Development of a hydrocyclone separation efficiency model using artificial neural networks

S Greyling

21818347

Dissertation submitted in fulfilment of the requirements for the degree *Magister* in Electrical and Electronic Engineering at the Potchefstroom Campus of the North-West University

Supervisor: Prof G van Schoor

Co-supervisor: Prof KR Uren

May 2016

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT

ABSTRACT

A hydrocyclone is an apparatus that is widely used throughout the mineral processing industry. Usually the hydrocyclone is used for the classification, desliming or dewatering of slurries. It is inexpensive, application-efficient and easily employed within different processes.

When classifying slurries, the separation efficiency (or the performance) of the hydrocyclone is described by the cut-size and the sharpness of classification coefficient, collectively referred to as a partition curve. These separation efficiency indicating parameters cannot be measured in real-time and are thus quantified by utilising models. Most of the available models are derived from experimentally obtained data and are therefore empirical in nature. Over the last two decades researchers have started employing alternative techniques in order to develop a separation efficiency model. These include updated empirical models, black-box approaches and Computational Fluid Dynamics (CFD) studies.

The main goal of this study was to develop an Artificial Neural Network (ANN) model that estimates the cut-size and sharpness of classification coefficient by using experimentally attained data. Such a model can be used in predicting the separation efficiency parameters in real-time, a soft sensor, subsequently lending itself to possible control of the hydrocyclone's performance in real-time.

It is important to note that an ANN's usefulness is directly related to the data that are used to train it. It was therefore imperative that high quality data were collected. Using Experimental Design (ED) a structured set of experiments, which included the entire operating range of the hydrocyclone, are described. An experimental procedure was planned and executed in order to obtain the necessary samples in an organised fashion. The experiments were taken on a 100 mm hydrocyclone test rig and the slurries consisted of fine silica with a maximum volumetric solid concentration of 3.125 %. The collected samples were then analysed using the Malvern Particle Size Analyser 2000. Finally the analysed data could be processed accordingly and then used to develop a specified ANN.

In order to determine the best possible ANN, many different variations were trained and then tested using data unknown to the ANN and comparing the obtained estimates to experimental data. Some of the ANN inputs include the pressure, volumetric solid concentration and the spigot opening diameter. To determine whether more inputs to the ANN might deliver better estimations, additional hydrocyclone variables (such as overflow flow rate and angle of discharge) were also used as inputs. The outputs were the separation efficiency indicating parameters. Firstly the cut-size and sharpness of classification coefficient as separate outputs were determined and

secondly the combined outputs thereof. In order to determine whether the ANN application is warranted, the ANN results were compared to a well-known empirical model from literature.

The study is concluded by meticulously reviewing the work that was done and the results that were attained, especially referring to the use of an ANN for estimating a hydrocyclone's separation efficiency compared to existing models from literature. It is evident that the more hydrocyclone variables that are used as ANN inputs, the better the ANN estimations become. Limited literature is available on estimating the sharpness of classification coefficient and this might be because of complex correspondence to the hydrocyclone variables. This study shows that the sharpness of classification coefficient estimations performs poorly, irrespective of the ANN architecture.

Some future work could focus on incorporating instrumentation on the test rig, in order to log certain measurements in real-time. This will also be useful for control purposes when a hydrocyclone model is used along with a control-valve. Another aspect that might be useful to investigate is the real-time processing of the angle of discharge. For this study the angle of discharge photos were only processed after the experiments were concluded. An on-line image processing aspect might be an interesting addition to the on-line measurements.

Keywords: Hydrocyclone, modelling, Artificial Neural Networks, cut-size, sharpness of classification.

ACKNOWLEDGEMENTS

"For You are my hope, O Lord God, You are my confidence from my youth. By You I have been sustained from my birth; my praise is continually of You."

I would like to sincerely thank the following persons and institutions, in no specific order, for the contribution they made to the completion of this dissertation:

- The North-West University Potchefstroom Campus for giving me the opportunity and financial support to enrol for my Master's degree and providing me with a world class education.
- Multotec, our industry partners, for the hydrocyclone test rig they sponsored.
- My supervisor, Prof George van Schoor, for his unstinting support, his knowledge, wisdom, leadership, reassurance and advice.
- My co-supervisor, Prof Kenny Uren, for his knowledgeable inputs, support and kind words throughout.
- Mr. Frikkie van der Merwe for his invaluable suggestions and guidance.
- Werner Greyling, my fiancé, for all his love and support, from assisting with the sampling to reading every draft.
- My father, mother and sister for their love, encouragement and guidance in all I do.
- Ms. Anrika Botha for her friendly assistance in procuring the necessary instruments.
- The Engineering Faculty Workshop for moving and modifying the test rig.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	Analysis of Variance
ANOVA	Artificial Neural Network
CCRD	Centrally Composite Rotatable Design
CFD	Computational Fluid Dynamics
CL	Confidence Level
df	Degrees of Freedom
ED	Experimental Design
MAE	Mean Absolute Error
MSE	Mean Square Error
PSD	Particle Size Distribution
P&ID	Piping and Instrumentation Diagram
RMSE	Root Mean Square Error
SE	Square Error
SS	Sum of Squares

LIST OF SYMBOLS

d_{50}	Cut-size
$C_{w\%}$	Solid concentration by weight
D_c	Hydrocyclone diameter
D_i	Inlet diameter
D_o	Vortex finder opening diameter
D_u	Spigot opening diameter
e	Experimental error
$e_{\%}$	Experimental error in percentage
F_1	Calibration factor of the d_{50}
F_2	Calibration factor of the m
F_4	Calibration factor of the S
h	Free vortex height
k	Number of variables
k_{plitt}	Hydrodynamic exponent (default value 0.5)
L_c	Length of cylindrical section of hydrocyclone
m	Sharpness of classification
M_o	Weight of overflow sample
M_u	Weight of underflow sample
n	Number of samples or observations
n_0	Number of centre samples or observations
P	Pressure
Q_i	Feed flow rate
Q_o	Overflow flow rate
Q_u	Underflow flow rate
r	Correlation coefficient
R^2	Coefficient of determination
$\bar{R}^2$	Adjusted coefficient of determination
s	Standard deviation
S	Volumetric flow split
t	Time
T	Temperature
y_i	Actual value or observation
$\hat{y}_i$	Estimated value or observation

α	Confidence Interval
β	Experimental Design coefficients matrix
γ	Gamma
ε	Error
η	Liquid viscosity
θ	Cone angle
ρ_l	Liquid density
ρ_o	Overflow density
ρ_p	Pulp density
ρ_s	Solid density
ρ_u	Underflow density
ϕ	Volumetric solid concentration
ω	Angle of discharge

CHAPTER 1 – INTRODUCTION

1.1 Background

A hydrocyclone is a stationary conical apparatus that is widely used throughout the mineral processing industry. It is utilised for the classification, desliming or dewatering of slurries. The separation within a hydrocyclone is based on sedimentation, where the swirl-motion is generated when the slurry is pumped through the tangential inlet of the hydrocyclone. Two vortices form within the hydrocyclone where the one vortex, moving downwards, transports the larger and coarser particles to the underflow. The second vortex, moving upwards, transports the finer particles and most of the water to the overflow. Figure 1-1 depicts a graphical representation of the explanation, showing the general shape of the hydrocyclone, important aspects and the relevant vortices. Hydrocyclones are usually inexpensive, adaptable and relatively small to employ [1], [2].

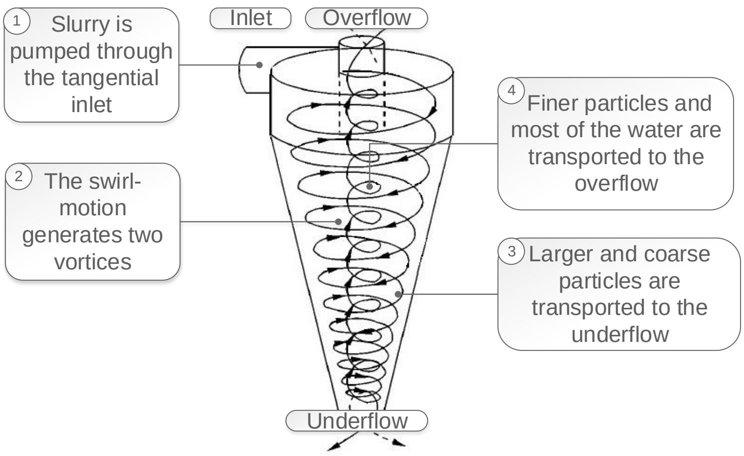


Figure 1-1: The hydrocyclone explained

Ever since the hydrocyclone became popular there have been researchers that worked on developing a model in order to quantify its separation efficiency. Models are incorporated to estimate the cut-size (d_{50}[1]) and sharpness of classification coefficient (m[2]), usually in the form of a partition curve, which is indicative of the hydrocyclone's performance. Figure 1-2 shows a partition curve with the cut-size and sharpness of classification respectively. Ideally a hydrocyclone is operated at a condition where a specific cut-size and sharpness of classification are obtained. The models' application proved useful as the two performance indicating parameters cannot be monitored in real-time. Most of the developed models, are based on experimentally obtained data. When using experimentally obtained data, especially of a system

[1] The cut-size describes the size of the particle that has a 50% probability of reporting either to the underflow or to the overflow.
[2] The sharpness of classification coefficient describes the gradient of the partition curve.

as complex and non-linear as the hydrocyclone; the models that are developed might not always be fully comprehensive. With the recent advances in computational power, hydrocyclone modelling techniques are extended to include Computational Fluid Dynamics (CFD).

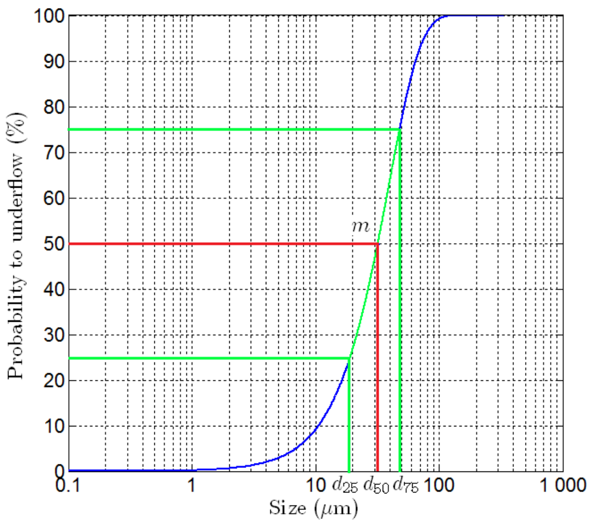


Figure 1-2: A partition curve indicating the cut-size and sharpness of classification

When employing hydrocyclones within the industry, the need exist to optimise the hydrocyclone's performance in terms of the cut-size and sharpness of classification. Fluctuations in the underflow and overflow streams might carry through to the down-stream processes and could potentially decrease the plant's global performance [3]. Seeing as the cut-size and sharpness of classification cannot be measured in real-time, a model that could supply possible estimations might be incorporated into a control scheme such as the one depicted in Figure 1-3. Such a controller would minimise the variations in the cut-size which would also minimise the effects carried through the plant. Figure 1-3 shows that a controller is linked to an error calculator which finds the difference between user-provided optimal[3] cut-size and sharpness of classification values and the current system's estimated cut-size and sharpness of classification. The controller then makes the appropriate changes to the hydrocyclone system in order to minimise the error.

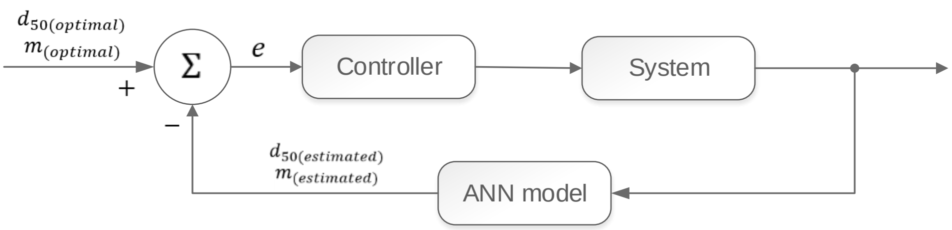


Figure 1-3: Possible use of an ANN model in a control scheme

[3] The user-specified optimal values are expected to always be realistic and viable.

1.2 Problem statement

The aim of this project was to develop a separation efficiency model for the university's hydrocyclone test rig. In other words a model that could estimate the hydrocyclone's cut-size and sharpness of classification coefficient. An Artificial Neural Network (ANN) modelling approach was taken, where the model was based on experimentally obtained data, making it empirical in nature. In order to acquire the experimental data, the test rig had to be instrumented accordingly.

1.3 Issues to be addressed and methodology

To ensure that the project was completed systematically, it was divided into four major phases as shown in Figure 1-4. Each one of these phases will be discussed in detail, endeavouring to describe the issue and the approach that was followed to resolve it.

1.3.1 Conceptual models

The first and most important element was to determine what type of model had to be developed. In other words defining the model's main concepts such as structure, inputs and desired outputs. In effect, specifying the type and scope of data that would be required. By extensively studying the available literature, it was concluded that an Artificial Neural Network (ANN) approach showed the better advantages and the most promising results. The conceptual models' inputs were based on the influential hydrocyclone variables as described in literature, specifically referencing the work done by H. Eren et al. [4], [5].

1.3.2 Experimental data acquisition

1.3.2.1 Experimental setup

The second phase of the project was the experimental data acquisition phase. With the model requirements having been identified, the hydrocyclone system needed to be set-up accordingly. This means that the necessary instruments were installed and the sampling requirements[4] prepared.

1.3.2.2 Design experiments

When working with a model that is based on experimental data only, it is crucial to collect useful measurements and samples as effectively as possible. In order to achieve that, experiments

[4] The requirements include, but are not limited to, the material needed for mixing slurries as well as sampling containers.

were pre-designed using Experimental Design (ED), stipulating the exact conditions, requirements and measurements of the samples that had to be gathered.

1.3.2.3 Experimental procedures

With the experiments defined, a step-wise experimental procedure was needed. By defining a procedure the sampling process remained constant throughout, ensuring that human errors were minimised and that time and the resources were efficiently utilised.

1.3.2.4 Experimental analysis

The samples collected during the experimental runs needed to be analysed; once again a step-wise analysis procedure was required. The analysis of the samples directly affect the outcome of the models and it was therefore important to systematically repeat the analysis process of every sample in the exact same fashion, minimising unforeseen effects such as erroneous analyses or the omission of samples.

1.3.2.5 Experimental processing

In order to render the data in a useable format, some additional processing was essential. Processing procedures were therefore defined in such a manner that it could easily be employed and repeated.

1.3.3 Model implementation

With the data in a useable format, the described conceptual models could be developed. With so many models that needed to be developed, a base ANN was created by utilising the Neural Network Toolbox command-line operations within MATLAB®. The base ANN was a feed-forward backpropagation network employing the Levenberg-Marquardt training algorithm. The various models could therefore be developed with ease by making only minor modifications to the base ANN's properties.

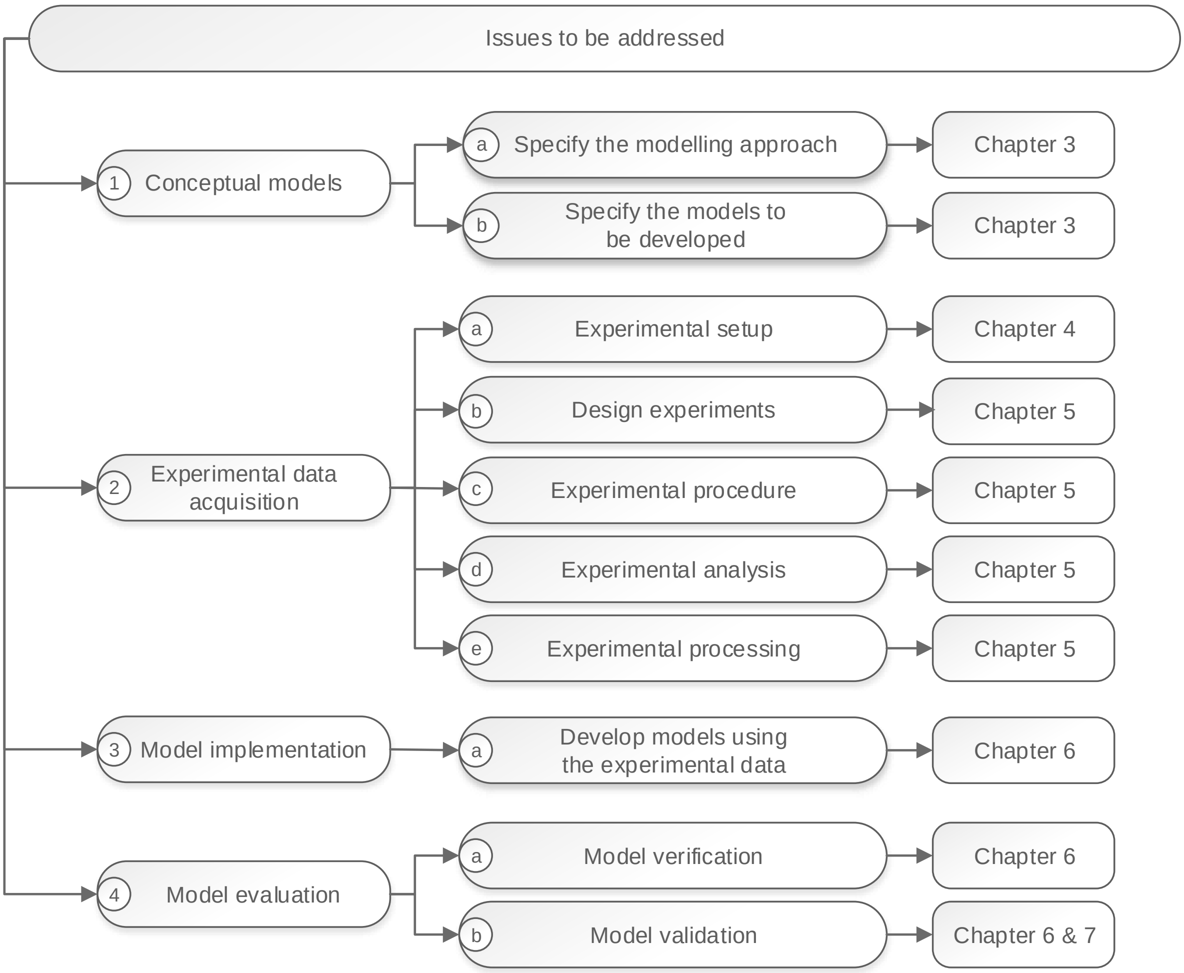


Figure 1-4: Summary of the project phases

1.3.4 Model evaluation

1.3.4.1 Model verification

In order to verify whether the conceptual models were successfully transformed into functional models, a verification flow diagram was created and utilised along with the Mean Square Error (MSE) graphs plotted during the ANN training. A second verification assessment was determining the statistical adequacy of the developed models by evaluating the Analysis of Variance (ANOVA) results of each.

1.3.4.2 Model validation

The model validation included four main measures, which assessed the accuracy of adequate models when they were utilised for their intended purpose. The first measure employed was evaluating the regression plots obtained when plotting the actual data against the predicted data. Next the models were visually assessed, by comparing the per sample plots of each. The third method of validation was the calculation of standard error metrics, where the errors were expected to be as small as possible. Lastly the best performing ANN model was compared to the popular, conventional model as developed by Flintoff et al.[5]; inspecting whether the developed model delivered comparable predictions to that of the conventional model.

1.4 Conference contributions

The following conference contributions emanated from the research, the full articles can be found in Appendix D.

1.4.1 Presented

S. van Loggenberg, G. van Schoor, K.R. Uren, and A.F. van der Merwe, “Hydrocyclone separation efficiency estimation using artificial neural networks”, SAUPEC 2015, Johannesburg, South Africa, 29 January 2015.

1.4.2 Under review

S. van Loggenberg, G. van Schoor, K.R. Uren, and A.F. van der Merwe, “Hydrocyclone cut-size estimation using artificial neural networks”, DYCOPS 2016, Trondheim, Norway, June 2016.

1.5 Dissertation overview

The dissertation will be discussed in the same order as described in subsection 1.3. Chapter 2 is used to describe the literature that was found to be vital to this project’s progress. The topics discussed include some basic hydrocyclone concepts, the relevant ANN aspects, introductory Experimental Design (ED) definitions and rudimentary statistical methods that are used in determining the adequacy and accuracy of the developed models. A brief overview of the Plitt-Flintoff mathematical model will also be discussed, detailing some of the most important aspects thereof. Chapter 3 will describe the modelling approach, specifically depicting the structure, input-output concepts and the scope of the expected data; i.e. detailing the conceptual models. The system realisation is communicated in Chapter 4, considering specifically the physical system and the instruments that were employed. The experimental, analysis and processing procedures used throughout, are also summarised. Chapter 5 details the Experimental Design (ED), describing the technique and showing how it was incorporated in designing the experiments. Chapter 6 relays the developed models’ particulars, adequacy analyses and the accuracy results. Chapter 7 compares the best performing model with the conventional mathematical model as developed by Flintoff et al. To conclude the dissertation, a final conclusion is given in Chapter 8, recapping some of the most important aspects and discussing the findings of the project.

[5] From here on in referred to as the Plitt-Flintoff model.

CHAPTER 2 – LITERATURE STUDY

2.1 Literature background

2.1.1 Chapter introduction

This chapter considers and examines the relevant literature, starting off with the most important hydrocyclone aspects and descriptions. Specific detail is given on the performance of hydrocyclones. An interesting research timeline is given in section 2.1.2.4. The next subsection gives an overview of what Experimental Design (ED) entails. From there on the most important statistical analyses that are relevant to this study are shown and explained. With the modelling approach being Artificial Neural Networks (ANN), some background is given on its history and applications. The chapter is concluded by critically reviewing an article that was found to reflect most of the objectives of this project.

2.1.2 Hydrocyclone overview

2.1.2.1 Hydrocyclone description

Figure 2-1 shows that a hydrocyclone consists of a conical container, with a spigot[6] that is connected to a cylindrical section that has a tangential inlet. The top part of the hydrocyclone is closed with a plate where an axially mounted overflow pipe passes through it. The overflow pipe extends into the hydrocyclone. This extension is called the vortex finder. The vortex finder prevents feed from short-circuiting directly to the overflow [6].

The variables that are associated with hydrocyclone performance are usually divided into two groups namely, design variables and operating variables. The design variables are dependent of the hydrocyclone size and proportions whereas the operating variables are independent of the size and proportions. It should be mentioned that these variables cannot be considered separately because of interactions with one another. Table 2-1 gives a summary of the related variables [1], [2], [5], [6].

Table 2-1: Hydrocyclone variables

Design variables		Operating variables	
Hydrocyclone diameter	D_c	Hydrocyclone throughput	Q
Feed inlet diameter	D_i	Feed pressure	P
Vortex finder diameter	D_o	Volumetric solid concentration	ϕ
Spigot opening diameter	D_u	Solid density	ρ_s
Free vortex height	h	Angle of discharge	ω
Cone angle	θ		

[6] Also called an apex.

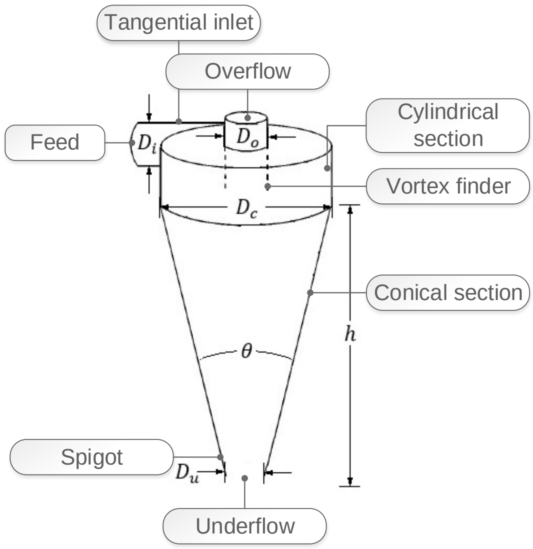


Figure 2-1: The hydrocyclone (adapted) [1], [5], [6]

2.1.2.2 General fluid flow

The most important flow pattern that is found in a hydrocyclone is the two vortices, the flow that reports to the underflow (primary vortex) and the flow that reports to the overflow (secondary vortex), as depicted in Figure 2-2. As mentioned before the vortices[7] are generated by the feed being fed through the tangential inlet. The primary vortex carries the coarse and larger particles to the underflow and the secondary vortex transport the fine particles and most of the water to the overflow [7].

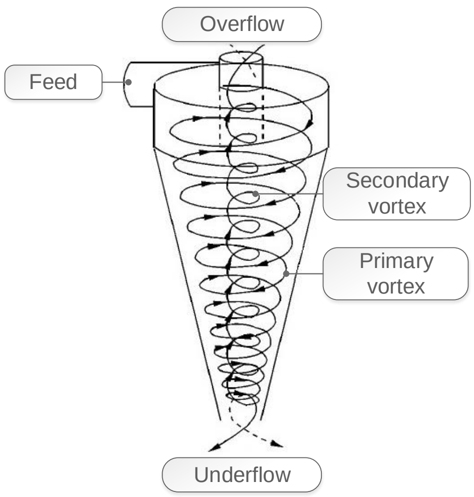


Figure 2-2: The two main vortices found within the hydrocyclone [2]

[7] Note that the two vortices revolve in the same direction.

Two additional flows exist within the hydrocyclone as shown in Figure 2-3. These are the short-circuit flow and the Eddy flow. The short-circuit flow is due to the hindrance created by the tangential velocity. In order to minimise the short-circuit flow, the vortex finder is incorporated. The Eddy flow occurs when the overflow opening cannot accommodate the secondary vortex [7].

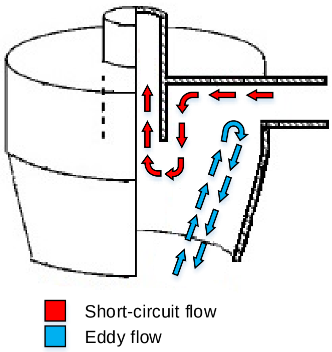


Figure 2-3: The two additional flow patterns found within the hydrocyclone

2.1.2.3 Hydrocyclone performance

When referring to the hydrocyclone's performance[8] there exist five major quantitative parameters which can be evaluated, they are said to be [1], [3], [5], [6], [8].

- Partition curves
- Cut-size (d_{50})
- Sharpness of classification coefficient (m)
- Pressure through-put relationship
- Split of water flow to products

A hydrocyclone's performance is usually described by a partition curve[9]. The partition curve is a graphical and quantitative representation of a hydrocyclone's particle size separation performance. It usually describes the weight fraction (or percentage) of each particle size in the feed which reports to underflow, shown on the y-axis, to the particle size, shown on the x-axis [1], [3], [5], [6], [8]. It is assumed that a fraction of the fine particles completely bypasses the hydrocyclone's classification process. This is called the bypass and it explains why the partition curve does not have an asymptote at zero. It is generally assumed that the bypass is equal to the fraction of water that reports to the underflow [8]. Thus the two types of partition curves that are generally discussed are the gross[10] partition curve, which does not take into account the water recovery, and the reduced[11] partition curve, which is adjusted to include the water recovery effect.

[8] Performance refers to the hydrocyclone's separation or classification efficiency.
[9] Also known as tromp curve, performance curve or efficiency curve.
[10] Also called uncorrected partition curve.
[11] Can also be referred to as corrected partition curve.

For this study, only the reduced efficiency curve is of interest. Figure 2-4 (a) illustrates the two types of partition curves as found throughout the literature [2].

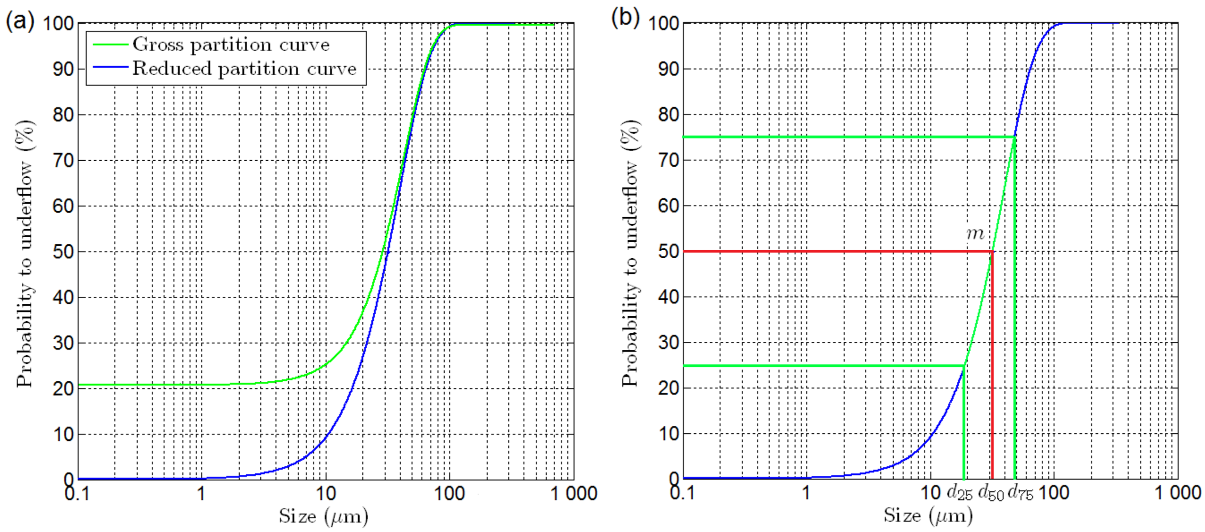


Figure 2-4: (a) The two types of partition curves seen in literature and (b) the partition curve explained

When studying Figure 2-4 (b) the cut-size (red line), indicated as d_{50}, is the size of the particle in the feed Particle Size Distribution (PSD) that has a 50% probability to either report to the underflow or the overflow [1]. The sharpness of classification (m) is a parameter that is used to quantitatively describe the hydrocyclone's classification by supplying a measure for the gradient of the partition curve (shown in green). The sharpness of classification is taken to be the gradient between d_{25} and d_{75} where an $m < 2$ signifies poor classification and $m > 3$ implies good sharpness of classification [5]. It is important to note that certain hydrocyclone variables will have a specific influence on the hydrocyclone's performance. Table 2-2 shows a summary of the expected effects on performance some variables will have [2].

Table 2-2: Design and operating variables effects on the hydrocyclone's performance [2]

Hydrocyclone variable		Cut-size		Sharpness of classification		Throughput	
Increase the pressure	P	Decrease	▼	ND[a]	-	Increase	▲
Increase the volumetric solid concentration	ϕ	Increase	▲	Decrease	▼	Increase	▲
Increase the spigot opening diameter	D_u	Decrease	▼	Decrease	▼	Increase	▲
Increase the hydrocyclone diameter	D_c	Increase	▲	Increase	▲	Increase	▲
Increase the feed inlet	D_i	Decrease	▼	Decrease	▼	Increase	▲
Increase the vortex finder diameter	D_o	Increase	▲	Increase	▲	Increase	▲
Increase the free vortex height	h	Increase	▲	Increase	▲	Increase	▲
Increase the cone angle	θ	Increase	▲	Increase	▲	ND[a]	-

[a] Not definable

2.1.2.4 Hydrocyclone research and modelling

Since as early as 1939 the hydrocyclone's versatility and success, within the mineral processing industry, created research opportunities for many. For the first two decades researchers worked on describing the generalities of the cyclone, focussing specifically on defining its operations. The most prominent contribution was that of Kelsall who studied and published the fluid flow patterns of the hydrocyclone in 1952 [5], [9]. In 1965 Bradley compiled a book that described the known fundamentals and theoretical equations of that time. It should be noted that these equations were not relevant to industrial hydrocyclones [1]. Literature suggest that the first comprehensive model applicable to industrial hydrocyclones was developed by Lynch & Rao in 1966 [5]. The approach was quickly adopted within the industry, leading to Lynch & Rao developing an up-scaling model in 1975. In 1976, Plitt developed a mathematical model by incorporating Lynch & Rao's database along with his own experiments. Plitt's work is one of the most referenced articles in hydrocyclone modelling. In 1978 Nageswararao developed the second general-purpose model, a slightly more complex version than that of Plitt [10]. In order to incorporate some of the findings of 1976 – 1987, Flintoff et al. revised the model Plitt developed. The most noticeable improvement being the addition of calibration factors. By 1996 H. Eren et al. were some of the first researchers that employed Artificial Neural Networks (ANN) to predict the hydrocyclone's Particle Size Distribution (PSD) and its cut-size under different operational conditions [4], [11]. Over the last decade interesting advances were made with the incorporation of Computational Fluid Dynamics (CFD) [10].

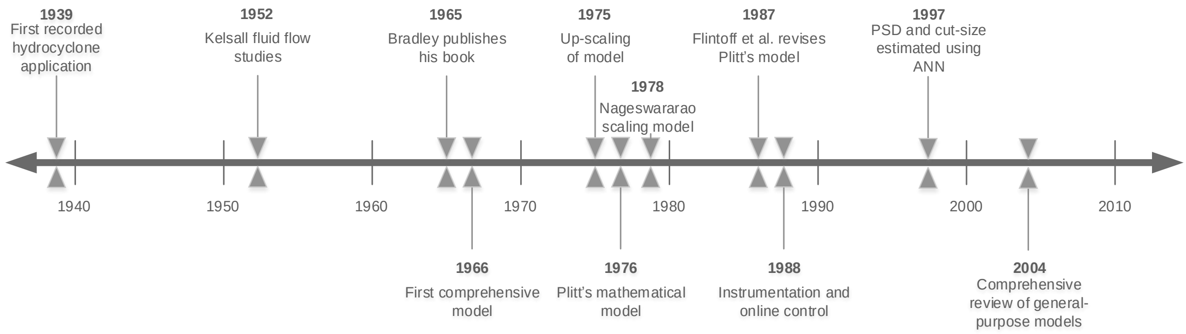


Figure 2-5: Timeline of the most important hydrocyclone research contributions

2.1.3 Experimental Design

Experimental Design (ED) can broadly be defined as the process of designing organised experimentation. In other words it is a technique that specifies how experiments should be conducted and how an influencing variable should be varied in order to obtain distinct and useful response results with as little experimental effort as possible [12], [13]. It is important to note that any conclusions drawn from an experiment will significantly depend on how that experiment was conducted. Usually ED is utilised to either understand more of the process variables (modelling) or to find combinations of the variables that delivers optimum responses (optimisation) [12].

A response variable is defined as a quantifiable characteristic of a process whether it is a product or only an aspect thereof. The variables are those independent factors that affect the process' overall response.

ED techniques have shown that the change-one-variable-at-a-time approach might not always be efficient or thorough enough. Firstly the change-one-factor-at-a-time approach requires more experimental runs than the ED techniques do. It also does not distinguish or convey the effects or interactions that two or more variables will have on the response. Lastly the change-one-factor-at-a-time approach cannot identify the specific levels of each variable that will optimise the response variable. Thus it becomes clear that the change-one-factor-at-a-time approach excludes important aspects of experiments, the variables and the responses that are being evaluated [13].

2.1.4 Useful statistics

In order to evaluate the adequacy, accuracy and validity of models, some basic statistical analyses are needed. The subsections will discuss the estimations of the experimental error, the employment of Analysis of Variance (ANOVA), correlations and regression and the calculation of two error metrics.

2.1.4.1 Experimental error

When measuring a physical feature, the measurement can never be error-free. This is because when a measurement is repeated, small variations occur within the measured quantity. These variations might be systematic or random of nature. Systematic errors are typically caused by definite erroneous elements, such as faulty calibration of instruments or incorrect measuring by the operator. Random errors are not as easily defined, but are said to be caused by unpredictable fluctuations or changes [13]. In order to determine an acceptable interval of variation an experimental error is calculated for the measured parameter.

Table 2-3 summarises the parameters and the relevant equations needed to calculate an expected experimental error. Start off by choosing a Confidence Level (CL). The CL indicates the level of certainty that is expected, directly defining an upper and lower bound. For instance a CL of 95% is chosen; it would signify that there is a 95% certainty that the measured or estimated value will lie within the lower and upper bounds, where about 5% fall beyond these bounds. Next, n repeated samples or measurements are required. By employing the rudimentary equations, the interval is eventually calculated. To determine the error percentage, the interval is converted to a percentage.

Table 2-3: Calculating experimental error

Parameter		
Confidence level	CL	-
Number of samples	n	-
Degrees of freedom	df	$n-1$
Mean	$\bar{y}$	$\frac{1}{n}\sum_{i=1}^{n} y_i$
Standard deviation	s	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$
t-value	t_{n-1}	t critical value table
Error	e	$t_{n-1}\left(\frac{s}{\sqrt{n}}\right)$
Interval	-	$[\bar{y}-e;\bar{y}+e]$
Error in percentage	$e_{\%}$	Convert interval to %

2.1.4.2 ANOVA - checking model adequacy

In order to check a developed model's adequacy, an Analysis of Variance (ANOVA) is done. For this study two types of models were developed: Experimental Design (ED) models and Artificial Neural Network (ANN) models. The ANOVA approach for each of these are quite similar, but include some differences when referring to the sources being investigated. Starting with the ED models' ANOVA, the sum of squares are firstly determined by using the formulae in Table 2-4. Here the y_i represents the actual data, $\hat{y}_i$ the estimated values, $\bar{y}$ the sample mean, y_{oj} the actual values of the central points and $\hat{y}_{0j}$ the estimated values of the central points. Next the corresponding degrees of freedom are determined, where the n is the number of samples and n_0 the number of central samples[12].

Table 2-4: Sum of square formulae

Description		Formula
Sum of squares due to regression	SS_X	$\sum(\hat{y}_i-\bar{y})^2$
Residual sum of squares	SS_R	$\sum(y_i-\hat{y}_i)^2$
Sum of squares relating the lack of fit	SS_L	SS_R-SS_E
Sum of squares due to pure error	SS_E	$\sum\left(y_{0j}-\hat{y}_{0j}\right)^2$
Total sum of squares	SS_T	SS_X-SS_R

[12] The central samples are those samples coded at 0.

With the values of Table 2-4 known, Table 2-5 can easily be completed. The calculated F-value of the regression and lack of fit is then compared to a critical F-value that is found within the standard critical F-value table contained in most statistics textbooks. The calculated and the critical F-value for both the regression and lack of fit are determined. Table 2-6 depicts the model's adequacy based on the comparison results.

Table 2-5: Experimental Design ANOVA

Source	Subscript	SS	df	MS	F
Regression	X	SS_X	n_0	$MS_X = \frac{SS_X}{\mathrm{df_X}}$	$\frac{MS_X}{MS_R}$
Residual	R	SS_R	$n - n_0 - 1$	$MS_R = \frac{SS_R}{\mathrm{df_R}}$	
- Lack of fit	L	SS_L	$\mathrm{df_R} - \mathrm{df_E}$	$MS_L = \frac{SS_L}{\mathrm{df_L}}$	$\frac{MS_L}{MS_E}$
- Error	E	SS_E	$n_0 - 1$	$MS_E = \frac{SS_E}{\mathrm{df_E}}$	
Total	T	SS_T	$n - 1$		

Table 2-6: Adequacy testing

Source	test F	Significance	Adequacy
Regression	$F_\alpha(\mathrm{df_X}, \mathrm{df_R})$ = critical F-value<calculated F-value	Significant @ α level	Model is adequate
Lack of fit	$F_\alpha(\mathrm{df_L}, \mathrm{df_E})$ = critical F-value>calculated F-value	Non-significant @ α level	
Regression	$F_\alpha(\mathrm{df_X}, \mathrm{df_R})$ = critical F-value>calculated F-value	Non-significant @ α level	Model is inadequate
Lack of fit	$F_\alpha(\mathrm{df_L}, \mathrm{df_E})$ = critical F-value>calculated F-value	Significant @ α level	

The same approach is followed with the Artificial Neural Network models. The ANOVA table only differs slightly. Table 2-7 shows how to complete the ANN ANOVA by utilising the same Sum of Square formulae as used with the ED ANOVA. Only now the $\hat{y}_i$ represents the ANN estimated values. For the degrees of freedom, k is the number of inputs incorporated into the ANN and n the number of samples evaluated. Determining the model's adequacy is much simpler as summarised in Table 2-8.

