UNIVERSITY OF BIRMINGHAM

ROLL COMPACTION OF PHARMACEUTICAL EXCIPIENTS AND PREDICTION USING INTELLIGENT SOFTWARE

By

RACHEL FRAN MANSA

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Chemical Engineering
School of Engineering
The University of Birmingham
July 2006

Summary

Roll compaction is a dry granulation method. In the pharmaceutical industry it assists in binding tablet ingredients together to form a larger mass. This is conducted to ease subsequent processing, decrease dust, improve flowability, improve material distribution, more suitable for moisture and heat sensitive materials than wet granulation methods, minimises operating space and suited for a continuous manufacturing set-up. In pharmaceutical roll compaction various types of powder material mixtures are compacted into ribbon that are subsequently milled and tableted. The aim of this research is to investigate the use of intelligent software (*FormRules* and *INForm* software) for predicting the effects of the roll compaction process and formulation characteristics on final ribbon quality. Firstly, the tablet formulations were characterised in terms of their particle size distribution, densities, compressibility, compactibility, effective angle of friction and angle of wall friction. These tablet formulations were then roll compacted. The tablet formulation characteristics and roll compaction results formed 64 datasets, which were then used in *FormRules* and *INForm* software training. *FormRules* software highlighted the key input variables (i.e. tablet formulations, characteristics and roll compaction process parameters). Next these key input variables were used as input variables in the model development training of *INForm*. The *INForm* software produced models which were successful in predicting experimental results. The predicted nip angle values of the *INForm* models were found to be within 5%, which was more accurate to those derived from Johanson's model prediction. The Johanson's model was not successful in predicting nip angle above the roll speed of 1 rpm due to air entrainment. It also over-predicted the experimental nip angle of DCPA and MCC by 200%, while the approximation using Johanson's pressure profile under-predicted the experimental nip angle of DCPA by 5-20% and MCC by 20%.

Dedicated to

Anh Tran

for your inspiration, understanding and friendship

Acknowledgements

Table of Contents   Page

1 Introduction

ABSTRACT

Roll compaction is a simple and cost effective dry granulation method used to bind smaller particles together to form a bigger mass to ease processing, decrease dust, improve flowability and improve material distribution. This section presents the basic process of roll compaction, the types of roll compaction assembly designs, the research objectives of this work and the research strategies employed.

The roll compaction process involves compressing powder material in between two counter rotating rollers. Roll compaction is traditionally used in the metallurgical industry to produce metal sheets, the mining industry to produce coal briquettes from waste coal powder, the mineral industry to crush rocks, the agricultural industry to make fertilizer pellets by briquettting, and the pharmaceutical industry to produce granules for tabletting. The last of these uses is of interest in this work. Figure 1.1 shows the role of roll compaction within the pharmaceutical tableting process.

Mixing of drug and excipients

↓

Roll Compaction

↓

Milling

↓

Tabletting
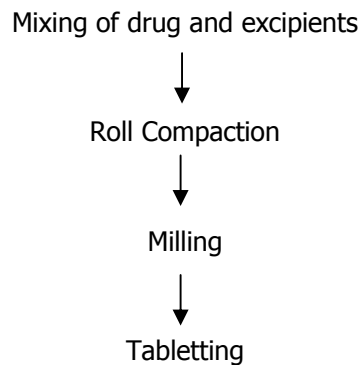
Figure 1.1 Tabletting process

1.1 Basics of the Roller Compactor Process

Figure 1.2 shows the schematic diagram of roll compactor. The process consists of three regions: the slip region, the nip region and the release region. In the slip region, the feed powder flows into the roll gap but the rollers are moving faster than the powder so that relative slip occurs. De-aeration also occurs and there comes a point at which the powder is gripped by the rollers (at the "nip angle", α) so that relative movement ceases between the powder and roll surface, and compaction

occurs. In the nip region the pressure may be very high (up to ~ 230 MPa for the roller compactor used here). A greater nip angle implies a larger maximum stress, so that to achieve acceptable compaction, the nip angle must be sufficiently large. In the release region, the compacted ribbon may show stress relaxation as it is released from the rolls. The resulting expansion of the compact is a function of the physical characteristics of the material, roll diameter, roll speed and roll pressure. Details of the press employed in this project are provided in section 4.2.1.
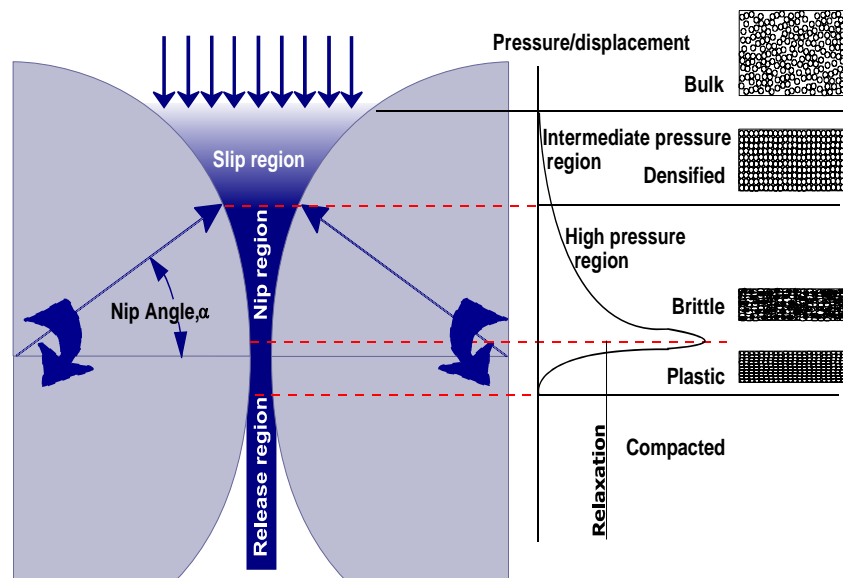
Figure 1.2 Schematic of the roller compaction process.

1.2 Possible Roll Compactor Designs

Various designs of the roll compactor are available. Basically the designs can be classified into five major characteristics of the design: i) roller assembly, ii) side sealing, iii) feeding mechanism, iv) roll surface type and v) roll layout.

1.2.1 Roller Assembly

The rollers can be fixed (usually instrumented and specifically made) or adjustable (e.g. Fitzpatrick L83 Chilsonator, Gerteis 3 W-Polygram etc). The fixed rollers do not move during compaction except for a slight give during higher pressure compaction and the compaction pressure is determined by the set roll gap and roll speed. On the other hand, adjustable rollers were made to be movable to control the roll compaction pressure during roll compaction. The fixed rollers are prone to powder jamming as opposed to the adjustable rollers.

1.2.2 Side Sealing

Sealing on either end of the roller gap is necessary to stop the powders from flowing away from the space between the two rollers. There can be two types of sealing; the side plate assembly and the rim roll assembly. The side plate is usually made of PTFE-platelet to avoid metal-metal friction and wear. The rim rolls consist of one roller with rims on both side and the other roll runs within the cavity of this rim. Imagine male and female fitting whereby one fits within the other. This allows the resistance to high transversal pressure without losing the sealing capability.

1.2.3 Feeding Mechanism

The feeding mechanism can be by gravity or by force feeding using a screw feeder. In gravity feeding method a hopper guides the powder material to the region between the two rollers and the hopper is filled up with powder material. The force feeding system is usually applied using a screw which is placed inside a hopper and the screw end is directed to the space between the two rollers. The screw pushes powder material inside the hopper towards the roll compaction region and this reduces the effect of air entrainment in the feeding region. The distance between the screw end and the roll compaction region will also determine the ribbon compact quality.

1.2.4 Roll Surface Type

The types of roller surface are smooth, rough, corrugated, ribbed, pocketed etc. The rough, ribbed or corrugated roll compactor surface assists the compaction process by providing grip on the powder material. A pocketed roller surface is used for briquetting.

1.2.5 Roll Layout

The fifth major characteristic design of the roll compactor is the layout of the rollers. There are three types of layouts; vertical, horizontal and inclined. The commonly

used layout is the horizontal layout. The powder material flows by gravity or screw feeder into the space between the rollers and undergoes compaction. The inclined layout allows easy harvesting of the ribbon compacts (e.g. Gerteis compactors inclined at 30°). The vertical layout might cause the roll compaction of the powder material to be unsuccessful because the draw angle is fairly small and the powder material remains at the slip zone since gravity is not overcome by friction with the roll surface.

1.3 Research Objectives

The overall aim of this research is to predict roll compaction output properties from tablet formulation and roll compaction process parameters using intelligent software (i.e. *Formrules* and *INForm* software). The specific aims are to characterise the tablet formulation and the ribbon characteristics. These results are related to the roll compaction processing conditions using the intelligent software. Next the intelligent software prediction is compared to traditional mathematical model predictive method (i.e. Johanson's Theory, (Johanson, 1965)). This was done to assess the strengths and weaknesses of the intelligent software predictions and to be compared against a well established mathematical theoretical model.

The novelty of this research lies in the fact that it relates the effects of varying the tablet formulations on the final ribbon quality within the intelligent software. Then

the use of the intelligent software provides the ability of predicting roll compaction process parameters for specific formulations and required ribbon quality.

1.4 Investigation Strategy

First of all it was important to measure the physical characteristics of the powder material (Chapter 3). The powder characterisation information was used for both the training of intelligent software and the calculations associated with the theoretical approach. The intelligent software was used in the form of a numeral representation of the powder according to certain physical characteristics, whereas, the theoretical approach used the Compressibility κ, angle of wall friction and effective angle of internal friction as independent variables.

The next step was to conduct roll compaction experiments on the various tablet formulations (Chapter 4). This was conducted to build a database of information which was to be examined by the intelligent software. Subsequently the prediction of the roll compaction output properties using a theoretical approach was conducted (Chapter 5). Then the key variables (i.e. most important powder characteristics) were investigated using a datamining intelligent software (Chapter 6). Information from Chapter 6 was then used in Chapter 7 to develop models. After that the model predictability was investigated and compared to the theoretical model.

2 Literature Survey

ABSTRACT

Roll compaction use in pharmaceutical application was first reported in 1966 by Jaminet and Hess, and Cohn *et al*. The first model to describe and predict the process of roll compaction was developed a year prior to that (Johanson, 1965). In the subsequent years rigorous roll compaction research was then conducted. This research mainly covered the effects of the process parameters on roll compaction, validation of existing models and novel methods for improving the roll compaction process. However, there is little information regarding the relationships between powder material characteristics and the quality of the final roll compacted ribbon. This is very important in the pharmaceutical industry because a wide range of formulations are used. In other fields, neural networks have been found to successfully produce predictive models for processes with a high number of input variables that are multivariate in nature. Thus, intelligent software was chosen here for predictive model development. This section gives a short history of the roll compaction process, previous research conducted on roll compaction and background on the individual programs which make up the intelligent software used here.

2.1 A Brief History of Roll Compaction

Henry Bessemer first used roll compaction to produce bronze flakes to make pigments for gilt decorations in 1843. Then in the 1900s, Hardy (1938) and Siemens and Hardy (1904), produced patents of rolling metal powders. Next Naeser and Zirm (1950) produced a systematic experimental study of the roll compaction of iron strip. Subsequently in 1966, the first pharmaceutical applications of roll compaction were published by Jaminet and Hess, 1966, and Cohn *et al.*, 1966.

In 1965, Johanson introduced a Rolling Theory for Granular Solids. It was the first complex and the only fully predictive mathematical model which predicts the nip angle and the pressure distribution after the powder is nipped by the rollers. After two decades, Katashinkii (1986) used the slab method to produce an analysis of the nip region. The analysis predicted the pressure distribution and roll separating force in metal rolling processes. Then one decade later, Inghelbrecht *et al.* used a Multilayer Feed-Forward neural network (MLF) to model the granule friability as a function of roll compaction parameters. Next, Turkoglu *et al.* (1999) used neural networks and genetic algorithms to predict and optimise the effect of binder type, binder concentration and the number of roll compaction passes on the properties of compressed tablets.

In 2001, Odagi *et al.* developed a 2-D simulation of the roll compaction process using Discrete Element Method software. Loginov *et al.* (2001) produced a

briquetting simulator, which is a roll compactor with deep pockets on the surface to produce pellets and also a scale up mathematical model. Two years later, Dec *et al.* expounded the Finite Element Method for predicting the powder behaviours in the nip region. Recently Zinchuk *et al.* (2004) presented a roll compaction simulation method which was then used in scale up.   The background for all the modelling methods developed over the past years will be summarised in the next section except for Johanson's Theory which is explained in Chapter 5.

2.1.1 Models Developed to Describe and Predict Roll Compaction

2.1.1.1   Slab Method

The slab method is generally used in the metal rolling process to predict the pressure distribution and roll separating force. It was first used by Katashinkii (1986) to predict pressures during the rolling process of metal powders. The model considers plane sections which are assumed to be constant as they pass through the rolls.

These plane sections are represented by the trapezoidal slabs shown in Figure 2.1. The force balance on the slab gives the equilibrium equation for the x-direction and is expressed as:

$$\frac{\partial(h\sigma_x)}{\partial x} + 2\left(p\tan\alpha - \tau_f\right) = 0 \tag{2.1}$$
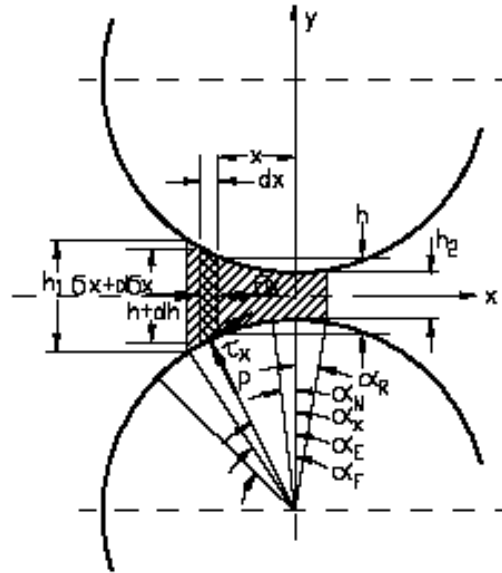
Figure 2.1 Stresses acting on the element in the roll nip region. (Figure is obtained from Dec *et al.*, 2003)

The frictional stress is expressed by:

$$\tau_f = Y(\rho): \quad for \mu(p)p \geq Y(\rho) \tag{2.2}$$

$$\tau_f = \mu(p)p: \quad for \mu(p)p < Y(\rho) \tag{2.3}$$

Where $Y(\rho)$ is the effective shear stress function and $\mu(p)$ is the coefficient of friction as a function of normal pressure determined from conducting a shear test on a shear tester.

The model requires the value of nip angle, neutral angle, initial stress conditions and density. The values of nip angle and neutral angle were taken from experimental data, while the initial stress conditions and density were assumed. The calculation

process was repeated until the result was equivalent to the compacted strip density. However the compacted material density for each subsequent calculation step was determined from compression test data for a corresponding mean stress.

2.1.1.2   Neural Network Modeling

Neural network modeling has been used by Inghelbrecht *et al.* in 1997 to predict the quality of compacts and granules from roll compaction of drum dried waxy starch. The process parameters were velocity of rolls (RS), horizontal screw speed (HS), vertical screw speed (VS) and air pressure (P_{air}). The neural network modeling was based on 80 roll compaction experiments. Out of 80 data sets, 60 were selected to be the training set. The next 20 data sets were used to test the performance of the neural network model. The multilayer feed forward network structure consisted of four inputs (i.e. 4 process parameters), five nodes in the hidden layer (a Sigmoidal transfer function was used) and one output (i.e. friability with a linear transfer function). The sigmoid transfer function was used to allow the prediction of non-linear relationships in the data.   The neural net was trained using the backpropagation learning algorithm (i.e. supervised learning).   The results were then compared to the results of quadratic modeling. The quadratic modeling was conducted using a central composite index and it is described by the equation below.

$$Y = b_o + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} x_i x_j + \sum_{i=1}^{n} b_{ii} x_i^2$$

$$\left(j > i; i, j = 1,...n\right)$$

(2.4)

Where Y is the response (e.g. friability), b is the regression coefficients, x the factors (e.g. compactor parameters) and n is the number of factors (n=4). The neural network was found to model the friability better than quadratic modeling.

2.1.1.3   Discrete Element Method (DEM)

In 2001, Odagi *et al.* developed a 2-dimensional discrete element method (DEM) simulation method for the flow properties of a powder undergoing roller compaction. DEM follows the conventional DEM scheme proposed by Tsuji, i.e. Hertz theory controls the particle-particle normal contacts and Mindlin theory describes the particle interaction in tangential direction (Tsuji *et al.*, 1992). However, Odagi introduced an additional adhesive force (f_a=100m$_g$g) which considers the effect of powder cohesiveness. The adhesive force was obtained from experimental measurements on tensile strength of compressed powder strips. The DEM simulations were based on assumptions that the particles are spherical and mono-sized. It requires detailed particle properties (Table 2.1) and experimental conditions (Table 2.2, which were obtained from Michel *et al.*, 1993). The simulation of the dynamic flow of the roll compaction of powder was 2 seconds.

The simulations were conducted with and without the adhesive forces. In the absence of adhesive force and low roll speed, there was little effect on the particle motion, particle flow rate between roll gap, and the powder density distribution. Furthermore the powder material failed to be compressed. In the presence of adhesive forces and increasing roll speed the powder flowrate increases linearly.

The results from the DEM simulation compare qualitatively with the experimental results. The DEM simulation pressure distribution deviated from the experimental results at roll speeds above 5 rpm. This was assumed to be due to the spherical approximation of the particle shape and the air effect which was neglected in the DEM simulations.

Table 2.1 Calculation conditions

Number of particles		30000
Particle diameter		100 µm
Particle density		2000 kg/m^3
Young's modulus		1x10^5 Pa
Poisson's ratio		0.26
Poisson's ratio (w)		0.33
Friction coefficient (p-p, p-w)		0.6
Time Step	$f_a/(m_g g)=0$	1 x 10^{-5} s
	$f_a/(m_g g)=100$	5 x 10^{-6} s

Table 2.2 Experimental conditions used for the DEM simulation (Michel *et al.*, 1993)

Roll diameter	100 mm
Length of roll	46 mm
Gap width between two rolls	1.3 mm
Particle density	2500 kg/m^3
Particle mean diameter	30 µm
Effective internal angle of friction	20.2^o
Effective wall angle of friction	35°

2.1.1.4   Finite Element Method (FEM)

The finite element method (FEM) has been used to analyse the roll compaction of powder. Dec *et al*. (2003) developed a two-dimensional model for roll compaction process using a commercially available ABAQUS finite element code. This simulation was conducted in order to evaluate the effect of the frictional coefficient at the roll/powder interface and the feed stress on basic process variables. The basic process variables are roll force, roll torque, nip angle and neutral angle. The rollers were represented as rigid elements of 100 mm diameter. The powder material was represented by a material mesh comprising of an array of 80 x 12 plane-strain continuum elements with reduced integration (CPE4R). The roll gap was 2.0 mm and the powder material entry angle was approximately 18°. The constitutive model of the powder was based on a pressure-dependent yielding plasticity model (modified Drucker-Prager/Cap model) with linear elasticity. This rate-independent model was calibrated on a series of mechanical tests, which were diametrical compression, simple compression and compaction in an instrumented die. Based on shear testing measurements the internal frictional angle was estimated to be 65°. The friction for the roll/material was assumed to follow the Coulomb friction law with a constant frictional coefficient. The simulation was conducted until steady state conditions were achieved. The steady state conditions were based on constant values of the roll force and roll torque. The results from the simulations were:

- Confirmation of the two regions expected in the interaction between the powder material and the roll surface, i.e. of the slip in the feed zone and sticking in the nip region.

- At coefficient of frictions of 0.35 and 0.50, the nip angle was approximately 8.5^{o} and 12° respectively.

- Increasing the coefficient of friction for a given feed stress increased the maximum roll pressure.

- Agreements to the expected increase in roll force and roll torque with increasing feed stress and frictional coefficient.

- The ribbon density increases with an increase in frictional coefficient and feed stress.

The advantages of FEM over previously employed modelling methods are:

- Models can be tailor made for a specific powder material via a process of hypothesis, numerical testing and reformulation.

- The model can predict compact densities, material flow, deformation energy, shear stress (roll torque), pressure distribution (roll force), nip angle and neutral angle.

- Adaptation to include a feeding process and roll surface geometry could be conducted.

However the problems in implementing FEM are from preparation of input data. This is because there needs to be a more accurate material model to represent the range of densities of compaction and using the appropriate friction model to describe the phenomena on the powder material and roll surface interface.

2.1.1.5   Roll Compactor Simulator

Zinchuk *et al.* (2004) proposed a method for simulation of the roller compaction process along with the techniques for quantitative evaluation of its products. This model included three major components:

i.   A compaction simulator producing a simulation of the compression events during roller compaction.

ii.   A demonstration of characterisation of powders at different stages of densification using of the material relative density.

iii. Evaluation of equivalency between "simulated" and real compaction products using relative density and tensile strength.

The method was based on using a batch process (i.e. direct vertical compaction of powder) to represent the process of a continuous process (i.e. roll compaction) (Figure 2.2). The displacement, *D* of the Upper and Lower punch is converted into the rotational speed of a tangential point on the roll surface via a sine function (Equation (2.5)).

$$Displacement, D = R\sin(\omega t)$$
(2.5)

Where *D* is displacement, *R* is roll radius, ω is roll rotation, *t* is time.

Solid fraction and tensile strength were identified by Zinchuk *et al.* (2004), as key indicators of ribbon quality and were used in evaluation of the simulation. Their results showed good mechanical properties agreement for equivalent solid fraction of real and simulated microcrystalline cellulose ribbons. They also claimed that this

proved the ability of this method to help in scale up and enable formulation of tablet

dosage forms beginning from the initial drug development process.
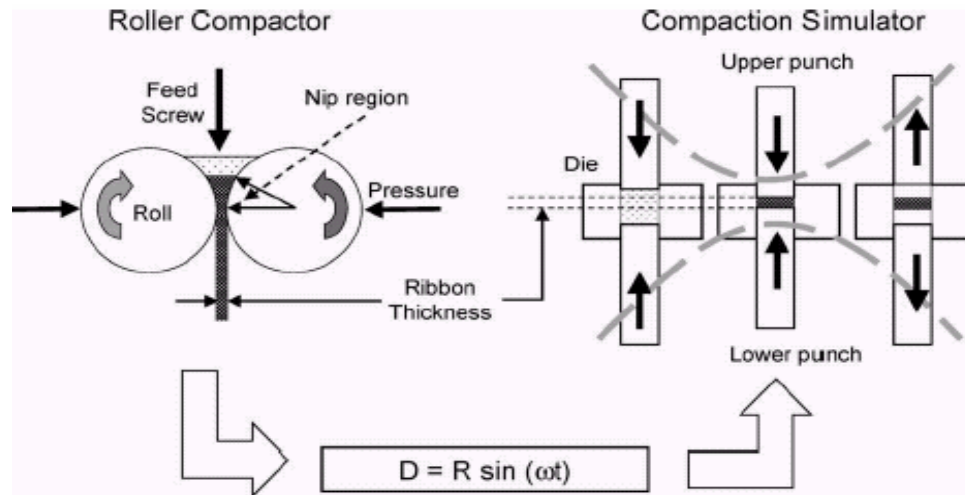
Figure 2.2 Schematic of the simulation of a roller compaction process using a compaction simulator
(Adapted from Zinchuk (2004)).

In 2001, Loginov *et al.* proposed a new roll briquetting simulator for understanding

the densification and performance of the roll compaction process. Furthermore, they

developed a mathematical model to relate the results obtained in the laboratory

scale to industrial scale.

2.1.2 Investigations on Roll Compaction

In this section, a summary of the previous research conducted on roll compaction is

presented in tabular form and divided into five major areas.

2.1.2.1   Previous roll compaction work on pharmaceutical excipients

Essentially there is a large amount of research conducted on the roll compaction of pharmaceutical excipients but it was rarely quantified (Table 2.3).   This section summarises the effects of binders on roll compacted material, comparison of fluidized bed granulation and roll compaction granulation and the effect of roll compaction on pharmaceutical excipients.   Parrot (1981) reported that roll compacted lactose and dibasic calcium phosphate granules did not improve granule mean particle size and bulk density.   Inghelbrecht and Remon (1998) conducted work on lactose of varying particle size in addition to manipulating the roll compaction process parameters.   They found that roll compaction of spray dried lactose was difficult because lactose was free flowing.   In addition to that they also discovered that the range of roll compaction parameter settings was influenced by the type, particle size, particle size distribution and bulk density of the powder material but the effects were not modeled.

2.1.2.2   Investigations into the effect of roll compaction process parameters

Table 2.4 shows a summary the work done on varying the roll compaction process parameters.   It covers research which varied the roll speed, roll gaps, feed pressure, vacuum feeding and cycles of roll compaction passes.

2.1.2.3   Validation of theoretical models

Table 2.5 shows a list of the validation conducted previously on Johanson's Theory and a force balance developed to relate shear to pressure. The Johanson's model was shown to be able to predict experimental results well by Hubert *et al.* (2000) and Yusof *et al.* (2005). While, Schonert and Sanders (2002) showed that their force balance exhibited two regions whereby, firstly the calculated shear value agrees very well with measured shear and secondly, the calculated shear value deviates remarkably with measured shear. This was explained as a result of material slip near the relaxation zone. It is interesting to note that they indicated that the increase in maximum pressure just above the gap neck, although no slipping occurs at this point, might be an indirect effect of the material slip near the relaxation zone.

2.1.2.4   Novel roll compaction methods and innovations

Table 2.6 shows a list of the novel roll compactions methods, such as using acoustic emission, a method to assess the feasibility of powder material for roll compaction and using ultrasound to enhance the performance of a roll compactor.

2.1.2.5   Previous roll compaction research conducted at The University of Birmingham

Table 2.7 shows the summary of previous roll compaction research conducted at the University of Birmingham. Michel (1994) highlighted the existence of neutral angle,

which was not accounted for in Johanson's model. He also showed that the throughput and the roll speed relationship stopped being linear at higher speeds due to deaeration problems in the feed region. This information was useful in adapting the model for comparison in this study. In 2001, Boursel proposed a derivation to include air entrainment effect into the Johanson's Theory. It was decided that for an initial comparison of the predictability of the software it would be wise to use a more established model like the original Johanson's model. Boursel also reported that roll compaction of varying particle sizes, at constant roll gap, varies the roll speed range greatly. This was noted and used as a guide for this research.

Perera (2004) found that higher nip angle resulted in longer compaction time and lubrication had an adverse effect on nip angle. He explained that nip angle was found to increase if the effective angle of internal friction, cohesion of powder was increased, the powder moisture content increased or the roll gap and the roll speed were decreased. It is interesting to observe the ability of the intelligent software model to predict the combined effect of the effective angle of wall friction, roll gap and roll speed. Bindhumadhavan (2004) reported on his validation of Johanson's model for the roller compaction of powders. He found that the peak pressure was predicted accurately by the model and that at lower roll speeds, the influence of roll speed on peak pressure was accurately predicted by the model. This was to be accounted for in the use of Johanson's model for comparison with the intelligent software prediction.

Table 2.3 Previous work on effects of roll compaction on pharmaceutical excipients

Investigators	Pharmaceutical Excipients/roll compactor	Objective	Finding
Jaminet and Hess, 1966	Lactose starch mixture/ Hutt	Investigated the effect of different binders on the properties of briquettes, granules and tablets.	• The addition of binders to lactose starch mixture resulted in the significant decrease of non compacted material. • The addition of ethylcellulose increased the briquette strength, however the addition of carbowax 4000 had the opposite effect. (The briquette strength was tested using a bending method). • The addition of a small amount of water decreased the amount of non compacted material and weakened the briquettes. • The roll compaction process parameter and briquette strength strongly determined the particle size distribution of the granules. • Proposed a starting material for roll compaction which contained corn starch (30%), lactose (63%), talc (2%), macrogol 4000 (1%) and MCC (4%).
Parrott, 1981	Lactose, dibasic calcium phosphate, magnesium carbonate, calcium carbonate/ Freund or vector - concavo-convex roll compactor (rim length 7mm and angle of 65°)	Investigations on the effects of roll compaction using a concavo-convex roll compactor on pharmaceutical powders.	• All the materials had an increase in granule mean particle size and bulk density except for lactose and dibasic calcium phosphate. Might be due to the brittle nature of the flakes which may have fractured in the oscillating granulator. This produced a smaller median diameter. • The roll compaction had a negative influence on flowability.
Li and Peck, 1990	Maltodextrin/ Fitzpatrick	Compare agglomeration by fluidized bed granulation and roll compaction dry granulation	• The fluidized bed granulation produced highly porous granules of low bulk density, whereas roll compaction produced granules with a significantly low degree of granular porosity with high bulk density. • The roll compacted granules showed a better flowability. In terms of gravimetric and volumetric flowrate. • Granules produced from roll compaction indicated a higher resistance to deformation.

Inghelbrecht and Remon, 1998	Four types of lactose of varying particle size/ Fitzpatrick L83 Chilsonator	Pressure, roll speed, vertical and horizontal screw speed	• Pressure had the most influence on the granule properties, followed by roll speed and then horizontal screw speed. Pressure refers to the air pressure used in controlling the hydraulic pressure between the rolls. • The type, particle size and density had an influence on the range of roll compaction parameter settings. • Spray dried lactose was reported to be difficult to handle. • Best compact quality was produced at high pressure and low screw speed.

Table 2.4 Previous work on effect of process parameters and powder properties on roll compaction

Investigators	Roll compaction equipment/protocol	Investigated roll compaction process parameters	Finding
Cohn *et al.*, 1966	Fitzpatrick / Roll compaction of Potassium chloride to investigate the effect of oil pressure on the amperage of roll compaction, amount of non compacted material and the hardness of the tablet.	oil pressure	• Leakage of powders caused large roll compaction problems. • No simple relationship existed between the input variables and the output variables, but the optimal settings of the output variables were found.
Spinov and Vinogradov, 1967	Roll compaction on a copper powder with a vacuum deaeration facility	Roll speed constant at 17.5 m/min, roll diameter of 120mm, material was fed using a hopper of length 80mm and inclination of 48°	• Vacuum deaeration during roll compaction minimized the pores and improved the quality of the compacted strips. • The bulk density and porosity of compacted strips were constant during vacuum deaeration roll compaction.
Funakoshi *et al.*, *1977*	Freund or vector - roll compactor designed as a concavo-convex pair/ roll compaction of lactose, and mixtures with small amount of riboflavin	Studied the factors affecting the compacting pressure distribution.	• The concavity of the roll rims were varied between 45, 65, 75 and 90°. This concavity adjusted the uniformity of the pressure over the whole width of the rolls. The best result was obtained at 65°.
Petit-Renaud *et al.*, 1998	Specially instrumented roll press by K.R. Komarek (B-100 QC)	Roll speed, screw speed and roll gap	• Increase in roll speed resulted in a linear increase of the mass throughput at constant roll gap and screw speed until a critical roll speed at which the mass throughput decreased. • Increase in roll speed resulted in a decrease of the normal pressure and nip angle.

Hirohata *et al.*, 2001	Roll compaction on electrolytic copper powder	Differential speed rolling (roll speed ratio 1.00, 1.20 and 1.33). Roll gap (0.05, 0.10 and 0.15 mm)	• The lower speed roller experiences a forward slip which is affected by the increase in roll speed ratio. While the higher speed roller experiences a decrease in roll speed at roll speed ratio of 1.33. • The rolling load becomes smaller with increasing roll speed ratio under constant compacting speed ratio. The rolling load is measured with a load cell of strain gage type mounted on a bearing supporting the rolls. The compacting speed ratio is the powder volume required to compact the powder per unit length and width. • The rolling load increases linearly with increasing compacting speed ratio. • Relative density increased with roll speed ratio for the same rolling load. • Compacted strip thickness for each initial roll gap becomes larger with increasing rolling load at constant roll speed ratio. • At decreasing rolling load the relative density variation of compacted strips rolled at different roll speed ratios decreased.
Bultmann, 2002	Gerteis 3 W-Polygram / roll compaction of MCC	Multiple compaction passes (up to 10 passes)	• Reduction of amount of fines with multiple compactions. • After roll compaction, powder flow properties improved and the mean granule size was increased. • Tablet tensile strength decreased after the first two compression passes.
Simon and Guigon, 2003	Roll compaction of lactose, alumina and sodium chloride on specially instrumented roll press by K.R. Komarek (B-100 QC)	Feeding conditions of single screw feed	• The local pressure in the feeding zone varies with time with the same period as the screw feeder rotation. • The compacted strip exhibits variation of density and strength according to the fluctuations of the screw feeding rate. • The powder-packing properties were also found to affect the heterogeneity of the compacted strip.
Lecompte *et al.*, 2005	Roll compaction of organic powder on a specially designed instrumented roll compactor	Roll speed, roll gap, press strength, rotation angle and the feed strength	• The pressure during roll compaction was constant and symmetrical because as opposed to Simon and Guigon's (2003) findings the screw feeder end was located further from the rolls. • Feeding-roll ratios below 1 result in the powder compaction mainly occurring in the centre of the roll width. • At high feeding-rolling ratio, homogenous pressure along the roll width is achieved. • Increasing feeding-rolling ratio results in a uniform pressure distribution and better compact quality.

Table 2.5 Validation of roll compaction theoretical models

Investigators	Validation work	Finding
Hubert *et al.*, 2000	Validated the Compressibility equation used in Johanson's theory using a die-press and roll press.	• Proposed a new set of equations derived from using material properties based on the modulus of volume transformation. • Johanson's theory was found to fit well with experimental results.
Schonert and Sander, 2002	A sensor was used to measure the pressure and shear forces in the roll compaction process of a particle bed. Based on a force balance, an equation was derived to relate shear to pressure. Assuming the proportionality between the transversal pressure in the bed and the normal pressure on the roller.	• The predicted and experimental results showed that there are two regions: a) the calculated and measured shear agrees very well b) the calculated and measured shear deviates remarkably. • The above finding was explained as due to material slip. Material slip only exists near the outlet in the relaxation zone, and no slip occurs around the shortest distance between the rollers (i.e. gap neck). • The material slip in the last part of the relaxation zone causes a tensile force on the material above the gap neck and increases due to the shear on the rollers. This tensile force was supposed and included in the force balance. • This tensile force could explain the increase in maximum pressure just above the gap neck, although no slipping occurs at this point.
Yusof *et al.*, 2005	A smooth roll compactor (diameter 0.08m and width 0.20m) was used to roller compact maize powder to validate the roll compaction parameters (roll force and roll torque) predicted by Johanson's Theory.	• The model predicted the experimental results (roll force and roll torque) reasonably well, despite being sensitive to the initial bulk porosity variations. • The roll power decreased as the roll gap and the roll speed decreased. • Throughput increased linearly with the increase in roll gap and roll speed. • The pressing time increased as the roll gap increased and the roll speed decreased. • Changes in the feed powder amount and friction ratio had little or no effect on the calculated roll power, throughput and pressing time.

Table 2.6 Novel roll compaction methods and innovations

Investigators	Novel Method	Findings
Hakanen *et al.*, 1993	Acoustic emission was used to characterise the roll compaction process.	• Over-compaction of microcrystalline cellulose could be detected using this method.
Salonen *et al.*, 1997	Acoustic relaxation emissions (ARE) from roll compaction of microcrystalline cellulose (MCC) and maize starch (MS) were detected using a microphone.	• ARE signal from compacting powder consists of short acoustic pulses of varying intensities. • Intensity of ARE was found to increase when applied compressive force was increased for roll compaction of both MCC and MS. • They suggested a possible relationship between the ARE and Young's modulus. • They also suggested the possible relationship of ARE and cohesive energy density of the powder, which could explain the parallels between ARE and Young's modulus (further work required).
Gereg and Cappola, 2002	A feasibility study was developed to identify powder material which could be roll compacted. It was conducted through a series of powder material characterisation, lab-scale hydraulic pressing of tablet compacts and industrial scale roll compaction to identify process parameters suitable to achieve the appropriate granule size and density.	• The material characteristics found for regular-grade and spray-dried Lactose were microscopy, bulk and tap density, Carr's Index, angle of repose, flow rate and sieve analysis. • Hydraulic pressing of tablet compacts and roll compaction produced a final tablet with similar density and hardness. Hence this suggested that the method developed was appropriate to determine the suitability of powders for roll compaction.
Gaete-Garreton *et al.*, 2003	Investigated the results of using ultrasound to enhance the performance of a roller mill.	• The use of ultrasonic energy resulted in a reduction in the total power required to operate the grinding in the roller mill. • The shaft torque required was found to be reduced with the presence of ultrasonic field. • Less abrasive wear on the grinding surface was observed with the use of ultrasonic field.

Table 2.7 Investigations conducted in the University of Birmingham.

Investigators	Roll compactor details	Material	Objectives	Findings
Michel, 1994	Specially made instrumented roll compactor with one pressure sensor on the surface of each of the two rollers. Gravity feeding Roller diameter: 100 mm Roller width: 46 mm	SH alumina (SH100 and SH500)	• To investigate the important parameters of roll compaction	• Compressibility; κ_{SH100} = 8.5 and κ_{SH500} = 12.5 • Effective angle of internal friction (^{o}); SH100 = 35, SH500 = 39 • Cohesion (kPa); SH100 = ~8, SH500 = ~0 • Angle of wall friction; SH100 = 24, SH500 = 22.5 • Very strong influence of the roll speed on the roll compactor performance and compact properties. The throughput and roll speed relationship stops being linear at higher speeds due to aeration in the feed region. • Showed that the pressure required to obtain a ribbon compact of given density can be estimated from uniaxial compaction tests using pressure-density equations. • Developed a model to relate the throughput to the neutral angle. The neutral angle was found to decrease with increasing roll speed and was independent of the roll gap. • At the same pressure at slip/nip transition, equivalent pressure profiles and resulting densities can be found at varying speeds. • The neutral angle decreased with increasing roll speed and independent of roll gap. • Pressure at the centre of the roller width was found to be higher than the pressure at the ends of the roller width.

Bourseul, 2001	Specially made instrumented roll compactor with one pressure sensor on one roll surface. Gravity feeding Roller diameter: 200 mm Roller width: 46 mm	SH alumina (SH100, SH150, SH300 and SH500)	• To gain a theoretical and experimental understanding of the physical phenomena controlling the operation of a roll press and the formation of agglomerates. • To establish predictive models using the intrinsic characteristics of the powder and the press. • To characterise the specificity of roll compaction as a forming process through characterisation of the compacts	• In roll compaction of varying particle sizes, at constant roll gap, the roll speed range varies greatly. • Investigation on 3 types of rollers (200 mm diameter smooth roller, 100 mm diameter smooth and rough rollers) and one type of powder material resulted in the results following a power law: $P_m/D = A.e^{-B.h.\omega}$, where P_m is the maximum pressure reached in the roll compactor, D the roll diameter, h the roll gap, ω the angular speed and A and B are constants. • Proposed a derivation to include air entrainment effect into the Johanson's Theory. The model predicted the influence of roll speed and roll gap correctly provided the appropriate permeability factor is used and above h/D>0.01. However the model overestimates the effect of h/D at h/D<0.01. • Attrition results indicated that roll compacted ribbon compacts were more resilient than tablet compacts produced from uniaxial compaction. • Ribbon compacts of different geometrical density appeared to display the same pore size distribution, as opposed to tablet compacts which only achieved the same mean pore size at very high compaction pressures. • This study showed that the way the compact density is achieved (i.e. the history of the particle bonds) outweighs the macroscopic end result (i.e. the final compact density).

| Perera, 2004 | Instrumented roll compactor with one pressure sensor on one roll surface.

Gravity feeding

Roller diameter: 200 mm

Roller width: 46 mm | Microcrystalline cellulose (Avicel: PH101, PH102, and PH105)

Lactose (regular and free flow) | • To understand and estimate the roll compaction properties of a given powder physical characteristics.

• To investigate the analogy between uniaxial compaction and roll compaction process | • The Kawakita correlation (empirical correlations for pressure-density relationships) showed the best fit to experimental data in both uniaxial compaction and roll pressing compared to Heckel and Cooper-Eaton equations.

• The results from the Kawakita correlation for uniaxial compaction and roll compaction showed that they corresponded well with each other. Hence it was proposed that uniaxial compaction could be used as a prediction for the performance of a given powder in roll compaction operations.

• Young's modulus of the ribbon compact was independent of particle size and inversely proportional to the roll gap, roll speed and lubricant content.

• Higher nip angle resulted in longer compaction time. Lubrication had an adverse effect on nip angle. Nip angle was found to increase if:

    1. the effective angle of internal friction and cohesion of powder was increased
    2. the powder moisture content increased
    3. the roll gap and the roll speed were decreased.

• Positron Emission Particle Tracking (PEPT) results showed that the existence of two different velocity regions in the roll compaction process. The point at which the particle velocity changed is assumed to be the nip angle.

• The PEPT nip angle value was compared to nip angle predicted from Johanson's theory. It was found that Johanson's theory overestimated the nip angle. |

30

Bindumadhavan, 2004	Instrumented roll compactor with one pressure sensor on one roll surface.
	
	Gravity feeding
	
	Roller diameter: 200 mm
	
	Roller width: 46 mm
		Microcrystalline cellulose (Avicel PH102)
		
		α-Lactose monohydrate
		
		Carbomers
			To understand the relationship between powder properties, roll compactor dimension, process parameters and product characteristics.	• An increase in fines content of powder material increased the peak pressure applied, compact bulk density and compact strength.
				
				• The peak pressure and nip angle increase with an increase in inter-particle friction and wall friction.
				
				• The performance of the roll compactor was found to depend strongly on the composition of the powder mixture.
				
				• An increase in the roll gap and roll speed decreases the compaction pressure.
				
				• PEPT results showed that nip angle could be estimated from particle velocity.
				
				• The PEPT estimated nip angle was comparable with the nip angle values from the pressure profiles. The PEPT result confirmed the assumptions of the Johanson's model, that there exist two different particle velocity zones.
				
				• X-ray microtomography results showed that the compacts were denser in the middle part compared to the sides of the ribbon compact. Indicating a non-uniform pressure distribution across the roll width.
				
				• Validation of Johanson's model for the roller compaction of powders showed that:
				
				   1. The measured roll pressures increased more rapidly than the predicted values as the powder passes through the nip region. The difference increases with decreasing minimum gap.
				
				   2. Peak pressure was predicted accurately by the model.
				
				   3. The roll gap size had a significant effect on the nip angles. The slip nip intersection showed that the nip angle was independent of roll gap, however the predicted pressure gradient showed that nip angle was comparable with experimental results.
				
				   4. At lower roll speeds, the influence of roll speed on peak pressure was accurately predicted by the model.

2.2 Intelligent Software

It has long been realised that the relationship between the pharmaceutical excipients, roll compaction process parameters and ribbon properties is non-linear. As Rowe and Roberts (1998) pointed out, the formulator must take into account the properties of the active ingredient as well as possible chemical interactions between it and the other ingredients. These are added to improve processibility and product properties, which may not necessarily contribute to product chemical stability. Furthermore there may be interactions between the added ingredients leading to physical instability. Other factors such as cost, market demand and the requirements of the regulatory bodies (FDA[*], MHRA[†] etc) are important influences on the ingredients and process used in the product formulation. Although models and simulations may be available, in many cases the formulation process has to be carried out in a design space that is multi-dimensional in nature and difficult to conceptualise (Rowe and Roberts, 1998). Thus there are many factors as summarised in Figure 2.3 which must be considered by the formulator.

The complexity of formulation design and the formulation process is a highly specialized task, requiring specific knowledge and often years of experience. To retain in-house expertise which may be lost as a result of employee movement,

---

[*] Food and Drug Administration

[†] Medicines and Healthcare Regulatory Authority

computer technology is used to harness this knowledge. Therefore in the last two decades the pharmaceutical industry has moved towards using artificial intelligence technology to predict these complex relationships and document it in a form available to all.
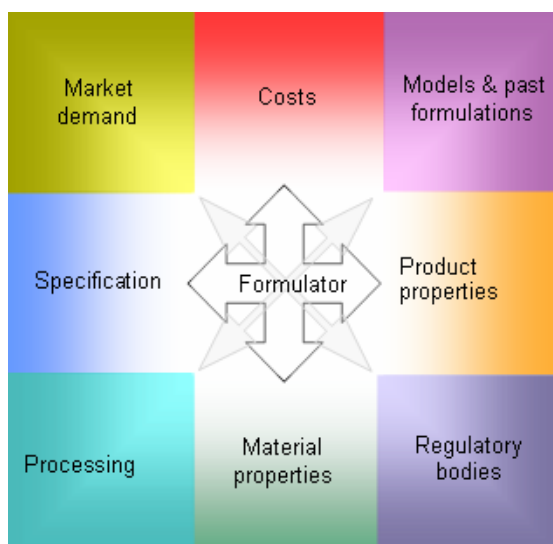
Figure 2.3 Complexity of formulation process. Adapted from Roberts and Rowe (1998).

In this research, intelligent software used were based on Artificial Neural Networks (ANNs), Genetic Algorithm (GA), fuzzy logic and neurofuzzy logic (see sections 2.2.1 to 2.2.4 for a description of each of these systems).   The advantages of this software include:

- ♦ The softwares were designed especially for pharmaceutical applications and personnel expertise in the individual systems programming is not required to perform any modelling work.

- ♦ The user can pick the software up and use the software to perform modelling on a database of input/output information. This gives the researcher ample time to collect information to build a database of information.

This research is different from Inghelbrecth *et al*. (1997) and Turkoglu *et al*. (1999) because here the inputs into the ANNs and genetic algorithm (GA) models include powder characteristics in addition to roll compaction process parameters. The following section gives a background of the computer programs applied in the intelligent software mentioned above.

2.2.1 Artificial Neural Networks (ANNs)

ANNs are biologically inspired computer programs designed to simulate the way in which the human brain processes information. These systems gather information by detecting the patterns and relationships in data and they learn (or are trained) through experience, not from programming (Agatonovic-Kustrin and Beresford, 2000). An ANN is formed from hundreds of single units, known as artificial neurons or processing elements, which are connected with coefficients (weights). These constitute the neural structure and are organised in layers. The layers consist of an input layer, one or more hidden layers and an output layer. In short, ANNs are mathematical systems that mimic the way in which the human brain processes information (Erb, 1993).

The first mathematical model of a biological neuron was presented by McCulloch and Pitts (1943). They suggested the unification of neuro-physiology with mathematical logic which paved the way for significant results in ANNs research. Then in 1949, Hebb proposed a learning rule derived from a model based on synaptic connections between nerve cells responsible for biological associative memory. In 1958,

Rosenblatt refined the rule developed originally by Hebb in the perceptron model (i.e. a simple model of a neuron). Then Willshaw and Von der Malsburg proposed the self organising map architecture based on competitive learning in 1976. A comprehensive review of the most significant steps in ANNs research can be found in Rajasekaran and Pai (2003).

The power of neural computations comes from connecting neurons in a network. Each processing element has weighted inputs, a transfer function and one output. The behaviour of a neural network is determined by three factors:

  i. the transfer functions of its neurons,
 ii. the architecture itself,
iii. the learning algorithm (rules).

An artificial neuron is also known as a processing element. A processing element has inputs, a transfer function and one output. It is essentially an equation which balances inputs and outputs. An artificial neuron has connection weights which represent the memory of the system. The weights, w_i are the adjustable parameters and in that sense a neural network is a parameterised system. The weighed sum of the inputs constitutes the activation of the neuron. The activation signal is passed through the transfer function to produce a single output of the neuron between zero to one. The type of output greatly depends on the type of transfer function used. In the example, the transfer function is sigmoidal. During

training, the weight connections are adjusted until the error in predictions is minimised and the network reaches the specified level of accuracy. Once the network is trained and tested it can be given new input information to predict the output.
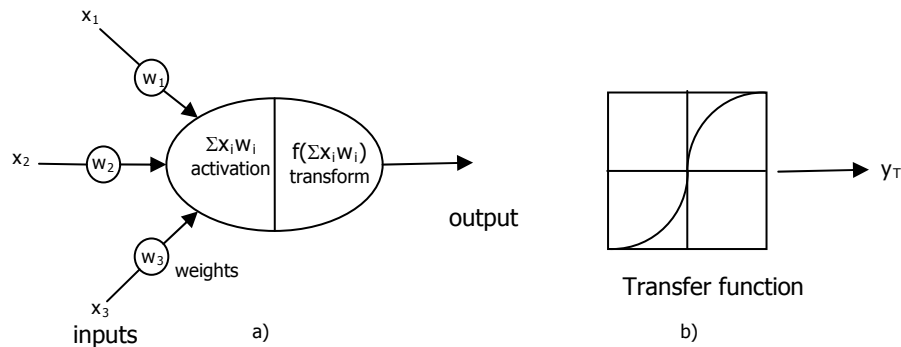
If inputs are $x_1=1$, $x_2=2$, $x_3=3$ and

Weights are $w_1=0.4$, $w_2=0.1$, $w_3=0.2$

Then Summations $S = \sum_{i}^{n} x_i w_i$

$S = 1(0.4) + 2(0.1) + 3(0.2) = 1.2$

If sigmoid transfer function is used,

i.e. $y_t = 1/(1+e^{-aS})$

since $a = 1$, then $y_t = 1/(1+e^{-1.2}) = 0.77$

Figure 2.4 Model of an Artificial Neuron processing a numeric data (Rowe and Roberts, 1998)

The four basic types of transfer functions are the thresholding function, piecewise-linear function, sigmoidal function and hyperbolic tangent function (Haykin, 1994; Rajasekaran and Pai, 2003). Table 2.8 illustrates the transfer functions and the types of output each of them gives. The thresholding function is the first and simplest form of transfer function developed by McCulloch and Pitts (1943). This transfer functions forms the neuron which is referred to in literature as the *McCulloch-Pitts model*. The piecewise-linear function can be viewed as a linear

combiner[‡] or as a threshold function. The most common form of activation function is the sigmoid function (Haykin, 1994). The slope parameter allows the sigmoid function to have varying slopes. It assumes values in the range of 0 to 1. According to Haykin (1994), the important feature of the neural network theory is that it is differentiable. Lastly, hyperbolic tangent function is a very flexible, non-linear, continuous and differentiable function (Bhadeshia, 1999). Chapter 7 shows that the transfer function is an important feature in the training of ANN models.
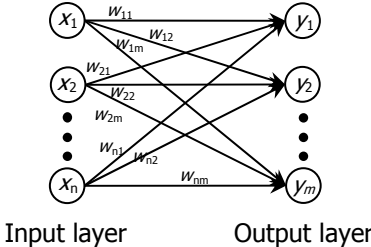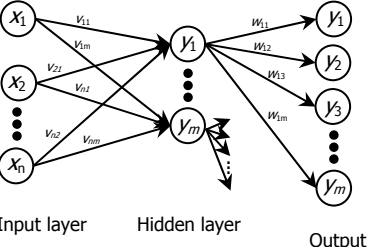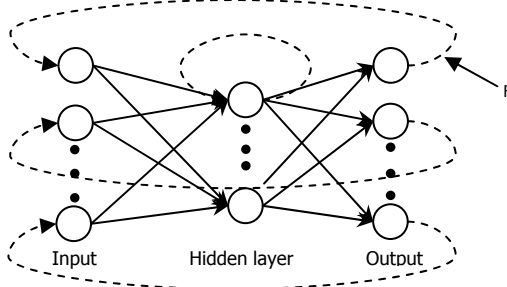
The way that the neurons are connected to each other has a significant impact on the network. Just like 'real' neurons, artificial neurons can receive either excitatory or inhibitory inputs. Excitatory inputs cause the summing mechanism of the next neuron to add while the inhibitory inputs cause it to subtract (Agatonovic-Kustrin and Beresford, 2000). A neuron can also inhibit other neurons in the same layer. This is called lateral inhibition. The network wants to 'choose' the highest probability and inhibit all others. This concept is also called competition. For instance in Multilayer Feedforward Network the number of nodes (neuron cells) in the hidden layer can affect the results of the model training. Having fewer nodes sometimes might be better than having more nodes, as shown in Chapter 7.

[‡] Linear combiner is also known as an adder for summing the input signals which is weighted by the respective synapses of the neuron.

Table 2.8 Types of transfer functions

Transfer function	Illustration	Output signal, y_t
Thresholding Function	y_t Output 1 0 θ Input	0 or 1
Piecewise-Linear Function	y_t Output 1 0 θ Input	$y_t = \begin{cases} 1, & \theta \geq \dfrac{1}{2} \\ y, & -\dfrac{1}{2} > \theta > \dfrac{1}{2} \\ 0, & \theta > -\dfrac{1}{2} \end{cases}$
Sigmoidal Function	y_t Output 0 Input	$y_t = 1/(1+e^{-aS})$ a = slope parameter $S = \displaystyle\sum_{i}^{n} x_i w_i$
Hyperbolic Tangent Function	y_t Output 0 Input The strength of each graph depends on the weights	$h = \tanh\left(\displaystyle\sum_{i} w_j^{(1)} x_j + \theta \right)$ *with* $y = w^{(2)} h + \theta^{(2)}$ Where w is weights, x is inputs and θ is another constant. (Bhadeshia, 1999)

Table 2.9 Three fundamentally different classes of Network Architecture. (Rajasekaran and Pai, 2003)

Type of NN architecture	Description	Illustration
Single layer feedforward network	It has two layers: input and output layer. Input layer transmits the signals to the output layer. The output layer performs the computation.	x_i : Input neurons y_i : Output neurons w_{ij} : Weights
Multilayer feedforward network (also known as multilayer perceptron, MLP)	It has three layers: input, hidden layer and output layer. The input layer is linked to the hidden layer. The hidden layer aids in performing useful intermediary computation directing the input to the output layer. The output layer performs computation. There could be more than one hidden layer.	x_i : Input neurons y_i : Hidden neurons z_i : Output neurons v_i : Input hidden layer weights w_{ij} : Output hidden layer weights
Recurrent neural network	Similar to the feedforward network above except that there is at least one feedback loop. A feedback loop can exist between the output layer and the input layer or neurons with its output fed back to itself as an input.	
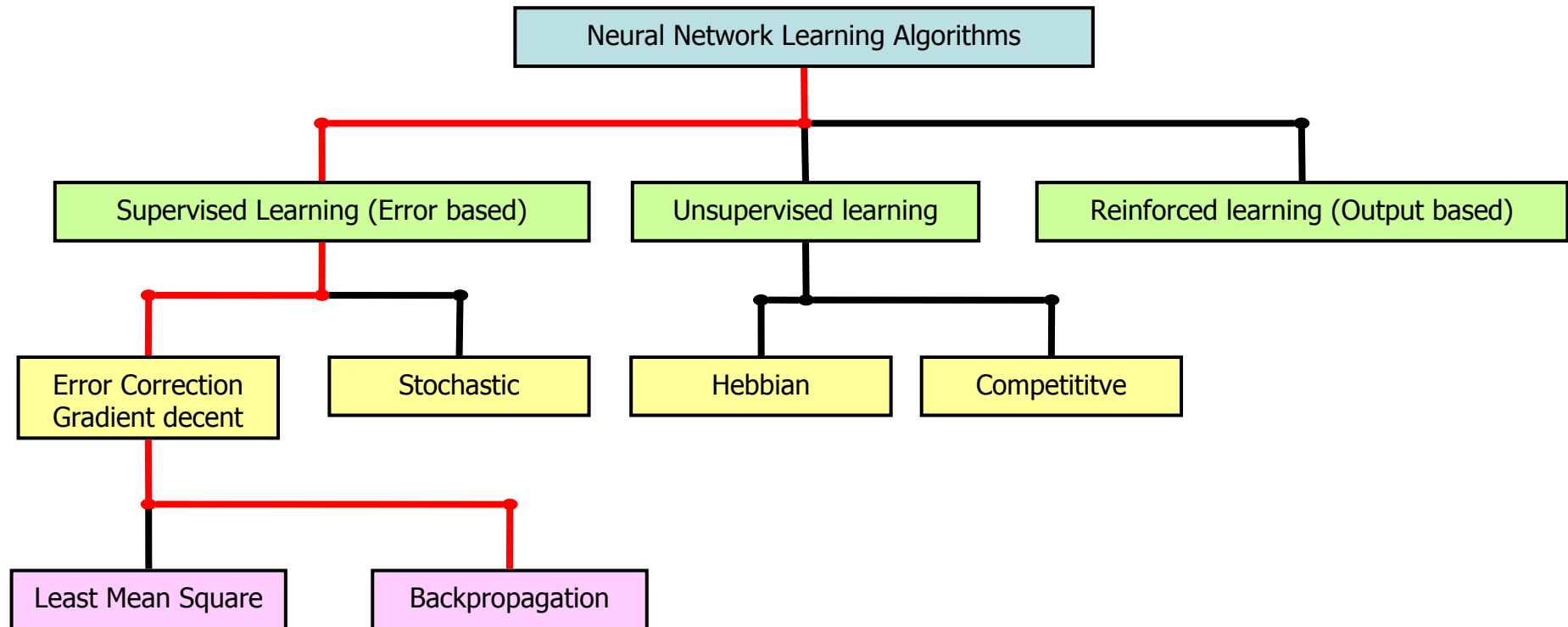
Figure 2.5 Classification of learning algorithm (Rajasekaran and Pai, 2003). The red linking lines show the classification of the learning algorithm within the intelligent software used within for this research.

Table 2.9 shows three fundamentally different classes of ANN. In the Single Layer Feedforward Network, the network links an output layer with an input layer, and information is stored in the network within the modifications made to the weight coefficients of the network. The Multilayer Feedforward Network is similar to the Single Layer Feedforward Network except for the added hidden layer. The hidden layer enables the network to acquire a global perspective despite its local connectivity. This is because the hidden layers give the network an extra set of synaptic connections and extra dimension of neural interactions (Churchland and Sejnowski, 1992).

The Recurrent Neural Network is different from the Feedforward Network because it would have at least one feedback loop. A neuron cell can have its output fed back into its input or the "output layer" output is fed back into the "input layer" input. This feedback loop has a large impact on the learning capability of the network and on its performance. The intelligent software used in this research applies the Multilayer Feedforward Network architecture.

The architecture of the ANN is very closely linked to the learning algorithm used to train the network. Figure 2.5 shows the classifications of learning algorithms and the red connecting lines show the classification of the learning algorithm used in this work. The figure shows a broad classification of ANNs into three basic types: supervised, unsupervised and reinforced. Supervised learning is a form of regression that relies on example pairs of data: inputs and outputs of the training

set. Comparison is made between the network's computed output and the correct expected output to determine the error. The error is then used to change the network parameters to improve model performance. Unsupervised learning means that the target output is not presented to the network. Thus the system learns on its own by discovering and adapting to structural features in the input patterns. In the reinforced learning, the system is presented with an indication of the correctness or incorrectness of the computed output. The information guides the networks learning process, whereby a reward is presented for a computed correct answer and a penalty is given for a computed incorrect answer. The explanation on these learning algorithms can be found in Haykin (1994) or Rajasekaran and Pai (2003). The backpropagation learning algorithm is explained in the subsection 2.2.1.1. It is a supervised learning algorithm which works on the principle of error correction gradient descent (see Section 9.5.3.5 in Appendix 5).

ANNs are a digitised model of a human brain. It is able to perform pattern recognition skills just like a brain, but its skill is limited because it is not as complex as the brain. ANNs rarely have more than a few hundred or a few thousands processing elements, while the human brain has ~100 billion neurons. The advantages of ANNs are (Rowe and Roberts, 1998):

i. ANNs are diverse as they are able to deal with complex, real world applications where data is fuzzy and non-linear.

ii. They are able to learn new relationships within the input data.

iii. Once trained a neural network can deal with unseen data and generate correct responses.

iv. ANNs are able to extract information from incomplete data or noisy data. Because they are statistical systems they are able to recognise underlying noise.

v. They are robust since there are many processing neurons in an ANN and damage to a few does not bring the system to a halt.

vi. Although training a neural network can be relatively slow and demanding of computer power, once trained ANNs are inherently fast.

vii. ANNs can adapt to new and changing environments and are easy to maintain.

2.2.1.1   Backpropagation

Backpropagation is a feedforward ANNs training strategy. It consists of an iterative optimisation of the error function. This error function represents a measure of the performance of the network (Haykin, 1994; Rumelhart *et al.*, 1986). This error function E is defined as the mean square sum of differences between the values of the output units of the networks and the desired target values (see equation (2.6))

$$E = \frac{1}{2} \sum_{p=1}^{P} \sum_{j=1}^{N_L} (t_j - a_j)^2$$

(2.6)

Where t_j and a_j are the target and actual response values of output neuron j, N_L is the number of output neurons and L is the number of layers. The error function depends on the size of the pattern set and the number of output neurons of the

specific network used and simulation performed. To allow consistent comparisons, some ANNs simulators divide the summation of the equation (2.6) by $N_L \cdot P$.

A pattern set is the input and corresponding target output. In a simulation the pattern error is calculated sequentially in an iterative manner. The weight corrections are performed during this training process to adapt the network to the desired behaviour. The weight corrections depend on the type of backpropagation strategies used (see section 9.5.3.5). The iteration continues until the weight values allow the network to represent the required relationships between input and output. Each presentation of the overall pattern set is named an epoch.

The minimisation of the error function depends on the strategies employed to manipulate the weight corrections. Five strategies which were used in the *INForm* software are presented in section 9.5.3.5.

2.2.2 Genetic Algorithms

Turban (1995) defines the genetic algorithm (GA) as a software program that learns from experience in a similar (simplified) manner to the way in which biological systems learn. Basically the GA is loosely based on the biological principles of genetic variation and natural selection, mimicking the principal ideas of evolution over many generations. It works with a population of individuals, each of which is a candidate solution to the problem. These individuals then reproduce through mating/mutation, all the time evolving new solutions to the problem. Eventually

after several generations an optimum solution will be found (Rowe and Roberts, 1998). Comparatively, the biological genetic evolution consist of an individual (or a population) of a biological species.   Whereas, the GA of an individual (or a population), is in the form of a string or ordered sequence of numbers representing a numerical solution and the mechanism of operation is based on logic and mathematics.

The use of evolutionary principles to compute problems and produce solutions was initiated by Friedberg in the 1950s. Then the methods that are still used today were developed in the 1960s.   The first two important methods were Evolutionary Programming and Evolutionary Strategies.   In 1965, Fogel invented Evolutionary Programming in which finite state automata are evolved. Then in 1965 and 1970, Schwefel and Rechenberg developed Evolutionary Strategies which are numerical optimisations on the basis of evolutionary principles. Next in 1965, Holland invented GA in which there is a representation of the solution which is varied and mapped onto the real solution by some function. Following that, the real solution is then tested on the problem to get the fitness of the individual. Further reading on the history of GA can be found in Rajasekaran and Pai (2003).

The GA software assists the neural network software by fine tuning (optimising) the model which has been developed.   The software can only perform its intended function by possessing the following features (Rowe and Roberts, 1998):

i. Fitness function – a numerical description to differentiate between a good or bad solution.

ii. Logical selection method – a logical method of selecting individual solutions to become parents of the next generation of solutions.

iii. Crossover and mutation operators – a logical mixing method analogous to the mixing of genes that accompanies reproduction and mutation chromosomes.

A diagrammatic form of the GA cycle is shown in Figure 2.6. The advantages of applying GA as an optimisation technique are listed below:

i. It is a stochastic algorithm, i.e. it relies on random elements in parts of its operation rather than being determined by specific rules.

ii. It is not susceptible to the initial starting point.

iii. It has the ability to find global maxima/minima.

iv. It is rapid and efficient.

v. The technique is effective for optimisation.

The GA is used in the optimisation section of the *INForm* software (Chapter 7). The optimisation can terminate under two conditions:

i. Once the specified number of iterations has been reached, regardless of whether the solutions actually meet the optimisation criteria the optimisations will stop. The final solutions will be the ones which best meet the criteria given within the constraints of the optimisation iterations.

ii. If the optimisation criteria are met.
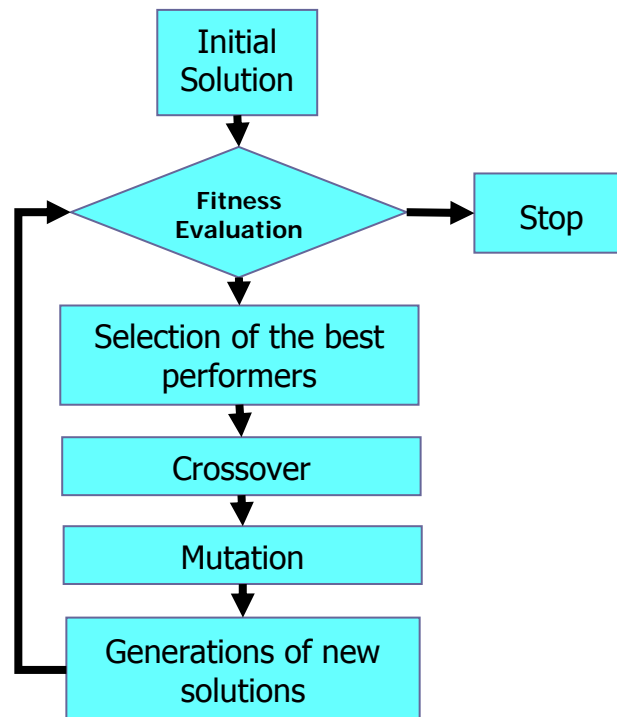
Figure 2.6 The GA cycle

2.2.3 Fuzzy Logic

Fuzzy logic is a problem solving technique with applications in control and decision making. It produces simple linguistically-expressed rules in the form of *IF (condition 1) AND (condition 2) AND (condition3), THEN (conclusion 1, with confidence factor x)*. The confidence factor associated with conclusion 1, describes the degree of membership to the extremes of the fuzzy set.   For example, in stating that something is "hot" with a confidence factor of 0.7, the membership of that property is 70% in the "hot" set and 30% in the "cold" set. The intelligent software strength lies in its ability to draw conclusions and generate responses based on vague, ambiguous, incomplete and imprecise information.   To simulate this process of

human reasoning it applies the mathematical theory of fuzzy sets first defined in the 1960s by Zadeh.

