

RHEOLOGICAL MODEL FOR PAINT PROPERTIES

by

Pieter Lafras Moolman

Dissertation presented for the Degree

of

DOCTOR OF PHILOSOPHY (Chemical Engineering)

in the Department of Process Engineering
at the University of Stellenbosch



Prof. J.H. Knoetze

STELLENBOSCH

March 2008

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:

Date:

ABSTRACT

The feasibility of predicting paint properties directly from the raw material formulation as well as the rheological data is investigated in this study. Although extensive work has been carried out on the prediction of paint properties in terms of the raw material data, very little research has been carried out on the prediction of paint properties in terms of the rheological data. Little is known about the relationship between fundamental rheological properties and real-world performance.

The paint under investigation consists of fourteen raw materials. These raw materials interact in a very complex manner to produce certain desired paint properties. Evaluation of these interactions in terms of constitutive equations is almost impossible and the relationships between paint properties, raw materials and rheology can only be modelled in a statistical way. Linear relationships are investigated with linear parameter estimation techniques such as multiple linear regression. However, it has been found that many of these relationships are non-linear and that linear modelling techniques are no longer applicable for certain situations, e.g. at very high concentrations of specific raw materials. Non-linear techniques such as neural networks are used in these situations.

The relationship between the raw materials, paint properties and rheology are evaluated using the following three models:

- **MODEL 1:** The relationship between rheology and raw materials
- **MODEL 2:** The relationship between paint properties and raw materials
- **MODEL 3:** The relationship between paint properties and rheology

MODEL 1 makes use of techniques such as principal component analysis and preliminary modelling to respectively reduce redundancy and to capture as much data as possible. **MODELS 2 and 3** make use of linear screening techniques in order to identify relevant raw materials and paint properties.

The validity of every model is checked to ensure that predictions and interpretations are unbiased and efficient. **MODEL 1** revealed that emulsion, extender particles, pigment, water, organic pigment and solvent are the six most important raw materials affecting the rheology of the specific paint. The rheology curves that are predicted most accurately by means of multiple linear regression are the “*Amplitude Sweep*” (AS), “*3-Interval-Thixotropy-Test*” (3-ITT) and the “*Flow Curve*” (FC). Non-linear rheological behaviour is encountered at high pigment volume concentrations (PVC) and volume solids (VS), due to the strong dependency of the rheology of the paint on these properties. It has been shown that neural networks perform better than multiple linear regression in predicting the rheological behaviour of these paint samples for which the raw materials vary by more than 20% from the standard formulation. On average, neural networks

improve predictability of the rheological parameters of these samples by 54%. The largest improvement in predictability is made on the rheological variable “*Extra Low Frequency*” value (C_{XLF}), where multiple linear regression resulted in relative errors of 59%, while neural networks resulted in errors of only 5%. Other predictions of rheology curves where neural networks have shown a major improvement on predictability are the “*Time Sweep*” (*TS*) – 68% increase in accuracy and “*Low Shear*” curve (*LS*) – 63% increase in accuracy. The smallest increase that the neural network had on the predictability of a rheology curve, was a 33% increase in accuracy of the “*Amplitude Sweep*” (*AS*) predictions.

Multiple linear regression models of **MODEL 2** predict the critical paint properties of *Opacity*, *Gloss*, *Krebs Viscosity* and *Dry Film Thickness* with relative errors smaller than 10%. It has been shown that 90% of all new predictions fall within the allowable error margin set by the paint manufacturer. Paint properties that can be predicted with an expected error of between 10% and 20% are *Dry and Wet Burnish*, *Open Time* and *Water Permeability*. Paint properties that are predicted the most inaccurately by MLR, that results in errors larger than 20% are *Dirt Pick-Up* and *Sagging*. Non-linear techniques such as neural networks are used to predict the paint properties of these paint samples for which the raw materials vary by more than 20% from the standard formulation. The neural networks show a major improvement on the predictability of the paint properties for those paint samples that vary more than 20% from the standard formulation. On average, neural networks improve predictability of the paint properties by 47%. The largest improvement in predictability is made on the *Wet Burnish20* prediction, where multiple linear regression resulted in relative errors of 66%, while neural networks resulted in errors of only 0.6%. Other paint property predictions where neural networks have shown a major improvement on predictability of 80% or more in accuracy are *Gloss* – 80% increase in accuracy and *Dry Film Thickness* – 81% increase in accuracy. The smallest increase that the neural network had on the predictability of a paint property, was a 33% increase in accuracy of the *Sag* predictions. **MODEL 2** makes it possible for the paint manufacturer to test tolerances around certain paint properties during manufacturing.

Rheology is still a very under-utilised tool for explaining certain paint properties. **MODEL 3** quantified the correlation between fundamental rheological properties and real world performance of a paint. It has been shown that rheological measurements can be used accurately to predict certain critical paint properties such as *Opacity*, *Krebs Viscosity*, *Dry Film Thickness* and *Gloss* within the allowable error margin given by the paint manufacturer. Multiple linear regression models predict the paint properties of *Opacity*, *Krebs Viscosity* and *Dry Film Thickness* with relative errors smaller than 10%, with rheology as input to the model. A neural network of **MODEL 3** was developed to predict the *paint properties* of those paint samples that vary more than 20% from the standard formulation, by using *rheology data* as input to the model. The neural networks perform better than multiple linear regression in predicting the paint properties of these paint samples. On average, neural networks that use rheology data as input, predict the paint properties 49% more accurate than equivalent multiple linear regression models. The greatest

improvement in model predictability is for *Water Permeability* - 73% increase in accuracy and *Gloss* - 70% increase in accuracy.

Two models **MODEL 2** (raw material composition as input) and **MODEL 3** (rheology data as input) have been developed for the prediction of paint properties. The table below shows a comparison between the linear models based on the average relative error that can be expected for new paint property predictions.

MODEL 2 AND MODEL 3 PAINT PROPERTY PREDICTIONS

Paint Property	Average Relative Error (%)	
	Model 2	Model 3
Dry Burnish (60°)	10.7	20.0
Wet Burnish (60°)	14.8	15.3
Dirt Pick-Up	43.6	48.0
Open Time	17.3	17.4
Dry Film Thickness	6.4	7.1
Gloss (60°)	7.2	15.3
Hiding Power	0.8	1.5
Krebs Viscosity	1.6	1.6
Opacity	0.8	1.1
Sag	43.2	38.9
Water Permeability	14.2	12.5

Akaike's information criterion (AIC) was also used to discriminate between the linear and non-linear models. The AIC for all three models indicate that, in general, the multiple linear regression models do better in condensing information compared to the neural network models of the paint samples for which the raw materials vary by more than 20% from the standard formulation.

This study has made some major contributions in the coatings industry, such as indicating the feasibility of predicting paint properties in terms of both raw material data and rheological data. The next step is to combine the two sets of data. This could possibly lead to even more accurate predictions in the paint properties. Cost factors for the raw materials can also be included to optimise the paint composition for the lowest cost, but still with acceptable paint properties. This study showed the feasibility of predicting paint properties in terms of raw material data and rheological data for a *single paint family*. This study can now be used as a basis for extending the current understanding of paint properties in terms of raw materials and rheology, to include a variety of different paint families.

OPSOMMING

Die uitvoerbaarheid daarvan om verfeienskappe direk vanaf 'n roumateriaalformulasie sowel as die reologiese data te voorspel, word in hierdie studie ondersoek. Breedvoerige navorsing aangaande die voorspelling van verfeienskappe in terme van roumateriaaldata is reeds uitgevoer, maar relatief min navorsing is al uitgevoer aangaande die bepaling van verfeienskappe in terme van reologiese data. Daar is min inligting beskikbaar oor die verhouding tussen fundamentele reologiese eienskappe van 'n verf en die werklike werksverrigting.

Die verf wat hier ondersoek word, bestaan uit veertien roumateriale. Die roumateriale reageer op mekaar op 'n besonders ingewikkelde wyse om sekere gewenste verfeienskappe tot gevolg te hê. Die evaluering van hierdie wisselwerkings in terme van koppelvergelykings is 'n byna onmoontlike taak en die verhoudings tussen verfeienskappe, roumateriale en reologie is grootendeels statisties van aard. Lineêre verhoudings word deur middel van lineêre parameter beramingstegnieke, insluitend meervoudige lineêre regressie, ondersoek. Dit is egter gevind dat hierdie verhoudings grotendeels nie-lineêr van aard is en dat lineêre modelleringstegnieke nie toepaslik is in sekere gevalle nie, bv. in gevalle waar baie hoë konsentrasies van spesifieke rou materiale voorkom. In hierdie gevalle word daar van nie-lineêre tegnieke, soos neurale netwerke, gebruik gemaak.

Die verhouding tussen die roumateriale, die verfeienskappe en reologie word geëvalueer deur van die volgende drie modelle gebruik te maak:

- **MODEL 1:** Die verhouding tussen reologie en roumateriale
- **MODEL 2:** Die verhouding tussen verfeienskappe en roumateriale
- **MODEL 3:** Die verhouding tussen verfeienskappe en reologie

MODEL 1 maak van tegnieke soos hoofkomponentanalise en voorlopige modellering gebruik om oortolligheid te verminder en om soveel as moontlik data vas te lê. **MODEL 2** en **3** maak van siftingstegnieke gebruik om sodoende die relevante roumateriale en verfeienskappe te identifiseer.

Die geldigheid van die drie modelle word gekontroleer om te verseker dat voorspellings en interpretasies objektief en doeltreffend is. **MODEL 1** toon aan dat die emulsie, aanvullerpartikels, pigment, water, organiese pigment en oplosmiddel die ses roumateriale is wat die grootste uitwerking op die reologie van die spesifieke verf het. Die drie reologiese krommes wat met die beste akkuraatheid deur middel van meervoudige lineêre regressie voorspel word is: die amplitude swaai, 3-Interval-Tiksotropie-Toets en die vloeikromme. Nie-lineêre reologiese gedrag word by hoë pigment volume konsentrasies en volume vastestowwe aangetref, as gevolg van die sterk

afhanklikheid van die reologie van verf van hierdie eienskappe. Dit is bewys dat neurale netwerke beter resultate as meervoudige lineêre regressie lewer in die voorspelling van die reologiese gedrag van verfmonsters waar die hoeveelhede van rou materiale meer as 20% verskil van die hoeveelhede in die standaardformulasie. Die gebruik van neurale netwerke verbeter die voorspelbaarheid van die reologiese parameters van hierdie monsters met gemiddeld 54%. Die grootste verbetering word gemaak op die reologiese veranderlike “Ekstra Lae Frekwensie” waarde (C_{XLF}). Meervoudige lineêre regressie het resultate met relatiewe foute van 59% gelever, terwyl neurale netwerke slegs 5% foute gelever het. Ander skattings van reologiekrommes, waar neurale netwerke 'n merkwaardige verbetering op voorspelbaarheid tot gevolg gehad het, sluit die tydskromming – 68% verbetering in akkuraatheid - en die lae skuiwingskromme – 63% verbetering in akkuraatheid – in. Die kleinste verbetering wat deur die neurale netwerk bewerkstellig is, is 'n 33% verbetering op die akkuraatheid van die voorspelling van amplitude swaai krommes.

Die meervoudige lineêre regressie modelle van **MODEL 2** voorspel die kritiese verfeienskappe van ondeursigtigheid, glans, Krebs viskositeit en droë laag dikte met relatiewe foute kleiner as 10%. Dit is bewys dat 90% van alle nuwe voorspellings binne die perke van die toelaatbare foutspelingsruimte wat deur die verfvervaardiger daargestel is, val. Die volgende verfeienskappe kan voorspel word met 'n verwagte fout van tussen 10% en 20%: droë en nat glansafwerking, ooptyd en water permeabiliteit. Die twee verfeienskappe wat die mees onakkuraat voorspel deur meervoudige lineêre regressie voorspel is, met foute groter as 20%, is vuilisoptelling en sakking. Nie-lineêre tegnieke soos neurale netwerke word gebruik om om die verfeienskappe van hierdie monsters, waar die hoeveelheid roumateriale met meer as 20% van die standaardformulasie verskil, te voorspel. Die neurale netwerk toon 'n merkwaardige verbetering op die voorspelbaarheid van die verfeienskappe vir hierdie tipe verfmonsters. Voorspelbaarheid word gemiddeld met 47% verbeter. Die grootse verbetering in voorspelbaarheid is dié van nat glansafwerking “20”, waar meervoudige lineêre regressie resultate met relatiewe foute van 66% gelever het, terwyl neurale netwerke foute van slegs 0.6% verkry het. Twee ander verfeienskappe waarvoor neurale netwerke 'n merkwaardige verbetering van 80% of meer op voorspelbaarheid gelever het, is glans – 80% verbetering in akkuraatheid – en droë laag dikte – 81% verbetering in akkuraatheid. Die kleinste verbetering wat neurale netwerke op die akkuraatheid van 'n verfeienskap gehad het, was 'n 33% vermeerdering in die akkuraatheid van die voorspelling van sakking. MODEL 2 maak dit vir die verfvervaardiger moontlik om toleransies rondom sekere verfeienskappe gedurende vervaardiging te toets.

Reologie is tot op hede steeds 'n baie onderbenutte stuk gereedskap wanneer dit by die verduideliking van sekere verfeienskappe kom. **MODEL 3** het die korrelasie tussen die fundamentele reologiese eienskappe en die werklike werksverrigting van 'n verf gekwantifiseer. Dit is aangetoon dat reologiese metings met groot akkuraatheid gebruik kan word om sekere kritiese verfeienskappe, soos ondeursigtigheid, Krebs viskositeit, droë laag dikte en glans te voorspel, binne die toelaatbare foutspelingsruimte wat deur die verfvervaardiger daargestel is. Meervoudige lineêre regressie modelle voorspel ondeursigtigheid, Krebs viskositeit en droë laag

dikte met relatiewe foute kleiner as 10%, wanneer reologiese parameters as insette vir die model gebruik word. 'n Neurale netwerk van **MODEL 3** is ontwikkel om die *verfeienskappe* van daardie verfmonsters waar die inhoud van roumateriale met meer as 20% van die standaardformulasie verskil, te voorspel deur *reologie data* as insette vir die model te gebruik. Die neurale netwerke verskaf meer akkurate resultate as meervoudige lineêre regressie wanneer dit by die voorspelling van die verfeienskappe van hierdie tipe monsters kom. Neurale netwerke wat reologie data as insette gebruik, voorspel gemiddeld verfeienskappe met 49% groter akkuraatheid in vergelyking met die ekwivalente meervoudige lineêre regresse modelle. Die grootste verbetering in model voorspelbaarheid is die voorspelling van water permeabiliteit - 73% vermeerdering in akkuraatheid – en glans – 70% vermeerdering in akkuraatheid.

Twee modelle: **MODEL 2** (roumateriaalkomposisie as invoer) en **MODEL 3** (reologie data as invoer) is ontwikkel vir die voorspelling van verfeienskappe. Die tabel hieronder vergelyk die lineêre modelle op grond van die gemiddelde relatiewe fout wat verwag kan word tydens die voorspelling van nuwe verfeienskappe.

MODEL 2 EN MODEL 3 VERFEIENSKAP VOORSPELLINGS

Verfeienskap	Gemiddelde Relatiewe Fout (%)	
	Model 2	Model 3
Droë glansafwerking (60°)	10.7	20.0
Nat glansafwerking (60°)	14.8	15.3
Vuilisoptelling	43.6	48.0
Ooptyd	17.3	17.4
Droë laag dikte	6.4	7.1
Glans (60°)	7.2	15.3
Wegsteekvermoë	0.8	1.5
Krebs viskositeit	1.6	1.6
Ondeursigtigheid	0.8	1.1
Sakking	43.2	38.9
Water permeabiliteit	14.2	12.5

Akaike se Inligtingskriterium (AIK) is ook gebruik om onderskeid te tref tussen lineêre en nie-lineêre modelle. Die AIK toon oor die algemeen vir al drie modelle aan dat meervoudige lineêre regressie modelle beter resultate as neurale netwerke vind, in terme van die kondensering van inligting, veral van daardie verfmonsters waar die hoeveelheid roumateriale met meer as 20% van die standaardformulasie verskil.

Hierdie studie het groot bydraes tot die verfindustrie gemaak, soos byvoorbeeld 'n aanduiding van die uitvoerbaarheid daarvan om verfeienskappe te voorspel in terme van beide roumateriaaldata en reologiese data. Die volgende stap sal die kombinasie van hierdie twee stelle data wees. Dit

kan moontlik lei na selfs groter akkuraatheid in die voorspelling van verfeienskappe. Koste faktore vir die roumateriale kan ook ingesluit word om byvoorbeeld verfkomposisie vir die laagste koste en aanvaarbare verfeienskappe te optimeer. Hierdie studie toon die uitvoerbaarheid daarvan om verfeienskappe te voorspel in terme van roumateriaaldata en reologiese data vir 'n enkele verffamilie. Die studie kan nou gebruik word as 'n basis om die huidige kennis en interpretasie van verfeienskappe in terme van roumateriale en reologie te verbreed deur 'n verskeidenheid verffamilies daarby in te sluit.

ACKNOWLEDGEMENTS

Professor Hansie Knoetze

Barloworld Plascon for its financial support.

THRIP for its financial support

The staff of Plascon Research Centre - Dr B Cooray, Dr John Engelbrecht, Mr Bertus Smit, Me Yolande Koen, Mr Johan Treurnicht for their assistance and valuable knowledge

Staff and friends at the Department of Process Engineering

My family - Johan, Alba, Marcelle and Jeanné

Carien Fouche

CONTENTS

Abstract	i
Opsomming	iv
Acknowledgements	viii
Contents	ix
1 Introduction	1
1.1 Background	1
1.2 Problem Statement	1
1.3 Challenges of this Work	2
1.4 Thesis Layout	2
1.5 References	4
2 Background and Literature Review	5
2.1 Rheology Fundamentals	5
2.2 Paint Rheology	5
2.2.1 Relevance of Rheology in the Coatings Industry	5
2.2.2 Rheology and Paint Chemistry	6
2.2.3 Rheological Measurements for the Study of Paint Properties	10
2.3 Modelling the Relationship between Paint Rheology and Raw Materials	19
2.3.1 Background	19
2.3.2 Linear Modelling	20
2.3.3 Non-Linear Modelling	22
2.4 Modelling the Relationship between Paint Properties and Raw Materials	24
2.4.1 Background	24
2.4.2 Linear Modelling	26
2.4.3 Non-Linear Modelling	27
2.5 Modelling the Relationship between Paint Properties and Paint Rheology ...	29
2.6 Nomenclature	29
2.7 References	30
3 Experimental Methods and Techniques	34
3.1 Introduction	34
3.2 Design of Experiments (DOE)	34
3.2.1 DOE for the Screening Process	35
3.2.2 DOE for the Modelling Process	38
3.3 Manufacturing of Paint Samples	46

3.3.1 The Raw Materials	46
3.3.2 Sample Preparation	47
3.3.3 Scientific Method	48
3.4 Paint Properties	52
3.5 Rheology	57
3.5.1 Introduction	57
3.5.2 Rheological Measurements	57
3.6 Nomenclature	61
3.7 References	62
4 Data Preparation	65
4.1 Introduction	65
4.2 General Data Preparation	68
4.3 Rheological Data Preparation	70
4.3.1 The Rheological Parameters	70
4.3.2 Preliminary Modelling	72
4.3.3 Redundant Rheological Data	75
4.4 The Screening Process	84
4.4.1 Raw Materials Screening	85
4.4.2 Paint Property Screening	95
4.5 Conclusions	102
4.6 Nomenclature	103
4.7 References	105
5 Linear Modelling	107
5.1 Introduction	107
5.1.1 Multiple Linear Regression (MLR)	108
5.1.2 Multicollinearity and its Effects in MLR	109
5.1.3 Model Validity	110
5.1.4 Model Predictability	113
5.2 Model 1: Relationship between Paint Rheology and Raw Materials	114
5.2.1 Background	114
5.2.2 Multiple Linear Regression (MLR) Modelling	116
5.2.3 Model Validity	119
5.2.4 Rheology Prediction	129
5.2.5 Model 1 Conclusions	139
5.3 Model 2: Relationship between Paint Properties and Raw Materials	140
5.3.1 Background	140
5.3.2 Multiple Linear Regression (MLR)	141
5.3.3 Model Validity	144
5.3.4 Paint Property Predictions	153

5.3.5 Model 2 Conclusions	160
5.4 Model 3: Relationship between Paint Properties and Paint Rheology	160
5.4.1 Background	160
5.4.2 Multiple Linear Regression (MLR)	161
5.4.3 Model Validity	163
5.4.4 Paint Property Predictions	167
5.4.5 Model 3 Conclusions	174
5.5 Model Interpretations	175
5.6 General Conclusions	179
5.7 Nomenclature	180
5.8 References	182
6 Non-Linear Modelling	184
6.1 Introduction	184
6.1.1 The Single Node (Neuron)	186
6.1.2 Developing the Neural Network Architecture	187
6.1.3 Neural Network Training	188
6.1.4 Neural Network Validation	189
6.1.5 Neural Network Prediction	190
6.2 Model 1: Relationship between Paint Rheology and Raw Materials	190
6.2.1 Background	190
6.2.2 Neural Network Modelling	192
6.2.3 Model 1 Conclusions	206
6.3 Model 2: Relationship between Paint Properties and Raw Materials	207
6.3.1 Background	207
6.3.2 Neural Network Modelling	211
6.3.3 Conclusions	218
6.4 Model 3: Relationship between Paint Rheology and Paint Properties	218
6.4.1 Background	218
6.4.2 Neural Network Modelling	220
6.4.3 Model 3 Conclusions	226
6.5 Model Interpretations	228
6.6 Model Discrimination	229
6.7 General Conclusions	231
6.8 Nomenclature	233
6.9 References	236
7 Conclusions and Future Work	237
7.1 Contributions	237
7.2 Conclusions	241
7.2.1 Conclusions Based on Linear Modelling.....	241

7.2.2 Conclusions Based on Non-Linear Modelling	242
7.3 Future Work.....	244
7.4 Nomenclature	245
8 Appendix A: Glossary and Terms	CD
9 Appendix B: Rheology Fundamentals	CD
10 Appendix C: Raw Data	CD
11 Appendix D: Results – Data Preparation	CD
12 Appendix E: Results – Modelling	CD

1 INTRODUCTION

1.1 BACKGROUND

For many, the beauty of science lies in the aesthetical. However, for many more, the beauty of science lies in the opportunities it provides us. In 2005, the president of the *Federation of Chemical Engineering* identified four main objectives for the future of chemical and process engineering. One of these objectives was identified as [1]:

“Product design and engineering with special emphasis on end-use properties of complex fluids and solids.”

Paint is a complex fluid, i.e. it consists out of more than one chemical type. In the past, a single viscosity value (Krebs viscosity) was used to make certain evaluations about the paint. Paint scientists soon identified the shortcomings of this method:

- i. A single viscosity measurement only evaluated the paint under a specific condition.
- ii. The flow behaviour of paint was not only governed by viscous behaviour, but also by elastic behaviour, i.e. paint is viscoelastic.

Recent advances in rheology (the study of flow and deformation of matter) have made it possible for scientists to evaluate the viscoelastic behaviour of paint over a wide range of shear conditions. It became evident that certain paint properties could be explained by certain viscoelastic flow behaviour. Paint scientists now had the tools to investigate the viscoelastic behaviour of paint, but the fundamental correlations between paint properties and paint rheology were still eluding the paint scientists.

1.2 PROBLEM STATEMENT

After it became apparent that rheology proved to be a useful tool in identifying certain paint property behaviour, albeit qualitatively, researchers such as Leskovsek, Tusar and Gasperlin started predicting certain paint properties in terms of the rheological behaviour. Although these researchers made a contribution in this field of study, their work was elementary and lacked the possibilities that are available today.

This study is an attempt in predicting paint properties of a specific paint as accurate as possible by means of rheological data and advanced modelling techniques. This will enable the paint

chemist to link fundamental rheological properties with real-world performance. At the same time an attempt is made to achieve accurate predictions regarding the rheology of the paint based on its raw material composition.

1.3 GOALS OF THIS WORK

- 1) The challenge exists to **make better use of rheological data** to link fundamental rheological properties to real-world performance, which is still seen as a major challenge in the coatings industry. One answer to this dilemma is advanced computer techniques, which utilises the fundamental rheological data as the required input.
- 2) An empirical model needs to be developed in order to **link fundamental rheological properties to real-world performance**. The challenge is to obtain paint property predictions that are as accurate as possible and within a specified range that is set out by the paint manufacturer. This model will make paint property predictions that are based on the rheological data possible.
- 3) An empirical model needs to be developed that links the **paint's composition to real-world performance**. The challenge is to obtain paint property predictions that are as accurate as possible and within a specified range that is set out by the paint manufacturer. This model will use the level of raw materials as the required input.
- 4) An empirical model needs to be developed that **links the paint's composition to fundamental rheological properties**. The challenge is to predict rheology curves that are as accurate as possible by using the level of raw materials as input to the model.

1.4 THESIS LAYOUT

The dissertation is written in such a way that each chapter can be read on its own, without having read the other chapters. Each chapter has a short introduction which highlights the background of the topics covered in the specific chapter, therefore some topics may overlap, although this has been kept to a minimum. Figure 1.1 shows a mind map of relevant topics of each chapter.

Challenge	Topic	Chapter
MAKE BETTER USE OF RHEOLOGICAL DATA	<i>Using principal component analysis and preliminary modelling techniques</i>	4
LINK BETWEEN PAINT COMPOSITION AND PAINT RHEOLOGY	<i>The effect of raw materials on the rheology of a paint family. Identifying the most important raw materials</i>	4
	<i>Linear modelling of rheological properties</i>	5
	<i>Non-linear modelling of the rheological properties of paint</i>	6
LINK BETWEEN PAINT COMPOSITION AND PAINT PROPERTIES	<i>The effect of raw materials on paint properties: Identifying most important raw materials</i>	3,4
	<i>Linear modelling of the relationship between paint properties and raw materials</i>	5
	<i>Non-linear modelling of the relationship between raw materials and paint properties</i>	6
LINK BETWEEN PAINT RHEOLOGY AND PAINT PROPERTIES	<i>Linear/non-linear modelling of the relationship between paint properties and paint rheology</i>	5,6

Figure 1.1 A Mind Map of Topics Covered in the Dissertation

1.5 REFERENCES

1. Carpentier, J. (2005). "Four main objectives for the future of chemical and process engineering mainly concerned by the science and technologies of new materials productions." Chemical Engineering Journal **107**: 3-17.

2 LITERATURE REVIEW

2.1 RHEOLOGY FUNDAMENTALS

Rheology is the science of deformation and flow. All forms of shear behaviour can be viewed as lying between two extremes: the flow of ideal viscous liquids on the one hand and the deformation of ideal elastic solids on the other. The behaviour of all real materials, such as paint, is based on the combination of both the viscous and the elastic portion and therefore, it is classified as viscoelastic materials.

Rheometry is the measuring technology used to determine rheological data and this includes technology such as the measuring systems, instruments and test and analysis methods. Viscoelastic materials can be investigated using rotational and oscillatory rheometers.

The understanding of coating performance in terms of rheology is becoming very important and modern rheological instrumentation and measurements are assisting in achieving this goal. The fundamentals underlying some of the basic rheological principals are given in Appendix B (*Chapter 9*).

2.2 PAINT RHEOLOGY

2.2.1 RELEVANCE OF RHEOLOGY IN THE COATINGS INDUSTRY

To a degree matched by few other areas, rheology determines the success of coatings [1]. A coating will usually not meet with success if the rheology is not adequate, even if all other properties are acceptable. Experienced formulators say that more than half the cost of new product development is consumed in “getting the rheology right”. Moreover, apparently “minor” changes in a raw material or process can cause significant and unexpected variability in product rheology. For all these reasons, rheological analysis is a vital and cost-effective tool for the coatings industry. Researchers identified the development of relationships between paint properties and the relevant rheological test methods as a major gap in the coatings industry [2], [3].

The large-scale move towards environmentally compliant coatings (waterborne, higher solids, reduced or zero volatile organic compounds (VOC's)) has in general resulted in more complex rheology, while reducing the number of formulating options and at the same time generating a host of performance/application problems. For many years, the flow behaviour of water-borne paints was inferior to solvent-borne paints. However, as a result of extensive research in order to comply with the latest European VOC regulations, this is no longer the case and the rheological performance of water-based and solvent-borne paints are comparable. For example, it has been recognised that control over the rheology of the paint is a key element in achieving extended open time and wet edge time [4].

When most scientists hear the word “rheology”, they immediately think of viscosity, and in fact, viscosity is one of the material properties that can be obtained from rheological measurements. However, viscosity is not a discrete measurement and depends on the conditions of measurement, for example the shear rate. This is indeed significant for materials such as paint, which is exposed to a variety of different shear rates (Figure 2.1), and hence may result in different apparent viscosities [5].

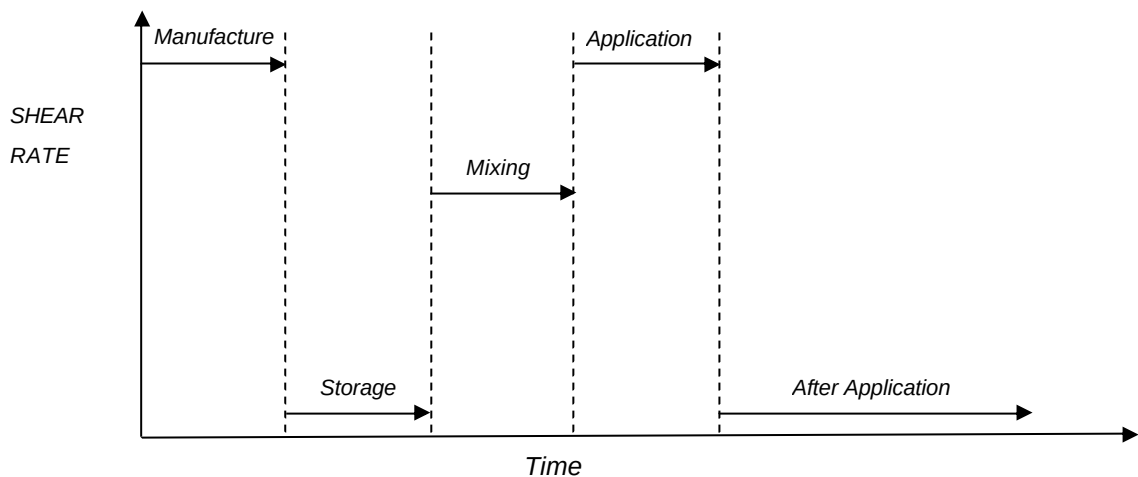


Figure 2.1 Shear Rate Profile over the Lifespan of a Coating

Paints are specifically formulated to achieve a balance in viscoelastic behaviour. It is up to the paint manufacturer to develop a paint so that the viscoelastic properties correspond to the end-use market it is designed for. The reason for this is that there is always a trade-off between some of the properties such as good levelling versus that of good anti-sagging behaviour, which requires opposing rheological behaviour. Therefore, there is no such thing as the “right” rheology profile. Take the rheological parameter $\tan \delta$ as an example. A higher $\tan \delta$ (G''/G') implies less structure, more fluid-like behaviour and may therefore tend to sag more and be more susceptible to pigment sedimentation. On the positive side, a paint with a higher $\tan \delta$ value will most probably level better due to the more fluid-like character. Furthermore, the behaviour of paint changes dramatically with changes in stress and frequency (different time scales) and it is therefore important to perform tests that cover a wide range of stresses and frequencies [5].

2.2.2 RHEOLOGY AND PAINT CHEMISTRY

A typical high performance architectural coating has been chosen for this study. Not only does the physical/chemical structure of the individual components affect the final rheology profile of the paint, but so do the various possible interactions between the paint components. It has become a common task for paint formulators to balance the levels of solvent, surfactant, rheology modifiers, etc. on experience, due to the various possible interactions [6]. However, the interactions among the ingredients of the modern latex paints are so numerous and complex [7] that it almost

becomes impossible to predict the exact response on rheology [8]. Literature sources can only give a faint idea of what to expect in terms of the rheological performance.

2.2.2.1 PIGMENTS AND EXTENDER PARTICLES

For the specific product under investigation, pigments and extender particles increase the volume fraction of solids in the system and induce a more solid-like character, e.g. strong elastic, shear-thinning behaviour. The rheological profiles can become quite complex for a paint with pigment volume concentrations (PVC) of between 21% and 43% [9]. The increase in solid-like character, or decrease in viscous contribution, decreases the open-time and therefore possibly the flow and levelling, but it also improves the viscosity in the high shear-rate region and supports sufficient film forming properties [9].

Interactions with rheology modifiers, dispersants and surfactant are often overlooked (mostly due to extremely complex adsorption/desorption behaviour at solid surfaces [7]), but have a substantial influence on the final coating rheology behaviour [10].

2.2.2.2 SOLVENTS

Solvents that were present in most aqueous paint formulations in the past, played an important role in influencing the rheology of the coating [4]. Currently low VOC decorative coatings are manufactured without solvents such as propylene glycol. When solvent is added to paint, a major drop in low-shear viscosity occurs without affecting the high-shear viscosity dramatically. The explanation for this is that the solvent increases the solvency of the aqueous phase for the hydrophobes and, consequently reduces the tendency for association. This method is often used by paint formulators to achieve the correct balance between high and low-shear viscosities of commercial paints [11]. It is with the addition of solvents that attempts have been made to improve the open time and wet edge time of water-borne dispersion paints. The use of propylene glycol at levels up to 25% by weight has been used to improve open time and wet edge time [4]. It is also known that solvents increase the scrub resistance of the paint and that phase separation may occur in extreme cases, giving rise to poor gloss [7].

2.2.2.3 SURFACTANTS

Surfactants are substances that reduce surface tension. Surfactant molecules act similar way as rheology modifiers by adsorbing onto the emulsion particles. This has a significant effect on the coating's rheology profile [12]. They actually compete for sites along with the rheology modifiers, resulting in a lower level of association of rheology modifier. The higher the surfactant level, the lower the degree of association of the rheology modifier. Excess surfactant will shift the equilibrium strongly to displace any associative thickener adsorbed onto the latex surface with adsorbed surfactant [13]. This results in a considerable drop in low-shear viscosity, whilst the high-shear viscosity remains essentially unchanged. This is in exact agreement to prediction: non-ionic surfactants will form more micelles of physical cross linking although the strength of each individual cross linked micelle will decrease somewhat [14]. The degree of viscosity drop depends on both the amount and type of surfactant. In aqueous paint formulations, surfactant levels as low as 0.1% – 0.2% can have a considerable effect on low shear viscosity [11]. Effects of surfactants

on the low-shear viscosity influences the gloss, hiding power and flow and levelling behaviour considerably [14], [11]. Inclusion of too much surfactant can result in foaming during manufacturing and application, and can also affect the water sensitivity of the dried film. Additionally, recent trends are towards lower and lower solvent levels.

2.2.2.4 DISPERSANTS

Dispersants are added to the coating formulation for stabilisation effects. Dispersants are amphiphiles: surfactant molecules of alternating hydrophobic and hydrophilic moieties. They stabilise the pigment and extender particles by adsorbing onto their surfaces. At the same time they interact with the binder dispersion and the rheology modifiers [9]. Therefore a careful selection is important for utilising the potential of associative thickeners [10]. The presence or absence of interaction between associative thickeners and pigment must be controlled carefully by correct selection of the dispersing agent [10]. Without the thickener interacting properly with the pigments, phase separation occurs on a micro scale [8] and flocculation will develop resulting in a loss of flow and gloss. If droplets containing separate phases of water and thickener merge and rise to the surface, syneresis can result.

Practical paint experience has shown that the selection and level of the dispersant play a major role in the rheology of pigmented systems containing associative thickeners [14].

2.2.2.5 EMULSIONS

Emulsions incorporate essential rheological properties in paint [9] and there is little doubt that the major viscosity contribution in a coating formulation comes from the volume fraction of the latex [15]. The interaction between associative thickener and latex is just as important in affecting the rheology profile of the paint [16]. The latex particle size and the efficiency of the associative rheology modifier go hand in hand. As the surface area of the latex increases, the efficiency of the thickener increases. It often happens that associative thickeners contribute too much to the mid-range viscosity when the latex particle size is small [16]. High performance emulsions tend to have a smaller particle size and may have the tendency to flocculate, limiting the gloss development in a paint [11].

2.2.2.6 COALESCENTS

Coalescents enable film formation of polymer dispersions. These are strongly involved in almost every associative process active in the system [9] and have a significant to large effect on the rheological behaviour during drying and the levelling and sagging properties [17]. In general, addition of coalescing agents to paints leads to higher viscosities. And it is suggested by some [17] that a coalescent can be used as a type of rheology modifier, in addition to thickeners, because of its significant effect on levelling and sagging.

2.2.2.7 DEFOAMERS

Defoamers increase the viscosity slightly, possibly as a result of solubilisation in the micelle, resulting in a higher volume and more interactions between micelles [14].

2.2.2.8 RHEOLOGY MODIFIERS

Rheology modifiers (sometimes referred to as associative thickeners) are polymers which have been hydrophobically modified. Essentially, apart from cellulose and ASE thickeners, there exists two types of rheology modifiers which act by both thickening the water phase and by associating with other species (pigments, dispersant, emulsions and surfactants) present in the paint [11], [16]:

HASE: Hydrophobically modified alkali-soluble emulsion

HEUR: Hydrophobically modified ethylene-oxide urethanes

Rheology modifiers affect only a part of the rheology profile of a paint and are known for affecting the following paint properties [11]:

- Flow
- Structure
- Sag resistance
- Spatter resistance
- Film build
- Brush drag
- Brush loading
- Open time

In the literature [4], the main route to achieve control over the rheology of water-borne systems is the use of associative thickeners. These paints show thixotropic behaviour, which is desirable because it combines excellent flow properties with little to no sagging and good pigment stability. There are three main rheological criteria for evaluating thickener efficiency [4]:

- 1) A thickener must have low-shear viscosities which give an aesthetically pleasing in-can image to the coating, non-dripping properties, good flow and levelling and minimum sagging tendencies.
- 2) An efficient thickener must also provide adequate coating viscosity at the application shear rate so that the required film thickness is obtained while also minimising brush or roller drag.
- 3) The extensional viscosity (elasticity) must be low so that spatter, which is caused by elastic strings, will be absent.

The HEUR type (hydrophobically modified ethylene urethane oxide) is used in this specific formulation. Performance of HEUR type rheology modifiers can be fine-tuned in the developmental stage by varying the size of the hydrophobe. Molecular mass in the order of 40 000 – 50 000 atomic units are typical. They can be thought of as “double ended surfactant molecules” which is a good analogy as they tend to be surfactant-like in character in some respects. The two ends of the molecule are very hydrophobic in nature, whilst the centre is quite hydrophilic. The hydrophobic groups associate with the emulsion particles through adsorption onto the particle surface. In addition they aggregate together in the water phase in a similar way to that in which surfactants form micelles. The network formation results in a major increase in viscosity. The network gives a uniform, stable dispersion of both pigment and emulsion particles, which give greater gloss potential than is possible with flocculated systems. It is obvious that the

surfactant properties of the rheology modifiers are very important for thickening in aqueous systems.

Although the competition between rheology modifiers and surfactants for the adsorption onto the pigment particles can clearly be seen at the low-shear viscosity behaviour, the high shear viscosity is unaffected and is assumed to be dependent on the concentration of the hydrophobe micelles of the rheology modifiers in the water phase and not the association with other particles [11]. Thus it is possible to control high-shear and low-shear viscosity independently with the combination of surfactants and rheology modifiers. More recently developed urethane rheology modifiers have been designed to have slightly less tendency to associate compared with earlier grades and, therefore, these are able to produce paints with very Newtonian* rheology profiles, without the need to add surfactant or solvent [14] and therefore it is often necessary to add a low level of a more pseudoplastic grade of rheology modifier to avoid the tendency to sag after application. By careful blending of a Newtonian and more pseudoplastic grade, it is possible to achieve the ideal rheology profile for many different applications [16], [18], [14].

Gloss levels, hiding power and flow and levelling properties of a very high order are achievable with urethane rheology modifiers [9]. HEUR thickener types tend to make paints slightly more water sensitive and this results in a slight loss of adhesion and greater tendency to early blistering under wet conditions. However, HEUR thickeners demonstrate better scrub resistance and are also resistant to microbial attack. It has also been proven, with rheological measurements, that open time of a water-borne paint improves at increasing concentration of HEUR thickener but at the same time may cause water sensitivity problems [19]. Addition of HEUR thickeners results in higher viscosities at high shear rates which leads to good film build per application with little or no spatter [4]. Also, their characteristically lower viscosities at low shear rates are desirable for improved flow and levelling and gloss [4].

With an understanding of the associative interactions that occur, a high degree of rheology control is possible over all parts of the shear rate spectrum. This enables different paint application properties to be optimised, largely independently of one another, with a high degree of control.

2.2.2.9 BACTERIACIDES AND FUNGICIDES

These have almost no influence on the rheology profile of the paint [14].

2.2.3 RHEOLOGICAL MEASUREMENTS FOR THE STUDY OF PAINT PROPERTIES

Currently there exist no standards for rheological evaluation of paints (not including viscosity measurements such as ICI viscosity or Krebs-Stormer viscosity). Rheological test methods do exist to investigate the rheological behaviour of the paint, many of which are only qualitative in nature, and these are far from standardised. These tests can be carried out in various ways and it is up to the rheologist performing these tests to extract as much data as possible from the rheology profiles. Very little work has been done in quantifying the rheological behaviour and

* See Chapter 9 (*Appendix B: Rheology Fundamentals*) for an explanation on this term

relating it to specific paint properties. A brief description of what is known about the relationship between paint properties and specific rheological test methods follows.

2.2.3.1 VISCOSITY CURVE (CSR[†] TEST)

This test gives information about paint behaviour over a wide range of shear rates (or shear stresses). A typical viscosity curve is illustrated in Figure 2.2.

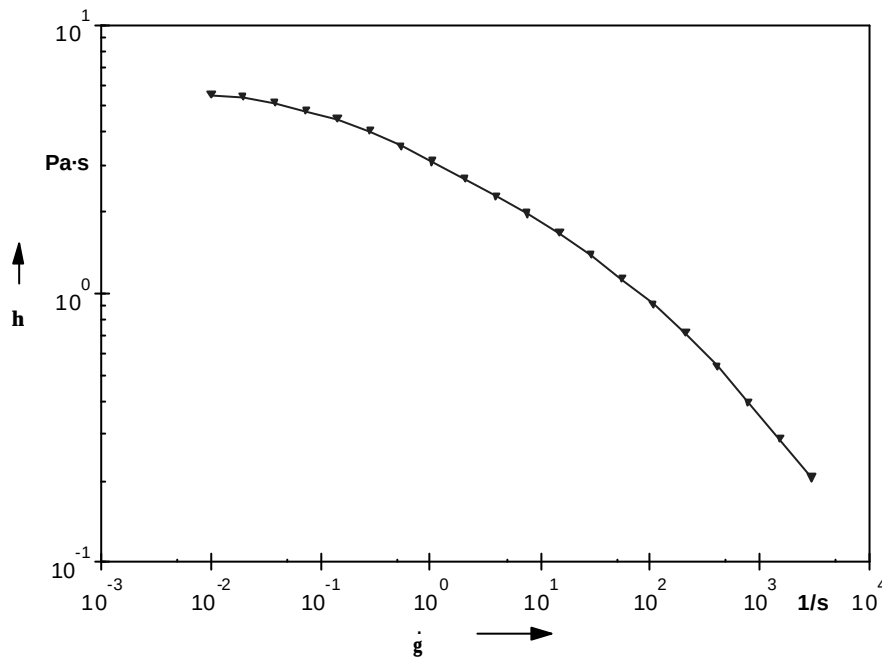


Figure 2.2 A Typical Viscosity Curve

Typical pseudoplastic behaviour is shown which is common for dispersion paints [4]. The following paint properties are evaluated by means of the viscosity curve:

- **Ease of application:** This term describes paint properties such as brush and roller drag. The shear rates associated with the common modes of application such as brushing and rolling are 10^3 s^{-1} to 10^6 s^{-1} . These shear rates might seem high, but are primarily the consequences of the thin films involved. For easy application, paints should show a decrease in viscosity with increasing shear rate (shear thinning behaviour) and the viscosity must be low for easy application at high shear rates [9]. A good viscosity for easy brush/roller application is in the region of 0.1 to 0.25 pascal seconds (Pa.s) [5], [19], [4]. Further insight into the interactive forces that are required to initiate flow during application can be obtained from the “apparent yield point” (or lack thereof) in the low shear rate region with shear rates $< 1 \text{ s}^{-1}$.
- **Pumping and dispensing:** Paint should be easy to pump. Therefore, viscosity values in the medium shear rate range (1 s^{-1} to 10^2 s^{-1}) should not be too high [5].

[†] CSR – Controlled shear rate: these type of tests are shear rate controlled under rotational movement

- **Sagging:** Sagging of paint is dependent on the viscosity of the paint at low shear rates [11]. Some sources state that at shear rates of 0.3 s^{-1} the coating will sag excessively if the viscosity is below 10 Pa.s [20].
- **Pigment settling and instability:** An idea of the degree of settling can also be estimated from the low-shear viscosity [11]. Water-borne dispersion coatings need to have high viscosities in the low shear region [9] to avoid settling and instability during storage and transport. There are other, more appropriate rheological ways, to investigate settling (see *Section 2.2.3.5*).
- **Levelling:** Levelling is evaluated by looking at the viscosity values at low shear rates and good levelling is typically obtained when the low-shear viscosity is in the range of 25 Pa.s to 100 Pa.s [11]. If the low-shear viscosity displays Newtonian behaviour and is low enough, flow and levelling allows the formation of film surfaces free of brush marks [9].
- **Brush loading and roller transfer:** This is the amount of paint that is “picked” up by the brush/roller and is affected by the medium shear rate (1 s^{-1} to 10^2 s^{-1}) viscosity [11].
- **Film thickness:** Viscosity at higher shear rates ($\dot{\gamma} > 1000 \text{ s}^{-1}$) affect performance under brush/roller, thus affecting the film build of the paint [11] and relative high viscosities in the high-shear area give sufficient film thickness of the coating [9].
- **Hiding power:** Definite correlations between the viscosity at high shear rates (10^4 s^{-1}) and the hiding power have been identified [15].

Note that paint properties, such as sagging and levelling, are investigated more accurately by simulating breakdown of the structure followed by a period of rest. The flow curve does not allow for structural regeneration and therefore, tests that simulate structural decomposition and regeneration are better in investigating the paint behaviour after application (see *Section 2.2.3.3* and *Section 2.2.3.7* for “*Structural Decomposition and Regeneration*” curves).

2.2.3.2 VISCOSITY / TIME CURVE (CSR TEST)

This test gives information about the behaviour of paint when subjected to a constant shear rate. A typical viscosity / time curve is illustrated in Figure 2.3.

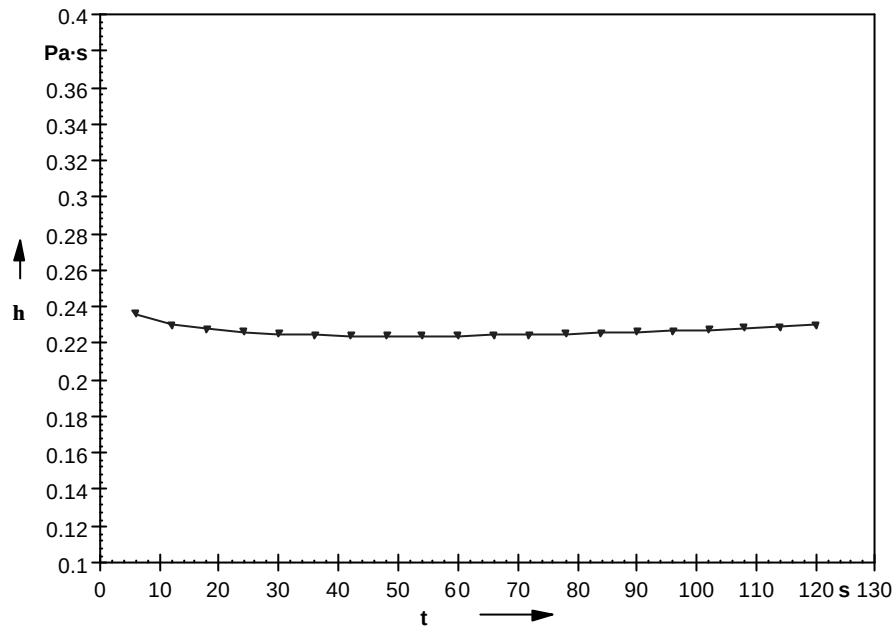


Figure 2.3 A Typical Viscosity/Time Curve at a constant shear rate

Viscosity/time curves can be performed at different shear rates in order to evaluate the following paint properties:

Pumping, dispensing and mixing: These actions occur at medium to high shear rates. The viscosity at these shear rates should not be too high to ensure that these actions are performed with ease [5].

Levelling: Levelling occurs at low shear rates and the viscosity should stay low enough and not show a steep increase over time [1].

Gloss: This paint property is dependent on the low shear viscosity. Gloss deficiency is caused by unfavourable rheology (too high viscosity at low shear rates) which leads to poor levelling [21].

Sagging: Sagging occurs at low shear rates. The viscosity should stay low enough for flow and levelling to occur, however, after application there should not be any further decrease in viscosity over time [1], but rather show an increase in viscosity so that excessive sagging is prevented.

2.2.3.3 STRUCTURAL DECOMPOSITION AND REGENERATION (CSR TEST)

This test gives information about the behaviour of the paint when subjected to small shear rates directly after a period of high shear. A typical structural decomposition and regeneration (thixotropy) curve is illustrated in Figure 2.4.

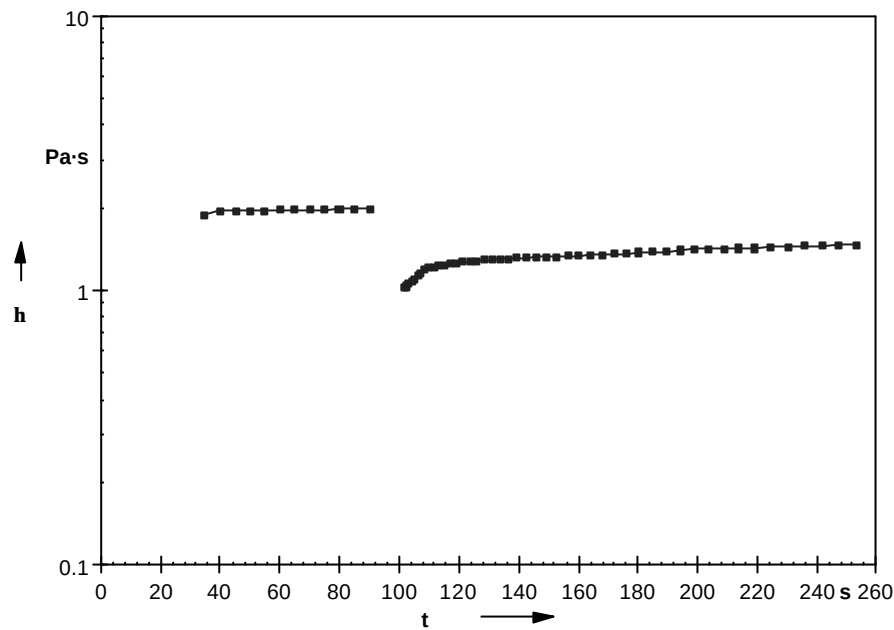


Figure 2.4 A Typical Structural Decomposition and Regeneration Curve (Rotational) at a Constant Low Shear Rate (e.g. 1 s^{-1}) for the 1st and 3rd Phases

The following paint properties are evaluated by the regeneration curve:

Levelling: Levelling of brush marks depends mainly on the levelling stress (t_L). Levelling ceases when t_L becomes equal to the yield stress (t_y) [20]. Slower structure formation after high shear results in better flow and levelling [9].