Table 2-7: Artificial Neural Network ANOVA

Source	Subscript	SS	df	MS	F
Model	X	$SS_T - SS_R$	k	$MS_X = \frac{SS_X}{\mathrm{df_X}}$	$\frac{MS_X}{MS_R}$
Error	R	SS_R	$n - k - 1$	$MS_R = \frac{SS_R}{\mathrm{df_R}}$	
Total	T	SS_T	$n - 1$		

Table 2-8: Adequacy testing

Source	test F	Significance
Model	$F_{\alpha}(\mathrm{df_X}, \mathrm{df_R})$ = critical F-value<calculated F-value	Significant @ α level
Model	$F_{\alpha}(\mathrm{df_X}, \mathrm{df_R})$ = critical F-value>calculated F-value	Non-significant @ α level

2.1.4.3 Correlation and regression

In order to determine how well the predicted values correlate to the actual values there exist three relevant coefficients that can be calculated. These three coefficients are the linear correlation coefficient (r), the coefficient of determination (R^2) and the adjusted coefficient of determination ($\bar{R}^2$). The correlation coefficient indicates the strength of the linear relationship between the actual and predicted values and can take on any value between -1 and 1. The r-value is calculated by utilising (2-1) where n is the number of samples, y_i is the actual values and $\hat{y}_i$ is the predicted values. Table 2-9 shows a summary of the r-value and its relevant meaning [13]. Generally it is said that a correlation greater than 0.8 describes as strong linear relationship while a correlation of less than 0.5 indicates a weak linear relationship.

$$r = \frac{n\sum y_i\hat{y}_i - (\sum y_i)(\sum \hat{y}_i)}{\sqrt{n(\sum y_i{}^2) - (\sum y_i)^2}\sqrt{n\left(\sum \hat{y}_i{}^2\right) - (\sum \hat{y}_i)^2}} \tag{2-1}$$

Table 2-9: Summary of r-value descriptions

r-value	Correlation	Meaning
$0 < r < 1$	Positive correlation	The larger the r-value the stronger the positive linear fit. If the y_i-value increases so does the $\hat{y}_i$-value.
$-1 < r < 0$	Negative correlation	The smaller the r-value the stronger the negative linear fit. If the y_i-value increases the $\hat{y}_i$-value decreases.
$r = 0$	No correlation	If no relationship is found i.e. weak correlation, the r-value is close to 0. There is thus a random relationship between the y_i-value and the $\hat{y}_i$-value.
$r = \pm 1$	Perfect correlation	Indicates a perfect linear relationship. All of the data points lie on the line, indicating that the $\hat{y}_i$-value = y_i-value.

The coefficient of determination (R^2) can be used to determine how well a model is expected to yield predictions. By using the scatter plot[13] like the one in Figure 2-6, with the actual values on the x-axis and the predicted values of the y-axis (points are indicated as blue markers), the R^2-value can easily be calculated. The easiest way to obtain the R^2-value is to utilise a software package like MATLAB® or Excel.

[13] The scatter plot is not required in order to calculate the coefficients, it is merely for illustrational purposes.

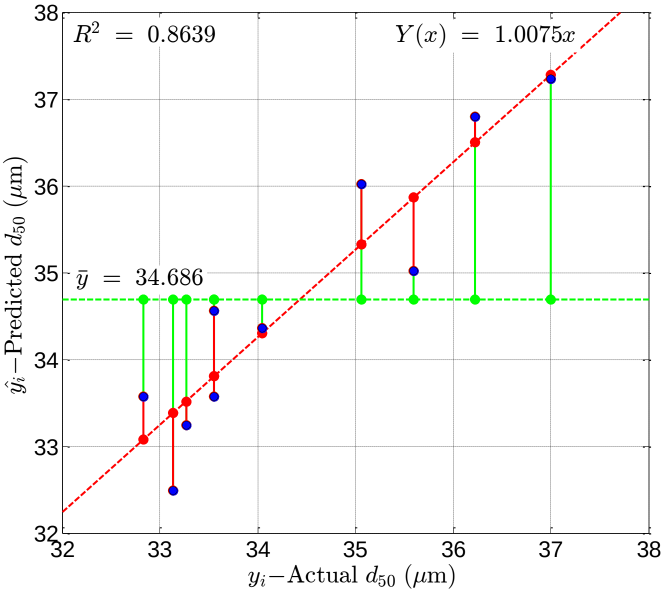


Figure 2-6: Scatter plot depicting how the R^2–value is determined

The coefficient of determination can also be calculated by hand. The diagonal red line is called the best fit linear regression line and it is positioned in such a way that the squared distance between the data points and the line is minimised. Its equation is in the $Y(x) = ax$ form, meaning the intercept is zero (0). The green horizontal line represents the mean of the samples, the mean being 34.686 for this example. The first value needed is the Sum of Squares of the regression (SS_R) which is obtained by finding the sum of the squared differences between the predicted values ($\hat{y}_i$) and the regression line values (Y_i) (see the vertical red lines on the scatter plot). Next the total Sum of Squares (SS_T) is determined by finding the sum of the squared differences between the predicted values ($\hat{y}_i$) and the mean ($\bar{y}$) (indicated as the vertical green lines). With the SS_R and the SS_T calculated, the R^2-value can be computed by utilising (2-2).

$$R^2 = 1 - \frac{SS_R}{SS_T} \tag{2-2}$$

It should be noted that the R^2-value usually increases when models incorporate more variables. Therefore one cannot directly compare models that differ by the number of their inputs. This is where the adjusted coefficient of determination ($\bar{R}^2$) comes to the aid. When looking at (2-3) it is seen that $\bar{R}^2$-value adjusts for the number of variables the model includes.

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1 - R^2) \tag{2-3}$$

Table 2-10 shows a summary of the example's scatter plot calculations, data point values r, R^2 and $\bar{R}^2$ results.

Table 2-10: Example actual and predicted data

i	Actual (y_i)	Predicted ($\hat{y}_i$)	$Y(x) = 1.0075x$	$(Y_i - \hat{y}_i)^2$	$(\hat{y}_i - \bar{y}_i)^2$	Calculated		
1	33.556	34.563	33.808	0.571	0.015			
2	32.831	33.574	33.077	0.247	1.236	n	k	$\bar{y}$
3	35.065	36.015	35.328	0.472	1.765	10	3	34.686
4	35.600	35.021	35.867	0.716	0.112			
5	33.268	33.237	33.518	0.079	2.102	SS_R		SS_T
6	34.047	34.358	34.302	0.003	0.108	3.044		22.364
7	37.000	37.233	37.278	0.002	6.486			
8	33.138	32.487	33.387	0.809	4.835	r	R^2	$\bar{R}^2$
9	33.553	33.574	33.805	0.053	1.238	0.9295	0.8639	0.8250
10	36.225	36.800	36.497	0.092	4.468			

2.1.4.4 Error metrics

In order to determine the performance of developed models two popular error metrics can be used. These two error metrics are the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE). When evaluating (2-4) it is evident that large deviations in the actual (y_i) and estimated ($\hat{y}_i$) values will result in a large error weight, making RMSE beneficial in penalising unwanted large deviations.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2-4}$$

The MAE in (2-5) scores the errors linearly. It should be noted that error metrics condense a set of errors into a single measure and can therefore only supply one type description of the model's error characteristics [13], [14]. Therefore, should the study require so, additional error analyses could be included to evaluate other aspects.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2-5}$$

2.1.5 Artificial Neural Networks overview

An Artificial Neural Network (ANN) is a system of simple processing units that are connected into a structured network by a set of weights [15]. The processing units, normally called neurons, are essentially the building blocks of ANNs [16], [17]. ANNs work especially well when employed for complex, non-linear systems. When working with ANNs it becomes clear that there are various aspects to its structure and processing capabilities. The following subsections will endeavour to discuss some of the important aspects.

2.1.5.1 Neurons

When studying Figure 2-7 it is seen that a neuron takes the sum of a bias[14] value b and a weight-multiplied input wp to deliver a resulting net input function n_{net}. The bias b is much like a weight except that it always has a constant input of 1. The net input function n is then used as an input to a specific activation[15] function f. Biases are beneficial in preventing a net input of zero and indirectly present an additional variable to the ANN. The weights w and b are adjustable parameters, thus it is said that the main concept of an ANN is to update and tune these parameters in such a way so as to obtain a desired output. The tweaking of these weight parameters are achieved through a process called training [17].

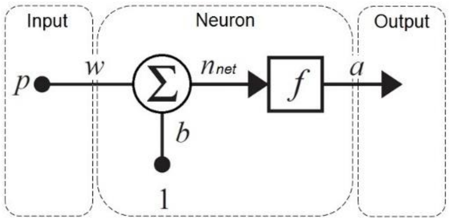


Figure 2-7: Single-input neuron (adapted) [18]

A neuron however is not restricted to only a single input, but can have multiple inputs as depicted in Figure 2-8.

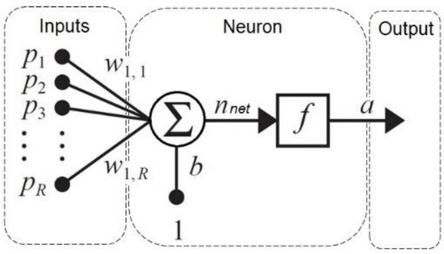


Figure 2-8: Multiple-input neuron (adapted) [18]

2.1.5.2 Layers

Most problems being investigated might need more than one multiple-input neuron. A layer is considered to be a collection of neurons all working in parallel. Thus the layer will comprise of all the weights, biases and activation functions of the included neurons. When developing ANNs one is not limited to only one layer of neurons, but can incorporate multiple layers as shown in the figure. Each element of input vector $\boldsymbol{p}$ is connected to each neuron via a weight matrix $\boldsymbol{W}$ and

---

[14] Also known as an offset.
[15] Some authors refers to the activation function as a transfer function.

every neuron has a bias b_i. Again it is seen that the net input functions n_{neti} is the sum of weighted inputs and a bias [18]. It is important to mention that the number of neurons need not equal the number of inputs.

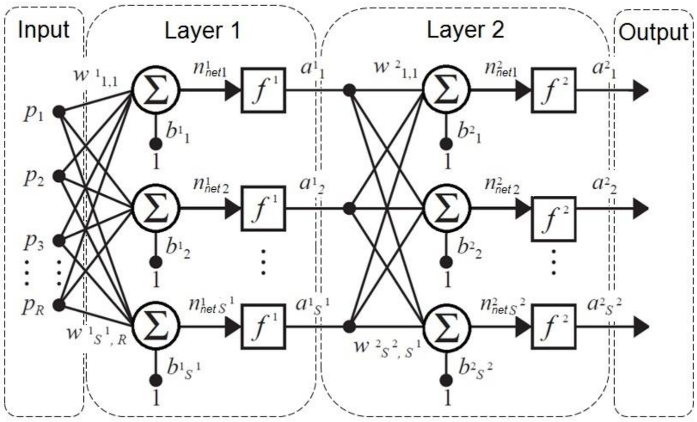


Figure 2-9: Multilayer Artificial Neural Network consisting of two layers (adapted) [18]

2.1.5.3 Architecture

An ANN's architecture is characterised by the types of neurons used and by their connections within the ANN. The two main architectures of ANNs are the Feed-forward ANN (FFANN) and the Recurrent ANN. The Feed-forward ANN's neurons receive only inputs from the preceding layer's neurons and presents the outputs only to the next layer's neurons [15], [16]. Thus the FFANN represents a function of its current inputs only. Recurrent ANNs have layers of neurons that might connect to neurons within the same layer or to any other layers' neurons [15], [17]. It is imperative to note that the architecture will mainly be determined by the nature of the investigation at hand [15], [16], [18].

2.1.5.4 Inputs

When looking at ANN inputs they can either be concurrent or sequential. Concurrent inputs are inputs that all take place on the same time or do not occur in an exact time sequence. Sequential inputs occur chronologically in time [17].

2.1.5.5 Scaling

Usually in practice the inputs of an ANN is transformed by a processing function to ensure that the input data is in a form which the ANN could manipulate and incorporate more efficiently. One of the most popular processing functions being used scales the input data into the interval of $[-1,1]$. The use of processing functions is not limited to only the ANN inputs. Targets provided by the user are also transformed for the same purposes [17].

2.1.5.6 Training

Training is the process of developing an ANN by adapting certain aspects thereof in such a way that the output obtained is as close as possible to the desired target. The modification of the ANN is achieved by employing mathematical algorithms. These algorithms might be either supervised or unsupervised. Supervised algorithms make use of known input-target pairs. The training process of supervised algorithms is thus governed by an external process which is able to determine whether the obtained outputs are suitable and able to calculate the error thereof. Supervised training is generally used when the investigation requires accurately defined input-output relationships. Unsupervised algorithms use only known inputs and has no method of knowing what the outputs might be. An ANN employing unsupervised algorithms is said to develop as it extends its understanding from previous inputs. Unsupervised training algorithms usually work better for pattern recognition problems [15]. Another important point to note is that there exist two training approaches. Incremental training which tunes the weights each time an input is given to the ANN or batch training where the weights are only tuned after all the inputs were given to the ANN [18].

2.1.5.7 Data division

In order to develop a supervised ANN the user-provided data sets of known inputs and targets are normally divided into three subsets. These three subsets are called the training, validation and testing data sets. The training data set is used to initially train the ANN by calculating the gradient and tuning the weights and biases appropriately. The validation data set is checked throughout the training process, specifically evaluating the error thereof. The ANN's weights and biases are adjusted up to the point where the validation error reaches a minimum. The testing data set is not used during the training process, but only afterwards in order to test and compare the developed ANN model.

There are four major data division techniques, each technique advantageous with different applications. The four are: Random division, block division, interleaved division and indexed division.

2.1.5.8 Activation functions

As mentioned previously the activation functions take the net input function n_{net} as an input. A activation function might be either linear or non-linear, its type mainly determined by the kind of problem that is investigated [17]

2.1.5.8.1 Linear activation function

The figure depicts a linear activation function clearly indicating that the output would be equal to the input.

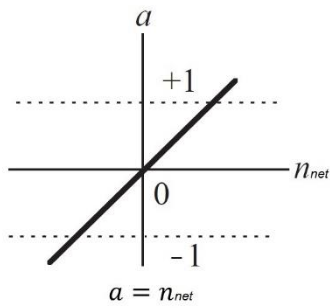


Figure 2-10: Linear activation function (adapted) [18]

2.1.5.8.2 Hard limit activation function

A hard limit activation function will take an input smaller than 0 and set the output to 0. Should the input be equal to or larger than 0 the output is set to 1. This function works especially well for problems which categorises the inputs into one of two distinct classes.

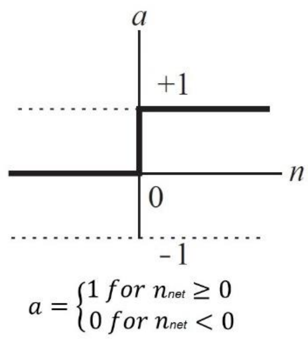
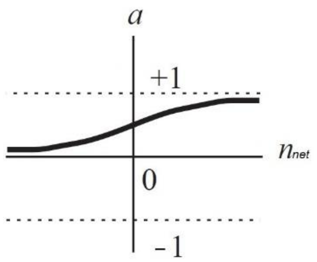


Figure 2-11: Hard limit activation function (adapted) [18]

2.1.5.8.3 Log-sigmoid activation function

The log-sigmoid activation function takes the input and transforms it into a value in the range $[0,1]$. It is usually used with ANNs employing the backpropagation algorithm.

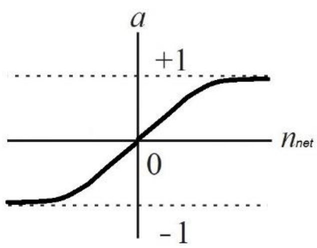


$$a = \frac{1}{1 + e^{-n_{net}}}$$

Figure 2-12: Log sigmoid activation function (adapted) [18]

2.1.5.8.4 Hyperbolic tangent sigmoid activation function

The hyperbolic activation function is shown in Figure 2-13. It is evident that by utilising this activation function an output would be given in the interval of $[-1,1]$.



$$a = \frac{e^{n_{net}} - e^{-n_{net}}}{e^{n_{net}} + e^{-n_{net}}}$$

Figure 2-13: Hyperbolic activation function (adapted) [18]

2.1.5.9 Performance evaluation

In order to evaluate the performance of the developed ANN it is necessary to make use of a performance method or function. The most widely used performance methods employed are the Mean Squared Error (MSE) and the Squared Error (SE). These methods are used to find the errors between the ANN outputs a_i and the expected targets t_i. The MSE of an ANN is consequently defined in (2-6)**.**

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(t_i - a_i)^2 \qquad \textbf{(2-6)}$$

When employing the Neural Network toolbox within MATLAB®, it performs the chosen performance calculations and graphing automatically, producing a graph as shown in Figure 2-14. During training the MSE decreases for the three data sets as the epochs proceed. Training is stopped when the green validation MSE stops decreasing. This is a method employed to ensure that the ANN does not over-train. The red line indicates how well the network will generalise for samples it had never seen [17].

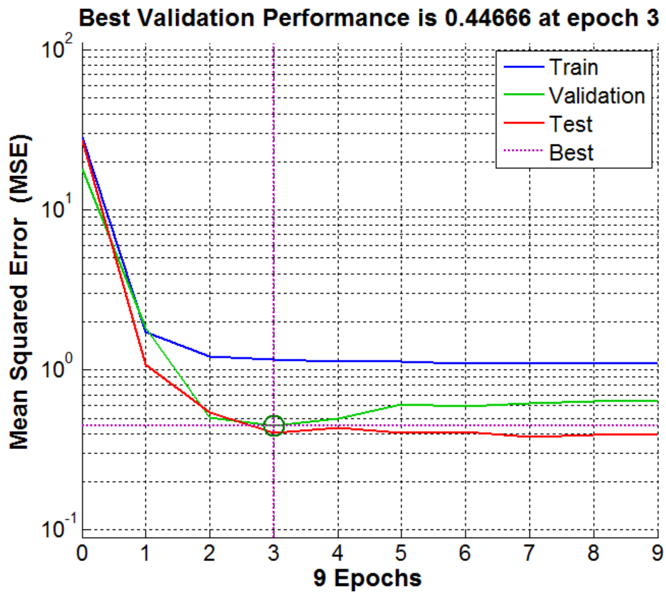


Figure 2-14: MSE performance plots

2.1.5.10 Layer assignment

Seemingly the best architecture to utilise is one sigmoid hidden layer and one linear output layer. The sigmoid layer will ensure that the non-linear relationship is learned and the linear layer is usually used along with function fitting or non-linear regression problems [17].

2.2 Critical literature review

As seen from section 2.1.2.4 many researchers have worked on and contributed to the development of hydrocyclone performance models. The articles that was found to closely correspond to this study, were: (1) the paper done by H. Eren et al. in 1997, titled: Artificial Neural Networks in Estimation of Hydrocyclone Parameter d_{50c} with Unusual Input Variables [4] and (2) Prediction of hydrocyclone performance using artificial neural networks done by M. Kamiri et al. in 2010 [19].

2.2.1 Artificial Neural Networks in Estimation of Hydrocyclone Parameter d_{50c} with Unusual Input Variables

As seen from section 2.1.2.4 many researchers have worked on and contributed to the development of hydrocyclone performance models. The article that was found to closely correspond to this study, was the paper done by H. Eren et al. in 1997, titled: Artificial Neural Networks in Estimation of Hydrocyclone Parameter d_{50c} with Unusual Input Variables [4]. This subsection is dedicated to critically reviewing the paper by firstly identifying the aim of the study, then examining the results shown and lastly by identifying possible inconsistencies.

The goal of the paper was to improve previous results[16] the authors had obtained when employing ANNs and when incorporating general[17] hydrocyclone variables [20]. In order to make improvements to the d_{50} estimations, the authors included unusual[18] variables with ANNs approach. Table 2-11 shows the results mentioned in the paper.

Table 2-11: ANN performance results obtained by H. Eren et al.

Description	r-value	R^2-value
Using general hydrocyclone variables	0.963	0.9666
Plitt model they compared the ANN with	0.895	0.8014
Using general hydrocyclone variables with 3 additional unusual variables	0.995	0.9890
Using 14 hydrocyclone variables (general and unusual)	0.995	0.9897

After evaluating the paper and its findings, some contentions arose:

1. The Plitt model with which the authors compared their ANN led to two issues.
   a. The outdated 1976 version of Plitt's model was used, which do not include any calibration factors.
   b. A contradictory statement was made in chapter I paragraph five whereby the authors mention that the same data that were applied to the authors' ANNs were applied to the Plitt model. Yet in the next sentence, they express that they could expect better Plitt results if they used their test rig's data.
2. The authors further stated that by incorporating 14 hydrocyclone variables the ANN results were further improved from that which they found from employing only three additional unusual variables. However when comparing the r-values and the R^2-values, no actual improvement was found (r-values $= 0.995$ and $R^2 \approx 0.9890$).
3. The authors seemingly found that the increasing R^2-values were of some importance. Usually the R^2-value will increase when more variables are added as mentioned in section 2.1.4.3.

Specific aspects that should be investigated that surfaced during the literature review.

1. Ensure that the Plitt-Flintoff model that includes the calibration factors is utilised.

[16] The focus was the estimations of the cut-size (d_{50}), no references were made to the sharpness of classification coefficient (m).
[17] These include the hydrocyclone variables that were generally used within literature.
[18] Some unusual variables would be the underflow/overflow flow rates and densities.

2. Determine whether the inclusion of unusual variables delivers better estimations or whether it eventually reaches a maximum r-value.
3. In order to successfully train their ANNs, over 200 samples were collected. Future studies should endeavour to acquire the same number of samples.

2.2.2 Prediction of hydrocyclone performance using artificial neural networks

The second article that was critically reviewed was the article called Prediction of hydrocyclone performance using artificial neural networks written by M. Kamiri et al. The aim of the authors' study was to employ ANNs to predict the corrected cut-size (d_{50}), the underflow flow rate (Q_u) and the overflow flow rate (Q_o). The authors conclude that the ANNs could be used for automatic control purposes. The inputs to the ANNs were pressure (P), feed solid per cent, the spigot opening diameter (D_u) and the vortex finder diameter (D_o). Two separate ANNs were developed. The first ANN had only the corrected cut-size as output and the second ANN had the two flow rates (Q_u and Q_o). The data set used to train, validate and test the ANNs comprised of 30 samples were the ratio was taken as 63:20:17.

When evaluating the results, it is seen that the authors used three different measures to determine whether the ANNs that were developed were accurate. The first approach was to calculate the error metrics (these include MSE, NMSE and MAE). The second measure the coefficient of determination (R^2) was established. Lastly the test data set's experimental (actual) values were compared to the predicted values. The d_{50c} ANN showed small error metrics and a very good coefficient of determination of 0.977. The actual versus predicted values show deviations. One might attribute these to the ANN errors and to possible experimental errors. The underflow and overflow flow rate ANN shows very promising results as the errors expected are small and the coefficient of determinations are very good (0.989 and 0.994 respectively). The actual and predicted values deviate only slightly and show that the ANN is sufficiently accurate.

One cannot deduce the effects of the experimental error as the authors do not supply any measure thereof. The ANNs were deemed as a promising aspect of automatic control of the hydrocyclone, yet no information is given thereafter. The conclusion is somewhat lacking as it only summarises the results already given in the article; it does not give any insight, recommendations or future work.

CHAPTER 3 – MODELLING APPROACH

3.1 Introduction

The first phase of the project was to research and identify possible modelling approaches specific to hydrocyclones. By choosing a modelling approach, the project was given direction and particular objectives could be established. This chapter starts off by detailing the types of modelling seen in literature by giving some background on them and by indicating some of the advantages and drawbacks to employing them. The chosen modelling approach is discussed and reasons why it was selected are given. With the modelling approach set, the model specifications in terms of the inputs and outputs are evaluated. In order to include the most influential variables as inputs to the model, the empirical models from literature are referenced and summarised. Based on those findings, nine models to be developed were specified; the outputs being the two efficiency indicating parameters. Having established the inputs and outputs of the nine models, the data that would be required to develop the models, could be outlined.

3.2 Types of modelling

From literature there are two main modelling approaches when working with hydrocyclones. The first approach, called empirical modelling, entails working with experimentally obtained data to derive statistical correlations between the variables and responses. Empirical models such as those developed by Plitt, Nageswararao, Lynch & Rao and Bradley were difficult to derive because the variables they incorporated had to be isolated and investigated separately, making it a complex and time-consuming exercise [1], [5], [10], [21]. Furthermore it is known that the experimental conditions of a hydrocyclone-set-up might undergo changes[19], delivering different results under seemingly similar operating conditions. These complications limited the number of variables their models could incorporate. Empirical models are therefore not always comprehensive and might not be representative when used with a different hydrocyclone system. The second approach, usually referred to as fundamental modelling, is used to describe the fluid flow and particle motion. This approach either directly incorporates basic fluid flow equations or makes use of Computational Fluid Dynamics (CFD) software packages [10]. Fundamental models are however extremely complex[20] and is said to be very demanding of computational power, signifying that the processing and simulations could become very time consuming [10], [22]. In order to avoid some of the shortcomings of the approaches mentioned, researchers started

[19] There occur frequent shifts in the feed Particle Size Distribution (PSD), slurry temperature and solid contents to name but a few.
[20] The fundamental models are represented by intricate three-dimensional, three-phase flows [22].

employing computational models such as Fuzzy models and Artificial Neural Networks (ANNs). Promising results were obtained, signifying quite a shift in the modelling of hydrocyclones [4], [20], [22].

After much deliberation, an Artificial Neural Network approach was chosen for this study. The approach presented many advantages and warranted favourable results. Some of these advantages include, but are not limited to:

1. The number of hydrocyclone variables one can use to develop ANN models are not as limited as with the conventional empirical models.
2. Research has shown that ANNs deliver better estimations than the conventional models [4], [20], [23].
3. Artificial Neural Networks handles complex and non-linear systems, such as the hydrocyclone, with relative ease.
4. ANNs have the potential to be used within control applications.
5. The computational requirements of ANNs are substantially less than that of CFD models.

There also exist disadvantages when employing ANNs when modelling hydrocyclones, these disadvantages include:

1. The size of the data set used for training the ANN should be comprehensive and large enough. In order to obtain such a data set many experiments need to be run, recorded, analysed and processed, which makes for a time-consuming approach.
2. If the ANNs' structure and training parameters are not chosen appropriately, under-fitting and over-fitting can easily occur, producing unusable ANNs.
3. Seeing as an ANN is a black-box approach additional investigations are required to understand which variables might be impacting the ANN's training and performance negatively.

3.3 Model specifications

With the approach identified, the model specifications in terms of the inputs and outputs were determined. As mentioned, the two efficiency indicating parameters (d_{50} & m) are the models' outputs. Empirical models from literature were extensively referenced in order to determine which variables would serve as feasible model inputs. Table 3-1 tabulates some of the most popular empirical models, their associated researchers and the variables that they considered influential specifically to the cut-size and sharpness of classification.

Table 3-1: Well-known empirical models and their variables [1], [4], [10], [20], [21], [24]

Researcher(s)	Variables	Year
	Cut-size	
Bradley[21]	$D_c, \eta, Q_i, \rho_s, \rho_l$	1965
Lynch & Rao	$D_o, D_u, C_{w\%}, Q_i$	1975
Flintoff et al.	$D_c, D_i, D_o, D_u, \eta, \phi, h, Q_i, \rho_s$	1987
Nageswararao	$D_c, \frac{D_o}{D_c}, \frac{D_u}{D_c}, \frac{D_i}{D_u}, \frac{L_c}{D_c}, \theta$	1978
H. Eren & A. Gupta	D_c, ρ_p, Q_i, h, T	1990
H. Eren et al.	$D_c, \rho_p, Q_i, h, t, Q_o, Q_u, \frac{Q_o}{Q_u}, \rho_o, \rho_u, \Delta\mathrm{P}$	1997
	Sharpness of classification coefficient	
Flintoff et al.	S, h, D_c, Q_i	1987

The first important aspect noticed was that not many researchers provided models that directly estimated the sharpness of classification coefficient. It will thus be interesting to see what results this study might deliver in terms of the sharpness of classification. When examining the cut-size models, it is seen that many of them included the design variables; such as hydrocyclone diameter (D_c), inlet diameter (D_i), the vortex finder diameter (D_o), the spigot opening diameter (D_u) and the free vortex height (h). The only non-constant[22] design variable that was relevant to this project's test rig is the spigot opening diameter (D_u). Most of the models deem the inlet flow rate (Q_i) influential, as such it would be imperative to also incorporate it into the new models. The popular ANN models - as developed by H. Eren et al. - suggest that by including unusual variables such as overflow and underflow flow rates (Q_o & Q_u) the models' estimation accuracy surpassed those of the conventional empirical models. These unusual variables will therefore be incorporated as well.

After considering the aspects mentioned, the main models that were to be developed are specified. These models are detailed in Figure 3-1, depicting the models' name, the inputs and the outputs. Basically three major groups were defined to be developed, one group with three inputs, the next group with five inputs and the last group utilising eight inputs. Each one of these groups has a subset of three models, one model where the cut-size is the only output, one where sharpness of classification is the only output and a third one where both cut-size and sharpness of classifications are outputs. This will investigate whether it is possible to model both the cut-size and sharpness of classification by incorporating the same variables as inputs; producing a single

[21] Bradley's model was theoretical rather than empirical.

[22] Constant inputs do not provide useful information to the network being trained and are therefore removed [17].

ANN that could predict the two outputs. Group 01XX consists of three inputs only; the three experimental condition variables that will be set during experimental runs (P, ϕ, D_u). The second group, 02XX, include the same first three inputs with the addition of two operating variables ($P, \phi, D_u, Q_i, \omega$[23]). This set-up is based on the assumption that the test rig could be instrumented with relatively inexpensive equipment[24]. The last group, referred to as 03XX, includes some of the unusual variables as suggested by H. Eren et al. The inputs are thus the three experimental condition variables, the inlet flow rate and angle of discharge and three relevant unusual variables ($P, \phi, D_u, Q_i, Q_o, Q_u, \rho_o, \omega$). It should be mentioned that this set-up will need to incorporate some expensive measuring instruments[25].

With the models specified the scope of the data that will be required also become more distinct. Table 3-2 summarises the measurements and samples that will need to be obtained in order to develop the models as described.

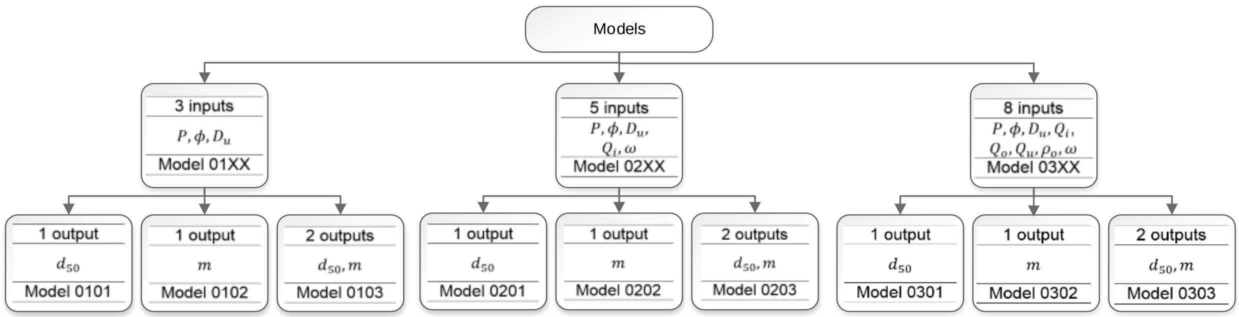


Figure 3-1: Model specification summary

[23] The angle of discharge is a relatively new operating variable used to evaluate the operating condition of the hydrocyclone system [27]; Chapter 4 discusses the angle of discharge in more detail.

[24] A flowmeter to measure the inlet flow rate and a digital camera to capture the angle of discharge can be utilised.

[25] In order to measure the overflow density on-line, a γ-ray density gauge would need to be installed. An additional flowmeter will also be required.

Table 3-2: Summary of parameters that will need to be measured

Variables to be measured		
Description	**Symbol**	**Unit**
Inputs		
Pressure	P	kPa
Volumetric solid concentration	ϕ	%
Spigot opening diameter	D_u	mm
Inlet flow rate	Q_i	$\mathrm{l/s}$
Density of overflow sample	ρ_o	$\mathrm{kg/m^3}$
Overflow flow rate	Q_o	$\mathrm{l/s}$
Underflow flow rate	Q_u	$\mathrm{l/s}$
Angle of discharge	ω	°
Outputs		
Cut-size	d_{50}	μm
Sharpness of classification	m	-

3.4 Conclusion

When assessing the possible modelling approaches, it seems that an Artificial Neural Network (ANN) approach would be ideal to apply when working with hydrocyclones, mainly because of its adaptability. It is also expected that ANNs will deliver better results when compared to other empirical models. By extensively studying the literature on empirical models that estimate the cut-size and sharpness of classification, it becomes clear which variables are seen as most influential. Based on these findings nine different models were specified. With the model specifications set, the data that will need to be acquired were also identified.

CHAPTER 4 – SYSTEM REALISATION

4.1 Chapter introduction

Chapter 4 details the realisation of the system and the obtainment of experimental data. The hydrocyclone test rig and the relevant aspects thereof are described first. Next the experimental procedure followed for this study is given. The methods used to analyse the collected data are summarised and finally the processing of the data is specified. To conclude the relevant chapter information is discussed and also recapped in a useful table.

4.2 Experimental set-up

4.2.1 The hydrocyclone test rig

The industry partners, Multotec, sponsored a refurbished hydrocyclone test rig as part of collaboration projects with the university; a snippet of the test rig schematic is depicted in Figure 4-1.

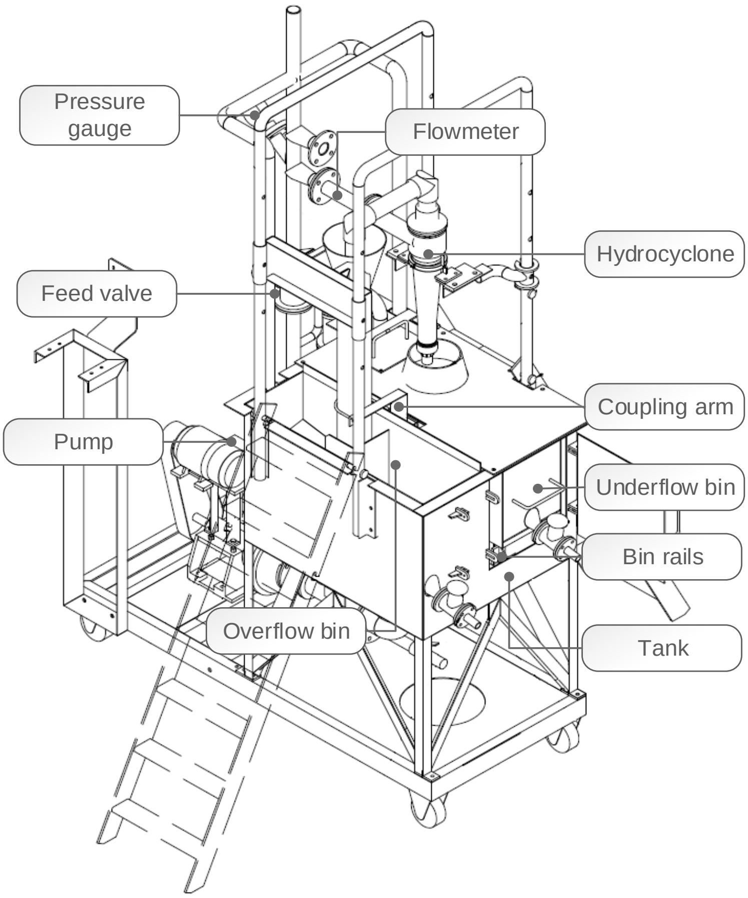
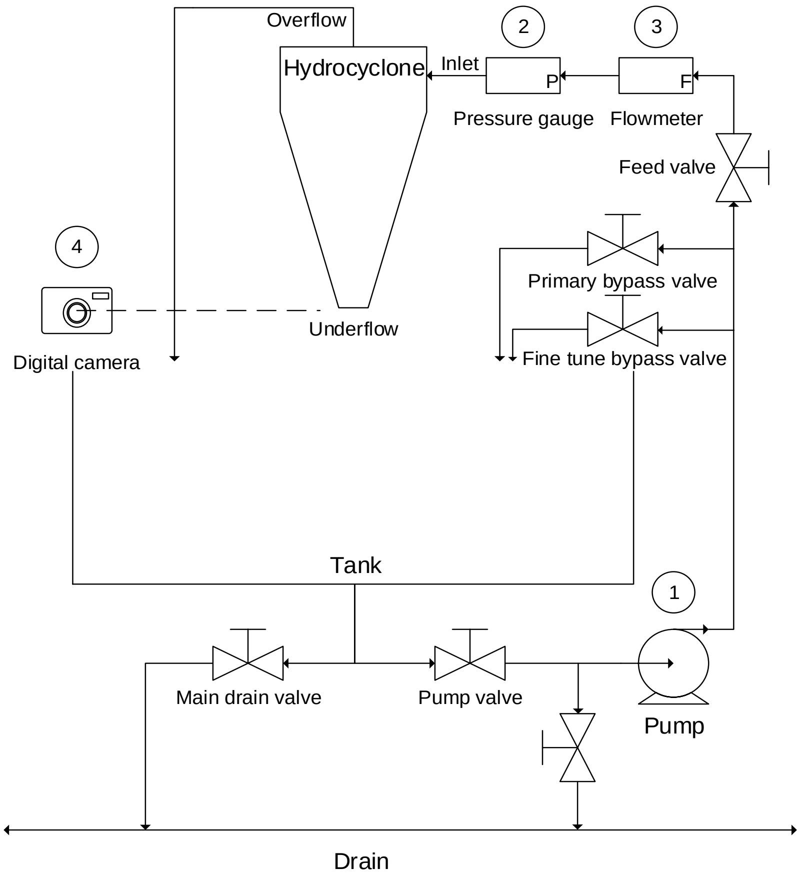


Figure 4-1: Hydrocyclone test rig schematic

The rig consists of a pump, a tank, a hydrocyclone, piping systems and two sampling bins for the underflow and overflow products. The underflow bin was built on a rail, making sampling more convenient. It is also directly coupled to an arm that simultaneously moves the overflow from the tank into its sampling bin. In other words the slurry circulates through the hydrocyclone and into the main tank until sampling is commenced.

As mentioned in Chapter 3, the instruments needed for the group 03XX models are quite expensive. The system was therefore not instrumented to measure all of the specified variables. The test rig was already fitted with an analog pressure gauge which measures the inlet pressure (P). Additionally a Doppler flowmeter was installed in order to continuously measure the inlet flow rate (Q_i). To capture the angle of discharge (ω), a digital camera was employed. The Piping and Instrumentation Diagram (P&ID), as derived from the actual system, is shown in Figure 4-2.



ITEM	QTY.	DESCRIPTION	MANUFACTURER
1	1	Pump	Bacuneer
2	1	Pressure gauge	WIKA
3	1	Flowmeter	Greyline instruments inc.
4	1	Digital camera	Samsung

Figure 4-2: P&ID of the hydrocyclone test rig

Table 4-1 summarises the most important hydrocyclone design variables. The inlet of the hydrocyclone is not circular but square, thus the square inlet dimensions were converted to a circular inlet diameter of the same area [10]. The conversion is shown in Appendix A.1.

Table 4-1: Hydrocyclone design variables

Description	Symbol	Size	
Hydrocyclone diameter	D_c	100	mm
Inlet diameter (circular)	D_i	30.3	mm
Vortex finder opening	D_o	34	mm
Free vortex height	h	531	mm
Spigot opening diameter	D_{u1}	15	mm
	D_{u2}	20	mm
	D_{u3}	25	mm
	D_{u4}	30	mm
	D_{u5}	35	mm
Cone angle	θ	8	°

4.3 Experimental procedure

The experimental procedure describes in detail the steps taken to prepare for and to execute the designed experiments. This subsection discusses the pre-experiment preparations, the experiment preparation, the sampling process, the post-experiments and the mixing of slurries.

4.3.1 Pre-experiment preparations

The pre-experiment preparations mostly consist of planning and calculations that were implemented before starting with the actual experimental runs. These steps are crucial as they define and organise the experiments as well as identify additional necessities beforehand. The pre-experiments involved calculating the mass of solids that would have been needed and preparing a large enough batch of uniform Particle Size Distribution (PSD) materials based on these calculations. The step-wise procedure is given in Appendix A.3.

4.3.2 Experiment preparations

After the experiments and related aspects were planned, a comprehensive procedure was developed that initiated the experiments and prepared for the sampling thereof. These preparations include filling the tank with the necessary water (210 litre), calibrating the Marcy scale, weighing off the needed silica for the specific volumetric solid concentration (ϕ) and inserting and securing the required spigot. A step-wise preparation procedure is discussed in Appendix A.3.

4.3.3 Sampling

The sampling steps describe in what manner and order the different measurements were taken. The step-wise procedure is given in detail in Appendix A.3. A summary of the measurements taken and their sequences are given in Table 4-2.

Table 4-2: Measurement parameters recorded during sampling

Description	Symbol	Unit	Acquisition instrument
Pressure	P	kPa	Pressure gauge
Volumetric solid concentration	ϕ	%	Marcy scale
Representative feed PSD sample	-	-	Sample taken
Inlet flow rate	Q_i	l/s	Flowmeter
Sampling time	t	s	Stopwatch
Angle of discharge	ω	–	Camera
Weight of underflow sample	M_u	kg	Scale
Weight of overflow sample	M_o	kg	Scale
Density of underflow sample	ρ_u	$\mathrm{kg/m^3}$	Marcy scale
Density of overflow sample	ρ_o	$\mathrm{kg/m^3}$	Marcy scale
Representative underflow sample	d_{50}, m	–	Sample taken

4.3.4 Post-sampling

The post-experiment procedure involves all the steps done after all of the experimental runs are completed. This ensures that no measurements were overlooked and that the test rig is cleaned out appropriately. The full post-experiments procedure is given in Appendix A.3.

