Particle Size Estimation of Hydrocyclone Overflow

by

Vincent Combemere Smith

Submitted to the University of Cape Town in fulfilment of the requirements for the degree of Master of Science in Engineering

April 2000

ACKNOWLEDGEMENTS

I would like to express my gratitude to the Measurement and Control Division at Mintek for inviting me to use this work as an MSc project.

A special word of thanks also to Professor Chris Swartz for his patience and excellent supervision and insight.

Thanks also to the following colleagues:

Agit Singh for his help in gathering plant data

Roger Henning and Gareth Smith for proof reading this dissertation

Janos Vida for dependable sample analysis

Dr Dave Hulbert for his encouragement and technical insight

A final word of gratitude goes to the staff of the South African gold plant who allowed me to use their plant for gathering of industrial data, without which this project would not have been possible.

SYNOPSIS

This dissertation describes the development of a robust hydrocyclone particle size estimation model that will form the basis of an industrial soft-sensor.

Aside from increased throughput, efficient product regulation constitutes the primary function of a milling circuit control system. Before the milling circuit product size can be regulated, it should be measured or estimated. The particle size estimation algorithm developed provides a reliable estimate of the product size and will compliment or replace conventional size measurement devices.

Mintek have previously developed the Particle Size Estimator (PSE), a soft-sensor utilising the hydrocyclone feed flowrate, feed density and underflow angle to provide an estimate of the product size. This work is an extension of previous efforts, focusing on the evaluation of various model types and combinations of input variables.

The research is based on data gathered from a large South African gold milling circuit. Extensive step tests and sampling formed the basis for obtaining data to allow subsequent model identification.

Altogether 4 model inputs and 2 model types are considered. It is clearly illustrated that use of the hydrocyclone underflow angle term does not benefit model performance in any way, while the inclusion of hydrocyclone overflow density dramatically improves model performance. Furthermore, neural network models proved to be more accurate than multiple linear regression models, but more difficult to calibrate and maintain.

The final model choice is a simple multiple linear regression model consisting of three inputs, viz. hydrocyclone feed flowrate, feed density and overflow density. To ensure availability when an input instrument fails, a family of three secondary, less accurate multiple linear regression models have been identified by ignoring each input sequentially. A calibration procedure, using simple offset correction has been developed to allow easy on-site maintenance by operational staff.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

NOMENCLATURE

Upper Case

A_i	Cyclone inlet area
CFD	Cyclone feed density
CFF	Cyclone feed flowrate
CM	Correlation matrix
COD	Cyclone overflow density
Cy_{50}	Characteristic cyclone number
CUA	Cyclone underflow angle (degrees)
D_c	Cyclone diameter (cm)
D_i	Cyclone inlet equivalent diameter (cm)
D_o	Vortex finder diameter (cm)
D_u	Spigot diameter (cm)
F	Solids mass flowrate in the feed
F_1, F_2, F_3, F_4	Calibration parameters for Plitt model
F_c	Centrifugal force
F_d	Drag force
HF	Vortex height factor
HUM	Hydrocyclone Underflow Meter
K_1 to K_8	Constant model parameters (Lynch and Rao model)
K_c	Constant parameter dependent on design and fluid characteristics
K_d, K_Q, K_v, K_w	Material specific constants
K_{uf}	Constant underflow angle correction factor in Mintek PSE
L	Length of cylindrical section of hydrocyclone
L_c	Total length of the hydrocyclone
MLR	Multiple linear regression
MSE	Mean square error
NNet	Neural network
O_j	Objective function
P	Cyclone inlet pressure (kPa)
PMD	Cumulative percent passing size D
PSE	Particle Size Estimator (% -size in μm)
Q	Cyclone feed flowrate (l/min)

QF	Flowrate factor
Q_o	Volumetric flowrate of the overflow
Q_u	Volumetric flowrate of the underflow
R_f	Bypass fraction, equal to water recovery in the underflow
R_s	Mass recovery of feed solids to the underflow
R_v	Recovery of feed volume to the underflow
S	Volumetric flow split between underflow and overflow (fraction)
T	Temperature
TF	Temperature factor
U	Solids mass flowrate in the underflow
V	Tangential velocity
W	Radial velocity

Lower Case

a, b, d, e	Constant model parameters in the Mintek PSE
d_{50}	Cut size (μm)
d_{50c}	Corrected cut size (μm)
d_i	Typical size of a particle in size class i (μm)
d_o	Size beyond which bypass is no longer significant
f_i	Mass fraction in size i in feed to the cyclone
g	Gravitational acceleration
h	Free vortex height of the cyclone (cm)
k	Exponent to account for solids density (k = 0.5 for laminar flow)
k_o	Maximum particle size on a sieve analysis plot
m	A measure of the sharpness of separation in the cyclone
n_1	Gradient of a sieve analysis plot
p_i	Partition factor for size i
p'_i	Corrected partition factor for size i
r	Radius of orbit at equilibrium
r_i	Solids bypass for size i
u_i	Mass fraction of solids of size i in the underflow
v	Axial velocity
w_{bias}	Bias weight
w_{ij}	Weight corresponding to index *ij*

Greek

α	Momentum term
β	Step size
Δ	Difference
δ	Local error at node
η	Viscosity of the carrier medium - water (centipoise)
φ	Percent solids in the cyclone feed on a volume basis (%)
φ_s	Mass percent solids in the feed
φ_v	Volumetric fraction of solids in the feed
μ	Viscosity of the carrier fluid
θ	Hydrocyclone cone angle
ρ	Density of the carrier fluid
ρ_p	Density of cyclone feed slurry (g/cm^3)
ρ_s	Density of the solid ore (g/cm^3)

CHAPTER 1. INTRODUCTION

In the minerals processing industry, milling is one of the most widely used comminution processes. Milling is an inherently energy and capital cost intensive process. The operating region of the milling circuit has a significant impact on downstream processes.

A common goal of a milling circuit is to produce a product of a specific passing size, while maintaining the required throughput, in a cost-effective manner. In the South African gold industry, most milling circuits are closed and use a hydrocyclone for classification. A measure or estimate of the particle size of the hydrocyclone overflow product and the subsequent control of particle size is beneficial to efficient operation and is a useful variable for grind versus throughput optimisation. Overgrinding results in higher operating costs and reduced throughput, with minimal return in downstream recovery, while undergrinding results in significant losses in recovery.

Conventional particle size measurement devices rely on taking a physical sample and then inferring the particle size from ultrasonic, x-ray or physical analysis. Due to the practical difficulties in working with slurry streams, these instruments tend to suffer from blockages and breakages leading to long periods of downtime. Furthermore, periods between adjacent sample analyses are in the order of minutes. Both downtime and lengthy analysis time detrimentally affect the performance of the associated particle size control system.

MINTEK has developed a technique for estimating the particle size of the hydrocyclone product stream. The instrument developed is known as the Particle Size Estimator (PSE). The PSE does not rely on a physical measurement of particle size. Hence, there is no need for the installation of a sampling system. It also means that the instrument requires little maintenance, which results in low down time. The PSE gives continuous estimates of particle size and is thus useful for implementation into an on-line control scheme. The PSE uses hydrocyclone feed flowrate, feed density and the angle of the underflow flare to predict particle size. For a fixed geometry hydrocyclone, there are three major operational factors affecting classification, viz. feed flowrate and density, and changes in the size distribution of the cyclone feed. Flowrate and density changes are measured, while the third degree of freedom, feed size changes, is accounted for by including a measurement of the hydrocyclone underflow angle as measured by a mechanical or ultrasonic device.

The PSE model was a first attempt at developing a soft-sensor for particle size estimation. This research will focus on improving the existing PSE model. In all cases, data were gathered on industrial rather than pilot scale by performing extensive test work and sampling on a large South African gold plant. Various model inputs and model types will be investigated to develop a more accurate and robust particle size prediction algorithm.

Chapter 2 reviews current literature available for hydrocyclone modelling. Fundamental, empirical and semi-empirical models are considered. Section 2.5 focuses on soft-sensors for particle size estimation and includes a detailed description of Mintek's current PSE algorithm.

Artificial neural network fundamentals are reviewed in Chapter 3. This will be important when neural network models are applied to the plant data.

Chapter 4 details the plant test site, experimental procedure, data analysis and system identification procedures.

Neural network and multiple linear regression models are compared for various inputs in Chapter 5. The final soft-sensor model structure and calibration are outlined here.

Finally, the findings of this research are summarised in Chapter 6.

CHAPTER 2. LITERATURE REVIEW

2.1 HISTORY AND DEVELOPMENT OF THE HYDROCYCLONE

Probably because of its simplicity, the hydrocyclone has been used for classification applications since before 1890. A patent was granted to Bretney (Bretney, 1891) for a device used to separate sand from water (practically identical to present-day desanding cyclones). Between 1900 and 1937, 30 additional patents were issued for the use of devices resembling the hydrocyclone of today. The first commercial application of the hydrocyclone came about in 1937, when hydrocyclones were sold for cleaning sand and dirt particles from fibre suspensions in the pulp and paper industry.

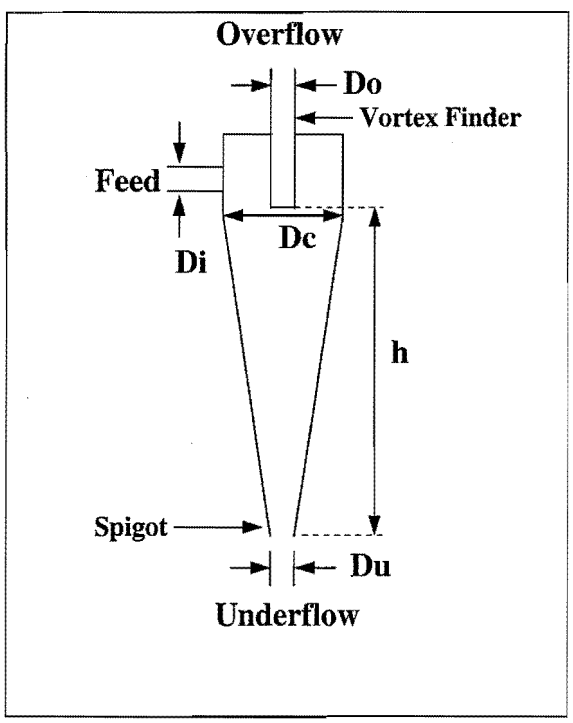
By 1940, hydrocyclones as we know them today, made an appearance in the paper and pulp as well as coal washing industries. At the same time, Driesen, of the Dutch State Mines was working on hydrocyclones for thickening and classification of slurry streams (Bradley, 1965). The Dutch State Mines hydrocyclone was first commercially implemented in the US in 1948.

Between 1950 and 1960, a large amount of research was conducted into the understanding of the basic operational principals of hydrocyclones. During this period, the use of hydrocyclones in the minerals industry began to increase rapidly. In the mid-1950s, hydrocyclones were applied in closed-circuit milling operations in South Africa.

Despite its mechanical simplicity, the hydrocyclone is one of the least understood items of equipment in the minerals industry. Over the last 30 years, there has been a great surge in research into hydrocyclone operation. Earlier researchers emphasised and developed theoretical models based on fluid flow characteristics, which gave subsequent researchers a fundamental understanding of the mechanism of classification. In the mid-1970s, researchers concentrated more on the empirical modelling of hydrocyclones. This period saw the birth of the most famous hydrocyclone models, the Lynch and Plitt models. Over the last 10 years, theoretical modelling, by solving the complex equations of fluid flow, has been rekindled.

2.2 DESCRIPTION OF HYDROCYCLONE OPERATION

Figure 1 shows a diagram of the hydrocyclone, indicating the design variables. As shown, the hydrocyclone consists of a cylindrical section, followed by a conical section. The feed slurry is introduced tangentially, and under pressure, to the cylindrical section. The fluid pressure creates a rotational motion of the fluid. This results in the generation of a centrifugal force, which acts upon the solid particles. Generally, the settling characteristics follow Stokes' law. The settling velocity is determined by the size and density of the solid particles. In a homogeneous slurry, where all particles have the same density, the settling rate (towards the wall of the hydrocyclone) is determined by particle size alone. In the hydrocyclone there exists a bulk movement of water towards the centre, which leads to the establishment of a second force, known as the drag force. The drag force (F_d) acts in the opposite direction to the centrifugal force (F_c). It is expected that large particles will move towards the hydrocyclone wall ($F_c > F_d$), while small particles will move towards the centre ($F_d > F_c$).



Where h = free vortex height
D_i = inlet diameter
D_c = hydrocyclone diameter
D_o = vortex finder diameter
D_u = spigot diameter

Figure 1. Hydrocyclone and design variables

When a large particle reaches the hydrocyclone wall, it slides down the cone and exits via an opening known as the spigot. The spigot discharge stream is known as the underflow stream,

and contains the coarse fraction of the feed with a high solids content. The small particles are entrained by the movement of water towards the centre where an upward rotating air core exists. The formation of the air core is due to the rotational motion of fluid within the hydrocyclone. The water and small particles form a second upward rotating spiral and exit via the vortex finder. This stream is known as the overflow and consists of fine particles at a low solids content. The overflow constitutes the product stream for a closed milling circuit, while the underflow is returned to the mill for further grinding.

2.3 THE PARTITION CURVE

One of the most popular methods of illustrating the classification performed by a hydrocyclone is by means of a partition curve (also known as an efficiency or Tromp curve). The efficiency curve relates the fraction of each particle size reporting to the underflow to that particle size. Mathematically, the curve is defined by equation 1.

$$p_i = \frac{Uu_i}{Ff_i} \tag{1}$$

where p_i = partition factor for size i

U = solids mass flowrate in the underflow
u_i = mass fraction of solids of size i in the underflow
F = solids mass flowrate in the feed
f_i = mass fraction of solids of size i in the feed

A typical partition curve is shown, together with the corrected version, in Figure 2. R_f represents the water recovery to the underflow. The solids bypass fraction (solids exiting in the underflow without undergoing classification) was defined by Kelsall (1953) to be equal to the water fraction recovered in the underflow for each size class (equation 2). This is based on the assumption that the water reporting to the underflow brings with it fine solids.

$$r_i = R_f \tag{2}$$

where r_i = solids bypass for size i

If the water split is known, the calculation of solids bypass is trivial. Kelsall's bypass model has been the most widely applied, however other models are mentioned for completeness (equation 3 and 4).

Austin and Klimpel (1981) :

$$r_i = a \tag{3}$$

where a = constant to be determined by experimentation and regression analysis

Finch et al. (1985) :

$$r_i = R_f\left(1 - \frac{d_i}{d_o}\right) \tag{4}$$

where d_i = characteristic size of size class i
d_o = particle size beyond which bypass is no longer significant

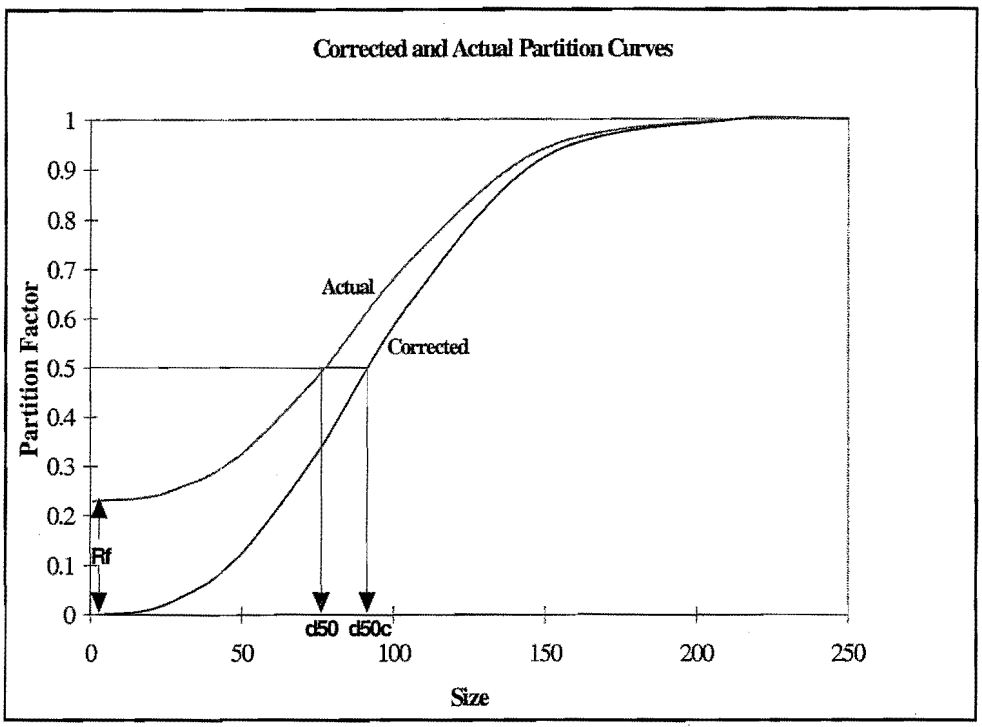


Figure 2. Partition curves for a hydrocyclone classifier illustrating d_{50}, d_{50c} and the water bypass (R_f)

In order to correct for the solids bypass, a corrected partition curve was defined. This curve shows only the truly classified solids. The corrected partition curve is calculated by equation 5 by using Kelsall's equation for solids bypass.

$$p_i' = \frac{p_i - R_f}{1 - R_f} \tag{5}$$

where p_i' = corrected partition function

The cut size (d_{50}) is defined as the particle size which has an equal possibility of reporting to the overflow or to the underflow (i.e. $p_i = 0.5$). The corrected cut size (d_{50c}) is determined by only considering the solids undergoing true classification (i.e. the size at which $p_i' = 0.5$). Figure 2 illustrates d_{50} and d_{50c}.

Partition curves have been found to display a characteristic shape and can be modelled by using sigmoidal growth functions. The various model equations for the determination of the corrected partition curve are shown in equation 6-8.

Exponential sum (Lynch and Rao, 1975) :

$$p_i' = \frac{e^{mX_i} - 1}{e^{mX_i} + e^m - 2} \tag{6}$$

Rosin Rammler (Plitt, 1971) :

$$p_i' = 1 - e^{-0.693X_i^m} \tag{7}$$

Logistic (Lilge and Plitt, 1968) :

$$p_i' = \frac{1}{1 + \left(\frac{1}{X_i}\right)^m} \tag{8}$$

where $X_i = d_i / d_{50c}$

m = sharpness of separation (higher m implies more efficient classification)

Using equation 5 and 7, for example, the actual partition curve can be empirically modelled.

$$p_i = R_f + (1 - R_f)\left(1 - e^{-0.693\left(\frac{d_i}{d_{50c}}\right)^m}\right) \tag{9}$$

From equation 9, it can be seen that R_f, d_{50c} and m completely describe the partition curve.

2.4 HYDROCYCLONE MODELLING

Hydrocyclone modelling has received a great deal of attention in the last 30 years. The earliest attempts at modelling considered the theoretical aspects of fluid flow within the hydrocyclone. The emphasis then shifted to modelling of an empirical nature. Lately, since the advent of more powerful computers, theoretical modelling has become popular once again. In this section, the various theoretical and empirical models will be reviewed. Since the most widely used and accepted models are empirical in nature, the emphasis will be on these models.

2.4.1 Theoretical Hydrocyclone Modelling of Cut Size (d_{50})

Three hypotheses have been used in theoretical attempts to relate d_{50} to the operating and design variables. These three approaches are discussed below.

(a) Equilibrium Orbit Hypothesis

As mentioned in section 2.2, two forces act upon solid particles in the horizontal plane. In the development of the equilibrium orbit theory, the d_{50} particle size was assumed to be the size for which there was a zero net velocity in the horizontal plane, i.e. the particle size for which $F_c = F_d$. By equating the drag and centrifugal forces at equilibrium, the relationship indicated in equation 10 was developed.

$$d_{50} = \left[\frac{18\mu W r}{(\rho_s - \rho)V^2}\right]^{0.5} \tag{10}$$

where μ = viscosity of carrier fluid
W = radial velocity
r = radius of orbit at equilibrium
V = tangential velocity
ρ_s = density of solid particles
ρ = density of carrier fluid

To solve for d_{50} in equation 10, it is necessary to determine W, V and r. Kelsall (1952) conducted research into the flow patterns within the hydrocyclone. This led to the identification of the envelope of zero vertical velocity, which is situated at the interface between the upward rotating spiral of the overflow and the downward movement of the underflow. Lilge (1962) suggested that the d_{50} particle size would fall on the envelope of zero vertical velocity and assumed that this occurred at a radius of 1/6 of the cyclone radius. This assumption led to the establishment of Lilge's cone force equation, shown in equation 11.

$$d_{50} = \left(\frac{39 D_i^{0.87} D_c^{1.13}}{\left(1 - \frac{D_i}{D_c}\right)^{0.8}}\right)\sqrt{\frac{(1 - R_v)\mu}{(\rho_s - \rho)Qh}} \tag{11}$$

where R_v = recovery of feed volume to the underflow
Q = volumetric flowrate of feed slurry

Equation 11 is semi-empirical in nature and is of limited practical use because of the difficulty in predicting slurry viscosities and the fact that it is only useful when applied to

slurries of low solids content. Bradley (1959) developed an equation along the same lines as the cone force equation. Bradley's equation is shown in equation 12.

$$d_{50} = \frac{3(0.38)^n D_i^2}{K_c}\left[\frac{\tan\left(\frac{\theta}{2}\right)\mu(1-R_f)}{D_c Q(\rho_s - \rho)}\right]^{0.5} \tag{12}$$

where K_c = a parameter dependent on the design and fluid characteristics
n = the power of the radius r in the tangential velocity distribution equation (Vr^n = constant)
θ = hydrocyclone cone angle

Practically, the only difference between equation 11 and 12, is the point at which the locus of zero vertical velocity was defined. For industrial hydrocyclones, both equations are of limited applicability because of the reasons previously mentioned.

(b) Retention Time Hypothesis

Rietema (1962) claimed that the equilibrium orbit approach was invalid, because of the short residence times in the hydrocyclone. He hypothesised that particles have insufficient time to reach equilibrium. Rietema proposed that the d_{50} particle is that particle which enters at the centre of the inlet, and just succeeds in reaching the wall at the spigot. Equation 13 was obtained by assuming laminar flow conditions and considering the movement of the d_{50} particle over a distance $0.5D_i$ in the available residence time.

$$d_{50}^2\left(\frac{(\rho_s - \rho)L_c P}{\rho\mu Q}\right) = \frac{36vD_c}{\pi VD_i} \tag{13}$$

where L_c = total length of the hydrocyclone
P = pressure drop over the hydrocyclone
v = axial (vertical) velocity

The right hand side of equation 13 is approximately constant and Rietema called it the cyclone characteristic number, Cy_{50}. Equation 13 only holds if the underflow discharge is small and the slurry concentration is low.

(c) Crowding Theory

The third approach to theoretical hydrocyclone modelling was developed by Fahlstrom (1963). He proposed that the spigot restricts the flow of solids and that under the effect of centrifugal force, the likelihood of a particle exiting through the spigot is determined by its size (in homogeneous slurries). This means that the largest particles are discharged first and that it becomes progressively more difficult for the discharge of smaller particles as the particle size becomes finer. The empirical model shown by equation 14 was obtained from

laboratory data after it was observed that d_{50} is inversely proportional to the recovery of solids to the underflow.

$$d_{50} = k_o\left(1-\frac{U}{F}\right)^{\frac{1}{n_1}} \tag{14}$$

for $0.4 < U/F < 1$

where k_o = maximum particle size on a sieve analysis plot
n_1 = gradient of sieve analysis plot

Equation 14 is more suitable for slurries of high solids content. Under these conditions, the crowding effect becomes more pronounced.

2.4.2 Empirical Hydrocyclone Modelling

The empirical models that have found the most application in industry are those proposed by Lynch and Rao (1975) and Plitt (1976). These two models will be discussed at length and a summary of other modelling approaches will be given.

(a) A summary of early empirical modelling approaches

Dahlstrom (1954) developed one of the earliest empirical expressions for d_{50}. Equation 15 was developed by using a 9 inch diameter hydrocyclone and a dilute feed slurry.

$$d_{50} = \frac{81(D_o D_i)^{0.68}}{Q^{0.53}}\left(\frac{1.73}{\rho_s-\rho}\right)^{0.5} \tag{15}$$

Yoshioka and Hotta (1955) used dilute slurries and small hydrocyclones to develop the expression shown in equation 16.

$$d_{50} = \frac{6300 D_c^{0.1} D_i^{0.6} D_o^{0.8} \mu^{0.5}}{Q^{0.5}(\rho_s-\rho)^{0.5}} \tag{16}$$

Both of the aforementioned equations are restricted to low feed solids content (< 20%), a condition hardly typical of an industrial application. Equation 15 can however be extended to slurries of solids content of up to 35% by including a term for the apparent slurry density and viscosity.

It is perhaps pertinent to note that all the correlations for d_{50} (theoretical and empirical) mentioned thus far, contain a common term, $(\rho_s - \rho)^{-0.5}$, indicating laminar flow conditions.

Another important parameter in determining hydrocyclone performance is the flow split, S, which is defined in equation 17.

$$S = \frac{Q_u}{Q_o} \tag{17}$$

where Q_u = volumetric flowrate of the underflow
Q_o = volumetric flowrate of the overflow

Most empirical relations for S are of the form of equation 18.

$$S = \left(\frac{D_u}{D_o}\right)^x \tag{18}$$

where $x = 4$ (Yoshioka and Hotta, 1955)

Dahlstrom (1954) developed the relation for flow split by including the volumetric feed flowrate term, shown in equation 19.

$$S = \frac{6.13}{Q^{0.44}}\left(\frac{D_u}{D_o}\right)^{4.4} \tag{19}$$

In practice, the flow split is controlled by adjusting the spigot to vortex finder ratio. Equation 18 and 19 are thus in agreement with what is physically expected.

Knowledge of the pressure drop across the hydrocyclone is necessary for calculating pumping requirements. Dahlstrom developed a correlation to calculate pressure drop (equation 20).

$$\frac{Q}{P^{0.5}} = 7.4 \times 10^{-6}\left(\frac{D_o}{D_i}\right)^{0.9} \tag{20}$$

Chaston (1958) developed an extremely simple expression, shown in equation 21, which according to Plitt (1976), predicts the pressure drop within 20%.

$$Q = 10 A_i P^{0.5} \tag{21}$$

where A_i = cross-sectional area of the inlet

Until 1976 there had been no attempt to mathematically express the sharpness of separation (exponent m in equation 6 to 8). However, it is clear that m is determined by the hydrocyclone dimensions, as well as the feed material. Fahlstrom (1963) noted that the sharpness of separation increases as the solids content of the slurry decreases. Fahlstrom also presented an empirical method for obtaining a specified cut size at the maximum possible efficiency. He called this cut size the nominal separating size and found that it was attainable for all hydrocyclones by adjusting the design variables.

(b) Hydrocyclone model by Lynch and Rao

Lynch and Rao (1975) conducted a series of 220 tests on 38.1 cm, 25.4 cm, 15.2 cm and 10.2 cm Krebs hydrocyclones to study the effect of operational and design variables on hydrocyclone performance. Three different feed size distributions were used (coarse, medium and fine). Depending on the feed size, a different relation was determined for each size. In the following section, only the general expressions are given. All expressions were determined by regression analysis.

Bypass (R_f)

Lynch and Rao assumed that Kelsall's expression (equation 2) was valid for the solids bypass. The resultant expression is shown in equation 22.

$$R_f = \frac{K_1 D_u}{(1-\varphi_v)Q} - \frac{K_2}{(1-\varphi_v)Q} - K_3 \qquad (22)$$

where φ_v = volumetric fraction of solids in the feed
K_1 to K_3 are constant model parameters

Corrected cut size (d_{50c})

The corrected cut size was modelled by the logarithmic expression shown in equation 23.

$$log(d_{50c}) = K_4 D_o - K_5 D_u + K_6 D_i + K_7 \varphi_s - K_8 Q \qquad (23)$$

where φ_s = mass percent solids in the feed
K_4 **to** K_8 are constant model parameters

Pressure drop / capacity relationship

Since only an approximate expression relating pressure to capacity is required for hydrocyclone design, equation 24 is of a general form (no tuning of constants required).

$$Q = 6 D_o^{0.73} D_i^{0.86} P^{0.42} \qquad (24)$$

Lynch and Rao chose to use the form of partition curve shown in equation 6 to model the hydrocyclone classification. It should be noted that the model proposed by Lynch and Rao does not contain an explicit relation for the sharpness of separation (m) used in equation 6. Tests thus have to be conducted on the required hydrocyclone to determine m as well as the calibration constants K_1 to K_8 in equation 22 and 23.

The model includes a solids concentration term in equation 22 and 23. This enables the model to be applied over a wide range of feed densities (not as restrictive as earlier models). Lynch and Rao chose to calculate R_f directly, thus eliminating the need to calculate the flow split (S). By using three different size distributions in the feed, it was found that depending on the size distribution, different constants (K_1 to K_8) were calculated by regression analysis. This implies that any model equation is limited to a narrow range in feed size distribution.

(c) Plitt model

Plitt (1976) performed tests using 1.25 inch, 2.5 inch and 6 inch diameter hydrocyclones. Plitt conducted his own tests and used experimental data from Lynch and Rao. Together this represented a data set of 297 hydrocyclone tests. In addition to the dimensional variables investigated by Lynch and Rao (D_i, D_o and D_u), the free vortex height (h) and hydrocyclone diameter (D_c) were also included. The same operational variables, namely, feed flowrate and density were considered. In all cases the relevant data were processed by using step-wise multiple linear regression.

