. While the reaction kinetics model is expanded to include the effect of bubble surface area flux S_b as well as a measure of froth stability, in the form of air recovery, α, any effects that either the collector or frother might have on the flotation kinetics or froth stability are only included in the empirically fitted constants. Their true flotation model is given by (Oosthuizen et al., 2021)

$$\frac{\mathrm{d}}{\mathrm{d}t}M_k^{i,j} = -K^{i,j}M_k^{i,j}S_{b_k}\alpha_k, \tag{2.7}$$

$$S_{b_k} = 6\frac{J_{g_k}}{D_{BP_k}}, \tag{2.8}$$

where $M_k^{i,j}$ is the mass of species i of particle class j in cell k, D_{BP} is the bubble size in the pulp and J_g is the superficial air velocity. The flotation rate constants, $K^{i,j}$ are fitted to plant data and form a distribution based on mineral species and class. They model α and the top of froth bubble size, D_{BF} as (Oosthuizen et al., 2021)

$$\frac{\mathrm{d}}{\mathrm{d}t}D_{BF_k} = \frac{K_{BS_{J_g}}J_{g_k} + K_{BS_\lambda}\lambda air_k - D_{BF_k}}{\lambda_{air_k}}, \tag{2.9}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha_k = \frac{K_{\alpha_{BF}}D_{BF_k} + K_{\alpha_{J_g}}J_{g_k} - \alpha_k}{\lambda_{air_k}}, \tag{2.10}$$

where $K_{BS_{J_g}}$, K_{BS_λ}, $K_{\alpha_{BF}}$ and $K_{\alpha_{J_g}}$ are model parameters that have to be estimated and λ_{air_k} is given by (Oosthuizen et al., 2021)

$$\lambda_{air_k} = \frac{h_{f_k}}{J_{g_k}}, \tag{2.11}$$

with h_f being the froth depth.

Importantly, Oosthuizen et al. (2021) have shown that their model is observable if the fitted parameters are added to the state vector without dynamics. This implies that the empirical parameters for the α_k and D_{BF_k} models can be estimated from available on-line measurements and adjusted over time. They also show that the flotation rate constants can be estimated from on-line measurements, at least for the steady-state case. This means that changes in reagent effects could be captured by the updating of the model parameters. Updating parameters, however, takes some time and is limited to small or very gradual changes in the parameters.

If the model of Oosthuizen et al. (2021) could be expanded with models that capture the effect of variation in reagent addition, with parameters that are observable (given realistically available on-line measurements), the problem of MPC model robustness could be overcome. The inclusion of reagent effects could also increase the accuracy and applicability of the model over larger operating ranges or even MPC-based reagent control strategies. As it is, their model could already inform manual reagent addition control through real time estimates of the effects of flotation *vs.* entrainment.

No examples of models where the effect of reagents on the froth phase and froth stability are explicitly modelled could be found. The froth phase and entrainment models in Neethling and Cilliers (2003) and Neethling and Cilliers (2009), for instance, mention chemical effects only in passing. While the population balance models of Alves Dos Santos et al. (2014) and Savassi (2005) recognise the importance of frother concentration in determining entrainment and froth recovery, none of the models used for the calculation of transfer rates make explicit provision for varying chemical conditions. It is assumed that in these models, the effects of chemical reagents would impact the empirically fitted parameters.

A couple of examples of fundamental models of aspects related to flotation were found in literature. The adsorption of collector to the mineral was modelled by Fairthorne et al. (1997) as a Langmuir kinetic model

$$[C]_{\text{ads}} = [S]_{\text{o}} \frac{\left(K_1[C]_{\text{eq}} + 2K_1K_2[C]_{\text{eq}}^2 + 3K_1K_2K_3[C]_{\text{eq}}^3 + nK_1K_2\ldots K_n[C]_{\text{eq}}^n\right)}{\left(1 + K_1[C]_{\text{eq}} + K_1K_2[C]_{\text{eq}}^2 + K_1K_2K_3[C]_{\text{eq}}^3 + K_1K_2\ldots K_n[C]_{\text{eq}}^n\right)}, \tag{2.12}$$

where $[C]_{\text{ads}}$ is the total concentration of adsorbed collector, $[C]_{\text{eq}}$ is the collector concentration remaining in the pulp, $[S]_{\text{o}}$ is the concentration of surface cites and K_n is the equilibrium constants for adsorption to the n^{th} layer. There are also some models that predict the critical coalescence concentration (CCC) of different frothers (Finch et al., 2008; Karakashev et al., 2020). However, examples of where these models were used in other models of flotation performance or other important flotation phenomena could not be found.

2.4 STATE ESTIMATION IN FLOTATION

As mentioned in Section 2.3.3, model-based control often relies on state feedback and typically not all states are measured. The unmeasured states must be estimated from the available measurements. There is a wealth of literature on linear (Kalman, 1960; Simon, 2006; Backi et al., 2020) and non-linear state estimation (Van der Merwe, 2004; Daum, 2005; Haseltine and Rawlings, 2005; Impraimakis and Smyth, 2022) as well as studies on the implementation of state estimation in various industries.

State estimation is commonly used in aeronautics, robotics (Rigatos, 2012), electrical power networks (Mathieu et al., 2013), processing industries (Kulikov and Kulikova, 2014, 2015) and many other fields (Simon, 2006; Jin et al., 2021). There are also examples of where state estimation is used in comminution (Le Roux et al., 2017; Olivier et al., 2012).

Yet, despite the importance of state estimation for model-based control, very few of the recent control-focussed modelling studies (Brooks and Koorts, 2017; Shean et al., 2018; Brooks et al., 2019; Quintanilla, Neethling, Navia and Brito-Parada, 2021) explicitly consider the observability of the proposed model states or parameters. The obvious exception is Oosthuizen et al. (2017, 2021) and Oosthuizen (2023) as already discussed. In the rest of this chapter, more examples of the use of state estimation in flotation research and/or industry are summarised. While not exhaustive, the list gives a good idea of the state of state estimation in flotation.

One of the earliest studies that considered state estimation in flotation modelling and control is Bascur Riquelme (1982). He highlights the importance of using state estimation to deal with unmeasured or infrequently measured variables, such as concentrate grade. An extended Kalman filter (EKF) was implemented for simulated responses of their calibrated model, showing good estimation of the grade between measurements for assay sampling time of less than 3.5 min. However, the estimation was not performed on experimental data where plant-model mismatch would have an impact.

More recently, Popli et al. (2015) developed a fundamental model for batch flotation cells and implemented an EKF to estimate the states and update parameters from machine vision measurements of the froth properties in order to perform fault and disturbance identification. They found that only one of three parameters was identifiable at a time and ran three EKFs in parallel, each estimating a different parameter. They were able to detect faults and disturbances on a laboratory scale batch flotation cell by comparing the different parameter estimates to heuristics they developed. While not completed on industrial scale, this study serves as a proof of concept that demonstrates the potential of non-linear state estimation in froth flotation, specifically regarding real-time parameter updates (Popli et al., 2015).

In a similar vein, Popli et al. (2018) used an EKF based on a population balance fundamental model of a quartz/galena laboratory batch flotation process along with on-line soft sensor measurements of the galena grade and overall solids recovery to estimate the relative entrainment and flotation contributions.

The soft sensors use support vector regression (SVR) to infer the grade and recovery measurements from the on-line image features provided by the VisioFroth package. The model focussed on flotation and entrainment phenomena. The EKF was used to update the entrainment, attachment and drainage rate parameters in real time from the measured model inputs (aeration rate and make-up water flow-rate), measured outputs (pulp height, froth height, froth velocity and bubble size) and the soft sensor output measurements. These are very promising results for the current study, but it should be noted that the experimental system was much simpler and better regulated than an industrial scale continuous flotation circuit. Furthermore, the availability of real-time grade and recovery measurements for each cell is not typical for real plants. Recovery is by definition a steady-state concept. Any "instantaneous recovery" measurement calculated from the feed and concentrate flow-rates and grades would only be a rough approximation of the real recovery. It would need to be averaged over a period (Oosthuizen, 2023) to approach the true overall recovery of the process.

Flotation has also expanded to other industries, notably coal enrichment. Joostberens et al. (2022) used a Kalman filter as an on-line system identification tool for a SISO linear model of the effect of the reagent dosage on the tailings ash content. The aim was to develop an on-line method for adjusting the model when a new operating point is implemented by the operators to allow for tailings ash control via reagent dosage around the operating point. They found that the on-line parameter fits for a dedicated step test yielded similar results to offline least square parameter fitting. It is unclear whether the method is meant to run continuously or only intermittently.

There are studies where state estimation is used to improve the measurements obtained from froth cameras rather than overall process monitoring. For example, Liu et al. (2013) use a Kalman filter to reduce the computational load of extracting froth velocity information for severely deformed and fragile froths via the scale-invariant feature transform (SIFT).

Another form of soft-sensor using state estimation was proposed by Maldonado, Desbiens, del Villar and Aguilera (2010) who investigated the use of multiple linear models of the effect of gas velocity on the bubble size in a flotation column collection zone, each calibrated at different known frother concentrations in an air-water-reagent system. A Kalman filter was built for each of the models and the set of estimators were run in parallel using measurements from a laboratory scale column with varying frother dosage rates. The frother concentration estimate is calculated as the sum of the calibration concentrations, weighted by the conditional probability of the associated model representing the actual

plant dynamics. While the frother concentration could successfully be estimated for the two-phase system, it is unclear what the effect of adding solids or higher concentrations of frothers would be. This method would likely only work for frother concentrations below the CCC, as it relies on changes in the pulp bubble size to differentiate between different concentrations.

2.5 CONCLUDING REMARKS

While the variables and fundamental phenomena affecting flotation are well-known and there exists many models of varying levels of detail, literature on the successful implementation of model-based control in flotation remains scarce. Recently, this has led to an increased focus on modelling specifically aimed at control (Oosthuizen et al., 2021; Quintanilla, Neethling and Brito-Parada, 2021). The model and parameter estimation approached suggested by Oosthuizen et al. (2021) appears to be very promising. However, while the model shows good steady-state correlation with plant data, both the model and state estimation approach are yet to be validated using dynamic industrial data. In general, state estimation has seen limited use in flotation literature.

The importance of reagents is also commonly recognised and improving the control of reagent addition has significant potential benefits (Brooks and Koorts, 2017). However, there appears to be a gap in the modelling of reagent effects. Often, constant reagent conditions are assumed and very few models make explicit provision for the effect of varying reagent conditions. While on-line parameter estimation using observable models could account for changing reagent effects (Oosthuizen, 2023), explicit models of the reagent effect are required for inclusion in model-based control strategies.

CHAPTER 3 METHODS

3.1 CHAPTER OVERVIEW

The highly interconnected and non-linear nature of the flotation process and related models necessitates the use of advanced analyses methods and state estimation algorithms. This chapter aims to give the necessary theoretical background of these techniques. Section 3.2 covers the Sobol sensitivity analysis, while Section 3.3 deals with the concept of non-linear observability. The non-linear state estimation algorithm, the MHE, is discussed in Section 3.4. An expansion of the well-known RMSE metric that divides the observed error into systematic and unsystematic (or random) error is discussed in Section 3.5.

3.2 SENSITIVITY ANALYSIS THEORY

In general a sensitivity analysis is aimed at quantifying the impact of model inputs or parameters (called factors) on the model outputs (Saltelli et al., 2019). The Sobol sensitivity analysis is a global variance based sensitivity analysis method that quantifies the relative contribution of the variance in each input factor, of arbitrary distribution and range, as well as interactions between input factors, to the total variance in the model output (Saltelli et al., 2010). This is done by calculating a first-order and total sensitivity index, S_i and S_{Ti} respectively, for each input factor, X_i. The fact that the Sobol sensitivity analysis accounts for interactions and evaluates the sensitivity of all factors simultaneously makes it particularly suited for dealing with the highly interconnected processes, such as the flotation models under consideration (see Section 4.4). A recent example of the Sobol sensitivity analysis being used in process control research is found in Wiid et al. (2021).

Saltelli et al. (2010) discusses how each of these sensitivity indexes can be calculated from function evaluations in a Monte-Carlo simulation and the use of the Sobol sequence, a low discrepancy quasi-random sampling method that aims to sample the input factor space more completely. Both the

sampling method and sensitivity analysis calculations are implemented in the `SALib` (Herman and Usher, 2017) library in Python. The basic calculations for the Sobol sensitivity analysis are summarised below, but the reader is referred to Saltelli et al. (2010) for more detailed discussions regarding the method.

If the scalar model output, Y, is given as

$$Y = f(\mathbf{X}), \tag{3.1}$$

where $\mathbf{X} = \{X_1, X_2, \ldots, X_k\}$ is the $N \times k$ matrix with N samples of the k input factors, the sensitivity indexes are (Saltelli et al., 2010)

$$S_i = \frac{\mathrm{V}_{X_i}\left(\mathrm{E}_{\mathbf{X}_{\sim i}}(Y|X_i)\right)}{\mathrm{V}(Y)} = \frac{\frac{1}{N}\sum_{j=1}^{N} f(\mathbf{B})_j \left(f\left(\mathbf{A}_{\mathbf{B}}^{(i)}\right)_j - f(\mathbf{A})_j\right)}{\mathrm{V}(Y)}, \tag{3.2}$$

and

$$S_{Ti} = \frac{\mathrm{E}_{\mathbf{X}_{\sim i}}\left(\mathrm{V}_{X_i}(Y|\mathbf{X}_{\sim i})\right)}{\mathrm{V}(Y)} = \frac{\frac{1}{2N}\sum_{j=1}^{N} \left(f(\mathbf{A})_j - f\left(\mathbf{A}_{\mathbf{B}}^{(i)}\right)_j\right)^2}{\mathrm{V}(Y)}. \tag{3.3}$$

where E and V respectively denote the expected value and variance operators and $\mathbf{X}_{\sim i}$ denotes the matrix of all factors except for X_i. The $\mathrm{V}_{X_i}\left(\mathrm{E}_{\mathbf{X}_{\sim i}}(Y|X_i)\right)$ term in (3.2) is the variance in the mean of Y, with all factors except X_i varied, taken over the entire range of X_i (Saltelli et al., 2010). The $\mathrm{E}_{\mathbf{X}_{\sim i}}\left(\mathrm{V}_{X_i}(Y|\mathbf{X}_{\sim i})\right)$ term in (3.3) is the expected variance of Y, with all factors except X_i fixed. For N simulations with k input factors the sampling matrices $\mathbf{A}$ and $\mathbf{B}$ are two independent $N \times k$ matrices generated using a Sobol sequence of size $(N, 2k)$ (Saltelli et al., 2010) divided into two halves. Each matrix contains N sampled values for the k factors in its columns. The matrix $\mathbf{A}_{\mathbf{B}}^{(i)}$ is the same as $\mathbf{A}$, except that the i-th column is replaced by the i-th column of $\mathbf{B}$. The notation $(\mathbf{A})_j$ denotes the j-th row of $\mathbf{A}$.

The value of S_i is the fraction of the variance in Y that can be attributed to the variation of X_i alone, while S_{Ti} is the fraction of variance in Y that can be attributed to any joint factor variation that includes X_i (Saltelli et al., 2010, 2019; Wiid et al., 2021).

3.3 OBSERVABILITY ANALYSIS THEORY

3.3.1 Non-linear observability condition

For the general non-linear state-space model (Villaverde et al., 2019)

$$\begin{aligned} \dot{\mathbf{x}}(t) &= f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}), \\ \mathbf{y}(t) &= g(\mathbf{x}(t), \mathbf{u}(t)), \end{aligned} \tag{3.4}$$

the states ($\mathbf{x}(t) \in \mathbb{R}^{n_\mathbf{x}}$) are said to be *observable* if they can be inferred from the known input ($\mathbf{u}(t) \in \mathbb{R}^{n_\mathbf{u}}$) and output ($\mathbf{y}(t) \in \mathbb{R}^{n_\mathbf{y}}$) history. Similarly, the unknown parameters ($\mathbf{p} \in \mathbb{R}^{n_\mathbf{p}}$) are *identifiable* if they can be inferred from the available $\mathbf{y}(t)$ and $\mathbf{u}(t)$ (Villaverde et al., 2019). These definitions are quite simplified, but capture the practical implications of these properties. Importantly, these properties only indicate that the states and parameters can be estimated in theory, but do not give information regarding the practical observability of these values. Note that the model assumes that the $\mathbf{p}$ does not vary with time. Henceforth the "function of time" notation (t) will be dropped, unless it is required for disambiguation.

The simplest check for observability of the system (3.4) is to check the observability (and/or identifiability) of a linearised version of the model given in deviation variables ($\delta\mathbf{z} = \mathbf{z} - \mathbf{z}_0$)

$$\begin{aligned} \delta\dot{\mathbf{x}} &= \mathbf{A}\delta\mathbf{x} + \mathbf{B}\delta\mathbf{u}, \\ \delta\mathbf{y} &= \mathbf{C}\delta\mathbf{x} + \mathbf{D}\delta\mathbf{u}, \end{aligned} \tag{3.5}$$

with

$$\begin{aligned} \mathbf{A} &= \left.\frac{\partial f(\mathbf{x},\mathbf{u},\mathbf{p})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0,\mathbf{u}_0}, & \mathbf{B} &= \left.\frac{\partial f(\mathbf{x},\mathbf{u},\mathbf{p})}{\partial \mathbf{u}}\right|_{\mathbf{x}_0,\mathbf{u}_0}, \\ \mathbf{C} &= \left.\frac{\partial g(\mathbf{x},\mathbf{u})}{\partial \mathbf{x}}\right|_{\mathbf{x}_0,\mathbf{u}_0}, & \mathbf{D} &= \left.\frac{\partial g(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}}\right|_{\mathbf{x}_0,\mathbf{u}_0}, \end{aligned} \tag{3.6}$$

taken around a nominal operating point ($\mathbf{x}_0$, $\mathbf{u}_0$). The states are linearly observable if

$$\operatorname{rank}(\mathcal{O}_L) = n_\mathbf{x}, \tag{3.7}$$

where

$$\mathcal{O}_L = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n_\mathbf{x}-1} \end{bmatrix}. \tag{3.8}$$

This is known as the linear observability-identifiability condition (LOIC). It is a sufficient, but not necessary condition for the observability of the non-linear model.

Should the LOIC not be satisfied, it is still possible that the states (and parameters) might be still be non-linearly observable, *i.e.,* inferable using non-linear state estimation algorithms, *e.g.,* MHEs. The non-linear observability-identifiability (NOIC) condition requires

$$\operatorname{rank}(\mathrm{d}\mathcal{O}(\mathbf{x}_0,\mathbf{u}_0)) = n_\mathbf{x}, \tag{3.9}$$

where

$$\mathrm{d}\mathcal{O}(\mathbf{x},\mathbf{u}) = \begin{bmatrix} \dfrac{\partial}{\partial \mathbf{x}} g(\mathbf{x},\mathbf{u}) \\ \dfrac{\partial}{\partial \mathbf{x}} \left(L_f g(\mathbf{x},\mathbf{u})\right) \\ \dfrac{\partial}{\partial \mathbf{x}} \left(L_f^2 g(\mathbf{x},\mathbf{u})\right) \\ \vdots \\ \dfrac{\partial}{\partial \mathbf{x}} \left(L_f^{n_\mathbf{x}-1} g(\mathbf{x},\mathbf{u})\right) \end{bmatrix}. \tag{3.10}$$

The i^{th} Lie derivative of g with respect to f is (Villaverde et al., 2019)

$$L_f^i g(\mathbf{x},\mathbf{u}) = \frac{\partial L_f^{i-1} g(\mathbf{x},\mathbf{u})}{\partial \mathbf{x}} f(\mathbf{x},\mathbf{u}) + \sum_{j=0}^{j=i-1} \frac{\partial L_f^{i-1} g(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}^{(j)}} \mathbf{u}^{(j+1)}, \tag{3.11}$$

with

$$L_f g(\mathbf{x},\mathbf{u}) = \frac{\partial g(\mathbf{x},\mathbf{u})}{\partial \mathbf{x}} f(\mathbf{x},\mathbf{u}) + \frac{\partial g(\mathbf{x},\mathbf{u})}{\partial \mathbf{u}} \mathbf{u}^{(1)}, \tag{3.12}$$

where the notation $\mathbf{u}^{(j)}$ denotes the j^{th} derivative of $\mathbf{u}$.

The identifiability of the parameters ($\mathbf{p}$) can be evaluated by expanding the state vector $\mathbf{x}$ with the parameters and modelling them with zero dynamics. The standard state observability tests above ((3.7) and (3.9)) can then be used for the expanded system

$$\dot{\tilde{\mathbf{x}}} = \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} f(\tilde{\mathbf{x}},\mathbf{u}) \\ \mathbf{0} \end{bmatrix}, \tag{3.13}$$

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{p} \end{bmatrix}. \tag{3.14}$$

3.3.2 Input excitation

The second term in (3.11) captures the impact of the input and its derivatives on the states and outputs. By assuming zero values for all input derivatives above a certain order, it is possible to determine the degree of input excitation required for observability (Villaverde et al., 2019).

The analysis in Oosthuizen (2023) assumed that $\mathbf{u}^{(j)} = 0 \; \forall \; j \geq 1$, effectively meaning that the analysis was done for constant model inputs (all input derivatives are zero). This assumption is also made here unless explicitly stated otherwise. In the context of flotation circuits this is a safe assumption as there are long periods of time for which the inputs remain relatively constant. This assumption is also a more conservative look at the observability of the model as higher order $\mathbf{u}^{(j)} \neq 0$ tend to improve the observability of the system.

3.3.3 Practical implementation

The analysis was achieved by defining the model equations and building the required $\mathrm{d}\mathcal{O}$ matrix symbolically using the `symengine` Python library, a group of SymPy-like Python wrappers of the C++ `SymEngine` library (Symengine Development Team, 2021). This allowed for much faster calculations than with SymPy or the Matlab Symbolic Math Toolbox, which is crucial for the large dimensionality of the model in question.

Even with this considerable improvement in performance the computational load of Lie derivatives of 5th or higher order increased drastically. As such, $\mathrm{d}\mathcal{O}$ was built recursively, evaluating the rank after adding the next order Lie derivative until full rank is achieved or the rank stops increasing (Villaverde et al., 2019). The process was terminated after $L_f^5 g$.

While this could lead to prematurely declaring the system unobservable if one of the higher order derivatives could yield full rank, any such observability would be practically limited. The Lie derivative necessary to achieve full rank indicates the lowest order of output derivatives which must have a non-zero value for observability. In practical applications reliably detecting such variation in the system outputs is very difficult, if not impossible, due to measurement noise and reduced output variance achieved through implemented controllers.

3.3.4 Identifying unobservable variables

If (3.9) is not satisfied and there is a rank deficiency of r_d, there are r_d unobservable states or unidentifiable parameters. In the case of parameters, state and parameter estimation of the remainder of the system can be made possible by assuming the unidentifiable parameters as constant.

The possibly unobservable variables can be determined as follows. The rank of $\mathrm{d}\mathcal{O}(\tilde{\mathbf{x}}_0, \mathbf{u}_0)$ (with $n_{\tilde{\mathbf{x}}}$ states and parameters) is taken as the maximum possible rank $r_{\max}$. That is

$$n_{\tilde{\mathbf{x}}} = r_{\max} + r_d. \tag{3.15}$$

Let $\mathrm{d}\mathcal{O}_i$ denote the ith column of $\mathrm{d}\mathcal{O}$, $i = k : l$ indicate a range of columns (inclusive and non-zero indexed) and $\tilde{x}_i$ indicate the ith state or parameter state. Next, let

$$\mathrm{d}\mathcal{O}'_j = \begin{bmatrix} \mathrm{d}\mathcal{O}_{1:j-1} & \mathrm{d}\mathcal{O}_{j+1:n_{\tilde{\mathbf{x}}}} \end{bmatrix}. \tag{3.16}$$

The variable $\tilde{x}_j$ is possibly unobservable/unidentifiable if

$$\text{rank}(\text{d}\mathcal{O}'_j) = r_{\max}, \tag{3.17}$$

as this indicates that the jth column did not contribute to the overall rank of the matrix (Villaverde et al., 2019). This is done for every $\tilde{x}_j$. Often (3.17) can hold for more than r_d variables, which implies that it is the linear combination of the variables that has a deficient rank. If r_d of these variables are assumed to be known constants (*i.e.,* not included in $\tilde{\mathbf{x}}$), an observable system results.

3.4 MOVING HORIZON ESTIMATOR

Typical estimators used for non-linear system models include the EKF, unscented Kalman filter, particle filters (Simon, 2006) and MHE (Lucia et al., 2017). Of these, the EKF has arguably received the most attention, especially in industrial application, with any more complex algorithm often described as unnecessarily complex (Rawlings and Bakshi, 2006). However, the EKF requires linear observability and identifiability, as it is merely applying the normal Kalman filter to a linearised model at each time-step. If a model does not satisfy the LOIC, such as the model of Oosthuizen (2023), the standard EKF cannot be implemented. The MHE can deal with such systems and has the added benefit that system constraints can easily be incorporated.

The MHE implementation of the `do-mpc` Python library (Lucia et al., 2017) is used. This uses a system of the form

$$\begin{aligned} \dot{\mathbf{x}}(t) &= f(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) + \mathbf{w}(t) \\ \mathbf{y}(t) &= g(\mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) + \mathbf{v}(t) \end{aligned}, \tag{3.18}$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{u}$ and $\mathbf{p}$ represent the states, measured outputs, inputs and parameters to be estimated respectively. The measurement noise ($\mathbf{v}$ or $\mathbf{v}_k$) and process noise ($\mathbf{w}$ or $\mathbf{w}_k$) are assumed to be Gaussian and additive (Lucia et al., 2017). For each horizon (i) of N time-steps, the optimisation problem for a discretised version of the non-linear state-space system is given by

$$\begin{aligned} &\min_{\mathbf{x}_0, \mathbf{p}, \{\mathbf{w}_k\}_{k=0}^{N-1}} \|\mathbf{x}_0 - \tilde{\mathbf{x}}_0\|^2_{\mathbf{P}_x} + \|\mathbf{p} - \tilde{\mathbf{p}}\|^2_{\mathbf{P}_p} + \sum_{k=0}^{N-1} \left(\|\mathbf{v}_k\|^2_{\mathbf{P}_v} + \|\mathbf{w}_k\|^2_{\mathbf{P}_w} \right), \\ &\text{subject to} \quad \left.\begin{aligned} \mathbf{x}_{k+1} &= f(\mathbf{x}_k, \mathbf{u}_k, \mathbf{p}, k) + \mathbf{w}_k \\ \mathbf{y}_k &= g(\mathbf{x}_k, \mathbf{u}_k, \mathbf{p}, k) + \mathbf{v}_k \\ h(\mathbf{x}_k, \mathbf{u}_k, \mathbf{p}, k) &\leq 0 \end{aligned} \right\} k = 0, \ldots, N, \end{aligned} \tag{3.19}$$

where $\|\mathbf{x}\|^2_{\mathbf{P}}$ denotes $\mathbf{x}^T\mathbf{P}\mathbf{x}$ and h is the function of non-linear constraints. This is solved numerically using `ipopt` in conjunction with the MA57 linear solver from the Harwell Subroutine Library (HSL) (HSL, 2013).

Only the initial states ($\mathbf{x}_{\{0|i\}}$) for each horizon are varied during optimisation, with deviation from the previous estimate of that state ($\tilde{\mathbf{x}}_0 = \mathbf{x}_{\{1|i-1\}}$) penalised by $\mathbf{P}_x$. The values of $\mathbf{p}$ are assumed to remain constant over the entire horizon, but are allowed to vary between horizons, with deviations from the previous estimate ($\tilde{\mathbf{p}}$) penalised by $\mathbf{P}_p$. The states are then propagated using f, with $\mathbf{w}_k$ considered an optimisation variable. The magnitude of the estimated $\mathbf{w}_k$ and observed $\mathbf{v}_k$ values over the entire horizon are penalised by $\mathbf{P}_w$ and $\mathbf{P}_v$ respectively. The tuning matrices ($\mathbf{P}_x$, $\mathbf{P}_p$, $\mathbf{P}_v$ and $\mathbf{P}_w$) are all specified as diagonal positive semi-definite matrices, with at least one positive definite matrix. Along with these matrices, the horizon length (N) and variable constraints play an important role in the solution of the optimisation problem and estimator performance.

3.5 PERFORMANCE METRICS

For the objective evaluation of state estimation performance, quantitative measures of performance are essential. The mean square error (MSE), or its square root (RMSE), is one of the most commonly used error metrics (Willmott, 1981; Armstrong and Collopy, 1992; Chai and Draxler, 2014; Botchkarev, 2019). For a single dimension with n $O_i|P_i$ pairs, the RMSE is calculated as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \tag{3.20}$$

where O and P denote the observed and predicted values respectively.