4.3.5 Mixing of the slurries

As specified in the Experimental Design (ED) the volumetric solid concentration (ϕ) was one of the three influential factors. In order to mix the slurries to the specified solid concentrations, some general slurry interrelation formulae were utilised. See Appendix A.3 for the detailed calculations.

4.4 Experimental analyses

4.4.1 Malvern particle size analyser

After completing the experimental runs the stored representative samples taken of the underflow need to be analysed in order to eventually extract the cut-size (d_{50}) and sharpness of classification (m) information. This analysis is done by using the Malvern Mastersizer 2000[26]. The Malvern is a laser diffraction instrument used to measure the Particle Size Distribution (PSD) of a presented

[26] The Malvern Mastersizer can analyse particles in the size range of 1 – 2000 $\mu\mathrm{m}$ [6]. When evaluating the silica's PSD profiles in section A.6 it is evident that the Malvern will be able to accurately measure the underflow samples.

sample. Sample particles are mixed into a dilute suspension which is circulated through an optical cell. Laser light is then shone through the suspension and particles. The light is scattered by the particles and detectors within the Malvern measure the intensity of the scattered light over a range of angles. The PSD is then calculated in terms of the light scattering pattern measured. The information is then processed and displayed on a computer software application. The principle is illustrated in Figure 4-3 [6].

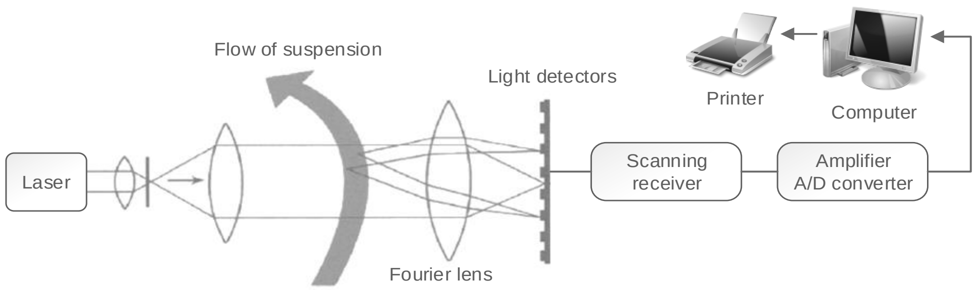


Figure 4-3: Laser diffraction instrument principle illustration (adapted) [6]

4.4.2 Malvern analysis procedure

Every one of the stored underflow samples were meticulously analysed by using the Malvern Mastersizer. The analysis procedure is described in Appendix A.4.

4.5 Processing of the experimental data

Some of the collected measurements still need to be processed in order to be applicable. This subsection details the processing procedures in regard to determining the cut-size and sharpness of classification coefficient, the underflow and overflow flow rate and the angle of discharge.

4.5.1 Cut-size and sharpness of classification from Malvern analysis

In order to obtain the cut-size and the sharpness of classification of each sample, the Malvern analyses were exported to Excel and processed by making use of an Excel macro. A detailed description of the processing is given in Appendix A.5.

4.5.2 Determining the underflow and overflow flow rate

Having had only one flowmeter installed, the only instrumentally measured flow rate was the inlet flow rate (Q_i). By utilising some of the other recorded measurements it is possible to calculate the volumetric flow rate of the underflow (Q_u) and of the overflow (Q_o). This then delivers two additional calculated variables. The conversion of these variables are shown in Appendix A.5.

4.5.3 Processing of the angle of discharge

The final variable that was processed, was the angle of discharge (ω). Photos were taken of the spigot and the underflow discharge directly after sampling the products. The sample's angle of discharge was extracted from the photo by utilising the Image Processing Toolbox in MATLAB®.

The diagram in Figure 4-4 show the chronological steps in processing the photos of the underflow discharge. The MATLAB® angle extraction method described in [25] was slightly altered in order to be applicable to the angle of discharge photos. The complete procedure is shown step-wise in Appendix A.5.

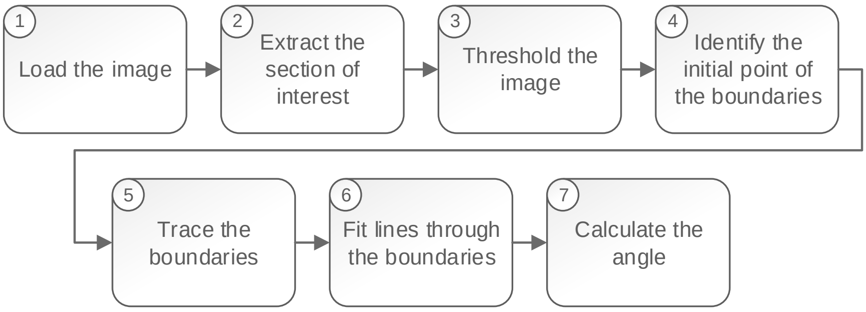


Figure 4-4: Processing steps in determining the angle of discharge from a photo

4.6 Chapter summary

This chapter detailed the various aspects of acquiring, analysing and processing the experimental data needed for this study. Firstly the sponsored test rig was discussed, detailing some of the most important aspects of its assembly and instrumentation. An in-depth description of the preparations and sampling procedures are given step-wise. The experiments, as set out during the Experimental Design phase, could be conducted and the measurements recorded accordingly. The collected underflow samples were analysed using the Malvern Mastersizer. In order to extract the cut-size (d_{50}) and sharpness of classification (m) coefficient, the Malvern analyses data were exported and processed using developed Excel macros. Some additional parameters such as the underflow (Q_u) and overflow flow rate (Q_o) were determined by employing standard conversion formulae. The last parameter that was processed, was the angle of discharge (ω). By utilising image processing functions within MATLAB® the angle of discharge for each sample could be extracted from the photos taken during sampling.

Table 4-3 summarises the obtained measured parameters, showing what instrument was used to obtain the measurement, whether (and in what manner) the measurement was analysed as well as the final processing technique used to eventually deliver useable data. This data can now be used to train and test the specified ANNs.

Table 4-3: Summary of collected experimental measurements

Measured variable			Details		
Description	**Symbol**	**Unit**	**Acquisition instrument**	**Analysis method**	**Processing technique**
Pressure	P	kPa	Pressure gauge	-	-
Volumetric solid concentration	ϕ	%	Formulae & Marcy scale	-	-
Spigot opening diameter	D_u	mm	Visual confirmation	-	-
Inlet flow rate	Q_i	l/s	Doppler flowmeter	-	-
Sampling time	t	s	Stopwatch	-	-
Weight of underflow sample	M_u	kg	Scale	-	-
Weight of overflow sample	M_o	kg	Scale	-	-
Density of underflow sample	ρ_u	$\mathrm{kg/m^3}$	Marcy scale	-	-
Density of overflow sample	ρ_o	$\mathrm{kg/m^3}$	Marcy scale	-	-
Cut-size	d_{50}	$\mu\mathrm{m}$	Sample taken	Malvern	Excel macro
Sharpness of classification	m	-	Sample taken	Malvern	Excel macro
Underflow flow rate	Q_u	l/s	-	-	Formulae
Overflow flow rate	Q_o	l/s	-	-	Formulae
Angle of discharge	ω	$^\circ$	Camera	-	MATLAB®

CHAPTER 5 – DESIGN AND ANALYSIS OF EXPERIMENTS

5.1 Introduction

For this study it was imperative to have useful experimental data seeing as the Artificial Neural Network is based on experimental data only. In order to acquire useful data it is necessary to design experiments beforehand to ensure that the entire system's operating range is included. It also ensures that the experimental effort is as effective as possible. Centrally Composite Rotatable Design (CCRD) is an Experimental Design technique that is incorporated to achieve just that. CCRD consists of well-defined experiments that are conducted in such a manner that comprehensive insight is gained with as little experimental effort as possible. The CCRD technique eventually delivers a mathematical model that can be used to further inspect the responses of interest. Figure 5-1 shows the essential steps of a CCRD. The chapter was also written in this order. The first step was to determine the factors that influenced the responses being investigated. Next the operating spectrum of the system was defined and assigned. The design matrix is done to outline what the experiments' conditions will be. These designed experiments were then conducted in order to record the responses. With the responses recorded the mathematical equation coefficients were determined by employing least square methods. The obtained model's adequacy was evaluated and if the model was deemed useful, it was employed to inspect characteristics of the responses. Only the most important aspects and calculations are shown and discussed.

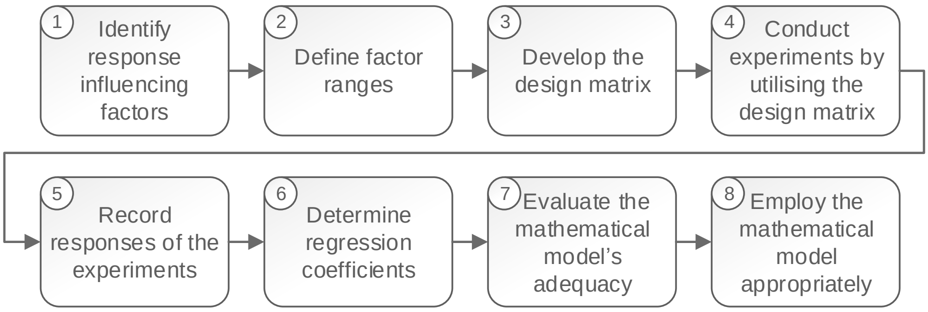


Figure 5-1: The steps taken with a CCRD approach

5.2 Factor identification

The first important step in Experimental Design is to determine the process variables that will significantly affect the response variables that are investigated. For this study, the response variables were the cut-size (d_{50}) and the sharpness of classification coefficient (m). Some additional response variables that were evaluated include the feed flow rate (Q_i) and the angle of discharge (ω).

The d_{50} influencing variables that were found to be most prominent throughout literature and that are relevant to the hydrocyclone test rig, were the feed pressure (P), the volumetric solid concentration (ϕ) and the spigot opening diameter (D_u) [1], [2], [10]. Table 5-1 summarises the response variables and the three independent variables, called factors, as discussed.

Table 5-1: Summary of the factors and response variables for the hydrocyclone

Factors		Response variables	
Factor name	**Notation**	**Response name**	**Notation**
Pressure	P	Cut-size	d_{50}
Solid concentration	ϕ	Sharpness of classification	m
Spigot opening diameter	D_u	Feed flow rate	Q_i
		Angle of discharge	ω

5.3 Ranges of factors

In order to utilise the CCRD technique, the factors first needed to be assigned practical operating condition ranges. It is of utmost importance that the ranges include the entire spectrum of operating conditions but that the ranges never exceed viable conditions [12], [13]. Some trial runs were carried out to determine the hydrocyclone test rig viable operating spectrum.

The hydrocyclone test rig was firstly filled with water only and was fitted with a 25 mm spigot. The valves were set appropriately and the lowest and highest possible pressures achieved were recorded. The pressure ranged between 35 and 120 kPa. Next a range of slurries were tested in order to determine the possible pressure ranges achievable with different volumetric solid concentrations. It was also vital to determine at what volumetric solid concentration the slurry exceeded acceptable[27] settling within the test rig tank. Starting off with a low volumetric solid concentration and sequentially adding more silica to the slurry, the lowest and highest pressures were recorded once more. The pressure ranged between 45 and 105 kPa over the varying solid concentration range. It was found that the highest acceptable volumetric solid concentration was around 3.125 %. The hydrocyclone has 5 spigots in sizes ranging from 15 to 35 mm which can be easily interchanged.

Having identified the factors' operating conditions, the factor range table could be developed, the completed table is shown in Table 5-3. For this CCRD approach the coded values were chosen

[27] When using silica slurry it is known that some settling within the sump will occur after the test rig pump is switched off. When switching the pump on after some time, the slurry being pumped will ensure that the settled silica is remixed into the slurry again where after sampling can continue. This specific test rig however only does so for slurries with low volumetric solid concentration and the pump is blocked if higher volumetric solid concentration slurries are used. Thus acceptable settling was described as a slurry that had a low enough volumetric solid concentration so that it could be remixed without blocking the pump when switched on.

to be -2, -1, 0, 1 and 2. These coded values correspond to range levels described as lowest, low, centre, high and highest. The first values that are assigned are the actual values of the low and high levels of each factor, as determined from the trial runs. Thus the factor's low value is coded to -1 and the high level to 1.

These actual values are chosen in such a manner that the lowest and highest levels (at -2 and 2 respectively) will remain within the viable operating condition ranges. Table 5-2 depicts the calculated g_{x_i} and t_{x_i} for the three variables by using (5-1) and (5-2).

$$g_{x_i} = \frac{high_{x_i} + low_{x_i}}{2} \tag{5-1}$$

$$t_{x_i} = \frac{high_{x_i} - low_{x_i}}{2} \tag{5-2}$$

Table 5-2: The calculated g_{x_i} and t_{x_i} values for the three variables

Variable	Unit	Symbol	low_{x_i}	$high_{x_i}$	g_{x_i}	t_{x_i}
P	kPa	x_1	62.0	85.0	73.5	11.5
ϕ	vol %	x_2	1.250	2.500	1.875	0.625
D_u	mm	x_3	20	30	25	5

With the g_{x_i} and t_{x_i} calculated for each factor, the remaining actual values for the lowest, centre and highest levels can be calculated using (5-3). The summary of the factors' coded and actual values are summarised in Table 5-3.

$$actual\ value = g_{x_i} + \left(coded\ value \cdot t_{x_i}\right) \tag{5-3}$$

Table 5-3: The actual and coded values of the variables

Variable		Unit	Symbol	Coded value				
				lowest	low	centre	high	highest
				-2	-1	0	1	2
Pressure	P	kPa	x_1	50.5	62.0	73.5	85.0	96.5
Solids concentration	ϕ	vol %	x_2	0.625	1.250	1.875	2.500	3.125
Spigot opening diameter	D_u	mm	x_3	15	20	25	30	35

5.4 Development of the design matrix

Now that the factors are coded and assigned actual values, the design matrix can be developed. The design matrix is made up out of specific sequences of the coded factors as prescribed for a CCRD approach with three variables ($k = 3$).

It is expected that there are $2^k = 2^{(3)} = 8$ factorial runs, $2k = 2(3) = 6$ axial runs and 6 centre runs [13]. The matrix is filled in Yates standard order and the result is depicted in Table 5-4. It shows the coded and actual values of the factors per experimental run. The centre runs are designed to be conducted under the same conditions, the responses of these repeated runs are then used to calculate the experimental error [12].

Table 5-4: The design matrix depicting the coded and actual values per experimental run

Run	**Coded value**			**Actual value**		
	x_1	x_2	x_3	x_1	x_2	x_3
			Factorial runs			
1	-1	-1	-1	62.0	1.250	20
2	1	-1	-1	85.0	1.250	20
3	-1	1	-1	62.0	2.500	20
4	1	1	-1	85.0	2.500	20
5	-1	-1	1	62.0	1.250	30
6	1	-1	1	85.0	1.250	30
7	-1	1	1	62.0	2.500	30
8	1	1	1	85.0	2.500	30
			Axial runs			
9	-2	0	0	50.5	1.875	25
10	2	0	0	96.5	1.875	25
11	0	-2	0	73.5	0.625	25
12	0	2	0	73.5	3.125	25
13	0	0	-2	73.5	1.875	15
14	0	0	2	73.5	1.875	35
			Centre runs			
15	0	0	0	73.5	1.875	25
16	0	0	0	73.5	1.875	25
17	0	0	0	73.5	1.875	25
18	0	0	0	73.5	1.875	25
19	0	0	0	73.5	1.875	25
20	0	0	0	73.5	1.875	25

5.5 Utilising the design matrix

After the design matrix was completed, it was employed in order to record the observed responses for each one of the 20 specified conditions. The runs from Table 5-4 were organised, and indirectly randomised[28], in order to improve sampling in terms of slurry mixing[29]. Table 5-5 states the new order in which the experiments were conducted, along with the responses recorded for

[28] Randomisation of runs ensures that the effects of unknown factors are minimised.

[29] The experiments were organised and grouped by the volumetric solid concentration. In other words, starting with the lower volumetric solid concentrations, it was only necessary to add silica to the existing slurry in order to run a preceding experiment with higher volumetric solid concentration. This ensured that minimal time, silica and water are wasted.

the cut-size (d_{50}), the sharpness of classification coefficient (m), feed flow rate (Q_i) and angle of discharge (ω).

Table 5-5: Summary of experimental run conditions and the response values

Run	Variables						Responses			
	Coded value			Actual value			Actual response value			
	x_1	x_2	x_3	x_1	x_2	x_3	d_{50}	m	Q_i	ω
11	0	-2	0	73.5	0.625	25	34.034	1.71	3.512	45.4
1	-1	-1	-1	62.0	1.250	20	36.430	1.67	3.222	43.2
2	1	-1	-1	85.0	1.250	20	35.817	1.68	3.607	43.6
5	-1	-1	1	62.0	1.250	30	32.000	1.69	3.561	54.1
6	1	-1	1	85.0	1.250	30	33.556	1.70	4.052	51.4
9	-2	0	0	50.5	1.875	25	33.428	1.74	2.997	48.3
10	2	0	0	96.5	1.875	25	36.122	1.60	4.083	43.9
13	0	0	-2	73.5	1.875	15	38.872	1.72	3.418	51.1
14	0	0	2	73.5	1.875	35	32.138	1.69	3.961	61.6
15	0	0	0	73.5	1.875	25	33.268	1.77	3.650	45.9
16	0	0	0	73.5	1.875	25	35.791	1.61	3.670	43.7
17	0	0	0	73.5	1.875	25	34.836	1.61	3.612	46.3
18	0	0	0	73.5	1.875	25	34.867	1.61	3.675	44.9
19	0	0	0	73.5	1.875	25	33.868	1.80	3.812	43.6
20	0	0	0	73.5	1.875	25	36.714	1.58	3.728	43.9
3	-1	1	-1	62.0	2.500	20	36.954	1.67	3.307	44.3
4	1	1	-1	85.0	2.500	20	36.346	1.65	3.950	41.9
7	-1	1	1	62.0	2.500	30	32.831	1.70	3.591	53.5
8	1	1	1	85.0	2.500	30	30.813	1.77	4.207	53.7
12	0	2	0	73.5	3.125	25	34.390	1.61	3.657	45.6

5.6 Mathematical model development

With the experimental runs completed and the responses of interest recorded, it was possible to develop mathematical models that describe the response variable in terms of the pressure, volumetric solid concentration and spigot opening diameter. For the CCRD approach the mathematical model consists of main effect terms of each factor, quadratic terms of each of the factors and first order interaction terms for each paired combination of factors. (5-4) shows the general form of the mathematical model equation. It can also be written in the form shown in (5-5).

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4{x_1}^2 + b_5{x_2}^2 + b_6{x_3}^2 + b_7x_1x_2 + b_8x_1x_3 + b_9x_2x_3 \tag{5-4}$$

$$Y = bX + \varepsilon \tag{5-5}$$

Here Y represents the matrix of actual response values. X is the matrix of the independent factors, as shown in (5-6). In this study the X matrix will stay the same because all the responses were taken under the same conditions. Matrix b represents the unknown coefficients matrix and ε the error matrix.

In order to determine the coefficient matrix, a least square method (lscov function) within MATLAB® was employed. In order to better organise the results each response's mathematical model, coefficients and discussions will be documented under its own subsection.

$$X = \begin{bmatrix}
1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 \\
1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\
1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
1 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\
1 & 0 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & -2 & 0 & 0 & 4 & 0 & 0 & 0 \\
1 & 0 & 0 & 2 & 0 & 0 & 4 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix} \tag{5-6}$$

5.6.1 The cut-size response

5.6.1.1 Cut-size mathematical model

The first response that was investigated, was the cut-size (d_{50}). For this study it was very important to understand what the cut-size might be under certain operating conditions. This would ensure that insightful additional experiments could be identified, conducted, analysed and eventually be used for the Artificial Neural Network training. It was thus a priority to develop an accurate mathematical model that could describe expected d_{50} values across the defined operating condition ranges.

As mentioned the least square function within MATLAB® was used to calculate the coefficients. The obtained coefficients' values and their corresponding terms are tabulated in Table 5-6. The mathematical equation is given in (5-7) and its performance is evaluated next.

Table 5-6: Cut-size equation coefficients

Coefficient		Corresponding term
Symbol	Value	
b_0	34.79934	constant
b_1	0.23156	x_1
b_2	-0.00919	x_2
b_3	-1.86344	x_3
b_4	-0.07458	${x_1}^2$
b_5	-0.21533	${x_2}^2$
b_6	0.10792	${x_3}^2$
b_7	-0.44613	x_1x_2
b_8	0.09488	x_1x_3
b_9	-0.37062	x_2x_3

$$\begin{aligned} y_{d_{50}} = 34.79934 + 0.23156x_1 - 0.00919x_2 - 1.86344x_3 - 0.07458{x_1}^2 \\ - 0.21533{x_2}^2 + 0.10792{x_3}^2 - 0.44613x_1x_2 + 0.09488x_1x_3 \\ - 0.37062x_2x_3 \end{aligned} \quad \textbf{(5-7)}$$

5.6.1.2 Summary of fit

Using the response equation in (5-7) the estimated values $\hat{y}_i$ were calculated for the same twenty experimental conditions. In order to evaluate the fit of the mathematical model, the estimated values on the y-axis were plotted against the actual response values on the x-axis as depicted in Figure 5-2.

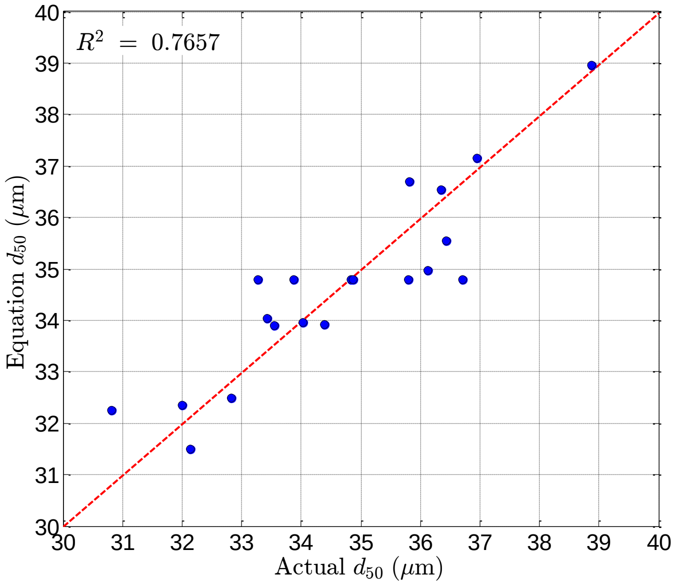


Figure 5-2: The actual cut-size versus the estimated cut-size

When the estimated and actual values are plotted against one another, the correlation coefficient (r) can be calculated along with the coefficient of determination (R^2). The r-value was calculated by employing (5-8). It was found to be 0.9 which is indicative of a strong positive linear correlation.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}} \tag{5-8}$$

To determine the R^2-value either (5-9) can be employed or MATLAB® functions can be utilised. In this study MATLAB® was used to determine the R^2-value and the results obtained were checked by employing (5-9). The R^2-value of 0.7657 indicates satisfactory fit.

$$R^2 = 1 - \frac{SS_X}{SS_T} \tag{5-9}$$

The adjusted R^2-value, idenoted as $\bar{R}^2$ in (5-10), is also calculated in MATLAB®. The R^2-value however, is not the only parameter that should be used to determine whether the model's fit is adequate, additionally an Analysis of Variance (ANOVA) is done as described in 5.6.1.3. The results obtained for the model's fit is summarised in Table 5-7.

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1 - R^2) \tag{5-10}$$

Table 5-7: The summary of fit for the cut-size response

Summary of fit	
r	0.9000
R^2	0.7657
$\bar{R}^2$	0.7218

5.6.1.3 Analysis of variance

In order to determine whether the model can be deemed adequate or not, analysis of variance (ANOVA) is done for the cut-size mathematical model. The F-value was determined for the regression as well as the lack of fit. Next the critical F-values for the regression and lack of fit were determined[30] for an $\alpha = 0.05$. The ANOVA results are summarised in Table 5-8. It is seen that the regression was found to be significant at the level 95 %. Furthermore the lack of fit was found to be non-significant at the level 95 %. It is thus concluded that the mathematical model represents the cut-size response adequately [26].

[30] The critical F-values were determined from the standard F Critical Values tables provided in [13].

Table 5-8: ANOVA for the cut-size mathematical model

Source	SS	df	MS	F	test F
Regression	61.056	8	7.632	5.864	$F_{0.05}(8,11) = 2.95 < 5.86$
Residual	14.317	11	1.302		Significant @ the level 95%
- Lack of fit	6.449	3	2.150	1.913	$F_{0.05}(3,7) = 4.35 > 1.91$
- Error	7.868	7	1.124		Non-significant @ the level 95%
Total	75.373	19			

5.6.1.4 Experimental error of the cut-size

Before the mathematical model can be utilised to gain insight into the response and its aspects, there is one last step to complete. An experimental error is calculated in order to determine the acceptable level of variation that occurs within the measurement of the system [13]. For this response, the experimental error was calculated for both the experiments and for the analysis done in the Malvern particle size analyser. The six centre point experiments that were conducted were repeated at the same factor conditions. The responses were recorded and used to determine the expected error. In order to determine the experimental error of the Malvern analyses, a random sample[31] was analysed three times. The results of the experimental errors are tabulated in Table 5-9. The Confidence Level (CL) was chosen at a level of 95%. The various parameters needed were calculated by using the formulae as summarised in Table 5-9. The experimental error for the experiments was found to be 2.95 % and 2.71 % for the Malvern analysis. The highest experimental error of the two is the experiments' error of 2.95 %; this will be used throughout the study to indicate the acceptable variation that is expected for the cut-size.

Table 5-9: Summary of experimental error of the cut-size response

Parameter			Experiments	Malvern
Confidence level	CL	-	95	95
Number of samples	n	-	6	3
Degrees of freedom	df	$n-1$	5	2
Average	$\bar{y}$	$\frac{1}{n}\sum_{i=1}^{n} y_i$	34.891	34.949
Standard deviation	s	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$	1.250	0.561
t-value	t_{n-1}	t critical value table	2.015	2.920
Error	e	$t_{n-1}\left(\frac{s}{\sqrt{n}}\right)$	1.0286	0.946
Interval	-	$[\bar{y}-e;\bar{y}+e]$	[33.862;35.919]	[34.003;35.895]
Error in percentage	$e_{\%}$	Convert interval to %	2.948	2.706

[31] Sample uf1_020 was analysed repeatedly.

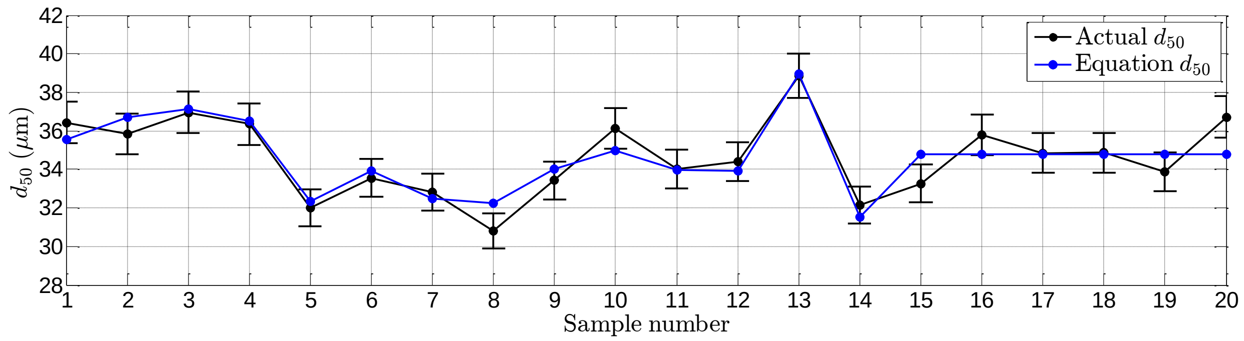


Figure 5-3: The actual and estimated cut-size shown per sample

In order to visualise the mathematical model's capabilities and error properties, the actual and estimated cut-size response values were plotted per sample as shown in Figure 5-3. The error bars depicting the acceptable levels of variation.

5.6.1.5 Insight into the cut-size response

The mathematical model was found to be adequate and with an experimental error established, the model can now be used to gain useful insight. Figure 5-4 shows the three three-dimensional response surface plots that are obtained by employing the mathematical model and plotting the relationship between the cut-size response on the y-axis and two of the three factors on the x-axes. The third factor that is not represented in the plot is kept at a constant 0 coded value. The topology of the response surface plots are illustrated by iso-contours as shown in Figure 5-5. The lines represent constant response values in a two factor plane. This is especially useful when investigating the effect variations in the two factors will have on the response. Figure 5-4 (a) and Figure 5-5 (a) depict the cut-size (d_{50}), the pressure (P) and the solid concentration (ϕ). It seems that the cut-size will increase the lower the solid concentration is and the higher the pressure becomes. Combinations of high solid concentration and high pressure or low solid concentration and low pressure will result in smaller cut-sizes.

Figure 5-4 (b) and Figure 5-5 (b) portray the cut-size (d_{50}), the pressure (P) and the spigot opening diameter (D_u). When studying these plots it appears that the biggest effect on the cut-size is the spigot size that is used. The cut-size increases as the spigot opening diameter becomes smaller. The pressure will have a minimal effect on the cut-size.

Figure 5-4 (c) and Figure 5-5 (c) show the cut-size (d_{50}), the solid concentration (ϕ) and the spigot opening diameter (D_u). The cut-size seems to increase the smaller the spigot size and the higher the solid concentration become. A high solid concentration along with a large spigot opening diameter will result in smaller cut-sizes.

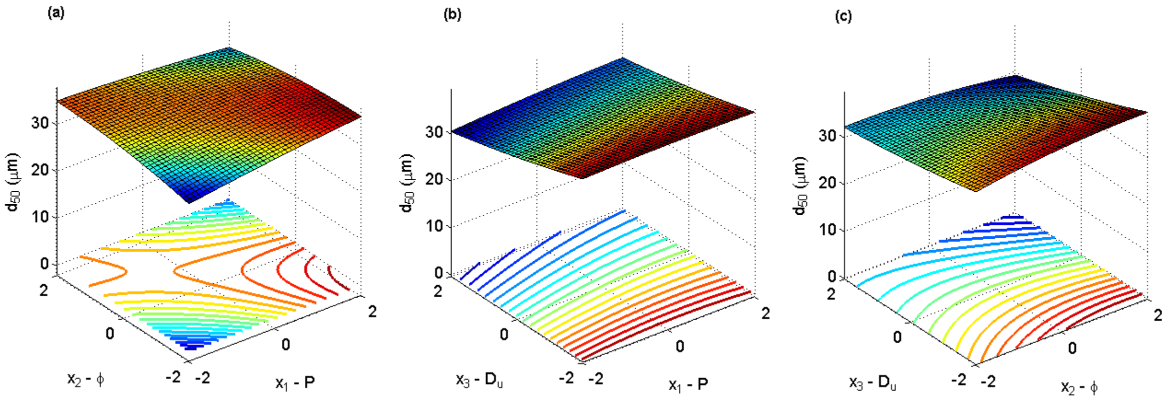


Figure 5-4: The cut-size response surface plots for (a) x_1x_2, (b) x_1x_3 and (c) x_2x_3

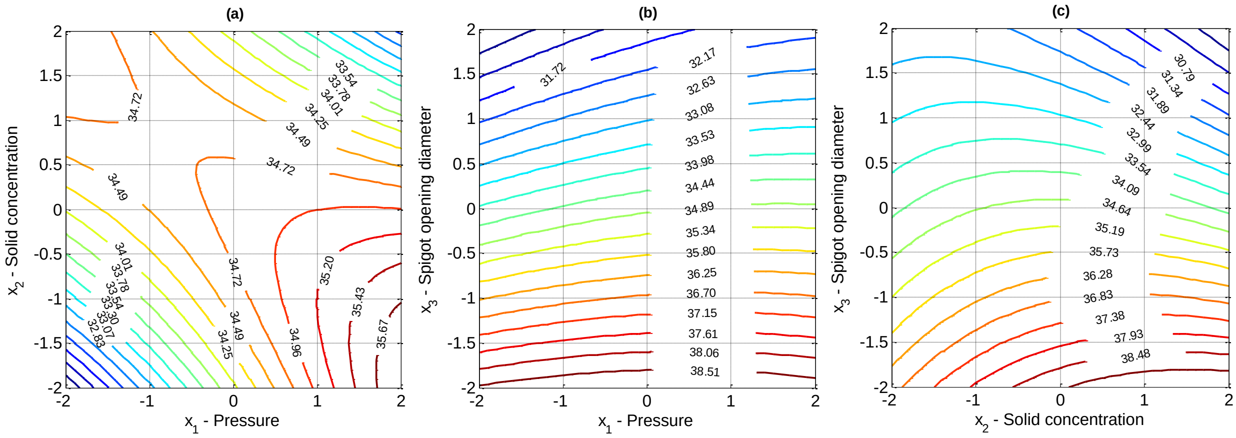


Figure 5-5: The cut-size contour plots for (a) x_1x_2, (b) x_1x_3 and (c) x_2x_3

The plots in Figure 5-6 represent the same information as given in Figure 5-4 and Figure 5-5 but in a third format. This format simplifies the observation of individual effects on the cut-size.

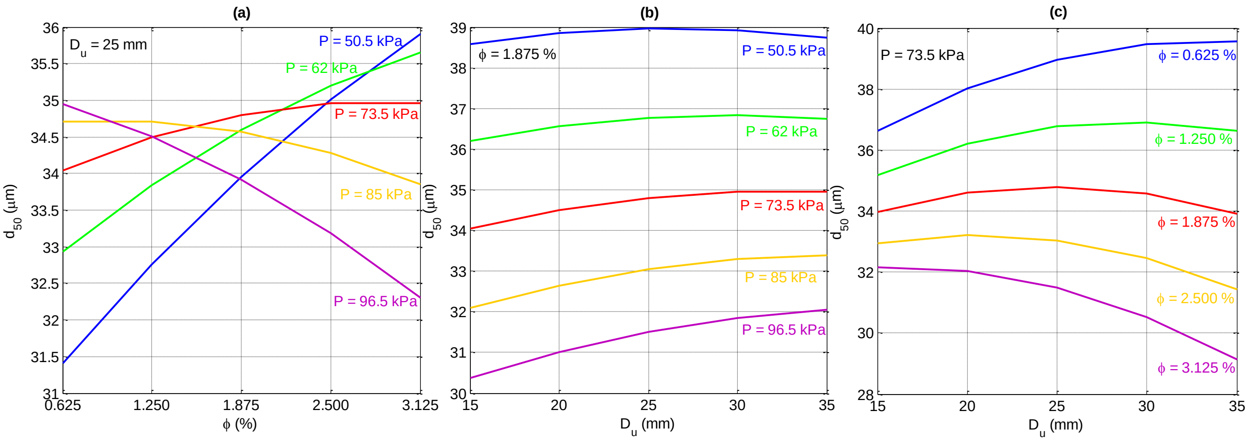


Figure 5-6: Expected effects of the factors on the cut-size

5.6.1.6 Additional experiments identified

In order to eventually train an Artificial Neural Network (ANN), some additional experimental runs are required. By utilising the contour plots twenty-one[32] new experimental conditions were identified. The new run conditions were taken across the spectrum at various points on the contour plots that seemed of note, in order to deliver a wide variety of cut-sizes and to exclude repetitions. This was to ensure that the ANN would be supplied with useful data and to keep the experimental efforts to a minimum. Figure 5-7 shows the contour plots and the twenty-one newly identified experimental conditions[33]. The new experimental runs and their conditions (coded and actual values) are tabulated in Table 5-10 and Table 5-11.

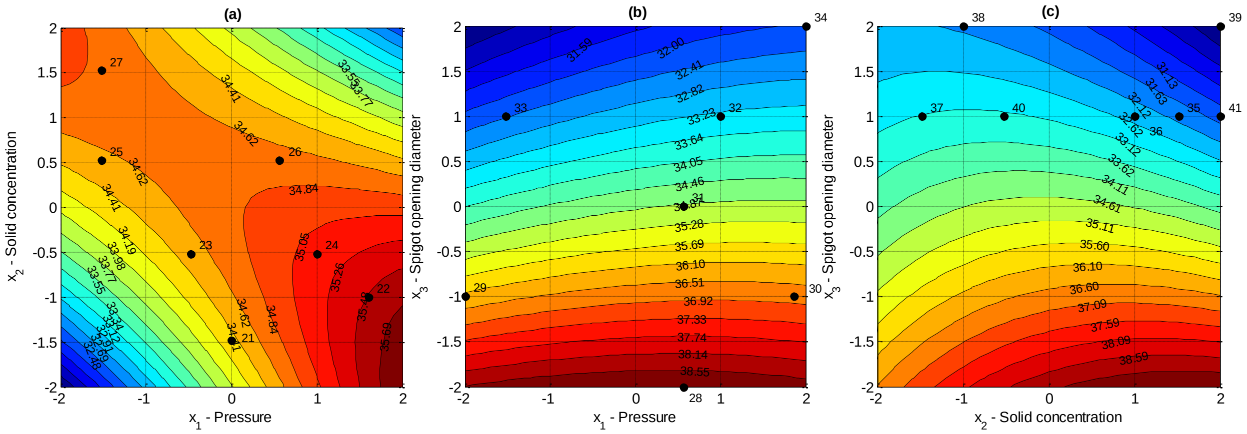


Figure 5-7: The cut-size contour plots depicting the additional experiments' conditions

Table 5-10: The coded and actual values of the newly identified experiments (experiment 21-35)

Run	Coded			Actual		
	x_1	x_2	x_3	P	ϕ	D_u
21	0	-1.48	0	73.5	0.95	25
22	1.61	-1	0	92	1.25	25
23	-0.48	-0.52	0	68	1.55	25
24	1	-0.52	0	85	1.55	25
25	-1.52	0.52	0	56	2.2	25
26	0.57	0.52	0	80	2.2	25
27	-1.52	1.52	0	56	2.825	25
28	0.57	0	-2	80	1.875	15
29	-2	0	-1	50.5	1.875	20
30	1.87	0	-1	95	1.875	20
31	0.57	0	0	80	1.875	25
32	1	0	1	85	1.875	30
33	-1.52	0	1	56	1.875	30
34	2	0	2	96.5	1.875	35
35	0	1.52	1	73.5	2.825	30

[32] Seven conditions per contour plot were identified.
[33] The new experiments were numbered from 21 to 41.

Table 5-11: The coded and actual values of the newly identified experiments (experiment 36-41)

Run	Coded			Actual		
	x_1	x_2	x_3	P	ϕ	D_u
36	0	1	1	73.5	2.5	30
37	0	-1.48	1	73.5	0.95	30
38	0	-1	2	73.5	1.25	35
39	0	2	2	73.5	3.125	35
40	0	-0.52	1	73.5	1.55	30
41	0	2	1	73.5	3.125	30

5.6.2 The sharpness of classification coefficient response

5.6.2.1 Sharpness of classification mathematical model

The second important response that was examined, is the sharpness of classification coefficient (m). The sharpness of classification coefficient is rarely discussed in literature. In order to gain some understanding in terms of which alternative hydrocyclone variables might affect the sharpness of classification, a mathematical model (much like the cut-size model) was developed. The same factors, experiments and methods that were used for the cut-size mathematical model were used to develop the sharpness of classification model. The least square function in MATLAB® was used once again in order to obtain the coefficients of the equation. These coefficients are given in Table 5-12. The equation coefficients were related to their corresponding terms and the mathematical model describing the sharpness of classification coefficient is shown in (5-11).

Table 5-12: Sharpness of classification equation coefficients

Coefficient		Corresponding term
Symbol	Value	
b_0	1.66636	constant
b_1	-0.01313	x_1
b_2	-0.00937	x_2
b_3	0.00813	x_3
b_4	0.00318	${x_1}^2$
b_5	0.00068	${x_2}^2$
b_6	0.01193	${x_3}^2$
b_7	0.00375	$x_1 x_2$
b_8	0.01125	$x_1 x_3$
b_9	0.01375	$x_2 x_3$

$$y_m = 1.66636 - 0.01313x_1 - 0.00937x_2 + 0.00813x_3 + 0.00318{x_1}^2 + 0.00068{x_2}^2 + 0.01193{x_3}^2 + 0.00375x_1x_2 + 0.01125x_1x_3 + 0.01375x_2x_3 \quad \textbf{(5-11)}$$

5.6.2.2 Summary of fit

Just as with the cut-size mathematical model, (5-11) will be evaluated by determining its fit. The estimated $\hat{y}_i$ values as obtained from the equation, are plotted against the actual values. The plot is shown in Figure 5-8 and the fit results are summarised in Table 5-13. The r-value was found to be 0.3984, which reflect poor positive correlation. MATLAB® calculated the R^2-value as -4.355. This indicates very poor fit of the model and the actual responses and these two coefficients already signifies that the model might be inadequate. To verify this assumption the Analysis of Variance is shown in 5.6.2.3.