Flow split (S) and bypass R_f

Plitt also assumed that equation 2 was valid for predicting entrained solids. Plitt chose to first calculate S and then use S to calculate R_f. The relations for S and R_f are shown in equation 25 and 26 respectively.

$$S = \left(\frac{18.62 \cdot \rho_p^{0.24} \cdot \left(D_u / D_o \right)^{3.31} \cdot h^{0.54} \cdot \left(D_u^2 + D_o^2 \right)^{0.36} \cdot e^{[0.0054 \cdot \varphi]}}{D_c^{1.11} \cdot P^{0.24}} \right) \tag{25}$$

$$R_f = \frac{\left(\frac{S}{1+S} \right) - \frac{R_s \varphi}{100}}{1 - \frac{\varphi}{100}} \tag{26}$$

where R_s = mass recovery of feed solids to the underflow
φ = volumetric percent solids in the feed

Since there is a dependence of R_s on R_f in equation 26, it requires an iterative solution.

Corrected cut size (d_{50c})

Plitt determined equation 27 to be the most suitable for predicting d_{50c}.

$$d_{50c} = \left(\frac{397 \cdot D_c^{0.46} \cdot D_i^{0.6} \cdot D_o^{1.21} \cdot \eta^{0.5} \cdot e^{0.063 \cdot \varphi}}{D_u^{0.71} \cdot h^{0.38} \cdot Q^{0.45} \cdot \left([\rho_s - 1] / 1.6 \right)^k} \right) \tag{27}$$

where $k = 0.5$ (assuming laminar flow)

It was discovered that the cut size was most sensitive to changes in the solids content of the feed (φ). Plitt attributed this observation to the crowding theory.

Pressure drop

It was found that $P \propto Q^{1.78}$ (equation 28). This suggests that the flow within the hydrocyclone follows the Blasius relationship which states that $P \propto Q^{1.75}$ for laminar flow.

$$P = \left(\frac{1.88 \cdot Q^{1.78} \cdot e^{0.0055 \cdot \varphi}}{D_c^{0.37} \cdot D_i^{0.94} \cdot h^{0.28} \cdot \left[D_u^{\,2} + D_o^{\,2} \right]^{0.87}} \right) \tag{28}$$

Sharpness of separation (m)

Surprisingly, Plitt's equation for m (equation 29) does not contain the variable φ. It seems reasonable that the sharpness of separation should decrease as the solids content increases due to crowding of the spigot. However, Plitt found that m was primarily influenced by the D_u / D_o ratio and the retention time within the hydrocyclone.

$$m = \left(1.94 \cdot \left[\frac{D_c^{\,2} \cdot h}{Q} \right]^{0.15} \cdot e^{\left[\frac{-1.58 \cdot S}{1+S} \right]} \right) \tag{29}$$

Plitt chose the partition curve shown in equation 7 to describe the classification. The actual classification curve is obtained by using equation 9 and the aforementioned relations for R_f, d_{50c} and m.

The Plitt model can be used (with fair accuracy) without adjusting any of the model parameters. In this way it has an advantage over the Lynch model. It also requires less tuning of model parameters if calibration tests are performed. Changes in feed size distribution are not taken into account, which could lead to errors in predicting hydrocyclone classification.

(d) Subsequent adaptations to the Plitt model

Luckie and Klimpel (1986) have shown that under certain conditions, the Plitt model does not guarantee a mass balance over the hydrocyclone. Plitt et al (1987) published a review of hydrocyclone literature to date. In this paper, the Plitt model equations (27, 29, 28, and 25) are pre-multiplied by a correction factor (F_1 to F_4 respectively) to facilitate the calibration of these equations for a particular industrial hydrocyclone. Plitt proposed the default value of each of these correction factors to be unity. However, this was shown to be incorrect by Cilliers and Hinde (1991), who found F_4, the correction factor for flow split to be a default of 0.176.

Hinde (1977) developed an analytic technique for calculating the water recovery to the underflow (R_f), instead of using the iterative approach of Plitt. Hinde's water recovery equation is shown in equation 30.

$$R_f = \left(\frac{\left[S/(1+S) \right] - \left(\varphi/100 \right) \cdot \left\{ 1 - \sum_{i=1}^{n} f_i \cdot e \left[-0.6931 \cdot \left(d_i / d_{50c} \right)^m \right] \right\}}{1 - \left(\varphi/100 \right) \cdot \left\{ 1 - \sum_{i=1}^{n} f_i \cdot e \left[-0.6931 \cdot \left(d_i / d_{50c} \right)^m \right] \right\}} \right) \qquad (30)$$

Later Plitt et al (1989) published a technical note for calculating the water recovery (R_f) in the same way as originally proposed by Hinde.

(e) Other empirical models for the hydrocyclone

Empirical modelling of the hydrocyclone has become relatively stagnant over the last 15 years. This can be attributed to the general acceptance of the model by Lynch and Rao and the Plitt model. These models have been incorporated into simulation packages and may be used for preliminary plant design.

An independent hydrocyclone model was developed by Nageswararao in 1978 as part of his Ph.D. thesis. All that has been published regarding this model is a brief technical note (Nageswararao, 1995). This model has been incorporated in JKSimMet, a steady state simulator developed by JKTech. Nageswararao postulated that industrial hydrocyclones can be described by a set of equations using scale factors [ ], design variables { }, and operating variables ( ). These expressions are illustrated in equations 31-34 below.

Pressure - Throughput relationship

$$Q = K_Q\left[D_c^{1.9}\right]\left\{\left(\frac{D_o}{D_c}\right)^{0.67}\left(\frac{D_i}{D_c}\right)^{0.45}\left(\frac{L}{D_c}\right)^{0.2}\left(\frac{1}{\theta}\right)^{0.1}\right\}\left\{\left(\frac{P}{\rho_p}\right)^{0.5}\right\} \quad (31)$$

where K_Q = material specific constant

L = length of cylindrical section of the hydrocyclone

Corrected Cut Size (d_{50c})

$$\frac{d_{50c}}{D_c} = K_d\left[\left(\frac{1}{D_c}\right)^{0.65}\right]\left\{\left(\frac{D_o}{D_c}\right)^{0.52}\left(\frac{D_u}{D_c}\right)^{-0.47}\left(\frac{D_i}{D_c}\right)^{-0.4}\left(\frac{L}{D_c}\right)^{0.15}\theta^{0.2}\right\}\left\{\left(\frac{P}{\rho_p g D_c}\right)^{-0.22}\varphi_v^{\,0.93}\right\} \quad (32)$$

where K_d = material specific constant

g = gravitational acceleration

Water recovery to the underflow (R_f)

$$R_f = K_w\left[D_c^{\,0.0}\right]\left\{\left(\frac{D_o}{D_c}\right)^{-1.19}\left(\frac{D_u}{D_c}\right)^{2.4}\left(\frac{D_i}{D_c}\right)^{-0.5}\left(\frac{L}{D_c}\right)^{0.22}\theta^{-0.24}\right\}\left\{\left(\frac{P}{\rho_p g D_c}\right)^{-0.53}\varphi_v^{\,0.27}\right\} \quad (33)$$

where K_w = material specific constant

Volumetric recovery of feed slurry to the underflow (R_v)

$$R_v = K_v\left[D_c^{\,0.0}\right]\left\{\left(\frac{D_o}{D_c}\right)^{-0.96}\left(\frac{D_u}{D_c}\right)^{1.83}\left(\frac{D_i}{D_c}\right)^{-0.25}\left(\frac{L}{D_c}\right)^{0.22}\theta^{-0.24}\right\}\left\{\left(\frac{P}{\rho_p g D_c}\right)^{-0.31}\right\} \quad (34)$$

where K_v = material specific constant

Most of the empirical hydrocyclone models of the last few years are simply adapted versions of the Lynch and Rao or Plitt models for a particular plant hydrocyclone. An example of the adaptation of the Lynch model for modelling the classification of complex nickel ore, is given by Gault and Nageswararao (1981).

Gupta and Eren (1990) developed one of the few independent empirical models. A novel test rig was set up, using a pneumatically adjustable vortex finder (the free vortex height could be varied automatically). The spigot diameter could also be adjusted pneumatically. The variation in d_{50c} was modelled by considering changes in D_u, h, φ, Q and the temperature (T).

A step-wise approach to modelling d_{50c} was used. One variable was varied at a time at discrete values of the other variables. In this way a family of curves were generated for each variable. It should be noted that temperature was considered in the final model (not previously used by Plitt and Lynch). The deviation in T from the normal temperature (T_n) was used. It was found that over a temperature range of 20°C, d_{50c} only varied by 2.24 μm. Experimental results indicated that d_{50c} was most sensitive to changes in D_u and φ and less sensitive to Q, h and T. For this reason, a basic equation for d_{50c} was developed as a function of D_u and φ at average values of the other variables. The deviation in d_{50c} predicted by the 2 variable model at values of Q, h and T other than the average was used to identify multiplicative adjustment factors for d_{50c}. The final model form is illustrated in equation 35.

$$d_{50c} = f(D_u, \varphi) \times QF \times HF \times TF \tag{35}$$

where QF, HF and TF are adjustment factors for Q, h and T respectively

2.4.3 Present and Future Hydrocyclone Modelling Trends

While empirical modelling of the hydrocyclone has taken a back-seat, theoretical modelling approaches have become more popular. However, it will be a few years before these models are capable of satisfactorily predicting hydrocyclone classification. In recent years, empirical modelling has taken an application specific approach. An example of this is the use of artificial intelligence techniques to model a particular hydrocyclone for the purposes of predicting the particle size in the overflow. This approach to modelling will be discussed in the next section, which deals with the development of soft-sensors.

White (1991) developed a simple hydrocyclone model, based on the crowding theory, for predicting particle size distributions in the overflow and underflow. Perhaps one of the most important advances in fundamental modelling has come from Monredon, Hsieh and Rajamani, who have concentrated on the physics of fluid flow within the hydrocyclone. Their research involved the physical identification of flow patterns by using a laser-Doppler velocimeter. The resultant velocity profile data were used to solve the hydrocyclone transport equations and incorporated into a fluid-flow model (Hsieh, 1988; Hsieh and Rajamani, 1988, 1991). Monredon, Hsieh and Rajamani (1992) used the fluid-flow model to describe the partition curve for a number of hydrocyclones of differing sizes. The model was found to predict very accurately for particle sizes below the cut size, while above the cut size, the predicted partition factor was always higher than the actual value. This deviation was attributed to the model not taking short-circuit flow into account (short-circuit flow carries

coarse particles to the overflow, i.e. these particles are not classified). This research has shown the large potential for hydrocyclone modelling from first principles.

2.5 SOFT SENSORS FOR PARTICLE SIZE ESTIMATION

As previously mentioned, the development of soft sensors that estimate the particle size in the hydrocyclone overflow, is becoming more popular. A soft sensor is a device which calculates its output from other measurement inputs rather than physically measuring the output. Since the aim of this Masters project is to develop a better estimation algorithm for MINTEK's Particle Size Estimator (*PSE*), a thorough review of the current *PSE* model will be given in the next section. This will be followed by an overview of other soft sensor approaches developed for the calculation of overflow particle size.

2.5.1 Mintek's Particle Size Estimator (PSE)

The Particle Size Estimator is a system, which as its name implies, estimates the particle size in the overflow product stream of classification hydrocyclones. The particle size estimation algorithm is based on a modified form of the Plitt model, and is empirical in nature. In this section, the particle size estimation algorithm incorporated within the *PSE* will be discussed.

(a) Industrial test work and implementation

The particle size estimation algorithm was developed by Viljoen (1993). Comprehensive test work and numerical analysis were performed to establish and validate the work. The final result was a model that predicted the cumulative percentage passing a particular particle size class for the hydrocyclone overflow stream. Subsequently the *PSE* has been patented by MINTEK and implemented on milling circuits in South Africa and in Chile.

Determining the significance of model parameters

Tests were conducted on a gold plant where on-line particle size measurement by means of a PSM was available. Initially, the measured particle size was modelled by a linear function of feed flowrate and feed density derived from a linearised Plitt model. This 2-parameter model is however limited since it accounts for only 2 degrees of freedom, viz. changing feed flowrate and density. Figure 3 illustrates the deviation between measured and predicted particle size over the long term (>12 hours) for the 2-parameter model. It was determined

that the discrepancy between the model prediction and measured particle size can be attributed to a change in the hydrocyclone feed size distribution, the third degree of freedom. To provide a model capable of reliably predicting particle size over the long-term it is necessary to include a third input variable. Test work was conducted to determine the influence of underflow angle on the particle size estimate of the hydrocyclone product stream. The underflow angle is measured by an instrument, developed by MINTEK, known as the Hydrocyclone Underflow Meter (*HUM*). The *HUM* is a mechanical device, of robust construction, which measures the angle of the underflow flare by contact with it. The *HUM* produces a standard 4-20 mA signal. It was found that the underflow angle measurement from the *HUM* could be modelled accurately by using a linear function of feed flowrate and feed density. Analogous to the 2-parameter particle size model, it was discovered that the underflow angle model was only applicable for short-term data (< 12 hours) as shown in Figure 4.

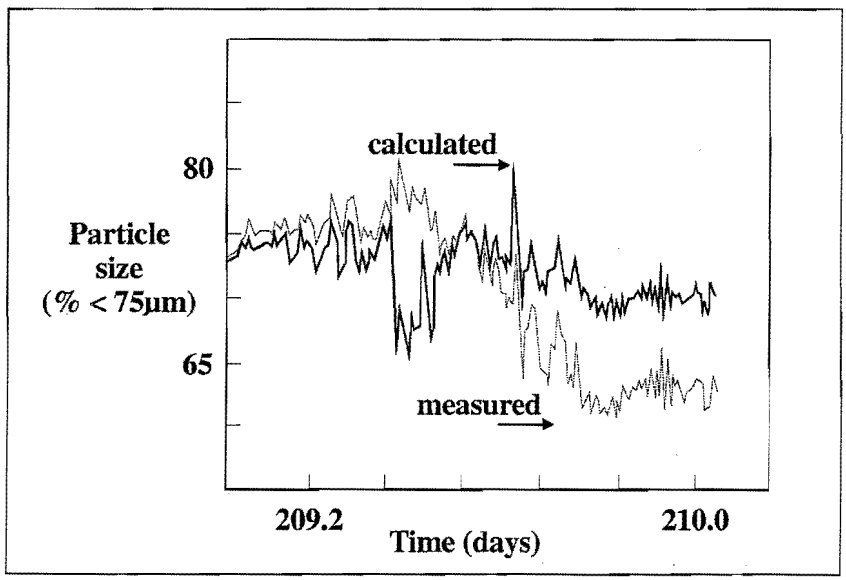


Figure 3. Particle size as a function of flowrate and density (Viljoen 1993)

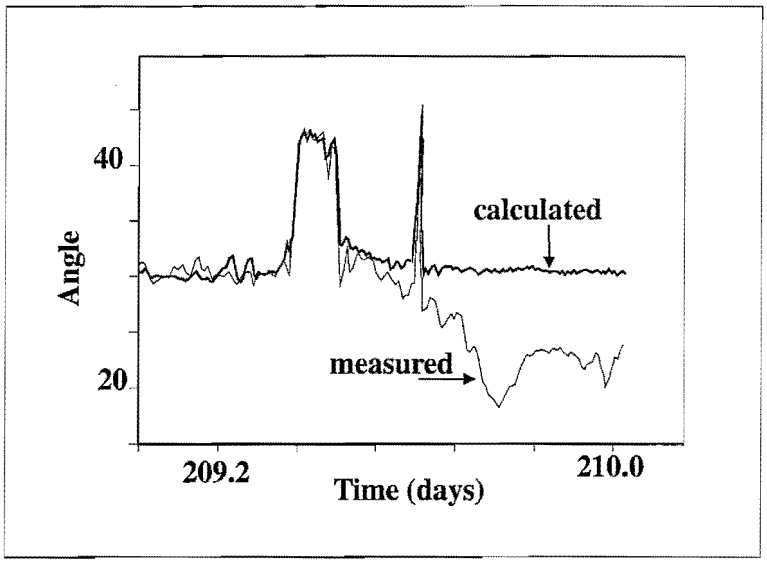


Figure 4. Model for underflow angle (Viljoen, 1993)

From a comparison between Figure 3 and 4 it appears that the deviation in angle corresponds to the deviation in product particle size. It was hypothesised that the deviation between the calculated and measured underflow angle could be used to correct the 2-parameter particle size model, thereby accounting for the third degree of freedom, viz. changing hydrocyclone feed size distribution.

Therefore, to predict particle size accurately over the long term, a 3-parameter model was proposed. In this way, particle size was modelled as a function of hydrocyclone feed flowrate, density and underflow angle. The improvement of this model over the 2-parameter model of Figure 3 is illustrated in Figure 5.

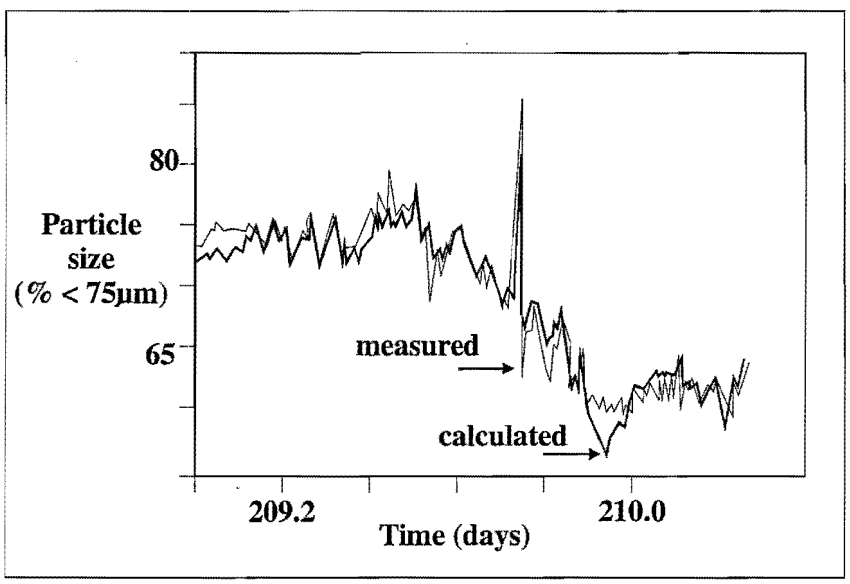


Figure 5. Particle size as a function of flowrate, density and underflow angle (Viljoen, 1993)

A description of the current *PSE* model and calibration technique

The inaugural step in calibrating the *PSE* is to conduct a mass balance around the hydrocyclone. This involves taking samples of the hydrocyclone feed, underflow and overflow streams. The samples are analysed by conventional sieve analysis (by using a standard sieve series ranging from 1000μm to 38μm) (see Chapter 4). A mass balance employing a Rossin-Rammler smoothing function is conducted for each size class, which enables a corrected partition curve to be derived. From the partition curve the following constants can be determined: R_f, m, d_{50c} and S. The previous 4 constants enable the calibration of the Plitt model equations 25 to 29, i.e. the estimation of the correction factors F_1 to F_4 (see Appendix A.2). The Plitt model is used to generate particle size data by changing feed flowrate and density which represent inputs into the model. Linear regression is used to derive a relationship of the following form (see section 5.1.1):

$$PSE_n(\ predicted\ ) = a \cdot CFF_n + b \cdot CFD_n \tag{36}$$

where ***CFF*** = cyclone feed flowrate
CFD = cyclone feed density
PSE = particle size estimate (% -75μm)

In equation 36, the subscript n is used to show that the variables are normalised i.e. the mean is deducted from each data point. Equation 36 can be used to estimate particle size provided that the only disturbances are due to feed flowrate and density.

In order to account for other variations, such as a change in the feed size distribution, it is necessary to use the cyclone underflow angle to modify the 2-parameter model of equation 36. Dynamic step tests, which involve changing cyclone feed flowrate and density so that the complete operational range is covered, are performed and overflow samples are taken. The overflow samples are analysed for the % -75μm (see Chapter 4). At the same time the plant data is logged, the most important variables are *CFF*, *CFD* and *CUA*. A zero order ARX model, calibrated on logged plant data, is used to find the relationship shown in equation 37 below (see section 5.1.2). A description of the ARX model can be found in Appendix A.5.

$$CUA_n(\ predicted\ ) = d \cdot CFF_n + e \cdot CFD_n \tag{37}$$

where ***CUA*** = Cyclone Underflow Angle

After conducting dynamic step tests on the plant, the residuals shown in equation 38 and 39 are calculated for each sample taken.

$$\Delta CUA_n = CUA_n(actual) - CUA_n(predicted) \quad (38)$$

$$\Delta PSE_n = PSE_n(actual) - PSE_n(predicted) \quad (39)$$

The residuals are plotted against each other (see Figure 6). The slope of the straight line in Figure 6 (determined by standard linear regression) determines the value of the constant K_{uf} (see section 5.1.3).

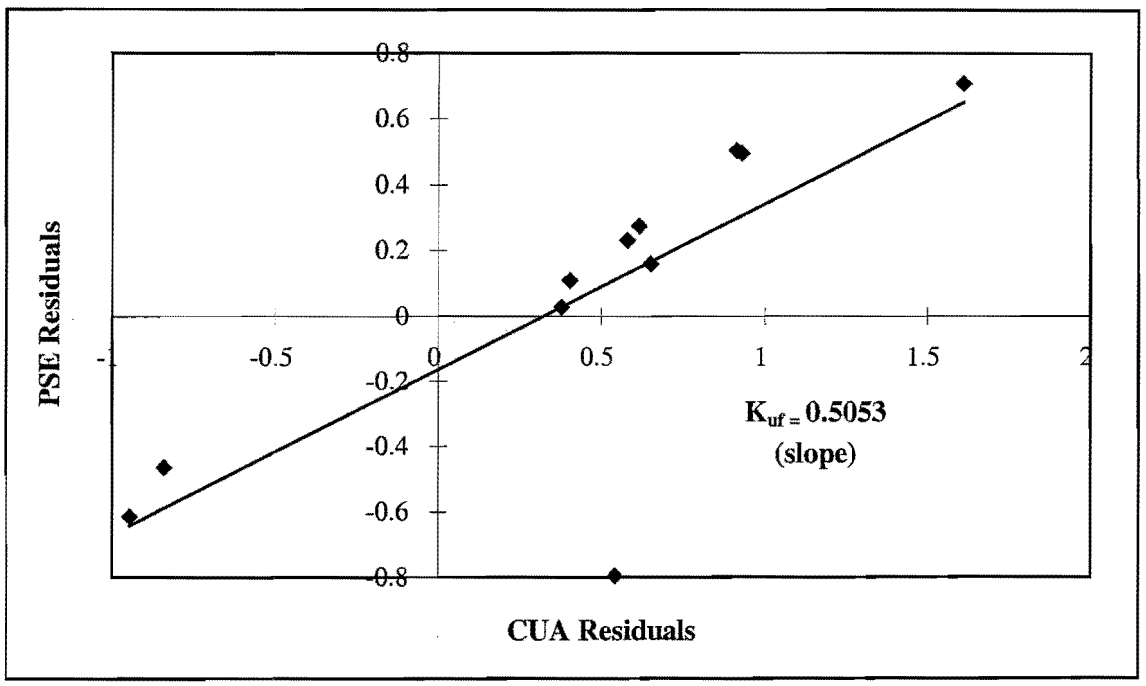


Figure 6. Residual plot for K_{uf} determination (Viljoen, 1993)

The final model for particle size estimation (shown in Figure 7) uses the underflow angle residual (from equation 38) to correct the particle size estimated by the 2 parameter model, the correction factor being ΔPSE_n. The magnitude of this adjustment is determined by the value of the constant K_{uf} (equation 40).

$$\Delta PSE_n = K_{uf} \cdot \Delta CUA_n \quad (40)$$

The final model incorporates hydrocyclone feed flowrate, feed density and underflow angle and is shown in broken-down form in equation 41.

$$PSE_n = a \cdot CFF_n + b \cdot CFD_n + K_{uf} \cdot \left[CUA_n - \left(d \cdot CFF_n + e \cdot CFD_n\right)\right] \quad (41)$$

(b) Summary of model structure and constants for the *PSE* model

The model structure and calculation method is illustrated in Figure 7. As shown, 5 constants are required to describe the model. The role of each of these constants is briefly summarised:

- ***a*** and ***b*** : These constants are derived by performing a statistical mass balance around the hydrocyclone and using the results in the modified version of the Plitt model. *a* is used to account for variations in feed flowrate, while *b* accounts for changes in feed density.
- ***d*** and ***e*** : The hydrocyclone underflow angle measurement from the Hydrocyclone Underflow Meter (*HUM*) is modelled by a linear function in 2 variables. *d* is used to quantify the relationship between feed flowrate and underflow angle, while *e* predicts the angle change due to variations in feed density.
- K_{uf} : This constant is used to correct the particle size estimated by the Plitt model, using the deviation between calculated and measured (actual) values of hydrocyclone underflow angle.

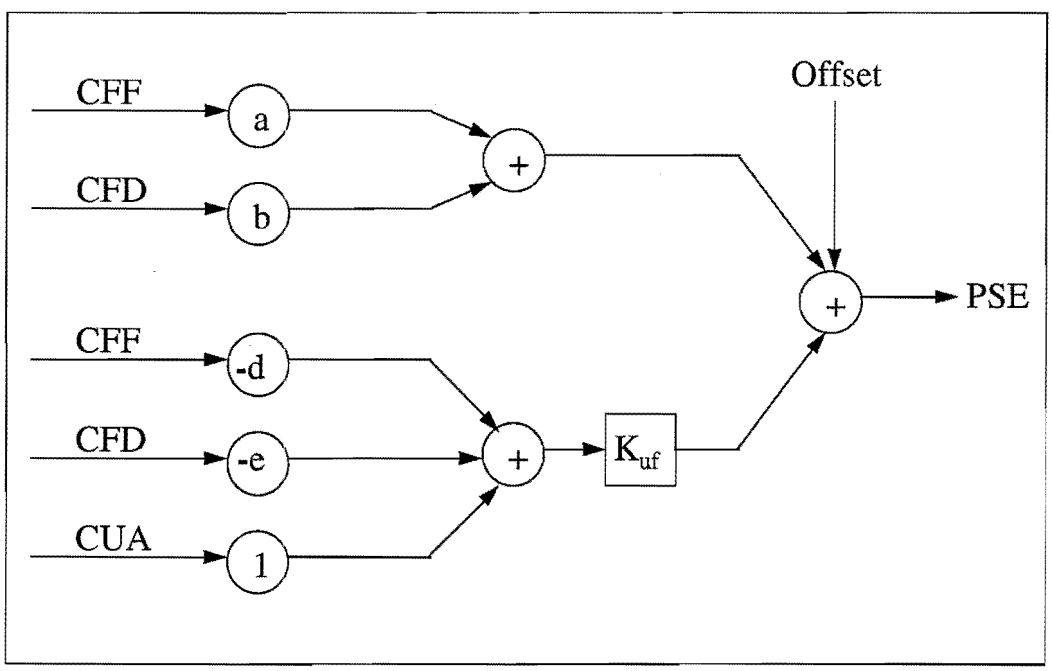


Figure 7. *PSE* Algorithm

The *PSE* algorithm also makes provision for measurement drift, by allowing an adjustable offset value. The offset is automatically calculated by calibrating the *PSE* once a week. The weekly calibration procedure involves sampling the overflow and determining the percentage passing a specific size class by conventional sieve analysis. The true value is then compared to the *PSE* prediction over the period at which the sample was taken, and the *PSE* is automatically adjusted.

(c) Problems associated with the *PSE* model

The results shown in Figure 6 could not be repeated in further industrial tests, implying that the use of the underflow angle residual term to correct the 2-parameter particle size estimate has resulted in negligible benefit in terms of model accuracy. This lack of correlation has re-occurred on several industrial implementations of the *PSE* and is an important reason for conducting this research. An example of this is given in section 5.1.3.

Furthermore, the Plitt model provides poor estimates of particle size when high hydrocyclone feed densities, typical on most plants, occur.

It seems more reasonable to replace the underflow angle residual term in equation 41 with the measured value of the underflow angle (from the HUM), since the predicted angle (equation 37) was calibrated by using the measured angle, and is likely to be affected strongly by noise on the flowrate and density measurements. Furthermore, the underflow angle term may not significantly contribute to enhancing model accuracy due to poor correlation with the particle size. Other measured variables should be assessed to determine if a better correlation with particle size could be obtained. Due to the poor performance of the Plitt model, a simple linear combination of the flowrate and density terms, calibrated by step test data, may provide more accurate results. This removes reliance on mass balance analysis and hence the Plitt model.

Chapter 5 will compare the existing PSE model to the above suggestions. Other model types and inputs will also be investigated.

2.5.2 Overview of other Soft Sensor Approaches

Most approaches in the estimation of hydrocyclone overflow particle size make use of existing models for the hydrocyclone. The Lynch and Plitt models form the basis of soft sensor estimation techniques, which make use of the classifier characteristics for on-line particle size analysis. These models describe the partition curve by providing relations for parameters such as d_{50c}. Seitz and Kawatra (1984) found from fundamentals, that the mass percent passing size D (PMD) can be expressed as:

$$PMD = \frac{\sum_{i=0}^{D} f_i\left(1 - p_i'\right)}{\sum_{i=0}^{i_{max}} f_i\left(1 - p_i'\right)} \times 100 \qquad (42)$$

where f_i = fraction of particles of size class i in the feed

The partition function, p_i' can be expressed as a function containing constants d_{50c} and m. The feed size distribution may be expressed as a Gaudin-Schuhmann distribution with distribution modulus α. From the above, a semi-empirical model for *PMD* can be developed. Seitz and Kawatra (1984) have used the previous analytical evaluation to estimate *PMD* in a more plant specific manner, by formulating a purely empirical expression shown in equation 43. Plant data is used to obtain m and d_{50c} as a function of operational variables as previously discussed. These expressions can then be substituted into equation 43 to obtain *PMD*.

$$\log PMD = c_0 - c_1 \log d_{50c} \qquad (43)$$

where c_0 and c_1 are functions of m and d_{50c}, the form of which is determined by regression analysis of experimental data

Plitt and Kawatra (1979) developed a technique for estimating the d_{50} of a hydrocyclone without performing overflow and underflow size analysis. The cut size was predicted within 10 % of the actual value by using the corrected solids recovery and feed size distribution. The inverse of this model was used by Osborne (1972) to predict the feed size distribution of a well instrumented hydrocyclone with known cut size. This procedure formed the basis of the London School of Mines particle size analyser.