Zadeh extended the traditional definition of a logic premise from having just two extremes (either completely true or completely false) to one in which there is a range in degree of truth from 0 to 100 per cent. Hence there is a range from partially true to partially false. Fuzzy logic thus extends traditional logic in two ways: firstly, sets can be labelled qualitatively using linguistic terms (e.g. hot, cold, warm); and secondly, the elements of these sets can be assigned varying degrees of membership called membership functions.

Fuzzy logic is explained by the example shown in Figure 2.7 (Rowe and Roberts, 1998). In this example the x axis is temperature with ranges for the fuzzy sets "cold", "cool", "warm" and "hot". The y axis represents the membership function and ranges from 0 to 100. It can be seen that the temperature $16^{o}C$ can be regarded as both cool and warm with membership functions of 80 and 20 respectively. This means $16^{o}C$ is cool to a greater degree than it is warm. The membership function is very subjective in nature and is a matter of definition rather than measurement.

For example in a simple conventional thermostat Fuzzy logic is expressed in a simple set of linguistic rules to control the on-off switch of a thermostat. If the thermostat allows minimum and maximum settings, for instance 20 - $22^{o}C$ and set at $21^{o}C$.
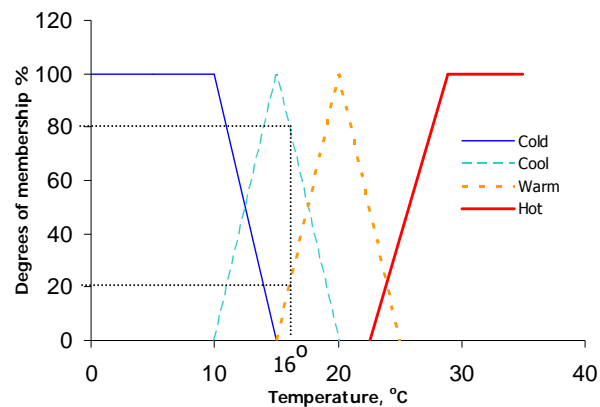
Figure 2.7 Fuzzy sets of temperature

This concept can be observed when the temperature rises to 22°C the heater and fan is turned off and when the temperature drops to 20°C the heater and fan is turned on. Thus the temperature is controlled to fluctuate around 21°C. The rules which control the thermostat may take the form:

♦ IF temperature is cold THEN fan_speed IS high

♦ IF temperature is cool THEN fan_speed IS medium

♦ IF temperature is warm THEN fan_speed IS low

♦ IF temperature is hot THEN fan_speed IS zero

The linguistic variables cold, cool, warm and high are labels which refer to the set of overlapping values shown in Figure 2.7. The labels are triangular input sets and are called membership functions (or linguistic membership functions). Figure 2.8 shows the stages in which a fuzzy controller works. Initially a crisp input, a temperature of 18°C, is translated into fuzzy truth-values in a process called fuzzification. This means that *IS cool* with truth value 0.7 (70%) and *IS warm* with truth value 0.3

(30%).   Then in "Rule Evaluation", a calculation is performed to determine the relevant rules and fuzzy output truth values are computed.   Thus in this case the Rule 3 which is the fan speed will be medium with a truth of 0.7 and Rule 4 which is fan speed will be low with the truth 0.3.   Finally defuzzification combines these two values to calculate the crisp output value for the fan speed, which falls between medium and low.
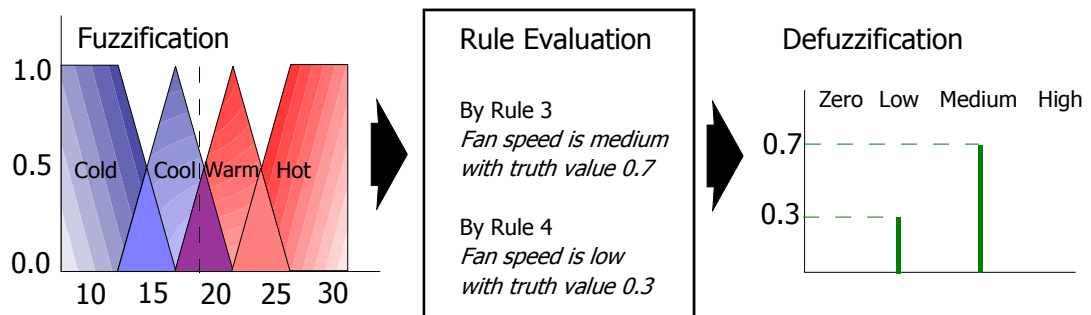
Figure 2.8 The system of fuzzy controller workings

This example illustrates the concepts of Fuzzy Logic and how it can be used to control a process.   In intelligent software system the same concept is applied to make decisions about various inputs (Chapter 6).   The next section describes how fuzzy Logic is combined with ANN to extract rules from data sets.

2.2.4 Neurofuzzy Logic

Neurofuzzy techniques unite the generality and flexibility of representation, a feature of fuzzy logic, with the powerful learning and adaptive capability of ANNs (Jang *et al.*, 1997).   It produces a similar result to fuzzy logic i.e. simple linguistically-expressed rules in the form of *IF (condition 1) AND (condition 2) AND (condition3),*

THEN (conclusion 1, with confidence factor x). The neural networks in neurofuzzy logic are not Multi-Layer Perceptron (MLP) networks but are Associative Memory Networks (AMNs). MLPs and AMNs have different structures. Figure 2.9 shows the basic structure of neurofuzzy logic.
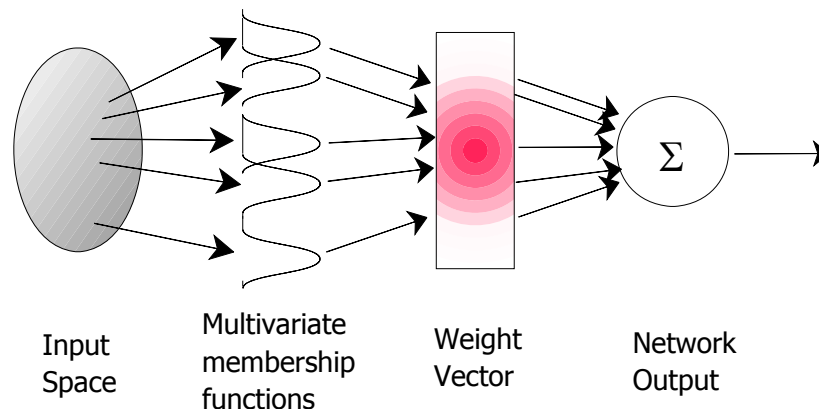
Figure 2.9 Basic structure of a neurofuzzy system (Adapted from Bossley, 1997)

Neurofuzzy networks are neural-processing structures that emulate the fuzzy logic functions. They are similar in architecture to the multilayer perceptron but are not adaptive; each must be constructed specifically for its intended application (Skapura, 1996). In the multilayer perceptron each element is completely interconnected between layers with each connection being a specific weight. However in the neurofuzzy network the connections are only between elements that require them for specific rules and there are no associated weights.

2.3 Summary

This chapter has covered a short history of roll compaction, a brief explanation of the major modelling techniques used for roll compaction, summary of previous research on roll compaction and a brief background on the individual software systems.The research conducted on roll compaction mainly comprised of the effects of the process parameters on roll compaction, validation of existing models and novel methods in improving the roll compaction process. There is a gap in knowledge about the ability to predict the relationship of powder material characteristics to the final roll compacted ribbon quality. Compared to the industries where roll compaction is mainly used, the pharmaceutical companies utilise a more varied powder material in their tablet formulations. It is important for the pharmaceutical industry to develop a model to understand the contributions of various types of pharmaceutical excipients on the final ribbon quality which in turn contributes to the final tablet quality.

Research conducted by Inghelbrecht *et al.* in 1997 showed that neural networks had been successful in predicting the relationship between roll compaction process parameters and the ribbon compacts and granules. The ANNs are flexible and able to process a high number of input variables which are multivariate in nature. The intelligent software mainly applying ANNs, was made with ease of use in mind and flexibility in organizing large amounts of data and input variables to predict output property. Therefore the intelligent software was chosen to be used as prediction software for model development.

The review on roll compaction modelling techniques highlighted the fact that Johanson's theory and DEM simulation were the only two models which had taken into account the powder material characteristics and had been validated to predict roll compaction outputs. However Johanson's theory was used more often than DEM modelling, thus Johanson's theory is a more established theory and was used in comparison to the predictability of models developed from the intelligent software.

Extensive work has been predominantly conducted to investigate effects of roll compaction processing parameters. However there is a lack of study on the contributions of the powder characteristics to the final roll compacted product (i.e. ribbon compact) quality. The intrinsic characteristics of the ribbon compact are largely due to the characteristics of the starting material. For instance, Johanson's theory has taken into account the Compressibility κ, effective angle of internal friction and the angle of wall friction of the powder material but it did not consider the contribution of the particle size distribution. Research conducted in the subsequent years did not consider relating the effects of the particle size distribution to the ribbon quality in a predictive model. Particle size distribution is only one of many methods which could be used for powder material characterisation. Chapter 3 describes the various powder characterisation methods employed here and the results from this chapter were used in the intelligent software model training in Chapters 6 and 7.

3 Characterisation of Pharmaceutical Excipients Used in Roll Compaction

ABSTRACT

The physical characteristics of pharmaceutical excipients contribute to the success or failure of pharmaceutical processing. The traditional methods of characterising pharmaceutical excipients are by size and morphology of individual particles (Aulton, 2002). However in terms of processing pharmaceutical products the bulk powder consists of millions of particles and it is important to be able to represent this as an integrated effect. Hence it is important to be able to represent this integrated effect in terms of bulk powder properties. The type of bulk material characterisation for this research was chosen from previous roll compaction modelling work.

This section describes the methods of determining the bulk powder material characteristics for five types of powder materials (i.e. pure and binary mixtures). The investigated powder material characteristics were particle size distribution, poured density, tap density, true density (i.e. particle density), Compressibility κ, compactibility, angle of internal friction, angle of wall friction and flow function.

3.1 Introduction

Pharmaceutical excipients are powder materials in a tablet which are not the active ingredients. They are added to ease the tabletting operation and to improve the quality of the tablet. Previous modelling methods revealed that certain powder material characteristics were important in predicting the roll compaction output property. For instance, Johanson's theory showed that the Compressibility κ, effective angle of internal friction and angle of wall friction were important variables in predicting the roll compaction pressure distribution and nip angle. In addition to that, DEM simulation required the particle density, particle mean diameter, effective angle of internal friction and angle of wall friction to model the roll compaction process. Hence, in this research work it was decided that the bulk powder material characterisation would include experiments to determine particle size distribution, poured density, tap density, true density (i.e. particle density), Compressibility κ, compactibility, angle of internal friction, angle of wall friction and flow function. In this section, the methods of bulk powder material characterisations are described and the results are presented.

3.2 Materials

Microcrystalline cellulose, dicalcium phosphate anhydrate and magnesium stearate were used as supplied and in the following mixtures: 1:1 and 2:1 ratios of MCC and DCPA, and MCC with 1% of MgSt.

Table 3.1 shows the basic material information. The interested reader can find further details of the main pharmaceutical uses and range of physicochemical characteristics in American Pharmaceutical Association and The Pharmaceutical Society of Great Britain (1986) and Cartensen and Ertell (1990). SEM photos of the three materials are shown in the next couple of pages.

Table 3.1 Material data

Material	Commercial name/Supplier	Physical appearance	Type of Excipient
Microcrystalline cellulose (MCC)	Comprecel M101/ Mingtai Chemical Co. Ltd, Taiwan from UNIVAR, UK	White, insoluble in water, non-reactive, free-flowing and needle-like structure.	Filler, binder, diluent, disintegrant, lubricant or glidant.
Dicalcium phosphate anhydrous (DCPA)	Anhydrous Emcompress /UNIVAR, UK	White, odourless, and tasteless powder.	Diluent or a direct compression excipient.
Magnesium stearate (MgSt)	Magnesium stearate/Mallinckrodt, USA	White powder and unctuous nature	Lubricant, glidant or antiadherent.

3.3 Apparatus and Methodology

3.3.1 Determination of Particle Size Distribution

Particle size distribution is a very important particle characteristic. Although using an average particle size is sometimes sufficient, it is usually very useful to know the width of a particle size distribution in agglomeration. As one obvious example of the importance of the size distribution, the void fraction will be affected by whether the "fines" in a powder are small enough to fill the voids between the larger particles. The higher the amount of pores the weaker the compact is, hence by filling the pores with fines the compact is strengthened.

The particle size distributions of the excipients were determined using laser light scattering (Sympatec Helos-Rodos T4.1). This equipment converts the angular distribution of the forward scattered light intensity from a multiparticle field measurement into a size distribution (Seville *et al.*, 1997). For particles that are much larger than the wavelength of light, any interaction with particles causes light to be scattered in a forward direction with only a small change in angle. This phenomenon is known as the Fraunhofer diffraction.

As the sample is fed into the observation region, it is illuminated using a parallel beam from a He/Ne laser. The scattered light is captured and focused on to a position-sensitive detector at the focal point of the lens. The distribution of scattered light is then converted into a particle volume distribution.

An injector size of 4 mm[§], vibration on the chute of 70% and height of funnel gap over the chute of 2 mm was selected and this automatically gave the settings for the VIBRI feeder. The lens numbers used were R4 (range 0 – 350 μm) or R5 (range 0 – 875 μm) depending on the size range of the particles. The minimum mass of sample used to obtain enough sample distribution was 25 g. The pressure at which the powder material was dispersed was chosen at 0.5 bar increments starting from 0.5 to 3 bar. The powder materials were dispersed at increasing

---

[§] The diameter of the injector should be selected based on the particle size, its diameter should be at least three times the particle diameter to prevent blockage.

pressure to obtain an optimum measuring pressure. This is done in order to ensure disaggregation without undue attrition. It should be noted that some samples may not exhibit an optimum pressure and these should be tested at an arbitrary pressure (e.g. 1bar) for comparison of results.

3.3.2 Determination of Poured and Tap Density

The poured density can be defined as the mass of the particles divided by the volume they occupy, including the space between the particles (American Society for Testing and Materials, 1994). Tap density is defined as the apparent powder density obtained under stated conditions of tapping (British Standards Institution, 1991). Tap density may be expected to depend on the shape, absolute size and size distribution, and surface characteristics of the ultimate particles, as well as the state of agglomeration of the powder (Veale, 1972). The poured and tap density measurement was conducted for all the powder materials mentioned in section 3.2.

3.3.2.1   Carr's Index

A simple test has been developed to evaluate the flowability of a powder by comparing the poured (fluff) density (ρ_b) and tapped density (ρ_t) of a powder and the percentage at which it packed down (Aulton, 2002). Carr found that the percentage compressibility of a powder is a direct measure of the potential powder arch or bridge strength and stability. It is calculated according to Equation (3.1):

$$Carr's\ index\ (\%) = \frac{Tapped\ density - Poured\ density}{Tapped\ density} \times 100 \qquad (3.1)$$

This index can be determined on small quantities of powder and may be interpreted as in Table 3.2. This table shows the generalized relationship between descriptions of powder flow and percent compressibility, according to Carr (Staniforth, 2002). Carr's Index is a one-point determination and does not reflect the rate at which the powder flows.

Table 3.2 Carr's Index as an indication of powder flow.

Carr's Index (%)	Type of flow
5-15	Excellent (free-flowing granules)
12-16	Good (free flowing powdered granules)
18-21	Fair (powdered granules)
23-28	Poor (very fluid powders)
28-35	Poor (fluid powders cohesive powders)
35-38	Very Poor (fluid cohesive powders)
>40	Extremely Poor (cohesive powders)

3.3.2.2   Hausner ratio

Hausner found that the ratio of tapped density to poured density was related to interparticle friction (Staniforth, 2002).   He showed that powders with low interparticle friction, such as coarse spheres had ratios of approximately 1.2, whereas more cohesive, less free-flowing powder such as flakes, have Hausner ratios greater than 1.6. Table 3.2 shows the Hausner ratio as an indication of powder flowability

Table 3.3 Hausner ratio as an indication of powder flowability

Hausner ratio	Type of flow
<1.25	Good flow
>1.25	Poor flow

Prior to analysis of poured density, 160 g of powder material was weighed and placed into a glass bottle. The glass bottle was not more than half full to allow blending of the composite by 10-end-over-end revolutions. The sample was then sieved. Next 25 g of the sieved powder material was weighed and placed into a glass jar. Subsequently the jar was tumbled by end to end motion for 10 revolutions in 20 seconds. After that the powder material was sieved again with a 1 mm sieve to remove lumps.

A clean 250 ml graduated cylinder was weighed. The powder material was transferred into the graduated cylinder; a paper funnel was placed on the cylinder mouth, then the cylinder was held at a 45° angle and with a metal spatula the powder material was fed into the cylinder. After 20 seconds the graduated cylinder was tapped onto the palm of the hand three times. Subsequently after 30 seconds an estimated average level in the cylinder to the nearest millilitre was recorded as V_o. Next the cylinder was reweighed. After that the bottom of the cylinder was layered with a piece of rubber cover to protect the glass cylinder and then placed on a Peschl shear tester sieve shaker (Figure 3.1) to be vibrated at maximum amplitude until the level of powder stopped decreasing. This means after tapping for a time, t or more, the volume in the graduated cylinder remained constant. Finally an estimated average level in the cylinder to the nearest millilitre was recorded as V_f. The steps above were repeated three times on each powder material.

Figure 3.1 a) Tapping device and b) Peschl Shear Tester Vibrating Sieve Shaker
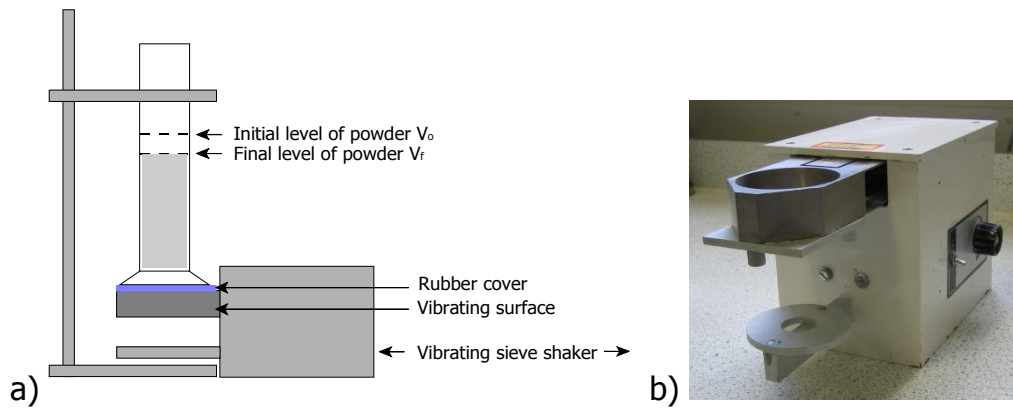
3.3.3 True Density Measurements – Hydrostatic Weighing

True density (also called the true particle density) is defined as the mass of a particle divided by its volume, excluding open pores and closed pores (British Standards Institution, 1991).   Hydrostatic weighing is also known as the displacement method.   This method involves determining the volume of a solid sample by comparing the weight of the sample in air to the weight of the sample immersed in a liquid of known density.   The volume of the sample is equal to the difference in the two weights divided by the density of the liquid.   Subsequently, the density of the sample can be determined by dividing the loss of weight of the immersed object by the volume found.   Water was used in this study as the liquid displacement medium.   All 5 powder materials are insoluble in water.   It is assumed that water penetrates all the pores.

First the empty density bottle and cover was weighed (g).   Secondly, the temperature of water was recorded ($^{\circ}$C).   Then the bottle was filled with water, covered and weighed (g).   Next the bottle was emptied; a powder material with

known mass (g) was poured into the bottle and refilled with water. Finally the bottle was weighed (g). This experiment was repeated five times for each powder material.

To calculate the true density of the solid sample; firstly the total volume of bottle (cm^3) was obtained by dividing the mass of water m_{W1} by the density ρ of water at temperature (^{o}C). Then the final mass of water (after adding the sample) m_{W2} was calculated by subtracting the mass of sample and the mass of empty bottle from the final bottle mass. Next the volume of water was obtained by taking a ratio of the final mass of water m_{W2} over the density of water. Hence the volume of sample is the volume of water subtracted from the total volume of the bottle. Finally the true density of sample can be obtained from the ratio of mass of sample over volume of sample.

3.3.4 Uniaxial Compaction – Compressibility and Compactibility Study

The aim of this investigation was to obtain a Compressibility value κ for each individual powder sample. The Compressibility value κ reflects the ability of a powder bed to decrease in volume under increasing pressure (Snow *et al.*, 1997). It is interesting to note that the compression of powder within the roll compaction process has itself been compared with the operation of uniaxial compaction (see Section 4.3). The powder sample was compressed in a punch and die (Figure 3.2). As the pressure increases the density of the powder bed increases (see Figure 3.3). The powder Compressibility κ is defined as the reciprocal of the slope

of the linear portion of the pressure profile. The density at pressure *P* is given by a compaction equation of the form:

$$\rho = \rho_o \left[\frac{P}{P_o} \right]^{1/\kappa} \tag{3.2}$$

where ρ_o is the density at an arbitrary standard pressure P_o. This equation has been shown in this study to apply approximately over the range 40 to 100 MPa.
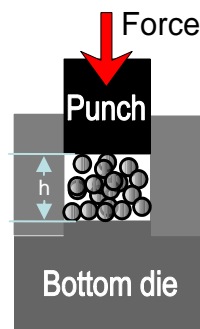
Figure 3.2 Schematic diagram of uniaxial compaction of the powder material.
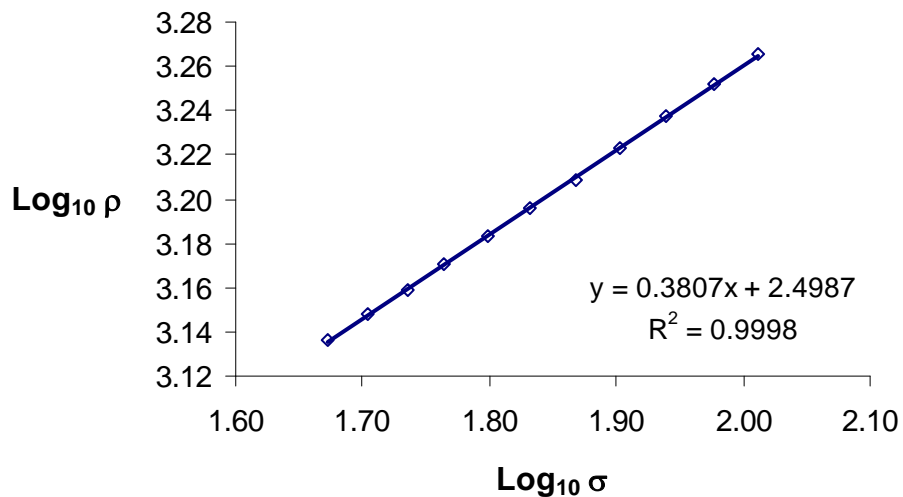
Figure 3.3 Graph of $\log_{10}$ of density (kg/m^3) against $\log_{10}$ of pressure (MPa) for 0.3 g of MCC (Pharmacel 101) compacted to maximum pressure of 100 MPa at compaction speed of 1 mm/sec.

An instrumented uniaxial press (30 kN, Lloyd Instruments 6000R, Fareham, England) was used to fabricate the compacts (Figure 3.4). The punch and die was

10 mm in diameter. Firstly the mass of powder material m_s which gave a 1 to 3 tablet height to diameter ratio at zero porosity was weighed. Next the punch and die were lubricated with 6%w/w of MgSt in methanol and allowed to dry prior to hand filling. The powder material of mass m_s was then filled into the 10 mm die diameter and aligned under the punch ready for compression (Figure 3.2). Next, the sample was compressed at pressures of 100 MPa at a speed of 1 mm/sec. This uniaxial compaction speed was calculated from the slowest roll compaction speed (i.e. 1 rpm for nip angle of 3° see section 9.1 in Appendix 1 for calculation) for comparison. Then the tablet thickness in the die was recorded. Finally the ejected tablet thickness was measured using an electronic vernier micrometer (Mitutoyo, Japan). This experiment was repeated five times for two pure samples and three binary mixtures.

A compliance test was conducted to validate the efficiency of the testing machine. The compliance test involved pressing the punch onto a flat surface (Figure 3.5). The flat surface must be the same material as the bottom punch and the punch was moved down at a speed of 1 mm/sec. The result was used to correct the uniaxial compaction profile (see Section 9.2 in Appendix 2).
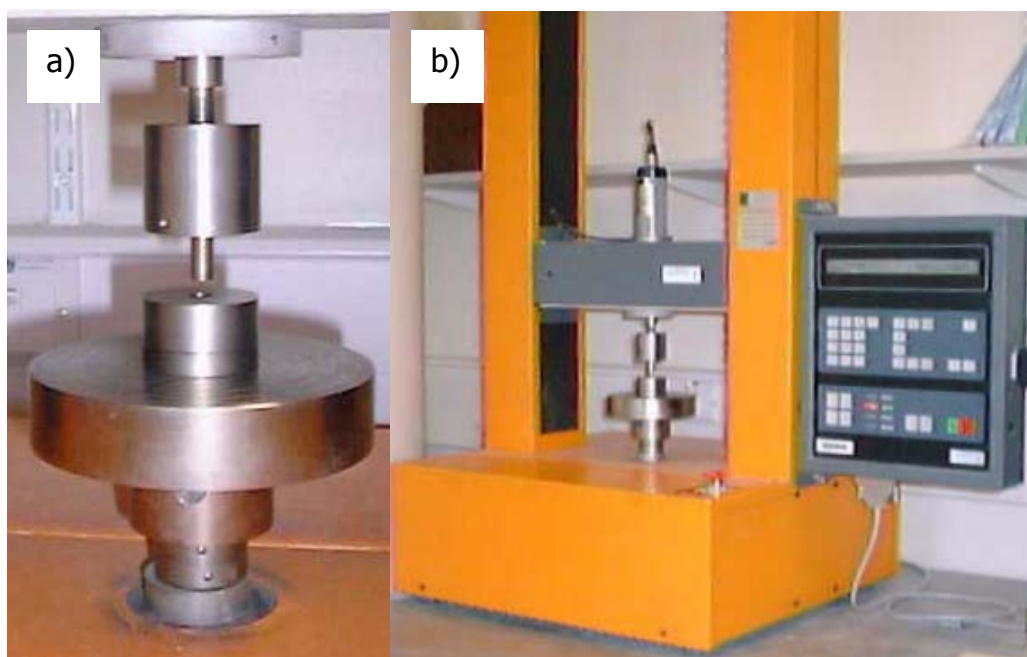
Figure 3.4 a) Punch and die, b) Lloyd universal testing machine.
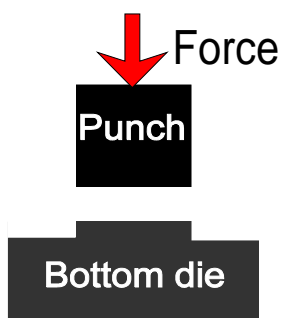
Figure 3.5 Schematic diagram of compliance test.

The tablets produced from the uniaxial compaction above were then mechanically tested to investigate the compaction behaviour of the powder material. The mechanical testing employed in this study is the diametrical compression test. This is an indirect method of determining the radial tensile strength of homogenous disk-shaped materials by the failure of the tablet (Edge *et al.*, 2000). The radial tensile strength refers to the maximum tensile stress a material will withstand prior

to fracture. This test will be used to approximate the compactibility of the tablets. Compactibility is defined as the ability of the powder to form a mechanically resistant tablet, described in terms of tablet strength against the applied compaction stress. The compactibility will be represented by the tensile strength calculated from the Equation (3.3), (Fell and Newton, 1970):

$$\sigma_x = {2F}/{\pi Dt} \tag{3.3}$$

where σ_x is the tensile strength F is the force required to break the tablet, D is the diameter of the tablet and t is the tablet thickness.
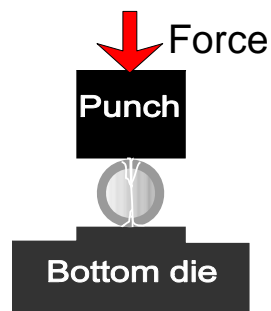
Figure 3.6 Schematic diagram of diametric tensile testing.

The compression test involved applying load onto the tablets along the tablet diameter (Figure 3.6).   The load was increased until a failure plane was established. The compact failed (i.e. breaks in tension) when the crack propagates across the specimen. This test was performed at a crosshead movement rate of 0.5 mm/sec and a load limit of 500N, using a 5kN load cell in a Lloyds Instruments LR30K machine, Fareham, England. Again the machine was remotely controlled by the computer via software. The software plotted a load against displacement graph

where the behaviour of the tablets being tested could be observed. This experiment was repeated five times for two pure samples and three binary mixtures.

3.3.5 Schulze Shear Testing – Powder Flow Investigations

The presence of inter-molecular forces results in a tendency for solid particles to stick to themselves (cohesion) and to other surfaces (adhesion). The cohesive forces acting between particles in a powder bed are composed mainly of short-range, non-specific van der Waals forces which increase as particle size decreases and vary with changes in relative humidity (Staniforth, 2002). The magnitude of cohesion and adhesion affects the frictional forces acting within a powder bed and resisting powder flow (Seville *et al.*, 1997).

The shear test has been designed to determine the flow properties of bulk solids that are required for establishing the functional and structural dimensions of silos, bins and hoppers. It is also used for quality control and research in the field of industrial process technology which requires the mechanics of bulk solids. The flow properties which arise from the interpretations of the shear testing are the *effective angle of internal friction*, the *angle of wall friction* and the *flow function* of the powder material investigated.

Effective Angle of Internal Friction

Knight (2003) explained that the powder state in the nip of a roll compactor is analogous to that in the converging section of a hopper, except that the powder becomes more compressed as it moves. It is therefore appropriate here to use a shear tester to measure the frictional properties of the powder which are of importance in determining the flow into the compactor

By using the Mohr-Coulomb relationship (Nedderman, 1992; Seville *et al.*, 1997) the effective angle of internal friction could be *evaluated* from the yield locus (also known as the Coulomb yield line) for a specific powder using a simple direct shear test. The Schulze shear cell is a simple approximation method to evaluate the yield locus of a powder material. It is an annular shear cell and it is one of three well known shear cells. The other two are the Jenike (Figure 3.7 a)) and the rotational (or Peschl) shear cell. The shear tester method of operation is rotational shearing in an annular trough.

The Coulomb model is a macroscopic yield criterion for initiating flow. It is represented in a two-dimensional stress representation. The third dimensional stress is not taken into account in the yield criterion. For a powder to start to shear, or yield, Coulomb proposed a linear relationship between the shear stress and normal stress (Seville *et al., 1997*) (Figure 3.7 b)):

$$\mu = \tan \phi \tag{3.4}$$

$$\tau = \mu\sigma + c \qquad\qquad (3.5)$$

where τ and σ are the shear and normal stresses μ is the internal friction coefficient and c is the cohesive shear stress or cohesion (i.e. the shear stress of the powder material under zero applied load). The internal friction coefficient μ is defined as the ratio of the shear force σA to the normal force τA (Figure 3.7 a)).
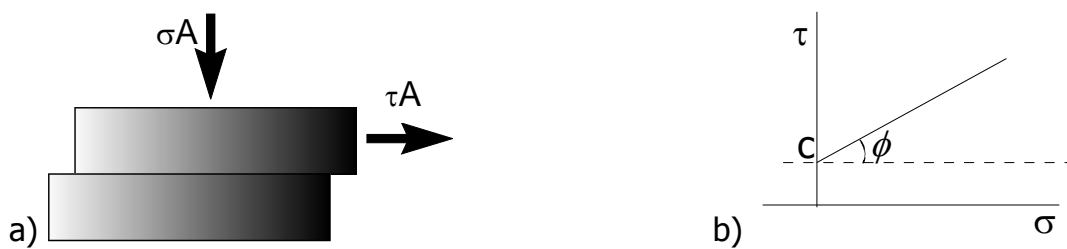
Figure 3.7 a) Schematic diagram showing the Jenike shear cell and b) the graphical representation of the Coulomb Model, where τ is shear stress, σ is normal stress, c is cohesive shear stress and $\mu = \tan \phi$ and a typical value of ϕ is 25 to 45 degrees.

The Schulze shear[**] testing equipment is shown in Figures 3.12 and 3.13. Firstly the sample was fed into the annular trough using a vibrating shaker to remove agglomerates. Next, it was consolidated with a predefined normal stress σ_{pre} acting onto the annular lid.   Then it was sheared while the normal stress was kept constant. The shear force increases with time t as indicated in the left diagram of Figure 3.8. After the shear stress reaches a constant value (i.e. the point at which the powder is in a state of steady flow), this was taken as the first point on the yield locus as shown on the right of Figure 3.8. Next the shearing is reversed until the shear stress is zero and then the normal stress is brought back to zero.

[**] It was commercially supplied by Dr.-Ing. Dietmar Schulze, Germany.
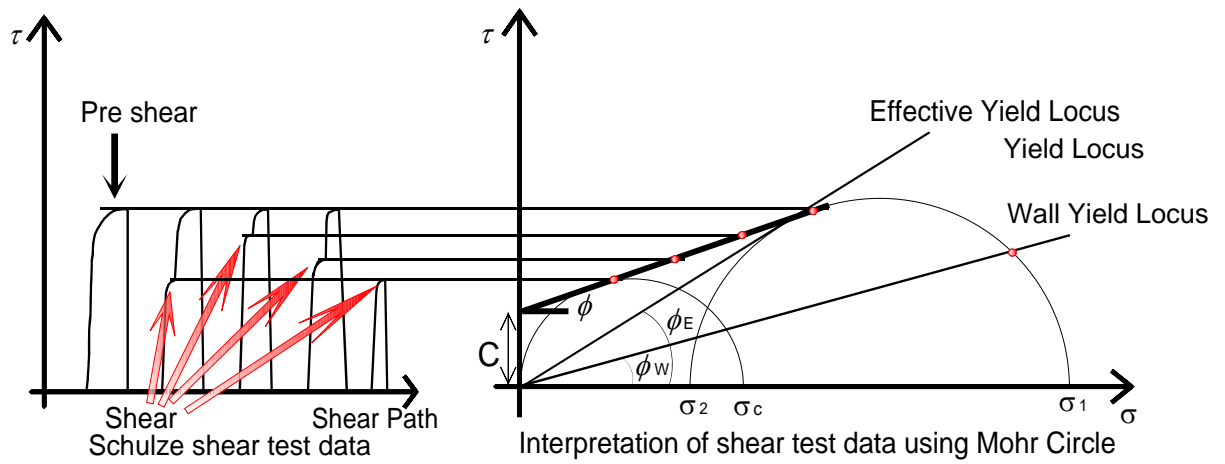
Figure 3.8 On the left is a plot of shear stresses and on the right is the yield locus.

Figure 3.9 Ring Shear Tester RST-XS. Automatic powder tester/flowability tester.

○ Cut-away view of the shear cell (series RST-01)   ○ Shear cell approx. 30 ml, 2 EURO coin
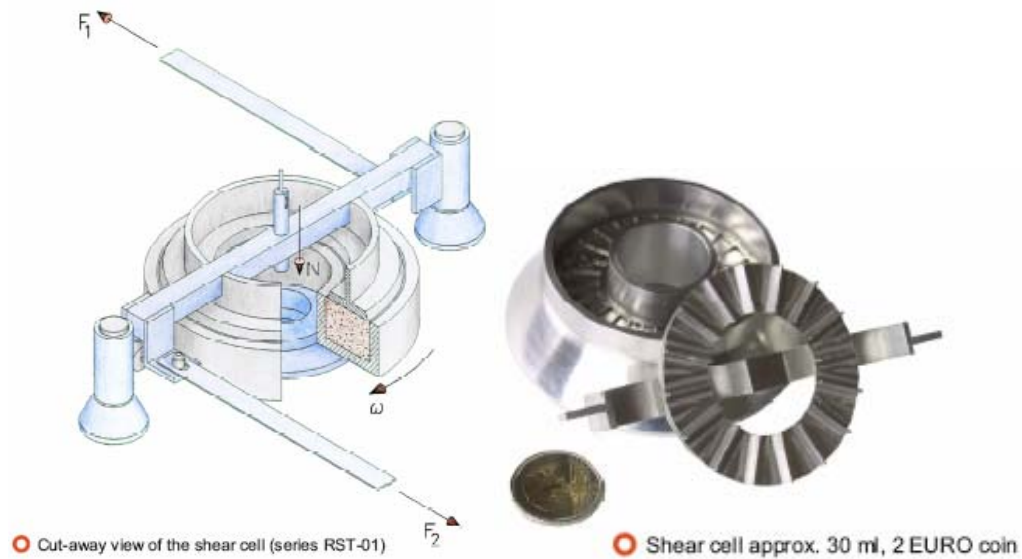
Figure 3.10 Diagrams and picture of Schulze RST-01 annular shear cell.

Following that, the sample is sheared again at a normal stress σ_{sh}, which is less than σ_{pre}. Table 3.4 shows the list of consolidation stresses and normal stresses used in the shear tests. Since the powder sample was then sheared at a smaller normal load than at preshear, it flowed at a lower shear stress. This lower shear stress corresponds to another point on the yield locus as shown in Figure 3.8. The following points were found in the same manner but at increasing normal stress σ_{sh} (but less than σ_{pre}), which then forms the yield locus.

Table 3.4 Consolidation stress and normal stress values used within the shear testing test.

Consolidation stress σ_{pre}	8 kPa	7 kPa	6 kPa
Normal stress σ_{sh} 1	1.6 kPa	1.4 kPa	1.2 kPa
Normal stress σ_{sh} 2	3.2 kPa	2.8 kPa	2.4 kPa
Normal stress σ_{sh} 3	4.8 kPa	4.2 kPa	3.6 kPa
Normal stress σ_{sh} 4	6.4 kPa	5.6 kPa	4.8 kPa
Normal stress σ_{sh} 5	1.6 kPa	1.4 kPa	1.2 kPa

Traditionally, it is assumed that the Mohr's circle describing the stress state during a consolidation or steady flow process touches the appropriate incipient yield locus at its end point, as shown in Figure 3.8 for the case of a Coulomb material (Nedderman, 1992). This interpretation gives rise to two important concepts, the *effective yield locus* and the *flow function*. The tangent of the larger Mohr circle running through the origin defines the effective yield locus and the angle to the horizontal is the effective angle of internal friction ϕ_e.

Angle of Wall Friction

The angle of wall friction ϕ_w is an approximation of the friction of powder material on a surface, which in this case is a stainless steel surface (Equation (3.6)) (Brown and Richards, 1970). The wall yield locus has been used to describe the relationship between the tangential and normal forces at the roll surface in the roll compaction process (Johanson, 1965). A linear wall yield locus is the ideal Coulomb Model for slip against a surface. It states that for a powder to slide against a surface, the following conditions must be met (Figure 3.11 b)):

$$\mu_w = \tan\phi_w \tag{3.6}$$

$$\tau = \mu_w\sigma + c_w \tag{3.7}$$

where τ and σ are the shear and normal stresses, μ_w is the internal friction coefficient and c_w is the cohesive shear stress or cohesion (i.e. the shear stress of the material under zero applied load). The wall yield locus is described by the angle of wall friction ϕ_w.
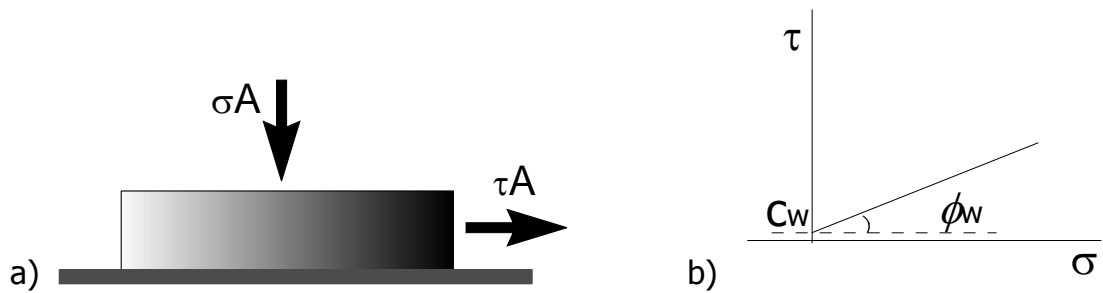
Figure 3.11 a) Schematic diagram showing the Jenike wall shear cell and b) the graphical representation of the Coulomb Model for wall friction, where τ is shear stress, σ is normal stress, c_w is cohesive shear stress and $\mu_w = \tan \phi_w$ and a typical value of ϕ_w is 15 to 35 degrees.

An approximate measurement of the angle of wall friction can also be carried out using the Schulze shear tester. The shear cell was replaced with a wall friction cell; see Figure 3.12. The wall material is stainless steel which is the same material as the surface of the rollers. The wall shear stresses τ_w required to move the powder material across the wall material are measured under different normal stresses σ_w. In the Schulze wall shear experiment four decreasing normal stresses were chosen and plotted on a graph of shear stress τ against normal stress σ (see Table 3.5). Connecting the points on the graph produces a wall yield locus. The angle of the wall yield locus to the horizontal is the angle of wall friction σ_w.

Figure 3.12 Schulze RST-03 annular wall shear cell.

Table 3.5 Normal stress values used in wall shear testing.

Normal stress σ_{sh} 1	6.0 kpa
Normal stress σ_{sh} 2	4.4 kpa
Normal stress σ_{sh} 3	2.8 kpa
Normal stress σ_{sh} 4	1.2 kpa

Flow Function

The flow function of a powder is the inverse of the gradient of the graph of unconfined yield stress σ_c against the consolidation stress σ_1. The larger the flow function, the better a bulk solid flows. Powders are classified in this approach as 'free flowing', 'easy flowing', 'cohesive', 'very cohesive' and 'non-flowing'. A 'non-flowing' powder builds up stable arches or pipes when discharged from a silo, or cakes at storage or transport.   In industries which involve bulk powder handling, it is essential to be able to describe the flowability of a powder or a bulk solid.

The relationship between consolidation stress σ_1 and unconfined yield stress σ_c could be shown through a simple uniaxial compression test as shown in Figure 3.13 (Schulze, 2003). Step a) in Figure 3.13 shows a cylinder with frictionless walls filled with a fine-grained, cohesive bulk solid.   Firstly the bulk solid was consolidated by the consolidation stress σ_1. Subsequently in b) the cylinder is removed and then in c) the cylindrical bulk solid sample is exposed to an increasing compressive stress until the specimen breaks (flows). The stress acting at failure is called the unconfined yield stress σ_c.
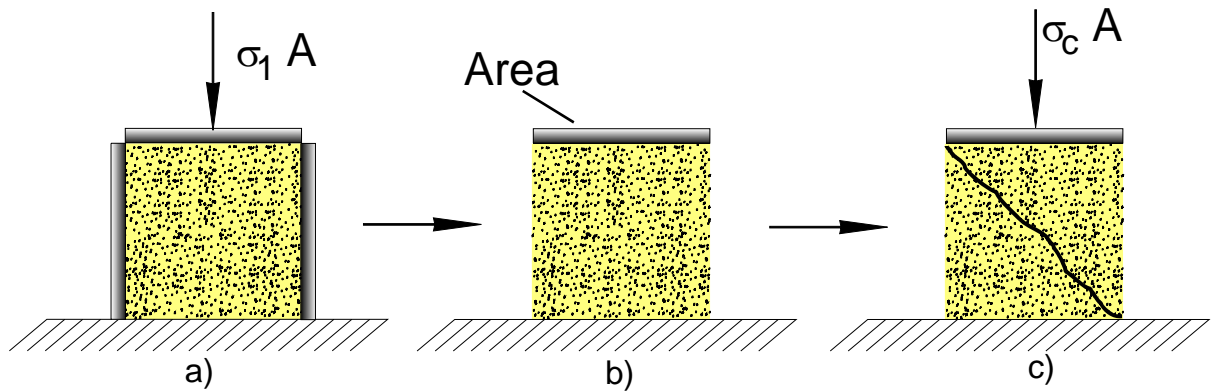
Figure 3.13 Uniaxial compression test to obtain the unconfined yield stress of a powder material.

The unconfined yield stress σ_c increases with increasing consolidation stress σ_1. Figure 3.14 shows that the unconfined yield stress σ_c is dependent on consolidation stress σ_1. The classification of the flow function numbers are as shown below:

Flow Function, FF	Type of Flow
FF<1	Non-flowing
1<FF<2	Very cohesive (to non-flowing)
2<FF<4	Cohesive
4<FF<10	Easy flowing
10<FF	Free flowing

The Schulze shear tester can be used to obtain the flow function value. These values were obtained from the interpretation of shear test data using a Mohr's Circle on Figure 3.8. In the shear testing experiment the consolidation stress σ_1 is a consequence of the normal stress σ_{pre}. Hence, by using three varied normal stresses $\sigma_{pre,}$ three respective consolidation stresses σ_1 were obtained. These were then plotted on the graph of unconfined yield stress σ_c against the consolidation stress σ_1 (see Figure 3.14).
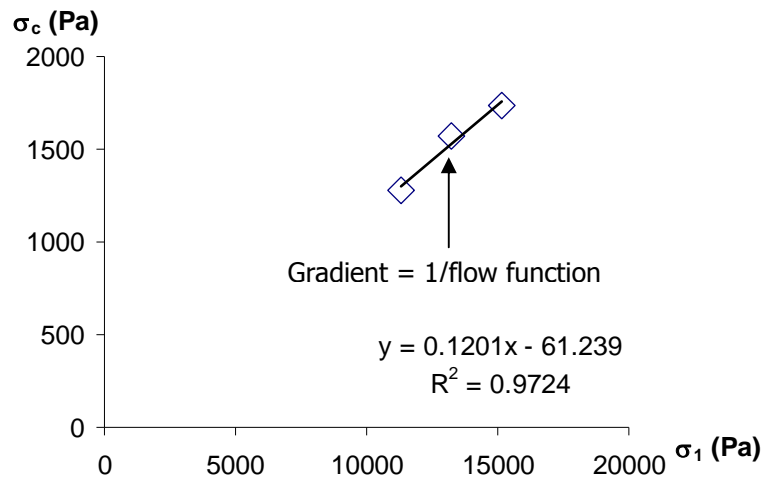
Figure 3.14 Graph of unconfined yield stress σ_c against consolidation stress σ_1 for the evaluation of DCPA powder flow function.

A well defined procedure was laid out in the operating manual (Schulze, 2002). The operation of the tester is limited to the filling of the shear cell (or the wall friction cell) with powder materials, the placement of the cell on the tester, entering the test parameters into the computer program RST-CONTROL 95 and starting the test and finally cleaning of the shear tester. Every step of the shearing is guided by the computer prompts.   The computer program conducts all the calculations to obtain the mechanical properties of the powder.   The underlying concepts and calculations are explained in Nedderman (1992).

3.4   Results and Discussions

3.4.1 The Particle Size Distribution for Pure Materials and Binary Mixtures

Figure 3.15 shows that DCPA has a narrow particle size distribution and comprises smaller particles than MCC. Both the pure powder materials have monomodal

particle size distributions. Mixing MCC into DCPA at 2:1 ratio and 1:1 ratio gives a bimodal particle size distribution, as would be expected.

Mixing 1% MgSt to pure MCC gives a monomodal particle size distribution which lies on top of the pure MCC frequency distribution. This means that the addition of 1% MgSt does not produce changes to the particle size distribution that are detectable using laser light diffraction.

The information on Table 3.6 is obtained from the particle size distribution figure above. The ratio of (d84-d16)/d50 will be used to represent the powder material PSD characteristics within the software simulation. Basically, on the extremes a high value of Ratio (d84-d16)/d50 represents a wide PSD and/or low PSD d_{50} and a low value of Ratio (d84-d16)/d50 represents a narrow PSD and/or high value of PSD d_{50}. However it could not differentiate between a narrow PSD over a small PSD d_{50} and a high PSD over a high PSD d_{50}. This will cause a problem if the types of powder material used were more varied. But since in this study we only have two extremes of high and low Ratio (d84-d16)/d50 and it is easy to determine which one has a wide PSD or narrow PSD therefore it is sufficient at this stage to represent the PSD.
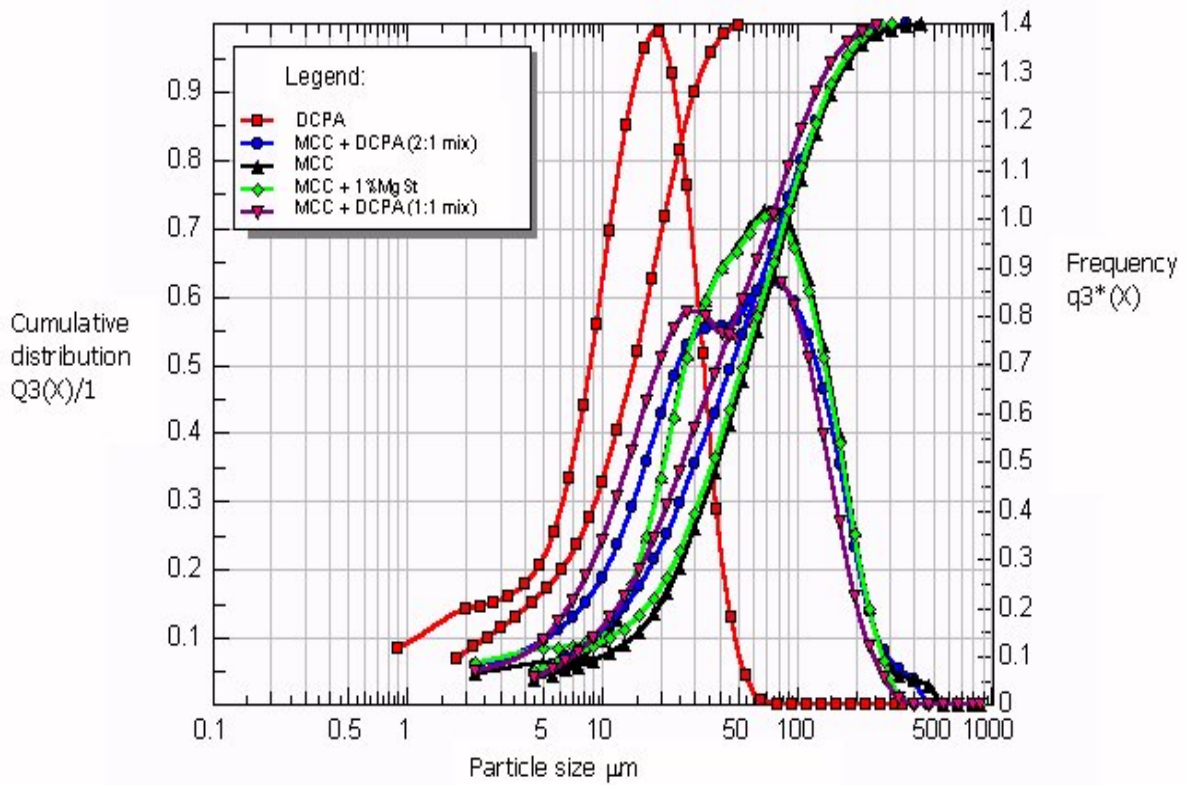
Figure 3.15 Graph showing the particle size distribution of pure DCPA, pure MCC, MCC with 1% MgSt, binary mixture of MCC + DCPA 1:1 mixture and MCC + DCPA 2:1 mixture. Left Y-axis represents the cumulative distribution of the particles. Right Y-axis represents the frequency distribution.

Table 3.6 Particle size distribution for pure and binary mixtures of pharmaceutical excipients

Sample	PSD d50 (µm)	PSD d16 (µm)	PSD d84 (µµm)	Ratio (d84-d16)/d50
DCPA	14.24	4.42	27.17	1.60
MCC + DCPA (1:1 Mix)	39.76	12.96	105.62	2.33
MCC + DCPA (2:1 Mix)	46.06	14.50	118.67	2.26
MCC	54.87	20.84	124.33	1.89
MCC + 1%MgSt	53.77	18.98	122.37	1.92

3.4.2 The Poured and Tap Density for Pure Materials and Binary Mixtures

Table 3.7 shows the poured and tap density of the powder materials studied. DCPA has a higher poured and tap density compared to MCC. The MCC particles have a needle like structure which makes them susceptible to bridging and arching. Bridging and arching of particles may lead to very low tap densities (Veale, 1972). This could explain why MCC has a lower tap density than DCPA. In contrast, DCPA particles have an almost cubic structure which eases the rearrangement of particles within a bed and decreases voidage, thus giving them a higher poured and tap density.

Binary mixtures of 1:1 and 2:1 of MCC and DCPA decreases in poured and tap density compared to pure DCPA. This is expected because the amount of DCPA per unit volume is decreasing. However the addition of 1% of MgSt to MCC is contributes to a very slight increase in poured and tap density. The tap density of MCC found in this study is comparable to the value (430 kg/m^3) found in the American Pharmaceutical Association and The Pharmaceutical Society of Great Britain (1986).

According to Carr's Index classification DCPA has poor flow for a very fluid powder, whereas the binary mixtures of MCC and DCPA, pure MCC and lubricated MCC have poor flow and are classified as fluid cohesive powders. However according to the Hausner ratio classification the powders are all poor flowing powders.

Table 3.7 Results from poured and tap density measurements

Sample	Poured density (kg/m^3)	Tap density (kg/m^3)	Carr's Index	Hausner's ratio
DCPA	866.00 ± 24.28	1189.90 ± 6.15	27.25	1.37
MCC + DCPA (1:1 Mix)	560.14 ± 7.18	865.15 ± 5.03	35.26	1.54
MCC + DCPA (2:1 Mix)	446.49 ± 3.26	658.98 ± 3.82	32.25	1.48
MCC	311.00 ± 2.23	450.47 ± 4.06	30.91	1.45
MCC + 1%MgSt	324.09 ± 7.35	454.65 ± 8.27	28.72	1.40

3.4.3 The True Densities for Pure Materials and Binary Mixtures

The true densities found in this section are all approximated values found using the water displacement method. The true densities for DCPA and MCC (grade PH101) (see Table 3.8) found in this study are comparable to the values found in literature American Pharmaceutical Association and The Pharmaceutical Society of Great Britain (1986), Roberts *et al.* (1995), which are 2873 kg/m^3 and 1510 kg/m^3 respectively.

Table 3.8 shows that DCPA has the highest true density value whereas MCC has the lowest true density value. Pure MCC and MCC with 1% of MgSt have comparable true densities. Addition of 1% of MgSt to MCC does not change the true density, which was expected. The true densities of 1:1 and 2:1 mixture of MCC and DCPA are decreasing compared to pure DCPA, which was also as expected.

Table 3.8 True densities from water displacement approximations

Sample	True density (kg/m^3)
DCPA	2849 ± 16
MCC + DCPA (1:1 Mix)	1820 ± 36
MCC + DCPA (2:1 Mix)	1750 ± 37
MCC	1509 ± 31
MCC + 1%MgSt	1500 ± 27

The true density was used in determining the density of ribbons produced from the roller compaction experiments in Chapter 4.

3.4.4 Compressibility κ Value and Tensile Strength of Powder Materials Obtained from Uniaxial Compaction and Diametrical Compaction

The Compressibility κ value reflects the reduction in powder volume under pressure. Figure 3.16 shows the corrected[††] uniaxial compaction of profile in the form of applied pressure against punch displacement.

The Compressibility κ value is the reciprocal of the gradient of the slope of $\log_{10}$ of density (kg/m^3) against $\log_{10}$ of pressure (MPa). Table 3.9 shows a high Compressibility κ value for DCPA and a relatively low Compressibility κ value for MCC. Hence MCC is more compressible than DCPA. The DCPA tablets produced were friable at the edges and this material is also widely known to be difficult to compress, whereas the MCC tablets were solid and compact. This might be due to

---

[††] The correction was conducted using the compliance test results.

the particle shape and the particle size distribution of the powder material. DCPA

has a small particle size distribution as opposed to the larger particle size

distribution of MCC, which gives MCC a higher capacity for pore filling.

Furthermore, MCC is soft and ductile (Roberts and Rowe, 1987), which means that

it yields more easily under compression.
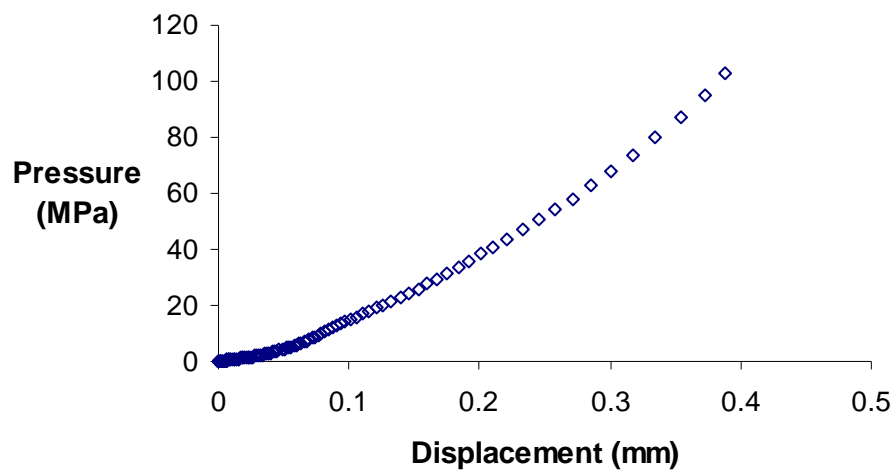
Figure 3.16 An example of the corrected uniaxial compaction profile of MCC at 1 mm/sec in a 10 mm die.

Table 3.9 Results from uniaxial compaction of powder material in 10 mm die, 100 MPa and 1 mm/s and diametrical test at 0.5 m/s

Sample	Compressibility κ	Tensile Strength (MPa)
DCPA	4.11 ± 0.13	0.52 ± 0.21
MCC + DCPA (1:1 Mix)	2.92 ± 0.04	3.25 ± 0.42
MCC + DCPA (2:1 Mix)	2.80 ± 0.02	4.27 ± 0.23
MCC	2.63 ± 0.02	5.90 ± 0.19
MCC + 1%MgSt	2.57 ± 0.04	6.13 ± 0.02

MCC has the highest tensile strength whereas DCPA has the lowest tensile strength (see Table 3.2). This means that MCC forms a strong compact and DCPA forms a comparatively weak compact. Mixing MCC with 1% MgSt does not change the tensile strength significantly. Increasing the ratio of MCC in a binary mixture of MCC and DCPA increases the tensile strength of the binary mixture. The needle-like structure of MCC may also contribute to its compactibility, which means the ability of the powder to remain in its compacted form.   DCPA on the other hand deforms by brittle fracture under compression and hence is less compactable.

3.4.5 Shear Testing and Flowability

The flow function data in Table 3.10 showed a significant difference between the pure excipients and binary mixtures. For instance, the flow function of MCC was 4.2, which means that it is a cohesive powder. Lubrication of the MCC powder with 1% of MgSt increased its flow function, which is reflected in its increased flowability. The flow function of DCPA is 8.9, which means that it is easy flowing. Binary mixtures of MCC and DCPA have an improved flow compared to pure MCC, although the flow was not related to the ratio of the binary mixtures.

The effective angle of internal friction did not show a significant difference as the consolidation stress was decreased from 8 kPa to 6 kPa. The effective angle of internal friction for DCPA is the lowest (about 36°) and MCC has the highest value (about 46°). Lubrication of MCC with 1% MgSt did not show an obvious change in effective angle of internal friction. However the binary mixtures of MCC and DCPA

show an increase in effective angle of internal friction as the proportion of MCC increases.

DCPA has a higher angle of wall friction on stainless steel than MCC, which means that DCPA has more resistance to flow on stainless steel surface compared to MCC. This is in agreement with a statement made by Brown and Richards, 1970 that the wall friction of particles smaller than 20 μm is about double that of the coarsest grade (-76+53 μm). Compared to pure MCC, lubricating MCC with 1% MgSt decreased the angle of wall friction,   MCC + 1% MgSt mixture's decrease in friction to the stainless steel surface. The angle of wall friction for binary mixtures of MCC and DCPA decreases as the ratio of MCC increases.

Table 3.10 Results from the shear testing

Sample	Effective angle of internal friction(^{o})			Angle of wall friction (^{o})	Flow Function
	consolidation stress 8kPa	consolidation stress 7kPa	consolidation stress 6kPa		
DCPA	36.7	36.3	36.3	19.1	**8.33** (Easy flowing)
MCC + DCPA (1:1 Mix)	41.0	41.3	41.3	17.7	**6.70** (Easy flowing)
MCC + DCPA (2:1 Mix)	43.0	43.3	44.0	10.8	**7.21** (Easy Flowing)
MCC	46.7	46.7	46.3	9.9	**4.20** (Cohesive)
MCC + 1%MgSt	46.3	47	46.7	7.8	**8.94** (Easy flowing)

3.5 Summary

In this section a description of the pure powder materials and description of the methods employed to characterise the pure powder materials and the binary mixtures have been presented. The particle size distribution experiment established that the DCPA had a smaller particle size distribution compared to MCC. It also showed that DCPA has a narrower particle size distribution compared to MCC. These two characteristics of the powder material were expressed in the Ratio (d84-d16)/d50.

The poured density and tap density of the powder materials were represented by Carr's Index and Hausner ratio. Carr's Index classified DCPA as a powder which has poor flow although it was classified by the shear testing as a very fluid powder, whereas MCC has a poor flow although it was classified by the shear testing as a fluid cohesive powder. The Compressibility κ value found from uniaxial compaction of a bed of the powder material indicated that dicalcium phosphate was less compressible than MCC. Diametrical compression showed that the MCC retained its compacted form better than DCPA.

Shear testing was used to find the flow functions, effective angle of internal friction and angle of wall friction of the powder materials. According to this test DCPA is an easy flowing powder whereas MCC is a cohesive powder. The shear test also gave an effective angle of wall friction which shows that DCPA has a higher angle of friction on a the stainless steel surface than MCC.

The bulk powder material characteristics are summarised in Table 3.11. These data provide the material characteristics to be used in the predictive intelligent software in Chapter 6 and 7.

Table 3.11 Powder Characteristics. The results shown are up to one standard deviation.

Sample	DCPA	MCC + DCPA (1:1 Mix)	MCC + DCPA (2:1 Mix)	MCC	MCC + 1%MgSt
Particle size distribution					
PSD d_{50}* (µm)	14.24	39.76	46.06	54.87	52.17
PSD d_{16}* (µm)	4.42	12.96	14.5	20.84	18.53
PSD d_{84}* (µm)	27.17	105.62	118.67	124.33	120.4
Ratio $(d_{84}-d_{16})/d_{50}$	1.6	2.33	2.26	1.89	1.95
Densities					
True density (kg/m^3)	2849 ± sd16	1820 ± sd36	1750 ± sd37	1509 ± sd31	1500 ± sd27
Tap Density (kg/m^3)	1190 ± sd6	865 ± sd5	659 ± sd4	450 ± sd4	455 ± sd8
Poured density (kg/m^3)	866 ± sd24	560 ± sd7	446 ± sd3	311 ± sd2	324 ± sd7
Carr's Index	27	35	32	31	29
Hausner's ratio	1.37	1.54	1.48	1.45	1.4
Uniaxial compression					
Compressibility κ**	4.11 ± sd0.13	2.92 ± sd0.04	2.80 ± sd0.02	2.63 ± sd0.02	2.57 ± sd0.04
Tensile Strength (MPa)	0.52 ± sd0.21	3.25 ± sd0.42	4.27 ± sd0.23	5.90 ± sd0.19	6.13 ± sd0.02
Shear testing					
Effective angle of friction (^{o})	36.3 ± sd0.6	41.3 ± sd0.6	44 ± sd0.0	46.3 ± sd0.6	46.7 ± sd0.6
Angle of Wall Friction (^{o})	19.1 ± sd0.4	17.7 ± sd0.4	10.8 ± sd0.2	9.9 ± sd1.4	7.8 ± sd0.5
Flow Function	8.33	9.16	7.21	4.16	8.94

4 Production of Ribbon Compacts by Roll Compaction

ABSTRACT

Roll compaction experiments on the powder materials were conducted to investigate the effects of varying roll compaction process parameters on pure and binary mixtures of the powder materials. The roll compactor was fed by gravity feeding. It was found that different types of powder materials can only be roll compacted within a specific roll speed range and over a particular roll gap range. It was also observed that DCPA formed friable ribbon compacts, but binary mixtures of DCPA and MCC produced stronger compacts. The pressure-density relationship for roll compaction was comparable to uniaxial compaction at lower pressures.

The key aim of this section was to build a database of roll compaction of various powder materials. A description of the roll compaction equipment and methods for ribbon production and ribbon density determination are presented.

4.1 Introduction

In the previous chapter, the powder materials were characterised. The next step in this research was to conduct roll compaction experiments on the powder materials to investigate the effects of varying roll compaction process parameters on pure and binary mixtures of the powder materials. In this section, a description of the roll compaction equipment is given, the method of ribbon compact production is explained, the technique of determining the ribbon density is established and the results are discussed.

The key aim of this section was to build a database of roll compaction of various powder materials. The results from this section will be used in Chapters 5, 6 and 7. In Chapter 5 the results will be compared with predictions from a theoretical model. In Chapter 6 the results will be used to determine the optimum inputs for the predictive intelligent software in Chapter 7.