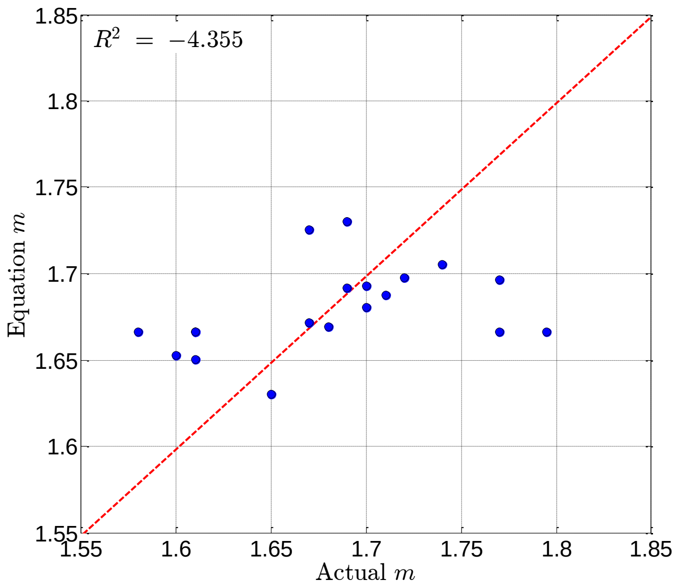


Figure 5-8: The actual versus the estimated sharpness of classification coefficient

Table 5-13: The summary of fit for the sharpness of classification coefficient response

Summary of fit	
r	0.3984
R^2	-4.3550
$\bar{R}^2$	-5.3591

5.6.2.3 Analysis of variance

With the r-value and R^2-value so poor, it is expected that the analysis of variance should confirm that the model is inadequate. Looking at the F-values of the regression and lack of fit and comparing them to the critical F-values for an $\alpha = 0.05$, it becomes clear that the regression is non-significant at the level 95 %. The ANOVA results are given in Table 5-14. Thus it is concluded that the mathematical model consisting of these specific factors are not representative and the model is deemed inadequate. The model can therefore not be used to provide any additional insight.

Table 5-14: ANOVA for the sharpness of classification mathematical model

Source	SS	df	MS	F	test F
Regression	0.011555	8	0.001444	0.256600	$F_{0.05}(8,11) = 2.95 > 0.26$
Residual	0.061915	11	0.005629		Non-significant @ the level 95%
- Lack of fit	0.017638	3	0.005879	0.929510	$F_{0.05}(3,7) = 4.35 > 0.93$
- Error	0.044277	7	0.006325		Non-significant @ the level 95%
Total	0.073594	19			

5.6.2.4 Experimental error of the sharpness of classification coefficient

Even though the mathematical model cannot be used, the experimental error will be calculated and used in other parts of this study. The experimental error of the sharpness of classification was calculated for the experiments and for the Malvern analysis. The CL was again chosen as 95 %. The same six experiments were used to determine the error of the experiments. It was found that the experimental error of the experiments was about 4.65 %. The Malvern analysis error was determined by using the random sample that was analysed three times. The expected experimental error of the Malvern is 2.06 %. The 4.65 % will be used throughout the study to indicate the acceptable variation of the sharpness of classification coefficient.

Table 5-15: Summary of experimental error of the sharpness of classification coefficient response

Parameter			Experiments	Malvern
Confidence Level	CL	-	95	95
Number of samples	n	-	6	3
Degrees of freedom	df	$n-1$	5	2
Average	$\bar{y}$	$\frac{1}{n}\sum_{i=1}^{n} y_i$	1.66	1.71
Standard deviation	s	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$	0.09401	0.02082
t-value	t_{n-1}	t critical value table	2.015	2.920
Error	e	$t_{n-1}\left(\frac{s}{\sqrt{n}}\right)$	0.077333	0.035094
Interval	-	$[\bar{y}-e;\bar{y}+e]$	[1.59;1.74]	[1.67;1.74]
Error in percentage	$e_{\%}$	Convert interval to %	4.65	2.06

Figure 5-9 shows the poor fit of the mathematical model. The error bars are also added to visually represent the acceptable variations that can be anticipated.

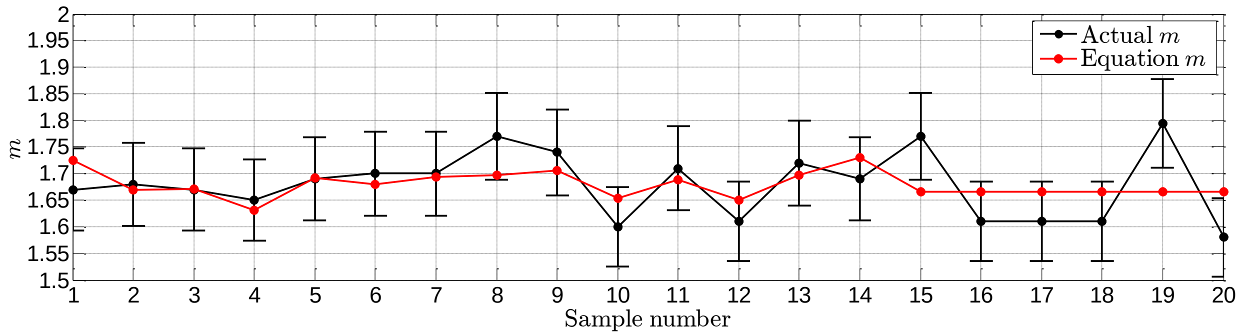


Figure 5-9: The actual and estimated sharpness of classification coefficient shown per sample

5.6.3 The feed flow rate response

5.6.3.1 Feed flow rate mathematical model

The first additional response that was investigated, is the feed flow rate (Q_i). The reason for examining this response was to establish a variable that could be used to verify whether the system was operating as it was expected to during experiments. In other words if the test rig is operating at a specific pressure, solid concentration and spigot configuration, the feed flow rate can be estimated and then be compared to the actual flow rate measurement.

Should there be a large deviation, the experiment would be discontinued and the system would be shut-off and checked for issues. An experimental run was only conducted when the expected feed flow rate coincided with the measured feed flow rate. The first twenty ED runs and their responses were recorded under the assumption that the system operated without any issues seeing as the mathematical model could only be developed and utilised after these runs were completed.

Using the least square function in MATLAB® the feed flow rate coefficients were calculated. The coefficients are tabulated in Table 5-16. The mathematical model obtained using these coefficients is given in (5-12).

Table 5-16: Feed flow rate equation coefficients

Coefficient		Corresponding term
Symbol	Value	
b_0	3.70223	constant
b_1	0.26919	x_1
b_2	0.05644	x_2
b_3	0.15069	x_3
b_4	-0.03226	${x_1}^2$
b_5	-0.02114	${x_2}^2$
b_6	0.00511	${x_3}^2$
b_7	0.04787	x_1x_2
b_8	0.00988	x_1x_3
b_9	-0.03038	x_2x_3

$$
\begin{aligned}
y_{Q_i} = 3.70223 + 0.26919x_1 + 0.05644x_2 + 0.15069x_3 - 0.03226{x_1}^2 \\
- 0.02114{x_2}^2 + 0.00511{x_3}^2 + 0.04787x_2 \\
+ 0.00988x_1x_3 - 0.03038x_2x_3
\end{aligned}
\qquad \textbf{(5-12)}
$$

5.6.3.2 Summary of fit

The estimated feed flow rate was determined by employing (5-12). The estimated values were then plotted against the actual feed flow rates as shown in Figure 5-10. The correlation coefficient (r) of 0.9842 shows very good positive correlation and the R^2-value that was found to be around 0.9675 indicates a good fit. To confirm that the mathematical model is adequate the analysis of variance is given in 5.6.3.3. The summary of fit of the feed flow rate model is also tabulated in Table 5-17.

Table 5-17: The summary of fit for the feed flow rate response

Summary of fit	
r	0.9842
R^2	0.9675
$\bar{R}^2$	0.9614

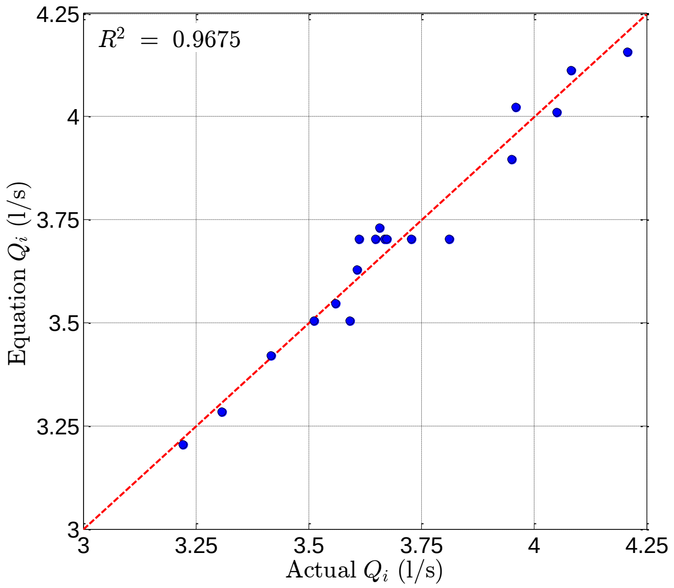


Figure 5-10: The actual feed flow rate versus the estimated feed flow rate

5.6.3.3 Analysis of variance

The ANOVA results for the feed flow rate mathematical model is summarised in Table 5-18. When comparing the critical F-values of the regression and lack of fit to the calculated F-values, it seen that the regression is seen as significant and the lack of fit is considered non-significant at an $\alpha = 0.05$. The model can thus be deemed adequate.

Table 5-18: ANOVA for the feed flow rate mathematical model

Source	SS	df	MS	F	test F
Regression	1.637	8	0.205	42.348	$F_{0.05}$(8,11) = 2.95 < 42.35
Residual	0.0532	11	0.00483		Significant @ the level 95%
- Lack of fit	0.0278	3	0.00927	2.558	$F_{0.05}$(3,7) = 4.35 > 2.56
- Error	0.0254	7	0.00362		Non-significant @ the level 95%
Total	1.691	19			

5.6.3.4 Experimental error of the feed flow rate

The experimental error of the feed flow rate was calculated and tabulated in Table 5-19. The expected experimental error was determined for the experiments only and was found to be around 1.56 %. Figure 5-11 shows the model estimated values and the actual measured feed flow rate per sample. The experimental error is also shown per sample.

Table 5-19: Summary of experimental error of the feed flow rate response

Parameter			Experiments
Confidence Level	CL	-	95
Number of samples	n	-	6
Degrees of freedom	df	$n-1$	5
Average	$\bar{y}$	$\frac{1}{n}\sum_{i=1}^{n} y_i$	3.691
Standard deviation	s	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$	0.0702
t-value	t_{n-1}	t critical value table	2.015
Error	e	$t_{n-1}\left(\frac{s}{\sqrt{n}}\right)$	0.0577
Interval	-	$[\bar{y}-e;\bar{y}+e]$	[3.633;3.749]
Error in percentage	$e_{\%}$	Convert interval to %	1.56

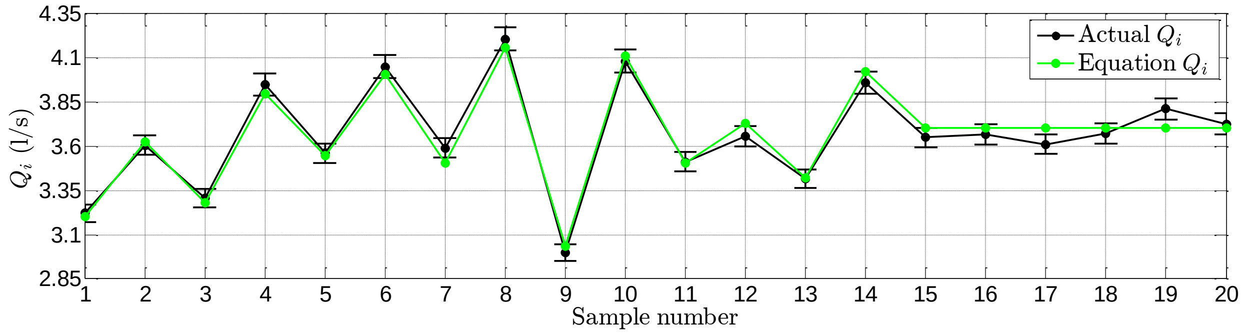


Figure 5-11: The actual and estimated feed flow rate shown per sample

5.6.3.5 Insight into the feed flow rate

The mathematical model adequately represents the feed flow rate of the system. The model can thus be used to firstly gain insight into what the effects might be of the factors on the feed flow rate response and to eventually validate the operation of the hydrocyclone during experimental runs. By employing the mathematical model given in (5-12), the three plots in Figure 5-12 were obtained. Figure 5-12 (a)[34] shows that the higher the solid concentration and the pressure become the higher the feed flow rate will become. Figure 5-12 (b)[35] depicts that an increase in pressure along with an increasing spigot size will substantially increase the feed flow rate. Figure 5-12 (c)[36] shows that the feed flow rate will increase as the spigot size and the solid concentration increase. The mathematical model seems to describe the effects of the flow rate as literature does [1], [2].

[34] The spigot size is kept constant at 25 mm.
[35] The solid concentration is fixed at 1.875 %.
[36] The pressure remains constant at 73.5 kPa.

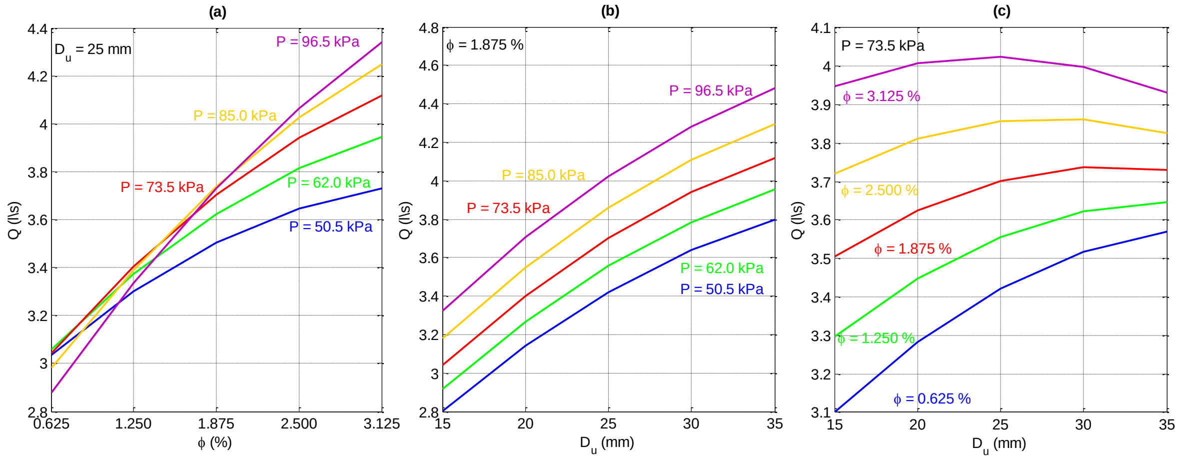


Figure 5-12: Individual effects of pressure, solid concentration and spigot opening diameter on the feed flow rate

5.6.4 The angle of discharge response

5.6.4.1 Angle of discharge mathematical model

The second additional response that was investigated, is the angle of discharge (ω). Only very recently have researchers been starting to investigate the angle of discharge. Literature suggest that the angle of discharge could be used as an additional parameter that conveys the operating health of a hydrocyclone [27]. So in order to determine how the three factors might influence the angle of discharge, a mathematical model for this response was also developed. Using the least square function in MATLAB® the coefficients were obtained and summarised in Table 5-20. The mathematical model is depicted in (5-13).

Table 5-20: Angle of discharge equation coefficients

Coefficient		Corresponding term
Symbol	Value	
b_0	44.72500	constant
b_1	-0.83125	x_1
b_2	0.09375	x_2
b_3	3.79375	x_3
b_4	0.35000	${x_1}^2$
b_5	0.20000	${x_2}^2$
b_6	2.91250	${x_3}^2$
b_7	0.01250	x_1x_2
b_8	-0.06250	x_1x_3
b_9	0.28750	x_2x_3

$$y_\omega = 44.72500 - 0.83125x_1 + 0.09375x_2 + 3.79375x_3 + 0.35000{x_1}^2 + 0.20000{x_2}^2 + 2.91250{x_3}^2 + 0.01250x_1x_2 - 0.06250x_1x_3 + 0.28750x_2x_3 \quad \textbf{(5-13)}$$

5.6.4.2 Summary of fit

By employing (5-13) the estimated angle of discharge values were obtained. These values were plotted against the actual angle of discharge values and the plot is shown in Figure 5-13. When assessing the r-value of 0.9651, the model shows a strong positive correlation. MATLAB® calculated the R^2-value and found it to be 0.9263. This signifies a model with a good fit. The summary of the fit is given in Table 5-21.

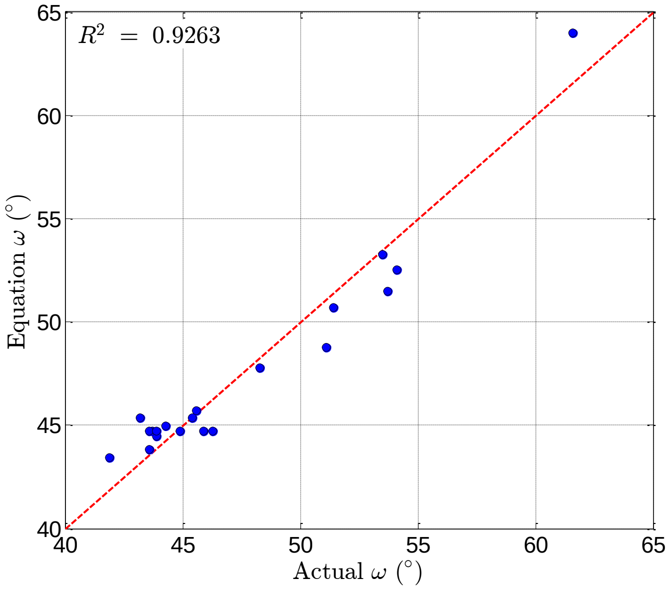


Figure 5-13: The actual angle of discharge versus the estimated angle of discharge

Table 5-21: The summary of fit for the angle of discharge response

Summary of fit	
r	0.9651
R^2	0.9263
$\bar{R}^2$	0.9125

5.6.4.3 Analysis of variance

To determine whether the mathematical model is adequate, an ANOVA is completed and given in Table 5-22. The critical F-value of the regression is much smaller than the computed F-value and was thus found to be significant at level 95 %. The lack of fit however, was found to be significant, subsequently rendering the mathematical model inadequate.

Table 5-22: ANOVA for the angle of discharge mathematical model

Source	SS	df	MS	F	test F
Regression	461.269	8	57.659	18.654	$F_{0.05}(8,11) = 2.95 < 18.65$
Residual	34.0006	11	3.09097		Significant @ the level 95%
- Lack of fit	27.112	3	9.0373	9.183	$F_{0.05}(3,7) = 4.35 < 9.18$
- Error	6.889	7	0.984		Significant @ the level 95%
Total	495.270	19			

5.6.4.4 Experimental error of the angle of discharge

The last step in the development of the angle of discharge mathematical model is assessing the experimental error. The calculated experimental error is summarised in Table 5-23. It shows that the expected experimental error of the angle of discharge is around 2.16 %. Again this calculated experimental error will be used throughout the rest of the study. The estimated angle of discharge and the actual values are plotted per sample and is shown in Figure 5-14 with the error bars indicating the acceptable experimental error that can be expected. The inadequate angle of discharge mathematical model cannot be used to gain further insight.

Table 5-23: Summary of experimental error of the angle of discharge response

Parameter			Experiments
Confidence Level	CL	-	95
Number of samples	n	-	6
Degrees of freedom	df	$n-1$	5
Average	$\bar{y}$	$\frac{1}{n}\sum_{i=1}^{n} y_i$	44.7
Standard deviation	s	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$	1.174
t-value	t_{n-1}	t critical value table	2.015
Error	e	$t_{n-1}\left(\frac{s}{\sqrt{n}}\right)$	0.966
	-	$[\bar{y}-e;\bar{y}+e]$	[43.7;45.7]
Error in percentage	$e_{\%}$	Convert interval to %	2.16

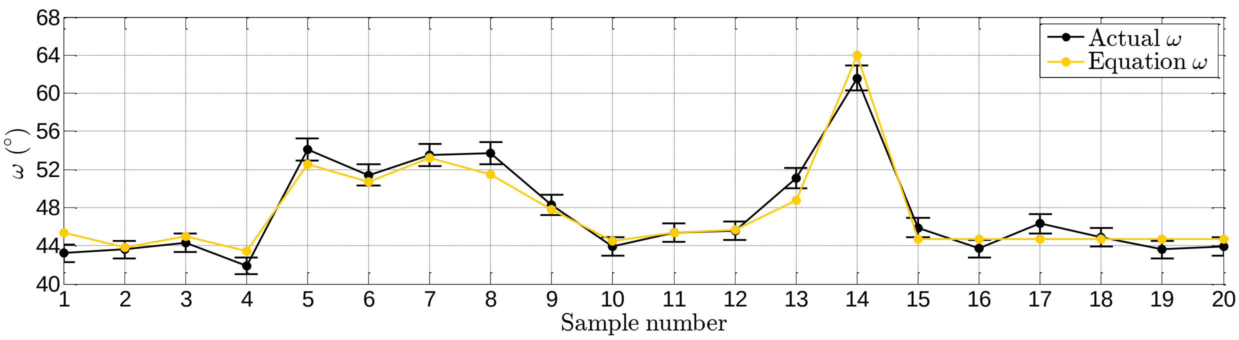


Figure 5-14: The actual and estimated angle of discharge shown per sample

5.7 Chapter summary

The CCRD approach was successfully employed to define comprehensive experiments and to use the recorded responses to ultimately obtain mathematical models. The cut-size mathematical model was found to be adequate and was utilised to identify twenty-one new experiments that included a wide range of cut-sizes and excluded repetitions. The sharpness of classification coefficient model proved to be inadequate, giving no information linking the response to the factors. It could thus not be used to gain further understanding of the performance or of the system. The two additional responses that were investigated were the feed flow rate and the angle of discharge. Only the feed flow rate mathematical model was found to be adequate. The feed flow rate mathematical model verified aspects that literature expressed and can be used to verify whether the hydrocyclone system is operating as expected. The angle of discharge mathematical model was deemed inadequate and could therefore also not be utilised.

CHAPTER 6 – ARTIFICIAL NEURAL NETWORK ESTIMATORS

6.1 Introduction

With the experimental data having been obtained and processed to be in a useable format, the models specified in Chapter 3 were developed. The chapter starts off by discussing the base ANN that was created as a first step in creating the specified models. The verification process that was used is described in detail. Next the validation procedures are applied, discussed and the appropriate results shown. By comparing the models that were developed, the best performing model is identified.

6.2 ANN development

6.2.1 The basic ANN

In order to transform the specified models into applicable ANNs, the ANN development diagram, as seen in Figure 6-1, was used as a reference. The diagram was compiled to ensure that the most important aspects of developing ANNs were identified. This was done to ensure that the necessary ANN properties are set-up accordingly during development. Chapter 3 discussed stages 1 and 2 of the development process, i.e. the problem being investigated and the model specifications. Stage 3, the model realisation, considers the various properties of the ANN. This includes the type of ANN, the training algorithm, the number of hidden layers, the number of hidden neurons, the activation functions of the layers, the layer processing functions, the sample division and the performance function. With so many specified models that needed to be developed, the best approach was to start off with a basic ANN. By creating a basic ANN most of the properties of the models stayed the same throughout; the different specified models obtained by only adjusting minor properties such as the number of neurons. MATLAB®'s Neural Network Toolbox command-line operations were utilised to develop the ANNs.

Table 6-1 summarises the basic ANN's properties. The type of ANN was chosen to be feed-forward backpropagation. Literature showed that these networks are the most widely used and should be considered the first type of network one should attempt to apply to a problem[37] [4],[28]. The backpropagation refers to the error method; error being calculated at the output layer and propagated back through the network to the input. The learning rule that dictates the backpropagation error was chosen to be Levenberg-Marquardt.

---

[37] Feed-forward backpropagation networks can be used for modelling, classification, predictions, control, data and image processing and pattern recognition [17].

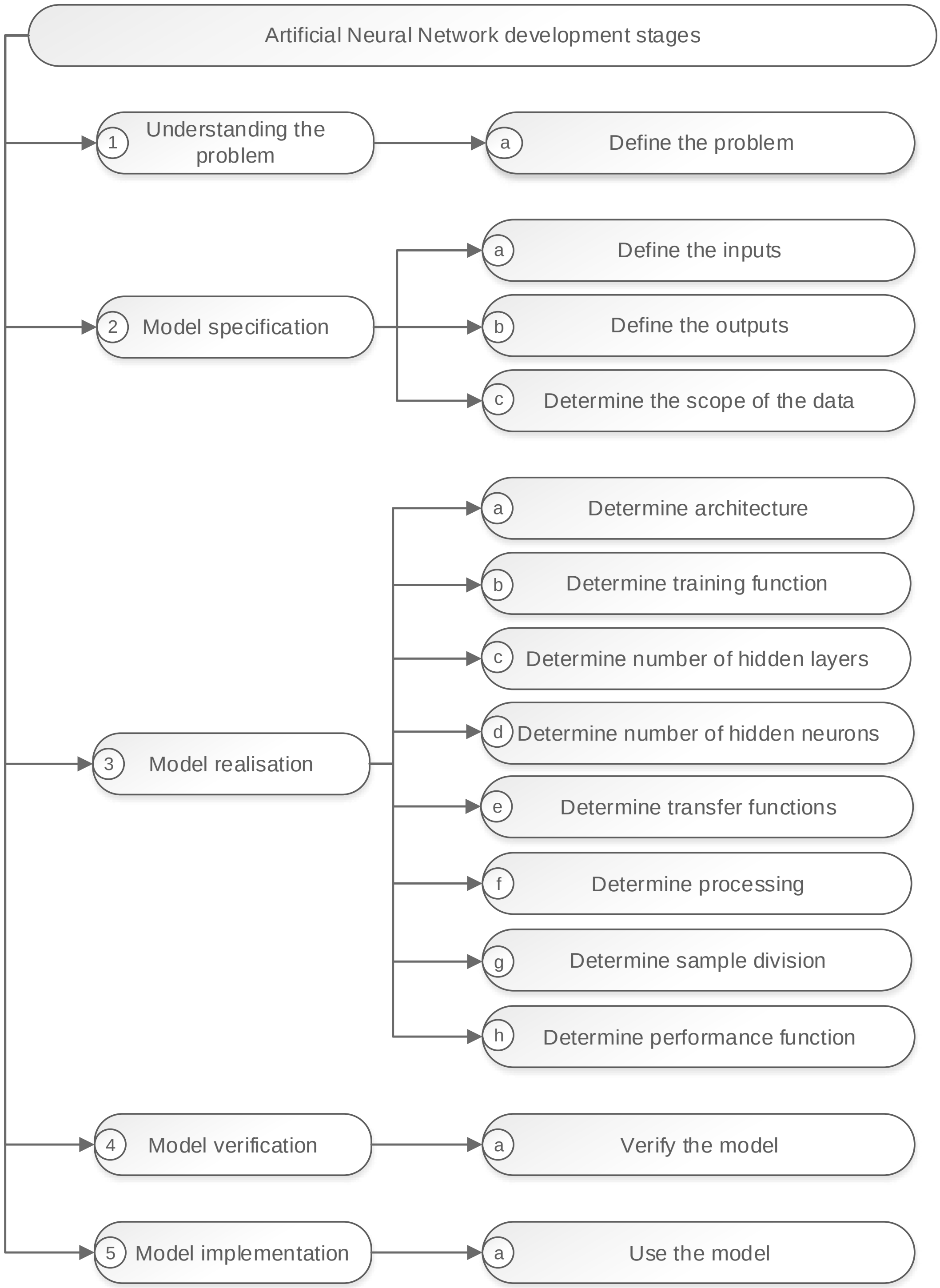


Figure 6-1: Development stages of an Artificial Neural Network [28]

Table 6-1: Summary of the base ANN's properties

Property	Description	Property	Description
Type of network	Feed-forward	Hidden layer activation function	tan-sigmoid
Training algorithm	Levenberg-Marquardt	Output layer activation function	linear
Training concept	Batch	Sample division	61:19.5:19.5
Number of hidden layers	1	Division mode	Random
Number of hidden neurons	1	Performance function	MSE
Processing functions	minmax		

Levenberg-Marquardt works especially well when applied to smaller networks and where memory usage is not a limitation. It is also said to be the best suited algorithm for function approximation [17]. These characteristics made Levenberg-Marquardt an ideal algorithm for the problem at hand. Initial trials of the base ANN indicated that the ideal number of hidden layers that worked well with the data and type of ANN, was one. Any more hidden layers seemed to over-fit the data. The default[38] number of neurons used, was determined by applying (6-1), where N indicates the number of neurons [29].

$$N = \frac{\frac{number\ of\ samples}{5} - outputs}{inputs + outputs + 1} \tag{6-1}$$

The hidden layer's activation function was tan-sigmoid and the output layer had a linear activation function as literature indicated this to be the best combination for function approximation applications. The architecture of the base ANN can be seen in Figure 6-2 depicting the input, one hidden layer and the output layer.

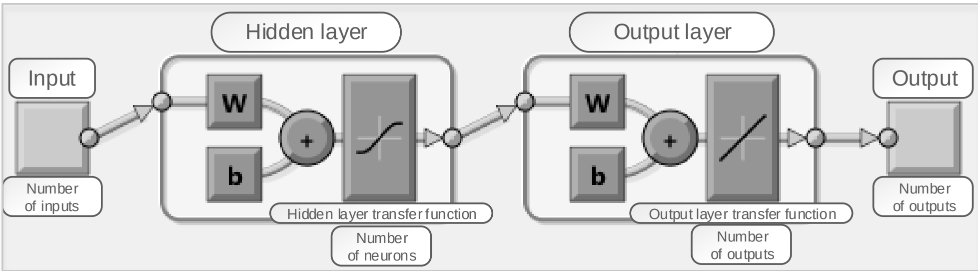


Figure 6-2: The base ANN's architecture (adapted) [17]

Some additional properties that were set were the processing functions of the inputs and outputs. The first processing property being the minmax property, used to transform the actual data values to a value between $[-1,1]$ and vice versa [17]. The removeconstantrow property was not set as no constant-value data were included as inputs. The 41 collected samples were divided into a data set for the training, validation and testing in a 61:19.5:19.5[39] ratio. This implies that 25 random samples were used to train the ANN, 8 random samples were used for validation and 8 random samples were completely withheld[40] from the ANN development. The base ANN's code is included in Appendix B.

[38] The default number of neurons is only set as a starting point, the number of neurons are adjusted during the verification loop as discussed in detail in section 6.2.2.
[39] With the limited number of samples, the ratio was found to work best.
[40] The samples that were withheld are referred to as unknown samples.

6.2.2 ANN training

In order to train the models that were specified, the basic ANN (as described in 6.2.1) was used along with a flow diagram as depicted in Figure 6-4. The main idea behind the flow diagram was to have a structured procedure and verification process to utilise during ANN training. The verification of a model is understandably one of the most important aspects of model development, as it evaluates the adequacy of the processes that went into developing the functional model [30]. In other words verification is the method of determining whether the conceptual model was accurately translated into a functional model [31].

The first task of the training procedure was to load the MATLAB® script containing the basic ANN. Next the model-appropriate inputs and targets were imported into the MATLAB® workspace as respective variables. The default number of neurons that were used with the first training attempt was always set to one. After the training of the ANN was completed the MSE performance graph[41] (see examples in Figure 6-3) was evaluated. It was expected that the MSE should decrease over each epoch that was executed, eventually delivering an acceptable minimum MSE. When working with Artificial Neural Networks it is hard to describe or define an acceptable MSE value.

By trial-and-error of the specific ANN a sense of the MSE characteristics are formed. When evaluating the two performance graphs in Figure 6-3 it is seen that both graphs' MSE decreased over the epochs as expected, but that the lowest MSE reached differs notably. This renders the larger MSE ANN of (b) unacceptable.

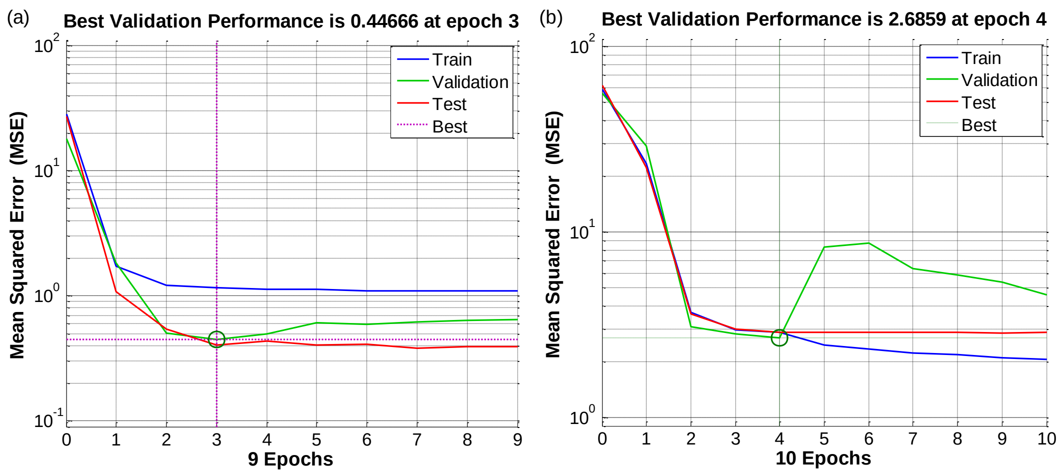


Figure 6-3: Examples of performance graphs (a) with acceptable MSE and (b) unacceptable MSE

[41] The Neural Network Toolbox automatically generates these performance graphs.

As seen from the flow diagram, should the MSE evaluation have indicated inadequacy, the training was repeated 10 times in order to find better performing ANNs. If the MSE still showed inadequacies after 10 attempts, the number of neurons was adjusted by adding one.

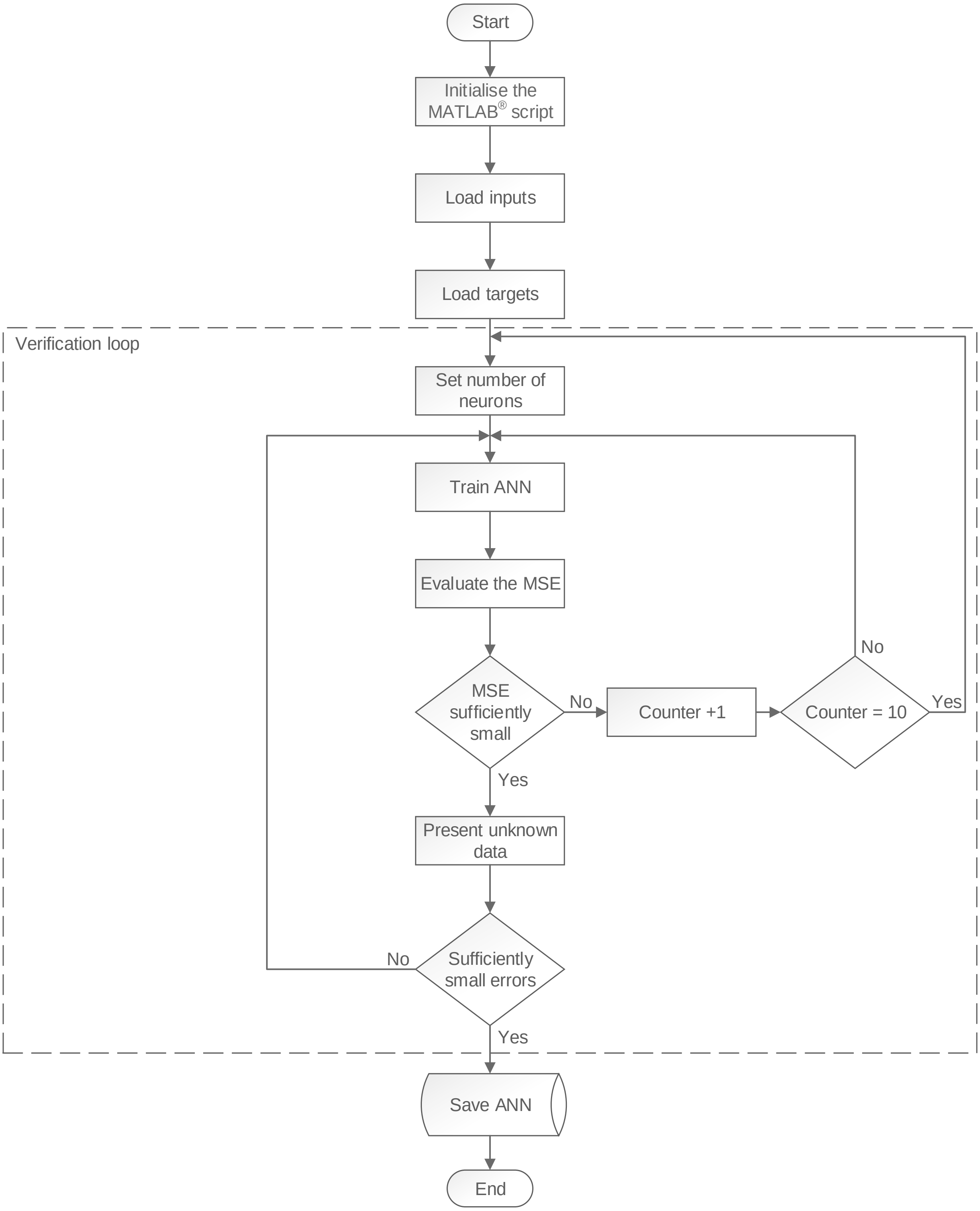


Figure 6-4: ANN development procedure and verification loop

Should the neuron addition deliver worse performing ANNs, the previous number of neurons were restored. If the MSE was found to be sufficiently small, the network was presented with unknown data, i.e. the 8 samples withheld from the ANN training (referred to as the testing data set). The MSE of the unknown samples were evaluated separately in order to determine whether the developed ANN generalised well for samples it had never seen. Only if the errors were sufficiently small[42] the developed ANN was saved. If the unknown samples delivered largely incorrect predictions, the ANN was retrained. By incorporating the described procedure, the nine models specified in Chapter 3 were developed and saved. Table 6-2 summarises the nine developed models' final specifications and ANN details.

Table 6-2: The developed models' specifications and ANN details

Model	Inputs		Output		Number of neurons
	#	Variables	#	Response	
0101	3	P, ϕ, D_u	1	d_{50}	2
0102	3	P, ϕ, D_u	1	m	4
0103	3	P, ϕ, D_u	2	d_{50}, m	4
0201	5	$P, \phi, D_u, Q_i, \omega$	1	d_{50}	4
0202	5	$P, \phi, D_u, Q_i, \omega$	1	m	6
0203	5	$P, \phi, D_u, Q_i, \omega$	2	d_{50}, m	5
0301	8	$P, \phi, D_u, Q_i, Q_o, Q_u, \rho_o, \omega$	1	d_{50}	3
0302	8	$P, \phi, D_u, Q_i, Q_o, Q_u, \rho_o, \omega$	1	m	4
0303	8	$P, \phi, D_u, Q_i, Q_o, Q_u, \rho_o, \omega$	2	d_{50}, m	4

6.2.3 Model adequacy

A second measure of validation was to determine whether the developed models were deemed statistically adequate. An Analysis of Variance (ANOVA) was completed for each one in order to determine whether the model was adequate or not. The calculated F-value of the models were compared to the appropriate critical F-value at an $\alpha = 0.05$. Should the calculated F-value have been larger than or equal to the critical F-value[43], the model was deemed significant and thereby considered adequate. Should the calculated F-value have been smaller than the critical F-value, the model was said to be inadequate and not representative of the system. Table 6-3 shows the ANOVA results for the cut-size and sharpness of classification. When considering the cut-size models, the calculated F-values for all the models were found to be larger than the critical F-values, affirming that all the cut-size models were deemed adequate. When examining the sharpness of classification models however, it was found that only Model 0102 could be deemed adequate, concluding the other models that were developed to be deficient and unusable.

---

[42] It must be emphasised that there does not exist a precise rule that stipulates an appropriate level of MSE. Only by trial-and-error can a perception be formed of the level of errors that are to be expected.

[43] The critical F-value is determined from the critical F-value tables in [13].