Sagging: It should be noted that in order to prevent sag, the viscosity must increase rapidly after application of the paint, but slowly enough for levelling to occur. Typically, 30 to 300 seconds must be allowed for proper levelling to occur [20]. The time dependent viscosity build-up is a result of thixotropy.

Film build: Film thickness depends on the degree and rate of structural recovery (the rate at which the viscosity increases after application) [22]. In general, one can expect that the faster the rate of viscosity-increase is directly after application, the greater the film thickness would be.

Gloss: Gloss and levelling go hand-in-hand and unfavourable rheology after application, leads to poor levelling and gloss deficiencies [21].

Storage stability: Thixotropic behaviour supports storage stability [9].

The opposing rheological behaviour in levelling versus sagging is typical for optimal paint properties. After film application, the coating viscosity must remain low enough for levelling to occur, but then the viscosity should increase steadily to prevent sagging [23].

2.2.3.4 AMPLITUDE SWEEP (CSD[‡] TEST)

This test gives information about the behaviour of paint when subjected to varying degrees of shear deformation. A typical amplitude sweep is illustrated in Figure 2.5.

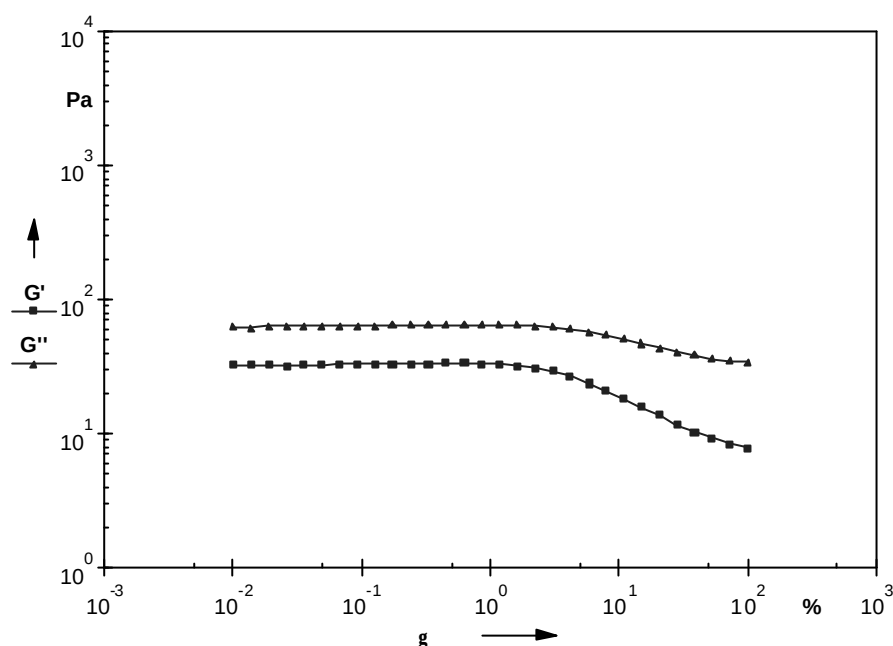


Figure 2.5 A Typical Amplitude Sweep

The following paint properties are evaluated by the amplitude sweep:

Pigment settling and storage stability: Although the amplitude sweep does not reflect conditions at rest, it can still give good insight about the structure of the sample and the type of interparticular interactions that might be present. A certain yield stress (t_y) needs to be overcome before flow occurs. Yield stresses prevent settling and storage instabilities [9].

Flow and levelling: The existence of a yield point affects the flow and levelling behaviour, however, it does not necessarily lead to imperfect levelling [22].

Sagging: Although the amplitude sweep does not give 100% representative data about the structure of the sample at rest (because the amplitude sweep is not necessarily performed at such low frequencies that represent rest), it still gives information about the structure of the sample in general [1]. And in general, if $G'' \gg G'$ in the LVER then the paint is more likely to sag, because the sample is predominantly liquid-like in character.

2.2.3.5 FREQUENCY SWEEP (CSD TEST)

This test gives information about the behaviour of paint when subjected to different time scale deformations. A typical frequency sweep is illustrated in Figure 2.6.

[‡] CSD - Controlled shear deformation: these type of tests are shear deformation controlled under oscillatory movement

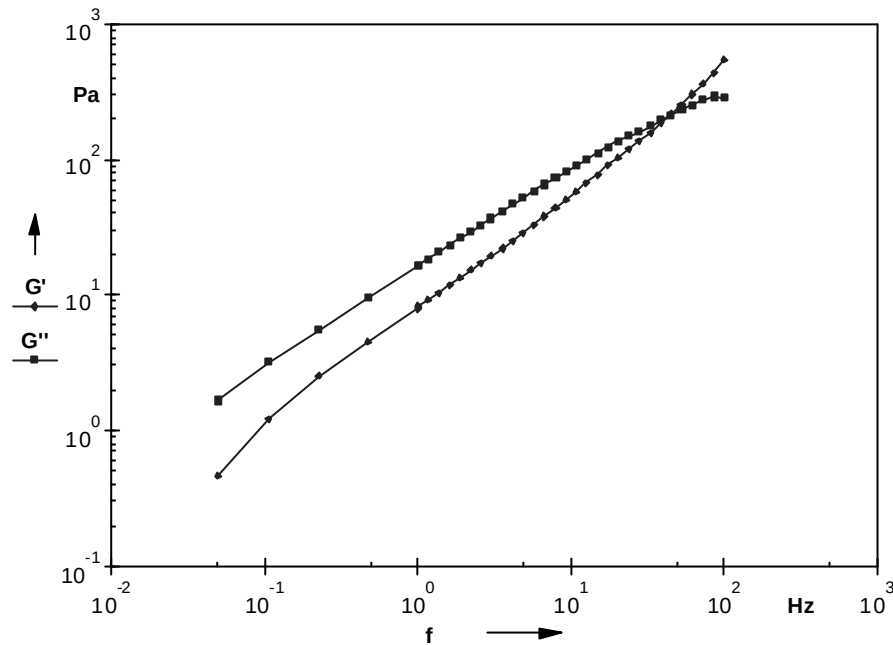


Figure 2.6 A Typical Frequency Sweep

The following paint properties are evaluated by the frequency sweep:

Spatter: Coatings with high elongation viscosity (elastic behaviour which is characterised with the rheological parameter G') form slender strings or thin fibres, which lengthen during application and eventually break. After breaking, the fibre snaps back onto the roller and the substrate. The release of energy stored in the elongated fibre results in the formation of small air-borne paint droplets, which produce spatter. Paint application occurs at fast rates (or short time scales) and the spatter behaviour can therefore be investigated at the high-frequency region of the frequency sweep. The frequency sweep makes it ideal to investigate the spatter behaviour because data about the elastic portion of the paint is obtained accurately. Spatter resistance is predicted by measuring the elastic modulus G' at high frequencies [4]. If $\tan \delta$ (G''/G') is very small in the high frequency region, then the paint behaves predominantly elastic and will tend to spatter [20].

Stability: Data from the frequency sweep at a deformation below the critical strain (g_{LVER}) is used to elaborate on the material's structure. $\tan \delta$, the ratio of energy dissipation and storage mechanism (G'/G''), at low frequencies, is a key parameter for investigating storage stability [20]. It has been found that the optimum storage stability of a paint is obtained for $1 < \tan \delta < 1.5$. [20]. Too high values of $\tan \delta$ indicate predominantly viscous flow – the particles are only weakly associated and the inter-particle forces cannot prevent sedimentation under gravity. Strong inter-particle interaction however decreases the $\tan \delta$ value (higher elasticity). The result of too low $\tan \delta$ values is coalescence and the formation of large aggregates, which settle with time.

Ease of application: Paint application occurs at fast rates (or short time scales) and the ease of application can therefore be investigated at the high-frequency region of the frequency sweep.

Application should be easy without the occurrence of slip (sliding of roller over surface without roller actually rolling) and therefore elastic behaviour should not dominate at high frequencies if the paint is to be applied with ease. In rheological terms, the parameter $\tan \delta$ should not be $\ll 1$ [1].

Flow and levelling: Flow and levelling can be evaluated by measuring the complex viscosity, (h^*) in the low frequency region. The complex viscosity is determined from the storage (G') and loss (G'') moduli. Paints with lower (h^*) have better flow and levelling properties. The complex viscosity at frequency ω is defined by [23] as:

$$h^* = \sqrt{\left(\frac{G'}{\omega}\right)^2 + \left(\frac{G''}{\omega}\right)^2}. \quad (2.1)$$

And therefore G' and G'' should be small. The behaviour of $\tan \delta$ in the low-frequency region also plays an important role in evaluating the flow and levelling behaviour. A paint levels better when it is possible to flow under conditions of rest and therefore $\tan \delta$ should be > 1 for the liquid-like character in the paint to dominate over the elastic part [1].

2.2.3.6 TIME SWEEP (CSD TEST)

This test gives information about the behaviour of paint when subjected to a single time scale deformation to see how the viscoelastic properties change over large periods of time. A typical time sweep is illustrated in Figure 2.7.

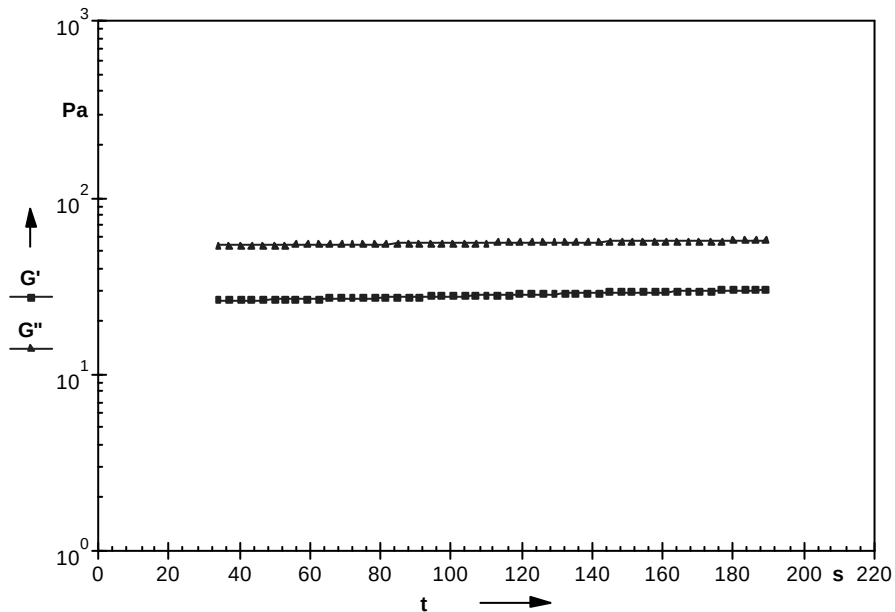


Figure 2.7 A Typical Time Sweep

The literature is not clear on what application information can be obtained from the time sweep. It is clear though that structural changes over large time-periods can be investigated with this method. Therefore, it should play a role in evaluating properties such as stability and phase separation.

2.2.3.7 STRUCTURAL DECOMPOSITION AND REGENERATION (CSD TEST)

This test gives information about the behaviour of the paint when subjected to a large degree of deformation followed by a small degree of deformation, simulating the application of paint followed by a rest phase. This test is performed in the oscillatory mode and therefore the elastic (G') and viscous (G'') portions of the paint can be evaluated. These parameters are very useful in investigating the behaviour of the paint directly after application.

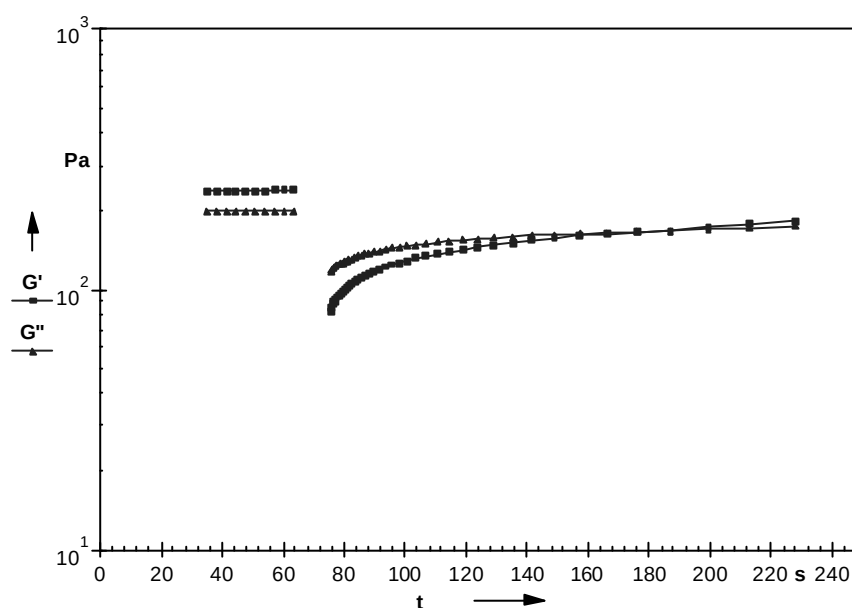


Figure 2.8 A Typical Structural Decomposition and Regeneration Curve (Oscillatory)

The ratio of the complex moduli G' and G'' is a good measure of the material's structure. The crossover point (if it exists) between G' and G'' (where $\tan \delta = 1$) characterises the transition between fluid (viscous) and solid (elastic) behaviour. The response of $\tan \delta$ in the 3rd interval (see Appendix B, Figure 9.20), after structural deformation, together with the occurrence of a cross-over point can be used to investigate the following properties:

Flow and levelling: When $G'' > G'$, the viscous contributions dominate and the paint is movable enough to allow good flow and levelling [9]. However if $G' > G''$, then the elastic properties are contributing the major part. These paints behave elastic due to high interparticular interactions that usually result in a yield point. A certain stress is needed for the yield point to be overcome before the paint starts to flow and may result in brush marks and poor levelling [9].

Pigment settling and storage stability: Although a too high yield point may cause brush marks due to poor flow and levelling, it is also this yield point that prevents pigments from settling and storage instabilities.

Sagging: A coating that sags is one that requires a long time for structural regeneration to occur after application [24]. As long as $G'' > G'$ the sample has a liquid-like character and the paint sags. However, if the curve of G' increases more steeply than the curve of G'' so that at some point during the third interval $G' = G''$ and then $G' > G''$, then the coating changes from liquid-like character to solid like character and sagging stops [24].

Film thickness: The elasticity after application has an influence on the film thickness [20] and the faster the structure recovers, the larger the film thickness will be [22].

Gloss: Unfavourable rheology after application leads to poor levelling with gloss deficiencies at the surface of the coating [21].

It can be seen from the rheological measurements and the information obtained from it, that it is important to consider the whole range over which the measurement is performed. This study combines all of the above test methods over their entire measurement ranges, to evaluate the paint properties. It is this thoroughness in rheological measurement for explanatory purposes, that is applied for the first time in such a study.

2.3 MODELLING THE RELATIONSHIP BETWEEN PAINT RHEOLOGY AND RAW MATERIALS

2.3.1 BACKGROUND

Most raw materials contribute in affecting the flow behaviour of the paint. The effect of specific raw materials on the rheology of the paint is *qualitatively* well established and it has become a common task for paint formulators to balance the levels of raw materials until a desired Krebs^s viscosity is obtained, based on experience [6]. However, the result only gives an optimal point in the formulation and not a model of the system [25]. Rheology control is important for many of the paint properties (levelling, spatter, sag, etc.) and therefore the effect that raw materials have on the rheology is very important. The effect of specific raw materials on the rheology has been studied extensively in practice, usually originating from a quality control perspective or identifying problems in paint properties caused by ill-conditioned rheological behaviour. It is therefore evident that the relationship between rheology and raw materials (MODEL 1) is only relevant as far as the relation between paint properties and rheology is concerned (MODEL 3).

The scope of the research is extended to the region between the lowest and highest values of any independent variable (raw material). Modelling techniques that have been used before in the coatings industry to model rheological behaviour include:

- response surface methodology [26]
- neurofuzzy techniques [27]

^s Krebs viscosity is a term frequently used in quality control of coatings. It is an international term for expressing the medium shear viscosity of a paint at a average rotational speed of 200 rpm and is often used as a specification for quality control.

- polynomial and neural network modelling [28], [29].

These techniques have almost exclusively been used for the modelling of *viscosity* values only [27], [30], [26]. This is largely due to the fact that viscosity is undoubtedly one of the most important properties in quality control in the coatings industry [21], [38]. However, modelling the *rheology* of paint, which entails far more than mere viscosity values has been an area of relatively little research and has mainly been studied by Leskovsek et al [31], [28], [29]. Leskovsek et al identified some of the advantages of empirical rheological modelling with linear and non-linear techniques:

1. With linear equations the main effects of factors on properties can be determined.
2. Explicit knowledge of constitutive equations and/or chemical structure of components is not required.
3. The same empirical model is suitable for predicting rheological properties of Newtonian and non-Newtonian fluids.
4. Neural networks can simultaneously simulate several properties.
5. The same neural network can be used for new observations.

However, although Leskovsek et al already identified certain advantages in the use of linear/non-linear techniques in the prediction of rheological properties, the following weaknesses in the work of Leskovsek et al are noted:

1. Rheological modelling is performed by using only data obtained from a frequency sweep. Although the frequency sweep gives a lot of information about the paint properties, other rheological curves (parameters) also contribute in explaining the behaviour. Other rheology curves give extra insight and are important in contributing to more accurate model predictions.
2. Modelling of the rheological response of the frequency is done at distinct frequencies (0.167 Hz, 1.00 Hz and 10.00 Hz) and not as a frequency sweep of a whole curve. The importance of evaluating the frequency sweep curve at higher frequencies and using the whole curve of the rheological measurements (preliminary modelling) is illustrated in Section 2.2.3.5

Initially, Gasperlin et al [32] also only used the frequency response at a specific frequency value (5.32 Hz) but understood the importance of using the whole curve instead and so introduced the technique of preliminary modelling in order to use the whole measured frequency sweep curve as a response to the model [33]. However, this work is related to the pharmaceutical industry and has not been applied in the coatings industry. Furthermore, Gasperlin et al [33] only used the frequency sweep.

2.3.2 LINEAR MODELLING

Multiple linear regression (MLR) is, in general, one of the most widely used modelling techniques. Other researchers [29] in the coatings industry have found it useful to express the rheological response (y_i) in terms of the raw materials (x_k) by means of the following equation:

$$y_i = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + e \quad (2.2)$$

This is a useful expression because the relative importance of each raw material on the rheological response is given by the regression coefficients (b_k) in equation 2.2. Leskovsek et al [29] illustrated the effect of pigment, binder, extender, matting agent and rheological additive on G' at 10 Hz of the frequency sweep, by plotting the bar graph of the regression coefficients (Figure 2.9).

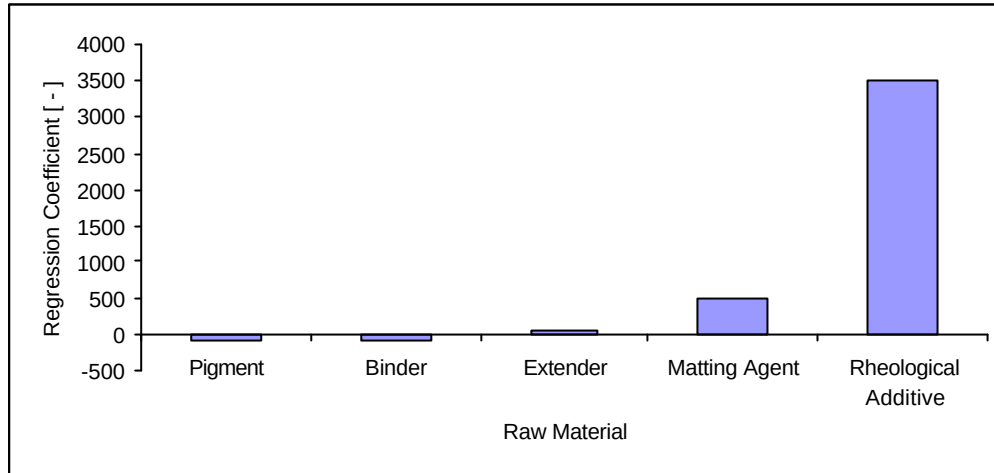


Figure 2.9 Values of Regression Coefficients for G' at 10 Hz from MLR [29]

The accuracy of prediction was determined with statistical validation and it was found that, according to statistical evaluation, the linear equation of the type shown in equation 2.2 is not complex enough to describe the selected rheological responses. It was established that interactions between the different raw materials do play a role and that models that take interactions into account (higher order polynomials and neural network models) are more suitable for these type of modelling problems.

Gasperlín et al [33], [34], who originally worked with Leskovsek et al, extended the rheological modelling techniques used in the coatings industry to that of the pharmaceutical industry. The observations made by Gasperlín et al in the pharmaceutical industry are closely related to that of the coatings industry. The influence of different ratios of individual components on the viscoelastic behaviour (illustrated as $\tan \delta$ curves of the frequency sweep) was studied at 5.32 Hz [32]. As mentioned earlier, evaluating the whole rheological curve over the entire range, and not just at one point (e.g. 5.32 Hz), is important to extract as much information as possible. Therefore, preliminary modelling techniques were used successfully to represent the complex viscosity (h^*) over the entire frequency sweep. It was also concluded that non-linear modelling techniques are better in prediction of rheological properties than linear models.

A major advantage of modelling the effect of raw materials with a linear model is that the main effects of the raw materials on the rheology can be determined. At the same time, this also has the disadvantage that it does not give any information about the interactions occurring between the raw materials.

2.3.3 NON-LINEAR MODELLING

Leskovsek et al [29], [28], [31], [35] used neural networks as non-linear modelling techniques to model some of the rheological properties of paint. The amount of pigment, binder, extender, matting agent and rheological additive were taken into account and used as input variables. The rheological responses included:

1. Infinite viscosity, h_{∞} (flow curve)
2. Yield point, t_{yp} (flow curve)
3. Storage modulus, G' at 0.17 Hz (frequency sweep)
4. Storage modulus, G' at 1.0 Hz (frequency sweep)
5. Storage modulus, G' at 10 Hz (frequency sweep)
6. Loss modulus, G'' at 0.17 Hz (frequency sweep)
7. Loss modulus, G'' at 1.0 Hz (frequency sweep)
8. Loss modulus, G'' at 10 Hz (frequency sweep)
9. Sagging (determined by measuring the film thickness at which the paint exhibits sagging from vertical objects).

Note from above rheological responses that only discrete points of the flow curve and frequency sweep were used as rheological responses.

A back propagation feed forward** neural network was used for training of the neural network. A two-layer neural network with five nodes in the input layer (corresponding to the five raw materials), six nodes in the hidden layer and nine nodes in the output layer (corresponding to the nine rheological properties) was used for modelling of the rheological properties. The architecture of this neural network is represented in Figure 2.10.

** In feed forward neural networks, information is propagated back through the net during the learning process, in order to update the weights. As a result, these neural networks are also known as back propagation feed forward neural networks.

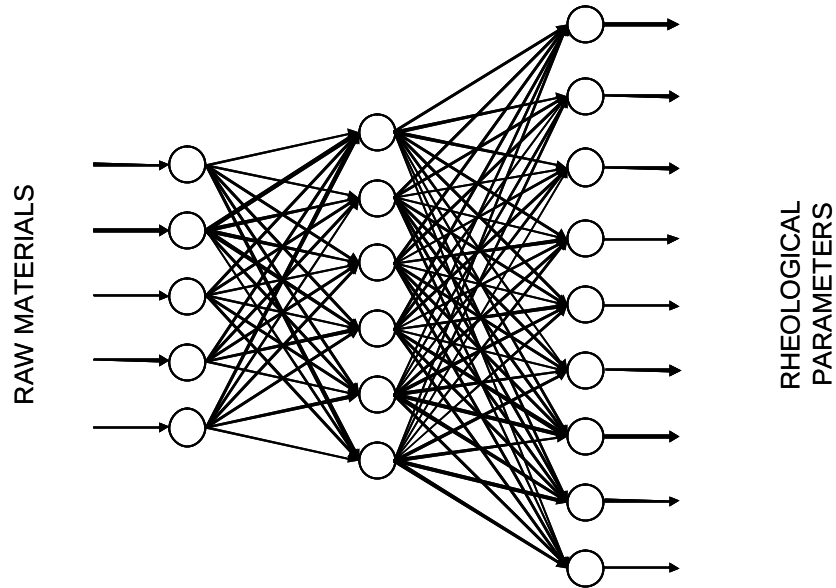


Figure 2.10 Neural Network Architecture used by Leskovsek et al

The neural network approach for modelling of the rheological properties allows the investigation of each response separately or all of them together. It was found that the matting agent and the rheological additive affect G' at 10 Hz the most. Differences between polynomial and neural network models were also observed.

Note that the work published by Leskovsek et al [31] included empirical models for predicting the relationship between paint composition and rheological behaviour only at *selected* frequencies (0.17 Hz, 1.0 Hz and 10 Hz) of the frequency sweep. Gasperlin et al [34] also investigated the dependence of a rheological parameter, the complex viscosity (η^*), on the composition of the system. However, the method of preliminary modelling (see Section 4.3.2 – ‘Preliminary modelling’) of rheological curves was introduced in this study. In other words, the entire range of the frequency was taken into account for modelling purposes and not only the complex viscosity at selected frequencies. Therefore, the complex viscosity (η^*) of the frequency sweep is represented by the following equation:

$$\ln \eta_i^* = a_i \ln f + b_i \quad (2.3)$$

Where a represents the slope of the preliminary model of the complex viscosity and b represents the offset of the preliminary model.

A two-layer neural network (one input layer and one hidden layer) was employed with the error back propagation supervised learning algorithm. The learning term was 0.5 and the momentum term 0.9. The contents of the individual emulsion components represent the input factors while the responses are respectively, values of parameters a_i and b_i of equation 2.3. Two neural networks were developed: one for parameters of a_i and one for parameters of b_i . The architecture for this network is illustrated in Figure 2.11.

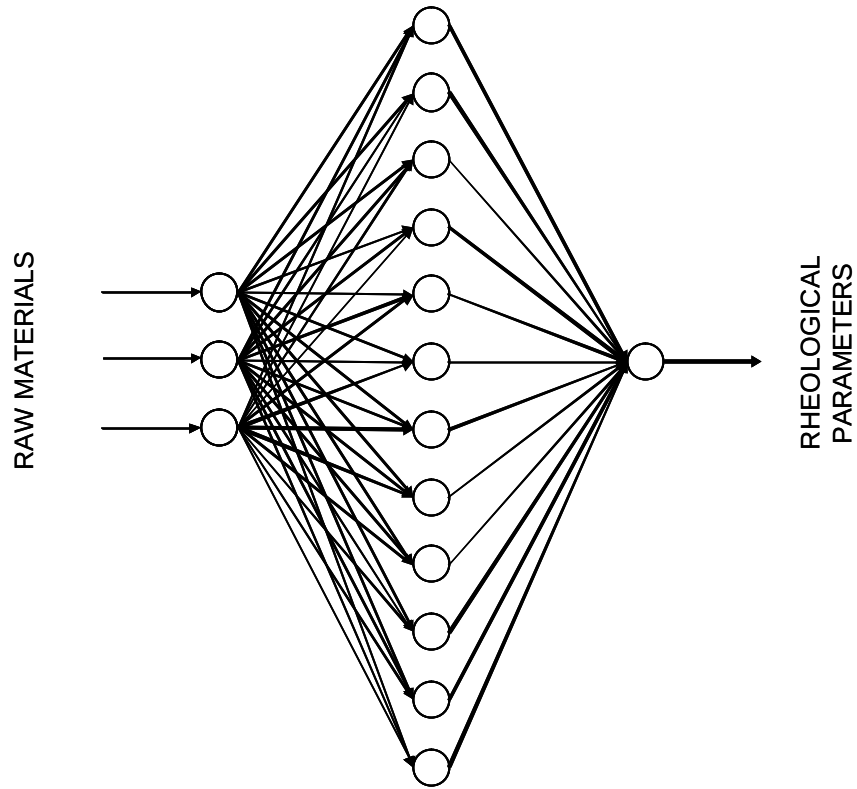


Figure 2.11 Neural Network Architecture used by Gasperlin et al

The performance of the neural network was compared to that of other modelling techniques including linear and higher order polynomials. It was found that the most accurate values of the parameters a_i and b_i are calculated with the two layer neural network.

2.4 MODELLING THE RELATIONSHIP BETWEEN PAINT PROPERTIES AND RAW MATERIALS

2.4.1 BACKGROUND

The general relationship between raw materials and the paint properties is well known and also documented [36]. Table 2.1 shows the major effect a paint raw material group has in a latex paint.

Table 2.1 The Major Effects of Raw Materials on the Paint Properties of a Latex Paint [37]

Raw Material Group	Major effects
Emulsion (binder)	Gives Durability, gloss
Pigments	Gives Hiding, colour, opacity, hardness, durability, corrosion resistance
Extenders	Gives Hiding power
Vehicle (water)	During Application
Solvents	Enhances Open time, adhesion, water resistance
Coalescents	Provides Film formation, defoamer
Dispersants	Ensures Stability (settlement, syneresis), colour acceptance
Surfactants	Gives Colour acceptance, stability, film formation
Defoamers	Prevents foaming
Thickeners	During Application
Biocides	Acts as Preservative (in-can and film)

In general, the raw materials in Table 2.1 contribute mainly to the specific paint property it was designed for. Numerous different products are available in each of the raw material groups all add to the specific paint property it is designed for. However, the contribution of each individual raw material to the final product is unique for every paint system, and so are the models which relate the raw materials to the rheological and mechanical properties. In Table 2.2, Schrickel [38] indicates how case-specific the *defoamer* performance is in the qualitative relationship between paint type and defoamer type.

Table 2.2 Defoamer Type versus Paint Type [38]

	Type 1	Type 2	Type 3	Type 4	Type 5
Matt Paint	++	+	+	0	++
Gloss Paint	0	++	0	0	+
Industrial Paint	0	+	++	+	+

Performance: ++ very good; + good; 0 fair; -low

Type 1: silicone free mineral oil defoamer

Type 2: silicone free, pharmaceutical grade oil defoamer

Type 3: polysiloxane emulsion defoamer

Type 4: silicone oil compound

Type 5: polysiloxane defoamer

Schricket [38] further indicates the major difference in defoamer performance when used in combination with different emulsions and dispersants. Only very general, straightforward qualitative ‘models’ can then be constructed as in Table 2.2.

However, this does not mean that a model which relates the paint properties to a specific raw material for a *specific* paint system is also useless. In fact, various quantitative models have been developed for specific paint systems. Empirical models have identified the effect of raw materials on general paint properties given in Table 2.1. For example, [29] modelled the effect of raw materials on i) hardness, ii) elasticity, iii) impact resistance, iv) adhesion, v) chemical resistance and vi) gloss. In another study, [29] modelled the effect of raw materials on i) sag, ii) gloss and iii) levelling. [39] modelled the effect of raw materials on opacity, as well as impact resistance and adhesion [25], while [36] modelled syneresis and settlement. The only drawback of these models is that they are case-specific and comparison between models is almost impossible (see *Chapter 7 Future Work*).

2.4.2 LINEAR MODELLING

The researchers above used linear modelling techniques, amongst others. [29] found that linear models are useful in explaining the main effects on the paint properties but indicated that these models are not complex enough to describe the selected properties.

Linear modelling was also used by Tusar et al [39] for determination of correlation between four raw materials (amount of binder, extender, additives and pigment) and the opacity. The values of the correlation coefficients of the MLR model were used to determine the relative importance of each of the raw materials. The model indicated the importance of raw materials as follows (Figure 2.12).

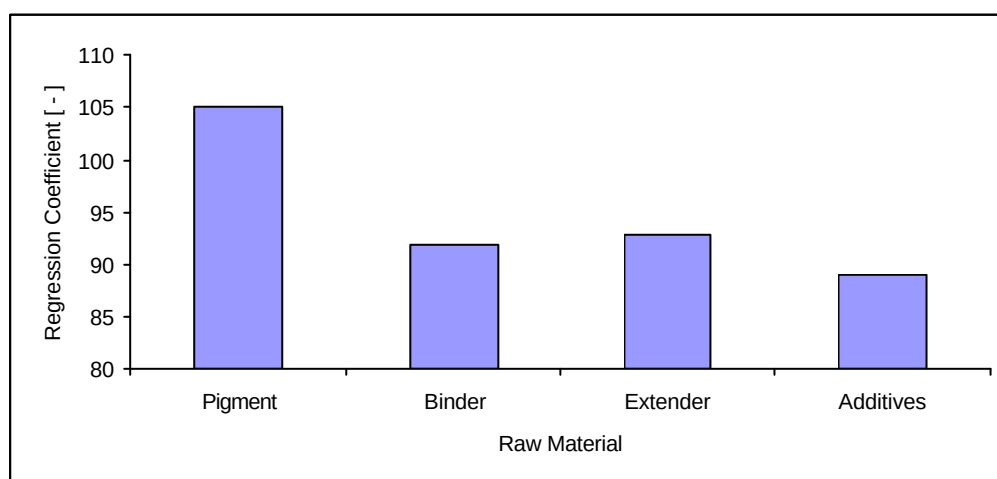


Figure 2.12 Regression Coefficients for Opacity Model [39]

The study also concluded that extrapolation outside the range of raw materials leads to incomprehensible results, stressing the importance of a comprehensive experimental design that will not require extrapolation of data.

2.4.3 NON-LINEAR MODELLING

Tusar et al (Section 2.4.2 – Linear Modelling) extended their linear models for opacity data to non-linear models, after it was concluded that the opacity behaves in a definite non-linear way with respect to the raw materials and that interaction terms definitely play a role in affecting the opacity. These models included higher order polynomials and neural networks [39]. It was found that the opacity values from the polynomial model were only accurate in a specific area of the experimental design while the predictions from the neural networks were accurate in the entire area of the selected factors. The architecture of the neural net used by Tusar et al is illustrated in Figure 2.13.

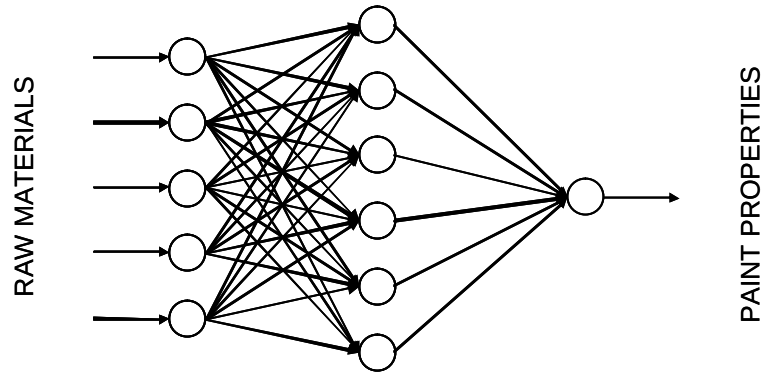


Figure 2.13 Neural Network Architecture Used by Tusar et al [25]

It can be seen that an extra input node was added to the existing four input factors. This node is for film thickness after it was found that the film thickness also affects the opacity drastically. A linear dependence between the root mean square error (RMSE)^{††} and the pigment was found. It was concluded that the neural network model was the most appropriate after analysing the errors between the actual and the predicted opacity values.

Tusar et al also investigated the non-linear behaviour of impact resistance and adhesion with respect to the raw materials of a clear coat [25]. Modelling with higher order polynomials and neural networks were used in a comparative study. It was found that both models are necessary to obtain a complete picture of correlation within the system. After validating the neural network model with a test dataset, it was concluded that different kinds of architecture gave both good and bad models and therefore, that optimisation of the neural network architecture was also very important. The optimum neural net architecture found by Tusar et al to model the impact resistance and adhesion the most accurately, is illustrated in Figure 2.14.

$$^{\dagger\dagger} RMSE = \sqrt{\frac{\sum_{i=1}^{\text{number of cases}} (\text{actual output} - \text{true output})^2}{\text{number of cases}}}$$

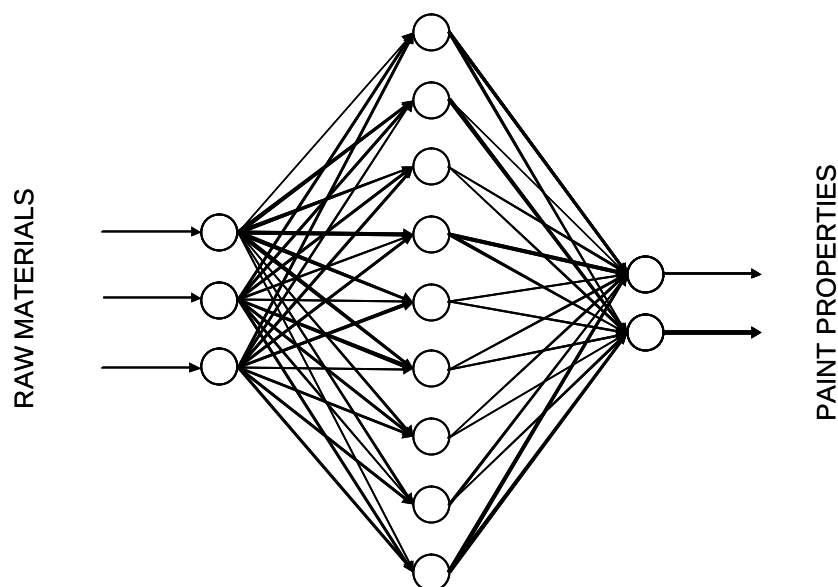


Figure 2.14 Neural Network Architecture Used by Tusar et al [25]

Figure 2.14 indicates that only three factors (polyester concentration, catalyst concentration and curing temperature) were used for model predictions. The advantage of using the specific architecture of Figure 2.14 (with two target nodes (impact resistance and adhesion), in a single neural network model) is that a general optimum for both factors is easier to locate.

According to the statistical evaluation, the neural network were accurate in prediction of impact resistance but not for adhesion values. It was recommended that a new set of experiments should be performed under different conditions. This included a larger training set or setting the target values of the output nodes to a narrower band (form 0.3 to 0.7 instead from 0.1 to 0.9). The advantage of a neural network is that new measurements can simply be added to the “training” set. According to Tusar et al, the neural network architecture does not have to be changed, even if the new measurements change the response surface. However, all of this once again indicates the importance of a thorough experimental design that can prevent the necessity of post-addition of extra measurements.

After statistical analysis of the neural network model, a formulation for a new clear coat was made based on the neural network predictions. The major advantage of obtaining a formulation in this way was that a formulation is available for every point (i.e. a product with exact defined properties) in the entire possible measurement space. This meant that the formulation could be easily modified for meeting the requirements and demands of the market.

Although not entirely applicable to the coatings industry, neural networks have also proven to be useful in prediction tablet coating properties in the pharmaceutical industry. It was shown by A.P. Plumb [43] and Hussein et al [41] that neural networks are superior in predicting the properties of a previously unseen formulation and are inherently more valuable as design tools in the development and optimisation of pharmaceutical formulations.

2.5 MODELLING THE RELATIONSHIP BETWEEN PAINT PROPERTIES AND PAINT RHEOLOGY

The understanding of coating performance in terms of rheology is far from complete [2]. Among the reasons are (a) the complexity of coating systems, which complicates the understanding of the role of rheology in process outcomes, and (b) difficulty in linking measured fundamental properties with real-world performance. Rheology has mainly been used as a quality *evaluation* mechanism, rather than a quality *prediction* mechanism, arising from typical problems such as: 'paint A spatters more than paint B, what is the difference in rheological behaviour between paint A and B and what causes it?' Of course, the rheological evaluation will indicate the differences in rheological behaviour, but at what point is differences in rheological behaviour significant in affecting a paint property? One answer to the dilemma is computer simulation of coating processes, which utilises the fundamental rheological data as the required input. However, although computer modelling is perhaps the ideal approach, it has not always been readily accessible [2]. Many coating problems were attempted to be solved from shear viscosity and viscoelastic data alone [23], [5], [24], [20] obtained from rheological measurements of paints. The data obtained from these measurements are mostly qualitative and can only be used for comparison between paints.

Modelling of product properties in terms of the rheological behaviour has been done in other areas of research, notably the pharmaceutical industry [32]. The pharmaceutical industry is technically very advanced and usually at the cutting edge of technology (due to the nature of the products), so the similarity between the modelling in this study and that done by others in the field of coatings rheology is a welcoming sign for the paint manufacturer.

It is clear that an empirical model, which relates the paint properties to the rheology is required and an article in *Progress in Organic Coatings*, 2004 [42] stated that one of the main challenges in the coatings industry that still remains, is to "Develop relationships between these properties (rheological) and relevant paint properties".

2.6 NOMENCLATURE

Abbreviations

CSD	Controlled shear deformation
CSR	Controlled shear rate
HEUR	Hydrophobically modified ethylene urethane oxide rheology modifier
LVER	Linear viscoelastic range
PVC	Pigment volume concentration
RMSE	Root mean square error
VOC	Volatile organic compound

Symbol	Description
f	Frequency (Hz)
G'	Storage modulus (Pa)
G''	Loss modulus (Pa)
t	Time (s)
x_k	Independent variable k
y_i	Dependent variable i

Greek Symbol	Description
a_i	Slope of preliminary model, ($i=HS, LS, 3-ITT_{(rot)}, etc$)
b_i	Offset of preliminary model, ($i=HS, LS, 3-ITT_{(rot)}, etc$)
b_k	Regression coefficient of the k 'th term
g	Deformation or strain (%)
g_{LVER}	Strain limit of the linear viscoelastic range (%)
$\dot{g}$	Shear rate (s^{-1})
δ	Phase shift angle, loss angle ($^\circ$)
$\tan \delta$	Loss factor, damping factor (-)
e	Error
h	Shear viscosity (Pa.s)
h^*	Complex viscosity (Pa.s)
h_∞	Infinite shear viscosity (Pa.s)
t_L	Levelling stress (Pa)
t_y	Yield stress (Pa)
t_{YP}	Yield point in terms of stress (Pa)
ω	Angular frequency (rad/s)

2.7 REFERENCES

1. Mezger, T. (2002). "Quality assurance - Rheological testing of the sagging behaviour of coatings." Stuttgart, Paar Physica.
2. Eley, R. R. (2001). Applied Rheology, ICI Paints. Strongsville, OH 44136
3. Schoff, C. K. (2005). "Organic coatings: The paradoxical materials." Progress in Organic Coatings **52**(1): 21-27.
4. Overbeek, A., Buckmann, F., Martin, E., Steenwinkel, E., Annable, T. (2003). "New generation decorative paint technology." Progress in Organic Coatings **48**: 125-139

5. Whittingstall, P. Paint evaluation using rheology, Rheometer Brochure, TA Instruments Publication RH001.
6. Rodriguez, C. L., Weathers, J., Corujo, B., Peterson, P. (2000). "Formulating water-based systems with propylene-oxide-based glycol ethers." Journal of Coatings Technology **72**(905): 67.
7. Hellgren, A., Weissenborn, P., Holmberg, K. (1999). "Surfactants in water-borne paints." Progress in Organic Coatings **35**: 79-87.
8. Kraus, S. S., Johnson, E. A. "The issue of compatibility between dispersants and associative thickeners." Brush Strokes 3(1)
9. Kastner, U. (2001). "The impact of rheological modifiers on water-borne coatings." Colloids and Surfaces A: Physiochemical and Engineering Aspects **183-185**: 805-821.
10. Svanholm, T., Kronberg, B., Molenaar, F. (1997). "Adsorption studies of associative interactions between thickener and pigment particles." Progress in Organic Coatings **30**: 167-171.
11. Prideaux, J. "Rheology modifiers and thickeners in aqueous paints." Journal of the Oil and Colour Chemists' Association, ISSN 0030-1337. Rohm and Haas.
12. Mahli, D. M., Steffenhagen, M. J., Xing, L., Glass, J.E. (2003). "Surfactant behaviour and its influence on the viscosity of associative thickeners solutions, thickened latex dispersions and waterborne latex coatings." Journal of Coatings Technology **75**(938): 39.
13. Glass, J. E. (1999). "Adsorption of hydrophobically-modified, ethoxylated urethane thickeners on latex and titanium dioxide disperse phase." Advances in Colloid and Interface Science **79**: 123-148.
14. Knoef, A. J. M., Slingerhand, H. (1992). "Urethane-based polymeric thickeners for aqueous coating systems." JOCCA **9**: 335-339.
15. Chen, M., Wetzel, W.H., Ma, Z., Glass, J.E. (1997). "Unifying model for understanding HEUR associative thickener influences on waterborne coatings: 1. HEUR interactions with a small particle latex." Journal of Coatings Technology **69**(867): 73.
16. Whitton, A. J., Van Doren, R. E. "Formulating with rheological additives for latex paints."
17. Molenaar, F., Svanholm, T., Toussaint, A. (1997). "Rheological behaviour of latexes in-can and during film drying." Progress in Organic Coatings **30**: 141-158.
18. Cloyd, J. D., Carico, K. C., Collins, M. J. (2001). Rheological study of associative thickener and latex particle interactions. New Orleans, International waterborne.
19. Reuvers, A. J. (1999). "Control of rheology of water-borne paints using associative thickeners." Progress in Organic Coatings **35**: 171-181.
20. Franck, A. Rheological characterisation of paints and coatings, TA Instruments.
21. Parsons, P. (1993). Surface Coatings: Raw Materials and Their Usage. London, Chapman and Hall.
22. Cohu, O., Magnin, A. (1996). "The levelling of thixotropic coatings." Progress in Organic coatings **28**: 89-96.
23. Hester, R. D., Squire, D. R. (1997). "Rheology of waterborne coatings." Journal of Coatings Technology **69**(864): 109.

24. Osterhold, M. (2000). "Rheological methods for characterising modern paint systems." Progress in Organic Coatings **40**: 131-137.
25. Tusar, L., Tusar, M., Leskovsek, N. (1995). "A comparative study of polynomial and neural network modelling for the optimisation of clear coat formulations." Surface Coatings International **10**: 427-434.
26. Anderson, M. J., Whitcomb, P.J. Computer-Aided Design of Experiments for Formulations, p 1-8. Mineapolis, Stat-Ease Manual, Inc.
27. Finding rules for rheology modifiers with formrules, Intelligensys Ltd. Belasis Technology Park, Teeside TS 23 4EA, UK
28. Leskovsek, N., Tusar, M., Tusar, L. "Study of rheological properties of paints with designing of experiments: a neural network approach." p. 709. Supplied by The British Library.
29. Gasperlin, M. (1995). "Empirical modelling of rheological and mechanical properties of paint." Rheology: 140.
30. Report: Optimising architectural coatings with inform. Intelligensys Ltd. Belasis Technology Park, Teeside TS 23 4EA, UK
31. Leskovsek, N., Tusar, M., Tusar, L. (1995). "Optimisation of rheological and mechanical properties of paint using methods of experimental design and neural network modelling (Polish)." p.258-259. Medvode, Ljubljana.
32. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Kristl, J. (1998). "Lipophilic semisolid emulsion systems: viscoelastic behaviour and prediction of physical stability by neural network modelling." International Journal of Pharmaceutics **168**: 243-254.
33. Gasperlin, M., Tusar, L., Smid-Korbar, J., Kristl, J. Viscous behaviour of pharmaceutical creams: characterisation by neural network and polynomial modelling." p. 14-20. University of Ljubljana, Faculty of Pharmacy, Slovenia.
34. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Zupan, J., Kristl, J. (2000). "Viscosity prediction of lipophilic semisolid emulsion systems by neural network modelling." International Journal of Pharmaceutics **196**: 37-50.
35. Leskovsek, N., Tusar, M., Tusar, L., Zupan, J. (1994). "Surface coating formulation using neural networks and statistical methods (Polish)." **28**(1).
36. Schaller, J. E. Introduction to latex paints, p. 1-57. Rohm and Haas Company.
37. (2003). Introduction to Paint Technology. Johannesburg, SA Paint Industry Training Institute: 1-48.
38. Davidson, G., Skuse, D.R. (1999). Advances in Additives for Water-Based Coatings. Cambridge, The Royal Society of Chemistry.
39. Tusar, L., Tusar, M., Barle, N., Leskovsek, N., Kunaver, M. (1995). "A study of the influence on hiding power of the composition of a paint and its film thickness." Surface Coatings International **11**: 473-481.
40. Gupta, A. (2001). "Optimization of product performance of a paint formulation using mixture experiment." Journal of Applied Statistics **28**(2): 199-213.
41. Hussain, A. S., Yu, X., Johnson, R. D. (1991). "Application of neural computing in pharmaceutical product development." Pharmaceutical Research **8**(10): 1248-1252.

42. Schoff, C. K. (2005). "Organic coatings: the paradoxical materials." Progress in Organic Coatings **52**(1): 21-27.
43. Plumb, A. P., Rowe, R. C., York, P., Doherty, C. (1999). "Aspects of training an artificial neural network for the modelling of tablet coating formulations." J. Pharm. Pharmacol **51**(92): 92.

3 EXPERIMENTAL METHODS AND TECHNIQUES

3.1 INTRODUCTION

Research in the paint industry usually requires extensive experimental work due to the practical nature of the product. The experimental methods (manufacturing and test work) are often quite tedious and very time consuming. Therefore, it is important that the data obtained from the experimental work should be adequate right from the start to prevent any additional experimental work. Various techniques are used in this study to ensure the adequacy of the data. Among these are:

- i. A systematic approach of designing the experiments so that valid and objective data are obtained from the results (see *Section 3.2 – Design of Experiment*).
- ii. International standard test methods are used for the measurement of paint properties. This ensures uniformity in the results but also provides the advantage that conclusions and interpretations obtained from the data can be compared internationally. Unfortunately, up to date, no standard test methods exist for the evaluation of the rheological behaviour of paint. However, the rheological methods used in this study can be seen as a contribution in establishing these standards.
- iii. A scientific method is followed to ensure that the effect of process variables are minimised.