However, it is only really sensible to use it for comparing different estimate sets as it is very difficult to determine what an acceptable RMSE for a single estimate should be. Willmott (1981) discusses an expansion of the MSE (and RMSE) by decomposing it into systematic (MSEs) and unsystematic (MSEu) components based on a linear line of best fit

$$\hat{P}_i = a + bO_i \tag{3.21}$$

on the parity plot of P *vs.* O. Essentially this decomposes the estimation error into a part due to a modelling bias (systematic) and the part caused by random noise (unsystematic).

The values of the two sub-metrics are calculated with

$$\text{MSEs} = n^{-1}\sum_{i=1}^{n}(\hat{P}_i - O_i)^2, \tag{3.22}$$

$$\text{MSEu} = n^{-1}\sum_{i=1}^{n}(P_i - \hat{P}_i)^2. \tag{3.23}$$

Furthermore

$$\text{MSE} = \text{MSEu} + \text{MSEs}. \tag{3.24}$$

The RMSEs and RMSEu are merely the square roots of the MSEs and MSEu. The relative value of these two metrics can give a good indication of the performance of a single estimate. If the systematic component is much larger than the unsystematic this could be a good indication of poor estimation performance.

As these metrics are not commonly in use, the prediction results of the α_2 parameter fits discussed later in Section 6.4 are analysed as a demonstration. The parity plot of the resulting predictions is given in Figure 3.1. The calculated metrics are given in Table 3.1.

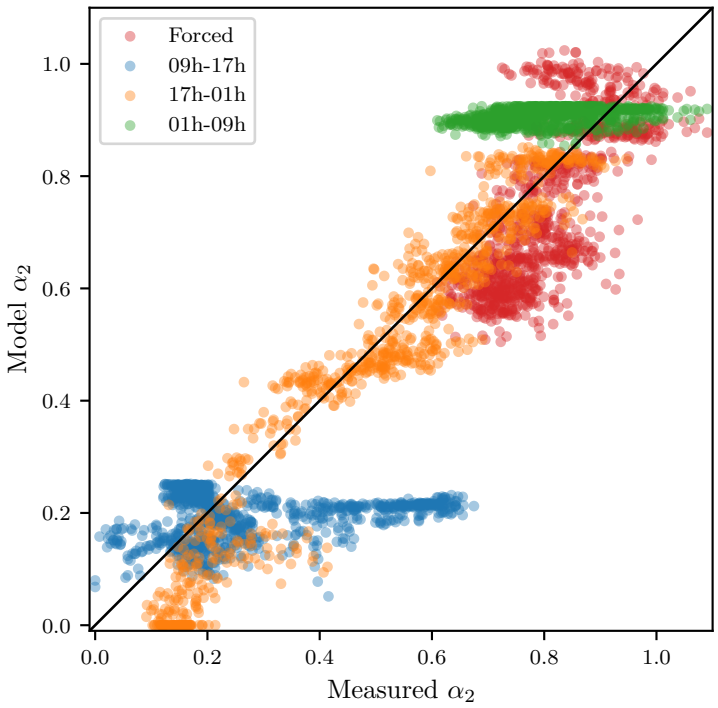

Figure 3.1. Parity plot of the prediction results for the α_2 parameter fits from Section 6.4, reproduced here for ease of reference. Each group of data-points contain the same number of observations.

The RMSE values for all four fits are quite large considering that the overall α_k range is 0–1. Nevertheless, the RMSE confirms that the parameters for the 17h–01h dataset resulted in the best prediction and that the "Modified" parameter values improved the prediction for the 01h–09h dataset. Importantly for this demonstration, the predictions of the 09h–17h and 01h–09h datasets in Figure 3.1 clearly failed to capture the entire dynamic range of the observed α_k. Correspondingly, the ratio of the RMSEs to RMSEu (S/U) is very high for these two datasets. Conversely, the parity plot of the 17h–01h dataset

Table 3.1. Prediction errors for α_2 parameter fits.

	09h–17h	17h–01h	01h–09h	Modified
RMSE	0.185	0.081	0.134	0.118
RMSEs	0.180	0.040	0.134	0.065
RMSEu	0.041	0.070	0.013	0.099
S/U ratio	4.40	0.57	10.15	0.65

seems to be a relatively good fit around the 1:1 line with some noise. This is reflected in the low RMSEs to RMSEu ratio.

3.6 CONCLUDING REMARKS

The analysis and state estimation methods used throughout the rest of this work were summarised in this chapter. The non-linear observability analysis is used in Chapter 4 and Chapter 6 to expand on the analysis of Oosthuizen (2023) and in Chapter 5 to evaluate the proposed model expansions. In Chapter 6 the observability analysis is also used in the model reduction based on the available industrial data. The Sobol sensitivity analysis is used in Chapter 5 to identify the dominant reagent-related parameters for model expansion. In Chapter 6 the MHE is implemented on industrial data and the expanded RMSE performance metrics are used to evaluate the estimation and prediction performance of the reduced model.

CHAPTER 4 FLOTATION MODELLING AND MODEL ANALYSIS

4.1 CHAPTER OVERVIEW

The model termed "original" in this chapter was published in Oosthuizen et al. (2021) and cited in subsequent publications (Quintanilla, Neethling, Navia and Brito-Parada, 2021; Venter et al., 2022). A slightly modified version, used for the reagent model expansion in Chapter 5 and published in Venter et al. (2022), is described in Section 4.2. Thereafter, the model has undergone further changes. The updated model from Oosthuizen (2023) is summarised in Section 4.3 and is used for the dynamic model analysis in Section 4.5 and MHE implementation in Chapter 6 in a reduced form. The causal flow of information between the variables in the updated model is discussed in Section 4.4. The observability of the models described in this chapter is considered in Section 4.5.

4.2 ORIGINAL MODEL

Oosthuizen et al. (2021) models each mechanical cell (subscript k) and concentrate tank (or hopper (subscript H_q)) in a flotation bank individually. (Note that this study accounts for multiple hoppers using the additional subscript q.) They also consider any number of mineral species (superscript i) and mineral classes (superscript j). The model variables, inputs and parameters are summarised in Table 4.1, Table 4.2 and Table 4.3. A simple diagram of a forced air flotation cell is shown in Figure 4.1.

4.2.1 Volume balances

The cell and hopper levels (L_k and L_{H_q}) are modelled using simple volume balances with the initial volumetric feed (Q_{F_1}) considered a measured disturbance and Q_{H_q} taken as a MV. The valve fractions ($v_{a,k}$) between tanks are also MVs that determine the flow of pulp from a cell (Q_{T_k}) according to

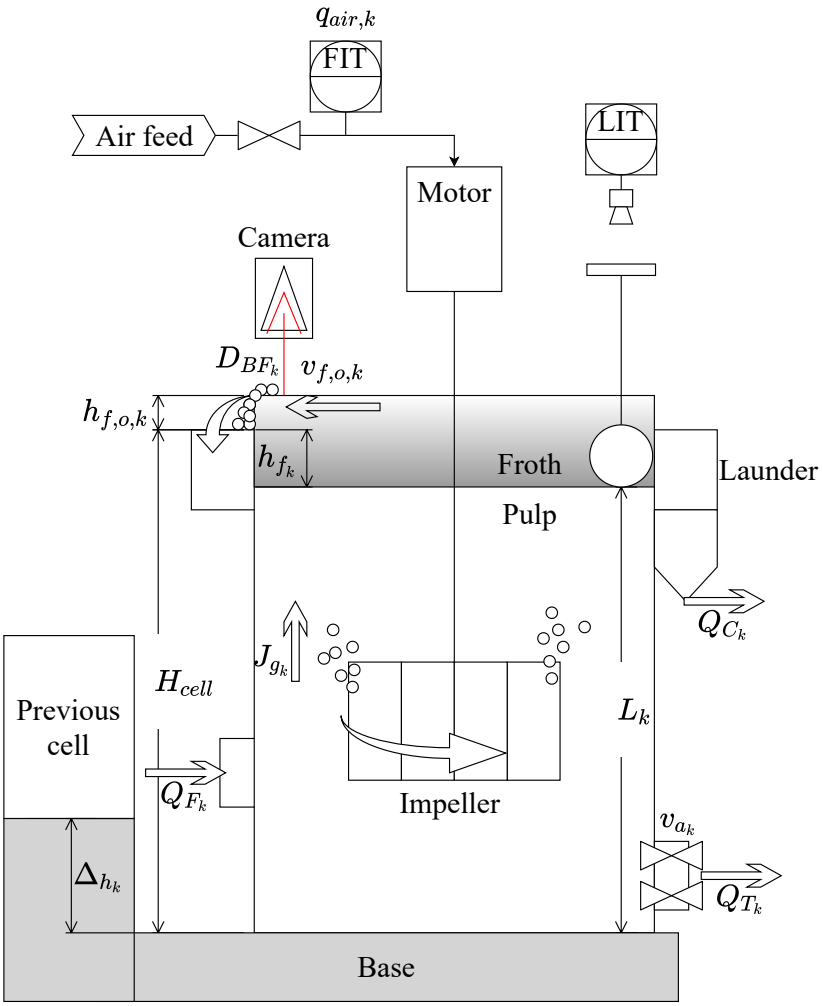

Figure 4.1. Diagram of forced air flotation cell indicating the measured variables relevant to a cell and cell dimensions.

equation

$$Q_{T_k} = v_{a,k} C_{v_k} \sqrt{L_k - L_{k+1} + \Delta_{h_k}}, \tag{4.1}$$

where C_{v_k} is the valve parameter that characterises the flow through the valve and Δ_{h_k} is the height difference between sequential cells in the bank (Oosthuizen, 2023).

The cell volume balances are

$$\frac{\mathrm{d}L_k}{\mathrm{d}t} = \left(Q_{F_k} - Q_{T_k} - Q_{C_k}\right)/A_k, \tag{4.2}$$

where A_k is the cell cross-sectional area and the volumetric feed to a cell (Q_{F_k}) is typically $Q_{F_k} = Q_{T_{k-1}}$. The concentrate volumetric flow-rate (Q_{C_k}) is calculated from the water recovery (Q_{w_k}) and the volume of the floated and entrained particles.

Table 4.1. Model states and intermediate variables.

Variable	Unit	Description
		Subscripts and superscripts
$\diamond$	–	Stream subscript (concentrate (C), tailings (T) and feed (F))
$\triangledown$	–	Tank subscript (hopper (H_q) or flotation cell (k))
i, j	–	Superscripts for mineral species and class
		States
α_k	–	Air recovery
D_{BF_k}	mm	Mean top-of-froth bubble size
$L_\triangledown$	m	Tank level (cell or hopper)
$M_\triangledown^{i,j}$	kg	Species mass in tank
		Auxiliary variables
$C_{s,\diamond}$, $C_{s,\triangledown}$	–	Mass fraction solids in stream or tank
$d_{prt}^{i,j}$	mm	Particle size for 50 % entrainment
$Ent_{Frac}^{i,j}$	–	Entrainment factor
h_{f_k}	mm	Froth height
λ_{air_k}	s	Froth residence time
Q_{w_k}	$\text{m}^3\,\text{h}^{-1}$	Water recovery flow-rate
$Q_{\diamond_\triangledown}$	$\text{m}^3\,\text{h}^{-1}$	Volumetric flow-rate of stream $\diamond$ to/from tank $\triangledown$
$\dot{M}_{\diamond_\triangledown}^{i,j}$	$\text{kg}\,\text{h}^{-1}$	Mass flow-rate in stream $\diamond$ to/from tank $\triangledown$
S_{b_k}	s^{-1}	Bubble surface-area flux for cell k

For the hoppers, the volume balances are

$$\frac{\text{d}L_{H_q}}{\text{d}t} = \left(\sum_{k=m}^{n} Q_{C_k} - Q_{H_q} \right) / A_{H_q}, \tag{4.3}$$

where m and n denote the range of cells from which the concentrate feeds into hopper q and A_{H_q} is the hopper cross-sectional area.

Table 4.2. Model inputs and exogenous variables.

Variable	Unit	Description
	Inputs	
J_{g_k}	$\mathrm{mm\,s^{-1}}$	Superficial gas velocity
Q_{H_q}	$\mathrm{m^3\,h^{-1}}$	Concentrate flow-rate from hopper
$v_{a,k}$	-	Valve fraction
	Exogenous variables	
$d_{p,\max}$	µm	Particle maximum diameter
$d_{p,\min}$	µm	Particle minimum diameter
$G_F^{0,0}$	–	Feed grade
Q_{F_1}	$\mathrm{m^3\,h^{-1}}$	Bank feed flow-rate
ρ_F	$\mathrm{kg\,m^{-3}}$	Feed density

4.2.2 Water recovery

Water recovery is calculated from the superficial gas velocity (J_{g_k}), air recovery (α_k), top-of-froth bubble size (D_{BF_k}) and fluid properties. The relevant equations are (Neethling and Cilliers, 2003)

$$\frac{Q_{w_k}}{A_k} = \begin{cases} \dfrac{J_{g_k}^2 \lambda_{\mathrm{out}_k}}{k_1} (1-\alpha_k)\,\alpha_k & 0 < \alpha_k < 0.5 \\ \dfrac{J_{g_k}^2 \lambda_{\mathrm{out}_k}}{4k_1} & \alpha_k \geq 0.5 \end{cases}, \tag{4.4}$$

$$\lambda_{\mathrm{out}_k} \approx \frac{6.81}{D_{BF_k}^2}, \tag{4.5}$$

$$k_1 = \frac{\rho g}{3\mu C_{PB}}, \tag{4.6}$$

where C_{PB} is the Plateau border drag coefficient, μ is the fluid viscosity, ρ is the fluid density and g is gravitational acceleration.

4.2.3 Mineral recovery

For ease of notation, let the concentration of species i, j in some tank $\triangledown$ be

$$C_{\triangledown}^{i,j} = \frac{M_{\triangledown}^{i,j}}{L_{\triangledown} A_{\triangledown}}. \tag{4.7}$$

Table 4.3. Model parameters.

Parameter	Unit	Description
A_k, A_{H_q}	m^2	Cell and hopper cross-sectional area
α_{OS_k}	–	Air recovery offset
Δ_{h_k}	m	Cell height difference
D_{OS_k}	–	Bubble size offset
D_{BP_k}	mm	Sauter mean bubble diameter in pulp for cell k
C_{PB}	–	Plateau border drag coefficient
C_{v_k}	–	Valve parameter
g	$\mathrm{m\,s}^{-2}$	Gravitational acceleration
H_{cell}	m	Cell height
k_1	–	Fluid properties combined
$K_{0,\alpha_{J_{g_k}}}$	–	Optimal J_{g_k} for $h_{f_k} = 0$
$K_{\alpha_{h_f}}$	–	Effect of h_{f_k} on PAR
$K_{\alpha_{J_g}}$	–	Effect of suboptimal J_{g_k} on α_k
$K_{BS_{J_g}}$	–	Effect of J_{g_k} on D_{BF_k}
K_{BS_λ}	–	Effect of λ_{air_k} on D_{BF_k}
$K^{i,j}$	–	Flotation rate constant for species i, class j
μ	Pa s	Fluid viscosity
Pe	–	Dispersion Peclet number
ρ	$\mathrm{kg\,m}^{-3}$	Fluid density
$\rho_s^{i,j}$	$\mathrm{kg\,m}^{-3}$	Solids density for species i and class j

The mineral masses in the cells ($M_k^{i,j}$) and concentrate tanks ($M_{H_q}^{i,j}$), are modelled using simple mass balances that assume a well-mixed pulp phase

$$\frac{\mathrm{d}M_k^{i,j}}{\mathrm{d}t} = \dot{M}_{F_k}^{i,j} - \dot{M}_{C_k}^{i,j} - C_k^{i,j} Q_{T_k}, \tag{4.8}$$

$$\frac{\mathrm{d}M_{H_q}^{i,j}}{\mathrm{d}t} = \sum_{k=a}^{b} \dot{M}_{C_k}^{i,j} - C_{H_q}^{i,j} Q_{H_q}, \tag{4.9}$$

where a and b denote the range of cells for which the concentrate is collected into tank H_q. Typically $\dot{M}_{F_k}^{i,j} = C_{k-1}^{i,j} Q_{T_{k-1}}$, except for the first cell in which case $\dot{M}_{F_k}^{i,j}$ is determined by the feed.

The mass flow-rates leaving a cell in the concentrate ($\dot{M}_{C_k}^{i,j}$) are calculated as

$$\dot{M}_{C_k}^{i,j} = K^{i,j} M_k^{i,j} S_{b_k} \alpha_k + Ent_{Frac}^{i,j} C_k^{i,j} Q_{w_k}, \tag{4.10}$$

where $K^{i,j}$ is the flotation rate constant and the bubble surface area flux (S_{b_k}) is

$$S_{b_k} = 6 \frac{J_{g_k}}{D_{BP_k}}. \tag{4.11}$$

Typically flotation rate constants are defined to include the effect of S_{b_k}, but as S_{b_k} varies with time, its effect is accounted for separately in (4.10) (Oosthuizen, 2023). Oosthuizen et al. (2021) takes the pulp bubble diameter (D_{BP_k}) as a known constant parameter. This means that it is not estimated on-line, *i.e.,* not included in the **p** of Section 3.4.

The second term in (4.10) quantifies recovery via entrainment using a simplified version of the entrainment factor ($Ent_{Frac}^{i,j}$) model developed by Neethling and Cilliers (2009). The entrainment depends on the minimum and maximum particle diameters of all modelled particle sizes ($d_{p,\min}$ and $d_{p,\max}$), the fluid properties (including C_{PB}), α_k, J_{g_k} and froth height (h_{f_k}). The froth height is taken as the difference between the cell depth (H_{cell}) and L_k

$$h_{f_k} = H_{\text{cell}} - L_k. \tag{4.12}$$

The particle size for which 50 % of the particles in a given mineral species and class are entrained, $d_{prt}^{i,j}$, is calculated as

$$d_{prt}^{i,j} = \sqrt[3]{\frac{\ln(0.5) J_{g_k}^2}{K_{ent}^{i,j} h_{f_k}}}, \tag{4.13}$$

with

$$K_{ent}^{i,j} = \left[\frac{1}{3} \frac{g\left(\rho_s^{i,j} - \rho\right)}{18\mu} \right]^{1.5} \frac{\sqrt{k_1(\sqrt{3} - \pi/2)Pe}}{\sqrt{\alpha_k(1-\alpha_k)}}, \tag{4.14}$$

where $\rho_s^{i,j}$ is the density of the mineral category i, j and Pe is the dimensionless Peclet number which gives the contribution of bulk mass transfer relative to diffusive transport phenomena.

An entrainment factor, $Ent_{Frac}^{i,j}$, for mineral group i, j is then calculated as

$$Ent_{Frac}^{i,j} = \frac{\ln\left(d_{prt}^{i,j}\right) - \ln\left(d_{p,\min}\right)}{\ln\left(d_{p,\max}\right) - \ln\left(d_{p,\min}\right)}, \tag{4.15}$$

where $d_{p,\min}$ and $d_{p,\max}$ are the minimum and maximum particle diameters of all modelled particle sizes (for any mineral species or class).

4.2.4 Froth model

The top of froth bubble size and air recovery are modelled as

$$\frac{\mathrm{d}D_{BF_k}}{\mathrm{d}t} = \frac{K_{BS_{J_g}} J_{g_k} + K_{BS_\lambda} \lambda_{\mathrm{air}_k} + D_{OS_k} - D_{BF_k}}{\lambda_{\mathrm{air}_k}}, \tag{4.16}$$

$$\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} = \frac{K_{\alpha_{J_g}} \left(J_{g_k} - K_{0,\alpha_{J_g}}\right)^2 + K_{\alpha_{h_f}} \left(h_{f_k} - K_{0,\alpha_{h_f}}\right)^2 - (\alpha_k - \alpha_{OS_k})}{\lambda_{\mathrm{air}_k}}, \tag{4.17}$$

where λ_{air_k} is a measure of the froth residence time

$$\lambda_{\mathrm{air}_k} = \frac{h_{f_k}}{J_{g_k}}. \tag{4.18}$$

The parameters $K_{BS_{J_g}}$, K_{BS_λ}, $K_{\alpha_{J_g}}$, $K_{\alpha_{h_f}}$, $K_{0,\alpha_{J_g}}$ and $K_{0,\alpha_{h_f}}$ characterise the effects of J_{g_k} and λ_{air_k} (alternatively h_{f_k}) on D_{BF_k} and α_k. All of these parameters are the same for all cells in the bank. The offset parameters (D_{OS_k} and α_{OS_k}) are assumed to be known constants, effectively making use of deviation variables for D_{BF_k} and α_k. The values of $K_{\alpha_{J_g}}$ and $K_{\alpha_{h_f}}$ are fitted to steady-state experimental data and are assumed to remain constant (are not estimated on-line). The $K_{0,\alpha_{J_g}}$ and $K_{0,\alpha_{h_f}}$ parameters correspond to the values of J_{g_k} and h_{f_k} for which PAR is achieved. Along with C_{v_k}, C_{PB} and $K^{i,j}$, the parameters $K_{BS_{J_g}}$, K_{BS_λ}, $K_{0,\alpha_{J_g}}$ and $K_{0,\alpha_{h_f}}$ are to be estimated from the available on-line measurements

4.2.5 Measurements

For the observer design, models for the plant outputs are required. The states α_k, D_{BF_k}, L_k and L_{H_q} are measured directly on-line. Solids fraction and grade measurements of the concentrate and tails streams are used in Oosthuizen (2023). The solids fraction in the concentrate and the tails (C_{s,H_q} and $C_{s,T}$) are calculated as

$$C_{s,H_q} = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{H_q}^{i,j}}{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{H_q}^{i,j} + \rho\left(A_{H_q} L_{H_q} - \sum_{i=0}^{m}\sum_{j=0}^{n} M_{H_q}^{i,j}/\rho^{i,j}\right)}, \tag{4.19}$$

$$C_{s,T} = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{f}^{i,j}}{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{f}^{i,j} + \rho\left(A_{f} L_{f} - \sum_{i=0}^{m}\sum_{j=0}^{n} M_{f}^{i,j}/\rho^{i,j}\right)}, \tag{4.20}$$

(total solids divided by total solids + water mass, with the water mass calculated from the pulp volume and solids volume). Here, f denotes the final cell in the bank.

The mass fraction desired mineral in concentrate and tails solids, ($G_{H_q}^{0,0}$ and $G_T^{0,0}$) are

$$G_{H_q}^{0,0} = \frac{M_{H_q}^{0,0}}{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{H_q}^{i,j}}, \tag{4.21}$$

$$G_T^{0,0} = \frac{M_{T,f}^{0,0}}{\sum_{i=0}^{m}\sum_{j=0}^{n} M_f^{i,j}}, \tag{4.22}$$

(mass desired in cell/total solids in cell).

4.3 UPDATED MODEL

The model in Section 4.2 has some limitations, which resulted in Oosthuizen (2023) proposing some modifications summarised here. The parameters in (4.17) are assumed to be the same for every cell in the bank and the α_{OS_k} and D_{OS_k} parameters are taken as known constants. While it is tempting to assume identical models for the different cells, it has been shown that in reality the optimal J_{g_k} and h_{f_k} to achieve PAR varies between cells (Hadler et al., 2010, 2012) and that the optimal J_{g_k} and achievable α_k vary rather quickly, even for nearly identical operating conditions (Phillpotts et al., 2021). The changing reagent concentration and floatable particle concentration variation down the bank is likely partly to blame (Wiese et al., 2005, 2011), but there are many other possible factors. Many of these factors also have an impact on the bubble size. Therefore, the adjusted α_k model in (4.23) which allows for varying PAR points and estimating the D_{OS_k} are likely much more realistic.

The updated model is the same as in Section 4.2 except for (4.17) and how α_{OS_k} and D_{OS_k} are treated. Rather than (4.17), the air recovery is modelled with

$$\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} = \frac{K_{\alpha_{J_g}}\left(J_{g_k} - K_{0,\alpha_{J_{g_k}}} - K_{\alpha_{h_f}} h_{f_k}\right)^2 + \alpha_{OS_k} - \alpha_k}{\lambda_{\mathrm{air}_k}}, \tag{4.23}$$

where λ_{air_k} is still a measure of the froth residence time

$$\lambda_{\mathrm{air}_k} = \frac{h_{f_k}}{J_{g_k}}. \tag{4.24}$$

The changed parabolic terms in (4.23) as compared to (4.17) are based on the observations of Hadler et al. (2012) who observed a peak in air recovery with varying h_{f_k} for a fixed J_{g_k}. The J_{g_k} at which the air recovery peaked was also observed to increase with increasing h_{f_k}. In (4.23), the PAR occurs at

$$J_{g_k} = K_{0,\alpha_{J_{g_k}}} + K_{\alpha_{h_f}} h_{f_k}. \tag{4.25}$$

Therefore, $K_{\alpha_{J_g}}$ characterises the impact of a suboptimal J_{g_k} on α_k.

The parameters $K_{BS_{J_g}}$, K_{BS_λ}, $K_{\alpha_{J_g}}$, $K_{\alpha_{h_f}}$, $K_{0,\alpha_{J_{g_k}}}$ characterise the effects of J_{g_k} and λ_{air_k} (alternatively h_{f_k}) on D_{BF_k} and α_k. Importantly, each cell has its own value for $K_{0,\alpha_{J_{g_k}}}$ which is also meant to be updated in real-time along with α_{OS_k}, D_{OS_k}.

Oosthuizen (2023) also included dynamic equations for J_{g_k} to obtain full non-linear observability. He uses

$$\frac{\mathrm{d}J_{g_k}}{\mathrm{d}t} = \frac{J_{g_k,SP} - J_{g_k}}{\tau_{J_g}}, \tag{4.26}$$

where $J_{g_k,SP}$ is the set-point for J_{g_k} and τ_{J_g} is the first-order time constant for the air valve, assumed as 1 (Oosthuizen, 2023). However, as discussed in Section 4.5, the observability rank gained by this addition disappears when J_{g_k} reaches the set-point. For the purpose of this study, J_{g_k} is assumed to be a measured input, as this reduces the number of states that must be reconciled during state estimation without loss of observability.

Finally, for the purpose of this study, the C_{s,H_q} and $C_{s,T}$ measurements used in Oosthuizen (2023) were replaced by concentrate (ρ_{H_q}) and tails (ρ_T) density measurements modelled as

$$\rho_{H_q} = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{H_q}^{i,j} + \rho\left(A_{H_q}L_{H_q} - \sum_{i=0}^{m}\sum_{j=0}^{n} M_{H_q}^{i,j}/\rho^{i,j}\right)}{A_{H_q}L_{H_q}}, \tag{4.27}$$

$$\rho_T = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n} M_{f}^{i,j} + \rho\left(A_{f}L_{f} - \sum_{i=0}^{m}\sum_{j=0}^{n} M_{f}^{i,j}/\rho^{i,j}\right)}{A_{f}L_{f}}. \tag{4.28}$$

where f denotes the final cell.

4.4 MODEL CAUSALITY

The interconnected nature of the model equations and relationships between variables from different parts of the model are not always clear from the model equations. Figure 4.2 shows the connections between the different variables, the causal flow of information within the model and the variable classification. This visualisation approach was based on the causality diagram in Steyn and Sandrock (2021) and should be seen as a visual aid for the observability analyses in Section 4.5.

From the number of connections with $\dot{M}_{C_k}^{i,j}$ it is clear that this variable is central to much of the model behaviour. The D_{BF_k} variables only really play a role in the water recovery (and consequently entrainment), yet it is affected by several empirical parameters. The disparity between the number of measurements "close to" D_{BF_k} and the parameters that need to be estimated could lead to issues with practical parameter identifiability. This is discussed in more detail in Section 4.5.2.