Table 6-3: Summary of ANOVA for the cut-size and sharpness of classification models

Source	df	SS	MS	F	test F
			Cut-size		
			Model 0101		
Model	3	89.469	29.823	34.656	$F_{0.05}(3,37) = 2.92 < 34.66$
Error	37	31.841	0.861		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0103		
Model	3	101.148	33.716	61.872	$F_{0.05}(3,37) = 2.92 < 61.87$
Error	37	20.162	0.545		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0201		
Model	5	83.291	16.658	15.335	$F_{0.05}(5,35) = 2.53 < 15.34$
Error	35	38.019	1.086		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0203		
Model	5	83.207	16.641	15.286	$F_{0.05}(5,35) = 2.53 < 15.29$
Error	35	38.103	1.089		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0301		
Model	8	102.699	12.837	22.073	$F_{0.05}(8,32) = 2.27 < 22.073$
Error	32	18.611	0.582		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0303		
Model	8	101.213	12.652	20.144	$F_{0.05}(8,32) = 2.27 < 20.14$
Error	32	20.097	0.628		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Sharpness of classification		
			Model 0102		
Model	3	0.051	0.017	4.696	$F_{0.05}(3,37) = 2.92 < 4.696$
Error	37	0.134	0.004		Significant @ the level 95%
Total	40	0.185			The model is deemed adequate
			Model 0103		
Model	3	-0.006	-0.002	-0.417	Non-conclusive
Error	37	0.192	0.005		
Total	40	0.185			
			Model 0202		
Model	5	0.025	0.005	1.068	$F_{0.05}(5,35) = 2.53 > 1.068$
Error	35	0.161	0.005		Non-significant @ the level 95%
Total	40	0.185			The model is deemed inadequate
			Model 0203		
Model	5	-0.011	-0.002	-0.403	Non-conclusive
Error	35	0.197	0.006		
Total	40	0.185			
			Model 0302		
Model	8	0.025	0.003	0.632	$F_{0.05}(8,32) = 2.27 > 0.63$
Error	32	0.160	0.005		Non-significant @ the level 95%
Total	40	0.185			The model is deemed inadequate
			Model 0303		
Model	8	-0.100	-0.013	-1.406	Non-conclusive
Error	32	0.286	0.009		
Total	40	0.185			

6.3 Model validation

6.3.1 Regression plots

With the models' adequacy known, the models were validated and compared with one another to determine which model delivered better predictions. Validation can be defined as the process of checking whether the model that was developed fulfils the original requirements and that it is accurate when employed for its intended use. In order to check the models' validity three different measures were employed, the first measure determining the coefficient of determination (R^2), the second measure visually inspecting the predicting capabilities and lastly by employing popular error metrics. The R^2-value is indicative of how well the predicted[44] values correspond to the actual[45] values. This is done by plotting the predicted values against the actual values. The dashed lines represent the best fit linear regression line between the actual and predicted values [17]. In general a higher R^2-value indicates a better model, but it is important to note that the R^2-value will always increase when more variables are included in the model. This renders the R^2-value impractical to use as is. This is where the adjusted R^2-value (denoted as $\bar{R}^2$) is used instead; it incorporates the R^2-value, but also adjusts for the number of variables included in the model. The $\bar{R}^2$-value of models can therefore be directly compared regardless of the number of variables. The results obtained are grouped under the two output parameters to better facilitate comparisons. The actual cut-size versus the predicted cut-size plots that were obtained for models 0101, 0103, 0201, 0203, 0301 and 0303 are depicted in Figure 6-5. The sharpness of classification models 0102, 0103, 0202, 0203, 0302 and 0303 plots are given in Figure 6-6. Table 6-4 summarises the models r-values, R^2-values and the $\bar{R}^2$-values and the performance ranking of each model. When comparing the $\bar{R}^2$-values of the cut-size models, it was found that Model 0103 performed the best and that the worst performing model with the lowest $\bar{R}^2$-value, was Model 0201. When looking at the sharpness of classification estimators the $\bar{R}^2$-values substantiate that the models were not adequate, since the negative $\bar{R}^2$-values imply very poor fit [32].

Table 6-4: Summary of the cut-size and sharpness of classification estimators' r, R^2 and $\bar{R}^2$

Cut-size					Sharpness of classification			
Model	Summary of fit			Ranking	Model	Summary of fit		
	r	R^2	$\bar{R}^2$			r	R^2	$\bar{R}^2$
0101	0.8619	0.6837	0.6581	4	0102	0.5571	-2.4610	-2.7416
0103	0.9163	0.8230	0.8086	1	0103	0.1288	-9.1170	-9.1170
0201	0.8412	0.5517	0.4877	6	0202	0.5291	-1.0710	-1.3669
0203	0.8430	0.6775	0.6314	5	0203	0.0728	-15.3100	-17.6400
0301	0.9211	0.8047	0.7559	3	0302	0.4027	-3.2210	-4.2763
0303	0.9150	0.8182	0.7728	2	0303	-0.2145	-6.3030	-8.1288

[44] Predicted values are often referred to as the outputs.
[45] Actual values are often referred to as the targets.

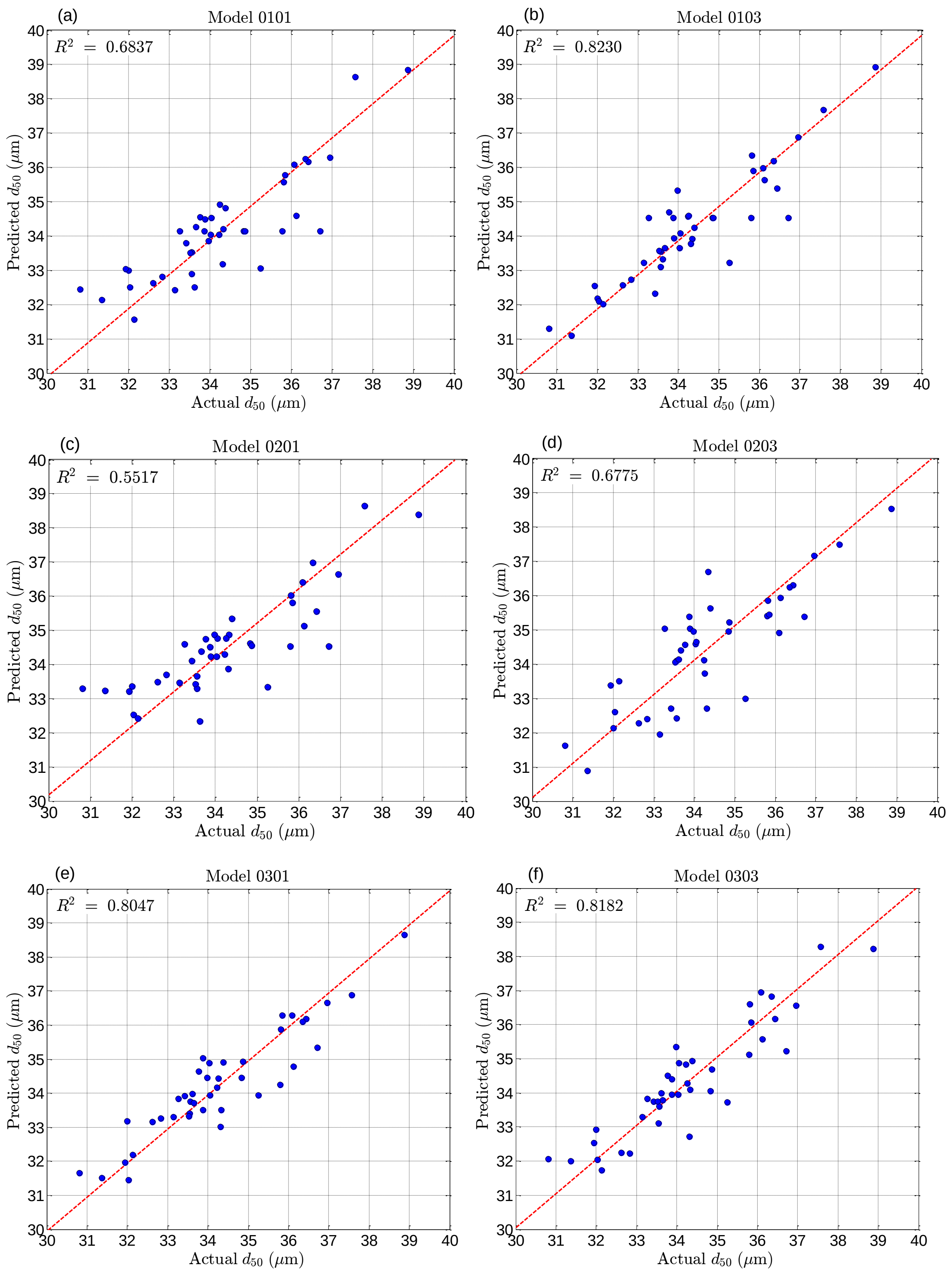


Figure 6-5: Actual versus predicted cut-size for (a) Model 0101, (b) Model 0103, (c) Model 0201, (d) Model 0203, (e) Model 0301 and (f) Model 0303

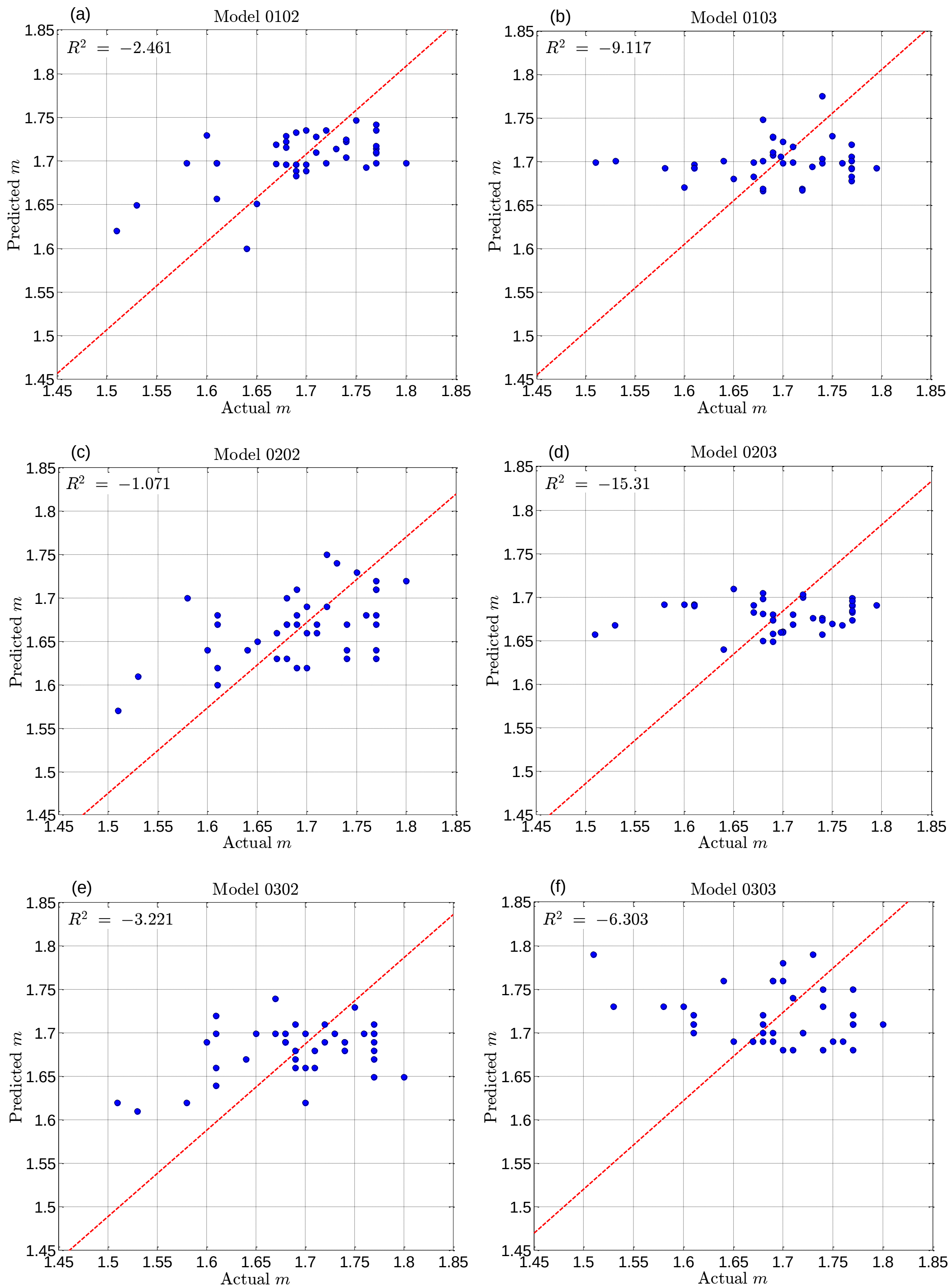


Figure 6-6: Actual versus predicted sharpness of classification of (a) Model 0102, (b) Model 0103, (c) Model 0202, (d) Model 0203, (e) Model 0302 and (f) Model 0303

6.3.2 Per sample plots

The second measure that was employed to check the validity of the models was to visually evaluate the prediction capabilities of the models. This was done by firstly plotting the actual values per sample for all forty-one samples for each one of the developed models; the models were once again grouped under the same output parameter. The calculated experimental error[46] was shown as error bars, indicating the acceptable errors that were to be expected. The models' predicted values were then plotted on the same graph. In this fashion it is easy to see what the actual value was supposed to be and what the models' predicted value was, also whether the predicted value falls within the experimental error interval. At the very least it was expected that the predicted values should follow the general trend of the actual data. The cut-size estimators' plots are depicted in Figure 6-7 and the sharpness of classification estimators' plots in Figure 6-8.

When observing the cut-size plots most of the models' predicted values seemed to correlate well to the actual values. Only very small differences are seen when comparing the models, but it was noted that models 0201 and 0203 correlated the worst of all the cut-size estimators, as it was expected to when referring to the $\bar{R}^2$-values. The sharpness of classification estimators' plots further promote the findings that most of the models were inadequate. Only Model 0102, which was found statistically adequate based on the ANOVA, seemed to vaguely follow the overall trend of the actual values. Based on the $\bar{R}^2$-value and the per sample plots of the sharpness of classification models, it is appropriate to say that the models were inadequate and could not be utilised; neither the one output ANNs nor the two output ANNs. Subsequently the focus was shifted to only investigate the cut-size estimators. With the sharpness of classification output parameter unusable the two-output ANNs serves no purpose, leaving only models 0101, 0201 and 0301. It should however be noted that the overall cut-size performance of the ANNs seem to improve when including the sharpness of classification as a second output (i.e. the XX03 models).

In order to view the remaining cut-size models' prediction capability of unknown samples in a separate fashion, the actual and predicted cut-size values of the withheld samples were also plotted per sample as shown in Figure 6-9. It is seen once again that Model 0201 delivered the worst correlation and that models 0101 and 0301 showed promising correspondence.

[46] The experimental error for both the cut-size and sharpness of classification were calculated in Chapter 5.

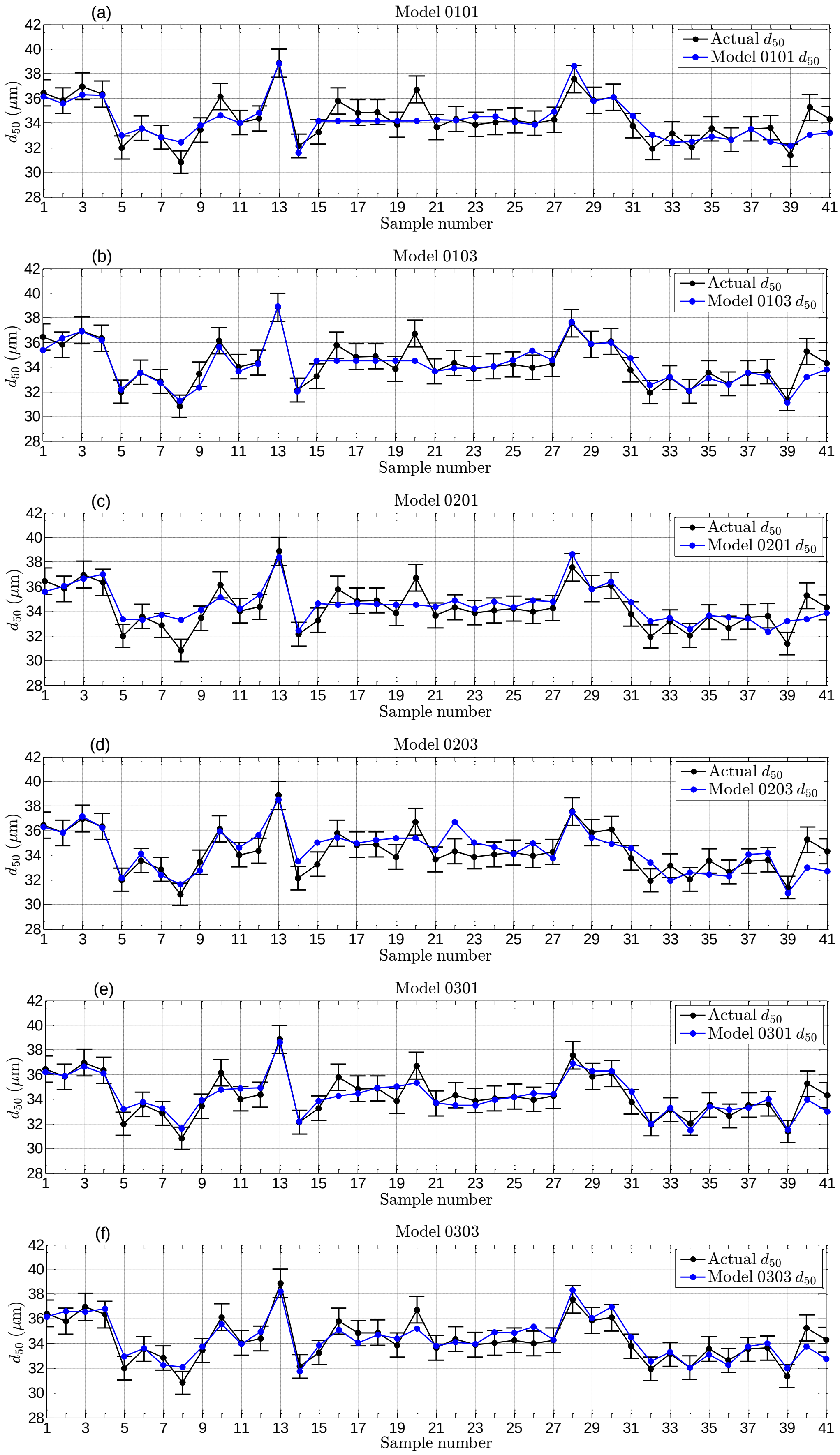


Figure 6-7: The actual and predicted cut-size of all the samples shown per sample for (a) Model 0101, (b) Model 0103, (c) Model 0201, (d) Model 0203, (e) Model 0301 and (f) Model 0303

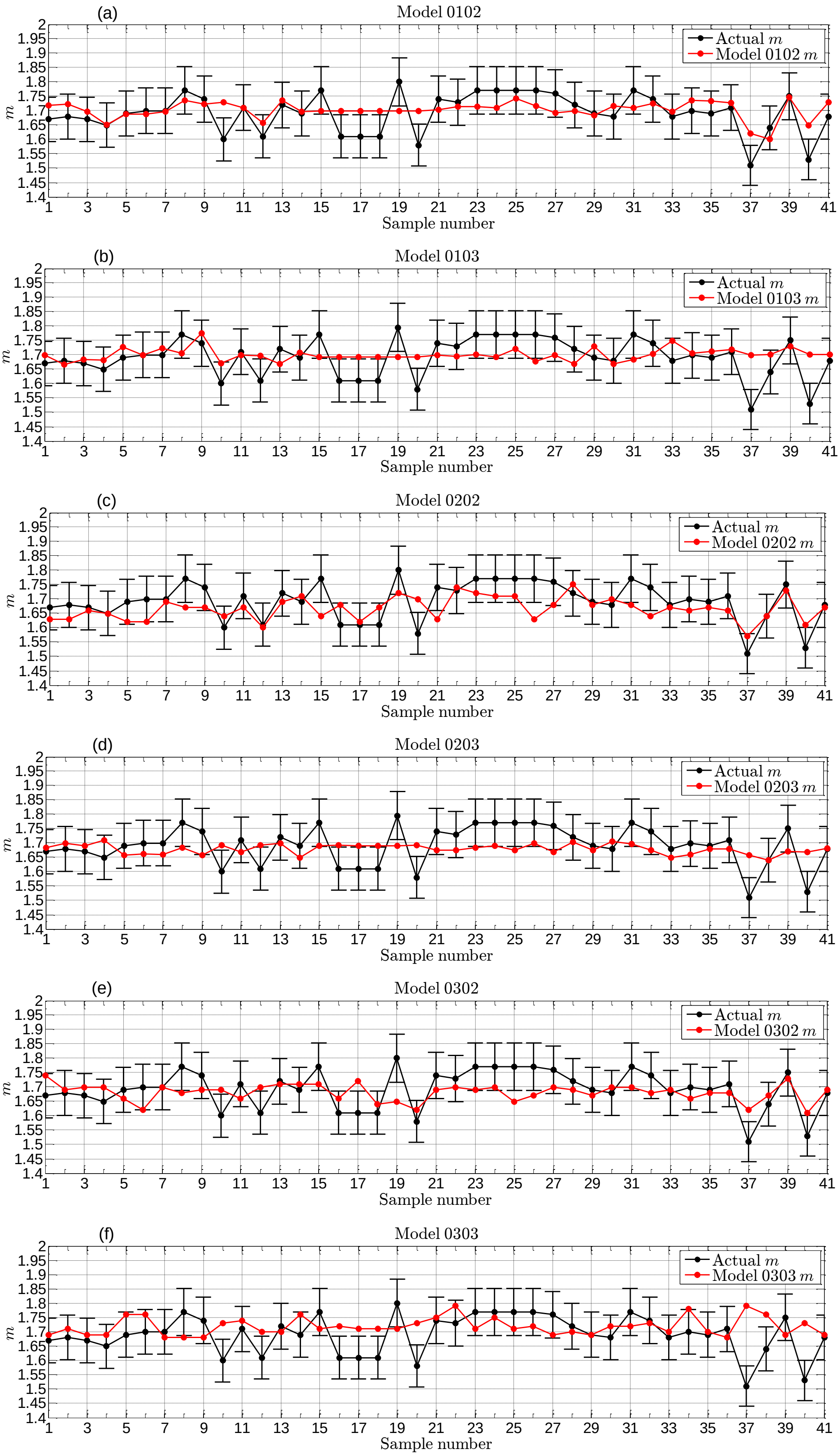


Figure 6-8: The actual and predicted sharpness of classification of all of the samples shown per sample for (a) Model 0102, (b) Model 0103, (c) Model 0202, (d) Model 0203, (e) Model 0302 and (f) Model 0303

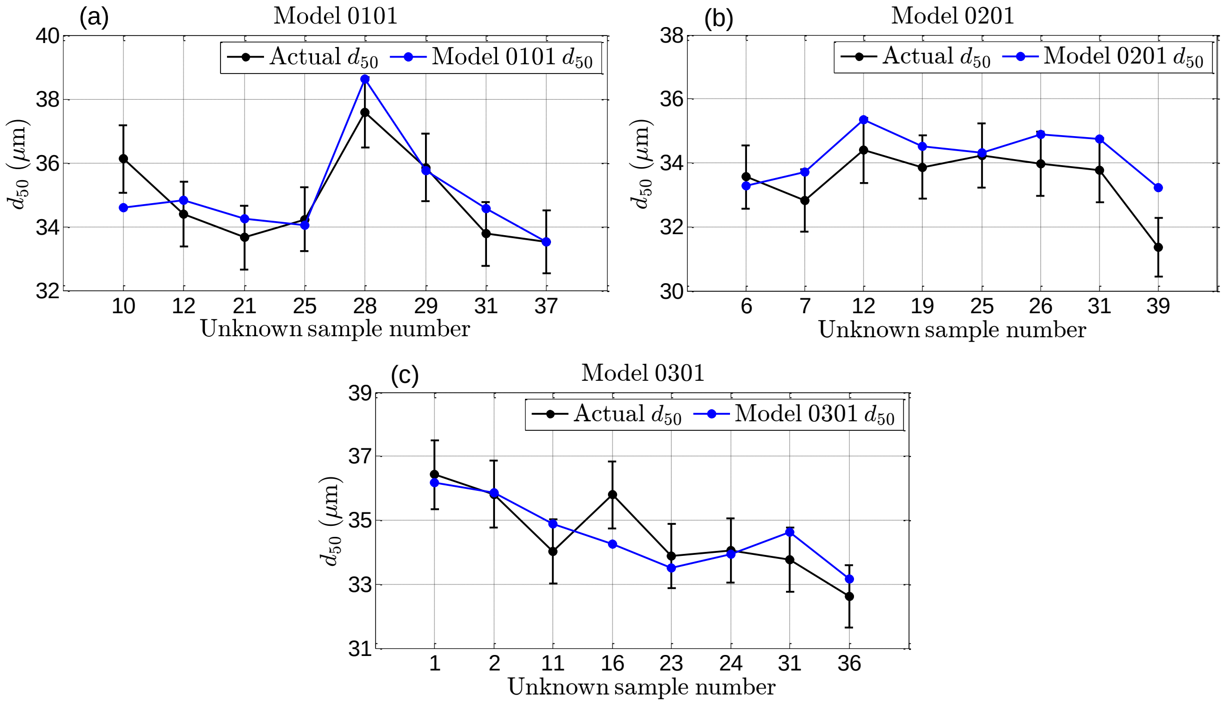


Figure 6-9: The actual and predicted cut-size for unknown samples shown per sample for (a) Model 0101, (b) Model 0201 and (c) Model 0301

6.3.3 Error metrics

To further investigate the cut-size models' performance to determine which one of the models would better predict the cut-size, two popular error metrics were calculated, assessed and compared. The metrics used were the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). The RMSE was determined by using (6-2) and the MAE calculated by using (6-3); this was done for all of the samples as well as the unknown samples.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \textbf{(6-2)}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \textbf{(6-3)}$$

When comparing the metrics, it is usually expected that the models' ranking order should be similar across all the errors. Therefore when evaluating the error metrics, they indicate that the models ranking best to worst performing, are Model 0301, Model 0101 and Model 0201. Figure 6-10 depicts these findings visually and Table 6-5 summarises the results, the calculated $\bar{R}^2$-value as previously shown and the models' rankings.

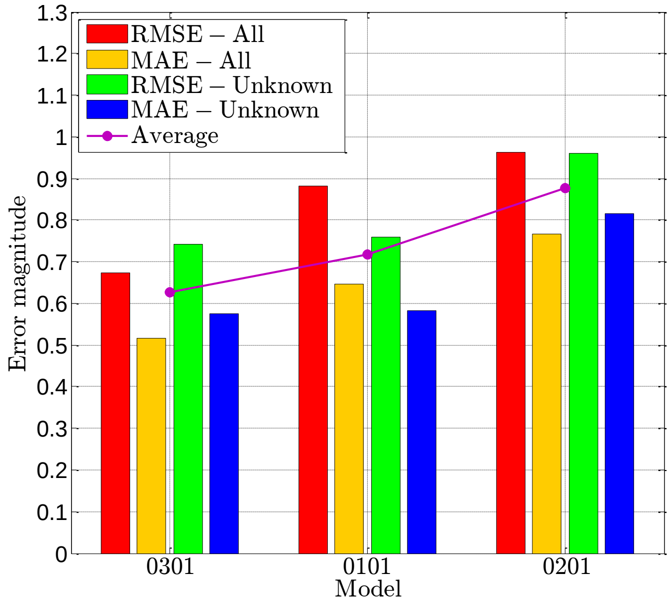


Figure 6-10: Visual representation of the error metrics of models 0301, 0101 and 0201

Table 6-5: Summary of error metrics for Model 0101, Model 0201 and Model 0301

Model	**Error metric**				**Average error**	$\bar{R}^2$	**Ranking**
	All samples		**Unknown samples**				
	RMSE	**MAE**	**RMSE**	**MAE**			
0101	0.8812	0.6460	0.7599	0.5826	0.7174	0.6671	2
0201	0.9630	0.7672	0.9592	0.8166	0.8765	0.5019	3
0301	0.6737	0.5170	0.7422	0.5757	0.6272	0.7633	1

When reviewing the validity of the models, it was found that the cut-size models were all valid and most of them sufficiently accurate. However only one of the sharpness of classification ANNs was found to be valid. This prompted the investigation to mainly pursue the cut-size ANNs, resulting in the sharpness of classification models being discarded (except Model 0102). With the sharpness of classification excluded, the three models that remained were the single-output cut-size ANNs. In order to determine which model would be the better model to incorporate, the models' error metrics were calculated and compared. The best performing ANN was found to be Model 0301 but it bettered Model 0101 only marginally. To determine whether all the unusual variables were needed for a well performing ANN such as Model 0301, additional cut-size ANNs were investigated. This aided in determining whether less of these unusual variables[47] might deliver comparable results for less of the effort.

[47] By incorporating less variables, the ANN architecture and the data acquisition are simplified.

6.4 Additional cut-size estimators

6.4.1 ANN training

As mentioned some additional estimators were developed in order to determine whether all eight of the inputs of Model 0301 were necessary. The five additional models that were specified are shown in Figure 6-11 depicting their name, inputs and output.

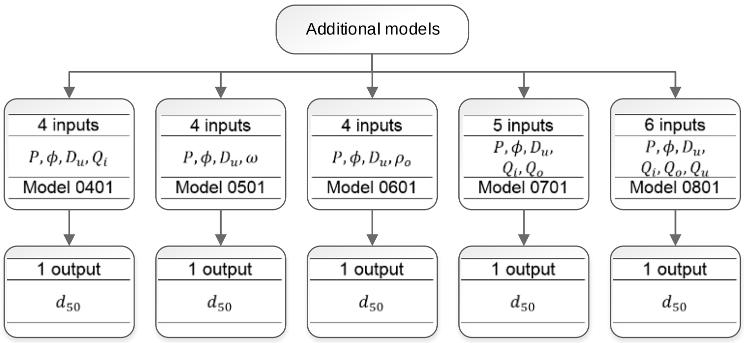


Figure 6-11: Summary of the additional models and their specifications

The first four-input ANN specified included the pressure (P), volumetric solid concentration (ϕ), spigot opening diameter (D_u) and the feed flow rate (Q_i). The second four-input ANN had the same base inputs (P, ϕ, D_u) and the angle of discharge (ω). These two models were evaluated in order to better understand the results obtained with Model 0201 ($P, \phi, D_u, Q_i, \omega$). Model 0601 investigated the effects of the overflow density (ρ_o). Literature, as discussed in Chapter 3, determines that the feed flow rate was deemed influential [3]. To inspect the inclusion of a second flow rate (overflow flow rate (Q_o)), a five-input ANN was developed (Model 0701). The last additional ANN, Model 0801, considered the effect of three flow rates namely the feed flow rate, the overflow flow rate and the underflow flow rate (Q_i, Q_o, Q_u). The same training and verification process as described in 6.2.2 was followed in order to develop and verify the additional ANNs. The specific details of the finally developed models are summarised in Table 6-6.

Table 6-6: Additional models' details and specifications

Model	Inputs		Output		Number of neurons
	#	Variables	#	Response	
0401	4	P, ϕ, D_u, Q_i	1	d_{50}	4
0501	4	P, ϕ, D_u, ω	1	d_{50}	2
0601	4	P, ϕ, D_u, ρ_o	1	d_{50}	3
0701	5	P, ϕ, D_u, Q_i, Q_o	1	d_{50}	3
0801	6	$P, \phi, D_u, Q_i, Q_o, Q_u$	1	d_{50}	5

6.4.2 Model adequacy

To check whether the models were statistically adequate, an ANOVA was completed for each one of the models. As discussed previously models that were found to have a larger calculated F-value than the identified critical F-value were deemed adequate. It is seen in Table 6-7 that all the additional models that were developed, were found to be adequate at an $\alpha = 0.05$. It could thus be concluded that the additional models were developed appropriately and could be put to use in determining their validity.

Table 6-7: Summary of ANOVA for the additional cut-size models

Source	df	SS	MS	F	test F
			Model 0401		
Model	4	93.603	23.401	30.404	$F_{0.05}$(4,36) = 2.69 < 30.40
Error	36	27.707	0.770		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0501		
Model	4	84.113	21.028	20.352	$F_{0.05}$(4,36) = 2.69 < 20.35
Error	36	37.197	1.033		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0601		
Model	4	83.707	20.927	20.034	$F_{0.05}$(4,36) = 2.69 < 20.034
Error	36	37.603	1.045		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0701		
Model	5	89.618	17.924	19.794	$F_{0.05}$(5,35) = 2.53 < 19.79
Error	35	31.692	0.905		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate
			Model 0801		
Model	6	100.173	16.695	26.856	$F_{0.05}$(6,34) = 2.42 < 28.86
Error	34	21.137	0.622		Significant @ the level 95%
Total	40	121.310			The model is deemed adequate

6.4.3 Model validation

6.4.3.1 Regression plots

With the models found to be suitable, the validation process could be implemented in order to determine the models' accuracy and performance ranking with regards to which variables made the most substantial contribution. The first measure was obtaining the regression plots of the models, shown in Figure 6-13, depicting the actual cut-size versus the predicted cut-size. The r-values, R^2-values and $\bar{R}^2$-values were obtained from the regression plots and are summarised in Table 6-8. When comparing the $\bar{R}^2$-values it seems that Model 0801, the model including the three flow rates, delivered the better predictions. This could be indicative that the combination of flow rates had a significant effect on the ANN performance and that it should be incorporated as far as possible. The least affecting variable seemed to be the overflow density used independently.

Table 6-8: Summary of the additional cut-size estimators' r, R^2 and $\bar{R}^2$

Model	Summary of fit			Ranking
	r	R^2	$\overline{R^2}$	
0401	0.8793	0.6853	0.6503	3
0501	0.8421	0.6135	0.5706	4
0601	0.8396	0.6131	0.5701	5
0701	0.8809	0.7630	0.7291	2
0801	0.9122	0.8124	0.7793	1

6.4.3.2 Per sample plots

To visually check the validity of the additional models, the actual and predicted cut-size values were plotted per sample, depicting all of the samples in Figure 6-14 and the unknown samples in Figure 6-15. The expected experimental error is also shown as vertical error bars. When assessing the actual and predicted values of all of the samples of the five models, only small variations are observed. The models that seemed to deviate the most were Model 0501 and Model 0601; as was expected when considering the $\bar{R}^2$-value. All of the models however seem to be satisfactorily accurate. The unknown samples show the models' prediction capabilities for the withheld samples, where models 0501 and 0801 seem to show the poorest correlation.

6.4.3.3 Error metrics

With the development of the additional models, it became evident that by including flow rates the ANN performance increased and that the individual inclusion of the overflow density did not have such a significant influence. In order to conclude whether eight inputs were really necessary to develop the best ANN, Model 0301 was compared to the additional models by considering their error metrics (see Figure 6-12 where 0101 and 0201 are also shown). When comparing the error metrics in Table 6-9 it is seen that Model 0301 still produced the lowest overall errors, only just surpassing the performance of Model 0401. This could indicate that a model employing only pressure, volumetric solid concentration, spigot opening diameter and feed flow rate as inputs, might be able to provide the same predictions as the eight input model provided that the $\bar{R}^2$-value of Model 0401 is improved.

Table 6-9: Summary of error metrics of the cut-size estimators

Model	Error metric				Average error	$\bar{R}^2$	Ranking
	All samples		Unknown samples				
	RMSE	MAE	RMSE	MAE			
0101	0.8812	0.6460	0.7599	0.5826	0.7174	0.6581	3
0201	0.9630	0.7672	0.9592	0.8166	0.8765	0.4877	6
0301	0.6737	0.5170	0.7422	0.5757	0.6272	0.7559	1
0401	0.8221	0.6384	0.6345	0.5672	0.6655	0.6503	2
0501	0.9525	0.7053	1.2630	1.0048	0.9814	0.5706	7
0601	0.9577	0.7946	1.1483	1.0362	0.9842	0.5701	8
0701	0.8792	0.6633	1.0108	0.7662	0.8298	0.7291	5
0801	0.7180	0.5687	1.0362	0.9279	0.8127	0.7793	4

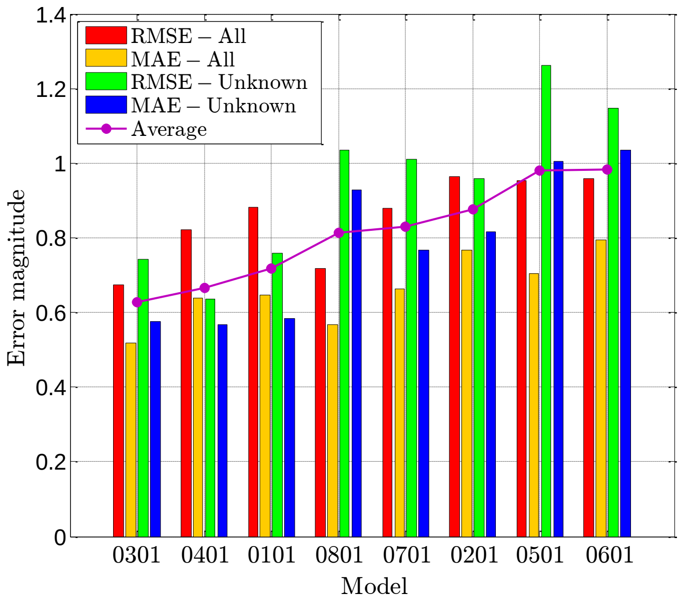


Figure 6-12: Visual representation of the error metrics of all the cut-size estimators

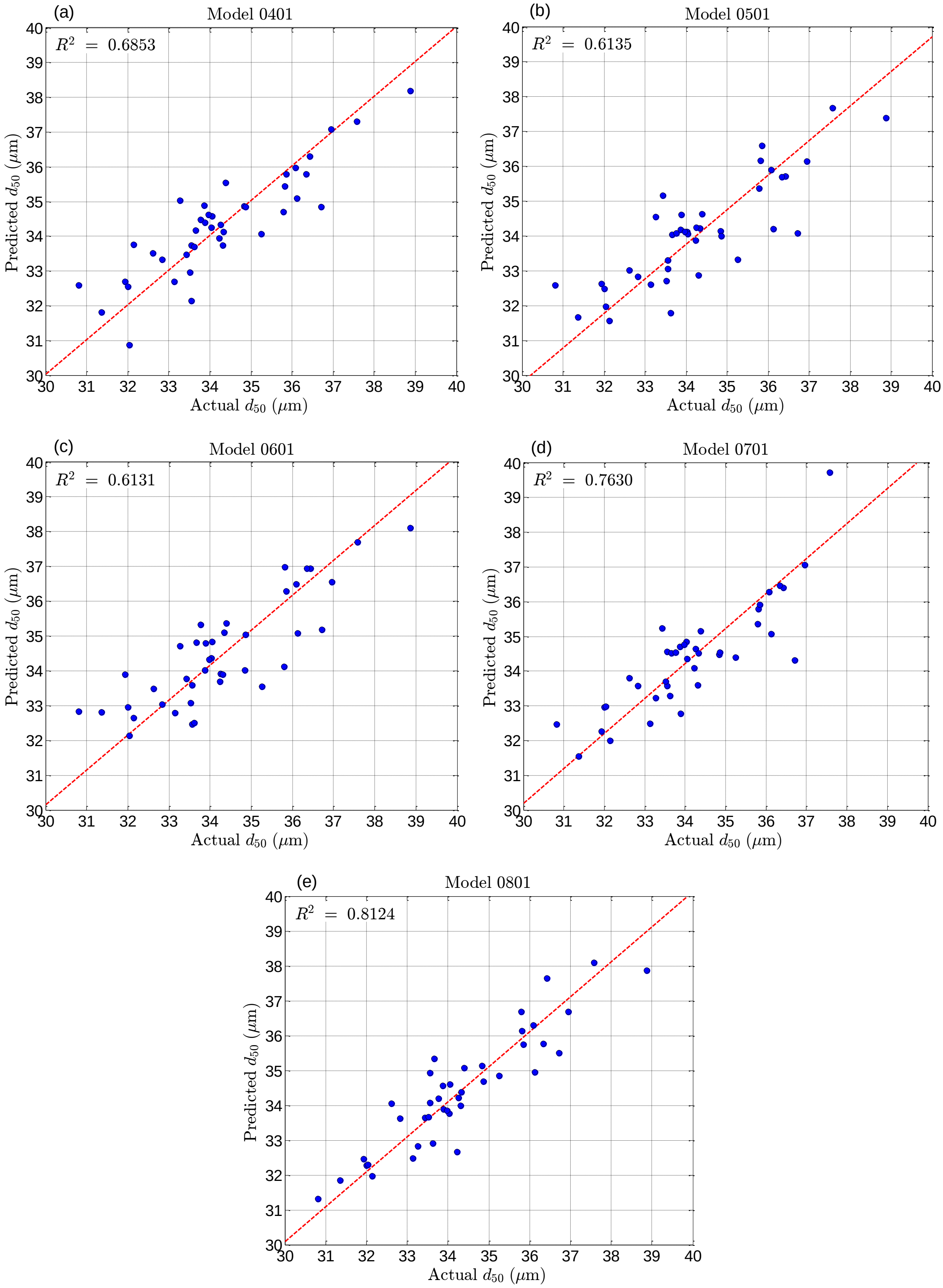


Figure 6-13: Actual versus predicted cut-size for (a) Model 0401, (b) Model 0501, (c) Model 0601, (d) Model 0701 and (e) Model 0801