3.2 DESIGN OF EXPERIMENTS (DOE)

Statistical design of experiments refers to the process of planning the experiment so that appropriate data can be analysed by statistical methods, resulting in valid and objective conclusions, in other words, the statistical approach is necessary if we wish to draw meaningful conclusions from the data. Choosing an experimental design to accommodate constraints on the sample size and experimental conditions, while also considering interactions, replications and non-linearities, is an important first step [1].

Empirical models are best developed with data that is obtained from traditional statistically designed experiments. The conclusions rely heavily on how the data were collected and therefore on the manner in which the experiments were designed and performed [2]. Amongst others, two of the main reasons why experimental design is important, are as follows [3]:

- i. Correlated data does not favour the development of an adequate model. Most data usually have strong correlations somewhere within the independent variables which can be minimised with proper experimental design.
- ii. Experimental design ensures that the dataset fully populates the n-dimensional feasible region of allowable process parameters.

Powerful desk-top software tools now make it easier to optimise paint formulations [4]. Statistically based design of experiments for mixtures has proved to be a viable method for making breakthrough improvements in cost and performance. A software package, Design-Expert[®] 6.0 from Stat-Ease, is used to generate the experimental design for two sets (screening and the modelling) of experiments^{††}. The screening process results in identifying the most important variables, which are then used for detailed modelling.

3.2.1 DOE FOR THE SCREENING PROCESS

Due to the complexity of the problem, a statistical screening process was required for evaluating the most important variables before the actual modelling could begin. These type of screening experiments have been successfully carried out before in the paint modelling process by other researchers [5]. Screening of many factors gives information about [6]:

- What the most important factors are
- Range of the factors
- Curvature (non-linear behaviour) in the response
- The next step to take.

Screening designs provide simple models with information about dominant variables, and information about ranges. In addition they result in a few experiments, which means that relevant information is gained in only a few experiments. Special screening designs are needed in constrained regions or mixture problems. Mixture problems are common in the chemical industry [7], [8], [9], [10], because the sum of components add up to 100%. This introduces a constraint on the design and must be handled with special tools and models. D-optimal designs are used when there are constraints put on the factors [11]. Linear models are sufficient for a screening design, since we are only interested in the major effects [6].

Screening designs in Design-Expert[®] are available for mixtures with 6 to 24 components and allows investigation of a large number of components in a minimal number of blends. The layout of the design that is used for the screening of the most important variables is as follows:

- *Type*: Mixture design (a type of design that is used where sum of components add up to unity)
- *Design*: Constrained design space (a type of design where many different vertices are possible and not only those on the outer edges of the design space)
- *Algorithm type*: Vertex-screening (an algorithm that uses the D-optimal algorithm to pick the best vertices to estimate a linear model)
- *Centroid replicates*: Replicates in the centre of the design-space that provide an estimate of pure error. The centroids also provide an estimate of curvature.

A paint formulator was consulted and it was suggested that a 20% variation in mass on the paint components are within practical limits and therefore the level of each variable (paint raw material)

^{††} Companies such as Ciba-Geigy[®], Duracell[®], IBM[®], Johnson & Johnson[®], Mobil[®], Sigma-Aldrich[®] and Unilever[®] have used Design-Expert[®] as part of their experimental design.

was set at + 20% and -20%. The standard formula with its upper (+20%) and lower (-20%) limits given in Table 3.1.

Table 3.1 Upper and Lower Limits of Paint Formulation Used in the Screening Process

Raw Material	Lower Limit (-20%)	Standard Formula (%)	Upper Limit (+20%)
Water	13.23	14.14	15.73
Soda Ash	0.12	0.15	0.18
Fungicide	0.16	0.20	0.24
Surfactant	0.18	0.23	0.28
Dispersant	0.94	1.18	1.42
Antifoamer	0.39	0.49	0.59
Solvent	3.26	4.08	4.89
Pigment	13.87	17.34	20.81
Extender	9.66	12.08	14.49
Emulsion/Binder	29.95	37.44	44.93
Organic Pigment	7.83	9.79	11.75
Coalescent	1.27	1.59	1.91
Rheology Modifier 1	0.82	0.82	0.82
Rheology Modifier 2	0.13	0.13	0.13

Note the upper and lower limits of Rheology Modifier 1 and 2. These levels are kept constant, regardless of the +20% upper and -20% lower limits. The idea behind this is that it is already known that the rheology modifiers have a major influence on the rheology. The idea of the screening process is to determine the most important variables, therefore the levels of the rheology modifiers are not varied because it is already known that they play a major role in the rheology of the paint.

There are several approaches in developing designs for constrained mixture experiments. A canaonical form of a quadratic mixture model as proposed by Scheffe (1958) has been used by other paint researchers [5] and was considered adequate to represent the variability in the responses of the paint properties.

The constraints on the levels of the paint components, given in Table 3.1, along with the design layout described above, were given to Design-Expert[®] and the program came up with a total of 29 runs. The variation (in % of the original fomulation) of each raw material is given below for each of the 29 runs. Figure 3.1 to Figure 3.12 illustrates the random variation of raw material levels for the respective runs. The actual paint formulation is given in *Chapter 10 – Appendix C*.

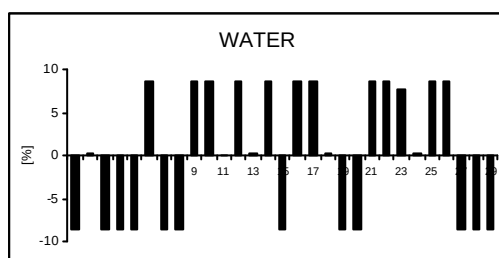


Figure 3.1 Variation of Water

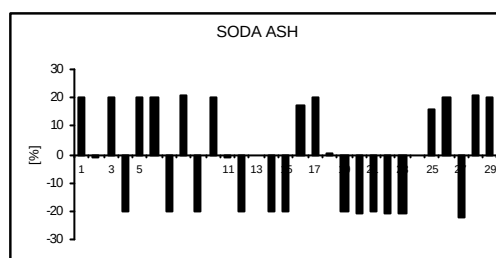


Figure 3.2 Variation of Soda Ash

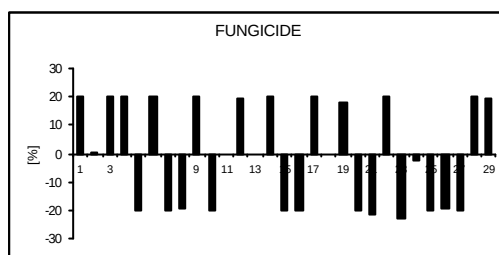


Figure 3.3 Variation of Fungicide

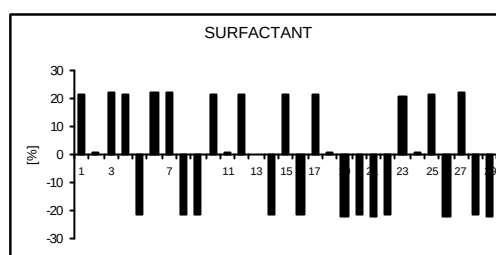


Figure 3.4 Variation of Surfactant

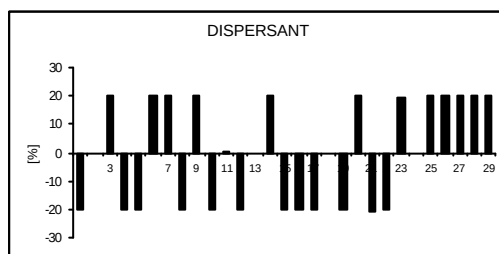


Figure 3.5 Variation of Dispersant

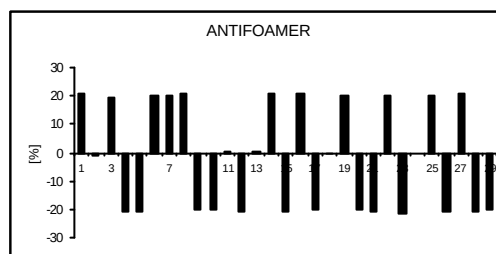


Figure 3.6 Variation of Antifoamer



Figure 3.7 Variation of Solvent

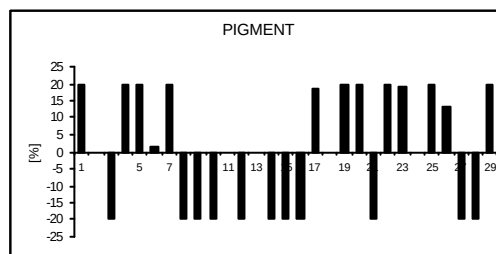


Figure 3.8 Variation of Pigment

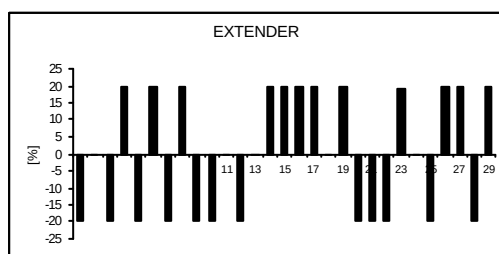


Figure 3.9 Variation of Extender

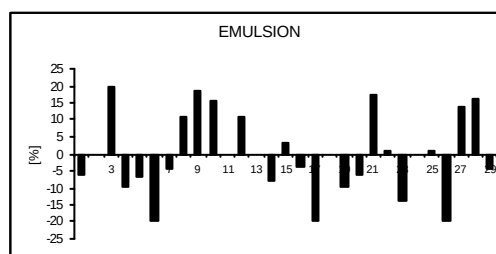


Figure 3.10 Variation of Emulsion

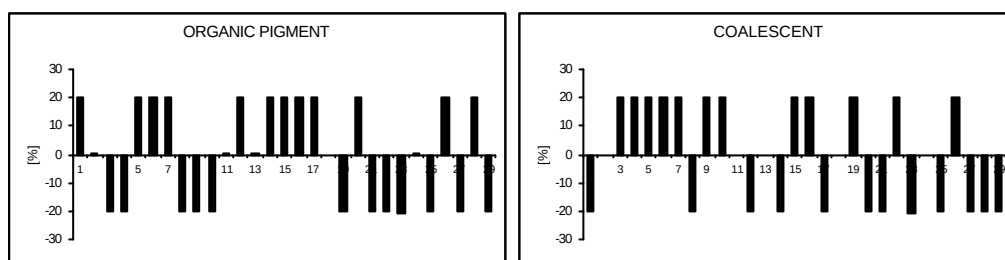


Figure 3.11 Variation of Organic Pigment Figure 3.12 Variation of Coalescent

The vertex-screening (D-optimal algorithm) ensures that the best vertices (ratios of paint components) are chosen so that a linear model could be fitted to the data. The linear model will not give much information, if any, about the interactions between the paint components and therefore screening only reveals the “big hitters”.

3.2.2 DOE FOR MODELLING PURPOSES

Experiments need to be performed in order to obtain data for three models. The experiments need to be of such nature that the correlation between:

- i) the raw materials and the rheological behaviour is obtained (MODEL 1)
- ii) the raw materials and the paint properties is obtained (MODEL 2)
- iii) the rheological parameters and paint properties is obtained (MODEL 3).

The main objective of this study is to predict the relationships between the raw materials and the rheological and mechanical properties of paint. The development of an empirical model aims at establishing internal validity (the approximate truth about inferences regarding causal relationships). Empirical models rely heavily on the experimental data, therefore careful attention is given to the design of the experimental space.

Various modelling techniques have been used in modelling of coating processes, ranging from linear [12], [13], polynomial [8] to neural networks [10], [12], [13], [14], indicating that linear as well as non-linear causal relationships exist. Similar relationships have also been found in preliminary work of this study (screening experiments). For the models under development to be of practical value, it is necessary that they should:

- i. establish accurate approximations regarding causal relationships (linear and non-linear)
- ii. be able to establish these approximations over a wide range of formulation possibilities (in other words, over the entire design space).

Statistical design of experiments can assist in achieving these goals. There are many types of experimental designs such as Box-Behnken and central composite designs, which map the centroid and surface of the model hypervolume respectively. This study uses computer-generated designs with extensive internal mapping to ensure that the most accurate approximations regarding the causal relationships are established. The experiments are designed based on modified distance-based (MDB) point selection techniques. These techniques ensure that:

- i. the chosen points are adequate to estimate the selected polynomial. This study makes use of the special cubic polynomial (equation 3.1) for point selection [2].

$$E(y) = \sum_{i=1}^p b_i x_i + \sum_{i<j}^p b_{ij} x_i x_j + \sum_{i<j<k} b_{ijk} x_i x_j x_k . \quad (3.1)$$

The special cubic polynomial facilitates linear as well as higher order terms. Higher order terms will typically be necessary in a study such as this because (1) the phenomena studied may be complex and (2) the experimental region is the entire operability region and is therefore large, requiring an elaborate model.

- ii. the set of points are spread (approximately) evenly over the entire feasible design region.

These techniques will assist in making the linear and non-linear models more adequate. However, non-linear models, such as neural networks, rely completely on empirical data to generalise the underlying trends in a system. The extent of nonlinearity or curvature are not known in advance, meaning that the required ability of the neural network to interpolate between the records is also not known. Systems with a high level of physical complexity need more data points to accommodate a higher degree of curvature [15]. The physical complexity of the system is not known in advance and therefore the number of data points required for adequate approximations is also not known. It is up to the computer-generated design of Design-Expert® to produce a number of records for the experimental design that offers extensive internal mapping of the hypervolume, as far as the special cubic polynomial allows. Design-Expert® uses the modified distance based design (MDBD). Rules of thumb do exist to establish a rough idea of the number of exemplars required for neural network modelling (see *Section 3.2.2.2 – Sample Size*).

3.2.2.1 MODIFIED DISTANCE BASED DESIGN (MDBD)

Distance based point selection chooses points from the candidate point set in a way that achieves maximum spread throughout the design region. The modified distance method checks each point as it is added to see if it estimates an additional coefficient in the model. The candidate point set is based on the model selected for the design. Design-Expert® offers an array of models and a model is chosen based on its ability to map non-linear functions the best. The point is not added to the design unless it contributes to coefficient determination.

The number of data points for the experimental design is determined by the number of input factors for the chosen model and the number of input factors are comprised of:

- the six most important factors as determined in the screening experiments (see Chapter 4)
- another two factors, which are the two types of rheology modifiers that were not included in the screening experiments
- a combination of the remaining six paint raw materials that do not play as important role in affecting the rheology (also determined from the screening analysis)

This results in a total of nine input factors which are used in the experimental design and Design-Expert[®] then calculates that a total of 134 runs will be sufficient in determining the coefficients for a cubic model. However, it must be noted that a cubic model is only being used as a type of model to generate data points that map the design space in the most non-linear way possible (that is available from the Design-Expert[®] software). This is being done because non-linear models will also be used to fit the data. The level of variation (in percentage) of the nine input factors (the eight most important raw materials plus the combination of the remaining six raw materials) for each run is given in the figures below. The actual paint formulations are given in *Chapter 10 – Appendix C*.

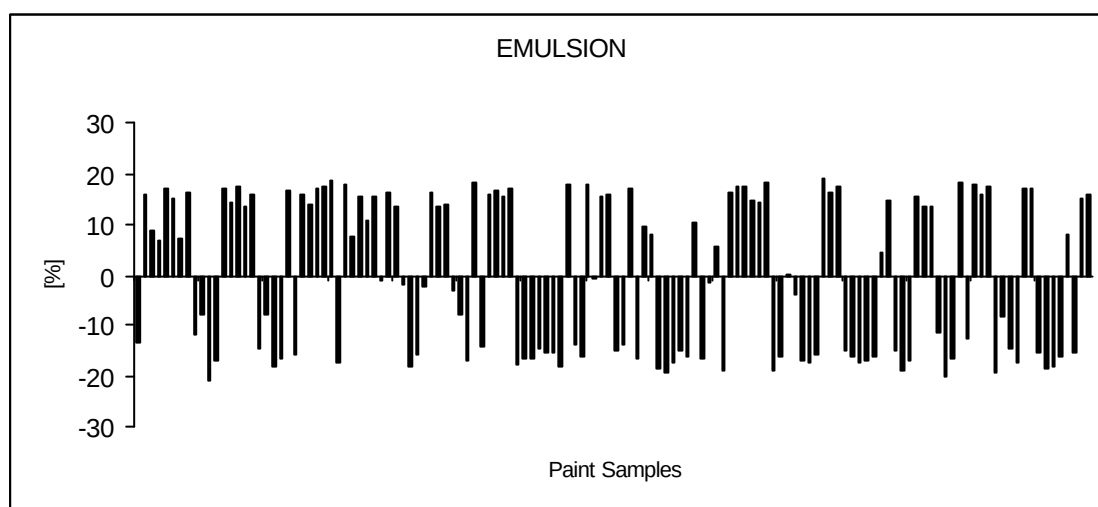


Figure 3.13 Variation (%) in Emulsion Levels on Standard Formulation

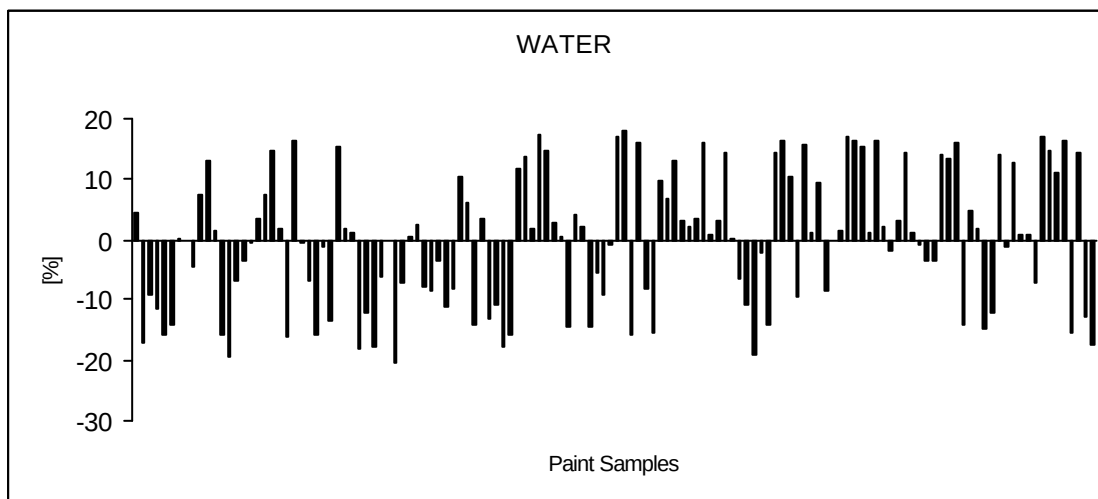


Figure 3.14 Variation (%) in Water Levels on Standard Formulation

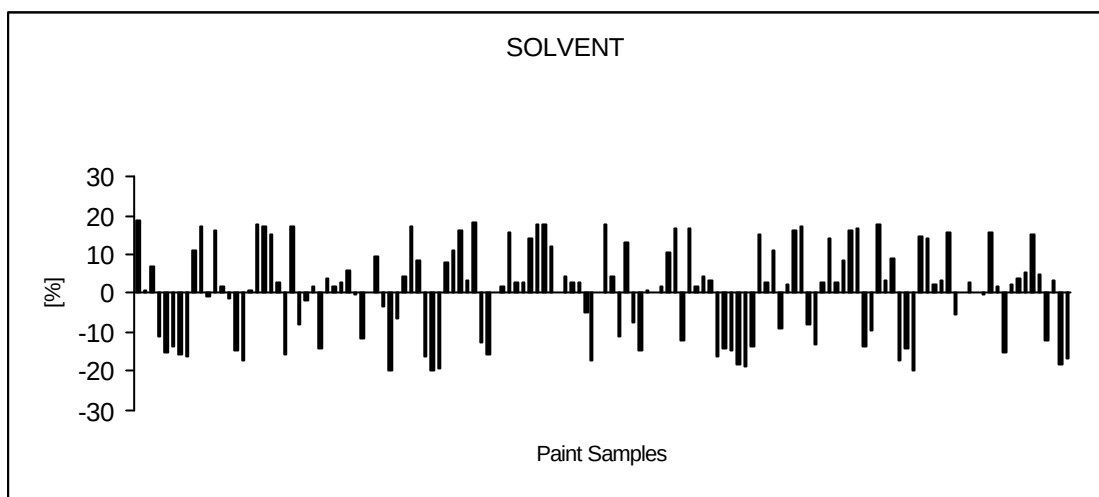


Figure 3.15 Variation (%) in Solvent Levels on Standard Formulation

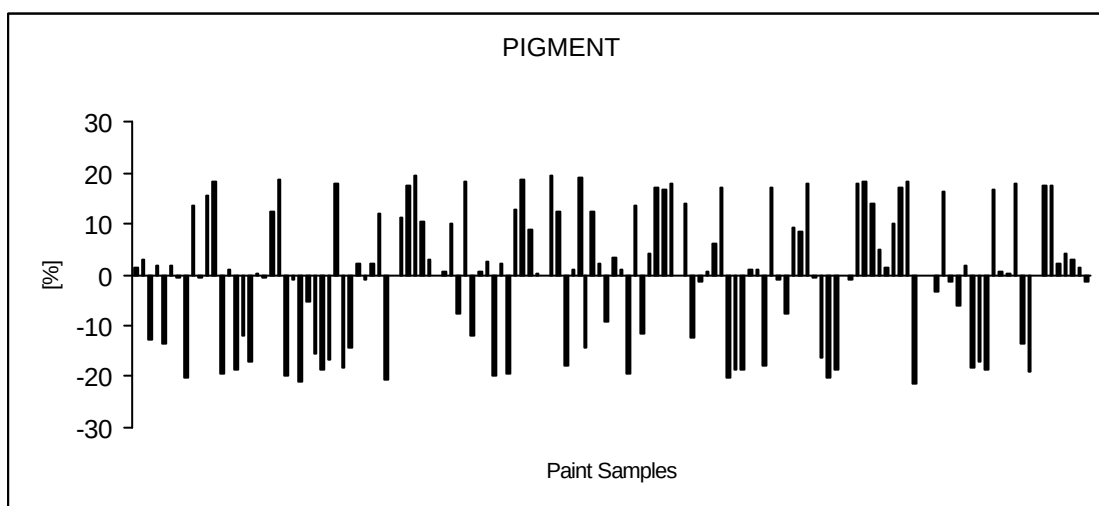


Figure 3.16 Variation (%) in Pigment Levels on Standard Formulation

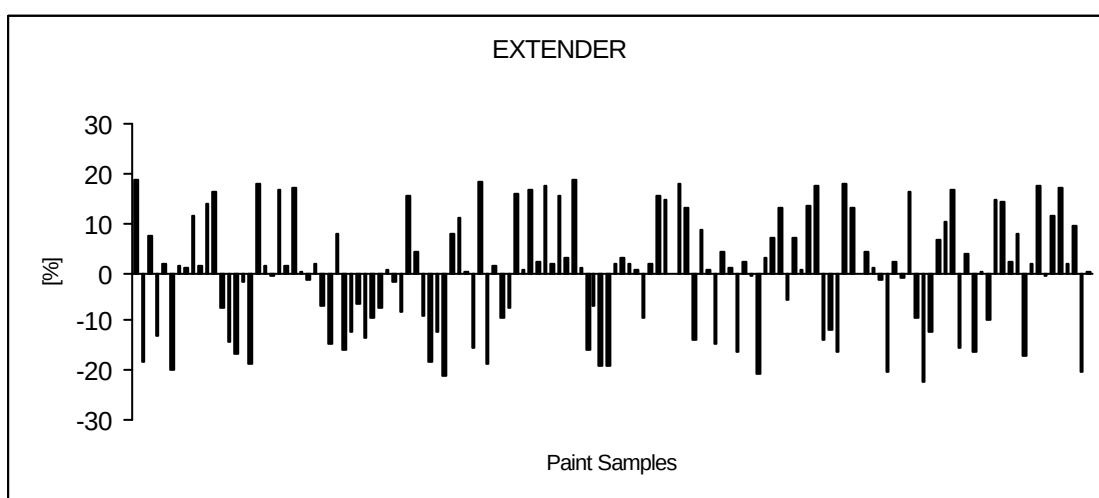


Figure 3.17 Variation (%) in Extender Levels on Standard Formulation

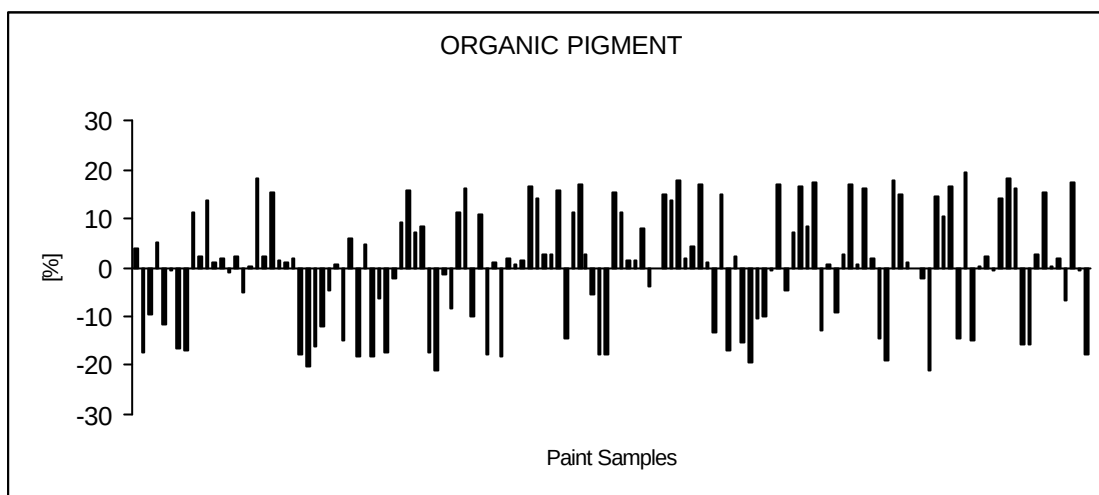


Figure 3.18 Variation (%) in Organic Pigment Levels on the Standard Formulation

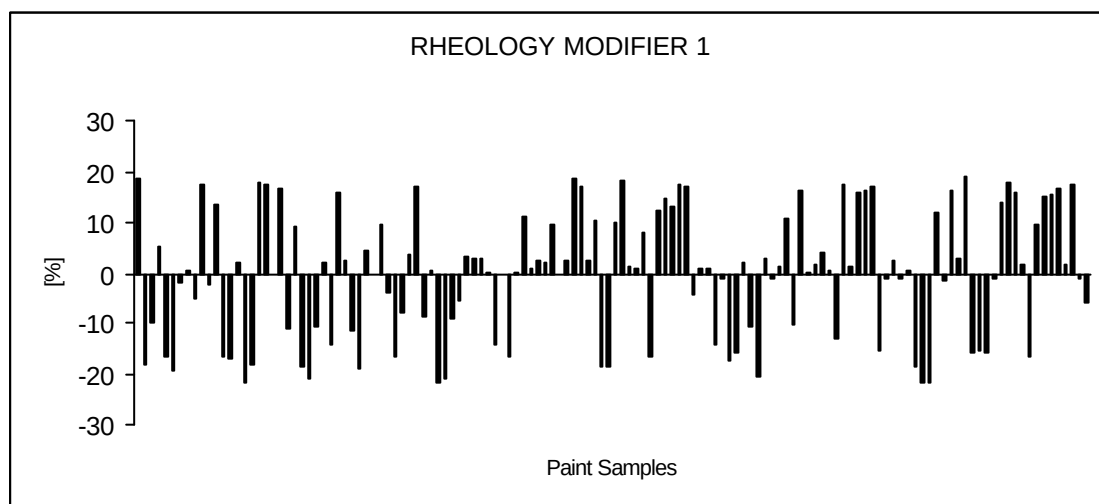


Figure 3.19 Variation (%) in Rheology Modifier 1 Levels on the Standard Formulation

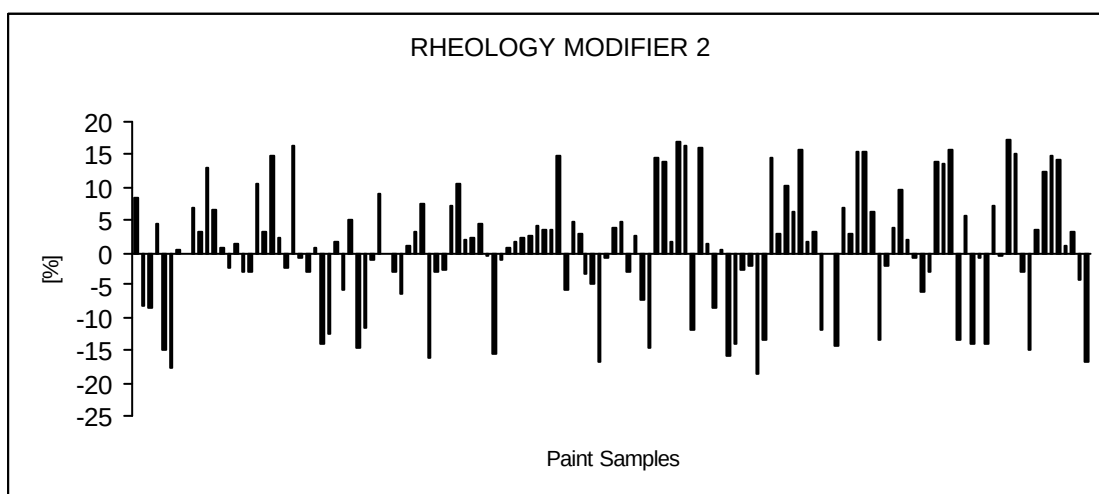


Figure 3.20 Variation (%) in Rheology Modifier 2 Levels on the Standard Formulation

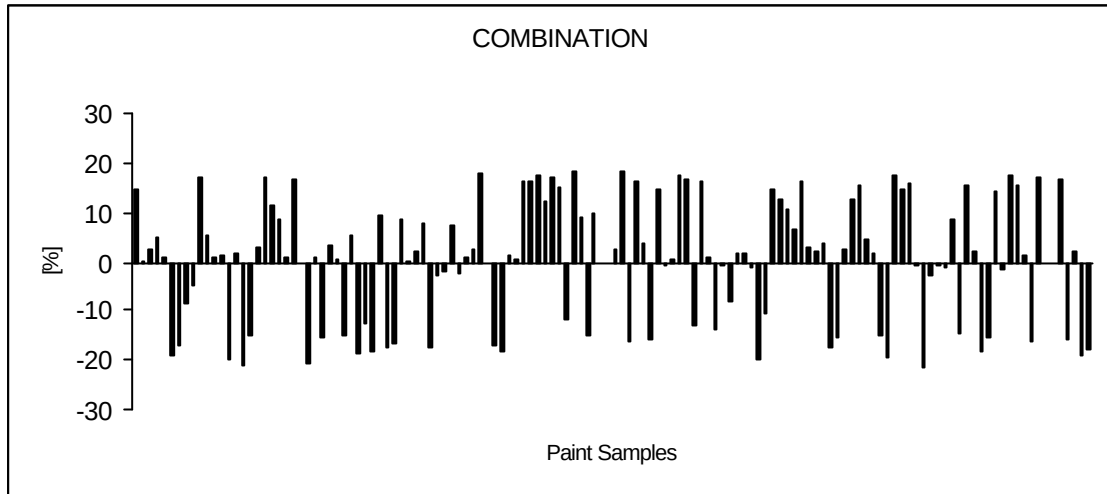


Figure 3.21 Variation (%) in the Combination of the six Remaining Raw Materials in the Standard Formulation

3.2.2.2 SAMPLE SIZE

It has been mentioned in *Section 3.2.2 – DOE for Modelling*, that the experimental design for modelling purposes is designed in such a way so that it assists in establishing causal approximations for both the linear and non-linear modelling techniques. The role of the sample size is of obvious concern and has always been a major issue when estimating what a sufficient size is for non-linear mapping such as the training of neural networks [1]. In general, the larger the sample size, the better the neural network performs at generalising the underlying trends [15], however a major problem in this study was the data acquisition process that, due to the nature of the coatings industry, is a tedious procedure. The size of the data sets depends on the complexity of the problem, the intrinsic dimensionality of the data, as well as the requisite resolution of the problem. An important question is “How large is large enough?” An exact answer can be obtained analytically [16], e.g. by means of Monte Carlo experiments, but these have limitations in that the results obtained pertain only to the environment in which the experiments are carried out and it is difficult to know whether any given data-generating mechanism is to any degree representative of the realistic empirical settings. Motivated by the desire to obtain distributional results, statisticians have developed so-called “resampling techniques” that permit accurate estimation of the sample distribution, but this is i) impractical in this situation and ii) the computational cost of such a procedure is impressive and today’s computers can not deal with the fundamental difficulty of the problem. Therefore, answering the question is extremely difficult and highly context dependent and there is no simple general answer and many rules of thumb [15]. One of these rules of thumb is given by Haykins [17] as:

$$N > \frac{Mp}{e} \quad (3.2)$$

N = size of exemplar training set
 M = hidden layer neuron count
 p = input nodes
 e = error of approximation

From equation 3.2 it is calculated that the required size of the exemplar training set is roughly 150 samples for MODEL 3 ($p = 3$) and 450 for MODEL 1 and MODEL 2. ($p = 9$). These estimations are based on no a priori knowledge about the complexity (dependence among factors) of the problem or the optimum network architecture and are therefore only used as a rough guide.

The literature is further consulted to give more insight on the number of exemplars required for adequate generalisation of the neural network. Previous studies [18] indicate that 102 records were sufficient for a neural network comprising of eight input factors and a single hidden layer of 5 nodes, to confidently predict the properties of a coating. Once again, this type of reference can only be used as an approximation. The 134 exemplars that were generated by the software of Design-Expert are sufficient.

The fact that the ratio of the number of samples to the number of weights of the neural network can also be used as a varying factor, meant that manipulation of the network architecture was also an option for better generalisation capabilities of the neural network [19].

3.2.2.3 PIGMENT VOLUME CONCENTRATION (PVC) AND VOLUME SOLIDS (VS)

A dried paint film is a three-dimensional structure and many performance attributes depend directly upon the volume proportions of the major ingredients. Weight relationships provide less useful guidance, except for the primary consideration of cost. A widely accepted way of viewing pigment content is the pigment volume concentration (PVC) of the dry paint film. This is the percentage of combined pigment volumes relative to the total volume of all non-volatile components.

$$PVC = \frac{\text{Volume of Pigment}}{\text{Volume of Binder} + \text{Volume of Pigment}} \times 100 \quad (3.3)$$

The percentage PVC at which there is just enough binder to bind the pigment particles is called the critical pigment volume concentration (CPVC) and depend on the binder type and the “binder demand” of the pigment.

The volume solids (VS) is the volume of the solids material in a given volume of wet paint, expressed as %.

$$VS = \frac{\text{Volume of Binder} + \text{Volume of Pigment}}{\text{Total Volume}} \times 100 \quad (3.4)$$

The volume solids determines a very important characteristic of the paint – how much dry film thickness (DFT) will remain from a particular wet film application.

From Figure 3.13 to Figure 3.21 it is clear that each factor varies between plus 20% and minus 20% on the standard formulation, as chosen. This results in the pigment volume concentration (PVC) for the designed experiments to be well below the critical pigment volume concentration (CPVC) and a volume solids (VS) to be relatively low as well.

It is well known that the volume solids affect the rheology of the paint and that PVC affect most properties of the paint and that many of them change abruptly at the CPVC [20]. Adding experimental data to the existing dataset of 134 data points will increase the size of the data set, but more importantly, will increase the information content of the data set. Experiments above the CPVC and with high VS will give valuable information about the system. This will also ensure that extrapolation outside the initial experimental region will not be necessary. Hence it is better to properly design an experimental space so that extrapolation is never required. The domain will include variable values (extreme values) that are not commonly seen during production levels.

Therefore, a further 30 data points with higher PVC and VS were added to the already existing 134 data points so that a total of 164 paint samples had to be manufactured. The PVC and VS for all the samples can be seen in the figures below. The actual paint formulation for these data points are given in *Chapter 10 – Appendix C*.

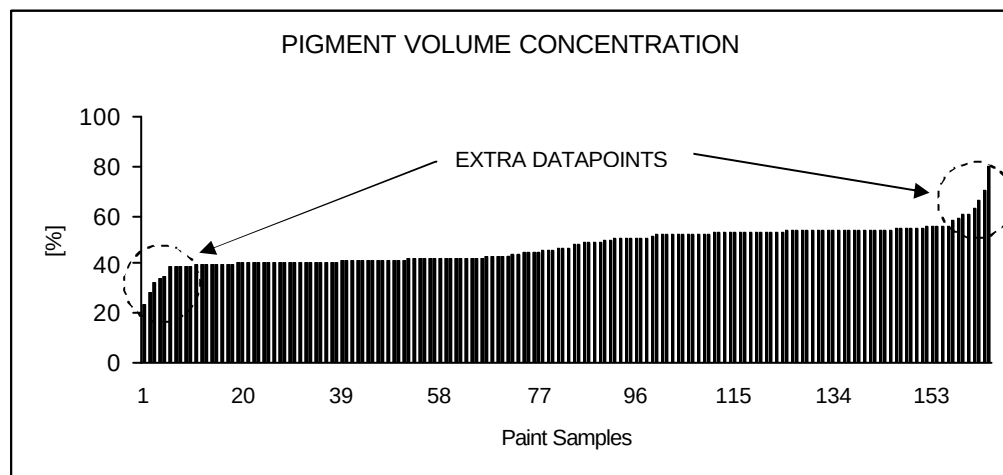


Figure 3.22 Pigment Volume Concentration of Paint Samples

The added experiments can be seen at the upper/lower limit of the pigment volume concentrations as indicated in Figure 3.22.

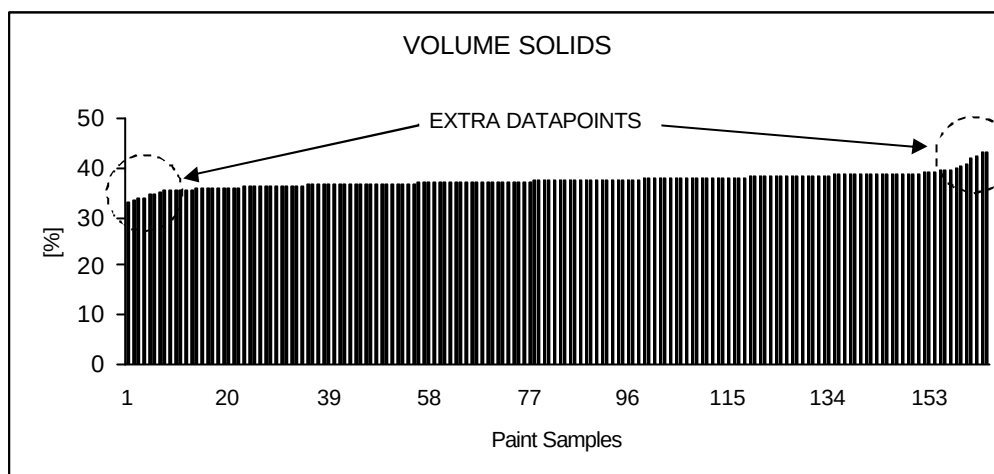


Figure 3.23 Volume Solids of Paint Samples

The extra points can again be seen at the upper/lower limit of the volume solids as indicated in Figure 3.23. The extra points added for the sake of PVC and VS are included randomly to the experimental design.

3.3 MANUFACTURING OF PAINT SAMPLES

3.3.1 THE RAW MATERIALS

An existing paint formula of a prominent South African paint company is used for investigation. This specific formulation has been chosen due to its unique/exceptional flow properties in the local market. According to experienced paint formulators this paint has “flow properties that are unmatched” by other paints and therefore it seemed worthwhile to investigate its rheological properties thoroughly. The paint consists of the following types of raw materials [24 – 35]:

- Pigment: Rutile pigment, wets out and disperses readily in aqueous and organic media, giving high opacity and excellent durability
- Extender Particles: Assist in reinforcement of paint film, reduction in the quantity of primary pigment, control gloss, improves durability and stability, enhances smoothness at the time of application
- Organic pigment: Improves efficiency of primary pigment, offers cost reduction and performance improvements
- Dispersant: Pigment dispersant, provides good gloss and hiding power. Good dispersion stability and film properties
- Emulsion: Pure acrylic emulsion, which is anionically stabilised. It lends excellent water and chemical resistance, excellent gloss, good flow properties and the ability to interact with associative thickeners, low coalescent demand and excellent adhesion properties
- Water: Forms basis/vehicle of the paint
- Coalescent: For improvement of paint properties such as minimum film forming temperature, during the film drying process

- Solvent: It has an important effect on the viscosity and stability of the coating due to its key property – balancing the hydrophilic-lipophilic character.
- Antifoamer: The antifoamer offers highly effective anti-foaming effects and storage stabilisation.
- Surfactant: A non-ionic surfactant is used as a wetting agent
- Fungicide: A microbiologically active wide-spectrum biocide is used.
- Rheology Modifiers: Two rheology modifiers are present in this paint:
 - 1) Thickener of type A: A nonionic HEUR thickener that is known for its almost Newtonian viscosity profile is used. It offers optimum film-build due to higher viscosities at high shear rates. Excellent water and alkali resistance is also obtainable
 - 2) Thickener of type B: A nonionic HEUR thickener was used in combination with that of type A. It thickens significantly at low shear rates and does not cause a change in viscosity at medium or higher shear rates. A pseudoplastic^{ss} rheology profile is obtained which lends an excellent balance between flow and sag to the paint

164 paint samples were manufactured, all with varying degrees of the above mentioned raw materials.

3.3.2 SAMPLE PREPARATION

The following formulation is followed during preparation of the paint samples:

❖ **In a 1 litre container (A), diameter = 11 cm:**

- 1) Pigment
- 2) Extender

❖ **In a 1 litre container (B), diameter = 11 cm:**

- 1) Water
- 2) Soda Ash
- 3) Fungicide
- 4) Surfactant
- 5) Dispersant
- 6) Antifoamer
- 7) Solvent

- Place B under mixer^{***} and start mixer at 150 rpm
- Add contents of A slowly to B (+/- 3 minutes)
- Increase mixer speed to +/- 500 rpm
- Mix for 30 minutes

❖ **In a 1 litre container (C), diameter = 11 cm:**

- 1) Emulsion
- 2) Organic pigment

^{ss} See Chapter 9, Appendix B (Section 9.1 - Flow Behaviour and Viscosity)

^{***} The Cowles blade is used as the impeller device, diameter = 9 cm

- 3) Water
 - Stop mixer
 - Remove B from mixer and place C under mixer
 - Start mixer at 150 rpm and add B to C (+/- 2-5 minutes)

❖ **Using B:**

- 1) Coalescent
- 2) Antifoamer
 - Mix content of B with a spatula
 - Add contents of B to C (+/- 30 seconds)
 - Mix at 150 rpm

❖ **In a small glass beaker**

Rheology modifier of type A

Rheology modifier of type B

Add rheology modifier (Type A) to C

Add rheology modifier (Type B) to C

Increase mixing speed to +/- 200 rpm

Mix for 10 minutes

3.3.3 SCIENTIFIC METHOD

The scientific method is a critical issue as results obtained and used in modelling work must be obtained scientifically and be reliable. In order to investigate the functionality of the raw materials, paint samples must be prepared in such a way that the processing conditions should be eliminated, e.g. effect of pH and the level of pigment dispersion. These processing variables are minimised in order to make accurate conclusions regarding the functionality of the raw materials.

3.3.3.1 PAINT pH

The effect of pH is minimised by ensuring that the pH of the manufactured samples are within the pH specifications of the standard paint formula (8.8 – 9.5)^{†††}. The pH of all the manufactured paint samples are $8.8 < \text{pH} < 9.2$ and therefore fall within this range and no pH adjustment was required. Further tests were carried out to ensure that the effect of pH within the range of manufactured paint samples does not have an effect on the paint properties and rheology. This involved measuring a paint property (opacity) and rheology of the standard paint formulation that has been adjusted to a range that covers the pH of the manufactured paint samples ($8.8 < \text{pH} < 9.2$). The results for the opacity values of these paint samples are shown in Table 3.2.

^{†††} pH of standard formulation obtained from paint manufacturer

Table 3.2: Opacity of Standard Paint Formula at different pH's

	pH	Opacity
Standard	8.82	0.96
	8.99	0.96
	9.20	0.96

Although the opacity values are lower than expected, they are still similar and therefore one can conclude that the pH variation (± 0.2) of the manufactured paint samples do not have a significant effect on the opacity.

The results of the rheology measurements of the three standard paint samples at different pH's are shown in Figure 3.24. The oscillation and rotational rheology measurements of all three samples are similar and no significant differences in the rheological behaviour occur. Therefore the pH variation (± 0.2) of the manufactured paint samples do not have a significant effect on the rheology.

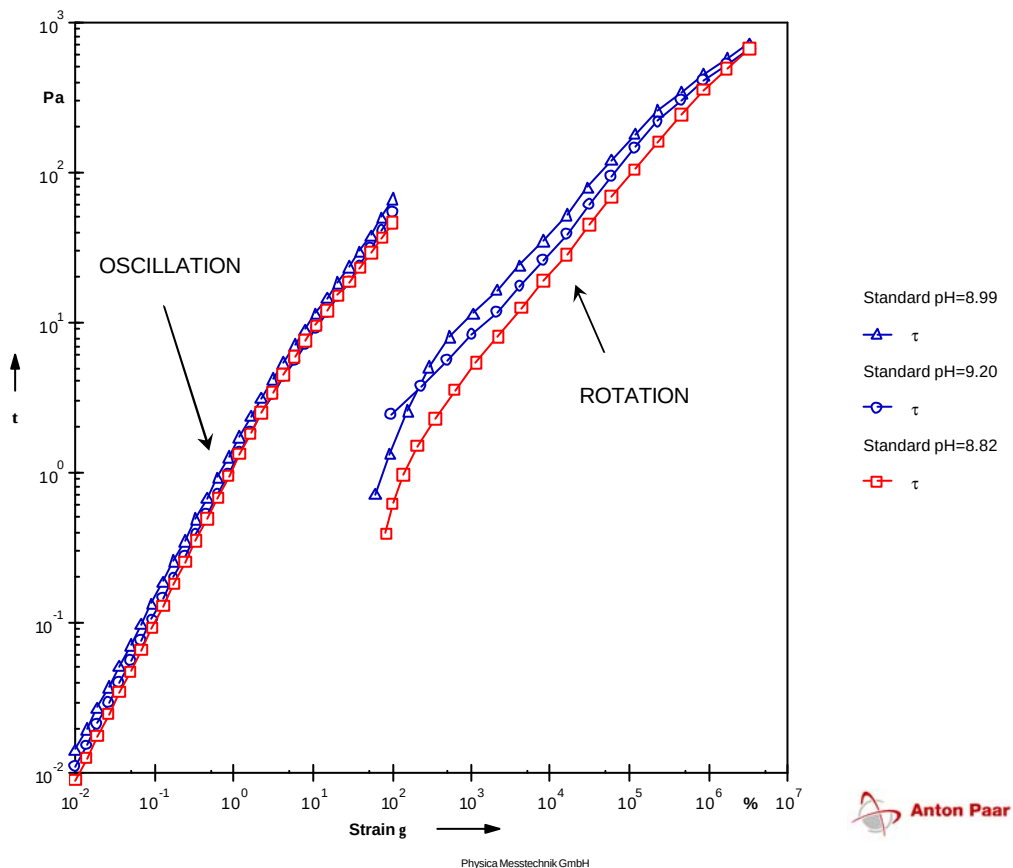


Figure 3.24: Rheology of Standard Paint Samples at different pH values

It has been shown that the pH variation of the manufactured samples does not play a significant role on opacity and the rheology.

3.3.3.2 PIGMENT DISPERSION

The opacity of a pigment can be related to the refractive index (RI). Other factors affecting opacity are the sizes of pigment particles, their distribution and the degree of dispersion achieved [36]. The effect of pigment dispersion on paint properties and rheology is minimised by dispersing the dry powder raw materials in constant ratios with respect to the water that is used in the dispersion phase. This is possible because water is also added in the letdown phase of the manufacturing process and therefore the amount of water that is added in the pigment dispersion phase can be chosen accordingly so that the ratio stays the same (constant) as that of the standard paint formulation. An example of the design of experiments of five paint samples of the dispersion phase is illustrated in Table 3.3.

Table 3.3: Design of Experiments of Dispersion Phase of Paint Sample Manufacturing

Sample ID	Pigment [%]	Water (dispersion) [g]	Pigment+Extender [g]	$\frac{\text{water}}{\text{pigment} + \text{extender}}$
Standard formula	17.3	82.3	294.2	0.280
95	17.1	75.2	268.9	0.280
125	20.9	91.9	328.6	0.280
105	15.0	71.9	257.1	0.280
150	8.7	63.71	227.6	0.280
90	14.6	70.0	250.3	0.280

Even though the level of pigment varies, the ratio of

$$\frac{\text{water}}{\text{pigment} + \text{extender}} = 0.280 \quad (3.5)$$

stays constant throughout the sample manufacturing process. This is only an attempt in minimising the variance of pigment dispersion, because other factors such as the level of dispersant, etc also plays a role. The effectiveness of this method is investigated by measuring the fineness of grind (FOG) and the rheology of the pigment grind.

The level of pigment dispersion is determined by measuring the FOG according to ISO 1524:2000 which is adopted with the permission of the international organisation for standardisation. The FOG specification of the standard formulation is < 25 microns⁺⁺⁺. Images that show the FOG of paint samples at manufactured at:

standard level of pigment

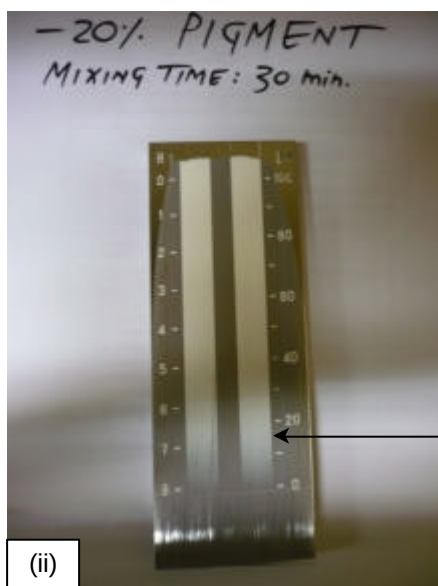
-20% pigment

+20% pigment is given in Figure 3.25.