4.2 Equipment and Methods

4.2.1 The Roll Compactor

The roll compactor was constructed at the University of Birmingham and is shown in Figure 4.1. The design is adapted from work by Michel (1994) and Bourseul (2001). The two fixed rollers of width 46 mm and diameter 200 mm are driven by two stepper motors. The stepper motors are controlled by a LabVIEW program (National Instruments, UK) and driven by a single quartz oscillator unit (McLennan Servo Supplies Ltd, UK: Model PM160). An external clock transmits the same

impulse frequency to both motors, which ensures a simultaneous rotation. The step length was fixed for this work at a 0.9° angle. Therefore, a roll speed of 10 rpm was equivalent to 3600° per minute or 60° per second or 66.67 impulses per second. The roller speed was varied between 0.5 and 9 rpm. The nominal roll gap was manually adjusted between 0 and 5 mm using the gap-setting screw before running the experiment.

During the experimental run both rollers are nominally fixed, but as a result of the compacting action, the actual roll gap exceeds the setting and is monitored continuously using two linear displacement transducers. The difference results from the elasticity of the press itself.
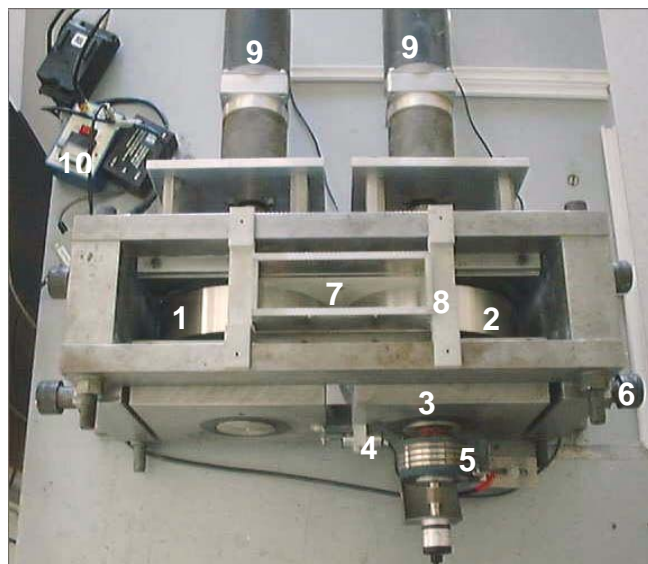
Legend:
1. Fixed roller
2. Movable axis roller
3. Mobile Bearing Block
4. Linear Displacement Transducer
5. Slip Rings Unit
6. Gap Setting Screw
7. Cheek Plate
8. Hopper Brace
9. Stepper Motor
10. Power Supply Unit (For the Pressure Sensors)

Figure 4.1 Photograph of Roller Compactor at the University of Birmingham.

There are two types of sensors on the equipment: two linear displacement transducers (Schlumberger DFG 5.0, RS Components Ltd., Corby, UK) and one piezo-electric pressure sensor which can detect a maximum pressure of 2,750 bars (PCB 105C33, Techni-Measure, Studley, UK). The piezo-electric pressure sensor is

located within one roller and enables the profiles of the pressure normal to the surface of the rollers to be recorded.

The nip angle defines the point at which the powder material starts getting roll compacted, therefore it is important to measure this value. Figure 4.2 shows the compaction profile of microcrystalline cellulose at a roll gap of 1.2 mm and a roll speed of 1 rpm. The nip angle could be estimated from the data by fitting straight line segments to the pressure profile as shown. The neutral angle is the angle at which the maximum roll compaction pressure occurs and represents the transition between the nip and the release region occurs. Note that, depending on the material properties, the peak pressure does not necessarily occur at the smallest distance between the two rollers (i.e. at the point of closest approach to the horizontal). Both Michel (1994) and Bourseul (2001) showed that the neutral angle decreases as the roll speed increases. There is also evidence that an increase in particle size results in a decrease in neutral angle. The main process parameters for the roll compaction are the roll gap, roll speed, roll surface and feed pressure.
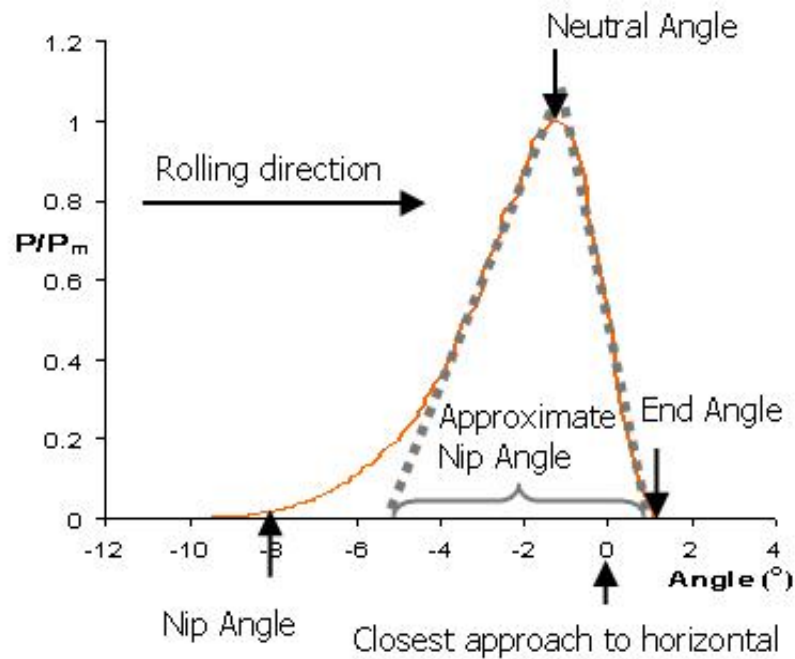
Figure 4.2 A typical roll compaction pressure profile. P is the pressure and P_m is the maximum pressure. The graph shows roller compaction of MCC, at a roll gap of 1.2 mm and a roll speed of 1 rpm.

4.2.2 Production of Ribbons

To operate the roll compactor, firstly the LabVIEW program "*roll press operation.vi*" was loaded from the desktop. Next the motor control, the power supply for the displacement transducers and the pressure sensor were switched on. Then the rollers were cleaned using laboratory paper towelling soaked with ethanol to remove any debris or powder from the roller surface. After that the roll gap was adjusted using the feeler gauge and gap setting screw (see Figure 4.1). The roll gap accross the roll width was kept uniform to avoid any leakage and non-uniform ribbon compact density. Subsequently the location of the pressure sensor was adjusted to the vertical position as a starting position using a low roll speed (1 rpm). Next the experiment data file was named to start the rolls rolling. After the pressure sensor

was set on the vertical position, the roll speed was set on the LabVIEW program. Roll speed and roll gap were varied in the roll compaction experiment as shown in Table 4.1.

To produce the ribbon compacts, firstly an empty bucket was weighed and placed under the roll compactor release region. Then the extractor fan was turned on. Next the powder material was fed into the hopper to the vibrating feeder. After that the rolls were set to start rolling. Next the rolls were allowed to achieve a full 360° turn to allow the pressure sensor reading to reach a steady minimum value. Then the vibrating feeder was set to feed powder material into the roll compactor hopper. Subsequently the roll compaction process was stopped by firstly switching off the vibrating feeder and then clicking on "END" button in the LabVIEW window after about 10-25 rotations. Next the bucket filled with sample is weighed.

The roll compaction process does not always succeed in forming ribbon compacts. This may be due to the rolls jamming or the powders failing to compact. The rolls can be jammed by cohesive or high bulk density powder material. Both these types of powders can choke the feeding area and impede deaeration. To maintain a constant roll compaction of cohesive or high bulk density powder materials the powder was fed carefully. This meant manually controlled feeding to avoid underfeeding or overfeeding the feed area of powder. A vibrating feeder was used to assist in maintaining a constant feed into the feed region, but as this does not totally solve the problem of overfeeding and underfeeding, extra attention to the feed region is required.

Table 4.1 Roll compaction experiments conducted for each material. The minimum amount of roller rotation achieved for certain powders were 8 turns and the maximum amount of roller rotation for other powders were 30 turns.

Experimental reference	Powder Materials	Roll Speed (rpm)	Roll Gap (mm)
1	DCPA	1	1.40
2	DCPA	1	1.20
3	DCPA	1	1.00
4	DCPA	1	0.80
5	DCPA	1.5	1.40
6	DCPA	2	1.40
7	DCPA	2	1.20
8	DCPA	2	1.00
9	DCPA	2	0.80
10	DCPA	3	1.20
11	DCPA	3	1.00
12	DCPA	3	0.80
13	MCC	1	1.00
14	MCC	1	1.10
15	MCC	1	1.20
16	MCC	1	1.40
17	MCC	3	0.70
18	MCC	3	0.80
19	MCC	3	1.00
20	MCC	3	1.20
21	MCC	5	0.70
22	MCC	5	0.80
23	MCC	5	1.00
24	MCC	5	1.20
25	MCC+DCPA (2:1 mix)	1	1.00
26	MCC+DCPA (2:1 mix)	1	1.20
27	MCC+DCPA (2:1 mix)	1	0.80
28	MCC+DCPA (2:1 mix)	1.5	1.00
29	MCC+DCPA (2:1 mix)	2	1.20
30	MCC+DCPA (2:1 mix)	2	0.80
31	MCC+DCPA (2:1 mix)	3	1.00

32	MCC+DCPA (2:1 mix)	3	1.20
33	MCC+DCPA (2:1 mix)	3	0.80
34	MCC+DCPA (2:1 mix)	5	1.00
35	MCC+DCPA (2:1 mix)	5	1.20
36	MCC+DCPA (2:1 mix)	5	0.80
37	MCC + 1%MgSt	1	0.80
38	MCC + 1%MgSt	1	0.60
39	MCC + 1%MgSt	1	0.40
40	MCC + 1%MgSt	3	0.80
41	MCC + 1%MgSt	3	0.60
42	MCC + 1%MgSt	3	0.40
43	MCC + 1%MgSt	4	0.40
44	MCC + 1%MgSt	4	0.50
45	MCC + 1%MgSt	5	0.80
46	MCC + 1%MgSt	5	0.60
47	MCC + 1%MgSt	5	0.40
48	MCC + 1%MgSt	5	0.50
49	MCC + DCPA (1:1 mix)	1	0.80
50	MCC + DCPA (1:1 mix)	1	1.00
51	MCC + DCPA (1:1 mix)	1	1.20
52	MCC + DCPA (1:1 mix)	1.5	1.20
53	MCC + DCPA (1:1 mix)	2	0.80
54	MCC + DCPA (1:1 mix)	2	1.00
55	MCC + DCPA (1:1 mix)	2	1.20
56	MCC + DCPA (1:1 mix)	3	0.80
57	MCC + DCPA (1:1 mix)	3	1.00
58	MCC + DCPA (1:1 mix)	3	1.20
59	MCC + DCPA (1:1 mix)	5	0.80
60	MCC + DCPA (1:1 mix)	5	1.00

4.2.3 Determination of Ribbon Density

The ribbon density was obtained by firstly taking samples across the ribbons as shown in Figure 4.3 a).   Each sample was weighed m_s and then coated with a thick layer of microcrystalline wax (M-Coat W-1, Vishay Micro-Measurements U.K.) and weighed again.   The wax was melted on a hot plate and applied onto the ribbon compact using a brush.   The coating kept the ribbon compact dry while it was immersed it in water (Figure 4.3 b)).   Next the density of wax and wax coated ribbon density was determined using the hydrostatic weighing method (explained in section 3.3.4).   The density of the ribbon was then obtained by compensating for the added wax (see Section 9.2 in Appendix 2).
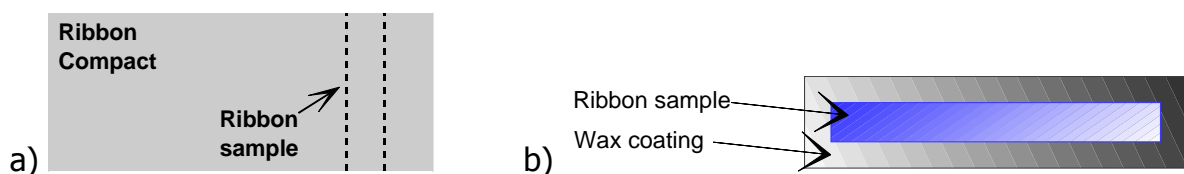
Figure 4.3   a) Schematic of a ribbon compact and the ribbon samples taken for density measurements. b) Schematic of a wax coated ribbon sample.

4.3 Results and Discussions

4.3.1 Experimental Observations

The success of roll compaction was dependent on the type of material being compacted.   It was observed that DCPA could only be roll compacted from 1 – 3 rpm, while the other four materials were compactable up to 5 rpm.   It is not clear why the roll compaction of DCPA at roll speeds higher than 3 rpm failed to produce ribbon compacts.   But from experimental observations, the combination of the DCPA being the heaviest bulk powder material combined with its ease of flowing did

disrupt the powder deaeration during roll compaction at high roll speed. The bulk powder flowed through the roller gap and thus was not compressed into a ribbon.

In addition to that the minimum and maximum roll gaps were also different for each of the powder materials. For example the lubricated MCC (MCC + 1% MgSt) could only be roll compacted at very low roll gaps. At roll gaps higher than 0.8 mm, it was observed to flow through the roll gap without experiencing compression.
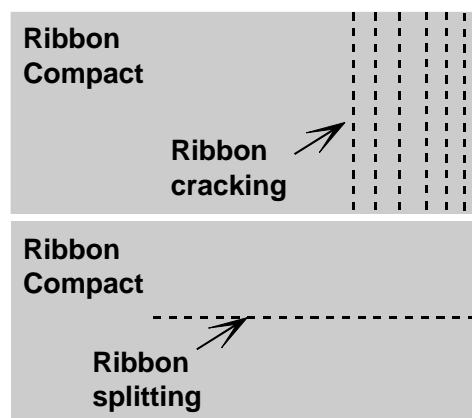
Figure 4.4 Schematic diagram of cracking and splitting along the dotted lines were observed on the DCPA ribbon compact.

Roll compaction of MCC or DCPA at low roll speed and low roll gap caused the roll compactor to jam. The MCC and lubricated MCC ribbon compacts ranged from hard to friable ribbons and no splitting of the ribbons occurred. The DCPA compacts were friable throughout the range of compaction pressures and showed cracking and splitting at higher pressures (Figure 4.4). The pure MCC produced a better ribbon compact than pure DCPA. The ribbon compact of MCC and DCPA 2:1 mixture produced a good compact similar to a ribbon compact from pure MCC.

Table 4.2 Summary of the tablet formulations characteristics and their respective roll compaction processing parameters.

Material Characteristics	DCPA	MCC + DCPA (1:1 mix)	MCC + DCPA (2:1 mix)	MCC	MCC + 1%MgSt
Particle size distribution (d_{16}, d_{50}, d_{84}) in (μm)	4, 14, 27 Narrow PSD	13, 40, 106 Wide PSD	15, 46, 119 Wide PSD	21, 55, 124 Wide PSD	19, 52, 120 Wide PSD
Bulk/Poured density (kg/m^3)	866 ± sd24 (Heavy material)	560 ± sd7 (average weight material)	446 ± sd3 (average weight material)	311 ± sd2 (Light material)	324 ± sd7 (Light material)
Compressibility κ	4.11 ± sd0.13 Least compressible	2.92 ± sd0.04	2.80 ± sd0.02	2.63 ± sd0.02	2.57 ± sd0.04 Most compressible
Tensile Strength (MPa)	0.52 ± sd0.21 Weakest compact	3.25 ± sd0.42	4.27 ± sd0.23	5.90 ± sd0.19	6.13 ± sd0.02 Strongest compact
Flow function	8 (easy flowing)	9 (easy flowing)	7 (easy flowing)	4 (cohesive)	9 (easy flowing)
Roll compaction speed range (rpm)	1 - 3	1 - 5	1 - 5	1 - 5	1 - 5
Roll compaction gap range (mm)	0.8 – 1.4	0.8 – 1.2	0.8 – 1.2	0.7 – 1.2	0.4 – 0.8

4.3.2 Relationships between Ribbon Density, Ribbon Porosity and Average Maximum Pressure

Figure 4.5 shows the expected relationship between increasing roll compaction pressure and the ribbon density (see equation (3.2)). The higher ribbon density for DCPA in Figure 4.5 reflects its higher poured density and particle density as opposed to MCC, which has a lower poured density and particle density. But in general the higher ribbon density was produced at lower roll speeds. Figure 4.6 shows the relationship between ribbon porosity and log P_m. The ribbon porosity generally decreases as the pressure increases for all the formulation. However each

formulation has varying slopes. DCPA has the sharpest decrease in slope and MCC

has the slowest decrease in slope with increasing pressure.
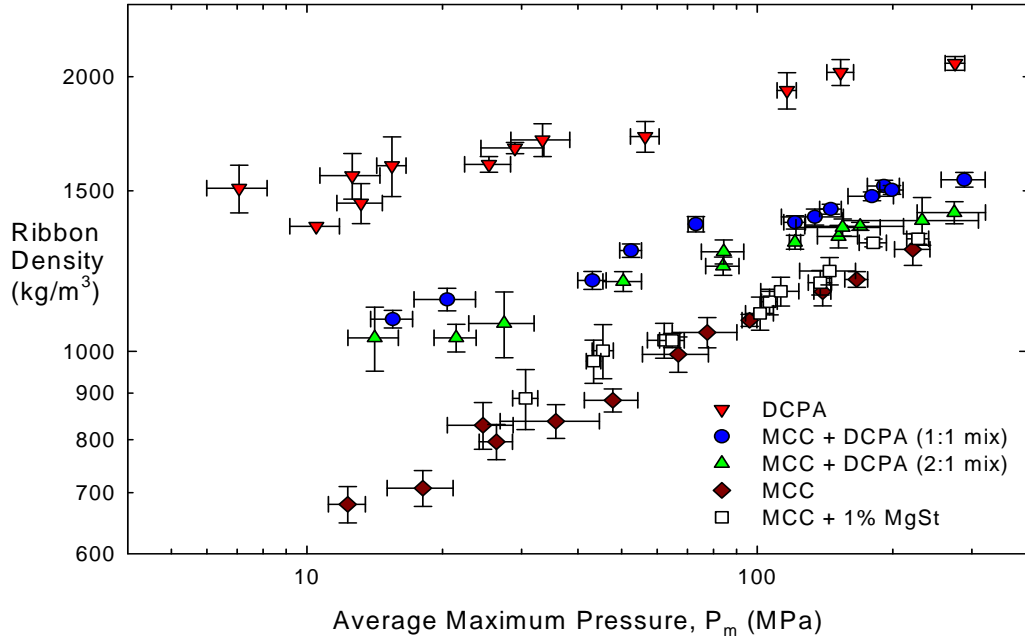
Figure 4.5 Log ρ as a function of log P_m. Results are from roller compaction of excipients at roll speeds of 1- 5 rpm and roll gaps of 0.5-1.4 mm.
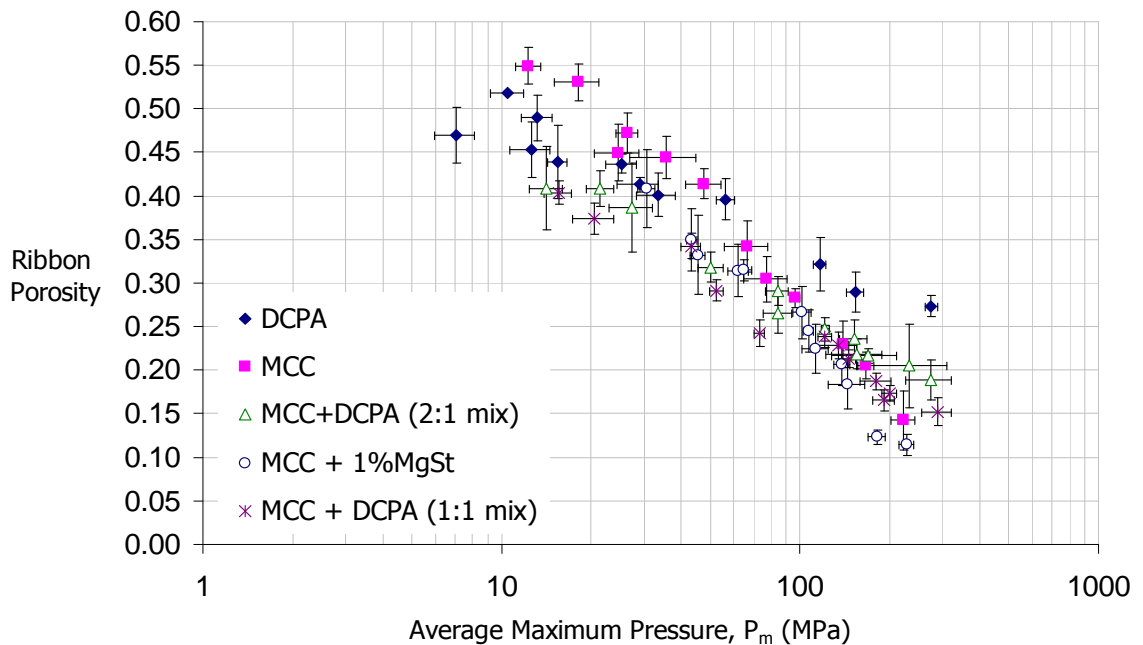
Figure 4.6 Ribbon Porosity as a function of log P_m. Results are from roller compaction of excipients at roll speeds of 1- 5 rpm and roll gaps of 0.5-1.4 mm.

4.3.3 Comparison between Roll Compaction and Uniaxial Compaction

The following figures in this section are to show a comparison between uniaxial compaction and roll compaction.   The uniaxial compaction was conducted at a compression speed of 1mm/sec and to a maximum pressure of 100MPa for five repeats (as shown in Section 3.3.4).   The graphical points for uniaxial compaction shown here is an average of five repeated compressions and the relationship was compared to roll compaction. Generally the uniaxial and roll compaction relationship shows divergence as the compaction pressure increases.   There is little agreement between the two methods in the case of the MCC + 1%MgSt mixture, which may reflect the different wall effects, since MgSt has a strong effect on wall friction.   It is to be noted that the results are quite different from those of Michel (1994) and Bourseul (2001), using inorganic materials, where good correspondence between the two methods was observed.
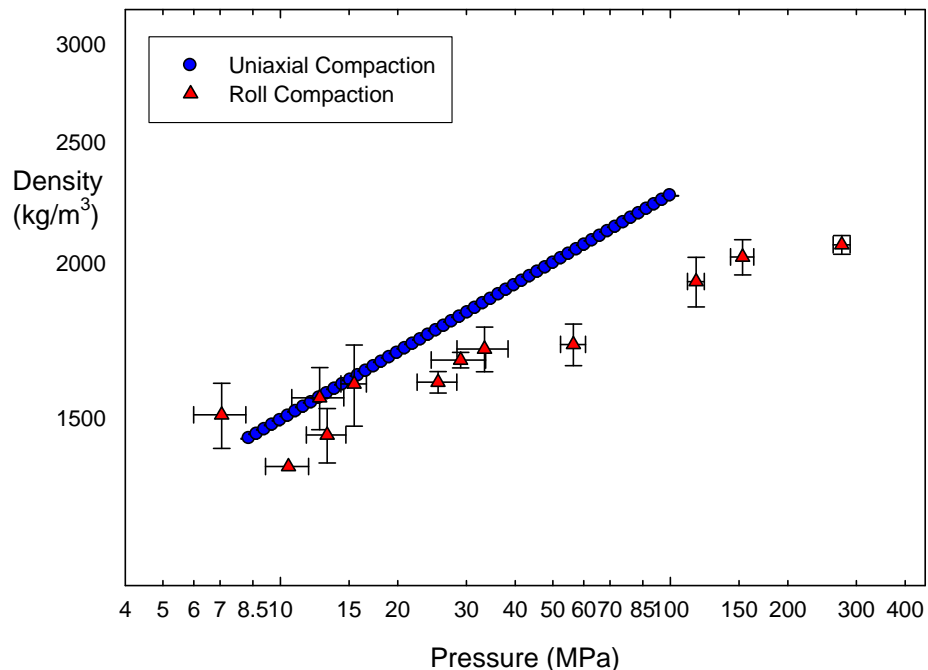
Figure 4.7 Comparison of uniaxial compaction with roll compaction for DCPA.
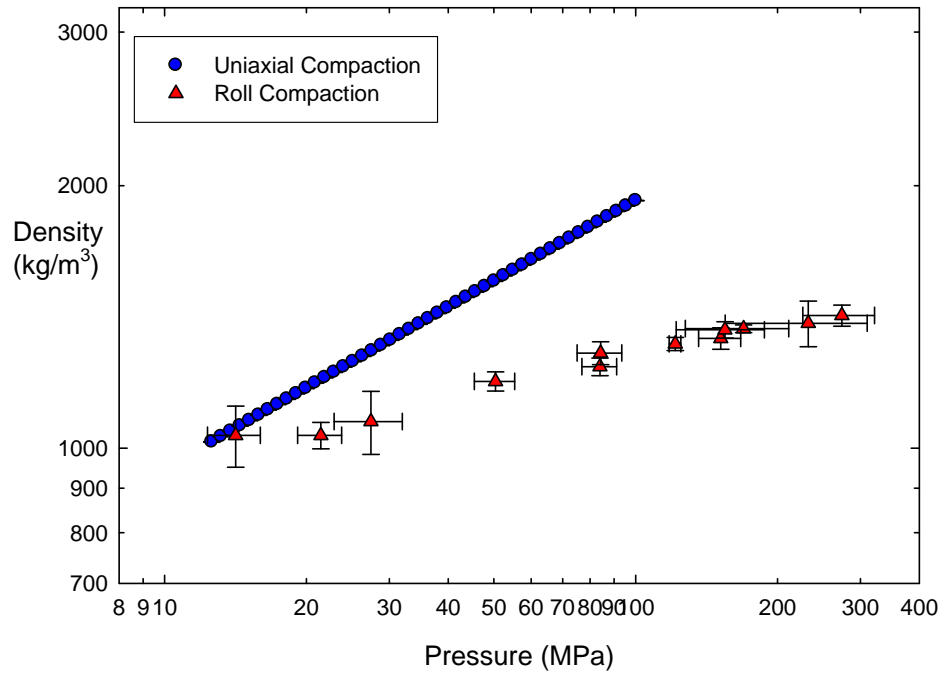
Figure 4.8 Comparison of uniaxial compaction with roll compaction for MCC + DCPA (2:1 mix).
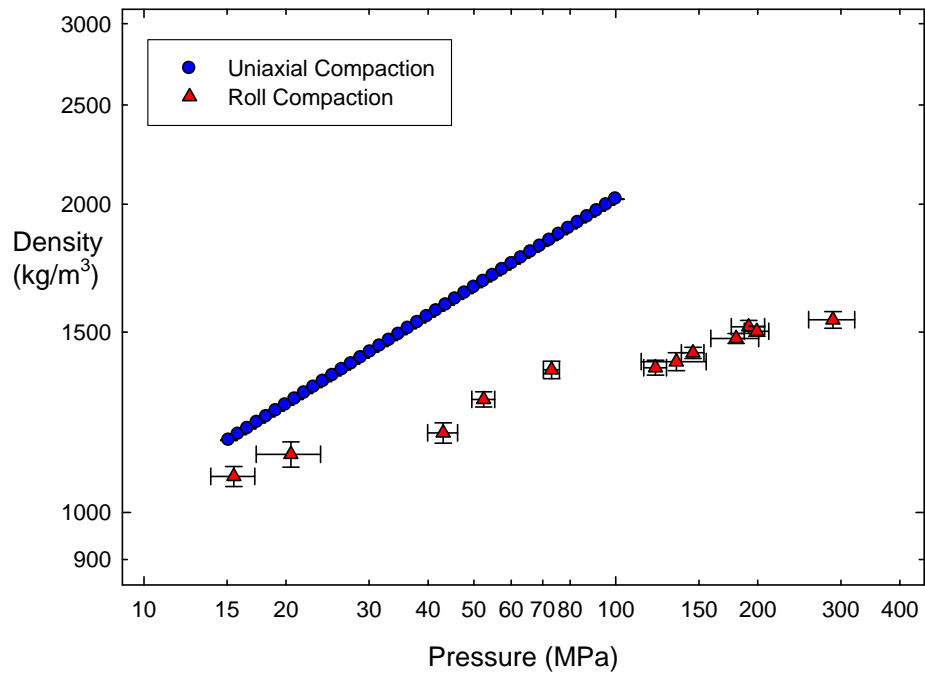
Figure 4.9 Comparison of uniaxial compaction with roll compaction for MCC + DCPA (1:1 mix).
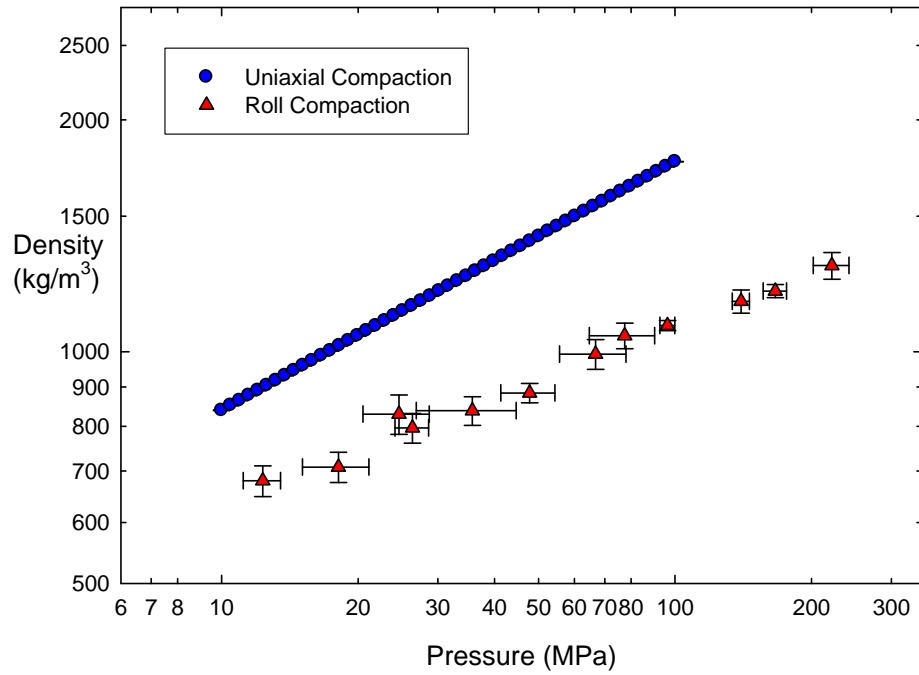
Figure 4.10 Comparison of uniaxial compaction with roll compaction for MCC.
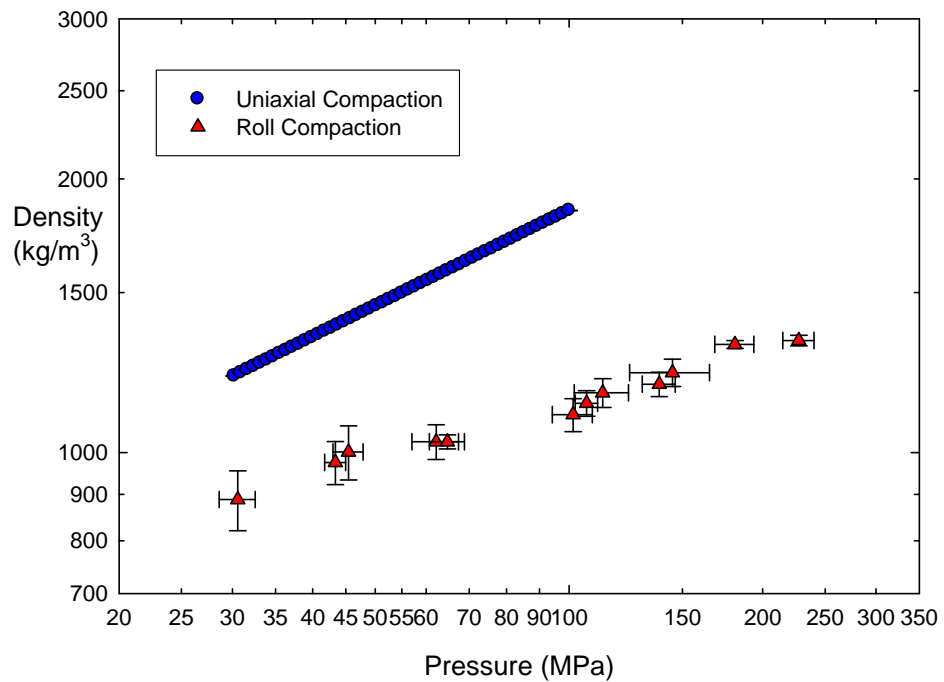
Figure 4.11 Comparison of uniaxial compaction with roll compaction for MCC + 1% MgSt.

4.3.4 The Effect of Roll Compaction Process Parameter Variation on Nip Angle

It is interesting to note that at constant roll speeds and increasing roll gap, the overlapping error in the nip angle may signify that the values may not be different from each other (Figure 4.13 to Figure 4.17). However, the roll compaction pressure profile in Figure 4.12 shows that for a constant roll gap, an increase in roll speed decreases the nip angle and this was repeated for all the powder materials. This trend was also observed by Michel (1994) and Bourseul (2001) in work on roll compaction (with 100 mm diameter rollers) of alumina (SH 150 which has a d_{50} of 11 μm). The nip angle for roll compaction of alumina with roll gaps of 0.85 - 1.2 mm at roll speeds between 0.5 – 1.5 rpm was between 3 and 6 degrees.
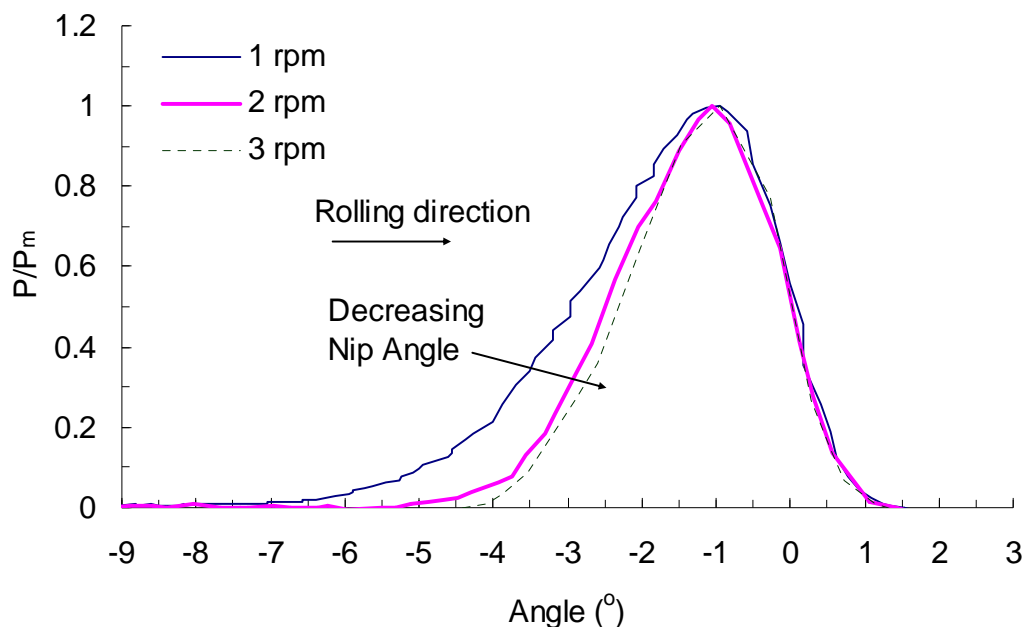
Figure 4.12 Pressure profile graph, where P is pressure and Pm is maximum pressure. This graph is a result of roll compaction of DCPA, at constant roll gap of 1.2 mm and roll speeds of 1 rpm, 2 rpm and 3 rpm.

The roll compaction of all the powder materials was conducted between 1 rpm and 5 rpm except for DCPA (i.e. 1-3 rpm). This is because above 3 rpm the DCPA

powder material flows through the rolls and no compaction was achieved. DCPA has a narrower PSD and smaller d_{50} compared to the other powder materials (Section 3.4.1). Michel (1994) and Boursel (2001) found that increasing the particle size, resulted in varied roll compaction operating roll speed ranges, as in the roll speed range was 0.5 to 1.1 rpm for particle having a d50 of 11 µm and the roll speed range was 5 to 20 rpm for particle having a d50 of 39 µm. Michel concluded that the influence of particle size on the roll speed was related to the conveying capacity of the press in the feeding zone and proposed that Johanson's theory which accounted for the effect of permeability could at least explain the tendency above qualitatively.

Perera (2004) found that nip angle increased if the effective angle of friction was increased, the cohesion of powder was increased, the roll gap was decreased or the roll speed was decreased. The present work was in agreement with all of the above except for roll gap. In this study the nip angle is observed to increase as the roll gap increases. The nip angle was found to increase as the effective angle of internal friction increased from DCPA, MCC + DCPA (1:1mix), MCC +DCPA (2:1) mix to MCC. The lubrication of MCC resulted in the decrease of the powder cohesivity, which in turn decreased the nip angle. Furthermore the lubricated MCC was roll compacted at a lower roll gap compared to pure MCC and all the other materials. It is also interesting to note that as the tensile strength increased the nip angle increased for all the powders except for lubricated MCC. Generally the nip angle increased for all the material as roll speed decreased and roll gap was increased.
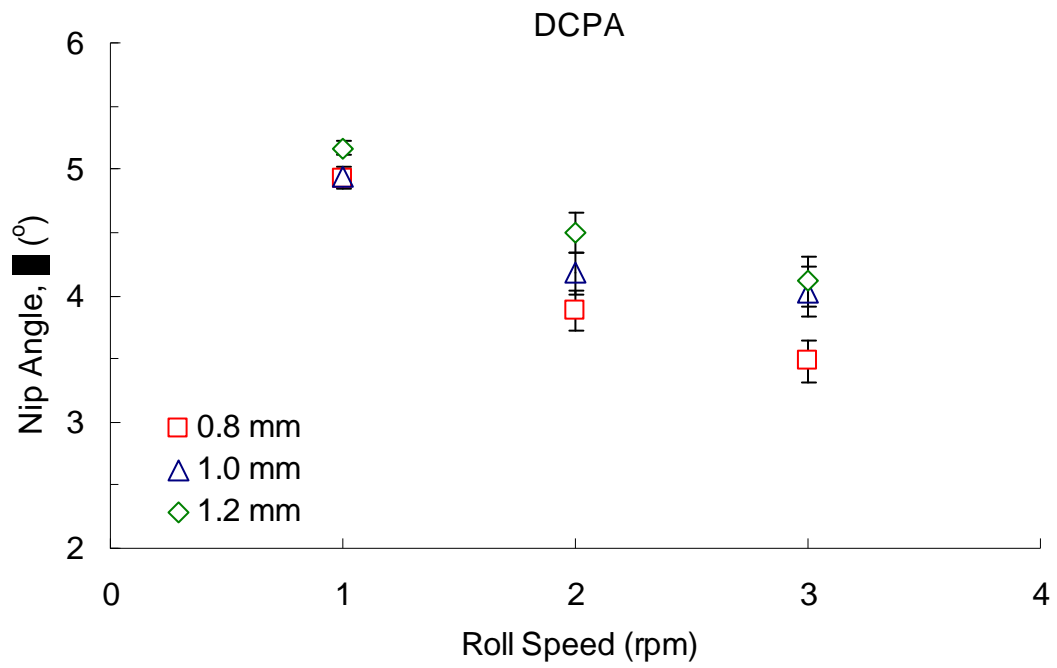
Figure 4.13 Nip angle as a function of roll speed for DCPA.
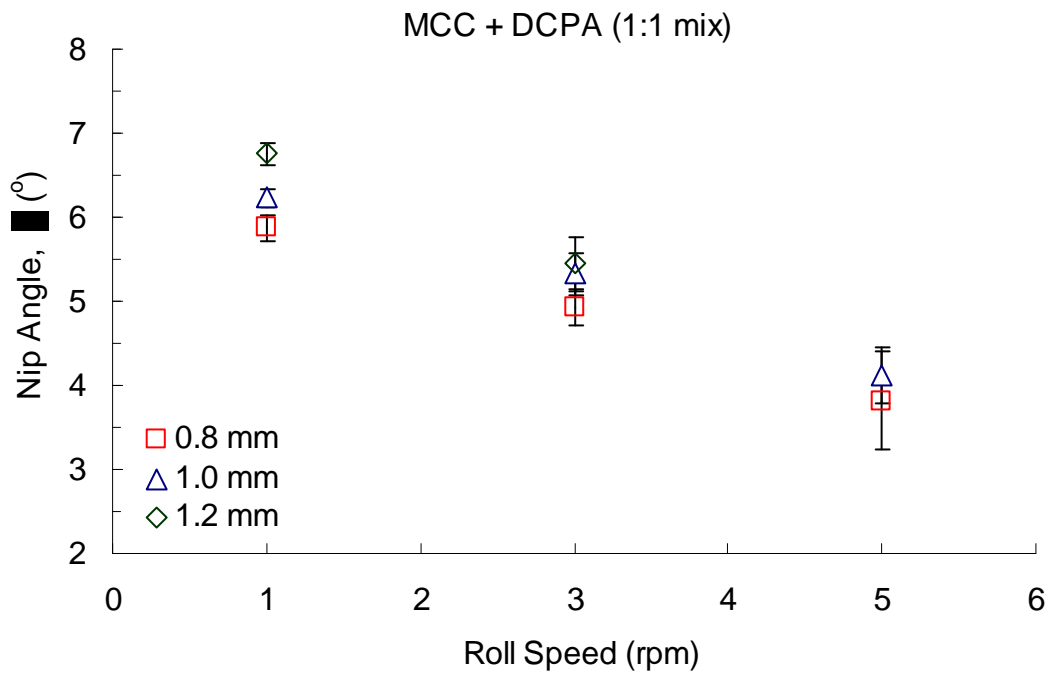
Figure 4.14 Nip angle as a function of roll speed for MCC + DCPA (1:1 mix).
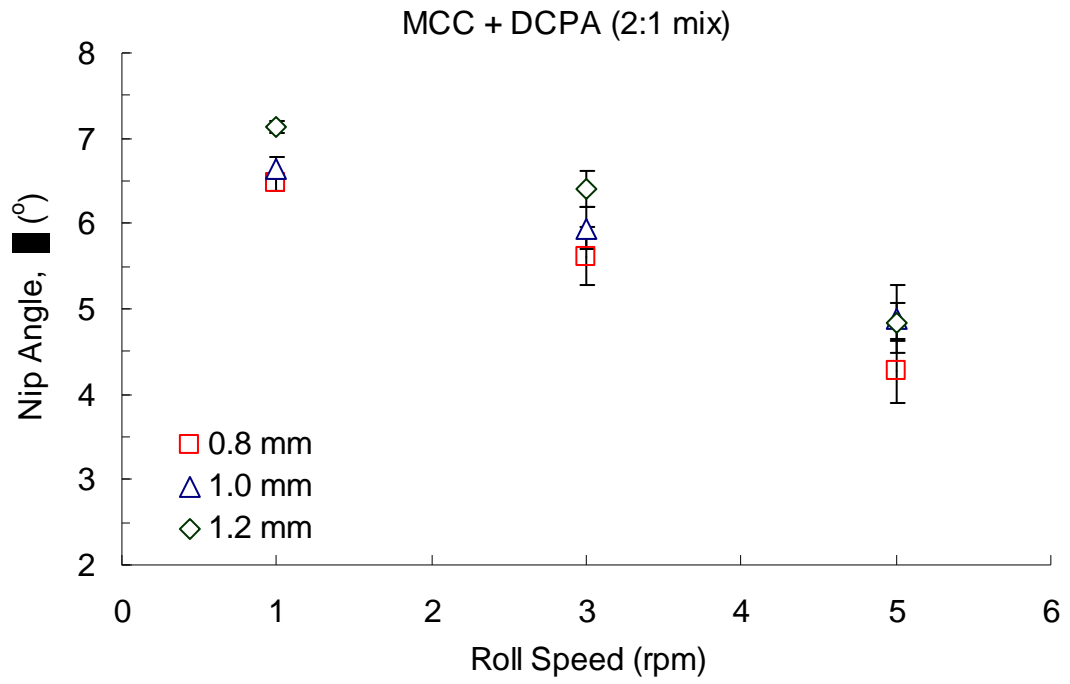
Figure 4.15 Nip angle as a function of roll speed for MCC + DCPA (2:1 mix).
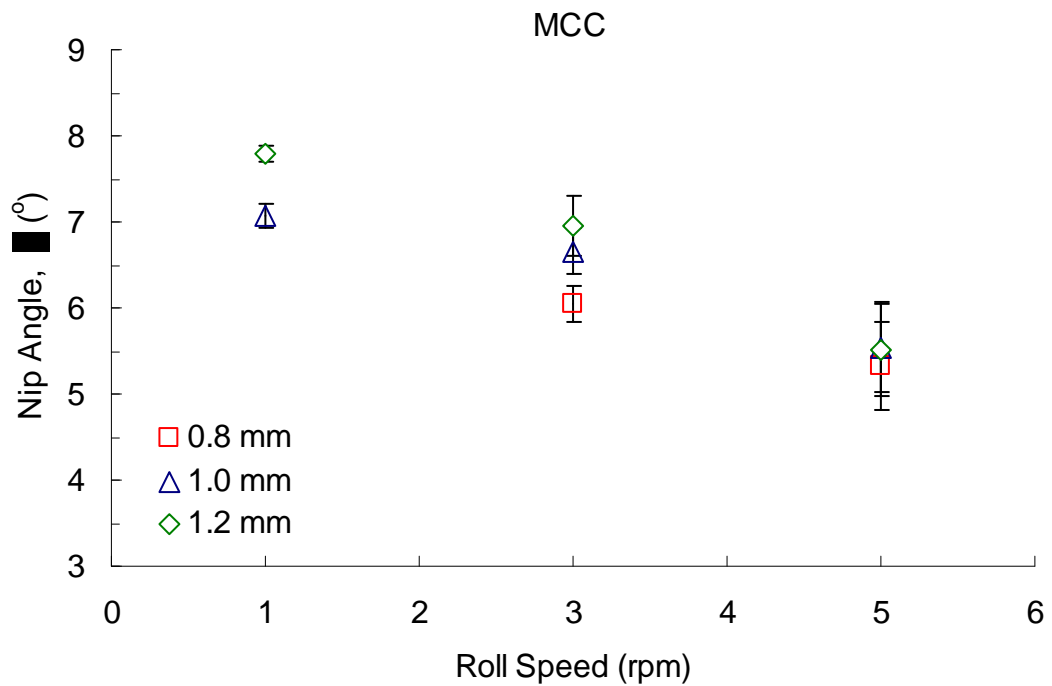
Figure 4.16 Nip angle as a function of roll speed for MCC + DCPA (2:1 mix).
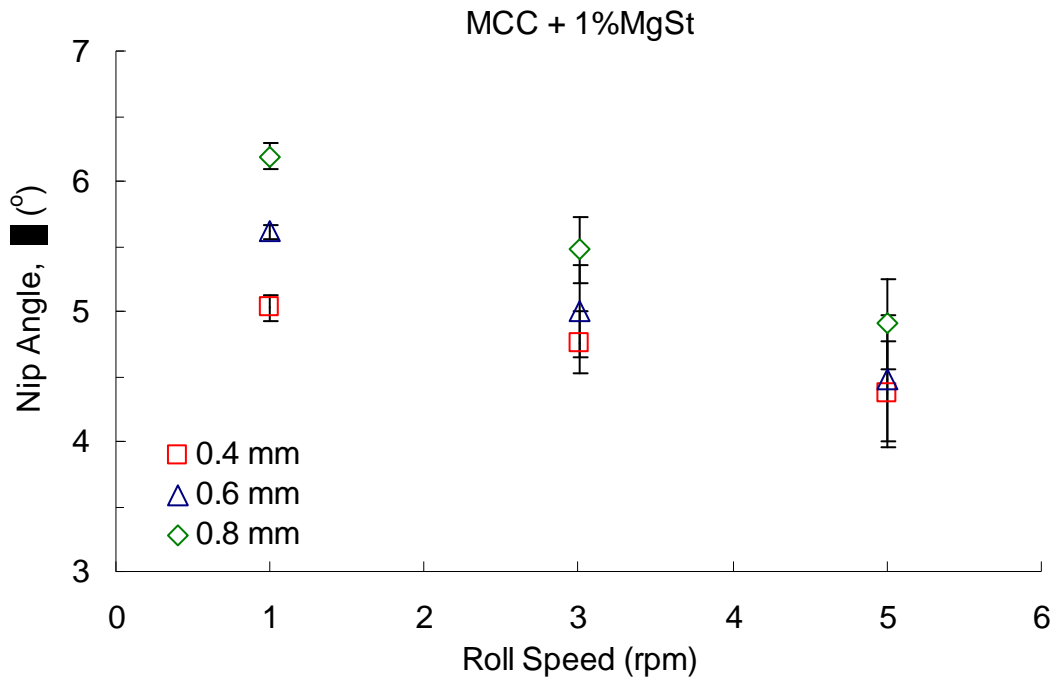
MCC + 1%MgSt

Figure 4.17 Nip angle as a function of roll speed for MCC + 1% MgSt.
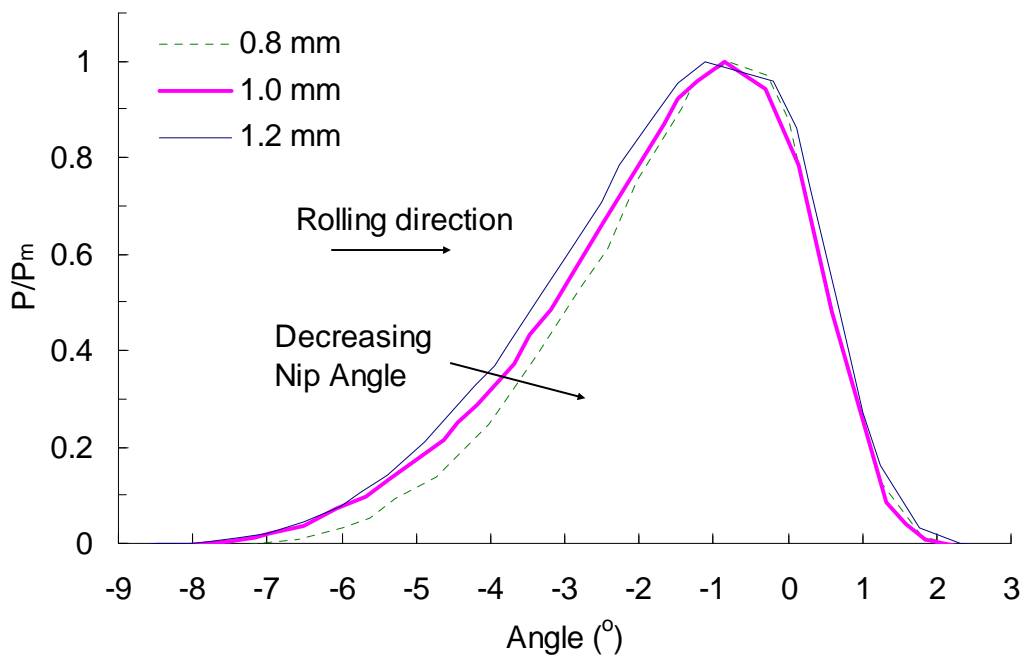
Figure 4.18 Pressure profile graph, where P is pressure and Pm is maximum pressure. This graph is a result of roll compaction of MCC + DCPA (2:1 mixture), at constant roll speed of 3 rpm and roll gaps of 0.8 mm, 1.0 mm and 1.2 mm.

Figure 4.18 shows that for a constant roll speed, an increase in the roll gap increases the nip angle. These trends can be observed for all the samples (Figure 4.23). Bourseul (2001) found that the nip angle increased with roll gap at smaller roll gaps but then showed a maximum. This was not investigated in the present work, as very large roll gaps, give poor compacts.
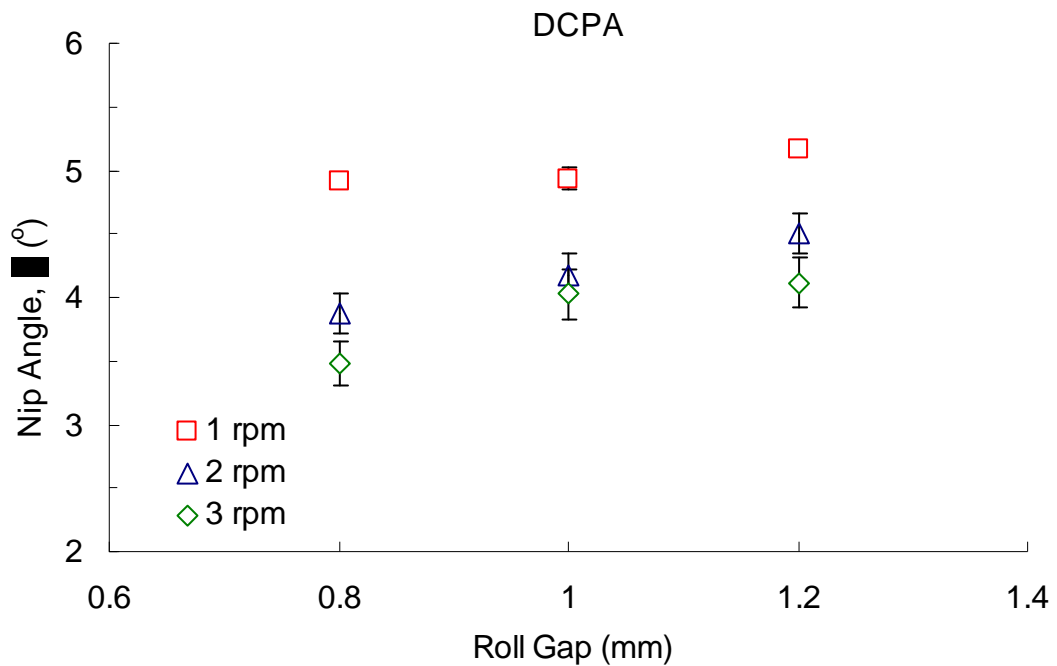
Figure 4.19 Nip angle as a function of roll gap for DCPA.
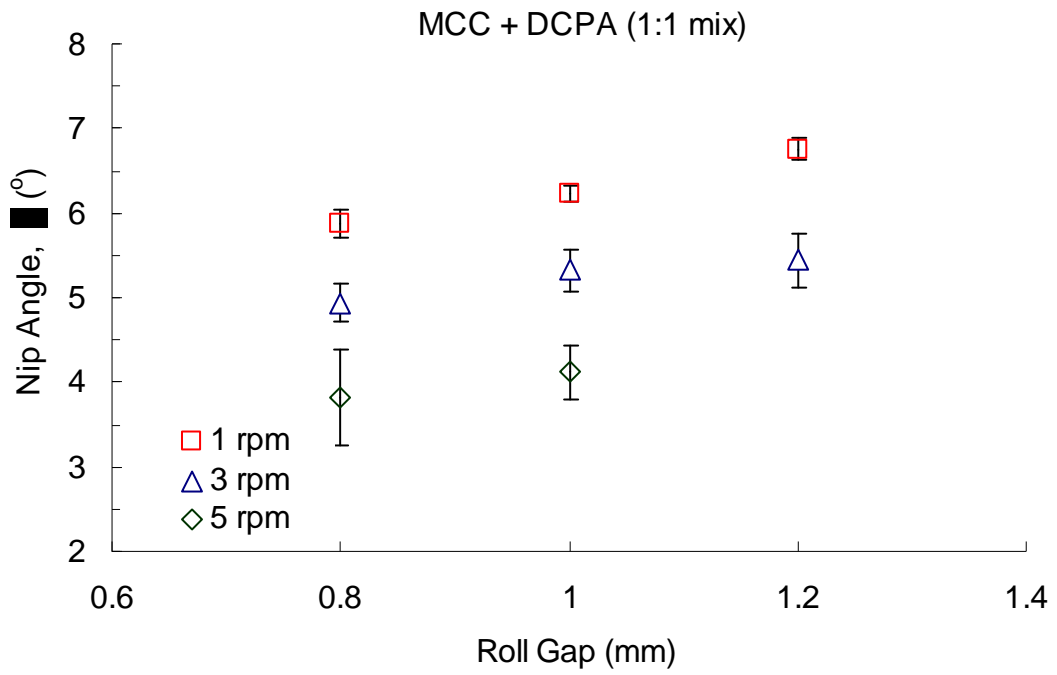
MCC + DCPA (1:1 mix)

Figure 4.20 Nip angle as a function of roll gap for MCC + DCPA (1:1 mix).
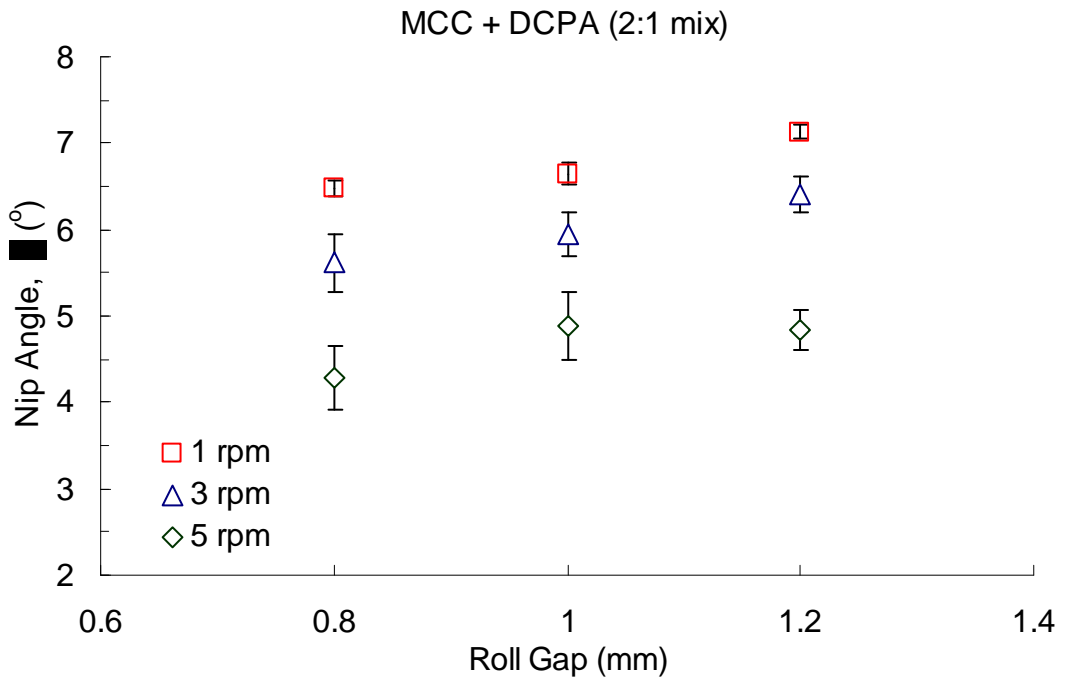
MCC + DCPA (2:1 mix)

Figure 4.21 Nip angle as a function of roll gap for MCC + DCPA (2:1 mix).
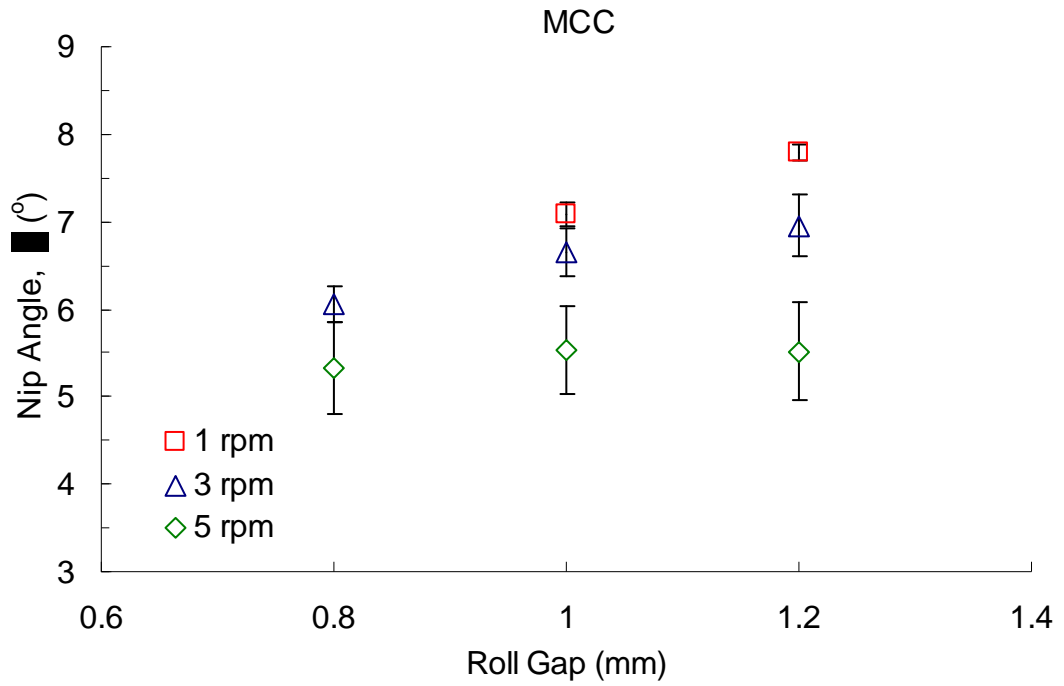
Figure 4.22 Nip angle as a function of roll gap for MCC.
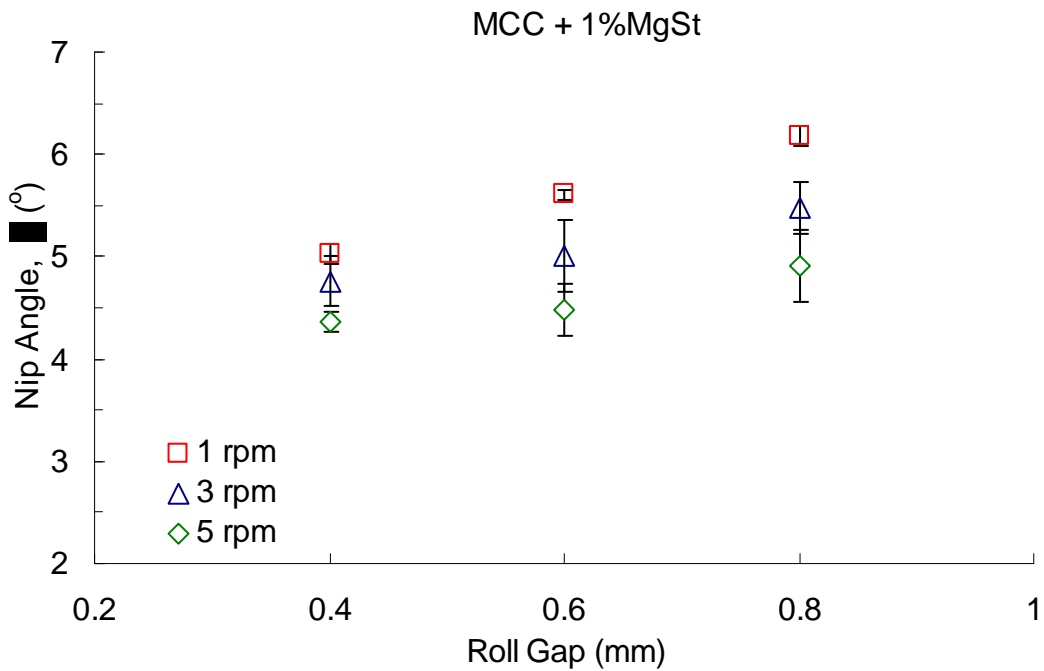
Figure 4.23 Nip angle as a function of roll gap for MCC + 1% MgSt.

4.3.5 The Effect of Roll Compaction Process Parameter Variation on Average Maximum Pressure

The maximum compaction pressure increased at constant roll speed and decreasing roll gap (Figure 4.24).   On the other hand if the roll gap was kept constant and the roll speed was decreased the maximum compaction pressure increased (Figure 4.25).
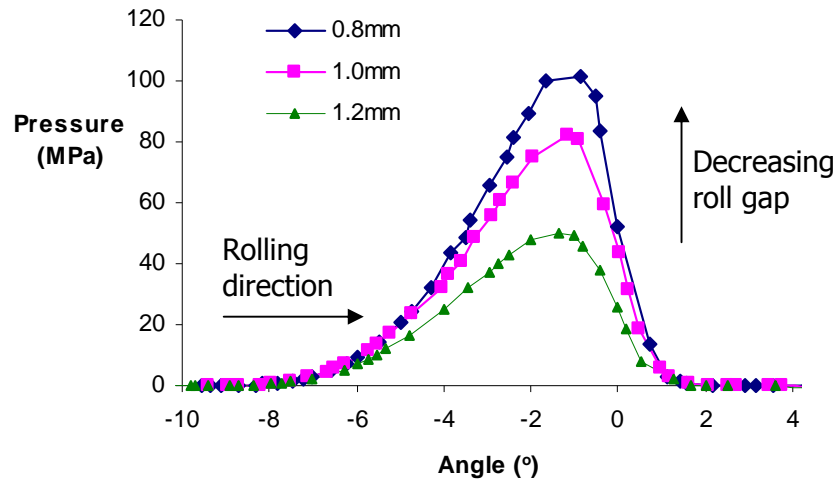
Figure 4.24 Plot of pressure profile. This graph is a result of roll compaction of MCC + DCPA (2:1 mixture), at constant roll speed of 3 rpm and roll gaps of 0.8 mm, 1.0 mm and 1.2 mm.
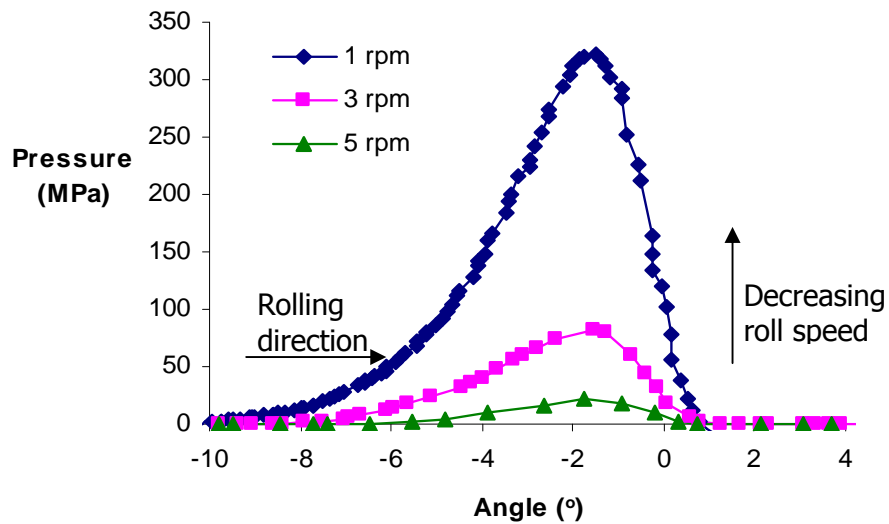
Figure 4.25 Plot of pressure profile. This graph is a result of roll compaction of MCC + DCPA (2:1 mixture), at constant roll gap of 1 mm and roll speeds of 1 rpm, 3 rpm and 5 rpm.