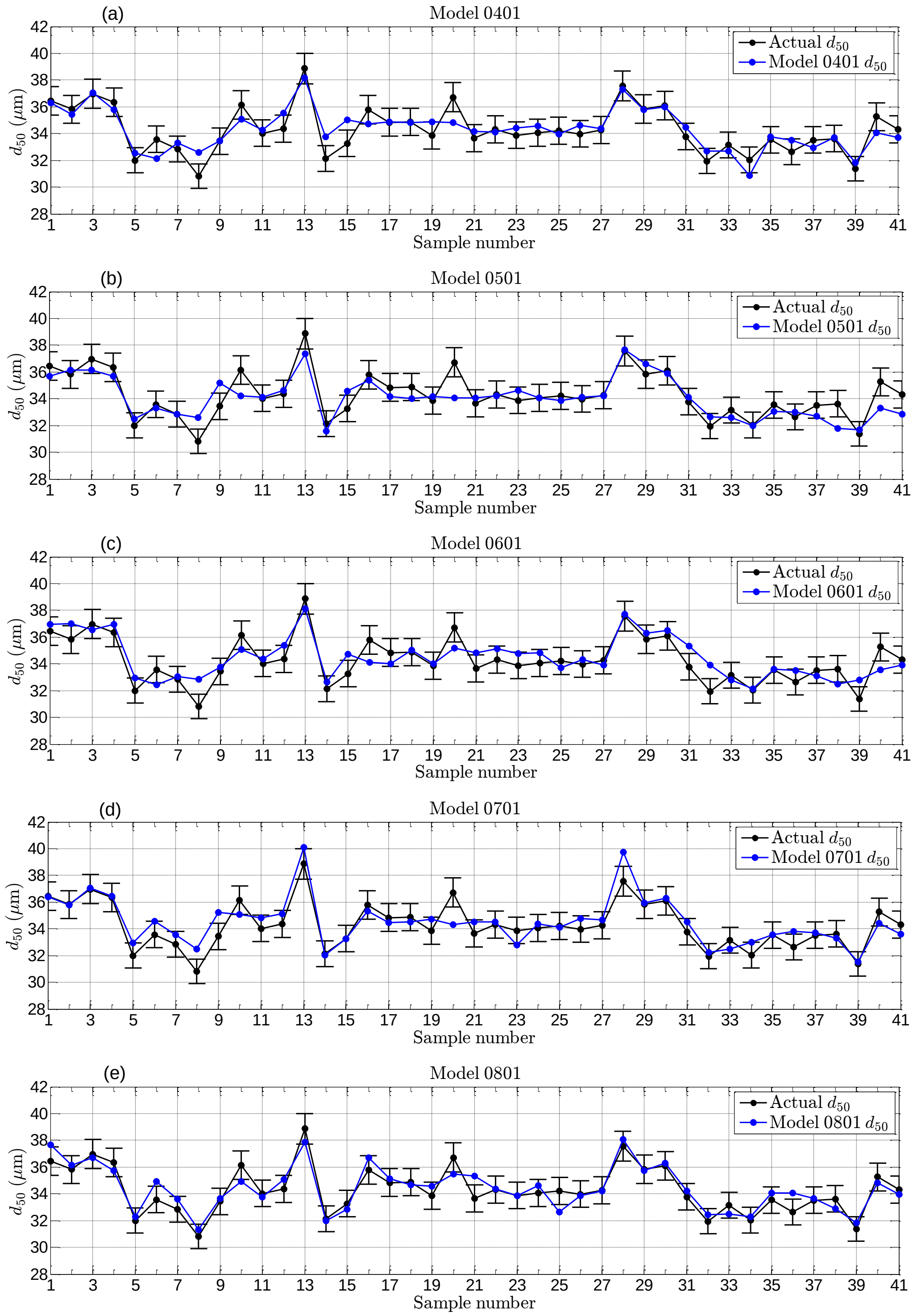


Figure 6-14: The actual and predicted cut-size for all samples shown per sample for (a) Model 0401, (b) Model 0501, (c) Model 0601, (d) Model 0701 and (e) Model 0801

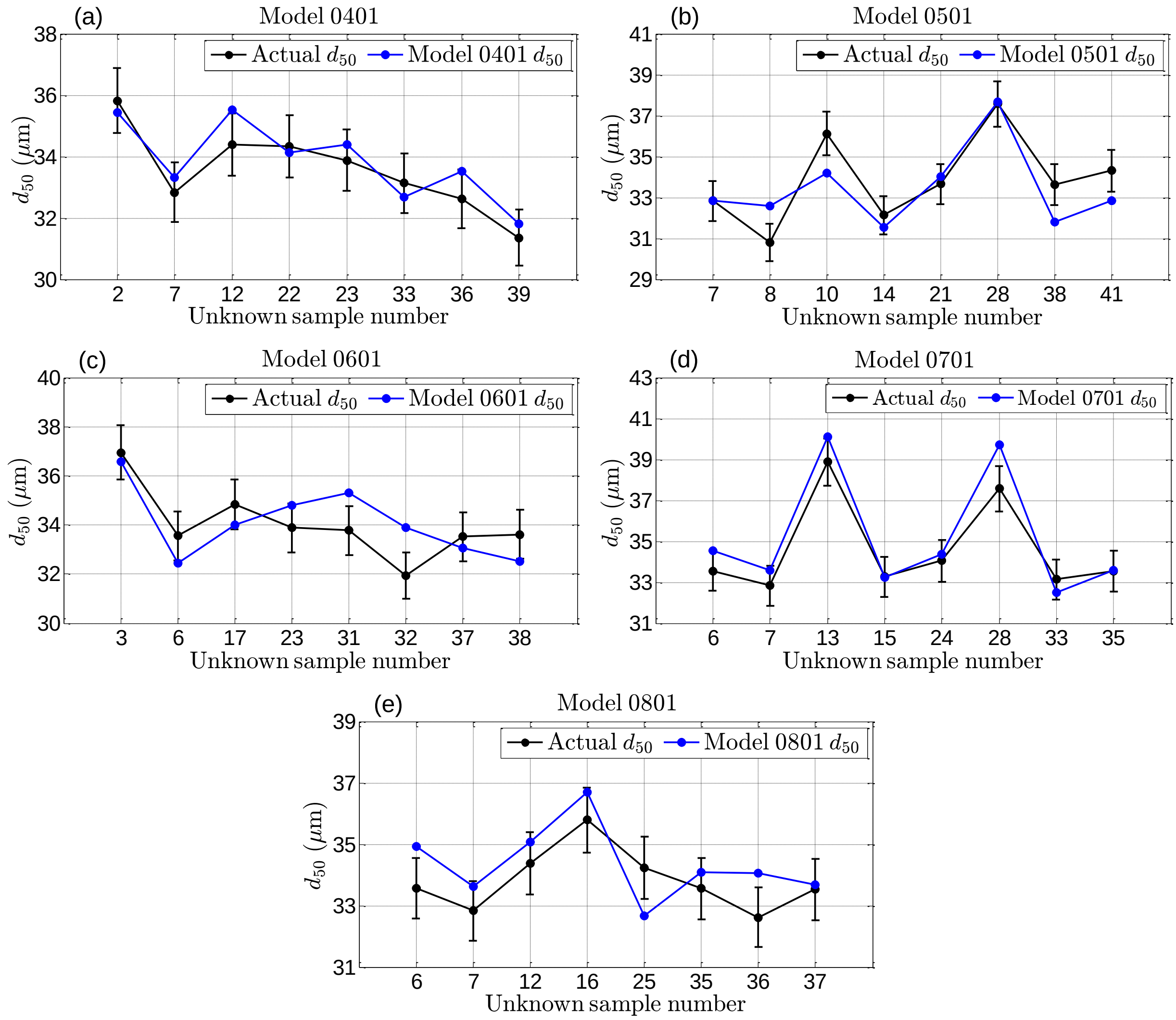


Figure 6-15: The actual and predicted cut-size for unknown samples shown per sample for (a) Model 0401, (b) Model 0501, (c) Model 0601, (d) Model 0701 and (e) Model 0801

6.5 Conclusion

By employing a basic ANN along with a verification flow diagram, the nine models that were specified in Chapter 3 were successfully transformed from conceptual models into functional models. The cut-size models were deemed adequate, but it was found that the sharpness of classification models were not statistically adequate (only Model 0102 was adequate). The inadequacy of the sharpness of classification models might indicate that the sharpness of classification coefficient cannot be comprehensively modelled by using the chosen variables; especially when referring to the variables Plitt incorporated in his mathematical model. The impractical sharpness of classification models were therefore discarded and the main focus shifted to the cut-size estimators. With the validation of the cut-size estimators, it became clear that the eight-input ANN (Model 0301) delivered the better predictions. In order to determine whether all eight of the inputs were necessary, additional cut-size estimators were developed and evaluated. It was found that by incorporating flow rates, the performance of the ANNs improved substantially. The individual inclusion of the angle of discharge (Model 0501) and overflow density (Model 0601) did not seem to improve the ANNs' performance. Even with the additional models that were developed, the eight-input model, Model 0301, seemed to deliver the best predictions. This might exhibit that combinations of the variables also influence the performance of the ANNs. The next step was to compare the predicting capabilities of the best ANN with that of Plitt's mathematical estimations as will be discussed in Chapter 7.

CHAPTER 7 – PLITT-FLINTOFF'S CONVENTIONAL MODEL

7.1 Introduction

The mathematical model L.R. Plitt published in 1976 is said to be one of the most popular conventional models that is used to estimate the cut-size and the sharpness of classification coefficient, amongst others [5]. The model is still used in some industries as an essential part of plant maintenance and in some hydrocyclone-design aspects. The researchers that worked on developing ANNs found that they delivered better predictions than that of the conventional models [4], [20]. In order to test whether this study's best performing ANN also delivered better predictions, a comparison was drawn and evaluated. The first step was to determine the Plitt-Flintoff mathematical model's calibration factors. Next the estimations using the mathematical model were evaluated by assessing their regression plots and the per sample estimations. Section 7.3 shows the comparison of the ANN approach and the Plitt-Flintoff, when looking at the prediction capabilities and the error metrics. From literature it is expected that the cut-size ANN should perform better than the Plitt-Flintoff model [4], [20].

7.2 The Plitt-Flintoff mathematical model

7.2.1 Calibration factors

In 1987 Flintoff et al. revised the mathematical model that Plitt developed. The first modification made was normalising the equation to a specific gravity of 2.6 in order to compensate for the effects of solid density. The equations were also adjusted by incorporating calibration factors that account for the application of different hydrocyclone systems. Equation (7-1) gives the revised form of the Plitt mathematical model for estimating the cut-size (d_{50}), (7-2) the volumetric flow split[48] (S) and (7-3) the sharpness of classification coefficient (m).

$$d_{50} = F_1 \frac{39.7 D_c^{0.46} D_i^{0.6} D_o^{1.21} \eta^{0.5} e^{0.063\phi}}{D_u^{0.71} h^{0.38} Q^{0.45} \left[\frac{(\rho_s - 1)}{1.6}\right]^k} \tag{7-1}$$

$$S = F_4 \frac{18.62 \rho_p^{0.24} \left(\frac{D_u}{D_o}\right)^{3.31} h^{0.54} \left({D_u}^2 + {D_o}^2\right)^{0.36} e^{0.0054\phi}}{D_c^{1.11} P^{0.24}} \tag{7-2}$$

[48] The volumetric flow split is the ratio of the volumetric flow of the underflow and the flow of the overflow. The sharpness of classification formula makes use of S.

$$m = F_2 1.94 \left(\frac{{D_c}^2 h}{Q}\right)^{0.15} e^{\left(\frac{-1.58S}{1+S}\right)} \quad \textbf{(7-3)}$$

By using the same operating and design variables of the different samples and by incorporating default calibration factors of 1, the Plitt-Flintoff values[49] were obtained. These estimated values usually differ from the actual experimental values and this is where the calibration factors become useful. In order to calculate the calibration factors (7-4) is used to determine the ratio of the actual experimental value and the value obtained when using the Plitt-Flintoff formulae.

$$Calibration\ factor = \frac{Actual\ value}{Plitt - Flintoff\ formula\ value} \quad \textbf{(7-4)}$$

Generally the factors are calculated by taking a few random samples and finding an average factor. In an attempt to improve the factors for this study though, the samples were grouped under their corresponding spigot size opening; as the spigot size opening is believed to be one of the most influential variables [3]. The factor for each of the grouped samples was calculated and an average obtained. These average factors were then assigned to each corresponding spigot opening diameter sample. Table 7-1 summarises the factor values (F_1, F_2 and F_4) as calculated per spigot opening diameter group. It seems that the cut-size and sharpness of classification factors tend to increase as the spigot opening diameter increases, where the volumetric flow split tend to decrease as the spigot opening diameter increases. These trends seem to indicate that the grouping approach might be beneficial.

Table 7-1: Factor values as assigned to the corresponding spigot opening diameters

Related parameter	Factor	Factor value per spigot opening diameter (mm)				
		15	20	25	30	35
d_{50}	F_1	2.372	2.745	3.112	3.435	3.885
m	F_2	0.999	1.067	1.039	1.133	1.273
S	F_4	0.808	0.377	0.151	0.101	0.074

7.2.2 Regression plots

By implementing the calibration factors, the Plitt-Flintoff values could be compared to the actual values. The first measure employed was once again obtaining the regression plots as shown in Figure 7-1. The cut-size described a R^2-value of only 0.2182 which is indicative of poor fit, with only a few samples reflecting correlation between estimated and actual values. When evaluating the sharpness of classification it was found that the R^2-value was -3.293 which suggests very

[49] This refers to the cut-size, volumetric flow split and the sharpness of classification.

poor fit; the regression plot resembling the plots found for the ANN models. These results already indicate that the Plitt-Flintoff formulae do not comprehensively represent the actual values.

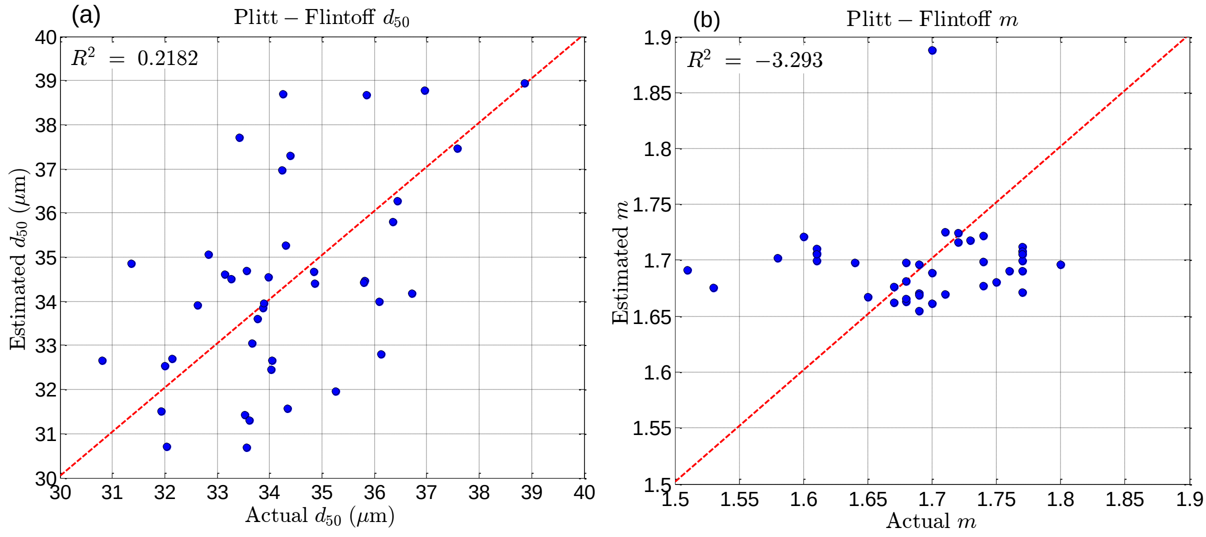
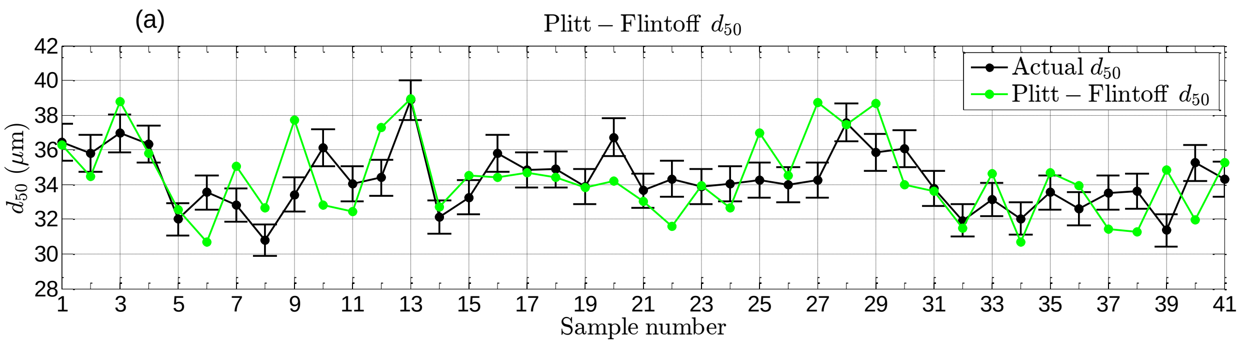


Figure 7-1: Actual versus estimated (a) cut-size and (b) sharpness of classification using the Plitt-Flintoff model

7.2.3 Per sample plots

In order to view the comparison of the estimated and the actual values, the values were plotted per sample as shown in Figure 7-2, once again showing the relevant experimental errors as error bars. The cut-size estimations show relatively poor correlation as expected, with only a few samples falling within the acceptable error margins. The sharpness of classification shows slightly better correspondence in terms of falling within the error interval, but does not seem to follow the trend[50] of the actual data. These results already indicated that the Plitt-Flintoff estimations are problematic. It should be noted that the grouped factor approach delivered exceedingly better per sample correspondence than just the factors that were determined by the average of a few random samples. To view the comparison, see Appendix C.



[50] The estimated sharpness of classification values do not seem to follow the actual values' trend but instead appear to only oscillate around a certain m-value.

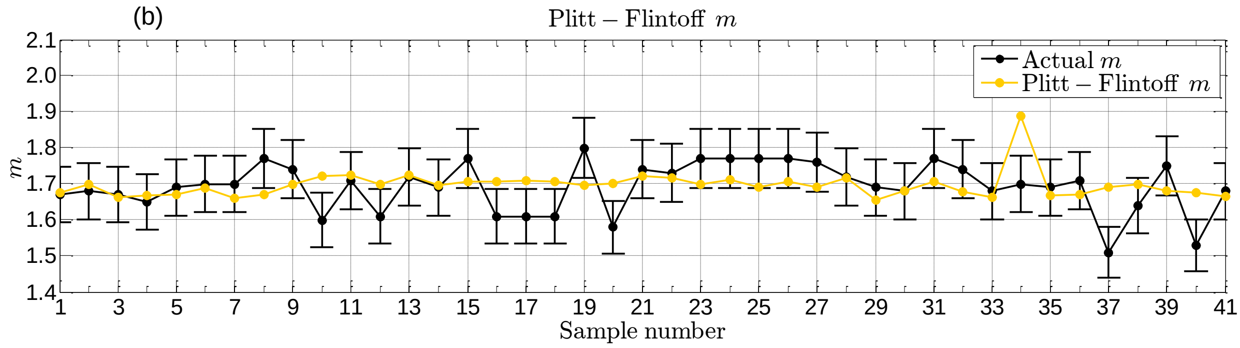


Figure 7-2: The actual and predicted (a) cut-size and (b) sharpness of classification for all samples shown per sample for the Plitt-Flintoff model

7.3 Comparison of the ANN and the Plitt-Flintoff models

To determine whether the best performing cut-size and sharpness of classification ANNs could deliver better predictions than the conventional models as developed by Plitt-Flintoff, the two modelling approaches were compared in terms of their per sample prediction capabilities and their error metrics. When assessing the cut-size per sample plot in Figure 7-3 (a) it can be seen that the predictions of Model 0301 performed significantly better than the Plitt-Flintoff estimations. Most of the ANN's predictions are well within the acceptable error intervals and seem to follow the general trend of the actual data exceedingly better than the Plitt-Flintoff model. Comparing the sharpness of classification as shown in Figure 7-3 (b), it is seen that the ANN and the Plitt-Flintoff model produced similar estimations. It is apparent that the error metrics will also not differ by much therefore indicating that the models' performance will be the same.

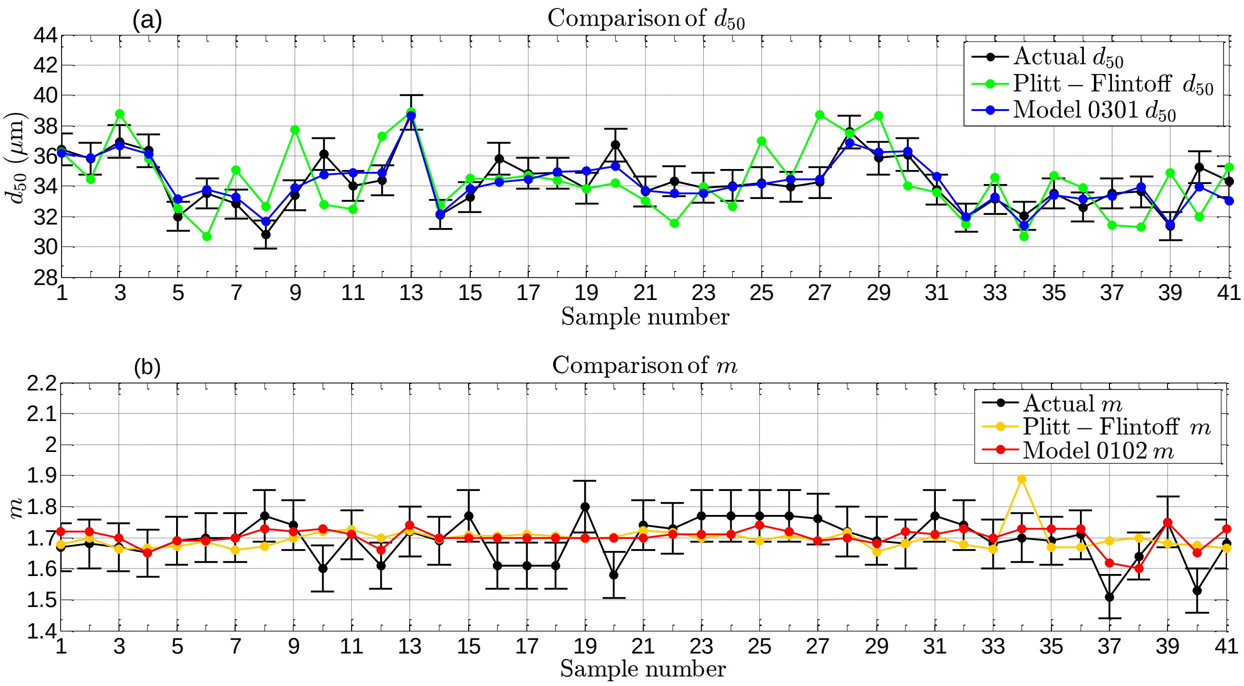


Figure 7-3: The actual and predicted (a) cut-size and (b) sharpness of classification for all samples shown per sample for the Plitt-Flintoff model and the ANN models

When comparing the unknown samples as depicted in Figure 7-4 (a), it is difficult to determine which model estimated the cut-size value better; it seems that Model 0301 might have performed better, but the error metrics will reveal the specifics thereof. Figure 7-4 (b) once again shows that the two models' estimations are alike; here also the error metrics will better pinpoint which model performed better, if at all.

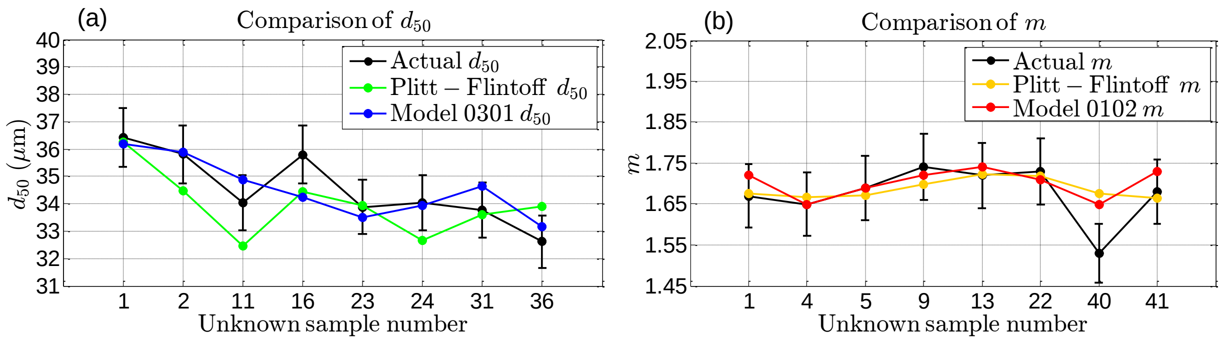


Figure 7-4: The actual and predicted (a) cut-size and (b) sharpness of classification for unknown samples shown per sample for the Plitt-Flintoff model and the ANN models

Figure 7-5 visually summarises the error metrics that were calculated for the ANNs and for the Plitt-Flintoff models. When comparing the cut-size errors it seems that Model 0301 delivers substantially smaller errors than the Plitt-Flintoff formula. This is true for all of the samples as well as the unknown samples. As the literature suggested these results also indicate that an ANN approach delivers better cut-size predictions with smaller overall errors compared to the Plitt-Flintoff conventional model. When evaluating the sharpness of classification however, Model 0102 delivers only marginally smaller erroneous predictions than those estimated with the Plitt-Flintoff formulae. The overall performance of the ANN demonstrates that the model is not representative of the system and is therefore not applicable.

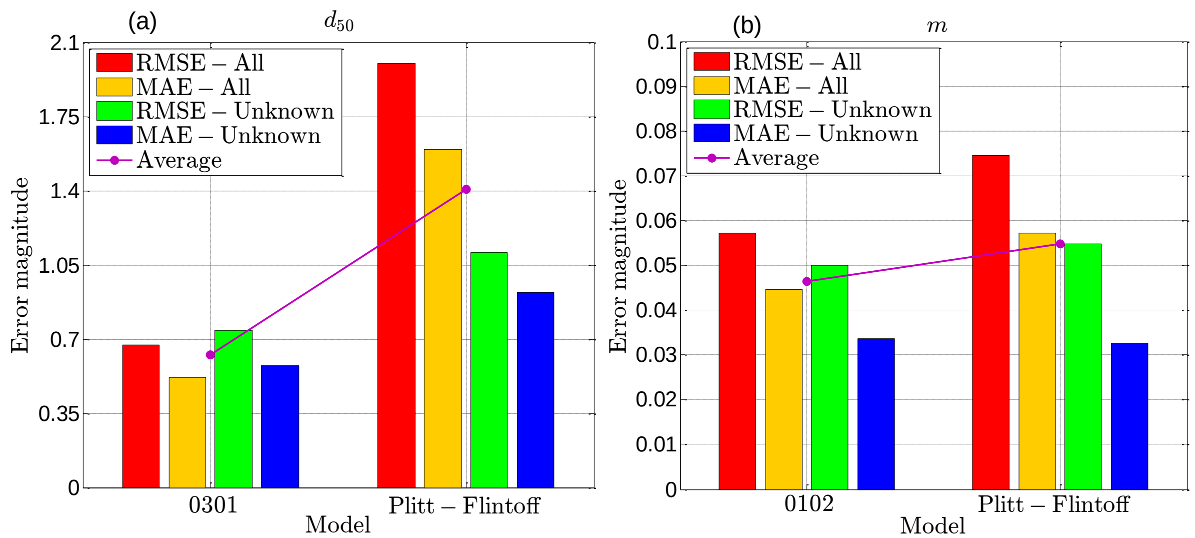


Figure 7-5: Visual representation of the error metrics of Plitt-Flintoff model and ANN models for (a) cut-size and (b) sharpness of classification

Table 7-2: Summary of error metrics of the Plitt-Flintoff model and ANN models for the cut-size and sharpness of classification

Model	Error metrics				Average error	$\bar{R}^2$	Ranking
	All samples		Unknown samples				
	RMSE	MAE	RMSE	MAE			
Cut-size							
Plitt-Flintoff	1.9989	1.5962	1.1087	0.9196	1.4059	0.0802	2
Model 0301	0.6737	0.5170	0.7422	0.5757	0.6272	0.7633	1
Sharpness of classification							
Plitt-Flintoff	0.07460	0.05720	0.05473	0.03256	0.05477	-3.6411	2
Model 0102	0.05720	0.04464	0.05003	0.03368	0.04639	-2.6432	1

7.4 Conclusion

The first order of business was to determine the calibration factors of the Plitt-Flintoff model in order to be able to draw a comparison between the performance of the developed ANNs and the mathematical estimations. It was found by grouping the samples under their spigot opening diameter the quality of the calibration factors could be enhanced. The estimations seem to have improved, but still delivered unreliable estimations. Some future work might investigate how to improve the calibration factors as the averaged values do not accurately convey all the information. When comparing the cut-size ANN (Model 0301) it was found that it performed better than the Plitt-Flintoff estimations, i.e. delivering better predictions with smaller errors. The sharpness of classification ANN's (Model 0102) predictions were almost identical to the estimations obtained via the Plitt-Flintoff method. The error metrics reflected the same order values. It would seem that the sharpness of classification is a difficult parameter to estimate.

CHAPTER 8 – CONCLUSION

8.1 Introduction

To conclude the dissertation, this final chapter provides a synopsis of some of the most important aspects and findings of the project, specifically focussing on the modelling results. Naturally the progress of the project highlighted pitfalls but it also identified interesting new and alternative investigations; these are given as recommendations in section 8.3. The closing paragraph marks the successful completion of the project and dissertation.

8.2 Conclusion

When working with hydrocyclones, it becomes evident that the two main performance indicating parameters, cut-size and sharpness of classification coefficient, cannot be observed or measured in real-time. To qualitatively describe these parameters some type of model is usually employed. This study's aim was to acquire experimental data and to use it to develop a model that could predict the cut-size and sharpness of classification. In order to collect the necessary data, the hydrocyclone test rig was instrumented accordingly and pre-designed samples were gathered. By incorporating Experimental Design (ED), not only were the experiments methodically planned, but valuable insights into the system were gained.

By extensively studying the literature, it was concluded that an Artificial Neural Network (ANN) approach showed the most favourable results in terms of the technique's flexibility, data integration advantages and level of intricacy. The most influential hydrocyclone variables, as determined from literature, were incorporated as model inputs. It was evident the flow rates were deemed extremely important. To ensure that an extensive study was done, nine different ANN models were developed; where some of the models had only one of the parameters as a single output and some models had both parameters as outputs. When evaluating the cut-size models' statistical adequacy, it was found that all of the models were adequate. Only one of the sharpness of classification models however was deemed adequate. The other models were all established to be faulty and unrepresentative. This inadequacy of the sharpness of classification results rendered the two-output models completely impractical, subsequently working with three cut-size models only (models 0101, 0201 and 0301). An additional investigation was presented after finding that the best performing cut-size model incorporated eight inputs. The investigation examined whether it was necessary to have all eight inputs, or if a smaller number of inputs might deliver the same results for less effort. Five new cut-size ANN models were therefore developed. After comparing all of the cut-size estimators, it was found that the eight input ANN still performed the best with the smallest average error of all the models, as shown in Table 8-1. This might be indicative that combinations of certain variables improve the ANNs prediction capabilities. In order to determine whether the developed ANNs would perform better than the conventional Plitt-Flintoff model, a comparison was drawn.

The usual method of calibrating the Plitt-Flintoff model is to calculate the average ratio of the Plitt-Flintoff estimation and the actual values. This average factor is then assigned to all the samples. It was however found that very poor estimations were obtained when employing this technique. The samples were therefore grouped under their respective spigot opening diameter, the average factor calculated for a few of the grouped samples and assigned to all the corresponding samples respectively. This improved the estimations substantially, but still gave some unreliable results.

With the Plitt-Flintoff model having been calibrated, the comparison between it and the developed ANNs could be drawn. It was found that the cut-size estimator (Model 0301) delivered about 124% better predictions than that of the Plitt-Flintoff model, when specifically comparing the average errors. The sharpness of classification ANN (model 0102) and the Plitt-Flintoff estimations were found to be very similar, having almost the same error-levels and seemingly oscillating around a specific m-value.

Table 8-1: Summary of the cut-size model results

Model	k	Summary of fit			All		Unknown		Average	Model 0301 improvement in %
		r	R^2	$\bar{R}^2$	RMSE	MAE	RMSE	MAE		
0301	8	0.862	0.684	0.658	0.674	0.517	0.742	0.576	0.627	-
0401	4	0.841	0.552	0.488	0.822	0.638	0.634	0.567	0.666	6.12
0101	3	0.921	0.805	0.756	0.881	0.646	0.760	0.583	0.717	14.40
0801	6	0.879	0.685	0.650	0.718	0.569	1.036	0.928	0.813	29.58
0701	5	0.842	0.614	0.571	0.879	0.663	1.011	0.766	0.830	32.32
0201	5	0.840	0.613	0.570	0.963	0.767	0.959	0.817	0.876	39.75
0501	4	0.881	0.763	0.729	0.952	0.705	1.263	1.005	0.981	56.48
0601	4	0.912	0.812	0.779	0.958	0.795	1.148	1.036	0.984	56.93
Plitt-Flintoff	10	0.467	0.218	0.080	1.999	1.596	1.109	0.920	1.406	124.16

8.3 Recommendations

A few recommendations that came forth from this study and that might be useful to future work include:

8.3.1 Experimental data acquisition

1. Ensure that all the practical test rigs, instruments and material are attained in a timely fashion. Delays in the construction of the hydrocyclone test rig postponed the progress of the entire project.
2. On this hydrocyclone test rig sampling was a two person task; it is therefore important to always have a project partner or assistant.

3. An assumption was made that each bag of silica used, had a uniform Particle Size Distribution (PSD), as indicated in the silica's experimental specifics sheet. A 150 samples were collected under this assumption. After analysing and processing about 90% (135) of those samples, major shifts[51] within the data were observed. By analysing the silica's PSD, it was found that every bag differed[52] from the other. With no way of knowing what the feed PSD of the samples were, they had to be discarded. Therefore to ensure the feed PSD is constant throughout, prepare a large quantity of well-mixed material.
4. The hydrocyclone test rig can be automated by installing a control valve and pressure transmitter[53] and connecting these instruments to a computer (for example via a CompactRIO).
5. Settling within the hydrocyclone tank was partly resolved by modifying the fine tune bypass valve so that it created a flow that mixed the silica into the slurry again. An improved mixing method might successfully minimise the settling.
6. Small spigots and high pressures direct the underflow-spray into the adjacent sampling bin (before sampling is commenced). When a shield is held between the sampling bin and the spray, the flow rate is altered. A permanent solution was not found, and therefore the sampling was initiated as quickly as possible.

8.3.2 Modelling

1. With the first 150 samples having been discarded, time constraints allowed for only 41 new samples to be collected. More samples will improve the training, testing and overall performance of the ANNs as using such small data sets can be partial to abnormalities. Additional testing samples will produce a clearer understanding of the ANNs' capabilities for samples the networks have never seen.
2. Report on the investigations and results when employing different types of ANNs with the same inputs.
3. Investigate and evaluate the results obtained when using the same variables[54] as Plitt-Flintoff as inputs to the sharpness of classification ANNs.
4. The grouping of the samples seem to improve some of the Plitt-Flintoff calibration issues. It might be interesting to investigate an alternative method of calculating the calibration factors.

---

[51] Large deviations in the cut-size values were observed for samples that were collected under the same operating conditions.
[52] Most bags only differed by a small degree, but a few were found that had large deviations.
[53] Both the control valve and the pressure transmitter were procured but not installed as yet.
[54] Variables such as feed flow rate (Q_i) and volumetric flow split (S).

5. The ANN model can be incorporated into a control scheme. The control application can be tested practically, should the control valve and pressure transmitter be installed appropriately.

8.4 Closure

The project was successful in developing an Artificial Neural Network model, based on experimentally obtained data, which could accurately predict a hydrocyclone's cut-size. The sharpness of classification coefficient could not be described successfully though.

Because changes within the test-rig occur over time, gradual deterioration of the ANNs' prediction accuracy will also occur. The ANNs as developed for this specific system will therefore not remain comprehensively applicable. When considering whether the ANN could be incorporated on industrial scale, it becomes evident that some additional work is first needed for the test-rig in order to rectify or compensate for the deviations. Only when the models are comprehensive in this regard and show promising results, can a strategy be devised to implement ANNs into industrial plants.

The specific skills that were honed include: mastering the research methodology, further improvement of professional and technical communication skills, working within a multidisciplinary environment and multi-faceted problem solving.

REFERENCES

[1] D. Bradley, *The Hydrocyclone - Bradley*. London: Pergamon Press Ltd., 1965.

[2] J. J. Cilliers, “Hydrocyclones for Particle Size Separation,” in *Particle Size Separation*, 2000, pp. 1819–1825.

[3] H. Eren and A. Gupta, “Instrumentation and on-line control of hydrocyclones,” in *International Conference on Control*, 1988, pp. 301–306.

[4] H. Eren, C. C. Fung, and K. W. Wong, “Artificial Neural Networks in Estimation of Hydrocyclone Parameter 50 with Unusual Input Variables,” *IEEE Trans. Instrum. Meas.*, vol. 46, no. 4, pp. 908–912, 1997.

[5] L. R. Plitt, “A Mathematical Model of the Hydroyclone Classifier,” *CIM Bull. Miner. Process.*, vol. 69, no. December, pp. 114–123, 1976.

[6] B. A. Wills and T. J. Napier-Munn, *Wills’ Mineral Processing Technology*, 7th Editio., no. October. Oxford: Butterworth-Heinemann, 2006.

[7] E. G. Kelly and D. J. Spottiswood, *Introduction to Mineral Processing*. Australia: John Wiley & Sons, 1982.

[8] M. Frachon and J. J. Cilliers, “A general model for hydrocyclone partition curves,” *Chem. Eng. J.*, vol. 73, no. February, pp. 53–59, 1999.

[9] D. F. Kelsall, “A Study of Motion of Solid Particles in a Hydraulic Cyclone,” *Chem. Eng. Res. Des.*, vol. 30, no. a, pp. 87–108, 1952.

[10] K. Nageswararao, D. . Wiseman, and T. . Napier-Munn, “Two empirical hydrocyclone models revisited,” *Miner. Eng.*, vol. 17, no. 5, pp. 671–687, May 2004.

[11] H. Eren, C. Fung, and K. Wong, “An application of artificial neural network for prediction of densities and particle size distributions in mineral processing industry,” in *IEEE Instrumentation and measurement Technical Conference*, 1997, pp. 1118–1121.

[12] J. J. Cilliers, R. C. Austin, and J. P. Tucker, “An evaluation of formal experimental design procedures,” in *International Conference on Hydrocyclones*, 1992, pp. 31–49.

[13] J. Devore and N. Farnum, *Applied Statistics for Engineers and Scientists*, 2nd ed. Brooks/Cole, Cengage, 2005.

[14] T. Chai, R. R. Draxler, and C. Prediction, “Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature,” no. 2005, pp. 1247–1250, 2014.

[15] M. N. Cirstea, A. Dinu, J. G. Khor, and M. McCormick, *Neural and Fuzzy Logic Control of Drives and Power Systems*. 2002.

[16] M. Nørgraad, O. Ravn, N. K. Poulsen, and L. K. Hansen, *Neural Networks for Modelling and Control of Dynamic Systems*. Springer, 2000.

[17] M. H. Beale, M. T. Hagan, and H. B. Demuth, *Neural Network Toolbox™ User’s Guide R2014b*. The Mathworks, Inc., 2014.

[18] M. Hagan, H. Demuth, and M. Beale, *Neural Network Design*, 2nd ed. 2014.

[19] M. Karimi, A. Dehghani, A. Nezamalhosseini, and S. Talebi, “Prediction of hydrocyclone performance using artificial neural networks,” *J. South African Inst. Min. Metall.*, vol. 110, no. 5, p. 207, 2010.

[20] H. Eren, C. C. Fung, and A. Gupta, “Application of artificial neural network in estimation of hydrocyclone parameters,” in *Diversity , the key to prosperity : the AusIMM 1996 Annual Conference*, 1996, pp. 225–229.

[21] A. J. Lynch and T. C. Rao, “Modelling and Scale-up of Hydrocyclone Classifiers,” *11th Int. Miner. Process. Congr.*, 1975.

[22] C. L. Karr, D. a Stanley, and B. McWhorter, “Optimization of hydrocyclone operation using a geno-fuzzy algorithm,” *Comput. Methods Appl. Mech. Eng.*, vol. 186, no. 2–4, pp. 517–530, Jun. 2000.

[23] H. Eren, C. C. Fung, K. W. Wong, and K. Street, “Back Propagation Neural Network in Determination of Parameter,” in *IEE TENCON - Digital Signal Processing Applications*, 1996, pp. 163–166.

[24] B. C. Flintoff, L. R. Plitt, and A. A. Turak, “Cyclone modelling: a review of present technology,” *CIM Bull.*, vol. 80, no. 905, pp. 39–50, 1987.

[25] I. The Mathworks, “Measuring Angle of Intersection,” 2015. [Online]. Available: http://www.mathworks.com/help/images/examples/measuring-angle-of-intersection.html.

[26] A. Kamoun, B. Samet, J. Bouaziz, and M. Châabouni, “Application of a rotatable orthogonal central composite design to the optimization of the formulation and utilization of an useful

plasticizer for cement," *Analusis*, vol. 27, pp. 91–96, 1999.