An important application of soft sensors for particle size estimation is to serve as a back-up system for physical size measurement instruments. This is useful, since control systems frequently rely on on-line particle size measurements (where available). The control system's performance deteriorates when the size signal is not available. Gonzalez and Meyer (1990) developed an estimation technique that uses the signal generated by the size measuring device as well as other input signals to predict the overflow particle size. The estimation method is based on an auto recursive identification method. When the measuring device is on-line, the model uses past values of both particle size (measured) and other variables to determine the best model parameters for size estimation. This process is continually updated, until the size measuring device fails, the model parameters are then frozen at their last value. The estimated particle size is then used in the control loop until the size measurement comes on-line again.

As previously mentioned, artificial intelligence techniques have become increasingly popular for modelling poorly understood process equipment. Van der Walt et al (1993) developed a neural network model to simulate hydrocyclone classification. Data were generated by using the model of Lynch and Rao, after which noise was added to the data. Inputs to the network were operational variables, viz. mass percent solids of the feed and feed flowrate, and design variables, viz. vortex finder, spigot and inlet diameters. The complete size distribution (consisting of 10 discrete sieve values) represented the output of the model. Three different models were investigated. A sigmoidal backpropogation neural network (SBNN), containing non-linear transfer functions in the hidden layer was compared to a SBNN with a single linear hidden node (similar to standard linear regression techniques). The third modelling approach used was based on multiple adaptive regression splines (MARS). The MARS model fits truncated cubic functions to the modelled data in a forward/backward stepwise regression procedure. The performance of each modelling approach was evaluated by using a test set of data. The non-linear SBNN outperformed the other modelling techniques. Although it is difficult to derive conclusions from a technique presented that uses different types of models to predict the output of another model (the data were generated by a model), the results indicate that the classification characteristics of a hydrocyclone are inherently non-linear. This becomes clear when the performance of the linear and non-linear SBNN's are compared.

Del Villar et al (1996) compared modelling the % -45μm in the hydrocyclone overflow by using a neural network, an auto regressive moving average (ARMA) model and a Kalman filter. Model inputs included flowrate and density of both the feed and overflow streams. The simulation package, DYNAFRAG was used to generate plant data. Noise was added to the data in 3 levels viz. no noise, moderate noise and extreme noise. In the presence of noisy data (typical of the plant environment) the ARMA model deteriorates markedly. A properly initialised Kalman filter indicated performance comparable to the neural network model; however the Kalman filter requires more expertise and effort on the part of the user. The neural network model was a feedforward network containing hyperbolic tangent transfer functions in the hidden layer. Both a static model, using only the most recent inputs, and a past input/output model, using the most recent inputs as well as the immediate past input and output data as inputs, were investigated. A simple network containing 3 hidden nodes was found to generalise well. The past input/output model outperformed the static model when the sampling horizon (time between samples) was large. Once again, the failure of the linear modelling technique (ARMA) indicates the need to model the hydrocyclone product size by non-linear methods.

As outlined previously, the PSE model has provided questionable results when applied to industrial data. This can probably be attributed to the underflow angle term being insignificant, or poor model structure. Subsequent chapters will deal with gathering of real plant data to be used for comparing and identifying models suitable for use in an industrial soft-sensor. The existing PSE model will be compared to other modelling techniques. More specifically, various forms of neural network and multiple-linear models will be assessed. Furthermore, the significance of model inputs will be examined. The objective is to develop a more accurate and robust soft-sensor, suitable for industrial application, to replace the PSE model.

CHAPTER 3. ARTIFICIAL NEURAL NETWORKS

3.1 HISTORY AND DEVELOPMENT OF NEURAL NETWORKS

The development of neural network modelling approaches was inspired by the study of the human brain. Neural networks are analogous to the brain in that they consist of neurons (or nodes) connected *completely* in parallel to all other nodes. Also analogous to the human brain is the fact that nodes are connected to each other by numerical linkages known as weights (or synapses in the case of the brain).

Research into the subject of neural networks began in the 1940s. Initial research efforts were pioneered by McCulloch and Pitts, and Hebb (Aldrich, 1995). In 1958, Rosenblatt presented the perceptron, an approach that is still used today (Roth, 1988). From 1969, research into the area of neural networks became less popular, especially after Minsky and Papert proved that a linear, 2-layered network structure, (the most popular structure of the time) was severely limited in terms of accuracy for certain applications (Roth, 1988).

Intensive neural network research was only revived again in the 1980s. Since then there has been almost an exponential growth in the development of artificial intelligence (neural networks and expert systems) applications, particularly for the modelling and interpretation of ill-defined systems.

3.2 THE MULTILAYER PERCEPTRON

The Multilayer Perceptron (MLP) is one of the most widely implemented neural network topologies. It has been estimated that of all neural network applications in the process industries, the MLP makes up approximately 87% (Aldrich, 1995). The MLP can be used to approximate virtually any modelling problem and for this reason they are discussed in detail in this review. Until recently, MLPs have been used mainly for pattern recognition and classification problems, however recent attempts to use MLPs as a modelling tool have shown great promise.

3.2.1 The Structure of the Multilayer Perceptron

A typical MLP is a feed forward neural network consisting of 3 distinct layers, viz. input, hidden and output layers. Each layer consists of processing elements or nodes. Depending on the complexity of the modelling problem, there may be several hidden layers. The number of nodes in the input and output layers is determined by the number of input and output variables respectively. Figure 8 shows the basic structure of a typical MLP containing 3 inputs, 1 hidden layer consisting of 4 nodes, and 2 outputs. In shorthand notation, this is referred to as a 3:4:2 network, referring to the number of nodes in each layer.

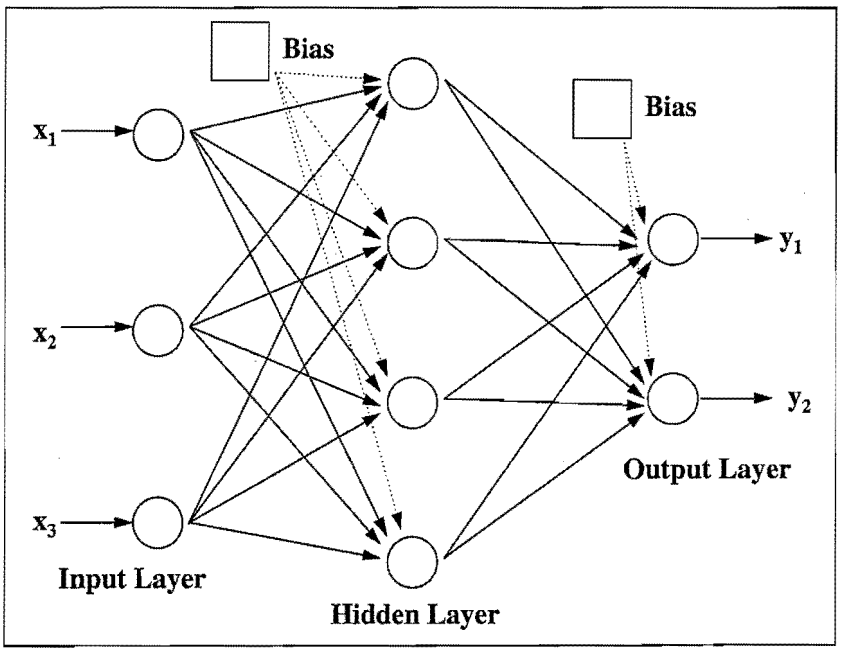


Figure 8. An example of a 3:4:2 MLP with bias nodes

Each node is connected to every other node in the next layer by adjustable numeric values known as weights. The nodes in the hidden and output layers perform 2 mathematical functions, viz. summation and transformation. These operations are illustrated by means of an example in Figure 9, using the inputs and outputs of node 2 in the hidden layer. The equations for the summation and transformation operations are shown in equation 44 and 45.

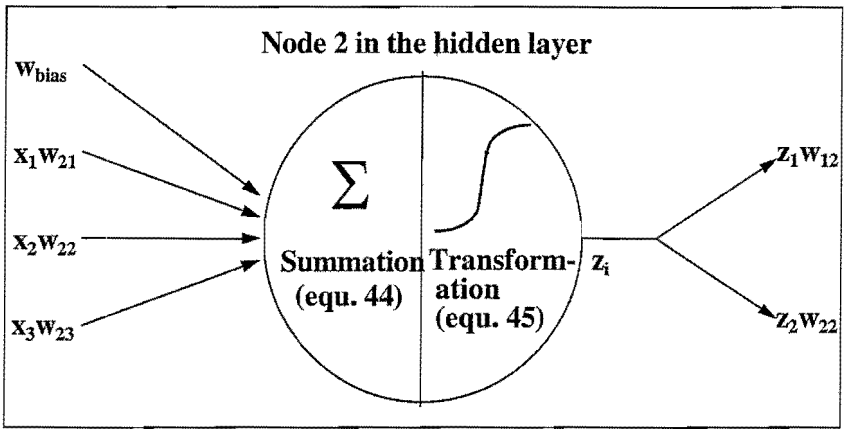


Figure 9. An example of all inputs, outputs and mathematical operations occurring at node 2 in the hidden layer

$$s_i = \sum_{j=1}^{n} w_{ij} x_j + w_{bias} \tag{44}$$

where w_{ij} = the weight connecting node i to node j in the previous layer
x_j = the input to node i originating from the output of node j in the previous layer

The transformation function, $f(s_i)$ in equation 45 and 46 may take on several forms from linear to non-linear functions. In most cases non-linear transformation functions are used. Generally, the form of the non-linearity is a smooth monotonically increasing and saturating function. Equations 45 and 46 show some of the most commonly used non-linear transformations.

$$z_i = f(s_i) = \tanh(s_i) \tag{45}$$

$$z_i = f(s_i) = \frac{1}{1 + e^{-s_i}} \tag{46}$$

Equation 46 is known as the sigmoidal or logistic function. Figure 10 illustrates the "squashing" nature of these functions.

After the input of node i in the hidden layer has been summed and transformed, it exits node i and is split into K branches, where K refers to the number of nodes in the output layer. A new matrix of weights is defined for connecting each node in the output layer with each node in the hidden layer (w_{ki}). The output layer also contains nodes with summation and transformation functions, which may or may not be of the same form as found in the hidden layer.

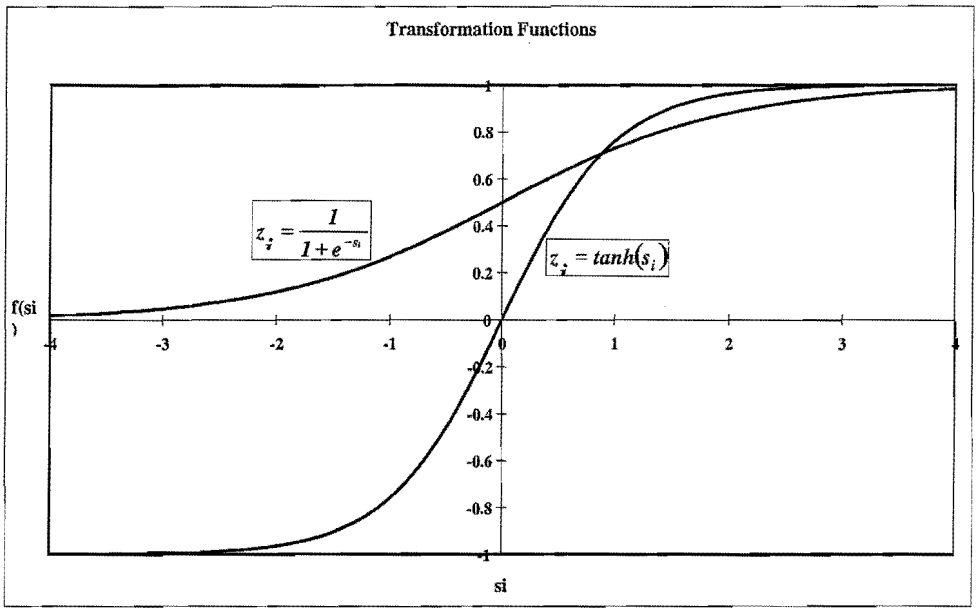


Figure 10. Illustration of most common transformation functions

A bias is shown in Figure 8. The network treats the bias node as an extra input with a value of 1. The bias node is connected to each node of the next layer by an adjustable weight (w_{bias}). The reason for using a bias is to add an extra degree of freedom, it is analogous to the *y intercept* in a linear equation, i.e. it allows the resultant plane in multi-dimensional space to be moved in another dimension.

3.2.2 Scaling of Input and Output Data

Due to the saturating nature of the transformation functions of equation 45 and 46, it is necessary to scale all input and output data presented to the network. When referring to Figure 10 it can be seen that both functions have lost almost all of their sensitivity when $s_i > 3$ or $s_i < -3$. Thus, to maximise the transfer of information between layers, it is important to scale the raw input data in such a way that the summed input (from equation 44) into the nodes of the hidden and output layer remains between -3 and 3. From the above discussion, it is clear that there is no hard rule for determining the extremes of input data, however, scaling the raw inputs independently for each variable between -1 and 1 is often a good starting point. Equation 47 represents a typical linear scaling function (NeuroSolutions User Manual, 1995). An equivalent linear scaling function was defined by Van der Walt, et al, 1993. If data are orders of magnitude apart, logarithmic scaling is recommended (Van der Walt, et al, 1993).

$$x_{sc} = \left(\frac{max_{sc} - min_{sc}}{max_{ac} - min_{ac}}\right) \cdot x_{ac} + max_{sc} - \left(\frac{max_{sc} - min_{sc}}{max_{ac} - min_{ac}}\right) \cdot max_{ac} \quad (47)$$

where x_{sc} = scaled variable
x_{ac} = actual value of variable
max_{sc} = maximum value for the scaled variable
min_{sc} = minimum value for the scaled variable
max_{ac} = maximum value of the actual variable
min_{ac} = minimum value of the actual variable

Suppose a hyperbolic tangent transformation function was used in the output layer. This function produces outputs (which are also outputs of the network) that range between -1 and 1. For calculation and backpropagation of the error, it is necessary to scale the actual outputs between -0.9 and 0.9 so that the network is not forced to saturation by attempting to minimise the error between the network and a value of -1 or 1. The desired (actual) output data are scaled by using equation 47 above.

3.2.3 Training and Testing of a Neural Network

Much like the human brain, neural networks have to be trained to perform optimally. MLP's are trained by backpropagating the error between the calculated network outputs and the actual or desired outputs. Training refers to the iterative procedure whereby the network weights are adjusted to minimise the output error.

The output error is defined in equation 48 as the difference between the desired output and the output as calculated by the network.

$$e_j(n) = d_j(n) - y_j(n) \quad (48)$$

where $e_j(n)$ = error at the j^{th} output node for the n^{th} data pattern
d = desired output
y = output as calculated by the network

It is often useful to calculate the error based on the complete data set (1 epoch). This error is called the cost, and is normally calculated as a mean square error (E) (equation 49).

$$E = \frac{1}{N}\left[\sum_{i=1}^{N}\left(d_j(i) - y_j(i)\right)^2\right] \quad (49)$$

where E = mean square error or cost
N = total number of data points or patterns

The cost is a function of the free parameters in the system (the weights and the bias). The cost function is a measure of the performance of the network. The weights of the network can thus be adjusted to minimise E, the cost function.

Gradient descent learning is the most widely applied principle for neural network training. The reason for this is that only basic computation is required to implement this method and that only local information is required to calculate the gradient. When plotted, the cost or error function (E) represents a surface. The gradient of a surface points to the direction of the maximum rate of change. A minimum (whether local or global) is found at the point where the gradient is zero (i.e. $dE/dw = 0$). Therefore, if the weights are moved in the opposite direction to the gradient, the system will approach points where the error surface is flatter (the cost function is minimised). This concept is illustrated in Figure 11.

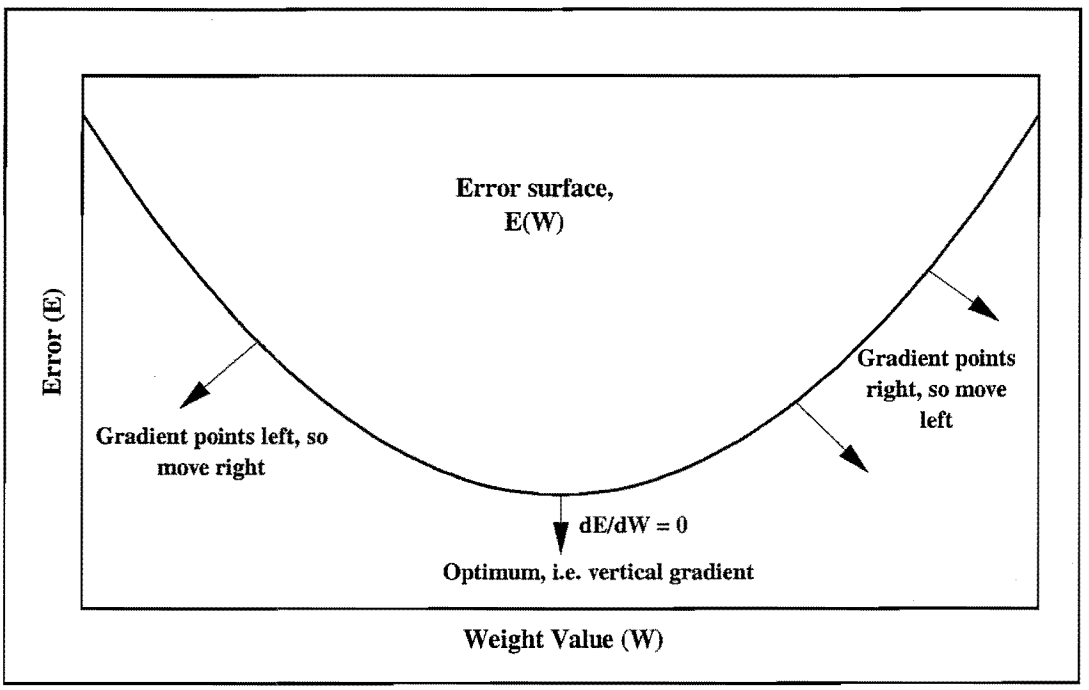


Figure 11. Illustration of the gradient descent procedure in 1-dimension

Due to the non-linearity contained within the nodes, the so-called Delta Rule is employed to calculate the gradient of the error function (Bhat and McAvoy, 1990 and Aldrich, 1995). This principle has been used for a long time in sensitivity analysis. The sensitivity (rate of change) of one quantity (E) with respect to another (w), related by a function $f(w)$, can be calculated using the chain rule, provided $f(w)$ is differentiable.

$$\frac{\partial E}{\partial w} = \frac{dE}{df}\frac{df}{dw} \tag{50}$$

The function $f(w)$ represents the non-linear processing element. The non-linearity of the processing element must be monotonic and smooth, otherwise the above equation cannot be evaluated. Once the error gradient has been determined, a new value for the weight can be calculated using Equation 51 below (Baldi, 1992 and NeuroSolutions User Manual, 1995).

$$w_{ij}(n+1) = w_{ij}(n) + \beta \cdot \delta_i(n) \cdot x_j(n) \tag{51}$$

where $\delta_i(n)$ = local error gradient at the node = $\frac{\partial E}{\partial w}$

$x_j(n)$ = input at weight $w_{ij}(n)$

β = step size (for the iterative procedure)

The above technique will regard a local minima in the error surface as a solution (i.e. $\delta_i(n)=0$). To overcome local minima entrapment, a momentum term is included in equation 51, the result is shown in equation 52 (Jacobs, 1988 and NeuroSolutions User Manual, 1995). The momentum term forces the training through any local minima.

$$w_{ij}(n+1) = w_{ij}(n) + \beta \cdot \delta_i(n) \cdot x_j(n) + \alpha\left(w_{ij}(n) - w_{ij}(n-1)\right) \tag{52}$$

where α = momentum (usually set between 0.1 and 0.9)

Momentum learning is an improvement to the straight gradient descent in the sense that a memory term is used (the past increment of the weight). Including this memory term speeds up and stabilises convergence (Jacobs, 1988).

The weights can either be updated by using the error calculated by introducing the network to a single input-output data pair (on-line learning), or the entire set of input and output data (an epoch) can be introduced to the network and the weights adjusted by calculating the average error (batch learning).

To begin training, the network weights need to be initialised. Unless other process knowledge exists, small random numbers, typically between -0.1 and 0.1, are best suited for initial weight values (Bhat and McAvoy, 1990 and Aldrich, 1995). Random numbers are used to avoid symmetry conditions that can trap the search algorithm. If the weight values are large, the transformation function used can saturate. If the transformation saturates, the change in error that propagates through the network becomes zero. The use of small initial weight values will ensure that every transformation function will operate in the linear region of the sigmoid at the beginning of learning.

It is essential that the network model be validated or tested either after or during training. A representative sample is taken from the total data set. This sample is not used during network training and is known as the test data set. Test data validates the network model and provides the best stop criteria during training. After an epoch of training data is sent through the network, the error is calculated for the training data based on the weight values obtained for that epoch. An epoch of test data is then passed through the network. The test error is calculated based on the weight values obtained during the training phase. A typical network performance curve is shown in Figure 12.

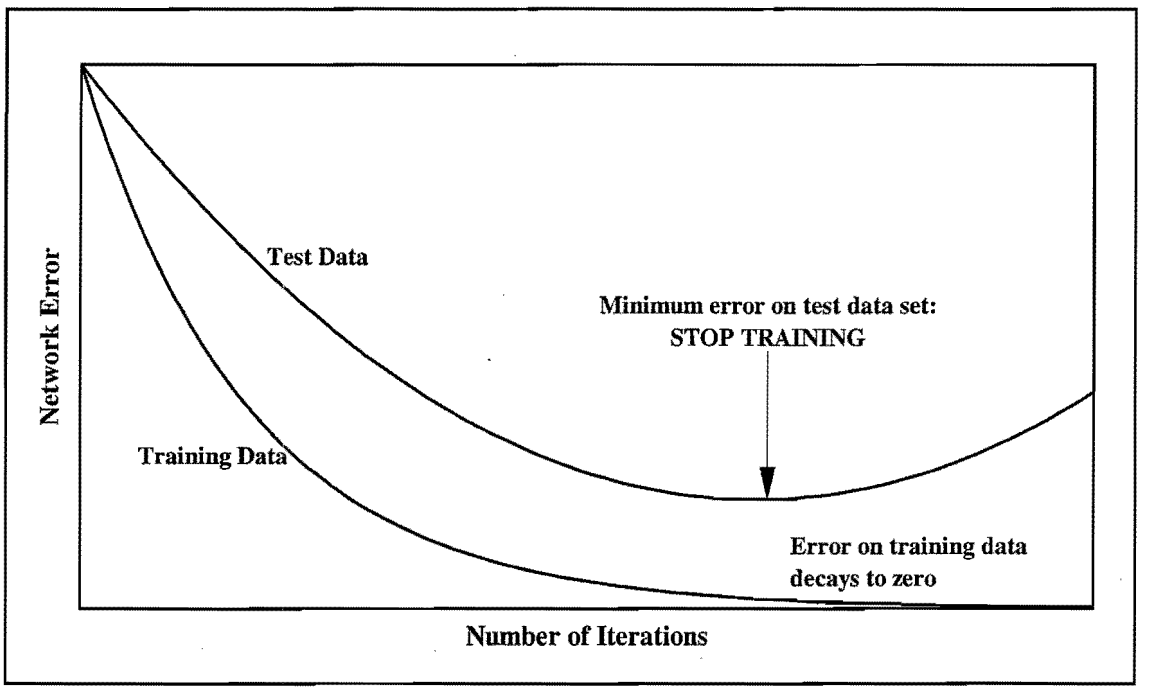


Figure 12. Typical training and test error curves illustrating the stop criteria

As the number of iterations increases, the training error continues to decrease. If the network is trained for to many iterations, model generalisation becomes poor, i.e. the network model is not able to interpolate accurately. Poor generalisation is represented in Figure 12 by an increase in the test error. For a network to have the best generalisation, training should be stopped at the minimum test error. This procedure is known as cross validation.

3.2.4 Practical issues regarding Training and Network Architecture

The accuracy of a neural network model is highly dependent on the number of independent data points presented to the network as well as the network architecture. Like all empirical

models, the final network model is only a compact representation of the training data, and hence is highly dependent on the quality and quantity of the training data. Equation 53 illustrates a rule of thumb for determining the approximate number of independent data points required to yield a desired degree of accuracy (NeuroSolutions User Manual, 1995).

$$N_d = \frac{N_w}{e} \tag{53}$$

where N_d = number of independent data points required
N_w = number of network weights
e = network error

Another rule of thumb for determining data requirements is that the network should contain at least 2 to 5 input-output data pairs per network weight (Aldrich, 1995).

The network architecture (or structure) for the input and output layers is determined by the number of inputs and outputs respectively. In a MLP it is possible to change the number of nodes in the hidden layer, as well as the number of hidden layers. Normally, when the number of nodes are increased, the training error decreases, however the performance should always be validated with a test data set to ensure generalisation. Choosing the correct network configuration is an iterative procedure. The performance of various network architectures should be analysed by comparing test errors. A good starting point for choosing the number of hidden nodes is shown by equation 54 for a network with a single hidden layer (Baum and Haussler, 1989 and Aldrich, 1995).

$$N_{hn} = N_{in} + 1 \tag{54}$$

where N_{hn} = number of hidden nodes
N_{in} = number of input nodes

Generally, the smallest network that provides an acceptable level of accuracy should be used to ensure good generalisation.

During network training, the weight values may be caught in a local minimum, for this reason, several training runs should be conducted, each with different initial random weight values. If there is no reduction in the network error and the error is unacceptably high, the network training has probably been caught in a local minimum. To force the training out of the local minima, the network weights should be randomised or the momentum term, α in equation 52, should be increased.

Training normally begins with network weights that are far from the optimum. To accelerate the initial stages of training, a large value of β (equation 51 and 52) should be used. As training progresses, β should be decreased to fine-tune the network weights.

A few practical implementations of neural networks pertaining to hydrocyclone modelling are given in section 2.5.2.

CHAPTER 4. EXPERIMENTAL PROCEDURE AND DATA ANALYSIS

4.1 DESCRIPTION OF MILLING CIRCUIT USED AS TEST SITE

All plant test work was conducted on a South African gold plant. The milling section consists of 3 milling circuits, each mill operating in closed loop with a hydrocyclone classifier. For consistency, only 1 mill was used throughout the test campaign. Figure 13 shows a flow diagram of the milling circuit and associated instrumentation.

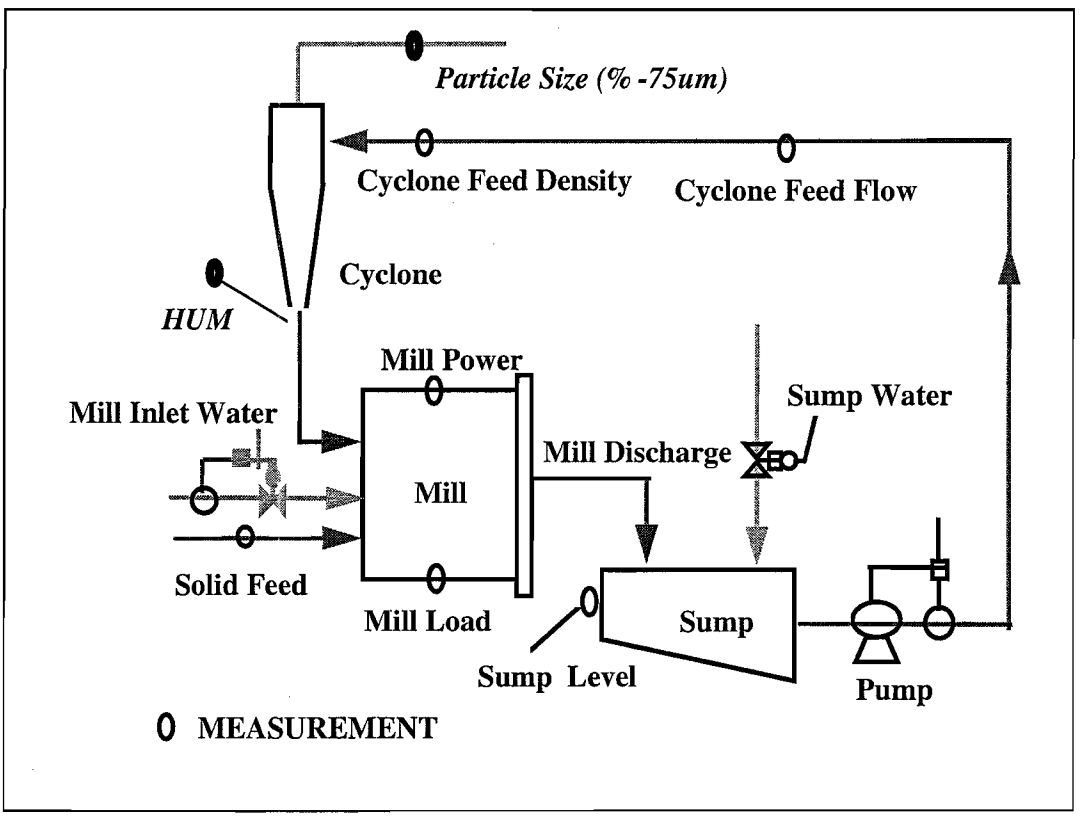


Figure 13. Milling circuit flow diagram and instrumentation

Included among the conventional measuring instruments is a hydrocyclone underflow meter (HUM), an instrument used to measure the angle of the underflow flare by means of a contact paddle. The paddle rests on the cone shaped flare and the angular movement of the paddle as the flare changes is converted via a transducer into a 4-20 mA analogue output signal.

A fixed geometry hydrocyclone, of dimensions shown in Table 1, was used to obtain classification data.