4.4 Summary

In this section a database of roll compaction performance on powder materials at various process parameters was successfully collected, and the effect of roll compaction pressure for each powder material density was found. Roll compaction process parameters was found to differ for each of the powder material, i.e. the roll speed and roll gap range were not the same for all the materials. Furthermore the compaction and density relationship for uniaxial compaction and roll compaction were compared. It was found that for DCPA the roll compaction and uniaxial compaction densities were in agreement at lower pressure and deviated at higher pressure. For the other four powder materials the relationships were not in agreement throughout the pressure range.

In general,

♦ Increase in powder material effective angle of internal friction and cohesivity increased the nip angle.

♦ In the case of powder material lubrication, decrease in powder material cohesivity decreases the operational roll gaps and nip angle.

♦ Decrease in powder material d50 size decreases the maximum roll speed operational range.

♦ Increase in roll speeds or decrease in roll gaps decreases the nip angle.

♦ Increase in the roll gap, decreases the maximum applied pressure in the nip.

♦ Increase in the roll speed, decreases the maximum applied pressure in the nip.

The results from this section will be used in Chapters 5, 6 and 7. In Chapter 5 this results will be compared with predictions from a theoretical model. In Chapters 6 and 7 they are used as inputs to an intelligent software model. The models find key variables from the inputs and find cause-and-effect result to predict outputs.

5 Johanson's Theory

ABSTRACT

Johanson's theory is the most established roll compaction theoretical model. This section gives the basic background on powder mechanics, i.e. the stresses and strain on a body, the Mohr's circle, the ideal coulomb material and Mohr-Coulomb failure analysis. The powder mechanic basics are important to understand the effective yield function (Jenike-Shield criterion). Johanson's theory was developed on the basis of the Jenike-Shield criterion. Although, Johanson's theory is theoretical, it has been validated numerous times in previous research. However its limitation is that it does not account for powder deaeration in the nip region (i.e. it does not account for varying roll speeds) and neutral angle.

This section also shows that Johanson's theoretical prediction of agreed well with the experimental data for all the powder material.

5.1   A Brief Explanation of Johansons's Theory

The Rolling Theory for Granular Solids developed by Johanson (1965) predicts the pressure distribution after the powder is nipped by the rollers. It also predicts the nip angle, which describes the point at which the powder is nipped by the rollers. The roll compacted material is assumed to be isotropic, frictional, cohesive and compressible. This is in parallel to the assumptions in the effective yield function proposed by Jenike and Shield (1959).

Sections 5.1.1 to 5.1.5 provide the fundamental background of the concept of soil mechanics and plasticity on which Jenike and Shield (1959) based the effective yield function. This was then used by Johanson to represent the plane-strain and plane-stress condition of the granular solid between the rolls. Sections 5.1.6 and 5.1.7 present Johanson's mathematical model to predict the pressure profile in the nip region and the nip angle.

5.1.1 Definition of Stress and Strain

Consider a prismatic body, under compression (Figure 5.1). It deforms as shown by the dashed lines. The deformation of the body depends on the force per unit area. The force per unit area is called stress, σ. The change in length divided by the original length is the strain, ε.
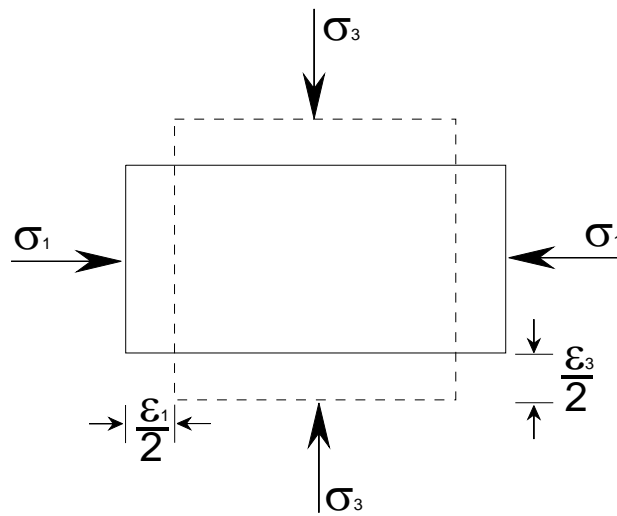
Figure 5.1 Stress and strain

where, σ_1 is the major principal stress in Pa, σ_3 is the minor principal stress in Pa, ε_1 is the strain caused by compression and ε_3 is the strain caused by expansion.

5.1.2 Mohr's Circle

Mohr's Circle is a method for representing the steady state stresses in a material as a function of material. It gives all possible combinations of normal stress σ and shear stress τ at a point. It is only applicable to two-dimensional situations.
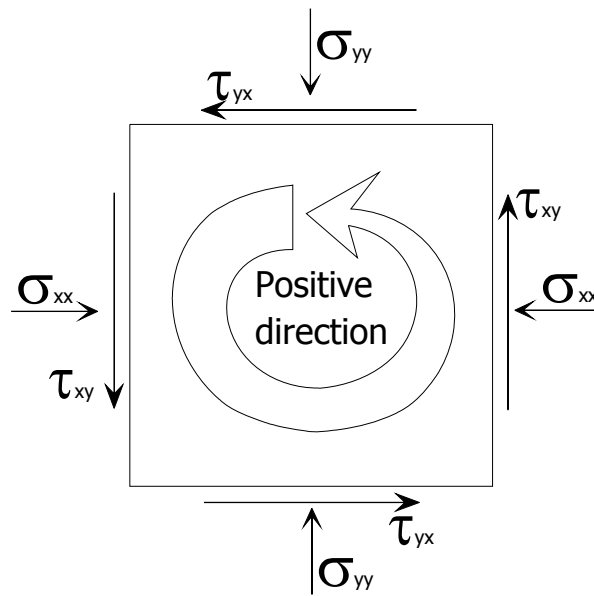
Figure 5.2 Definition of normal and shear stresses.

The compressive stresses are taken to be positive (Figure 5.2). In the use of Mohr's circle, the shear stresses are taken to be positive when acting on an element in an anticlockwise direction. For the element to be stable $\tau_{xy} = - \tau_{yx}$. Since the basic definitions have now been presented, the stress components can be calculated for a wedge-shaped body of unit depth normal to the paper.
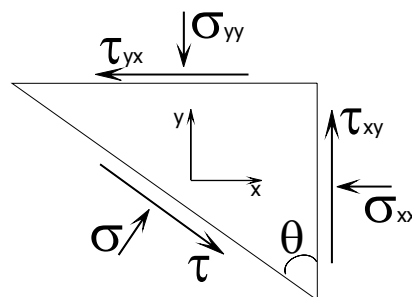
Figure 5.3 Stresses on a wedge-shaped body.

Take the area on the hypotenuse AC to be unity (Figure 5.4). Thus the area of AB and BC faces are cos θ and sin θ respectively.
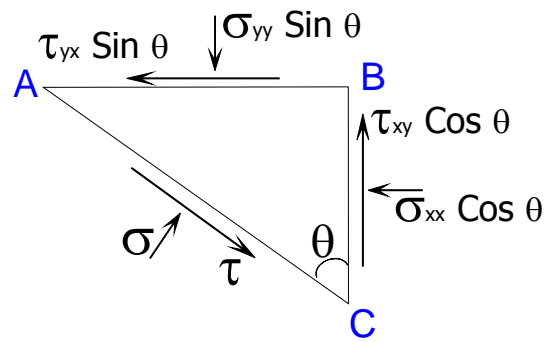
Figure 5.4 Forces on the wedge-shaped body.

Resolving forces in the direction of the normal stress:

$$\sigma = \sigma_{xx} \cos\theta \cos\theta - \tau_{xy} \cos\theta \sin\theta + \sigma_{yy} \sin\theta \sin\theta + \tau_{yx} \sin\theta \cos\theta \qquad (5.1)$$

Resolving the shear forces:

$$\tau = \sigma_{xx} \cos\theta \sin\theta + \tau_{xy} \cos\theta \cos\theta - \sigma_{yy} \sin\theta \cos\theta + \tau_{yx} \sin\theta \sin\theta \qquad (5.2)$$

Substituting $\tau_{xy} = - \tau_{yx}$ and recalling $\cos 2\theta = 1 - 2 \sin^2\theta = 2 \cos^2\theta - 1$, and that $\sin 2\theta = 2 \sin\theta \cos\theta$:

$$\sigma = \frac{1}{2}\left(\sigma_{xx} + \sigma_{yy}\right) + \frac{1}{2}\left(\sigma_{xx} - \sigma_{yy}\right)\cos 2\theta - \tau_{xy} \sin 2\theta \qquad (5.3)$$

and

$$\tau = \frac{1}{2}\left(\sigma_{xx} - \sigma_{yy}\right)\sin 2\theta + \tau_{xy} \cos 2\theta \qquad (5.4)$$

117

Let

$$p = \frac{1}{2}\left(\sigma_{xx} + \sigma_{yy}\right) \tag{5.5}$$

$$R^2 = \left(\frac{\sigma_{xx} - \sigma_{yy}}{2}\right)^2 + \tau_{xy}^2 \tag{5.6}$$

$$\tan 2\lambda = \frac{2\tau_{xy}}{\left(\sigma_{xx} - \sigma_{yy}\right)} \tag{5.7}$$

Thus,

$$\sigma = p + R\cos\left(2\theta + 2\lambda\right) \tag{5.8}$$

$$\tau = R\sin\left(2\theta + 2\lambda\right) \tag{5.9}$$

The equations (5.8) and (5.9) give the definitions of a circle on (σ, τ) axes, and known as Mohr's Circle.
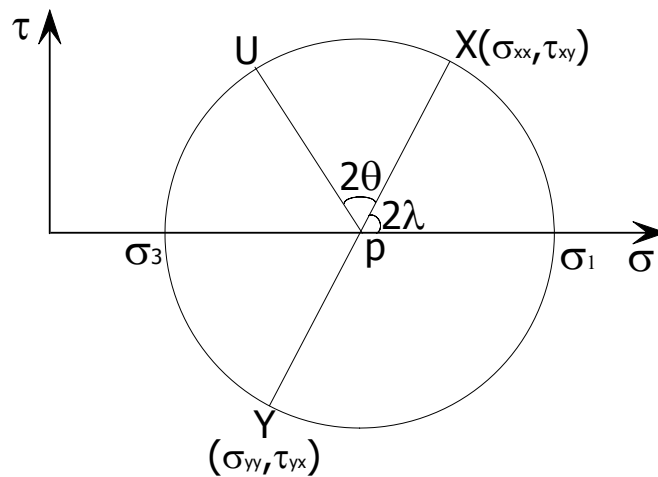
Figure 5.5 Mohr's Circle for stresses

The centre point of the circle (Figure 5.5) is p and the radius is R. The stresses on x and y planes, (σ_{xx}, τ_{xy}) and (σ_{yy}, τ_{yx}) are marked by the points X and Y respectively.

Every point on the circle represents a combination of normal stress σ and shear stress τ on some plane. The two planes of most interest are the planes on which the shear stress is zero, which are known as the major principal plane and the minor principal plane. The stresses acting on these planes are known as the major principal stress σ_1 and the minor principal stress σ_3. Mohr's circle was used together with the Coulomb yield criterion to form the Mohr Coulomb Failure Analysis, as described below.

5.1.3 Ideal Coulomb material

The Coulomb yield criterion describes the behaviour of bulk solids.

$$\tau = \mu\sigma + c \tag{5.10}$$

where μ is the coefficient of friction and is equivalent to the tangent of the angle of inclination φ, and c is cohesion. The ideal Coulomb material is presented as a straight line on Figure 5.6.
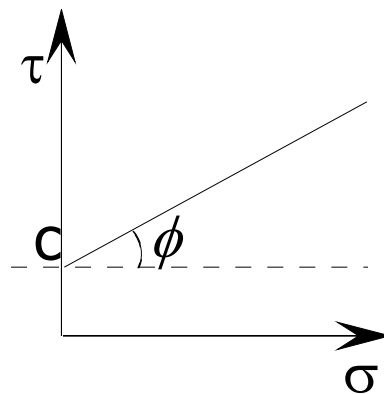
Figure 5.6 Shear stress/normal stress behaviour for an ideal Coulomb Material.

For material with $\tau < \mu\sigma + c$, no motion in the bulk solid will occur. For $\tau = \mu\sigma + c$, a slip plane will be formed. However $\tau > \mu\sigma + c$, cannot occur.

5.1.4 Mohr-Coulomb Failure Analysis

By plotting the Coulomb yield criterion on a Mohr's Circle diagram we have 3 possibilities of powder flow behaviour.
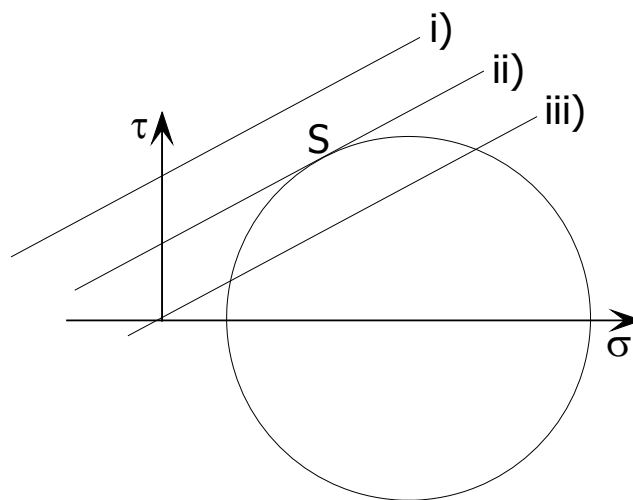
Figure 5.7 Mohr's Circle and the coulomb line

In condition i) the material is in static equilibrium since $\tau < \mu\sigma + c$. No powder flow occurs. Next in condition ii) the powder slips at plane S, a point at which $\tau = \mu\sigma + c$. For all other points on condition ii) $\tau < \mu\sigma + c$ and no powder flow occurs. However in condition iii) the material shows $\tau > \mu\sigma + c$ and this condition is physically impossible.

5.1.5 Effective Yield Function

The roll compaction of pharmaceutical excipients involves continuous shear deformation of the powder material into a solid mass. The material is assumed to be isotropic, frictional, cohesive, compressible and obey the effective yield function proposed by Jenike and Shield (1959).
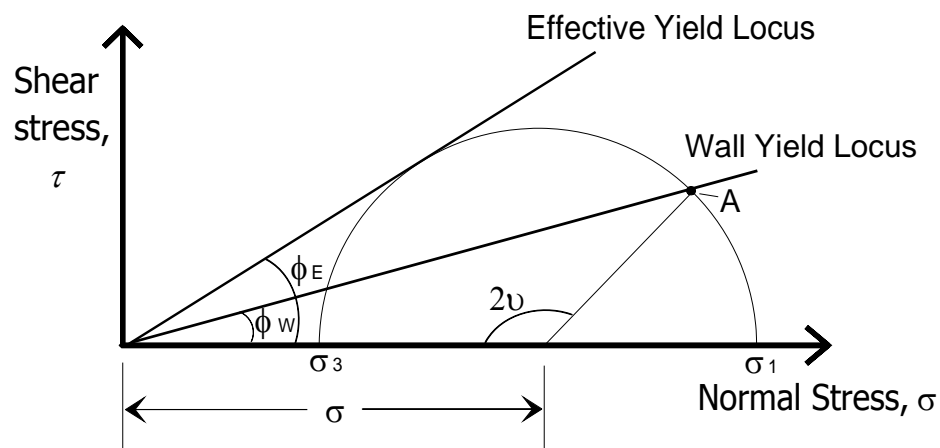
Figure 5.8 Effective and wall yield loci

The effective yield function for plane strain and plane stress condition of the powder material between the rolls can be represented in the Figure 5.8, where ϕ_E is the effective angle of friction, ϕ_w is the angle of wall friction and υ is the acute angle between the tangent to the roll surface and the direction of the major principal stress σ_1 (Figure 5.9).

The effective yield equation (5.11) was combined with the equations of equilibrium in plane strain and axial symmetry to form a hyperbolic system of partial differential

equations. This system can be solved if provided with sufficient boundary conditions (Johanson, 1964).
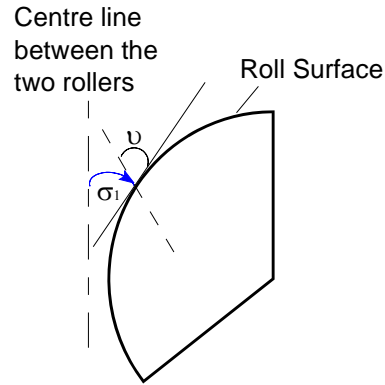
Figure 5.9 Acute angle υ

$$\sin \phi_E = \frac{\sigma_1 - \sigma_3}{\sigma_1 + \sigma_3} \tag{5.11}$$

The condition of slip on the surface of the rollers gives one boundary condition. At the roll surface, the relationship between the tangential and the normal forces is given by the wall yield locus (Figure 5.8). The angle of wall friction ϕ_w describes this locus but, for calculation purposes, it is more convenient to use the acute angle υ. The shear stress and the normal stress at the surface of the rollers are described by Point A in Figure 5.8.

$$2\upsilon = \pi - \arcsin\left[\frac{\sin \phi_w}{\sin \phi_E}\right] - \phi_w \tag{5.12}$$

The effective angle of friction ϕ_E and the angle of wall friction ϕ_w are experimented inputs.

5.1.6 Determination of Pressure Profile in the Nip Region
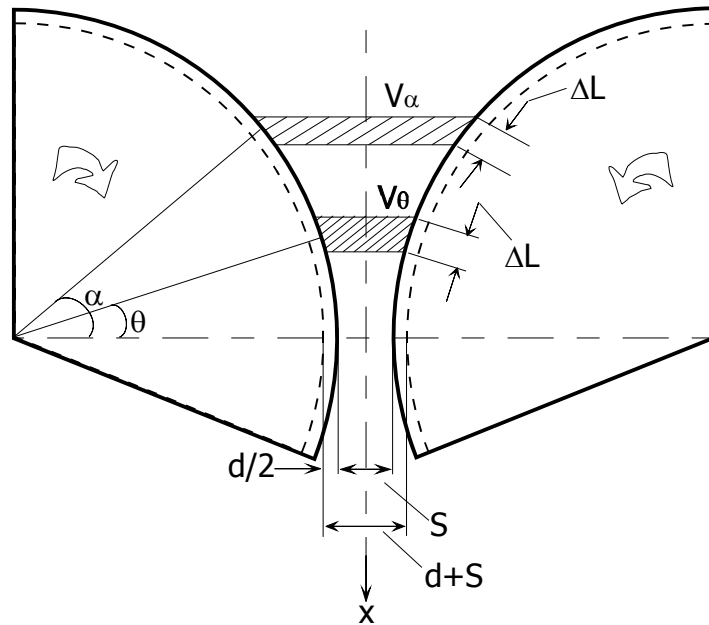
Figure 5.10 Region of nip in a roll press

The material in the nip region does not experience slip along the roll surface. It must be compressed to the final roll gap dimension. This means volume V_α must be compressed to volume V_θ between the same arc-length segments. Continuity requires that the bulk densities ρ_α and ρ_θ in volumes V_α and V_θ be related by,

$$\frac{\rho_\alpha}{\rho_\theta} = \frac{V_\alpha}{V_\theta} \qquad (5.13)$$

Pressure σ_θ at any angle $\theta < \alpha$ can be determined as a function of the pressure σ_α at angle $\theta = \alpha$ by a pressure density relationship. Experimental evidence such as conducted in Section 4.3, indicates that for increasing pressures the log density is a linear function of log pressure. This relationship is presented by

$$\frac{\sigma_\alpha}{\sigma_\theta} = \left(\frac{\rho_\alpha}{\rho_\theta}\right)^\kappa \qquad (5.14)$$

Substituting density with volume we get

$$\frac{\sigma_\alpha}{\sigma_\theta} = \left(\frac{V_\alpha}{V_\theta}\right)^\kappa \qquad (5.15)$$

where κ is a material property, the Compressibility value. This is taken as constant and in practice this is an approximation, as shown in Section 3.4.4.

Volume V_θ in the arc-length segment ΔL is

$$(5.16)$$

$$V_\theta = \Delta L W [S + D(1 - \cos\theta)]\cos\theta$$

Where D is the diameter of the rollers and W is the roll width. Combining the equations (5.14) and (5.15) gives

$$\sigma_\theta = \sigma_\alpha \left[\frac{\left(1 + \dfrac{S}{D} - \cos\alpha\right)\cos\alpha}{\left(1 + \dfrac{S}{D} - \cos\theta\right)\cos\theta}\right]^\kappa \qquad (5.17)$$

Hence the pressure distribution between the rolls can be obtained provided the nip angle is known. Equation (5.17) is only applicable for angles $\theta \le \alpha$.

5.1.7 Determination of Nip Angle

The principle of the determination of nip angle is that at that point the non-slip pressure gradient is the same as the pressure gradient for slip. Consider the pressure gradient dσ/dx, where x is the vertical coordinate in Figure 5.10. By assuming that slip occurs along the roll surface, the pressure gradient dσ/dx would be determined by equation (5.12) and the Jenike-Shield yield criterion (i.e. effective yield function). By first order approximation we obtain:

$$\frac{d\sigma}{dx} = \frac{4\sigma\left(\frac{\pi}{2} - \theta - \upsilon\right)\tan\phi_w}{\frac{D}{2}\left[1 + \frac{S}{D} - \cos\theta\right]\left[\cot(A - \mu) - \cot(A + \mu)\right]} \tag{5.18}$$

where
$$A = \frac{\theta + \upsilon + \frac{\pi}{2}}{2}$$
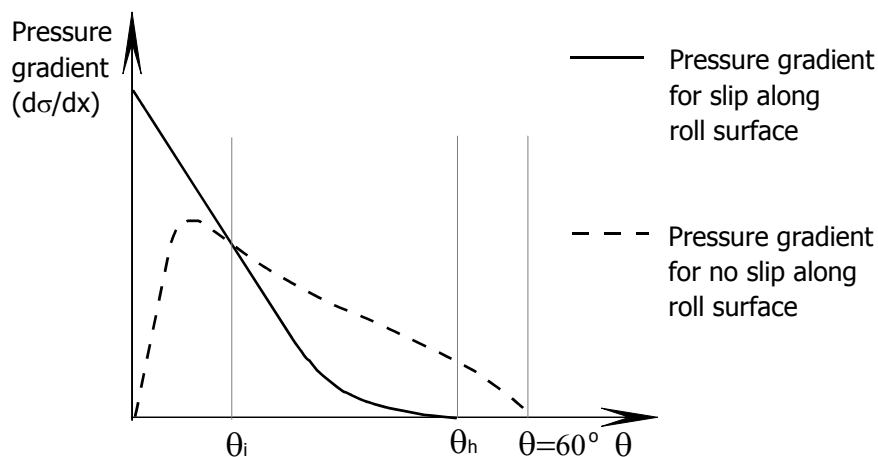
Figure 5.11 Vertical pressure gradient versus angular position in roll bite. θ_h is the angle which describes the height of powder material in between the rolls.

The solid line in the figure above is a typical pressure gradient profile given by the Equation (5.18) for the condition of slip along the roll surface. Near $\theta = \theta_h$, the pressure gradient $d\sigma/dx$ is zero. However when slip does not occur along the roll surface, pressure σ is given by equation (5.17) and the pressure gradient is

$$\frac{d\sigma}{dx} = \frac{\kappa\sigma_\theta\left(2\cos\theta - 1 - \dfrac{S}{D}\right)\tan\theta}{\dfrac{D}{2}\left[\dfrac{d}{D} + \left(1 + \dfrac{S}{D} - \cos\theta\right)\cos\theta\right]} \tag{5.19}$$

This pressure gradient is zero for angles $\theta = 0^{\rm o}$ and for $\theta = 60^{\rm o}$. Johanson (1965) proposed that the intersection of the two pressure gradient curves is a point at which the nip angle occurs. Hence $\theta_i = \alpha$ and can be found by equating equations (5.18) and (5.20).

$$\frac{4\sigma\left(\dfrac{\pi}{2} - \theta - \upsilon\right)\tan\phi_w}{\dfrac{D}{2}\left[1 + \dfrac{S}{D} - \cos\theta\right]\left[\cot(A - \mu) - \cot(A + \mu)\right]} = \frac{\kappa\sigma_\theta\left(2\cos\theta - 1 - \dfrac{S}{D}\right)\tan\theta}{\dfrac{D}{2}\left[\dfrac{d}{D} + \left(1 + \dfrac{S}{D} - \cos\theta\right)\cos\theta\right]} \tag{5.20}$$

Equation (5.20) is used to calculate the nip angle. This nip angle is then used to give a pressure profile within the nip region using equation (5.17). This pressure profile will then allow for an approximation of a new nip angle value. Both these nip angle values will be compared to the nip angle obtained from the roll compaction experiments (Chapter 4). The Compressibility value κ, angle of wall friction ϕ_w and effective angle of friction ϕ_E were found and reported in Sections 3.4.4 and 3.4.5.

Table 5.1 shows a summary of the material properties used as inputs to this calculation.

Table 5.1 Material properties

Powder Material	Effective Angle of Friction, ϕ_E (°)	Angle of wall friction, ϕ_w (°)	Compressibility κ (Compressed at 100 MPa at a speed of 1 mm/sec)
DCPA	36.3	19.1	4.11
MCC + DCPA (1:1 mix)	41.3	17.7	2.92
MCC + DCPA (2:1 mix)	44.0	10.8	2.79
MCC	46.3	9.9	2.63
MCC + 1% MgSt	46.7	7.8	2.57

5.2 Johanson's Prediction of Nip Angle
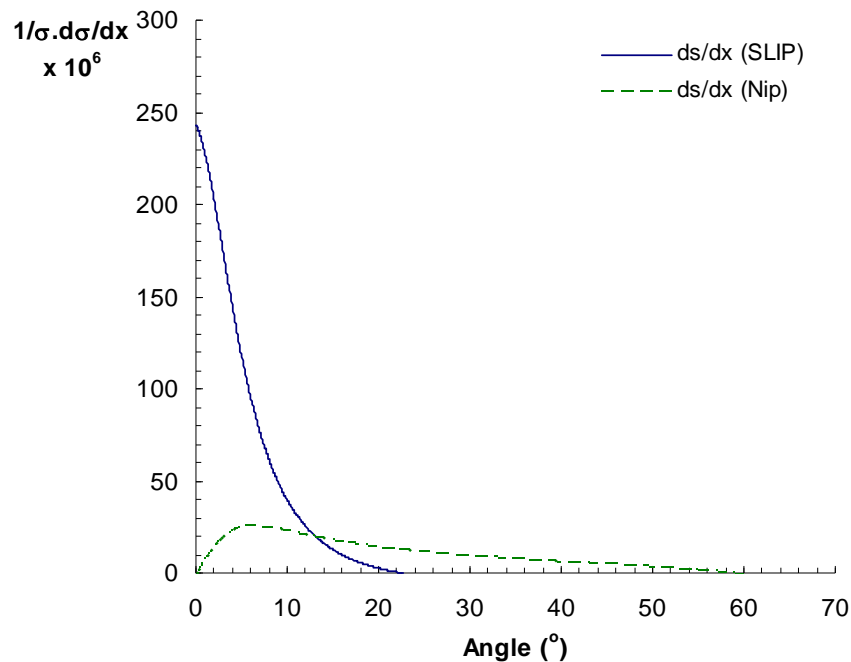
Figure 5.12 Pressure gradients against angular position for MCC + DCPA (1:1 mix) roll compacted at 1 rpm and 1.2 mm.

The nip angle found from solving the equation (5.20) was fed into equations (5.18) and (5.19) to produce the pressure gradient profiles shown in Figure 5.12.

Figure 5.13 presents a comparison of Johanson's pressure profile prediction with experimental results at various roll gaps and at a roll speed of 1 rpm within the nip region. The lowest roll speed was used for comparison as Johanson's theory does not account for roll speed and neglects deaeration effects. By choosing the slowest speed the deaeration effects on the experimental results are minimised. Good agreement between theory and experiment is demonstrated.
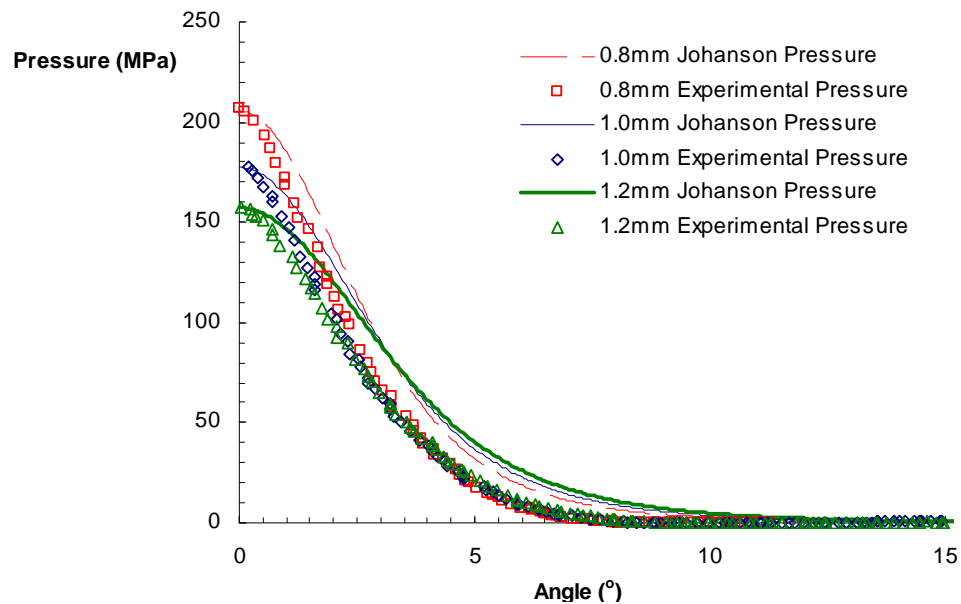
Figure 5.13 Experimental and Johanson Pressure Profile within the nip region. The powder material used in this is MCC + DCPA (2:1 mix) roller compacted at 1 rpm.

Here, the nip angle found from equation (5.20) was used in equation (5.17) to produce a pressure profile. In previous work by Bindhumadhavan (2004), it was found that the pressure profile needed to be fitted to the maximum pressure achieved in the experiment. This approach was also adopted in the present work. Johanson's approach does not predict the neutral angle, which as observed does not

always occur at $\theta=0$. For purposes of comparison, therefore the angle at which the maximum pressure occurs is taken here to be zero.

A nip angle was approximated from Johanson's pressure profile above by using the tangent to the steepest section of the profile. This is done because this is the same method which is used to get the nip angle from experimental pressure profiles. The theoretical nip angle from equation (5.20) is always larger. In the subsequent graphs it will be called 'Approximation using Johanson Pressure Profile'. Figure 5.14 presents a comparison of the experimental nip angles with the prediction from Johanson's model for all five powder materials. For DCPA and 1:1 mixture of MCC and DCPA the prediction from equation (5.20) deviates hugely from the experimental nip angle i.e. double the experimental nip angle. However for 2:1 mixture of MCC and DCPA, MCC and MCC with 1%MgSt, the deviation decreases as S/D ratio increases. Subsequently the 'Approximation using Johanson Pressure Profile' showed very small deviation from the experimental nip angle, to within 1° to 2°. Overall the theoretical prediction agreed well with the experimental data for all the materials.
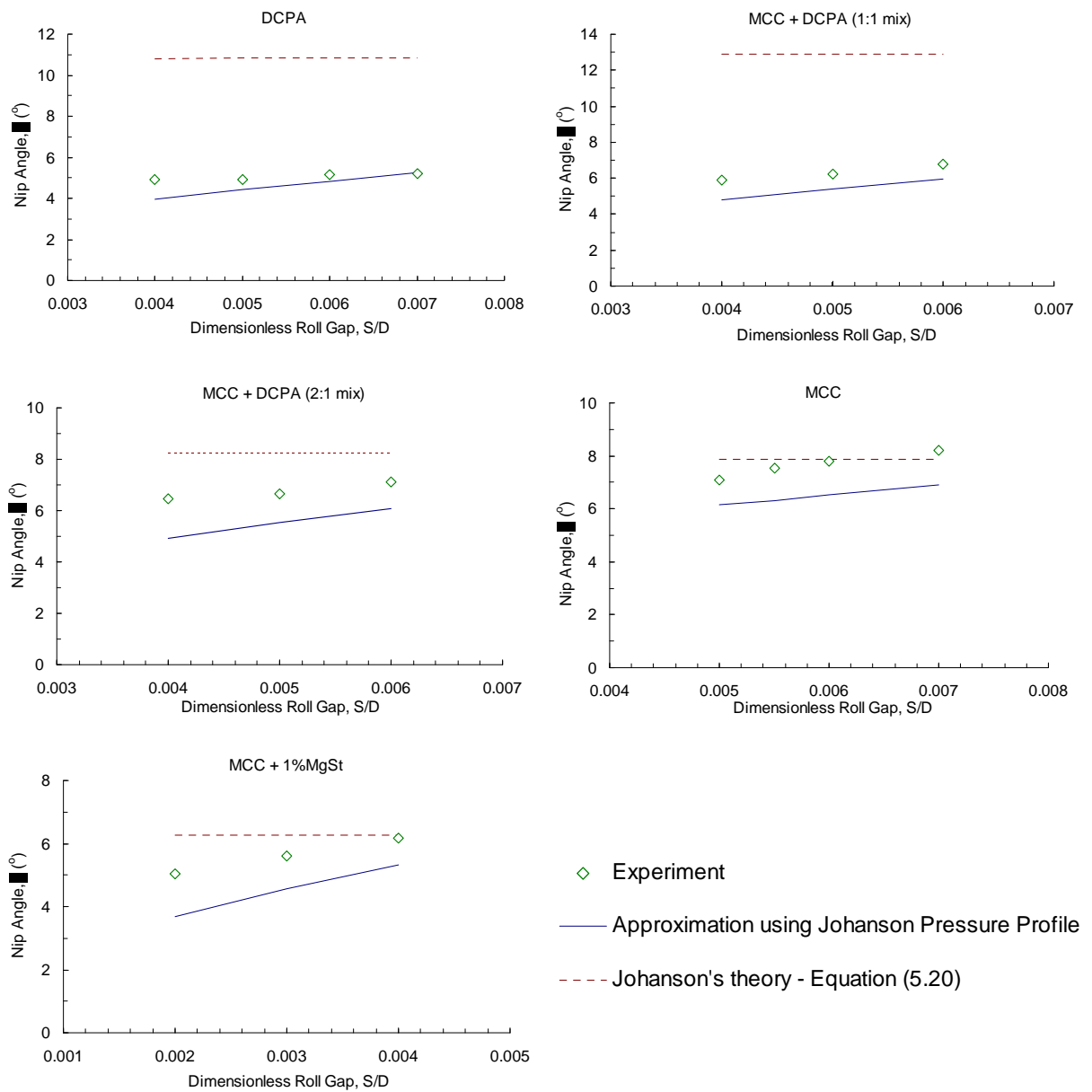
Figure 5.14 Variation of nip angle with dimensionless roll gap for powder material compacted at 1 rpm.

5.3 Concluding Remarks

This section has presented the powder mechanics background required to understand the theory behind the Johanson's theory. This section also shows that

Johanson's theory prediction agreed well with the experimental data for all the powder material. However it must be noted that the Johanson's theory does not account for powder deaeration in the nip region and the neutral angle. Hence for the purpose of comparison with experimental results, the experimental results were produced from roll compaction at 1 rpm roll speed and varying roll gaps.

Furthermore the experimental pressure profile was normalised to zero angle at the maximum pressure. Note that the Compressibility κ values used here were obtained from experiments conducted at 100 MPa compression pressure and 1 mm/sec compaction speed because Compressibility κ varies with maximum uniaxial compaction pressure. The theoretically predicted nip angle in this section was compared to the values predicted by the intelligent software in Section 7.

6 Intelligent Software: Formrules – Finding rules

ABSTRACT

FormRules software is specific example a 'data-mining' intelligent software program which produces simple linguistically-expressed rules in the form of *IF (condition 1) AND (condition 2) AND (condition3), THEN (conclusion 1, with confidence factor x)*. It highlights the key variables which are important to a specific output property. These key variables will then be used as inputs in the *INForm* software training. It is important, however, for the user to evaluate the practicality and relevance of the inputs used and highlighted in the software because the software is not designed to do this. If this aspect is neglected unrealistic relationships may be predicted. Ultimately this software is able to process a large amount of information quickly and efficiently; as soon as the database is set up inside *FormRules*, different types of trials are possible.

This section describes how the software was trained to develop models and presents the results of training 11 trials which were conducted with varying types of input variables.

FormRules is based on neurofuzzy logic which combines adaptive (learning) capabilities of ANN and the linguistic capabilities of fuzzy logic (Jang *et al.*, 1997). Neurofuzzy logic is a combination of ANN and fuzzy logic and was explained in Section 2.2.4. *FormRules* software is used in sequence with *INForm* software. Firstly, *FormRules* is used to determine the key variables that relate the properties of the roll compacted ribbons to the initial formulation in the form of rules. Having established this, the *INForm* software is then used to provide quantitative predictions (Lindberg and Colburn, 2004). *INForm* software is presented in Chapter 7.

6.1 Practical Application

The main aim of using *FormRules* software is to determine the key variables relating initial formulation to roll compaction product properties and to discover the minimum number of inputs required to produce satisfactory predictive model.

6.1.1 Training FormRules

The powder characteristics of the formulation are listed in Table 3.11. These powder characteristics along with roll compaction process parameters and outputs were then fed into the *FormRules*. Trials were conducted to relate the powder characteristics and roll compaction process parameter to the output. The trials were also used to investigate the minimum number of inputs required to develop a good model. A general *FormRules* software training flow chart is shown in Figure 6.1.

See Appendix 4 for brief user-software interface tutorial. For this investigation the inputs and outputs listed in Figure 6.2 were used in the *FormRules* software training. Note that the average maximum pressure and the average nip angle are intermediate output properties of the roll compaction process. The *FormRules* manual (Intelligensys, 2002) states that the minimum amount of data required by the neurofuzzy system to produce a reliable model is 2 or 3 data sets per input. In this case a sufficient number of data sets supplied: 64 data set for, 18 inputs, which equates to 3.5 data sets per input.

After the data were imported into the software, the inputs and outputs of the model were set. These are inputs and outputs in the input space and output space of Figure 2.9. The first model training was conducted with default settings. If the f-ratio falls below 4 and the trained data set R^2 is less than 80 hence the model requires retraining. Retraining involves changing the training parameters (Section 9.4.3, Appendix 4) to fulfil the requirement of a good model. A good model is a model which gives an f-ratio of more than 4 and a trained data set R^2 of more than 80 (Section 6.1.2). Following successful model training, the neurofuzzy results diagram can be displayed to highlight the key variables found for each output (Figures 6.3 to 6.6). Then the linguistically-expressed rules (e.g. Table 6.1) can be obtained in a report form so as to aid in understanding the specific relationships between the inputs and the outputs.

```
         ┌─────────────┐
         │  Entering   │
         │    data     │
         └─────────────┘
                │
                ▼
    ┌───────────────────────────┐
    │ Setting inputs and outputs│
    └───────────────────────────┘
                │
                ▼                        ┌──────────────┐
    ┌───────────────────────┐           │    Change    │
    │  Developing the model │◄──────────│   Training    │
    └───────────────────────┘           │   Parameter  │
                │                        └──────────────┘
                ▼                              ▲
           ╱─────────╲         No              │
          ╱ Are the   ╲────────────────────────┘
         ╱ f-ratio > 4 ╲
         ╲ and the      ╱
          ╲ Trained R² ╱
           ╲  > 80?   ╱
            ╲────────╱
                │ Yes
                ▼
    ┌───────────────────┐
    │  Using the Model  │
    └───────────────────┘
                │
                ▼
         ┌─────────────┐
         │  Generate   │
         │   reports   │
         └─────────────┘
```
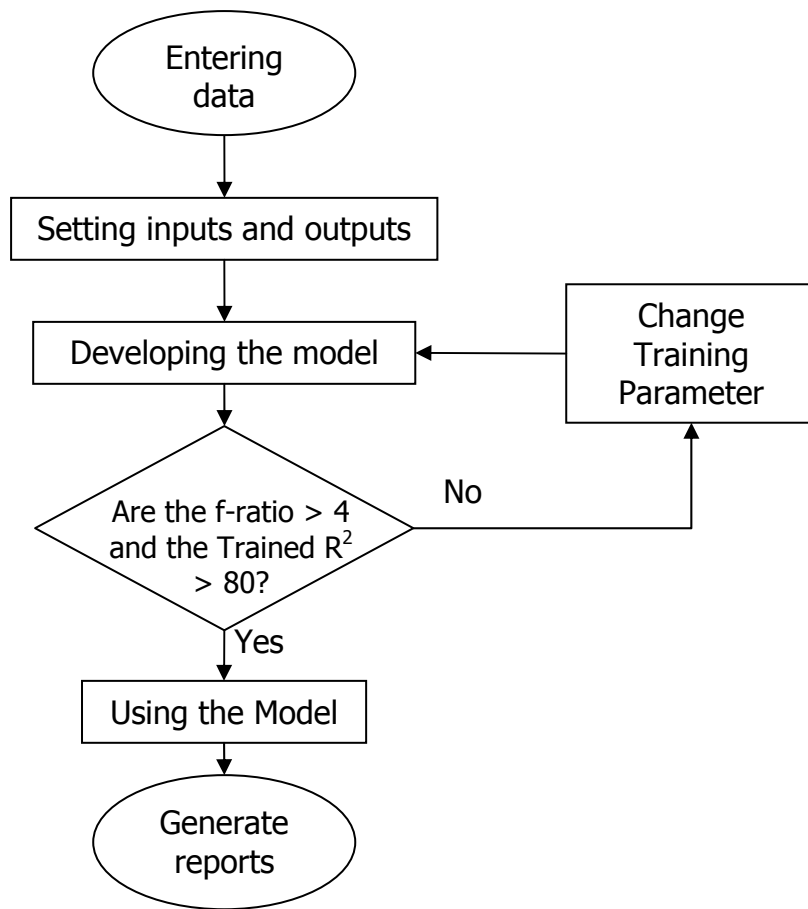
Figure 6.1 FormRules Training flowchart
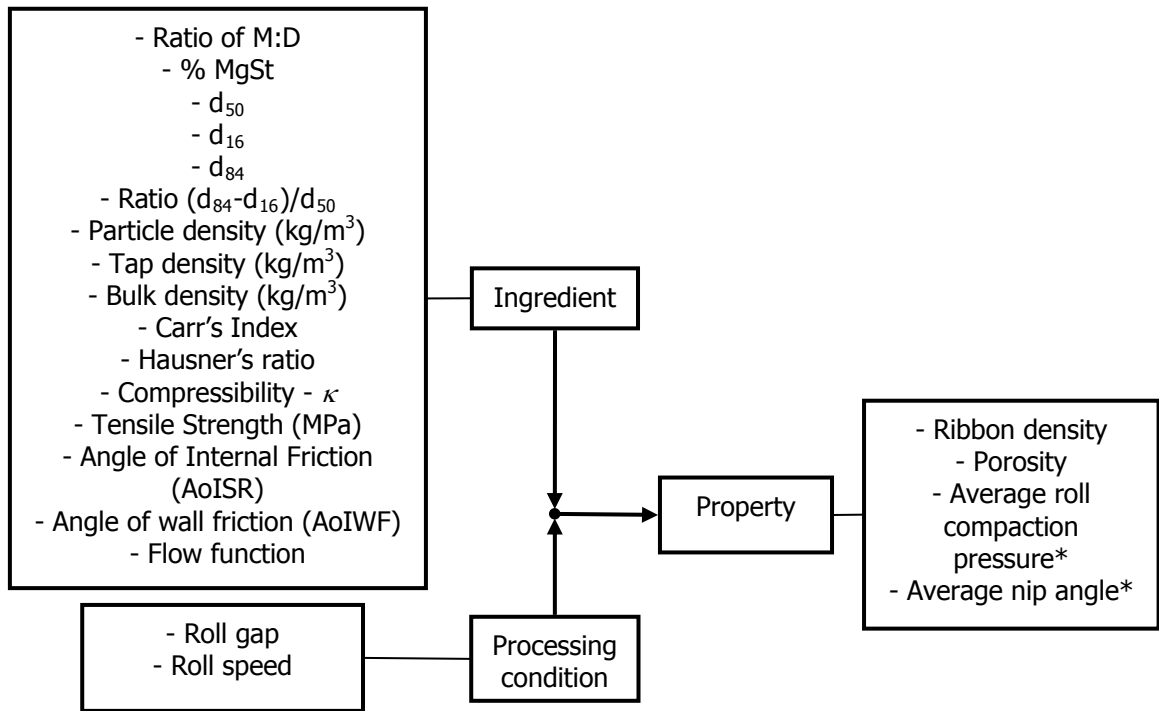
Figure 6.2 List of inputs to be investigated using Formrules against the output values required. *Average roll compaction pressure and average nip angle are intermediate output properties of the roll compaction process. The ratio of M:D stands for ratio of MCC:DCPA.

6.1.2 f-ratio and FormRules R^2 Explained

The statistics system used to assess the quality of the model in the *INForm* software

is called ANOVA (Analysis of Variance). The results of ANOVA are expressed in the

form of f-ratio and R^2. The f-ratio is the ratio of the mean square of the model, over

the mean square of the error.

$$f-ratio = \frac{mean\ squares\ for\ the\ \mathrm{mod}el}{mean\ squares\ for\ the\ error} = \frac{SSR/k}{SSE/(n-k-1)} \qquad (6.1)$$

where SSR is the sum of squares for the model, SSE is the sum of squares for the

error, k is the number of weights and biases (i.e. degrees of freedom for model), n

is the number of datasets and n-k-1 is the degrees of freedom for the error.

For the model to be of acceptable quality the mean squares for the model need to be at least four times greater than the mean squares of the error, i.e. f-ratio > 4.

The R^2 in *INForm* is calculated from

$$R^2 = \left(1 - \frac{SSE}{SST}\right) \times 100$$

(6.2)

For linear regression the total variance in a single output variable can be expressed as

SST = SSR+ SSE

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

(6.3)

where SST is the total sum of squares, $\bar{y}$ is the mean of the dependent variable, and $\hat{y}$ is the predicted value from the model.

In *FormRules* software training, the R^2 value should be at least 80% for a good quality model.

6.1.3 Results and Discussion for Trial A

6.1.3.1   Neurofuzzy Diagrams

Figure 6.3 show the neurofuzzy results for ribbon density. This figure highlights that the PSD Ratio (d84-d16)/d50, the Roll Speed and Roll Gap are important input variables for ribbon density. Only these three inputs were predicted by *FormRules* software to affect the ribbon density.   In the display diagram this relationship is shown by the box labeled submodel 1.   In the next section, the linguistically-expressed rules for ribbon density output are explained.
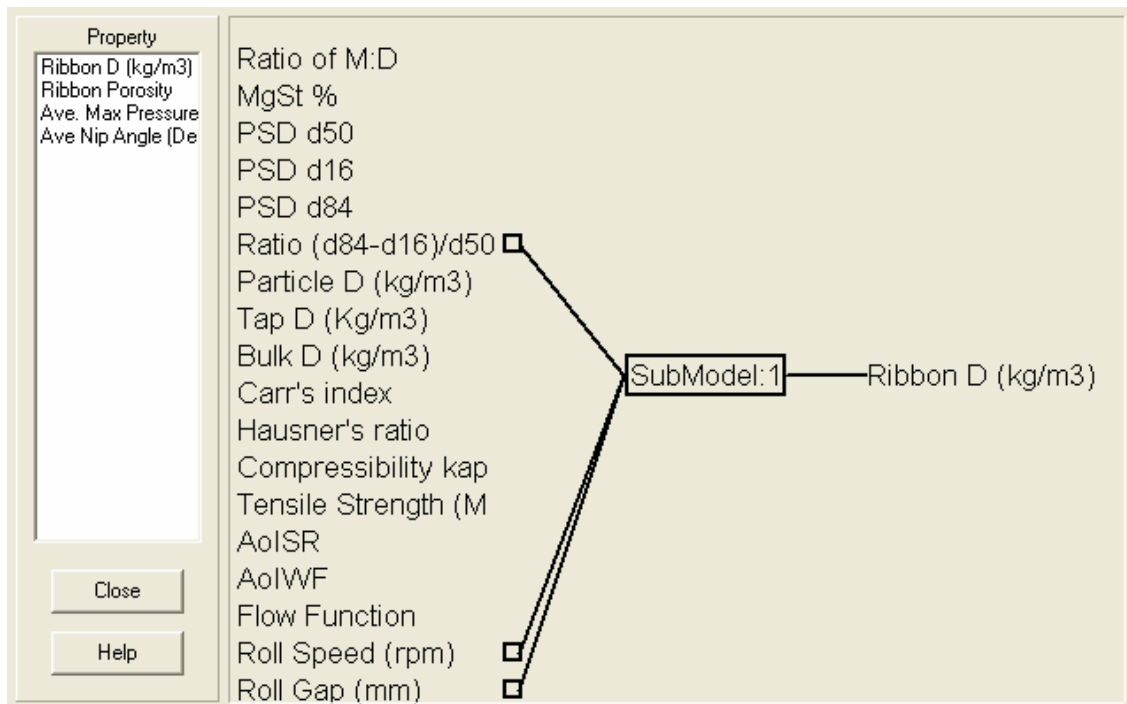
Figure 6.3 Neurofuzzy Results for ribbon density (kg/m^3)

Figures 6.4 to 6.6 show the neurofuzzy results for ribbon porosity, average maximum pressure and average nip angle. The neurofuzzy models developed for each one contain more than one submodel.   This is because the software is detecting multivariate effects of the inputs on each of the outputs.   In the earlier

case of neurofuzzy results for ribbon density, the submodel 1 is observed to be combining the effects of three input variables, whereas in the neurofuzzy results for ribbon porosity, submodel 1 combines the effect of 2 input variables and submodel 2 and 3 consist of only one input variable.   Each of these submodels has an independent effect on the ribbon porosity.   This can be observed more intricately within the linguistically-expressed rules.
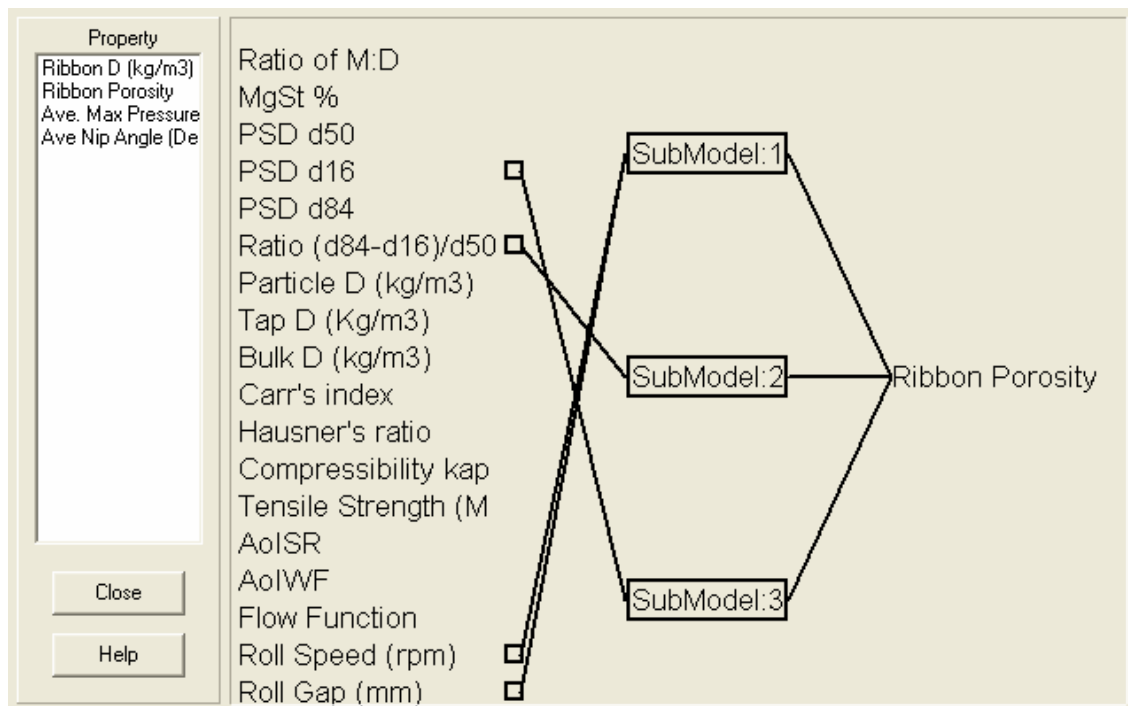
Figure 6.4 Neurofuzzy results for ribbon porosity
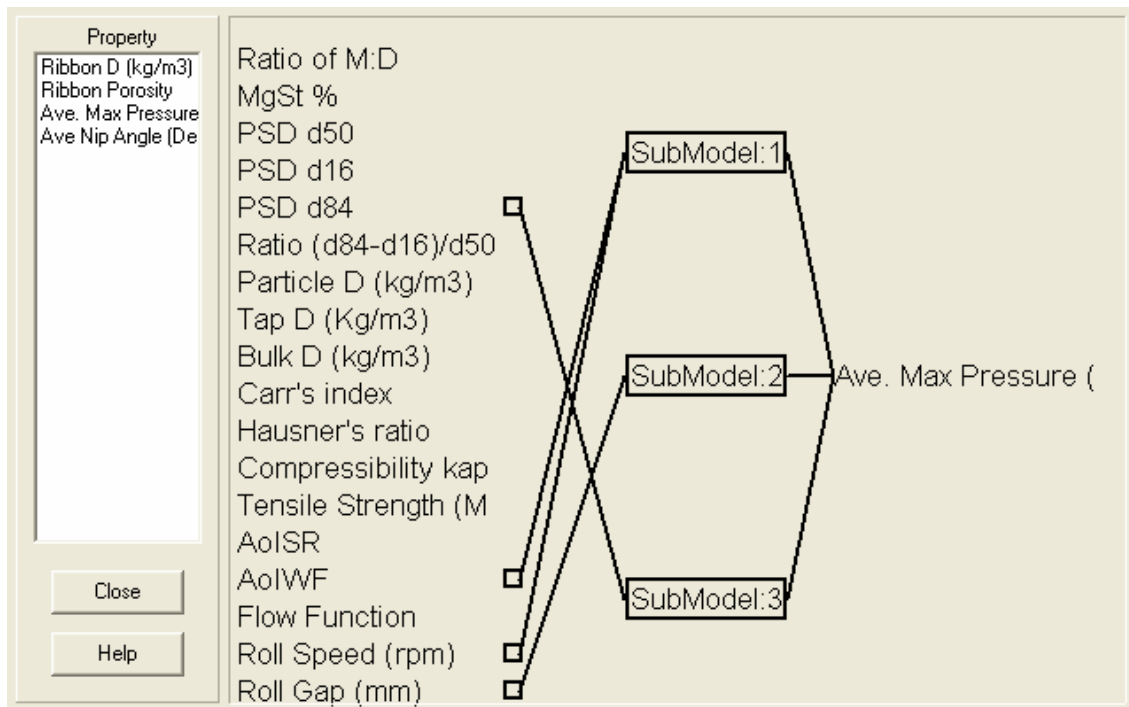
Figure 6.5 Neurofuzzy results for average maximum pressure (MPa)
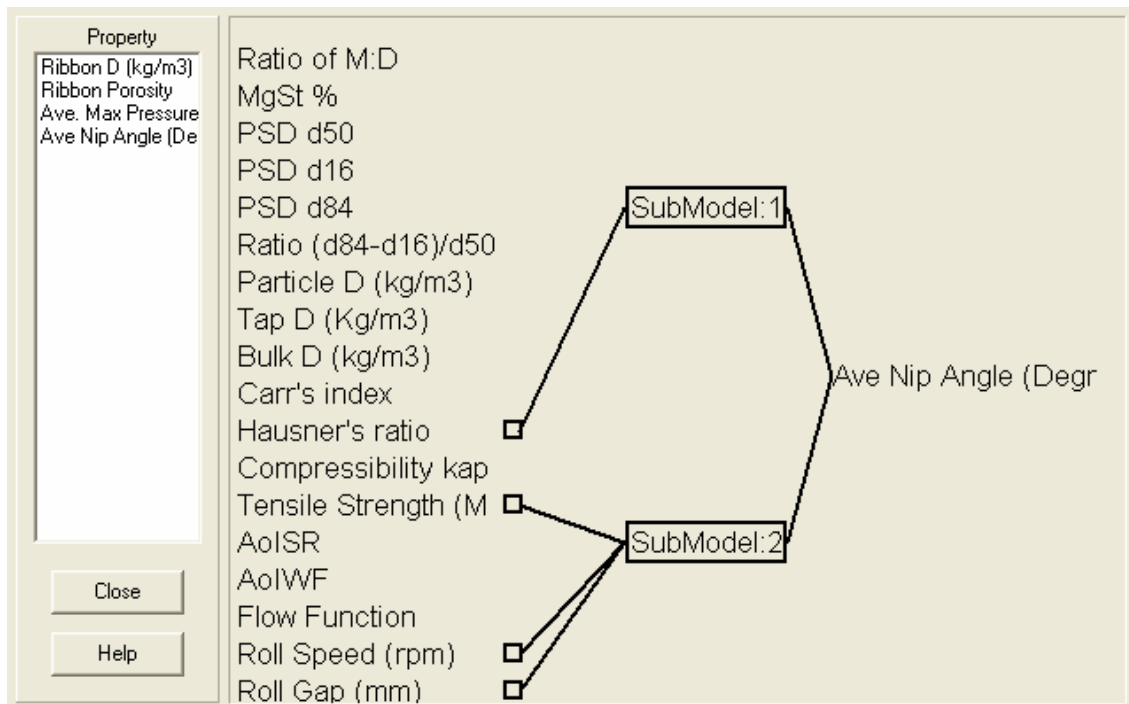
Figure 6.6 Neurofuzzy results for average nip angle (°)

6.1.3.2   Linguistically-Expressed Rules

Table 6.1 shows the neurofuzzy rules developed by the *FormRules* software for ribbon density in Trial A. Note that the rules are numbered to ease discussion. It can be seen that the relationships between the inputs and the output are not linear. Referring to rules 1 – 3, at low Roll Speeds and Roll Gaps the Ribbon Density is high except when the PSD Ratio (d84-d16)/d50 is in mid range. In this case the ribbon density is low with membership function of 58%. This membership function indicates that the ribbon density is 58% low and it also means that the ribbon density is 42% high. This is a fuzzy relationship linguistic expression rule which sounds ambiguous but could be useful when defuzzified in steps shown in Section 2.2.3.

Rules 4 – 6 state that at low Roll Speeds and high Roll Gaps, the Ribbon Density is mainly high with the exception again that at PSD Ratio (d84-d16)/d50 in Mid the ribbon density is low with membership function of almost 90%. This means that it is more definite for powder materials with PSD Ratio (d84-d16)/d50 in mid range, roll compacted at low Roll Speed and high Roll Gap to be of low Ribbon Density. It can also be seen from rules 4 and 6 that their membership functions are lower hence the chances of obtaining a high ribbon density are less. Rules 7 to 12 show that at high Roll Speeds, regardless of Roll Gap and PSD Ratio (d84-d16)/d50, the Ribbon Density will be low but at varying membership functions. The Ribbon Density is definitely low for rules 7 and 11.

Table 6.1 Neurofuzzy rules for ribbon density (kg/m^3) submodel 1 for Trial A.

Rule 1 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is HIGH (1.00)
Rule 2 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.58)
Rule 3 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is HIGH (0.69)
Rule 4 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is HIGH (0.75)
Rule 5 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.87)
Rule 6 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is HIGH (0.54)
Rule 7 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
Rule 8 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.63)
Rule 9 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is LOW (0.63)
Rule 10 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is LOW (0.64)
Rule 11 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (1.00)
Rule 12 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is LOW (0.72)

Table 6.2 shows that although ribbon porosity is calculated from ribbon density the neurofuzzy rules developed from *FormRules* software are very different for both outputs.

Table 6.2 Neurofuzzy rules for ribbon porosity submodel 1, 2 and 3 for Trial A.

SubModel:1
Rule 1 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW THEN Ribbon Porosity is LOW (1.00)
Rule 2 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is MID THEN Ribbon Porosity is LOW (1.00)
Rule 3 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH THEN Ribbon Porosity is LOW (1.00)
Rule 4 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW THEN Ribbon Porosity is LOW (1.00)
Rule 5 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is MID THEN Ribbon Porosity is HIGH (1.00)
Rule 6 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH THEN Ribbon Porosity is HIGH (1.00)
SubModel:2
Rule 1 - IF Ratio (d84-d16)/d50 is LOW THEN Ribbon Porosity is HIGH (1.00)
Rule 2 - IF Ratio (d84-d16)/d50 is HIGH THEN Ribbon Porosity is LOW (0.93)
SubModel:3
Rule 1 - IF PSD d16 is LOW THEN Ribbon Porosity is HIGH (0.77)
Rule 2 - IF PSD d16 is HIGH THEN Ribbon Porosity is LOW (0.53)

The submodel 1 and 2 results are expected. However the submodel 3 is not expected; the rules 1 and 2 indicate that at low PSD d16 the ribbon porosity is high and at high PSD d16 the ribbon porosity is low. It may be assumed that the

FormRules software has detected that pharmaceutical formulations with large amounts of smaller particles would produce ribbon of lower porosity because of the higher possibility of pore filling.

Table 6.3 shows the neurofuzzy rules for the average maximum pressure output, which has three submodels and the first one detected the combined effect of Roll Speed and Angle of Wall Friction. In submodel 1 only rules 2 and 3 give a high maximum pressure. These rules states that at low Roll Speeds and from mid to high range of Angle of Wall Friction, a high roll compaction pressure can be obtained. The importance of angle of wall friction to the average maximum pressure is in agreement with Johanson's Theory as shown in Chapter 5. Submodel 2 has stated the obvious, however submodel 3 has stated that PSD d84, the upper range of the PSD is important to the Average Maximum Pressure.   This is logical as larger particles would be compacted at higher pressure than smaller particles at constant volume.

Table 6.4 shows the neurofuzzy rules for average nip angle (^{o}) with two submodels. The first submodel relates Hausner's Ratio to the nip angle. As mentioned in section 3.3.2, Hausner's Ratio is related to interparticle friction and low Hausner's Ratio means the powder flows well, while higher ratio means poor flow. The rules 1 to 3 in submodel 1 are logical, considering that for good flowing powders the nip angle is lower and for poor flowing powders the nip angle is higher. Submodel 2 combines the effect of Roll Speed, Roll Gap and Tensile Strength on the Average Nip Angle.

Table 6.3 Neurofuzzy rules for Average Maximum Pressure (MPa) submodel 1, 2 and 3 for Trial A. AoIWF is Angle of Wall Friction.

SubModel:1
Rule 1 - IF Roll Speed (rpm) is LOW AND AoIWF is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
Rule 2 - IF Roll Speed (rpm) is LOW AND AoIWF is MID THEN Ave. Max Pressure (MPa) is HIGH (1.00)
Rule 3 - IF Roll Speed (rpm) is LOW AND AoIWF is HIGH THEN Ave. Max Pressure (MPa) is HIGH (1.00)
Rule 4 - IF Roll Speed (rpm) is MID AND AoIWF is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
Rule 5 - IF Roll Speed (rpm) is MID AND AoIWF is MID THEN Ave. Max Pressure (MPa) is LOW (0.95)
Rule 6 - IF Roll Speed (rpm) is MID AND AoIWF is HIGH THEN Ave. Max Pressure (MPa) is LOW (1.00)
Rule 7 - IF Roll Speed (rpm) is HIGH AND AoIWF is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
Rule 8 - IF Roll Speed (rpm) is HIGH AND AoIWF is MID THEN Ave. Max Pressure (MPa) is LOW (1.00)
Rule 9 - IF Roll Speed (rpm) is HIGH AND AoIWF is HIGH THEN Ave. Max Pressure (MPa) is LOW (1.00)
SubModel:2
Rule 1 - IF Roll Gap (mm) is LOW THEN Ave. Max Pressure (MPa) is HIGH (1.00)
Rule 2- IF Roll Gap (mm) is HIGH THEN Ave. Max Pressure (MPa) is LOW (1.00)
SubModel:3
Rule 1 - IF PSD d84 is LOW THEN Ave. Max Pressure (MPa) is LOW (0.96)
Rule 2 - IF PSD d84 is HIGH THEN Ave. Max Pressure (MPa) is HIGH (0.70)

The Roll Speed and Roll Gap were explained in Section 4.3.4 and tensile strength was explained in Section 3.3.4. The experimental results show that for increasing roll gap and decreasing roll speed the nip angle increases. The opposing effects of the roll gap and roll speed on nip angle are hard to quantify. Yet Submodel 2 highlighted that for Rules 1, 2, and 5 to 8 the average nip angle were low. Whereas, Rules 3 and 4 gave a high average nip angle, which means that the combination of low Roll Speeds, high Roll Gap and Tensile Strength are important to, regardless of the Tensile Strength being high or low. Generally, the rules 1 to 8 do not give any indication that the Tensile Strength has a large effect on Average Nip Angle except for rules 3 and 4 however it was observed in Chapter 4 that increasing tensile strength resulted in an increase in nip angle except for lubricated

MCC. Thus this is just one of the examples where the *FormRules* software has highlighted inputs which are multivariate in nature.