[27] M. J. J. Van Vuuren, C. Aldrich, and L. Auret, "Detecting changes in the operational states of hydrocyclones," *Miner. Eng.*, vol. 24, no. 14, pp. 1532–1544, 2011.

[28] I. A. Basheer and M. Hajmeer, "Artificial neural networks : fundamentals , computing , design , and application," *J. Microbiol. Methods*, vol. 43, pp. 3–31, 2000.

[29] C. C. Klimasauskas, "Applying Neural Networks," *PCAI*, vol. 5, no. 3, pp. 20–24, 1991.

[30] B. J. Taylor, *Methods and Procedures of Verification and Validation of Artificial Neural Networks*. Springer, 2006.

[31] R. G. Sargent, "Verification and Validation of simulation models," in *Proceedings of the 2005 Winter Simulation Conference*, 2005, pp. 130–143.

[32] J. Frost, "The Minitab Blog." [Online]. Available: http://blog.minitab.com/blog/adventures-in-statistics/multiple-regession-analysis-use-adjusted-r-squared-and-predicted-r-squared-to-include-the-correct-number-of-variables.

APPENDIX A – PROCEDURES

A.1 Inlet conversion

In order to convert the square inlet dimensions to a circular diameter, the area of the square inlet was calculated by using (A-1).

$$A_{square} = l_i \times b_i = 45 \times 16 = 720 \text{ mm}^2 \tag{A-1}$$

The calculated area of the square inlet is taken as the area of the circular inlet and the inlet's radius r_i is computed using (A-2).

$$r_i = \sqrt{\frac{A_{circle}}{\pi}} = \sqrt{\frac{720}{\pi}} = 15.14 \text{ mm} \tag{A-2}$$

Finally the inlet diameter D_i is obtained by using (A-3).

$$\therefore D_i = 2r_i = 2(15.14) = 30.28 \approx 30.3 \text{ mm} \tag{A-3}$$

A.2 Mixing of slurries

As specified in the Experimental Design (ED) the volumetric solid concentration (ϕ) was one of the three influential factors. In order to mix the slurries to the specified solid concentrations, some general slurry interrelation formulae were utilised. In this study silica was used as the solid material and water as the liquid medium of the slurry. The first formula that was used, is shown in (A-4). It was employed to calculate the slurry density (ρ_p) by referring to the chosen volumetric solid concentration[55] (ϕ), the solids density (ρ_s) and the liquid density (ρ_l). Note that the silica density (ρ_s) is taken as 2650 kg/m^3 and the water (ρ_l) as 1000 kg/m^3.

$$\rho_p = \frac{\phi(\rho_s - \rho_l)}{100} + \rho_l \tag{A-4}$$

Having determined the slurry density (ρ_p) the weight solid concentration ($C_{w\%}$) can be calculated next by using (A-5).

$$C_{w\%} = \frac{\rho_s(\rho_p - \rho_l)}{\rho_p(\rho_s - \rho_l)} \times 100 \tag{A-5}$$

[55] Note that some textbooks refer to volumetric solid concentration as $C_{v\%}$ rather than ϕ.

Now that the weight solid concentration ($C_{w\%}$) is known, the weight of the water as a percentage can be determined by employing (A-6).

$$Water\ wt\ \% = (100 - C_{w\%}) \quad \textbf{(A-6)}$$

The total mass (M_T) of the slurry could thus be calculated by utilising (A-7) where the mass of the water (M_w) is equal to 210 kg[56].

$$M_T = \left(\frac{100}{Water\ wt\ \%}\right)(M_w) \quad \textbf{(A-7)}$$

Finally the mass of solids (M_s) can be determined by calculating the difference between the total mass (M_T) and the mass of water (M_w) as depicted in (A-8).

$$M_s = M_T - M_w \quad \textbf{(A-8)}$$

These four formulae were used throughout the study to calculate the mass of solids required whenever mixing a slurry of a different composition. In order to verify that the slurry was mixed correctly, a representative sample was taken from the tank and weighed using the Marcy scale.

A.3 Experimental procedures

Pre-experiment preparations

The pre-experiment preparations mostly consist of planning and calculations that were implemented before starting with the actual experimental runs. These steps are crucial as they define and organise the experiments as well as identify additional necessities beforehand.

The steps were executed in the following manner:

1. Prepare an experimental specifics sheet[57] based on the design matrix (see Table 5-4).
2. Calculate[58] the mass of the solids required for each run and insert it into the experimental specifics sheet.
3. Based on step 2, prepare a sufficient amount of silica by extensively mixing several bags to ensure that a constant feed PSD will be used throughout.

---

[56] The tank is filled to the 210 litre mark, thus 210 l = 210 kg.
[57] The experimental specifics sheet specifies a run's operating conditions and provides spaces to record the measurements obtained during the experiments.
[58] The calculations are described in section A.2.

Experiment preparations

After the experiments and related aspects were planned, a comprehensive procedure was developed that initiated the experiments and prepared for the sampling thereof. The steps taken in order to achieve that are summarised as follows:

1. Always wear the appropriate Personal Protective Equipment[59] (PPE) when entering the laboratory.
2. Ensure that the test rig is clean[60] after being previously used.
3. Fill the tank with water up to the 210 litre mark.
4. Always ensure that the Marcy scale is calibrated by using clean water.
5. Referring to the experimental specifics sheet, weigh-off the required amount of silica to attain the specified volumetric solid concentration.
6. Also check (or change) the spigot size as specified on the experimental specifics sheet.
7. Ensure that the feed valve is fully-closed and that the bypass valve is fully-open.
8. Start the hydrocyclone pump.
9. Add the weighed solids into the tank water by gradually[61] mixing it.
10. Sampling can commence when the water and silica have been mixed together sufficiently.

Sampling

The sampling steps describe in what manner and order the different measurements were taken. The steps are:

1. Fully-close the bypass valve.
2. Open the feed valve until the required pressure, as specified in the experimental specifics sheet, is obtained.
3. Take a representative sample of the slurry in order to determine the feed PSD[62] later. This study's feed PSD profile is shown in Appendix A section A.4.
4. Ready the stopwatch and sampling bins.
5. Sampling can be initiated as soon as the hydrocyclone operates at steady state.
6. Start the stopwatch as the sampling bins engage the products.
7. Simultaneously note the feed flow rate (Q_i) as displayed on the flowmeter.
8. Stop the stopwatch when the sampling bins are removed from the products.

---

[59] The PPE that was worn throughout the experiments were a dust mask, safety glasses, dust jacket, long pants and steel-toed shoes.

[60] Should the test rig not be clean, add 210 litre water into the tank. Fully-open the feed valve and fully-close the bypass valve. Start the pump and let it run for a while. When the rig is deemed clean, switch the pump off and drain the water from the tank.

[61] This ensures that no clumps are formed and that instant settling is avoided.

[62] The Malvern analyses should reflect that the feed PSD was constant for every experimental run.

9. Record the sampling time (t) on the experimental specifics sheet.
10. Take a photo of the spigot and underflow discharge directly after the sampling time was recorded (The photo will be used to later determine the angle of discharge (ω)).
11. Switch the pump off.
12. Drain the sampling bins into smaller buckets in order for them to be weighed.
13. Record the samples' mass on the experimental specifics sheet respectively (M_o and M_u).
14. Determine the samples' density using the Marcy scale and record the measurements on the experimental specifics sheet (ρ_o and ρ_u).
15. Take a representative underflow sample[63] for analyses later on (d_{50} and m).
16. After checking whether all the necessary measurements were conducted and recorded, the remaining of the samples in the buckets can be mixed back into the tank.
17. Refer to the experimental specifics sheet in order to determine the next run's volumetric solid concentration, operating conditions, spigot size, etc.
18. Always ensure that the settled silica is mixed well into the slurry after the pump is switched on again.

Post-experiments

After all the experimental runs are completed, be sure to follow the subsequent steps:

1. Ensure that all the runs were conducted and that no measurements were overlooked.
2. Drain the slurry from the tank.
3. Hose out visible silica from the tank.
4. Fill the tank to the 210 litre mark with clean water.
5. Run the pump for a while to clean out any remaining silica within the pipes and hydrocyclone body.
6. Switch the pump off and drain the water from the tank.
7. Ensure that the area around the test rig is tidy and safe when leaving the laboratory.

A.4 Analysis procedure

Malvern analysis procedure

Every one of the stored underflow samples were meticulously analysed by using the Malvern Mastersizer. The analysis procedure is described in the steps as follow:

1. Switch on the Malvern Analyser.

[63] A 35 ml container was used to store the sample.

2. Switch on the computer that is connected to the analyser. Open the Malvern software application and start a new measurement file. Ensure that the software measurement settings are set-up correctly. The settings used for this study's analyses are tabulated in Table A-1.

Table A-1: Malvern software application settings

Sample material	
Name	Silica 0.0
Refractive index	1.544
Dispersant	
Name	Water
Refractive index	1.33

3. Place clean water in a beaker under the pump. Lower the pump into the water and start the pump. This is to ensure that the Malvern is rinsed at least once before starting any analysis. The pump is set to 2200 rpm throughout the analyses.
4. Leave the pump on for about 15 seconds.
5. Switch the pump off, lift the pump out of the water and allow the water to drain from the analyser's pipes and cell.
6. Remove the beaker from the Malvern and discard the water appropriately.
7. Fill the beaker with new clean water, place under the pump, lower the pump into the water and start the pump once more.
8. Add a very small amount of dissolved dishwashing liquid to the beaker. This helps with keeping the particles separate during circulation.
9. Start the Background measurement on the software application. This is to determine what the sample's threshold is. Always check whether the sample's analysis file name corresponds to the sample's label that is being analysed.
10. When the Background measurement is completed, the software application moves on to the Add Sample measurement. This measurement indicates the amount of the sample that's needed for an accurate analysis. This is done by determining the level of light being obscured by the sample particles already added to the medium. It is measured continuously as small, representative quantities of the sample are added. An adequate level of light obscuration is between 15 – 18 %. When the level of light obscuration is reached, no more of the sample should be added. The light obscuration averaged 17.3 % for this study.
11. With an adequate light obscuration level, the Start button within the software application is clicked to initiate the measurement. The Malvern measures the sample three times and calculates an average. To ensure that the analysis had no major issues, the three

measurements and their average are visually evaluated by the user. Should the user find large deviations between the measurements, the sample is re-analysed[64].

A.5 Processing procedure

Some of the collected measurements still need to be processed in order to be usable. This subsection details the processing procedures in regard to determining the cut-size and sharpness of classification coefficient, the underflow and overflow flow rate and the angle of discharge.

Cut-size and sharpness of classification from Malvern analysis

In order to obtain the cut-size and the sharpness of classification of each sample, the Malvern analyses were processed. The steps that should be executed in order to easily obtain the necessary data is summarised here:

1. From the Malvern software application be sure to electronically export the analyses done as text files.
2. Using Excel, import the analysis' text file. Do this for all the analyses, placing each text file's data on a new sheet within the same Excel book[65].
3. In a new Excel book[66] first copy and paste the cut-size.
4. Next copy and transpose the particle size and volume in percentage data into rows.
5. Calculate the cumulative volume percentage within an additional column.
6. Plot the cumulative volume percentage against the size utilising a semilog x-axis (a partition curve is obtained).
7. Extract the gradient of the partition curve by locating the data between the d_{25} and d_{75} and fitting a straight line through it. The gradient of the fitted line is thus taken as the gradient of the partition curve. Figure A-1 depicts the graphical representation of the method.

---

[64] Redo the analysis procedure starting at step 3.
[65] The imported text file data will be cluttered and unorganised; the data having been placed into columns.
[66] A macro was developed to automatically complete steps 3 through 7 to ensure precision and to save time.

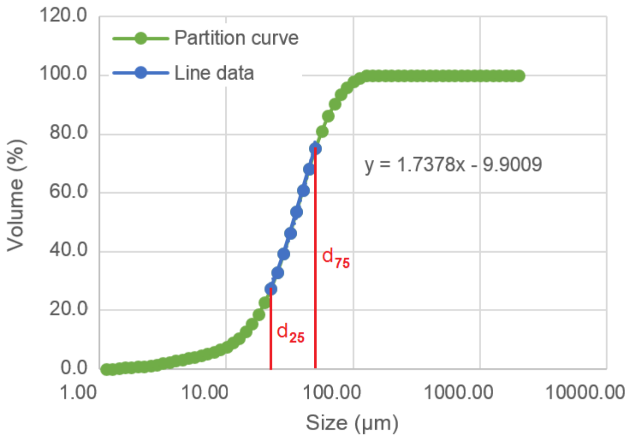


Figure A-1: Malvern analysis partition curve

Determining the underflow and overflow flow rate

Having had only one flowmeter installed, the only instrumentally measured flow rate was the inlet flow rate (Q_i). However by utilising some of the other recorded measurements it is possible to calculate the volumetric flow rate of the underflow (Q_u) and of the overflow (Q_o). This delivers two additional measured parameters. When looking at (A-9) it shows that the volumetric flow rate could be calculated using the weight of the samples that were recorded along with the density of the sample and by utilising the sampling time (t). Some basic unit conversions are shown in (A-10) which eventually produce a formula that can directly relate the measured weight and density of a sample to a volumetric flow rate characterised in $\mathrm{l/s}$ [6].

$$volumetric\ flow\ rate = \frac{weight\ of\ sample}{density\ of\ sample} \qquad \textbf{(A-9)}$$

$$Q = \frac{x\ [\mathrm{kg/s}]}{y[\mathrm{kg/m^3}]} = \frac{1000x}{y}[\mathrm{l/s}] \qquad \textbf{(A-10)}$$

Processing of the angle of discharge

The final parameter that was processed, was the angle of discharge (ω). Photos were taken of the spigot and the underflow discharge directly after sampling the products. The sample's angle of discharge was extracted from the photo by utilising the image processing toolbox in MATLAB®.

Literature suggest that the angle of discharge should be measured where the underflow discharge forms a flat downward slope as shown in region (b) in Figure A-2. When examining region (a) it is seen that as the discharge initially breaches the spigot opening a curved spray is formed. Region (c) indicates where gravity might start deforming the spray profile [27]. In order to accurately extract the angle of discharge, the photos were all edited prior to processing by adding a white line parallel to the spigot opening where the flat downward sloping region begins.

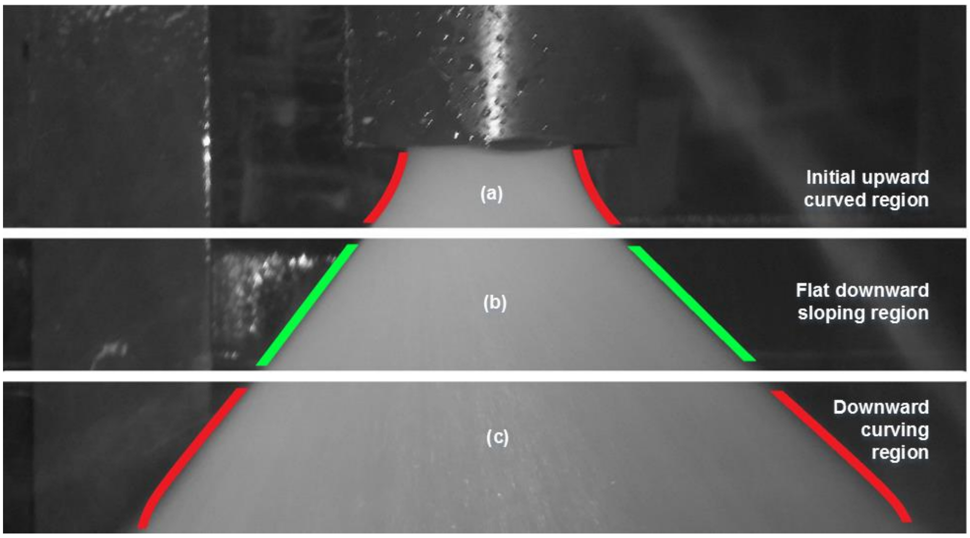


Figure A-2: Discharge spray profile regions

In order to obtain the angle of discharge from the photo, execute these steps:

1. From within MATLAB® load the specific photo of interest.
2. In order to identify and extract the desired boundaries, crop the small section containing the necessary information. Figure A-3 shows the cropped section of the original photo.

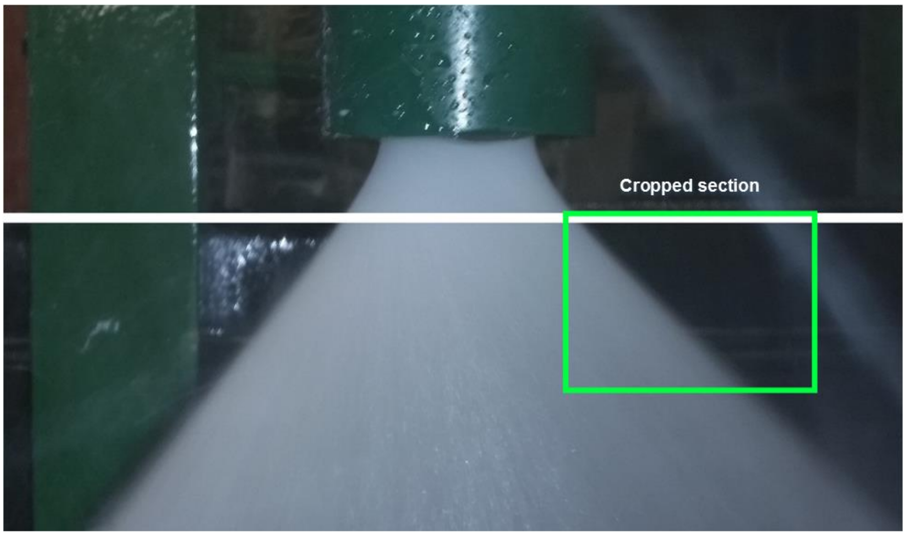


Figure A-3: Cropped section from the original photo

3. Convert the cropped image to a black and white image in order to distinguish the background from the spigot line and discharge (objects). Shown in Figure A-4 (b) that the background is white and the objects are black.

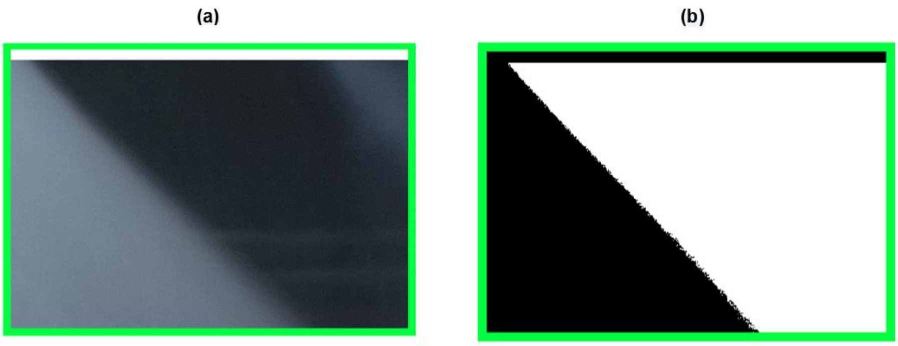


Figure A-4: The cropped section in (a) being converted to black and white in (b)

4. For MATLAB®'s bwtraceboundary function the user needs to specify a single point on the boundary where the function can start tracing. In other words a transition pixel is identified and supplied. A transition pixel is where an object (black) pixel goes to a background (white) pixel. This is done for both the spigot line and the discharge line.
5. The bwtraceboundary function, extracts (X, Y) coordinates, indicating the pixels that were found to be indicative of object boundaries. The user further assists the function by specifying a tracing direction. For this study the spigot line was traced East and the discharge line was traced South-East.
6. It is obvious that the extracted boundary coordinates will not lay in a straight line. Therefore lines are fitted through the traced pixels. A line equation is calculated for both the traced boundaries.
7. These fitted lines are now used to determine the angle of discharge. By using the line equations, the direction vectors are calculated. (A-11) shows the formula used in determining the angle. $\underline{l_1}$ is the direction vector of the spigot line and $\underline{l_2}$ the direction vector of the discharge line as shown in Figure A-5. $\underline{l_1} \cdot \underline{l_2}$ is the dot product of the two vectors and $\left|\underline{l_1}\right|$ and $\left|\underline{l_2}\right|$ is the magnitude of the direction vectors. The angle of discharge is displayed on the image and then recorded to the experimental specifics sheet accordingly.

$$\omega = \cos^{-1}\left(\frac{\underline{l_1} \cdot \underline{l_2}}{\left|\underline{l_1}\right| \left|\underline{l_2}\right|}\right) \qquad \textbf{(A-11)}$$

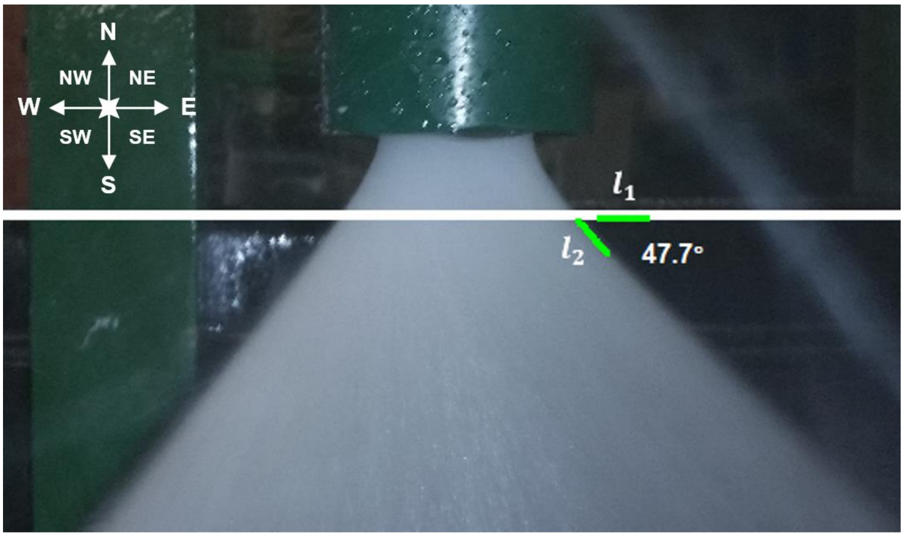


Figure A-5: Traced boundaries shown on the original photo along with ω

A.6 Feed PSD profiles

As mentioned it was expected that each experimental run's feed PSD should be similar. After analysing the PSD samples, it was found that the feed PSD was indeed constant throughout. Evaluating the PSD profiles of each sample depicted in Figure A-6, the largest deviation that was seen was that of PSD_006, having shifted slightly right.

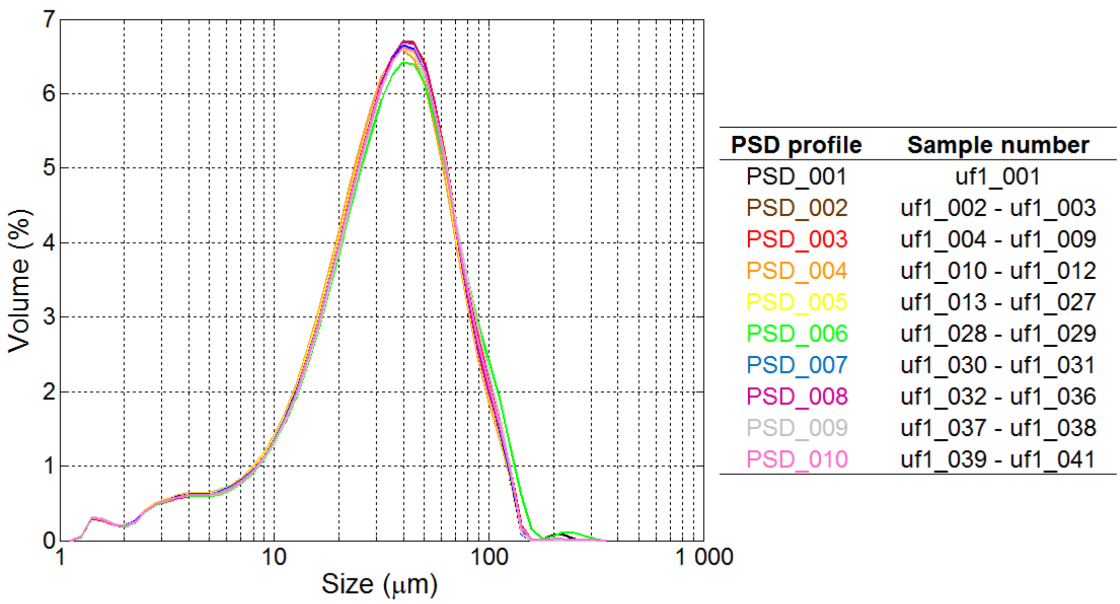


Figure A-6: The feed PSD profiles of each sample

APPENDIX B – DATA

B.1 Final data

B.2 Raw data

B.3 MATLAB® code

B.4 Models developed

B.5 Models ANOVA

B.6 Plitt-Flintoff model

APPENDIX C – PLITT-FLINTOFF CALIBRATION FACTORS

Usually the Plitt-Flintoff calibration factors are calculated by taking the average factors of a few random samples and assigning those averaged factors to all of the data. This approach however delivered very poor cut-size and sharpness of classification estimations as seen in Figure C-1 (a) and Figure C-2 (b). In order to improve the estimations, the samples were grouped under the five different spigot opening diameters and the factor averages calculated for each group. The averaged factors were then assigned to all the samples with that corresponding spigot opening diameter. The estimations improved as shown in Figure C-1 (b) and Figure C-2 (b), but still showed unsatisfactorily correlation. The factor averages that were calculated for this study is tabulated in Table C-1.

Table C-1: Factor values for ungrouped and grouped approaches

Related parameter	Factor	Factor value					
		Ungrouped	Grouped per spigot opening diameter				
			15	20	25	30	35
d_{50}	F_1	3.217	2.372	2.745	3.112	3.435	3.885
m	F_2	1.331	0.999	1.067	1.039	1.133	1.273
S	F_4	0.209	0.808	0.377	0.151	0.101	0.074

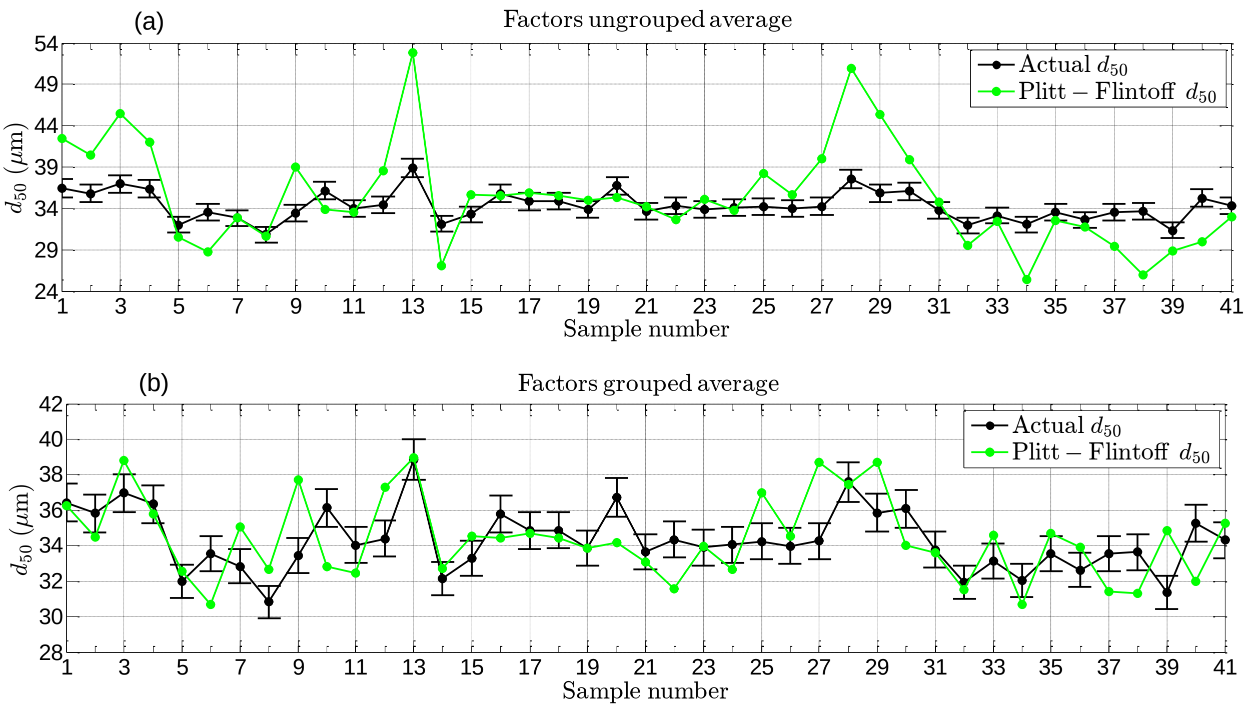


Figure C-1: The cut-size shown per sample when employing the (a) ungrouped factors and (b) grouped factors

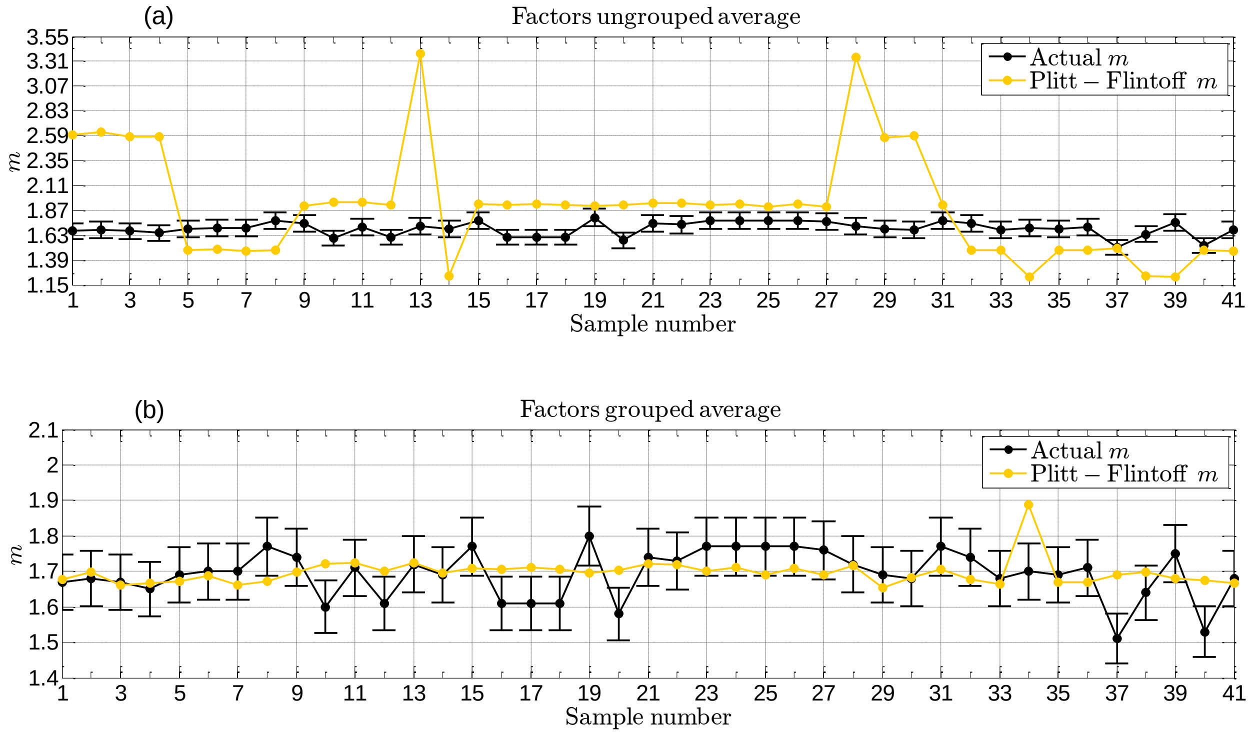


Figure C-2: The sharpness of classification shown per sample when employing the (a) ungrouped factors and (b) grouped factors

APPENDIX D – CONFERENCE CONTRIBUTIONS

D.1 Hydrocyclone separation efficiency estimation using artificial neural networks

This conference article was presented at the Southern African Universities Power and Energy Conference (SAUPEC) in Johannesburg on the 29th of January 2015. The full article is follows on page 109.

D.2 Hydrocyclone cut-size estimation using artificial neural networks

This article was submitted to the Symposium on Dynamics and Control of Process Systems (DYCOPS) and is currently under review. Should the article be accepted, the symposium will be held in Trondheim, Norway, during June 2016. The full article is set out from page 115.

HYDROCYCLONE SEPARATION EFFICIENCY ESTIMATION USING ARTIFICIAL NEURAL NETWORKS

S. van Loggenberg*, G. van Schoor, K.R. Uren* and A.F. van der Merwe*****

** School of Electrical, Electronic and Computer Engineering, North-west University Potchefstroom, Potchefstroom 2520, South Africa E-mail:21818347@nwu.ac.za/Kenny.Uren@nwu.ac.za*
*** School of Electrical and Electronic Engineering, North-west University Potchefstroom, Potchefstroom 2520, South Africa E-mail:George.VanSchoor@nwu.ac.za*
**** School of Chemical and Mineral Engineering, North-west University Potchefstroom, Potchefstroom 2520, South Africa E-mail:Frikkie.VanDerMerwe@nwu.ac.za*

Abstract: The hydrocyclone is widely used throughout the mineral processing industry when working with slurries for either classifying, desliming or dewatering. Hydrocyclones are inexpensive, application-efficient and relatively small to employ. In order to quantify its separation efficiency, models are incorporated to estimate the cut-size and sharpness of classification, usually in the form of a partition curve. Most models are based on experimentally obtained data and are therefore not universally applicable. Over the last decade researchers have started employing Artificial Neural Networks (ANN) in order to obtain just such a model. This study endeavoured to use experimentally acquired data to develop a model that predicts the cut-size and sharpness of classification. A control phase is also discussed.

Keywords: Hydrocyclone, Artificial Neural Network, partition curve, sharpness of classification, cut-size.

1. INTRODUCTION

Ever since the hydrocyclone became popular in the mineral processing industry, there have been researchers that worked on developing a model for it. In 1964 Bradley published a book in which the known hydrocyclone fundamentals and research of that time were documented. Bradley also compiled an extensive list of theoretical equations estimating the cut-size, among others. These equations however were not always relevant to industry type hydrocyclones [1], [2]. Lynch and Rao were the next researchers that made an important contribution to hydrocyclone modelling. The research was mainly focused around developing empirical equations for hydrocyclones used in industry [3]. In 1976 Plitt published a paper on his mathematical model. Up to this day it is seen as one of the most popular and widely referenced mathematical models. Plitt developed the model by utilising empirical data along with hydrocyclone variables that were deemed important in describing the hydrocyclone's operation [2]. Plitt and Flintoff revised the mathematical model and published a reviewed article on it in 1987 [4]. Throughout the literature mentioned, it becomes clear that because the process and fundamental principles of the hydrocyclone are so complex, the models that were developed were not always fully comprehensive.

With the computing power that is available presently, the modelling techniques have improved significantly, resulting in a better understanding of the hydrocyclone principles. These models now include Computational Fluid Dynamics (CFD), System Identification and statistical correlations. In 1997 H. Eren et al developed Artificial Neural Networks (ANNs) that were employed to predict the Particle Size Distributions (PSD) and cut-size of various hydrocyclones [5], [6].

A hydrocyclone's separation efficiency is mainly described by the cut-size (d_{50}) and the sharpness of classification (m). These two parameters are used within a partition curve to quantitatively depict the separation efficiency of the hydrocyclone. Ideally a hydrocyclone is operated at conditions where a specific cut-size and sharpness of classification is achieved. These parameters cannot however be monitored in real-time. Thus this study aimed in developing an ANN, based on experimentally obtained data, which can predict the cut-size and sharpness of classification parameters at certain operating conditions.

This article will endeavour to describe some basic concepts of the hydrocyclone. A control perspective is briefly investigated. Next the developed Artificial Neural Network application and its specifics will be discussed. The results that were obtained is given and considered and finally a comprehensive conclusion binds the findings of the article.

2. HYDROCYCLONE OVERVIEW

2.1 A general description

A hydrocyclone is a stationary conical-apparatus that was developed to classify or separate solids from water, better known as slurries, and is generally used in mineral processing applications where classifying, dewatering or desliming is concerned. The separation of slurries within hydrocyclones is based on centrifugal sedimentation, where the necessary swirl motion is generated by the slurry that is fed into the hydrocyclone by means of a pump [1]. A hydrocyclone is relatively small, inexpensive and efficient when used within specific conditions and applications. Over the years it has however become evident that the hydrocyclone and related investigations are very complex and therefore the analyses thereof is still not completely comprehensive [7].

2.2 Design and operating variables

Two groups of hydrocyclone variables exist. Design variables, which are dependent of the hydrocyclone size and proportions and the operating variables that are related to the operating conditions of the hydrocyclone and independent of its size and proportions. These variables however, cannot be considered separately because they interact with one another. Table 1 summarises the main variables that fall within the two groups mentioned [8]. Figure 1 depicts a hydrocyclone and its design variables.

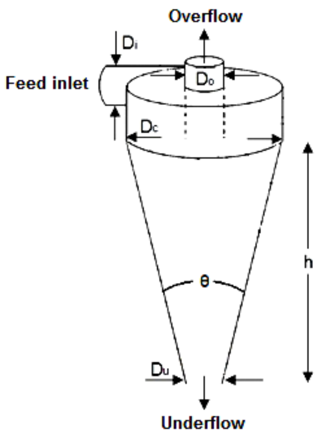


Figure 1: Graphical representation of a hydrocyclone with relevant design variables (adapted) [1]

Table 1: Hydrocyclone design and operating variables

Design variables	
Hydrocyclone diameter	D_c
Feed inlet diameter	D_i
Vortex finder diameter	D_o
Spigot diameter	D_u
Cone angle	θ
Free vortex height	h
Operating variables	
Hydrocyclone through-put	Q
Pressure drop	P
Solid concentration	ϕ
Solid density	ρ_S
Liquid medium velocity	V

2.3 Hydrocyclone performance

When describing the hydrocyclone's performance[1] the five major quantitative parameters are said to be [1], [2], [8]–[10]:

- Partition curves
- Cut-size (d_{50})
- Sharpness of classification (m)
- Pressure through-put relationship
- Split of water flow to products

For this study it was decided that only the cut-size, sharpness of classification and partition curves would be investigated. A brief discussion will follow describing these chosen parameters.

Partition curves: A partition[2] curve is a graphical and quantitative representation of a hydrocyclone's particle size separation performance. It usually describes the weight fraction (or percentage) of each particle size in the feed which reports to underflow, shown on the y-axis, to the particle size, shown on the x-axis. It is however assumed that a fraction of the fine particles completely bypasses the hydrocyclone's classification process. This is called the bypass and it explains why the efficiency curve does not have an asymptote at zero. It is generally assumed that the bypass is equal to the water that reports to the underflow. Thus the two types of partition curves that are generally discussed are the gross[3] partition curve, which does not take into account the water recovery, and the reduced[4] partition curve, which is adjusted to include the water recovery effects [10]. Figure 2 illustrates the two types of partition curves.

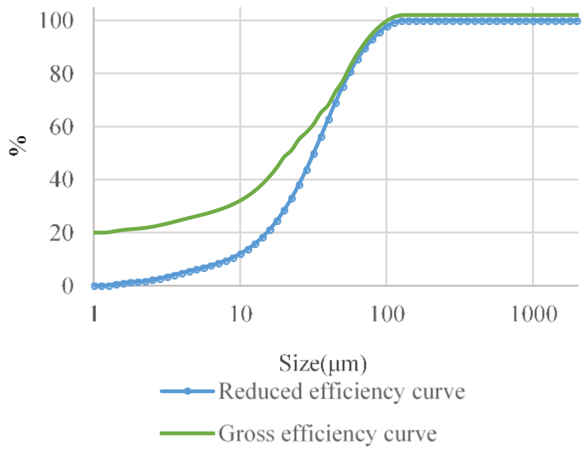


Figure 2: The two types of partition curves

[1] Performance refers to the hydrocyclones separation or classification efficiency.
[2] Also known as efficiency curves, performance curves or tromp curves.
[3] Also called uncorrected partition curve.
[4] The same as corrected partition curve.

Cut-size: The cut-size, also indicated as d_{50}, is said to be the size of the particle in the feed Particle Size Distribution (PSD) that has a 50% probability to either report to the underflow or the overflow of the hydrocyclone [1]. The cut-size is shown in Figure 3.

Sharpness of classification: The sharpness of classification is a parameter that is used to quantify the hydrocyclones classification or separation efficiency by supply a measure for the gradient of the partition curve. A $m < 2$ would signify poor classification and a $m > 3$ would imply good sharpness of classification [2].

Figure 3 also depicts the sharpness of classification indicated as m.