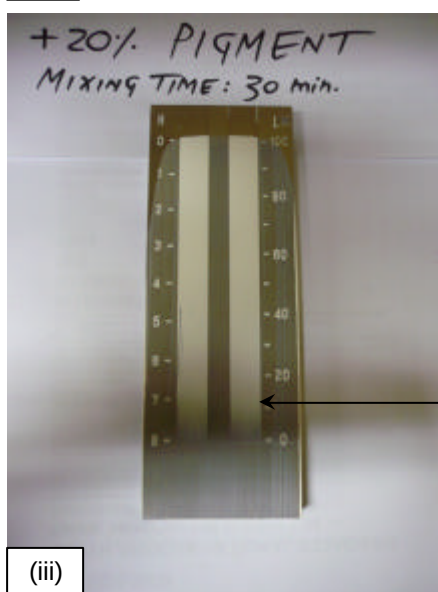
⁺⁺⁺ FOG of standard formulation obtained from paint manufacturer



FOG < 25 micron



FOG < 25 micron



FOG < 25 micron

Figure 3.25: FOG of Pigment Grind at (i) Standard Pigment (ii) -20% Pigment and (iii) +20% Pigment Levels

The standard formulation, the paint sample with +20% pigment and the paint sample with -20% pigment, have similar FOG values. The rheology of these pigment grinds are shown in Figure 3.26. Similar rheological behaviour is observed in the oscillation and rotational rheology curves.

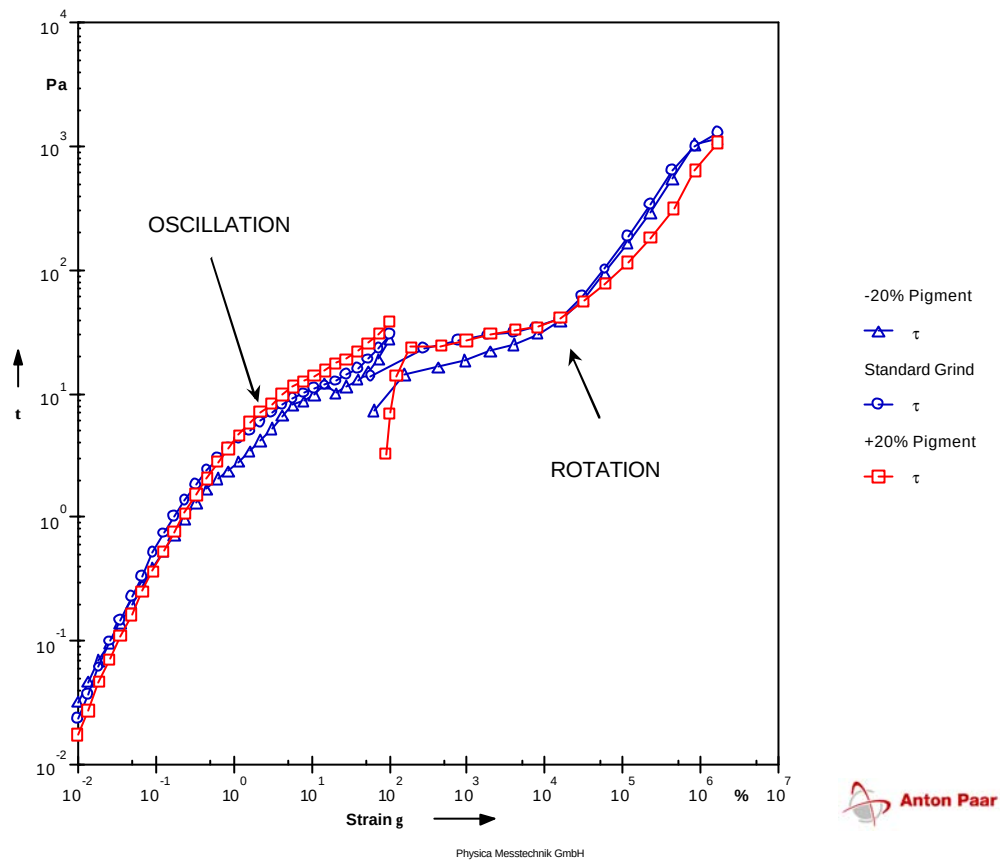


Figure 3.26: Rheology of Pigment Grind at Standard Pigment, -20% Pigment and +20% Pigment Levels

It is therefore scientifically possible to compare paint samples with different levels of pigment and extender by ensuring that the pigment dispersion process stays the same by means of controlling the ratio of dry powder materials and water in the grinding phase.

3.4 PAINT PROPERTIES

Unlike rheological test methods, there exist many measuring techniques for measuring paint properties, some of them which are developed in-house by paint companies and some of which are combinations of standards, etc. One of the most important criteria for choosing a test method was that the results had to be comparable with other results. This is only obtainable by using standard test methods as far as possible, e.g. American Standard Test Methods (ASTM). On the one hand, comparable results are very important, but on the other hand, one also wants results that are of practical importance and can be used by the paint formulator, e.g. Coatings Research Group International (CRGI). This resulted in a range of different test methods. Refer to Table 3.4 for a list of test methods used.

Table 3.4 Standard Test Methods for Evaluating Paint Properties

Paint Property	Standard Test Method
Dry Burnish	CRGI 34
Wet Burnish	CRGI 33
Chemical Resistance NH ₃	CRGI 36
Scrub Resistance	CRGI 30
Dirt Pick-Up	In-house
Open Time	ASTM D1640 – 95
Dry Film Thickness	ASTM D1005 – 95
Wet Film Thickness	ASTM D1212 – 91
Gloss	ASTM D523 – 89
Settling	ASTM D869 – 85
Hiding Power (Roller)	CRGI 43
Hiding Power (Brush)	CRGI 43
Krebs Viscosity	ASTM D562 – 01
Opacity	ASTM D2805 – 96a
Sagging	ASTM D4400 – 99
Spreading Rate	CRGI 38
Stain Removal	CRGI 32
Water Permeability	ASTM D1653 – 93
Water Resistance	ASTM D870 – 02
Adhesion	ASTM D6677 – 01
Brush Drag	ASTM D4958 – 97
Ease of Stir	In-house
Levelling	ASTM D4062 – 99
Roller Drag	ASTM D4958 – 97
Spatter	CRGI 39
Syneresis	ASTM D1849 – 95
Microbial Attack	ASTM D3274 – 95

A brief description follows which states the basic principle followed in each measurement.

❖ **DRY BURNISH**

- Standard: CRGI 34
- Summary: A drawdown of a paint film (175 mm) is allowed to age for one week at room temperature. A cheesecloth-wrapped sponge is used to burnish the film under controlled, reproducible conditions. The 85° gloss of the film is measured at several points in the path of the cloth-wrapped sponge, before and after burnishing. The increase in average gloss is reported

❖ **WET BURNISH**

- Standard: CRGI 33
- Summary: A drawdown of a paint film (175 mm) on a black plastic panel is allowed to age for one week at room temperature. A brittle brush is used to burnish the film under controlled, reproducible conditions. A soap solution is applied on the film whilst burnishing occurs. The 85° gloss of the film is measured at several points in the path of the brittle brush, before and after burnishing. The increase in average gloss is reported

❖ **CHEMICAL RESISTANCE NH3**

- Standard: CRGI 36
- Summary: A paint film (150 mm) is drawn down on a glass panel and air dried for one week. A pool of dilute ammonium hydroxide (1.4% NH3) is placed in contact with the film and the film is rated for resistance to blistering and wrinkling.

❖ **SCRUB RESISTANCE**

- Standard: CRGI 30
- Summary: A test film (150 mm) is drawn down on a black plastic scrub panel and air dried for one week. It is put into a Sheen® abrasion machine and scrubbed with a cleanser solution. The test is terminated when the film in scrub path has been completely abraded away. Results are reported in number of cycles to completely abrade away the film.

❖ **DIRT PICK-UP**

- Standard: In-house method
- Summary: Dirt (sieved and 'from the area', according to the standard) is sprinkled over a paint film (150 mm) and left to stand for 30 minutes. The film is then washed to remove the dirt. Gloss values are taken before and after and is used to calculate the dirt collection. Categorical evaluation is also done to determine the degree of dirt collected on the film.

❖ **OPEN TIME**

- Standard: ASTM D1640 - 95
- Summary: To determine set-to-touch time, lightly touch the test film (150 mm) with the tip of a clean finger and immediately place the fingertip against a piece of clean, clear glass. Observe if any coating is transferred to the glass. Report the time when no more paint is transferred to the glass.

❖ **DRY FILM THICKNESS**

- Standard: ASTM D1005 - 95
- Summary: Close the gauge of a micrometer slowly until contact is made with the uncoated test panel and record the reading. Open the gauge and lay the coated test panel under the gauge. Close the gauge slowly and take the reading. The difference in readings is the thickness of the film. Take average values.

❖ **WET FILM THICKNESS**

- Standard: ASTM D1212 - 91

- Summary: A draw-down (150 *mm*) of the paint is made and the wet film thickness is determined as quickly as possible to reduce shrinkage due to solvent loss. A gauge with a convex lower surface is pushed into the wet film until the centre touches the substrate. Read the point at which the coating first makes contact with the eccentric wheel and determine the mean which is considered as one reading.

❖ **GLOSS**

- Standard: ASTM D523 -91
- Summary: A gloss meter makes measurements at 20, 60 and 85 degrees at the surface of a dry paint film (150 *mm*). Take the average after three readings.

❖ **SETTLING**

- Standard: ASTM D869 85
- Summary: Place paint in a container, close the container and hold undisturbed for shelf aging for six months. After six months, open the container and use spatula to determine the extent of portions of the paint that may have separated during storage by moving the spatula at the bottom of the container. Note that this is a very subjective method.

❖ **HIDING POWER**

- Standard: CRGI 43
- Summary: Measure the reflectance of the paint on the black and white painted surfaces with a reflectance-measuring instrument. Calculate mean values of reflectance of each test area at four locations.

❖ **KREBS VISCOSITY**

- Standard: ASTM D562 - 01
- Summary: The load required to produce a rotational frequency of 200 rpm for an offset paddle rotor immersed in a paint is determined.

❖ **OPACITY**

- Standard: ASTM D2805 – 96a
- Summary: Measure the reflectance of the paint film (150 *mm*) on black and white hiding power charts with a reflectance-measuring instrument. Calculate mean values of reflectance of each test area at three locations.

❖ **SAGGING**

- Standard: ASTM D4400 - 99
- Summary: The paint is drawn down on a non-porous surface using a special segmented blade. The panel is then hung in a vertical position with the stripes formed by the segmented blade arrayed horizontally, and the film is rated for sag tendency.

❖ **SPREADING RATE**

- Standard: CRGI 38
- Summary: After the weight per volume of the test paint is determined, the coating is applied onto a surface of known area. The net weight of paint used is determined, then the spreading rate in terms of surface area per volume is calculated by conversion of the weight of paint to its equivalent volume.

❖ **STAIN REMOVAL**

- Standard: CRGI 32
- Summary: A test film (175 *mm*) is drawn down on a scrub panel. After one week it is stained with specified staining media, put into a washability machine (same machine used for testing of burnish and scrub resistance - Sheen® Abrasion Tester) and scrubbed with a cleaning solution. Reflectance properties are measured with a colour computer before and after scrubbing.

❖ **WATER PERMEABILITY**

- Standard: ASTM D1653 - 93
- Summary: The test specimen (paint film of 150 *mm* on filter paper (grade 589/1, or “fast”)) is sealed to the open mouth of a cup containing water, and the assembly is placed in a test chamber with controlled humidity. Periodic weighing of the cup is made to determine the rate of water vapour permeating through the paint film.

❖ **WATER RESISTANCE**

- Standard: ASTM D870 - 02
- Summary: Paint coated glass panels (paint film of 150 *mm*) are partially immersed in water. The exposure conditions are controlled by selecting the temperature of the water. Water permeates the coating at rates that are dependent upon the characteristics of the coating and upon the temperature of the water. The time it takes for softening of the coating to occur is reported.

❖ **ADHESION**

- Standard: ASTM D6677 - 01
- Summary: Adhesion is determined by making an “X” cut into the coating film (150 *mm*) to the substrate and lifting the coating with an adhesive material. Adhesion is evaluated qualitatively on a 0 to 10 scale.

❖ **BRUSH DRAG**

- Standard: ASTM D4958 - 97
- Summary: A polyester brush (width = 50 *mm*) is used to apply paint on a test area. The degree of brush drag is evaluated subjectively and given a rating.

❖ **EASE OF STIR**

- Standard: In-house Method
- Summary: The ease of stir is evaluated with a stirring apparatus and given a subjective rating.

❖ **LEVELLING**

- Standard: ASTM D4062 - 99
- Summary: A special levelling test blade designed to lay down a film with parallel ridges simulating brush marks is used for applying the paint to a substrate. After allowing the draw-down to dry in a horizontal position, levelling of the test paint is rated by comparing the contrast of lightness and shadow caused by the paint ridges under the same lighting conditions.

❖ **ROLLER DRAG**

- Standard: ASTM D4958 - 97
- Summary: Follows the same procedure as for brush drag, except for a sheep skin roller.

❖ **SPATTER**

- Standard: CRGI 39
- Summary: The paint is applied by roller to a test surface. A black plastic sheet is placed on a small horizontal shelf directly beneath and at a right angle to the test surface. The paint is applied in a certain manner and the degree of spatter is rated according to the intensity of the droplets that have accumulated on the black plastic sheet.

❖ **SYNERESIS**

- Standard: ASTM D1849 - 95
- Summary: The paint is poured into a test tube and left to stand for six months. A visual evaluation is done after 6 months to determine the degree of phase separation. A subjective rating is given.

❖ **MICROBIAL ATTACK**

- Standard: ASTM D3274 - 95
- Summary: A thin film (150 mm) of paint is applied to a glass panel. The panel is left to stand in open air conditions. A visual evaluation is done after 6 months to determine the degree of microbial attack on the film.

3.5 RHEOLOGY

3.5.1 INTRODUCTION

The ideal situation for the paint formulator would be to adjust low, medium and high shear rate behaviour independently, allowing optimisation of each property in turn [21]. For coatings, the process of application and film formation obviously requires not only a large total deformation, but also a high degree of control of flow in order to achieve success. Flow cannot be controlled unless it can be properly measured. The objective for the applied rheologist, therefore, is to develop methods of rheological characterisation that (a) yield accurate data for complex fluids, and (b) are relatable to the critical processes that paints undergo. The latter objective requires characterisation methods that cover a wide range of stresses and time scales [22].

3.5.2 RHEOLOGICAL MEASUREMENTS

Various types of viscometers are in common use. For example, the high-shear viscosity of paint is often measured with the ICI Cone and Plate viscometer, whilst for low and medium shear rate measurements, the Brookfield, Stormer and ICI Rotoviscometers are commonplace. These viscometers have the drawback that they are only capable of measuring single viscosity points. Modern sophisticated rheometers are capable of measuring various rheological parameters over the complete range of shear rates/deformation, allowing for the graph to be plotted over the entire input range.

3.5.2.1 RHEOMETER

Rheometers are able to measure more than mere viscosity values are called rheometers. All the rheological properties of the paint samples were measured with a Paar Physica® Modular Compact Rheometer (MCR 300). The measuring drive of the rheometer is able to adjust speeds and to measure the torques that are generated - it is then operated in the controlled shear rate mode (CSR) for rotational tests or controlled shear deformation mode (CSD) for oscillatory tests. The rheometer is also able to adjust the torques and to measure the speed that is generated – it is then operated in the controlled shear stress (CSS) mode. The rheometer software carries out the conversion to rheological variables after identifying the measuring system and its geometrical dimensions.

3.5.2.2 MEASURING SYSTEMS

Different measuring systems are available, depending on the type of measurement that is to be performed.

1) Cone and plate (CP)

This measuring system consists of an upper cone and a lower plate. It is designed for high shear rates and therefore has a very narrow shear gap. The geometrical dimensions of the cone are determined by the cone radius R and the cone angle α (angle at which the cone is truncated with respect to the horizontal). The cone and plate (CP) measuring system has the following advantages [23]:

- There are homogenous shear conditions present, because the shear rate is constant in the entire conical gap
- Only a very small amount of the sample is required
- Possible air bubbles are pressed out of the conical shear gap before the test starts
- Cleaning is easy and safe.

Disadvantages of the cone and plate (CP) measuring system are [23]:

- It limits the maximum particle size allowed in the sample
- For high-viscosity samples, it can cause shear even before the test has started
- Skin formation and solvent evaporation can occur at the edge of the upper plate
- It is recommended that CP measurements should be run at one temperature only, due to the strong influence of equipment-dependent thermal expansion on the narrow gap.

Despite the many possible disadvantages, the first advantage outweighs them all and should be used under shearing conditions, unless conditions dictate otherwise.

2) Plate and plate (PP)

This measuring system consists of two plates. The surfaces of both the upper and lower plate are flat. The geometry of the plate is determined by the plate radius R . The plate and plate (PP) measuring system has the following advantages [23]:

- Oscillatory tests within the LVER can be performed
- It is possible to set the gap size and therefore it is possible to measure dispersions with relatively large particles
- The shear rate can be changed easily by changing the gap size

- Temperature effects due to expansion or contraction are less pronounced using a large gap.

Disadvantages of the plate and plate (PP) measuring system are [23]:

- There is no constant shear gradient in the gap because the shear rate increase from zero at the centre of the plate to a maximum at the edge. This can greatly influence samples which are viscoelastic. However, at very low deformation values, within the LVER, this effect is negligible and can therefore be used for oscillatory tests.

3.5.2.3 MEASUREMENTS

Measurements were performed both in the rotational (CSR) and oscillatory (CSD) modes. Table 3.5 gives a detailed description of the measurements that are performed on the paint samples.

Table 3.5 Rheological Measurements

Type	Measuring System	Mode	Presetting	Measurement
Flow Curve (FC)	CP	CSR	Shear rate profile $\dot{g}(t)$	Shear stress $t(\dot{g})$
High Shear (HS)	CP	CSR	Constant shear rate $\dot{g} = \text{constant}$	Shear stress $t(t)$
Low Shear (LS)	CP	CSR	Constant shear rate $\dot{g} = \text{constant}$	Shear stress $t(t)$
3-ITT (rot)	CP	CSR	Shear rate profile $\dot{g}(t)$	Shear stress $t(t)$
Amplitude Sweep (AS)	PP	CSD	$g(t) = g_A \cdot \sin(wt)$ $w = \text{constant}$ $g_A = \text{variable deformation amplitude}$	$t(t) = t_A \cdot \sin(wt + d)$
Frequency Sweep (FS)	PP	CSD	$g(t) = g_A \cdot \sin(wt)$ $w = w(t)$ $g_A = \text{constant}$	$t(t) = t_A \cdot \sin(wt + d)$
Time Sweep (TS)	PP	CSD	$g(t) = g_A \cdot \sin(wt)$ $w = \text{constant}$ $g_A = \text{constant}$	$t(t) = t_A \cdot \sin(wt + d)$
3-ITT (osc)	PP	CSD	$g(t) = g_A \cdot \sin(wt)$ $w = \text{constant}$ $g_A = \text{constant}$	$t(t) = t_A \cdot \sin(wt + d)$
Extra Low Frequency Sweep	PP	CSD	$g(t) = g_A \cdot \sin(wt)$ $w = \text{constant}$ $g_A = \text{constant}$	$t(t) = t_A \cdot \sin(wt + d)$

The last three tests of Table 3.5 produced different test results although the presetting and measurement seemed the same. These tests differed in measurement stages and measurement duration.

All the tests were performed at 25 °C and with presettings as described in Table 3.5. The cone and plate (CP) measuring system was used for all the tests that were performed in the controlled shear rate (CSR) mode. The truncated upper cone ensured that the shear rate stays constant across the entire radius of the cone so that a similar shear was applied to the whole paint sample. The plate-and-plate (PP) measuring system was used for all the tests that were performed in the

controlled shear deformation (CSD) mode. Of all the controlled shear deformation tests, the amplitude sweep was performed first in order to determine the linear viscoelastic range (LVER) of each paint sample. The rest of the controlled shear deformation tests (Frequency Sweep, Time Sweep, 3-ITT and Extra Low Frequency Sweep) were all performed within the LVER ($g < g_{LVER}$).

3.6 NOMENCLATURE

Abbreviations

3-ITT	Three-interval-thixotropy-test
AS	Amplitude sweep
ASTM	American standard test methods
FC	Flow curve
CPVC	Critical pigment volume concentration
CRGI	Coatings research group international
CSD	Controlled shear deformation
CSR	Controlled shear rate
CSS	Controlled shear stress
FS	Flow curve
HS	High shear
LS	Low shear
LVER	Linear viscoelastic range
PVC	Pigment volume concentration
TS	Time sweep
VS	Volume solids
XLFF	Extra low frequency

Symbol

Description

E	Expected value operator
M	Hidden layer neuron count (-)
N	Size of exemplar training set (-)
p	Number of components (equation 3.1); number of input nodes (equation 3.2) (-)
R	Cone and plate radius (m)
t	(s)
x_i	Value of variable i
x_j	Value of variable j
x_k	Value of variable k
y	Random variable

Greek Symbol	Description
α	Cone angle (°)
b_i	Regression coefficient associated with the variable i
b_{ij}	Regression coefficient associated with the interaction term of variables i and j
b_{ijk}	Regression coefficient associated with the interaction term of variables i, j and k
g	Deformation (%) or strain (1)
g_A	Strain (%) or deformation (1) amplitude
g_{LVER}	Strain at limit of the linear viscoelastic range (%) or (1)
$\dot{g}$	Shear rate (1/s)
δ	Phase shift, loss angle (°)
e	Error of approximation
τ	Shear stress (Pa)
τ_A	Shear stress amplitude (Pa)
ω	Angular frequency (rad/s)

3.7 REFERENCES

- Coit, D. W., Jackson, B. T., Smith, A. E., (1998). "Static neural network process models: considerations and cases." Int. J. Prod. Res **36**(11): 2953-2967.
- Montgomery, D. C. (2005). Design and Analysis of Experiments, John Wiley and Sons Inc.
- Coit, D. W., Smith, A. E. (1995). Using Designed Experiments to Produce Robust Neural Network Models of Manufacturing Processes. Proceedings of the Fourth Industrial Engineering Research Conference (IERC) May 1995.
- Anderson, M. J., Whitcomb, P.J. Computer-Aided Design of Experiments for Formulations, p 1-8. Mineapolis, Stat-Ease Manual, Inc.
- Gupta, A. (2001). "Optimization of product performance of a paint formulation using mixture experiment." Journal of Applied statistics **28**(2): 199-213.
- Trygg, J., Wold, S. (2002). Introduction to statistical experimental design. Homepage of Chemometrics, August 2002.
- Bourquin, J., Schmidli, H., Van Hoogevest, P., Leuenberger, H. (1998). "Pitfalls of artificial neural networks (ANN) modelling technique for data sets containing outlier measurements using a study on mixture properties of direct compressed dosage form." European Journal of Pharmaceutical Sciences **7**: 17-28.
- Gasperlin, M., Tusar, L., Smid-Korbar, J., Kristl, J. "Viscous behaviour of pharmaceutical creams: characterisation by neural network and polynomial modelling." p. 14-20. University of Ljubljana, Faculty of Pharmacy, Slovenia.

9. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Zupan, J., Kristl, J. (2000). "Viscosity prediction of lipophilic semisolid emulsion systems by neural network modelling." International Journal of Pharmaceutics **196**: 37-50.
10. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Kristl, J. (1998). "Lipophilic semisolid emulsion systems: viscoelastic behaviour and prediction of physical stability by neural network modelling." International Journal of Pharmaceutics **168**: 243-254.
11. Design Expert Manual (2004), Stat-Ease
12. Gasperlin, M. (1995). "Empirical modelling of rheological and mechanical properties of paint." Rheology: 140.
13. Leskovsek, N. Tusar, M., Tusar, L. "Study of rheological properties of paints with designing of experiments: a neural network approach." p. 709. Supplied by The British Library.
14. Leskovsek, N. Tusar, M., Tusar, L., Zupan, J. (1994). "Surface coating formulation using neural networks and statistical methods (Polish)." **28**(1).
15. Aldrich, C. (1997). Neural Networks for the Process Industries. Unpublished Manuscript, Stellenbosch.
16. White, H. (1989). Learning. San Diego, Department of Economics, University of California: 425-464.
17. Haykins, S. (1994). Neural Networks: A Comprehensive Foundation, Chapter 6. New York, Macmillan College Publishing Company.
18. Plumb, A. P., Rowe, R. C, York, P., Doherty, C. (2002). "The effect of experimental design on the modelling of a tablet coating formulation using artificial neural networks." European Journal of Pharmaceutical Sciences **16**: 281-288.
19. Wu, W., Walczack, B., Massart, D. L., Heurding, S., Erni, F., Last, I. R., Prebble, K. A. (1996). "Artificial neural networks in classification of NIR spectral data: Design of the training set." Chemometrics and Intelligent Laboratory Systems **33**: 35-46.
20. Schaller, J. E. Introduction to latex paints, p.1-57. Rohm and Haas Company.
21. Prideaux, J. "Rheology modifiers and thickeners in aqueous paints." Journal of the Oil and Colour Chemists' Association, ISSN 0030-1337. Rohm and Haas.
22. Eley, R. R. (2001). Applied Rheology, ICI Paints. Strongsville, OH 44136
23. Mezger, T. (2002). The Rheology Handbook, Chapter 1-6. Hannover, Vincent Verlag.
24. Huntsman Tioxide, Tioxide TR94, Technical Data Sheet
25. Rohm and Haas, (1999). Ropaque Ultra, Opaque Pigment Slurry, Technical Data Sheet
26. Huntsman Tioxide, Tioxide R-CR40, Technical Data Sheet
27. DOW Surfactants, Triton-X100 Nonionic Surfactant, Technical Data Sheet
28. Talc de Luzenac, Steabright, Technical Data Sheet
29. Barloworld Plascon SA, AE 846:Emulsion, Product Data
30. Barloworld Plascon SA, Dispersant Evaluation, Interim Report
31. Rohm and Haas, (1996). Orotan 731-K, Versatile Pigment Dispersant, Technical Data Sheet
32. Arkem (PTY) Ltd, Alcosperse, Technical Bulletin
33. Acti-Chem SA (PTY) Ltd, Acticide HF, Product Information
34. Rohm and Haas, (1996). Acrysol RM-8W Rheology Modifier

35. Rohm and Haas, (1996). Acrysol RM-2020 Rheology Modifier
36. Parsons, P. (1993). Surface Coatings: Raw materials and Their Usage, Chapter 26. London, Chapman and Hall.

4 DATA PREPARATION

4.1 INTRODUCTION

It is the nature of empirical data to be noisy, poorly distributed and spotty. These issues, if not properly addressed, can prevent even the most sophisticated and robust methods from yielding stable models and reliable results. Many strategies and techniques are useful in gathering, cleaning up and adding value to data. However, an important point to remember is that odd patterns in data do not necessarily indicate that the data should be modified. These patterns may point to important relationships within the data and may in fact be a resource to the modeller. This chapter has a closer look at these patterns and its aim is to gather as much information about the system as possible before the actual modelling begins. However, even basic modelling techniques are required at this point in order to learn more about the structure of the data.

Figure 4.1 shows the interactions between various generated sets of data and the models associated with each.

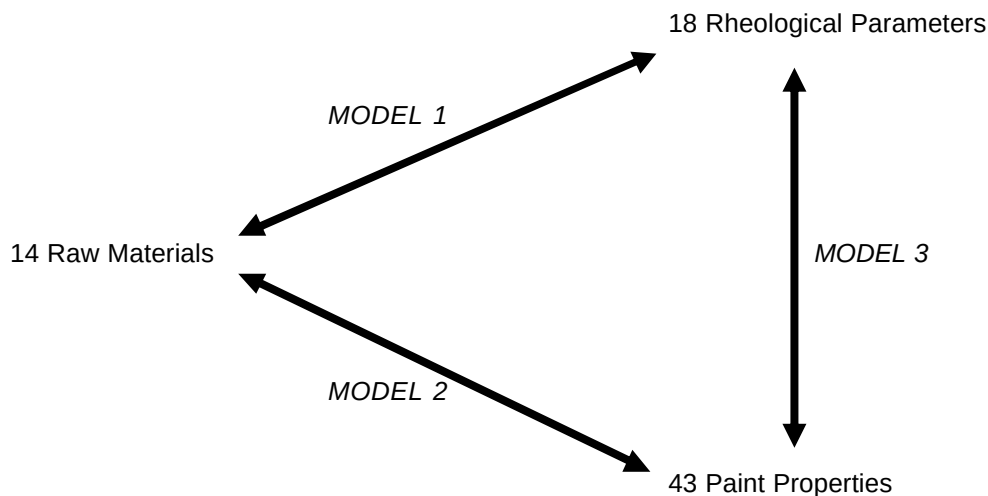


Figure 4.1 Modelling Summary

As mentioned before, the aim of this study is to develop 3 models shown in Figure 4.1. They are:

- MODEL 1: *Rheological Behaviour* = f (*Raw Materials*)
- MODEL 2: *Paint Properties* = f (*Raw Materials*)
- MODEL 3: *Paint Properties* = f (*Rheology Data*)

To be able to develop these three models a lot of data is needed. This not only requires a careful design of the experiments (DOE), but also extensive data preparation before the actual modelling process can begin. Figure 4.2 illustrates the process of data preparation that precedes the modelling stage.

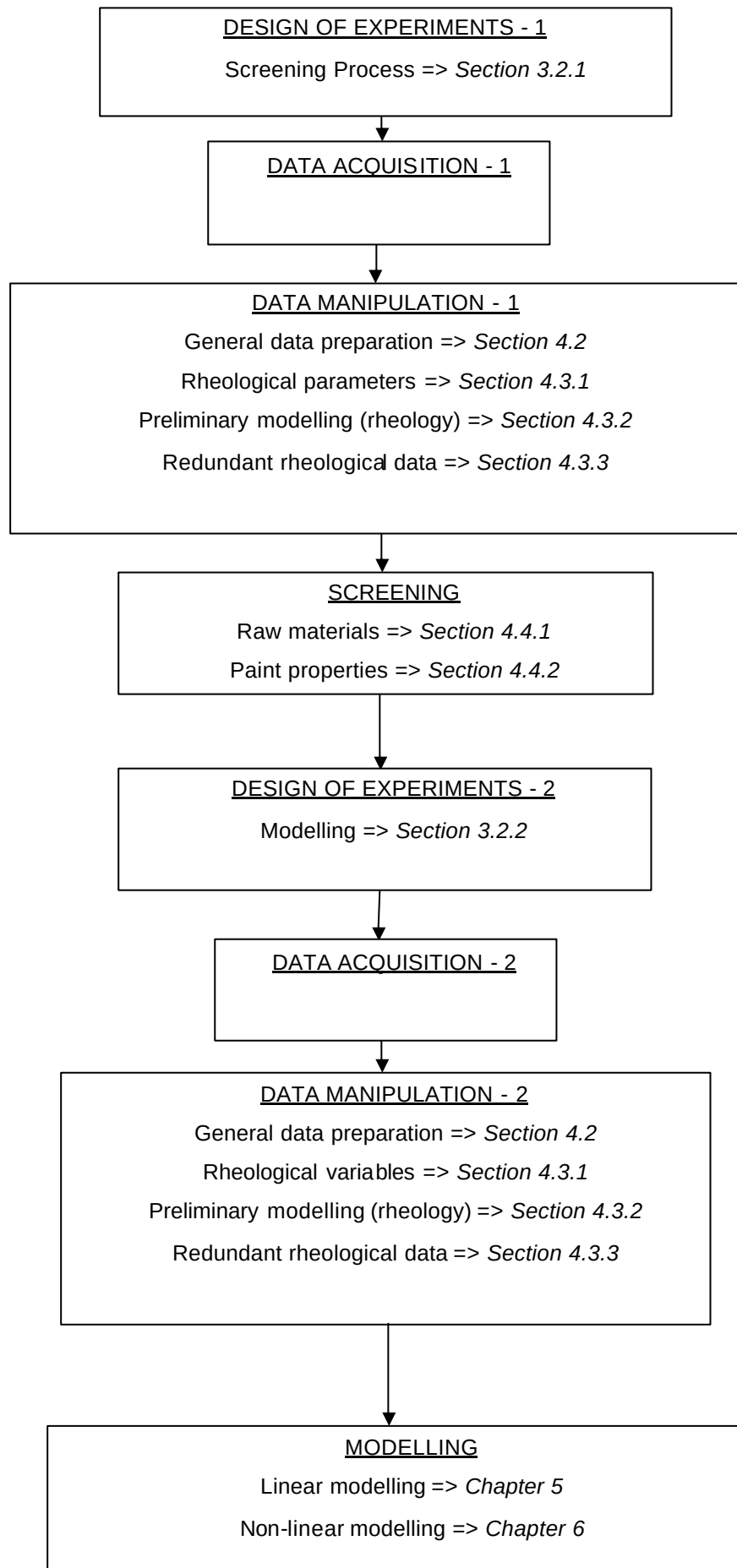


Figure 4.2 Flow Diagram of Modelling Process

4.2 GENERAL DATA PREPARATION

This type of data preparation is done routinely on any data set and includes:

- Removing outliers

Anomalous outliers can be one of the most disruptive influences on a quantitative model. An outlier is an extreme data point that may have an undue influence on a model and should be removed. Outliers are often, but not always, caused by erroneous data cases. Ninety-five percent of normally distributed data lies within two standard deviations of the mean [1]. Disregarding values outside that range is a simple method for removing outliers, which due to their abnormal size can have a serious effect on the model. However, not all outliers are the result of erroneous data. The outlier could be indicating important information – perhaps the process is not linear. In this study outliers are not discarded without careful examination. The presence of outliers plays a very important role in non-linear modelling techniques such as artificial neural networks (ANN) [2]. Due to the extreme flexibility of artificial neural networks, the presence of an outlier cannot be detected. Contrary, the outlier gets integrated into the artificial neural network, resulting in unreliable predictions. This is a clear weakness of ANN and therefore possible outliers are handled with great care.

Recall that a further 30 data points with higher pigment volume concentration (PVC) and volume solids (VS) were added to the already existing 134 data points so that a total of 164 paint samples had to be manufactured. These paint samples are added with the specific intention to add to the information content of the data set, because it is also known that VS and PVC can change the rheological character and paint properties abruptly. Therefore data points (rheological and paint properties) with values > 2 standard deviations were not considered to be outliers *only* if they were one of the 30 extra data points with extra high/low PVC and solids. These data points were kept with the hope that they would add extra information to the non-linear nature of the problem. An illustration of possible outliers for a rheological parameter is given in Figure 4.3.

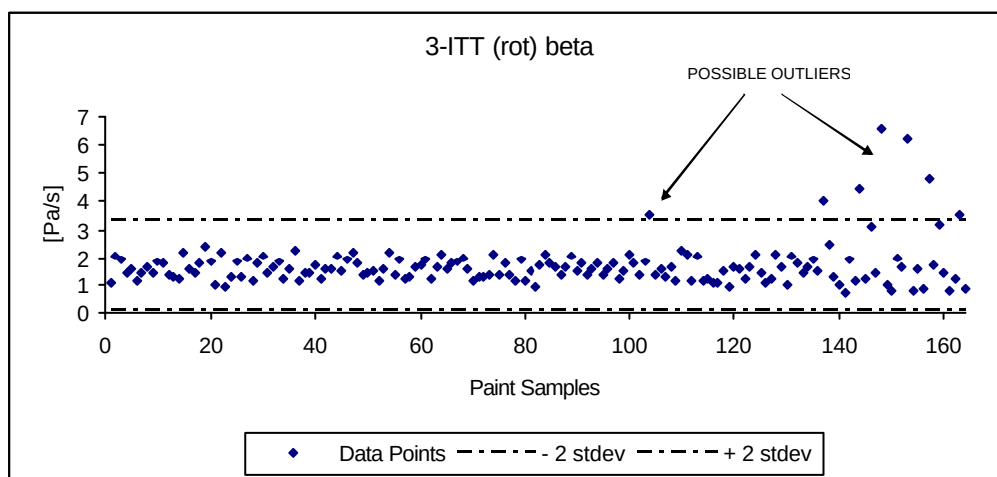


Figure 4.3 Data Points with Possible Outliers of the 3-ITT (rot) Beta Parameter

Most of the measured data points for this rheological parameter lie within 2 standard deviations from the average, except for seven of these data points, as can be seen in Figure

4.3. These seven data points, according to the definition given in the previous paragraph, are possible outliers. However, distinction needs to be made between possible outliers which can not add to the information content and data points which can. In this case, only paint sample #104 is considered to be an outlier and not the others (paint samples #137, #144, #146, #148, #157 and #163). These samples are not considered to be outliers, even though their values are larger than 2 times the standard deviations, because of their extreme high/low PVC and VS which may lead to abnormal rheological behaviour. The value for 3-ITT (rot) β for paint sample #104 is replaced with the average value of the 3-ITT (rot) β . Outliers in paint properties are removed on the same basis as described above.

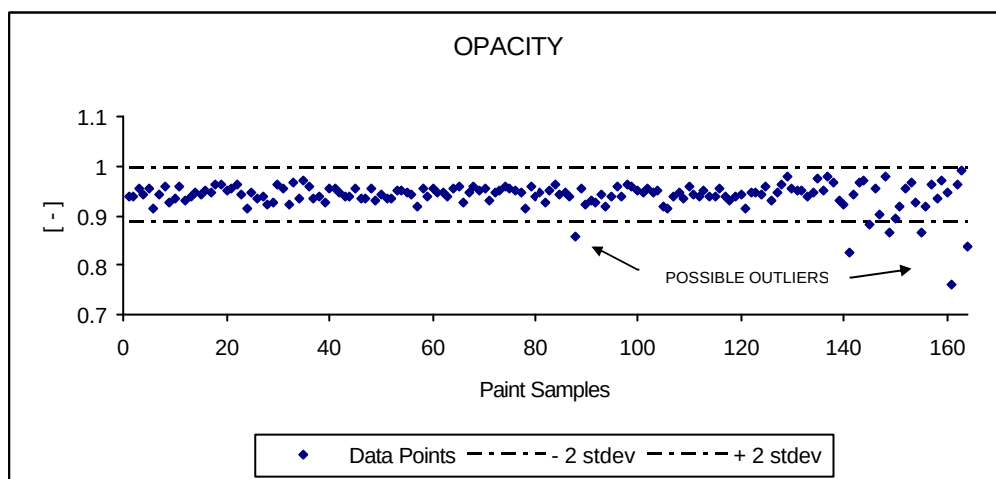


Figure 4.4 Data Points with Possible Outliers in Opacity Data

Once again, note all the possible outliers in Figure 4.4. Only paint sample #88 is considered to be an outlier as the others are samples with extreme high/low PVC and VS. The opacity value for paint sample #88 is replaced with the average opacity value.

Figure 4.4 shows that the degree of variance in the opacity values is below expected. The aspect of opacity has been the subject of much theorising and speculation as stated by Parsons [18]: "In common with much of paint technology, theory trails practical application and can only partially explain observed phenomena, but as predictive tool, it is overwhelmed by the complexities of real life and humbled by the simplicities of direct experimentation". It is therefore difficult to give the exact reason for the low level of variance experienced for this specific paint. However, the opacity of paint can be related to its refractive index. Rutile pigment (TiO_2) contributes the most to the refractive index (RI) of the paint. Other factors affecting opacity are the size of the pigment particles, their distribution and the degree of dispersion achieved. It is important to note that the only relevant size (mean, distribution and spacing) consideration for the pigment particles is in the final dry coating. A dry paint film is a three-dimensional structure and the attributes depend directly upon the volume proportions of the major ingredients. A widely accepted way of viewing pigment content is the pigment volume concentration (PVC) of the dry paint film. The low variance in opacity, even though the pigment is varied by $\pm 20\%$ by mass, is possibly due to the fact that the standard paint formulation has a high PVC and therefore changes in the level of pigment do not have a

significant effect on the opacity. The opacity is therefore insensitive to changes in pigment concentration.

In total, eight data points were considered to be outliers, which is a small amount considering the number of measurements (3618 measurements (2412 paint property measurements and 1206 rheological measurements)).

- Standardising

Various advantages exist in standardising the data:

- One of the first decisions to be taken before the application of principal component analysis (see Section 4.3.3 Redundant Rheological Data), concerns the data on which the analysis are to be performed. The presence of considerable differences among the variables (e.g. variables representing vastly different entities, or having different units or scales of measurement, etc.) result in variables with large variances. If the data is not standardised, the principle components will simply correspond to the original variables with the largest variance and little will be learnt from the data.
- Neural networks perform better when data is standardised. Since correlations are unaffected by scaling data, it is both sufficient and convenient to deal with standardised variables. A standardised variable is obtained by subtracting the average value for each observation and dividing by the standard deviation of the observations:

$$X_{\text{standardised}} = \frac{X_i - X_{\text{avg}}}{X_{\text{standard deviation}}} \quad (4.1)$$

The average value of a standardised variable is zero. The variance is equal to 1 and the covariance between standardised variables is the same as the correlations of the unstandardised data.

- Scaling

This is accomplished by mapping the minimum and maximum values of the actual input and output data linearly to the respective minimum and maximum values of the network ranges. This is once again important for data used in neural networks. Before data can be presented to a neural network, it has to be scaled to within ranges that the neural network can accommodate. The sigmoidal activation function produces outputs in the range of -1 to +1.

4.3 RHEOLOGICAL DATA PREPARATION

4.3.1 THE RHEOLOGICAL PARAMETERS

Viscosity is undoubtedly one of the most important rheological parameters characterising paints [3]. This is mostly due to the historical perspective - it was easy to characterise the flow behaviour of paint with a single viscosity measurement. However, paints are viscoelastic materials,

consisting of both viscous and elastic characteristics and therefore it is more appropriate to characterise/model a paint considering both of these properties. Using only the viscosity value as a model input is no longer sufficient and other researchers have also started to model viscoelastic materials based on the viscous/elastic properties of the material.

Knowledge of the shear stress τ , the deformation γ and the phase shift δ , allows one to *completely* describe the viscoelastic properties of a material [4]. In most rheological analyses, these parameters are manipulated into new parameters which are more understandable, e.g., the storage modulus (elastic part):

$$G' = \left(\frac{t_A}{g_A} \right) \cdot \cos d. \quad (4.2)$$

This is a measure of the material's elasticity (where t_A and g_A are the shear stress and the shear strain amplitudes, respectively). The loss modulus:

$$G'' = \left(\frac{t_A}{g_A} \right) \cdot \sin d \quad (4.3)$$

is a measure of the material's energy dissipation (viscous part). The limiting cases of $\delta = 0^\circ$ and $\delta = 90^\circ$ correspond to purely elastic and purely viscous behaviour, respectively. These moduli are components of the complex shear modulus, which is another useful parameter that is often used to express paint behaviour and is defined as:

$$G^* = G' + iG''. \quad (4.4)$$

Researchers such as Gasperlin et al [5] and Leskovsek et al [6] used the viscous (G'') and the elastic (G') part of the complex shear modulus for modelling parameters.

Gasperlin et al [7] also used the ratio:

$$\frac{G''}{G'} = \tan d \quad (4.5)$$

which is a measure of the ratio of elastic to viscous parts of the sample, as a model parameter for modelling the rheological response. From the complex shear modulus (G^*) one can calculate the complex viscosity :

$$|h^*| = \frac{|G^*|}{\omega} \quad (4.6)$$

which is basically the viscosity determined in the controlled shear deformation (CSD) mode in the linear viscoelastic region. Gasperlin et al [3] also used the complex viscosity as a model parameter.

The rheological parameters in equation 4.2 to equation 4.6 are all very useful because they all make sense. They are all manipulations of the fundamental rheological properties which are not always so intuitively understandable, but which could also be used for modelling purposes. Fundamentally, the flow behaviour can be completely described by the shear stress τ (as a function of shear rate), while knowledge of the shear stress (τ), the deformation ($\dot{\gamma}$) and the phase shift (δ), allows one to *completely* describe the viscoelastic properties of a material [4]. See Table 4.1 for the fundamental rheological parameters

Table 4.1 Fundamental Rheological Parameters

	Test Presetting	Result
Flow Properties	$\dot{\gamma}$	τ
Viscoelastic Properties	$g(t) = g_A \cdot \sin(\omega t)$	$t(t) = t_A \cdot \sin(\omega t + d)$

The model parameters are therefore the shear stress (τ) from the rotational tests (flow behaviour) and a combined parameter ($t_A \cdot d$) from the oscillatory tests (viscoelastic behaviour).

4.3.2 PRELIMINARY MODELLING

It is important to measure the response of paint over as wide a range as possible in order to obtain a true picture of the rheology profile. A single measurement is insufficient and can be misleading [8]. In the past a single viscosity measurement (ICI viscosity/Krebs viscosity) was used to evaluate the flow behaviour of paint. However, it became apparent that even though two paints may display the same Krebs viscosity (Figure 4.5), they behave different in terms of certain other flow properties. The reason for this is that paint properties need to be evaluated at different shear rates on the flow curve, e.g. pigment settling at low shear rates, mixing abilities at medium shear rates and application properties at high shear rates. Modern rheological measuring instruments make it possible to measure over an entire shear rate / shear deformation range.

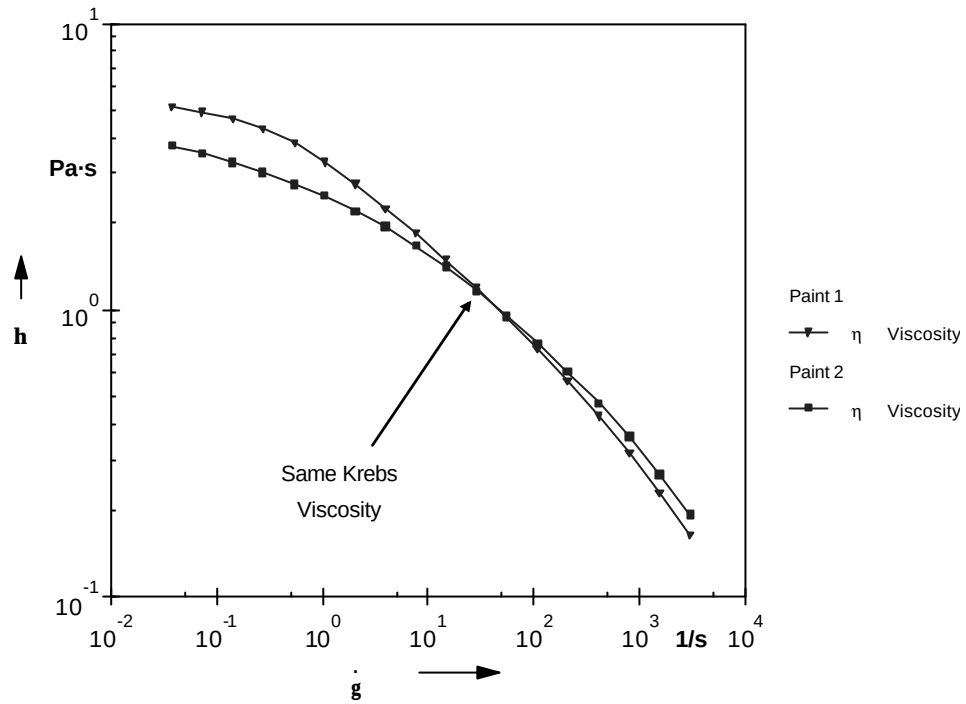


Figure 4.5 Viscosity Curve of Two Paints With the Same Krebs Viscosity [9]

Preliminary modelling is used as a technique to express the entire range of rheological data simply by fitting a trendline to the data and then using the parameters of the trendline for further modelling purposes. This technique has been used with success by other researchers to model the complex viscosity of paint over a frequency range [3], [5]. No references have been found where preliminary modelling is applied to other rheological curves (besides the frequency sweep), but in principle it is also possible and carried out in this study. Preliminary modelling is therefore performed on all the rheological curves (except the 'extra low frequency', or XLF measurement, which is a single point) in the following manner.

- 1) Calculate the natural logarithm of the rheological parameter (τ for CSR tests and $\tau \delta$ for CSD tests) and fit the data with a trendline.
- 2) Obtain the parameters α (for the slope of the curve) and β (for the offset of the curve)

Figure 4.6 illustrates a flow curve (CSR test) with

$$y = \beta x^\alpha \quad (4.7)$$

as the fitted trendline and

$$\alpha = 0.7336$$

$$\beta = 2.9354$$

for the trendline parameters.

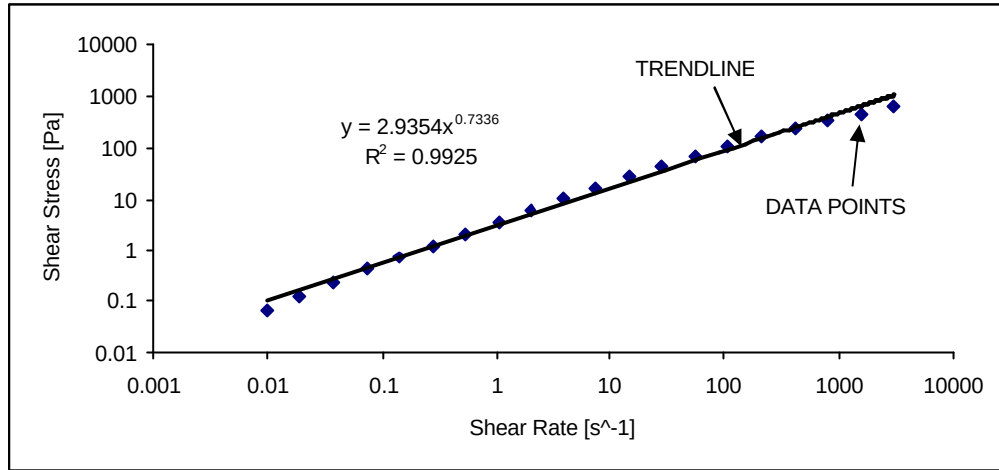


Figure 4.6 Preliminary Modelling for a Flow Curve (CSR test)

Previously preliminary modelling has only been applied to data obtained from the frequency sweep [3], [5]. In this study, the method of preliminary modelling for the extraction of α and β values are extended to all the rheological curves. The method of preliminary modelling is applied to tests performed in the CSR mode, as well as to tests that are performed in the CSD mode. Figure 4.7 illustrates an example for the 3-ITT (OSC) test with

$$y = \beta x^\alpha \quad (4.8)$$

as the fitted trendline and

$$\alpha = 0.1554$$

$$\beta = 1.886$$

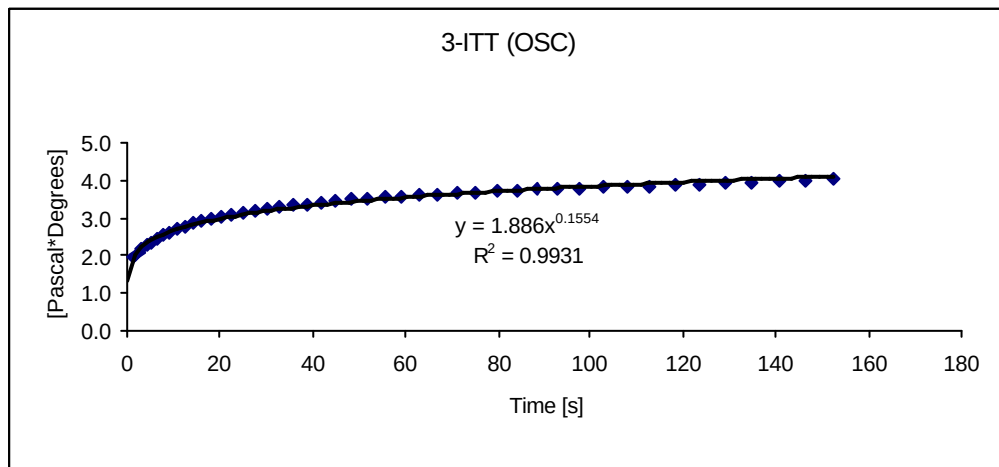


Figure 4.7 Preliminary Modelling for the 3-ITT (oscillatory, CSD test)

The α and β values of all the rheological tests are combined to form *A* and *B* matrices of α and β values respectively. The *A* and *B* matrices are used for further modelling purposes.