Table 6.4 Neurofuzzy rules for average nip angle (°), submodel 1 and 2 for Trial A.

SubModel:1
Rule 1 - IF Hausner's Ratio is LOW THEN Ave Nip Angle (°) is LOW (0.78)
Rule 2 - IF Hausner's Ratio is MID THEN Ave Nip Angle (°) is HIGH (0.67)
Rule 3 - IF Hausner's Ratio is HIGH THEN Ave Nip Angle (°) is HIGH (0.53)
SubModel:2
Rule 1 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Tensile Strength (MPa) is LOW THEN Ave Nip Angle (°) is LOW (0.79)
Rule 2 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Tensile Strength (MPa) is HIGH THEN Ave Nip Angle (°) is LOW (0.69)
Rule 3 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Tensile Strength (MPa) is LOW THEN Ave Nip Angle (°) is HIGH (0.53)
Rule 4 - IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Tensile Strength (MPa) is HIGH THEN Ave Nip Angle (°) is HIGH (1.00)
Rule 5 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Tensile Strength (MPa) is LOW THEN Ave Nip Angle (°) is LOW (1.00)
Rule 6 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Tensile Strength (MPa) is HIGH THEN Ave Nip Angle (°) is LOW (1.00)
Rule 7 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Tensile Strength (MPa) is LOW THEN Ave Nip Angle (°) is LOW (1.00)
Rule 8 - IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Tensile Strength (MPa) is HIGH THEN Ave Nip Angle (°) is LOW (0.54)

6.1.3.3   Concluding Remarks for Trial A

The *FormRules* program was found to be user friendly and straightforward to use. The default settings were found to be suitable for the data sets obtained from roll compaction. In Trial A all 18 inputs were fed into the *FormRules* software and 8 inputs were highlighted as the key variables. The key variables for the roll compaction output property and intermediate output property are summarized in Table 6.5.

Trial A showed that the *FormRules* software was successful in detecting the importance of the roll compaction process parameters as input variables in relation to the output properties.   Generally, the software detected sound multivariate

relationships between inputs and outputs. However, it is important to be aware of the relevance of the key variables which this software highlights as important in relation to the output property.

Table 6.5 Summary of key variables found by *FormRules* software model for Trial A.

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio (d84-d16)/d50, Roll Speed (rpm) and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio (d84-d16)/d50 Submodel 3 : PSD d16	Ribbon Porosity
Submodel 1 : Angle of Wall Friction (AoIWF) and Roll Speed (rpm) Submodel 2 : Roll Gap (mm) Submodel 3 : PSD d84	Average Maximum Pressure (MPa)
Submodel 1 : Hausner's Ratio Submodel 2 : Tensile strength (MPa), Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle (^{o})

In the following section, 11 Trials were presented to demonstrate the ability of the *FormRules* software to highlight key variables from lesser amounts of inputs and produce workable neurofuzzy rules.   Then *FormRules* software was used in sequence with *INForm* (See Chapter 7). The chosen trials in this chapter were used in the next chapter to train *INForm*.

6.2 Results and Trials

Combinations of various numbers of inputs relating to roll compaction process conditions were tested in 11 trials.   The model assessment for Trials A – K are shown in Table 6.6, which summarises the f-ratio and the Mean Square (MS) Error

for each output property. The table also shows the MS Error for each model and the system structure developed for the specific trial output property. The f-ratio values are all above 4 and the test R^2 values are all above 85%.

Table 6.6 Summary model assessments for Trials A – K.

Trial	Output Property	MS Error	System structure	f-ratio	Test R^2
A	Ribbon D (kg/m³)	0.0011	17(2)18(2)6(3)	178.57	97.72
	Ribbon Porosity	0.0075	17(2)18(3)+6(2)+4(2)	51.85	88.48
	Ave. Max Pressure (Mpa)	0.0114	17(3)15(3)+18(2)+5(2)	26.50	85.11
	Ave Nip Angle (°)	0.0011	11(3)+17(2)18(2)13(2)	253.17	97.99
B	Output Property	0.0011	12(2)13(2)5(3)	178.57	97.72
	Ribbon D (kg/m³)	0.0075	12(2)13(3)+5(2)+3(2)	51.85	88.48
	Ribbon Porosity	0.0114	12(3)11(3)+13(2)+4(2)	26.50	85.11
	Ave. Max Pressure (Mpa)	0.0011	7(3)+12(2)13(2)9(2)	253.17	97.99
C	Ave Nip Angle (°)	0.0012	1(2)5(2)3(2)+6(2)2(2)	172.77	97.39
	Output Property	0.0053	5(2)6(2)2(2)+3(2)+1(2)	58.77	91.87
	Ribbon D (kg/m³)	0.0069	3(2)+6(2)2(2)5(3)	37.91	90.96
	Ribbon Porosity	0.0016	5(2)4(3)+3(2)6(2)	190.61	97.00
D*	Ave. Max Pressure (Mpa)	0.0022	1(3)3(2)+4(2)2(2)	118.91	95.28
	Ave Nip Angle (°)	0.0074	3(2)4(3)+1(3)	52.34	88.58
	Output Property	0.0064	4(2)3(3)2(2)+1(3)	37.53	91.63
	Ribbon D (kg/m³)	0.0016	3(2)4(3)+1(3)+2(2)	187.93	96.96
E	Ribbon Porosity	0.0011	1(2)6(2)4(3)+7(2)2(2)	128.76	97.62
	Ave. Max Pressure (Mpa)	0.0074	6(2)7(3)+1(3)	52.34	88.58
	Ave Nip Angle (°)	0.0116	6(3)4(3)+7(2)+3(2)	25.97	84.85
	Output Property	0.0011	4(3)+6(2)1(2)7(2)	245.52	97.93
F	Ribbon D (kg/m³)	0.0012	1(2)5(2)3(2)+6(2)2(2)	172.77	97.39
	Ribbon Porosity	0.0053	5(2)6(2)2(2)+3(2)+1(2)	58.77	91.87
	Ave. Max Pressure (Mpa)	0.0069	3(2)+6(2)2(2)5(3)	37.91	90.96
	Ave Nip Angle (°)	0.0016	5(2)4(3)+3(2)6(2)	190.61	97.00

	Output Property	0.0016	1(2)8(2)7(2)+9(2)2(2)	132.47	96.62
G	Ribbon D (kg/m³)	0.0074	8(2)9(3)+1(3)	52.34	88.58
	Ribbon Porosity	0.0064	9(2)8(3)2(2)+6(3)	37.81	91.69
	Ave. Max Pressure (Mpa)	0.0011	7(3)+8(2)9(2)1(2)	247.04	97.94
	Ave Nip Angle (°)	0.0011	7(2)8(2)6(3)	178.57	97.72
H	Output Property	0.0075	7(2)8(3)+6(2)+4(2)	51.85	88.48
	Ribbon D (kg/m³)	0.0068	8(2)7(3)2(2)+5(2)	38.53	91.09
	Ribbon Porosity	0.0010	7(2)8(2)6(3)+5(2)	198.86	98.14
	Ave. Max Pressure (Mpa)	0.0012	1(2)7(2)5(2)+8(2)2(2)	172.77	97.39
I*	Ave Nip Angle (°)	0.0084	7(2)8(3)+3(2)	52.85	87.06
	Output Property	0.0114	7(3)6(3)+8(2)+5(2)	26.50	85.11
	Ribbon D (kg/m³)	0.0011	4(3)+7(2)8(2)1(2)	247.04	97.94
	Ribbon Porosity	0.0012	1(2)5(2)4(2)+6(2)2(2)	172.77	97.39
J*	Ave. Max Pressure (Mpa)	0.0084	5(2)6(3)+3(2)	52.85	87.06
	Ave Nip Angle (°)	0.0069	4(2)+6(2)2(2)5(3)	37.91	90.96
	Output Property	0.0011	5(2)3(3)+4(2)6(2)+1(2)5(2)	192.32	97.88
	Ribbon D (kg/m³)	0.0012	1(2)5(2)3(2)+6(2)2(2)	172.77	97.39
K	Ribbon Porosity	0.0053	5(2)6(2)2(2)+3(2)+1(2)	58.77	91.87
	Ave. Max Pressure (Mpa)	0.0114	5(3)4(3)+6(2)+3(2)	26.50	85.11
	Ave Nip Angle (°)	0.0011	4(3)+5(2)1(2)6(2)	245.52	97.93

Table 6.7 summarises the inputs fed into each trial. Note that the greyed cells are inputs which were fed into the *FormRules* program and the numbered items represent the frequency at which the inputs were highlighted in each of the Trials. Trial A was conducted to investigate the ability of *FormRules* software to process and highlight key variables from the maximum amount of inputs. Then the highlighted key variables from Trial A as well as the formulation variables and 3 inputs (i.e. Carr's Index, Compressibility κ and effective angle of internal friction) which were thought to be important were used as inputs in Trial B to investigate the

importance of these variables when compared to each other or not. Next, Trial C was conducted to determine if the *FormRules* software would be successfully trained without the Formulation inputs (i.e. Ratio M:D and MgSt%) but with only PSD d16, PSD d84, Ratio (d84-d16)/d50, Hausner's Ratio, tensile strength and Angle of Wall Friction. These inputs were characteristics which were highlighted in earlier trials.

Following that, Trial D was conducted to investigate the minimum amount of inputs that would be sufficient to form neurofuzzy rules. Then Trials E to H were conducted to study if feeding the *FormRules* software with inputs from each characterisation group (i.e. shear testing, uniaxial compaction, densities and PSD) would be successful. Next the highlighted key variables from Trial B were used as inputs in Trial I with the exception of adding the Ratio M:D and replacing tensile strength with Compressibility κ. Tensile strength was switched to Compressibility κ because in practice it is experimentally more straightforward to obtain this value than to obtain tensile strength. This trial was conducted to investigate whether *FormRules* software would show that the formulation would be picked up as an important variable, to allow for easy association of which sample it is in the *INForm* model prediction and if Compressibility κ would be highlighted as important.

After that, Trial J was conducted to investigate whether the *FormRules* software would be successfully trained with only two characteristic inputs, which are Ratio (d84-d16)/d50 and Compressibility κ on top of the Formulation and Process Condition. These two inputs are most directly characterised and in practice the

industry would want to be able to perform the most representative, simplest and quickest characterisation method for the formulation. Lastly, Trial K was conducted to check if the *FormRules* software could be successfully trained with only two characteristic inputs that are Compressibility κ and Angle of Wall Friction on top of the Formulation and Process Condition. These two inputs are characteristics of the powder material which were used in Johanson's Theory in Chapter 5 to predict the nip angle value. This trial is important because the model trained from this input will be used in comparison to Johanson's Theory. However within the model trained for predicting the average nip angle, it was found that the Angle of Wall Friction was highlighted as a key variable but not Compressibility κ. The Compressibility κ is a very important factor in Johanson's model to predict nip angle. Although this trial was tested in the *INForm* model training, it was not reported in Chapter 7 because it does not provide a better comparison to the traditional model than the Trials D, I or J.

Only 7 out of 11 trials were used in *INForm* software training. This was done because Trials A and B had similar outputs hence there was no reason to repeat the same trial on *INForm* software training. The Trials E, G and H, were not tested on *INForm* software training because these tests were only conducted to observe the results of training a model with inputs derived mainly from one method of characterisation (e.g. Trials H inputs were all from particle size characterisation only) and the inputs were only highlighted for certain numbers of outputs or not all. Trial F was trained for *INForm* because the Compressibility κ was highlighted as key

variable for all the outputs and it was interesting to see if this trial would be successfully produce a good model.

6.2.1 Neurofuzzy result summary

In this subsection the results of the neurofuzzy modelling are summarised. The key variables (i.e. the submodel inputs) for each of the output are shown in tabular form.

Trial A There were 18 inputs altogether and the inputs were classified under Formulation, Particle Size Distribution (PSD), Densities, Uniaxial Compaction, Shear Testing and Process Condition. The trial successfully highlighted 8 inputs as important variables and successfully formed good predictive models in *INForm.* Table 6.5 shows the summary of the key variables highlighted by *FormRules* software modeling.

Trial B 13 inputs were fed into the *FormRules* program. Although the trial was successful, it showed that only 8 inputs were highlighted as important variables instead of all 13 inputs which activated as inputs. It was noted that in this trial the Formulation inputs were not highlighted as important variables. Table 6.8 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.7 *FormRules* trials

TRIALS		A	B	C	D*	E	F	G	H	I*	J*	K
Formulation	Ratio of M:D				4	3	2	3		2	2	3
	MgSt%				3	1	3	2	1	1	2	2
PSD	d50											
	d16	1	1	1					1			
	d84	1	1	1					2			
	Ratio (d84-d16)/d50	2	2	2					3	1	2	
Densities	Particle Density (kg/m^3)											
	Tap Density (kg/m^3)											
	Bulk Density (kg/m^3)											
	Carr's Index							1				
	Hausner's Ratio	1	1	1				2		1		
Uniaxial compaction	Compressibility κ						4			2	3	3
	Tensile Strength (MPa)	1	1	1			1					
Shear test	Angle of internal friction (AoISR)					1						
	Angle of Wall Friction (AoIWF)	1	1	1		3				1		2
	Flow function											
Process condition	Roll Speed	4	4	5	4	4	4	4	4	4	5	4
	Roll Gap (mm)	4	4	4	4	4	4	4	4	4	4	4
Success of modelling in *FormRules* software		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tested on *INForm* rules		Yes	No	Yes	Yes	No	Yes	No	No	Yes	Yes	Yes
Success of modelling in *INForm* rules		Yes	n/a	Yes	Yes	n/a	Yes	n/a	n/a	Yes	Yes	Yes

Note: The greyed cells are inputs which were fed into the *FormRules* software and the numbered items represent the frequency at which the inputs were highlighted in each of the trials.

* Represent trials which had been chosen to be optimised and experimentally validated in *INForm*

Table 6.8 Summary of Neurofuzzy results from *FormRules* software model training for Trial B

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio (d84-d16)/d50, Roll Speed (rpm) and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio (d84-d16)/d50 Submodel 3 : PSD d16	Ribbon Porosity
Submodel 1 : Angle of Wall Friction and Roll Speed (rpm) Submodel 2 : Roll Gap (mm) Submodel 3 : PSD d84	Average Maximum Pressure (MPa)
Submodel 1 : Hausner's Ratio Submodel 2 : Tensile Strength (MPa), Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle (^{o})

Trial C 8 inputs were used and the trial was successful. 8 inputs were highlighted as important variables. This trial did undergo *INForm* model training and was successful. Table 6.9 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.9 Summary of Neurofuzzy results from *FormRules* software model training for Trial C

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio (d84-d16)/d50, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio (d84-d16)/d50 Submodel 3 : PSD d16	Ribbon Porosity
Submodel 1 : Angle of Wall Friction and Roll Speed (mm) Submodel 2 : Roll Gap (mm) Submodel 3 : PSD d84	Average Maximum Pressure (MPa)
Submodel 1 : Hausner's Ratio Submodel 2 : Tensile Strength (MPa), Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle (^{o})

Trial D 4 inputs were fed into the *FormRules* program. They were classified under Formulation and Process Condition. The trial was successful, 4 inputs were highlighted as important and successfully formed good predictive models in *INForm.* Table 6.10 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.10 Summary of Neurofuzzy results from *FormRules* model training for Trial D

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : MgSt%, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio of M:D	Ribbon Porosity
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio of M:D	Average Maximum Pressure (MPa)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio of M:D Submodel 3 : MgSt %	Average Nip Angle (^{o})

Trial E 7 inputs were fed into the *FormRules* program, classified under Formulation, Shear Testing and Process Condition. The trial was successful and 6 inputs were highlighted as important variables. This trial did not undergo *INForm* model training. Table 6.11 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.11 Summary of Neurofuzzy results from *FormRules* software model training for Trial E

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D, Angle of Wall Friction and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio of M:D	Ribbon Porosity
Submodel 1 : Angle of Wall Friction and Roll Speed (rpm) Submodel 2 : Roll Gap (mm) Submodel 3 : Angle of Internal Friction	Average Maximum Pressure (MPa)
Submodel 1 : Angle of Wall Friction Submodel 2 : Ratio of M:D, Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle (^{o})

Trial F 6 inputs were fed into *FormRules* program, classified under Formulation, Uniaxial compaction and Process Condition. The trial was successful and 6 inputs were highlighted as important variables. This trial did undergo *INForm* model training and was successful. Table 6.12 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.12 Summary of Neurofuzzy results from *FormRules* software model training for Trial F

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D, Compressibility κ and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : MgSt%, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Compressibility κ Submodel 3 : Ratio of M:D	Ribbon Porosity
Submodel 1 : Compressibility κ Submodel 2 : MgSt%, Roll Speed (rpm) and Roll Gap (mm)	Average Maximum Pressure (MPa)
Submodel 1 : Tensile Strength (MPa) and Roll Speed (mm) Submodel 2 : Compressibility κ and Roll Gap (mm)	Average Nip Angle (^{o})

Trial G   9 inputs were fed into the *FormRules* program, classified under Formulation, Densities and Process Condition. The trial was successful and 6 inputs were highlighted as important variables. This trial did not undergo *INForm* model training. Table 6.13 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.13 Summary of Neurofuzzy results from *FormRules* software model training for Trial G

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D, Hausner's Ratio and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio of M:D	Ribbon Porosity
Submodel 1 : MgSt%, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Carr's Index	Average Maximum Pressure (MPa)
Submodel 1 : Hausner's Ratio Submodel 2 : Ratio of M:D, Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle (^{o})

Trial H   9 inputs were fed into the *FormRules* program, classified under Formulation, Densities and Process Condition. The trial was successful and 6 inputs were highlighted as important variables. This trial did not undergo *INForm* model training. Table 6.14 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.14 Summary of Neurofuzzy results from *FormRules* software model training for Trial H

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio (d84-d16)/d50, Roll Speed (rpm) and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio (d84-d16)/d50 Submodel 3 : PSD d16	Ribbon Porosity
Submodel 1 : MgSt%, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : PSD d84	Average Maximum Pressure (MPa)
Submodel 1 : Ratio (d84-d16)/d50, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : PSD d84	Average Nip Angle (^{o})

Trial I 8 inputs were used and the trial was successful. 8 inputs were highlighted as important variables. This trial did undergo *INForm* model training and was successful. Table 6.15 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.15 Summary of Neurofuzzy results from *FormRules* software model training for Trial I

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D, Compressibility κ and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio (d84-d16)/d50	Ribbon Porosity
Submodel 1 : Angle of Wall Friction and Roll Speed (rpm) Submodel 2 : Roll Gap (mm) Submodel 3 : Compressibility κ	Average Maximum Pressure (MPa)
Submodel 1 : Hausner's Ratio Submodel 2 : Ratio of M:D, Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle (^{o})

Trial J   6 inputs were used and the trial was successful.   6 inputs were highlighted as important variables. This trial did undergo *INForm* model training and was successful. Table 6.16 shows the summary of the key variables highlighted by *FormRules* software.

Table 6.16 Summary of Neurofuzzy results from *FormRules* software model training for Trial J

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D, Compressibility κ and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Ratio (d84-d16)/d50	Ribbon Porosity
Submodel 1 : Compressibility κ Submodel 2 : MgSt%, Roll Speed (rpm) and Roll Gap (mm)	Average Maximum Pressure (MPa)
Submodel 1 : Ratio (d84-d16)/d50 and Roll Speed (rpm) Submodel 2 : Compressibility κ and Roll Gap (mm) Submodel 3 : Ratio of M:D and Roll Speed (rpm)	Average Nip Angle ($^\circ$)

Trial K   6 inputs were used and the trial was successful.   6 inputs were highlighted as important variables. This trial did undergo *INForm* model training and was successful. Table 6.17 shows the summary of the key variables highlighted by *FormRules* software.

158

Table 6.17 Summary of Neurofuzzy results from *FormRules* model training for Trial K

Key Variables	Roll Compaction Output Property
Submodel 1 : Ratio of M:D, Compressibility κ and Roll Speed (rpm) Submodel 2 : MgSt% and Roll Gap (mm)	Ribbon Density (kg/m^3)
Submodel 1 : MgSt%, Roll Speed (rpm) and Roll Gap (mm) Submodel 2 : Compressibility κ Submodel 3 : Ratio of M:D	Ribbon Porosity
Submodel 1 : Angle of Wall Friction and Roll Speed (mm) Submodel 2 : Roll Gap (mm) Submodel 3 : Compressibility κ	Average Maximum Pressure (MPa)
Submodel 1 : Angle of Wall Friction Submodel 2 : Ratio of M:D, Roll Speed (rpm) and Roll Gap (mm)	Average Nip Angle ($^{\circ}$)

6.3 Discussions

The *FormRules* program was found to be successfully trained for all the combinations of inputs in Trials A – K. Table 6.7 shows that inputs which were highlighted as important in the earlier trials would repetitively be highlighted as important in subsequent trials. These inputs are Ratio (d84-d16)/d50, Hausner's Ratio, Compressibility κ, Angle of Wall Friction and processing condition. The processing condition inputs were expected to be important and were always highlighted as a key variable by the *FormRules* trials. However the formulation inputs (i.e. Ratio M:D and 1%MgSt) were not always highlighted as important variables (as shown in Trial A and Trial C).

A closer look at the Ratio (d84-d16)/d50 (Trials A – C and Trials H - J) input showed that it was a key variable for the models developed for ribbon density, ribbon

porosity and average nip angle. This is a good indication that the software could detect the importance of the PSD in the production of ribbons. Neurofuzzy logic from Trial B (Table 6.18) was used as an example to show the neurofuzzy rules developed to relate Ratio (d84-d16)/d50 to the outputs. The rules report that at high Roll Speed the Ribbon Density is low regardless of the Roll Gap and the Ratio (d84-d16)/d50 and this is in agreement with real experimental results. Low Roll Speed resulted in high Ribbon Density regardless of the Roll Gap but at mid range of Ratio (d84-d16)/d50 the Ribbon Density was low. This could be due to the roll compaction of MCC + 1%MgSt which was one of the two values which were in the mid range of the extremes for Ratio (d84-d16)/d50. Lubrication of the powder material results in low Ribbon Density in the end product. However, it showed that a combination of Roll Speed, Roll Gap and Ratio (d84-d16)/d50 affected the Ribbon Density in a non-linear way.

Table 6.19 shows the neurofuzzy rules developed for ribbon porosity in Trial B. This trial showed that at high Ratio (d84-d16)/d50 roll compaction produced low ribbon porosity, while at low Ratio (d84-d16)/d50 roll compaction produced high ribbon porosity. There is a larger amount of fines in the high Ratio (d84-d16)/d50 than at low Ratio (d84-d16)/d50. Hence the reported effects on ribbon porosity might be due to the amount of fines within the compact. A larger amount of fines in the powder material produced ribbon compacts with lower porosity.

Table 6.18 Neurofuzzy rules for Trial B Ribbon Density output.

SubModel:1
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.58)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is HIGH (0.69)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is HIGH (0.75)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.87)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is HIGH (0.54)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.63)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is LOW (0.63)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is LOW (0.64)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is LOW (0.72)

Table 6.19 Neurofuzzy rules for Trial B Ribbon Porosity output.

SubModel:2
IF Ratio (d84-d16)/d50 is LOW THEN Ribbon Porosity is HIGH (1.00)
IF Ratio (d84-d16)/d50 is HIGH THEN Ribbon Porosity is LOW (0.93)

Table 6.20 shows the neurofuzzy rules developed for the average nip angle in Trial B. It shows that combinations of Roll Speed, Roll Gap and Ratio (d84-d16)/d50 affected the average nip angle in a non-linear way.

Table 6.20 Neurofuzzy rules for Trial B Average Nip Angle (°) output.

SubModel:1
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ave Nip Angle (°) is HIGH (0.97)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ave Nip Angle (°) is LOW (1.00)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ave Nip Angle (°) is LOW (1.00)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ave Nip Angle (°) is HIGH (1.00)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ave Nip Angle (°) is HIGH (0.64)
IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ave Nip Angle (°) is HIGH (0.70)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ave Nip Angle (°) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ave Nip Angle (°) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ave Nip Angle (°) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ave Nip Angle (°) is LOW (0.53)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ave Nip Angle (°) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ave Nip Angle (°) is LOW (1.00)

The next highlighted key variable was Hausner's Ratio for the outputs of ribbon density and average nip angle. Trial H was chosen to exhibit the neurofuzzy rules developed for the outputs ribbon density and average nip angle (see Table 6.21 and Table 6.22).

Table 6.21 Neurofuzzy rules for Trial H Ribbon Density (kg/m^3) output.

SubModel:1
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND Hausner's Ratio is LOW THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND Hausner's Ratio is HIGH THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND Hausner's Ratio is LOW THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND Hausner's Ratio is HIGH THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND Hausner's Ratio is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND Hausner's Ratio is HIGH THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND Hausner's Ratio is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND Hausner's Ratio is HIGH THEN Ribbon D (kg/m3) is HIGH (1.00)

Table 6.22 Neurofuzzy rules for Trial H Average Nip Angle (^{o}) output.

SubModel:1
IF Hausner's Ratio is LOW THEN Ave Nip Angle (^{o}) is LOW (0.79)
IF Hausner's Ratio is MID THEN Ave Nip Angle (^{o}) is HIGH (0.66)
IF Hausner's Ratio is HIGH THEN Ave Nip Angle (^{o}) is HIGH (0.52)

Then Compressibility κ was a key variable for ribbon density, ribbon porosity, average maximum pressure and average nip angle. Trial G was chosen to exhibit the neurofuzzy rules developed for all the outputs affected by Compressibility κ (see Table 6.23 to Table 6.26).

Table 6.23 Neurofuzzy rules for Trial G Ribbon Density (kg/m^3) output.

SubModel:1
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND Compressibility κ is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND Compressibility κ is HIGH THEN Ribbon D (kg/m3) is HIGH (0.57)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND Compressibility κ is LOW THEN Ribbon D (kg/m3) is LOW (0.84)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND Compressibility κ is HIGH THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND Compressibility κ is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND Compressibility κ is HIGH THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND Compressibility κ is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND Compressibility κ is HIGH THEN Ribbon D (kg/m3) is LOW (1.00)

Table 6.24 Neurofuzzy rules for Trial G Ribbon Porosity output.

SubModel:2
IF Compressibility κ is LOW THEN Ribbon Porosity is LOW (1.00)
IF Compressibility κ is HIGH THEN Ribbon Porosity is HIGH (1.00)

Table 6.25 Neurofuzzy rules for Trial G Average Maximum Pressure (MPa) output.

SubModel:1
IF Compressibility κ is LOW THEN Ave. Max Pressure (MPa) is LOW (0.61)
IF Compressibility κ is HIGH THEN Ave. Max Pressure (MPa) is LOW (1.00)

Table 6.26 Neurofuzzy rules for Trial G Average Nip Angle (°) output.

SubModel:2
IF Compressibility κ is LOW AND Roll Gap (mm) is LOW THEN Ave Nip Angle (°) is LOW (1.00)
IF Compressibility κ is LOW AND Roll Gap (mm) is HIGH THEN Ave Nip Angle (°) is LOW (1.00)
IF Compressibility κ is HIGH AND Roll Gap (mm) is LOW THEN Ave Nip Angle (°) is HIGH (1.00)
IF Compressibility κ is HIGH AND Roll Gap (mm) is HIGH THEN Ave Nip Angle (°) is HIGH (1.00)

It is interesting to note that the Angle of Wall friction was a key variable for ribbon density, average maximum pressure and average nip angle. Moreover it was most frequently highlighted as a key variable for the average maximum pressure. Trial F was chosen to exhibit the neurofuzzy rules developed for all the outputs affected by Angle of Wall Friction (see Table 6.27, Table 6.28 and Table 6.29).

Table 6.27 Neurofuzzy rules for Trial F Ribbon Density (kg/m^3) output. AoIWF is Angle of Wall Friction.

SubModel:1
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND AoIWF is LOW THEN Ribbon D (kg/m3) is LOW (0.52)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND AoIWF is MID THEN Ribbon D (kg/m3) is LOW (0.50)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is LOW AND AoIWF is HIGH THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND AoIWF is LOW THEN Ribbon D (kg/m3) is HIGH (0.89)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND AoIWF is MID THEN Ribbon D (kg/m3) is HIGH (1.00)
IF Ratio of M:D is LOW AND Roll Speed (rpm) is HIGH AND AoIWF is HIGH THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND AoIWF is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND AoIWF is MID THEN Ribbon D (kg/m3) is HIGH (0.87)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is LOW AND AoIWF is HIGH THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND AoIWF is LOW THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND AoIWF is MID THEN Ribbon D (kg/m3) is LOW (1.00)
IF Ratio of M:D is HIGH AND Roll Speed (rpm) is HIGH AND AoIWF is HIGH THEN Ribbon D (kg/m3) is LOW (0.99)

Table 6.28 Neurofuzzy rules for Trial F Average Maximum Pressure (MPa) output. AoIWF is Angle of Wall Friction.

SubModel:1
IF Roll Speed (rpm) is LOW AND AoIWF is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
IF Roll Speed (rpm) is LOW AND AoIWF is MID THEN Ave. Max Pressure (MPa) is HIGH (1.00)
IF Roll Speed (rpm) is LOW AND AoIWF is HIGH THEN Ave. Max Pressure (MPa) is HIGH (1.00)
IF Roll Speed (rpm) is MID AND AoIWF is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
IF Roll Speed (rpm) is MID AND AoIWF is MID THEN Ave. Max Pressure (MPa) is LOW (0.75)
IF Roll Speed (rpm) is MID AND AoIWF is HIGH THEN Ave. Max Pressure (MPa) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND AoIWF is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND AoIWF is MID THEN Ave. Max Pressure (MPa) is LOW (1.00)
IF Roll Speed (rpm) is HIGH AND AoIWF is HIGH THEN Ave. Max Pressure (MPa) is LOW (1.00)
SubModel:3
IF AoISR is LOW THEN Ave. Max Pressure (MPa) is LOW (1.00)
IF AoISR is HIGH THEN Ave. Max Pressure (MPa) is HIGH (0.87)

Table 6.29 Neurofuzzy rules for Trial F Average Nip Angle (°) output. AoIWF is Angle of Wall Friction

SubModel:1
IF AoIWF is LOW THEN Ave Nip Angle (°) is LOW (0.81)
IF AoIWF is MID THEN Ave Nip Angle (°) is HIGH (0.90)
IF AoIWF is HIGH THEN Ave Nip Angle (°) is LOW (0.63)

The next stage would be to choose the trials which would be brought into *INForm* model training, prediction and experimental validation. The chosen trials are Trial D,

Trial I and Trial J. Trial D was chosen because it was interesting to observe the capabilities of the software when given only the formulation and roll compaction process parameters as inputs to produce predictive models. Trial I was chosen to form predictive models to relate the material characteristics which the pharmaceutical industry are interested in to the output property of the roll compaction process. Trial J was chosen because it minimised the amount of inputs, the two characterisation experiments were the most straightforward and they are frequently obtained experimentally in the pharmaceutical industry.

6.4 Summary

The success of each trial proved that the *FormRules* software can take any amount of inputs within the rule of thumb of 2-3 inputs per data set. It can produce neurofuzzy rules which highlight the key variables which affect the output property. However the user should keep in mind the practicality and relevance of the inputs used and highlighted in the software because the suggested relationships are not always physically appropriate. As soon as the database is set up inside *FormRules*, different types of trials are possible. The flexibility of the software allows the user to set up trials rapidly according to what is required by the user.

This chapter shows how *FormRules* is used for the initial formulation and selection of the roll compaction process condition to give the desired roll compaction output

property via neurofuzzy rules.   11 types of trials on the different combinations of input variables were also presented.   3 of the trials were chosen for *INForm* model training, prediction and experimental validation.   They are Trial D, Trial I and Trial J, which will be discussed in the following chapter.

7 Intelligent Software: *INForm* – Intelligent Formulation

ABSTRACT

INForm software is an example of intelligent software which detects cause-and-effect relationships from experimental data sets. It develops predictive models and subsequently optimises these models. It can be used for conducting "what if" experiments. This means the consequence of changing the formulation or the processing condition within a given range on the output property can be obtained. Furthermore, the "optimisation function" on *INForm* model can be used to search for the combination of formulation and process conditions which will give the closest value to the output properties required by the user. The roll compaction process parameters were predicted for specific tablet formulation and output properties. Experimental work validated that the predictions were accurate. Then the ability of the *INForm* models to predict output properties was investigated and compared to Johanson's model for prediction of the nip angle. The *INForm* model showed a better agreement with the experimental results compared with Johanson's model prediction.

This section describes the method of training *INForm* software to develop models, shows how the model could be of use and compares the *INForm* model prediction to Johanson's model prediction.

The *INForm* programme utilises ANNs, GA, fuzzy logic and neurofuzzy logic (see Sections 2.2.1 to 2.2.4). As mentioned earlier ANNs are mathematical systems that mimic the way in which the human brain processes information (Erb, 1993). They are used to generate 'black box' models that link inputs to outputs from experimental data sets and detect relationships. Subsequently GA and fuzzy logic are employed for optimisation in the multidimensional space.

This chapter describes the method of using the software for roll compaction of pharmaceutical excipients listed in Chapter 3. Alongside this, the rationale behind the training parameters for the models is presented. Then as stated in Chapter 6 the key inputs for Trial D, Trial I and Trial J will be fed into three separate models for training and optimisation of each of the models. The trained model is then used in a case study to search for the combination of process conditions for a required output property of a formulation. A comparison between Johanson's theory (Chapter 5) and the *INForm* model was conducted.

7.1 Practical Application

The aim of using *INForm* software was to develop models designed to predict roll compaction process parameters for a specified formulation and roll compaction output property. Trial D was used as an example for the training of *INForm*.

7.1.1 Training *INForm* - Trial D

Before the software training could be conducted, the bulk data obtained from the roll compaction experiments (Chapter 4) on five different types of formulation were divided into three groups. The first group was used as the validation data. The validation data were a data set which was withheld from the software and consisted of 20% of the bulk data set (12 data sets). The validation data sets were carefully chosen from the spread of data so as not to disrupt and form large missing regions in the experimental space. The second group of data was called the test data and consisted of 20% of the balance of the bulk data (test data = 10 data sets and training data = 41 data sets) after the validation data had been withheld. The same care in choosing the data set was practiced.

The test data was the data set used by the software to check, if the errors in the training of the model were decreasing with every iteration cycle conducted. If the errors did increase the software calculation stopped and the calculation was restarted to obtain a better model. The test data set could be manually chosen or could be automatically chosen by the software. To ensure that the model would train well, there should be at least 2-3 training data set per input. In the *INForm* training of trial D there were 10 training data sets per input, which is a good ratio to model with.

Figure 7.1 shows the general steps in the training of the *INForm* programme. The steps through the software interface are explained in Appendix 5. Initially the data

need to be in a text file to be fed into the software. Secondly the highlighted inputs/outputs from Trial D were set in *INForm*. Next the model was trained separately for each output. This was done to allow an independent prediction for each output property. Then the model was developed by choosing the correct set of test data from the bulk data and also using the right set of training parameters (training parameters are explained in Section 9.5.3). Initially the test set data was chosen randomly from the bulk data set, while training parameters were used as default. Next the model was trained. Figure 9.32 shows the window which contains the neural network training results. Then the model was assessed for its ability to predict using the steps explained in Section 7.1.2.

Once the model was found to be predicting data successfully, the model was used to construct surface response graphs (3D graphs), make "what if" predictions and for optimisation of Trial D, Trial I and Trial J. The response surface graphs show how well the data covers different regions of space, while the "what if" predictions can be used to predict the output properties for specified formulation and roll compaction process conditions.
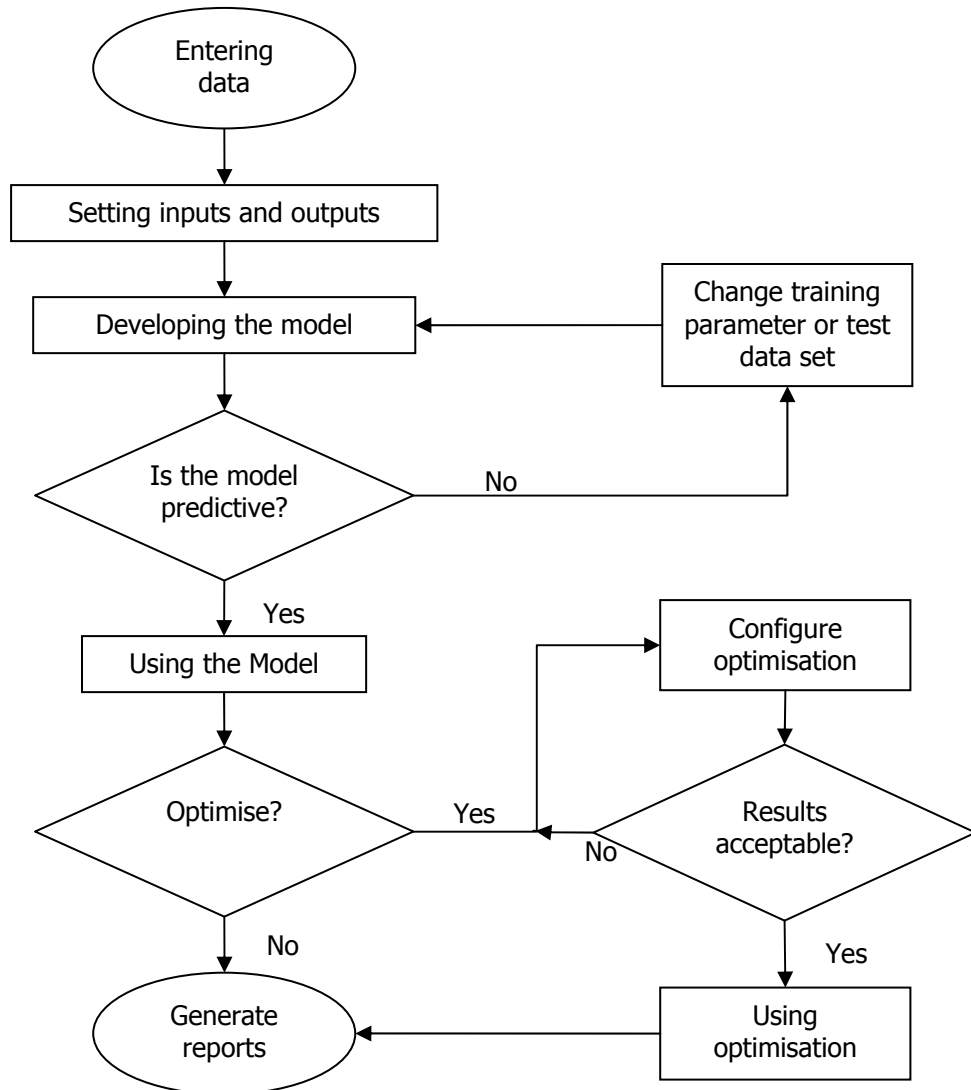
Figure 7.1 Flowchart showing the steps in training and optimisation of *INForm* programme.

7.1.2 The Model Assessment – Is the Model Predictive?
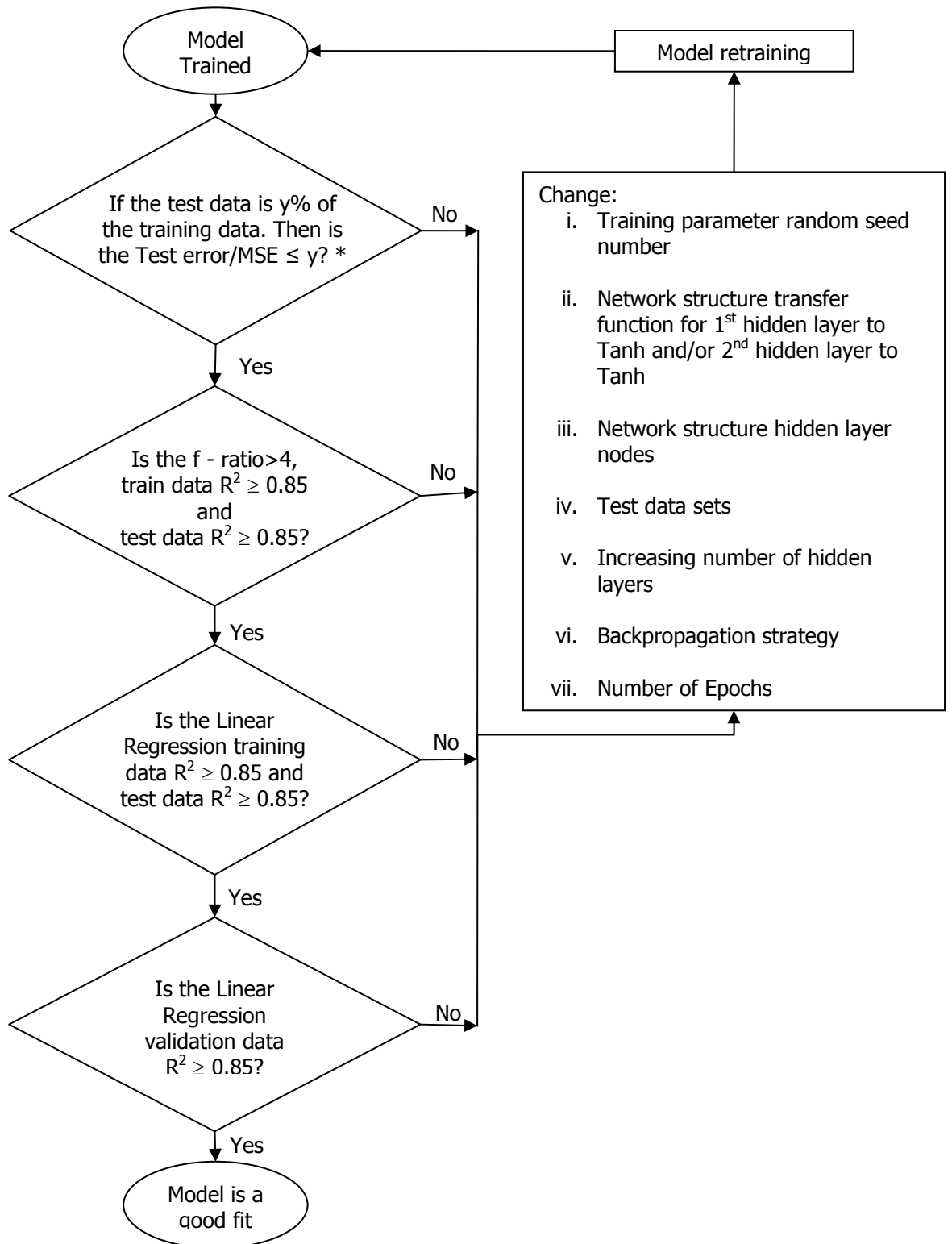
Figure 7.2 Flow of trained model assessment and readjustment of the network structure. *This condition depends on the size of the bulk data and the size of the test data taken out (see text for explanation).

After the model was trained, it was assessed for ability to predict. Figure 7.2 shows the sequence of model assessment. Firstly, the MS error and the test error were observed. It is very rare that the software will be able to achieve an error value of 10^{-4}. Therefore Colburn (2006) stated a couple of conditions which could be used to ensure a model was sufficiently trained. The basic condition is that if the test data consists of 10% of the training data, then the ratio of test error to MS error should be less than 10 to obtain an acceptable trained model. However, this condition depends on the size of the bulk data set and the size of the test data set taken out. If the size of the bulk data set was extremely large then 50% of the test data would not affect the coverage of space in the training data. In this case the Test Error should be equal to or just a little more than the MS Error to obtain a sufficiently trained model.

Next the model statistics were checked for positive results. Table 7.1 showed a positive result except for the test data R^2 for average maximum pressure. A positive result was indicated by an f-ratio of more than 4, "train data R^2" greater than 0.85 and the "test data R^2" greater than 0.85. A high R^2 value does not really mean that the model has been trained exceptionally well. It can mean that it is well trained or was overtrained. An overtrained model will not be able to predict unseen data because it has memorized the training and test data and has lost its ability to predict. The f-ratio and the *INForm* R^2 values were explained in section 6.1.2.   If the results were negative the model was retrained by changing either the test set data or the training parameters. The adjustment to the training parameter was

explained in the section 9.5.3. If the model gave a positive result (as shown in Table 7.2) the next stage would be to assess the predictability of the model on the training and test data.

Table 7.1 Summary of Model Statistics from the *INForm* model training result.

Output Property	f-ratio	Train data R^2	Test data R^2
Ribbon Density (kg/m^3)	25.19	0.99	0.93
Ribbon Porosity	8.07	0.97	0.93
Average Maximum Pressure (MPa)	14.30	0.98	0.75
Average Nip Angle (^{o})	89.79	1.00	0.91

Table 7.2 Summary of Model Statistics from the *INForm* model retraining result.

Output Property	f-ratio	Train data R^2	Test data R^2
Ribbon Density (kg/m^3)	33.13	0.99	0.94
Ribbon Porosity	21.29	0.99	0.92
Average Maximum Pressure (MPa)	12.53	0.98	0.85
Average Nip Angle (^{o})	43.20	0.99	0.91

This was conducted by constructing a linear regression line for the predicted against observed data. Figure 7.3 and Figure 7.4 show examples of this linear regression line for Trial D. The slope of the linear regression line should be unity and the R^2 should be equal to or more than 0.85. If this requirement was not fulfilled the model was retrained. But as the example shows the requirement was fulfilled. Following this the model was assessed to ensure that the model trained was not memorizing the train and test data.

This involved constructing a similar regression line for validation data (Figure 7.5). The same requirement applies to the slope of this graph and the R^2. If these requirements were fulfilled then models can be used for prediction. If it was not fulfilled than the model will be retrained. The linear regression R^2 values for all the output properties were above 0.85. This means that the model trained for the Trial D using default parameters except for the 1st hidden layer transfer function changed to Tanh was a good model.
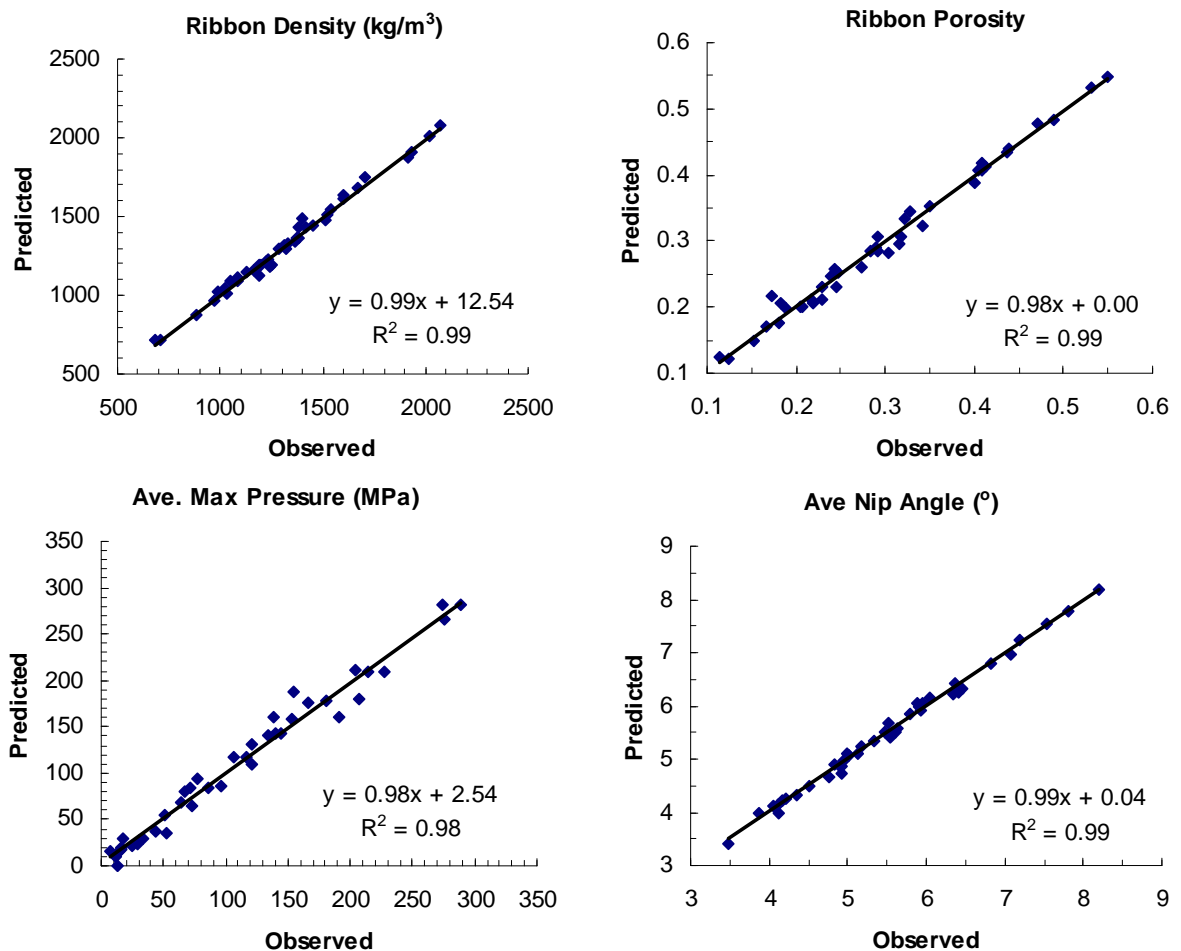
Figure 7.3 Output property linear regression for training data predicted against observed for Trial D.
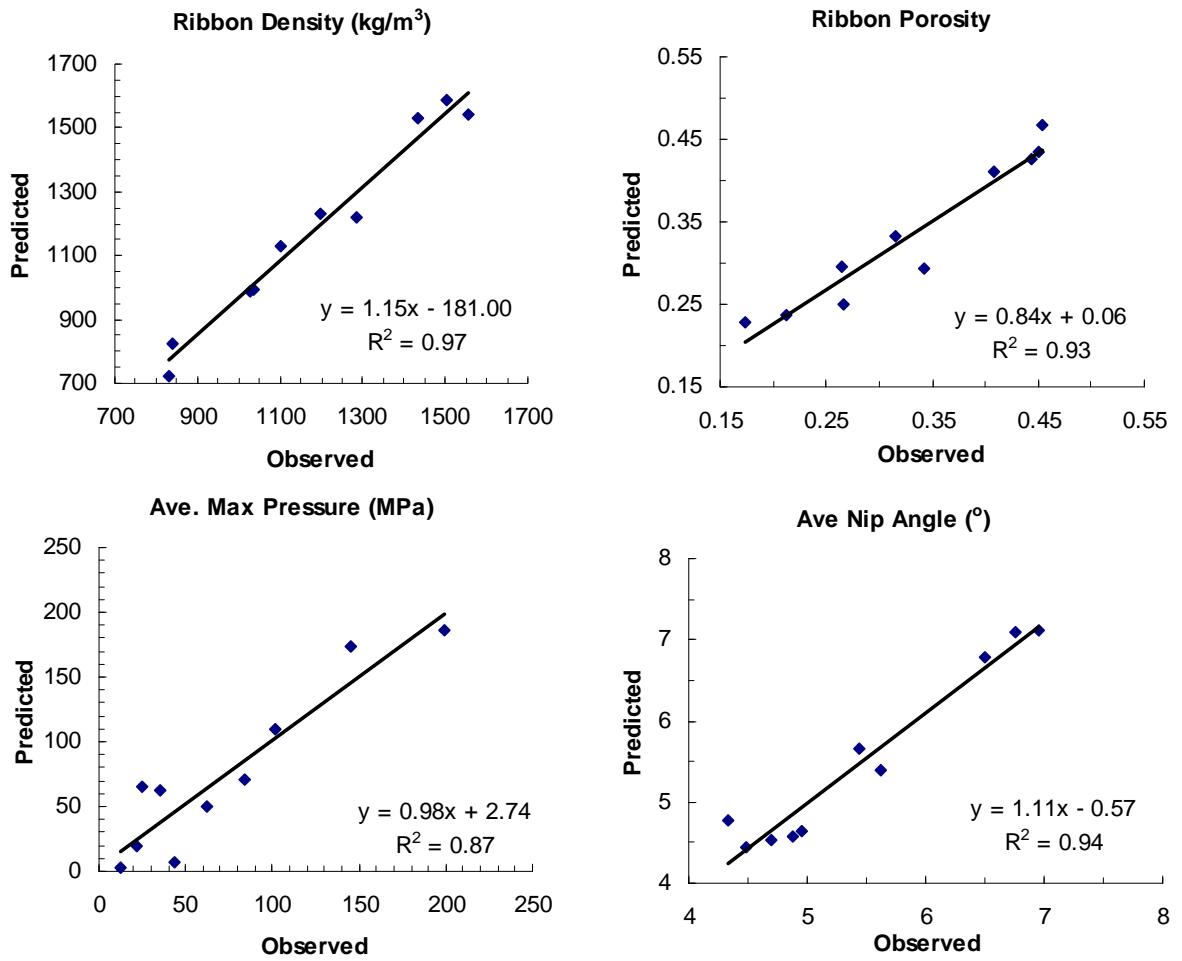
Figure 7.4 Output property linear regression for test data predicted against observed for Trial D.
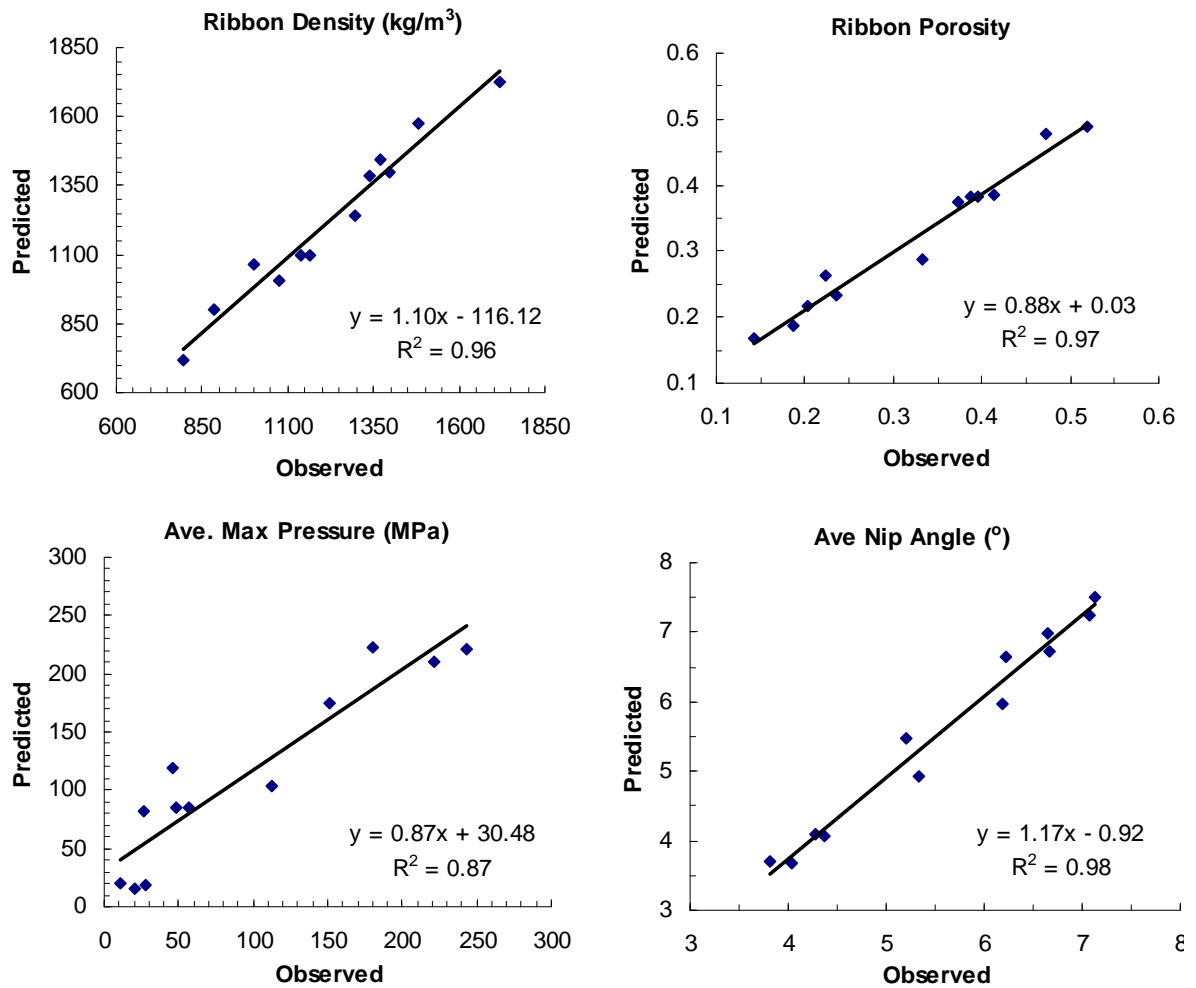
Figure 7.5 Output property linear regression for validation data predicted against observed for Trial D.

7.1.2.1 f-ratio and *INForm* R^2

Similar with *FormRules* f-ratio and R^2, hence please refer to Section 6.1.2. However in *INForm* software training, the R^2 value should be at least 75% for a good quality model, this condition was increased to 85% for roll compaction. This is because the standard deviations within the training and testing datasets were small. Hence, the user could expect a better quality model which could give predictions with a high R^2 value.

7.1.3 Using the Model

Section 7.1.1 described the training and assessment of the trained model. Once this model was trained it could be used for "what if" predictions or for optimisation. "What if" predictions are a feature of this software which can be used to predict the output properties for specified formulation and roll compaction process parameters. On the other hand, optimisation is a feature which can be used to predict the roll compaction process parameters for a specified formulation and output property or the formulation for a specified roll compaction process and output property.

7.1.3.1   "What if" Predictions

"What if" predictions can be used to predict the output properties resulting from specific formulation and roll compaction process parameters.   Before "what if" predictions can be conducted it is advisable to observe the response surface graphs. The response surface graphs show how well the data covers different regions of space, hence allowing the user to be able to use the "what if" predictions within the range of roll compaction process parameters to obtain reliable output predictions. It is important to note that *INForm* software does not extrapolate results well.

Appendix 5, Section 9.5.3 shows the method of obtaining the response surface graphs and conducting "what if" predictions. In brief, a set of data for a formulation was chosen to consult. The roll compaction process parameters were then fed into the "given" column. Prediction of the output properties can then be conducted.

7.1.3.2   Optimising the Formulation
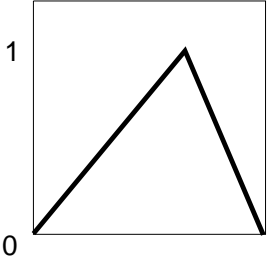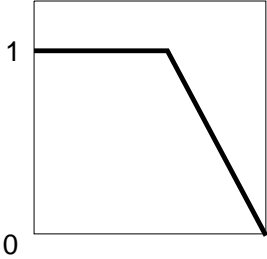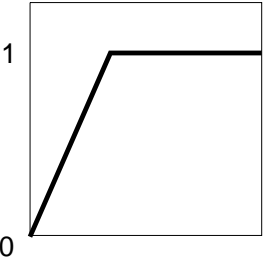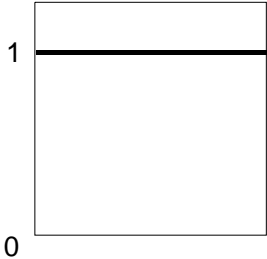
Optimisation can be used either to predict the roll compaction process parameters from a specified formulation and output property or to predict the formulation from a specified roll compaction process property and output property. In this study the roll compaction process parameters were predicted from a specified formulation and output property. The process of optimisation is shown in Figure 7.6. Section 9.5.4.3 (Appendix 5) shows a simplified tutorial-like method to work with the software interface.

A data set close to the desired output property was chosen for consultation. This speeds up the optimisation process. The next stage involves adjusting the optimisation configuration. This feature allows the user to specify certain objectives for the software optimisation. The weight can be set on a scale of 0 to 10. This shows a relative importance between conflicting objectives. For example the user is trying to predict the roll compaction process parameters required for a combination of output properties. The user can then specify in what order the output properties should be satisfied.

Next the desired output property is fed into the Mid1 and Mid2 columns (Figure 9.42). Both these values should be the arithmetic mean of the Minimum and Maximum values. At default, Min and Max would display the global minimum and maximum values. But after the Mid1 and Mid2 are changed these values should be adjusted to have a value within the global minimum/maximum range and to keep

Mid1 and Mid2 as the arithmetic mean. These values for Min, Mid1, Mid2 and Max are used to direct the optimisation program to calculate a result within a desired range. The desirability function further assists the optimisation program to focus on either the lower end, higher end, near mid values or just no preference. There are four types of desirability functions; 'tent', 'down', 'up' and 'flat' and they are described in Table 7.3.

Table 7.3 Description for each desirability function

Desirability function	Sketch	Description
Tent		The values between Mid1 and Mid2 are of 100% desirability and Min/Max are of 0% desirability. The desirability increases linearly from Min to Mid1 and decreases from Mid2 to max.
Down		The values below Mid2 are of 100% desirability and Max is 0% desirability. The desirability decreases linearly from Mid2 to Max.
Up		The values above Mid1 are of 100% desirability and Min is 0% desirability. The desirability increases linearly from Min to Mid1.
Flat		All values have equal desirability.
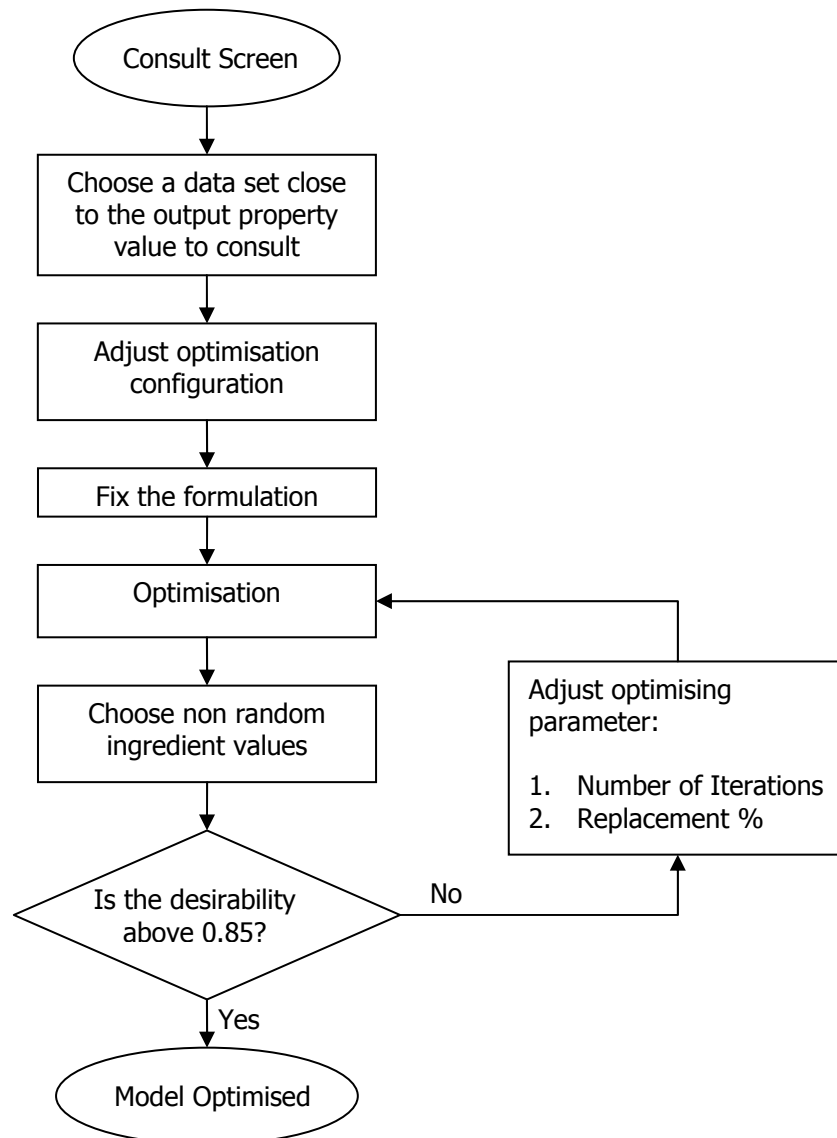
Figure 7.6 Flowchart for optimisation of roll compaction process.