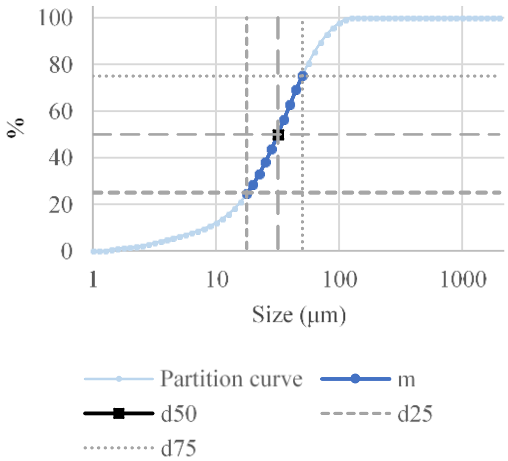


Figure 1: Partition curve displaying the d_{50} point as well as the m

In order to define the functionality of the partition curves there exist three distribution functions that closely fits the hydrocyclone's experimental performance. These distribution functions include the Exponential function, the Rosin-Rammler function as well as the Logistic function. The Rosin-Rammler function was said to deliver the best fit of the three [11]. Equation (1) shows the form in which the Rosin-Rammler function is used [2]. The sharpness of classification is also obtainable by using (2) [8].

$$y = [1 - e^{-0.693\left(\frac{d}{d_{50}}\right)^m}] \times 100 \tag{1}$$

Where y is the distribution function, d is the PSD range in μm, d_{50} the cut-size in μm and m the sharpness of classification.

$$m = \frac{d_{75} - d_{25}}{2d_{50}} \tag{2}$$

Where m is the sharpness of classification, d_{75} the size of the particle that has a 75% probability of reporting to the underflow, d_{25} the size of the particle that has a 25% probability of reporting to the underflow and d_{50} the cut-size.

2.4 A control perspective

There exist the need, especially within the industry, to optimise the hydrocyclone's performance so that it is operated as efficient as possible when referring to the cut-size and partition curve. Therefore a model will be very useful in a control applications where the cut-size and sharpness of classification needs to be optimised. Figure 4 shows a control scheme that could be implemented along with the developed model.

The hydrocyclone variables that might change during operation are the pulp density (ρ_p) and therefore the solids concentration (ϕ). The control system would have to account for these changes by adjusting the pressure and the flow rate so that the optimal cut-size and sharpness of classification is obtained.

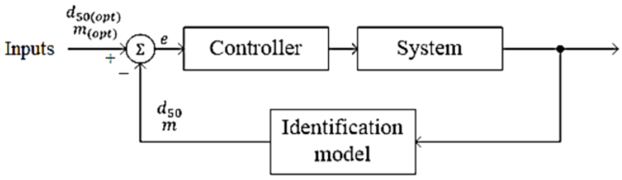


Figure 4: Control scheme

3. ARTIFICIAL NEURAL NETWORK INFERENCE MEASUREMENT

3.1 Artificial Neural Network overview

Artificial Neural Networks (ANNs) are used to describe input-output relationships of systems. Usually when using ANNs, there is no a priori knowledge needed of the system, which would make it an ideal modelling method to use with the hydrocyclone, because of the complex fundamentals that are involved. It also works very well when the model is solely based on experimentally acquired data, as with this study. Previous works have shown that the accuracy with which the ANN model predicts separation efficiency are comparable, in some cases even better, to the conventional models [6], [12]. ANN models are especially convenient should the model include alternative hydrocyclone variables, such as overflow and underflow flow rates and overflow density [6], [12].

3.2 The developed Artificial Neural Network

For this study a hydrocyclone's separation efficiency was predicted by developing an ANN model that is based on measurements obtained from experiments done on a hydrocyclone test-rig [2], [5], [12], [13]. The design configuration used for this study was fixed as portrayed in Table 2. A next phase would implement some sequential changes to the vortex finder diameter and spigot diameter.

Table 1: Experimental details

Design variable dimensions	
Hydrocyclone diameter	100 mm
Feed inlet diameter	48 mm
Overflow diameter	60 mm
Underflow diameter	15 mm
Free vortex height	700 mm

When investigating conventional models, it becomes evident that certain hydrocyclone variables are more influential than other in the estimation of the separation efficiency. The hydrocyclone variables that are seen as the most influential will be used as inputs to the ANN model along with the desired outputs.

Table 2 summarises the 8 hydrocyclone variables that were chosen as model inputs and Table 3 the 2 efficiency-indicating outputs.

Table 2: Artificial Neural Network model inputs

Design variables		
Hydrocyclone diameter	D_c	cm
Feed inlet diameter	D_i	cm
Overflow diameter	D_o	cm
Underflow diameter	D_u	cm
Free vortex height	h	cm
Operating variables		
Inlet flow rate	Q	l/s
Solid concentration	ϕ	%
Feed density	ρ_p	$\mathrm{kg/m^3}$

Table 3: Artificial Neural Network model outputs

Hydrocyclone performance parameters	
Cut-size	d_{50}
Sharpness of classification	m

3.3 The Artificial Neural Network details

After the model inputs and outputs were determined and the experimental procedure was defined, the experimental sampling was done and analysed using a Malvern particle size analyser. In total 44 underflow samples were taken and analysed. Of these, 2 samples were withheld from the network development. The 2 samples were only used after the ANN was trained, validated and tested in order to have results that could be compared to the Malvern analysis results. Thus 42 samples were used for training, validation and testing purposes and 2 were used to depict the estimation of the ANN model in comparable output-parameters.

Using the nntool in Matlab® a feed-forward backpropagation network was used, employing the Levenberg-Marquardt training function. The Levenberg-Marquardt training function uses the Levenberg-Marquardt optimisation technique to update the weights and bias values. It is one of the fastest backpropagation methods available although the memory requirements are somewhat high. Literature suggests that the number of neurons could be determined by using (3) [14],

$$\frac{Number\ \ of\ samples}{5} = (i+1)N + (N+1)o \qquad (3)$$

where i is the number of inputs, N the number of neurons and o the number of outputs.

When applying (3) it was found that 2 neurons would be sufficient. After incorporating 2, 4, 5 and 10 neurons within the ANN, it became clear that the 5 neuron ANN delivered the best results for the configuration as summarised in Table 5. Figure 5 shows the network scheme diagram indicating the number of inputs, number of hidden layers, neurons used and the number of outputs that was used in the final ANN.

Table 4: Artificial Neural Networks specifics summary

Sample division %		Number of samples	Hidden layers	Neurons
Training	70	29		
Validation	15	7	1	5
Testing	15	6		

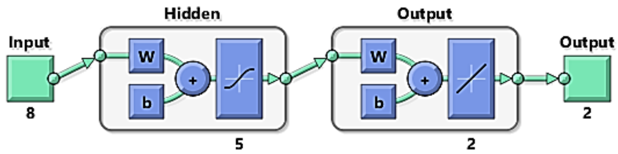


Figure 1: Artificial Neural Network diagram

4. EVALUATION

In order to evaluate the performance of the ANN the Mean Squared Error (MSE) plot and Regression plots were examined. The MSE plot, as depicted in Figure 6, shows how the MSE decreases for the three data sets (training, validation and testing) as the iterations proceeds. Training is stopped when the green validation line stops decreasing. This ensures that the network does not over-train. The red test line shows how well the network will generalise for new data.

Figure 7 shows the three data sets regression as well as the overall network regression. The regression plot illustrates the networks outputs versus the targets. The data markers should preferably be as close as possible to the dashed line. The regression lines (coloured lines) shows how well the network outputs are centred around the targets. The R-value states the average scatter around the dashed line, of which the ideal R-value would be 1. From Figure 7 it is seen that the overall network R-value is 0.9979 and a very high correlation is achieved.

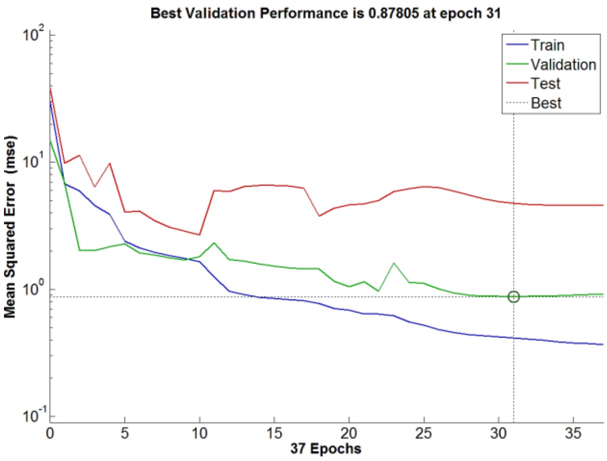


Figure 1: MSE performance plot

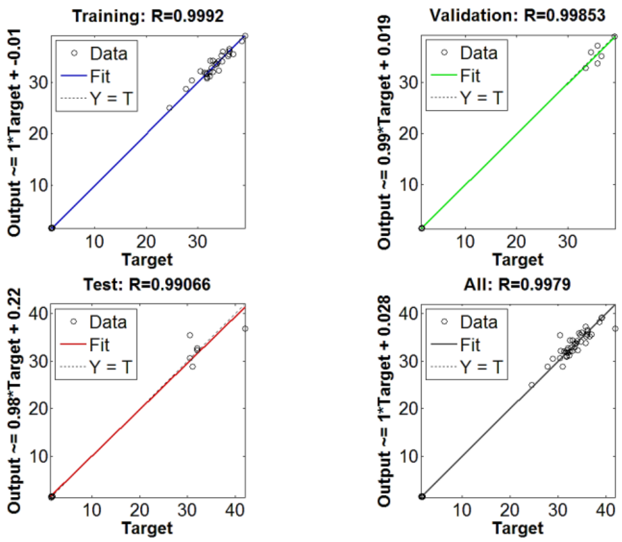


Figure 2: Regression plots

The 2 samples that were withheld from the network during development were now used to determine whether the network could accurately estimate comparable outputs when unknown inputs were introduced. Table 6 shows the 2 samples' experimental results; in other words, the ANN would be deemed accurate should it be able to predict these values with marginally small errors. Table 7 tabulates the predicted results that were obtained by simulating the ANN, the percentage error of the values are also given.

Table 1: The Malvern experimental analysis results

Experimental analysis	
$d_{50(1)}$	39.943
$d_{50(2)}$	30.636
$m_{(1)}$	1.4506
$m_{(2)}$	1.7454

Table 2: Artificial Neural Network estimation of the cut-size and sharpness of classification of the 2 samples

	Predicted value	**error (%)**
$d_{50(1)}$	35.0493	12.25
$d_{50(2)}$	32.3457	5.58
$m_{(1)}$	1.4263	1.68
$m_{(2)}$	1.5844	9.22

For each of the 2 withheld samples, the predicted values were plotted against the experimental analysis. The distribution function given in (1) was used along with the predicted sample values in order to obtain partition curves that could be compared to the experimentally obtained partition curves. The estimated samples deviate slightly as expected by the % error indicated. The overall shape however is satisfactory as shown in Figure 8.

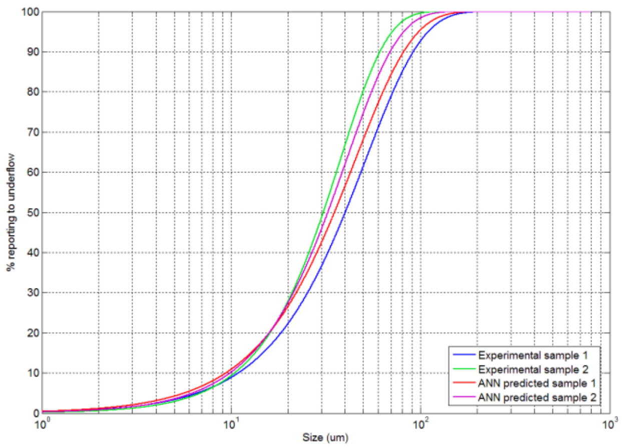


Figure 8: Experimental partition curve versus the predicted samples' partition curve

5. CONCLUSION

It was found that by using influential hydrocyclone variables as inputs along with a feed-forward backpropagation Levenberg-Marquardt ANN, a model could be developed that accurately estimated the separation efficiency parameters of a hydrocyclone. The ANN model errors were found to be marginally small, with an overall R-value of 0.9979. The next phase of the study would be to investigate the model estimations when compared to conventional models, the estimations when design variations are introduced and to determine the application of the developed model when used for control purposes.

ACKNOWLEDGEMENTS

The authors would like to thank the School of Chemical and Mineral Engineering for the acquisition and analysis of the data, specifically R. Kok and H. Mulke.

REFERENCES

[1] D. Bradley, *The Hydrocyclone - Bradley*. London: Pergamon Press Ltd., 1965.

[2] L. R. Plitt, “A Mathematical Model of the Hydroyclone Classifier,” *CIM Bull. Miner. Process.*, vol. 69, no. December, pp. 114–123, 1976.

[3] A. J. Lynch and T. C. Rao, “Modelling and Scale-up of Hydrocyclone Classifiers,” *11th Int. Miner. Process. Congr.*, 1975.

[4] B. C. Flintoff, L. R. Plitt, and A. A. Turak, “Cyclone modelling: a review of present technology,” *CIM Bull.*, vol. 80, no. 905, pp. 39–50, 1987.

[5] H. Eren, C. Fung, and K. Wong, “An application of artificial neural network for prediction of densities and particle size distributions in mineral processing industry,” in *IEEE Instrumentation and measurement Technical Conference*, 1997, pp. 1118–1121.

[6] H. Eren, C. C. Fung, and K. W. Wong, “Artificial Neural Networks in Estimation of Hydrocyclone Parameter 50 with Unusual Input Variables,” *IEEE Trans. Instrum. Meas.*, vol. 46, no. 4, pp. 908–912, 1997.

[7] E. G. Kelly and D. J. Spottiswood, *Introduction to Mineral Processing*. Australia: John Wiley & Sons, 1982.

[8] B. A. Wills and T. J. Napier-Munn, *Wills' Mineral Processing Technology*, 7th Editio., no. October. Oxford: Butterworth-Heinemann, 2006.

[9] H. Eren and A. Gupta, “Instrumentation an on-line control of hydrocyclones,” in *International Conference on Control*, 1988, pp. 301–306.

[10] M. Frachon and J. J. Cilliers, “A general model for hydrocyclone partition curves,” *Chem. Eng. J.*, vol. 73, no. February, pp. 53–59, 1999.

[11] S. Hore and S. Das, “Data-Based Performance Modelling of Hydrocyclone for Processing Iron Ore Fines,” in *International Seminar on Mineral Processing Technology*, 2011, pp. 1–6.

[12] H. Eren, C. C. Fung, and A. Gupta, “Application of artificial neural network in estimation of hydrocyclone parameters,” in *Diversity , the key to prosperity : the AusIMM 1996 Annual Conference*, 1996, pp. 225–229.

[13] H. Eren, C. C. Fung, K. W. Wong, and K. Street, “Back Propagation Neural Network in Determination of Parameter,” in *IEE TENCON - Digital Signal Processing Applications*, 1996, pp. 163–166.

[14] C. C. Klimasauskas, “Applying Neural Networks,” *PCAI*, vol. 5, no. 3, pp. 20–24, 1991.

Hydrocyclone cut-size estimation using artificial neural networks

S. van Loggenberg*, G van Schoor*, K.R. Uren* and A.F. van der Merwe**

** School of Electrical, Electronic and Computer Engineering, North-West University, Potchefstroom, South Africa (e-mail:sarita.svl@gmail.com/george.vanschoor@nwu.ac.za/kenny.uren@nwu.ac.za).*
*** School of Chemical and Mineral Engineering, North-West University, Potchefstroom, South Africa (e-mail:10212361@nwu.ac.za).*

Abstract: The hydrocyclone is widely used throughout the mineral processing industry when working with slurries. It is either used for classifying, desliming or dewatering. Hydrocyclones are inexpensive, application-efficient and relatively small to employ. In order to quantify its separation efficiency, models are utilised to estimate the cut-size and sharpness of classification coefficient, usually in the form of a partition curve. Most models are based on experimentally obtained data and are therefore not always universally applicable. Over the last decade researchers have started employing Artificial Neural Networks (ANNs) in order to obtain a dynamic model. This study endeavoured to use experimentally acquired data to develop models that predict the cut-size. The models are discussed and evaluated in detail and the best predicting model was compared to a conventional model from literature.

Keywords: Artificial Neural Network, Modelling, Hydrocyclone, Cut-size, Partition curve, Plitt-Flintoff.

1. INTRODUCTION

Ever since the hydrocyclone became popular in the mineral processing industry, there have been researchers that worked on developing a model describing its separation efficiency. Bradley (1965) published a book in which the known hydrocyclone fundamentals and research of that time were detailed. He also compiled an extensive list of theoretical equations estimating the cut-size and sharpness of classification, among others. These equations however were not always relevant to industrial hydrocyclones. The next important contributions made regarding the modelling of hydrocyclones were by Lynch & Rao (1975). Their research was mainly focused around developing empirical equations for industrial hydrocyclones. Some time later, in 1976, Plitt published a paper on his mathematical model of a hydrocyclone. His model is one of the most popular and most extensively referenced models. Plitt developed the model by utilising empirical data along with hydrocyclone variables that were deemed important in describing the hydrocyclone's operation. Flintoff et al. (1987) later reviewed the mathematical model and published a revised article on it where useful calibration factors were added to the mathematical model to further improve estimations.

With the advance in computational capability models now include Computational Fluid Dynamics (CFD), System Identification and expanded statistical correlations. In 1997 H. Eren et al were some of the first researchers that employed Artificial Neural Networks (ANNs) to predict the Particle Size Distributions (PSD) and cut-size of various hydrocyclones (H Eren et al. 1997; Halit Eren et al. 1997).

A hydrocyclone's separation efficiency is mainly described by the cut-size (d_{50}) and the sharpness of classification (m). These two parameters are used within a partition curve to quantitatively depict the separation efficiency of the hydrocyclone. Ideally a hydrocyclone is operated at conditions where a specific cut-size and sharpness of classification is achieved. These parameters cannot however be monitored in real-time (Frachon & Cilliers 1999). This paper thus aimed at developing ANNs, based on experimentally obtained data, which could predict the cut-size and sharpness of classification parameters at certain operating conditions. Throughout the literature mentioned, one thing becomes abundantly clear and that is the complexity of the hydrocyclone's fundamental principles.

A brief overview of a hydrocyclone is given in Section 2 in terms of what it is, where it is used, variables associated with it and how the performance is described. Section 3 focusses on the ANN models that were developed, discussing their inputs, architecture and sample division. In order to check the adequacy of the developed ANN models, Analysis of Variance (ANOVA) studies were done and is shown in Section 4.1. To determine which one of the models performed the best, three popular error metrics were utilised. Finally the best ANN model was compared to Plitt-Flintoff's mathematical model estimations, in order to determine whether the ANN would perform better than the mathematical model, given in Section 4.2. The paper is concluded by outlining the work done and discussing the most important aspects of the study's findings.

2. HYDROCYCLONE OVERVIEW

2.1 A general description

A hydrocyclone is a static, conical apparatus that is generally used within the mineral processing industry to separate solids from water, better known as slurries. The separation of the slurries is based on sedimentation, where the necessary swirl motion is generated by the slurry being fed into the hydrocyclone by means of a pump (Bradley 1965). Two vortices form within the hydrocyclone referred to as the primary vortex and the secondary vortex as depicted in Fig. 1 (a). The primary vortex moves downwards and carries the coarse particles to an opening called the underflow. The secondary vortex carries the lighter particles, along with most of the water, upwards to an opening called the overflow (Frachon & Cilliers 1999).

2.2 Hydrocyclone variables

When working with hydrocyclones, two groups of variables are observed. The design variables (shown in Fig. 1 (b)) include variables that are dependent on the hydrocyclone's size and design proportions. The operating variables are independent of the hydrocyclone's design and solely relate to the operating conditions of the hydrocyclone. It should however be noted that these two groups of variables cannot be considered separately because of interactions that occur between them. Table 1 summarises the main variables allocated to the two groups.

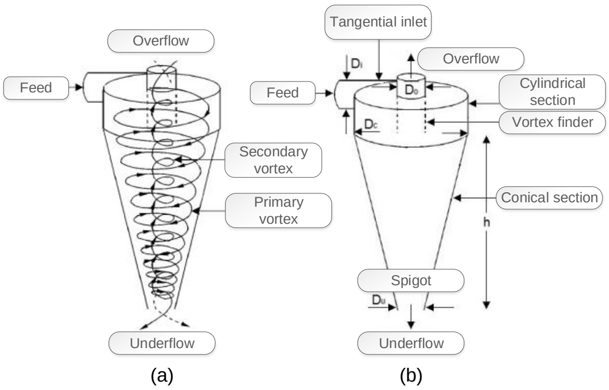


Fig. 1. A graphical representation of a hydrocyclone depicting the (a) vortices and (b) relevant design variables (adapted) (Frachon & Cilliers 1999).

2.3 Performance of a hydrocyclone

A hydrocyclone's performance is mainly described by its separation efficiency which is directly quantifiable by a partition curve. The partition curve describes the weight (or percentage) fraction of each particle size in the feed that might report to the underflow on the *y*-axis, to the specific particle size, on the *x*-axis. An example of a partition curve is given in Fig. 2 showing how the two efficiency indicating parameters, cut-size and sharpness of classification coefficient, are related to it. The cut-size, indicated as d_{50}, is defined as the size of the particle in the Particle Size Distribution (PSD) that has a 50% probability of reporting to either the underflow or the overflow of the hydrocyclone (Bradley 1965). The sharpness of classification coefficient (m) is a parameter that supplies a measure for the gradient of the partition curve. Ideally an m > 3 is required in order to obtain sufficiently sharp separation (Plitt 1976).

Table 1. Hydrocyclone variables

Design variables	
Hydrocyclone diameter	D_c
Feed inlet diameter	D_i
Vortex finder diameter	D_o
Spigot opening diameter	D_u
Cone angle	θ
Free vortex height	h
Operating variables	
Inlet flow rate	Q_i
Overflow flow rate	Q_o
Underflow flow rate	Q_u
Pressure	P
Volumetric solid concentration	ϕ
Solid density	ρ_S
Overflow density	ρ_O
Angle of discharge	ω

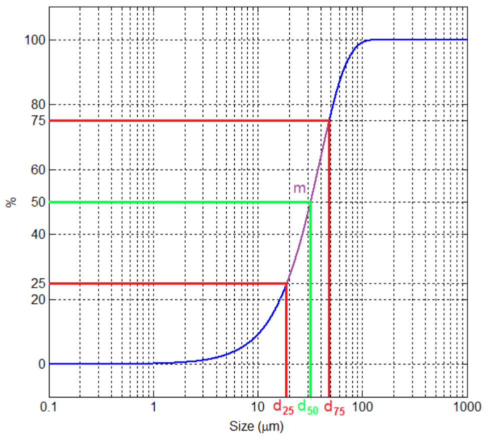


Fig. 2. A partition curve displaying the cut-size and sharpness of classification coefficient.

2.4 Experimental setup

A refurbished hydrocyclone test-rig, shown in Fig. 3, was instrumented and used in order to acquire the necessary experimental data. The hydrocyclone's dimensions are tabulated in Table 2. An analog pressure gauge was utilised to measure the inlet pressure and a Doppler flow meter to measure the inlet flow rate.

Table 2. The hydrocyclone dimensions

Design variable dimensions	
Hydrocyclone diameter	100 mm
Feed inlet diameter	33.4 mm
Overflow diameter	34.0 mm
Free vortex height	531 mm

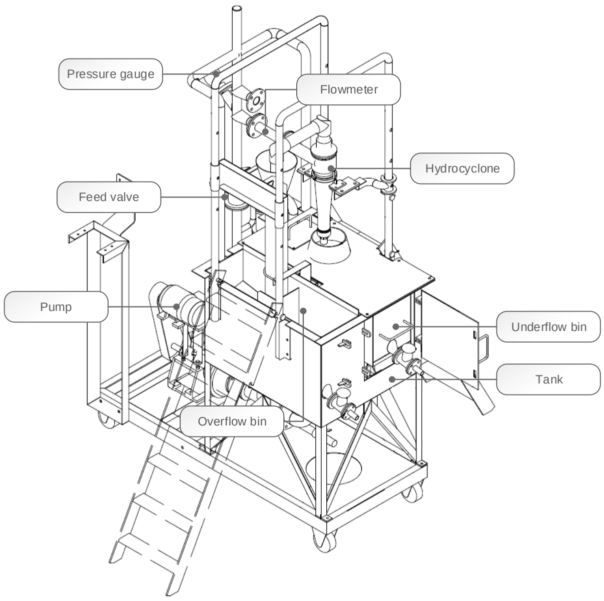


Fig. 3. The schematic of the hydrocyclone test-rig.

3. ARTIFICIAL NEURAL NETWORKS

3.1 Artificial Neural Network overview

Artificial Neural Networks (ANNs) are employed to describe input-output relationships of systems. Usually when working with ANNs no *a priori* knowledge is needed of the system. ANNs are therefore an ideal modelling method to use for the hydrocyclone system because of the complex dynamics that are involved. Previous works that incorporated ANNs found that the models could deliver comparable, and in some cases even better, predictions to the conventional models. The ANN models proved especially useful when alternative variables were included such as overflow and underflow flow rates and overflow density (Eren et al. 1996; Halit Eren et al. 1997).

3.2 Developed Artificial Neural Networks

The main goal of this study was to develop an Artificial Neural Network, using experimentally obtained data, which could predict the hydrocyclone's cut-size (d_{50}) and sharpness of classification coefficient (m). In order to obtain the experimental data, 41 experiments and their conditions were defined by utilising the Centrally Composite Rotatable Design (CCRD) as investigated by Cilliers et al. (1992). These experimental runs were conducted and the samples were analysed accordingly, by using a Malvern Mastersizer 2000 particle size analyser. In investigating the conventional models it became evident that some hydrocyclone variables are considered more influential than others in estimating the separation efficiency. Plitt (1976) deemed the design variables, the inlet flow rate and the solid density as most influential and developed a mathematical model around these variables. When studying Eren & Gupta (1988) the spigot opening diameter is said to affect the performance the most. Unusual variables, such as overflow and underflow flow rates and overflow density were additionally incorporated in the models developed by Halit Eren et al. (1997).

To ensure that an accurate and comprehensive model was developed for this study, three different ANNs were created and evaluated. For each one of the three models the variables used as inputs were adjusted and the performance of the ANN evaluated. Model 1 used only the 3 inputs that were set during the experimental runs. Model 2 had two additional operating variables, the inlet flow rate and the angle of discharge, in order to evaluate whether the angle of discharge might improve the ANN's performance. The final model, Model 3, incorporated some of the unusual variables that were relevant to this study as suggested by Halit Eren et al. (1997). Table 3 summarises the models' specifications in terms of their inputs, outputs, network architecture and sample division. All three the models were based on the same Artificial Neural Network architecture as shown in Fig. 4 varying only the number of neurons.

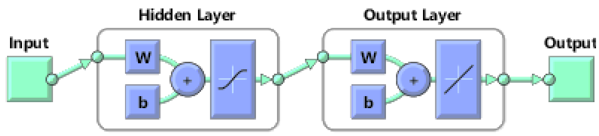


Fig. 4. The Artificial Neural Network architecture.

It was found that the sharpness of classification ANNs were all inadequate when evaluating their ANOVA results, delivering mostly unusable predictions. When examining the literature, very little information is available on the modelling of the sharpness of classification coefficient. This might indicate that the m cannot be modelled comprehensively using the chosen variables, especially when examining the variables Plitt (1976) used to model the sharpness of classification. The results discussed in this paper will therefore only reflect the cut-size models and their relevant outcomes.

4. EVALUATION

4.1 Model adequacy and comparison

The first step of the model evaluation was to check whether the developed models could be deemed adequate or not. To achieve that, an Analysis of Variance (ANOVA) was completed for each of the models. The calculated F-value of the models were compared to the appropriate critical F-value at an $\alpha = 0.05$. The $\alpha = 0.05$ implies that a confidence level of 95% is evaluated and it is expected that about 5% of the samples might yield erroneous results. Should the calculated F-value be larger than the critical F-value, the model was said to be significant and thereby considered adequate (Devore & Farnum 2005). The ANOVA of the models are tabulated in Table 4. When evaluating the table it is seen that the F-values of all three models are satisfactorily higher than the identified critical F-value and it can thus be concluded that all three models are adequate.

With all three developed models found adequate, the models were compared with one another to determine which one would better predict the cut-size. The first measure employed

Table 3. Model details and specifications

Name	Inputs		Output		Sample division						Hidden layers	Number of neurons
					Training		Validation		Testing			
	#	Variables	#	Variable	%	Samples	%	Samples	%	Samples		
Model 1	3	P, ϕ, D_u	1	d_{50}	60%	25	20%	8	20%	8	1	2
Model 2	5	$P, \phi, D_u,$ Q_i, ω	1	d_{50}	60%	25	20%	8	20%	8	1	4
Model 3	8	$P, \phi, D_u, Q_i,$ Q_u, Q_o, ρ_o, ω	1	d_{50}	60%	25	20%	8	20%	8	1	3

Table 4. Summary of ANOVA for Model 1, Model 2 and Model 3

Source	df[a]	SS[b]	MS[c]	F	test F[d]
			Model 1		
Model	3	89.469	29.823	34.656	$F_{0.05}(3{,}37)$ = 2.92 < 34.656
Error	37	31.841	0.861		Significant @ the level 95%
Total	40	121.31			∴The model is deemed adequate
			Model 2		
Model	5	83.291	16.658	15.335	$F_{0.05}(5{,}35)$ = 2.53 < 15.335
Error	35	38.019	1.086		Significant @ the level 95%
Total	40	121.31			∴The model is deemed adequate
			Model 3		
Model	8	102.699	12.837	22.073	$F_{0.05}(8{,}32)$ = 2.27 < 22.073
Error	32	18.611	0.582		Significant @ the level 95%
Total	40	121.31			∴The model is deemed adequate

[a] df - degrees of freedom; [b] SS - Sum of Squares; [c] MS - Mean Square; [d] F_α(df Model, df Error).

was to calculate the coefficient of determination (R^2). This was done by plotting the ANN predicted cut-size, denoted as $\hat{y}_i$, against the actual cut-size y_i. The process was repeated for all three the models and the results are shown in Fig. 5. The dashed line represents the best fit linear regression line between the actual and predicted cut-sizes. An $R^2 = 1$ signifies an exact linear relationship, thus the higher the R^2-value the stronger the linear relationship is expected to be. When comparing the three models' R^2, it would seem that Model 3 performed the best. This was however the first assessment and some additional evaluation is needed, as the R^2 should never be the only measure examined (Devore & Farnum 2005).

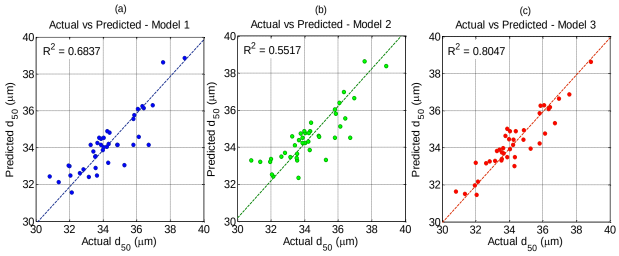


Fig. 5. The actual cut-size versus the predicted cut-size for (a) Model 1, (b) Model 2 and (c) Model 3.

The second measure used to evaluate the models was to visually compare the predicted cut-sizes. The actual cut-size was plotted per sample depicting the expected experimental error, calculated as 2.95%, as error bars. The predicted cut-size per sample was then plotted on the same graph. The models' predicted cut-size was expected to at least follow the trend of the actual cut-size. The graphs obtained for the three models are given in Fig. 6. Very small differences are observed between the three models but it is noted that the Model 2 shows the worst correlation of the three. During the training of the ANNs were trained, some samples were completely withheld from the ANN. These samples were then later used to test the performance of the ANN model when presented with unknown data. These unknown samples' actual and predicted cut-size are visually portrayed in Fig. 7. Once more it was found that the correlation of Model 2 was the worst.

To further investigate the models' performance, three popular error metrics were calculated, assessed and compared. The metrics used in this study was the Mean Square Error (MSE), the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). When comparing the three models in terms of the three metrics, one would like to see the same ranking order

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \quad (3)$$

y_i denotes the actual cut-size, $\hat{y}_i$ the predicted cut-size and n the number of observations evaluated (Devore & Farnum 2005).

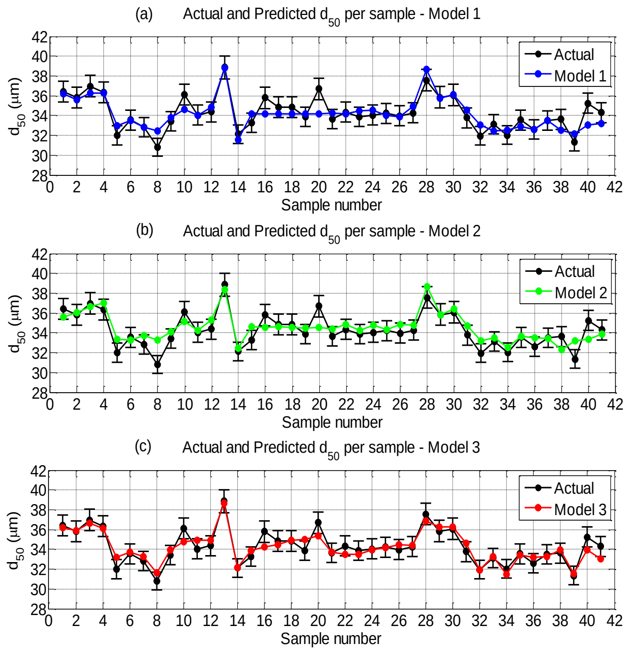


Fig. 6. The actual and the predicted cut-size shown all samples for (a) Model 1, (b) Model 2 and (c) Model 3.

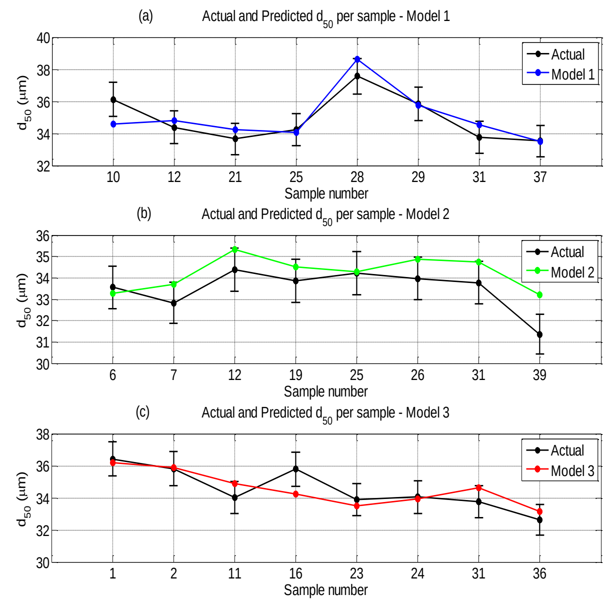


Fig. 7. The actual and predicted cut-size of unknown samples for (a) Model 1, (b) Model 2 and (c) Model 3.

The metric assessment results computed for all the samples and unknown samples are tabulated in Table 5. When comparing the metrics all three indicate that the models, ranking best performing to worst performing, are Model 3, Model 1 and Model 2. Model 3 only just surpass Model 1 when unknown samples are considered. When reviewing the model evaluations, it is seen that Model 3 performed the best in terms of its R^2-value, the visual comparison and the error metrics. Model 3 should therefore be used to ensure sufficient predictions of the cut-size.

Table 5. Summary of error metrics when evaluating all and unknown samples for Model 1, Model 2 and Model 3

Name	Error functions		
	MSE	RMSE	MAE
All samples			
Model 1	0.7766	0.8812	0.6460
Model 2	0.9273	0.9630	0.7672
Model 3	0.4539	0.6737	0.5170
Unknown samples			
Model 1	0.5774	0.7599	0.5826
Model 2	0.9200	0.9592	0.8166
Model 3	0.5508	0.7422	0.5757

4.2 The best ANN model versus a conventional model

In order to evaluate whether the developed ANN model could substitute the conventional model, a comparison was drawn and examined. The conventional model that was investigated was the mathematical model Plitt and Flintoff revised in 1987. Equation (4) shows the mathematical model that was used to calculate the cut-size (d_{50}) when the hydrocyclone design variables and operating conditions are known. F_1 is a calibration factor which is used to improve the cut-size estimation (Flintoff et al. 1987). It is usually computed by finding the factor-ratio between the actual measured cut-size and the model calculated cut-size.

$$d_{50} = F_1 \frac{39.7 D_c^{0.46} D_i^{0.6} D_o^{1.21} \eta^{0.5} e^{0.063\phi}}{D_u^{0.71} h^{0.38} Q^{0.45} \left[\frac{(\rho_S - 1)}{1.6}\right]^k}, \quad (4)$$

After computing the calibration factor for known samples, (4) was employed to estimate the cut-size for the same unknown samples used to evaluate Model 3. Fig. 8 shows the actual and predicted cut-sizes for unknown samples. It is difficult to distinguish exactly which model might perform better as some samples of Model 3 seem better than the Plitt-Flintoff estimation and vice versa. In order to differentiate which model performs better the same three error metrics were calculated and compared. The metric results, shown in Table 6, indicate that Model 3 predicts significantly better than the Plitt-Flintoff mathematical model. It is therefore concluded that not only can an ANN model be developed and used to predict the cut-size at specific operating conditions, but that an ANN model could substitute the conventional mathematical model.

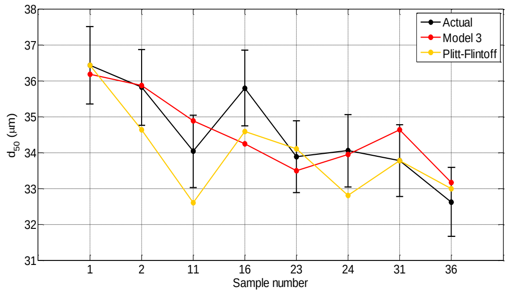


Fig. 8. Comparing the actual and predicted cut-size of unknown samples employing Model 3 and Plitt-Flintoff's mathematical model.

Table 6. Summary of error functions when evaluating the unknown samples for Model 3 and Plitt-Flintoff model

Name	Error functions		
	MSE	RMSE	MAE
Model 3	0.5508	0.7422	0.5757
Plitt-Flintoff	0.8292	0.9106	0.7073

5. CONCLUSION

This paper found that it was possible to develop Artificial Neural Networks (ANNs) comprising of different hydrocyclone variables as inputs. Using variables that were deemed important in literature, three different models were developed. All three these models were found to be adequate in predicting the cut-size at various viable operating conditions, falling mostly within the acceptable experimental error of 2.95%. When comparing their performance however, Model 3 with its 8 inputs seemed to deliver the most accurate predictions. It should be noted that it only just performs better than Model 1. Model 2 with the inlet flow rate and the angle of discharge as additional inputs, seems to perform the worst of the three. In wanting to investigate whether Model 3 might be able to substitute the conventional mathematical model of Plitt and Flintoff (1978), the two models' prediction capabilities were compared by presenting them with unknown samples. The error metrics that were used, implied that Model 3 would be better at predicting cut-size than the Plitt-Flintoff model. It could thus be concluded that Model 3 would be able to replace the conventional mathematical model and that it would deliver more accurate predictions. Future work could involve investigating whether all the unusual variables are needed to improve the ANN. In other words determine if less inputs could deliver the same or better predictions. This could simplify the ANN as well as minimise the data acquisition effort (measuring less variables).

ACKNOWLEDGEMENTS

This work is based on the research support in part by the National Research Foundation of South Africa (Grant Number 91093).

REFERENCES

Bradley, D., 1965. *The Hydrocyclone - Bradley*, London: Pergamon Press Ltd.

Cilliers, J.J., Austin, R.C. & Tucker, J.P., 1992. An evaluation of formal experimental design procedures. In *International Conference on Hydrocyclones*. pp. 31–49.

Devore, J. & Farnum, N., 2005. *Applied Statistics for Engineers and Scientists* 2nd ed., Brooks/Cole, Cengage.

Eren, H., Fung, C. & Wong, K., 1997. An application of artificial neural network for prediction of densities and particle size distributions in mineral processing industry. In *IEEE Instrumentation and measurement Technical Conference*. pp. 1118–1121. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=612374 [Accessed April 21, 2014].

Eren, H., Fung, C.C. & Gupta, A., 1996. Application of artificial neural network in estimation of hydrocyclone parameters. In *Diversity , the key to prosperity : the AusIMM 1996 Annual Conference*. Perth, Western Australia, pp. 225–229.

Eren, H., Fung, C.C. & Wong, K.W., 1997. Artificial Neural Networks in Estimation of Hydrocyclone Parameter 50 with Unusual Input Variables. *IEEE Transactions on Instrumentation and Measurement*, 46(4), pp.908–912.

Eren, H. & Gupta, A., 1988. Instrumentation and on-line control of hydrocyclones. In *International Conference on Control*. Oxford: IET, pp. 301–306.

Flintoff, B.C., Plitt, L.R. & Turak, A.A., 1987. Cyclone modelling: a review of present technology. *CIM bulletin*, 80(905), pp.39–50. Available at: http://cat.inist.fr/?aModele=afficheN&cpsidt=7683497 [Accessed April 21, 2014].

Frachon, M. & Cilliers, J.J., 1999. A general model for hydrocyclone partition curves. *Chemical Engineering Journal*, 73(February), pp.53–59.

Plitt, L.R., 1976. A Mathematical Model of the Hydroyclone Classifier. *CIM Bulletin Mineral Processing*, 69(December), pp.114–123.