Geometric Variable	Dimension (cm)
Hydrocyclone diameter, D_c	91.4
Spigot diameter, D_u	9
Vortex finder diameter, D_o	30
Equivalent inlet diameter, D_i	24
Free vortex height, h	266.8

Table 1. Hydrocyclone dimensions

Typical operating conditions for the mill and hydrocyclone are shown in Table 2.

Operating variable	Operating range
Mill solids feed rate (t/hr)	100-120
Mill power (kW)	2500-2700
Hydrocyclone feed flowrate (m^3/hr)	350-500
Hydrocyclone feed density (t/m^3)	1.30-1.35

Table 2. Typical plant operating conditions

4.2 DATA SAMPLING AND RECORDING REQUIREMENTS

Data sampling refers to the process of recording available plant measurements over the period of the test work. The test work assumes the form of perturbation step tests, where a test structure is designed and implemented as an experiment. In simple terms, the scope of the experiment requires that an input variable (cyclone feed flow or sump feed water) be stepped up or down by a known quantity, and the effects in cyclone feed flowrate, density and underflow angle be logged at equal, 10 second intervals. A 10 second interval is a sufficiently small time interval to detect the system dynamics. For the purposes of test work, the following operational variables were logged at 10 second intervals:

- Mill solids feed rate (t/hr)
- Mill power (kW)
- Mill load (t)
- Sump level (% full)
- Cyclone feed pressure (kPa)
- Cyclone feed flowrate (m^3/hr)
- Cyclone feed density (t/m^3)
- Cyclone underflow angle (degrees)

A typical section of logged data is shown in Appendix A1.

Step tests to change the feed flowrate (by changing the pump speed) and feed densities (by changing the discharge-sump dilution water flowrate) were performed with the sump in open-loop, i.e. the level was allowed to move between 0 and 100 % full. The length of the steps must be long enough to approach steady state (i.e. well settled state). In the case of integrator (sump level) responses, the length and size of the step change is restricted by the capacity of the process unit (sump). Hence, the step length and step size, in either pump speed or water flow, must be traded off to achieve the maximum size step, while still allowing the output variable, particle size, to approach steady state within the step period.

4.3 REPRESENTATIVE SAMPLING

The term representative sampling refers to the reliability of the sampling procedure in providing objective physical data. Typically the process streams of interest are devoid of automatic sampling devices, and thus impartiality has to be approximated by employing proven techniques and observing certain rules in an attempt to eliminate bias. The samples are taken from the following streams:

- Mill discharge stream
- Cyclone underflow stream
- Cyclone overflow stream

The mill discharge stream represents the cyclone feed flow stream prior to water addition in the sump. The mill discharge stream is sampled since it comprises the particle size profile of the cyclone feed stream, while being more accessible in terms of sampling. The sample is typically cut from the stream using a sample cutter.

The following subsection provides a guide to important factors to note with regard to physical sampling of a flowing slurry stream:

- The sampling device must intercept the entire cross section of the stream at any given time, during the sampling process.
- It is important to retain equal percentages of slurry from the entire stream, this necessitates that the sample cutter be moved through the stream at a uniform rate, while maintaining the blades at 90° to the flow, and ensuring that the sampler be moved completely across and out of the stream.
- The distance between the cutting edges of the sampling device should be great enough to allow the largest particles free passage between them. Typically, the distance between

the blades is set to at least 3 times the diameter of the largest particle expected in the stream. This can be estimated by taking an initial sample at the maximum aperture.

- The speed at which the sampler is moved through the stream should not be great enough to knock particles away which should be retained as part of the sample. Generally, the sample cutter movement should be uniform, at approximately 2 m/s.
- The cutting edges should be sharp (cutting edge at 45° to the plane of the blade), so that particles have an equal chance of falling to either side of the cutting edge.

4.3.1 Mass Balance Sampling Requirements

Samples are taken while the milling circuit is at steady state, which is usually determined rather subjectively, where the steady state criterion relies on a visual inspection of the graphical trends of plant variables. Ideally, the operational variables shown in Table 2 should show little variation in relation to the mean values, and the rate of change for each variable should be close to zero.

The steady state criterion is important due to the physical logistics of the sampling procedure. Sampling is usually performed by one person. This limitation means that each stream must be sampled individually, and hence separately, thus necessitating that the composition and flowrate of each stream be constant (steady state).

4.3.2 Step Test Sampling Requirements

Samples are taken from the cyclone overflow and underflow at the end of each step, to represent the mean particle size and density of the overflow as well as the underflow density as a result of changes caused by the step change. This implies that the sample is taken momentarily before the subsequent step change is initiated.

4.4 SAMPLE PREPARATION

The sample preparation procedure followed is outlined below in a few precise steps:

- The masses of the wet samples are recorded.
- To prevent clogging and damage to the sieves, the mill discharge and cyclone underflow samples are washed through a 1mm screen. All the +1mm particles are dried and the dry

mass is recorded. The -1mm sample is then filtered and dried, and the dry mass is recorded. By using simple ratios, the +1mm fraction mass corresponding to the sample size cut (see below) is added back (mathematically) after complete sample analysis (see 4.5).

- The -1mm sample is passed through a 1mm screen again, with a soft rubber stopper used to break any lumps (conglomerates of previously independent particles). The sample is then blended.
- Approximately 100g-150g of sample is cut and used for the sizing analysis.

4.5 SAMPLE ANALYSIS

The sample (100g-150g) is deslimed (washed on a vibrating 38μm screen), which removes the bulk of the -38μm particles. The sizing sieves are typically a stack of 8 sieves (425μm, 300μm, 212μm, 150μm, 106μm, 75μm,53μm and 38μm) and a receiver pan below the final 38μm sieve. The process of sizing proceeds as follows :

- Weigh and record the mass of each clean sieve prior to sizing.
- Place the stack of sieves in decreasing order of mesh size on the sieve shaker, and place the deslimed +38μm sample on the top (425μm) sieve. The sieve stack is secured and vibrated for 25 minutes.
- Weigh and record the mass of each sieve including the retained sample. The -38μm fraction will be the difference between the sample mass cut for sizing and the deslimed sample, together with any -38μm material remaining on the receiver pan after the sieve stack is vibrated.

It should be noted that the above procedure is only valid for mass balance samples where a complete size distribution is required. The sample preparation procedure for the overflow samples resulting from step changes is identical to the procedure outlined in section 4.4, however sample analysis only needs to be conducted for one size class. The procedure for determining the particle size of the overflow stream is as follows:

- Wash the 100-150g cut sample over a vibrating 75μm sieve (the sieve mass is recorded).
- The sieve and +75μm particles remaining on the sieve are dried and weighed. This allows for the determination of the -75μm mass by difference.

During the step test procedure, underflow and overflow samples are taken for density analysis. The density is determined as a mass percent solids by the following procedure:

- Record the wet mass of the entire sample.
- Dry the entire sample and record the dry mass.

The percent solids can then be calculated by the following formula:

$$\%solids = \frac{drymass}{wetmass} \times 100 \tag{55}$$

4.6 DATA PRE-TREATMENT

Both mass balance and step test results need to be treated to ensure that the information gained will be of use in the modelling phase.

4.6.1 Mass Balance Smoothing

All experimental techniques, especially sampling, result in the generation of scattered data. Scatter in sampled data becomes more evident when partition numbers are calculated by equation 1, since the scatter of the data compounds itself. Performance parameters (d_{50c}, m, S and R_f) need to be interpolated from the partition curve, which may lead to large errors if scattered data is used.

Woollacott et al (1992) developed a mass balance smoothing procedure and incorporated it into a computer program known as *Backfill*. The *Backfill* software employs a model based smoothing technique that fits an S-shaped partition curve to the unsmoothed mass balance data. Either the Lynch or the Plitt model can be chosen. In all cases, the Rosin-Rammler smoothing function (Plitt), shown in equation 7 , was used. The use of this equation resulted in best fit estimates of the parameters shown in Table 3. The actual mass balance results, smoothing procedure and calculation of Plitt model correction factors are summarised in Appendix A.2.

Solids Flowrate	
• Feed	1
• Underflow	0.523
• Overflow	0.477
% Solids (by mass)	
• Feed	38.339 %
• Underflow	76.063 %
• Overflow	24.834 %
Hydrocyclone Partition Parameters	
• d_{50c}	87.803 μm
• m	3.1997
• R_f	10.233 %
• S	0.2211

Table 3. Best fit hydrocyclone parameter estimates

The smoothed corrected partition curve for the *Backfill* mass balance is shown in Figure 14.

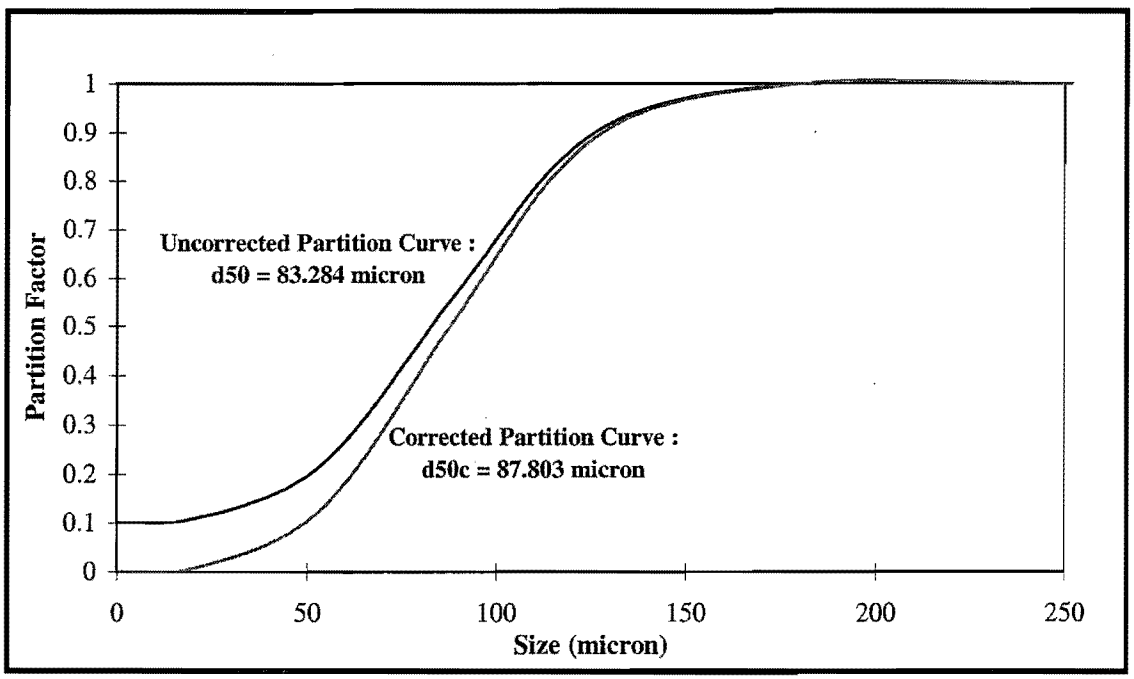


Figure 14. Partition curves from mass balance results

4.6.2 Identification of Step Tests

For modelling purposes, it is extremely important to associate the correct values of the independent, logged (10 second) variables, viz. hydrocyclone feed flowrate, density and

underflow angle, with the overflow particle size and density, as well as the underflow density, derived from physical sampling.

After each step change in feed flowrate or density, the hydrocyclone is allowed a suitable period of time to reach steady state. The steady state condition is maintained for a period of approximately 10 minutes if possible. Since system dynamics are typically in the order of seconds, it is not necessary to take account of time lags between independent and dependent variables.

The time at which each sample was taken is recorded. Since time is logged, it is possible to associate the exact values of logged variables (flowrate, density and underflow angle) with sampled data (overflow particle size and density and underflow density). Since most measurements are inherently noisy, and because time lags between variables become important when using spot values, the instantaneous logged values corresponding to sample times are not used. Instead, an approximately 5 minute steady state period is identified graphically (by using custom written Matlab m-files). The average values and standard deviation is calculated for each steady state stretch. The results of the step tests performed are shown in Appendix A.3.

4.6.3 Dividing Data into Training and Test Data Sets

As mentioned in section 3.2.3, it is necessary to test or validate neural network performance by using an independent test data set that is not used in the training phase. The test data set will also be used to compare the performance of the various modelling approaches that will be investigated in Chapter 5, hence providing a useful model selection criteria.

The total 71 sample data set shown in Appendix A.3, was separated into a training set (50 samples) and a test set (21 samples). To ensure unbiased allocation of data to the test data set, the following procedure was utilised to ensure that both training and test data originated from the same population as the total data set:

1. 21 random integers between 1 and 71 were generated. The total data set was numbered from 1 to 71 as shown in Appendix A.3. Row numbers corresponding to the 21 random numbers generated were used to constitute the test data set, while the remaining 50 rows represented the training data.

2. To ensure that the training and test data are representative samples of the total population, t-tests were performed. In all cases *CFF*, *CFD*, *CUA* and *PS* (where *PS* represents the sampled particle size in % -75μm) were tested to determine whether their respective means (μ) are similar at a 97.5% confidence interval by using a 2 tail t-test. The Null Hypothesis is as follows:

$$\mu_{test/train} \neq \mu_{population}$$

The Null Hypothesis is rejected when:

$$t < t_{0.0125} \text{ or } t > -t_{0.0125}$$

Table 4 and 5 show some of the important statistical parameters.

	CFF	*CFD*	*CUA*	*PS*
Population Mean	402.18	1.2924	42.391	74.260
Sample Mean	396.43	1.2960	43.555	73.561
Population Variance	5899.9	0.0042	85.959	66.472
Sample Variance	6975.9	0.0048	70.563	68.012
$t_{0.0125}$	1.9840	1.9837	1.9813	1.9828
t	0.3850	-0.2823	-0.7190	0.4611

Table 4. Statistical parameters - training data

	CFF	*CFD*	*CUA*	*PS*
Population Mean	402.18	1.2924	42.391	74.260
Sample Mean	415.86	1.2840	39.619	75.924
Population Variance	5899.9	0.0042	85.959	66.472
Sample Variance	3279.7	0.0029	116.52	61.897
$t_{0.0125}$	2.0167	2.0227	2.0452	2.0322
t	-0.8843	0.5966	1.0662	-0.8442

Table 5. Statistical parameters - test data

In all cases the Null Hypothesis was rejected, proving that both training and test data are representative of the total population for all variables tested. The resultant training and test data are shown in Appendix A.4.

CHAPTER 5. DEVELOPMENT OF A MODEL FOR PARTICLE SIZE PREDICTION

5.1 CALIBRATING THE MINTEK PSE MODEL

It is necessary to calibrate the MINTEK Particle Size Estimator (*PSE*) to establish a base case for evaluating model performance. In all cases subsequent models will be evaluated relative to the *PSE* model, and each other, by comparing the respective test mean square errors (*MSE*) and coefficient of determination or correlation coefficients (R^2). The *MSE* was defined previously by equation 51 and the R^2 is defined in Appendix A.6.

Since a mass balance and numerous step tests were performed at the test site, adequate data were available to allow calculation of the *PSE* model constants.

5.1.1 Calculation of Constants a and b

Mass balance sampling of hydrocyclone feed, overflow and underflow streams at steady state enabled calibration of the Plitt model equations (equations 25-30). Smoothed mass balance results were used, as provided by the *Backfill* software. The best-fit hydrocyclone parameter estimates are shown in Table 3. This information allowed the calculation of the Plitt model calibration constants F_1, F_2, F_3 and F_4 by using the *Eureka* software to simultaneously solve the system of 5 non-linear equations. The mass balance smoothing technique and Plitt model calibration is further explained in Appendix A.2. Table 6 shows the Plitt model calibration constants.

F_1	0.44514
F_2	0.91820
F_3	1.10066
F_4	0.93124

Table 6. Plitt model calibration constants

To determine the respective relationships between operating variables, hydrocyclone feed flow and density, and particle size, expressed as % -75μm, it is necessary to generate particle size data at discrete flowrate and density values by using the Plitt model. The partition curve defined by equation 7 is determined by using the corrected Plitt model equations (equations 25-30) to calculate R_f, d_{50c} and m. The partition curve is re-determined for various flowrate

values, at a constant density, and vice versa so that the full range of plant operation is covered. A constant hydrocyclone feed size distribution, determined from mass balance samples, is assumed throughout. Equation 42 is used to calculate the % -75μm particle size in the overflow corresponding to each value of flowrate and density. Linear regression is performed on the 2 sets of resultant particle size data, viz. constant density, changing flowrate; and constant flowrate, changing density. The slope of the respective regression curves shown in Figure 15 and 16 determines the value of constants *a* and *b*.

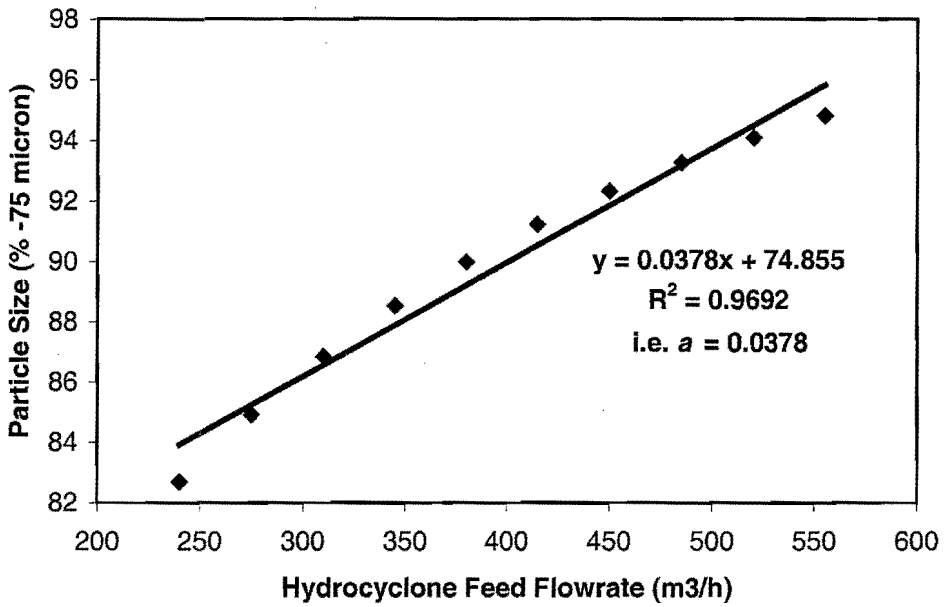


Figure 15. Plitt model flowrate fit

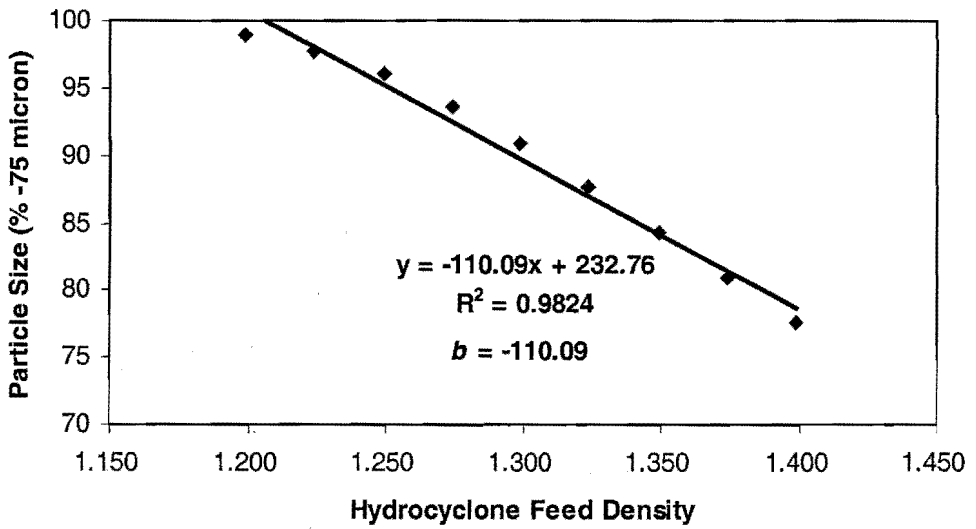


Figure 16. Plitt model density fit

5.1.2 Calculation of Constants d and e

Dynamic data, resulting from 10 second logged step test data, is used to identify the relationship between underflow angle and flowrate and density respectively. The underflow angle is measured on-line by the HUM. A zero order ARX model is used to fit the data in the MATLAB simulation environment. The dynamic data is first detrended (normalised in such a way that the mean for each variable is zero). After a graphical inspection of the data, a delay of 10 seconds for flowrate and 40 seconds for density with respect to angle is evident. These delays are included in the ARX model. The ARX model fitting procedure is described in Appendix A.5. Figure 17 shows a comparison between the calculated angle and measured angle by using the values of d and e shown below.

$$d = -0.0913$$

$$e = 18.3595$$

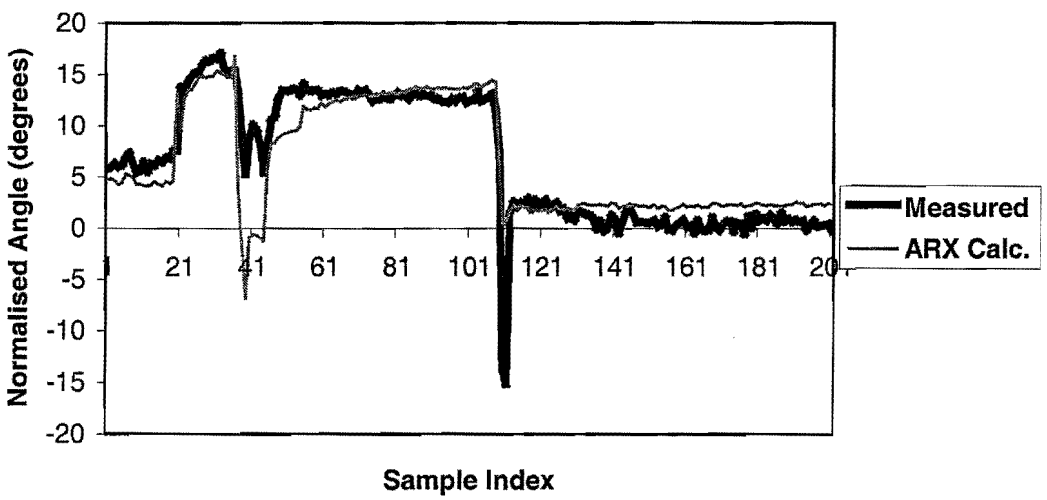


Figure 17. ARX model angle relationship using d and e above

5.1.3 Residual Analysis - Calculating K_{uf}

Now that a, b, d and e are known, equation 38 and 39 can be used to calculate the angle and particle size residuals. The particle size residual is plotted against the underflow angle residual. The slope of the best fit linear regression through the residual data determines the magnitude of the *PSE* model constant, K_{uf}. The angle residual is used to correct the Plitt estimate of particle size by taking account of particle size disturbances in the feed. The residual plot is shown in Figure 18. From the small R^2 it is apparent that the correlation

between residuals is poor. The lack of correlation has re-occurred on several occasions, and was one of the main reasons for undertaking this research. The lack of correlation indicates that angle appears to play a largely insignificant part in accurate particle size calculation, hence essentially leaving the Plitt model to provide a reliable estimate.

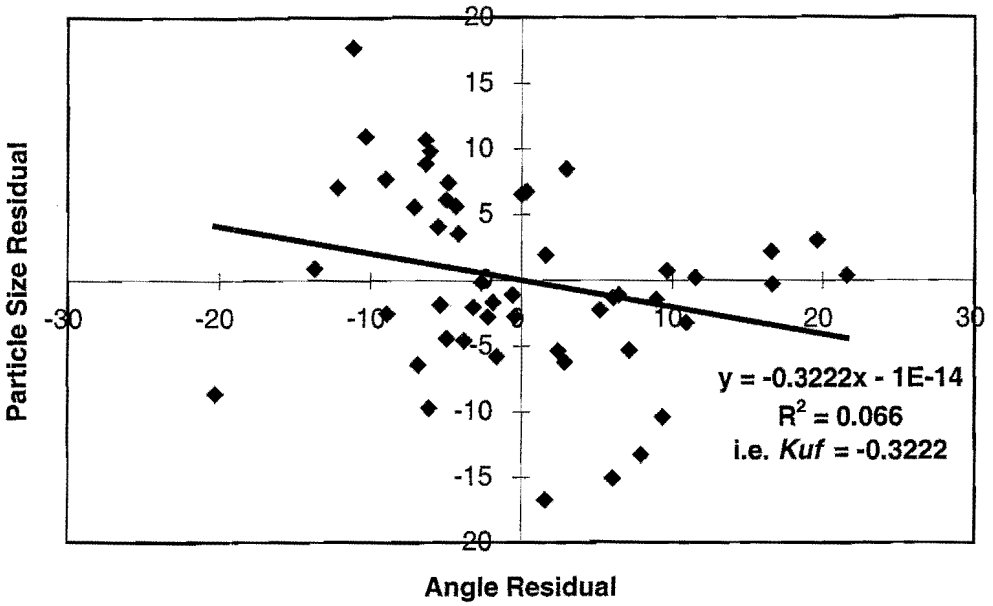


Figure 18. Residual plot for K_{uf} determination

5.1.4 Establishing a Base Case: PSE Error Analysis

Mass balance results were used to identify the initial 2-parameter particle size estimate. To calibrate the underflow angle correction to the model, only training data were used. This is decribed in the previous sections. Figure 19 shows the actual and calculated particle size for each sample. A *MSE* of 46.28 was achieved on the training data. The test data set was presented to the PSE model, and will be used in subsequent sections so that model performance can be fairly assessed. On the 21 sample test data, the *MSE* was calculated to be 37.54 as shown in Figure 19. The test *MSE* will be used as a basis of comparison for the remainder of this chapter.

Another useful technique for accessing the performance of a variety of models is to plot the calculated output against the desired response. For a perfect model, all sample points should fall on the $y = x$ line, i.e. the slope should be unity. An *XY* plot is shown in Figure 20. Also shown on Figure 20 is the R^2 value, another useful parameter for gauging model performance.

The correlation coefficient or R^2 will also be used extensively for determining goodness of model fit and is defined and further explained in Appendix A.6. As R^2 approaches unity, the model fit becomes better. The correlation coefficient quantifies the proportion of the variability in y explained by x assuming a linear relationship between x and y (where y is the actual output and x is either the predicted output or a model input).

Residual analysis is another useful technique for assessing model performance since all the information on lack of fit is contained in the residuals. As illustrated in Figure 21, a plot of model residual versus the predicted value, is a useful diagnostic tool. Residual analysis is best done by graphical inspection. Ideally, residuals should be randomly scattered as in Figure 27. If any trends are observed on the residual plot, it is likely that the model is not adequate, viz. higher order terms or more input variables are required.

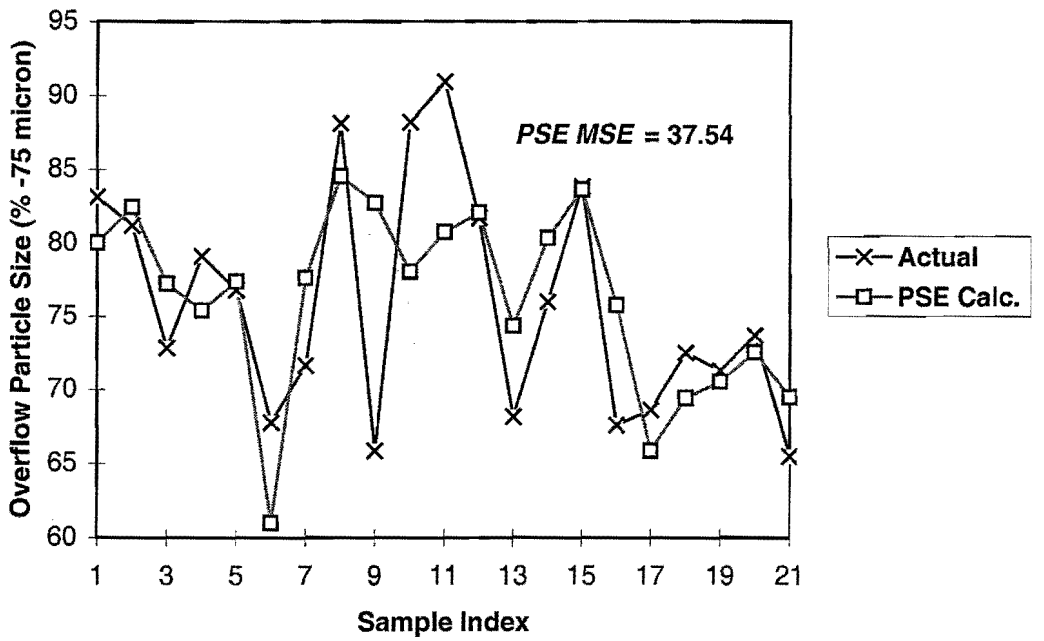


Figure 19. *PSE* model fit (test data)

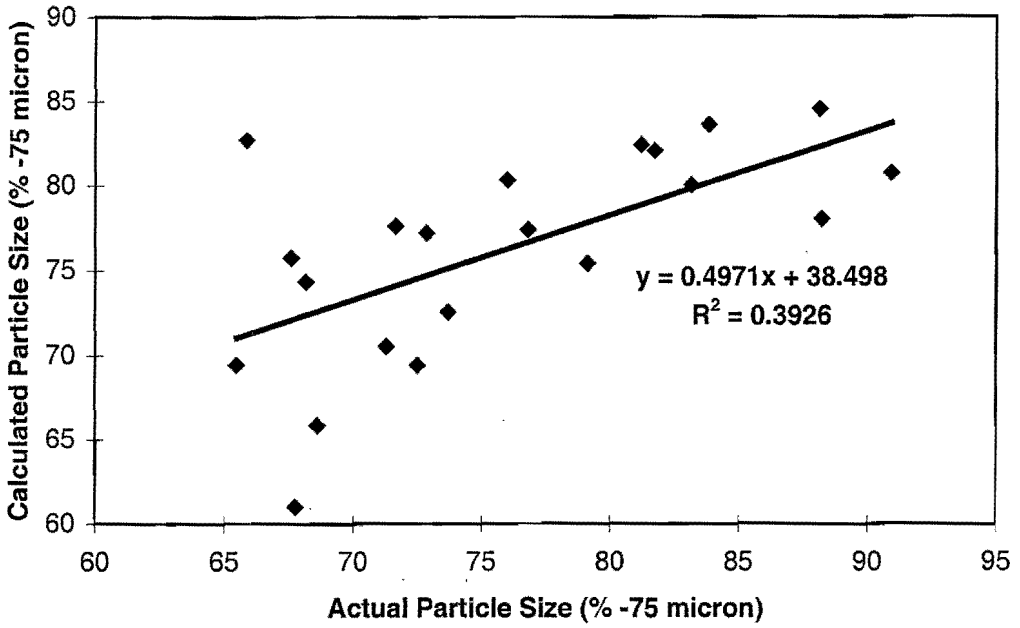


Figure 20. *XY* plot for *PSE* model (test data)

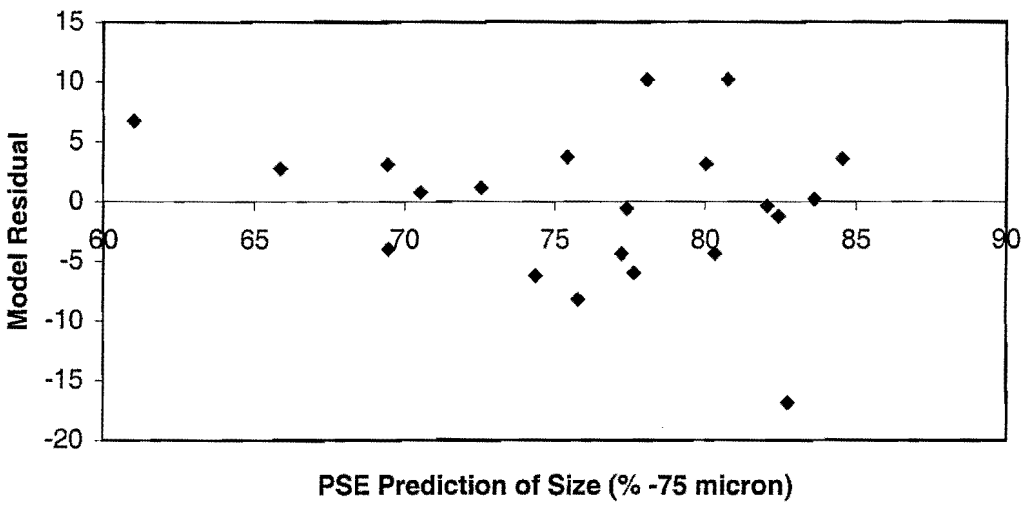


Figure 21. Residual plot for *PSE* model (test data)

A summary of some of the important "goodness of fit" criteria is shown in Table 7. These values will be used later for comparing subsequent models.