The α_k values impact entrainment, flotation rate and water recovery. J_{g_k} not only impacts the α_k differential equation, but also determines the measured α_k value, which could lead to discontinuities in the state measurement if there are step changes in J_{g_k}. In that regard, treating J_{g_k} as a state with

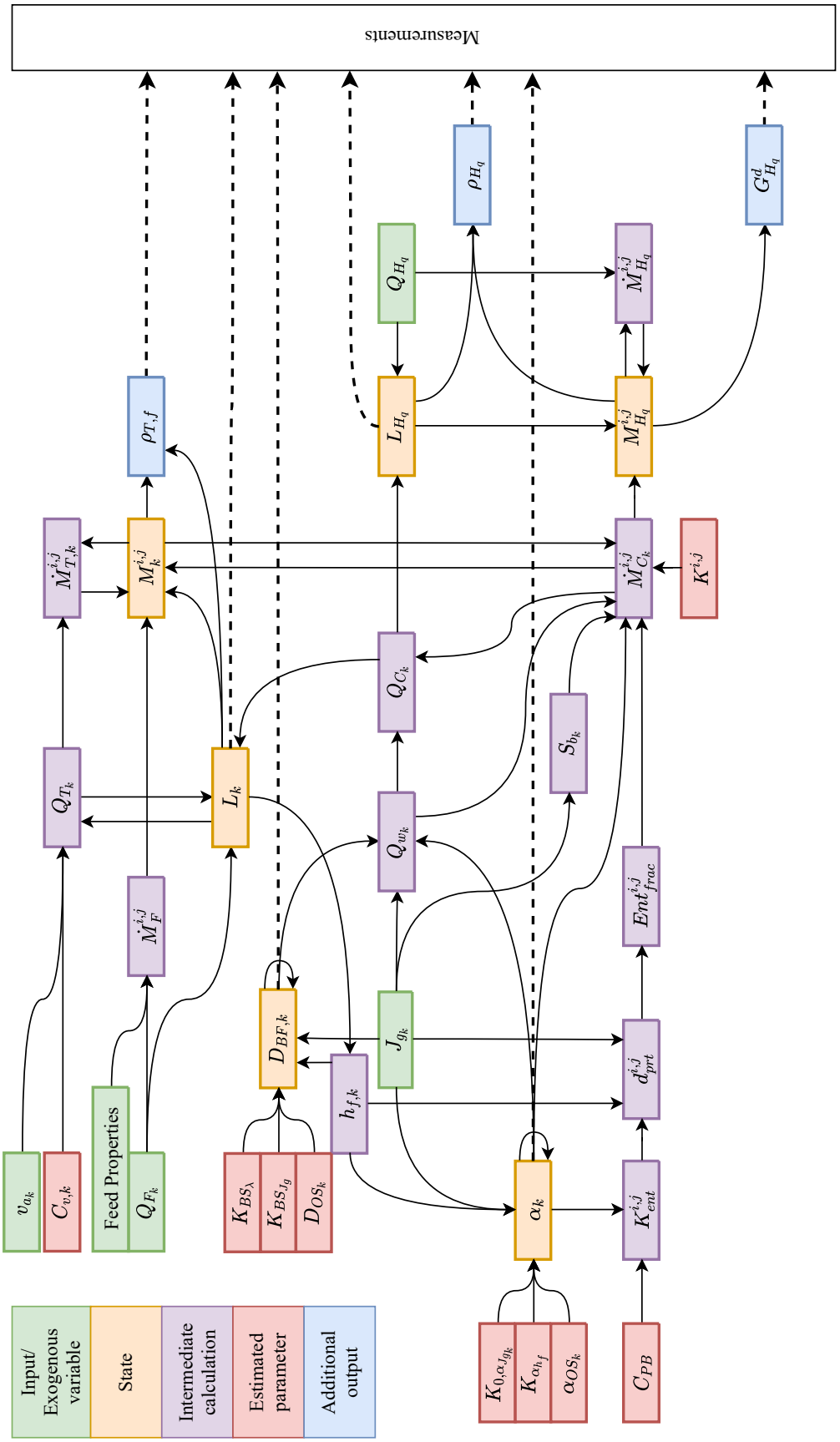

Figure 4.2. Model information flow diagram. This shows the interaction between the various parameters and variables in the Oosthuizen (2023) model.

set-point input as in (4.26) makes sense. Similar arguments regarding parameter identifiability hold as with D_{BF_k}.

4.5 OBSERVABILITY

4.5.1 Existing observability analyses

The observability analyses available in literature (Oosthuizen et al., 2021; Oosthuizen, 2023) for the models described in Section 4.2 and Section 4.3 are summarised in Section 4.5.1.1 and Section 4.5.1.2.

The analyses were conducted using the data from the four cell rougher circuit described in Hadler et al. (2010). Oosthuizen et al. (2021) grouped all the minerals into either “desired" or “gangue". The i, j superscripts are replaced by $0,0$ or $1,0$ for the two groups respectively. On-line measurements of the feed, tailings and concentrate desired mineral content were assumed to be available.

4.5.1.1 Original model

For the original model in Section 4.2, the resulting system of states ($\mathbf{x}$), estimated parameters ($\mathbf{p}$), inputs ($\mathbf{u}$) and outputs ($\mathbf{y}$) is

$$\mathbf{x} = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad M_{H_q}^{i,j}, \quad M_k^{i,j}\right]^T, \tag{4.29}$$

$$\mathbf{p} = \left[C_{PB}, \quad C_{v_k}, \quad K^{i,j}, \quad K_{0,\alpha_{J_g}}, \quad K_{0,\alpha_{h_f}}, \quad K_{BS_{J_g}}, \quad K_{BS_\lambda}\right]^T, \tag{4.30}$$

$$\mathbf{u} = \left[J_{g_k}, \quad v_{a,k}, \quad Q_{H_q}\right]^T, \tag{4.31}$$

$$\mathbf{y} = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad G_{H_q}^{0,0}, \quad G_T^{0,0}, \quad C_{s,H_q}, \quad C_{s,T}\right]^T. \tag{4.32}$$

Here (as well as all subsequent instances) the i, j superscripts and k and H_q subscripts imply the full range of species, mineral classes, flotation cells and concentrate tanks, unless given a specific value (as is the case with $G_{H_q}^{0,0}$). Using the nominal variable and parameter values in Oosthuizen et al. (2021), the linear observability matrix (see Section 3.3) for the system was shown to have full rank.

More detailed observability analyses were conducted for the reagent model expansions in Chapter 5. As these results are very similar to the original results in Oosthuizen et al. (2021), no further observability analyses on the “original” model are reported on here. All analyses in the remainder of the chapter are based on the updated model as described in Section 4.3.

4.5.1.2 Updated model

Oosthuizen (2023) performed an observability analysis on the model in Section 4.3 and found that the model states

$$\mathbf{x}_1 = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad J_{g_k}, \quad M_{H_q}^{i,j}, \quad M_k^{i,j}\right]^T, \tag{4.33}$$

and the parameters

$$\mathbf{p}_1 = \left[\alpha_{OS_k}, \quad C_{PB}, \quad C_{v_k}, \quad D_{OS_k}, \quad K^{i,j}, \quad K_{0,\alpha_{J_{g_k}}}, \quad K_{\alpha_{h_f}}, \quad K_{BS_{J_g}}, \quad K_{BS_\lambda}\right]^T, \tag{4.34}$$

are non-linearly observable and identifiable given the inputs

$$\mathbf{u}_1 = \left[J_{g,SP}, \quad Q_H, \quad v_{a,k}\right]^T, \tag{4.35}$$

and measurements

$$\mathbf{y}_1 = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad G_{H_q}^{0,0}, \quad G_T^{0,0}, \quad C_{s,H_q}, \quad C_{s,T}\right]^T. \tag{4.36}$$

Note that $K_{\alpha_{J_g}}$ was assumed to be a known constant as it was unidentifiable with the given model structure and available measurements. They also point out that $K_{BS_{J_g}}$ is only observable if the J_{g_k} dynamics (described in (4.26)) are included.

4.5.2 Additional analysis and results

As the observability analyses in Section 4.5.1 are numeric, they can only evaluate the local observability (*i.e.,* around an operating point) and it is possible that there are unobservable or unidentifiable operating points (Villaverde et al., 2019). However, given the possibility of different plant configurations (number of cells, concentrate tanks and recyle streams) and the complexity of the symbolic calculations involved, evaluating the generalised observability of the models in this chapter is impractical, if not impossible. Implementing the model and parameter estimation approach of Oosthuizen et al. (2021) on other plants, will therefore require additional observability checks based on the particular plant configuration and likely operating points.

In order to investigate the possible limitations that could arise at specific operating points and the plant configuration encountered at MNC, the existing observability analyses were expanded as described in the following three sections (4.5.2.1, 4.5.2.2 and 4.5.2.3). The analyses were done for the MNC rougher bank as shown in Figure 4.3. The observability results were confirmed at at least three different operating points (*i.e.,* state and parameter values) obtained from the MNC plant historian data. The parameters in (4.34) were fitted to plant data as far as possible. Values had to be assumed for C_{PB}, $K_{\alpha_{J_g}}$, $K_{\alpha_{h_f}}$, $K_{BS_{J_g}}$ and $K^{i,j}$ based on the values in Oosthuizen (2023) and Steyn and Sandrock (2021). The

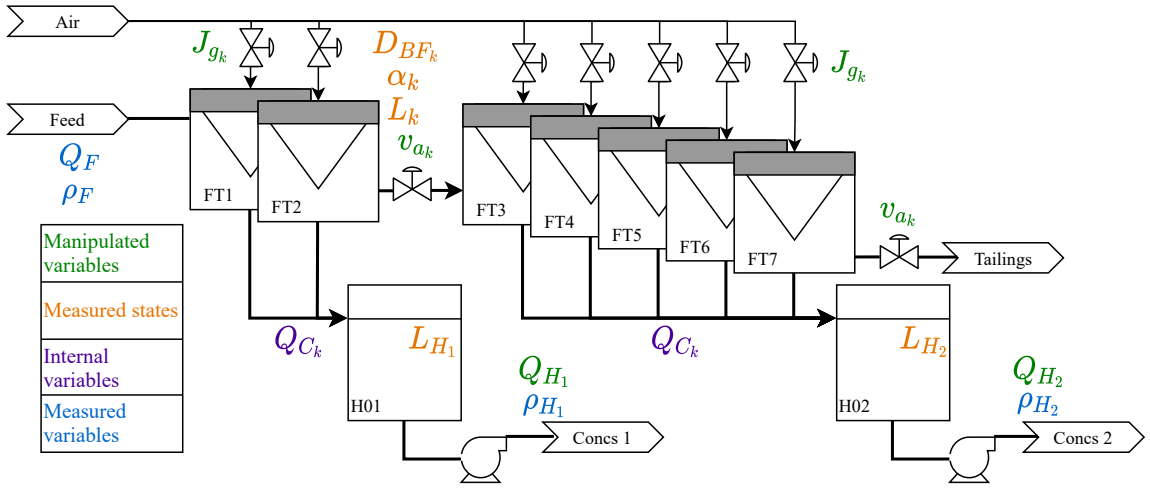

Figure 4.3. Second rougher bank at Mogalakwena North Concentrator. Variables that are applicable to multiple cells are indicated with a subscript k.

remaining parameter values in Table 4.3 were either obtained from plant personnel (A_k, A_{H_q}, Δ_{h_k} and H_{cell}), or given assumed values based on a mineralogy report of the plant ($\rho_s^{i,j}$) and literature (μ, ρ, Pe) (Hadler et al., 2010; Oosthuizen, 2023). The bank profiles in Steyn and Sandrock (2021) along with steady-state assumptions were used to estimate reasonable state values for the component masses. Note that at no point was the LOIC satisfied for the updated model.

4.5.2.1 Superficial gas velocity dynamics

As mentioned in Section 4.5.1.2, Oosthuizen (2023) found that the updated model was only fully observable when J_{g_k} dynamics were included (see (4.26)). However, his observability analysis was conducted at an operating point where $J_{g_k} \neq J_{g,SP}$. If the $J_{g,SP}$ are held constant, the J_{g_k} values eventually reach their set-points. Re-evaluating the observability of the system in Section 4.5.1.2 with $J_{g_k} = J_{g,SP}$ results in a d0 rank deficiency of 1, thus loss of observability. The $K_{BS_{J_g}}$ and D_{OS_k} were identified as the possibly unidentifiable parameters (PUPs) (see Section 3.3.4).

The observability analysis in Section 4.5.1.2 was repeated without the J_{g_k} dynamics (J_{g_k} was treated as a known input), which assumes that J_{g_k} reaches $J_{g,SP}$ immediately. When the derivatives of $\mathbf{u}$ were taken as 0, the system was unobservable with a rank deficiency of 1 and $K_{BS_{J_g}}$ and D_{OS_k} were again identified as PUPs. Full observability was achieved when a non-zero first derivative $\mathbf{u}^{(1)}$ for any of the J_{g_k} values was specified (see Section 3.3.2) or if at least one of the PUPs are assumed to be a known constant (*i.e.,* not included in $\mathbf{p}$).

For all subsequent analyses, the J_{g_k} dynamics are neglected and the modified state and input vectors

$$\mathbf{x}_2 = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad M_{H_q}^{i,j}, \quad M_k^{i,j}\right]^T, \tag{4.37}$$

$$\mathbf{u}_2 = \left[J_{g_k}, \quad Q_{H_q}, \quad v_{a,k}\right]^T, \tag{4.38}$$

are used. A reduced parameter vector

$$\mathbf{p}_2 = \left[\alpha_{OS_k}, \quad C_{PB}, \quad C_{v_k}, \quad D_{OS_k}, \quad K^{i,j}, \quad K_{0,\alpha_{J_{g_k}}}, \quad K_{\alpha_{h_f}}, \quad K_{BS_\lambda}\right]^T, \tag{4.39}$$

is also used. The value for $K_{BS_{J_g}}$ was assumed to be a known constant as it would require J_{g_k} input excitation of at least a ramp to be observable. In Section 5.3 $K_{BS_{J_g}}$ is shown to have very low sensitivity indices relative to K_{BS_λ} and D_{OS_k}, making its exact value less important. This assumption makes all D_{OS_k} identifiable without requiring input excitation.

4.5.2.2 Concentrate grade measurements

Both Oosthuizen et al. (2021) and Oosthuizen (2023) assume that on-line compositional analyses of the desired mineral is available for the feed, concentrate and tailings. Such measurements are possible using XRF analysers or BlueCube analysers, yet are not commonly implemented on industrial flotation plants (Shean and Cilliers, 2011; Oosthuizen et al., 2017). For example, at MNC there are shiftly composite assays of Pt, Pd and Rh content (3E) of the feed and tailings to and from the rougher banks and there is the possibility of XRF analysis of these two streams. However, there are no samples (composite or on-line) of the two concentrate streams leaving each rougher bank. The impact of the missing concentrate grade measurements is investigated below.

Removing the concentrate grade measurements for both concentrate streams of the MNC model, yields a rank deficiency of 2 with all four $M_{H_q}^{i,j}$ (assuming two mineral species and two concentrate tanks) as the possibly unobservable variables at the $L_f^5 g$ level. No first-order input excitation could make $M_{H_q}^{i,j}$ observable and the rank deficiency remained 2.

In order to determine the minimum measurement requirement for observability of the full updated model, a combined concentrate grade measurement ($G_{C,t}^{0,0}$) was added to the measurements with

$$G_{C,t}^{0,0} = \frac{\sum_{q=1}^{2} Q_{H_q} C_{H_q}^{0,0}}{\sum_{q=1}^{2} Q_{H_q}\left(C_{H_q}^{0,0} + C_{H_q}^{1,0}\right)}. \tag{4.40}$$

The output vector used for the analysis is

$$\mathbf{y}_2 = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad G_{C,t}^{0,0}, \quad G_T^{0,0}, \quad \rho_{H_q}, \quad \rho_T\right]^T. \tag{4.41}$$

The addition of $G_{C,t}^{0,0}$ ensures that all $M_{H_q}^{i,j}$ are observable. Table 4.4 summarises the number of variables after expanding the vectors $\mathbf{x}_2$, $\mathbf{y}_2$, $\mathbf{p}_2$ and $\mathbf{u}_2$ for the seven cells ($k \in \{1,2,\ldots,7\}$), two mineral groups ($i \in \{0,1\}$ and $j = 0$) and two concentrate tanks ($q \in \{1,2\}$). The rank results of the observability matrix for increasing Lie derivative order are summarised in Table 4.5.

Table 4.4. Number of variables.

n_x	n_p	n_y	n_u
41	33	28	16

Table 4.5. Observability analysis results.

Lie derivative	With $G_{C,t}^{0,0}$		Without $G_{C,t}^{0,0}$	
	d$\mathcal{O}$ Rank	Rank deficiency	d$\mathcal{O}$ Rank	Rank deficiency
$L_f^0 g$	28	46	27	47
$L_f^1 g$	56	18	54	20
$L_f^2 g$	73	1	71	3
$L_f^3 g$	74	0	72	2
$L_f^4 g$	74	0	72	2

The order of $L_f^i g$ at which d$\mathcal{O}$ achieves full rank gives and indication of the minimum order of non-zero output derivatives required for observability. Table 4.5 shows that at least third-order non-zero output derivatives are required for observability. At $L_f^2 g$, there is a rank deficiency of 1, with all $K_{\alpha_{h_f}}$, $K_{0,\alpha_{J_{g_k}}}$ and α_{OS_k} parameters identified as PUPs. For $L_f^1 g$ the rank deficiency is 18, with most of the mineral masses, D_{BF_k} model parameters and C_{v_k} added to the list of potentially unobservable states and parameters. This indicates possible practical observability, and consequently estimation, issues if the system outputs do not vary enough. These issues are expected to be encountered first in the α_k model parameters, as they required higher order derivatives.

One of the main goals of control is to limit variability in the outputs of a system. By stabilising the plant, the higher order derivatives are also reduced. This is observed in the MNC data, with various measurements (*e.g.*, L_k and ρ_T) remaining fairly constant over long periods of time. This could make effective state and parameter estimation quite difficult, as the estimation algorithm would have to rely on very small changes.

Even if there were sufficient excitation and movement in the outputs, the results in Table 4.5 show that the state and parameter estimation of the model as described in Section 4.3 would not be possible without on-line (or relatively frequent) compositional measurements of the concentrate from the bank. It also seems extremely unlikely that any form of feed-back model-based control approach that accounts for the recovery of different mineral species (as would be desirable for flotation grade and recovery optimisation (Oosthuizen et al., 2017; Quintanilla, Neethling and Brito-Parada, 2021; Steyn and Sandrock, 2021; Oosthuizen, 2023)) could succeed without compositional measurements.

4.5.2.3 Multiple species

In both existing analyses in Section 4.5.1, only two mineral species (desired and gangue minerals, labelled $0,0$ and $1,0$) were considered. However, Steyn and Sandrock (2021) indicated that there are three clear floatability groups in the minerals processed at MNC. The observability of the model was, therefore, also evaluated for three mineral categories. The same system and data as in Section 4.5.2.2 was used, with desired and gangue minerals. However, the gangue content in the system was reduced and substituted with a third, fictional, species ($i = 2$) of intermediate floatability. The compositional measurements used for the numeric operating point were synthesised using steady-state mass balances.

Full observability was only achieved if at least two out of the three component compositions were measured on-line in the tails and combined concentrate. An example of a valid measurement vector is

$$\mathbf{y}_2 = \left[\alpha_k, \quad D_{BF_k}, \quad L_{H_q}, \quad L_k, \quad G_{C,t}^{0,0}, \quad G_{C,t}^{2,0}, \quad G_T^{0,0}, \quad G_T^{2,0}, \quad \rho_{H_q}, \quad \rho_T\right]^T. \tag{4.42}$$

As already mentioned in Section 4.5.2.2, the available compositional analyses at MNC currently measure only the 3E content. The 3E minerals are typically very closely related and associate with the same minerals Steyn and Sandrock (2021), which makes it impossible to construct compositional measurements for three mineral groups from only the 3E content of a stream. As such, any implementation of full updated model at MNC would currently be limited to two species. If additional compositional measurements that can differentiate between more minerals are made available, more species could be included in an observable model implementation.

4.6 CONCLUDING REMARKS

The two flotation circuit model versions in Oosthuizen et al. (2021) and Oosthuizen (2023), used throughout the rest of this work, were summarised along with the results of existing observability

analyses for both. It was shown that adding the J_{g_k} dynamics used in Oosthuizen (2023) does not improve the practical observability of the system, as observability is lost when J_{g_k} reaches its set-point. The observability analysis of the "updated" model was also expanded in order to investigate the limitations imposed by removing compositional measurements.

It was shown that the absence of on-line compositional analysers for the concentrate streams severely limit the observability of the updated model. However, if multiple concentrate streams are present, observability is possible without requiring analysers on each stream as a combined concentrate compositional measurement is sufficient. It was also shown that if three mineral species are to be modelled, the concentrate and tailings content for at least two of the three species would need to be measured for observability.

For all the observability analyses in this chapter, the order of the Lie derivative at which observability was achieved highlighted possible issues with practical observability and identifiability. High order output derivatives were required for full observability and identifiability, which is counter to the typical goal of control: the reduction of output variability.

Interestingly, the grouping of PUPs (as in Section 4.5.2.1) was quite a common phenomenon during the observability analyses conducted for this study (not all reported here). Depending on which measurements were removed, the rank deficiency of $\mathrm{d}\mathcal{O}$ would increase by one or two, but a whole group of parameters would become possibly unidentifiable. If one parameter in the group is assumed to be a known constant, the overall dimension of the problem would reduce by one, while the rank of $\mathrm{d}\mathcal{O}$ remained the same (decreasing the rank deficiency). This meant that all other parameters in that particular group became identifiable. Often, multiple groups would be present, typically related to different equations in the model, *e.g.,* the α_k or D_{BF_k} model parameters would group separately. To the knowledge of the author, this phenomenon is not well-known in literature. While not particularly relevant to the current study, further formal investigation could be valuable. For instance, this could help identify sections of a plant that are under-instrumented or determine control-critical measurements which could cause the entire control approach to fail. It could also help to identify sub-models that are identifiable in isolation, which would allow a more robust distributed state estimation approach or the use of, for example, dual estimation (Olivier et al., 2012).

CHAPTER 5 FROTHER MODEL FEASIBILITY

5.1 CHAPTER OVERVIEW

The main body of work in this chapter has already been published in Venter et al. (2022). In this work a variant of the model in Oosthuizen et al. (2021) is used as described in Section 4.2. Venter et al. (2022) demonstrates that, under certain conditions, obtaining observable empirical models of the frother effect is possible. The parameters that would be most affected by reagent conditions are identified (Section 5.2) and a Sobol sensitivity analysis is used to determine the section of the model to expand (Section 5.3). The proposed model expansions are discussed in Section 5.4 with the simulation set-up and results for these models discussed in Section 5.5. A similar model expansion of the updated model in Section 4.3 was also investigated and the results are included in Addendum A.

5.2 IMPACT OF REAGENTS

Collectors are used to preferentially enhance the floatability of the desired mineral, but can also increase the floatability of gangue if used in excess (Wills and Finch, 2015). Wiese et al. (2011) have shown that changes in bubble loading due to collector variations can also have an impact on froth stability. However, for this study the collector concentration is assumed to mainly impact the floatability rate constants, $K^{i,j}$.

Frother addition is aimed at reducing the pulp bubble size and improving froth formation and stability (Harris et al., 2013; Wills and Finch, 2015). Increasing the frother dosage only reduces D_{BP} up to a point known as the critical coalescence concentration (CCC), beyond which D_{BP} remains a function of the aeration rate and dispersion mechanism (Finch et al., 2008; Harris et al., 2013; Wills and Finch, 2015). While the frother dosage typically surpasses the CCC (Harris et al., 2013) and D_{BP} is not expected to vary much, it is included in the sensitivity analysis in Section 5.3 for completeness.

The effects of frother dosage on froth stability would likely be modelled best as part of the D_{BF_k} and α_k state equations (see (4.16) and (4.23)), as these states are related to the rates of bubble coalescence, bursting and froth structure (Oosthuizen et al., 2021). Unfortunately, due to the empirical nature of some of the parameters in these equations ($K_{BS_{J_g}}$, K_{BS_λ}, $K_{\alpha_{J_g}}$, $K_{\alpha_{h_f}}$, $K_{0,\alpha_{J_g}}$ and $K_{0,\alpha_{h_f}}$) and the lack of experimental data with varying reagent dosages for the simulated circuit, it is difficult to determine which of these parameters would change with reagent concentration and to what extent.

This implies modelling one or more of the parameters in (4.16) or (4.23) ($K_{BS_{J_g}}$, K_{BS_λ}, $K_{\alpha_{J_g}}$, $K_{\alpha_{h_f}}$, $K_{0,\alpha_{J_g}}$ and $K_{0,\alpha_{h_f}}$) as functions of frother concentration. For this reason, the model expansion for the frother effects will be guided by the sensitivity analysis and observability results.

While it is possible that the reagents, especially the frother, could affect the Pe or μ values used in the calculation, typical reagent concentrations are in the range 5 ppm to 30 ppm (Finch et al., 2008; Harris et al., 2013; Wills and Finch, 2015). It is unlikely that changes in this concentration would impact these parameters significantly, if at all and therefore they were excluded from the current study.

The Plateau border drag coefficient is also considered. Together with D_{BF_k} and α_k, C_{PB} plays a role in determining the water recovery, Q_{C_k}, and entrainment of particles (Neethling and Cilliers, 2009; Oosthuizen et al., 2021). For the purpose of this study, C_{PB} will be considered in the sensitivity analysis as being affected by the frother concentration.

5.3 SENSITIVITY ANALYSIS RESULTS

A Sobol sensitivity analysis was conducted on the model parameters identified above. The parameters considered as input factors, their nominal values and sensitivity indexes for steady-state grade and recovery are summarised in Table 5.1. Only two mineral groups were considered: the desired mineral, $\{0,0\}$, and gangue, $\{1,0\}$. In the absence of data on parameter values over a larger operating range or other reagent concentrations, the parameters were all allowed to vary by 10 % of their nominal value, based on plant data in Hadler et al. (2010).

In PGM processing, there is usually a large difference in the concentrations of the desired mineral and gangue. For the plant under consideration, the desired mineral mass fraction in a cell is less than about 0.015. The desired mineral floatability constant ($K^{0,0}$) is also much larger than that of the gangue ($K^{1,0}$). These factors result in true flotation dominating the recovery of desired mineral in the

Table 5.1. Sensitivity indexes (in percentage).

Parameter	Nominal value	Grade		Recovery	
		S_1	S_T	S_1	S_T
C_{PB}	12.0	13.0	13.4	0.0	0.0
D_{BP}	0.6 mm	0.9	1.2	5.1	5.1
$K^{0,0}$	2.3	1.5	1.7	5.0	5.1
$K^{1,0}$	2.14×10^{-4}	0.0	0.0	0.0	0.0
$K_{0,\alpha_{J_g}}$	8.42	23.3	23.7	14.5	16.7
$K_{0,\alpha_{h_f}}$	94.9	49.3	49.8	70.7	74.6
$K_{\alpha_{J_g}}$	-1.9×10^{-2}	0.1	0.4	0.8	0.4
$K_{\alpha_{h_f}}$	-1.0×10^{-4}	0.9	1.2	1.8	2.5
$K_{BS_{J_g}}$	0.067	6.2	6.7	0.0	0.0
K_{BS_λ}	0.032	3.4	3.4	0.0	0.0

concentrate ($\dot{M}_{C_k}^{0,0}$) and entrainment dominating the mass flow rate of gangue in the concentrate ($\dot{M}_{C_k}^{1,0}$). As the recovery is mostly determined by true flotation, it is expected to be sensitive only to factors that influence true flotation. The grade will additionally be influenced by all factors that have an impact on entrainment.

The parameter with the highest sensitivity for both grade and recovery determination is $K_{0,\alpha_{h_f}}$, followed by $K_{0,\alpha_{J_g}}$ (these parameters have the highest S_1 and S_T values in Table 5.1). This implies that air recovery is probably the dominating factor in the water recovery, entrainment and flotation models.

The sensitivity indexes for $K^{1,0}$, $K_{\alpha_{J_g}}$ and $K_{\alpha_{h_f}}$ were small enough to justify assuming constant values for these parameters. While C_{PB}, $K_{BS_{J_g}}$ and K_{BS_λ} were the next most important for grade, they had no real impact on recovery as they only affect entrainment directly. The grade sensitivity towards C_{PB} warrants future investigation into the impact reagents might have on this parameter.

5.4 PROPOSED MODEL EXPANSIONS

The observability and identifiability of the model expansions were evaluated as in Section 4.5. Although not used in the observability analysis by Oosthuizen et al. (2021), the frother addition rate to the first

cell ($\dot{M}^f_{F_1}$) is assumed to be measured on-line and will be used in the current observability analysis as an MV.