Once the optimisation configuration has been adjusted, the optimisation can be conducted on the model using given values (also known as the values used for consultation).   Next the formulation values are fixed and non-random starting ingredient values are used for optimisation.   The optimisation desirability result should be above 85%.   If this requirement is not fulfilled, then the optimisation parameters can be changed.   These parameters control the GA's operation.   The

number of iterations is the number of calculation cycles which are performed before the optimisation is halted. This value can be increased if the calculation is found to stop too early. The replacement % is the number of 'genetic material' (calculated values) which will be replaced in each iteration. A large replacement % will mean a large sample is taken from the data spread and there will be a longer calculation time. A smaller replacement % is used to focus the calculation to the best value.

7.2 Results and Discussions

7.2.1 Discussing the Challenges of using the Connections Tab to use the Key Variables Highlighted by *FormRules* Software

The connections tab is shown by Figure 9.30 in the Appendix 5, Section 9.5.3.6. It allows the user to break the connection a particular input has with the specific output model network. If the input's connection to the specific output is broken, then it is no longer used in building the model for that respective output. This gives the user more options of building a better model which employs the results produced by the *FormRules* software.

Table 7.4 Example 1 for *INForm* model training example of Trial I for the connectivity function.

Example	Output property	Training parameters which were adjusted (RS is random seed, TF is transfer function)	Activated input connection	Epoch	Train set mean sq error	Test set mean sq error	Test error/MSE	f-ratio	Train Set R-squared	Test Set R-squared	system structure
Example 1	Ribbon density (kg/m³)	RS=9823; nodes=2.	Ratio M:D, Compressibility κ, roll speed, MgSt% and Roll Gap	992	0.0037	0.0020	0.5360	1.26683	-51.7093	-19.0516	I(5) - HL(2) - O(1)
	Ribbon porosity	RS=500; nodes=2.	Roll Speed, Roll Gap and Ratio (d84-d16)/d50	167	0.0096	0.0124	1.2932	1.40112	-182.431	17.0702	I(3) - HL(2) - O(1)
	Ave. Max Pressure (MPa)	RS=500, nodes=2.	Angle of Wall Friction, Roll Speed, Roll Gap and Compressibility κ	999	0.0215	0.0096	0.4459	2.59824	24.609	22.2349	I(4) - HL(2) - O(1)
	Ave Nip Angle (°)	RS=500; nodes=2.	Hausner's Ratio, Ratio M:D, Roll Speed and Roll Gap	982	0.0019	0.0116	5.9897	114.206	98.2167	84.679	I(4) - HL(2) - O(1)

Table 7.5 Example 2 *INForm* model training example of Trial I for the connectivity function.

Example	Output property	Training parameters which were adjusted (RS is random seed, TF is transfer function)	Activated input connection	Epoch	Train set mean sq error	Test set mean sq error	Test error/MSE	f-ratio	Train Set R-squared	Test Set R-squared	system structure
Example 2	Ribbon density (kg/m^3)	Random seed=5765; nodes=3.	Ratio M:D, Compressibility κ, Roll Speed, MgSt% and Roll Gap	652	0.0015	0.0060	4.1222	0.586779	-33.8567	-31.8859	I(5) - HL(3) - O(1)
	Ribbon porosity	Random seed=2345; 1st hidden layer transfer function=Tanh; output transfer funciton=Tanh; nodes=3.	Roll Speed, Roll Gap and Ratio (d84-d16)/d50	616	0.0036	0.0078	2.1651	0.684784	-178.851	1.44392	I(3) - HL(3) - O(1)
	Ave. Max Pressure (MPa)	Random seed=1; 1st hidden layer transfer function=tanh, output transfer function=tanh; nodes=3.	Angle of Wall Friction, Roll Speed, Roll Gap and Compressibility κ	989	0.0112	0.0394	3.5238	1.1284	6.76101	41.2454	I(4) - HL(3) - O(1)
	Ave Nip Angle ($^\circ$)	Random seed=500, nodes=2.	Hausner's Ratio, Ratio M:D, Roll Speed and Roll Gap	982	0.0019	0.0116	5.9897	114.206	98.2167	84.679	I(4) - HL(2) - O(1)

Table 7.6 Example 3a for *INForm* model training example of Trial I for the connectivity function.

Example	Output property	Training parameters which were adjusted (RS is random seed, TF is transfer function)	Activated input connection	Epoch	Train set mean sq error	Test set mean sq error	Test error/ MSE	f-ratio	Train Set R-squared	Test Set R-squared	system structure
Example 3a	Ribbon density (kg/m^3)	RS=5230, 1st hidden layer TF=Tanh, Output TF=tanh, Nodes=3. Standard incremental	Ratio M:D, Compressibility κ, Roll Speed, MgSt% and Roll Gap	439	0.0105	0.0085	0.8060	11.0146	89.8041	86.6748	I(5) - HL(2) - O(1)
	Ribbon porosity	Same as above	Same as above	230	0.0218	0.0185	0.8487	8.01836	83.0834	81.6824	I(5) - HL(2) - O(1)
	Ave. Max Pressure (MPa)	Same as above	Same as above	258	0.0126	0.0068	0.5426	18.5488	92.2128	91.8507	I(5) - HL(2) - O(1)
	Ave Nip Angle (^{o})	Same as above	Same as above	325	0.0041	0.0081	1.9678	44.7334	96.2313	89.3625	I(5) - HL(2) - O(1)

Table 7.7 Example 3b for *INForm* model training example of Trial I for the connectivity function.

Example	Output property	Training parameters which were adjusted (RS is random seed, TF is transfer function)	Activated input connection	Epoch	Train set mean sq error	Test set mean sq error	Test error/ MSE	f-ratio	Train Set R-squared	Test Set R-squared	system structure
Example 3b	Ribbon density (kg/m^3)	RS=5230, 1st hidden layer TF=Tanh, Output TF=tanh, Nodes=3. Standard incremental	Ratio M:D, Compressibility κ, Roll Speed, MgSt% and Roll Gap	439	0.0105	0.0085	0.8060	0.818794	-205.445	-228.645	I(5) - HL(2) - O(1)
	Ribbon porosity	RS=5230, 1st hidden layer TF=Tanh, Output TF=tanh, Nodes=3. Standard batch	Roll speed, Roll Gap and Ratio (d84-d16)/d50	624	0.0100	0.0118	1.1821	31.0527	92.2511	88.3136	I(3) - HL(2) - O(1)
	Ave. Max Pressure (MPa)	Same as above	Same as above	999	0.0245	0.0070	0.2863	15.1446	84.8313	91.6239	I(3) - HL(2) - O(1)
	Ave Nip Angle (º)	Same as above	Same as above	769	0.0057	0.0167	2.9276	46.8081	94.7577	77.9888	I(3) - HL(2) - O(1)

However it was found out that using this feature did not produce positive results. For Trial I, the *FormRules* software showed that for each of the output properties different input parameters were important. These key parameters were connected to the specific output property in the *INForm* model training. The model training proved to be complicated and difficult to use. Although the interactive training mode which can be chosen is able to give flexibility in changing the training parameters during training, the comparison of MS Error and Test Error does not always guarantee a good model. The f-ratio and the R^2 can only be seen after all the models had been trained and it is difficult to gauge the acceptability of the model without observing these values.

This feature can be observed from the *INForm* model training examples of Trial I shown in Tables 7.4 to 7.7. The training was conducted separately and interactively for each output property. This meant that training was conducted for each output sequentially after another, but without being allowed to check the training result of the currently trained output property. The key input variables for each output property were activated for the respective output property (see Table 6.15 for the key input variables). Example 1 (Table 7.4) shows the training parameters adjusted for each of the output properties. The ratio of test error to MSE was less than 10 for all the output properties, yet only the Average Nip Angle model produced an f-ratio of more than 4, Train R^2 more than 85 and Test R^2 of 85. The model training for the other 3 outputs failed i.e. for ribbon density, ribbon porosity and average maximum pressure. Example 2 (Figure 7.5) shows the next simulation test on the

same Trial I. The training parameters for Average Nip Angle were kept constant, but were adjusted for the first 3 outputs. Example 2 was conducted to achieve a higher ratio of test error over MSE. However it still gave failed outputs for the first 3 output properties.

It was difficult to conduct the training without being able to observe the immediate training result. Hence, to make the training of this model easier and to allow the user to observe the results of the training as soon as possible the training method was changed. Initially only the first output property was trained using a specific training parameter combination (Table 7.6). The results for the rest of the output properties were ignored. After the training parameters for the first output property were found to be successful, the second output property training was conducted while keeping the successful training parameters for the first output property constant (Table 7.7).

Although the first training parameters were kept constant for the first output property and changed for the second output property the result produced was unsuccessful for the first output property, while it was successful for the second output property. The results indicated that each of the four output property training events had some influence on each of the others. This makes the model complicated and difficult to train because each of the output properties will have different key inputs according to *FormRules* models, but the results of the *INForm* training showed that all four output property models will be inter-related. Thus the

different combinations of training parameters of each output property need to be found and this will take a lengthy time with a lot of effort required from the user. If this feature was not used, i.e. the inputs were all activated for the model training (as was reported) the training process was usually simpler, quicker and successful.

7.2.2 The *INForm* Assessment Summary on Trial D, Trial I and Trial J

Table 7.8 shows the summary of the assessments of the Trials D, I and J. The table contains the MS error and the Test error which are part of the first step of model training. The model statistics (i.e. f-ratio, Training R^2 and Test R^2) are summarised and that Trials D, I and J were trained successfully. The regression lines are also summarised in this table. It shows that all three trials were assessed and validated to be good models.

These models can be safely used for "what if" predictions and optimisations. The Trial D system structure for all the output properties are all similiar with 4 inputs, 5 hidden layers and 1 output. The Trial I system structure for all output properties except for ribbon porosity has 8 inputs, 2 hidden layers and 1 output. The Trial I system structure for ribbon porosity had 8 inputs, 3 hidden layers and 1 output. The Trial J system structures for all the output properties are all similar with 6 inputs, 3 hidden layers and 1 output.

Table 7.8 The summary of assessments for Trials D, I and J.

Trials	Output property	Number of Epochs	MS Error	Test Error	System structure	Model Statistics			Linear regression lines		
						f-ratio	Train R^2	Test R^2	Train R^2	Test R^2	Validation R^2
Trial D	Ribbon Density	998	0.0009	0.0040	I(4) - HL(5) - O(1)	33	0.99	0.94	0.99	0.97	0.96
	Ribbon Porosity	984	0.0017	0.0086	I(4) - HL(5) - O(1)	21	0.99	0.92	0.99	0.93	0.97
	Average Max Pressure	375	0.0037	0.0127	I(4) - HL(5) - O(1)	13	0.98	0.85	0.98	0.87	0.87
	Average Nip Angle	344	0.0007	0.0069	I(4) - HL(5) - O(1)	43	0.99	0.91	0.99	0.94	0.98
Trial I	Ribbon Density	408	0.0022	0.0044	I(8) - HL(2) - O(1)	42	0.98	0.93	0.98	0.94	0.86
	Ribbon Porosity	176	0.0030	0.0102	I(8) - HL(3) - O(1)	12	0.98	0.90	0.98	0.91	0.89
	Average Max Pressure	153	0.0098	0.0096	I(8) - HL(2) - O(1)	14	0.94	0.89	0.94	0.94	0.88
	Average Nip Angle	998	0.0011	0.0042	I(8) - HL(2) - O(1)	86	0.99	0.94	0.99	0.95	0.96
Trial J	Ribbon Density	999	0.0011	0.0029	I(6) - HL(3) - O(1)	54	0.99	0.95	0.98	0.97	0.94
	Ribbon Porosity	239	0.0044	0.0084	I(6) - HL(3) - O(1)	17	0.97	0.92	0.96	0.93	0.87
	Average Max Pressure	999	0.0083	0.0059	I(6) - HL(3) - O(1)	11	0.95	0.93	0.97	0.90	0.93
	Average Nip Angle	342	0.0007	0.0067	I(6) - HL(3) - O(1)	87	0.99	0.91	1.00	0.93	0.96

7.2.3 "What if" Predictions

The response surface graphs for trial D are shown in Figures 7.7 to 7.10. The data points are represented by the red triangles. The response graphs display the extrapolation of the experimental ranges of the formulations. This is because the software can understand the global maximum and global minimum from the total data sets which it has been trained on, but cannot differentiate between the specific maximum and minimum ranges for the individual formulations. Hence the graphs show the way the data covers different regions of space as well as giving an extrapolation of the results to cover the global maximum and minimum of the experimental range. The extrapolated regions should not be taken as reliable information for the "what if" predictions.

Certain graphs show negative values for example in the ribbon porosity 3D graph for MCC (Figure 7.8) and average maximum pressure 3D graph for MCC + 1%MgSt (Figure 7.9). These negative values were outside the range of experiments. The model has conducted an extrapolation of the results and has concluded improbable results.

Table 7.9 shows the "what if" predictions. Three examples were given. Each example was given a best match for "Ingredients" and "Properties". This command instructs the model to interpolate and find the closest result to either the "ingredients" or "properties". Using the best match for "ingredients" allows one to

observe what possible output property would be expected if one conducted a certain experiment on a specified formulation. The best match for "properties" gave the formulation and roll compaction process parameter values which result in the most similar output values.

In Example 1 (Table 7.9), the DCPA formulation was chosen and interpolation of the data was conducted. The prediction was in agreement with the expected trends in the roll compaction process. Comparing the best match ingredient to the "what if" result, the constant roll gap and increasing roll speed would produce a ribbon of lower density, higher porosity, lower roll compaction pressure and lower nip angle. The best match properties found the closest match to the output results from the data base and given the formulation and inputs related to it.

Both examples 2 and 3 (Table 7.9), show the extrapolation of the roll compaction process parameters. In example 2 the roll compaction roll speed is exaggerated (25 rpm) to show that the prediction of the model was not limited by the range of the roll compaction process parameter. However it is important to note that the results are not reliable in this case. In practice the roll compactor would not be able to compact at that roll compaction speed, but if it did the ribbon density should be lower not higher, while the ribbon porosity should be higher, than the best match "ingredients" result. Example 3 showed that even a relatively small extrapolation out of the range of process parameters gives an incorrect result, as shown by the roll compaction pressure results.

The response surface can be used to plan the "what if" predictions and also plan the experiment. Next we can deduce that the predictions that are well within the process parameters are reliable. But extrapolations out of the process conditions are not reliable. The best match ingredients command is useful when we need to check if the predicted values are reliable. The best match property command can be used to search for the formulation and roll compaction process parameters contributing to it. But the formulation could not be specified to obtain the roll compaction process parameters. This could be conducted in the optimise section (Section 7.1.3.2).
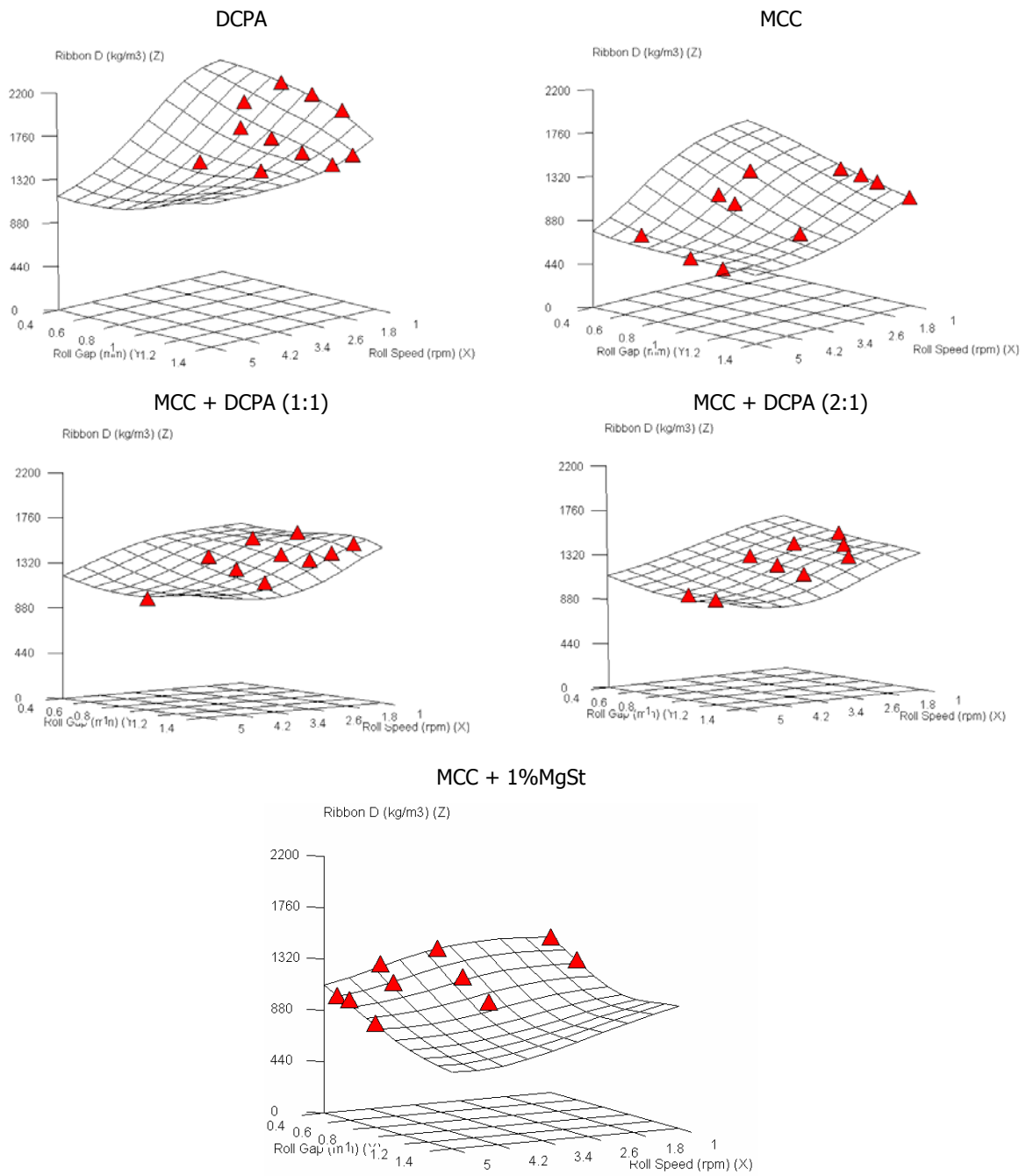
DCPA

MCC

MCC + DCPA (1:1)

MCC + DCPA (2:1)

MCC + 1%MgSt

Figure 7.7 The response surfaces for ribbon density. The red triangles are points from the train and test data set.

DCPA

MCC

MCC + DCPA (1:1)
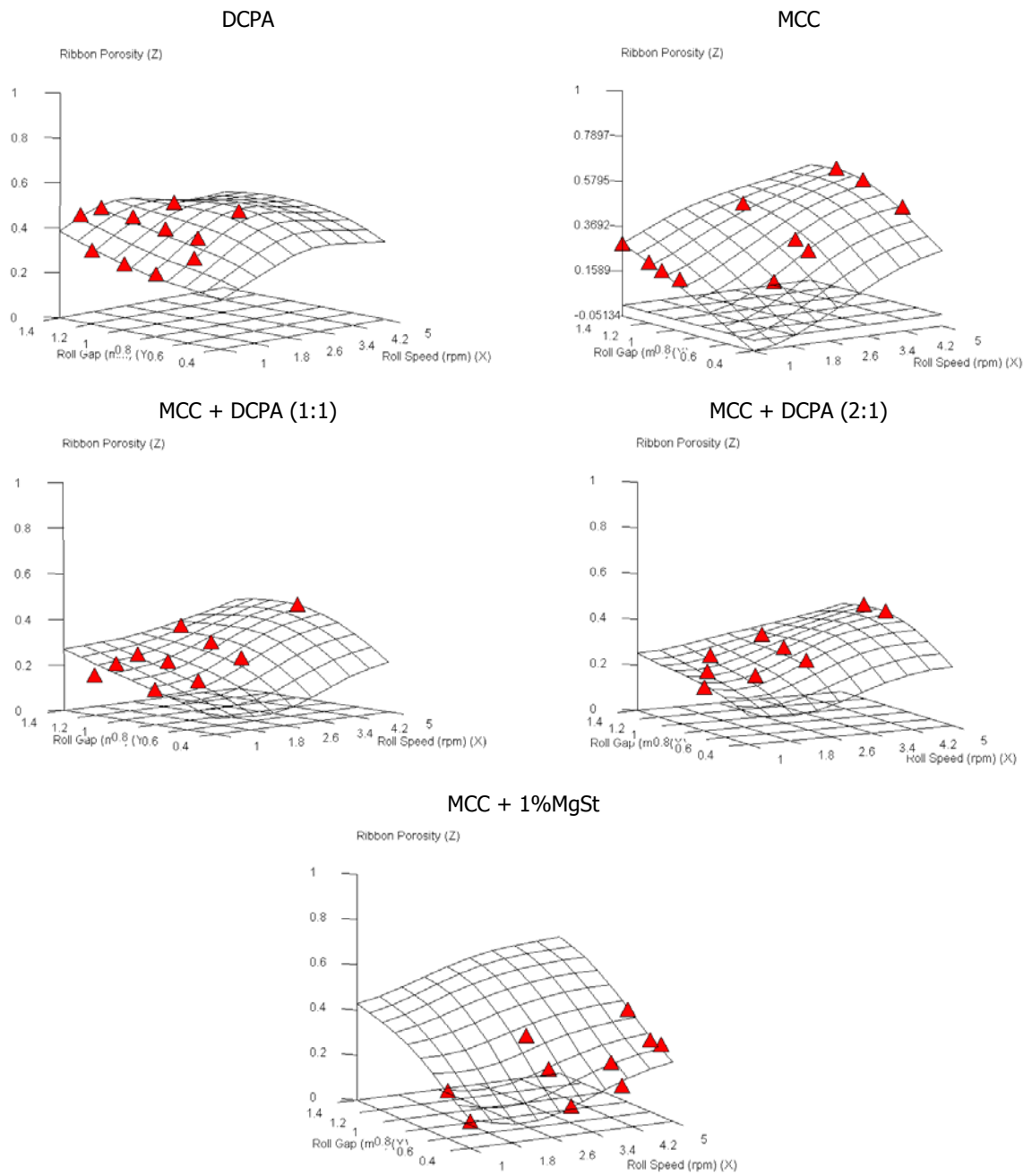
MCC + DCPA (2:1)

MCC + 1%MgSt

Figure 7.8 The response surfaces for ribbon porosity. The red triangles are points from the data sets. The red triangles are points from the train and test data set.
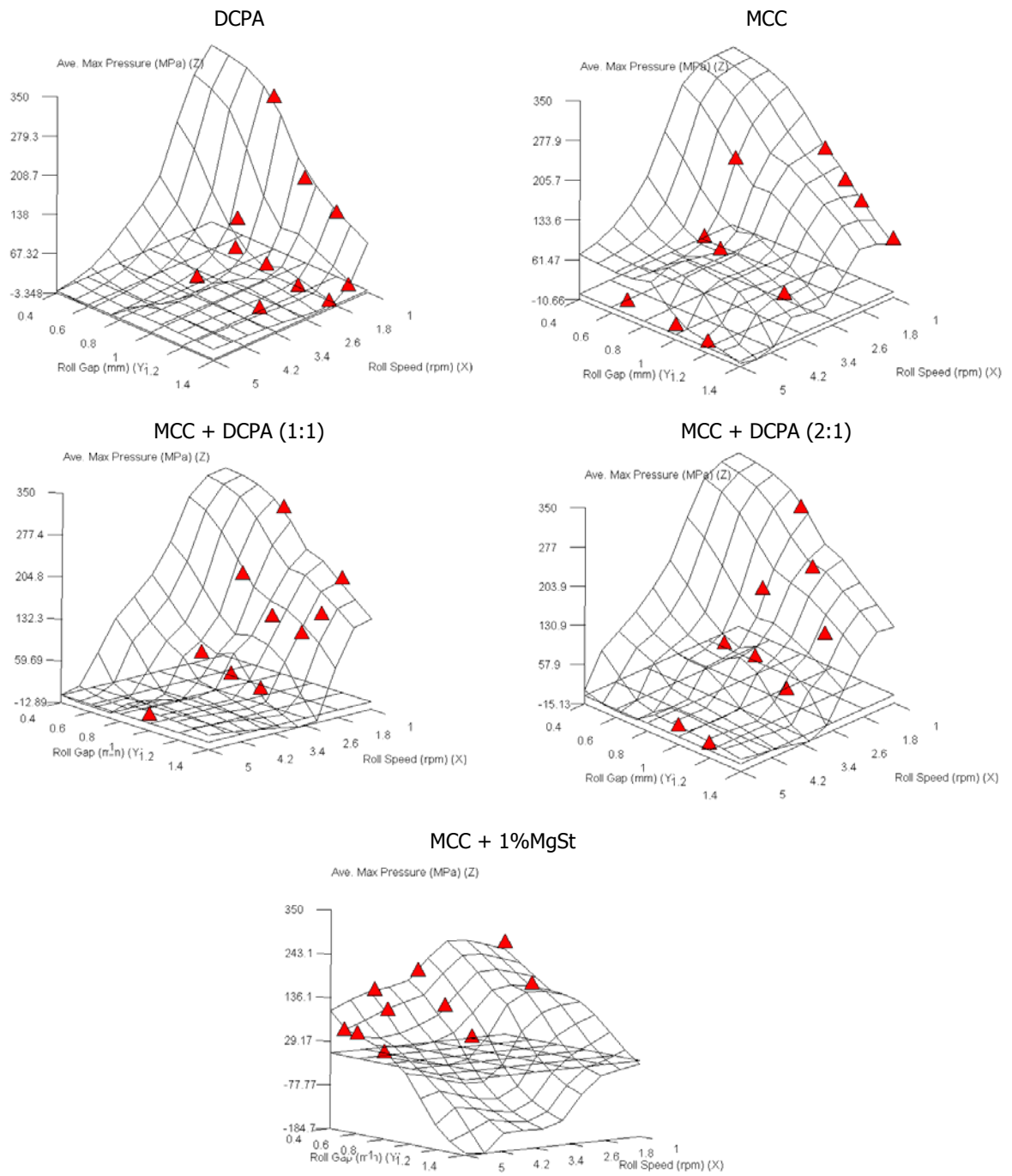
Figure 7.9 The response surfaces for average maximum pressure (MPa). The red triangles are points from the train and test data set.
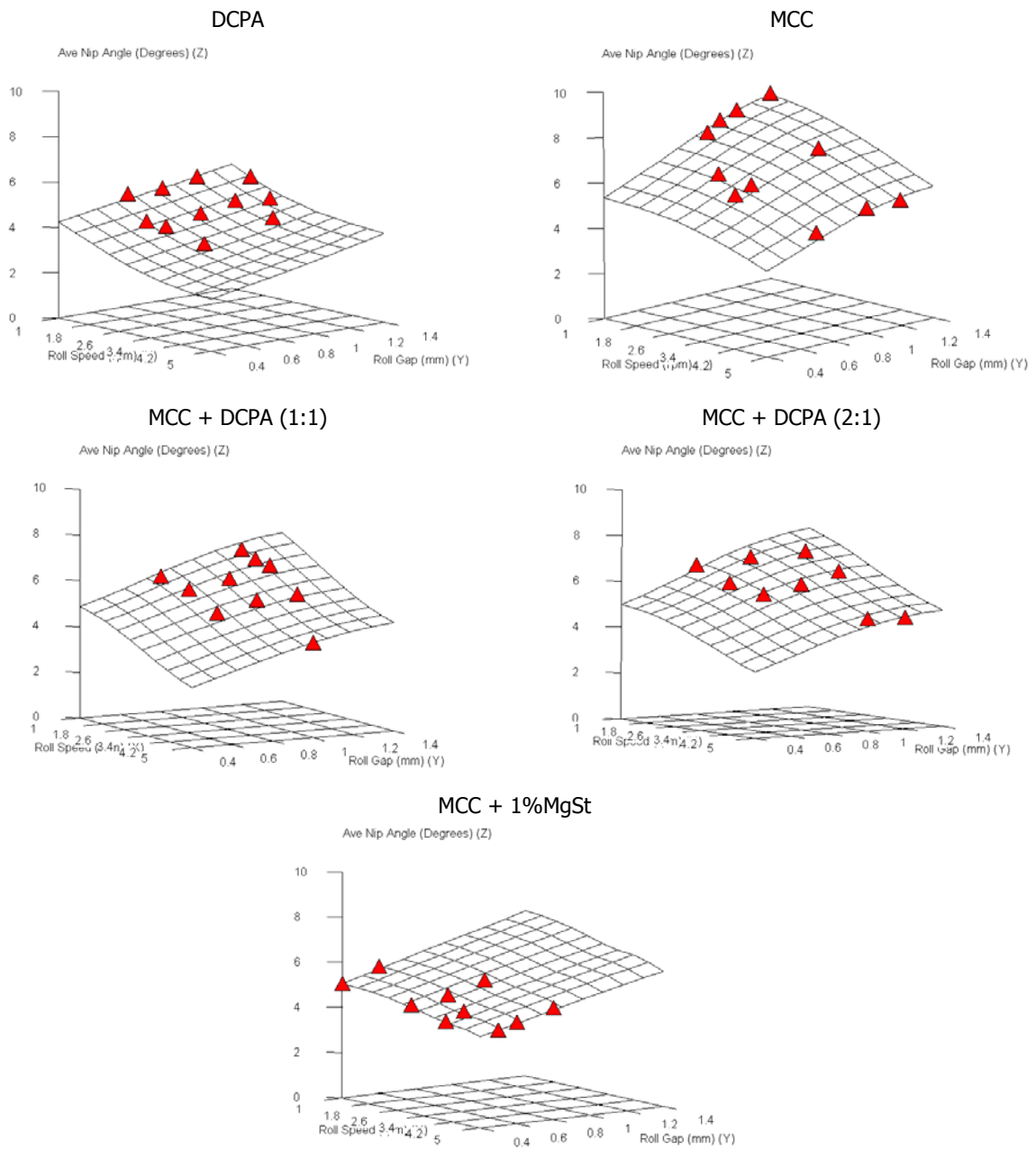
Figure 7.10 The response surfaces for average nip angle (^{o}). The red triangles are points from the train and test data set.

Table 7.9 "What if predictions" examples

Inputs	Example 1			Example 2			Example 3		
	"What if"	Best Match Ingredient	Best Match Properties	"What if"	Best Match Ingredient	Best Match Properties	"What if"	Best Match Ingredient	Best Match Properties
Ratio of M:D	0.00	0.00	0.00	1.00	1.00	0.00	1.00	1.00	1.00
MgSt %	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00
Roll Speed (rpm)	2.50	2.00	3.00	25.00	5.00	2.00	4.00	4.00	5.00
Roll Gap (mm)	1.00	1.00	1.20	2.00	1.20	1.20	1.00	0.40	1.00

Outputs	Predicted	Best Match Ingredient	Best Match Properties	Predicted	Best Match Ingredient	Best Match Properties	Predicted	Best Match Ingredient	Best Match Properties
Ribbon D (kg/m3)	1548.42	1670.66	1509.32	1581.41	679.67	1598.24	783.02	1224.79	708.12
Ribbon Porosity	0.46	0.41	0.47	0.43	0.55	0.44	0.50	0.18	0.53
Ave. Max Pressure (MPa)	19.98	28.98	7.07	22.61	12.33	15.45	-56.05	144.80	18.10
Ave Nip Angle (Degrees)	3.93	4.18	4.12	5.35	5.52	4.50	5.62	4.35	5.53

7.2.4 Optimisation

The optimisation process predicted results which were compared with experimental results and are shown in Table 7.10. It is shown that the optimisation for trials D, I and J were all successful in predicting the experimental parameters by setting ribbon porosity output. It should be noted that the optimisation procedure would not reproduce the same roll compaction results if the same optimisation was repeated.

Table 7.10 Summary of optimisation results compared to experimental results.

Trials	Data fed into optimisation			Data given by optimisation		Results obtained from experiments		
	Ratio M:D	MgSt %	Ribbon Porosity	Roll Speed (rpm)	Roll Gap (mm)	Roll Speed (rpm)	Roll Gap (mm)	Ribbon Porosity
Trial D	0	0	0.35	1.94	0.61	1.94	0.60	0.34±0.01
Trial D	1	0	0.25	1.00	1.27	1.00	1.25	0.24±0.01
Trial I	0	0	0.35	1.06	1.23	1.06	1.25	0.39±0.01
Trial I	1	0	0.25	1.20	1.21	1.20	1.20	0.26±0.02
Trial J	0	0	0.35	1.73	0.72	1.73	0.70	0.37±0.02
Trial J	1	0	0.25	2.07	0.40	2.07	1.00	0.24±0.04

7.2.5 Comparison between Johanson's Theory and "What if" Predictions in predicting the Nip Angle

Figure 7.11 shows Johanson's theory predictions and also *INForm* software "what if" predictions. There are three models developed in *INForm*. They are Trial D, Trial I and Trial J. These trials produced predictions which were in better agreement with the experimental results than Johanson's theory prediction. The *INForm* models

were able to predict the experimental nip angle to within an accuracy of 5%. Basically any of the three trials could be used to predict the Nip Angle. Whereas the Johanson's model was not successful in predicting nip angle above the roll speeds of 1 rpm due to air entrainment. It also over-predicted the experimental nip angle of DCPA and MCC by 200%, while the approximation using Johanson's pressure profile under-predicted the experimental nip angle of DCPA by 5-20% and MCC by 20%. Hence this shows that the *INForm* software was very flexible and could produce models from different combinations of inputs which can be used to make good predictions.

In addition to that, the *INForm* software was also easier to use compared to Johanson's theory. The types of input variables can be varied whereas for Johanson's theory the independent variables are set. The *INForm* software models can also be set to predict different types of outputs in one calculation, whereas Johanson's theory can only predict the Nip Angle in one calculation. Further calculation is required to obtain other required outputs. Moreover, Johanson's theory does not account for the effects of varying roll speed due to deaeration in the nip region and *INForm* software could use any available data to give an accurate prediction. Nonetheless, the setting up of the *INForm* software model did take time in terms of collecting experimental data and setting it up in the software but the usage of the model to predict required outputs was very quick.

As long as the model has been tested to predict for data which it has not seen before (validation data) then it could be used to safely predict output properties within the region of space for which experimental data has been conducted. It may be concluded that the model is robust within the experimental space but not out of it whereas Johanson's model is robust for any type of sample material and roll compaction process parameters. Furthermore the software is a very useful tool to organise a large amount of data on different types of pharmaceutical excipients and using those data to contribute in predicting the required output properties of the products.

However the model developed from this software does not have a mathematical form unlike Johanson's model. The *INForm* model is also not based on physics of materials but cause-and-effect within the dataset. In addition to that the effects of each input variable were not clear in *INForm* software models unlike Johanson's models.

7.3 Summary

This work concludes that the *INForm* software could be used to develop models to predict the effects of tablet formulation and roll compaction process parameters on final ribbon quality. This investigation implies that the software could be used to develop models with the minimum input variables, such as formulation contents ratio and percentage of lubricant. It can also develop models to include input

variables which are characteristics of the tablet formulation such as Compressibility κ or PSD Ratio $(d_{84}-d_{16})/d_{50}$ such as shown in Trial I and J. It has been proven in this work that the results from the 'what if predictions' and the optimisation feature could be trusted within the space of experimental data. The results from the "what if" predictions were also compared to Johanson's theory and proved to be give a more agreeable prediction.
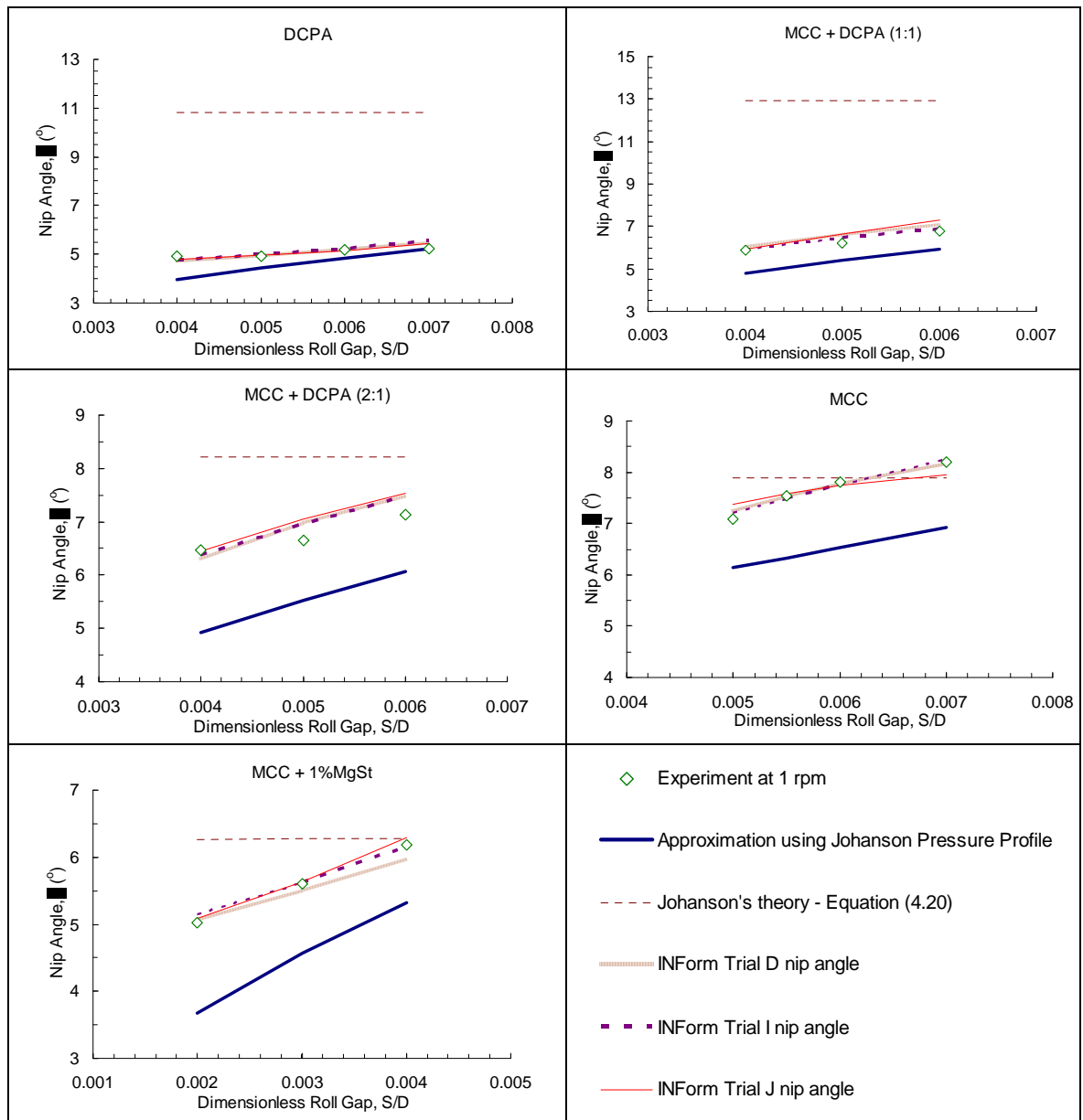
Figure 7.11 Comparison of the Johanson's theory prediction and *INForm* software prediction of nip angle for powder material compacted at 1 rpm.

It can be suggested that this software be used as a form of database to collect large amounts of information for the Pharmaceutical Industry, which will then benefit the industry as a predictive tool to obtain answers to help solve problems. It was discovered that as long as the software was fed a sufficient amount of datasets, it was very useful to predict a large amount of output properties from a limited amount of material characteristics.

However this software should not be used to predict for new untried formulations or roll compaction process parameters. This is because this software works on a basis of cause and effect and can only predict for experimental ranges which it has been trained on. It does not do well when used for extrapolating results. In such cases as predicting new untried formulations, the traditional engineering mathematical models should be used. This is because for new formulations, the basic physical characteristics will be accounted for within the traditional mathematical models and not accounted for in the *INForm* models.

Trials on the *INForm* software to investigate the benefit of using key variables found from FormRules were unfruitful. It was because it made the model training more complicated as separate models for each output needed to be developed. In addition there was no easy way to assess the quality of each trained model as soon

as one is trained. An assessment on the training and test MS error was insufficient to confirm that the model was of a good quality.

Furthermore, one can say that the *INForm* software models are "just a fancy form of regression" and could one not get results which are just as good by regression alone? It can also be hypothesized that in a linear relationship the 'what if prediction' is comparable to the common regression method. For instance the roll compaction output property could be obtain by working forward from the specific formulation and roll compaction process parameter to intra-polate on the output property using the regression relationship.

Following that the optimisation results could also be reproduced using the regression method by working back using the output property to obtain the roll compaction process parameter for a specific formulation. However, this has not been investigated within this research and it is unknown what advantages the *INForm* software has over the regression method in a non-linear input output relationship. Nevertheless it is interesting to note that Inghelbrecth *et al.* (1997) reported that the Multilayer Feed-Forward Network neural network (MLF) predicted the granule friability more accurately than a second order polynomial. Moreover the comparisons can be made between the concepts within ANN and regression analysis (Table 7.11). The ANNs are essentially equations interrelated to each other nodes within a network and is easy to model non linear relationships, while regression

analysis is limited by polynomial equations and more challenging to model non linear relationships.

Table 7.11 Summarised comparison between neural networks and regression analysis.

Neural Networks	Regression Analysis
Inputs	Independent variables
Outputs	Dependent variables
Calculates using mathematical equation	Calculates using mathematical equation
Weights	Coefficient
Improves model by evaluation of prediction error.	Improves model by evaluation of prediction error.

Investigations comparing regression to neural network, which did not involve roll compaction have been conducted in the past and showed two things (De Veaux *et al.*, 1993; Timofei *et al.*, 1997; Subramanian *et al.*, 2004). De Veaux *et al.* reported that the multivariate adaptive regression splines (MARS) were more accurate and made faster calculations than ANNs, while Timofei *et al.* reported that the ANNs performed slightly better than Multiple Linear Regression analysis. Subramanian *et al.* (2004) demonstrated that ANNs showed less error than multiple regression analysis and provided a more accurate prediction. De Veaux *et al.* (1993) found that MARS produces a model where basis functions are grouped together based on the order of interactions of the predictor variables they contain, whereas neural networks produce a black box. MARS was also found to generally give a better prediction than neural networks at a lower number of data points. As the number of data points increased, the two methods were expected to asymptotically reach the same level of accuracy.

Tomefei *et al.* (1997) found that the MLR approach lead to a better interpretation of the contribution of individual terms, but neural networks can extract more information from the data than statistical methods, especially where non-linear relationships are involved. Subramanian *et al.* (2004), reported that the normalised error from ANNs prediction was less than the multiple regression analysis, ANNs easily handled large amounts of input variables and became more helpful when the number of experiments were greater. In the multiple regression analysis, higher number of input variables led to a polynomial with more coefficients, which involves tedious computation. In view of this, the ANNs were more flexible to work with when historical data was available.

Nevertheless this software is superior to the simple regression method in terms of handling large amounts of data and when used to obtain specific values from the data in a time efficient way. This is because regression can only be done to calculate for one value, whereas the INForm software could be used to calculate several different outputs property, formulation or roll compaction process parameter. It must be noted that this modelling method does not describe how each input contributes to the output properties whereas regression equation would.

8 Final Conclusions and Future Work

8.1 Main Conclusions

Research on the roll compaction as a dry granulation method has been conducted since the 1960's. Although the effects of roll compaction of various types of powder materials on the final ribbon compact or tablet quality have been reported, they have rarely been quantified. The aim of this work was therefore to develop predictive models using intelligent software which relate the tablet formulations and roll compaction process parameters to the final ribbon quality. The intelligent software model was compared to a traditional theoretical model and from this, a list of strengths and weaknesses of the present models were determined. Below, the general conclusions associated with the roll compaction of a variety of tablet formulations are presented, along with conclusions relating to the predictive models.

8.1.1 Production of Ribbon Compacts from Tablet Formulations

Gravity feeding roll compaction experiments conducted on five types of tablet formulation revealed that different types of powder material characteristics required specific roll compaction roll gap and roll speed ranges. In comparing uniaxial compaction to roll compaction, it was found that for DCPA the densities were in agreement at lower pressure and deviated at higher pressure. For the other four tablet formulations the relationships were not in agreement over the investigated pressure ranges.

The roll compaction of a variety of tablet formulations resulted in the following conclusion:

♦ The nip angle generally increases if: i) maximum applied pressure in the nip increases, ii) the powder material effective angle of internal friction increases, iii) the powder material cohesivity increases, iv) the roll speed decreases, or v) the roll gap increases. However in the case of lubricated powder, the decrease in powder material cohesivity decreased the nip angle as well as the operational roll gaps.

♦ The upper limit of the roll speed operational range was found to decrease as the powder material d50 decreased.

♦ The maximum applied pressure in the nip decreases if the roll gap increases or if the roll speed decreases.

8.1.2 Predictive Capability of Models

FormRules software training was found to produce good neurofuzzy models which highlighted the key input variables which affected the output properties. The appropriateness of the key inputs should be carefully considered because the suggested relationships are not always practical and physically significant. The *FormRules* models were found to be flexible; as soon as the database is set up inside the software, rapid and varied trials are possible.

The highlighted key input variables from *FormRules* models were used in *INForm* model development. Three trials were chosen to be investigated using the *INForm*

software. Those three trials were all successfully trained to produce good predictive models. The *INForm* models were found to be useful in two ways:

i.   "What if" predictions could be used to find the roll compaction output property for a particular tablet formulation and roll compaction at specific process parameters. In industrial practice, this would save a lot of experimentation time, provided that the tablet formulation and the roll compaction process parameters are within the scope of the database.

ii. The "Optimisation" function could be used to predict the roll compaction process parameters for a specific tablet formulation, at a known output property. The predicted roll compaction process parameters were in agreement with experimental results.

The *INForm* models were used to predict nip angle values and were compared to experimental results and Johanson's model prediction. This showed that the *INForm* model predictions were in better agreement with the experimental results compared to Johanson's model. Table 8.1 shows a comparison of the two models. This research does not claim that an *INForm* model is superior to Johanson's model, but it does show the strengths and weaknesses of using the *INForm* model for prediction.

The ability of *INForm* software to handle large amounts of data to detect "cause and effect" relationships can be of great use in the pharmaceutical industry. This would enable the industry to save time by decreasing the number of experiments for new

tablet formulations plus the software provides a database for an organised reference, and if the model was trained with good datasets, it could ease the validation process essential for good manufacturing practice.

Table 8.1 Summary of the comparison of *INForm* model and Johanson's Model.

INForm Model	Johanson's Model
Does not have any obvious mathematical equations to relate inputs to outputs. Usually known as a "Black box" model.	Mathematical Model based on established powder mechanics theories.
Does account for the increasing Roll Speed, as the data are considered in detecting the "cause and effect" relationships between input and output.	Does not account for effects of increasing roll speed due to the air entrainment within the nip region.
Could not predict outside the scope of data that it has been fed. i.e. fails when used to extrapolate outside the range of tablet formulation or roll compaction process parameters which it was trained on.	Could predict nip angles for any material provided that the model is fed sufficient material characteristics. (i.e. Compressibility κ, effective angle of friction and angle of Wall Friction.)
Once the model is trained, the model could be used to process large amounts of information of various tablet formulations and process parameters to predict required output property quickly.	Could only be conducted for one tablet formulation and process parameter to obtain one output property at any one time.
Produces very accurate nip angle predictions.	Predictions of nip angles showed very small deviation from the experimental nip angle, which is within 1° to 2°.
Flexible – Predictions improve with volume and type of data	Rigid – Fixed performance

8.2 Future Work

Based on the experimental research presented in this thesis, the subjects listed below are suggested avenues for future work:

i.   To produce tablets from the roll compacted ribbons and to study the disintegration and dissolution times of the tablets. These values can then be used in comparison with tablets produced via wet granulation.

ii. To find a better representation of PSD or to prove that PSD Ratio (d_{84}-d_{16})/d_{50} is a good indication of PSD width and median. PSD has an obvious effect on the roll compaction of powder materials. It would benefit the *INForm* models if PSD was indicated by a less ambiguous representation.

iii. To add into the present intelligent software data base. By varying the type of tablet formulation, which highlights each the bulk powder characteristics to a different extent.

iv. To compare the *INForm* models prediction with regression analysis prediction and Johanson's model modified by Boursel (2001) to accommodate air entrainment.

9 Appendices

9.1 Appendix 1

Comparison of roll compaction to uniaxial compaction.

Converting the angular speed of the roller compactor to horizontal speed.
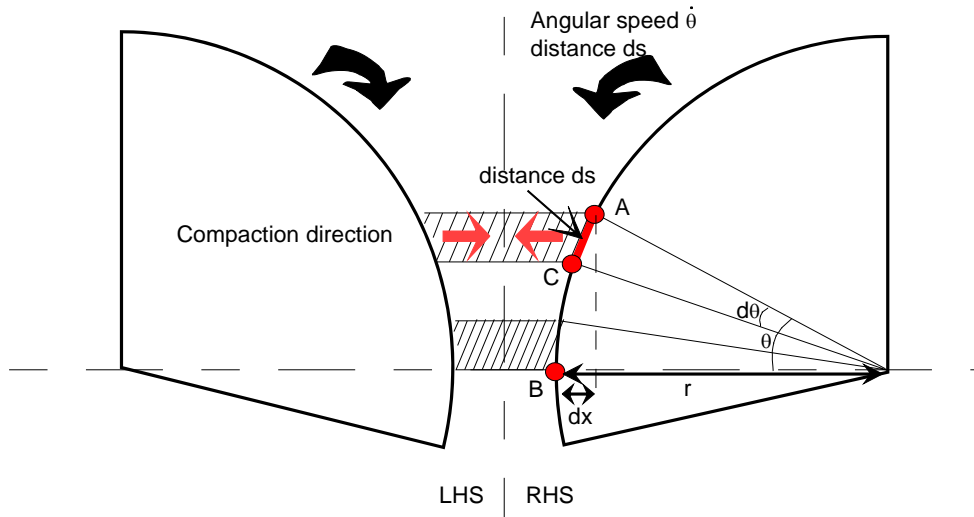
Since LHS = RHS

Consider RHS

Where length of arc, $AC = ds = rd\theta$

Hence $dx = ds\sin\theta$

$$dx = rd\theta\sin\theta$$

$$\frac{dx}{dt} = r\frac{d\theta}{dt}\sin\theta$$

Because there are two compressions occurring simultaneously, thus the compaction speed is:

$$2\frac{dx}{dt} = 2r\frac{d\theta}{dt}\sin\theta$$

Radius, r =   100.00   mm

Table 2 The conversion of roll speed to horizontal compaction speed.

Nip Angles (°)	roll speed (rpm)	roll speed (rad/s)	Horizontal compaction speed (mm/s)	Horizontal compaction speed (mm/min)
3	**1**	**0.1047**	**1.10**	**65.75**
4	1	0.1047	1.46	87.64
5	1	0.1047	1.83	109.50
6	1	0.1047	2.19	131.33
7	1	0.1047	2.55	153.12

9.2 Appendix 2

Compliance results and an example of punch displacement according to the equipment compliance result
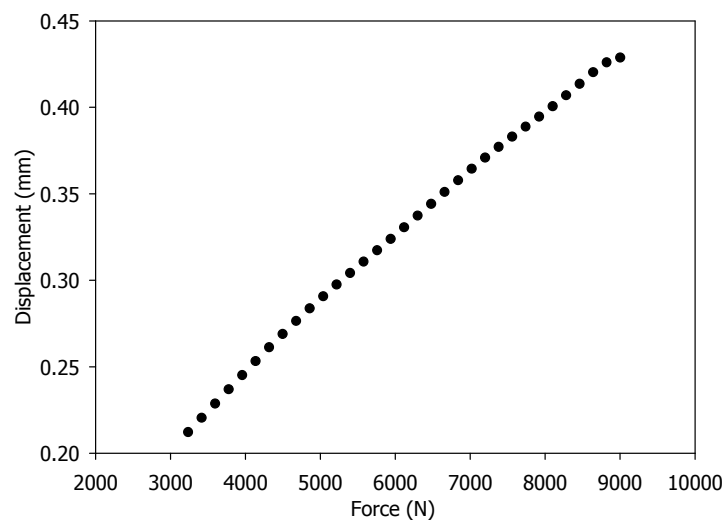
Figure 9.1 Compliance test result for the 30 kN load cell at 1mm/s vertical speed.

The graph above shows the compliance. The graph produces a relationship represented by the equation below:

$$y_c = 2.28 \times 10^{-4} x - 3.38 \times 10^{-7} x^2 + 3.46 \times 10^{-10} x^3 - 2.02 \times 10^{-13} x^4 + 7.24 \times 10^{-17} x^5$$
$$- 1.64 \times 10^{-20} x^6 + 2.35 \times 10^{-24} x^7 - 2.08 \times 10^{-28} x^8 + 1.03 \times 10^{-32} x^9 - 2.20 \times 10^{-37} x^{10}$$

where y_c is the displacement of the machine, and x is the load. The equation above is used to correct the original displacement of the uniaxial compaction results using the equation below.

Compliance corrected displacement = Original displacement − y_c

9.3 Appendix 3

Ribbon density calculations

To calculate the ribbon density; firstly the total volume of bottle (cm^3) was obtained by dividing the initial mass of water (g) by the density ρ of water at temperature (^{o}C). The initial mass of water (g) was obtained by subtracting mass of empty bottle (including cover) from mass of bottle and cover filled with water. Then the final mass of water (after adding the sample) was calculated by subtracting the mass of wax coated ribbon compact and the mass of empty bottle from the final bottle mass. Next the volume of water was obtained by taking a ratio of the final mass of water over the density of water. Hence the volume of sample is the volume of water subtracted from total volume of bottle. Then the mass of wax was obtained by subtracting mass of ribbon from mass of wax coated ribbon. Since the volume of wax coated sample is equal to the sum of volume of ribbon sample and Volume of wax. Replacing the volumes for the ratio of mass over densities we obtained, volume of wax coated ribbon sample = mass of ribbon sample / density of

ribbon sample + mass of wax / density of wax. Rearranging this equation would give us the ribbon density = mass of ribbon sample (1/ (volume of wax coated ribbon sample − mass of wax / density of wax)). Finally ribbon porosity was calculated from, ε = 1- (ribbon density / true density).

9.4 Appendix 4

FormRules software tutorial

9.4.1 Entering Data and Setting Inputs and Outputs

Initially the 64 data sets were filled into excel and saved as a text file (FR803a.txt). The first row was filled with the input/output name and the first column was filled with the formulation name. Next the text file was imported into the FormRules by clicking on "Data Sheet" button and clicking on "Import" (Figure 9.2).
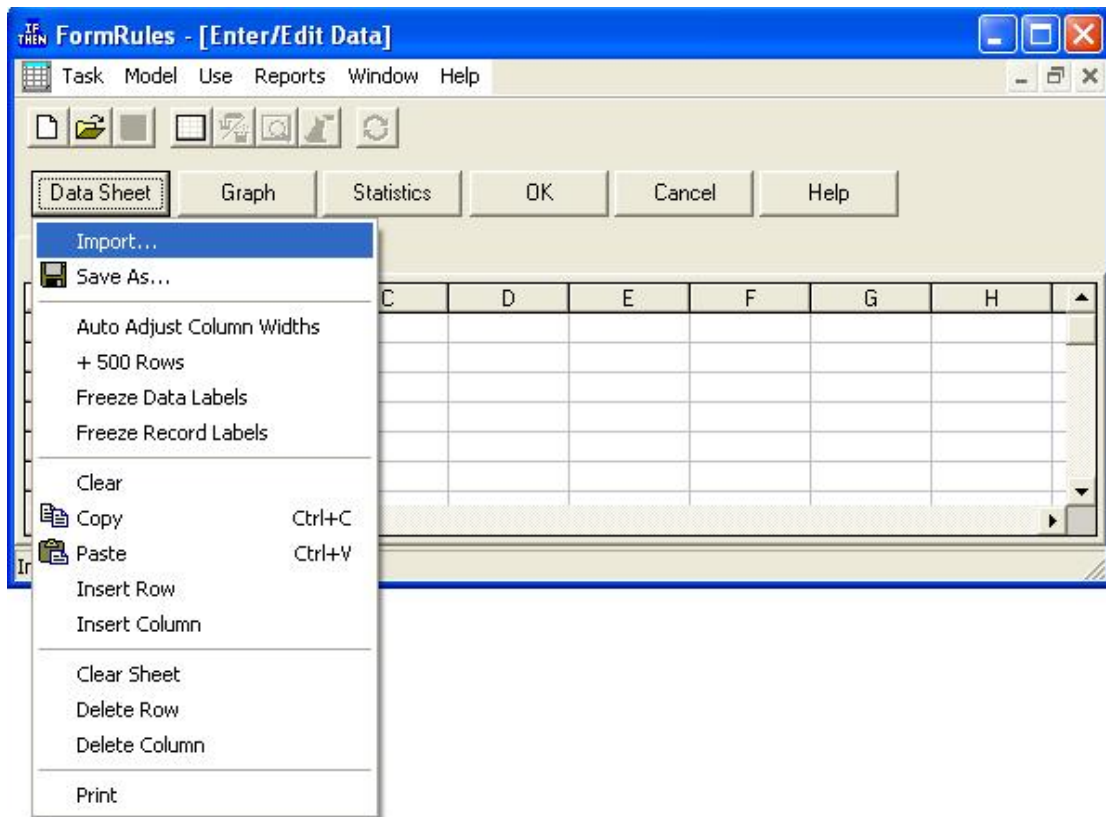
Figure 9.2 Importing data into *FormRules*

Figure 9.3 shows the imported data sets within *FormRules*.
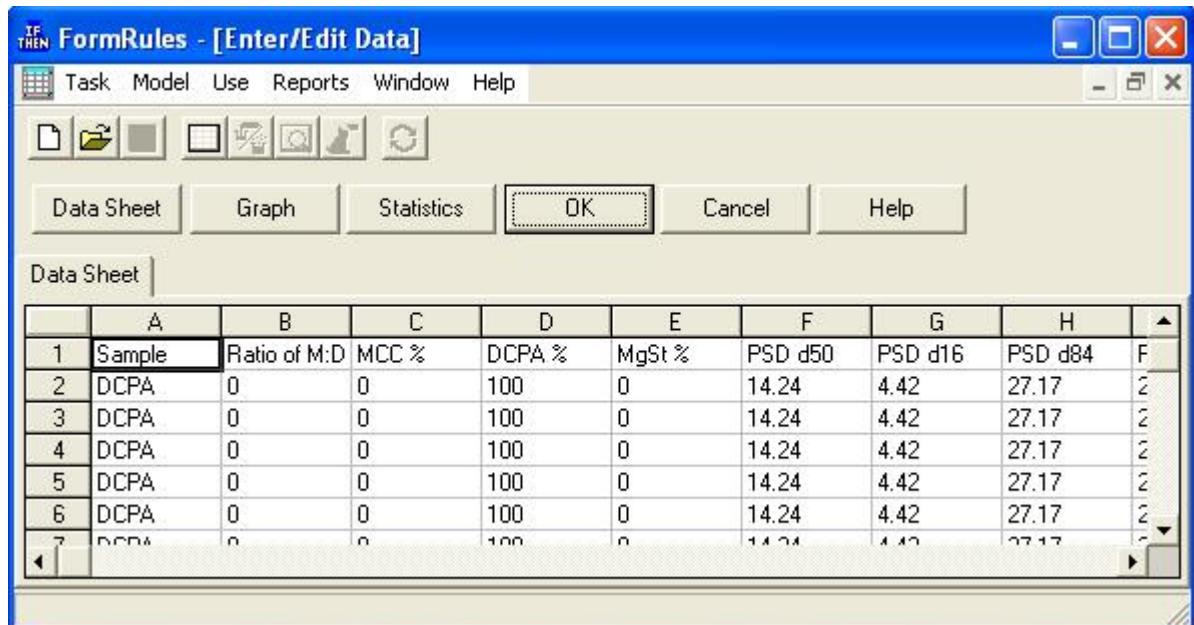
Figure 9.3 Data set has been imported into *FormRules*

Then "Ok" was clicked on to bring the user into setting field types for the inputs and outputs (Figure 9.4). In this window the user was required to set the inputs and outputs for the model training. The inputs are known as ingredients and processing condition, while the output was known as property within this program. Hence the formulation and formulation characteristics were set as "ingredient", the process parameters were set as "processing condition" and the roll compaction output and intermediate output property was set as "property". Note that inputs which were not used were set as "not used".
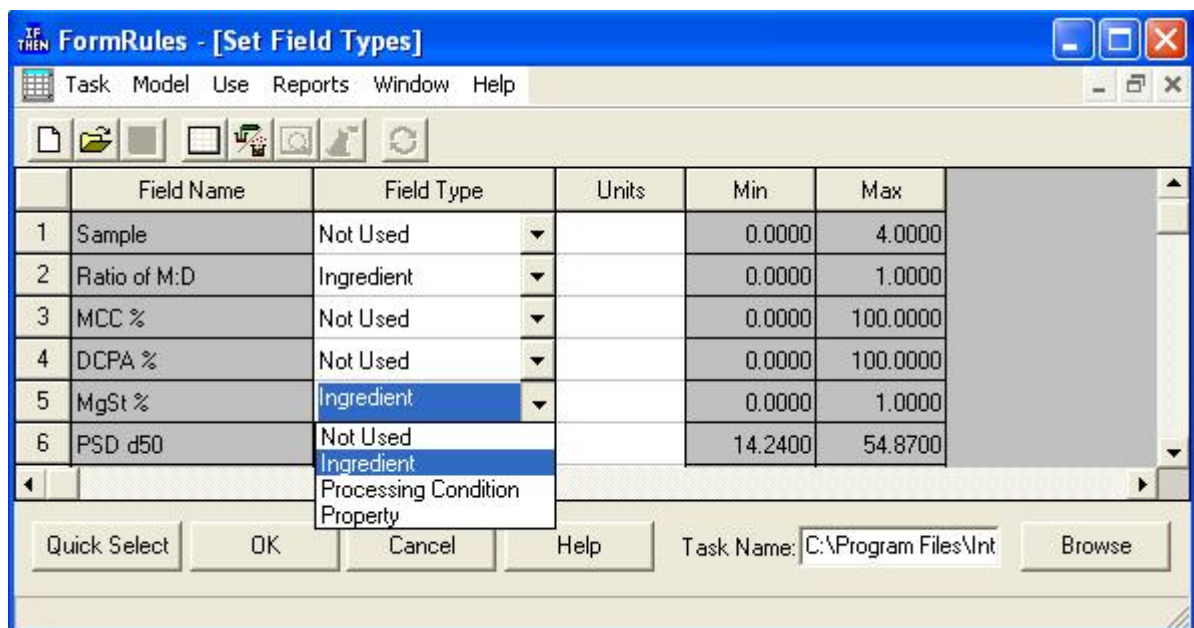
Figure 9.4 Setting field types

After the field types were set the data analysis window will appear (Figure 9.5). Although the possibility of analyzing the data has been provided, it is not essential to conduct the analysis unless the user has doubts on the data coverage, or needs to check on the simple statistics or to look for outliers.
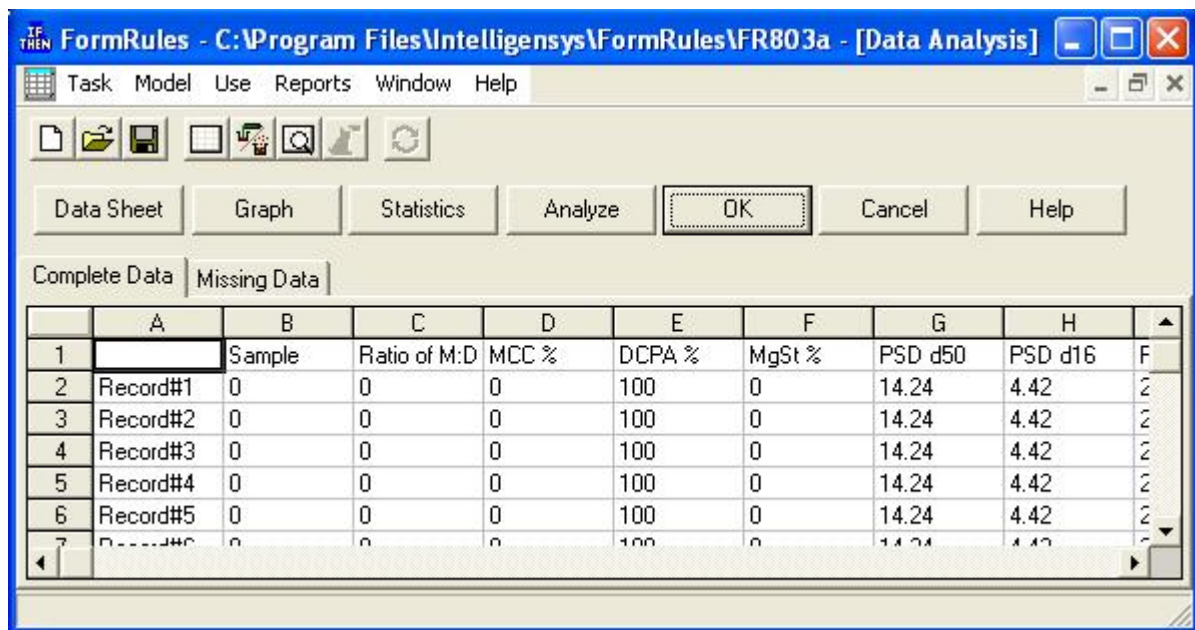
Figure 9.5 Data analysis

On the data analysis window, the "Ok" button was clicked and this brings the user to the training window which is explained in the next section (Figure 9.6).