	Training Data	Test Data
MSE	46.28	37.54
R^2	0.401	0.393

Table 7. Summary statistics for *PSE* base case

5.2 MULTIPLE LINEAR REGRESSION MODELLING

Due to its simplicity, multiple linear regression models will be investigated as a first alternative to the *PSE* model described above. Equation 56 defines the general structure of multiple linear regression models.

$$\bar{y} = b_0 + b_1 x_1 + b_2 x_2 + ... + b_r x_r \qquad (56)$$

where $\bar{y}$ = predicted output

x_i = ith model input

b_i = coefficient corresponding to the ith input

The coefficients, b_0 to b_r are solved in a least squares sense by minimising the objective function defined by equation 57.

$$O_j = \sum_{i=1}^{n} (y_i - \bar{y}_i)^2 \qquad (57)$$

where n = number of data points available

y_i = actual value of the ith output

The *Statistica* software package was used to solve for MLR model constants. All models were identified by using the training data set, while testing data were reserved to allow assessment of model performance. The MLR solution procedure used for identification of all MLR model constants is outlined in Appendix A.7.

Two multiple linear regression (MLR) models were examined and are discussed in the ensuing sections.

5.2.1 3-Parameter MLR Model

To serve as a direct comparison with the *PSE* model, a MLR model of the order shown in equation 58 was examined.

$$\overline{PSE} = b_0 + b_1 CFF + b_2 CFD + b_3 CUA \qquad (58)$$

where $b_0 = 100.90$

$b_1 = 0.05098$

$b_2 = -43.041$

$b_3 = 0.18907$

Model fit, *XY* and residual plots are shown in Figures 22, 23 and 24 respectively.

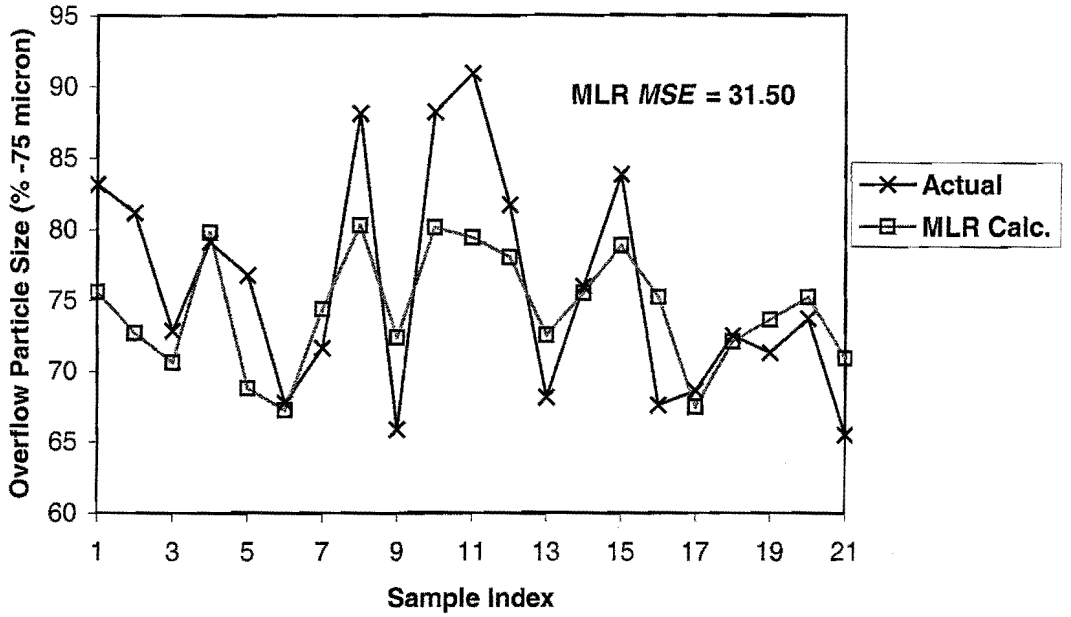


Figure 22. MLR model fit (test data)

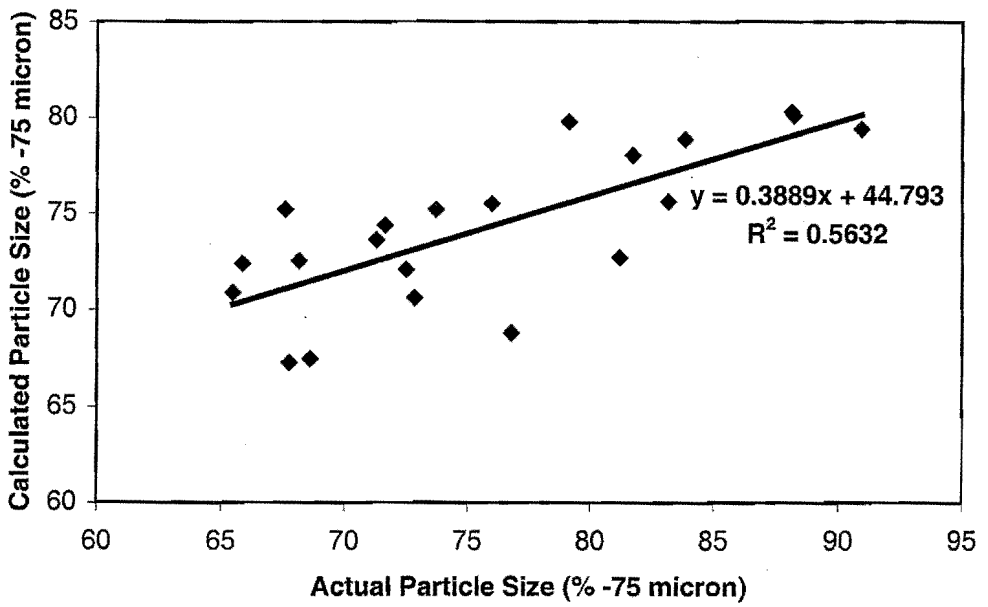


Figure 23. *XY* plot for MLR model (test data)

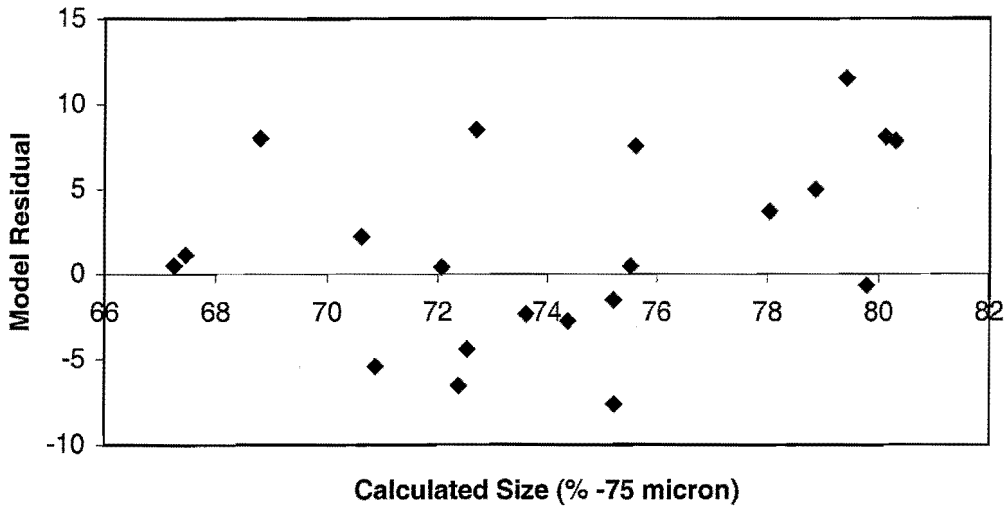


Figure 24. Residual plot for MLR model (test data)

From a precursory examination of the plots and model fit statistics, shown in Table 8, it can be seen that the MLR model gives a slight improvement over the *PSE* model.

	Training Data	Test Data
MSE	31.90	31.50
R^2	0.521	0.563

Table 8. Summary statistics for *MLR*

5.2.2 2-Parameter MLR Model

CM below is a correlation matrix showing the sample correlation between each input and output variable for the training data set. The sample correlation indicates how well input x is correlated to output y and is defined by the equation shown in Appendix A.6.

$$CM = \begin{bmatrix} 1.00_{CFF-CFF} & -0.44_{CFF-CFD} & -0.50_{CFF-CUA} & 0.58_{CFF-PSE} \\ & 1.00_{CFD-CFD} & -0.16_{CFD-CUA} & -0.62_{CFD-PSE} \\ & & 1.00_{CUA-CUA} & -0.01_{CUA-PSE} \\ & & & 1.00_{PSE-PSE} \end{bmatrix}$$

Subscript i-j indicates that the corresponding matrix entry quantifies the correlation between variable i and j. It can be seen that there exists a fairly strong linear relationship between flowrate and size (the nearer the sample correlation to 1 or –1, the stronger the linear relationship), and density and size respectively, while there is a very weak linear relationship

between angle and size (sample correlation is near 0). Furthermore, there exists a relatively strong relationship between angle and flowrate. This implies that most of the angle variation can be attributed to changes in flowrate. Moreover, the relative insensitivity of size to angle variation (when considered on a linear basis) suggests that a 2-parameter MLR model be identified by leaving out the angle term. The resultant model should account for a similar amount of the variance as the 3-parameter model and should thus provide comparable levels of accuracy. The proposed model is shown in equation 59.

$$\overline{PSE} = b_0 + b_1 CFF + b_2 CFD \qquad (59)$$

Where the coefficients were identified to be:

$b_0 = 128.43$

$b_1 = 0.03749$

$b_2 = -53.805$

Some important summary statistics are shown in Table 9 for later comparison.

	Training Data	Test Data
MSE	33.30	30.55
R^2	0.500	0.557

Table 9. Summary statistics for *MLR*

A graphical analysis of model performance is shown in Figure 25 to 27.

From an examination of the *MSE* and R^2 values, it can be seen that the exclusion of the underflow angle term has not sacrificed model accuracy. Hence it may be concluded that there is no statistical benefit to model performance by including the angle term in a linear form of the model. Later sections will focus on non-linear modelling techniques where the benefit of including the angle term will be re-assessed. A more rigorous statistical analysis of model performance will also be conducted.

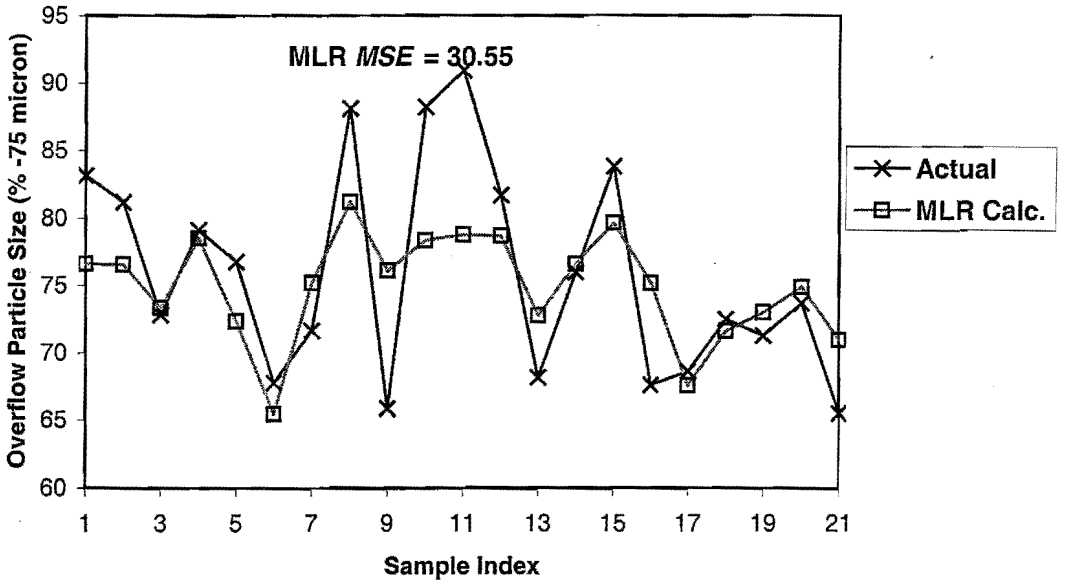


Figure 25. MLR model fit (test data)

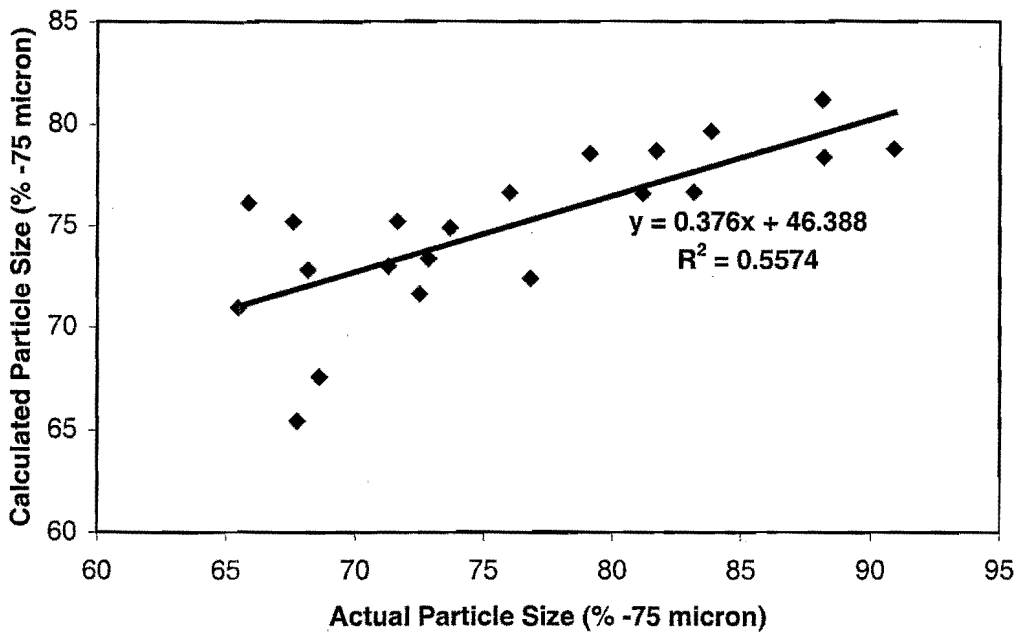


Figure 26. *XY* plot for MLR model (test data)

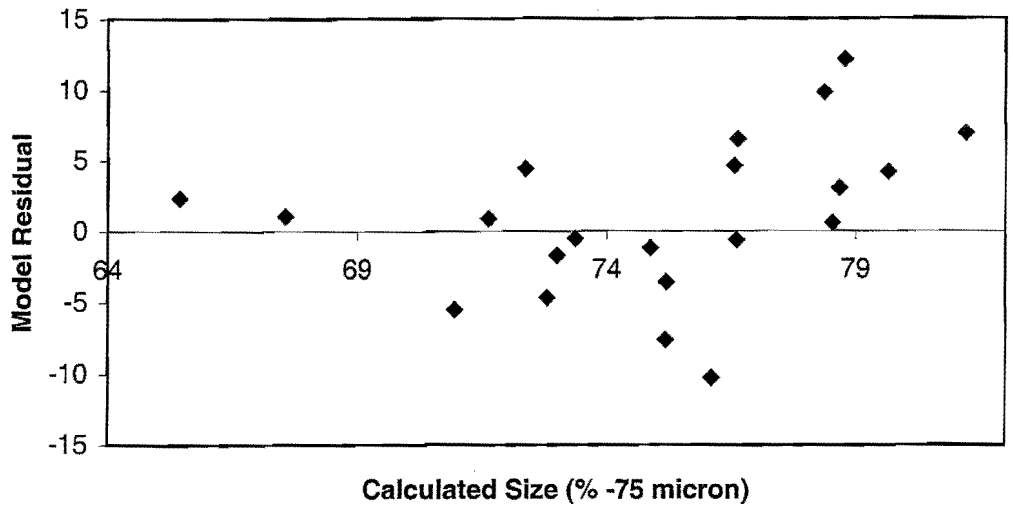


Figure 27. Residual plot for MLR model (test data)

5.3 NEURAL NETWORK MODELLING

Since a neural network model is a massively parallel connected model structure containing non-linear transfer functions, it is ideal for a first attempt at characterising some of the expected non-linearities in hydrocyclone classification. The advantage of using a neural network modelling approach is that the benefit of introducing non-linearity's can be assessed without spending much time identifying non-linear combinations of variables.

All neural network models investigated are of the multi-layer perceptron (MLP) type as described in Chapter 3. Backpropagation of the prediction error was used to adjust the weights by using the gradient descent optimisation procedure. In all cases, input and output variables were linearly scaled between –0.9 and 0.9. A single hidden layer was used, while all nodes in the hidden and output layers used the *tanh* transfer function. Different network topologies were investigated by altering the number of hidden nodes in a trial and error sense. To ensure good generalisation and hence preventing overtraining, training was terminated when the *MSE*, applied to the test data set, started to increase. Optimal weights were determined by using a commercially available neural network package (NeuroSolutions), as well as by implementing the entire network in a spreadsheet package (Excel), and solving for the optimum weights by using the built-in solver routine. The Excel spreadsheet solution procedure is outlined in Appendix A.8. Both training techniques gave similar results, however NeuroSolutions converged faster, while Excel enabled more interaction with the user. This was due to the degree of user interaction required for supervised training, i.e.

training was manually terminated when a consistent increase in the test data occurred. Approximately 10 training runs were performed for each topology. Analogous to the MLR models, 2 classes of neural network models will be discussed in the following sections.

5.3.1 3-Input Neural Network Model

As mentioned previously, a neural model relating particle size to the inputs feed flowrate, feed density and underflow angle, would be useful for assessing the benefit of including the angle term as a model input in an inherently non-linear sense. The performance of networks containing between 1 and 5 hidden nodes were compared. The results are summarised in Table 10 and Figure 28.

Network Structure	Training Data		Test Data	
	MSE	R^2	*MSE*	R^2
3:1:1	31.64	0.526	27.61	0.617
3:2:1	30.86	0.537	26.84	0.644
3:3:1	20.65	0.696	24.00	0.664
3:4:1	25.02	0.626	26.51	0.682
3:5:1	10.56	0.843	41.97	0.403

Table 10. Summary of various network performance indices

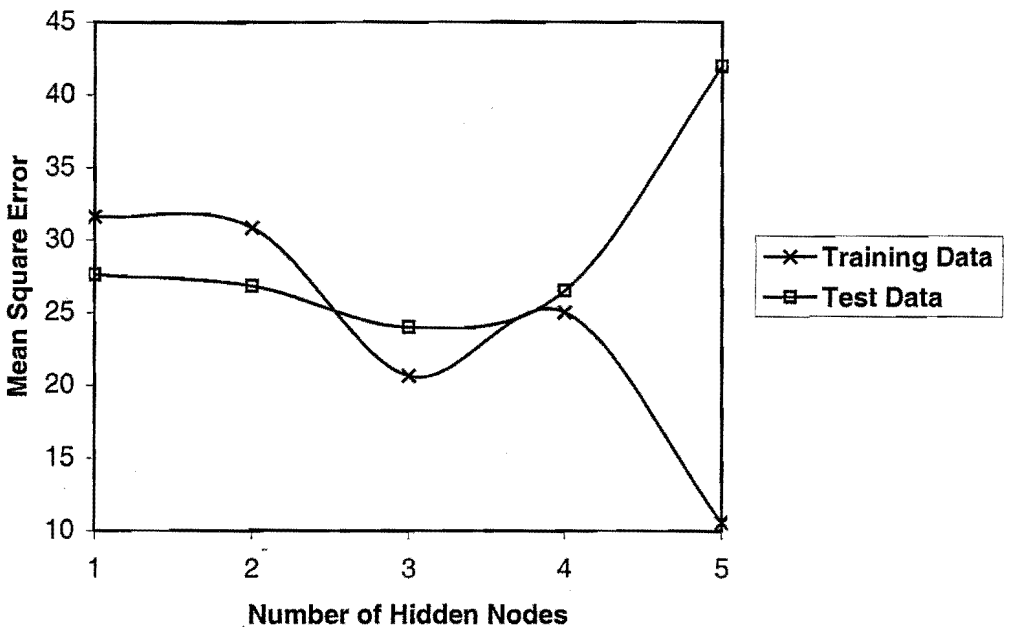


Figure 28. Performance of various network structures

From Figure 28 it is clear that a network containing 3 hidden nodes (and a bias node for both hidden and output layers) provides the minimum test *MSE* and hence the best interpolative capability. The weight values for the optimal *3:3:1* network are defined in the matrices below.

$$\begin{bmatrix} -0.950 & -0.164 & 1.764 & 2.010 \\ 1.081 & -0.311 & 0.775 & 0.858 \\ 0.223 & -2.118 & 0.787 & -2.232 \end{bmatrix}$$

where: the i^{th} row indicates the weight values for the i^{th} hidden node

the j^{th} column indicates the weight values for the j^{th} input where

$j = 1 \Leftrightarrow$ cyclone feed flowrate

$j = 2 \Leftrightarrow$ cyclone feed density

$j = 3 \Leftrightarrow$ cyclone underflow angle

$j = 4 \Leftrightarrow$ bias node input

$$\begin{bmatrix} -0.528 & 0.563 & 0.631 & 0.261 \end{bmatrix}$$

where: the j^{th} column indicates the weight values from the j^{th} hidden node

w_{14} represents the bias node connection to the output node

The performance of the neural network model is graphically illustrated in Figures 29 to 31.

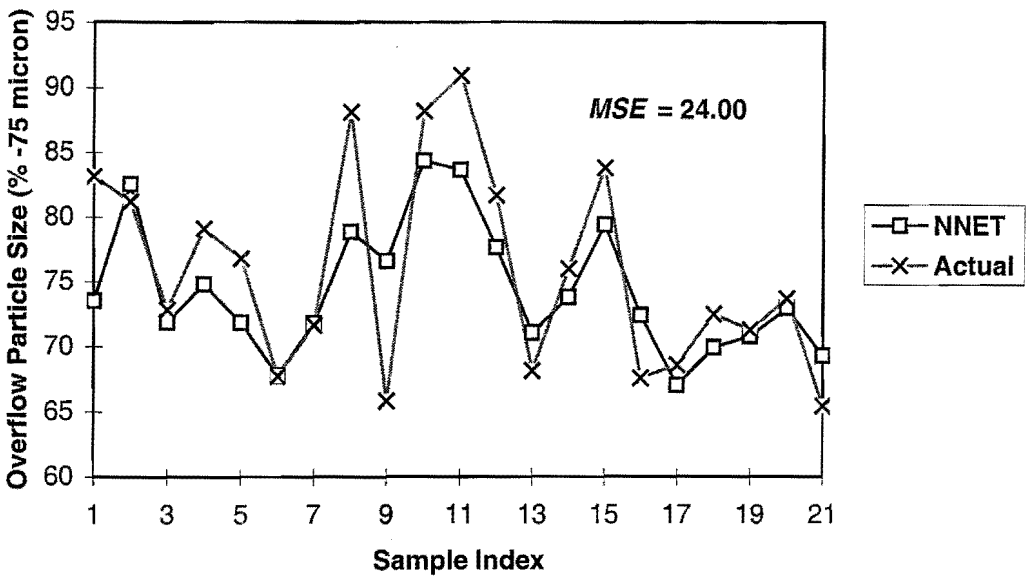


Figure 29. Neural network model fit (test data)

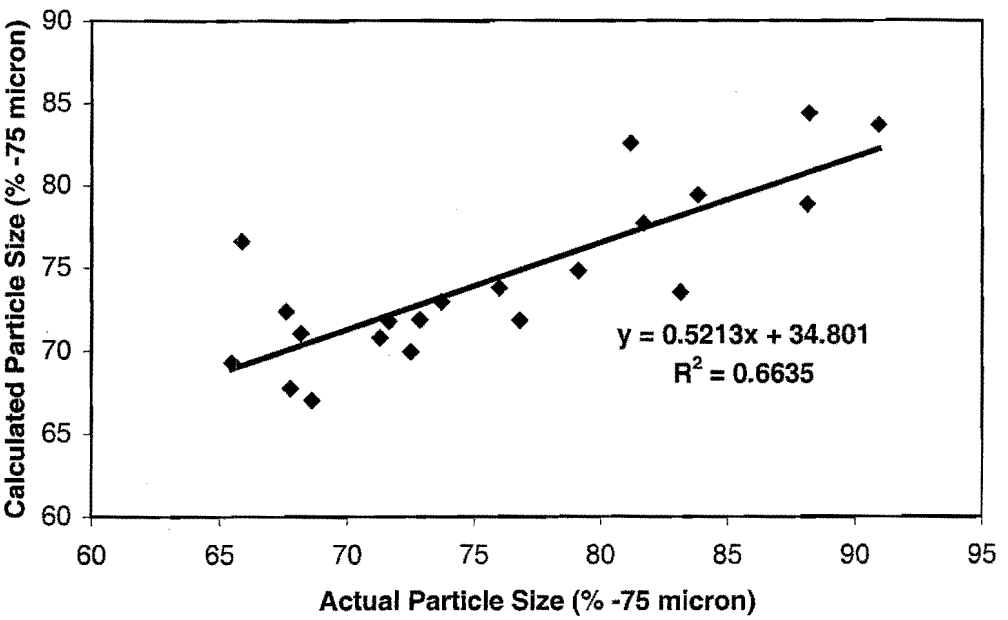


Figure 30. *XY* plot for neural network model (test data)

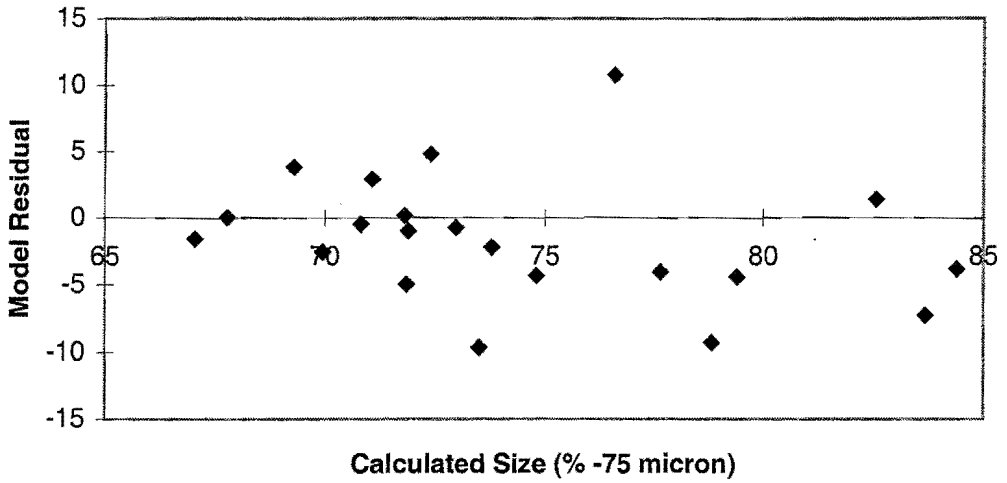


Figure 31. Residual plot for neural network model (test data)

5.3.2 2-Input Neural Network Model

The 2-input neural network model uses cyclone feed flowrate and density as inputs. A constant bias was also introduced to the hidden and output layers respectively.