The frother mass in each cell (M^f_k) was modelled using simple mass balances

$$\frac{\mathrm{d}M^f_k}{\mathrm{d}t} = \dot{M}^f_{F_k} - C^f_k(Q_{T_k} + Q_{C_k}), \tag{5.1}$$

where the frother mass in-flow rate ($\dot{M}^f_{F_k}$) will typically be the out-flow rate from the previous cell. The exception is the first cell in the bank, where $\dot{M}^f_{F_1}$ is an additional MV. This assumes that the frother leaves the cell in the concentrate and tailings based on the concentration in the pulp, C^f_k, which is

$$C^f_k = \frac{M^f_k}{L_k A_k}. \tag{5.2}$$

While this is certainly not the case, the lack of data regarding the split of frother between the concentrate and tailings prevents the development of a more accurate, yet relatively simple, model fit for the current analysis.

As the most influential parameters identified in the sensitivity analysis ($K_{0,\alpha_{h_f}}$ and $K_{0,\alpha_{J_g}}$) are primarily influenced by the frother, the collector concentrations were not modelled. This results in the $\mathbf{x}$ and $\mathbf{u}$ vectors in (4.29) and (4.31) being updated to

$$\mathbf{x}_3 = \left[\alpha_k, \quad D_{BF_k}, \quad L_k, \quad L_H, \quad M^{i,j}_k, \quad M^{i,j}_H, \quad M^f_k\right]^T, \tag{5.3}$$

$$\mathbf{u}_3 = \left[J_{g_k}, \quad v_{a,k}, \quad Q_H, \quad \dot{M}^f_{F_1}\right]^T. \tag{5.4}$$

The measured outputs in (4.32) remain the same.

5.4.1 Model expansion 1 (ME1)

At first a simple linear expansion

$$K_{0,\alpha_{h_f},k} = K_{CF} C^f_k + K_{CF,0}, \tag{5.5}$$

of the most important parameter, $K_{0,\alpha_{h_f}}$, is investigated. Here K_{CF} quantifies the impact of C^f_k on $K_{0,\alpha_{h_f}}$ and $K_{CF,0}$ is the value of $K_{0,\alpha_{h_f}}$ when the frother concentration is 0. This is a very naive model as there is no supporting literature (to the best knowledge of the author) that the frother concentration would change the optimal froth height for peak air recovery. However, there is literature (such as McFadzean et al. (2016)) that shows the effect of frother on maximum froth height. Furthermore, due to the empirical nature of the model, it is not impossible for the frother to have an effect on $K_{0,\alpha_{h_f}}$.

All states (including frother concentrations (C_k^f)) are observable and all estimated parameters (compare to (4.30))

$$\mathbf{p}_3 = \left[C_{PB},\quad C_{v_k},\quad K^{i,j},\quad K_{BS_{J_g}},\quad K_{BS_\lambda},\quad K_{0,\alpha_{J_g}},\quad K_{CF},\quad K_{CF,0}\right]^T, \tag{5.6}$$

are identifiable from the available measurements and known MVs. The linear observability analysis results are summarised in Table 5.2. A distinction is made between the PUPs and the possibly unobservable states (PUXs). While the LOIC is satisfied, full rank is only reached at $\mathbf{CA}^2$, which could indicate possible observability issues if the higher order output derivatives are small.

Table 5.2. ME1 Linear observability results.

Order	Rank	Deficiency	PUXs	#PUXs	PUPs	#PUPs
$\mathbf{CA}^0$	17	22	$M^{i,j}_{k\in\{1:3\}}, M^f_k$	10	$C_{PB}, C_{v_k}, K_{0,\alpha_{J_g}}, K_{BS_{J_g}}, K_{BS_\lambda}, K_{CF}, K_{CF,0}, K^{i,j}$	12
$\mathbf{CA}^1$	32	4	$M^{i,j}_{k\in\{1:3\}}, M^f_k$	10	$C_{PB}, C_{v_{k\in\{1:3\}}}, K_{0,\alpha_{J_g}}, K_{CF}, K_{CF,0}, K^{i,j}$	9
$\mathbf{CA}^2$	39	0	–	0	-	0

5.4.2 Model expansion 2 (ME2)

An expansion less naive than ME1 must consider process knowledge. In the proposed expansion the steady-state value of α_k is modified by adding $K_{F\alpha}(C_k^f - K_{F\alpha,0})$ to the numerator of (4.17)

$$\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} = \frac{K_{\alpha_{J_g}}\left(J_{g_k} - K_{0,\alpha_{J_g}}\right)^2 + K_{\alpha_{h_f}}\left(h_{f_k} - K_{0,\alpha_{h_f}}\right)^2 + K_{F\alpha}(C_k^f - K_{F\alpha,0}) - (\alpha_k - \alpha_{OS_k})}{\lambda_{\mathrm{air}_k}}. \tag{5.7}$$

The frother concentrations remain observable and both of the additional parameters ($K_{F\alpha}$ and $K_{F\alpha,0}$) are identifiable in addition to the original parameters, resulting in

$$\mathbf{p}_4 = \left[C_{PB},\quad C_{v_k},\quad K^{i,j},\quad K_{BS_{J_g}},\quad K_{BS_\lambda},\quad K_{0,\alpha_{J_g}},\quad K_{0,\alpha_{h_f}},\quad K_{F\alpha},\quad K_{F\alpha,0}\right]^T. \tag{5.8}$$

While it is unlikely that the effect of concentration on the steady-state air recovery would actually be linear, real-time parameter estimation could update this simple model to track the operating region of the system. The linear observability analysis results are summarised in Table 5.3. Overall, the results are almost identical to that of ME1 and again full rank is reached at $\mathbf{CA}^2$, with the same implications.

The fact that modelled frother concentrations are observable implies that such concentrations can be estimated without any concentration measurements. This is very promising as reagent concentrations

Table 5.3. ME2 linear observability results.

Order	Rank	Deficiency	PUXs	#PUXs	PUPs	#PUPs
$\mathbf{CA}^0$	17	23	$M^{i,j}_{k\in\{1:3\}}, M^f_k$	10	C_{PB}, C_{v_k}, $K_{0,\alpha_{h_f}}$, $K_{0,\alpha_{J_g}}$, $K_{BS_{J_g}}$, K_{BS_λ}, $K_{F\alpha}$, $K_{F\alpha,0}$, $K^{i,j}$	13
$\mathbf{CA}^1$	32	8	$M^{i,j}_{k\in\{1:3\}}, M^f_k$	10	C_{PB}, $C_{v_{k\in\{1:3\}}}$, $K_{0,\alpha_{h_f}}$, $K_{0,\alpha_{J_g}}$, $K_{F\alpha}$, $K_{F\alpha,0}$, $K^{i,j}$	10
$\mathbf{CA}^2$	40	0	–	0	-	0

in flotation plants are notoriously difficult to measure and on-line reagent concentration measuring techniques have yet to be developed (Gélinas and Finch, 2007; Maldonado, Desbiens, del Villar and Aguilera, 2010). While Knight and Knights (2011) report some success with an on-line xanthate probe measuring the residual collector concentration, any such probe has yet to reach the market. The lack of such measurements makes any control strategy involving reagents very challenging. With concentration estimates, it is possible to include frother addition rate into a model-based control strategy.

5.5 SIMULATION

5.5.1 Simulation set-up

The expanded models of Sections 5.4.1 and 5.4.2 were simulated using the variable values in Oosthuizen et al. (2021), as summarised in Table 5.4 and Table 5.5. The nominal values for the estimated parameters are given in Table 5.1. The bank of cells and most important variables are shown in Figure 5.1. A solids content of $444\,\mathrm{kg\,m^{-3}}$ in the feed was assumed. This translates to a solids mass fraction of 35 wt%. To achieve a frother concentration in the slurry around a typical 10 ppm (mass basis) (Harris et al., 2013), a frother dosage of between $0.01\,\mathrm{kg\,m^{-3}}$ and $0.02\,\mathrm{kg\,m^{-3}}$ is required. An initial concentration of $C^f_k = 0.01\,\mathrm{kg\,m^{-3}}$ was chosen for all cells and the feed composition. Without data containing varying frother dosages, the parameters for the model expansions could not be fitted properly. Instead, their nominal values (see Table 5.5) were chosen such that the model steady-state remains the same as in Oosthuizen et al. (2021) for the chosen nominal concentration. The following simulation results should therefore only be interpreted qualitatively to determine whether the model expansions have predictive potential.

Each run was simulated for a process time of 60 min with step changes in $\dot{M}^f_{F_1}$ as shown in Figure 5.2. As

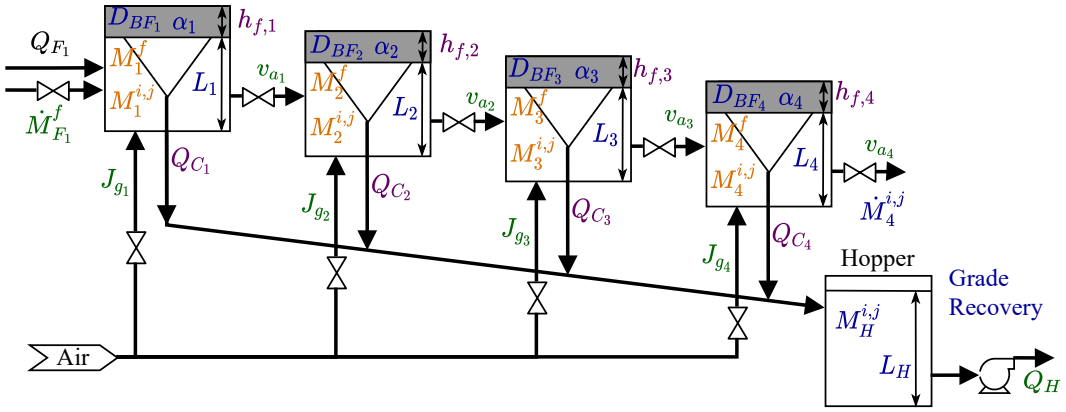

Figure 5.1. Four-cell rougher circuit, described in Hadler et al. (2010), adapted from Oosthuizen (2023). Green: manipulated variables. Blue: measured variables. Orange: variables not measured. Purple: intermediate variables.

the cells are assumed to be well-mixed and there are no transport delays in the model, the concentration in all cells react immediately, albeit with dynamics which become slower as one progress down the bank. The first cell reaches its steady-state concentration within 5 min, while the final cell takes about 10 min. The step durations mean that for the first step all cells reach steady-state, but that only Cell 1 is able to reach its steady-state value before the third step is implemented. The change in all other cells is interrupted by the third step. All further results are shown for Cell 4, as it has the slowest, most complex dynamics.

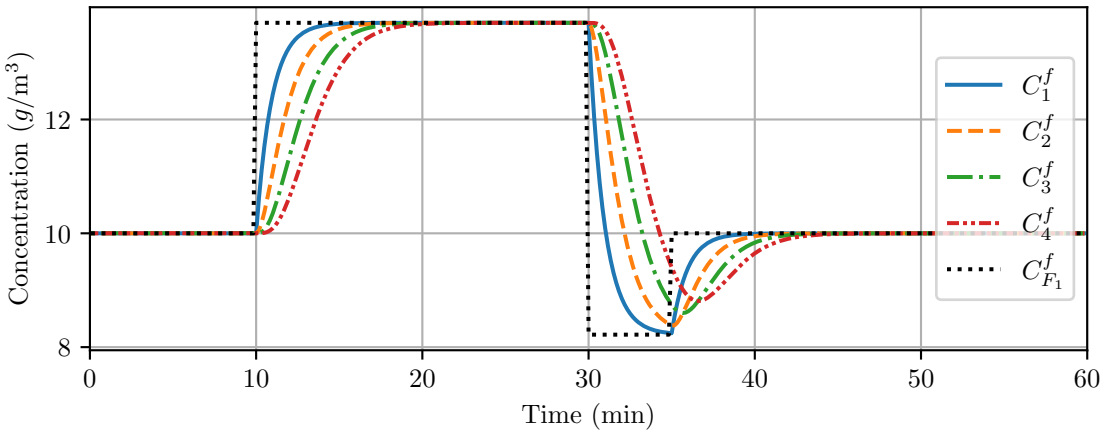

Figure 5.2. Frother concentration in all cells with changes in frother addition rate.

The models were simulated for three different constant pulp-levels, 1.28 m, 1.30 m and 1.32 m, which

Table 5.4. Initial values of model variables for the flotation cells and hopper (Oosthuizen et al., 2021).

Variable	Unit	Cell 1	Cell 2	Cell 3	Cell 4	Hopper
			Model states			
α_k	–	0.21	0.19	0.14	0.10	–
D_{BF_k}	mm	1.83	1.54	1.91	1.43	–
L	m	1.28	1.28	1.28	1.28	1.00
$M_k^{0,0}$	kg	51.8	29.4	16.8	16.8	235.8
$M_k^{1,0}$	kg	5118	5118	5118	5118	82.5
M_k^f	kg	0.11	0.11	0.11	0.11	–
			Manipulated variables			
J_{g_k}	$\mathrm{mm\,s^{-1}}$	7.45	8.45	9.51	10.02	–
Q_{T_k}	$\mathrm{m^3\,h^{-1}}$	729.3	728.1	727.4	726.4	3.6

Table 5.5. Nominal values of model variables for the flotation bank.

Variable	Unit	Value	Variable	Unit	Value
C_0^f	$\mathrm{kg\,m^{-3}}$	0.01	$K_{F\alpha}$	–	3.0
$d_{p,\min}$	μm	2.77	$K_{F\alpha,0}$	–	0.01
$d_{p,\max}$	μm	23.2	$\dot{M}_{F_1}^f$	$\mathrm{kg\,h^{-1}}$	7.3
K_{CF}	–	3000	Q_F	$\mathrm{m^3\,h^{-1}}$	730.0
$K_{CF,0}$	–	64.9			

correspond to froth heights of 120 mm, 100 mm and 80 mm respectively. Henceforth, these simulations will be referred to as Runs 1, 2 and 3 for ME1 and Runs 4, 5 and 6 for ME2.

5.5.2 Model expansion 1 (ME1) results

An increase in froth stability is expected with higher frother concentrations, leading to higher air recoveries. However, during the ME1 simulation the calculated $K_{0,\alpha_{h_f},k}$ is either larger or smaller than the froth height, depending on the simulation run, as shown in Figure 5.3. ME1 incorrectly predicts an inversion in the effect of frother if the calculated $K_{0,\alpha_{h_f},k}$ is larger than the froth height (see the simulation results in Figure 5.4). This is especially clear in Run 2, where $K_{0,\alpha_{h_f},k}$ moves past the

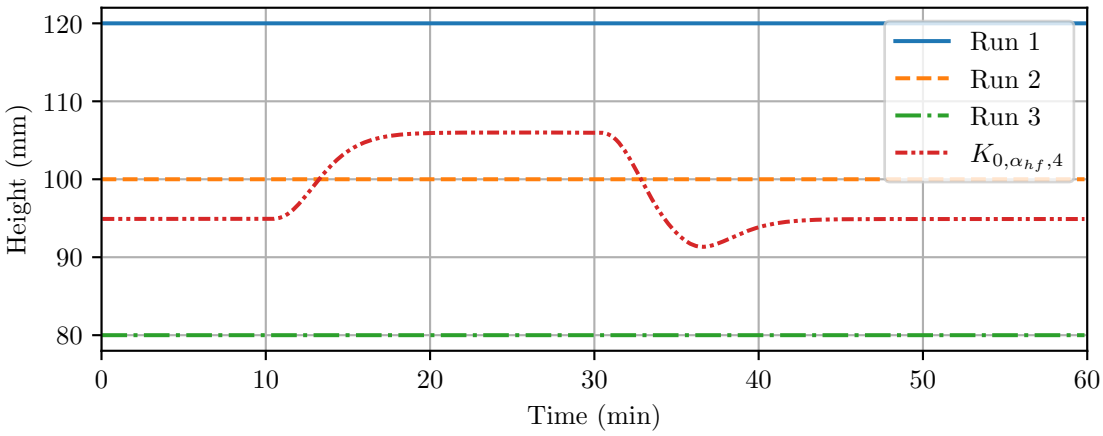

Figure 5.3. Calculated $K_{0,\alpha_{h_f},4}$ compared to the froth heights over the different simulation runs. Note that $K_{0,\alpha_{h_f},4}$ crosses the froth height in Run 2, resulting in an opposite air recovery gradient.

froth height half-way through the response and the air recovery trajectory changes direction (compare Figure 5.3 and Figure 5.4).

5.5.3 Model expansion 2 (ME2) results

Next, consider the results for ME2 shown in Figure 5.5. The direction of the frother effect is more in line with what is expected from literature (Wiese et al., 2011; Wills and Finch, 2015). Increasing the frother concentration improves the froth stability resulting in higher air recoveries. An increase in air recovery leads to an increase in mineral recovery and heights closer to the PAR height (94.9 mm) are shown to have higher air recoveries. This matches the observations of Hadler et al. (2012).

Deeper froths result in higher grades due to an increase in drainage of entrained particles. Increased air recovery also leads to increased recovery through true flotation (see (4.10)). While more desired mineral is recovered through both mechanisms (drainage and true flotation), the increase in entrained gangue recovery causes the grade to decrease during periods with increased frother concentration.

The predicted magnitude of variation in air recovery, grade and recovery is solely due to the specific model parameters chosen for the simulations. To obtain more realistic results and evaluate the predictive capabilities of the model expansion, industrial data is required. Most importantly, data is required where the frother concentration is varied.

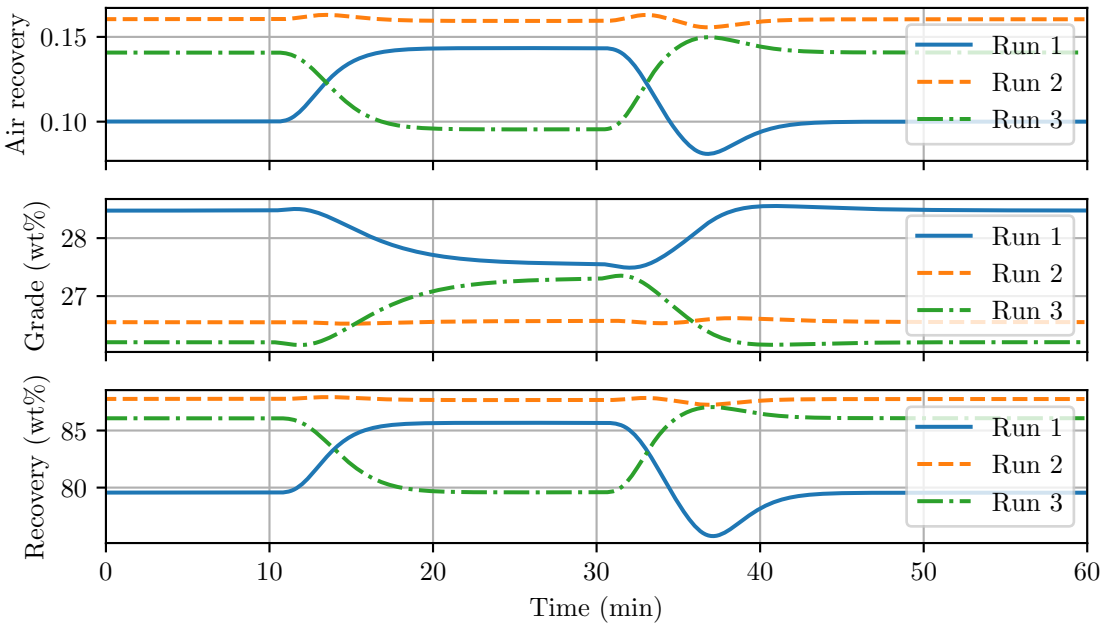

Figure 5.4. Cell 4 results for the three simulations of ME1. Results show the incorrect inversion of frother effect.

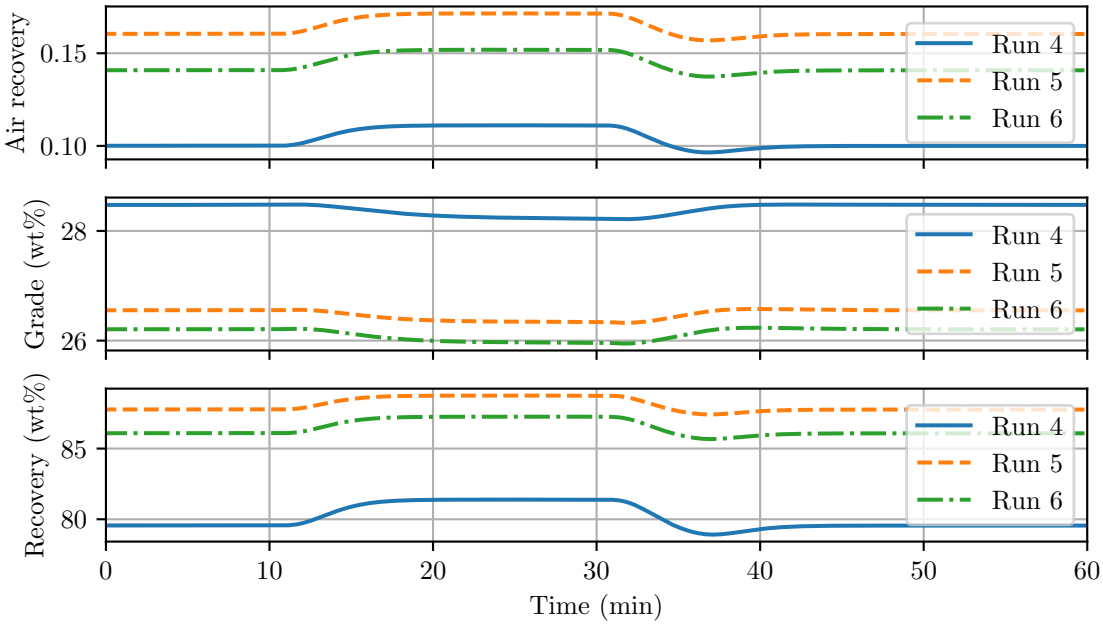

Figure 5.5. Cell 4 results for the three simulation runs of ME2. Results show the model responding as expected.

5.6 CONCLUDING REMARKS

The work presented here serves as a proof of concept for the inclusion of reagent effect models in observable and identifiable dynamic modelling frameworks of froth flotation circuits. Of the two model expansions investigated, ME2 was found to be the most appropriate frother effect model. The fact that the concentrations can be estimated from the model without any concentration measurement is very promising indeed. Based on the results shown here, real time model-based control that includes reagent addition as a MV could be possible, though there is still a substantial amount of work to be done. Future work should include a sensitivity analysis using more appropriate variable ranges, as well as the validation and/or improvement of the model expansion. Both avenues rely heavily on obtaining plant data where the reagent dosage is varied. The current approach should also be expanded to consider the effect of the collector on the flotation parameters, as there should be a very clear link between the measured variables and the collector dosage.

CHAPTER 6 MHE IMPLEMENTATION ON REAL WORLD DATA

6.1 CHAPTER OVERVIEW

As discussed in Chapter 2, the identifiable parameter modelling approach proposed by Oosthuizen (2023) can potentially address the issue of long-term model maintenance with limited on-line measurements for use in model-based control strategies. This, however, requires that the model can accurately replicate the plant dynamics and track parameter changes given the measurements available on the plant. While Oosthuizen (2023) has shown good parameter tracking results in simulation and fitted parameters to steady-state plant data from Hadler et al. (2010), the prediction accuracy of the model has not been validated on dynamic plant data yet.

This chapter investigates whether a non-linear state estimator algorithm, specifically an MHE, is able to update the model states and parameters from real industrial plant data such that the k-step ahead prediction of the model is sufficiently accurate for control. The MNC plant and available measurements are discussed in Section 6.2. Given the available measurements and the related observability limitations discussed in Chapter 4, the implementation and fair evaluation of the full model and MHE on the plant data was not possible. A reduced version of the model (Section 6.3) is shown to be observable and identifiable with the available measurements.

Data from the MNC historian is used to fit the initial model parameters for the reduced model (see Section 6.4). The model and parameter values are used to design and tune an MHE (Section 6.5) which is implemented using the historian data as on-line measurements. The state and parameter estimation results are discussed in Section 6.6. The predictive capability of the model and estimated parameters are also evaluated.

6.2 PLANT AND MEASUREMENT DESCRIPTION

6.2.1 Plant overview

MNC is the largest single stream Platinum Group Metals (PGM) concentrator in the world (Steyn and Sandrock, 2021). It has been the subject of several studies (Venkatesan et al., 2014; Finfer, 2018; Steyn and Sandrock, 2021; Horn et al., 2022) and a very detailed plant description is given in Steyn and Sandrock (2021). Here follows only a brief overview of the section relevant to this study.

After being crushed in the dry section, the ore is fed to the primary SAG mill which feeds into two parallel rougher banks of 7 cells each. The second rougher bank (RB2) will be the focus of this work, as it only uses Dorr-Olivier forced air cells (labelled FT1–FT7 for the purposes of this study) and no naturally aspirated cells. Figure 6.1 shows a simplified diagram of RB2 with the available measurements and MVs.

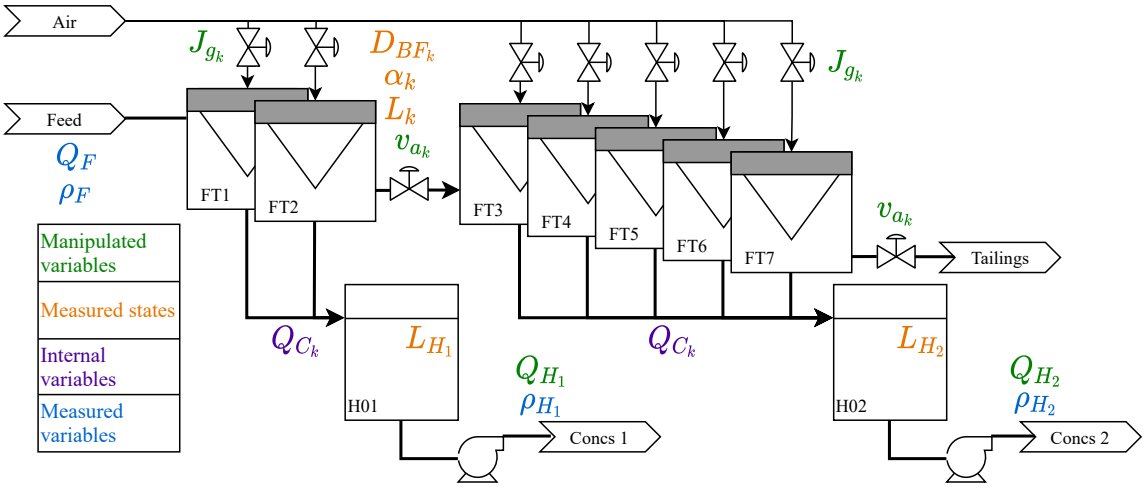

Figure 6.1. Second rougher bank at Mogalakwena North Concentrator. Variables that are applicable to multiple cells are indicated with a subscript k.

The bank is fed from a surge tank, which serves both rougher banks. The concentrates from the first two and last five cells are combined into two separate concentrate tanks (subscripts H_1 and H_2). Variable speed pumps on the tank outlets are used to pump the concentrates to downstream processes. The concentrate from the first two cells can either go directly to the final concentrate or the cleaners bank. The combined concentrate of the last five cells is sent to the cleaners for further upgrading. The tailings from both RB1 and RB2 are combined in a single sump, from which it is sent to the scavenger section via a secondary comminution section.

6.2.2 Available dynamic data

The data used in the remainder of this chapter were obtained directly from the MNC plant historian, except for the pulp sensor data in Section 6.2.2.2, which were provided by StoneThree[1]. Between 25 and 26 June 2019, step tests of the air feed to the first cell (FT1) and RB2 frother dosage were conducted (Horn et al., 2022). The pulp sensor data during this time period were carefully monitored and were made available for this study. The data used for this study were therefore limited to the period between 09:00 25 June 2019 and 09:00 26 June 2019.

In the following four sections (6.2.2.1, 6.2.2.2, 6.2.2.3 and 6.2.2.4), the available on-line measurements are discussed. Figure 6.2 and Figure 6.3 show the dynamic data collected for FT1 and FT2 over the entire period considered. The J_{g_1} step tests occur on 25 June, 13:00–17:00. The second half of the data (21:00 25 June – 09:00 26 June) represent normal plant operation.

6.2.2.1 Level, density and flow-rate

The cell pulp levels (L_k) are measured with a ball-float and sonar plate set-up. The float has a limited motion range and the level is reported as a percentage of this range. Dimensions from the plant were used to convert the supervisory control and data acquisition (SCADA) level percentages to a physical pulp level in the cell. The concentrate tank levels (L_{H_q}) are measured with a reflective float and sonar level transducer over the entire height of the tank. While physical distances were used throughout, as required by the model, any level data will be reported as deviation variables for commercial reasons.