9.4.2 Model Training

In this training window the "parameter" button can be clicked on to reveal the model training parameters. There are three separate tabs shown by Figure 9.7, Figure 9.8 and Figure 9.9. Initial training would use the default settings but subsequent training would require the training parameters to be readjusted. Each of the training parameters is explained in Section 9.4.3. In the training of the *FormRules* for data set FR803a.txt the default settings for the training parameters were used. To proceed to training the "Train" button was clicked.
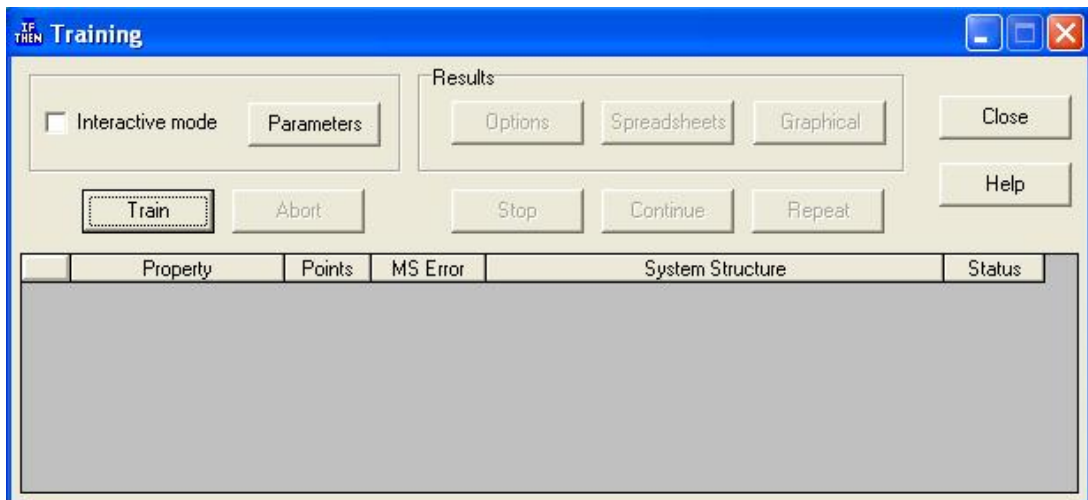
Figure 9.6 Training window
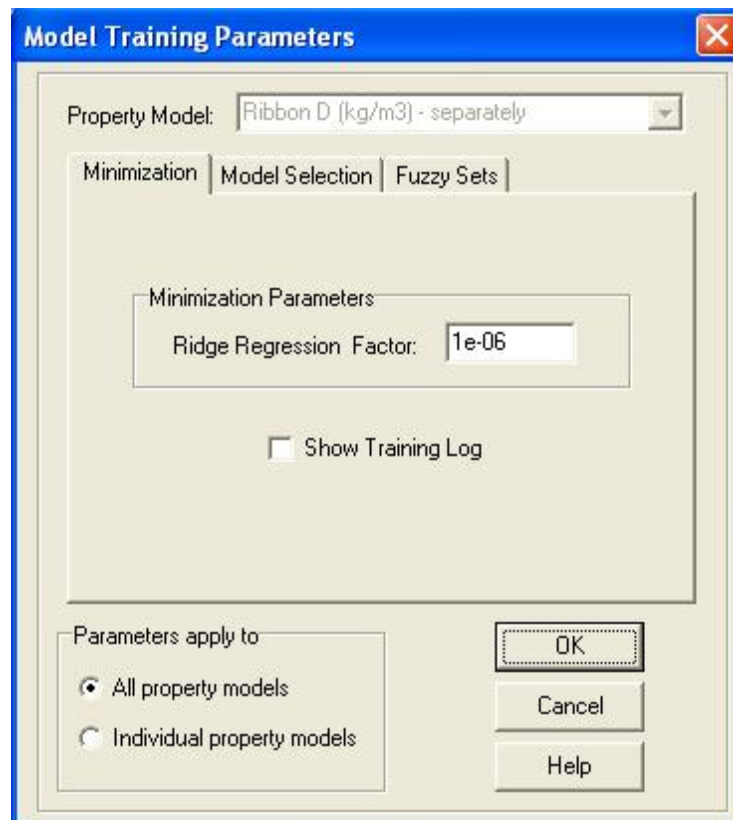
Figure 9.7 Model Training Parameter: Minimisation tab
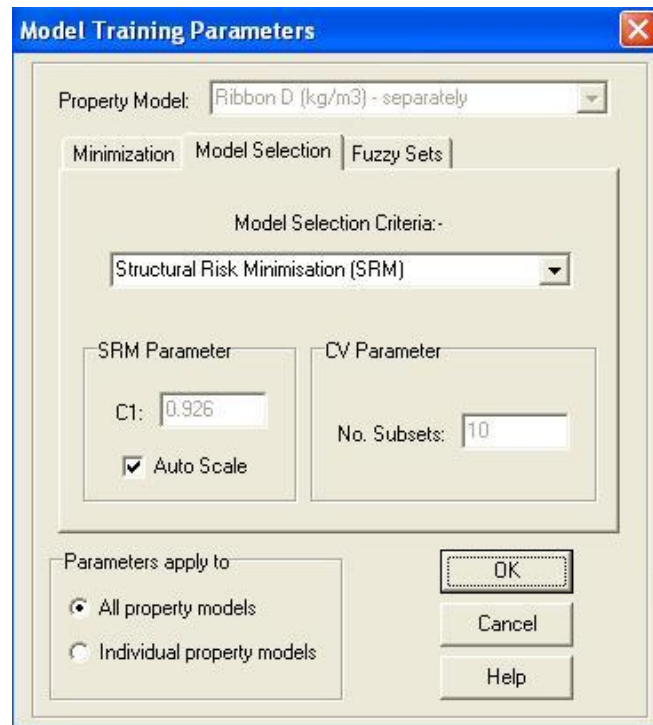
Figure 9.8 Model Training Parameter: Model Selection
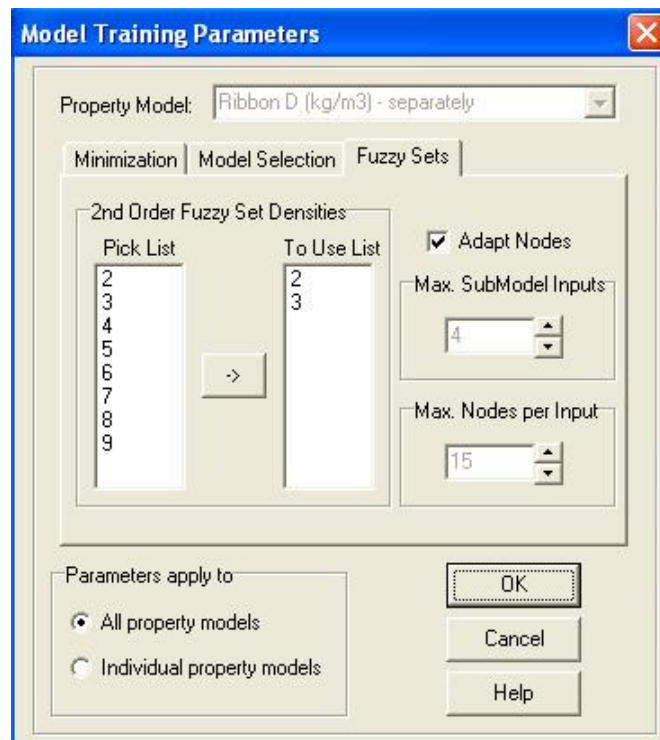
Figure 9.9 Model Training Parameter: Fuzzy sets

9.4.2.1 Training complete

After training was completed a small window indicating the completion of training will appear as shown in Figure 9.10. Click "Ok" on the "Training is complete!" window. Note that in the FormRules model training each output property has been trained independently from each other. Note that in the FormRules model training each output property has been trained independently from each other.
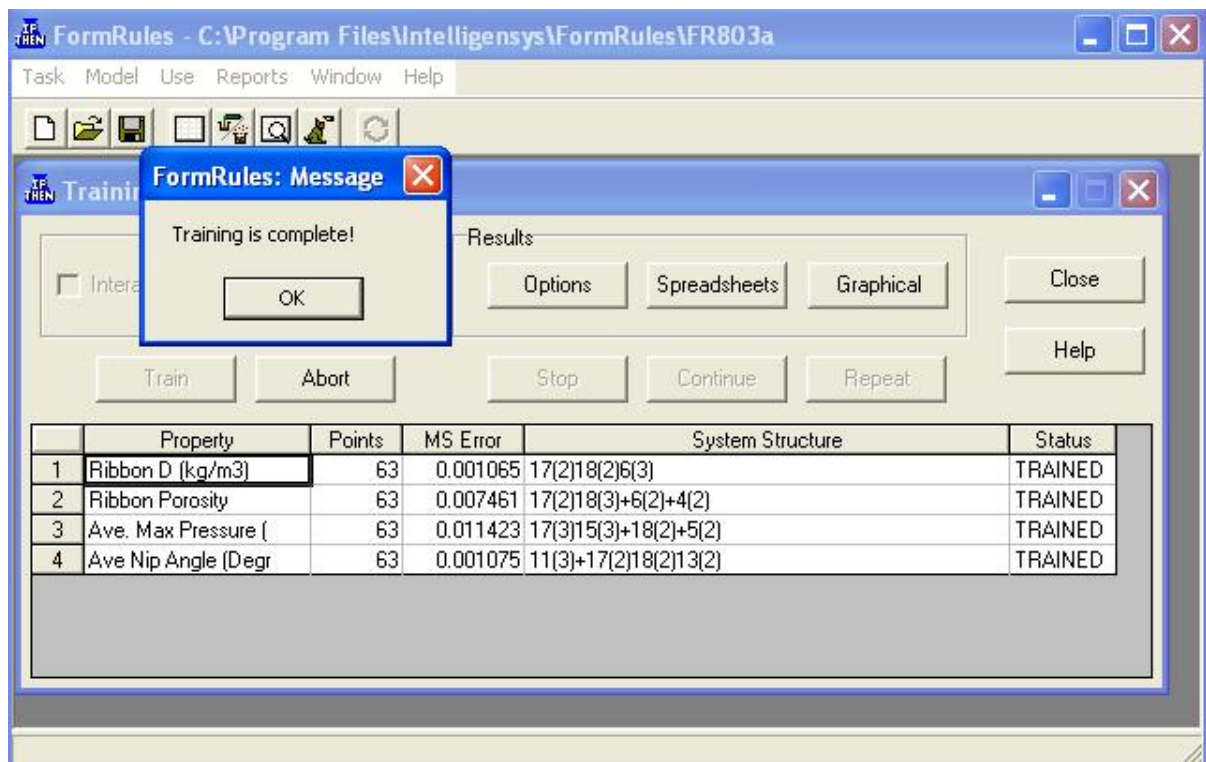
Figure 9.10 Model training completed

9.4.2.2 Assessing the trained model

The trained model was assessed by evaluating the f-ratio and R^2 value within the model statistics tab which could be opened by clicking on the "Spreadsheet" button.

A good model would show computed f-ratio above 4 and Train Set R^2 values of above 80. As shown in Figure 9.11 the model was successfully trained.

	A	B	C	D	E
1	Train Set				
2	Ribbon D (kg/m3)				
3	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
4	Model	5.5472e+06	12	462267	178.568
5	Error	129437	50	2588.74	
6	Total	5.67668e+06	62		
7					
8	Train Set R-squared	97.7198			
9					
10	Ribbon Porosity				
11	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
12	Model	0.682515	8	0.0853144	51.8469
13	Error	0.0888573	54	0.00164551	
14	Total	0.771374	62		
15					
16	Train Set R-squared	88.4807			
17					
18	Ave. Max Pressure (				
19	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
20	Model	325791	11	29617.4	26.5026
21	Error	56993.9	51	1117.53	
22	Total	382787	62		
23					
24	Train Set R-squared	85.1108			
25					
26	Ave Nip Angle (Degr				
27	Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	Computed f ratio
28	Model	73.2162	10	7.32162	253.171
29	Error	1.50382	52	0.0289196	
30	Total	74.7203	62		
31					
32	Train Set R-squared	97.9874			
33					

Figure 9.11 Model statistics on which the models can be assessed for acceptability

9.4.2.3   Examining the trained models

Clicking on the "Graphical" button gives the neurofuzzy results window as shown in Figure 9.13 which is a graphical display of the submodels for each property, and the highlighted inputs to each submodel. The figure shows that the inputs Ratio (d84-d16)/d50, roll speed (rpm) and roll gap (mm) are key variables in relation to ribbon density (kg/m^3). There are two ways in which the fuzzy rules can be obtained. Firstly, by clicking on the "Spreadsheet" button and choosing the "Rules" tab (Figure 9.12).

Secondly by left-clicking on the "Submodel: 1" box in Figure 9.13 and choosing "Show Rules". The fuzzy rules for ribbon density submodel: 1 are listed below:

IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is HIGH (1.00)

IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.58)

IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is HIGH (0.69)

IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is HIGH (0.75)

IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.87)

IF Roll Speed (rpm) is LOW AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is HIGH (0.54)

IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is LOW (1.00)

IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (0.63)

IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is LOW AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is LOW (0.63)

IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is LOW THEN Ribbon D (kg/m3) is LOW (0.64)

IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is MID THEN Ribbon D (kg/m3) is LOW (1.00)

IF Roll Speed (rpm) is HIGH AND Roll Gap (mm) is HIGH AND Ratio (d84-d16)/d50 is HIGH THEN Ribbon D (kg/m3) is LOW (0.72)

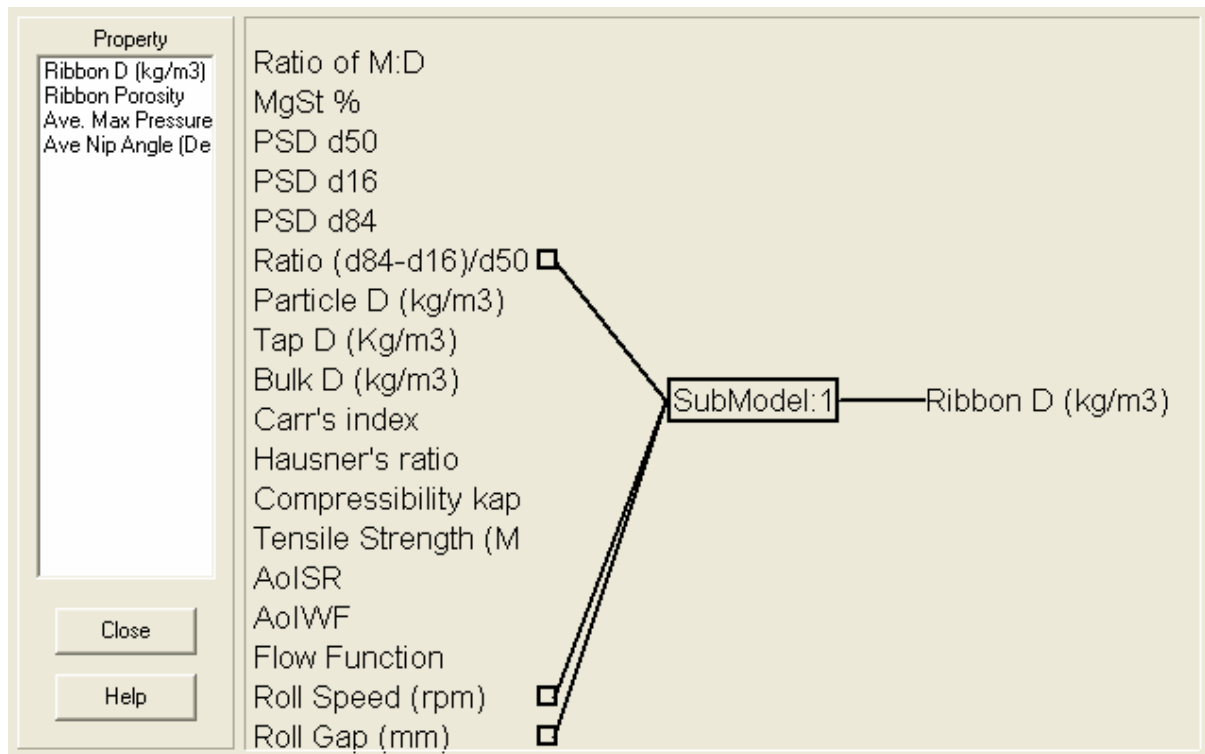Figure 9.12 *FormRules* model training results

Figure 9.13 Neurofuzzy Results for ribbon density (kg/m^3)

Submodel is the name of the set of rules relating a set of inputs to one output. It can be seen that for some outputs, the software generates up to three submodels. By highlighting the output property name in the left hand column, the neurofuzzy result for that output would be displayed. As shown in Figure 9.14, Figure 9.15 and Figure 9.16.
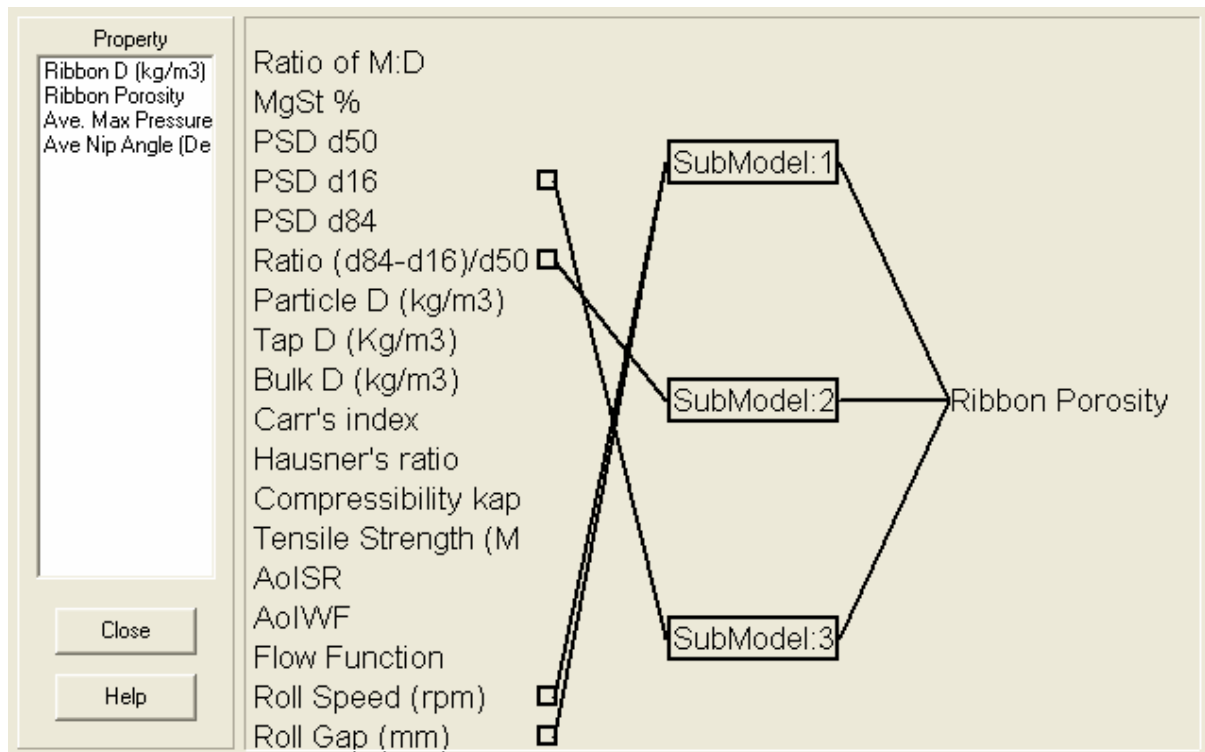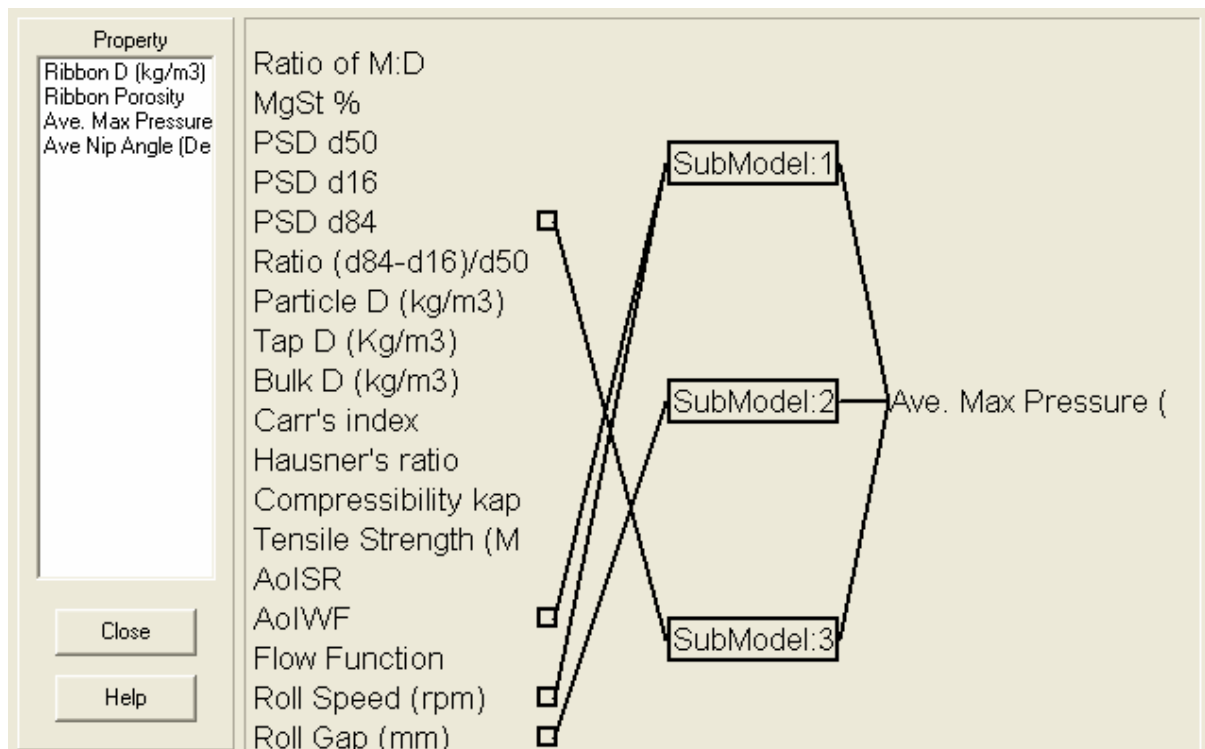
Figure 9.14 Neurofuzzy results for ribbon porosity

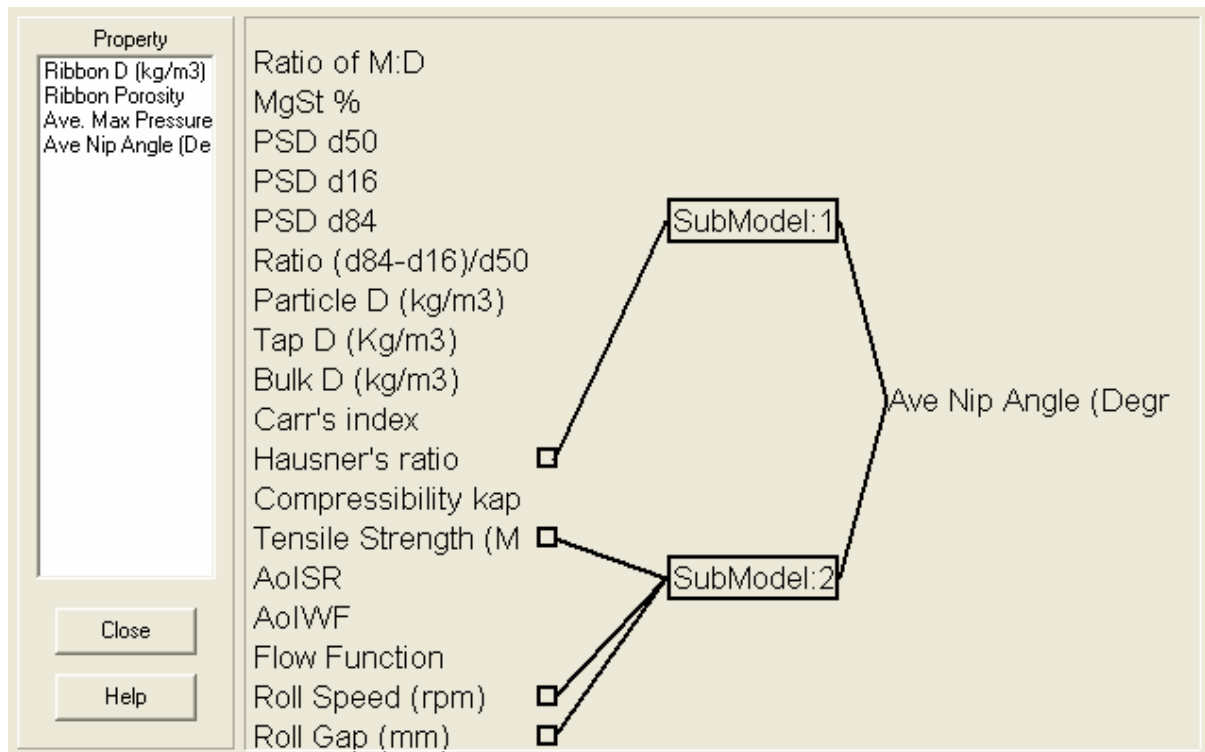Figure 9.15 Neurofuzzy results for average maximum pressure (MPa)

Figure 9.16 Neurofuzzy results for average nip angle (°)

9.4.3 Model Training Parameters

9.4.3.1   Minimisation Parameter

Figure 9.7 shows the minimisation parameter which controls the gradient descent minimisation of the neurofuzzy system. The minimisation parameter used in this system is known as the ridge regression factor. Ridge regression is a method originally developed for solving badly conditioned linear regression problems (Hoerl and Kennard, 1970). It adds a term that minimises the magnitude of the regression coefficients and makes the calculation more stable. In *FormRules* model training a small ridge regression parameter is desirable as it reduces the chance of over-training.

9.4.3.2   Model Selection Criterion

The choice of model selection criterion is important because it determines the most appropriate model and the quality of the trained model. The type of model selection criterion is affected by the size of the data set, the amount of noise in the data and the complexity of models developed. Figure 9.8 shows the default model selection criterion which is Structural Risk Minimisation (SRM; see below). There are two types of model selection criteria: the first is the validation and the second is the statistical significance method. Examples of the former are Cross Validation (CV) and Leave-one-out Cross Validation (LOOCV). Examples of the later are SRM, Minimum Descriptor Length (MDL), Bayesian Information Criteria (BIC) (see below for further explanation.

The effect of using different models has not been tested in this investigation. *FormRules* uses SRM as a default and this generally produces the best balance between simplicity of models and predictability of the results.

Structural Risk Minimization (SRM)

SRM is a complex machine learning and statistical inference. The neural network with fixed architecture and α corresponding to the weights and biases is considered as a learning machine (Burges, 1998). The principle of SRM (Vapnik, 1979) was summarised in Burges (1998). The SRM was introduced to find the subset of

functions which minimizes the bound on an actual risk (or problem). It is a structure which arranges chosen set of functions so that the Vapnik Chervonenkis (VC) dimension *h* varies smoothly. The structure divides the entire class of functions into nested subsets (Figure 9.17).
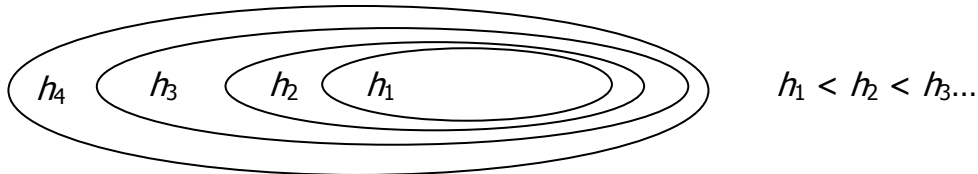

Figure 9.17 Nested subsets of functions, ordered by VC dimension.

For a trained machine the expectation of the test error which is also known as the actual risk is therefore:

$$R(\alpha) = \int \frac{1}{2} \mid y - f(x, \alpha) \mid dP(x, y)$$

(9.1)

P(x,y) is assumed to be some unknown probability distribution and neural network is the machine that needs to learn the mapping of $x_i \mapsto y_i$. The "empirical risk" R_{emp} (α) is defined to be the mean error rate on the training set and depicted as shown below:

$$R_{emp}(\alpha) = \frac{1}{2l} \sum_{i=1}^{l} \mid y_i - f(x_i, \alpha) \mid$$

(9.2)

The quantity $\frac{1}{2} \mid y_i - f(x_i, \alpha) \mid$ is called the loss and in the case it can only take the values 0 and 1. With probability 1-η, the following bound holds (Vapnik, 1995):

231

$$R(\alpha) \le R_{emp}(\alpha) + \sqrt{\frac{h(\log(2l/h)+1) - \log(\eta/4)}{l}} \tag{9.3}$$

Where h is a non-negative integer called the Vapnik Chervonekis (VC) dimension. The right hand side of (9.3) is the "risk bound".

The SRM used in *FormRules* software was based on the risk minimization theory explained above. Assuming function:

$$SRM = MSE\left[\frac{1}{1-X}\right]_{\infty} = [z]_{\infty} \tag{9.4}$$

Here

$$X = C_1 * \sqrt{\frac{((1-N)\ln(2P) - \ln(1+N)!) + C_2}{P}} \tag{9.5}$$

where P is the number of data patterns, N is the number of independent parameters in the model structure and $1+N$ is the VC dimension h.

And $[z]_\infty$ = z if z ≥ 0

$[z]_\infty$ = ∞ if z < 0

MSE = Mean Square Error is a sum of the overall output errors. It measures how well the model fits into a particular set of training data.

$$MSE = \frac{1}{P}\sum_{p=1}^{P} E_P = \frac{1}{P}\sum_{p=1}^{P} \left(R^{(p)}tget - R^{(p)}predicted\right)^2 \qquad (9.6)$$

where $R^{(p)}tget$ is the target value of property in the pattern p, $R^{(p)}predicted$ is the predicted (current) value of property in pattern p, p is the pattern or data record number and P is the total number of patterns.

FormRules SRM model selection criteria have two parameters which are called C_1 and C_2. The C_2 parameter is fixed but C_1 can be manually modified by unchecking the auto-scale. Decreasing C1 will give more complex models. Auto-scale was checked on to allow the automatic scaling of C1 from 0.8 to 1.0. If the number of data points is above 100 then a value of 1.0 is used for C1. If there are fewer than 100 data points, then the C1 value is scaled linearly between 0.8 and 1.0 and for very small data sets the value is closer to 0.8. Decreasing C_1 will give more complex models, and even quite a small change can have a significant effect.

Minimum Descriptor Length (MDL)

Minimum Description Length (MDL) is an information theoretical model selection principle. MDL seeks to minimize the number of bits[‡‡] needed to describe the data over the available models. Rissanen (1978), formulated the MDL idea using essentially the formal equivalent of negative-log probabilities (Rissanen, 1978). He developed his ideas on MDL further in 1987, 1989 and 1996 (Rissanen, 1987;

[‡‡] Bits is short for binary digit which is each 0 or 1 in the binary system of a computer language.

Rissanen, 1989; Rissanen, 1996). A review of the fundamentals and latest development in MDL is presented by Grunwald (2005).

Grunwald stated that MDL does embodies a preference for 'simpler' models rather than a best fitting model as an inference strategy useful in a very complex environment. He concluded that MDL is a versatile method for inductive inference which can be interpreted in four different ways and still give a reasonable result. It is asymptotically consistent and achieves good rates of convergence. Hansen and Yu reported excellent MDL behaviour in a regression context (Hansen and Yu, 2000; Hansen and Yu, 2001). Allen *et al*. (2003), Kontkanen *et al*. (1999) and Modha and Masry (1998) reported excellent behaviour of predictive coding in Bayesian network model selection and regression. However if none of the distributions under consideration represents the data generating machinery very well, then MDL may sometimes not find the 'best' approximation within this area. This has been observed in practice (Clarke, 2003; Kearns *et al*., 1997; Pednault, 2003).

In *FormRules* software MDL is a function taking the form shown below:

$$MDL = \ln(MSE) + N\frac{\ln(P)}{P} \tag{9.7}$$

Bayesian Information Criteria (BIC)

Bayesian Information Criteria (BIC) is derived from an extension of Schwarz's (1978) basic idea. BIC is used in model selection to provide a measure of the weight of evidence favoring one model over another (or the Bayes factor) (Weakliem, 1999). Jeffreys firstly developed Bayesian methods for inductive inference in 1939 (Jeffreys, 1939). Then in 1992 Bayesian approach in the neural network literature was introduced by Mackay (Mackay, 1992). Kass and Wasserman (1995) report a review of the techniques of Bayesian inference.

In the *FormRules* software the BIC has a function of the form:

$$BIC = MSE\left[\frac{P + (\ln(P) - 1)N}{P - N}\right]_\infty = MSE[z]_\infty \tag{9.8}$$

where $\quad [z]_\infty = z$ if $z \geq 0$

$\qquad\qquad [z]_\infty = \infty$ if $z < 0$

Cross Validation (CV)

Cross Validation (CV) is used to select between competing model alternatives. First CV divides the dataset into a training set and a test set (Friedl and Stampfer, 2002). Then the test set is used for testing the model which was built on the training set. These two steps were repeated for a number of divisions of the data. Then the test

procedure results are combined and used for the next model selection and assessment. The idea for CV was first developed by Stone (1974).

In *FormRules* software the CV function partitions the data into a number of subsets. Only one subset is used as a test set and the remaining data are used as a training set. This is repeated for every subset and the performance of the model was determined from the average adjusted Mean Square Error (MSE(adj)) of the tests. Mean Square Error was adjusted to take into account the complexity of the network compared to the number of data patterns:

$$MSE(adj) = \frac{1}{P-N} \sum_{p=1}^{P} \left(R^{(p)} tget - R^{(p)} predicted \right)^2 \qquad (9.9)$$

where $R^{(p)}tget$ is the target value of the property in the pattern p, $R^{(p)}predicted$ is the predicted (current) value of property in pattern p, p is the pattern or data record number and P is the total number of patterns.

The problem with this method was that each model being tested must be trained the same number of times as there are subsets. Hence the number of subsets (CV parameter) should be kept below 10. Increasing it would cause the training time to be longer. The CV parameter (i.e. "No. Subsets") belongs to CV and will only need to be changed if the CV model selection criterion is chosen (Figure 9.18).
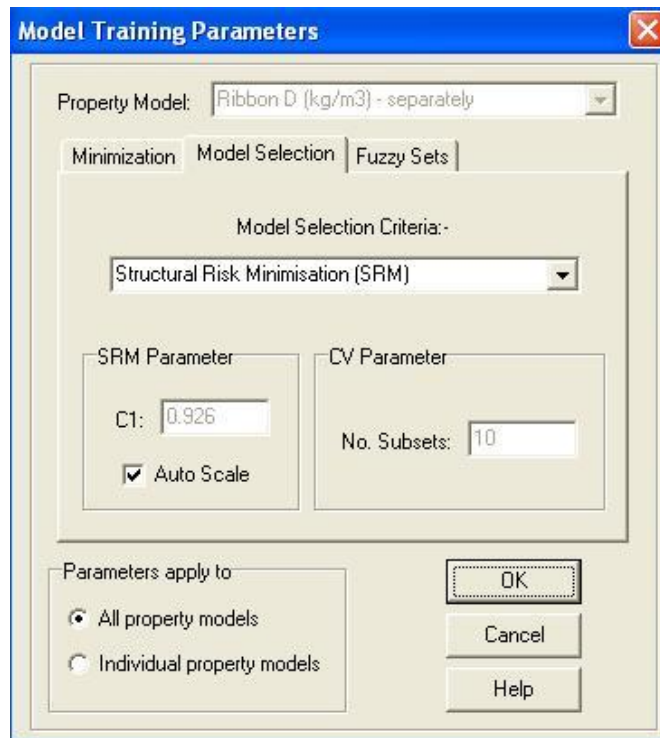
Figure 9.18 Model Training Parameter: Model Selection

Leave One Out Cross Validation (LOOCV)

Leave One Out Cross Validation (LOOCV) is a variation of CV. In *FormRules* software LOOCV, each subset corresponds to only one data record. The process of training as described above was not conducted because it would be impractical. Hence an analytical solution was used which determined the MSE if the process was conducted. LOOCV MSE calculation does not go through minimization during the model selection process but only once a solution has been found. The disadvantage of using LOOCV is that it involves a slow inversion of a matrix, it is significantly slower than the statistical methods and it has the highest tendency to overtrain.

$$LOOCV = \frac{1}{N} y \, \mathrm{P}(diagP)^{-2} \mathrm{P} \, y \tag{9.10}$$

P is the projection matrix calculated via singular value decomposition (SVD) of the auto-correlation matrix. It can be used to calculate the network weights. N is the number of data patterns and y is the vector of output values.

9.4.3.3   Fuzzy Set Parameter

Figure 9.19 shows the fuzzy set tab within the model training parameter. This tab explains the options in manipulating the structure of the neurofuzzy logic network. *FormRules* model training occurs in phases. The initial phase involves the system trying out a series of models with a fixed number of densities (i.e. nodes) per input. After the best model was found, a refining stage occurs in which the model adapts the number of nodes per input. Inputs fed in as text are encoded, the model will use 1st Order Fuzzy Set Densities and the number of sets are defined by the number of text values. Changes on this tab do not alter the results for text inputs.

Inputs fed in as numeric data will use 2nd Order Fuzzy Set Densities. These are linear or triangular in shape as explained in Section 2.2.3. The number of sets which are required to obtain a good match between actual and predicted values depends on the relationship between input and output. A linear relationship requires the minimum of 2 sets. A non-linear relationship would require 3 or more. By default sets of 2 and 3 set densities are used. Higher set densities require longer

to train and produce more complicated rules. The default settings were used in training.

The "adapt nodes" are left as checked to allow for adaptation to occur after the final model has been found to develop better models. The "maximum number inputs" per submodel can be set from 1 to 4. By default a maximum of 4 inputs are used per submodel. Reducing this may reduce the complexity of the rules, but in general will result in the match between predicted and actual results to be less good. The "maximum number of nodes per input" can be set from 10 to 15. By default a maximum of 15 nodes are used per input. Reducing this will reduce the complexity of the rules and increase the speed of model minimisation; however this will result in the match between predicted and actual results being reduced in quality.
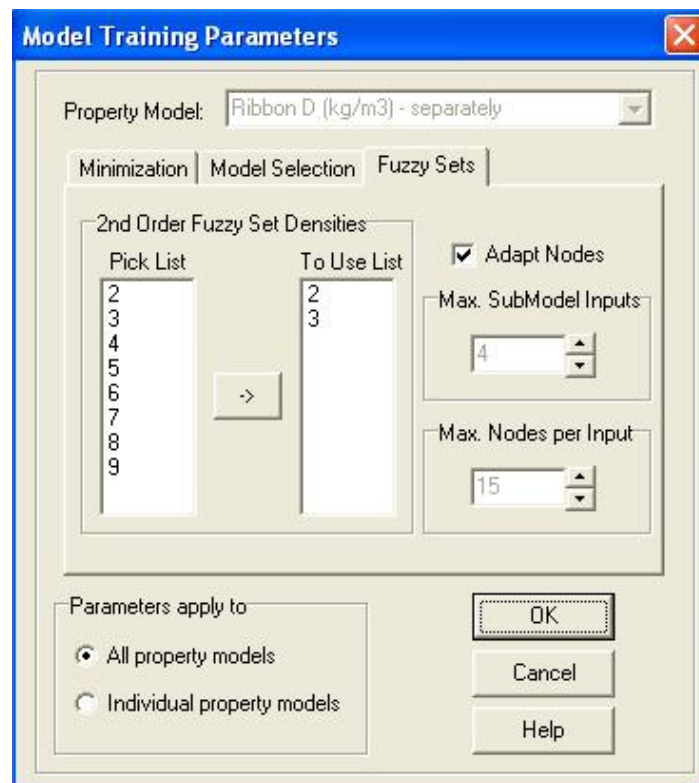
Figure 9.19 Model Training Parameter: Fuzzy sets

9.5 Appendix 5

INForm software tutorial – Example Trial D

9.5.1 Entering data and setting inputs/outputs

Initially the data was filed into The Excel Spreadsheet. The first row was filled with the input/output name and the first column was filled with the formulation name. Then 42 data sets were stored in a file named INF1208a.txt. This file was imported into the *INForm* program by clicking on "Data Sheet" button and then on "Import" (Figure 9.20). Figure 9.21 shows the imported data sets within *INForm*. Then "Ok"

was clicked to open up the next window, which is setting field types window (Figure 9.22). In this window the inputs and outputs were set. The inputs were labeled in two classifications "Ingredients" and "Processing Conditions". The output was labeled as the "Property". The "Not Used" label was used for inputs or outputs which were not activated within the modeling. For example the Hausner's Ratio will be assigned as "Not Used". The data analysis window (Figure 9.23) will replace the "Set Field Types" window after the inputs/outputs were set. No data analysis was conducted on this window. Next "Ok" was clicked to open up the following "Training" window which will be explained in the next section.
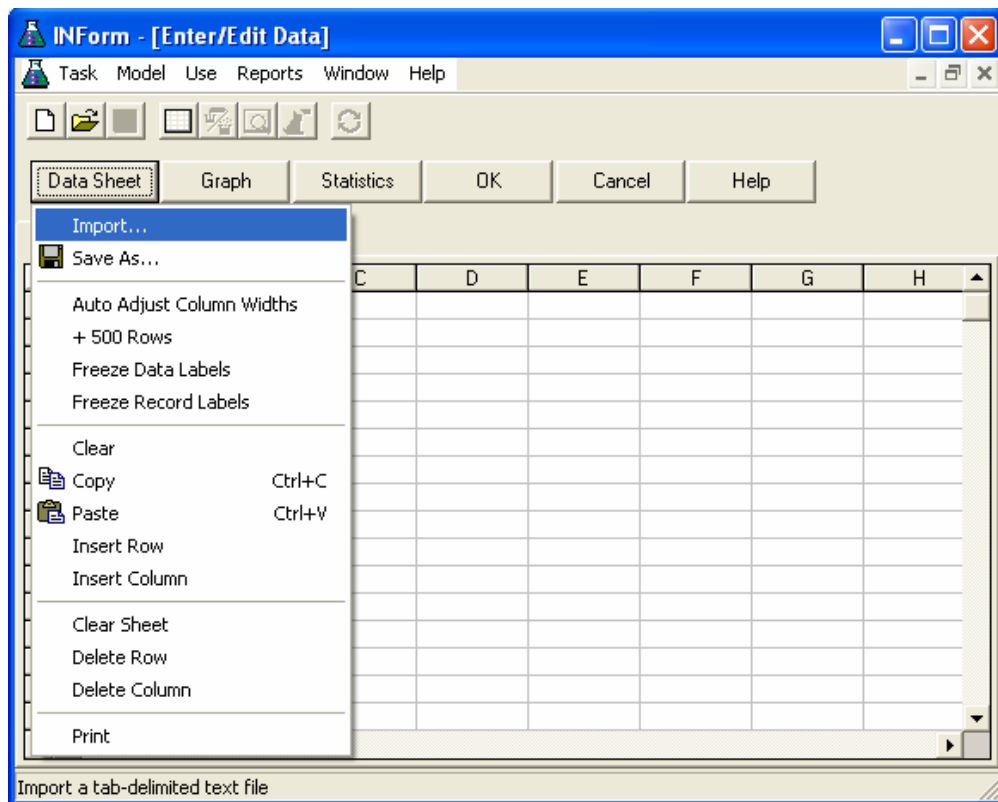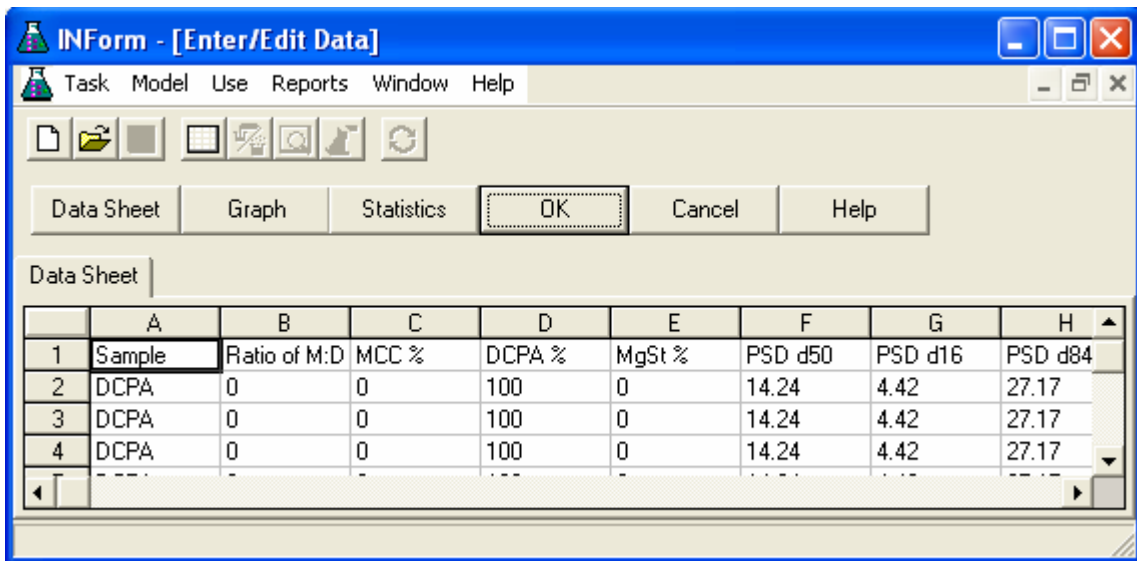

Figure 9.20 Importing data into *INForm* software.

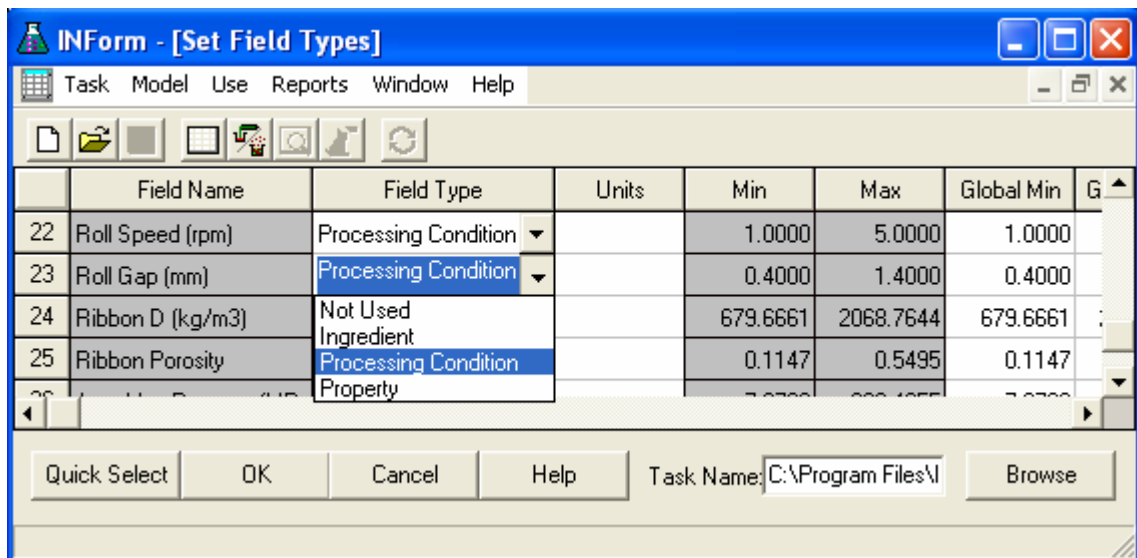Figure 9.21 Imported data set.

Figure 9.22 Setting inputs and outputs.

Figure 9.23 Data analysis.

9.5.2 Model training: Model development and assessment

Figure 9.24 shows the training window. The "Outputs trained" box shows the choices to train the model together in one model or separately. The outputs were trained separately. Next the test data was set by clicking on "Options" button on Figure 9.25. Test data was imported from INF1201testdata.txt. Following this the Model Training Parameters window was opened by clicking on "Parameters". Initially the default training parameters were used (Figure 9.26 to Figure 9.30). The training was initiated by clicking on "Train" button on the training window. Once the training was completed the "Results" box will be activated (Figure 9.31) and the quality of the model will be assessed using the steps shown in Figure 9.36.

In the example shown, clicking on "View Results" button will show us the training results window (Figure 9.32). The data from this page is summarized in the Table

7.1.   The f-ratios are above 4 and the train data R^2s are above 0.85 for all the output property. However for test data the R^2 value was above 0.85 for all except average maximum pressure, hence this model was not a good model and needs to be retrained. The transfer function for the 1st hidden layer will be changed from Asymmetric sigmoid to Tanh.

The retrained model results are shown by Table 7.2. The retrained model shows an acceptable model was trained. Next the linear regression lines for train data and test data were evaluated. Figure 7.3 and Figure 7.4 shows that the linear regression R^2 values were all above 0.85 and this means that the model trained was acceptable. To ensure that the model trained was not memorizing the train and test data, the validation data was predicted and the results are shown in Figure 7.5. The linear regression R^2 values for all the output property were above 0.85. This means that the model trained for the Trial D using Tanh function as the 1st hidden layer transfer function and default parameters was a good model.
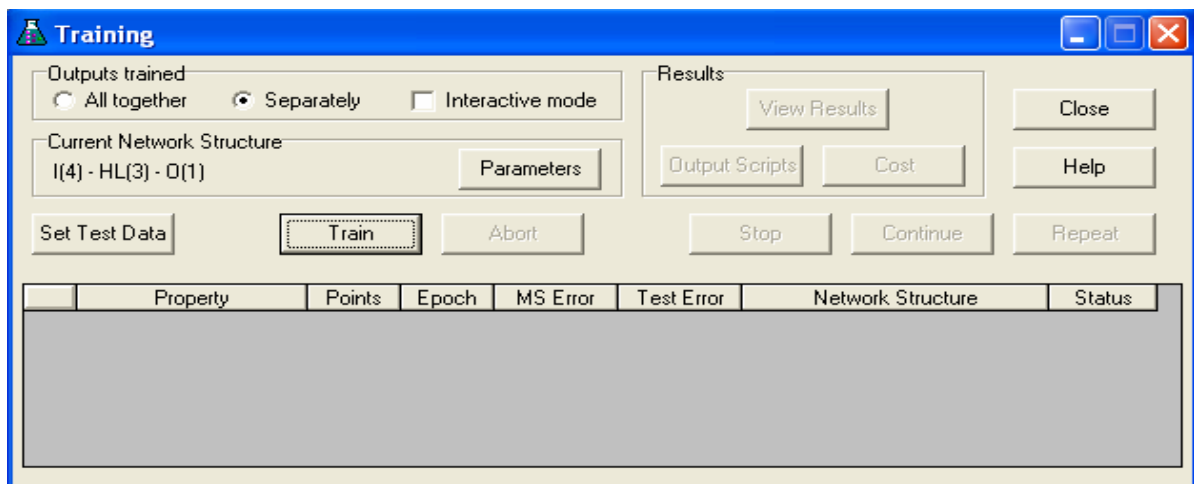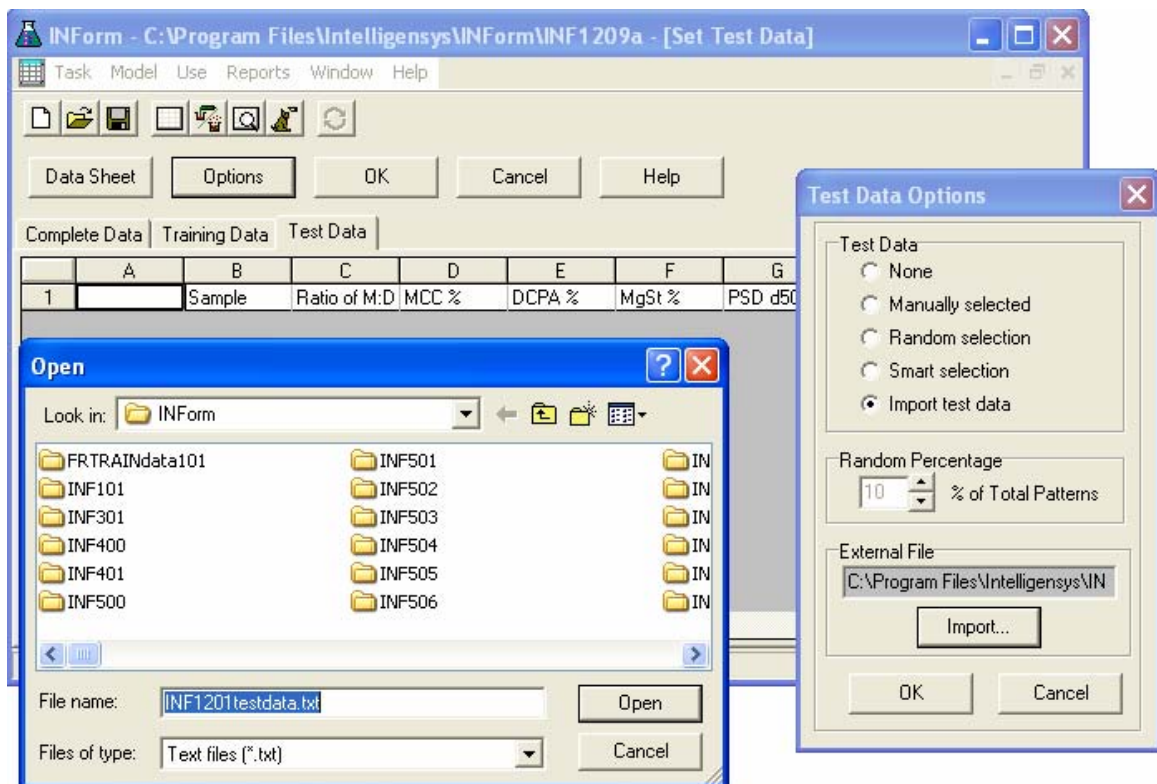
Figure 9.24 Training window.

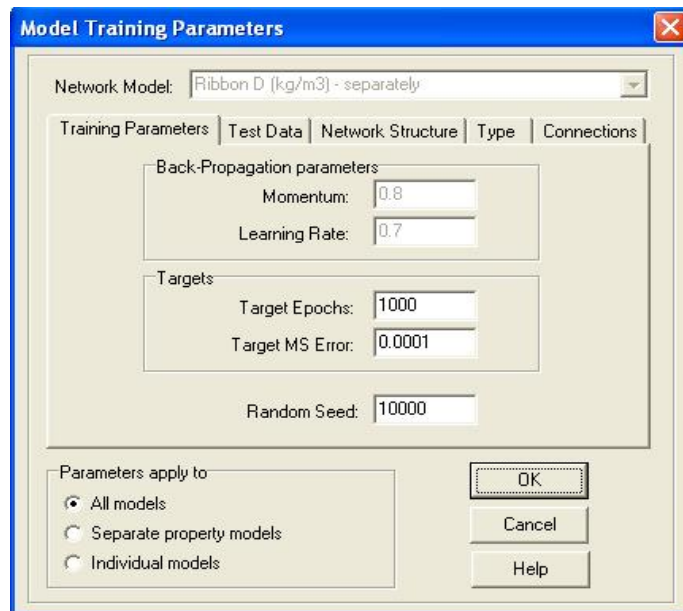Figure 9.25 Setting test data.

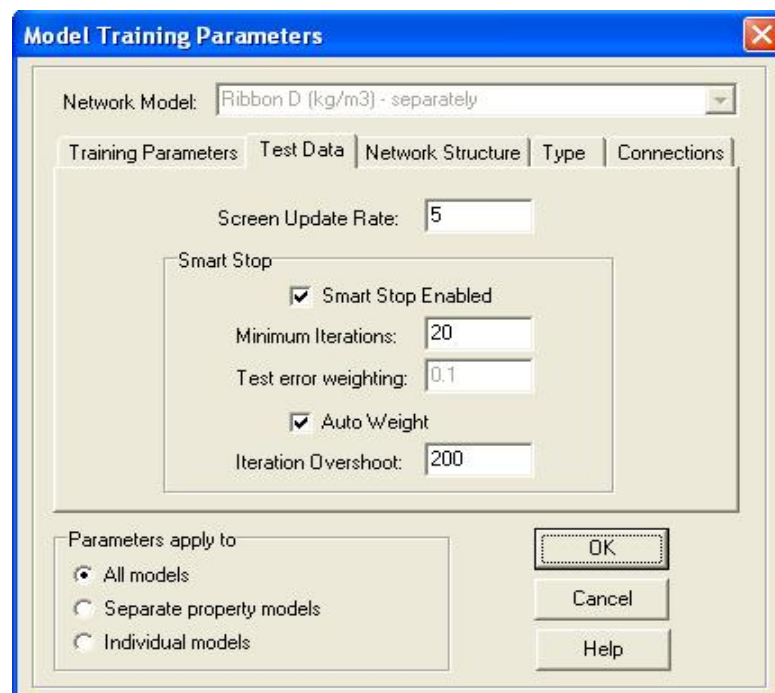Figure 9.26 Model Training Parameters : Training Parameters Tab.

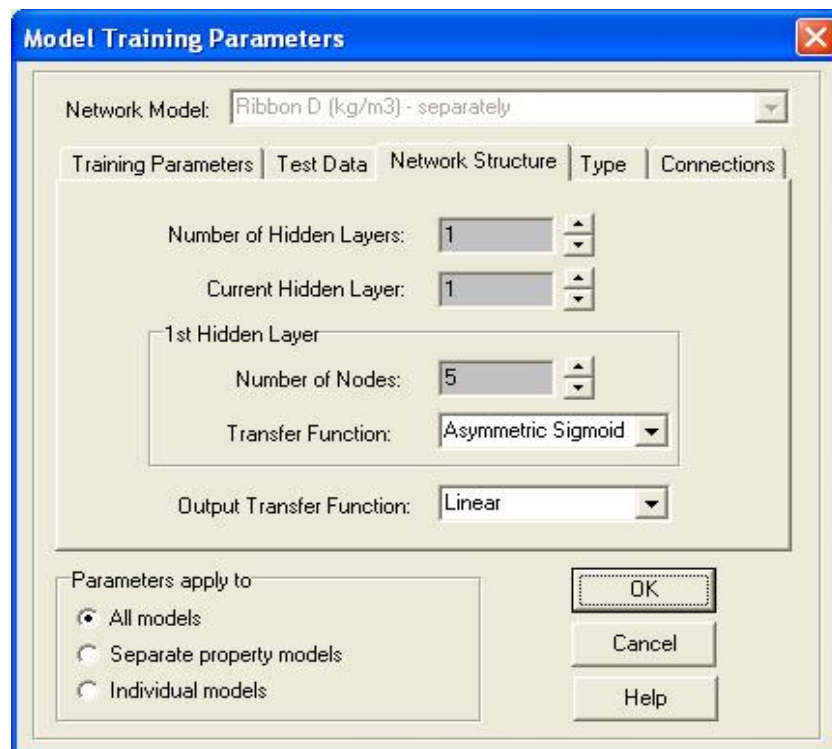Figure 9.27 Model Training Parameters : Test Data Tab.

Figure 9.28 Model Training Parameters : Network Structure Tab.
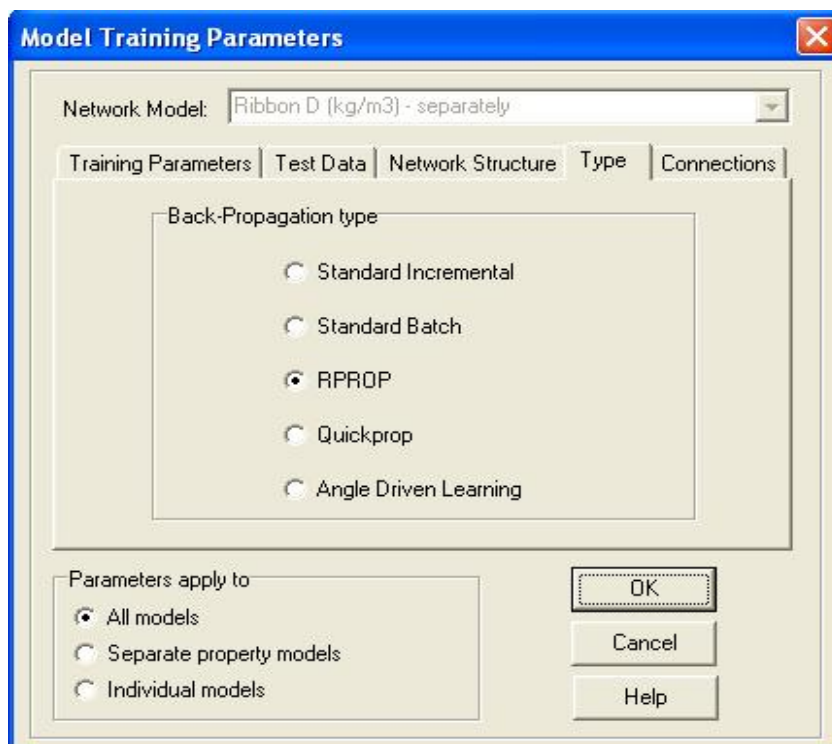
Figure 9.29 Model Training Parameter : Type Tab.
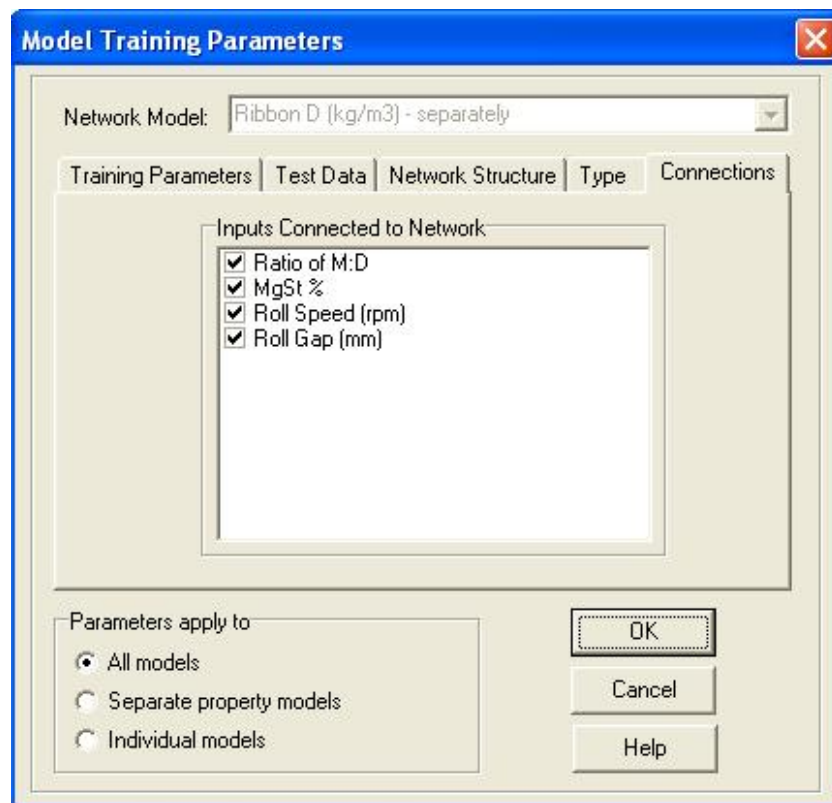
Figure 9.30 Model Training Parameters : Connections Tab.

Figure 9.31 The training results

Figure 9.32 Training Results: Model Statistics

9.5.3 The Training Parameters are explained

The training parameters were explained in this section. However the order in which the parameters are presented is not the order in which the parameters are readjusted before retraining. The stages of readjustments of the network are presented in Figure 7.2.

9.5.3.1   Training parameters: Backpropagation parameters and Targets

On the training parameters tab (Figure 9.26) the options for changing the learning rate, momentum, target epochs and target MS error are given. These training parameters are linked to the backpropagation strategies (explained in section 9.5.3.5). If either the *standard incremental* or *standard batch* was chosen the learning rate and momentum parameters requires adjusting during training. If the *angle driven learning* option was chosen an initial value of the learning rate and momentum needs to be set. However if the *Quickprop* was chosen then only an initial learning rate needs to be set. Otherwise no initial value of the learning rate and the momentum is required for the default *INForm* backpropagation strategy (i.e. *RPROP*).

The target epochs are the number of iterations that had been used in training the network. If the model was unsuccessfully trained there is a possibility that the number of epochs will need to be increased. Note that the number of epochs will not improve the results of the MS error and the test error if the neural network set up was not suitable for the type of data chosen or the type of test data chosen. The number of epochs should be changed only if one of two conditions occurs. The first is if the size of the total data sets which failed to train was extremely large. The second is if the test error and MS Error were observed to be approaching the target values but was stopped before the target convergence could be achieved by a limitation in the number of iterations. Otherwise this parameter is usually the last to be adjusted. The target MS error is 10^{-4} and this value is rarely achieved.

The *random seed* is a function which initializes the network with different small weights. Changing this number would generate new sets of weights for the network and this could effectively improve or worsen the model training. However, this is a good first step to take if the model was not training well.

9.5.3.2   Smart stop: minimum and overshoot iteration

The smart stop ensures that the training stops at the optimum combination of test error and MS error. This parameter is only activated when the test data is used to assist in training the network. According to Colburn (2006), smart stop compares the values of test error against the product of weight multiplied with MS error. The weight can be specified or can be automatically specified by *INForm*. The weight is basically the ratio of test data to the MS error, which normalises the value of MS error to be compared to test error.

If the normalised MS error is greater than the test error, the network training will continue. Otherwise if the normalised MS error is less than the test error, the network training will cease. However the smart stop, allows the calculation cycle to continue beyond this point until the specified number of iterations (iteration overshoot) is reached. Then if the conditions are still not met, the smart stop will restart the calculations from the point at which the test error first exceeded normalised MS error. This can be observed from the Figure 9.33.