The R^2 values in Figure 4.6 and Figure 4.7 indicate how good the fit of the trendlines to the natural logarithm of the data is.

Table 4.2 R^2 Values for Preliminary Modelling (Screening and Modelling Data)

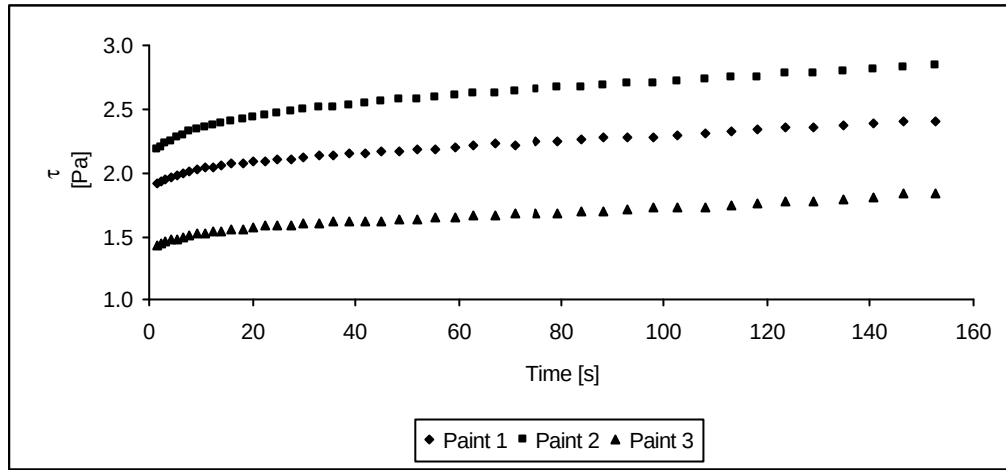
Test Type	R^2_{avg} (screening data)	R^2_{avg} (modelling data)
	29 runs	164 runs
Flow Curve (FC)	0.99	0.99
High Shear (HS)	0.86	0.97
Low Shear (LS)	0.94	0.86
3-ITT (rot)	0.98	0.97
Amplitude Sweep (AS)	0.99	0.98
Frequency Sweep (FS)	0.99	0.99
Time Sweep (TS)	0.99	0.98
3-ITT (osc)	0.95	0.94

The high R^2_{avg} values in Table 4.2 indicate that the trendlines represent the data well and that the α and β parameters obtained from the preliminary modelling can be used with confidence in further modelling steps. These parameters for the preliminary modelling are presented in Appendix D (for the screening rheological data) and Appendix E (for the modelling rheological data). Both of these appendices are in electronic format on the CD at the end of the dissertation.

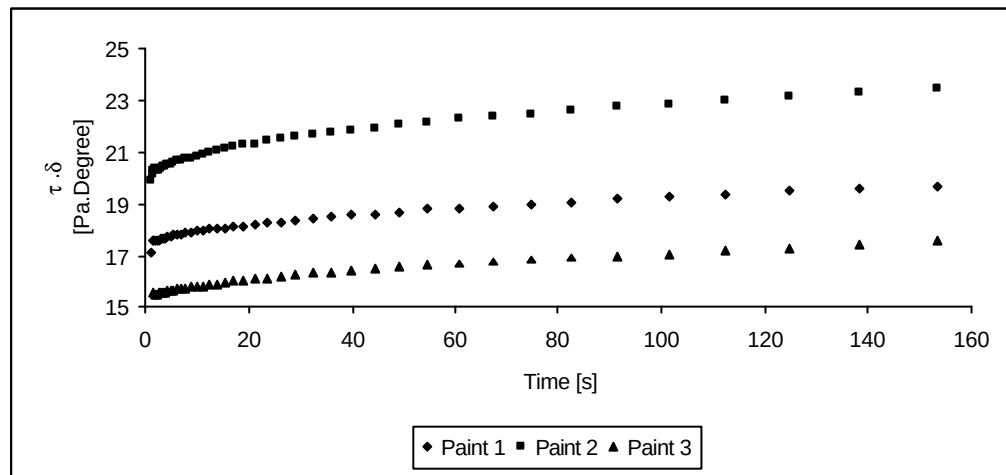
4.3.3 REDUNDANT RHEOLOGICAL DATA

4.3.3.1 BACKGROUND

Nine rheological tests are performed on each paint sample (see Table 3.5). Some of these tests are related to each other and therefore the rheological data will also be correlated to a certain degree and many of the data obtained from these tests will be redundant. A good example of this is the 3-ITT tests (rotational and oscillatory). These tests are performed in the controlled shear rate (CSR) test mode as well as in the controlled shear deformation (CSD) test mode, which are rotational and oscillatory tests respectively. Figure 4.8 shows the thixotropic behaviour of three different paints measured with both the 3-ITT tests.



(a)



(b)

Figure 4.8 3-ITT Tests: (a) Rotational (CSR) and (b) Oscillatory (CSD)

Figure 4.8 (a) and (b) almost look identical in terms of the shape and the relative placement of the curves. There is a definite correlation in the measurements between the two different test types (rotational and oscillatory). The correlation can be quantitatively observed by the correlation coefficients of the preliminary model parameters, α (gradient of preliminary model) and β (offset of preliminary model) of the rheological measurements performed on 164 paint samples. This is given in Table 4.3.

Table 4.3 Correlation Coefficients of 3-ITT Test Parameters (Rotational and Oscillatory)

	$a_{3-ITT(rot)}$	$b_{3-ITT(rot)}$
$a_{3-ITT(osc)}$	0.88	0.16
$b_{3-ITT(osc)}$	-0.07	0.73

The thixotropic behaviour obtained from these test types are highly correlated (high correlation coefficients between $\alpha_{\text{oscillatory}}$ and $\alpha_{\text{rotational}}$ as well as $\beta_{\text{oscillatory}}$ and $\beta_{\text{rotational}}$). This seems intuitive because the measurements are basically of the same type and it was just the test mode (oscillatory vs. rotational) that differed. The high level of multicollinearity (see Chapter 8 - Appendix A: Glossary and Terms) is therefore obvious. However, not all collinear effects are that obvious and statistical techniques are required for the evaluation of multicollinearity.

4.3.3.2 MULTICOLLINEARITY IN THE RHEOLOGICAL DATA

Table 4.4 illustrates the level of correlation between the 18 rheological variables of the modelling data (164 paint samples).

Table 4.4 Correlation Coefficients of Rheological Parameters

	RP1	RP2	RP3	RP4	RP5	RP6	RP7	RP8	RP9	RP10	RP11	RP12	RP13	RP14	RP15	RP16	RP17	RP18
RP1	1.00																	
RP2	-0.07	1.00																
RP3	0.73	0.16	1.00															
RP4	-0.07	0.88	-0.04	1.00														
RP5	0.75	0.24	0.89	0.16	1.00													
RP6	0.03	-0.10	0.02	-0.10	-0.03	1.00												
RP7	0.64	0.30	0.77	0.26	0.82	0.01	1.00											
RP8	-0.65	-0.51	-0.80	-0.42	-0.86	0.05	-0.77	1.00										
RP9	0.57	0.47	0.80	0.36	0.86	0.00	0.90	-0.83	1.00									
RP10	-0.43	-0.49	-0.62	-0.39	-0.70	0.04	-0.58	0.76	-0.81	1.00								
RP11	0.61	0.51	0.81	0.42	0.86	-0.01	0.79	-0.87	0.87	-0.71	1.00							
RP12	-0.60	-0.48	-0.81	-0.40	-0.85	-0.01	-0.85	0.87	-0.90	0.69	-0.96	1.00						
RP13	0.67	0.38	0.78	0.29	0.86	-0.02	0.63	-0.78	0.73	-0.62	0.84	-0.77	1.00					
RP14	0.13	-0.05	0.15	-0.03	0.14	0.01	0.17	-0.16	0.15	-0.12	0.13	-0.16	0.12	1.00				
RP15	-0.20	0.08	-0.12	0.01	-0.14	-0.01	-0.09	0.12	-0.07	0.07	-0.14	0.13	-0.12	-0.17	1.00			
RP16	0.47	0.61	0.65	0.50	0.67	-0.01	0.83	-0.74	0.88	-0.61	0.78	-0.80	0.62	0.09	-0.01	1.00		
RP17	0.83	0.14	0.75	0.13	0.79	0.00	0.68	-0.71	0.69	-0.52	0.79	-0.79	0.70	0.10	-0.24	0.58	1.00	
RP18	0.48	0.29	0.60	0.23	0.62	0.03	0.70	-0.60	0.72	-0.47	0.64	-0.68	0.50	0.13	-0.09	0.69	0.58	1.00

$$RP1= b_{3-ITT(osc)} \text{ of } \ln y_{3-ITT(osc)} = a_{3-ITT(osc)} \ln x_{3-ITT(osc)} + b_{3-ITT(osc)} \quad (4.9)$$

$$RP10= a_{FC} \text{ of } \ln y_{FC} = a_{FC} \ln x_{FC} + b_{FC} \quad (4.13)$$

$$RP2= a_{3-ITT(osc)} \text{ of } \ln y_{3-ITT(osc)} = a_{3-ITT(osc)} \ln x_{3-ITT(osc)} + b_{3-ITT(osc)} \quad (4.9)$$

$$RP11= a_{HS} \text{ of } y_{HS} = a_{HS} x_{HS}^2 + b_{HS} x_{HS} + g_{HS} \quad (4.14)$$

$$RP3= b_{3-ITT(rot)} \text{ of } \ln y_{3-ITT(rot)} = a_{3-ITT(rot)} \ln x_{3-ITT(rot)} + b_{3-ITT(rot)} \quad (4.10)$$

$$RP12= b_{HS} \text{ of } y_{HS} = a_{HS} x_{HS}^2 + b_{HS} x_{HS} + g_{HS} \quad (4.14)$$

$$RP4= a_{3-ITT(rot)} \text{ of } \ln y_{3-ITT(rot)} = a_{3-ITT(rot)} \ln x_{3-ITT(rot)} + b_{3-ITT(rot)} \quad (4.10)$$

$$RP13= g_{HS} \text{ of } y_{HS} = a_{HS} x_{HS}^2 + b_{HS} x_{HS} + g_{HS} \quad (4.14)$$

$$RP5= b_{AS} \text{ of } \ln y_{AS} = a_{AS} \ln x_{AS} + b_{AS} \quad (4.11)$$

$$RP14= b_{LS} \text{ of } \ln y_{LS} = a_{LS} \ln x_{LS} + b_{LS} \quad (4.15)$$

$$RP6= a_{AS} \text{ of } \ln y_{AS} = a_{AS} \ln x_{AS} + b_{AS} \quad (4.11)$$

$$RP15= a_{LS} \text{ of } \ln y_{LS} = a_{LS} \ln x_{LS} + b_{LS} \quad (4.15)$$

$$RP7= b_{FS} \text{ of } \ln y_{FS} = a_{FS} \ln x_{FS} + b_{FS} \quad (4.12)$$

$$RP16= a_{TS} \text{ of } y_{TS} = a_{TS} x_{TS} + b_{TS} \quad (4.16)$$

$$RP8= a_{FS} \text{ of } \ln y_{FS} = a_{FS} \ln x_{FS} + b_{FS} \quad (4.12)$$

$$RP17= b_{TS} \text{ of } y_{TS} = a_{TS} x_{TS} + b_{TS} \quad (4.16)$$

$$RP9= b_{FC} \text{ of } \ln y_{FC} = a_{FC} \ln x_{FC} + b_{FC} \quad (4.13)$$

$$RP18= C_{XLF} \text{ of } y_{XLF} = C_{XLF} \quad (4.17)$$

A high level of pairwise correlation (values > 0.8) is observed in the correlation matrix. The advantage of the correlation matrix is that it also identifies some of the correlations between variables that are not that intuitive. However, multicollinearity (multiple correlations) can exist without being identified by the correlation matrix (pairwise correlation coefficients) and the statistical method of computing the *condition number*, C, is used to determine the level of multicollinearity.

$$C = \sqrt{\frac{\text{maximum eigenvalue of the correlation matrix}}{\text{minimum eigenvalue of the correlation matrix}}} \quad (4.18)$$

Large values of C indicate evidence of strong collinearity. The harmful effects of collinearity become strong when the values of the condition number, C, exceeds 15 [10].

Table 4.5 Level of Multicollinearity in Rheological Data

	Rheology Data (Screening)	Rheology Data (Modelling)
Condition Number C [-]	4.48×10^6	2.67×10^6

It is clear from Table 4.5 that high levels of multicollinearity ($C \gg 15$) exist between the variables. Multicollinearity has a negative effect on the modelling process and should be kept at a minimum if possible. Principal component analysis is a useful way of preventing the negative effects that multicollinearity has on modelling.

4.3.3.3 PRINCIPAL COMPONENT ANALYSIS (PCA) ON THE RHEOLOGICAL DATA

Both Table 4.4 and Table 4.5 indicate that multiple correlations between the rheological parameters exist. One can therefore expect that the dimensionality of the problem could be decreased significantly. A statistical technique, principal component analysis, is used to simplify the correlated multidimensional dataset to a dataset that is i) uncorrelated and ii) lower in dimension so that further analysis is simplified and more significant. A small set of uncorrelated variables is easier to understand and work with, compared to a large set of correlated variables. Principal component analysis linearly transforms a set of variables into a substantially smaller set of uncorrelated variables containing most of the information of the original set of variables [11]. Technically speaking, PCA is a linear transformation that transforms the data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. Geometric representation of the steps in PCA is illustrated in Figure 4.9

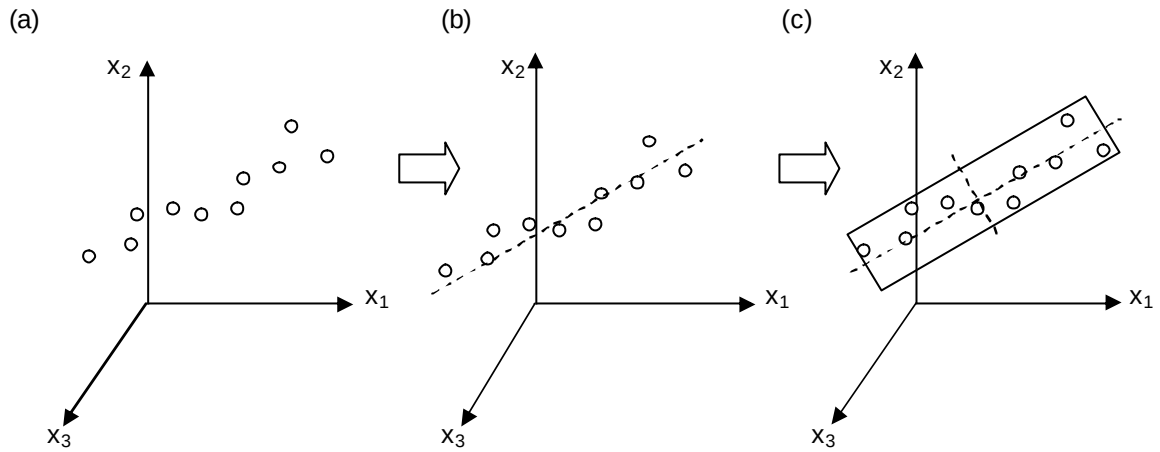


Figure 4.9 Geometric Representation of PCA: (a) Datapoints In the Observation Space, (b) 1st PC, (c) Plane Defined by First 2 PC's

PCA can be used for dimensionality reduction in a dataset while retaining those characteristics of the dataset that contribute most to its variance, by keeping lower-order principal components and ignoring higher order ones. Such low-order components often contain the “most important” aspects of the data.

Principal component analysis is applied to both the screening dataset and the modelling data set by using the Matlab[®] Statistics Toolbox.

Since variance forms the basis on which principal components are derived, it is necessary to ensure that the measured variables are at least comparable in their magnitude of variance and their units of measurements. Therefore, it is necessary to standardise the rheological data before principal component analysis is applied to the dataset.

Principal Component Analysis (Screening Data)

Principal component analysis is performed on the first set of rheological data (screening data) because of its high level of multicollinearity ($C=4.48 \times 10^6$). The Pareto plot (Figure 4.10) illustrates the relative importance of each principal component (bar graph) and the cumulative variance explained by the principal components (line graph).

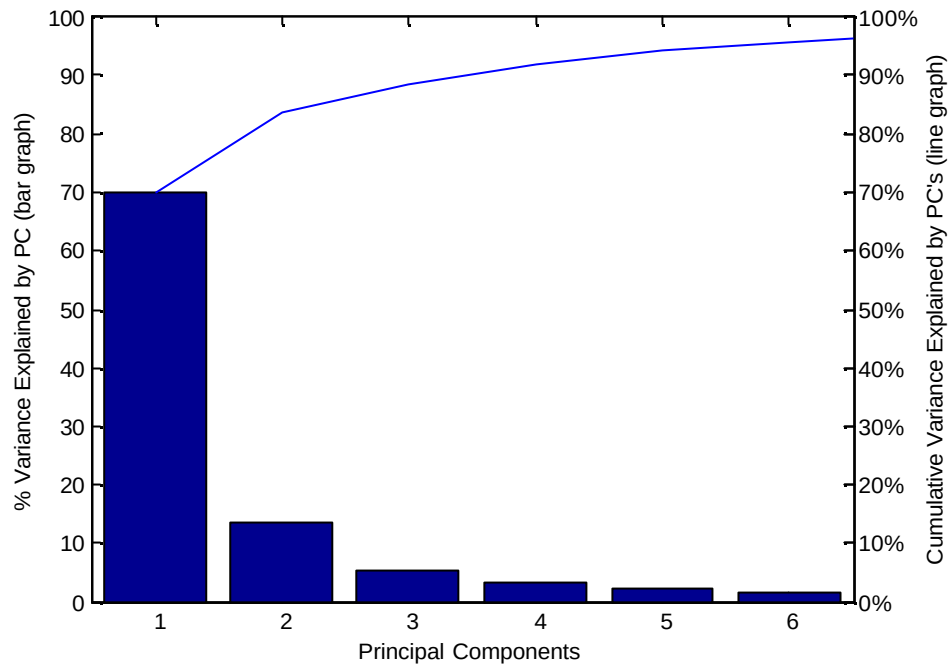


Figure 4.10 Pareto Plot for Principal Components of Rheological Screening Data

Figure 4.10 shows that the 1st six principal components (PC's) explain more than 95% of the variance in the data and the 1st two PC's explain more than 80% of the variance. The 1st principal component is very significant and explains 70% of the variance in the data. PCA has therefore reduced dimensionality of the rheological dataset significantly from 18 highly correlated variables to just a few that explain the largest part of the variation in the rheological behaviour.

The cumulative variance explained (line graph) shows an exponential growth in variance explained by each additional PC and therefore the lower order PC's explain more of the variance. There is no clear 'cut-off' point in terms of how many PC's to use for further modelling purposes but the software package suggests that extraction of the number of PC's that explain 95% of the variance is satisfactory. The cumulative variance explained by the first 6 PC's is 95.61%

A scatterplot of the actual values of the 1st two transformed variables, t1 and t2 (Figure 4.11) illustrates that they are uncorrelated, because there is no pattern in the structure of the data.

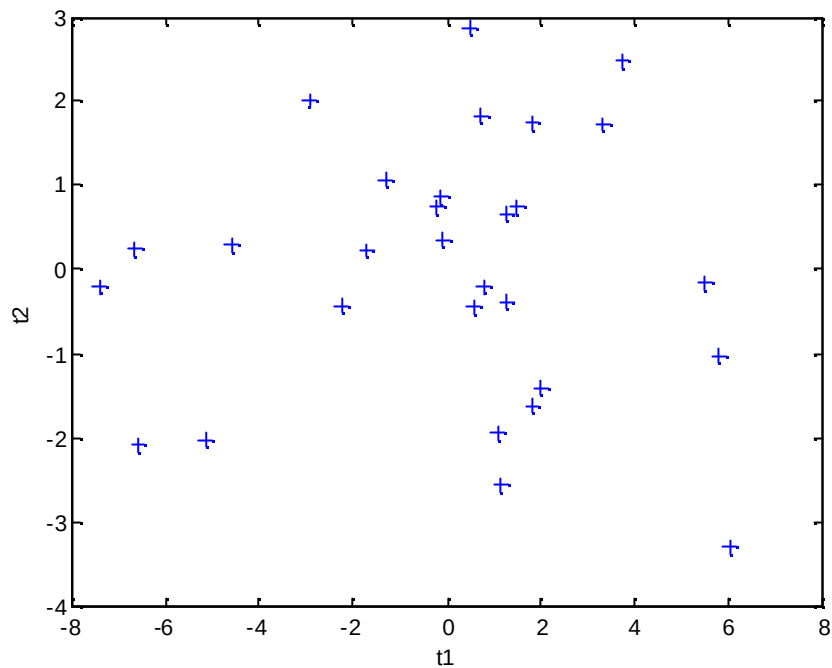


Figure 4.11 Scatterlot of 1st Two Transformed Variables

Principal Component Analysis (Modelling Data)

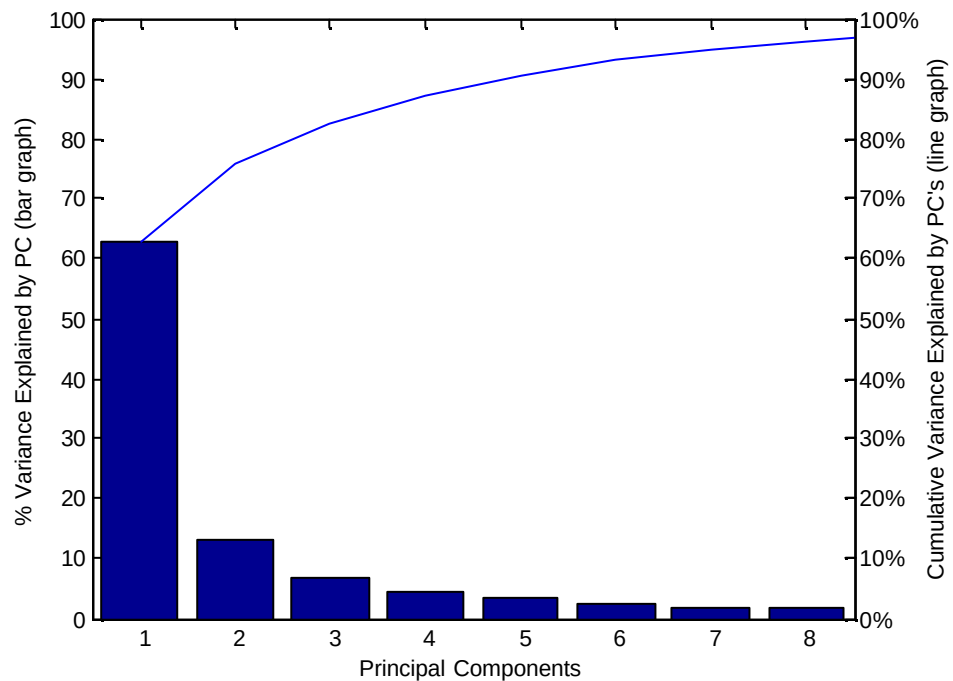


Figure 4.12 Pareto Plot for Principal Components of Rheological Modelling Data

From Figure 4.12 it is clear that principal component analysis reduces the complexity of the rheological data considerably. In this case the 1st three PC's explain more that 80% variance in

the data (compared to 2 PC's in the screening data) and 95% variance in the data is explained by the 1st 8 PC's (compared to 6 PC's in the screening data).

The modelling data therefore requires more PC's to explain the same amount of variance as the screening data. This is not surprising because the modelling data has more variance in its structure due to samples that have extreme PVS/VIS values.

A scatterplot of the values of the 1st two transformed variables, t1 and t2, indicate that there is no pattern in the structure and therefore one can conclude that they are uncorrelated.

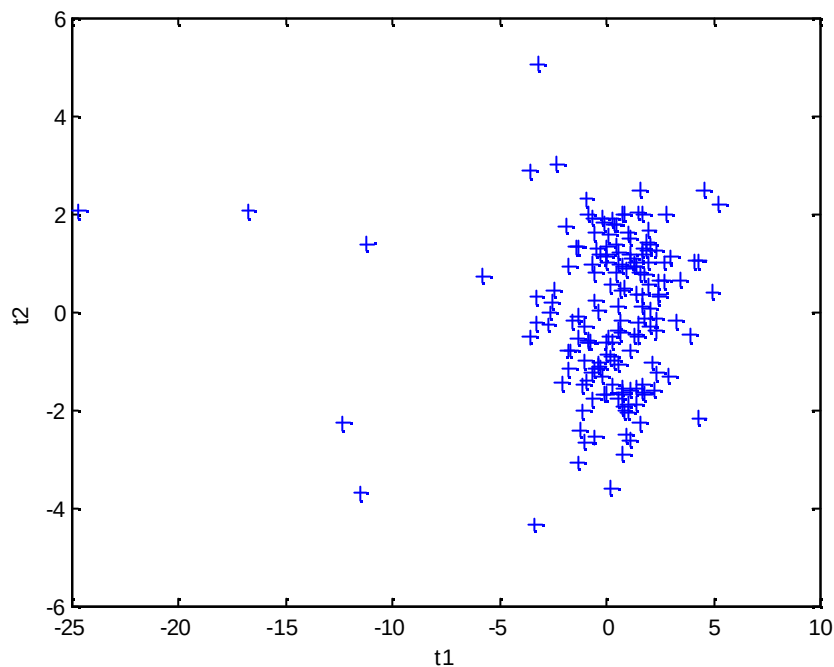


Figure 4.13 Scatterlot of 1st Two Transformed Variables

It is a difficult decision to make on how many PC's to keep for further modelling purposes. In a model one would like to explain as much variance as possible with as few explanatory variables as possible. It is therefore clear that some trade-off exists in the amount of variance that is explained by the chosen PC's, and the complexity of the modelling process.

Principal component analysis has the following advantages:

1. It ensures that the variables are uncorrelated, which is very useful when these variables are used for further modelling
2. It reduces the dimensionality of the dataset, making modelling simpler.

However, a major drawback of principal component analysis is that the principal components lack simple interpretation since each is, in a sense, a mixture of the original variables. They provide the possibility for improvements on estimation/prediction techniques but lack the characteristic of causal interpretation.

4.4 THE SCREENING PROCESS

Part of the novel contribution of this study, is to model the paint properties in terms of the rheological behaviour of the paint (MODEL 3). The rheology can not explain the variance in *all* the paint properties and it will be useless to try and model a certain paint property that is not related to the rheological behaviour of the paint. In the same manner, not *all* raw materials affect the rheology of the paint and it will be useless to include these raw materials in the model. A screening process is used to determine:

- i. The most important raw materials that influence the rheology (MODEL 1)
- ii. The paint properties that are affected most by these specific raw materials (MODEL 2)

Screening (Definition): Sifting through a large number of factors with the fewest number of trials

Effectively, the screening process determines the most important raw materials that influence the rheology. Furthermore, the screening process then determines which paint properties are affected most by these specific raw materials that were rendered most important in influencing the rheological behaviour. The thought process behind the screening is shown in Figure 4.14.

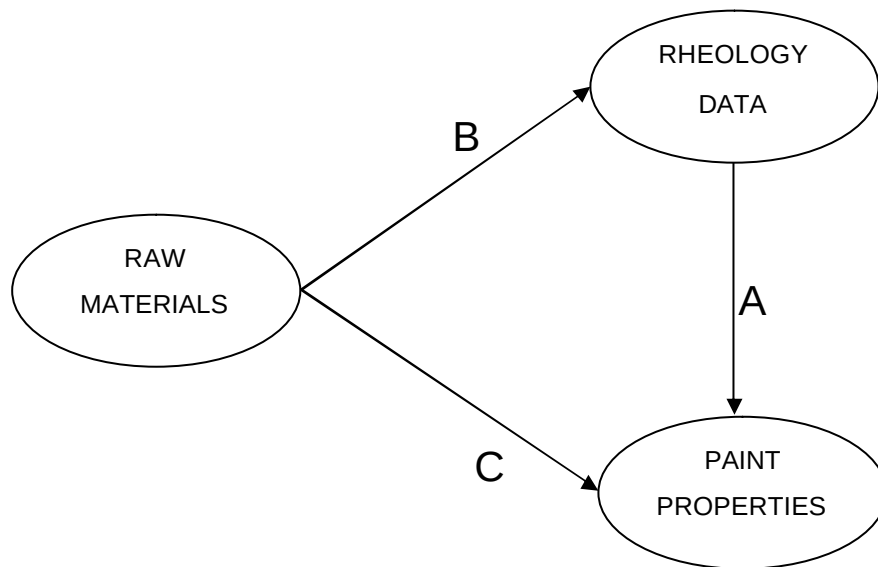


Figure 4.14 Methodology of Screening Process

The thought process in Figure 4.14 can be summarised as follows:

OBJECTIVE: Model the paint properties in terms of the rheological behaviour of paint (*path A*).

PROBLEM: Rheology do not influence all the paint properties and the set of 43 possible paint properties need to be reduced to a set that is influenced by the rheological behaviour of the paint. Part of the problem is that the paint properties are determined by the paint formulation which is controlled by the level of raw materials, which also influences the rheological behaviour.

SOLUTION: Determine the raw materials that influence the rheological behaviour (*path B*), so that the variance which occurs in the rheological data will be able to explain the variance in paint properties. Then, determine the paint properties which are affected by these specific raw materials (*path C*). It should then also be able to model the paint properties (in terms of the rheological data) that were selected by following paths B and C.

There is a strong interplay between design of experiments and the regression analysis [12] and therefore the use of special design of experiment software was required. The software package Design Expert[®] estimated that 29 runs are sufficient to determine the linear correlations between the 14 raw materials and a measured property.

The screening procedure is performed by a statistical technique called multiple linear regression (MLR). The regression coefficients of the MLR model indicate the relative importance of the regressor variables (raw materials) and are determined by the method of least squares^{sss} in the MATLAB[®] environment.

4.4.1 RAW MATERIAL SCREENING (*PATH B*)

OBJECTIVE: Identify most important raw materials that affect the rheological data.

Some raw materials play an important role in the rheology of paint (e.g. pigment and emulsion), while others play a less important role (e.g. fungicide and defoamer). No clean cut decisions can be made to the extent of influence a certain raw material has on the rheology of that paint and therefore a statistical technique such as MLR is used to determine the most influential raw materials.

4.4.1.1 MULTIPLE LINEAR REGRESSION FOR SCREENING OF RAW MATERIALS

Multiple linear regression (MLR) is used as the technique to determine the causal relation between the raw materials and the set of transformed rheological data (principal components)

$$y_i = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k + e \quad (4.19)$$

Where $y_i = \text{PC}_i$ and $x_k = \text{level of raw material } k$. Transformation of the original rheological variables into principal components (PC's) is:

- i) useful, because models are required only for the PC's that explain most of the variance in the data and
- ii) possible, because the MLR is used to obtain a causal model that explains the relative importance of the raw materials and prediction or interpretation of the response is not required.

^{sss} In the most general terms, the method of least squares aims at minimizing the sum of squared deviations of the observed values for the dependent variable from those predicted by the model.

The first 6 PC's of Figure 4.10 explain 95% of the variance in the data and therefore it would be useful to construct 6 models, one model for each PC. The importance of each model is related to the amount of variance that the corresponding PC explains in the rheological data. Therefore, the MLR model for PC1 is much more important than the MLR model for PC6, because PC1 explains 70% of the variance and PC6 only 2% of the variance in the rheological data.

Figure 4.15 illustrates the model fit for the 1st principal component obtained from the MLR model.

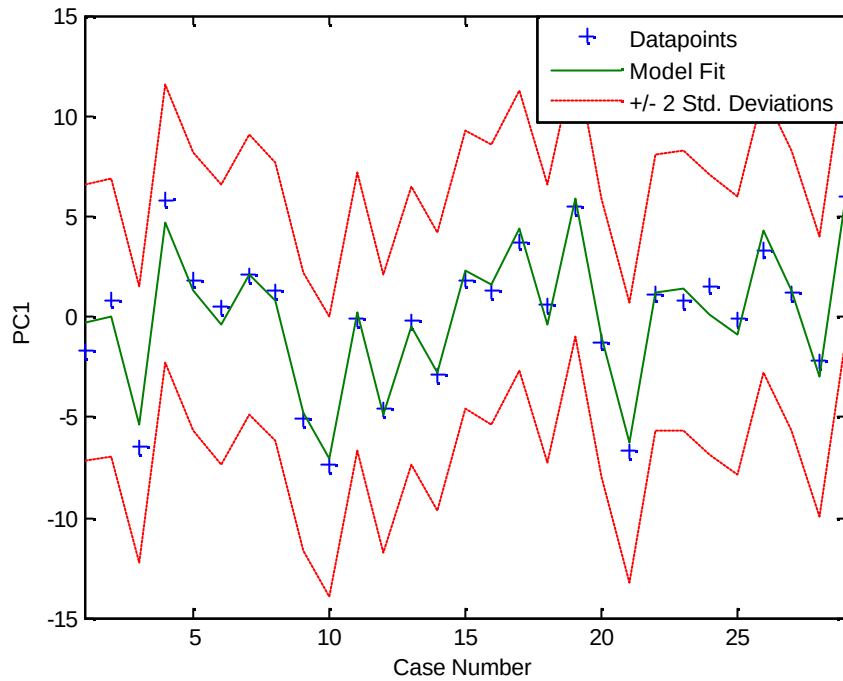


Figure 4.15 Model Fit (MLR) for 1st Principal Component

The immediate impression of Figure 4.15 is that the MLR model fits the data well because all the fitted datapoints lie within two standard deviations of the model. The high R^2 value of 0.9598 also gives an indication that the model fits the data well. The goodness of the model fit is accepted once the validity of the model is verified (Section 4.4.1.3).

The importance of each regressor variable (raw material) is given by the magnitude of the regression coefficient $b_1, b_2, \dots, b_k$. Figure 4.16 illustrates the relative importance (absolute values) of each raw material on the 1st principal component of the rheology data.

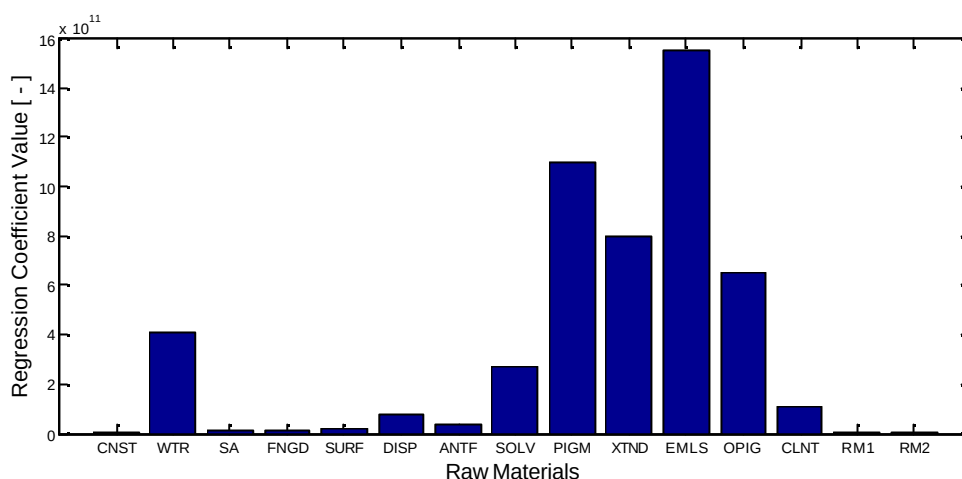


Figure 4.16 Relative Importance of Regression Coefficients for 1st Principal Component

(With CNST=constant, WTR=water, SA=soda ash, FNGD=fungicide, SURF=surfactant, DISP=dispersant, ANTF=antifoamer, SOLV=solvent, PIGM=pigment, XTND=extender, EMLS=emulsion, OPIG=organic pigment, CLNT=coalescent, RM1=rheology modifier 1, RM2=rheology modifier 2)

Multiple linear regression (MLR) models are developed for all the principal components (PC1 – PC18) to determine the relative importance of the raw materials on each of the principal components. Similar graphs for all the PC models, as that of Figure 4.15 and Figure 4.16, are presented in Chapter 11, Section 11.2: - *Raw Materials Screening*.

The variance that each PC explains is given in Table 4.6. The table also includes the computed p-value (significance level). The p-value approach for hypothesis testing rejects the null hypothesis H_0 if the p-value is less than $\alpha=0.05$. That is, it can be said with 95% certainty that there is at least one variable that plays a significant role in explaining the variance. Table 4.6 shows that the null hypothesis H_0 can be rejected for PC1, PC2 and PC16 (p-values < 0.05) with 95% certainty, therefore accepting these models. This implies that at least one of the raw materials contribute significantly to the specific model in explaining the variance. (See Section 8.3.1 of Appendix A for a detailed description of the null hypothesis)

Table 4.6 Regression Analysis for MLR (Rheology Data)

Principal Component	Variance in Rheological Data Explained by PC (%)	R ² for MLR Model Fit	p-value (significance level)	Reject H_0 ($p < 0.05$)
PC1	70.10	0.9790	0.0000	YES
PC2	13.33	0.9235	0.0000	YES
PC3	5.05	0.6024	0.4185	NO
PC4	3.12	0.6591	0.2604	NO
PC5	2.33	0.6574	0.2657	NO
PC6	1.69	0.3411	0.9603	NO
PC7	1.51	0.6892	0.1885	NO
PC8	0.94	0.5254	0.6423	NO
PC9	0.61	0.7190	0.1287	NO
PC10	0.42	0.5260	0.6409	NO
PC11	0.26	0.6039	0.4142	NO
PC12	0.22	0.5253	0.6430	NO
PC13	0.17	0.5289	0.6339	NO
PC14	0.09	0.6052	0.1754	NO
PC15	0.09	0.7284	0.1126	NO
PC16	0.04	0.8161	0.0203	YES
PC17	0.03	0.6292	0.3410	NO
PC18	0.01	0.6441	0.2996	NO

The total average contribution of the raw materials in explaining the rheology is considered to be the sum of the weighted influence of each of the raw materials contributing to the 1st two principal components of the rheological data (PC16 is not included because of its insignificant contribution in explaining variance (0.04%) in the rheological data). In mathematical terms, the raw material contribution (RMC) in explaining the variance in rheological data can be expressed as:

$$RMC = \sum_{i=1}^2 VARIANCE\ EXPLAINED_{PC_i} \bullet REGRESSION\ COEFFICIENTS_{RawMaterial} \quad (4.20)$$

Graphically, the combined raw material contribution in explaining 83.43% of the variance in the rheological data (the 1st two principal components), is illustrated in Figure 4.17.

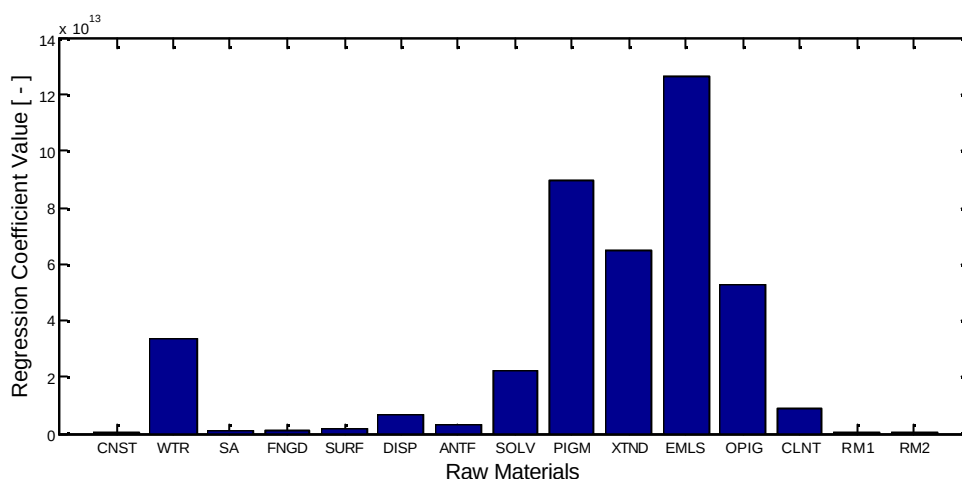


Figure 4.17 Combined Contribution of Raw Materials to 1st two Principal Components of the Rheological Data

Therefore the relative importance of the raw materials on the rheological data is as follows:

Emulsion (EMLS) > pigment (PIGM) > extender particles (XTND) > organic pigment (OPIG) > water (WTR) > solvent (SOLV) > coalescent (CLNT) > dispersant (DISP) > antifoamer (ANTF) > surfactant (SURF) > fungicide (FNGD) > soda ash (SA) > rheology modifier 1 (RM1) > rheology modifier 2 (RM2)

Some remarks about the relative importance of the raw materials in Figure 4.17:

- It is well known that volume solids play a major role in the rheological behaviour of paint [13]. Therefore it is not surprising that the model indicates that the 5 most influential raw materials are the emulsion, pigment, extender particles, organic pigment and water – all which contribute in a major way to the level of volume solids in the paint.
- Figure 4.17 indicates that the rheology modifiers contribute the least in explaining the variance in the rheological data. At first this might seem strange, because it is known that rheology modifiers influence the rheology in a critical way. But, the fact is that the rheology modifiers were kept at a constant level in the screening experiment. The rheology modifiers are already included in the model and were not varied in the screening procedure.
- Figure 4.17 indicates that the raw materials *soda ash* and *fungicide* have the smallest influence on the rheology (excluding the rheology modifiers which are ignored in this case). This correlates with the literature and paint scientists' views.

Before accepting the results in Figure 4.17, validity of the model(s) have to be established.

4.4.1.2 EFFECT OF MULTICOLLINEARITY IN THE RAW MATERIAL SCREENING DATA

Multicollinearity plays an important role when MLR models are used to evaluate the causal relationship between the explanatory variables and the fitted response. The main disadvantage of

multicollinearity is the increase in error of the regression coefficients. It follows that confidence in the regression coefficients decreases and it becomes harder to find statistically significant coefficients.

A constrained mixture design is used to determine the experimental design space for the raw materials. These type of designs often lead to multicollinearity [14]. Design Expert® automatically designs the experimental space so that a minimum level of collinearity exists. However, the constraints put onto each raw material were stringent (plus and minus 20% for each raw material on the standard formulation) and therefore multicollinearity is definitely expected. The level of multicollinearity in the raw material data is given by the condition number, C of Equation 4.18 as:

$$C_{RawMaterials} = 14.96$$

This value is still below the critical value ($C_{critical} = 15$) where the harmful effects of multicollinearity are not serious [10].

4.4.1.3 VALIDITY OF RAW MATERIAL SCREENING PROCEDURE

The models have to be validated first before conclusions can be accepted as correct. This means that the validity of the assumptions that were made during MLR, needs to be checked. These assumptions need to be correct for the method on which MLR is based (method of least squares estimates) to be valid. The validity of the model is not influenced by multicollinearity but by the following factors [17]:

1. Linearity

Usually a scatter plot of the data gives an initial indication of the degree of linearity that is experienced between the response ($\hat{y}_i$) and the explanatory variables (x_i). In this case a scatter plot of the response (PC_i) against a raw material (x_i) does not give much information about the linearity of the data because the raw materials are only varied at three different levels which make it difficult to determine the degree of linearity.

However, the null hypothesis, which is a test for significance of regression to determine whether a linear relationship between the response and a regressor variable exists, is used. The p-values for the MLR models of PC1 and PC2 in Table 4.6 indicate that at least one of the variables definitely contributes significantly to the *linear* model and therefore there exists a linear relationship between at least one of the raw materials and the 1st two principal components.

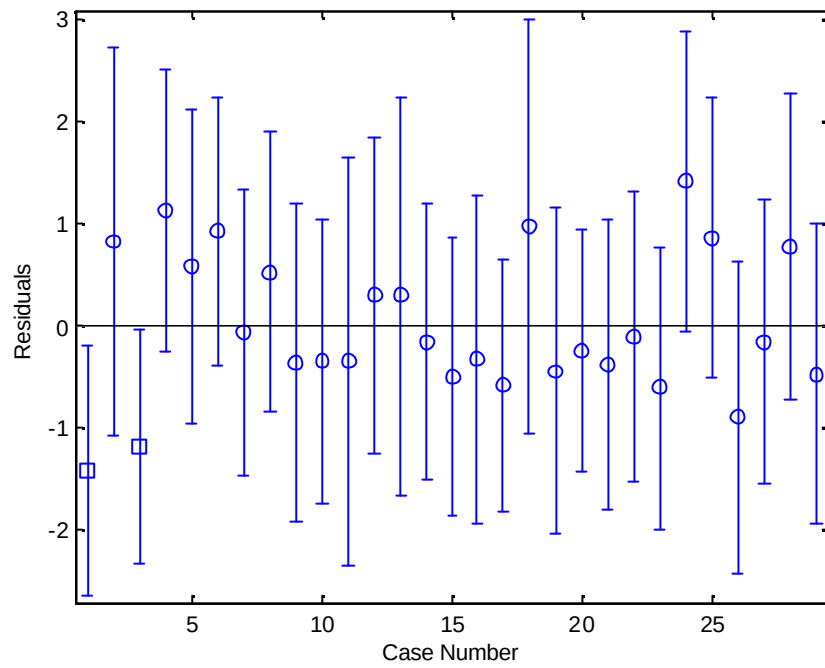
2. Mean independence

The assumption of mean independence assumes that the independent variables are unrelated to the random error ε of equation 4.19. This is one of the most critical assumptions [15]. The mean errors, or residuals, of the models are practically zero for both MLR models (PC1 and PC2) as can be seen in Table 4.7.

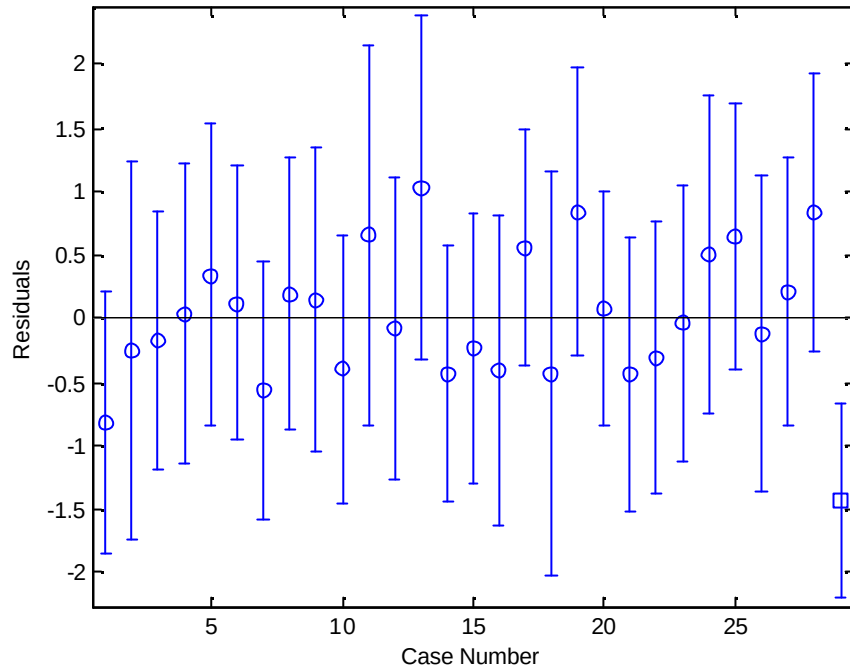
Table 4.7 Mean of Residuals for MLR on PC1 and PC2

Mean for Residuals for PC1 model	-1.23×10^{-4}
Mean for Residuals for PC2 model	-1.71×10^{-4}

The residual case order plots (RCOPLLOT) in Figure 4.18 (a) and (b) for PC1 and PC2 respectively, also indicate that the residuals are centred round zero.



(a)



(b)

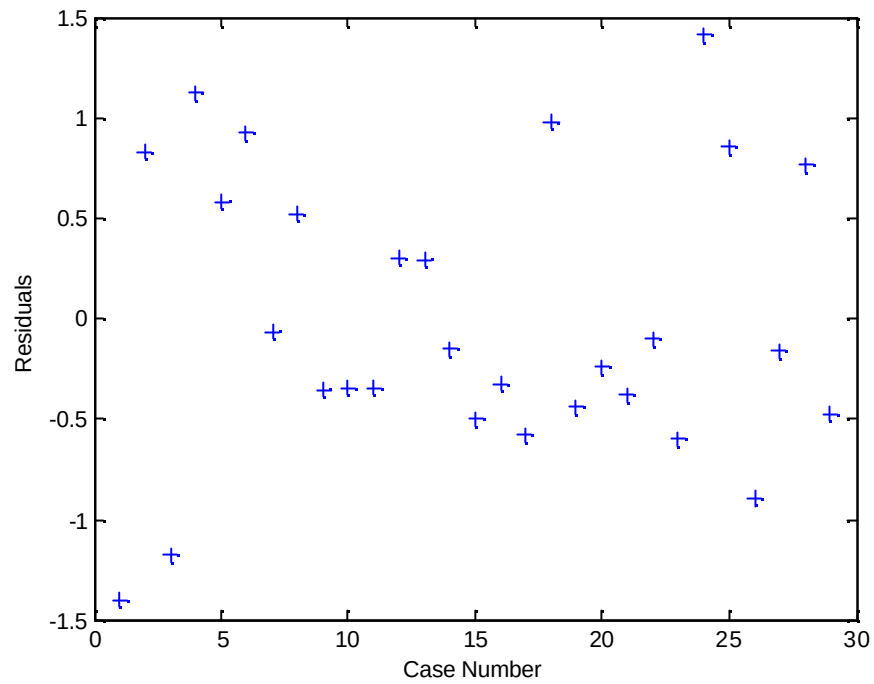
Figure 4.18 RCOplot for (a) PC1 and (b) PC2

These plots also indicate the 95% confidence intervals about the residuals, plotted as error bars. In plot (a) case number 1 and 3 (paint ID = 8 and paint ID = 7 respectively, indicated by markers) fall outside the 95% confidence interval and are considered as outliers. In plot (b) case number 29 (paint ID = 14, indicated by the marker) is considered as an outlier.