The flow-rates between cells (Q_{T_k}) are controlled using dual dart-valves with split-range control and are not measured. The plant level controller specifies a total valve fraction ($v_{a,k}$) which is then divided between the two valves. Only the total $v_{a,k}$ value is available in the historian and will be used as the model valve fraction MV value in (4.1).

The volumetric flow-rates for the feed (Q_{F_1}), the two combined concentrate streams (Q_{H_q}) and combined tailings (Q_T^{Tot}) are measured on-line. The bulk densities of these streams (ρ_F, ρ_{H_q} and ρ_T respectively) are also measured on-line. As the measured tailings is a combination of the tails from RB1 and RB2,

[1] StoneThree are an industrial IoT consulting company based in the Western Cape, that develops smart sensors aimed at process monitoring and health and safety. They also provide engineering consulting services to their clients, leveraging data science and machine learning to improve operational performance. One particular field of their expertise is the monitoring of flotation froths.

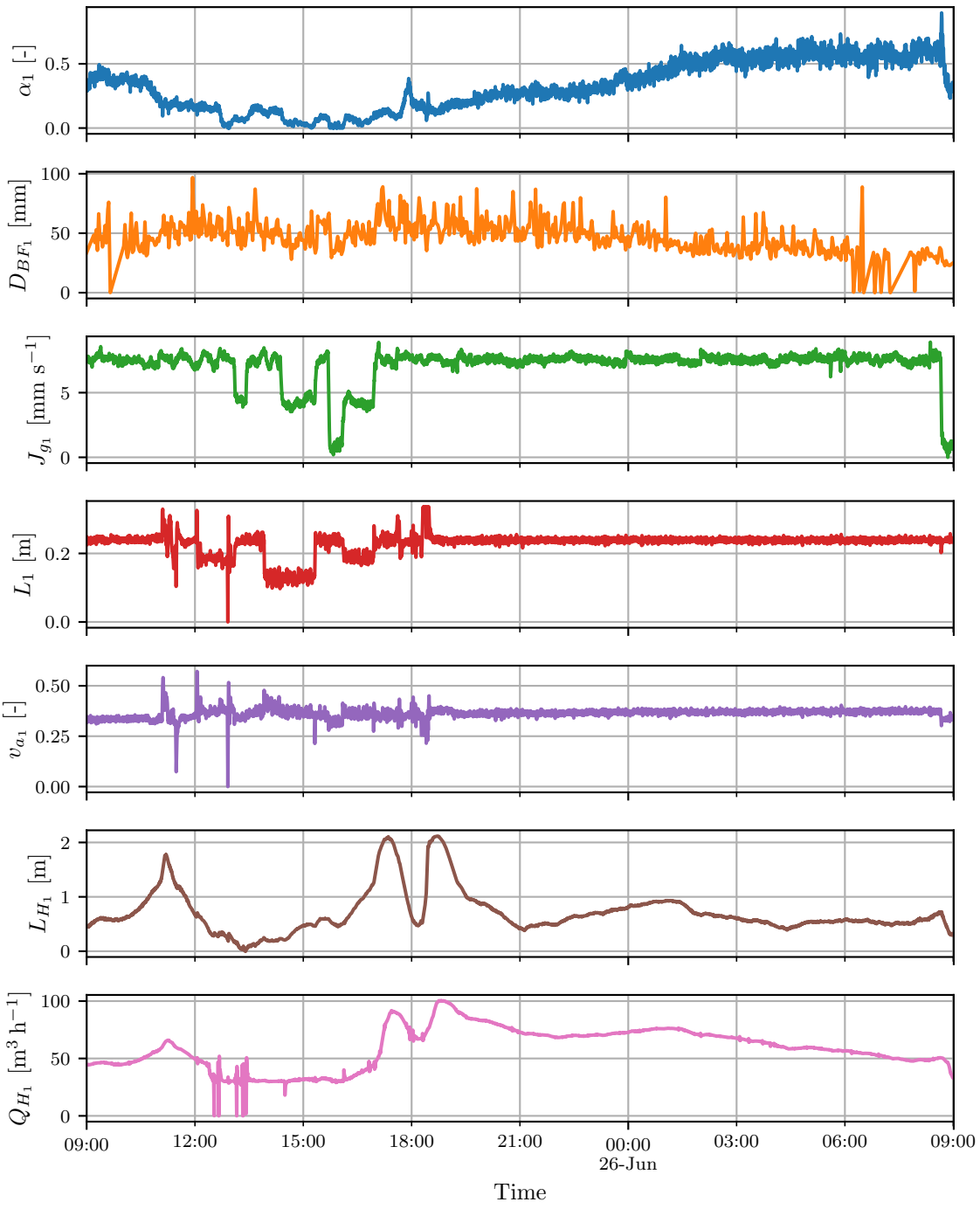

Figure 6.2. Available dynamic data for FT1. As the densities are not used for the MHE implementation, they were omitted. The data has been normalised for commercial reasons. Step tests of J_{g_1} were conducted during 13:00–17:00 , 25 June 2019.

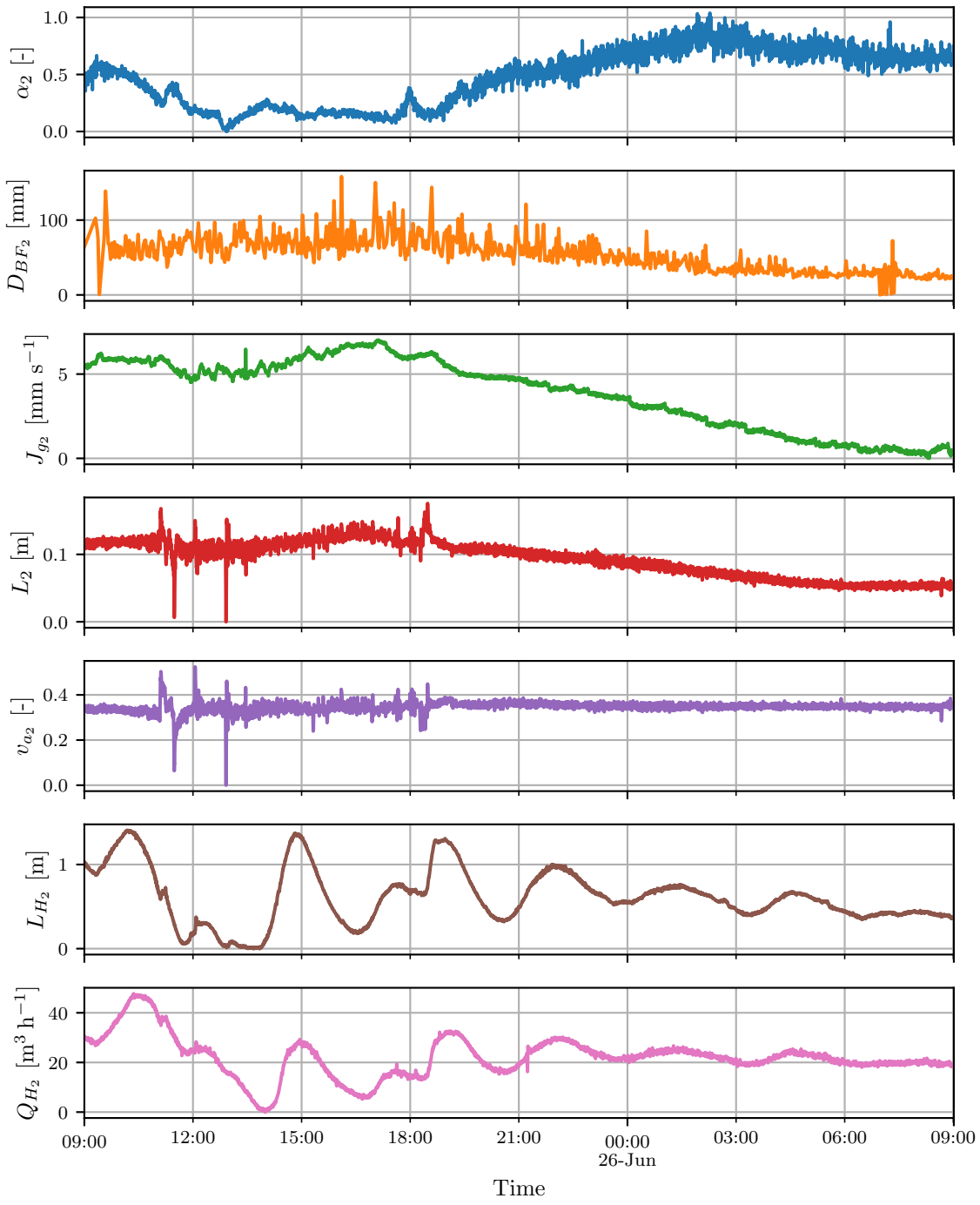

Figure 6.3. Available dynamic data for FT2. As the densities are not used for the MHE implementation, they were omitted. The data has been normalised for commercial reasons.

there is a high likelihood that the ρ_T density measurement is not representative of the RB2 contribution and will definitely include dynamics from the tailings sump, which are not included in the model in Chapter 4. However, for sections where both banks are operated at similar conditions and the variability of the measurements is not very high, the density measurement should give a sufficient approximation for the purposes of this study.

During the time period under consideration, RB2 was fed directly from the distribution box from the primary mill and not the surge tank. This meant that the Q_{F_1} flow-rate measurement was not available, which introduced additional uncertainty during state estimation (see Section 6.6.1).

6.2.2.2 Superficial gas velocity

Although the aeration rate or J_{g_k} in each cell is a crucial operating variable and is present in many of the phenomenological flotation models, it is very difficult to measure on-line (Shean and Cilliers, 2011). Rather, the linear air velocity in the air feed to each cell ($q_{air,k}$) is measured and reported to the SCADA in $\mathrm{m\,s^{-1}}$.

Previous studies (Venkatesan et al., 2014; Steyn and Sandrock, 2021) at MNC obtained linear relationships between J_{g_k} and $q_{air,k}$, using the Anglo Platinum Bubble Sizer (APBS) to manually measure J_{g_k}. Figure 6.4 shows the mean values of the J_{g_k} profile between 25 and 26 June 2019 obtained by applying the Steyn and Sandrock (2021) conversion

$$J_{g_k} = m_{J_g,k} q_{air,k} + c_{J_g,k}, \tag{6.1}$$

on the SCADA data as is. Here, $m_{J_g,k}$ and $c_{J_g,k}$ are linear parameter fits to the measurements, as reported by Steyn and Sandrock (2021).

If the $q_{air,k}$ measurements are reliable, obtaining J_{g_k} by merely converting $q_{air,k}$ using the ratio of the cell (A_k) and inlet pipe (A_{air}) cross-sectional areas is expected to be reliable. However, using this conversion on the available data yielded J_{g_k} values that were not only larger than the Steyn model values, but also had a different profile shape in the fourth and sixth cells (see Figure 6.4).

Recently, StoneThree installed image-based on-line pulp sensors on the first two cells (FT1 and FT2 in Figure 6.1) of RB2 at MNC (Horn et al., 2022) that measure J_{g_k} directly. Although they contained high levels of noise, these pulp measurements proved to be invaluable in calibrating the $q_{air,k}$ to J_{g_k}

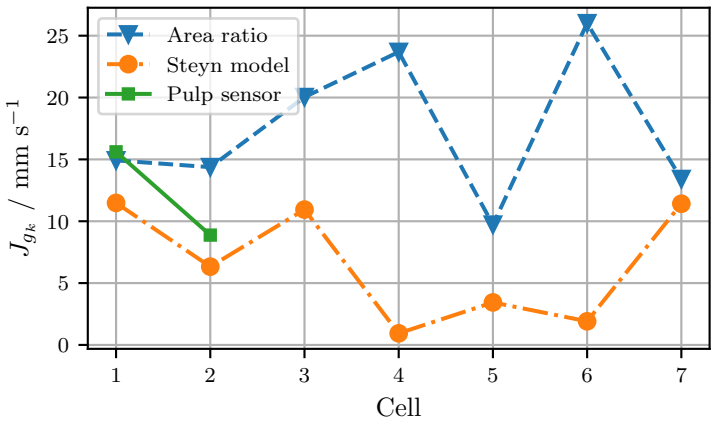

Figure 6.4. SCADA to J_{g_k} conversion options. All J_{g_k} values, are the mean J_{g_k} over the period 25–26 June 2019. The values shown here have been offset from the actual values for commercial reasons.

conversion. The means of the pulp sensor measurements for the same two day period are also shown on Figure 6.4.

The pulp sensor J_{g_1} measurements correlate well with the J_{g_1} obtained using the area ratio conversion, but J_{g_2} did not. Instead, the difference between the pulp sensor J_{g_1} and J_{g_2} was very close to the difference between the values obtained using the Steyn and Sandrock (2021) conversion. New conversion ratios (no $c_{J_g,k}$ terms) for FT1 and FT2 based on the available on-line measurements were determined.

While there is a significant amount of noise in the pulp sensor measurements, the correlation is good for a period of roughly twelve hours of data. If the same conversions are to be determined for the remaining cells, similar data for FT3–FT7 are required.

The marked difference between the various possible J_{g_k} values is somewhat concerning if these measurements are to be used in phenomenological models. Without more information, it is difficult to determine which of the three conversions in Figure 6.4 is more accurate. Possible explanations include faulty velocity measurement calibration or incorrect pipe diameter. Alternatively, the APBS and pulp sensor measurements could also have measurement biases.

The fact that J_{g_k} measurements from three independent studies (Venkatesan et al. (2014), Steyn and

Sandrock (2021) and Horn et al. (2022)) do not completely match the "Area ratio" J_{g_k} values seem to indicate some inaccuracy in this approach. However, the $q_{air,k}$ measurements are far less noisy than the pulp sensor measurements. Given the lack of a clearly more accurate approach, it was decided to use the $q_{air,k}$ measurements converted using the ratios fitted to the available pulp sensor measurements.

6.2.2.3 Air recovery

Air recovery (α_k) is the fraction of air fed to the cell (Q_{a_k}) recovered as overflowing froth, calculated as (Oosthuizen, 2023; Phillpotts et al., 2021)

$$\alpha_k = \frac{w h_{f,o,k} v_{f,o,k}}{Q_{a_k}}. \tag{6.2}$$

where w is the launder lip length, the height of the froth over the lip ($h_{f,o,k}$) is measured by a StoneThree laser froth height sensor and the froth overflow velocity ($v_{f,o,k}$) is measured by the StoneThree froth camera. Note that $h_{f,o,k}$ used in (6.2) is different from the h_{f_k} used in the model equations ((4.16) and (4.23)) (see Figure 4.1). Note: Q_{a_k} is calculated as

$$Q_{a_k} = A_k J_{g_k}, \tag{6.3}$$

using the J_{g_k} values as determined in Section 6.2.2.2 and not using A_{air}.

The cell diameter D_{cell} and radial launder length ($d_{l,r}$) were obtained from the plant personnel and were used to calculate w as

$$w = \pi D_{\text{cell}} + 2 n_{l,r} d_{l,r}, \tag{6.4}$$

where $n_{l,r}$ is the number of radial launders. This assumes that the froth overflows the entire launder length at the same velocity and the same height. While unlikely, there is no way to confirm the homogeneity of the froth from one camera per cell. The general trend of the air recovery should, however, remain relatively unchanged with this assumption, especially when approaching steady-state operation.

The composite nature of the α_k measurements (in (6.2)) means that they are subject to very high uncertainty, as the uncertainties of all the constituent measurements are combined. High measurement noise could obscure the required dynamic information required to estimate the α_k model parameters, especially as the parameters in (4.23) only become identifiable at higher order Lie derivatives (see the discussion in Section 4.5.2.1, Table 4.5 and Table A.2). Errors in the measured α_k values (as would be introduced by incorrect J_{g_k} measurements (Section 6.2.2.2) or non-homogeneity of the overflowing conditions) could also pose a problem to the use of phenomenological models such as (4.4).

6.2.2.4 Froth bubble size

The top of froth bubble size for each cell (D_{BF_k}) is measured using the StoneThree froth cameras. The mean bubble size is reported at a sampling period of 2 min. This sampling time is much slower than the other on-line measurements used, which are sampled every 10 s. To achieve the same sampling-rate, the D_{BF_k} measurements were up-sampled by the historian using polynomial fit interpolation. The expected speed of the D_{BF_k} dynamics (Oosthuizen, 2023) is much faster than the 2 min sampling time. The amount of dynamic information contained in the D_{BF_k} measurements is, therefore, expected to be severely limited, which could complicate parameter estimation for the D_{BF_k} model parameters in (4.16). The possible implications are, however, shown to be limited in Section 6.3.2.

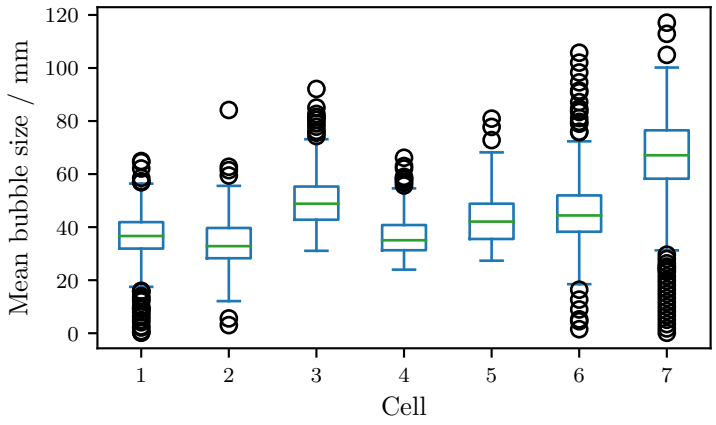

Figure 6.5. Distribution of the raw mean D_{BF_k} measurements over a normal 8 h shift (00:00–08:00, 26 June 2019), sampled every 2 min. Each box represents 240 data-points. Outliers are shown as black circles.

If a mean cannot be assigned to the distribution (*e.g.,* when it is a bi-modal distribution or when the bubble identification algorithm fails) a `nan` value is entered in the historian. Figure 6.5 shows the distribution of mean D_{BF_k} measurements as recorded in the historian for the 8 h shift from 00:00–08:00, 26 June 2019. The D_{BF_k} measurement algorithm had trouble assigning a distribution mean value near the end of the shift (see Figure 6.2 and the large number of low outliers in Figure 6.5). From the available plant data, no obvious anomalous operation could be found during this period. The D_{BF_k} measurement failure could either be the result of some unmeasured disturbance or (though less likely) bubble conditions that are not adequately dealt with in the measurement algorithm. The state estimation algorithm might struggle to deal with such unknown disturbances.

6.2.3 Static data

Currently, 8 h shiftly composite assays, sampled every 10 min and then combined, are available for the feed of rougher section and the individual bank tailings. The assays report on Pt, Pd and Rh (3E) content for the tails and the feed. The feed assays additionally report on the Fe content and particle size distribution. No assays of the concentrate streams are available.

Oosthuizen et al. (2021) assumes on-line grade measurements for the feed, concentrate and tailings. Without these on-line compositional measurements, the possible utility of the approach suggested by Oosthuizen et al. (2021) is limited (see Section 4.5.2). Unfortunately, on-line XRF analysers are maintenance-intensive, especially given the harsh, heterogeneous conditions of the rougher feed stream and are consequently only occasionally used for the roughers at MNC.

6.3 REDUCED MODEL

In Section 4.5.2, the limitation placed on observability of the full updated model when no concentrate grade measurements are available were discussed. Given the available measurements in Section 6.2.2 and the absence of compositional assays of the concentrate streams (Section 6.2.3), synthesising the necessary $G_{C,t}^{0,0}$ and $G_T^{0,0}$ at a reasonable frequency is not possible. The implementation of an MHE based on the full model described in Section 4.3 was therefore not possible. Instead, a reduced model was developed that is observable with the available measurements (Section 6.3.1 and Section 6.3.2).

6.3.1 Model summary

The first two cells in the bank are the most instrumented with on-line J_{g_k} data and their concentrates are combined into a single cell. Furthermore, Steyn and Sandrock (2021) reports that the first two cells alone achieve 61 % recovery, while the total recovery of the bank is around 64 %. This makes the first two cells a natural subsystem for a reduced model.

All solids mass balance related equations (flotation, entrainment, component balances *etc.*) are removed, as their inclusion requires compositional measurements. It is assumed that the k_1 in (4.4) not only classifies the water recovery, but also accounts for the volume of the recovered solids. The new parameter is called k_m. Thus, Q_{C_k} is given by

$$\frac{Q_{C_k}}{A_k} \approx \begin{cases} \dfrac{6.81 J_{g_k}^2}{k_m D_{BF_k}^2} (1-\alpha_k)\,\alpha_k & 0 < \alpha_k < 0.5 \\ \dfrac{6.81 J_{g_k}^2}{4 k_m D_{BF_k}^2} & \alpha_k \geq 0.5 \end{cases} . \tag{6.5}$$

The remaining equations are summarised in Table 6.1 and are the same as the corresponding equations in Section 4.3.

Table 6.1. Reduced model equations.

	Unit	Description	Equation	
			States	
α_k	-	Air recovery	$\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} = \frac{K_{\alpha_{J_g}}\left(J_{g_k} - K_{0,\alpha_{J_{g_k}}} - K_{\alpha_{h_f}} h_{f_k}\right)^2 + \alpha_{OS_k} - \alpha_k}{\lambda_{\mathrm{air}_k}}$	(4.23)
D_{BF_k}	mm	Froth bubble size	$\frac{\mathrm{d}D_{BF_k}}{\mathrm{d}t} = \frac{K_{BS_{J_g}} J_{g_k} + K_{BS_\lambda}\lambda_{\mathrm{air}_k} + D_{OS_k} - D_{BF_k}}{\lambda_{\mathrm{air}_k}}$	(4.16)
L_{H_q}	m	Hopper level	$\frac{\mathrm{d}L_{H_q}}{\mathrm{d}t} = \left(\sum_{k=m}^{n} Q_{C_k} - Q_{H_q}\right)/A_{H_q}$	(4.3)
L_k	m	Cell level	$\frac{\mathrm{d}L_k}{\mathrm{d}t} = \left(Q_{F_k} - Q_{T_k} - Q_{C_k}\right)/A_k$	(4.2)
			Auxiliary variables	
h_{f_k}	mm	Froth height	$h_{f_k} = 1000(H_{\mathrm{cell}} - L_k)$	(4.12)
λ_{air_k}	s	Froth residence time	$\lambda_{\mathrm{air}_k} = \frac{h_{f_k}}{J_{g_k}}$	(4.18)
Q_{C_k}	$\mathrm{m^3\,h^{-1}}$	Concentrate flow-rate	$\frac{Q_{C_k}}{A_k} \approx \begin{cases} \frac{6.81 J_{g_k}^2}{k_m D_{BF_k}^2}(1-\alpha_k)\alpha_k & 0 < \alpha_k < 0.5 \\ \frac{6.81 J_{g_k}^2}{4 k_m D_{BF_k}^2} & \alpha_k \geq 0.5 \end{cases}$	(6.5)
Q_{T_k}	$\mathrm{m^3\,h^{-1}}$	Tank outlet flow-rate	$Q_{T_k} = v_{a,k} C_{v_k}\sqrt{L_k - L_{k+1} + \Delta_{h_k}}$	(4.1)

6.3.2 Observability results

The reduced model variable vectors used in the observability analysis are

$$\mathbf{x}_3 = \left[\alpha_k,\ D_{BF_k},\ L_{H_q},\ L_k\right]^T, \tag{6.6}$$

$$\mathbf{p}_3 = \left[\alpha_{OS_k},\ C_{v_k},\ D_{OS_k},\ K_{0,\alpha_{J_{g_k}}},\ K_{\alpha_{h_f}},\ K_{BS_\lambda},\ k_m\right]^T, \tag{6.7}$$

$$\mathbf{u}_3 = \left[J_{g_k},\ Q_{H_q},\ v_{a,k}\right]^T, \tag{6.8}$$

$$\mathbf{y}_3 = \left[\alpha_k,\ D_{BF_k},\ L_{H_q},\ L_k\right]^T, \tag{6.9}$$

which are expanded for two cells ($k \in 1,2$) and one concentrate tank ($q \in 1$). The resulting variable counts are given in Table 6.2. Note that $K_{BS_{J_g}}$ is assumed to be a known constant (as was done in Section 4.5.2). The observability analysis results for the reduced model are summarised in Table 6.3. Full rank is achieved at $L_f^3 g$ level, which is already better than for the full model. If necessary, the chances of getting reasonable on-line parameter estimates can be improved further by assuming $K_{\alpha_{h_f}}$ is a known constant.

Table 6.2. Reduced number of variables.

n_x	n_p	n_y	n_u
7	11	7	5

Table 6.3. Reduced model observability results.

$L_f^i g$ order	d$\mathcal{O}$ rank	Deficiency	PUPs	#PUPs
0	7	11	$\alpha_{OS_k}, C_{v_k}, D_{OS_k}, K_{0,\alpha_{J_{g_k}}}, K_{\alpha_{h_f}}, K_{BS_\lambda}, k_m$	11
1	14	4	$\alpha_{OS_k}, D_{OS_k}, K_{0,\alpha_{J_{g_k}}}, K_{\alpha_{h_f}}, K_{BS_\lambda}$	8
2	17	1	$\alpha_{OS_k}, K_{0,\alpha_{J_{g_k}}}, K_{\alpha_{h_f}}$	5
3	18	0	–	0

The possible impact of only steady-state D_{BF_k} measurements (as discussed in Section 6.2.2.4) was investigated by assuming that D_{BF_k} can be modelled as an algebraic equation, *i.e.,* (4.16) in Table 6.1 is replaced by

$$D_{BF_k} = K_{BS_{J_g}} J_{g_k} + K_{BS_\lambda} \lambda_{\text{air}_k} + D_{OS_k} \tag{6.10}$$

and D_{BF_k} is removed from $\mathbf{x}_3$ in (6.6). This reduces the n_x to only 5 and the required d$\mathcal{O}$ rank to 16. The observability analysis results are shown in Table 6.4. Full state observability and parameter identifiability is again achieved at the $L_f^3 g$ level. Compared to the results in Table 6.3, D_{OS_k} and K_{BS_λ} become identifiable sooner when the (6.10) is used. Therefore, the slower D_{BF_k} sampling time does not appear to affect the parameter identifiability negatively. However, the fictitious dynamic data added by the upsampling of D_{BF_k} is still expected to have a detrimental impact on the accuracy of the parameter estimation.

Table 6.4. Reduced model with steady-state D_{BF_k} model observability results.

$L_f^i g$ order	d$\mathcal{O}$ Rank	Deficiency	PUPs	#PUPs
0	7	9	$\alpha_{OS_k}, C_{v_k}, D_{OS_k}, K_{0,\alpha_{J_{g_k}}}, K_{\alpha_{h_f}}, K_{BS_\lambda}, k_m$	11
1	13	3	$\alpha_{OS_k}, K_{0,\alpha_{J_{g_k}}}, K_{\alpha_{h_f}}$	5
2	15	1	$\alpha_{OS_k}, K_{0,\alpha_{J_{g_k}}}, K_{\alpha_{h_f}}$	5
3	16	0	$--$	0

6.4 PROCESS PARAMETER FITS

The following three subsections discuss how the initial parameter estimates for the model were obtained. The exact value of these parameters are not as important as getting them close enough to the “correct" value for the estimator to be effective.

The cell dimensions (A_k, Δ_{h_k}, H_{cell}) and the measurements needed to convert the SCADA level percentages to pulp levels were obtained from MNC personnel.

6.4.1 Air recovery model parameters

Selecting enough steady-state data to get a good picture of the system is a very tedious process. This is mostly because the identification of steady-state periods in the data is done manually. As the aim is to obtain reasonable initial values, the dynamic J_{g_k}, α_k and h_{f_k} data were resampled every 30 s and then assumed to be steady-state data-points. The model inputs (J_{g_k} and h_{f_k}) typically remain constant for much longer than the relatively fast α_k dynamics. In addition, froth residence times are typically 20 s to 65 s (Steyn and Sandrock, 2021), while J_{g_k} typically remains relatively constant for well over 10 min. Therefore, the dynamic periods included in the fitting data with this assumption, should have a small enough impact on the parameter fits.

The 24 h dataset in Figure 6.2 and Figure 6.3 was divided into three 8 h sections: 09h–17h, 17h–01h and 01h–09h. A 1.5 h section of data where there is some movement in the J_{g_k} was selected in each section and the parameters in (4.23) were fitted using only the data in these selected subsections by minimising the RMSE between the predictions and observations. The parameter values are given in Table 6.5.

Table 6.5. Fitted parameters for α_k model at different times.