The minimum iterations determine the minimum number of calculation cycles the network conducts before the smart stop is implemented. The iteration overshoot ascertains the maximum number of calculation cycles in the 'forward look' before smart stop can be implemented. If the network model training was observed to be stopped too early, then this value should be increased. This is indicated by the quick rise and fall of the test error and the MS error in the graph of mean square error against number of epochs in Figure 9.33.
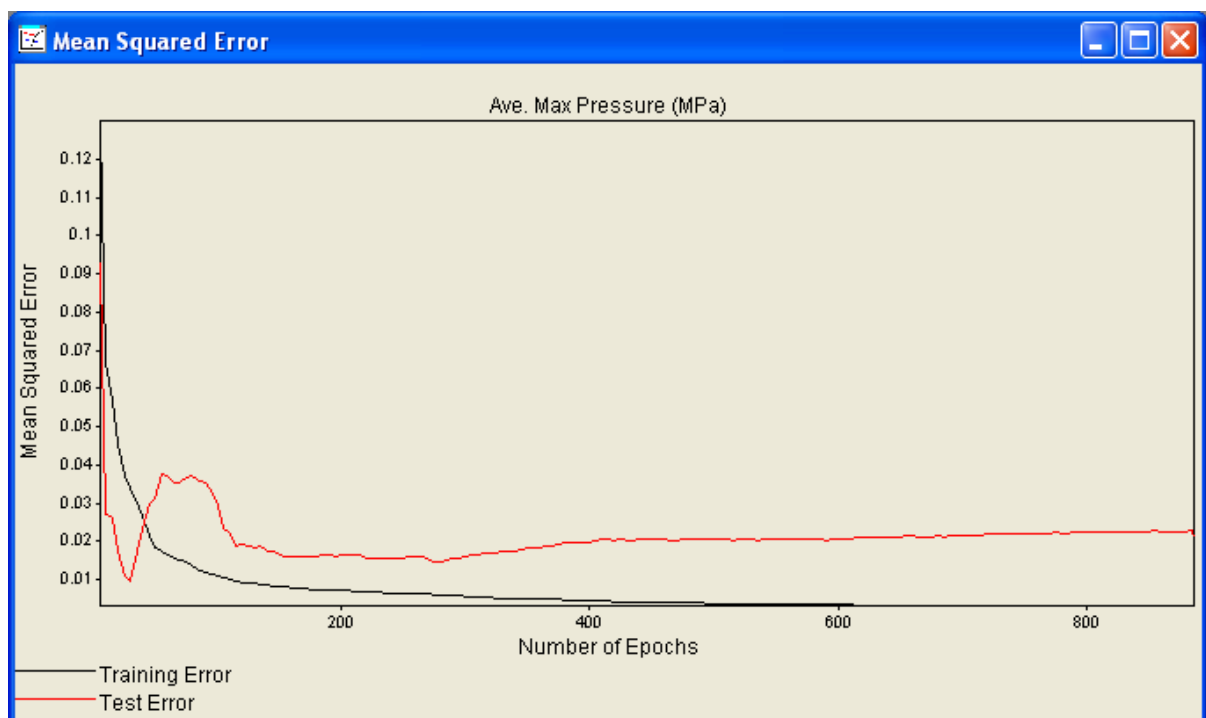

Figure 9.33 Graph of mean square error against number of epochs for average maximum pressure neural network training.

9.5.3.3　Transfer function

On the network structure tab (Figure 9.28), the transfer function for the hidden layers could be changed to Asymmetric sigmoid, Tanh or Symmetric sigmoid.

Sigmoidal and Tanh transfer functions are in the shape of an "S". This Asymmetric and Symmetric transfer function converts the neuron's net input of negative to positive infinity into outputs of values 0 to 1. Whereas Tanh transfer function converts the net input into values between -1 and 1. The output transfer function has the same choices as the hidden layers except for linear transfer function. The linear transfer function will convert the neuron's net input to any value.

Asymmetric sigmoid is the most commonly used transfer (Haykin, 1994) function in the ANN training hence it is used as a default in the *INForm* software for the hidden layer. However the linear transfer function is used for the output layer. The type of transfer function can change the way in which the neural network trains. Thus if the model failed to train under default settings, the user is advised to change the transfer function. It was reported that Tanh (also known as Hyperbolic Tangent Function) is a very flexible, non-linear, continuous and (Bhadeshia, 1999). Hence Tanh has been used in previous trials which showed it seems to work best for most of the data. Note that the hidden layer transfer function would be the second network structure to be changed and followed by the output transfer function (Figure 7.2).

9.5.3.4   Hidden layers

The network structure can be specified on the network structure tab (Figure 9.28). The number of hidden layers is one. This should always be kept at minimum to

avoid developing complex models for a simple input – output relationship. The number of hidden layers can be increased if the model failed to train or the expected relationship between input and output is non-linear and multivariate. The number of nodes in the hidden layer can be changed. Initially during training the default number of nodes was used (5 nodes). However if the model does not train well, then by trial and error the amount of nodes can be adjusted. It is best to use the least amount of nodes to avoid overtraining the network.

9.5.3.5   The backpropagation strategies

The backpropagation strategy can be specified on the type tab (Figure 9.29). These strategies are *standard incremental*, *standard batch*, *angle driven learning*, *resilient backpropagation* (*RPROP*) and *Quickprop*. The subsequent paragraphs will briefly explain these strategies. Backpropagation is one of the most famous training algorithms for multilayer perceptrons (MLP) (Schiffman *et al.*, 1993). It is a gradient descent technique which minimises the error between the target and actual output values to train a neural network MLP.

In MLP training, an initial weight was assumed for the nodes in the hidden layers. Then using these weights an output was calculated. Initially this actual output usually differs largely from the target value. Thus the weights are adjusted to minimise the error function E (also known as Mean Square error (MS error) in the *INForm* program). This error function represents a measure of the performance of

the network (Haykin, 1994; Rumelhart *et al.*, 1986).   A more comprehensive

explanation of the basics of backpropagation is presented in section 2.2.1.1. Figure

9.34 shows the flow of calculation for standard incremental backpropagation.

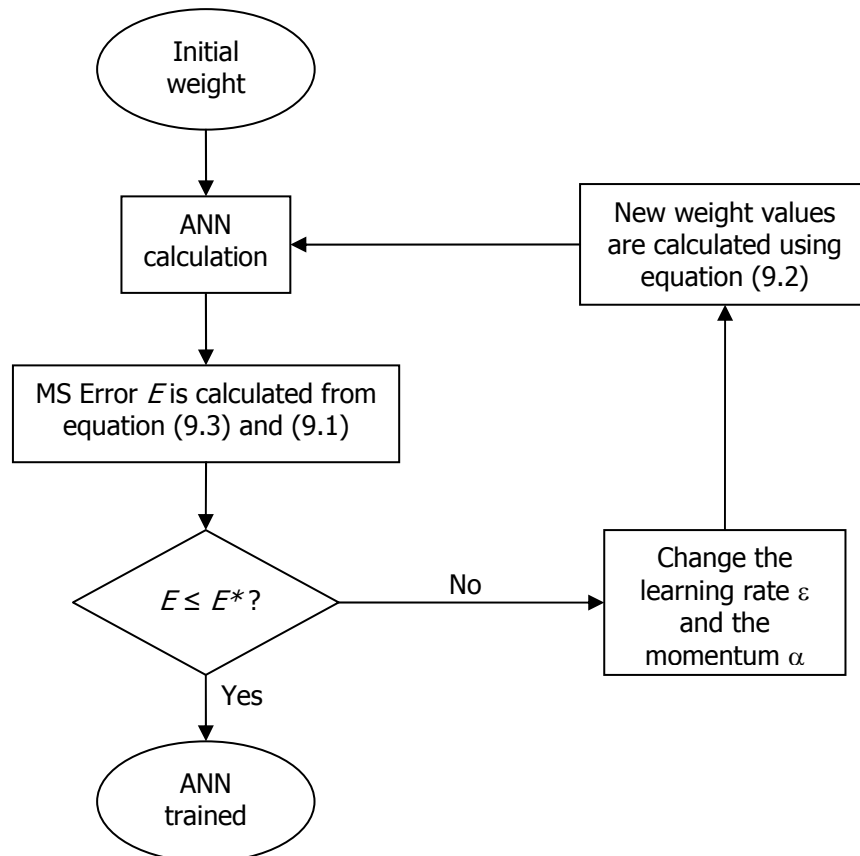Figure 9.34 Flow of calculation for standard incremental backpropagation. E^* is set at 10^{-4} in *INForm*.

This error function E is defined as the mean square sum of differences between the

values of the actual output units of the networks and the desired target values (see

equation (9.1)).

$$E = \frac{1}{P}\sum_{p=1}^{P} E_P = \frac{1}{2N_L}\sum_{p=1}^{N_L}\sum_{j=1}^{N_L} b^P \left(t_j - a_j\right)^2$$

(9.1)

t_j and a_j are the target and actual response values of output neuron j in the pattern p. Note that actual response is the currently calculated output. N_L is the number of output neurons, L is the number of layers and P is the total number of patterns. In the *INForm* programme the square of the difference between target and actual response was multiplied by a belief factor of record pattern p represented by b^p. In terms of formulation, the pattern p is referring to a single experimental recipe. There are two basic backpropagation strategies for updating the weights; *standard incremental* and *standard batch*.

Standard incremental

In *standard incremental* the new weights are calculated each time a new data record is presented to the neural network (Figure 9.34). The calculation is conducted by equation (9.2) (Rumelhart *et al.*, 1986).

$$\Delta w_{ik}(n) = -\varepsilon \frac{\delta E_p}{\delta w_{ik}} + \alpha * \Delta w_{ik}(n-1)$$

(9.2)

Here $\Delta w_{ik}(n)$ is the change in weight for the node connecting node i and node k for the nth iteration of the process. ε is the learning rate and α is the momentum. E_p is the difference between the calculated and observed outputs of the pth pattern.

$$E_P = \frac{1}{2N} \sum_{j=1}^{N_L} b^P (t_j - a_j)^2$$

(9.3)

256

In these backpropagation strategies the learning rate and momentum are two adjustable parameters (non-negative numbers) (Rumelhart *et al.*, 1986). The learning rate allows control of the average size of the weight changes and the momentum function is used to speed up training in very flat regions of the error surface, and it suppresses weight oscillations in step "valleys".

Standard batch

The second alternative given by the *INForm* software is *standard batch* in which the new weights are calculated only after all the patterns have been presented to the neural network. The weight was calculated by replacing the E_p in equation (9.3) with E calculated from equation (9.1). It is important to note that the sum of weights can become very large in this strategy causing the weight change to be large. Following this the weights and the MS error can diverge leading to unsuccessful training. Thus the learning rate must be kept small to avoid this divergence. However the learning rate is dependent on the size and nature of the data set, hence it is important to use trial and error to find the right learning rate value for different sets of data.

Angle driven learning rate

Another backpropagation strategy was to adjust the learning rate during training. It is called the angle-driven learning rate approach proposed by Chan and Fallside (Chan and Fallside, 1987). It is an adaptation of angle between gradient directions in consecutive iterations. The calculation method considered that previous weight

update $\Delta \vec{w}_{n-1}$ and the current gradient descent $\nabla \vec{E}(w_n)$ vector directions gave an angle between θ them. In successive iterations this angle θ gives information about the properties of the error surface. If these two vectors gave similar directions, it indicated stability of the search procedure, and then the learning rate value can be increased. However if a noticeable difference between these directions was detected (shown by a change in the angle value between iterations), it suggests that there are irregular regions present on the error surface. This situation requires a reduction in the learning rate. The angle information is calculated from the cosine rule (equation (9.4)):

$$\cos \theta_n = \frac{-\nabla \vec{E}(w_n) \cdot \Delta \vec{w}_{n-1}}{\| \nabla \vec{E}(w_n) \| \| \Delta \vec{w}_{n-1} \|}$$

(9.4)

Note that the angle between $-\nabla \vec{E}(w_n)$ and $\Delta \vec{w}_{n-1}$ is equal to the angle between $\Delta \vec{w}_n$ and $\Delta \vec{w}_{n-1}$ only when momentum is not used. Otherwise the effect of momentum on the learning rule, combined with the direction of the previous gradient vector $-\nabla \vec{E}(w_{n-1})$ results in a weight update $\Delta \vec{w}_{n-1}$ that follows a direction different from the gradient one. Basically this strategy uses the principle that if the error value decreased to a small value then the learning rate is increased, while if the error value increases, the learning rate is decreased.

The value of the learning rate for the current iteration is obtained by combining its value in the previous iteration with the current value of cos θ (see equation (9.5)).

$$\varepsilon_n = \varepsilon_{n-1}\left(1 + \frac{1}{2}\cos\theta_n\right)$$

<div align="right">(9.5)</div>

Subsequently, once the vectors are in parallel, the adaptation on the momentum was proposed based on the fact that it should be proportional to the value of the learning rate $\alpha_n = \lambda_n \varepsilon_n$. Further details on the momentum could be obtained from Chan and Fallside (1987) or briefly in Moreira and Fiesler (1995).

Quickprop

The next backpropagation strategy given in the *INForm* option is Quickprop, developed by Fahlman (1988). It is mainly composed of a set of heuristics and was adapted from two basic approaches. The first approach was to dynamically adjust the learning rate (based heuristically on the history of the computation) and the second approach was the second-order method (based on Newton's method).

According to Moreira and Fiesler (1995) the main principle of Quickprop is to perform independent optimisation for each weight. It minimises an approximation of its curve by using a parabola whose arms turn upwards (Figure 9.35). At each step, the parabola is minimised using the inclination of the corresponding weight dimension of the error surface in the current and previous step. The weight adjustment equation for this method is

$$\Delta w_{ij}(n) = \frac{\nabla E(w_n)}{\nabla E(w_{n-1}) - \nabla E(w_n)} \Delta w_{ij}(n-1)$$ (9.6)

This equation has no learning rate, but the step size could still be controlled using a "maximum growth factor" µ to avoid the situation at which the parabola determines a very large or infinite step. The upper limit is $\Delta w_{ij}(n) \le \mu \, \Delta w_{ij}(n-1)$ and µ was suggested to be 1.75.
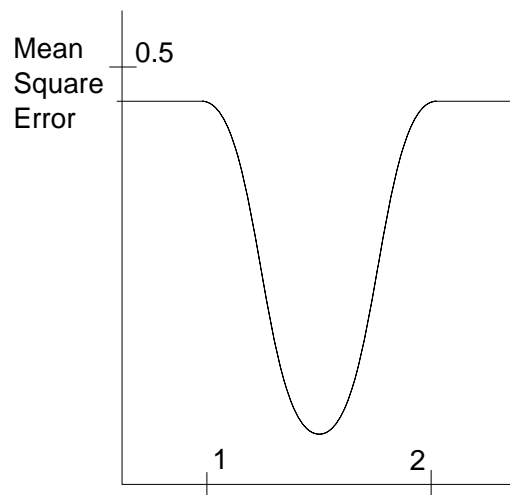
Figure 9.35 Mean square error versus weight value

For the first step and cases where the previous weight step is zero, quickprop uses simple gradient descent (i.e. delta rule, see Moreira and Fiesler (1995)) along with the learning rate. The quickprop was improved by combining both update rules which results in always adding the delta rule update to the quadratic rule, mentioned above. However, when the sign of the steepness in the current step is opposite to the previous one, only the quadratic rule is used.

Further improvements involve adding a weight decay term to each weight to prevent it from growing exceedingly and causing overflow. Finally, the flat-spot elimination technique is proposed. A flat-spot problem causes a weight too small to update which is to be avoided in training. This technique involves always adding a constant of 0.1 to the derivative of the actual output. Quickprop requires a learning rate and momentum to be fed into it to initiate or to restart (if it has been set to zero) training.

RPROP

The default backpropagation strategy given in *INForm* is RPROP. It is an adaptive version of the "Manhattan-Learning" rule (Riedmiller and Braun, 1992). RPROP is based on the sign of the local gradient in consecutive iterations and the weight update is done directly without the use of the derivative or the learning rate. In each iteration, the update rate $\Delta_{ij}(n)$ was added to each weight. The value of $\Delta_{ij}(n)$ is adapted by the gradient sign technique.

$$\Delta_{ij}(n) = \begin{cases} 1.2 * \Delta_{ij}(n-1) & \text{if } \dfrac{\partial E(n-1)}{\partial w_{ij}(n-1)} \dfrac{\partial E(n)}{\partial w_{ij}(n)} \quad \rangle \quad 0 \\[3mm] 0.5 * \Delta_{ij}(n-1) & \text{if } \dfrac{\partial E(n-1)}{\partial w_{ij}(n-1)} \dfrac{\partial E(n)}{\partial w_{ij}(n)} \quad \langle \quad 0 \\[3mm] \Delta_{ij}(n-1) & otherwise \end{cases} \tag{9.7}$$

If the current local gradient is positive, the update-value is subtracted to the weight. On the other hand, if the current local gradient is negative, the update-value is added to the weight. However with a gradient of zero, the weight remains

unchanged. To start the calculation a small initial update value $\Delta_{ij}(n)$ was chosen.

To stop oscillation and arithmetic underflow of floating point values the step sizes were bound by upper and lower limits. The RPROP does not require an input for the learning rate or momentum parameter.

Advantages and disadvantages of the backpropagation strategies

Previous research has shown that Quickprop and RPROP have comparable training speeds (Schiffmann *et al.*, 1994). However Quickprop, RPROP does not require the user to feed it any values for the learning or momentum parameter. Hence for simple data and novice users the RPROP was the best choice for a default parameter. In saying that RPROP has its faults such as it has convergence problems, resulting in a poor mean square error or fails to train the model. In such cases another backpropagation strategy is tried.

The disadvantage of the *standard incremental* and *standard batch* strategies are that both uses the learning rate and momentum parameter, thus in developing the model the user must choose "interactive mode" to change these parameters while training. This can be problematic as it is difficult to adjust these parameters even for relatively uncomplicated problems. Thus for beginners it is not advisable to use these two strategies but if it were attempted, the *INForm* manual can be referred to for possible methods and tips.

Quickprop is different to *standard incremental* and *standard batch* strategy such that it is less dependent on the initial starting value, in that Quickprop adjusts the values of the learning rate during the course of the training. However a large learning rate can cause problems which can be alleviated using a smaller learning rate value.

Angle driven learning rate is similar with Quickprop as it adjust the values of the learning rate during the course of the training. Previous work by Moreira and Fiesler (1995) uncovered that it was not advantageous over the fixed parameter methods, in terms of number of iterations, it could not converge in several situations and it was the least performing of all the adaptive methods studied. In fixed parameter methods the parameters will be maintained at their defined values during training and fewer parameters are adapted to the data.

In general they are no hard rules about which other backpropagation methods should be used if the default fails because it will always depend on the type of data fed into the neural network. So trial and error is used to obtain the best backpropagation method for any data used.

9.5.3.6 Connections

The connections tab is shown by Figure 9.30. It allows the user to break the connection a particular input has with the network. It displays a list of inputs the user has set in the Set Field Types. Each input has a check-box next to its name. By default all the check-boxes are ticked on, which means that each input is

connected to the network. If the box is un-ticked then that input is no longer used in building the model.

There is a subtle difference between setting that input as not used in the Set Field Types screen and breaking the connection that an input has to the network. If set as not used then that input is not used in any property models and the data for that input is not displayed in the results tables. According to the *INForm* Manual, breaking an input's connection to the network can be done for specific output property models at occasions whereby the user knows which inputs are important variables and which ones are not. The input's connection for the non important variables may be disconnected. By excluding inputs that are not relevant, the models are likely to be simpler and easier to interpret. The data for the excluded input is still included in results tables because it may have been used by some of the property models.

9.5.4 Using the model

9.5.4.1   Response surface graph – 3D graph

The response graphs were constructed by first choosing a data set for a specified formulation. This was done by clicking on "View Data" button on Model Consult Window (Figure 9.36), highlighting a specified formulation data set and clicking on "To Consult" button (Figure 9.37). Then clicking on the "3D Graph" button opened

the 3D Graph Setup window (Figure 9.38).   The response graphs can then be

constructed as preferred.

9.5.4.2   "What if predictions"

The "what if predictions" was conducted by filling in values for a specified

formulation and roll compaction process condition in the given column and clicking

on the "Predict" button.   The predicted results appeared in the "Found" column

(Figure 9.39).   These values were compared with the closest input data set by

clicking on the "Properties" button and choosing "Use Found Values" (Figure 9.40).

Then the "Best Match" button was clicked on and "Ingredients" was chosen (Figure
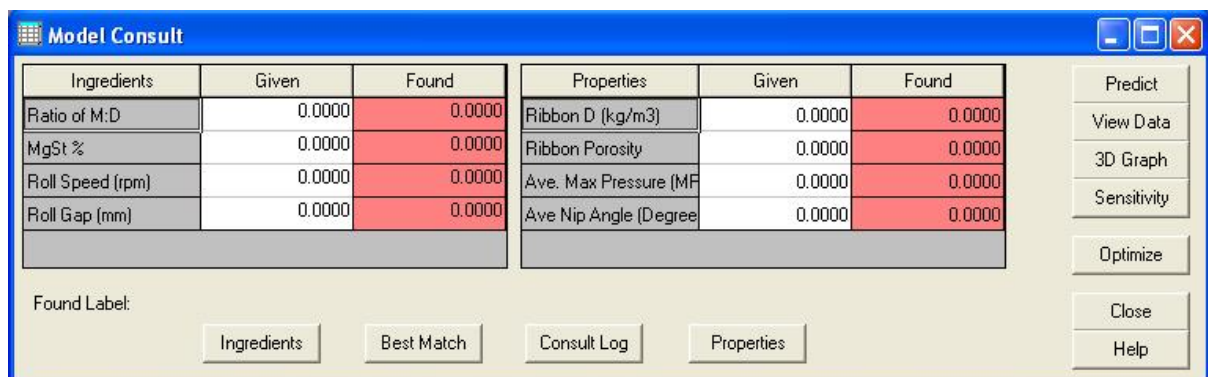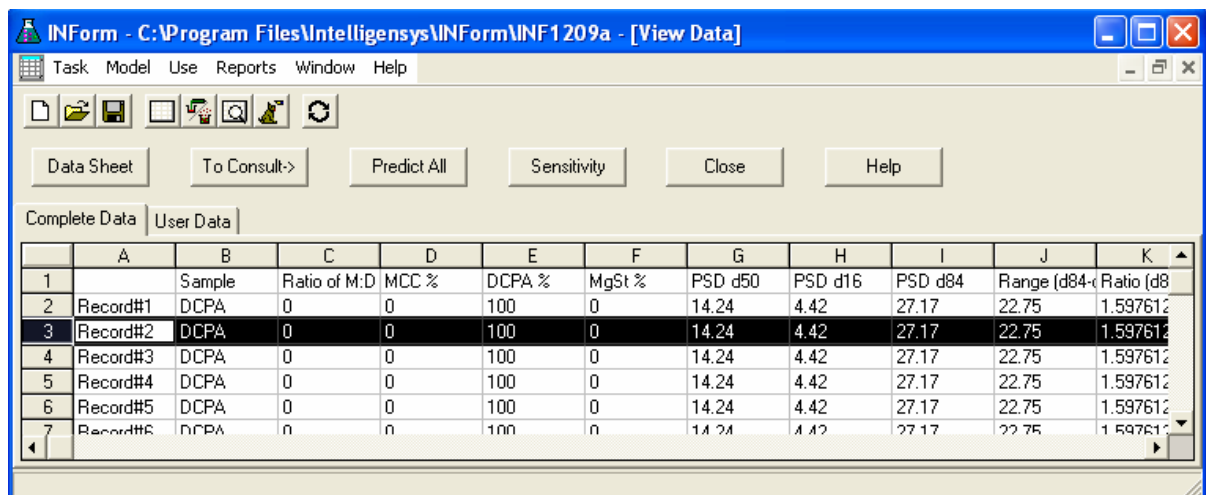
9.41).

Figure 9.36 Model Consult window
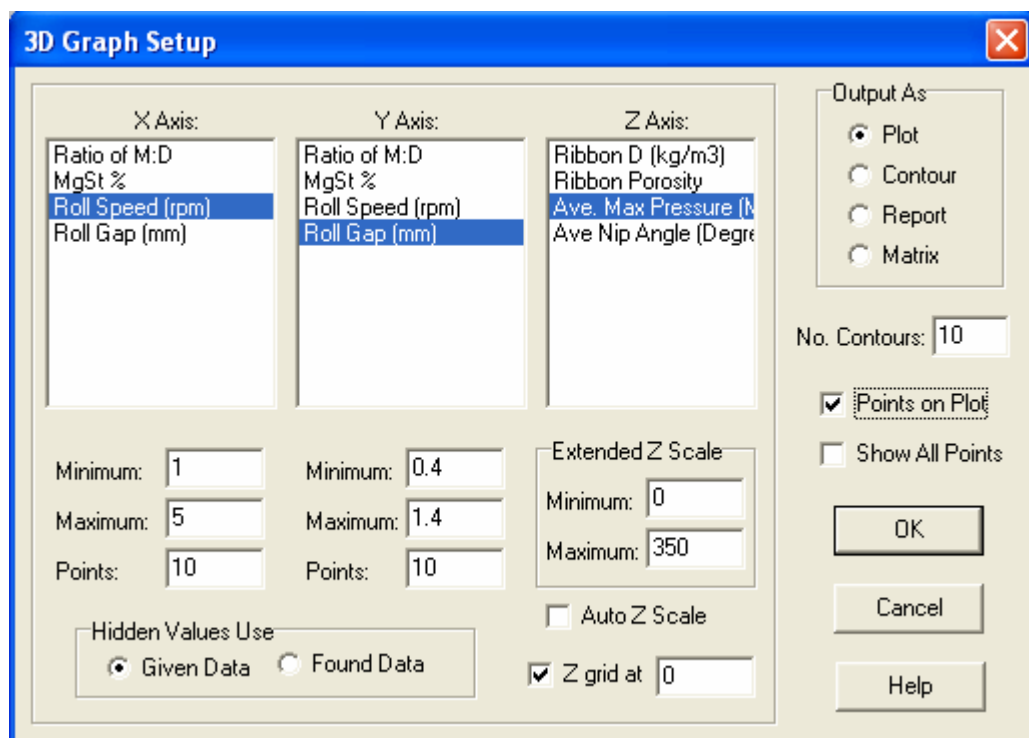
Figure 9.37 Choosing a formulation "To Consult".

Figure 9.38 3D Graph Setup for average maximum pressure against roll speed (rpm) against roll gap (mm).
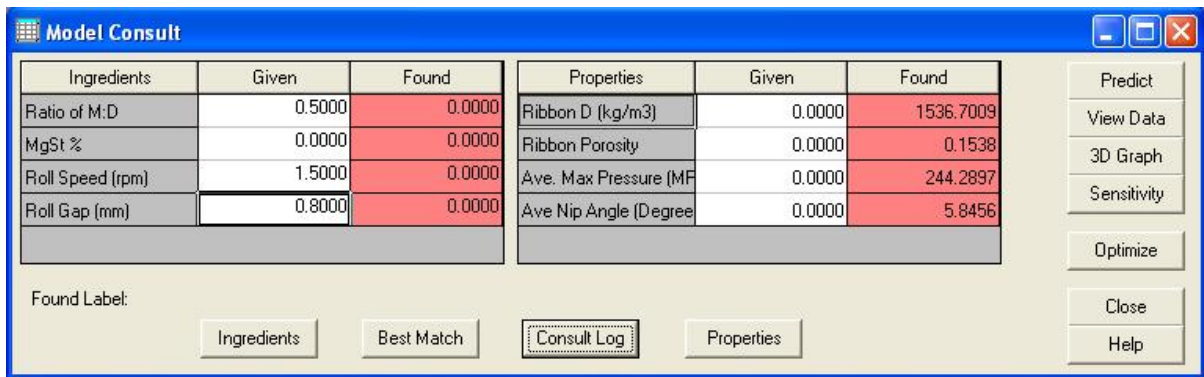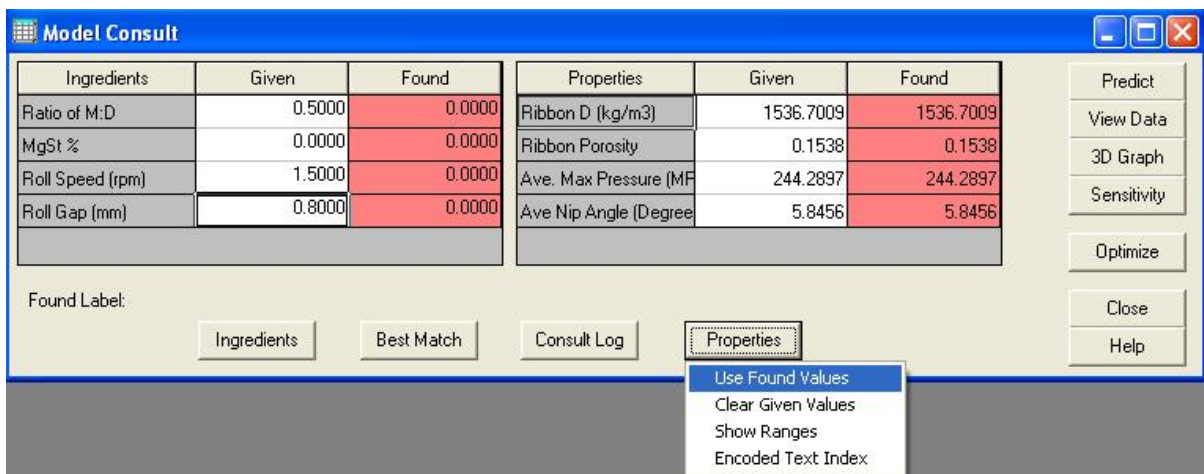
Figure 9.39 Predicted "What if" result.

Figure 9.40 To move the predicted "What if" values into the Given Column from Found Column.
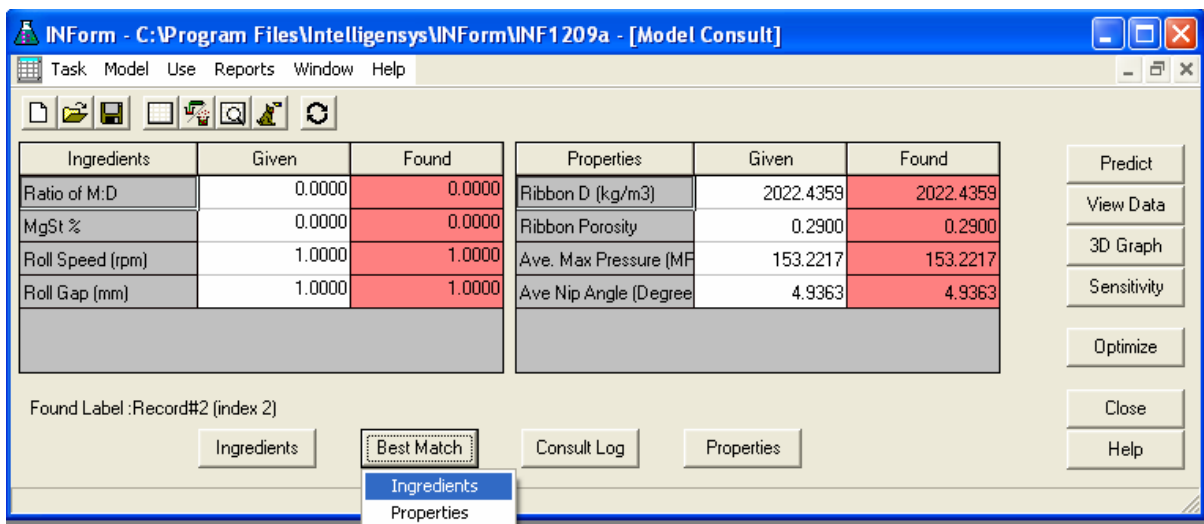
Figure 9.41 To choose best match for either ingredients or properties.

9.5.4.3   Optimising the formulation

To conduct optimisation on the trained model initially a set of data corresponding to the specific formulation and the closest value to the output value should chosen for consulting from the view data option. The optimisation of the model can then be conducted by clicking on "Optimise" button.   The Figure 9.42 shows the optimiser configuration window where the optimisation can be adjusted to obtain an objective.

Trial D will be used as an example in this optimisation.   The objective of this example was to use the optimisation feature to obtain the roll compaction process parameter which produces an MCC ribbon porosity of 0.25.   In this example the ribbon porosity of 0.25 was required and hence a weighting of 10, Mid1 and Mid2 are 0.25 and a desirability function of "tent" was chosen (Figure 9.43).   Next the value of min and max for ribbon porosity were adjusted to be equidistant in value from the Mid1 and Mid2 value because these values should be the arithmetic mean between the min and max points.

The desirability function chosen for ribbon porosity was the "tent" to show that the values between Mid1 and Mid2 are of 100% desirability (Figure 9.44).   The values for the Mid1 and Mid2 are the values which we would like to obtain for that output property.   The weighting, Mid1 and Mid2 for the other output properties were left at default values.   The desirability function was changed to "flat" because the output property values had equal desirability (Figure 9.45).

After the optimizer configuration was set, the "Ok" button on Figure 9.43 was clicked. The Optimize window follows (Figure 9.46). Next click on the "Use Given" button to fill in the "Value" columns with data which was initially chosen in the consult window because by guiding the optimisation by giving it the closest value to the objective value, the user would assist the software in obtaining the desired output faster. Then the in the "Fixed" column, the boxes for the Ratio M:D and MgSt% was ticked. This is because these values represent the formulation type and it is to be kept fixed.

Next the "Optimize" button was clicked on and the user will be asked "Start the optimisation using random ingredient values?" The user should choose "No", since specific values have been fed into the "Value" column. This will initiate the optimisation of the models. After the optimisation is completed the results can be assessed. The desirability for all the outputs should be above 85%.
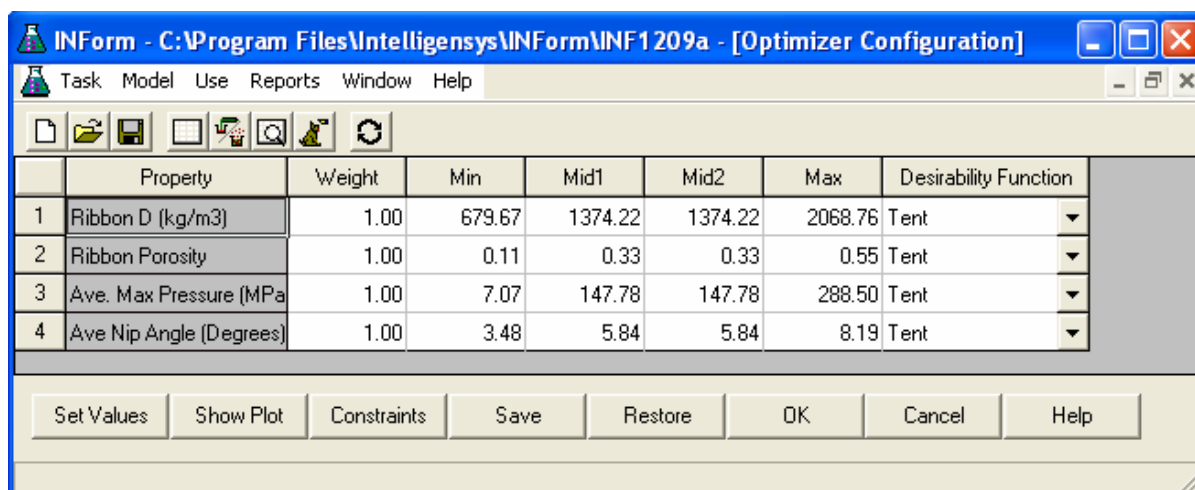
Figure 9.42 Optimizer configuration window
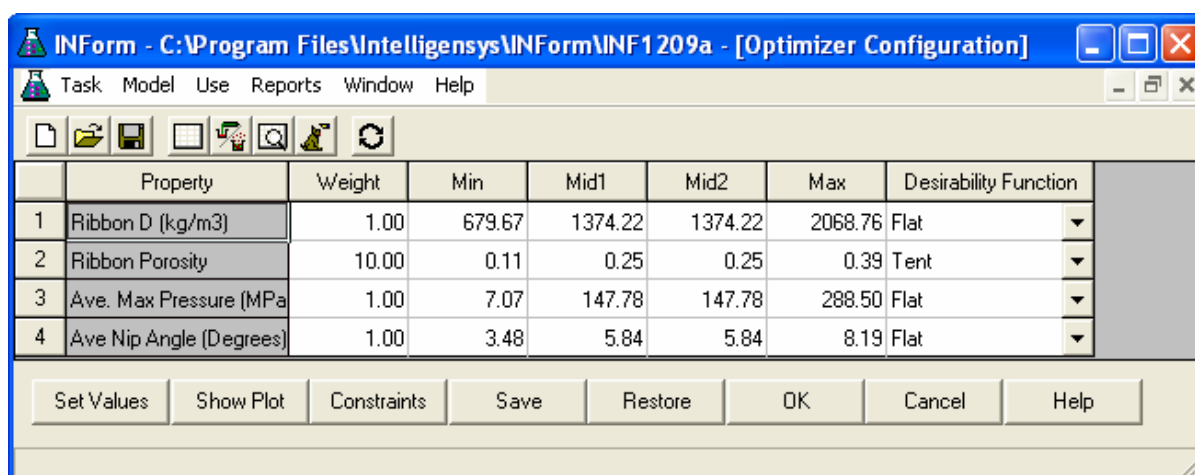
Figure 9.43 Optimizer configuration with maximum ribbon porosity desirability. Showing the need to find the roll compaction process parameter required to produce a ribbon of porosity of 0.25.
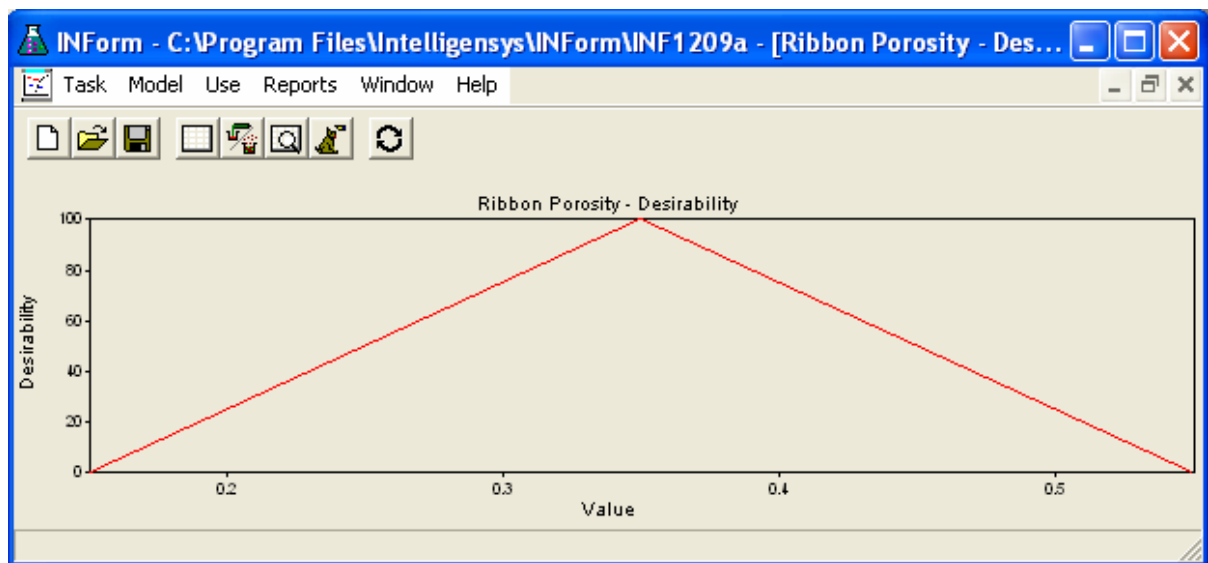
Figure 9.44 Graph of ribbon porosity desirability against values. Which shows the "tent" desirability function.
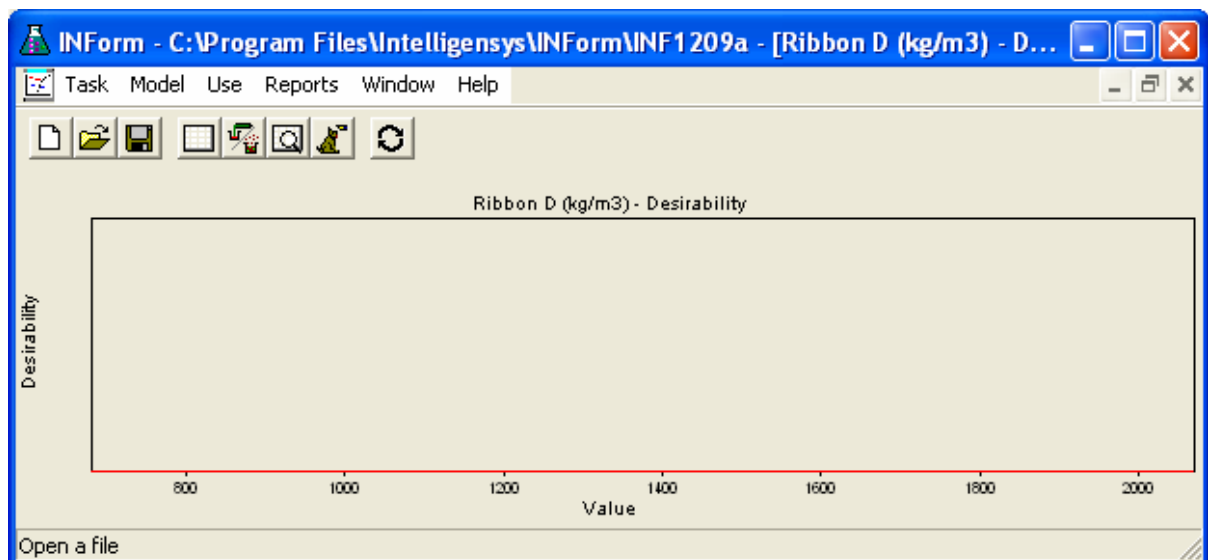
Figure 9.45 Graph of ribbon density desirability against values.   Which is the "flat" desirability function.
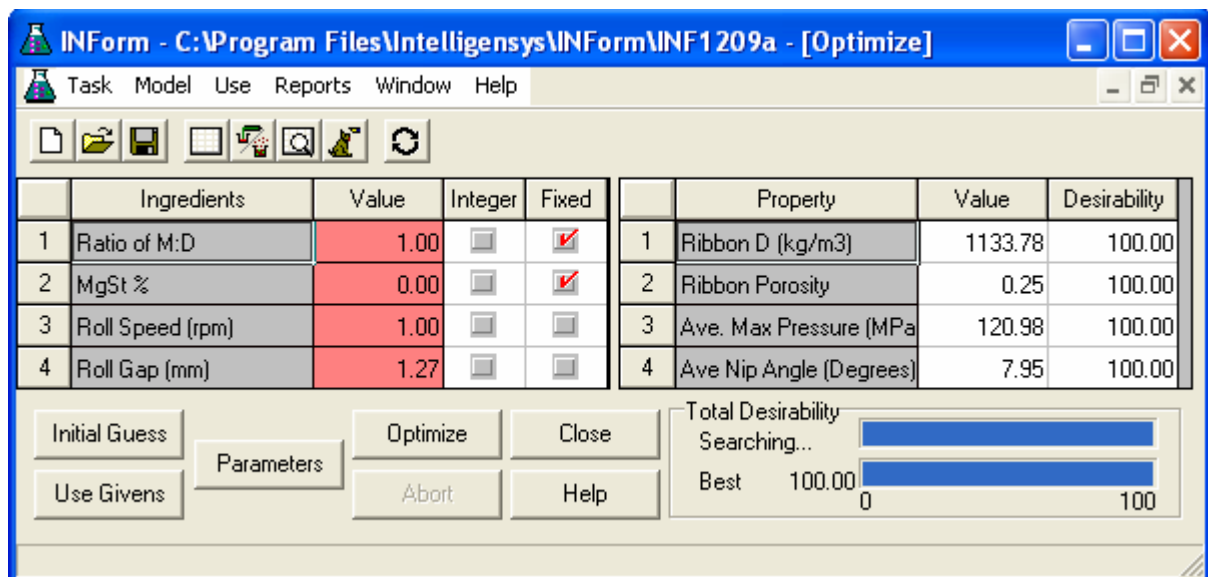
Figure 9.46   Result of optimizing the Trial E for MCC ribbon porosity of 0.25.

9.6 Appendix 6 - List of publications

Rachel Mansa, Rachel Bridson, Richard Greenwood, Jonathan Seville and Helen Barker, (16-18th March, 2005), Using Intelligent Software to Predict the Effects of Formulation and Processing Parameters on Roller Compaction, 8th International Symposium on Agglomeration, Bangkok, Thailand, page 167 – 178.

Rachel Mansa, Rachel Bridson, Richard Greenwood, Jonathan Seville and Helen Barker, (16-18th March, 2004), Using Intelligent Software to Predict the Effects of Formulation and Processing Parameters on Roller Compaction, International Congress for Particle Technology, Partec2004, Nuremberg Germany, page 24.9.

10 References

Agatonovic-Kustrin, S.; Beresford, R. , Basic Concepts of Artificial Neural Network (ANN) Modelling and its Application in Pharmacuetical Research. *Journal of Pharmaceutical and Biomedical Analysis,* **2000**, *22*, 717-727.

Allen, T. V.; Madani, O.; Greiner, R. , Comparing Model Selection Criteria for Belief Networks. *Under submission.,* **2003**.

American Pharmaceutical Association and The Pharmaceutical Society of Great Britain, Handbook of Pharmaceutical Excipients
AMERICANPHARMAC1986 The Pharmaceutical Press:London, 1986; 30-32.

American Society for Testing and Materials , , *Compilation of ASTM Standard Definitions;* 8th ed.; Philadelphia, 1994.

Aulton, M. E. , *Pharmaceutics: The Science of Dosage Form Design;* 2nd ed.; Churchill Livingstone: Edinburgh, 2002; pp. 133.

Bessemer, H., Autobiography. 1905; 68.

Bhadeshia, H. K. D. H. , Neural Networks in Material Science. *ISIJ International,* **1999**, *39* (10), 966-979.

Bindumadhavan, G. , *Roll Compaction of Pharmaceutical Powders*. PhD Thesis The University of Birmingham, 2004.

Bossley, K. M. , *Neurofuzzy Modelling Approaches in System Identification*. PhD Thesis University of Southampton, 1997.

Bourseul, F. R. G. , *Investigation on Roll Pressing as a Forming Operation*. PhD Thesis University of Birmingham, 2001.

British Standards Institution , , *British Standard BS 2955 Glossary of Terms Relating to Particle Technology;* London, 1991.

Brown, R. L.; Richards, J. C. , Principles of Powder Mechanics. 1st ed.; Pergamon Press Ltd: Oxford, 1970; pp 111-112.

Bultmann, J. M. , Multiple Compaction of Microcrystalline Cellulose in a Roller Compactor. *Eur. J. Pharm. Biopharm.,* **2002**, *54*, 59-64.

Burges, C. J. C. , A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery,* **1998**, *2* (2), 121-167.

Cartensen, J. T.; Ertell, C. , Physical and chemical properties of calcium phosphates for solid state pharmaceutical formulations. *Drug Development and Industrial Pharmacy,* **1990**, *16*, 1121-1133.

Chan, L. W.; Fallside, F. , An Adaptive Training Algorithm for Backpropagation Networks. *Computer Speech and Language,* **1987**, *2*, 205-218.

Churchland, P. S.; Sejnowski, T. J. , *The Computational Brain;* MIT Press: Cambridge MA, 1992.

Clarke, B. , Comparing Bayes and non-Bayes Model Averaging when Model Approximation Error Cannot be Ignored. *J. Machine Learning Res.,* **2003** (4), 683-712.

Cohn, R.; Heilig, H.; Delorimier, A. , Critical Evaluation of the Compactor. *Journal of Pharmaceutical Sciences,* **1966**, *55*, 328-331.

Colburn, E. , Personal Communication, Feb 21, 2006.

De Veaux, R. D.; Psichogios, D. C.; Ungar, L. H. , A Comparison of Two Nonparametric Estimation Schemes: MARS and Neural Networks. *Computers Chem Engng,* **1993**, *17* (8), 819-837.

Dec, R. T.; Zavaliangos, A.; Cunningham, J. C. , Comparison of various modeling methods for analysis of powder compaction in roller press. *Powder Techonology,* **2003**, *130*, 265-271.

Edge, S.; Fraser Steele, D.; Ansong, C.; Tobyn, M. J.; Staniforth, J. N. , *International Journal of Pharmaceutics,* **2000**, *200*, 67-72.

Erb, R. J. , Introduction to Backpropagation Neural Network. *Pharmaceutical Research,* **1993**, *10* (2), 165-170.

Fahlman, S. E. , *An Empirical Study of Learning Speed in Backpropagation Networks;*CMU-CS-88-162; Technical Report: 88.

Fell, J. T.; Newton, J. M. , Determination of Table Strength by the Diametrical-Comperssion Test. *Journal of Pharmaceutical Sciences,* **1970**, *59* (5), 688-691.

Friedberg, R. M. , A Learning Machine - part I. *IBM Journal of Research and Development,* **1958**, *2* (1), 2-11.

Friedl, H. and Stampfer, E., Cross Validation, in Encyclopedia of Environmetrics. El-Shaarawi, A. and Piegorsch, W., 2002; 1, 452-460.

Funakoshi, Y.; Asogawa, T.; Satake, E. , Use of a Novel Roller Compactor with a Concave-convex Roller Pair to Obtain Uniform Compacting Pressure. *Drug Development and Industrial Pharmacy,* **1977**, *3*, 555-573.

Gaete-Garreton, L.; Vargas-Hernandez, Y.; Chamayou, A.; Dodds, J. A.; Valderama-Reyes, W.; Montoya-Vitini, F. , Development of an Ultrasonic High-Pressure Roller Press. *Chemical Engineering Science,* **2003**, *58* (19), 4317-4322.

Gereg, G. W.; Cappola, M. L. , Roller Compaction Feasibility for New Drug Candidates: Laboratory to Production Scale. *Pharmaceutical Technology,* **2002**, *26*, 14-23.

Grunwald, P.; Myung, I. J.; Pitt, M. , *Advances in Minimum Description Length: Theory and Applications;* MIT Press: 2005; pp. 5-80.

Hakanen, A.; Laine, E.; Jalonen, H.; Linsaari, K.; Jokinen, J. , Acoustic Emission during Powder Compaction and its Frequency Spectral Analysis. *Drug Development and Industrial Pharmacy,* **1993**, *19*, 2539-2560.

Hansen, M.; Yu, B. , Wavelet Thresholding via MDL for Natural Images. *IEEE Transactions on Information Theory,* **2000**, *46*, 1778-1788.

Hansen, M.; Yu, B. , Model Selection and the Principle of Minimum Description Length. *Journal of the American Statistical Association,* **2001**, *96 (454)*, -746.

Hardy, C.   US Patent No.21314336, 1938.

Haykin, S.   , *Neural Networks; A comprehensive Foundation;* Macmillan College Publishing Company / IEEE Press: New York, New York, 1994.

Hebb, D.   , *The Organization of Behaviour;* Wiley: New York, 1949.

Hirohata, T.; Masaki, S.; Shima, S.   , Experiment on Metal Powder Compaction by Differential Speed Rolling. *J. Materials Processing Technol.,* **2001**, *111* (1-3), 113-117.

Hoerl, A. E.; Kennard, R. W.   , Ridge regression: biased estimation for nonorthogonal problems. *Technometrics,* **1970**, *12* (1), 55-82.

Hubert, M.; Molnar, A.; Jasso, I.   , Contribution to the Theory of Roll Press Design. 2000; p 2.29-2.34.

Inghelbrecht, S.; Remon, J.; Aguiar, P. F.; Walczak, B.; Massart, D. L.; Velde, F. V. D.; De Baets, P.; Vermeersch, H.; De Backer, P.   , Instrumentation of a roll compactor and the evaluation of the parameter settings by neural networks. *International Journal of Pharmaceutics,* **1997**, *148*, 103-115.

Inghelbrecht, S.; Remon, J. P.   , The Roller Compaction of Different Types of Lactose. *International Journal of Pharmaceutics,* **1998**, *166*, 135-144.

Intelligensys., FormRules Manual. 2002.

Jaminet, F.; Hess, H.   , Untersuchungen uber Kompaktierung und Trockengranulierung. *Pharmaceutica Acta Helvetiae,* **1966**, *41*, 39-58.

Jang, J. S. R.; Sun, C. T.; Mizutani, E.   , *Neuro-Fuzzy and Soft Computing;* Prentice Hall: Englewood Cliffs, New Jersey, 1997.

Jeffreys, H.   , *Theory of Probability;* Clarendon Press: Oxford, 1939.

Jenike, A. W. and Shield, R. T., On the Plastic Flow of Coulomb Solids Beyond Original Failure. Journal of Applied Mechanics, Vol.26, Trans.ASME 1959; 81[Series E], 599-602.

Johanson, J. R.   , Stress and Velocity Fields in the Gravity flow of Bulk Solids. *Journal of Applied Mechanics, Vol. 31, Trans. ASME,* **1964**, *86* (Series E), 499-506.

Johanson, J. R.   , A Rolling Theory for Granular Solids. *J. Applied Mechanics,* **1965**, 842-848.

Kass, R. E.; Wasserman, L.   , A Reference Bayesian Test for Nested Hypotheses with Large Samples. *Journal of the American Statistical Association,* **1995**, *90*, 928-934.

Katashinskii, V. P.   , Analytical determination of specific pressure during the rolling of metal powders (in Russian). *Soviet Powder Metal Ceram.,* **1986**, *10* (6), 765-772.

Kearns, M.; Mansour, Y.; Ng, A.; Ron, D.   , An Experimental and Theoretical Comparison of Model Selection Methods. *Machine Learning,* **1997**, *27*, 7-50.

Knight, P.   , Personal Communication, Apr 4, 2003.

Kontkanen, P.; Myllymaki, P.; Silander, T.; Tirri, H.   , On Supvervised Selection of Bayesian Networks. Laskey, K., Prade, H., Eds.; Morgan Kaufmann Publishers: 1999.

Lecompte, T.; Doremus, P.; Thomas, G.; Perier-Camby, L.; Thiesse, J. C. L.; Masteau, J. C. L.; Debove, L. , Dry Granulation of Organic Powders - Dependence of pressure 2D-distribution on different process parameters. *Chemical Engineering Science,* **2005**, *60* (14), 3933-3940.

Li, L. C.; Peck, G. E. , The Effect of Agglomeration Methods on the Micromeritic Properties of a Maltodextrin Product. Maltrin 150. *Drug Development and Industrial Pharmacy,* **1990**, *16*, 1491-1503.

Lindberg, N. L.; Colburn, E. , *Pharmaceutical Technology Europe,* **2004**, *16* (5), 35-39.

Loginov, Y.; Bourkine, S. P.; Babailove, N. A. , Cinematics and Volume Deformations During Roll-Press Briquetting. *Journal of Materials Processing Technology,* **2001** (118), 151-157.

Mackay, D. J. C. , *Bayesian Methods for Adaptive Models*. PhD Thesis Caltech, 1992.

McCulloch, W. S.; Pitts, W. , A logical calculus of the ideas imminent in nervous activity. *Bulletin of Mathematical Biophysics,* **1943**, *5*, 115-153.

Michel, B. , *Contribution à l'étude de l'agglomération des poudres en press à rouleaux lisses*. PhD Thesis Université de Technologie de Compiègne, France, 1994.

Michel, B.; Seville J.P.K.; Guigon, P.; Sidawy, C. , Experimental Study of the Roll Compaction of Powders. JAPAN, 1993.

Minsky, M.; Papert, S. , *Perceptrons;* MIT Press: Cambridge, MA, 1969.

Modha, D. S.; Masry, E. , Prequential and Cross-Validated Regression Estimation. *Machine Learning,* **1998**, *33* (1), 5-39.

Moreira, M.; Fiesler, E. , *Neural Networks with Adaptive Learning Rate and Momentum Terms*;Number: 95-04; IDIAP Technical Report: Martigny, Valais, Suisse, 95.

Naeser and Zirm., *Stahl Eisen*. 1950; 70, 995-1004.

Nedderman, R. M. , *Static and Kinematics of Granular Materials;* Cambridge University Press: Cambridge, 1992; pp. 35-160.

Odagi, K.; Tanaka, T.; Tsuji, Y. , Compressive flow property of powder in roll-type presses - Numerical simulation by discrete element method (in Japanese). *Journal of the Society of Powder Technology,* **2001** (38), 150-159.

Parrott, E. L. , Densification of powders by concavo-convex roller compactor. *Journal of Pharmaceutical Sciences,* **1981**, *70*, 288-291.

Pednault, E., Personal Communication. 2003.

Perera, L. N. , *Roll Compaction of Pharmaceutical Excipients*. PhD Thesis The University of Birmingham, 2004.

Petit-Renaud, A.; Laroche, C.; Guigon, P. , Experimental Study of the Roll Compaction of Powders. Brigthon, U.K., 1998.

Rajasekaran, S.; Pai, G. A. V. , *Neural Networks, Fuzzy Logic, and Genetic Algorithms: Synthesis and Applications;* Prentice Hall of India: New Delhi, 2003.

Rechenberg, I. , *Optimierung technischer Systeme nach Prinzipien der biologischen Evolution (in german), PhD thesis*. Technical University of Berlin, 1970.

Riedmiller, M. and Braun, H., RPROP - A Fast Adaptive Learning Algorithm. 1992.

Rissanen, J. , Fisher Information and Stochastic Complexity. *IEEE Transactions on Information Theory,* **1996**, *42*, 40-47.

Rissanen, J. , Stochastic Complexity in Statistical Inquiry. *Singapore: World Scientific,* **1989**.

Rissanen, J. , Stochastic Complexity. *Journal of the Royal Statistical Society, B,* **1987**, *49*, 223-239.

Rissanen, J. J. , Modelling by the Shortest Data Description. *Automatica-J. IFAC,* **1978**, *14*, 465-471.

Roberts, R. J.; Rowe, R. C. , The Compaciton of Pharmaceutical and other Model Materials - A Pragmatic Approach. *Chemical Engineering Science,* **1987**, *42* (4), 903-911.

Roberts, R. J.; Rowe, R. C.; York, P. , The relationship between the fracture properties, tensile strength and critical stress intensity factor of organic solids and their molecular structure. *International Journal of Pharmaceutics,* **1995**, *125* (1), 157-162.

Rosenblatt, F. , The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review,* **1958**, *65*, 386-408.

Rowe, R. C.; Roberts, R. J. , *Intelligent Software for Product Formulation;* Taylor & Francis: London, 1998; pp. 3.

Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. , *Parallel Distributed Processing : Explorations in the Microstructure of Cognition;* MIT Press, Cambridge, Massachusetts, 1986; Vol. 1 pp. 318-362.

Salonen, J.; Salmi, K.; Hakanen, A.; Laine, E.; Linsaari, K. , Monitoring the Acoustic Activity of a Pharmaceutical Powder during Roller Compaction. *International Journal of Pharmaceutics,* **1997**, *153*, 257-261.

Schiffman, W.; Joost, M.; Werner, R. , Proceeding of the European Symposium on Artificial Neural Networks. 1993; pp 97-104.

Schiffmann, W., Joost, M., and Werner, R., Optimization of the Backpropagation Algorithm for Training Multilayer Perceptrons. 29-9-1994.

Schonert, K.; Sander, U. , Shear Stresses and Material Slip in High Pressure Roller Mills. *Powder Technology,* **2002**, *122*, 136-144.

Schulze, D. , *Ring Shear Tester RST-XS: Operating Instructions;* Schulze: Wolfenbuttel, 2002.

Schulze, D. Fundamentals of Bulk Solid Mechanics. *http://www. dietmar-schulze. de/grdle1. html* **2003**, 3-4.

Schwefel, H. P. , *Kybernetische Evolution als Strategie der Experimentellen Forschung in der Stromungstechnik, Diploma thesis*. Technical University of Berlin, 1965.

Seville, J. P. K.; Tuzun, U.; Clift, R. , *Processing of Particulate Solids;* Blackie Academic & Professional: London, 1997; pp. 35-37.

Siemens; Halske , German Patent No. 154998, 1904.

Simon, O.; Guigon, P. , Correlation between Powder-Packing Properties and Roll Press compact heterogeinity. *Powder Technology,* **2003**, *130* (1-3), 257-264.

Skapura, D. M. , *Building Neural Networks;* Addison-Wesley: Reading, MA, 1996.

Snow, R. H.; Allen, T.; Ennis, B. J.; Litster, J. D. , Size Reduction and Size Enlargement. In *Perry's Chemical Engineers Handbook*, 7th ed.; Perry, R. H., Green, D. W., Maloney, J. O., Eds.; McGraw-Hill: 1997; pp 20-70.

Spinov, V. A.; Vinogradov, G. A. , Influence of Air on the Rolling of Powders (in Russian). *Poroskaya Metallurgyia,* **1967**, *8* ((56)), 96-99.

Staniforth, J. , Powder Flow. In *Pharmaceutics: The Science of Dosage Form Design*, 2nd ed.; Aulton, M. E., Ed.; Churchill Livingstone: Edinburgh, 2002; pp 197-210.

Stone, M. , Cross-Validatory Choice and Assessment of Statistical Predictions (with discussion). *Journal of the Royal Statistical Society, Series B,* **1974**, *36*, 276-278.

Subramanian, N.; Yajnik, A.; Murthy, R. S. R. , Artificial Neural Network as an Alternative to Multiple Regression Analysis in Optimizing Formulation Parameters of Cytarabine Liposomes. *AAPS PharmSciTech 2004,* **2004**, *5* (1), 1-9.

Timofei, S.; Kurunczi, L.; Suzuki, T.; Fabian, W. M. F.; Muresan, S. , Multiple Linear Regression (MLR) and Neural Network (NN) Calculations of some Disazo Dye Adsorption on Cellulose. *Dyes and Pigments,* **1997**, *34* (3), 181-193.

Tsuji, Y.; Tanaka, T.; Ishida, T. , Lagrangian Numerical Simulation of Plug Flow of Cohesionless Particles in a Horizontal Pipe. *Powder Technology,* **1992**, *71*, 239-250.

Turkoglu, M.; Aydin, I.; Murray, M.; Sakr, A. , Modelling of a roller-compaction process using neural networks and genetic algorithms. *Eur. J. Pharmaceutics and Biopharmaceutics,* **1999**, *48*, 239-245.

Vapnik, V. , The Nature of Statistical Learning Theory. *Springer-Verlag,* **1995**.

Vapnik, V. , *Estimation of Dependences Based on Empirical Data (in Russian);* Nauka: Moscow, 1979.

Veale, C. R. , *FINE POWDERS: Preparation, Properties and Uses;* Applied Science Publisher Ltd: Essex, 1972; pp. 138.

Weakliem, D. L. , A Critique of the Bayesian Information Criterion for Model Selection. *Sociological Methods & Research,* **1999**, *27* (3), 359-397.

Willshaw, D. J.; Von der Marlsburg, C. , How Patterned Neural Connections can be set up by Self-Organization. *Proc. of hte Royal Society of London Series B,* **1976**, *194*, 431-445.

Yusof, Y. A.; Smith, A. C.; Briscoe, B. J. , Roll Compaction of Maize Powder. *Chemical Engineering Science,* **2005**, *60*, 3919-3931.

Zadeh, L. A. , Fuzzy Sets. *Information and Control,* **1965**, *8* (3), 338-353.

Zinchuk, A. V.; Mullarney, M. P.; Hancock, B. C. , Simulation of roller compaction using a laboratory scale compaction simulator. *International Journal of Pharmaceutics,* **2004**, *269* (2), 403-415.

Space used so that the printing would produce the thesis without the quotes in the text and in the form required.. due to difficulty in using reference manager.

Not to be sent in

(Bessemer, 1905)   Hardy, 1938; Siemens and Halske, 1904   (Naeser and Zirm, 1950)   Johanson, 1965   Katashinskii, 1986   Inghelbrecht *et al*, 1997 Turkoglu *et al*, 1999   Odagi *et al*, 2001   Loginov *et al*, 2001   Dec *et al*, 2003   Zinchuk *et al*, 2004   Rowe and Roberts, 1998   McCulloch and Pitts, 1943; Hebb, 1949   Rosenblatt, 1958   Minsky and Papert, 1969   Willshaw and Von der Marlsburg, 1976   Friedberg, 1958   Rechenberg, 1970; Schwefel, 1965   Zadeh, 1965   Cartensen and Ertell, 1990 Bindumadhavan, 2004   Kass and Wasserman, 1995   Cohn *et al*, 1966; Jaminet and Hess, 1966   Roberts *et al*, 1995   (Michel, 1994; Bourseul, 2001).   Johanson, 1965   Jenike and Shield, 1959   Grunwald *et al*, 2005   Stone, 1974   Cohn *et al*, 1966   Tsuji *et al*, 1992   Michel *et al*, 1993   Funakoshi *et al*, 1977   Petit-Renaud *et al*, 1998   Hirohata *et al*, 2001   Lecompte *et al*, 2005   Hubert *et al*, 2000   Yusof *et al*, 2005 Hakanen *et al*, 1993   Salonen *et al*, 1997   Gaete-Garreton *et al*, 2003 Haykin, 1994; Rumelhart *et al*, 1986 Jang *et al*, 1997   Seville *et al*, 1997   Snow *et al*, 1997   Edge *et al*, 2000   De Veaux *et al*, 1993; Timofei *et al*, 1997; Subramanian *et al*, 2004   Allen *et al*, 2003; Kontkanen *et al*, 1999; Modha and Masry, 1998 Clarke, 2003; Kearns *et al*, 1997; (Pednault, 2003)   Schiffman *et al*, 1993 (Schiffmann *et al*, 1994)