The 2-input model is able to serve as a direct comparison with the 3-input neural model, enabling the benefit of the inclusion of the underflow angle term to be assessed by comparing the resultant non-linear model performance. Another advantageous attribute to identifying this model is that the merit of including a non-linear combination of feed flowrate and density can be assessed when the performance is compared to the 2-parameter MLR model.

The performance of networks containing between 1 and 4 hidden nodes was assessed and is summarised in Table 11 and Figure 32.

Network Structure	Training Data		Test Data	
	MSE	R^2	*MSE*	R^2
2:1:1	31.78	0.532	23.90	0.679
2:2:1	28.77	0.568	21.13	0.737
2:3:1	29.47	0.558	23.34	0.699
2:4:1	28.43	0.574	22.58	0.700

Table 11. Summary of various network performance indices

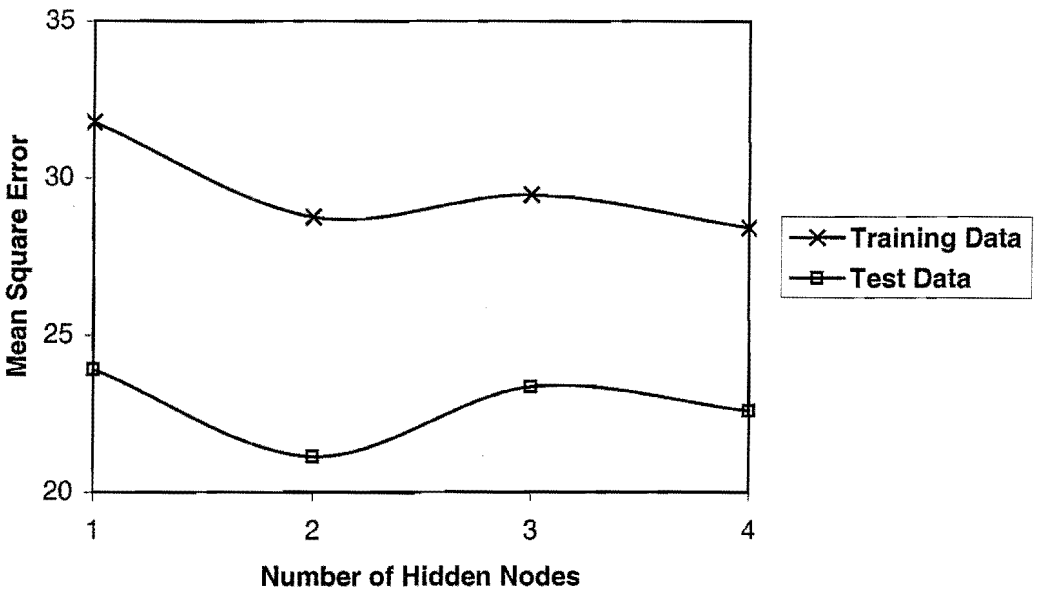


Figure 32. Performance of various network structures

Figure 32 clearly illustrates that a network containing 2 hidden nodes represents the optimum network structure in terms of obtaining the minimum test *MSE*. The weight matrices for the optimally trained *2:2:1* network are defined below.

$$\begin{bmatrix} \mathbf{0.511} & \mathbf{2.198} & \mathbf{2.137} \\ \mathbf{1.532} & \mathbf{-1.219} & \mathbf{-1.378} \end{bmatrix}$$

where: the i^{th} row indicates the weight values for the i^{th} hidden node

the j^{th} column indicates the weight values for the j^{th} input where

$j = 1 \Leftrightarrow$ cyclone feed flowrate

$j = 2 \Leftrightarrow$ cyclone feed density

$j = 3 \Leftrightarrow$ bias node input

$$\begin{bmatrix} \mathbf{-0.564} & \mathbf{0.424} & \mathbf{0.358} \end{bmatrix}$$

where: the j^{th} column indicates the weight values from the j^{th} hidden node

w_{13} represents the bias node connection to the output node

Figures 33 to 35 illustrate the performance of the optimally trained *2:2:1* neural network model.

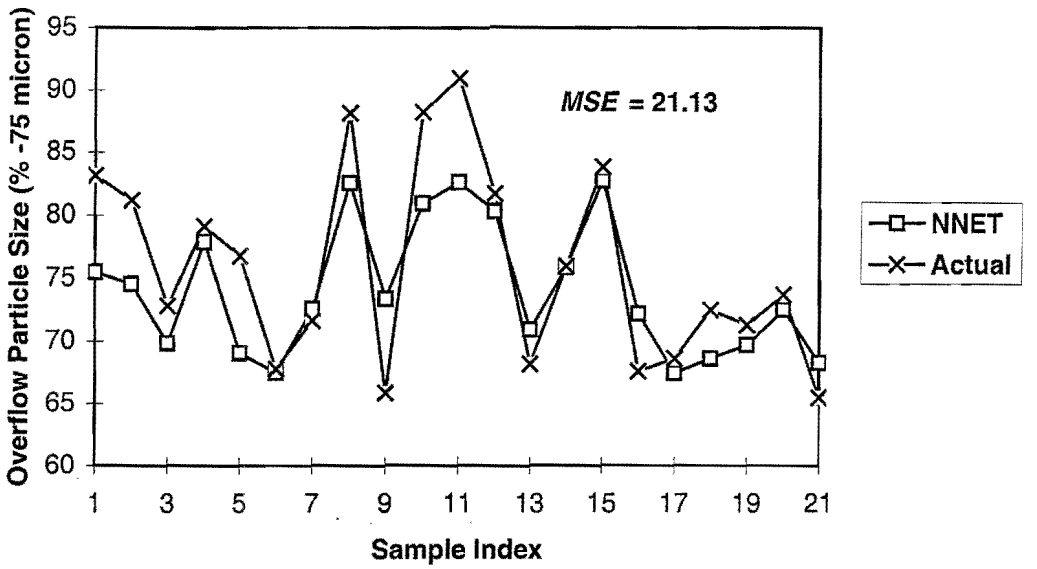


Figure 33. Neural network model fit (test data)

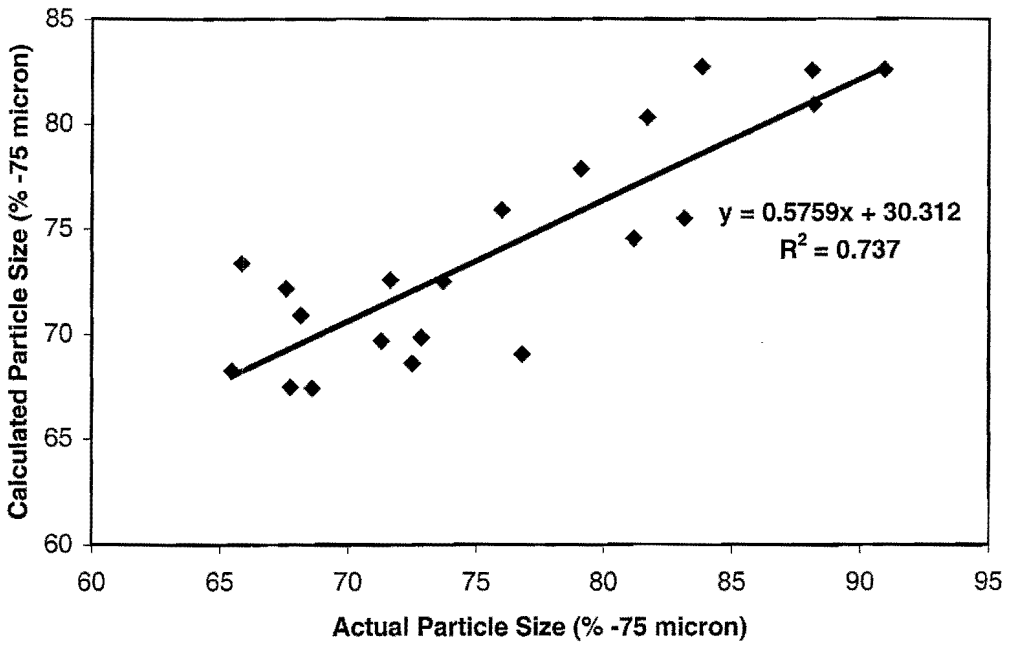


Figure 34. *XY* plot for neural network model (test data)

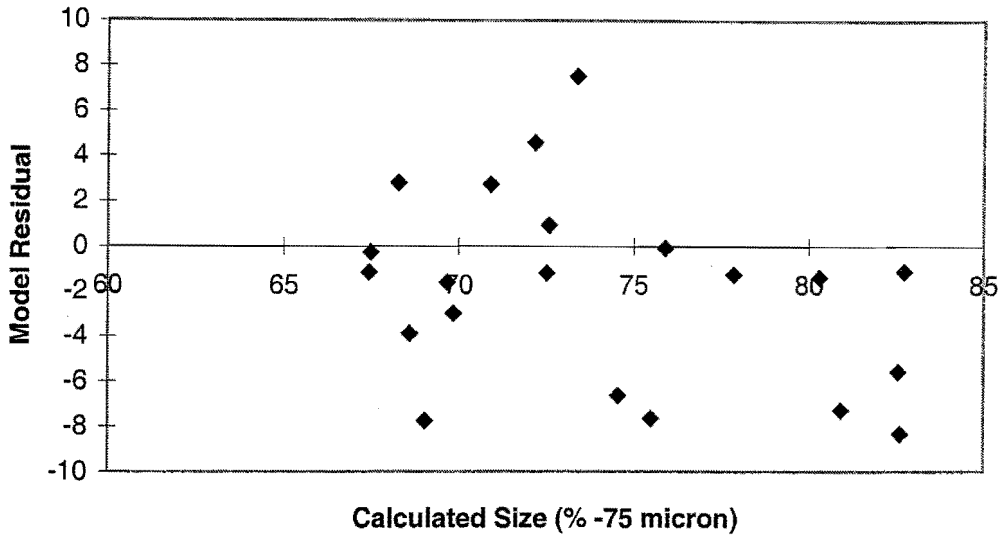


Figure 35. Residual plot for neural network model (test data)

When the 2 input neural network is compared to the 3 input neural network it can be seen that, analogous to what was observed in regression modelling, there is no significant contribution due to the underflow angle term. The statistical significance of neglecting the underflow angle term will be discussed later after performing rigorous statistically sound tests to evaluate model performance.

5.4 INCLUDING OVERFLOW DENSITY AS A MODEL INPUT

As mentioned previously, overflow samples were analysed for both particle size (% -75 μm) and density (mass % solids). After a precursory graphical inspection of the experimental data, it is observed that there appears to be a significant linear relationship between particle size and hydrocyclone overflow density. Figure 36 illustrates the relationship for the complete data set.

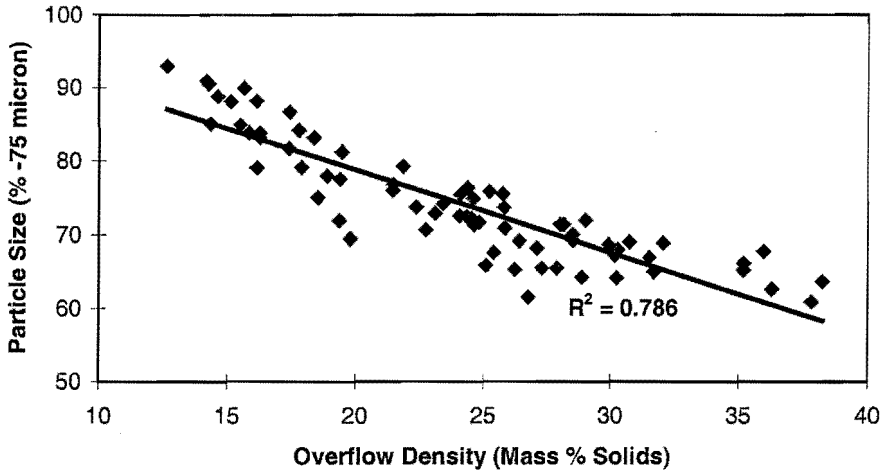


Figure 36. Correlation between particle size and overflow density (entire 71 sample data set used)

The significance of the linear relationship is further enforced by a correlation coefficient, R^2, of 0.786, which is of considerable consequence bearing in mind experimental errors associated with sampling and sample analysis.

Due to the strong influence of overflow density on particle size, it was decided to include it as an input variable in future modelling approaches.

5.4.1 Measurement of Overflow Density

Before overflow density can be used as a model input for a size estimation soft sensor, it must be measured so that it is available as an input signal.

Generally, overflow density measurement is available for most plants in the South African platinum industry, since it is important in characterising the feed to the subsequent flotation circuit. Typically, the standard nucleonic type densitometer is used for density measurement. It is necessary to install the nucleonic densitometer on a pipeline that is free of air. This is typically accomplished by installing it on a vertical pipe section after a sump/pump combination.

South African gold plants, however, do not normally measure overflow density since the cyclone overflow product is normally gravity fed, via an open channel, to a thickener. Hence, to install a densitometer would require large infrastructural changes, viz. the addition of an extra sump and pump as well as closed conduit piping.

To enable the incorporation of the overflow density signal into a soft sensor model, it is important to find a reliable method of measuring density, not requiring large infrastructural changes, for use on plants where this measurement is lacking. For this reason, a market survey of available density measurement devices was conducted. The coriolis massflow meter was obtained on a trial basis from a local supplier to assess its ability to measure slurry density. The following short section describes some of the findings and explains the operating principal of the coriolis massflow meter.

The coriolis massflow meter

The measuring tube within the coriolis meter is vibrated at its natural frequency by supplying an external voltage. As slurry flows through the measurement tube, a force is imparted which disrupts the natural vibrational frequency. Inductive sensors pick up the change in motion of the tube, which is converted, to a mass flowrate by appropriate internal signal processing.

What makes the device so attractive for use in abrasive slurry streams is that there are no moving parts exposed to the slurry that may clog or abrade. Also, the pressure drop across the device is minimal, making the system energy efficient. For a typical plant installation, it is suggested that a representative bleed stream of approximately 150 l/min be taken from the full overflow stream and passed through the coriolis meter. The bleed stream can flow through the meter by gravity and be returned to a lower point in the open channel. A 37.6 mm inner diameter flowmeter was tested, although various sizes are available to allow measurement for a large range of flowrates. At design flowrates the instrument proved to be accurate (average absolute error of <1%) and sufficiently robust for industrial use.

5.4.2 MLR Modelling

To assess the benefit of including overflow density as an input variable, a MLR model was considered as a first approach. A 3-parameter model was developed using *CFF*, *CFD* and *COD* (cyclone overflow density). *CUA* was not considered to be a significant model input

(see section 5.2.2) and was therefore not included. The model structure and constants are shown in equation 60 below.

$$\overline{PSE} = b_0 + b_1 CFF + b_2 CFD + b_3 COD \qquad (60)$$

where $b_0 = 114.06$

$b_1 = -0.04368$

$b_2 = 12.977$

$b_3 = -1.6177$

Some important summary statistics are shown in Table 12 for later comparison.

	Training Data	Test Data
MSE	10.46	7.71
R^2	0.843	0.877

Table 12. Summary statistics for *MLR* with COD

The Figures below illustrate the benefit of including the overflow density term.

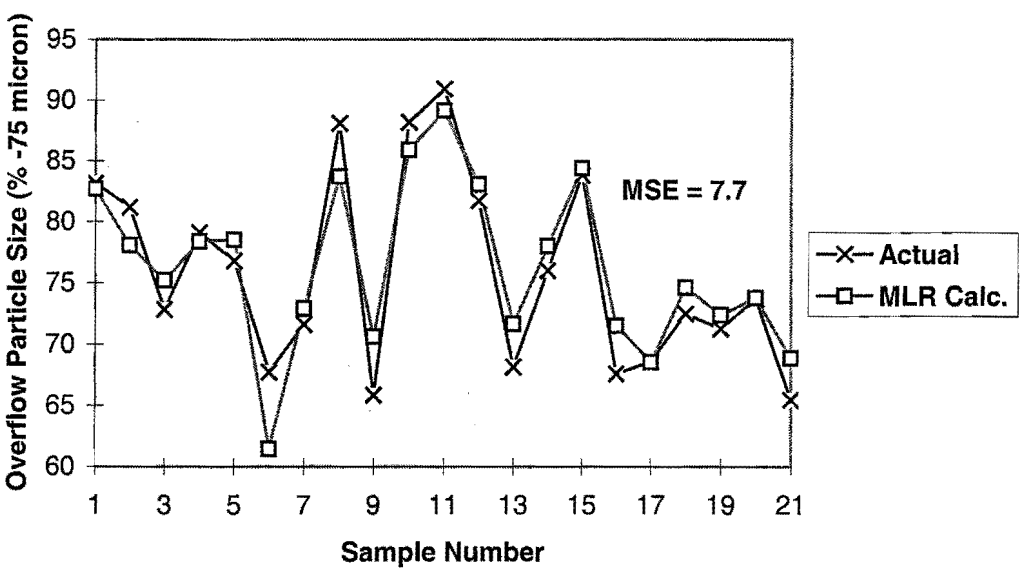


Figure 37. MLR with COD model fit plot (test data)

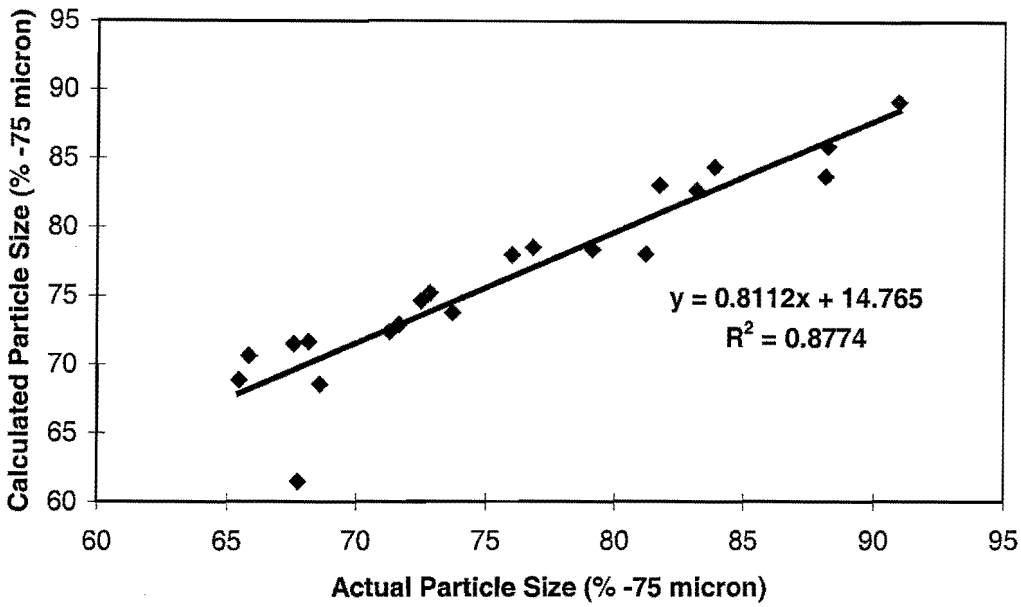


Figure 38. *XY* plot for MLR model with COD (test data)

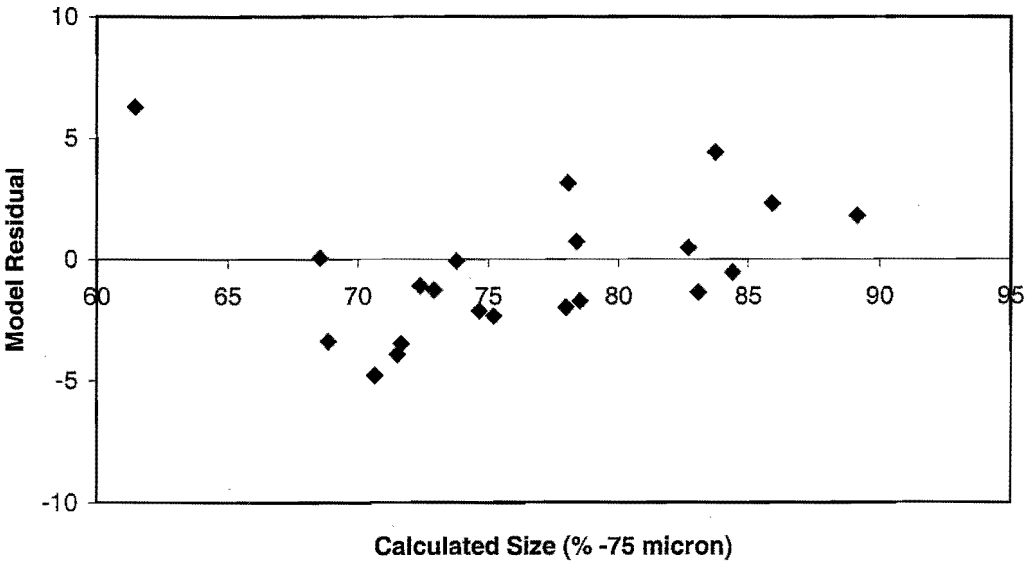


Figure 39. Residual plot for MLR with COD model (test data)

As expected, the inclusion of the overflow density in a simple regression model significantly improves the accuracy of the size estimate.

5.4.3 Neural Network Modelling

The neural network model considered here consists of 3 inputs, viz. cyclone feed flowrate, feed density and overflow density. Once again a constant bias was introduced to the hidden and output layers respectively.

A variety of network structures were trained until completion – minimum test *MSE*. The results are shown in tabular and graphic form below.

Network Structure	Training Data		Test Data	
	MSE	R^2	*MSE*	R^2
3:1:1	6.85	0.899	3.91	0.934
3:2:1	4.99	0.925	3.42	0.946
3:3:1	4.92	0.926	2.74	0.955
3:4:1	5.43	0.919	3.87	0.939
3:5:1	4.87	0.927	3.71	0.941

Table 13. Summary of various network performance indices

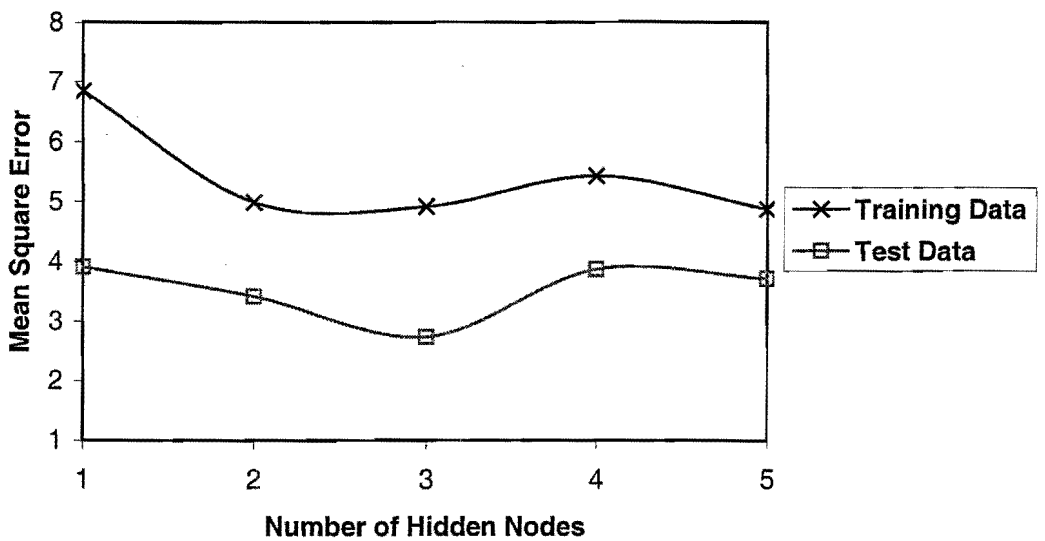


Figure 40. Performance of various network structures

From the above analysis, it is clear that a *3:3:1* network structure represents the best compromise between accuracy and generalisation. The weight matrices for the optimal network structure are described by:

$$\begin{bmatrix} -0.070 & -0.031 & 0.147 & -0.115 \\ -0.022 & 0.010 & -0.899 & -1.599 \\ 1.070 & -0.445 & 1.584 & -1.504 \end{bmatrix}$$

where: the i^{th} row indicates the weight values for the i^{th} hidden node

the j^{th} column indicates the weight values for the j^{th} input where

$j = 1 \Leftrightarrow$ cyclone feed flowrate

$j = 2 \Leftrightarrow$ cyclone feed density

$j = 3 \Leftrightarrow$ cyclone overflow density

$j = 4 \Leftrightarrow$ bias node input

$$\begin{bmatrix} 0.052 & 4.110 & -1.964 & 1.749 \end{bmatrix}$$

where: the j^{th} column indicates the weight values from the j^{th} hidden node

w_{14} represents the bias node connection to the output node

The performance of the *3:3:1* network is analysed and illustrated by the Figures that follow.

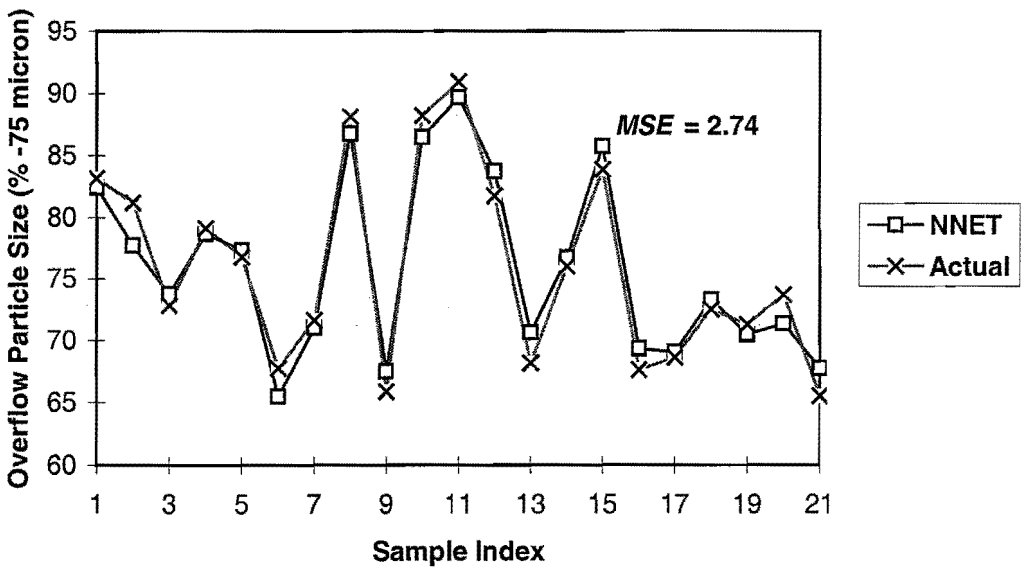


Figure 41. COD Neural network model fit (test data)

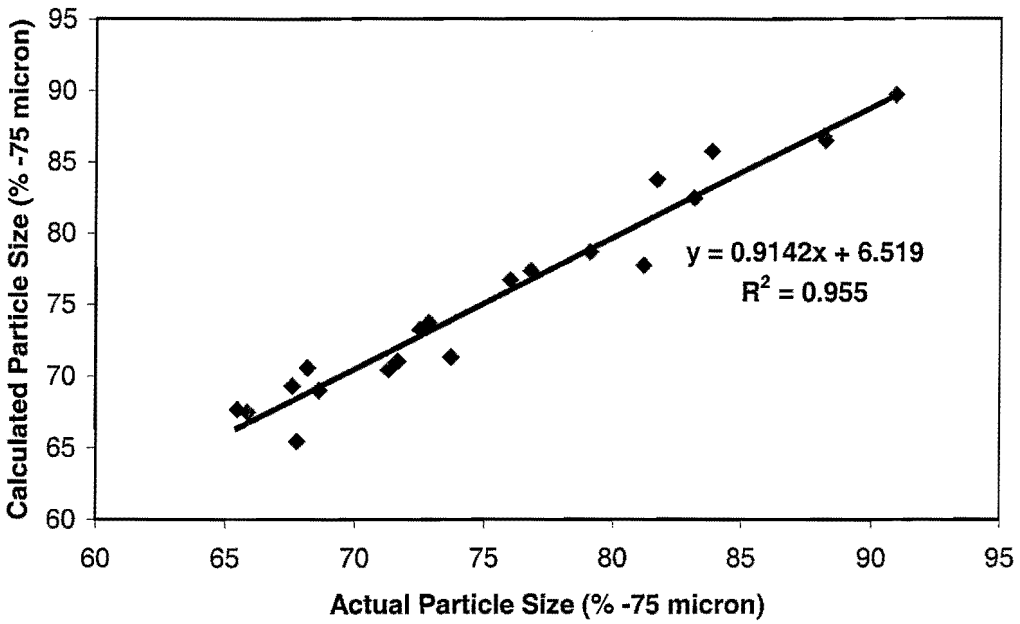


Figure 42. *XY* plot for COD neural network model (test data)

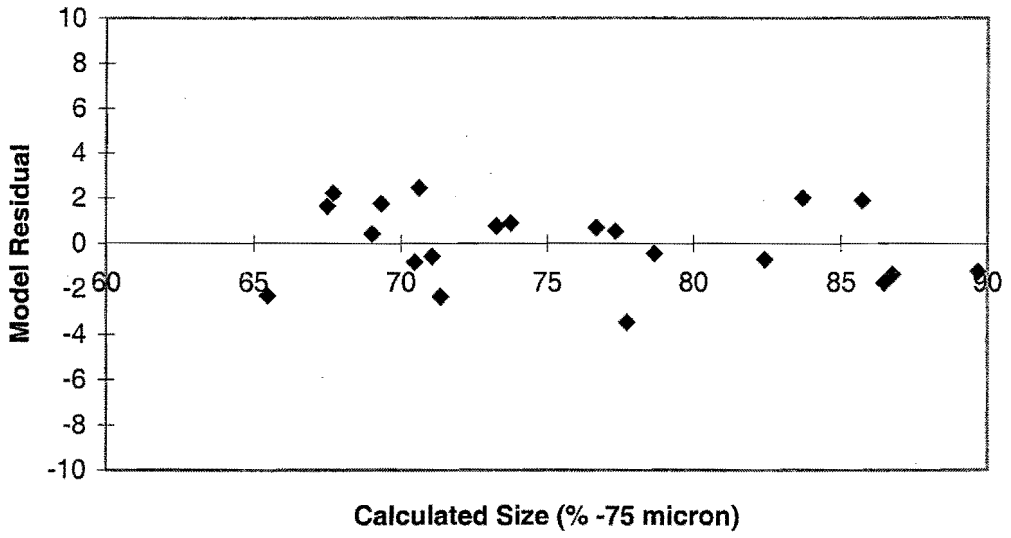


Figure 43. Residual plot for COD neural network model (test data)

5.5 ANALYSIS OF MODEL PERFORMANCE AND FINAL MODEL STRUCTURE

From the preceding sections it is clear, after residual and correlation analysis, that the inclusion of overflow density dramatically improves model performance. Furthermore, benefit is realised from the inclusion of non-linear modelling techniques. The purpose of this

section is to more rigorously assess model performance from a statistical point of view. This, together with practical issues will be used to choose the final model.