Section	$K_{\alpha_{J_g}}$	$K_{\alpha_{h_f}}$	$K_{0,\alpha_{J_{g_1}}}$	$K_{0,\alpha_{J_{g_2}}}$	α_{OS_1}	α_{OS_2}
09h–17h	-2.48×10^{-2}	1.02×10^{-2}	19.40	14.64	0.135	0.252
17h–01h	-2.48×10^{-2}	1.02×10^{-2}	18.82	7.34	0.392	1.082
01h–09h	-2.48×10^{-2}	1.02×10^{-2}	20.93	8.95	0.538	0.925
Modified	-2.48×10^{-2}	1.02×10^{-2}	18.08	12.39	0.700	1.100

As in Oosthuizen et al. (2021), a constant $K_{\alpha_{J_g}}$ was assumed. The same value (-2.48×10^{-2}) was used, as this was found to fit the shape of other peak air recovery curves in literature (Phillpotts et al., 2021) and worked for the current dataset. Hadler et al. (2012) found that the J_{g_k} for PAR is higher for deeper froths. The Oosthuizen (2023) optimal J_{g_k} is at

$$J_{g_k} = K_{0,\alpha_{J_{gk}}} + K_{\alpha_{h_f}} h_f. \tag{6.11}$$

Therefore, the $K_{\alpha_{h_f}}$ is expected to be positive (deeper froth, higher air at PAR).

The fitted $K_{\alpha_{h_f}}$ varied significantly depending on which 1.5 h subsection of the corresponding 8 h section was used, often reaching the negative values. However, when the $K_{\alpha_{h_f}}$ of the best fit (FT2 for the 17h–01h section) was fixed for all sections, the resulting fits did not differ much from those obtained when allowing $K_{\alpha_{h_f}}$ to vary. Considering the potential challenges to observability (see Section 6.3.2), it was decided to keep $K_{\alpha_{h_f}}$ constant at 1.02×10^{-2}.

The fitted parameters of each section in Table 6.5 were used to predict the steady-state α_k over the entire corresponding 8 h section. The results of these predictions are shown in Figure 6.6. Apart from the 17h–01h section of FT2 all the predictions tended to underestimate the variance in α_k, with reasonable prediction correlation only near the 1.5 h fitted section.

The value of α_{OS_k} places a hard upper bound on the possible α_k prediction. Unless the fitting section contained considerable variance, the resulting parameters tended towards assigning α_{OS_k} a value in, or very close to, the range of observed α_k values in the section. The $K_{0,\alpha_{J_{g_k}}}$ values then tended to predict optimal J_{g_k} values very close to the J_{g_k} values in the section. As shown in Figure 6.6, this leads to horizontal grouping of the the predicted values, rather than the required 45° angle.

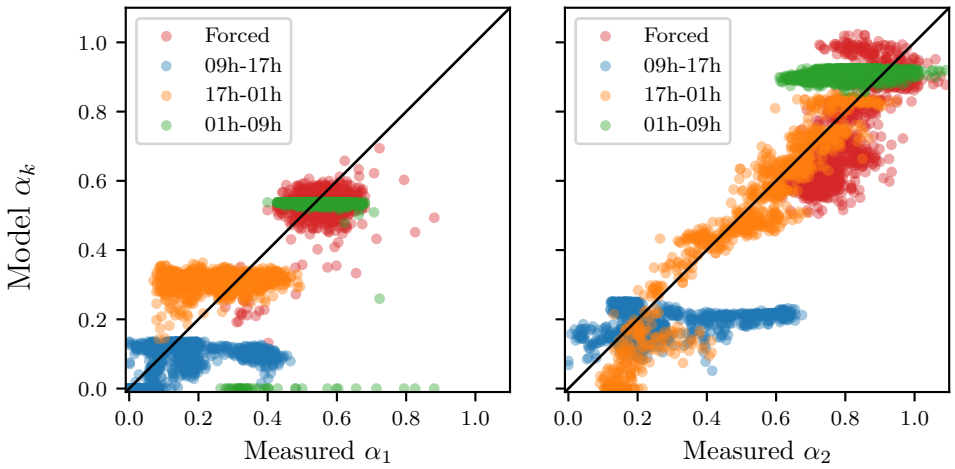

Figure 6.6. Parity plots of the α_k model predictions over 8 h periods. Note how the "Modified" parameter set captures more of the variance than the 01h–09h predictions.

It was found that by artificially increasing α_{OS_k}, the variance captured over the entire 8 h period could be increased, improving the correlation slightly, while slightly increasing the RMSE over the fitting section. This approach is shown for the 01h–09h section with the coerced parameter set in Table 6.5 and resulting predictions (Figure 6.6) labelled "Modified".

6.4.2 Bubble size model parameters

The same basic steady-state parameter fitting procedure as for α_k was followed to fit parameters in (4.16). The fitted parameters are given in Table 6.6 and the resulting steady-state prediction parity plots are shown in Figure 6.7. The values of K_{BS_λ} were constrained to positive values. A negative K_{BS_λ} leads to an increased D_{BF_k} when L_k decreases. This in turn causes an increase in the predicted Q_{C_k} (see (6.5)) which decreases the level further. This could lead to instability in the model.

Table 6.6. Fitted parameters for D_{BF_k} model at different times.

Section	$K_{BS_{J_g}}$	K_{BS_λ}	D_{OS_1}	D_{OS_2}
09h–17h	2.0	0	8.1	38.6
17h–01h	7.6	0	-118.6	-51.9
01h–09h	11.9	0	-231.3	-114.9

Again, as shown in Figure 6.7, the correlation was much better for the second cell than for the first cell. The model typically fails to capture all the variability in the measured values for FT1, characterised by

the horizontal grouping of the data points.

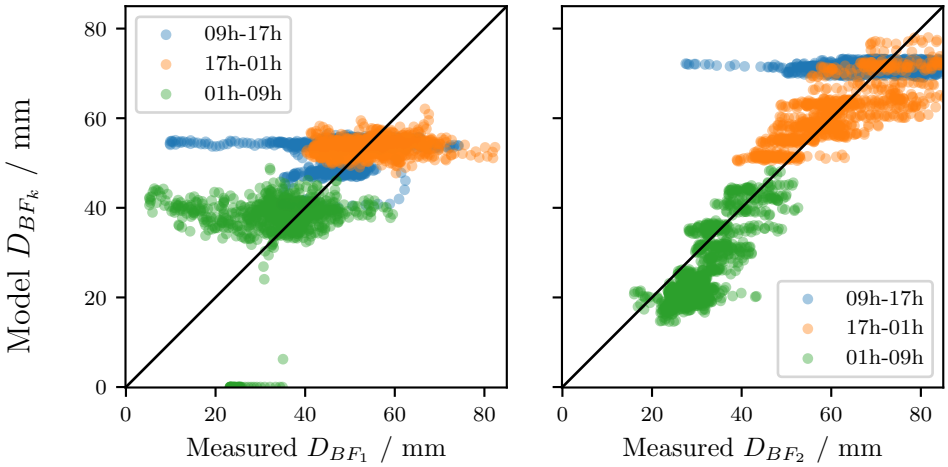

Figure 6.7. Top of froth bubble size parameter fit predictions. Parameters to be used as initial guesses for the MHE.

The lack of clear or consistent correlation of the α_k and D_{BF_k} predictions with the measurements in Figure 6.6 and Figure 6.7 is definitely concerning. If the α_k and D_{BF_k} models ((4.23) and (4.16)) fail to approximate the dynamics of the plant, the overall MHE implementation is likely to fail.

However, the periods over which the model is tested are extremely large and include dynamic periods and high levels of noise. The models were fitted and evaluated assuming that each data point is at steady-state, purely in order to obtain reasonable initial values for the MHE. The J_{g_k} and frother dosage step tests (Section 6.2.2) in FT1 contained in the data also shift the operating point around considerably. The fact that there are sections of data for which the crudely fitted parameters produce good trends, is promising.

6.4.3 Volume balance parameters

Initial k_m and C_{v_k} values were estimated by taking a section of data with relatively constant tank levels (L_k and L_{H_q}) and assuming steady-state. With known Q_{H_q}, k_m was solved from (4.3) and (6.5) for a zero derivative, *i.e.*, $\frac{\mathrm{d}k_m}{\mathrm{d}t} = 0$. With k_m fixed, Q_{C_k} values were known and Q_{T_k} was solved from (4.2) and an assumed Q_{F_1}. With the known differences in L_k, (4.1) was solved for C_{v_k}. The parameter values are given in Table 6.7.

Table 6.7. Volume balance parameter values.

Parameter	Value
C_{v_1}	2446
C_{v_2}	2569
k_m	710

6.5 MHE DESIGN

Several different combinations of tuning parameters, initial conditions and data sections were investigated. The design (MHE3) outlined in the following five subsections is representative of the best results obtained. As this already serves to highlight the strengths and weaknesses of the model, state estimation and data, the other designs are not reported on fully. However, some general observations regarding the effect of alternative design choices are mentioned where a noticeable difference was observed.

Where a comment is made regarding increased or decreased performance, both the realism of the estimated states as well as the prediction performance (see Section 6.6) are considered. In many cases absolute values are not reported due to the sensitive nature of the industrial data.

6.5.1 Basic implementation

The plant model, MHE and simulators used for the studies below were all implemented using `do-mpc` as discussed in Section 3.4. The continuous model equations are defined symbolically using `Casadi` and discretised automatically by `do-mpc` to the form required for (3.19) using collocation.

The necessary plant data to be used as measurements were obtained as described in Section 6.2. As no Q_{F_1} measurements were available (see Section 6.2.2.1), a constant value was assumed based on typical operating conditions and information from the comminution section. This is a significant source of disturbance in the L_k models and would mostly impact the C_{v_k} parameter estimation. The use of $\mathbf{w}_k$ would also help compensate for this assumption.

At each time step, the past measurements and inputs are passed to the MHE, which solves (3.19) and the state and parameter estimates are stored in a `do-mpc` data-structure. The MHE and simulators

never have access to future measurements or inputs when calculating a given time-step.

6.5.2 Sampling time and horizon

Oosthuizen (2023) uses a 10 s interval for his MHE and MPC implementations. He argues that this is fast enough to capture the fast (D_{BF_k}, α_k, L_k and L_H) dynamics while not requiring long prediction horizons to capture steady-state behaviour during prediction. He also mentions λ_{air_k} in the D_{BF_k} model corresponds to a 16 s froth residence time, while the residence time in pulp cells is about 60 s. Steyn and Sandrock (2021) report froth residence times between 20 s and 65 s that increase down the bank, which corresponds well with the data obtained from MNC. The MNC pulp residence time (around 300 s) is also larger than for the system in Oosthuizen (2023), due to the significantly larger cells. Figure 6.8 shows the poles of linearised versions (see (3.5) and (3.6)) of the reduced model in Section 6.3 fitted to MNC data and the full updated model (see Section 4.3) implemented on the data from Hadler et al. (2010) as was done in Oosthuizen (2023). While there are more MNC poles closer to the origin, indicating overall slower dynamics, there are also several comparable poles. The overall timescale covered is roughly the same and the same sampling time of 10 s will be used.

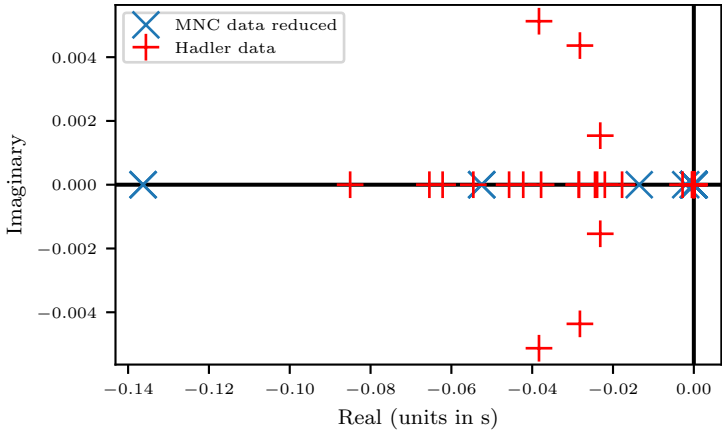

Figure 6.8. Poles of the updated and reduced models implemented on the two available datasets.

A shorter horizon allows for faster tracking of changing parameters, while a longer horizon lends more noise rejection (Oosthuizen, 2023). A horizon of $N = 10$ was used in (Oosthuizen, 2023). For the current dataset, this horizon was found to be too short and the MHE adjusted the parameters too much. A horizon of $N = 20$ was used, but similar results were obtained for horizons up to $N = 40$, after which the estimator became sluggish and predictions were no longer reliable.

For the particular set of parameters used to plot the poles in Figure 6.8, a large J_{g_1} value led to a small right-hand plane (RHP) pole. This is due to a disparity in the average magnitudes of J_{g_1} and J_{g_2} leading

to a much larger Q_{C_1} via (6.5). It is possible that the J_{g_k} measurements discussed in Section 6.2.2.2, are incorrect. However, testing the same MHE3 (Section 6.5.3) with a reduced J_{g_1} (and the necessary adjustments to initial parameters) yielded decreased estimation and predictive performance. As the linear poles do not capture all the interconnected dynamics of the non-linear model, it is possible that the non-linear model has stable dynamics while a linearised version of the model is unstable at the particular point used for linearisation.

6.5.3 Parameters and noise

In the preceding sections of this chapter several sources of uncertainty were identified that could inform the choice of $\mathbf{P}_w$ for $\mathbf{w}_k$ in (3.19):

- Slow D_{BF_k} sampling (Section 6.2.2.4)
- Dual dart values modelled as a single valve (Section 6.2.2.1)
- Unknown Q_{F_1} for the available time period (Section 6.2.2.1)
- Uncertainty regarding the correct $q_{air,k}$ to J_{g_k} conversion (Section 6.2.2.2)
- Combined nature of the α_k measurements increasing uncertainty (Section 6.2.2.3)

As the MHE is based on a semi-phenomenological model and the volume balances ((4.2) and (4.3)) play a central role in many of the other model equations (see Section 4.4 and Table 6.1), any unmodelled or unmeasured disturbances could have a severe impact on the estimation accuracy. The process noise, $\mathbf{w}_k$, should compensate for unmeasured and unmodelled disturbances to reduce error in the state and parameter estimates. However, if the $\mathbf{P}_w$ weights are too low the optimisation algorithm could converge to unrealistically large $\mathbf{w}_k$ values, effectively attributing actual parameter or state changes to random noise. This would lead to unresponsive parameter estimation and poor prediction performance, as the predictor does not have access to the future $\mathbf{w}_k$ values.

The observability analysis in Section 4.5 showed that $K_{BS_{J_g}}$ is unidentifiable for constant J_{g_k}. An intermediate value of 0.75 was chosen from the parameter fits in Table 6.6 and $K_{BS_{J_g}}$ was excluded from the $\mathbf{p}_3$ vector in (6.7). The α_{OS_k}, $K_{0,\alpha_{J_{g_k}}}$ and $K_{\alpha_{h_f}}$ parameters are still possibly unidentifiable at $L_f^2 g$, but become identifiable at $L_f^3 g$ (see the PUPs in Table 6.3 which were determined as discussed in Section 3.3.4). This could indicate difficulty in estimating these four parameters if the higher-order (third-order and above) output derivatives are too small. Removing $K_{\alpha_{h_f}}$ from $\mathbf{p}_3$ in (6.7) (assuming a known constant value) could potentially improve the estimation of the other four parameters. Furthermore, the parameter fits for K_{BS_λ} in Section 6.4.2 returned the same value for all time periods.

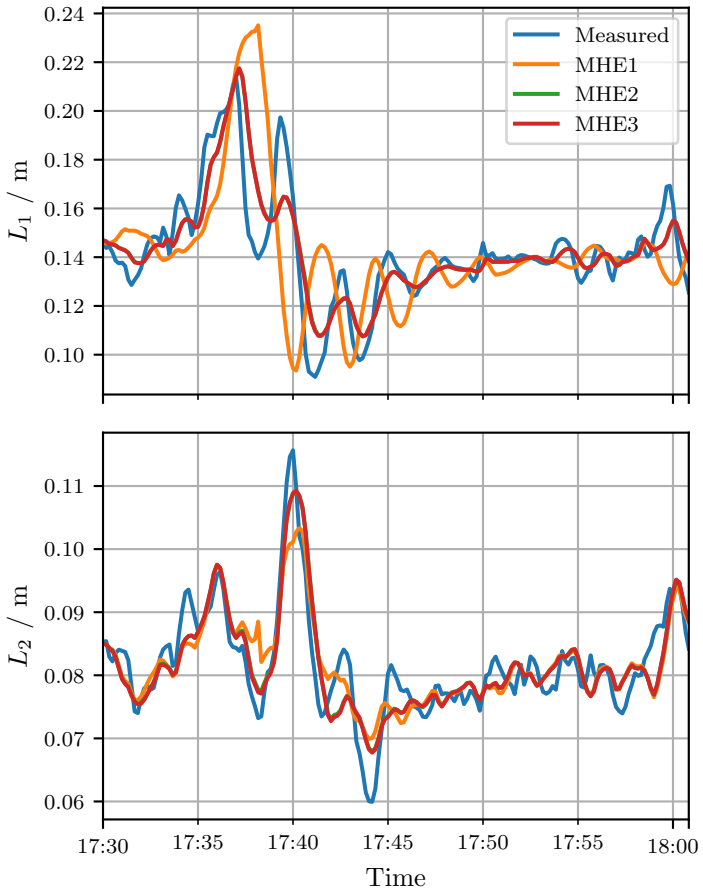

Figure 6.9. Extract of L_k estimation results for the three different MHEs. Note that MHE2 and MHE3 are virtually indistinguishable here. The lines were offset from their actual values to protect sensitive information.

Fixing this value could also improve the observability of the parameters given low the variance in the plant data.

Three versions of a similarly tuned MHEs were tested in order to investigate the use of $\mathbf{w}_k$ and estimation of $K_{\alpha_{J_g}}$ and K_{BS_λ}. The first (MHE1) excluded $K_{\alpha_{h_f}}$ and K_{BS_λ} from $\mathbf{p}_3$ and did not include $\mathbf{w}_k$. MHE2 also excluded the two parameters, but included heavily penalised $\mathbf{w}_k$ for each state. MHE3 made use of the same $\mathbf{P}_w$ tuning as MHE2, but estimated $K_{\alpha_{h_f}}$ and K_{BS_λ}. Figure 6.9 and Figure 6.10 show small representative extracts of the L_k and α_k estimates for the three estimators and the corresponding estimate RMSE values are given in Table 6.8 and Table 6.9. The results were similar, yet less pronounced, for the D_{BF_k} and L_{H_q} estimates.

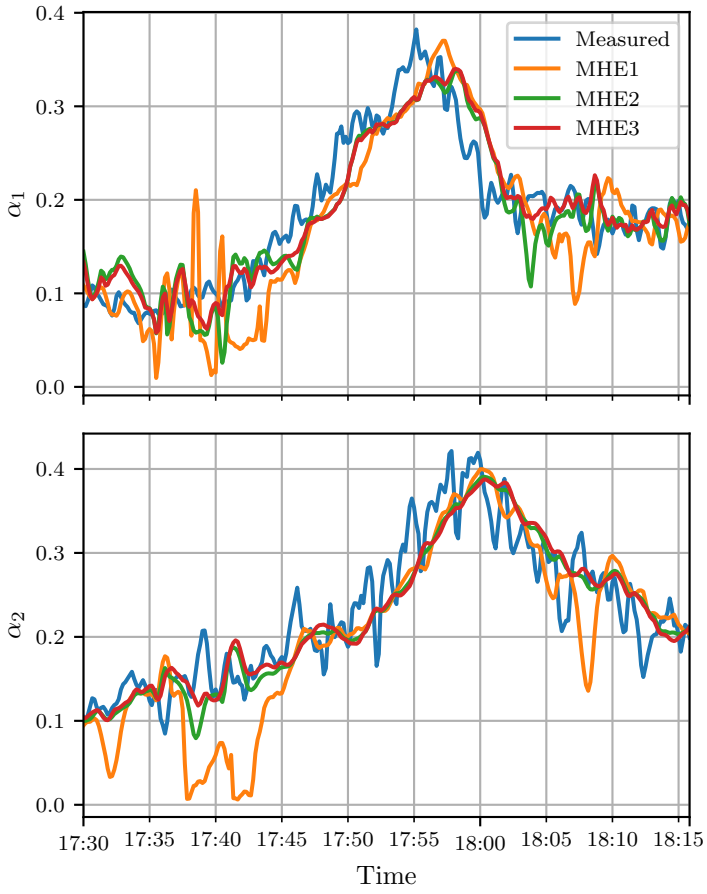

Figure 6.10. Extract of α_k estimation results for the three MHEs.

Table 6.8. L_k estimation performance metrics for three MHE designs.

Metric	L_1			L_2		
	MHE1	MHE2	MHE3	MHE1	MHE2	MHE3
RMSE$\times 10^3$	14.93	7.14	7.12	5.64	5.27	5.18
RMSEs$\times 10^3$	4.23	2.02	2.00	0.78	0.66	0.68
RMSEu$\times 10^3$	14.31	6.85	6.84	5.59	5.23	5.14
S/U	0.30	0.29	0.29	0.14	0.13	0.13

Table 6.9. α_k estimation performance metrics for three MHE designs.

Metric	α_1			α_2		
	MHE1	MHE2	MHE3	MHE1	MHE2	MHE3
RMSE$\times 10^3$	37.82	39.87	34.61	51.92	48.55	47.95
RMSEs$\times 10^3$	6.31	7.87	6.18	6.59	7.43	7.73
RMSEu$\times 10^3$	37.29	39.08	34.06	51.50	47.98	47.32
S/U	0.17	0.20	0.18	0.13	0.15	0.16

The use of $\mathbf{w}_k$ through a reasonable choice of $\mathbf{P}_w$ significantly improved the estimation of L_k. Note how in Figure 6.9 MHE2 and MHE3 track the rather oscillatory L_1 quite well, while MHE1 is almost completely out of phase. The MHE2 and MHE3 RMSE values in Table 6.8 of the L_1 estimation compared to the measurements are less than half that of MHE1. The improved state estimation also led to improved prediction accuracy, as the model could be initialised more accurately for a given time-step. The fact that α_k, D_{BF_k} and L_{H_1} prediction accuracy increased when $\mathbf{w}_k$ was included without the predictors (see Section 6.6.2) having access to $\mathbf{w}_k$ indicates that the $\mathbf{w}_k$ estimates did not overcompensate and decrease the parameter estimation accuracy significantly.

There was very little noticeable difference between MHE2 and MHE3 in the L_k estimates, but MHE3 showed some improvement over MHE2 in the estimation and prediction accuracy for α_k. The additional parameters resulted in smoother state estimates and less variable parameter estimates. This in turn provided a better starting state and more consistent parameters between time-steps which improved prediction accuracy slightly. The MHE3 tuning parameters were used in the final estimation results in Section 6.6 and are summarised in Table 6.10.

6.5.4 Initialisation

The fitted parameter values and nominal state estimates described in Section 6.4 were used as initial estimates for the MHE. Using significantly different initial state estimates had very little effect, as the states were measured directly. Using different initial parameters occasionally lead to significantly different results. This was typically the case if the $\mathbf{P}_p$ weights were higher, leading to slow changes in parameter estimates. However, typically the parameters quickly converged to the same values regardless of the initial value.

6.5.5 Tuning

The weightings of tuning matrices are specified in Table 6.10 (before normalisation). The final value of the weighting matrix entry for each of these is determined by dividing the specified value by the square of the operating range of the variable.

Additional non-linear constraints were required to obtain realistic parameter estimates. The steady-state air recovery ($\alpha_{k,ss}$)

$$\alpha_{k,ss} = K_{\alpha_{J_g}} \left(J_{g_k} - K_{0,\alpha_{J_{g_k}}} - K_{\alpha_{h_f}} h_{f_k} \right)^2 + \alpha_{OS_k}, \tag{6.12}$$

frequently converged to negative or much too large values and was constrained to positive values. This could be considered "over-training" the model, but the aim was to give the MHE the best possible chance.

Table 6.10. MHE tuning parameters.

Variable	Range	Lower Bound	Upper Bound	$\mathbf{P}_x$	$\mathbf{P}_p$	$\mathbf{P}_v$	$\mathbf{P}_w$
			States				
α_k	0.1	0	1	0.01	-	0.1	1×10^6
D_{BF_k}	20	10	120	0.0001	-	0.1	1×10^7
L_{H_q}	0.1	-	-	1	-	100	1×10^5
L_k	0.1	-	-	1	-	100	1×10^5
			Parameters				
α_{OS_k}	0.2	0.05	1.2	-	0.1	-	-
D_{OS_k}	20	-300	150	-	10	-	-
C_{v_k}	1000	1000	4000	-	0.1	-	-
$K_{0,\alpha_{J_{g_k}}}$	1	10, 2	25, 18	-	1	-	-
$K_{\alpha_{h_f}}$	0.001	-0.02	0.02	-	0.1	-	-
K_{BS_λ}	0.1	0	0.02	-	100	-	-
k_m	500	0.1	3000	-	0.001	-	-

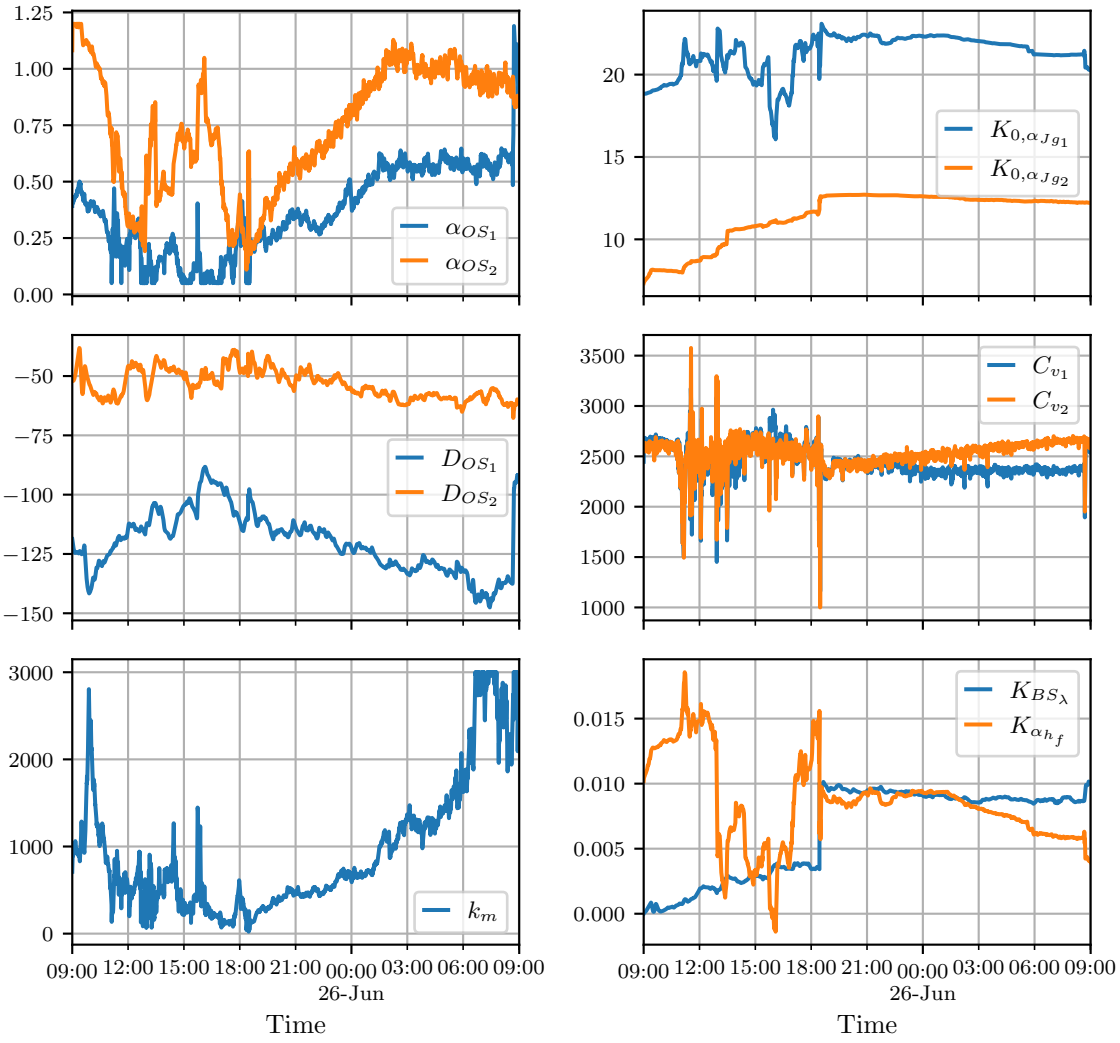

Figure 6.11. Estimated parameter values for MHE3 over the entire 25–26 June dataset. Note the much more gradual parameter changes after 21:00, which represents normal operation.