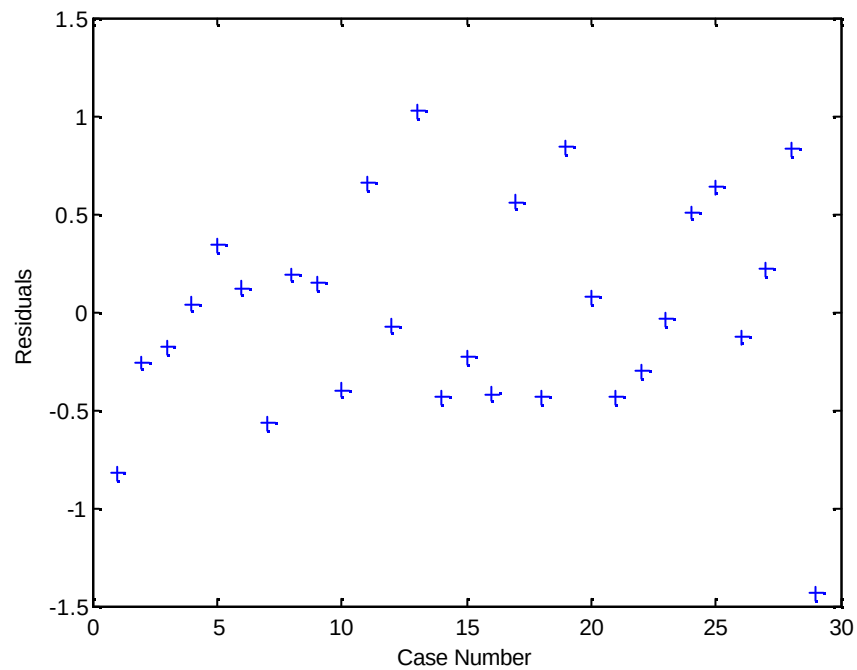
3. Homoscedasticity

It is difficult to determine from the scatter plot of residuals against each of the raw materials whether the variance in the residuals are independent (no systematic patterns) of the value of a specific raw material, because the raw materials are only varied at three different levels.

The fitted values of PC1 and PC2 show no systematic pattern in the case order (Figure 4.19 (a) and (b) respectively) and therefore the residuals are independent of the run order. This is expected because the experiments are performed in random order.



(a)



(b)

Figure 4.19 Scatterplot of Residuals of (a) PC1 and (b) PC2

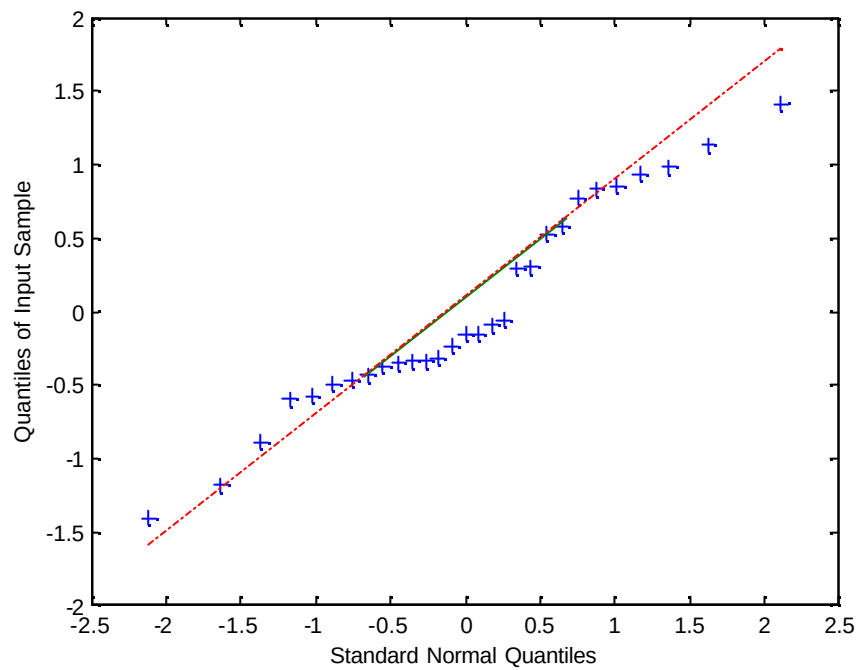
4. Uncorrelated errors

The individual contribution of each of the raw materials to the error term ε of equation 4.19 must be uncorrelated. It is difficult to diagnose correlated disturbances. However, it is known that the issue of correlated disturbances is strongly affected by the experimental design and that if the design space is chosen randomly that it is unlikely that correlated disturbances will be a problem

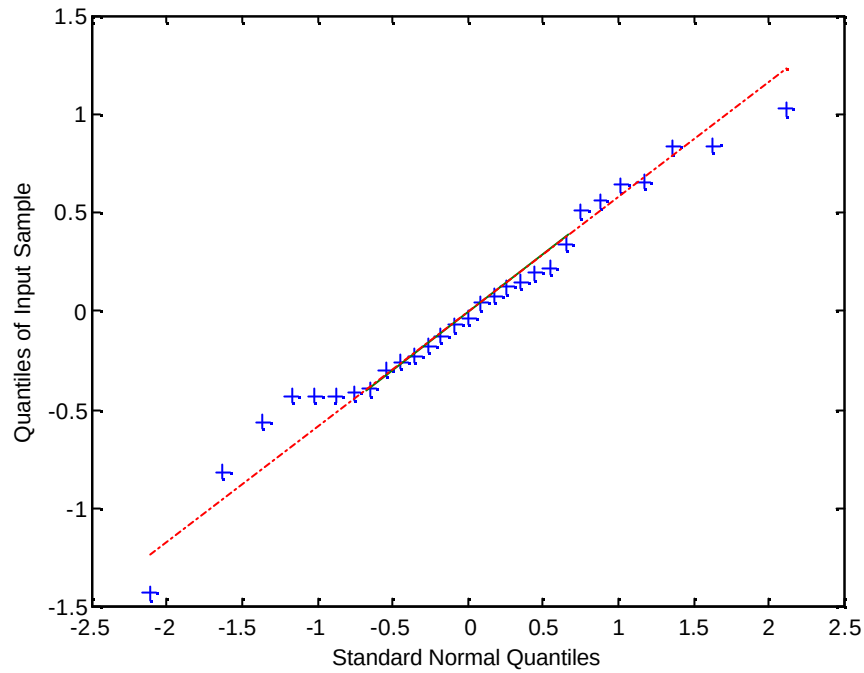
[15]. Design Expert® randomly chooses levels of raw materials for the experimental design, as far as the constraints of the mixture design allows it to, and therefore it is assumed that correlated errors of each raw material do not occur.

5. Normal disturbance

The quantile plot of sample data for the residuals of PC1 and PC2 versus the normal standard quantiles is illustrated in Figure 4.20 (a) and (b) respectively. No severe deviations from normality (identified as a definite S-shaped curve) occur and therefore the data is drawn from a more or less normal distribution (indicated by the straight line).



(a)



(b)

Figure 4.20 QQPLOT for (a) PC1 and (b) PC2

All justifications above indicate that the assumptions made during the MLR process are valid and therefore the conclusions that are drawn from the models are accurate.

4.4.2 PAINT PROPERTY SCREENING (*PATH C*)

OBJECTIVE: Identify the paint properties that are significantly affected by the 6 most important raw materials that affect the rheology the most.

4.4.2.1 MULTIPLE LINEAR REGRESSION FOR SCREENING OF PAINT PROPERTIES

Multiple linear regression is once again used to obtain the correlation between the paint properties and the 6 raw materials identified as the most influential in terms of affecting the rheological behaviour. These raw materials in order of importance are:

- Emulsion
- Pigment
- Extender particles
- Organic pigment
- Water
- Solvent

Figure 4.21 illustrates the model fit for one of the 43 paint properties, Gloss 85 obtained from the MLR model with these 6 raw materials as predictor variables.

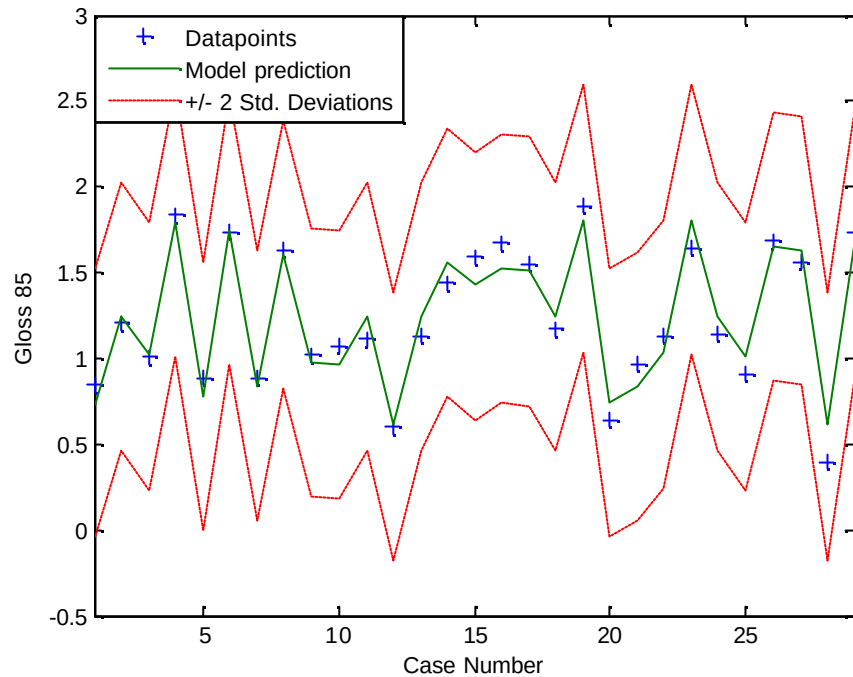


Figure 4.21 Model Fit (MLR) for Gloss 85

Figure 4.21 illustrates that the model fits the data well. The model can explain 93.68% of the variance in the Gloss 85 data. Therefore, the MLR model for Gloss 85, with only 6 raw materials, is accurate in explaining the variance in the data. This is not surprising, because it is known that the levels of pigment and emulsion play major roles in the gloss of a paint [16]. Gloss is a very important paint property as it classifies the paint into a certain product range, e.g. high gloss, medium gloss (sheen), or low gloss (matt), see Figure 4.22. Gloss is strongly affected by the PVC.

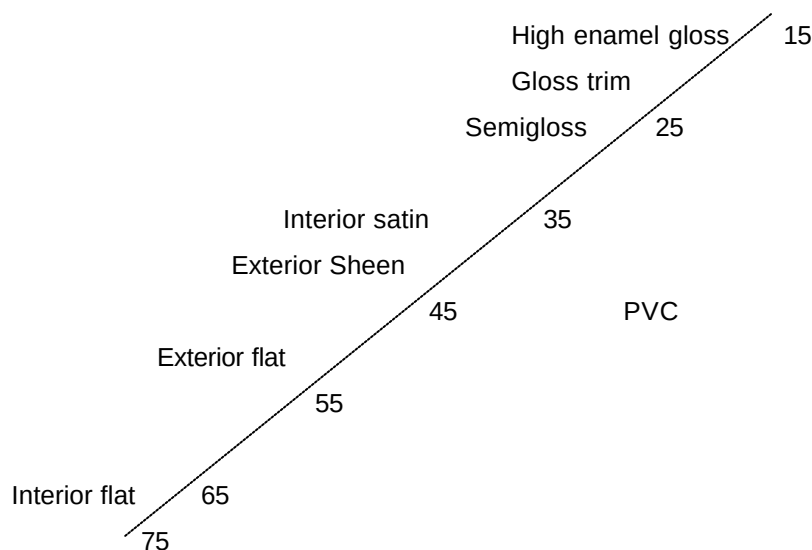


Figure 4.22 Formulation Types in Terms of Pigment Volume Concentration (PVC) [13]

Figure 4.23 illustrates the relative importance of the raw materials on the Gloss 85 property of the paint as determined from the regression coefficients of the MLR model.

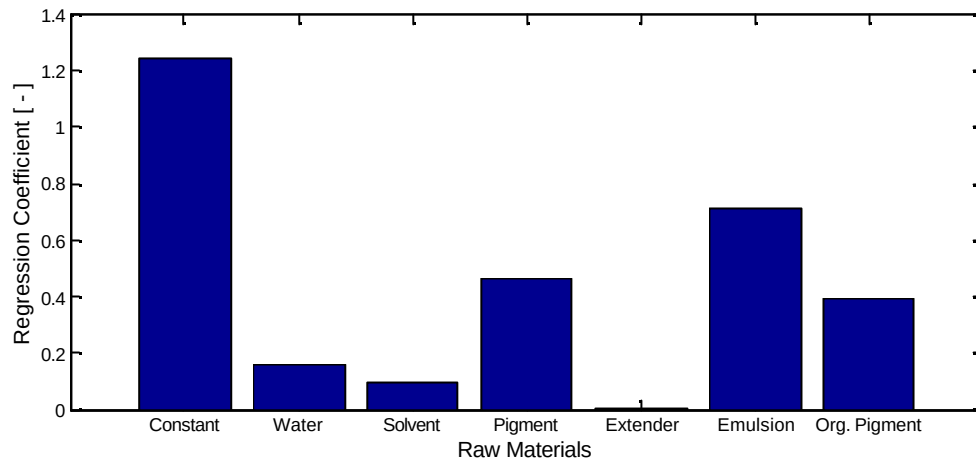


Figure 4.23 Relative Contribution of Raw Materials to Gloss 85

Ignoring the constant of the model, the MLR model for the Gloss 85 paint property indicates that the emulsion has the biggest influence on the gloss value of paint, followed by pigment, organic pigment, water, solvent and extender. This correlates with practical experience and literature sources [16].

Multiple linear regression (MLR) is used to develop predictive models of all the paint properties. The idea is to determine which of the 43 paint properties is best predicted by the 6 raw materials that are used as predictor variables. Two extra predictor variables (PVC and VS) are added to the existing six to assist in predicting the variability in the paint properties. Multicollinearity plays a minor role when the main goal of the regression analysis is to predict the dependent variables [15], [17], [11], [10]. Therefore, even though PVC and VS are correlated with the existing 6 predictor variables, they are added to the regression equation in order to explain more of the variance of each paint property. The variance in each paint property, explained by the MLR model is given in Table 4.8. Note that variance (R^2) in each paint property is explained better with the inclusion of PVC and VS. However, the standard errors of the predicted paint properties may increase a bit, but at the moment it is only important to see which of the paint properties' variance can be best explained by at least one of the regressor variables.

The output also includes the computed p-values (called the significance level). The p-value approach for hypothesis testing rejects the null hypothesis H_0 if the p-value is less than $\alpha=0.05$. If $p < 0.05$ then there is a 95% probability that at least one of the explanatory variables is significant in predicting the variance in the paint property.

Table 4.8 Regression Analysis for MLR (Paint Properties)

Paint Property	R ² (with PVC and VS)	R ² (without PVC and VS)	p-value (with PVC and VS)	p-value (without PVC and VS)	Reject H_0 **** (p < 0.05)
Dry Burnish (20)	0.7091	0.6115	0.0005	0.0010	YES
Dry Burnish (60)	0.7204	0.6413	0.0003	0.0004	YES
Dry Burnish (85)	0.4876	0.3116	0.0554	0.1782	YES ^{††††}
Wet Burnish (20)	0.5434	0.5185	0.0228	0.0079	YES
Wet Burnish (60)	0.7898	0.7474	0.0000	0.0000	YES
Wet Burnish (85)	0.7070	0.6083	0.0005	0.0001	YES
Chemical Resistance Blistering NH3	0.2371	0.2322	0.6274	0.3889	NO
Dirt Pick-up (delta gloss 85)	0.5683	0.5482	0.0146	0.0043	YES
Open Time - Glass	0.4820	0.3552	0.0601	0.1060	YES ^{††††}
Dry Film Thickness	0.6446	0.6023	0.0029	0.0012	YES
Wet Film Thickness	0.3429	0.2849	0.2968	0.2373	NO
Gloss (20)	0.9514	0.9478	0.0000	0.0000	YES
Gloss (60)	0.9821	0.9815	0.0000	0.0000	YES
Gloss (85)	0.9475	0.9368	0.0000	0.0000	YES
Hiding Power Roller - 1st coat	0.6002	0.5736	0.0078	0.0025	YES
Hiding Power Brush - 1st coat	0.1379	0.0995	0.9074	0.8677	NO
Hiding Power Roller - 2nd coat	0.3634	0.3309	0.2454	0.1428	NO
Hiding Power Brush - 2nd coat	0.1753	0.0177	0.8191	0.9986	NO
Krebs1	0.9386	0.9386	0.0000	0.0000	YES
Krebs2	0.9521	0.9487	0.0000	0.0000	YES
Krebs3	0.9333	0.9275	0.0000	0.0000	YES
Krebs4	0.9286	0.9250	0.0000	0.0000	YES
Opacity	0.5188	0.3515	0.0344	0.1111	YES
Anti-sag Index	0.7452	0.7392	0.0002	0	YES
Spreading Rate 1st coat	0.1121	0.1094	0.9508	0.8368	NO
Spreading Rate 2nd coat	0.1534	0.0977	0.8743	0.8729	NO
Stain Removal	0.2764	0.2525	0.4963	0.3251	NO
Water Permeability	0.8594	0.8317	0.0000	0	YES
Water Resistance	0.8053	0.6992	0.0000	0.0001	YES

**** The null hypothesis is rejected when $p < 0.05$ based on the p-values with PVC and VS

†††† The null hypothesis is rejected even though $p > 0.05$, because the probability of the null hypothesis being rejected is still relative large (94.46%, $p=0.0554$)

††† The null hypothesis is rejected even though $p > 0.05$, because the probability of the null hypothesis being rejected is still relative large (93.99%, $p=0.0601$)

X-hatch Adhesion(glass)	0.1768	0.1366	0.8150	0.7421	NO
X-hatch Adhesion(panel 2A)	0.4179	0.4010	0.1377	0.0571	NO
Brush Drag1st coat	0.5381	0.5055	0.0250	0.0101	YES
Brush Drag 2nd coat	0.6197	0.5721	0.0051	0.0025	YES
Chemical ResistanceWater Softness	0.1713	0.1350	0.8299	0.7480	NO
Dirt Pick-up	0.3747	0.3523	0.2198	0.1100	NO
Ease of Stir	0.8384	0.8176	0.0000	0	YES
Fibre Levelling 1st coat	0.2025	0.1593	0.7396	0.6565	NO
Fibre Levelling 2nd coat	0.2166	0.1685	0.6949	0.6209	NO
Levelling (Leneta)	0.6255	0.6094	0.0045	0.0015	YES
Roller Drag 1st coat	0.3438	0.3092	0.2944	0.1831	NO
Roller Drag 2nd coat	0.6339	0.5938	0.0037	0.0015	YES
Spatter	0.4987	0.3840	0.0471	0.0724	YES
Synerisis	0.7510	0.6594	0.0001	0.0001	YES
Scrubs ^{ssss}	NA	NA	NA	NA	NA
Settlement ^{*****}	NA	NA	NA	NA	NA
Microbial Attack ⁺⁺⁺⁺⁺	NA	NA	NA	NA	NA

All the paint properties for which the null hypothesis, H_0 , gets rejected, is used for further modelling. Note that, as can be expected, adding PVC and VS as explanatory variables to the regression analysis, make the chances that “at least one explanatory variables play a significant role in explaining the variance in behaviour”, so much more and therefore paint properties such as *opacity* and *spatter* would not have been selected for further modelling (p-values > 0.05) if PVC and VS were not included as predictor variables..

Only paint properties with p values >> 0.05 are rejected, because p values are indications of probability and therefore *Dry Burnish* and *Open Time* are also selected for further modelling because their p values ≈ 0.05.

4.4.2.2 EFFECT OF MULTICOLLINEARITY IN THE PAINT PROPERTY SCREENING DATA

As indicated earlier, multicollinearity plays a minor role when the main goal of the regression analysis is to predict the dependent variables. Addition of PVC and VS, although adding to the level of collinearity, is not considered to be a problem.

^{ssss} No variance in results: all scrubs > 10 000 cycles

^{*****} No variance in results: no settlement in any of the paints

⁺⁺⁺⁺⁺ No variance in results: no microbial attack notices on any of the paint films

4.4.2.3 VALIDITY OF PAINT PROPERTY SCREENING DATA

(Only *Gloss85* is used as an example to illustrate validity of the model. 'Mean of residuals', 'RCOPLOTS', 'Scatterplots of Residuals' and 'QQPLOTS' for all the paint properties are given in Chapter 11, Appendix D)

The following are evaluated to check for the validity of the models:

1. Linearity

A scatterplot of the *Gloss85* values against the level of each raw material would have given a good indication of the level of linearity that exists between a raw material and the paint property. Each raw material is varied at just three levels and therefore it is difficult to spot any trends in the data. The null hypothesis, which tests for significance of regression, to determine whether a linear relationship between the response and the regressor variables exist, is used. The p-values for the MLR model of *Gloss85* in Table 4.8 indicate that at least one of the variables (raw materials, PVC and VS) definitely contributes significantly to the *linear* model and therefore there exists a linear relationship between at least one of the raw materials and the *Gloss85* values.

The assumption of linearity only needs to be *approximately* true, because at this point, the idea of the models is not to be able to predict the paint properties absolutely accurate, but to see whether it is *possible* to predict a certain paint property with the set of predictor variables.

2. Mean independence

The assumption that the independent variables are unrelated to the random error ε is validated by the fact that the average of the residuals is 1.82×10^{-16} and for all practical purposes is equal to zero. The scatterplots and RCOPLOTS of Appendix D indicate that the average for all the residuals of the paint properties are for all practical purposes equal to zero. Therefore the independent variables are unrelated to the random error explained by the MLR model for each paint property.

The residual case order plot (RCOPLOT) in Figure 4.24 for the *Gloss85* values also indicate that the residuals are centred round zero.

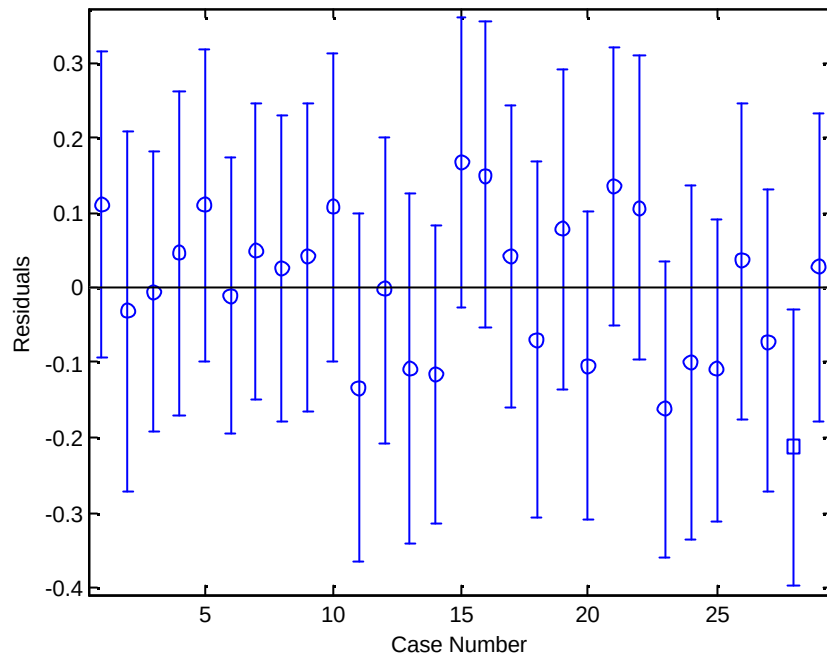


Figure 4.24 RCOplot for Gloss85

Figure 4.24 indicates that case number 29 can be considered as an outlier ($\square$ marker).

3. Homoscedasticity

The scatterplot of residuals against case number in Figure 4.25 indicate no systematic pattern and therefore the residuals are independent of the run order.

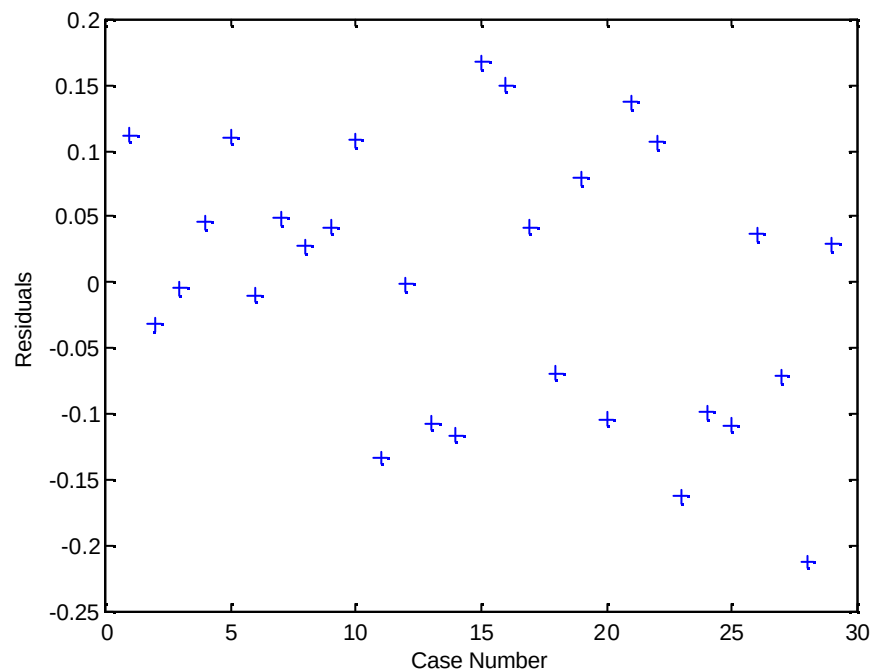


Figure 4.25 Scatterplot of Residuals for Gloss85

4. Uncorrelated errors

It is difficult to diagnose the contribution in error of each explanatory variable to the error term ε in the regression equation. Design Expert[®] determines specific levels of raw materials so that errors are uncorrelated, as far as the constraints allow. In general, it is unlikely for errors to be correlated when runs are performed with random levels of variables [15].

5. Normal disturbance

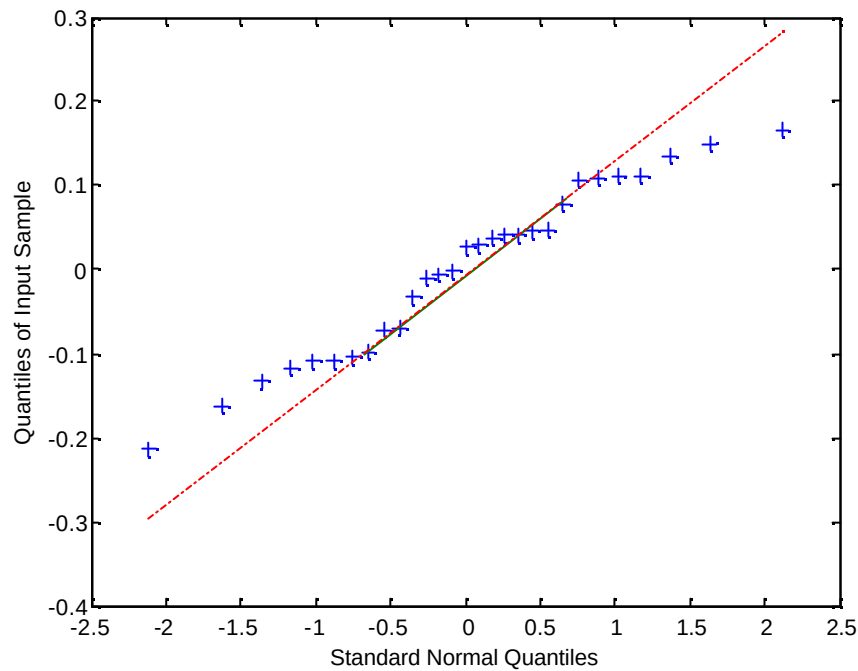


Figure 4.26 QQPLOT for Gloss85

The QQPLOT in Figure 4.26 indicates no severe nonnormality in the residual data. Nonnormality is identified by a definite pattern in the residual data. The tendency of the QQPLOT to bend slightly at the upper and lower ends is common and the plot is not grossly nonnormal.

4.5 CONCLUSIONS

1. The rheological parameters of shear stress (τ) from the rotational tests (flow behaviour) and a combined parameter ($t_A \cdot d$) from the oscillatory tests (viscoelastic behaviour) are used successfully as input variables to the rheological model. Although these rheological variables do not intuitively make sense, it simplifies the modelling process and result in accurate model predictions.
2. Preliminary modelling (the fitting of a trendline to the rheological data) accurately represents the rheology curves of this specific paint. The high R^2 values of the trendlines ($R^2 > 0.94$) show that the trendline parameters can be used with confidence to describe the rheology curve. It is only the trendline of the Low Shear curve which R^2 -value is smaller than 0.9.

3. Collinearity in the rheology data is abundant. This is because many of the tests determine the same behaviour but in a different test mode - rotational vs oscillatory. Rheology variables that are the most highly correlated (correlation values > 0.9) are:

i) b_{FS} and b_{FC}

ii) a_{HS} and b_{HS} .

There are 18 other correlations between rheology variable which are also considered to be high (>0.8). Rheology variables that are least correlated (correlation values = 0) are:

i) a_{AS} and b_{FC}

ii) a_{AS} and b_{TS} .

4. Principal component analysis on the highly correlated rheology data significantly reduces dimensionality of the dataset. Three principal components explain more than 80% of the variance in all the rheology data.
5. The most important raw materials that affect the rheology (excluding the rheology modifiers) are *emulsion > pigment > extender particles > organic pigment > water > solvent > coalescent > dispersant > antifoamer > surfactant > fungicide > soda ash*.
6. All paint properties change as the level of the raw materials vary. However, the level of certain paint properties did not vary enough for the models to explain the variance accurately. These paint properties are:
 - chemical resistance (NH₃, H₂O)
 - wet film thickness
 - hiding power (brush/roller)
 - spreading rate (1st and 2nd coat)
 - stain removal
 - adhesion
 - fibre levelling
 - roller drag (1st coat)
 and therefore not considered for further detailed modelling.

4.6 NOMENCLATURE

Abbreviations

3-ITT	Three-interval-thixotropy test
3-ITT (osc)	Three-interval-thixotropy test in oscillatory mode
3-ITT (rot)	Three-interval-thixotropy test in rotational mode
ANTF	Antifoamer
AS	Amplitude sweep

CLNT	Coalescent
CNST	Constant
C_{XLF}	Constant of extra low frequency
DISP	Dispersant
EMLS	Emulsion
FC	Flow curve
FNGD	Fungicide
FS	Frequency sweep
HS	High shear
H_0	Null hypothesis
LS	Low shear
MLR	Multiple linear regression
OPIG	Organic pigment
PC	Principal component
PIGM	Pigment
PVC	Pigment volume concentration
RM1	Rheology modifier 1
RM2	Rheology modifier 2
RMC	Raw material contribution
SA	Soda ash
SOLV	Solvent
SURF	Surfactant
TS	Time sweep
VS	Volume solids
WTR	Water
XLF	Extra low frequency
XTND	Extender

Symbol	Description
C	Condition number (-)
G'	Storage modulus (Pa)
G''	Loss modulus (Pa)
G^*	Complex modulus (Pa)
t	Time (s)
x	Independent variable
y	Dependent variable
x_{avg}	Average value of variable
x_i	Value of variable
$x_{standard deviation}$	Standard deviation of variable

$X_{s \text{ standardised}}$

Standardised value of variable

Greek Symbol

Description

a

Preliminary model parameter (gradient)

a_i

Gradient of preliminary model where $i=HS, LS, 3-ITT_{(rot)}$, etc

$a_{oscillatory}$

Gradient of oscillatory measurement

$a_{rotational}$

Gradient of rotational measurement

b

Preliminary model parameter (offset)

b_i

Preliminary model parameter (offset) where $i=HS, LS, 3-ITT_{(rot)}$, etc

$b_{oscillatory}$

Offset of oscillatory measurement

$b_{rotational}$

Offset of rotational measurement

b_0

Constant associated with regression analysis

b_k

Regression coefficient associated with variable $k=1,2,3,\dots$

g

Deformation [%] or strain [-]

g_A

Deformation [%] or strain [-] amplitude

$\dot{g}$

Shear rate [1/s]

d

Phase shift, loss angle [°]

e

Error of approximation

h

Shear viscosity [Pa.s]

h^*

Complex viscosity [Pa.s]

t

Shear stress [Pa]

t_A

Shear stress amplitude [Pa]

w

Angular frequency [rad/s]

4.7 REFERENCES

1. Swingler, K. (1994). Applying Neural Networks. A Practical Guide. Chapter 1-3. San Diego, Academic Press Inc.
2. Bourquin, J., Schmidli, H., Van Hoogevest, P., Leuenberger, H. (1998). "Pitfalls of artificial neural networks (ANN) modelling technique for data set containing outlier measurements using a study on amixture properties of direct compressed dosage form." European Journal of Pharmaceutical Sciences **7**: 17-28.
3. Gasperlin, M., Tusar, L., Smid-Korbar, J., Kristl, J. "Viscous behaviour of pharmaceutical creams: characterisation by neural network and polynomial modelling." p. 14-20. University of Ljubljana, Faculty of Pharmacy, Slovenia.

4. Osterhold, M. (2000). "Rheological methods for characterising modern paint systems." Progress in Organic Coatings **40**: 131-137.
5. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Zupan, J., Kristl, J. (2000). "Viscosity prediction of lipophilic semisolid emulsion systems by neural network modelling." International Journal of Pharmaceutics **196**: 37-50.
6. Leskovsek, N., Tusar, L., Tusar, M. (1995). "Empirical modelling of rheological and mechanical properties of paint." Rheology: 140.
7. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Kristl, J. (1998). "Lipophilic semisolid emulsion systems: viscoelastic behaviour and prediction of physical stability by neural network modelling." International Journal of Pharmaceutics **168**: 243-254.
8. Prideaux, J. "Rheology modifiers and thickeners in aqueous paints." Journal of the Oil and Colour Chemists' Association. ISSN 0030-1337. Rohm and Haas.
9. Moolman, P. L. (2003). Rheology of Coating Systems. Department Chemical Engineering. University of Stellenbosch.
10. Chatterjee, S., Price, B. (1991). Regression Analysis by Example. Second Edition, Chapter 6. New York, John Wiley & Sons, Inc.
11. Aldrich, C. (2002). Introduction to exploratory data analysis and empirical modelling. Unpublished Manuscript, Stellenbosch.
12. Montgomery, D. C. (2005). Design and Analysis of Experiments, John Wiley and Sons Inc.
13. Schaller, J. E. Introduction to latex paints, p.1-56. Rohm and Haas Company.
14. Engineering statistics handbook, NIST. **2006**.
(<http://www.itl.nist.gov/div898/handbook/pmd/section1/pmd142.htm>)
15. Allison, P. D. (1999). Multiple Regression. A Primer, Chapter 1. California, SAGE Publications Ltd.
16. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Zupan, J., Kristl, J (2000). "Viscosity prediction of lipophilic semisolid emulsion systems by neural network modelling." International Journal of Pharmaceutics **196**: 37-50.
17. Neter, J., Kutner, M. H., Nachtsheim, C. J., Wasserman, W. (1996). Applied Linear Statistical Models, Chapter 8. Tom Casson.
18. Parsons, P. (1993). Surface Coatings: Raw materials and Their Usage, Chapter 26. London, Chapman and Hall.

5 LINEAR MODELLING

5.1 INTRODUCTION

A linear model is based on a linear equation and the graph of the equation explaining the relationship between the actual and the predicted value is a straight line [1].

There is a considerable volume of literature available from coatings raw material suppliers (Rohm and Haas®, Huntsman®, Noveon®, etc) that describe modern paint formulating and optimisation methodology that involves design of experiments and basic and advanced modelling techniques [2], [3].

Figure 5.1 shows the relationship between raw materials, paint rheology and paint properties. Quality, as well as various paint properties is directly determined by the raw materials in the paint formulation (MODEL 1 and MODEL 2). Indirectly, some of the paint properties are affected by the rheology of the paint (MODEL 3). Interaction between the raw materials makes the use of constitutive equations almost impossible in explaining the paint properties. This makes models ideal in assisting the paint scientist to obtain an approximation of a property at any level of a raw material (although still within the range used in the model).

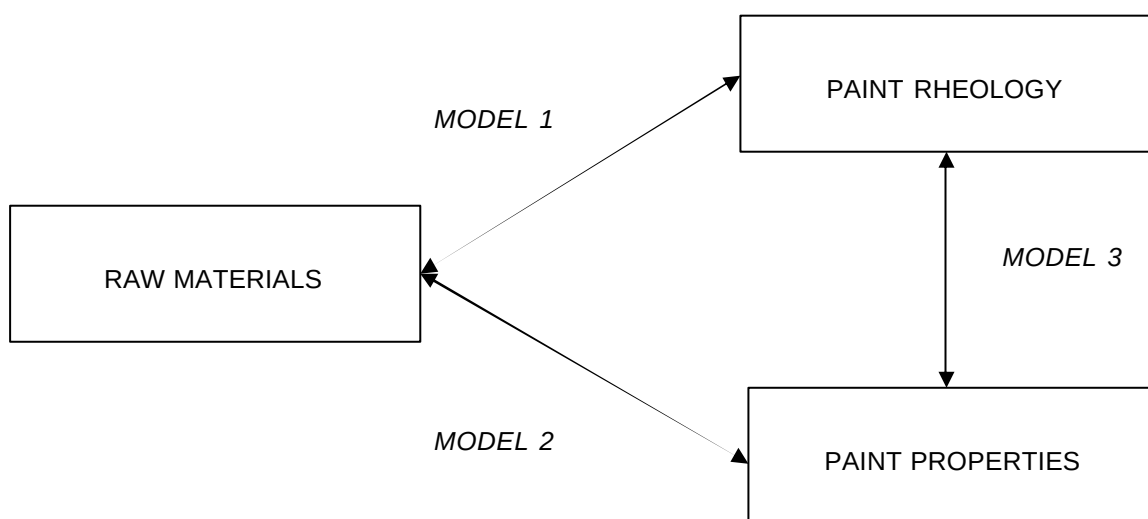


Figure 5.1 Relationship Between Raw Materials, Paint Rheology and Paint Properties

By following the principle of Occam's razor⁺⁺⁺⁺, linear modelling is a good starting point if little is known about the behaviour of a system. Although it is obvious that interactions occur between

⁺⁺⁺⁺ Occam's razor states that the explanation of any phenomenon should make as few assumptions as possible, eliminating, or "shaving off" those that make no difference in the observable predictions of the explanatory hypothesis or theory. In short, when given two equally

raw materials, linear modelling assumes that none of them do. This is not necessarily a 'bad' thing, because a lot of information can be obtained from linear modelling:

- The contribution each raw material makes to the rheology and paint properties
- The level of non-linearity or the interactions between the different raw materials.

It might also happen that a linear model is sufficient in explaining some of the paint properties within the allowed error margin and then more complicated models are not required.

5.1.1 MULTIPLE LINEAR REGRESSION (MLR)

In experimental work, it is often desirable to find whether, and by how much, the level of one quantity changes with changes in the level of some other quantities. One of the best ways to accomplish this is by the use of regression analysis.

Multiple regression is a statistical method for studying the relationship between a single dependent variable and several independent variables [4].

There are two major uses of multiple regression, both of which are used in this study [4]:

1) Causal

In causal analysis, the independent variables are regarded as causes of the dependent variable. The aim is to determine whether a particular independent variable *really* affects the dependent variable, and to estimate the magnitude of that effect, if any. This type of analysis is performed in the screening procedure of the raw materials and the linear modelling of selected raw materials. Several raw materials might be related to a dependant paint property, however without running an experiment and collecting data, one does not know which variables will indeed be strong predictors of the response variable. Thus the data will be collected on all k predictors and the response. Then multiple regression analysis is used to determine those predictors that are less effective in explaining the behaviour of the dependent variable.

2) Prediction

In the prediction study, the goal is to develop a formula for making predictions about the dependent variable, based on the observed values of the independent variables. Confidence in the predictions is also important. To account for this a 95% confidence interval for a future value (95% prediction interval) is constructed.

In general, the response variable y may be related to k regressor variables. The model

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_kx_k + e \quad (5.1)$$

valid explanations for a phenomenon, one should embrace the less complicated formulation. ('*entia non sunt multiplicanda praeter necessitatem*' – entities should not be multiplied beyond necessity)

is called a multiple linear regression model with k regressor variables. The parameters $b_j, j = 0, 1, \dots, k$, are called the regression coefficients and indicate the relative importance (positive or negative) of the relating regressor variable. In multilinear regression the regression coefficients are only valid if the error on the independent variables are negligible when compared with the error of the dependent variable. Large errors on the independent variables can result in inaccuracies in the model. The function is linear because no regression coefficients are raised to a power greater than one. Note that the error term ε is present, because individual variation will make it impossible to predict the response exactly accurately. This model describes a hyperplane in the k -dimensional space of the regressor variables $\{x_j\}$. The parameter b_j represents the expected change in response y per unit change in x_j .

Methods for estimating the parameters in multiple linear regression models are called model fitting. The statistical method of least squares^{sssss} is typically used to estimate the regression coefficients in a MLR model.

5.1.2 MULTICOLLINEARITY AND ITS EFFECTS IN MLR

Interpretation of the multiple linear regression equation (equation 5.1) depends implicitly on the assumption that the explanatory variables are not strongly interrelated [5]. The interpretation of the relationships between the predictor or explanatory variables and the response variable are of particular interest. Interest lies in the following questions:

1. What is the relative importance of the effects of the different predictor variables?
2. What is the magnitude of the effect of a given predictor variable on the response variable?
3. Can any predictor variable be dropped from the model because it has little or no effect on the response variable?

The interpretation of these questions may not be valid if there are strong linear relationships among the explanatory variables. When there is a complete absence of linear relationships among the explanatory variables, they are said to be *orthogonal*. In most regression applications, the variables are not orthogonal and the condition of severe nonorthogonality is referred to as *multicollinearity* [1].

Usually, the lack of orthogonality is not serious enough to affect the analysis. However, in some situations the explanatory variables are so strongly interrelated that the regression results are ambiguous and it can have the following effects [1]:

1. The fact that some or all of the predictor variables are correlated among themselves does not, in general, inhibit the ability to obtain a good fit nor does it tend to affect inferences about mean responses or predictions of new observations, provided these inferences are made within the region of observations.

^{sssss} In the most general terms, the method of least squares aims at minimizing the sum of squared deviations of the observed values for the dependent variable from those predicted by the model.

2. The estimated regression coefficients tend to have large sampling variability when the predictor variables are highly correlated. As a result, only imprecise information may be available about the individual true regression coefficient. Many of the estimated regression coefficients individually may be statistically not significant even though a definite statistical relation exists between the response variable and the set of predictor variables. Therefore the effect of multicollinearity is only a minor concern for models whose primary goal is prediction. The effect of multicollinearity is more severe in the case of a causal model [4] and the most serious danger of multicollinearity is concluding that *none* of the collinear variables has an effect on the dependent variable when, in fact, any one of them alone has a very strong effect. Instead of looking at the significance of each variable alone, one should look at the joint significance of all the independent variables so that it can be determined whether at least one variable affects the response.

Multicollinearity should be prevented by careful design of experiments so that the scientist has control over the values of the independent variables. Levels of the independent variables can then be chosen in such a way that all correlations among the predictor variables are zero. This is at least true for *unconstrained* experimental designs (no constraints are placed on the level of the independent variables in the experimental design), however constrained mixture designs may create multicollinearity. The use of pseudo components is recommended for constrained mixture designs with high levels of multicollinearity [6].

Another way that the effects of multicollinearity can be prevented is by the *principal component* approach. Principal components, which are linearly independent, are extracted from the collinear data and used as independent variables for the regression equation (equation 5.1).

Pairwise correlation between predictor variables is often helpful; however, serious multicollinearity can exist without being disclosed by pairwise correlation coefficients [1], e.g. correlation between a group of predictor variables and a single predictor variable [2]. Therefore a better method for diagnosing multicollinearity is the measure of *overall* multicollinearity of the variables. This can be obtained by computing the condition number of the correlation matrix by using equation 4.18. The condition number will always be greater than 1. A large condition number indicates evidence of strong collinearity. The harmful effects of collinearity in the data become strong when the values of the condition number exceed 15. This value for C is determined empirically and not theoretically. Corrective action should always be taken when $C > 30$ [4].

5.1.3 MODEL VALIDITY

Validity of using multiple linear regression (MLR) as a modelling technique depends on whether certain assumptions are satisfied. There is enormous confusion about just what those assumptions are and there is no single set of assumptions on which everyone agrees. Some assumptions involve stronger conditions than others and it is not surprising that the models with stronger assumptions usually lead to stronger conclusions. There is also the tendency to treat all assumptions as equally important when in fact some are much more critical than others [4].

The best known standards of performance and for testing the validity of a statistical method are:

- i. *Bias*: A model that is unbiased is preferred. A model is unbiased if there is no systematic tendency to produce estimates that are either too high or too low
- ii. *Efficiency*: Efficiency is an estimate of how much variation there is around the true value and is measured by the standard error. Efficient models have standard errors that are as small as possible.

Allison [4] presents five assumptions for a valid MLR model:

- 1) *Linearity*: The dependent variable y is a linear combination of the independent variables x_i , plus a random error ϵ , as is illustrated by equation 5.1. The assumptions of linearity will be only approximately true.
- 2) *Mean independence*: The most important assumption that is made about ϵ is that its mean, or average value, does not depend on the independent variables x_i . More specifically, it is assumed that the mean of ϵ is always zero. The assumption of mean independence assumes that the independent variables are unrelated to the random error ϵ . The assumption of mean independence is the most critical assumption of all because:
 - Violations can produce severe bias in the estimates
 - There are often reasons to expect violations
 - There is no way to test for violations without additional data.

Randomisation of the experimental runs ensures that the unmeasured characteristics of the subjects are not related to the treatment variable and therefore if the data is produced by a randomised experiment, the violations are unlikely to occur.

- 3) *Homoscedasticity (variance independent)*: The variance of ϵ must be independent on the values of x_i . A systematic pattern in the residuals plotted against any of the explanatory (independent) variables suggests an inadequate model.
- 4) *Uncorrelated errors*: The error term ϵ in the linear equation of equation 5.1 is actually a different random variable for every individual in the sample. It can also be interpreted as the combined effects of all the causes of y that are not directly included in the equation. Assumption 4 states that the individual contribution of each independent variable to the error term ϵ must be uncorrelated with the individual contribution of any other independent variable. Although it is possible to diagnose correlated disturbances by examining the data, there are not many convenient ways to do it. Generally the issue of correlated errors is strongly affected by the experimental design.
- 5) *Normal disturbance*: The error term ϵ must have a normal distribution. The normality assumption is probably the least important of the five assumptions. The criteria for unbiasedness and efficiency do not depend at all on this assumption. The *central limit theorem* tells us that, if the sample size is moderately large enough, the model will be a good approximation even if ϵ is not normally distributed.

If all these assumptions are met, ordinary least squares (a statistical method for estimating the regression coefficients of MLR) has several desirable properties:

- The assumptions of linearity (1) and mean independence (2) imply that the method of least squares is unbiased.
- The additional assumptions of homoscedasticity (3) and uncorrelated errors (4) imply that the method of least squares is efficient.
- The assumption of normality (5) implies that valid accurate p-values are obtained for hypothesis testing.

As stated earlier, one of the major uses of MLR is to obtain a causal relationship between the predictor variables (independent variables) and the response (dependent variable). In other words, the relative importance of each predictor variable is important in determining a causal relationship between a specific predictor variable and the response. When MLR is used to determine the causal relationship between specific predictor variables and the response, multicollinearity plays an important role. A high degree of multicollinearity means that at least two of the independent variables in the regression equation of equation 5.1 are highly correlated. Multicollinearity is not a violation of any of the assumptions discussed above and it does not prevent the calculation of the regression coefficients. However, it does make it more difficult to reliably estimate the regression coefficients of those variables that are collinear.

When the primary goal of MLR models are used for prediction purposes, then multicollinearity is only a minor concern [4].

The following analyses can be performed on the data to investigate the validity of the assumptions mentioned above.

5.1.3.1 ANALYSIS OF SIGNIFICANCE [7]

In multiple linear regression (MLR) certain tests of hypotheses about the model parameters are helpful in measuring the usefulness of the model. The test for significance of regression is a test to determine whether a *linear* relationship exists between the response variable y and a subset of regressor variables $x_1, x_2, \dots, x_k$. The appropriate hypotheses are

$$\begin{aligned} H_0 : b_1 = b_2 = \dots = b_k = 0 \\ H_1 : b_j \neq 0 \quad \text{for at least one } j \end{aligned} \tag{5.2}$$

The hypotheses in equation 5.2 are discussed in detail in Chapter 8, Section 8.3.1 The Null Hypothesis.

5.1.3.2 ANALYSIS OF MODEL-FIT RESIDUALS

Residuals can be plotted in the following ways:

1. Plot of residuals against observations (various possibilities):
 - The plot of residuals against the fitted values should not indicate any systematic deviations from the response plane, nor indicate that the variance of the error terms varies with the level of the predicted value, thereby testing for *homoscedasticity* (variance

independence). The plot will also give an indication on the level of *mean independence* that, when the independent variables are unrelated to the random error, should tend to zero.

- Residual case order errorbar plot (RCOPlot) of confidence intervals on residuals from regression. The plot shows the residuals plotted in case order. The 95% confidence intervals about these residuals are plotted as error bars. An observation is an outlier if its error bar does not cross the zero reference line. The residuals should also be centered round zero.
- Residual quantile-quantile plot (QQPlot). The plot displays a quantile-quantile plot of the sample quantiles of the residuals versus theoretical quantiles from a normal distribution. If the distribution of the residuals is normal, the plot will be close to linear.

2. Plot of residuals against each of the predictor variables:

Each of these plots can provide further information whether curvature exists and therefore evaluate the assumption of linearity. It also evaluates the model for variance independency (homoscedasticity).

3. Plot of residuals against interaction terms:

Possible interaction terms not included in the model can be evaluated.

5.1.3.3 ANALYSIS OF REGRESSION R^2 STATISTIC

The regression R^2 statistic is a measure of how well the variation of the fitted values of the model is explained by the independent variables. The closer the R^2 values are to 1, the greater is the variance explained by the regressor variables [8].

5.1.4 MODEL PREDICTABILITY

Note that R^2 values do not necessarily indicate the predictability of the model, even if R^2 is high. Therefore model predictions need to be evaluated in different ways. Predictability of the model can be evaluated in two different ways:

- i. without the availability of future observations – using prediction intervals
- ii. with the availability of new observations – calculating average errors.

5.1.4.1 ANALYSIS OF CONFIDENCE INTERVALS

The MLR models will be used for predictive purposes in predicting the value of an observation that is yet to be chosen from the population. It is useful to have a 95% confidence interval produced by the model so that the prediction for a specific observation can be assumed accurate at a certain level (95%) of confidence. The *prediction interval* gives information on individual predictions of the dependent variable. That is, a *prediction interval* for a predicted value of the dependent variable gives us a range of values around which an additional observation of the dependent variable can be expected to be located [9].