5.5.1 Statistical Analysis of Model Residuals

With the inclusion of the existing PSE model, 7 models have been identified and calibrated in previous sections. A statistical analysis of model residuals on the test data will be used to identify the most accurate model. Figure 44 below summarises the results of the aforementioned models in terms of their *MSE*.

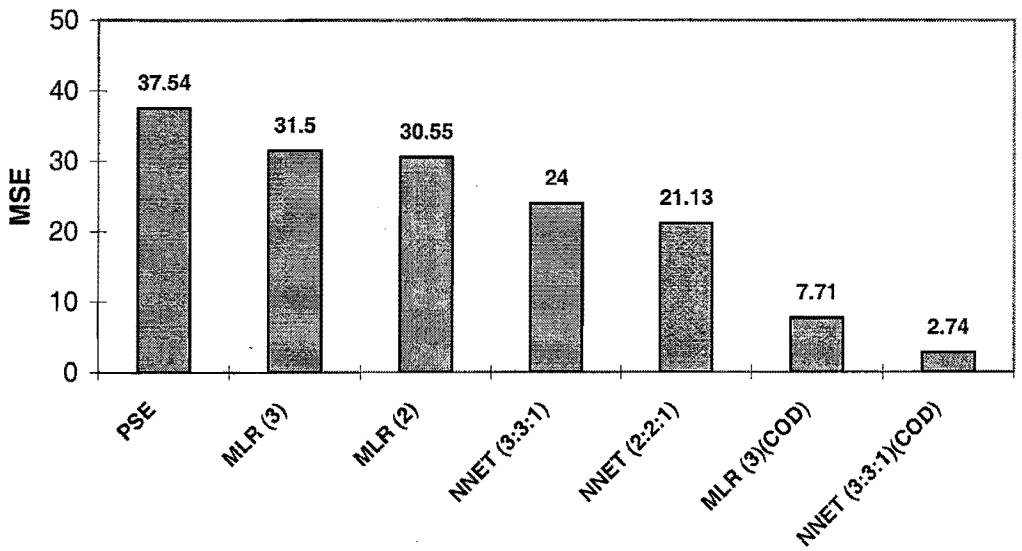


Figure 44. Summary of model *MSEs*

Since we are dealing with a distribution of errors, quoting a single average value may be misleading. A more correct technique would be to establish the range over which each model's residuals are likely to be distributed. To add further significance to the range of errors obtained, confidence intervals may be used. The limits of the distribution for each model's residual mean is obtained for a 98% confidence interval. The following residuals are considered:

- Model error based on the test data set.
- The absolute value of the model error on the test data set.

It is useful to look at the absolute value of the model error to provide an estimate of the average error obtained for each modelling technique. When actual errors are considered, the

effect of positive and negative deviation cancel and the true average model error is not obtained. Table 14 and 15 below summarise important statistical values for each case above.

	PSE	*MLR3*	*MLR2*	*NNET3*	*NNET2*	*MLR3 COD*	*NNET3 COD*
Mean	-0.316	1.594	1.195	1.540	1.884	-0.429	-0.008
Median	0.208	0.496	1.067	1.551	1.251	-1.097	-0.419
Standard Deviation	6.270	5.511	5.574	4.766	4.296	2.810	1.696
Confidence (98.0%)	3.459	3.040	3.075	2.629	2.370	1.550	0.936
Max Limit of Mean	3.143	4.635	4.270	4.169	4.254	1.121	0.927
Min Limit of Mean	-3.775	-1.446	-1.880	-1.089	-0.486	-1.979	-0.944

Table 14. Analysis of model residuals (test data)

	PSE	*MLR3*	*MLR2*	*NNET3*	*NNET2*	*MLR3 COD*	*NNET3 COD*
Mean	4.642	4.570	4.412	3.810	3.653	2.252	1.442
Median	3.702	4.380	4.385	3.822	2.797	1.987	1.371
Standard Deviation	4.098	3.334	3.482	3.156	2.859	1.663	0.832
Confidence (98.0%)	2.261	1.839	1.921	1.741	1.577	0.917	0.459
Max Limit of Mean	6.903	6.409	6.333	5.551	5.230	3.170	1.901
Min Limit of Mean	2.382	2.730	2.492	2.069	2.076	1.335	0.983

Table 15. Analysis of absolute model residuals (test data)

The median is included in Tables 14 and 15 to serve as a comparison to the mean. The advantage of using the median is that, since it is defined as the centre of a data set arranged according to size, it is not biased by extreme values as the mean may be.

The standard deviation is by far the most general measure of variation. Its benefit comes from it providing the average deviation from the mean in the same units as the observations.

The confidence interval for the mean, which is a data-based interval of numbers thought likely to contain the mean with a stated probability-based confidence of 98%, was calculated by using the *Statistica* software package. The confidence interval for the mean (given by the maximum and minimum limits of the mean in Tables 14 and 15) was calculated by adding and subtracting the result from the mean, respectively. This technique assumes that the population is normally distributed.

The mean, median, standard deviation and confidence interval are defined in Appendix A.6.

From the above analysis it is possible to make the following comments:

- There is a large difference between the mean and the median for the PSE and 2-input neural network model. Since this is not the case for other models, it indicates that these models do not fit all points. This can be attributed to the model not accounting for enough of the input variance.
- A clear benefit is realised when the overflow density term is included.
- Inclusion of cyclone underflow angle does not benefit model performance in any way. This is true for both linear (MLR) and non-linear (NNET) model structures.
- Some accuracy benefit is obtained when highly non-linear models are used.

Figure 45 illustrates the mean (centre square) and 98% confidence limits (vertical line) of the absolute value of the residuals.

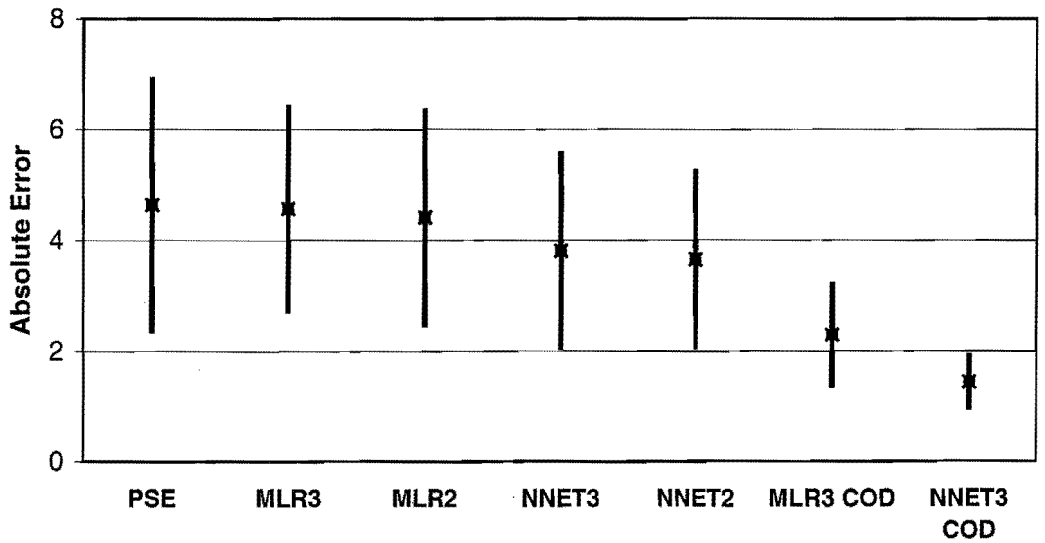


Figure 45. Absolute errors and 98% confidence limits

More specifically the neural network model with overflow density term yielded the most accurate estimate of particle size with an absolute error of *0.983<error<1.901.* The multi-linear regression model with overflow density provided an error of *1.335<error<3.170.* These models are significantly better than the all other models considered, clearly demonstrating the benefit of including overflow density. No significant improvement over the PSE base case is observable when overflow density is not included, even when non-linear neural network models are employed.

From Figure 45 and Tables 14 and 15, it can be concluded that the neural network and MLR models containing overflow density as an input, provide a statistically significant improvement over the other models. For the neural network model with overflow density, it is observed that the maximum error of the mean (for absolute residuals) is smaller than the minimum value of the mean for all other models not containing overflow density at 98% confidence. The overflow density MLR model provides a maximum absolute error (at 98% confidence) smaller that the mean error for all models not containing overflow density, indicating that, at worst, it is more accurate than the average error obtained by the other models. No other statistically significant improvement is observed when the models not containing overflow density are compared.

5.5.2 Choosing the Optimal Model for Soft-Sensor Implementation

From the previous section it can be seen that overflow density should be included as an input, implying that the final choice of model is between the regression and neural network models. On performance alone, the neural network model is a clear winner. However, the final model will form the basis of a soft-sensor to be implemented in an industrial environment. This implies that the practical issues shown in Table 16 need to be carefully evaluated for each model.

Criteria	MLR	NNET
Data requirements	Only 4 model constants required, implying an over-determined system (n>4). Requires very few data points to calibrate accurately.	Requires 16 network weights, implying that at least 16 data points are required for calibration. In addition, a test data set is also required. Normally 2-5 data points are required per weight, implying high data requirements.
Maintenance	Easy to maintain and recalibrate due to simple model structure.	Requires specialist knowledge to maintain model accuracy. On-site maintenance is therefore difficult.
Robustness	Regression models may be extrapolated with moderate confidence. Also due to their simplicity, robust performance is expected.	Neural network models are notoriously bad at predicting beyond the boundaries of the calibration data. Over training may also result in poor performance due to poor generalisation. These factors affect the robustness of the model.

Table 16. Practical considerations impacting on model choice

To produce a commercially viable soft-sensor for particle size estimation it is important to consider commissioning time. A model providing an adequately accurate estimate, but requiring modest quantities of data, and hence less sampling and analysis, requires shortened commissioning time. This, together with the factors listed in Table 15, result in the multi-linear regression model, with *CFF*, *CFD* and *COD* as inputs, being the final choice of model.

5.5.3. Incorporation of the Chosen Model into an Industrial Soft-sensor

Since it is envisaged that the final model will later be incorporated into a soft-sensor for use in the production environment, it is considered essential, in the interest of robustness, to handle input errors correctly. Standard procedures have been developed for the detection of basic errors in input instruments. The following error checks were identified:

1. Instrument OFF or power failure (0 mA)
2. Instrument output stuck at minimum (<4.2 mA)
3. Instrument output stuck at maximum (>19.8 mA)

4. Instrument output stuck within range (standard deviation below pre-defined threshold)

In the event of an input instrument failure, a secondary, less accurate model is used to ensure availability. In all cases, only 1 input failure is allowed, resulting in a family of 3 secondary models, each composed by neglecting alternate input variables. All secondary models are of the multi-linear regression type and are shown below with *MSE* and R^2 values for the test data.

1. **If overflow density measurement fails:**

$$\overline{PSE} = 128.43 + 0.03749CFF - 53.805CFD \qquad (61)$$

$MSE = 30.55$

$R^2 = 0.557$

2. **If feed density measurement fails:**

$$\overline{PSE} = 125.93 - 0.03960CFF - 1.4829COD \qquad (62)$$

$MSE = 8.21$

$R^2 = 0.869$

3. **If feed flowrate measurement fails:**

$$\overline{PSE} = 100.76 - 0.34165CFD - 1.0823COD \qquad (63)$$

$MSE = 13.05$

$R^2 = 0.801$

The previous findings regarding the benefit of overflow density inclusion are confirmed by the low error values on secondary models containing *COD*.

For industrial implementations, the following calibration procedure will be used:

1. A series of step changes in model inputs should be made. The entire operating range should be covered. Once steady state operation has been achieved after each step, a physical overflow sample should be taken. 20-30 different operating points should be identified.
2. All data should be logged at 10 second intervals. Overflow samples should be analysed for % -75 μm and associated with steady state averages of input variables.

3. Multi-linear regression model constants should be calculated and implemented in the software module or hardware unit provided.

Furthermore, it is important to note that the soft-sensor model proposed requires routine re-calibration. To allow easy re-calibration it is proposed that the model takes the form of equation 64:

$$\overline{PSE} = b_0 + b_1 CFF + b_2 CFD + b_3 COD + offset \qquad (64)$$

where $offset$ = an adjustable parameter that is calculated during re-calibration

The re-calibration procedure to be followed is outlined below:

1. Ensure that milling circuit and cyclone in particular is running at steady state.
2. Take 5 overflow samples at 3 minute intervals, i.e. samples correspond to a 15 minute time period.
3. The soft-sensor calibration mode should be activated. This stores the latest 15 minute average of the particle size as calculated by the soft-sensor (PSE_{recal}) and corresponds to the physical samples taken.
4. The 5 overflow samples should be analysed for % -75μm and the results averaged (PSE_{samp}).
5. The PSE_{samp} value should be entered into the soft-sensor. The offset is calculated as shown in equation 65.

$$offset = PSE_{samp} - PSE_{recal} \qquad (65)$$

6. Equation 64 is now corrected. This procedure should be repeated every 2 weeks.

CHAPTER 6. CONCLUSIONS

The aim of this project was to develop a reliable hydrocyclone particle size estimation model. This model is to form the basis of a soft-sensor for particle size estimation on full-scale milling circuits.

It is important to note that a database of overflow particle size responses to step changes in sump water addition (cyclone feed density), pump speed (cyclone feed flowrate) and mill feed rate has been collected after extensive test work on an industrial gold plant. Furthermore, due to the duration of the tests (3 months) gradual ore changes were also accounted for.

In all cases, real plant data has been used to identify suitable model inputs and model types. This enables the models to easily be extended to other plants, since no, often problematic, scale-up from pilot to industrial scale is required.

The complete data set is divided into training and testing sets. In all cases models are identified for the training data, while the test data is used to compare resultant model performance and assess the benefit of including or removing model inputs.

It is clearly shown that the inclusion of the underflow angle term does not benefit the accuracy of size prediction. This implies that the hypothesis on which the previous PSE model is based is incorrect. The inclusion of both linear and non-linear combinations of the underflow angle term were analysed and confirmed the above.

Cyclone feed density and flowrate capture most of the normal variation likely to affect cyclone classification and are useful in providing a basis from which to derive a more complete model. Cyclone classification can be affected by changes in the feed flowrate and density, as well as changes in the cyclone feed size distribution. It is therefore necessary to identify and make use of a third input variable to accurately model the cyclone's classification behaviour. As previously discussed, the underflow angle does not provide information that is useful for characterising the third degree of freedom, viz. changes in the feed size distribution, affecting cyclone classification. For this reason, another model input variable, cyclone overflow density, was assessed to determine the benefit of its inclusion.

Overflow density is measured on most platinum plants by using standard nucleonic densitometers. On most gold plants, however, it is not measured. A simple, inexpensive density measurement device, the Coriolis Mass Flow Meter, was extensively tested on a pilot scale. This instrument is suitable and robust enough to measure overflow pulp density of a suitably sized bleed stream on plants where this measurement is not available.

An extremely strong relationship is observed between particle size and overflow density. This observation is given even more impetus when the benefit of its inclusion in linear and non-linear models is assessed. Typically, the inclusion of the overflow density term reduces model error by a factor of 2.

Multiple linear regression and neural network models were compared for various combinations of model inputs. Neural network modelling resulted in the most accurate models, while multiple linear regression models provided comparable levels of accuracy. The most accurate model was a *3:3:1* neural network model using cyclone feed flowrate, feed density and overflow density as inputs. The multiple linear regression model using the same inputs provided similar levels of accuracy. The similar performance of highly non-linear neural network models and linear regression models illustrates that very little benefit is realised by including non-linear modelling elements.

Due to its simplicity, ease of calibration and maintenance, small identification data requirements and comparable accuracy, a simple multiple linear regression model represents the final model choice. Model inputs are cyclone feed flowrate, feed density and overflow density.

To ensure robustness and availability in the industrial environment, a family of three secondary, less accurate models has been identified. Each secondary model contains only two inputs, alternating between neglecting each of the three possible inputs in turn. To allow for input instrument drift and changes in calibration as well as ore changes, a re-calibration procedure is suggested. An offset variable is added to the model. The offset is calculated by obtaining the difference between the actual sampled particle size and the calculated size corresponding to a user defined 15 minute steady state period of operation.

This study has shown that the implementation of a robust soft-sensor is possible in the industrial environment. Comparable accuracy and higher availability than conventional

particle size measurement devices imply that the soft-sensor model may replace or supplement conventional equipment.

REFERENCES

Aldrich, C., *Neural Networks for the Process Industries*, Course Notes, 1995.

Austin, L.G., and Klimpel, R.R., 'An Improved Method of Analysing Classifier Data', *Powder Technology*, 29(1981), p. 277.

Baldi, P., 'Gradient Descent Learning Algorithms: A General Overview' Submitted to IEEE Trans. Neural Networks, 1992.

Baum, E.B., and Haussler, D., 'What Size Net Gives Valid Generalization?', *Neural Comp.*, 1(1989), pp. 151-160.

Bell, M.S., *Particle Size Estimator Calibration Manual,* MINTEK, Randburg, 1995.

Bhat, N., and McAvoy, T.J., 'Use of Neural Nets for Dynamic Modelling and Control of Chemical Process Systems', *Computers chem. Engng.*, 14(4/5)(1990), pp. 573-582.

Bradley, D., and Pulling, D.J., 'Flow Patterns in the Hydraulic Cyclone and their Interpretation in terms of Performance', *Trans.Instn.Chem.Engrs.*, 37(1959), pp. 34-45.

Bradley, D., *The Hydrocyclone*, Pergamon Press, Oxford, 1965.

Bretney, E., 'Water Purifier', *US Patent No. 453105*, 1891.

Chaston, I.R.M., 'A Simple Formula for Calculating the Approximate Capacity of a Hydrocyclone', *Trans.IMM*, 67(1958), p. 203.

Cilliers, J.J., and Hinde, A.L., 'An Improved Hydrocyclone Model for Backfill Preparation', *Minerals Engineering*, 4(1991), pp. 683-693.

Dahlstrom, D.A., 'Fundamentals and Applications of the Liquid Cyclone', *Chem.Eng.Prog.*, Symposium Series No.15, Mineral Engineering Techniques, 50(1954), p. 41.

References

Del Villar, R.G., Thibault, J., and Del Villar, R., 'Development of a Softsensor for Particle Size Monitoring', *Minerals Engineering*, 9(1996), pp. 55-72.

Fahlstrom, P.H., 'Studies of the Hydrocyclone as a Classifier', *Mineral Processing*, Proceedings 6th Int.Congress, Cannes, 1963.

Finch, J.A., Laplante, A.R., and del Villar, R., 'Modelling Cyclone Performance Curves with a Size Dependent Correction Factor', *Unpublished*, 1985, Cited by Plitt L.R. et al., 'Cyclone Modelling : A Review of Present Technology', *CIM Bulletin*, 80(1987), p. 41.

Gault, G.A., and Nageswararao, K., 'Hydrocyclone Classification Studies of Complex Nickel Ore', *Proc.Australas.Inst.Min.Metall.*, No.280(1981), pp. 61-65.

Gonzalez, G.D., and Meyer, W.H., 'Replacement of a Particle Size Distribution Sensor in Grinding Plant Control by an Estimator', *Fourth Conference on Control Engineering*, Gold Coast, Australia, 1990.

Gupta, A., and Eren, H., 'Mathematical Modelling and On-line Control of Hydrocyclones', *Aus.IMM Proceedings*, No.1(1990), pp. 31-41.

Hinde, A.L., 'Control of Milling Circuits Using Simple Feedback Loops', *Grinding Theory and Practice*, SAIMM Vacation School, Johannesburg, (1977), pp. 192-214.

Hsieh, K.T., 'Phenomenological Model of the Hydrocyclone', Ph.D. Thesis, University of Utah.

Hsieh, K.T., and Rajamani, K., 'Phenomenological Model of the Hydrocyclone : Model Development and Verification for Single Phase Flow', *Int.J.Mineral.Process.*, 22(1988), pp. 223-237.

Hsieh, K.T., and Rajamani, K., 'A Mathematical Model of the Hydrocyclone Based on Physics of Fluid Flow', *AIChE J.*, 37(1991), pp. 735-746.

Jacobs, R., 'Increased Rates of Convergence Through Learning Rate Adaption', *Neural Networks*, 1(1988), pp. 295-307.

References

Kelsall, D.F., 'A Study of the Motion of Solid Particles in a Hydraulic Cyclone', *Trans.Inst.Chem.Eng.*, 30(1952), p. 87.

Kelsall, D.F., 'A Further Study of the Hydraulic Cyclone', *Chem.Eng.Sci.*, 2(1953), p. 254.

Lilge, E.O., 'Hydrocyclone Fundamentals', *Trans.IMM*, 71(1962), p. 285.

Lilge, E.O., and Plitt L.R., 'Cone Force Equation and Hydrocyclone Design in Minerals Technology - An Inter-American Approach', *Am.Soc.Mech.Eng.*, (1968), p. 108.

Luckie, P.T., and Klimpel, R.R., 'Classification and its Interaction with other Mineral Processing Operations', *Advances in Mineral Processing*, ed., Somasundaran, P., (1986).

Lynch, A.J., and Rao, T., 'Modelling and Scale Up of Hydrocyclone Classifiers', *Proc.11th Int. Min.Proc.Cong.*, Cagliari, (1975).

Monredon, T.C., Hsieh, K.T., and Rajamani, K., 'Fluid Flow Model of the Hydrocyclone : An Investigation of Device Dimensions', *Int.J.Mineral.Process.*, 35(1992), pp. 65-83.

Nageswararao, K., 'Technical Note - A Generalised Model for Hydrocyclone Classifiers', *AusIMM Proceedings*, No.2(1995), p. 21.

NeuroSolutions User Manual, 3rd Edition, 1995.

Osborne, B.F., 'A Complete System for On-Stream Particle Size Analysis', *CIM Bull.*, 65(1972), pp. 97-107.

Plitt, L.R., 'The Analysis of Solid-Solid Separations in Classifiers', *CIM Bull.*, 64(1971), pp. 42-47.

Plitt, L.R., 'A Mathematical Model of the Hydrocyclone Classifier', *CIM Bull.*, 69(1976), pp. 114-123.

Plitt, L.R., and Kawatra, S.K., 'Estimating the Cut (d_{50}) Size of Classifiers Without Product Particle Size Measurement', *Int.J.Mineral.Process.*, 5(1979), pp. 369-378.

Plitt, L.R., Flintoff, B.C., and Turak, A.A., 'Cyclone Modelling : A Review of Present Technology', *CIM Bull.*, 80(1987), pp. 39-50.

Plitt, L.R., Conil, P., and Brousaud, A., 'An Improved Method of Calculating the Water Split in Hydrocyclones', *Min.Eng.*, 3(1990), pp. 533-535.

Rietema, K., *Cyclones in Industry*, Elserivier Press, 1962, Chapter 4.

Roth, M.W., 'Neural Network Technology and its Applications', *Johns Hopkins APL Technical Digest*, 9(3)(1988), pp. 242-253.

Seitz, R.A., 'Further Studies on the Use of Classifiers for the Control of Wet Grinding Circuits', *Int.J.Mineral.Process.*, 12(1984), pp. 239-249.

Van der Walt, T.J., Van Deventer, J.S.J., and Barnard, E., 'Neural Nets for the Simulation of Mineral Processing Operations : Part I. Theoretical Principles', *Minerals Engineering*, 6(1993), pp. 1127-1134.

Van der Walt, T.J., Van Deventer, J.S.J., and Barnard, E., 'Neural Nets for the Simulation of Mineral Processing Operations : Part II. Applications', *Minerals Engineering*, 6(1993), pp. 1135-1153.

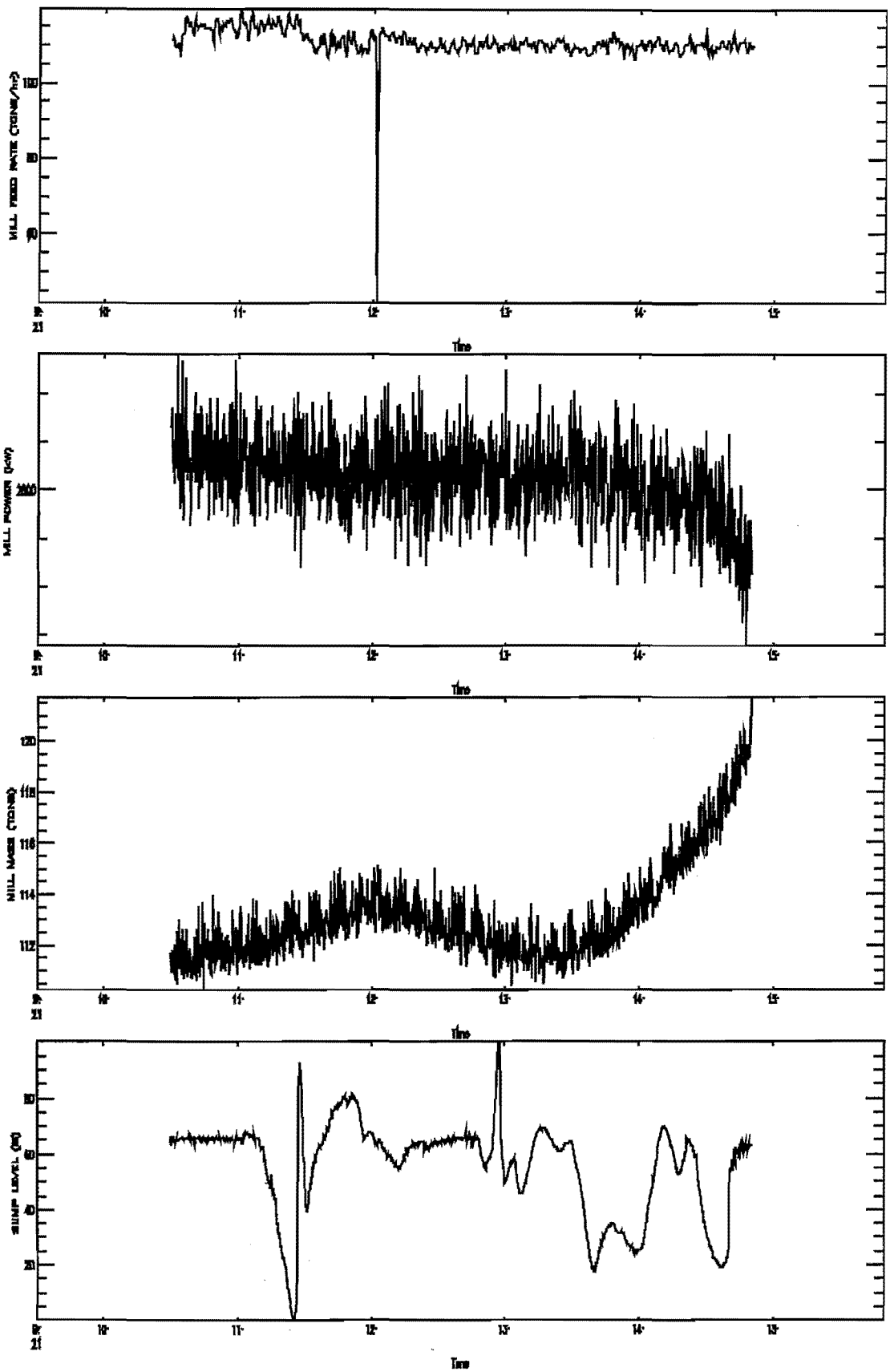
Viljoen, A.T., MINTEK Internal Communication, 1993.

Woollacott, L.C., Moys, M., and Hinde, A.,'The use of Material Balance Smoothing in the Evaluation of Backfill Operations', *J.S.Afr.Inst.Min.Metall.*, 92(5)(1992), pp. 121-129.