6.6 ESTIMATION AND PREDICTION PERFORMANCE

6.6.1 Estimation results

In the following sections the measured state values are used as the O_i values for the calculation of the error metrics (see Section 3.5), given in Table 6.11. The parameter estimates are shown in Figure 6.11. For the step test period (12:00–18:00), the parameter estimates vary considerably more and less gradually than during normal operation (after 18:00). This is especially obvious in the α_k model parameters ($K_{\alpha_{h_f}}$, α_{OS_k} and $K_{0,\alpha_{J_{g_k}}}$) and C_{v_k}.

The changes in $K_{0,\alpha_{J_{g_k}}}$ and α_{OS_k} are particularly correlated with the J_{g_k} steps. This indicates that the α_k model does not extrapolate very well to different operating conditions. The estimates for K_{BS_λ} and D_{OS_k} were significantly more consistent. This could indicate that the D_{BF_k} captures the investigated operating range quite well.

The large variance in k_m reflects the many different recovery mechanisms contained in this single parameter. Without more detailed knowledge about the true mass and water recoveries, further discussion regarding the realism of these estimates is difficult.

The large changes in C_{v_k} correspond to changes in L_1, with C_{v_2} lagging slightly behind C_{v_1}. With the uncertainty regarding Q_{F_1} (see Section 6.2.2.1) it is difficult to attribute these changes to either disturbances in Q_{F_1}, incorrect Q_{C_k} estimates or modelling error of Q_{T_k} in (4.1) brought about by assuming a single valve fraction for the two split-control dart valves (see Section 6.2.2). The fact that both C_{v_k} values change in a similar fashion is indicative of density or Q_{F_1} disturbances, but it is then somewhat strange that it only occurs during the first half of the data.

Due to higher weightings, the changes in $K_{0,\alpha_{J_{g_k}}}$ are very gradual. This raises the concern that parameter convergence might take very long. For example, $K_{0,\alpha_{J_{g_2}}}$ tends gradually downward after 19:00 in Figure 6.11. When the MHE is initiated at 21:00 with a $K_{0,\alpha_{J_{g_2}}}$ of around 7 (much lower), the estimate remains relatively constant. Attempts were made to allow faster convergence over the entire period, if the lower value was indeed a better fit for the second half of data (21:00–09:00). However, this merely resulted in much more variability in the state and parameter estimates while $K_{0,\alpha_{J_{g_2}}}$ was again estimated at similarly high values. Further tuning might change the outcome, but it is currently unclear which changes would need to be made.

The L_k and L_{H_1} state estimates followed the measurements quite closely throughout (see Figure 6.9 MHE3 for a representative section) resulting in low estimation RMSE values for these states (see Table 6.11). As these measurements are expected to be relatively reliable, this was deemed acceptable. Due to the measurement uncertainty of α_k and D_{BF_k}, "good" estimates of these states are expected to reject a considerable amount of variance as measurement noise. Sample sections of the α_k and D_{BF_k} estimates are given in Figure 6.12, Figure 6.13 and Figure 6.14.

Considering the frequent and quite large step changes in the plant operating conditions during the

Table 6.11. Estimation and prediction error metrics.

	Estimation				Prediction			
State	RMSE	RMSEs	RMSEu	S/U	RMSE	RMSEs	RMSEu	S/U
L_1	0.007	0.002	0.007	0.29	0.059	0.003	0.059	0.05
L_2	0.005	0.001	0.005	0.13	0.026	0.001	0.026	0.04
L_{H_1}	0.008	0.000	0.008	0.03	0.112	0.022	0.110	0.20
α_1	0.035	0.006	0.034	0.18	0.048	0.009	0.047	0.19
α_2	0.048	0.008	0.047	0.16	0.077	0.017	0.075	0.22
D_{BF_1}	0.996	0.639	0.764	0.84	1.063	0.646	0.843	0.77
D_{BF_2}	1.299	0.726	1.077	0.67	1.353	0.763	1.117	0.68

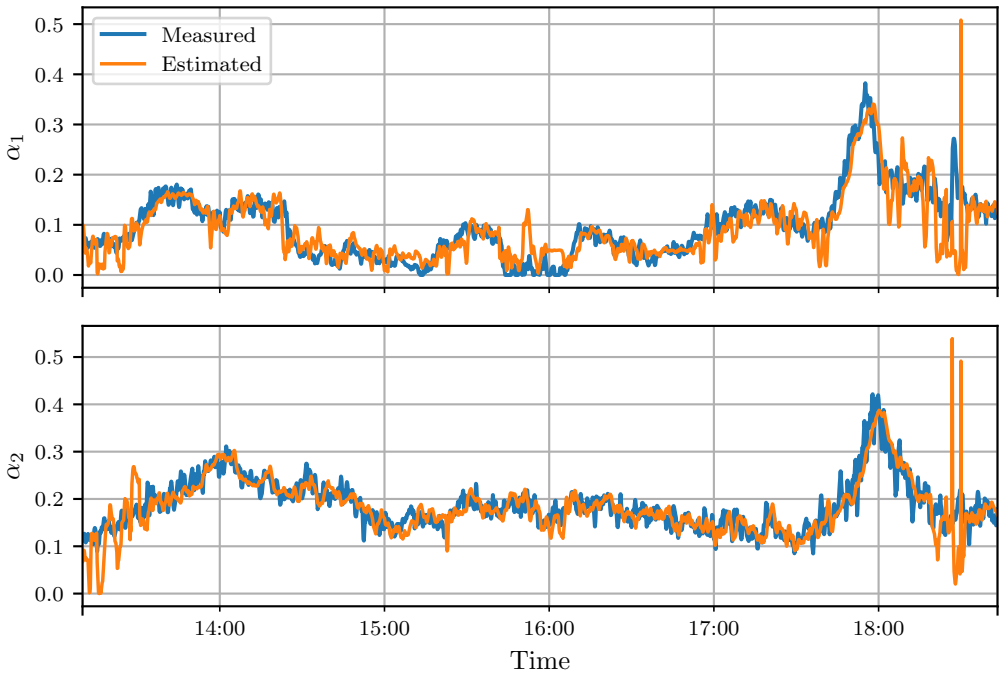

Figure 6.12. State estimates for α_k during the J_{g_k} and frother step tests.

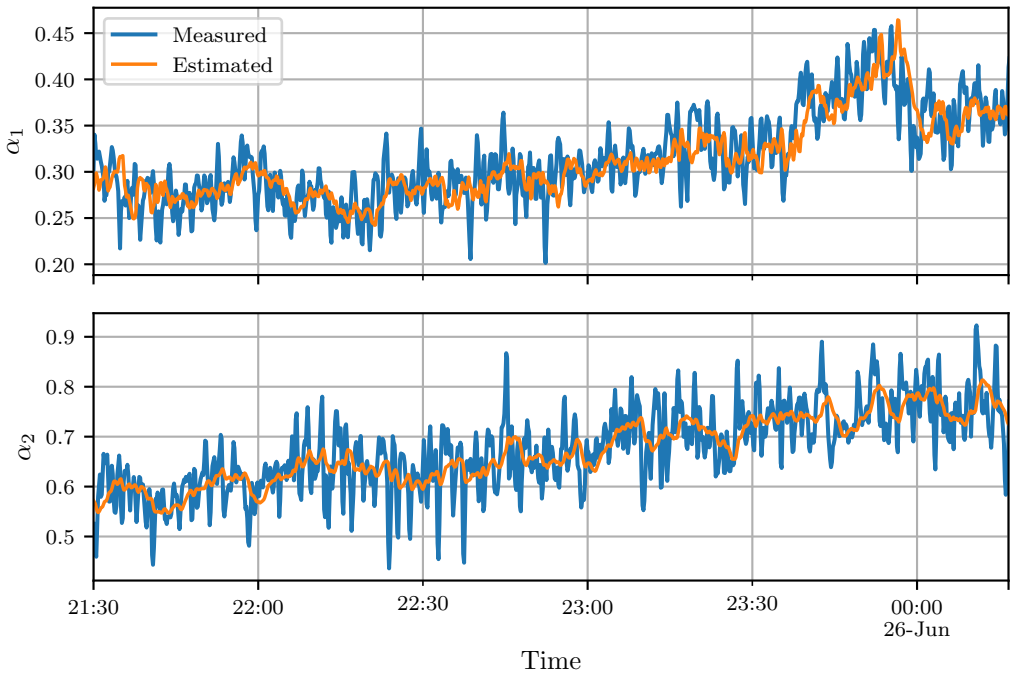

Figure 6.13. State estimates for α_k during normal operation with less input variance.

period shown in Figure 6.12, the α_k estimates are reasonably good. Occasionally the estimates nearly mirrors the movement of the measurement (*e.g.,* α_1 from 18:00–18:20) or deviate significantly from the measured values (*e.g.,* α_1 just before and after 16:00 and both α_k values at around 13:20 and 18:20–18:40). It is possible that deficiency in the modelling of α_k and/or Q_{C_k} forces estimates further from the measurements in order to better fit the level measurements and volume balances. While the α_k estimates in Figure 6.13 appear to reject more noise than in Figure 6.12, the estimated dynamics lag slightly behind the measured dynamics. This is due to the more gradual parameter estimates during 21:00–09:00 (see Figure 6.11). This could indicate that the model is being adjusted to account for unmodelled factors causing the movement in α_k. This is confirmed by the predictions during 21:00–09:00, as discussed in Section 6.6.2.

The D_{BF_k} estimates correlate relatively well with the observed dynamics. Results similar to those shown in Figure 6.14 were obtained over the entire time period. The estimate of a step change sometimes overshoots the measured change, *e.g.,* the downward steps in D_{BF_1} around 14:30 and 15:45 in Figure 6.14. Assuming a constant $K_{BS_{J_g}}$ (Section 6.5.3) and a very small K_{BS_λ} (Figure 6.11) mean that the input effects of J_{g_k} are overestimated. The estimator compromises between the overshooting model predictions and measurements based on the relative certainties (Table 6.10).

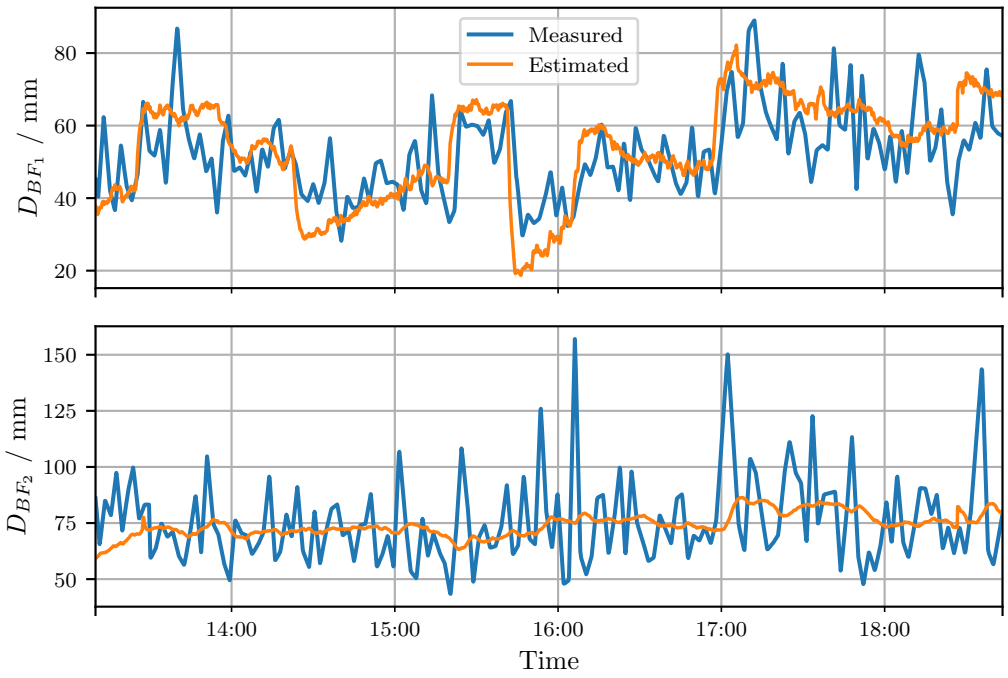

Figure 6.14. State estimates for D_{BF_k} during the step tests. Similar results were obtained for the other sections of data.

Apart from D_{BF_k}, the S/U ratios of the state estimates (Table 6.11) were quite small, which points to reasonable estimates which reject noise effectively (Willmott, 1981). The higher values for D_{BF_k} are likely caused by the overshooting D_{BF_k} estimates described above. Due to the longer D_{BF_k} sampling time (see Section 6.2.2.4), it is more difficult to determine which D_{BF_k} movement is purely due to measurement noise rather than actual dynamics. The interpolation required for resampling also created many additional data-points which would skew the error metrics. This would also increase the S/U ratio. As D_{BF_k} plays such an integral role in the Q_{C_k} model, future investigation into the measurements and model are warranted.

The problems of a fixed $K_{BS_{J_g}}$ could be addressed by using dual state estimation (Olivier et al., 2012). This could allow for two separate parameter subsets to be estimated in an alternating manner. The use of an EKF for this purpose could also be made possible if the subsets are chosen such that the LOIC is satisfied.

6.6.2 Prediction results

Oosthuizen (2023) uses an MPC horizon spanning 5 min. While the dynamics for MNC are slower than in his case, this was selected as the minimum prediction horizon for which the model should be able to predict. For the model prediction to be deemed sufficiently accurate for feedback control, it should be able to predict at least directional changes over this period.

The prediction accuracy over the entire section of data was tested by initialising a simulator for every 5 min section of data using the state and parameter estimates at the start of that section. The states were simulated over the rest of the section using the real inputs. No further state or parameter updates were made over the course of the simulation and the simulator did not have access to $\mathbf{w}_k$ estimates. The prediction results for the step test section are shown in Figure 6.15, Figure 6.16 and Figure 6.17. The RMSE, RMSEs and RMSEu for the various predictions are given in Table 6.11.

First consider the L_k predictions. There are sections for which the predictions are excellent, *e.g.,* the first two steps in L_1 in Figure 6.15 and large sections of L_2, especially around 15:00. However, quite often the prediction diverges substantially from the measured values, *e.g.,* the downward step in L_1 around 16:00 and the steps after 18:00. A possible cause of these inaccuracies are changes in the unknown Q_{F_1} that would have a significant impact on both levels, but especially L_1. The C_{v_k} estimates change significantly around these points (see Figure 6.11) which means that a single C_{v_k} value cannot predict correctly over 5 min.

Another very likely possibility is incorrect Q_{C_k} estimates, though this is not always the case. For one of the more obvious deviations from both L_1 and L_{H_1} (at 15:20), L_{H_1} (see Figure 6.17) is under-predicted indicating too low Q_{C_k} values. However, L_1 is also under-predicted which would require a too high Q_{C_1} value. Combined with a good prediction for L_2 at this point, it is more likely that underestimating Q_{F_1} would have caused the observed effect.

Regardless of the cause, the inclusion of $\mathbf{w}_k$ in the estimation model would have compensated somewhat for the model inaccuracies. However, the prediction model does not have access to $\mathbf{w}_k$ over the future prediction period, which certainly decreased the prediction model performance.

The predictions of all other states were repeated with the estimated and measured L_k values used at every step in the prediction in order to limit the possible impact of incorrect L_k predictions. The

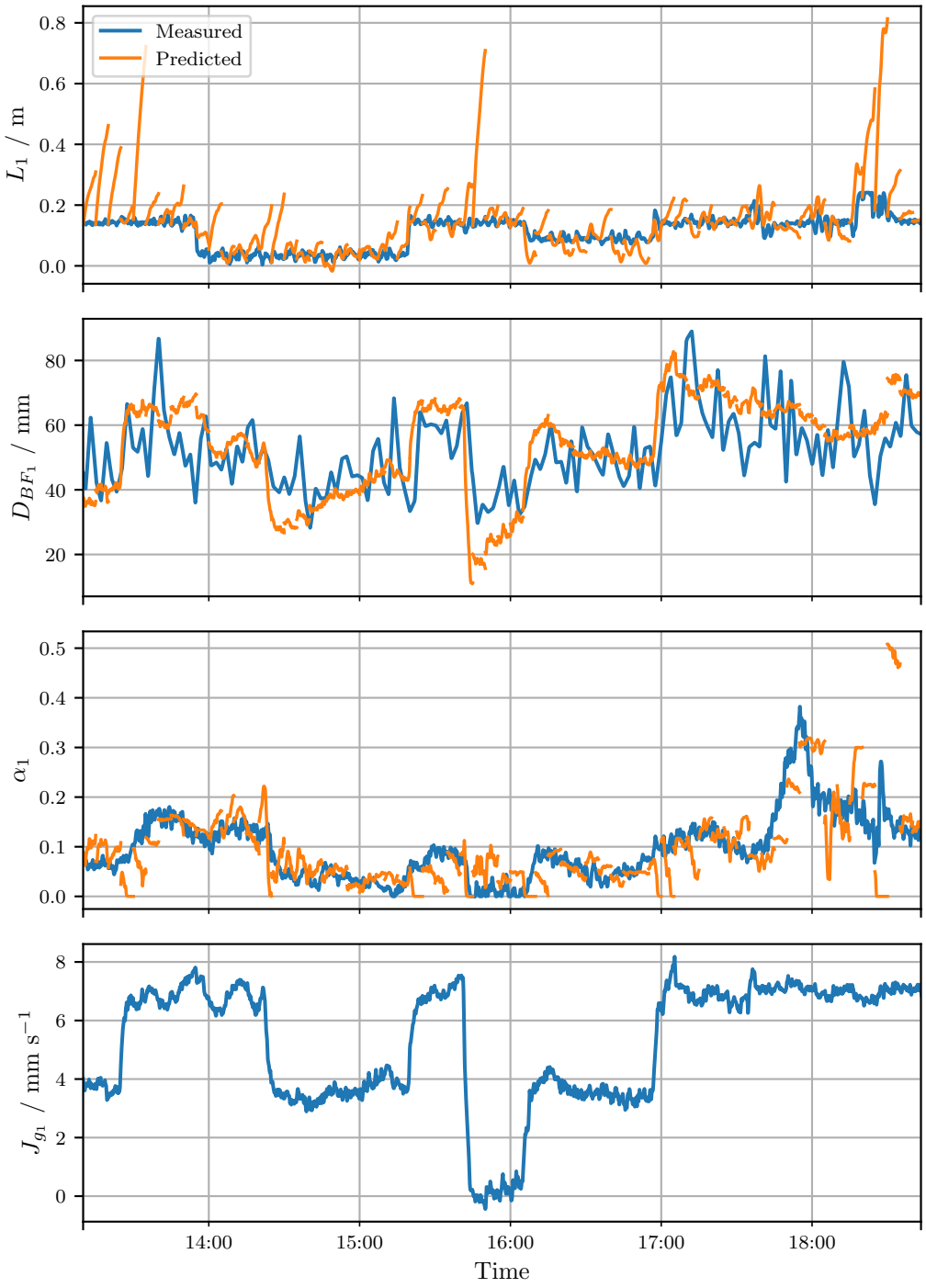

Figure 6.15. FT1 state predictions initialised every 5 min in the step test time period. L_1 and J_{g_1} were offset from their actual values to protect sensitive operational information.

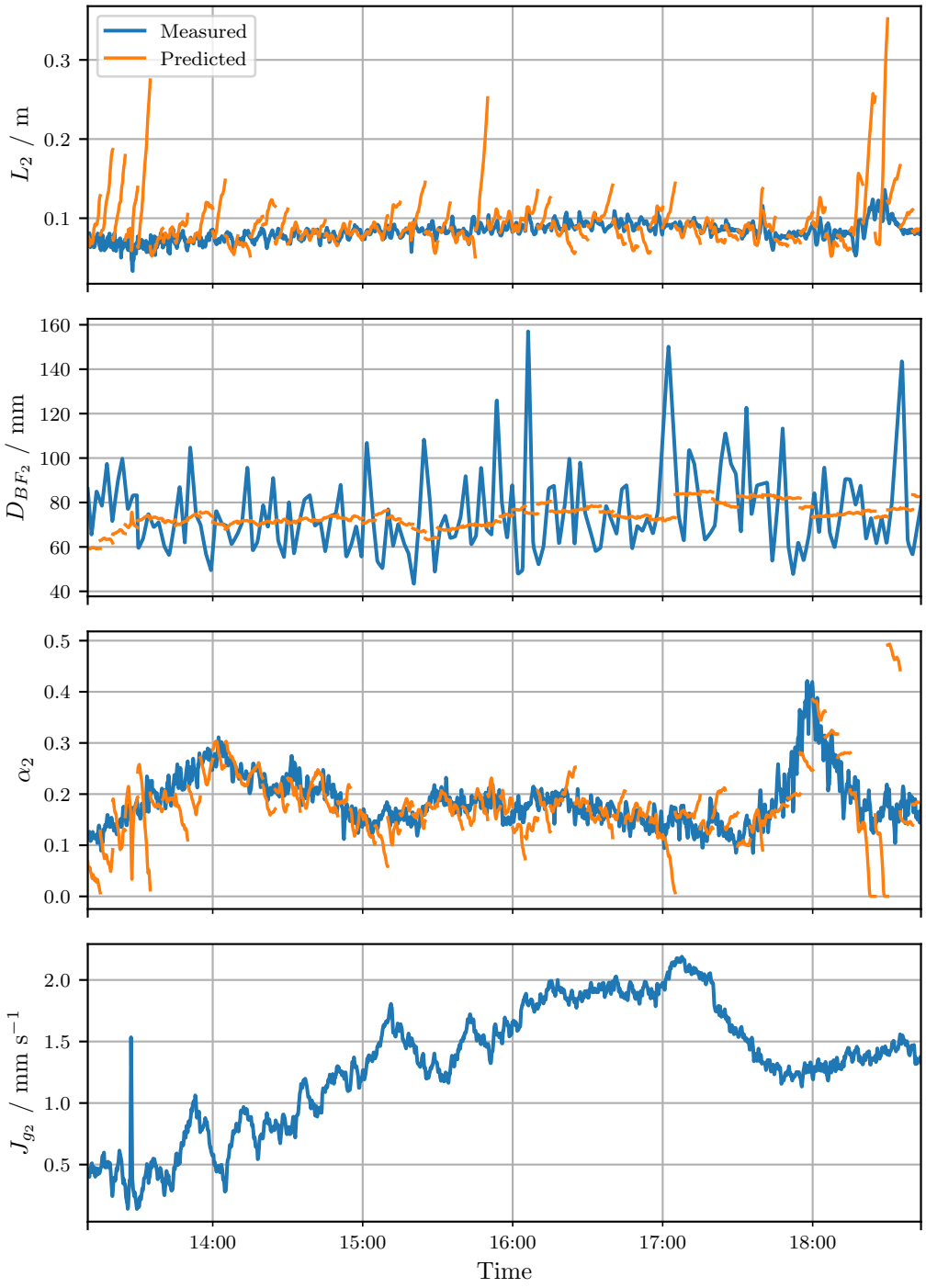

Figure 6.16. FT2 state predictions initialised every 5 min in the step test time period. L_2 and J_{g_2} were offset from their actual values to protect sensitive operational information.

results were visually very similar whether using the L_k measurements or predictions and there were only small reductions in the prediction error metrics in Table 6.11 when the L_k measurements were used. Therefore, only the results including predictions for L_k are shown without loss of the overall observations and arguments made.

The J_{g_k} steps had a clear impact on D_{BF_1} and the predictions in Figure 6.15 were very good, apart from the overshoot discussed in Section 6.6.1. The RMSE values showed that there was very little difference between the estimated and predicted D_{BF_k}. One issue with the D_{BF_k} model is seen in the last two hours of the D_{BF_2} predictions in Figure 6.16. If the J_{g_k} remains the same, the model prediction over the whole prediction horizon will be the same. If D_{BF_k} is changing due to some unmodelled factor, this results in the disjointed horizontal lines observed in the predictions. The only tracking then achieved between horizons is due to the update of the state estimate and reinitialisation of the prediction model. A very similar issue is observed in the later sections (not shown) of α_k, which is consistent with the lagging state estimates described earlier.

The prediction quality of the α_k values in Figure 6.15 and Figure 6.15 are somewhat inconsistent. For example, the α_1 predictions between 13:30 and 15:10 were really quite good and at least captured the directionality of α_1 changes. In contrast, after 17:00 the direction of α_1 is more often predicted incorrectly than not. The α_1 and α_2 prediction accuracy is especially bad after 17:50. Incorrect α_k predictions would result in incorrect Q_{C_k} values, which then has a cascading effecting throughout the model. The α_k model parameter estimates in Figure 6.11 vary significantly. While changes in the plant dynamics are certainly expected due to the changing frother dosages, it is also quite possible that the parameter estimates are compensating too much for unmodelled dynamics. The S/U ratios for the α_k predictions in Table 6.11 are quite low, which is promising yet difficult to reconcile with some of the observed prediction errors.

It is relatively easy to evaluate the accuracy of the levels, α_k and D_{BF_k} models by comparing the predicted values to the measured values. However, without measurements of Q_{C_k} the validity of (6.5) has to be evaluated based on the L_{H_1} predictions, shown in Figure 6.17. Some of the estimation and prediction inaccuracies could be attributed to the uncertainty regarding the actual tank dimensions.

The model predicts well on sections where the gradient of L_{H_1} is relatively constant (*e.g.,* 14:00–15:00 and 17:50–18:00), as the model would only need to keep predicting roughly the estimated Q_{C_k} value

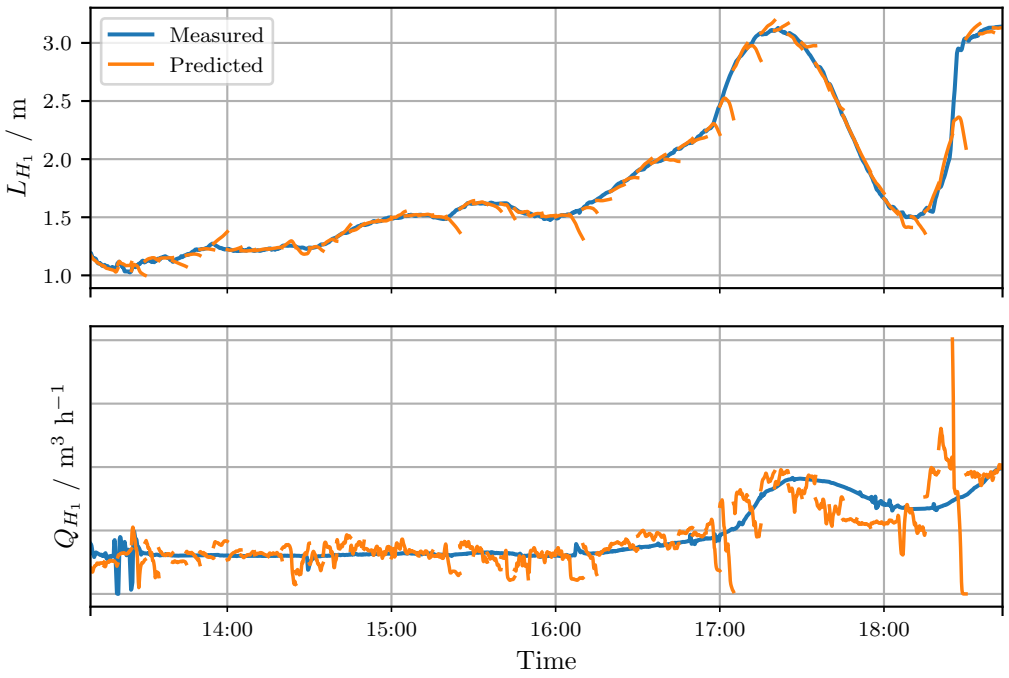

Figure 6.17. L_{H_q} predictions during the step test period, initialised every 5 min and the input Q_{H_1} along with the predicted $\sum Q_{C_k}$.

at the start of the section. However, in general the predictions around a drastic change in gradient of L_{H_1} are quite poor (*e.g.,* at around 15:20 and during 17:00–17:30) and the L_{H_1} prediction goes in the opposite direction from the measured values. Where the S/U ratio for α_k and D_{BF_k} stayed roughly the same and decreased for L_k, L_{H_1} saw a drastic increase in its prediction RMSEs and S/U compared to the estimation metrics (see Table 6.11).