5.1.4.2 PREDICTION OF NEW OBSERVATIONS

The predictability of the model can be evaluated by calculating the average relative error between model predictions and new observations.

$$relative\ error = \frac{|actual - predicted|}{actual} \quad (5.3)$$

The "error" is the amount by which the predicted value differs from the quantity that is measured.

5.2 MODEL 1: RELATIONSHIP BETWEEN PAINT RHEOLOGY AND RAW MATERIALS

5.2.1 BACKGROUND

Historically, the lack of a quantitative relationship between the raw materials and the rheology led to the development of rheological models [10], [11]. However, these models still have certain weaknesses (discussed in Chapter 2, Section 2.3.1) and the aim of this part of the study is to overcome these deficiencies by the following means:

- Incorporation of preliminary modelling for rheological evaluation, introduced by Gasperlin et al, ensures that the rheological behaviour gets evaluated over the entire curve, instead of distinct points on the curve. This is important because a raw material may introduce different behaviour in different regions of the rheology curve. This has previously only been applied in the pharmaceutical industry and not in the coatings industry.
- Each paint sample gets evaluated by a set of nine different rheological measurements instead of only one. In this manner the effect that a raw material has on the rheology becomes much more evident.

The above two methods ensure that all possible effects of the raw materials on the rheology are captured. Incorporation of these two methods into the rheological evaluation of the paint leads to the eighteen rheological parameters in Table 5.1.

Table 5.1 Rheology Parameters from Preliminary Models

MEASUREMENT	PRELIMINARY MODEL	RHEOLOGICAL PARAMETERS
1. FLOW CURVE (FC)	$\ln t_{FC} = a_{FC} \ln \dot{g}_{FC} + b_{FC}$	a_{FC} b_{FC}
2. HIGH SHEAR (HS)	$t_{HS} = a_{HS} t_{HS}^2 + b_{HS} t_{HS} + C_{HS}$	a_{HS} b_{HS} C_{HS}
3. LOW SHEAR (LS)	$\ln t_{LS} = a_{LS} \ln t_{LS} + b_{LS}$	a_{LS} b_{LS}
4. 3-ITT (rot)	$\ln t_{3-ITT(rot)} = a_{3-ITT(rot)} \ln t_{3-ITT(rot)} + b_{3-ITT(rot)}$	$a_{3-ITT(rot)}$ $b_{3-ITT(rot)}$
5. AMPLITUDE SWEEP (AS)	$\ln td_{AS} = a_{AS} \ln g_{AS} + b_{AS}$	a_{AS} b_{AS}
6. FREQUENCY SWEEP (FS)	$\ln td_{FS} = a_{FS} \ln f_{FS} + b_{FS}$	a_{FS} b_{FS}
7. TIME SWEEP (TS)	$td_{TS} = a_{TS} t_{TS} + b_{TS}$	a_{TS} b_{TS}
8. 3-ITT (osc)	$\ln td_{3-ITT(osc)} = a_{3-ITT(osc)} \ln t_{3-ITT(osc)} + b_{3-ITT(osc)}$	$a_{3-ITT(osc)}$ $b_{3-ITT(osc)}$
9. EXTRA LOW FREQUENCY (XLF)	$td_{XLF} = C_{XLF}$	C_{XLF}

The rheological parameters in Table 5.1 represent the gradient (a) and the offset (b) of the preliminary models (except for HS and XLF) fitted to the rheological data. The high R^2 -values of the preliminary models (see Chapter 4, Section 4.3.2 Preliminary Modelling, Table 4.2) indicate that these models fit the data well and therefore the preliminary models can be seen as good representations of the *whole* rheological curve for each measurement. The aim is to model the rheological parameters in Table 5.1 as accurately as possible with the use of the main raw materials (determined in the screening procedure) so that a new rheological curve can be constructed from the a and b values obtained from the predicted responses. The raw materials that are used to construct the predictive linear models are:

1. Emulsion (binder)
2. Pigment
3. Extender Pigment
4. Organic Pigment

5. Water
6. Solvent
7. Rheology Modifier (type 1)
8. Rheology Modifier (type 2)
9. Combination of remaining raw materials (as a relative ratio).

The first six raw materials were determined as the most important raw materials affecting the rheology (see *Chapter 4, Section 4.4.1 Raw Material Screening*). The rheology modifiers are also included because it is known that they influence the rheology. Although the remaining raw materials (coalescent, antifoamer, fungicide, soda ash, surfactant and dispersant) were determined as 'less' important, they are included as a single variable because their combined effect might still be important.

A high level of collinearity exists between the rheological parameters (condition number, $C \gg 15$). Principal component analysis reveals that more than 80% of the variance in the rheological behaviour is explained by the first three principal components (see Table 5.2).

Table 5.2 Variance in Rheological Data Explained by the 1st Three Principal Components

PRINCIPAL COMPONENT	VARIANCE IN RHEOLOGICAL DATA [%]
PC1	62.75
PC2	13.06
PC3	6.87

Models that predict the effect of raw materials on the principal components (PC1 to PC3) would have simplified the modelling process considerably. Only three models, instead of eighteen would have been constructed. However, a major drawback of principal component analysis is that the principal components lack simple interpretation since each is, in a sense, a mixture of the original variables and reconstruction of the PC's to the eighteen original rheological parameters is almost impossible (advanced statistical methods do exist but is beyond the scope of this study). In modern coating rheology, it is still very much standard practise to represent the rheological characterisation of a coating with an entire set of rheological curves ([12], [13], [14]) and not only a few principal rheological parameters.

5.2.2 MULTIPLE LINEAR REGRESSION (MLR) MODELLING

Multiple linear regression is used as a modelling technique to determine:

- i. The *causal relationship* between the raw materials and the rheological behaviour
- ii. The *predictive relationship* between the raw materials and the rheological behaviour.

Although the causal relationship between the raw materials and the rheological behaviour is of great importance to the paint formulator, this study concentrates more on the predictive capabilities and therefore more attention is given to the predictive models.

5.2.2.1 MLR – CAUSAL MODEL

It is interesting to see what the most important factors are in explaining 80% of the variance in the rheological data. This is illustrated in Figure 5.2

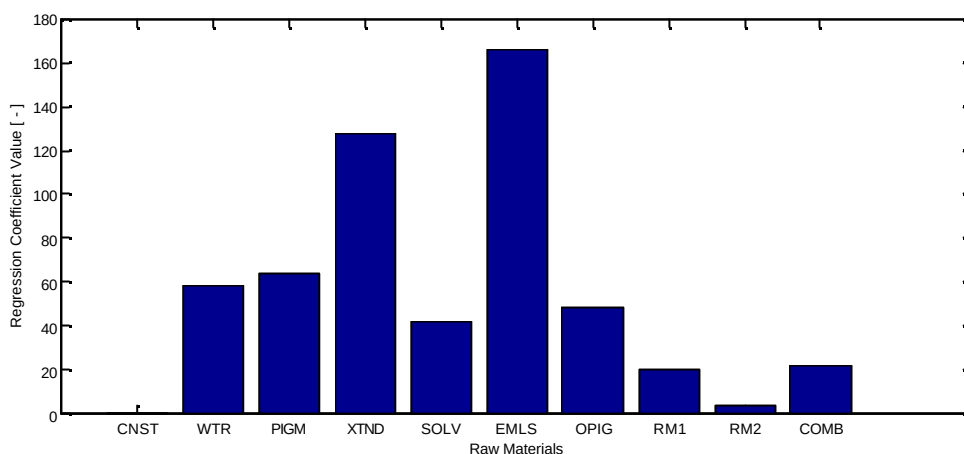


Figure 5.2 Relative Contribution to 80% Variance in Rheological Response

(CNST=constant, WTR=water, PIGM=pigment, XTND=extender, SOLV=solvent, EMLS=emulsion, OPIG=organic pigment, RM1=rheology modifier1, RM2=rheology modifier2, COMB=combination.)

It is interesting to note from Figure 5.2 that the five most important raw materials affecting the rheology, is also the five raw materials that affect the volume solids of a coating the most. The strong correlation between volume solids and the rheology of a coating is well known [15] and it confirms the findings in Figure 5.2.

One of the recent developments in the coatings industry is towards low volatile organic compounds (VOC) coatings. For example, the paint under investigation is currently being formulated without any VOC's (solvents) compared to solvent levels of $\pm 5\%$ when the study was undertaken. Figure 5.2 indicates that the solvent also plays a relative important role in affecting the rheology. It is evident that rheology control has become very important for paints that are currently being manufactured.

The effect of rheology modifiers, RM1 and RM2 is unexpectedly small. A reason for this might be because although one thinks of rheology modifiers as having a major effect on the rheology, they actually do not. The 'basic' rheology is set up by the five raw materials mentioned above and rheology modifiers are added in the manufacturing process to 'modify' or 'fine-tune' the rheology. The inclusion of the rest of the paint components as a variable (*combination*) paid off, because its relative importance in affecting the rheology seems significant as it affects most of the rheology to the same extent as RM1.

5.2.2.2 MLR – PREDICTIVE MODEL

Multiple linear regression (MLR) models are developed to predict the rheological behaviour in terms of the preliminary model parameters $a_{FC}, b_{FC}, a_{HS}, \dots, C_{XLF}$ of the rheology curves

given in Table 5.1. The set of nine raw materials, combined with the pigment volume concentration (PVC) and volume solids (VS) are used as predictor variables. For example, the regression model for the mean response of the gradient of the frequency sweep (a_{FS}) is expressed as follows:

$$\begin{aligned} a_{FS} = & b_0 + b_{emulsion}x_{emulsion} + b_{pigment}x_{pigment} + b_{extender}x_{extender} + \\ & b_{org.pigment}x_{org.pigment} + b_{water}x_{water} + b_{solvent}x_{solvent} + b_{RM1}x_{RM1} + \\ & b_{RM2}x_{RM2} + b_{combination}x_{combination} + b_{PVC}x_{PVC} + b_{VS}x_{VS} + e \end{aligned} \quad (5.4)$$

The coefficients of determination (β_i) are only valid if the error of the independent variables (x_i) are negligible when compared with the error on the dependent variable (α_i). The independent variables of equation 5.4 are fixed points and therefore their variance are negligible compared to the variance of the dependent variables.

A total of 98 experimental runs (cases) are used for model fitting. The reason for using 98 experimental runs originates from the training data set that is used to train the neural network (see Chapter 6, Section 6.2.2 Neural Network Modelling). The aim is to, in the end, compare the *linear* modelling techniques (MLR) to the *non-linear* modelling techniques (NN) and for this reason, the models have to be fitted/trained on the same data.

Figure 5.3 illustrates the MLR model fit for a_{FS} on the 98 data points.

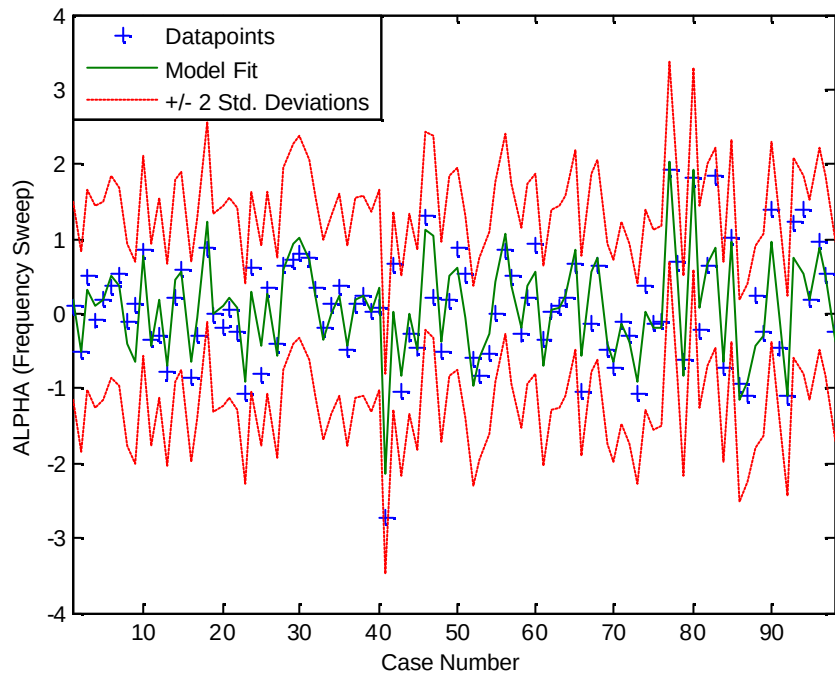


Figure 5.3 Model Fit (MLR) for ALPHA (Frequency Sweep)

All the points lie within plus or minus two standard deviations (std. deviations) of the model and therefore the model seems to explain the variance in the gradient of the frequency sweep (a_{FS}) well. However, the model needs to be validated before definite conclusions can be drawn.

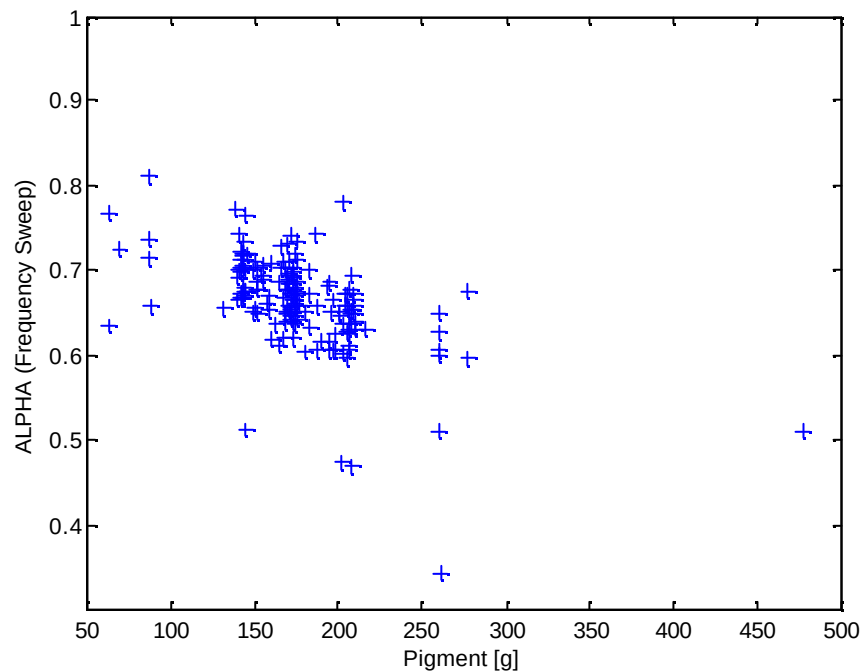
Model fits are constructed for all the preliminary model parameters (See Chapter 12, Appendix E)

5.2.3 MODEL VALIDITY

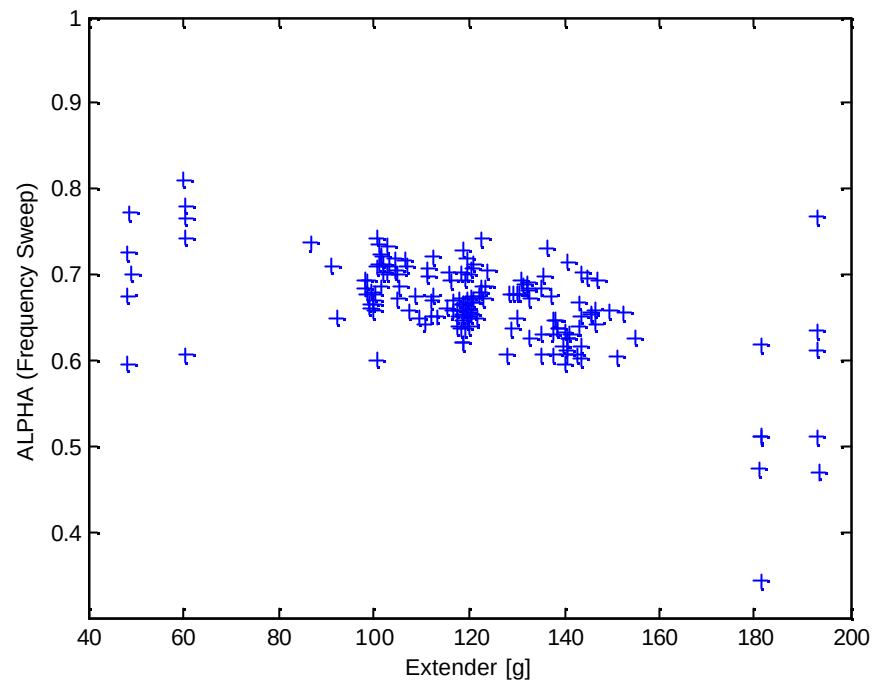
A valid model ensures that predictions are unbiased and efficient. Multiple linear regression is based on certain assumptions (see *Section 5.1.3 Model Validity*). It is important to remember that some assumptions rely stronger on certain conditions than others. Stronger assumptions lead to stronger conclusions. Ultimately, the model is validated by the average relative error that is made between the actual and predicted value of unseen data.

1. Linearity

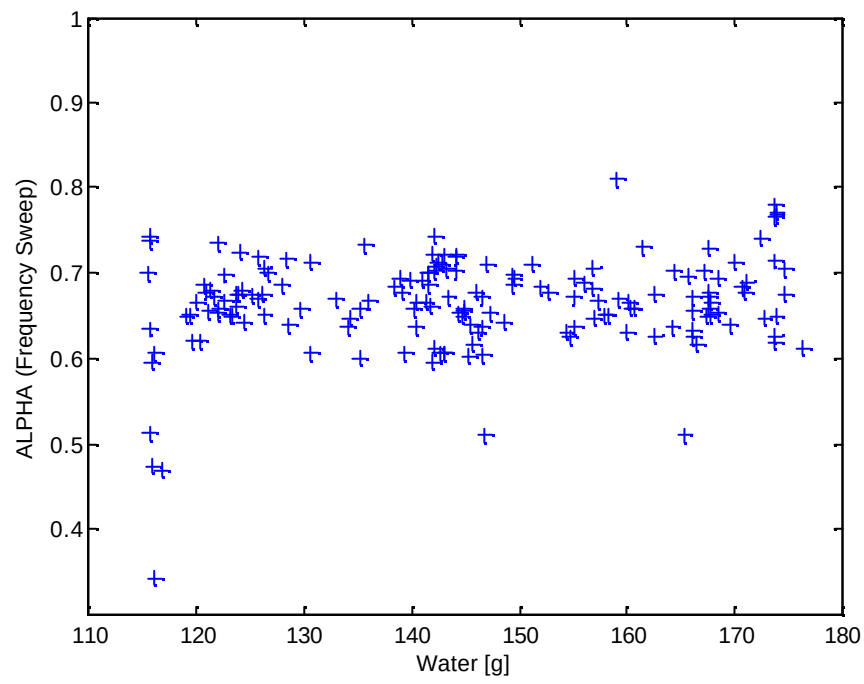
The models are based on the assumption that each raw material contributes in a linear way in predicting the rheology, although it is known that raw materials interact to produce certain rheology in the paint. These interactions might be small in comparison to the contribution of each raw material to the rheology, making the assumption of linearity not grossly impossible. An estimation of the degree of linearity can be obtained from the scatterplot of the measured response as a function of the predictor variable, x_i (j = raw materials, PVC and VS).



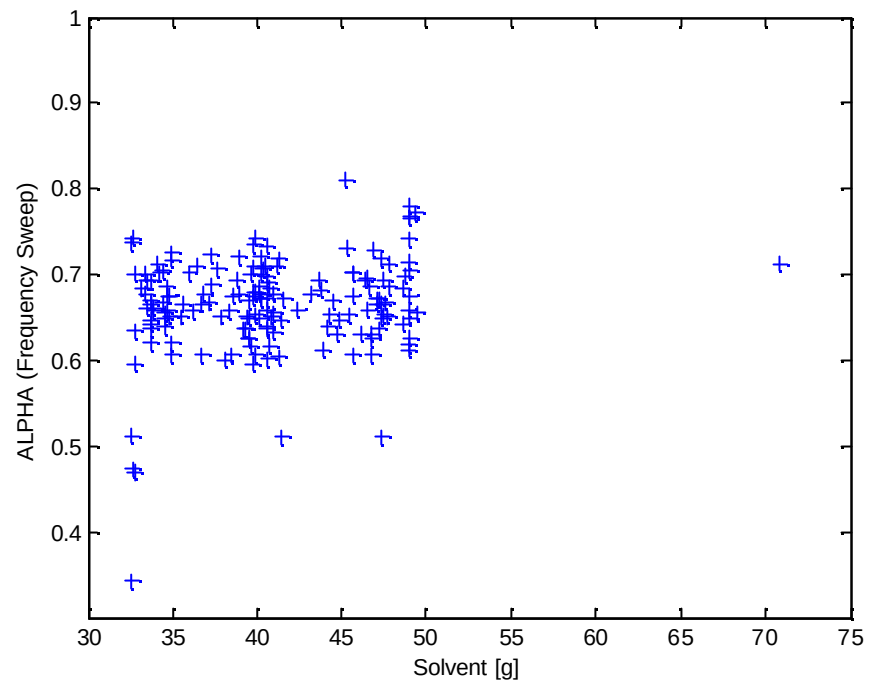
(a)



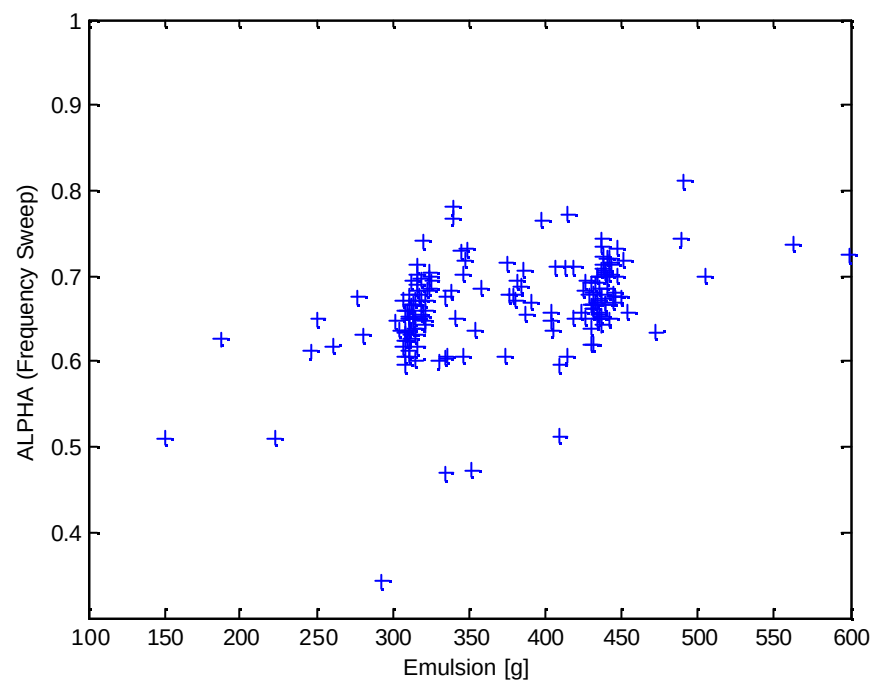
(b)



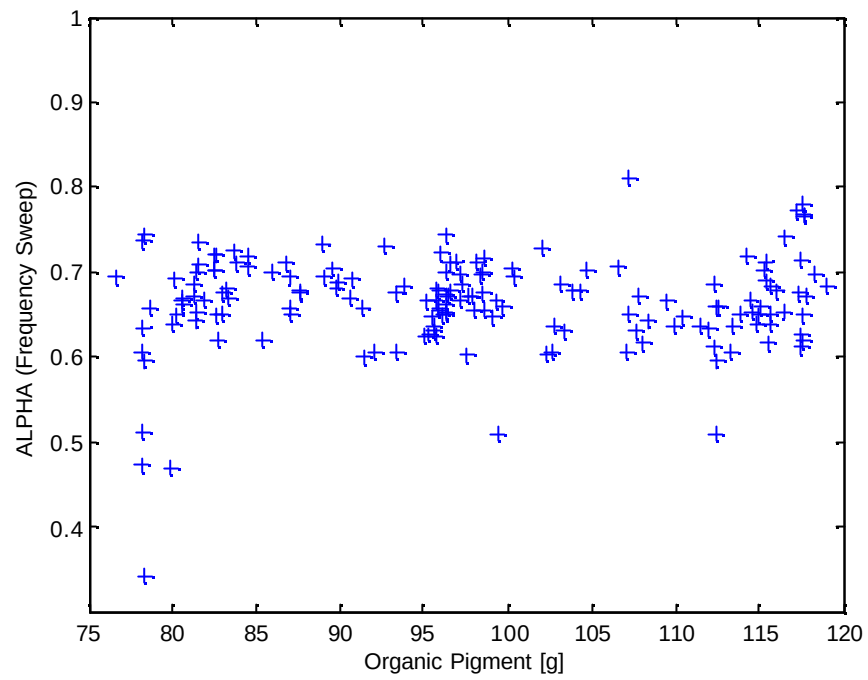
(c)



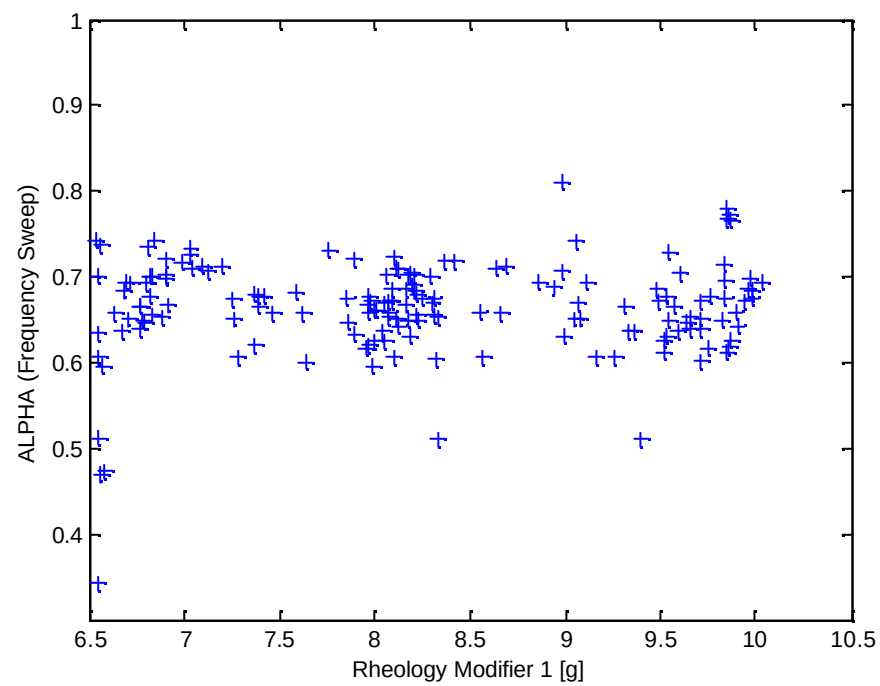
(d)



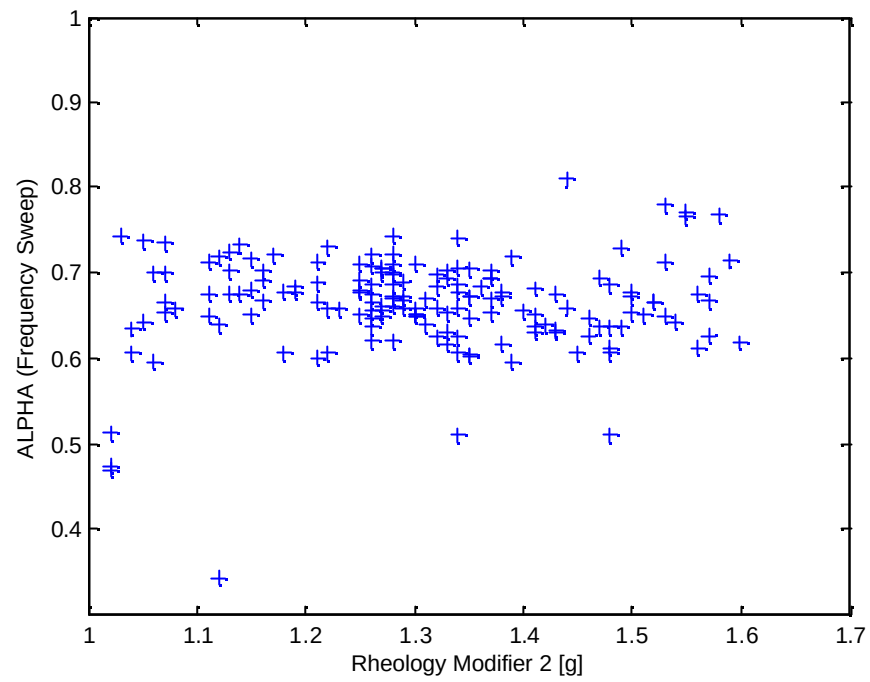
(e)



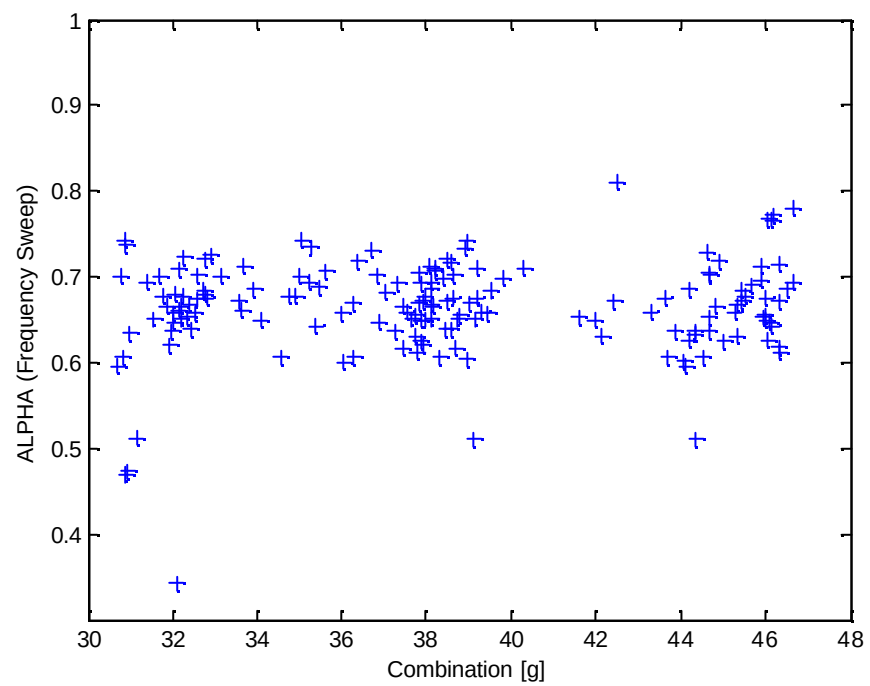
(f)



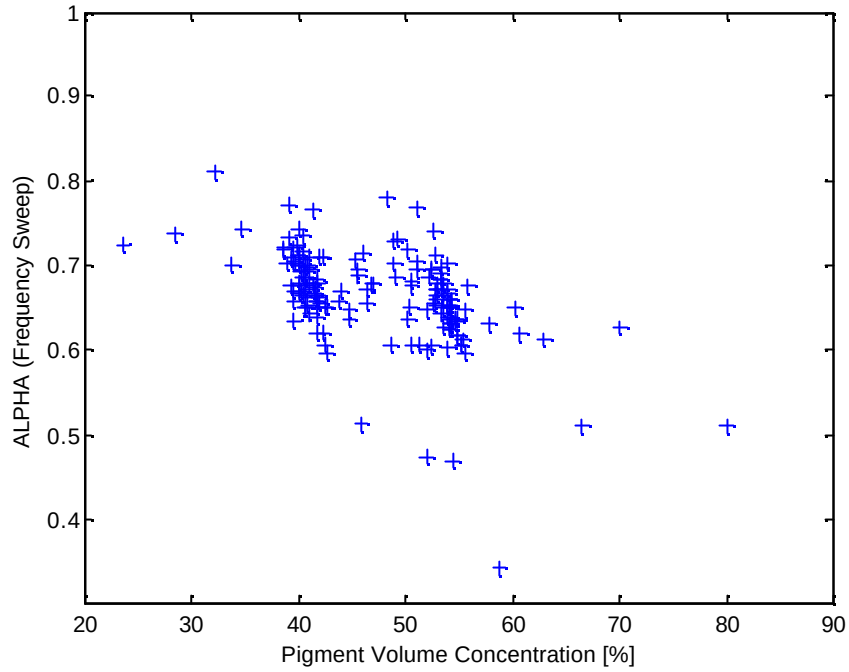
(g)



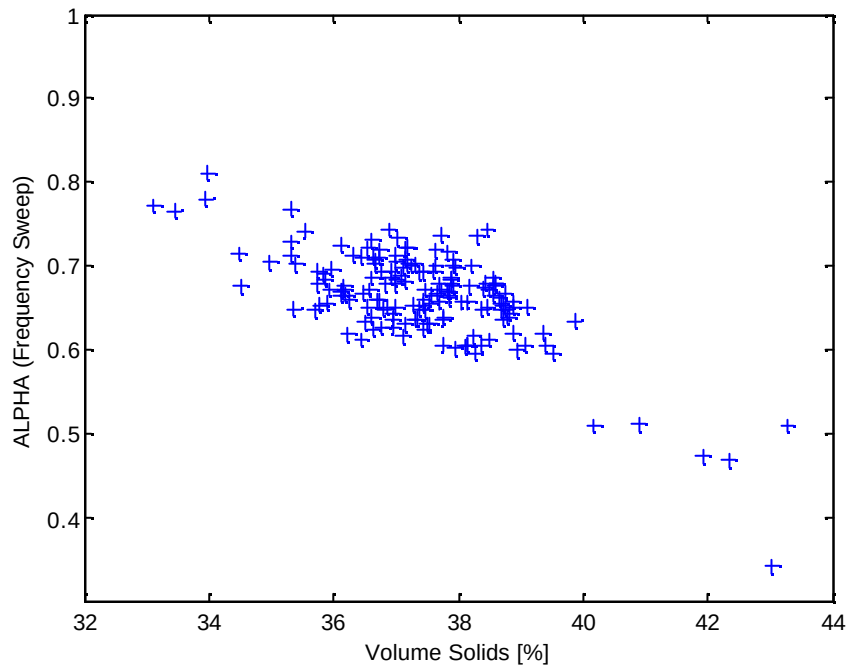
(h)



(i)



(j)



(k)

Figure 5.4 Scatterplot of ALPHA (Frequency Sweep) with Predictor Variables

Figure 5.4 (a) to (k) illustrate that some linear assumptions are more obvious for some of the predictor variables than for others, for example, the linear relationship between *volume solids* and a_{FS} (Figure 5.4 (k)) is more evident than for *pigment volume concentration* and a_{FS} (Figure 5.4 (j)). Some of the predictor variables indicate non-linearities at the extreme high and low ranges of

the variable (e.g. organic pigment, Figure 5.4 (f)), however, in most cases severe curvature does not exist and the assumption of linearity is not grossly violated.

Another indication that at least one regressor variable contributes significantly in explaining the linear relationship between the response and the regressor variables, is by rejecting the null hypothesis, that not *one* of the regressor variables is significant in explaining the variance in the response. The p-values in Table 5.3 ($p=0.0000$) indicate that there is definitely at least one regressor variable that plays a role in explaining the variance in the response. Therefore there exists at least one linear relationship between the rheological response and the raw materials.

Table 5.3 Evaluation of MLR Models for Rheology Data

Rheological Curve	Model Parameters	R^2	p-values
Flow curve	a_{FC}	0.5384	0.0000
	b_{FC}	0.7351	0.0000
High Shear	a_{HS}	0.7969	0.0000
	b_{HS}	0.8084	0.0000
	g_{HS}	0.9100	0.0000
Low Shear	a_{LS}	0.4288	0.0000
	b_{LS}	0.8418	0.0000
3-ITT (rot)	$a_{3-ITT(rot)}$	0.8639	0.0000
	$b_{3-ITT(rot)}$	0.8267	0.0000
Amplitude Sweep	a_{AS}	0.4825	0.0000
	b_{AS}	0.8499	0.0000
Frequency Sweep	a_{FS}	0.8317	0.0000
	b_{FS}	0.6689	0.0000
Time Sweep	a_{TS}	0.6520	0.0000
	b_{TS}	0.5508	0.0000
3-ITT (osc)	$a_{3-ITT(osc)}$	0.9676	0.0000
	$b_{3-ITT(osc)}$	0.5595	0.0000
Extra Low Frequency	C_{XLF}	0.4540	0.0000

In some cases the regression models explain a lot of the variance ($R^2 > 0.8$) in the rheological behaviour while in other it explains relatively little variance ($R^2 < 0.5$). It must be remembered that this is not necessarily an indication of what the predictive capabilities of the models are.

2. Mean Independence

The assumption of mean independence assumes that the independent variables are unrelated to the random error ε in equation 5.4. This is one of the most critical assumptions [4]. The scatterplot of residuals in Figure 5.5 also illustrates that the residuals for a_{FS} are centred round zero. The calculated average value of the residuals $\approx$ zero for all the models (see residuals of scatterplots and RCOPLOTS of Chapter 12, Appendix E).

3. Homoscedasticity

The plot of residuals of a_{FS} in Figure 5.5 also indicates that errors occur randomly and are not linked to patterns in the experimental run-order.

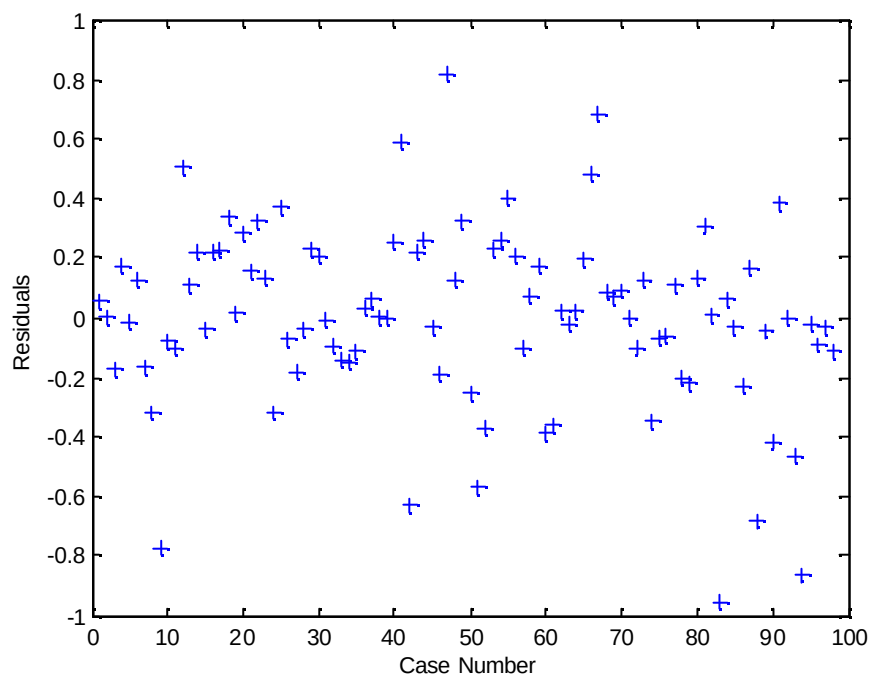


Figure 5.5 Plot of Residuals for ALPHA (Frequency Sweep)

4. Uncorrelated Errors

There is no easy way to determine whether contributions in the error term ε of equation 5.4 are correlated. However, it has been shown that the issue of correlated disturbances is strongly affected by sampling design [4] and that if samples are chosen at random order from an experimental space, that it is unlikely that correlated disturbances will be a problem. Furthermore, Design Expert[®] (the software package used for designing experiments) estimated the set of experiments so that multicollinearity between variables is kept at a minimum, reducing the possibility of correlation between error terms.

5. Normal Distribution

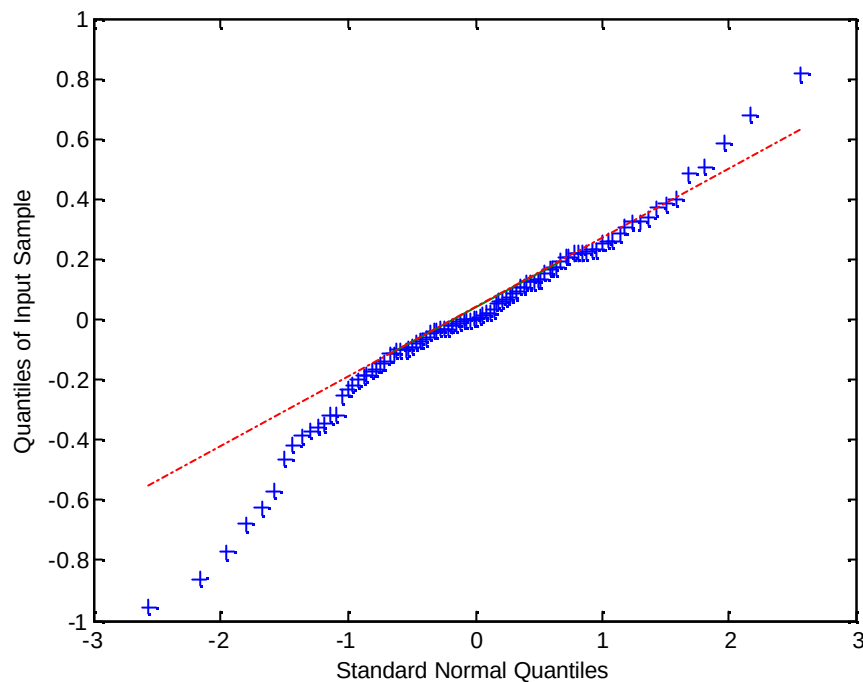


Figure 5.6 QQPLOT for ALPHA (Frequency Sweep) Data

The quantile-quantile plot of residuals in Figure 5.6 illustrates a definite S pattern which is an indication of nonnormality. The data points deviating from normality are either outliers or influential data points. It is unlikely that they are outliers as they do not occur randomly but deviate from normality in a definite S pattern. If these points are influential data points then a linear model is probably not the most adequate model. However, the assumption of normality seems valid for the majority of the data points. The occurrence of influential data points is not unlikely, because samples with extreme pigment volume concentrations (PVC) and volume solids (VS) are included in the design for the reason of being influential. Therefore, thirty extra samples, all with PVC and VS values far beyond that of the normal formula are included in the design to give insight into the paint properties that do not show a lot of variance in behaviour within the normal range of PVC and VS. It is expected that these samples show a certain degree of nonlinearity.

6. Influential Data Points

Influential data points affect the results that are included in the linear analysis. The residual case order errorbar plot (RCOPlot) of confidence intervals on residuals from the MLR, indicate the residuals that are considered 'abnormally' large (in terms of a normal distribution) and which are not fitted accurately with the MLR model (indicated with in Figure 5.7). These are either outliers or influential data points.

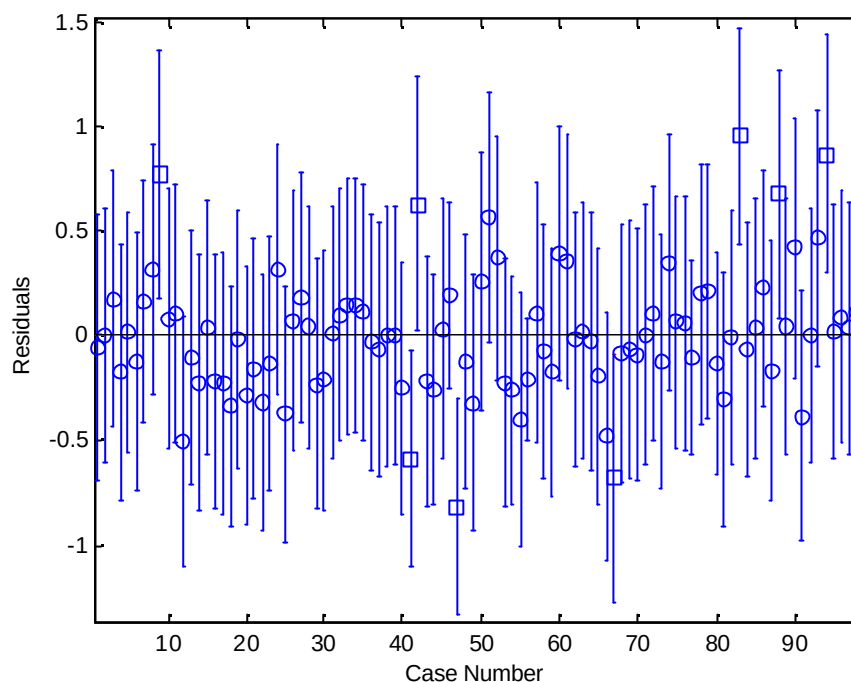


Figure 5.7 RCOPLLOT for AlphaFS with Square Markers that Indicate Outliers

Further investigation of Figure 5.7 reveals that 5 of the 8 (63%) residuals that lie outside the confidence interval are of paint samples with excessive PVC/VS loadings. Paint samples with excessive PVC/VS only make up 17% of the sample size and therefore the 63% occurrence of these samples outside the confidence interval clearly indicates that PVC and VS affect the model fit and therefore the results. Table 5.4 indicates the occurrence of paint samples with excessive PVC/VS as outliers according to the confidence interval of the MLR model.

Table 5.4 Occurrence of Extreme PVC/VS Paint Samples Outside of Confidence Interval

Rheology Curve	Model Parameters	Occurrence [%]
Flow curve	a_{FC}	33
	b_{FC}	0
High Shear	a_{HS}	75
	b_{HS}	80
	g_{HS}	0
Low Shear	a_{LS}	60
	b_{LS}	33
3-ITT (rot)	$a_{3-ITT (rot)}$	45
	$b_{3-ITT (rot)}$	67
Amplitude Sweep	a_{AS}	0
	b_{AS}	0
Frequency Sweep	a_{FS}	63
	b_{FS}	50
Time Sweep	a_{TS}	56
	b_{TS}	33
3-ITT (osc)	$a_{3-ITT (osc)}$	38
	$b_{3-ITT (osc)}$	33
Extra Low Frequency	C_{XLF}	33

It is evident that influential data points are abundant in all of the models. The MLR models struggle to fit the influential data points accurately which are most probably caused by nonlinearities in the data.

5.2.4 RHEOLOGY PREDICTION

Table 5.5 shows the regression coefficients used for rheology predictions. These regression coefficients are used in equations similar to equation 5.4 for predictions.

Table 5.5 Regression Coefficients for Model Parameters (MODEL 1)

Rheology Curve	Model Parameters	β_{constant}	β_{pigment}	β_{extender}	β_{water}	β_{solvent}	β_{emulsion}	$\beta_{\text{org.pigment}}$	β_{RM1}	β_{RM2}	$\beta_{\text{combinatio}}$	β_{PVC}	β_{vs}
FC	a_{FC}	0.0352	0.9055	0.8794	-0.8425	-0.2131	0.5127	0.5588	0.1104	0.0664	-0.3581	-0.9191	-1.7486
	b_{FC}	-0.0844	-1.8202	-1.8894	1.9686	0.5571	-0.3461	-0.6289	0.0579	0.0106	0.7064	1.0246	3.5326
HS	a_{HS}	0.0004	-2.5647	-2.8694	3.2835	0.9823	-0.4885	-0.8093	0.1539	-0.0431	1.1913	1.1759	5.7951
	b_{HS}	0.0008	3.3664	3.5128	-3.5716	-1.0529	0.7000	1.0743	0.0198	0.0591	-1.3508	-1.7946	-6.3242
	g_{HS}	0.0148	-0.0105	-0.6727	1.6459	0.5226	-0.5912	-0.0413	0.6174	0.1196	0.6958	-1.4406	3.3756
LS	a_{LS}	0.0307	-6.1485	-5.5253	0.6149	-0.0318	-1.8396	-2.0277	0.0477	0.0516	0.3792	6.3514	3.0137
	b_{LS}	-0.0452	-1.4435	-1.7700	2.7763	0.8531	1.2870	-0.4787	0.0615	0.0140	0.8945	1.5580	4.2789
3-ITT (rot)	$a_{3-ITT (rot)}$	0.0559	-3.0743	-2.8381	1.0983	0.3051	-0.5174	-1.1434	-0.1685	-0.2651	0.5060	4.1452	2.4009
	$b_{3-ITT (rot)}$	-0.0201	-0.9564	-1.4118	2.5252	0.6787	0.8001	-0.3130	0.2115	0.2203	0.8635	0.3603	4.2867
AS	a_{AS}	-0.0116	-0.3117	-0.8294	0.8541	0.1708	0.3183	0.0219	0.3403	0.1166	0.4469	-0.5676	1.4255
	b_{AS}	0.0354	-0.9061	-1.4392	2.7113	0.7193	1.0282	-0.4351	0.3182	0.1721	0.9461	0.5782	4.5662
FS	a_{FS}	-0.1329	2.9153	3.3398	-4.3469	-1.1319	-2.0297	1.0206	0.0080	-0.0861	-1.4177	-3.7037	-6.9758
	b_{FS}	-0.0430	-1.4255	-1.6483	2.2889	0.6066	0.4184	-0.4437	0.0639	0.0449	0.7231	0.9145	3.6700
TS	a_{TS}	-0.0859	-1.3623	-1.4156	1.0869	0.3114	0.8992	-0.5828	0.0193	0.0146	0.3100	2.4447	1.8227
	b_{TS}	0.0010	2.1882	1.4090	1.3037	0.2705	0.3435	0.6468	0.2506	0.1505	0.4529	-3.4339	1.5535
3-ITT (osc)	$a_{3-ITT (osc)}$	0.0034	-1.8025	-1.5825	0.8301	0.2466	0.4974	-0.7670	-0.1485	-0.0753	0.2555	3.5393	1.4413
	$b_{3-ITT (osc)}$	0.0423	3.3933	2.2488	1.6057	0.3008	1.0647	1.1016	0.3122	0.2772	0.3922	-4.7101	1.5693
XLF	C_{XLF}	-0.0726	-2.4736	-2.4978	2.4159	0.6790	0.1006	-0.7137	0.0009	0.0055	0.8219	1.8860	4.1712

5.2.4.1 MODEL PREDICTABILITY

The main purpose of the regression models is for prediction. The aim is to be able to construct rheological curves from unseen raw material data that closely resemble what the actual curve would have looked like. The raw materials in the regression equation 5.4 are used as predictor variables, $x_{emulsion}, x_{pigment}, x_{extender}, \dots, x_{VS}$. Figure 5.8 and Figure 5.9 give the 95% prediction interval which indicates with 95% certainty where future observations for the gradient (a_{FS}) and the offset (b_{FS}) of the frequency sweep are expected to be.