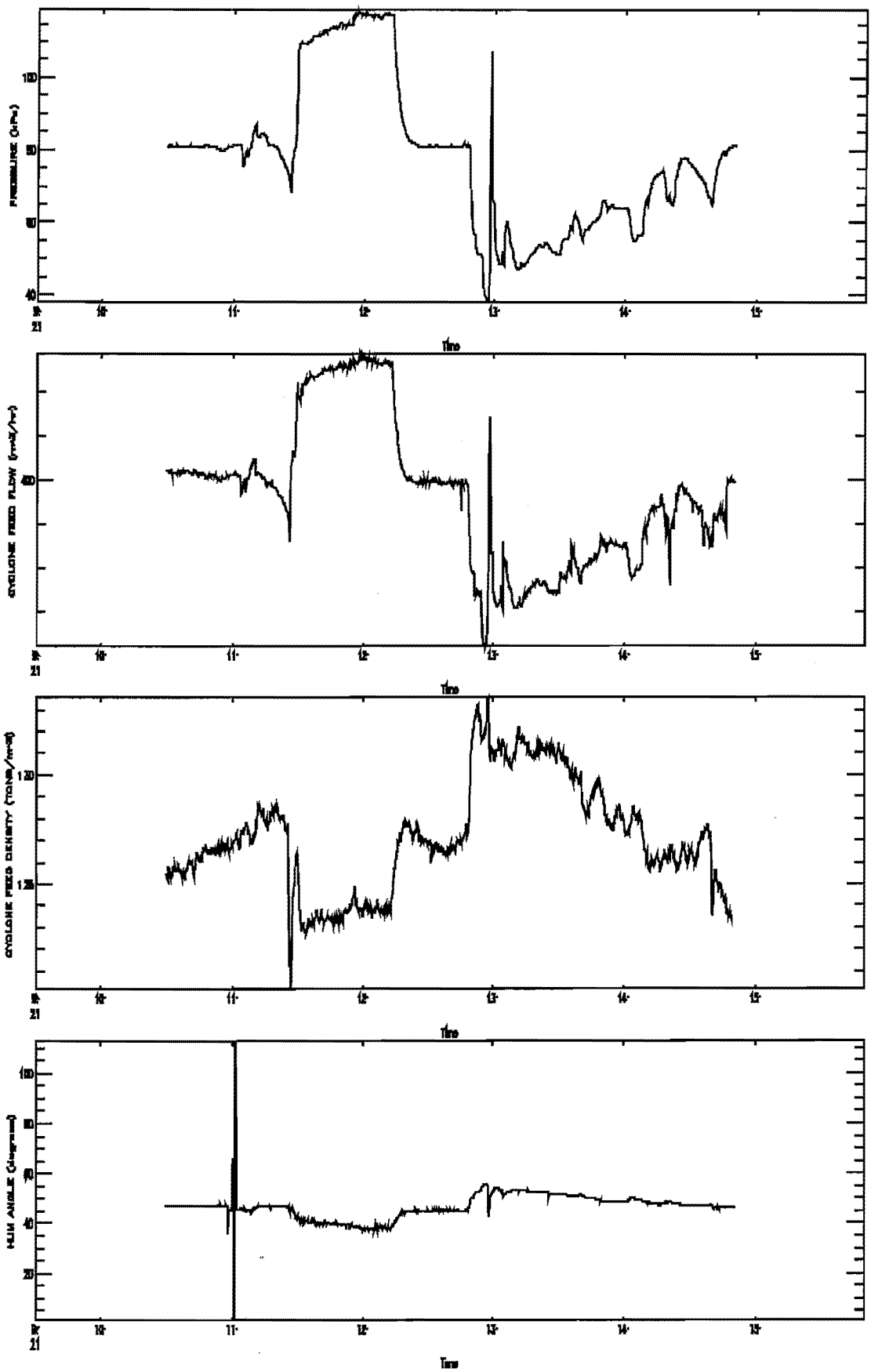
White, D.A., 'Efficiency Curve Model for Hydrocyclones Based on Crowding Theory', *IMM*, Section C, 100(1991), pp. C135-C138.

Yoshioka, N., and Hotta, Y., 'Liquid Cyclone as a Hydraulic Classifier', *Chem.Eng.*, 19(1955), p. 652.

APPENDIX A.1. AN EXAMPLE OF 10 SECOND LOGGED PLANT DATA



Appendix A.1. An Example of 10 Second Logged Plant Data



APPENDIX A.2. MASS BALANCE PROCEDURE AND CALIBRATION OF THE PLITT MODEL

Mass balance sampling and analysis procedure

As described in section 4.3, the mill discharge stream, cyclone overflow stream and cyclone underflow stream are sampled while the milling circuit is at steady state. Samples are prepared and analysed according to the procedure outlined in sections 4.4 and 4.5. The table below gives the uncorrected mass balance results used for calibration of the PSE model.

WDL Mass Balance: Mill 3 **11/2/97**

	Mill Discharge Sample		Cyclone Underflow Sample		Cyclone Overflow Sample	
Wet Mass	6804 g		2559.8 g		2425.2 g	
+1mm	265.8 g		241.8 g		0 g	
Dry Mass	3956.1 g		1527.8 g		528.7 g	
Cut Mass	140.0 g		134.3 g		133.7 g	
Pan	1.05 g		0.46 g		1.11 g	
Sieve	**Mass (g)**	**%**	**Mass (g)**	**%**	**Mass (g)**	**%**
1000	9.41	6.2957	21.26	13.6641	0.00	0.0000
425	5.10	3.4135	6.91	4.4422	0.02	0.0150
300	4.53	3.0320	7.12	4.5772	0.00	0.0000
212	9.10	6.0908	16.26	10.4529	0.02	0.0150
150	12.77	8.5472	26.46	17.0100	0.16	0.1197
106	18.18	12.1682	38.63	24.8336	3.81	2.8497
75	13.77	9.2165	17.96	11.5457	12.93	9.6709
53	12.45	8.3330	6.99	4.4936	17.46	13.0591
38	10.02	6.7065	3.42	2.1986	14.88	11.1294
Fines	54.08	36.1966	10.55	6.7822	84.42	63.1414
Total	149.41	100.00	155.56	100.00	133.70	100.00
% Solids		62.05		69.13		21.80

Mass balance smoothing

Mass balance smoothing involves fitting a corrected partition curve through the mass balance data. This enables the extraction of cyclone performance parameters, viz. d_{50c}, m, S and R_f. The general formulation of the mass balance problem is as follows:

$$Q = \sum_{k=1}^{K} w_k \left(Y_k - \hat{Y}_k\right)^2$$

where Y_k represents an actual value of an adjusted variable, viz. total mass flow of each stream, mass of each size class in each stream and water content of each stream.

$\hat{Y}_k$ is the improved estimate of Y_k.

w_k is a weighting factor, normally set to a magnitude of $1/Y_k^2$ to prevent data with numerically smaller values from being adjusted by a greater extent than data with higher values.

In addition to the simple set of incremental mass balance equations, the Rosin-Rammler (equation 7) partition function is added to reduce the size of the minimisation problem. The inclusion of the partition curve function allows partition numbers and water recoveries to be derived from the parameters in the cyclone model (equation 7). These can be substituted into the material balance equations. The minimisation procedure now only has to find values for d_{50}, m, R_f, S and the properties of the feed stream.

The Backfill software package is used to obtain best estimates for the parameters required for determining the partition curve. The software uses a Gauss-Newton non-linear regression technique.

Best-fit partition curve parameters are shown in Table 3 of section 4.6.

Calibrating the Plitt model equations

Eureka is a numerical environment for solving simultaneous non-linear equations. The Plitt model equations, equations 25-29, together with d_{50}, m, R_f, S, operational and geometric characteristics of the cyclone are entered in the Eureka program. The 5 Plitt model equations are solved simultaneously to provide an estimate of the 4 Plitt model correction factors, F_1 to F_4. The Plitt model is now calibrated to provide particle size estimates for changes in the operational variables, viz. feed flowrate and density. The values of the correction factors are given in Table 6 of section 5.1.

APPENDIX A.3. LOG SHEETS FOR RECORDING PLANT DATA

Sample Number	CFF average	CFF std	CFD average	CFD std	CUA average	CUA std	U/F Density	O/F Density	O/F %-75 mic
1	316.241	6.5221	1.3348	0.0039	46.7204	1.9965	71.1808	25.8094	73.658
2	314.983	8.0268	1.3373	0.0021	45.3922	0.6895	71.8500	24.5796	74.818
3	396.348	12.1643	1.3228	0.0031	39.8539	1.3447	73.9807	23.4155	74.143
4	404.683	4.4236	1.2444	0.0150	40.4144	3.4685	70.8669	18.3700	83.148
5	365.369	3.2046	1.3823	0.0092	50.9291	0.8279	70.2402	28.1519	71.327
6	371.184	2.8847	1.3320	0.0020	46.8328	0.7132	71.9499	24.0814	75.369
7	372.309	2.1375	1.3298	0.0021	46.3547	0.5375	69.9098	25.7404	75.444
8	486.621	5.9738	1.3028	0.0022	16.2052	5.9690	72.7065	19.4690	81.182
9	335.046	5.2027	1.3394	0.0034	43.8739	0.5053	70.1708	24.3711	76.280
10	421.611	6.2716	1.3172	0.0052	26.0170	1.1866	71.7518	23.1134	72.843
11	536.138	4.4081	1.3008	0.0026	39.8967	2.0205	75.9490	17.8993	79.113
12	408.911	4.2910	1.3269	0.0017	22.0457	1.9999	71.8091	21.4841	76.794
13	273.719	2.9848	1.3614	0.0066	58.1848	0.5958	70.2777	35.9781	67.755
14	538.736	4.6619	1.2658	0.0053	23.6396	6.9208	72.4872	12.6414	92.976
15	511.375	4.4365	1.2809	0.0029	22.4495	3.5435	72.6978	14.2687	90.531
16	395.830	3.2232	1.2650	0.0022	40.9898	0.8265	71.7528	24.8094	71.643
17	493.347	3.2997	1.2214	0.0024	36.1101	0.4316	74.4238	15.1285	88.124
18	412.024	1.1567	1.2966	0.0015	21.8627	0.3314	77.0388	27.2991	65.448
19	281.397	2.7870	1.3318	0.0017	49.7627	0.3002	66.8167	37.8336	60.846
20	399.172	2.7440	1.2841	0.0032	36.9502	0.4714	75.0436	26.2261	65.266
21	339.936	1.1012	1.3042	0.0015	42.0109	0.5323	71.9291	30.2394	64.116
22	441.914	2.7393	1.2801	0.0023	21.4520	0.4371	76.6301	25.0849	65.857
23	402.977	1.6889	1.2112	0.0014	57.2173	0.1219	66.9364	16.1484	88.204
24	450.446	2.1651	1.2159	0.0016	53.8007	0.1516	69.1634	17.7937	84.156
25	364.316	1.7561	1.2352	0.0021	56.9030	0.1426	64.6993	21.8539	79.240
26	306.477	1.6710	1.2512	0.0016	60.5889	0.5904	61.7377	25.2245	75.813
27	433.432	3.4091	1.2095	0.0009	55.1638	0.2649	66.1197	17.4323	86.669
28	471.481	3.4345	1.2003	0.0021	54.5906	0.3901	68.1790	15.6606	89.965
29	507.725	4.1401	1.1991	0.0012	53.2234	0.1702	67.7813	14.6274	88.802
30	241.816	2.4733	1.2760	0.0051	63.7436	0.7391	62.6423	32.0603	68.827
31	391.383	1.1489	1.2298	0.0017	42.3550	0.1162	69.4149	18.8896	77.921
32	362.971	2.6232	1.2337	0.0017	43.5621	0.1670	66.5576	19.3862	77.493
33	293.392	4.1757	1.2659	0.0014	51.8098	0.1471	52.3974	29.0042	71.871
34	398.639	3.6524	1.2005	0.0020	52.1916	0.2339	63.4710	14.1903	90.923
35	507.769	5.6716	1.2118	0.0019	42.7904	0.4010	72.2220	15.8537	83.832
36	490.844	3.7908	1.2107	0.0024	42.2989	0.3666	71.2235	14.3434	85.040
37	423.903	3.0172	1.2200	0.0022	42.4731	0.6743	69.8694	17.4124	81.701
38	403.229	2.4489	1.2665	0.0016	45.2500	0.3641	70.8019	25.8348	70.848
39	381.218	9.8861	1.2817	0.0026	47.0984	0.1590	70.7887	28.5125	69.191
40	527.902	2.0997	1.2336	0.0012	39.7748	0.3165	73.1312	18.5394	74.996

Appendix A.3. Log Sheets for Recording Plant Data

Sample Number	CFF average	CFF std	CFD average	CFD std	CUA average	CUA std	U/F Density	O/F Density	O/F %-75 mic
41	534.3672	2.4871	1.2382	0.0017	37.8425	0.5012	74.052331	19.367694	71.857355
42	272.7173	0.9217	1.3112	0.0024	51.9276	0.423	66.564912	35.187925	65.166909
43	294.624	3.7558	1.28	0.0033	49.8845	0.1887	65.951904	28.019276	71.32321
44	342.5018	13.4944	1.2725	0.0027	47.3045	0.1676	68.453982	27.101855	68.156618
45	397.0473	1.7125	1.2395	0.0022	40.8699	0.3885	70.845959	21.44521	75.986842
46	431.9449	4.3535	1.2075	0.001	41.8778	0.5764	68.896271	16.280414	83.826607
47	479.8273	2.1497	1.2009	0.002	39.4078	0.646	70.203702	15.504066	84.944056
48	516.4435	2.9485	1.2017	0.002	38.0227	0.2073	69.264999	16.282421	83.173145
49	383.432	2.2338	1.2494	0.0024	42.2178	0.2631	69.795278	24.330939	72.408339
50	323.52	2.8724	1.2621	0.0035	45.7585	0.4446	65.856368	28.502865	69.977843
51	409.0114	1.4018	1.2744	0.0024	44.0259	0.1941	72.168675	25.396046	67.590361
52	440.2674	5.2585	1.2683	0.003	40.6166	0.6465	74.814132	26.774194	61.514936
53	514.1736	1.981	1.237	0.0011	37.7963	6.1604	75.582187	19.806289	69.434103
54	262.9545	2.1953	1.3583	0.0027	43.1567	0.086	65.353352	36.295823	62.58832
55	348.8976	3.6179	1.3149	0.0079	39.3095	0.3408	68.429232	30.16824	67.174377
56	526.4673	15.8509	1.3081	0.0048	33.5015	0.3299	72.038353	16.158974	79.071322
57	288.828	5.5609	1.3841	0.0057	42.7995	0.1824	66.926308	31.511539	66.873598
58	336.744	4.0044	1.3659	0.0039	43.324	0.0577	66.179921	29.923618	68.601504
59	406.5876	4.6729	1.339	0.002	42.752	0.0728	69.38293	24.046068	72.494593
60	435.381	2.7695	1.3335	0.0013	41.856	0.0785	70.907217	24.611337	71.286904
61	506.7141	24.1788	1.3485	0.0035	40.5113	0.0765	74.001504	22.739602	70.631104
62	399.072	3.6133	1.3652	0.0045	39.606	0.2356	70.768667	24.529181	72.144374
63	354.7985	37.6796	1.384	0.004	44.5945	0.163	70.282401	30.2884	67.940177
64	486.8218	3.4769	1.3344	0.0049	36.6421	0.1488	72.0736	22.365326	73.692762
65	398.6788	2.6828	1.3461	0.0057	40.1494	0.2102	69.205655	27.88956	65.472561
66	445.4363	1.5964	1.3279	0.0055	37.4218	0.0653	70.920378	26.406541	69.121622
67	305.501	0.7211	1.4253	0.0032	45.2145	0.258	64.257649	38.247713	63.596554
68	365.7113	4.3228	1.3639	0.0028	42.0722	0.1419	66.782985	30.7369	68.987915
69	339.4604	1.6943	1.3945	0.0034	43.4695	0.0954	67.134558	35.197049	66.117479
70	388.9571	0.8012	1.5328	0.0085	37.4058	0.0486	71.23866	31.703454	64.934966
71	441.0275	2.9124	1.3798	0.0054	35.6353	0.0514	72.673528	28.873214	64.185351

APPENDIX A.4. TRAINING AND TEST DATA SETS

TRAINING DATA SET (50 SAMPLES)

Sample Number	CFF average	CFD average	CUA average	U/F Density	O/F Density	O/F %-75 mic
1	316.2410	1.3348	46.7204	71.1808	25.8094	73.658
2	314.9826	1.3373	45.3922	71.8500	24.5796	74.818
3	396.3477	1.3228	39.8539	73.9807	23.4155	74.143
4	365.3691	1.3823	50.9291	70.2402	28.1519	71.327
5	371.1840	1.3320	46.8328	71.9499	24.0814	75.369
6	372.3093	1.3298	46.3547	69.9098	25.7404	75.444
7	335.0461	1.3394	43.8739	70.1708	24.3711	76.280
8	538.7356	1.2658	23.6396	72.4872	12.6414	92.976
9	511.3746	1.2809	22.4495	72.6978	14.2687	90.531
10	412.0240	1.2966	21.8627	77.0388	27.2991	65.448
11	281.3973	1.3318	49.7627	66.8167	37.8336	60.846
12	399.1724	1.2841	36.9502	75.0436	26.2261	65.266
13	339.9360	1.3042	42.0109	71.9291	30.2394	64.116
14	450.4460	1.2159	53.8007	69.1634	17.7937	84.156
15	364.3161	1.2352	56.9030	64.6993	21.8539	79.240
16	306.4773	1.2512	60.5889	61.7377	25.2245	75.813
17	433.4324	1.2095	55.1638	66.1197	17.4323	86.669
18	471.4812	1.2003	54.5906	68.1790	15.6606	89.965
19	507.7246	1.1991	53.2234	67.7813	14.6274	88.802
20	241.8160	1.2760	63.7436	62.6423	32.0603	68.827
21	391.3833	1.2298	42.3550	69.4149	18.8896	77.921
22	362.9713	1.2337	43.5621	66.5576	19.3862	77.493
23	293.3918	1.2659	51.8098	52.3974	29.0042	71.871
24	507.7692	1.2118	42.7904	72.2220	15.8537	83.832
25	490.8439	1.2107	42.2989	71.2235	14.3434	85.040
26	403.2288	1.2665	45.2500	70.8019	25.8348	70.848
27	381.2184	1.2817	47.0984	70.7887	28.5125	69.191
28	527.9017	1.2336	39.7748	73.1312	18.5394	74.996
29	534.3672	1.2382	37.8425	74.0523	19.3677	71.857
30	272.7173	1.3112	51.9276	66.5649	35.1879	65.167

Appendix A.4. Training and Test Data Sets

TRAINING DATA SET (CONTINUED)

Sample Number	CFF average	CFD average	CUA average	U/F Density	O/F Density	O/F %-75 mic
31	294.6240	1.2800	49.8845	65.9519	28.0193	71.323
32	479.8273	1.2009	39.4078	70.2037	15.5041	84.944
33	516.4435	1.2017	38.0227	69.2650	16.2824	83.173
34	383.4320	1.2494	42.2178	69.7953	24.3309	72.408
35	323.5200	1.2621	45.7585	65.8564	28.5029	69.978
36	440.2674	1.2683	40.6166	74.8141	26.7742	61.515
37	514.1736	1.2370	37.7963	75.5822	19.8063	69.434
38	262.9545	1.3583	43.1567	65.3534	36.2958	62.588
39	348.8976	1.3149	39.3095	68.4292	30.1682	67.174
40	526.4673	1.3081	33.5015	72.0384	16.1590	79.071
41	288.8280	1.3841	42.7995	66.9263	31.5115	66.874
42	506.7141	1.3485	40.5113	74.0015	22.7396	70.631
43	399.0720	1.3652	39.6060	70.7687	24.5292	72.144
44	354.7985	1.3840	44.5945	70.2824	30.2884	67.940
45	445.4363	1.3279	37.4218	70.9204	26.4065	69.122
46	305.5010	1.4253	45.2145	64.2576	38.2477	63.597
47	365.7113	1.3639	42.0722	66.7830	30.7369	68.988
48	339.4604	1.3945	43.4695	67.1346	35.1970	66.117
49	388.9571	1.5328	37.4058	71.2387	31.7035	64.935
50	441.0275	1.3798	35.6353	72.6735	28.8732	64.185

Appendix A.4. Training and Test Data Sets

TEST DATA SET (21 SAMPLES)

Sample Number	CFF average	CFD average	CUA average	U/F Density	O/F Density	O/F %-75 mic
1	404.6832	1.2444	40.4144	70.8669	18.3700	83.148
2	486.6211	1.3028	16.2052	72.7065	19.4690	81.182
3	421.6114	1.3172	26.0170	71.7518	23.1134	72.843
4	536.1382	1.3008	39.8967	75.9490	17.8993	79.113
5	408.9107	1.3269	22.0457	71.8091	21.4841	76.794
6	273.7193	1.3614	58.1848	70.2777	35.9781	67.755
7	395.8302	1.2650	40.9898	71.7528	24.8094	71.643
8	493.3469	1.2214	36.1101	74.4238	15.1285	88.124
9	441.9136	1.2801	21.4520	76.6301	25.0849	65.857
10	402.9768	1.2112	57.2173	66.9364	16.1484	88.204
11	398.6394	1.2005	52.1916	63.4710	14.1903	90.923
12	423.9033	1.2200	42.4731	69.8694	17.4124	81.701
13	342.5018	1.2725	47.3045	68.4540	27.1019	68.157
14	397.0473	1.2395	40.8699	70.8460	21.4452	75.987
15	431.9449	1.2075	41.8778	68.8963	16.2804	83.827
16	409.0114	1.2744	44.0259	72.1687	25.3960	67.590
17	336.7440	1.3659	43.3240	66.1799	29.9236	68.602
18	406.5876	1.3390	42.7520	69.3829	24.0461	72.495
19	435.3810	1.3335	41.8560	70.9072	24.6113	71.287
20	486.8218	1.3344	36.6421	72.0736	22.3653	73.693
21	398.6788	1.3461	40.1494	69.2057	27.8896	65.473

APPENDIX A.5. THE ARX MODEL (Auto-Regression with eXternal input)

The structure of the model for a single input is given by a simple linear difference equation:

$$y(t) + a_1y(t-1) + \ldots + a_{na}y(t-na) = b_1u(t-nk) + \ldots + b_{nb}u(t-nk-nb+1) + e(t)$$

Where na and nb refer to the order of input and output respectively

nk is the time delay in sample interval units (i.e. if logging interval is 10 s, then nk = 1 refers to a time delay of 10 s)

vectors **a** and **b** are the model constants to be determined

e is the model error

The above equation is for a single input model. Further inputs can be added as shown below for the 2-input zero order ARX model used for underflow angle modelling.

$$y(t) = b_{11}u_1(t-nk_1) + b_{21}u_2(t-nk_2) + e(t)$$

Or more specifically,

$$CUA(t) = dCFF(t-1) + eCFD(t-4) + e(t)$$

It should be noted here that CUA lags CFF by 1 sample period (10 seconds) and the delay with respect to CFD is 4 sample periods (40 seconds).

The model relates the current output y(t) to a finite number of past outputs y(t – k) and inputs u(t – k).

In the above equation the only unknown parameters are the vectors **a** and $\mathbf{b_1}$ **to** $\mathbf{b_n}$ (for n inputs) where

$$\mathbf{a} = [a_1\ a_2\ \ldots\ a_{na}] \quad \text{and} \quad \mathbf{b_n} = [b_1\ b_2\ \ldots\ b_{na}]$$

Solution of Equations

Least Squares – Minimises the sum of squares of the right-hand side minus the left-hand side of the expression above, with respect to each element in **a** and **b**. This creates a system of linear equations in terms of the parameters above.

Appendix A.5. The ARX Model (Auto Regression with eXternal input)

Least squares formulation:

$$MinP = \frac{1}{N} \sum_{t=1}^{N} [e(t)]^2$$

P is minimum where

$$\frac{\partial P}{\partial b_1} = \frac{\partial P}{\partial b_2} = 0$$

Matlab has a built-in ARX function that computes Least Squares estimates of ARX-models. The syntax used is:

>> th = arx(z,nn)

th: returned as the estimated parameters of the ARX model along with the structure of the model. **th** represents **theta** which is a matrix containing information about the model structure, estimated parameters and their estimated accuracy.

z: the output-input data **z = [y u]** with **y** and **u** as column vectors. Multiple inputs (u's) can be specified.

nn: **nn = [na nb nk]** as given above. Where nb and nk are repeated successively for multiple inputs.

APPENDIX A.6. DEFINITION OF USEFUL STATISTICAL CRITERIA

Sample Correlation

The sample correlation indicates how well input x is correlated to output y and is defined by the equation below:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}}$$

where: the summation refers to the entire data set considered

subscript i refers to the i^{th} data point

$\bar{x}$ and $\bar{y}$ refer to the mean of the input and output vectors respectively.

The sample correlation always lies in the interval from –1 to 1. When $r = -1$ then x is perfectly correlated with y in a negative sense, viz. as x increases, y decreases. When $r = 1$ then x is perfectly correlated with y in such a way that an increase in x represents an increase in y. Perfect correlation implies that all (x,y) data points fall on a straight line. r is a measure of the strength of the linear relationship between x and y. A strong linear relationship is indicated by r-values near to 1 and –1, while near zero r-values indicate a lack of linear relationship.

Coefficient of Determination

The coefficient of determination is a quantification of the quality of the fitted equation and is given by:

$$R^2 = \frac{\sum (y_i - \bar{y})^2 - \sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$$

where: $\hat{y}$ represents the predicted value or model output

R^2 provides information on the fraction of the raw variation in y accounted for using the fitted equation. High values of R^2 (with a maximum of 1) indicate that the model obtained accounts for a significant amount of the variance in y. Low R^2 values imply that the model is

not able to account for enough of the variance in y and normally implies that a poor fit results.

Sample Mean

The sample mean ($\bar{x}$) is given by:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

where: n is the number of data points

x_i is the value of the i^{th} sample

Median

The median is defined as the "middlemost" data value when the data are arranged according to size. More precisely, if the observations are arranged according to size and n is an odd number, the median is the value of observation $(n+1)/2$; if n is an even number, the median is defined as the mean of the observations numbered $n/2$ and $(n+2)/2$.

Standard Deviation

The standard deviation (s) is given by:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

Confidence Interval

Suppose that a large sample (n>30) with population mean μ and known variance s^2 exists. Then the equation below defines the confidence limits of μ:

$$\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Appendix A.7. MULTIPLE LINEAR REGRESSION SOLUTION PROCEDURE

In multiple linear regression, the aim is to fit a surface described by:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + ... + b_r x_r$$

where $\hat{y}$ = predicted output

x_i = ith model input

b_i = coefficient corresponding to the ith input

The coefficients, b_0 to b_r are solved in a least squares sense by minimising the objective function defined by:

$$O_j = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left(y_i - \left(b_0 + b_1 x_1 + b_2 x_2 + ... + b_r x_r\right)\right)^2$$

where n = number of data points available

y_i = actual value of the ith output

Setting the partial derivatives with respect to b_i equal to 0 gives $r + 1$ linear equations in $r + 1$ unknowns, $b_0, b_1,...,b_r$. The solution of the equation set defined below provides the best estimate of $b_0, b_1,...,b_r$.

$$nb_0 + \left(\sum x_{1i}\right)b_1 + \left(\sum x_{2i}\right)b_2 + ... + \left(\sum x_{ri}\right)b_r = \sum y_i$$

$$\left(\sum x_{1i}\right)b_0 + \left(\sum x_{1i}^{\,2}\right)b_1 + \left(\sum x_{1i}x_{2i}\right)b_2 + ... + \left(\sum x_{1i}x_{ri}\right)b_r = \sum x_{1i}y_i$$

$$\left(\sum x_{2i}\right)b_0 + \left(\sum x_{1i}x_{2i}\right)b_1 + \left(\sum x_{2i}^{\,2}\right)b_2 + ... + \left(\sum x_{2i}x_{ri}\right)b_r = \sum x_{2i}y_i$$

$$\left(\sum x_{ri}\right)b_0 + \left(\sum x_{1i}x_{ri}\right)b_1 + \left(\sum x_{2i}x_{ri}\right)b_2 + ... + \left(\sum x_{ri}^{\,2}\right)b_r = \sum x_{ri}y_i$$

APPENDIX A.8. DETERMINATION OF NEURAL NETWORK WEIGHTS IN EXCEL

The neural network was implemented in Excel to solve for the network weights. The Excel Solver was used to obtain a solution. The solution procedure is outlined in a step-wise sense below:

1. A weight matrix (which includes weights for bias nodes associated with the hidden and output layers) is entered so that it contains random numbers between –0.1 and 0.1.
2. Input and output data are added in columns for the training and test data separately.
3. The input data are scaled between –1 and 1 and the output data are scaled between –0.9 and 0.9 according to equation 48.

$$x_{sc} = \left(\frac{max_{sc} - min_{sc}}{max_{ac} - min_{ac}}\right) \cdot x_{ac} + max_{sc} - \left(\frac{max_{sc} - min_{sc}}{max_{ac} - min_{ac}}\right) \cdot max_{ac}$$

4. The scaled inputs, together with the hidden layer weights are then used to calculate the input into each hidden node according to equation 44.

$$s_i = \sum_{j=1}^{n} w_{ij} x_j + w_{bias}$$

5. The inputs from 4 above are then transformed by equation 46 to obtain the output from each hidden node.

$$z_i = \tanh(s_i)$$

6. The hidden layer outputs, combined with the associated output weights are then used to calculate the input to the single hidden layer by using the same procedure as shown in 4 above.
7. These inputs are then transformed by the tanh function described in 5.
8. The calculated network outputs are then "unscaled" by applying the inverse of equation 48.
9. The MSE is calculated by equation 51.

$$E = \frac{1}{N}\left[\sum_{i=1}^{N}\left(d_j(i) - y_j(i)\right)^2\right]$$

10. The Excel solver is then used to update the network weights in such a way that the minimum training MSE is obtained.
11. Training is manually terminated when the test MSE begins to increase. This procedure requires continual monitoring of the test MSE.

More specifically, the objective function is defined below:

$$O_j = \frac{1}{N}\left[\sum_{i=1}^{N}\left(d_j(i) - y_j(i)\right)^2\right]$$

This function is minimised by the Excel Solver by setting the objective function equal to 0 – this causes the network weights to be adjusted. It should be noted that only training data are presented to the network for calculation of the MSE.

The Solver employs a quasi-Newton method using central derivatives and quadratic extrapolation. Precision and convergence criteria can also be specified.

Without taking part in training, the test data set is used for calculation of the test MSE. These test MSE values are monitored during training and serves as a stop criteria.

The entire training data set is used for each iteration, i.e. weights are updated only after presentation of the complete data set. This is known as batch training.