6.7 CONCLUDING REMARKS

Given the absence of on-line compositional measurements and the resulting observability limitations discussed in Chapter 4, a reduced model of the first two cells in the bank was shown to be fully observable from the available data, which included experimental StoneThree pulp sensors. The reduced model has improved observability characteristics which should make successful on-line parameter estimation more achievable.

An MHE implementation was designed based on the reduced model. Overall, the state estimation results were quite reasonable, but lagging values indicated that the parameter estimation was likely compensating for a considerable amount of unmodelled dynamics or unmeasured disturbances. The C_{v_k}

estimates in particular varied quite drastically during the step test period, likely indicating disturbances in the unknown Q_{F_1}, density and/or gas hold-up. It is also possible that (4.1) does not accurately approximate the effect of the dual dart-valves. There were some issues with the parameter estimation of $K_{0,\alpha_{J_{g_k}}}$ and without more information regarding the actual concentrate flow-rates it is unclear whether the k_m estimates are realistic.

Due to the slower sampling time for D_{BF_k} (see Section 6.2.2.4), there still remains the question of whether the smaller oscillatory variance are actual dynamics not captured by the model or merely the result of measurement noise. However, the D_{BF_k} model captures the effect of J_{g_k} very well. There are certainly neglected factors which cause some variance in D_{BF_k}, but these do not appear to be as important or have such fast dynamics, as updating the D_{BF_k} model parameters accounts for most of the large scale variance.

Keeping in mind that the results shown above are the best achieved over many different data sections and MHE design iterations, the model and MHE in their current form are most likely not yet ready for actual control implementation on this particular industrial site, given the limitations and uncertainty of the available data. The Q_{C_k} model especially, whether in (6.5) or as a cascade effect of incorrect α_k predictions, has significant difficulty in predicting dynamic changes over a 5 min window.

However, these results show that the modelling and parameter estimation approach has some merit. Despite clear limitations in the α_k and D_{BF_k} models, reasonable predictions were obtained for several sections of data at very different operating conditions, showing that the real-time parameter estimates compensate for the simplicity of the models. Furthermore, the L_{H_1} predictions performed very well over large sections of data, but struggled at points where the gradient changed significantly. Alternative state estimation algorithms (*e.g.,* dual estimation (Olivier et al., 2012)), better data of possible feed disturbances, confirmation of J_{g_k} and α_k measurement quality and α_k and Q_{C_k} model improvements are all promising avenues of future investigation. The faster sampling of D_{BF_k} should also be considered.

CHAPTER 7 CONCLUSION

7.1 GENERAL DISCUSSION

This study expanded on the work of Oosthuizen (2023) in an attempt to validate his proposed model on dynamic plant data. While limitations in the available data hindered definitive conclusions regarding the validity of the model and parameter estimation approach, several valuable insights were obtained that can guide future model validation attempts and model improvements.

A Sobol sensitivity analysis showed that the α_k model is a dominant factor in both grade and recovery. A very simple inclusion of frother concentration in the α_k model was shown to be fully observable and identifiable. Importantly, the frother concentrations in the cells were found to be observable using only measurements available on modern flotation circuits. While this model is unlikely to have proper predictive capability, this demonstrates that including reagent dosage as an MV in a model-based control strategy is theoretically possible and it might even be possible to develop soft sensors for reagent concentration if their effect on measurable variables is modelled.

An expanded observability analysis showed that non-zero high order derivatives of the outputs are required for full parameter identifiability and that input excitation significantly improves the observability and identifiability of the model. As typical plants are operated such that the output variability is minimised and adding additional input excitation is generally considered undesirable, this indicates potential issues with the practical identifiability of the model parameters.

Several aspects of the available industrial data were identified that would also limit the possible estimation and prediction performance. There was some uncertainty regarding the actual values of J_{g_k} as the conversion between the air feed rate and J_{g_k} values measured using the StoneThree pulp sensor did not appear to remain constant. However, the pulp sensor data proved invaluable in calibrating the

inputs to the model. As J_{g_k} plays an integral part in nearly all the recent phenomenological flotation models, the further development and validation of such pulp sensors will likely be necessary for future model-based control implementation.

The need for dedicated, or at least calibrated α_k measurements was also identified. Constructing these measurements from several other plant measurements without a reference for calibration proved quite challenging. As α_k also has a significant impact on the model, this likely introduced significant uncertainty that the state estimation had to contend with.

The lack of on-line compositional analysers, especially for the concentrate grade, was shown to severely limit the potential scope of the model and parameter estimation. Shiftly assays were found to be unsuitable substitutes for these measurements due to the loss of dynamic information, especially since there are no assays of the concentrate grade. This prevented the implementation of the full model and a reduced model had to be developed. The reduced model was fully observable and identifiable, but required several simplifying assumptions that likely reduced the model predictive capability and range considerably. While such measurements are quite expensive and maintenance intensive (Brooks and Koorts, 2017), it is highly unlikely that any model-based control strategy for flotation would be successful without some form of on-line compositional measurement.

Slightly lagging state estimates and considerable variance in the parameter estimates indicated that the parameter estimation was compensating for a considerable amount of unmodelled dynamics or unmeasured disturbances. This was supported by steady-state parameter fits of the α_k and D_{BF_k} models which fitted certain sections of data very well (especially for the second cell), while failing to capture the dynamics of other periods.

The updated model predictions over a 5 min horizon also contained sections where the dynamic changes of the states were captured very well and sections where the model completely diverged from the observed values. It is possible that eliminating the uncertainty in measurements (especially Q_{F_1} and α_k) discussed above and improved parameter estimation could result in significant improvement in the prediction accuracy and even prediction horizon. While the model and MHE in their current state are not suitable for model-based control, this approach shows enough potential to warrant further investigation.

7.2 RECOMMENDATIONS FOR FUTURE WORK

Based on the limitations and results discussed in Section 7.1, there are several possible avenues for future investigation that could lead to more definitive conclusions.

- Most importantly, the current study should be repeated after eliminating the uncertainty of the concentrate tank dimensions, J_{g_k} and D_{BF_k} measurements and Q_{F_1} values. Furthermore, the model should be tested on a circuit where the necessary on-line compositional measurements are available.
- Updating identifiable parameters in real-time shows enough potential that future modelling work should at least consider the observability and identifiability of the model and parameters. Currently, this is mostly overlooked or at least not explicitly reported.
- Investigating other state estimation algorithms, such as dual state estimation (Olivier et al., 2012), should also be considered. Dual estimation in particular could yield much improved parameter estimation. The parameters are split into identifiable subsets which are estimated in an alternating fashion. Carefully choosing the subsets such that linear observability is obtained would allow the use of EKF or simply improve the estimation behaviour. The high order derivatives required for observability of the model indicate potential practical estimation problems which could be avoided using this approach. Better parameter estimates could in turn lead to better predictions or at the very least help to identify sections of the model which must be improved. The use of multi-rate state estimator should also be investigated given the different sampling times of some of the measurements (particularly D_{BF_k} and composition measurements).
- The observability of frother concentrations in the model expanded to include reagent effects is very promising. The development of soft-sensors for reagent concentration should be investigated.
- Nearly all aspects of the work done would benefit greatly from automated data classification similar to what was proposed in Shardt and Brooks (2018). Automating the process of identifying good sections of data suitable for model validation would reduce the effort and error involved considerably.

REFERENCES

Aldrich, C., Marais, C., Shean, B. J. and Cilliers, J. J. (2010). Online monitoring and control of froth flotation systems with machine vision: A review, *International Journal of Mineral Processing* **96**(1–4): 1–13.

Alves Dos Santos, N., Savassi, O., Peres, A. E. C. and Martins, A. H. (2014). Modelling flotation with a flexible approach – integrating different models to the compartment model, *Minerals Engineering* **66–68**: 68–76.

Armstrong, J. S. and Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons, *International Journal of Forecasting* **8**: 69–80.

Backi, C. J., Gravdahl, J. T. and Skogestad, S. (2020). Combined state and parameter estimation for not fully observable dynamic systems, *IFAC Journal of Systems and Control* **13**: 100103.

Bascur, O. and Kuyvenhoven, R. (2010). A flotation model framework for dynamic performance monitoring, *Proceedings of the VII International Mineral Processing Seminar*, Vol. 1, pp. 328–339.

Bascur Riquelme, O. A. (1982). *Modelling and computer control of a flotation cell*, PhD thesis, The University of Utah.

Bergh, L. and Yianatos, J. (2013). Control of rougher flotation circuits aided by industrial simulator, *Journal of Process Control* **23**(2): 140–147.

Bicak, O., Ekmekci, Z., Bradshaw, D. J. and Harris, P. J. (2007). Adsorption of guar gum and CMC on pyrite, *Minerals Engineering* **20**(10): 996–1002.

Botchkarev, A. (2019). A new typology design of performance metrics to measure errors in machine learning regression algorithms, *Interdisciplinary Journal of Information, Knowledge & Management* **14**: 45.

Brooks, K. S. and Koorts, R. (2017). Model predictive control of a zinc flotation bank using online X-ray fluorescence analysers, *IFAC-PapersOnLine* **50**(1): 10214–10219.

Brooks, K., Westcott, M. and Bauer, M. (2019). A combined MPC for milling and flotation – a simulation study, *IFAC-PapersOnLine* **52**(14): 24–29.

Bulatovic, S. M. (2007). *Handbook of flotation reagents: Chemistry, theory and practice: Volume 1: flotation of sulfide ores*, Elsevier.

Casali, A., Gonzalez, G., Agusto, H. and Vallebuona, G. (2002). Dynamic simulator of a rougher flotation circuit for a copper sulphide ore, *Minerals Engineering* **15**(4): 253–262.

Chai, T. and Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?: Arguments against avoiding RMSE in the literature, *Geoscientific Model Development* **7**(3): 1247–1250.

Daum, F. (2005). Nonlinear filters: Beyond the Kalman filter, *IEEE Aerospace and Electronic Systems Magazine* **20**(8): 57–69.

Fairthorne, G., Fornasiero, D. and Ralston, J. (1997). Interaction of thionocarbamate and thiourea collectors with sulphide minerals: A flotation and adsorption study, *International Journal of Mineral Processing* **50**(4): 227–242.

Finch, J. A., Nesset, J. E. and Acuña, C. (2008). Role of frother on bubble production and behaviour in flotation, *Minerals Engineering* **21**(12–14): 949–957.

Finfer, D. (2018). Flotation process metering of concentrate, slurry, air and water flows using non-intrusive fibre-optic sensing, *Proceedings of the Copper Cobalt Africa, 9th Southern African Base Metals Conference*, pp. 181–191.

Fu, Y. and Aldrich, C. (2019). Flotation froth image recognition with convolutional neural networks, *Minerals Engineering* **132**: 183–190.

Gélinas, S. and Finch, J. A. (2007). Frother analysis: Some plant experiences, *Minerals Engineering* **20**(14): 1303–1308.

Gupta, A. and Yan, D. S. (2016). *Mineral Processing Design and Operations: An Introduction*, 2nd edn, Elsevier, Amsterdam.

Hadler, K. and Cilliers, J. J. (2009). The relationship between the peak in air recovery and flotation bank performance, *Minerals Engineering* **22**(5): 451–455.

Hadler, K., Greyling, M., Plint, N. and Cilliers, J. J. (2012). The effect of froth depth on air recovery and flotation performance, *Minerals Engineering* **36–38**: 248–253.

Hadler, K., Smith, C. D. and Cilliers, J. J. (2010). Recovery vs. mass pull: The link to air recovery, *Minerals Engineering* **23**(11–13): 994–1002.

Harris, A., Venkatesan, L. and Greyling, M. (2013). A practical approach to plant-scale flotation optimization, *Journal of the Southern African Institute of Mining and Metallurgy* **113**(3).

Haseltine, E. L. and Rawlings, J. B. (2005). Critical evaluation of extended Kalman filtering and moving-horizon estimation, *Industrial & Engineering Chemistry Research* **44**(8): 2451–2460.

Herman, J. and Usher, W. (2017). SALib: An open-source Python library for sensitivity analysis, *The Journal of Open Source Software* **2**(9): 97.

Hodouin, D., Bazin, C., Gagnon, E. and Flament, F. (2000). Feedforward–feedback predictive control of a simulated flotation bank, *Powder Technology* **108**(2–3): 173–179.

Horn, Z. C., Haasbroek, A. L., Nienaber, E. C., Auret, L. and Brooks, K. S. (2022). Comparison of online and offline pulp sensor metrics in an industrial setting, *IFAC-PapersOnLine* **55**(21): 91–96.

HSL (2013). A collection of Fortran codes for large scale scientific computation, Oxford, England. http://www.hsl.rl.ac.uk/. Last accessed on 22 February 2023.

Hu, W. (2014). *Flotation circuit optimisation and design*, PhD thesis, Imperial College London.

Hu, W., Hadler, K., Neethling, S. J. and Cilliers, J. J. (2013). Determining flotation circuit layout using genetic algorithms with pulp and froth models, *Chemical Engineering Science* **102**: 32–41.

Impraimakis, M. and Smyth, A. W. (2022). An unscented Kalman filter method for real time input-parameter-state estimation, *Mechanical Systems and Signal Processing* **162**: 108026.

Jin, X., Jeremiah, R. J. R., Su, T., Bai, Y. and Kong, J. (2021). The new trend of state estimation: From model-driven to hybrid-driven methods, *Sensors* **21**(6): 2085.

Joostberens, J., Rybak, A. and Rybak, A. (2022). Identification of the dynamic properties of the coal flotation process as a control object with the use of the Kalman filter, *Energies* **15**(21): 7926.

Jovanović, I. and Miljanović, I. (2015). Contemporary advanced control techniques for flotation plants with mechanical flotation cells – a review, *Minerals Engineering* **70**: 228–249.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *Journal of basic Engineering* **82**(1): 35–45.

Karakashev, S. I., Grozev, N. A., Batjargal, K., Guven, O., Ozdemir, O., Boylu, F. and Çelik, M. S. (2020). Correlations for easy calculation of the critical coalescence concentration (CCC) of simple frothers, *Coatings* **10**(7): 612.

Knight, J. W. and Knights, B. D. H. (2011). Industrial trial of Mintek's xanthoprobe at the Eland platinum concentrator, *Proceedings of the 6th Southern African Base Metals Conference*, pp. 87–98.

Kulikov, G. Y. and Kulikova, M. V. (2014). Accurate state estimation in the Van der Vusse reaction, *2014 IEEE Conference on Control Applications (CCA)*, pp. 759–764.

Kulikov, G. Y. and Kulikova, M. V. (2015). State estimation in chemical systems with infrequent measurements, *2015 European Control Conference (ECC)*, pp. 2688–2693.

Le Roux, J. D., Steinboeck, A., Kugi, A. and Craig, I. K. (2017). An EKF observer to estimate semi-autogenous grinding mill hold-ups, *Journal of Process Control* **51**: 27–41.

Liu, J. P., Gui, W. H. and Tang, Z. H. (2013). Flow velocity measurement and analysis based on froth image SIFT features and Kalman filter for froth flotation, *Turkish Journal of Electrical Engineering and Computer Sciences* **21**: 2378–2396.

Lucia, S., Tătulea-Codrean, A., Schoppmeyer, C. and Engell, S. (2017). Rapid development of modular and sustainable nonlinear model predictive control solutions, *Control Engineering Practice* **60**: 51–62.

Maldonado, M., Desbiens, A., del Villar, R. and Aguilera, R. (2010). On-line estimation of frother concentration for flotation processes, *Canadian Metallurgical Quarterly* **49**(4): 435–446.

Maldonado, M., Desbiens, A., Villar, R. D., Poulin, E. and Riquelme, A. (2010). Nonlinear control of bubble size in a laboratory flotation column, *IFAC Proceedings Volumes* **43**(9): 19–24.

Mathieu, J. L., Koch, S. and Callaway, D. S. (2013). State estimation and control of electric loads to manage real-time energy imbalance, *IEEE Transactions on Power Systems* **28**(1): 430–440.

McFadzean, B., Marozva, T. and Wiese, J. (2016). Flotation frother mixtures: Decoupling the sub-processes of froth stability, froth recovery and entrainment, *Minerals Engineering* **85**: 72–79.

Neethling, S. J., Brito-Parada, P. R., Hadler, K. and Cilliers, J. J. (2019). The transition from first to zero order flotation kinetics and its implications for the efficiency of large flotation cells, *Minerals Engineering* **132**: 149–161.

Neethling, S. J. and Cilliers, J. J. (2003). Modelling flotation froths, *International Journal of Mineral Processing* **72**(1–4): 267–287.

Neethling, S. J. and Cilliers, J. J. (2009). The entrainment factor in froth flotation: Model for particle size and other operating parameter effects, *International Journal of Mineral Processing* **93**(2): 141–148.

Olivier, L. E., Huang, B. and Craig, I. K. (2012). Dual particle filters for state and parameter estimation with application to a run-of-mine ore mill, *Journal of Process Control* **22**(4): 710–717.

Oosthuizen, D. J. (2023). *A dynamic flotation model for real-time control and optimisation*, PhD thesis, University of Pretoria.

Oosthuizen, D. J., Craig, I. K., Jämsä-Jounela, S. L. and Sun, B. (2017). On the current state of flotation modelling for process control, *IFAC-PapersOnLine* **50**(2): 19–24.

Oosthuizen, D. J., Le Roux, J. D. and Craig, I. K. (2021). A dynamic flotation model to infer process characteristics from online measurements, *Minerals Engineering* **167**: 106878.

Pérez-Correa, R., González, G., Casali, A., Cipriano, A., Barrera, R. and Zavala, E. (1998). Dynamic modelling and advanced multivariable control of conventional flotation circuits, *Minerals Engineering* **11**(4): 333–346.

Phillpotts, D., Whitehead, B. and Ramatsoma, S. (2021). Monitoring of air recovery for froth flotation optimisation on an industrial circuit, *Proceedings of the XXX International Mineral Processing Congress*, pp. 3348–3357.

Popli, K., Afacan, A., Liu, Q. and Prasad, V. (2018). Real-time monitoring of entrainment using fundamental models and froth images, *Minerals Engineering* **124**: 44–62.

Popli, K., Sekhavat, M., Afacan, A., Dubljevic, S., Liu, Q. and Prasad, V. (2015). Dynamic modeling and real-time monitoring of froth flotation, *Minerals* **5**(3): 570–591.

Pu, Y., Szmigiel, A., Chen, J. and Apel, D. B. (2020). FlotationNet: A hierarchical deep learning network for froth flotation recovery prediction, *Powder Technology* **375**: 317–326.

Putz, E. and Cipriano, A. (2015). Hybrid model predictive control for flotation plants, *Minerals Engineering* **70**: 26–35.

Quintanilla, P., Neethling, S. J. and Brito-Parada, P. R. (2021). Modelling for froth flotation control: A review, *Minerals Engineering* **162**: 106718.

Quintanilla, P., Neethling, S. J., Mesa, D., Navia, D. and Brito-Parada, P. R. (2021). A dynamic flotation model for predictive control incorporating froth physics. Part II: Model calibration and validation, *Minerals Engineering* **173**: 107190.

Quintanilla, P., Neethling, S. J., Navia, D. and Brito-Parada, P. R. (2021). A dynamic flotation model for predictive control incorporating froth physics. Part I: Model development, *Minerals Engineering* **173**: 107192.

Rawlings, J. B. and Bakshi, B. R. (2006). Particle filtering and moving horizon estimation, *Computers & Chemical Engineering* **30**(10): 1529–1541.

Rigatos, G. G. (2012). Nonlinear Kalman filters and particle filters for integrated navigation of unmanned aerial vehicles, *Robotics and Autonomous Systems* **60**(7): 978–995.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S. and Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices, *Environmental Modelling & Software* **114**: 29–39.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, *Computer Physics Communications* **181**(2): 259–270.

Savassi, O. N. (2005). A compartment model for the mass transfer inside a conventional flotation cell, *International Journal of Mineral Processing* **77**(2): 65–79.

Seborg, D. E., Mellichamp, D. A., Edgar, T. F. and Doyle III, F. J. (2011). *Process Dynamics and Control*, 3rd edn, John Wiley & Sons.

Shardt, Y. A. W. and Brooks, K. (2018). Automated system identification in mineral processing industries: A case study using the zinc flotation cell, *IFAC-PapersOnLine* **51**(18): 132–137.

Shean, B., Hadler, K. and Cilliers, J. J. (2017). A flotation control system to optimise performance using peak air recovery, *Chemical Engineering Research and Design* pp. 57–65.

Shean, B., Hadler, K., Neethling, S. and Cilliers, J. J. (2018). A dynamic model for level prediction in aerated tanks, *Minerals Engineering* **125**: 140–149.

Shean, B. J. and Cilliers, J. J. (2011). A review of froth flotation control, *International Journal of Mineral Processing* **100**(3–4): 57–71.

Simon, D. (2006). *Optimal state estimation: Kalman, H_∞ and Nonlinear Approaches*, Wiley-Interscience, New Jersey: Hoboken.

Steyn, C. and Sandrock, C. (2021). Causal model of an industrial platinum flotation circuit, *Control Engineering Practice* **109**: 104736.

Symengine Development Team (2021). Symengine. https://symengine.org/index.html. Last accessed on 22 February 2023.

Taguta, J., O'Connor, C. T. and McFadzean, B. (2017). The effect of the alkyl chain length and ligand type of thiol collectors on the heat of adsorption and floatability of sulphide minerals, *Minerals Engineering* **110**: 145–152.

Urbina, R. H. (2003). Recent developments and advances in formulations and applications of chemical reagents used in froth flotation, *Mineral Processing and Extractive Metallurgy Review* **24**(2): 139–182.

Van der Merwe, R. (2004). *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*, PhD thesis, Oregon Health & Science University.

Venkatesan, L., Harris, A. and Greyling, M. (2014). Optimisation of air rate and froth depth in flotation using a CCRD factorial design – PGM case study, *Minerals Engineering* **66–68**: 221–229.

Venter, J., Le Roux, J. D. and Craig, I. K. (2022). Observable dynamic models of reagent effects for model-based froth flotation control, *IFAC-PapersOnLine* **55**(21): 102–107.

Ventura-Medina, E. and Cilliers, J. J. (2002). A model to describe flotation performance based on physics of foams and froth image analysis, *International Journal of Mineral Processing* **67**(1–4): 79–99.

Villaverde, A. F., Tsiantis, N. and Banga, J. R. (2019). Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models, *Journal of The Royal Society Interface* **16**(156): 20190043.

Wang, G., Ge, L., Mitra, S., Evans, G. M., Joshi, J. B. and Chen, S. (2018). A review of CFD modelling studies on the flotation process, *Minerals Engineering* **127**: 153–177.

Wang, L., Peng, Y., Runge, K. and Bradshaw, D. (2015). A review of entrainment: mechanisms, contributing factors and modelling in flotation, *Minerals Engineering* **70**: 77–91.

Wiese, J., Harris, P. and Bradshaw, D. (2005). The influence of the reagent suite on the flotation of ores from the Merensky reef, *Minerals Engineering* **18**(2): 189–198.

Wiese, J., Harris, P. and Bradshaw, D. (2011). The effect of the reagent suite on froth stability in laboratory scale batch flotation tests, *Minerals Engineering* **24**(9): 995–1003.

Wiid, A. J., Le Roux, J. D. and Craig, I. K. (2021). Pressure buffering control to reduce pollution and improve flow stability in industrial gas headers, *Control Engineering Practice* **115**: 104904.

Willmott, C. J. (1981). On the validation of models, *Physical Geography* **2**(2): 184–194.

Wills, B. A. and Finch, J. (2015). *Wills' Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery*, Butterworth-Heinemann.

Xie, R., Zhu, Y., Liu, J. and Li, Y. (2021). The flotation behavior and adsorption mechanism of a new cationic collector on the separation of spodumene from feldspar and quartz, *Separation and Purification Technology* **264**: 118445.

Xie, Y., Wu, J., Xu, D., Yang, C. and Gui, W. (2017). Reagent addition control for stibium rougher flotation based on sensitive froth image features, *IEEE Transactions on Industrial Electronics* **64**(5): 4199–4206.

Xie, Y., Wu, J., Xu, D., Yang, C. and Gui, W. (n.d.). Reagent addition control for stibium rougher flotation based on sensitive froth image features, **64**(5): 4199–4206.

Yianatos, J. and Contreras, F. (2010). Particle entrainment model for industrial flotation cells, *Powder Technology* **197**(3): 260–267.

Zhang, J., Tang, Z., Xie, Y., Ai, M., Zhang, G. and Gui, W. (2021). Data-driven adaptive modeling method for industrial processes and its application in flotation reagent control, *ISA Transactions* **108**: 305–316.

Zhu, J., Gui, W., Yang, C., Xu, H. and Wang, X. (2014). Probability density function of bubble size based reagent dosage predictive control for copper roughing flotation, *Control Engineering Practice* **29**: 1–12.

ADDENDUM A REAGENT MODEL EXPANSION REVISIT

Since the completion of the work in Venter et al. (2022), the original model described in Section 4.2 has undergone some further modifications and the study was attempted for the latest version of the model as described in Section 4.3. The results for the updated study are discussed below.

A model expansion similar to ME2 (Section 5.4.2) was attempted for (4.23). Once again, the sensitivity analysis (see Table 5.1) showed that the air recovery model parameters dominate the recovery variance. As seen in Section 5.5, adjusting the steady-state α_k value (effectively modifying α_{OS_k}) resulted in the most realistic qualitative predictions. Therefore, (4.23) will be modified to

$$\frac{\mathrm{d}\alpha_k}{\mathrm{d}t} = \frac{K_{\alpha_{J_g}}\left(J_{g_k} - K_{0,\alpha_{J_{g_k}}} - K_{\alpha_{h_f}} h_{f_k}\right)^2 + K_{F\alpha} C_k^f + \alpha_{OS_k} - \alpha_k}{\lambda_{\mathrm{air}_k}}, \tag{A.1}$$

with the C_k^f modelled as before (5.1). Let this be Model Expansion 3 (ME3).

Unfortunately, the resulting state space system

$$\mathbf{x} = \left[\alpha_k,\ D_{BF_k},\ L_k,\ L_{H_q},\ M_k^{i,j},\ M_{H_q}^{i,j},\ M_k^f\right]^T, \tag{A.2}$$

$$\mathbf{p} = \left[\alpha_{OS_k},\ D_{OS_k},\ C_{PB},\ C_{v_k},\ K_{0,\alpha_{J_{g_k}}},\ K_{\alpha_{h_f}},\ K_{BSJ_g},\ K_{BS_\lambda},\ K_{F\alpha},\ K^{i,j}\right]^T, \tag{A.3}$$

$$\mathbf{u} = \left[J_{g_k},\ v_{a,k},\ Q_{H_q},\ \dot{M}_{F_1}^f\right]^T, \tag{A.4}$$

$$\mathbf{y} = \left[\alpha_k,\ D_{BF_k},\ L_k,\ L_H,\ G_{H_q}^{0,0},\ G_T^{0,0},\ C_{s,H},\ C_{s,T}\right]^T, \tag{A.5}$$

is no longer linearly observable. Table A.1 summarises the number of parameters resulting from the state-space system expanded for the four-cell rougher circuit in Figure 5.1. The non-linear observability analysis results are given in Table A.2.

Table A.1. Number of variables.

n_x	n_p	n_y	n_u
27	23	17	10

Table A.2. Observability analysis results for ME3.

Lie Derivative	Rank(d$\mathcal{O}$)			
Input excitation	None	J_{g_1}	$\dot{M}^f_{F_1}$	J_{g_1} & $\dot{M}^f_{F_1}$
$L^0_f g$	17	17	17	17
$L^1_f g$	34	34	34	34
$L^2_f g$	43	44	43	44
$L^3_f g$	47	48	47	48
$L^4_f g$	48	49	49	50
$L^5_f g$	48	49	49	50
Rank deficiency	2	1	1	0

At $L^4_f g$ level, there is a rank deficiency of 2. As in Section 4.5, input excitation of one or more J_{g_k} is required for $K_{BS_{J_g}}$ to be identifiable. The remaining possibly unobservable variables are C^f_k, α_{OS_k} and $K_{F\alpha}$. First-order input excitation of $\dot{M}^f_{F_1}$ results in full rank for these variables. Even then, full rank is only achieved at $L^4_f g$ level and the resulting d$\mathcal{O}$ matrix has a conditioning number of 3×10^{11}, both of which indicate severe challenges to observability. Even if $K_{F\alpha}$ is assumed to remain relatively constant and is adjusted manually based on historic data and/or step tests, the state estimation algorithm will likely struggle to obtain good C^f_k estimates. Alternative models or additional measurements could alleviate this and further investigation is recommended.