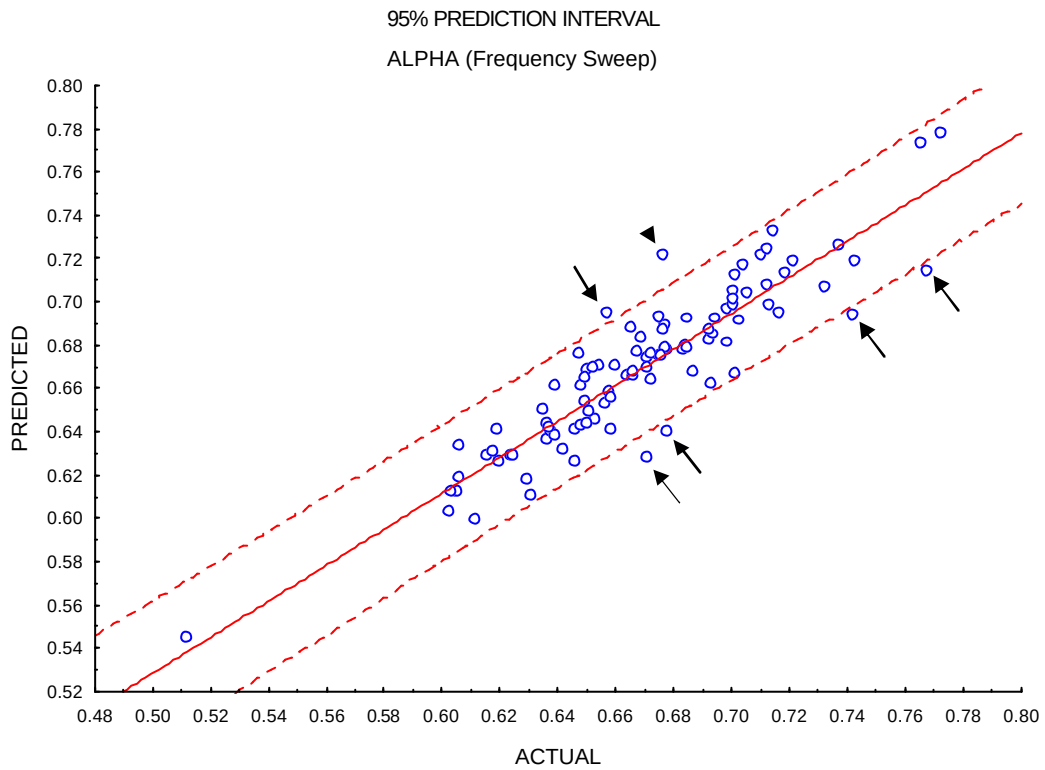


Figure 5.8 Prediction intervals of MLR model for *ALPHA*-values of the Frequency Sweep

One can therefore expect that most of the future predictions from unseen raw material data would fall within the 95% prediction interval as indicated in Figure 5.8. However, $\pm 6\%$ (5 out of 98 samples) of predicted data points will fall outside of the 95% prediction interval.

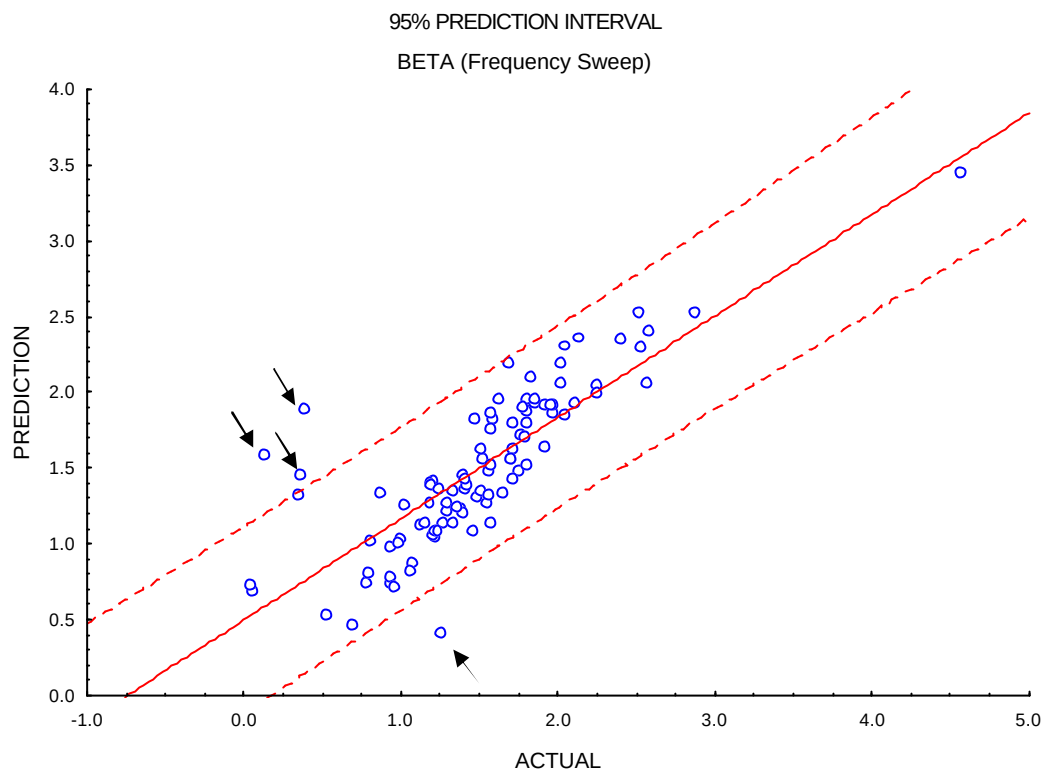


Figure 5.9 Prediction Intervals of MLR Model for $BETA$ -values of the Frequency Sweep

Once again, although the majority of predictions for b_{FS} will fall within the 95% prediction interval, Figure 5.9 indicates that $\pm 4\%$ (4 of 98 samples) of predicted values can be expected to fall outside the 95% prediction interval.

Closer inspection of these 'outliers' reveals something that was already suspected after evaluation of the RCO PLOT in Figure 5.7. That is, four of the six (67%) data points that fall outside the 95% prediction interval in Figure 5.8 are samples with extreme PVS/VS. In Figure 5.9 one out of four (25%) data points outside the 95% prediction interval has extreme PVC/VS. In both cases this is more than the 17% occurrence of samples with extreme PVS/VS in the design space and therefore it can be concluded that paint samples with extreme PVC/VS are more likely to fall outside of the 95% prediction interval. The RCO PLOT of Figure 5.7 has indicated that these data points with extreme PVC/VS lead to large residuals. Therefore, large residuals can be expected for points that fall outside the 95% prediction interval.

All the models are presented with 'unseen' raw material data that have not been used in estimating the model fit. The 'unseen' raw material data fall within the same range as the fitted data and therefore it can be expected that the models will predict the data points in a way similar to the model fit and the prediction intervals. Therefore, it can be expected that the models will not be accurate in predicting the rheological response of samples with extreme PVC/VS. Figure 5.10 gives the predicted a_{FS} values against the actual a_{FS} values of 33 'unseen' samples.

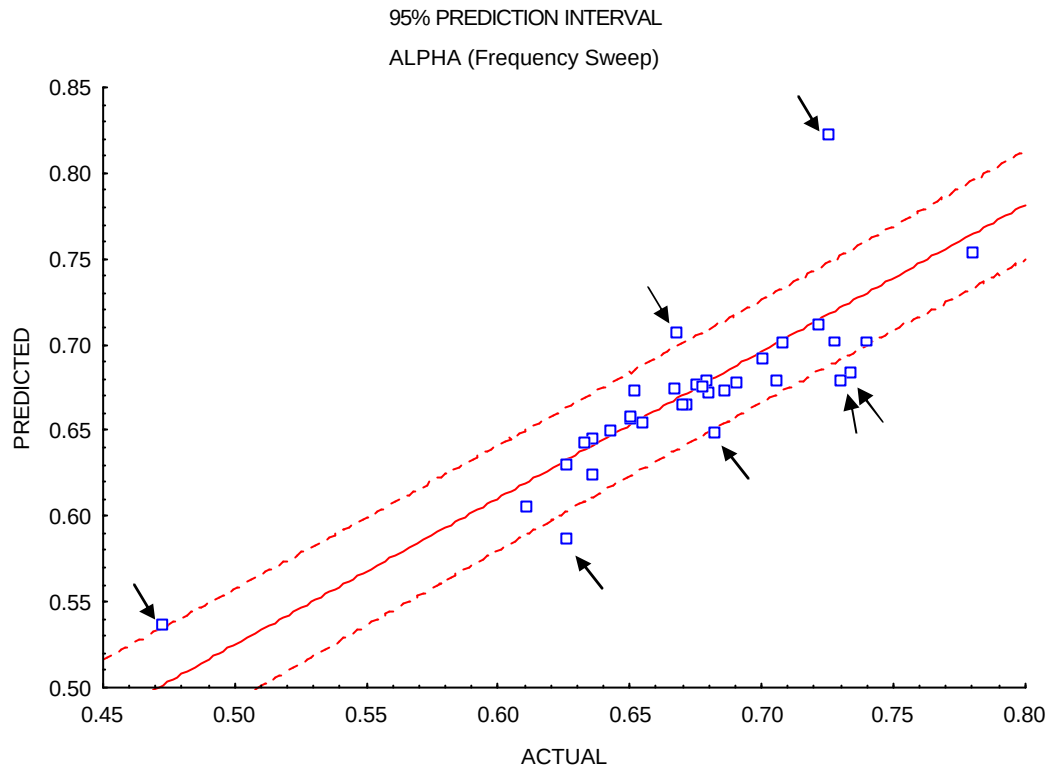


Figure 5.10 Model Predictions for ALPHA (Frequency Sweep) for Unseen Data

The following becomes clear under closer inspection of Figure 5.10:

- Most of the data points ($\pm 80\%$) are predicted within the 95% prediction interval
- Seven predicted data points ($\pm 20\%$) can be considered as possible outliers (outside the 95% prediction interval)
- Closer inspection of the data reveals that 3 of the 7 “outliers” are samples with extreme PVC/VS
- The unseen data only have 4 samples with extreme PVC/VS and therefore 75% of samples with extreme PVC/VS is predicted as “outliers” outside the 95% prediction interval. The “outliers” are therefore influential data points that the linear model cannot predict accurately.

Figure 5.11 gives the predicted b_{FS} values against the actual b_{FS} values for the 33 ‘unseen’ samples.

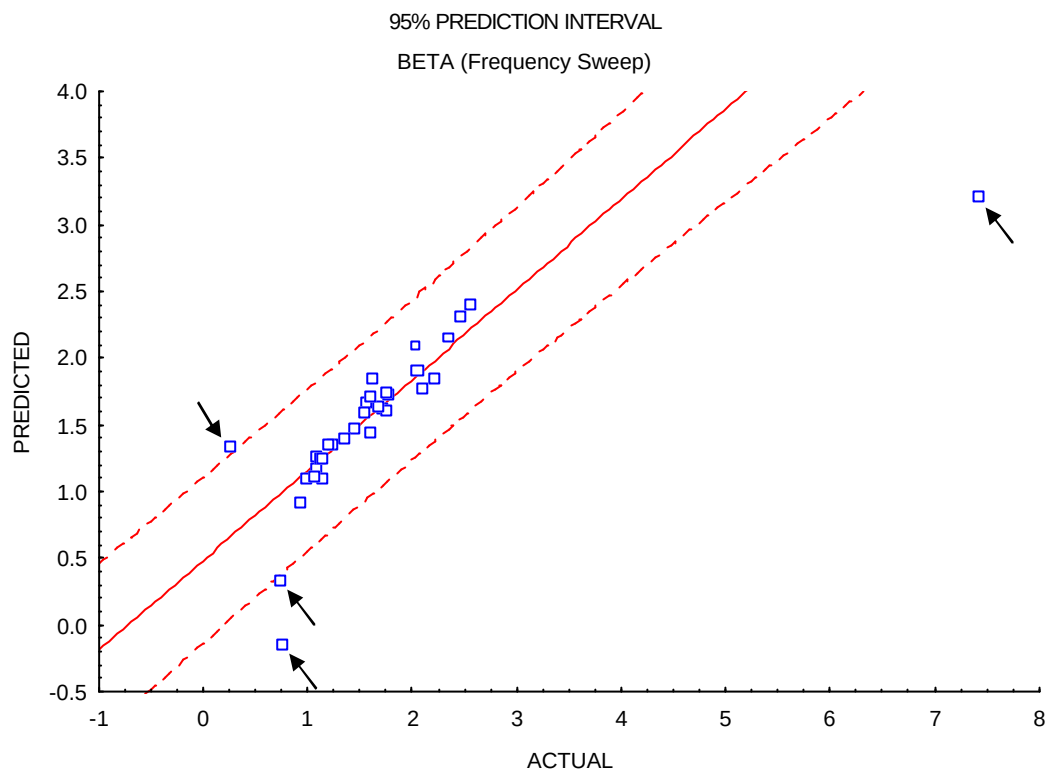


Figure 5.11 Model Predictions for BETA (Frequency Sweep) for Unseen Data

The following is revealed under closed inspection:

- Most of the data points ($\pm 87\%$) are predicted within the 95% prediction interval
- Four predicted data points ($\pm 13\%$) can be considered as possible “outliers”.
- Three of these four “outliers” (75%) are samples with extreme PVC/VS.

It is clear that linear models can accurately predict the values of a_{FS} and b_{FS} as long as the raw materials fall in the normal plus or minus 20% range of the standard formulation. However, paint samples with extreme PVC/VS are influential data points which the MLR models do not predict accurately and non-linear modelling techniques seem to be necessary to model these points accurately. In Chapter 6 a comparison is drawn between the magnitude of error obtained from linear versus non-linear modelling techniques for the data points that are considered to be outliers according to the MLR models.

5.2.4.2 MODEL ERRORS

Average errors for the prediction of the 33 ‘unseen’ paint samples are given in Table 5.6. In general the MLR models predict the rheological behaviour with relative small relative errors, considering that inaccurate predictions of samples with extreme PVC and VS contribute significantly in affecting the relative error. The error will reduce considerably if the samples with extreme PVC/VS can also be predicted accurately. This is one of the main reasons for implementation of non-linear modelling (Chapter 6).

It is important to know whether these errors in prediction of rheological parameters are significant in affecting the paint performance. Unfortunately, seeing that this type of study is attempted for the first time, no error margins exist which state what margin of error is allowed in the rheological behaviour before the paint performance gets affected. This is one of the reasons why MODEL 3 is developed – to establish a quantitative model/scale so that these type of errors can be evaluated quantitatively and determine whether they are acceptable in terms of the sensitivity of the paint properties.

Table 5.6 Relative Errors for Unseen Data

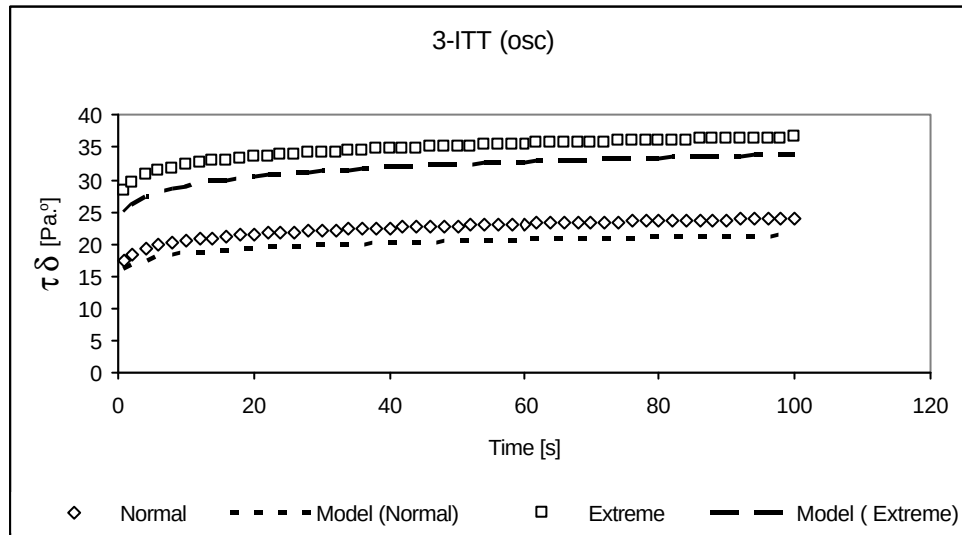
Rheology Curve	Rheological Parameter	Relative Error [%]	
		PVC/VS included as predictors	PVC/VS excluded as predictors
Flow Curve	a_{FC}	3.24	3.42
	b_{FC}	14.7	21.33
High Shear	a_{HS}	10.83	12.44
	b_{HS}	14.34	18.73
	g_{HS}	2.88	2.93
Low Shear	a_{LS}	51.62	46.86
	b_{LS}	15.46	19.45
3-ITT (rot)	$a_{3-ITT(rot)}$	10.87	14.82
	$b_{3-ITT(rot)}$	8.32	8.36
Amplitude Sweep	a_{AS}	0.42	0.43
	b_{AS}	5.89	6.31
Frequency Sweep	a_{FS}	2.97	3.03
	b_{FS}	24.81	25.74
Time Sweep	a_{TS}	21.13	18.64
	b_{TS}	21.92	25.75
3-ITT (osc)	$a_{3-ITT(osc)}$	8.03	14.82
	$b_{3-ITT(osc)}$	23.41	20.12
Extra Low Frequency	C_{XLF}	44.22	49.54

It was mentioned in Chapter 4 that the inclusion of pigment volume concentration (PVC) and volume solids (VS) as predictor variables add to the multicollinearity of the predictor variables. It is known that multicollinearity does not play a major role when a regression model is used for prediction purposes [1], [4], [16]. In Table 5.6 it can be seen that inclusion of PVC and VS as predictor variables actually leads to better predictions with 14 of the 18 predicted parameters having smaller errors than without the inclusion of PVC and VS.

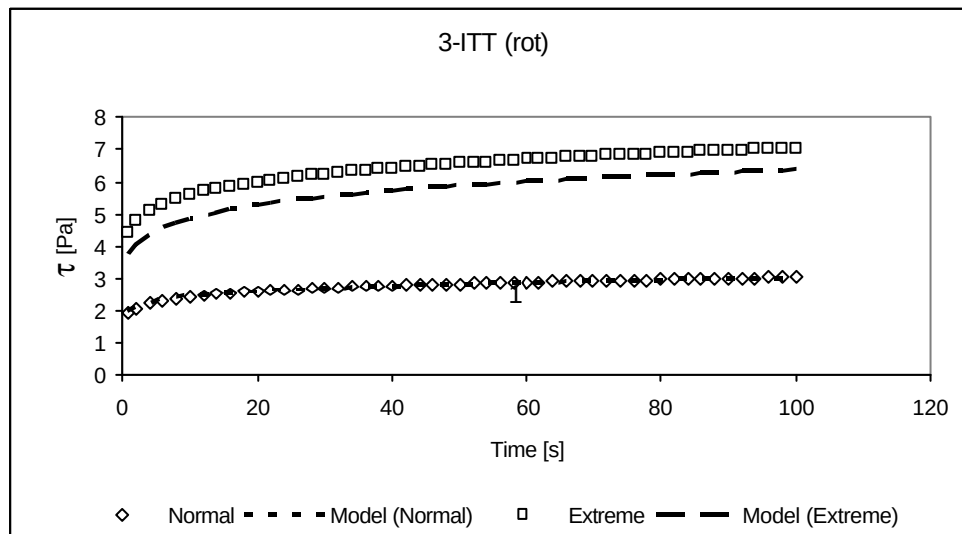
5.2.4.3 PRACTICAL APPLICABILITY OF MODEL

The rheological parameters obtained from the model fit can be used to predict the response to previously “unseen” raw material data. “Unseen” raw material data are data that is not used to develop the model fit, but rather to test the model parameters or to see how well the model can predict the rheological response given only the raw material levels. Measured and model predictions for the rheology curves are shown in Figure 5.12 (a) to (i). The entire model prediction is obtained from the predicted a_i and b_i values.

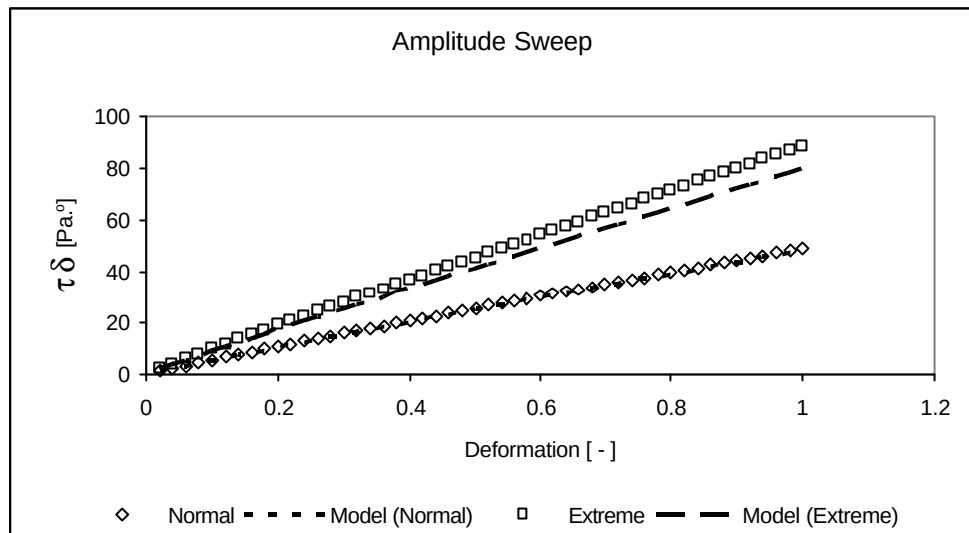
Each figure illustrates the model prediction for a paint sample with ‘normal’ variation in raw material levels (<20%) and a model prediction for a paint sample with ‘extreme’ variation in raw material levels (>20%). It is clear from all the graphs that the model predictions for ‘normal’ variations in raw material levels are much more accurate than for model predictions where extreme variations in raw material levels occur. In other words, the model is more accurate in predicting the rheological response for normal variations than for extreme variations. It is clear that another model is required for extreme variations. Attention to this fact is given in Chapter 6.



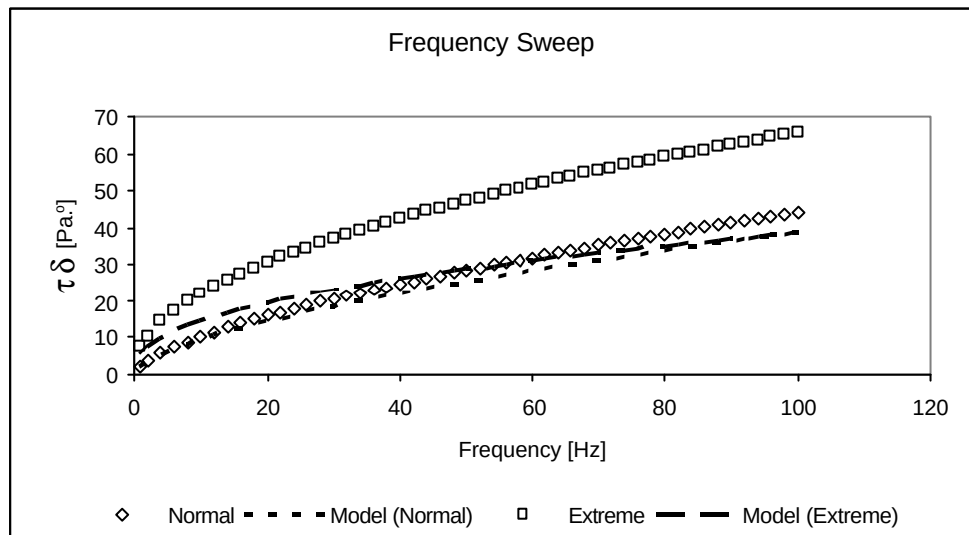
a)



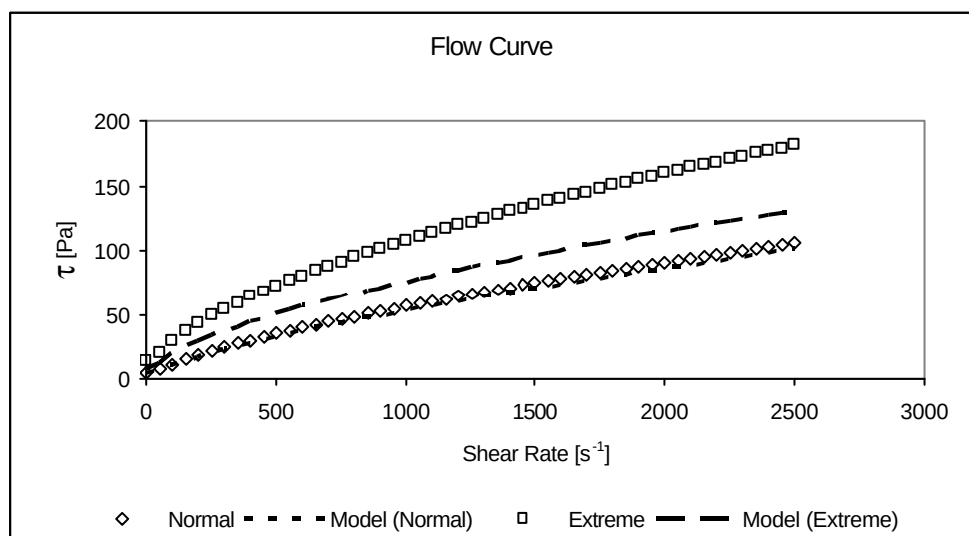
b)



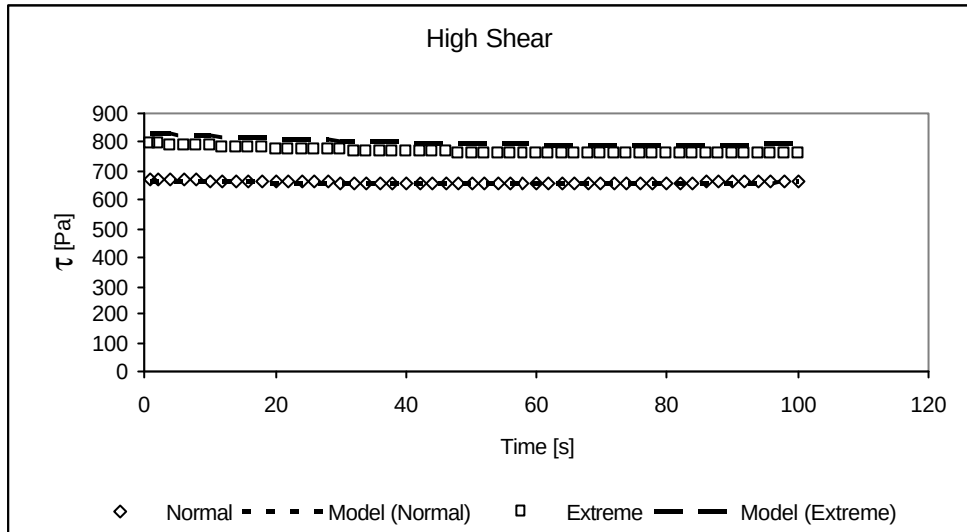
c)



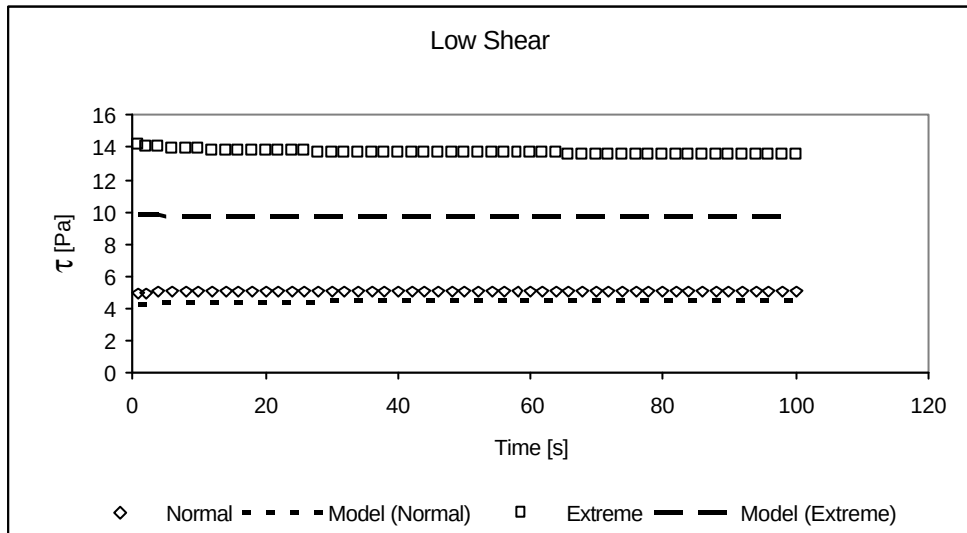
d)



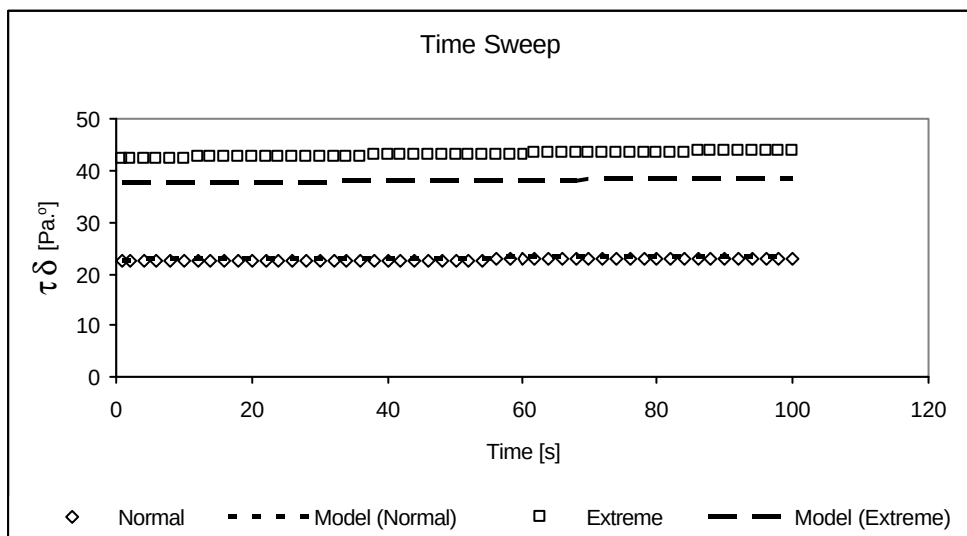
e)



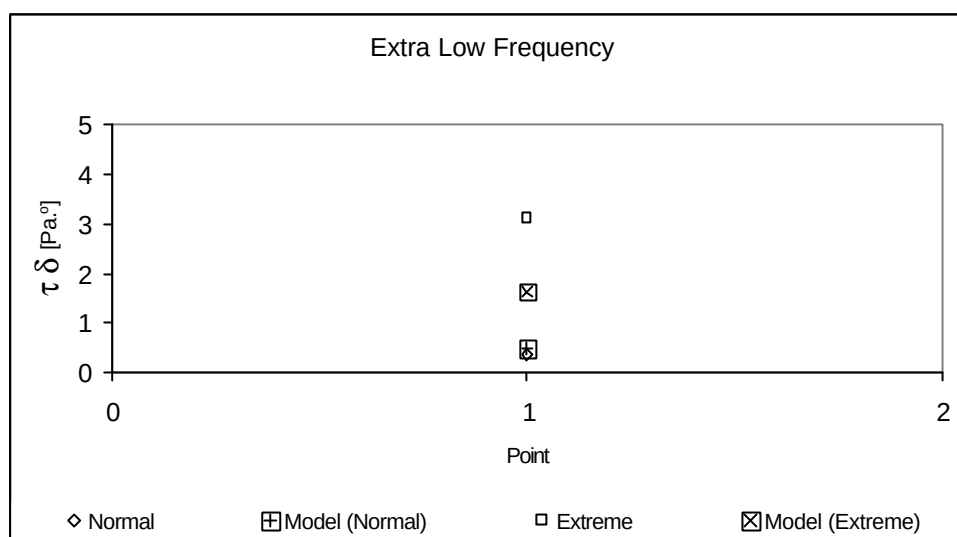
f)



g)



h)



i)

Figure 5.12 Predicted Response for Rheology Curves

5.2.5 MODEL 1 CONCLUSIONS

The following conclusions can be drawn from the previous discussion:

- Rheology data is highly correlated and three principal components explain 80% of the variance in the rheological data that was originally characterised by 18 parameters.
- The most important raw materials affecting the rheology are (in order of importance): emulsion, extender, pigment, water, organic pigment, solvent, combination of remaining raw materials, type 1 rheology modifier (RM1), type 2 rheology modifier (RM2).
- Multiple linear regression (MLR) is a valid method for rheological modelling of paint.
- Extreme PVC/VS have an effect on the predictability of the linear models. The MLR models predict the rheological behaviour of paint with normal ranges of raw materials accurately, but cannot predict the rheological behaviour of samples with extreme PVC/solids accurately. The rheological behaviour of paint is therefore relatively linear as far as the normal range (plus or minus 20% variation in a raw materials from the standard formulation) of raw materials is concerned, but it seems that nonlinearities start to play an important role at wider ranges of raw materials. Therefore, beyond the normal range of raw materials, linear models are not accurate in predicting the rheological behaviour. The use of linear models for extrapolation beyond the normal ranges can not be used accurately.
- Table 5.5 gives the matrix of regression coefficients that can be used in equation 5.4 to predict the rheological behaviour of this specific paint.
- The rheological behaviour of more than 80% of new paint formulations can be predicted with 95% confidence by using MLR, as long as the raw materials do not vary more than 20% from the standard formulation. Different modelling techniques are required for the models that vary more than 20% from the standard formulation.

- Table 5.6 gives the error that can be expected in predicting the rheology curves by using MLR. The smallest error (0.42%) occurs in predicting the gradient of the amplitude sweep while the largest error (44.22%) can be expected in predicting the extra low frequency point. The rheology curves that are predicted most accurately by means of MLR are the amplitudes sweep, 3-ITT in the rotational mode and the flow curve. This is very good, because the most essential rheological information can be obtained from these curves:
 - i) Amplitude Sweep: Gives information about the structure of the paint.
 - ii) 3-ITT: Gives information about the way the paint behaves after it has been applied to a surface.
 - iii) Flow Curve: Gives the basic flow behaviour of the paint over a range of shear rates.
 The rheology curves that are predicted most inaccurately by MLR are the low shear curve, the time sweep and the extra low frequency point. This is not a such a big concern because these rheology curves give less valuable information about the paint and its behaviour:
 - i) Low Shear Curve: Gives information about the paint under low shear conditions, but this can also be obtained from the flow curve.
 - ii) Time Sweep: Gives information of paint when subjected to long stationary periods, which is mostly only during storage.
 - iii) Extra Low Frequency point: Gives information when subjected to very slow time-scale behaviour (which is usually long after application when paint is already dry).

5.3 MODEL 2: RELATIONSHIP BETWEEN PAINT PROPERTIES AND RAW MATERIALS

5.3.1 BACKGROUND

The relationship between raw materials and the paint properties are well established, albeit qualitatively, but it is surprising to note how little quantitative modelling has been done on paint properties. Due to all the possible interactions between the raw materials, the effect of raw materials on paint properties are unique for every paint system, and the major effects in Table 5.7 can only be used as a indication.

Table 5.7 Major Functionality of Raw Materials

Raw Material	Major effects
Emulsion (binder)	Durability, gloss
Pigments	Hiding, colour, opacity, hardness, durability
Extenders	Hiding power
Vehicle (water)	Application
Solvents	Open time, adhesion, water resistance
Coalescent	Film formation, defamer
Dispersants	Stability (settlement, syneresis), colour acceptance
Surfactants	Colour acceptance, stability
Defamers	Defaming
Thickeners	Application
Biocides	Preservative (in-can and film)

5.3.2 MULTIPLE LINEAR REGRESSION (MLR)

Multiple linear regression (MLR) is also used as a first technique to investigate the linear relationship between the raw materials and the paint properties. This is because i) other authors have successfully (to a certain extent) used linear modelling techniques such as MLR to model paint properties (see Chapter 2, *Section 2.4.2 Linear Modelling*) and ii) MLR gives a good first approximation of the degree of linearity that exists between the paint properties and the raw materials.

Furthermore, the uniqueness of each paint formula makes comparison of models between this study and others almost impossible. The aim is therefore to develop the quantitative relationship (model) between the raw materials and paint properties for this *specific system*. It can then be extended to other paint systems (*Chapter 7, Future Work*).

In this part of the study a model is mainly developed for predictive purposes and therefore the causal relationship between the paint properties and the raw materials is only briefly discussed.

Gloss is important in paints, particularly where high aesthetic appeal is required and in high-quality paints [17]. The relationship between gloss and the raw materials is therefore important and it is this property (gloss) that is used as an example for the modelling process. The paint under investigation is classified as a 'medium gloss' paint and therefore *Gloss60* (gloss measured at a reflectance angle of 60°) is ideal in expressing the gloss values.

5.3.2.1 MLR – CAUSAL MODEL

The degree of multicollinearity that exists between the raw materials is still acceptable ($C < 15$ [5]) for the MLR model to be used as a causal model. However, when variables such as pigment volume concentration (PVC) and volume solids (VS) are added, multicollinearity becomes a problem ($C \gg 15$) for the MLR model to explain the causal relationship between the explanatory variables and the paint properties. Therefore the causal relationship between *Gloss60* and the raw materials is expressed as follows (without PVC and VS as regressor variables).

$$\begin{aligned} a_{Gloss60} = & b_0 + b_{emulsion}x_{emulsion} + b_{pigment}x_{pigment} + b_{extender}x_{extender} + \\ & b_{org.pigment}x_{org.pigment} + b_{water}x_{water} + b_{solvent}x_{solvent} + b_{RM1}x_{RM1} + \\ & b_{RM2}x_{RM2} + b_{combination}x_{combination} + e \end{aligned} \quad (5.5)$$

Figure 5.13 illustrates the absolute values of the regression coefficients which indicate their importance in affecting the *Gloss60* paint property.

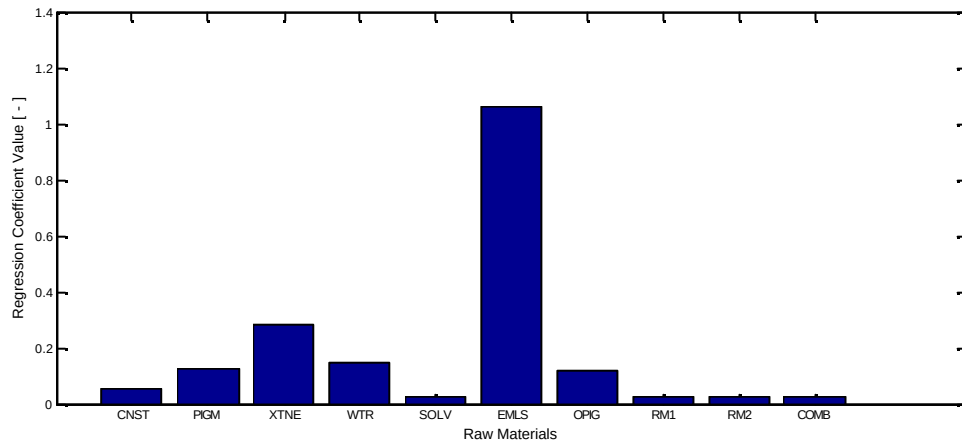


Figure 5.13 Regression Coefficients (Absolute) for Gloss60 MLR Model

(CNST=constant, WTR=water, PIGM=pigment, XTND=extender, SOLV=solvent, EMLS=emulsion, OPIG=organic pigment, RM1=rheology modifier1, RM2=rheology modifier2, COMB=combination.)

The importance of each raw material in Figure 5.13 is given by its relative height. It can be seen that the importance of the emulsion overshadows that of the other raw materials. According to paint scientists, gloss is mostly affected by the combination of resin with the pigments and any or all of the following can cause gloss deficiency [17]:

- Poor level of dispersion and the presence of very large pigment particles or pigment agglomerates.
- Insufficient dispersion stability which leads to flocculation of pigment particles.
- Pigment concentration which is too high.

- An unfavourable rheology that leads to poor levelling.

Figure 5.13 also illustrates the importance of the emulsion, pigments (organic and inorganic), extender and water in affecting the gloss of the paint.

5.3.2.2 MLR – PREDICTIVE MODEL

Multiple linear regression (MLR) is used again to model the relationship between a specific paint property (*Dry Burnish 20*, *Dry Burnish 60*, etc.) and the set of 11 input variables. For example, the regression model for the mean response of *Gloss 60* is expressed as follows:

$$\begin{aligned} a_{Gloss60} = & b_0 + b_{emulsion}x_{emulsion} + b_{pigment}x_{pigment} + b_{extender}x_{extender} + \\ & b_{org.pigment}x_{org.pigment} + b_{water}x_{water} + b_{solvent}x_{solvent} + b_{RM1}x_{RM1} + \\ & b_{RM2}x_{RM2} + b_{combination}x_{combination} + b_{PVC}x_{PVC} + b_{VS}x_{VS} + e \end{aligned} \quad (5.6)$$

Similar to equation 5.4, the independent variables of equation 5.6 are fixed points and therefore their variance are negligible compared to the variance of the dependent variables.

Note that x_{PVC} and x_{VS} are also used as predictor variables. As mentioned earlier, the effect of multicollinearity is minor when the regression equation is used for predictive purposes. The model fit of the 98 samples can be seen in Figure 5.14

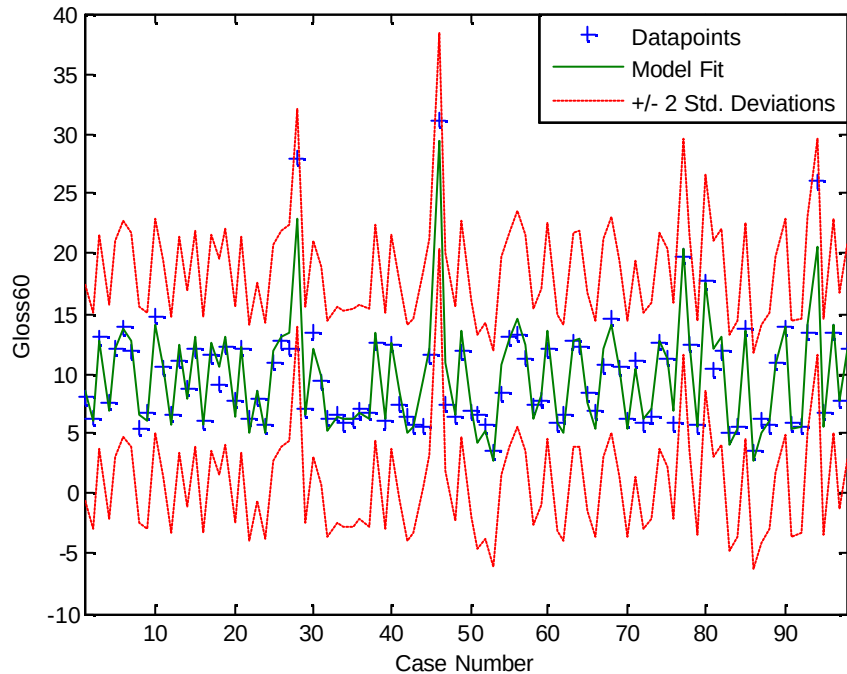


Figure 5.14 Model Fit (MLR) for Gloss 60

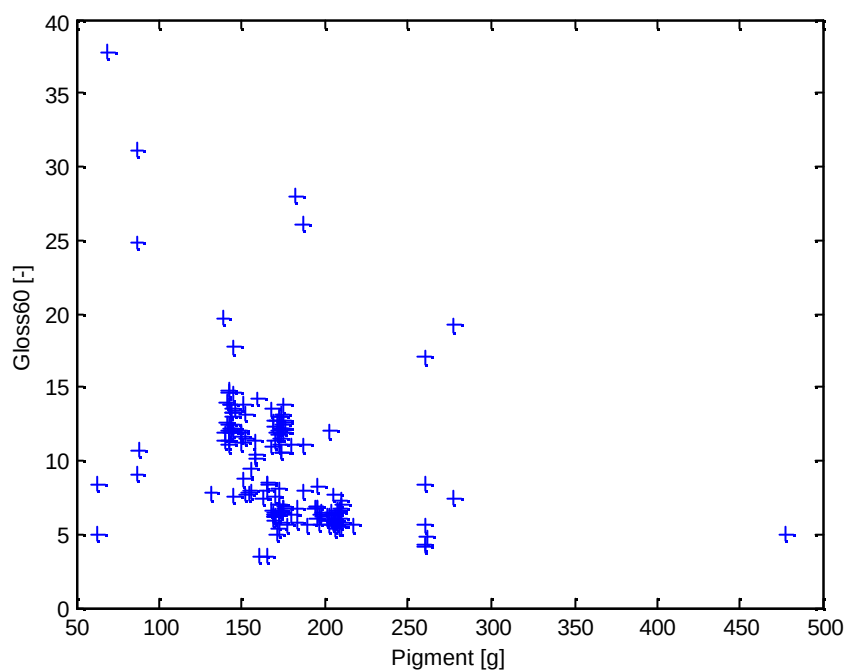
Figure 5.14 gives an indication of how accurate the model fits the data. Except for three extreme points, the model seems to fit the data well. The model assumptions need to be validated first before more serious conclusions can be made about the model fit and its predictions.

5.3.3 MODEL VALIDITY

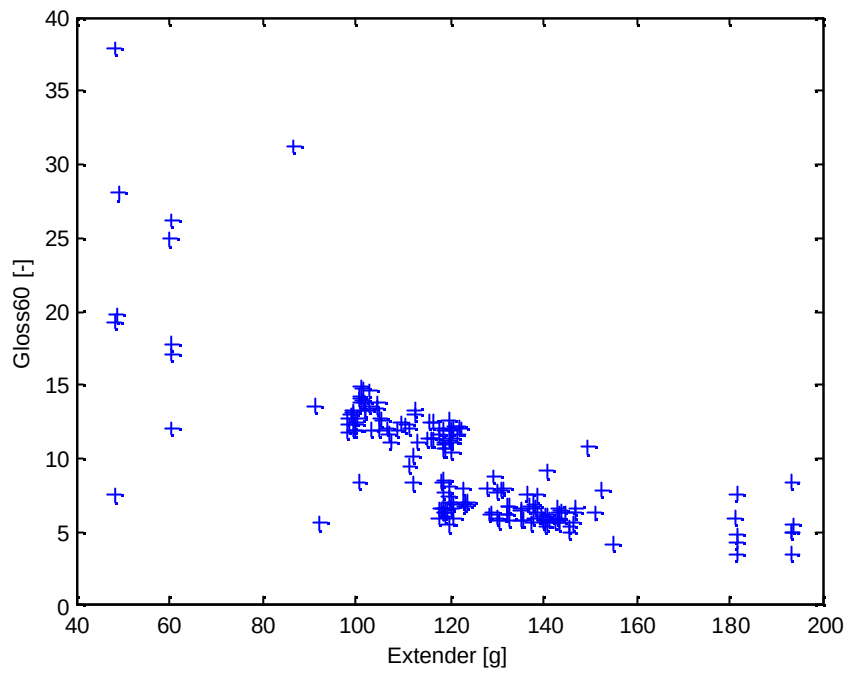
The model validity is based on the same assumptions as in Section 5.2.3.

1. Linearity

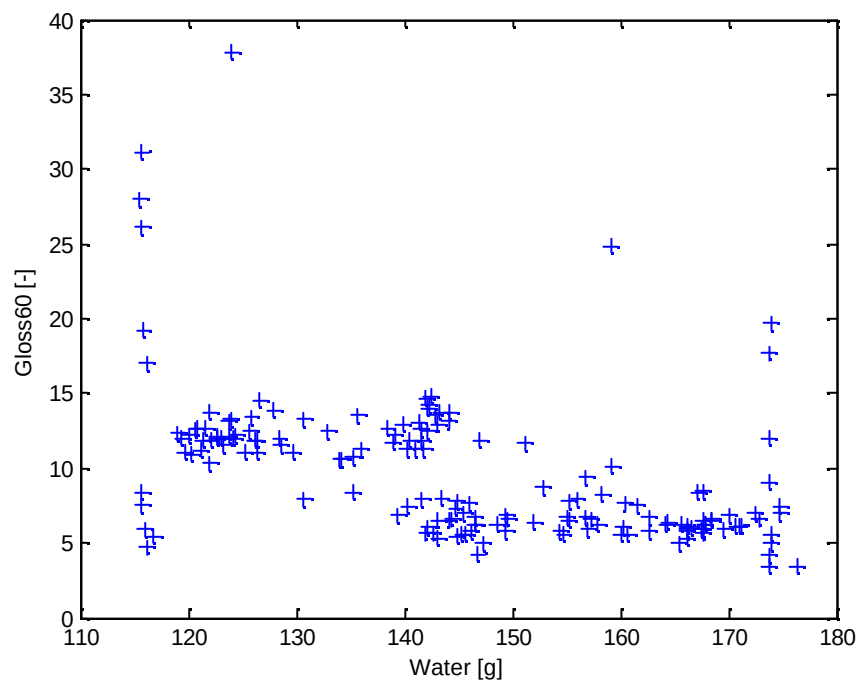
A scatterplot of *Gloss60* against each of the predictor variables (Figure 5.15 (a) to (k)) gives an indication of the linear relationship between the gloss of the paint and each of the predictor variables.



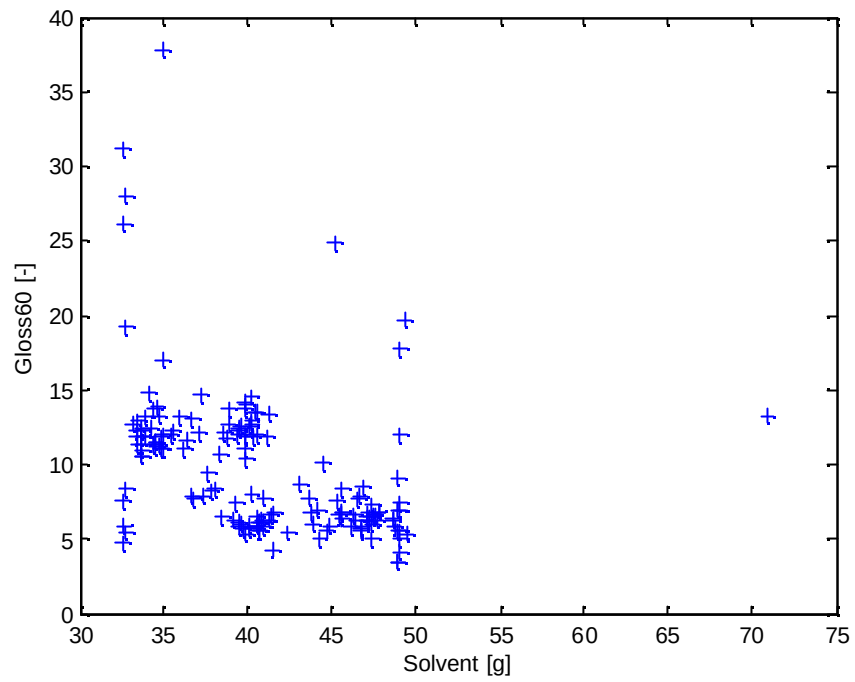
(a)



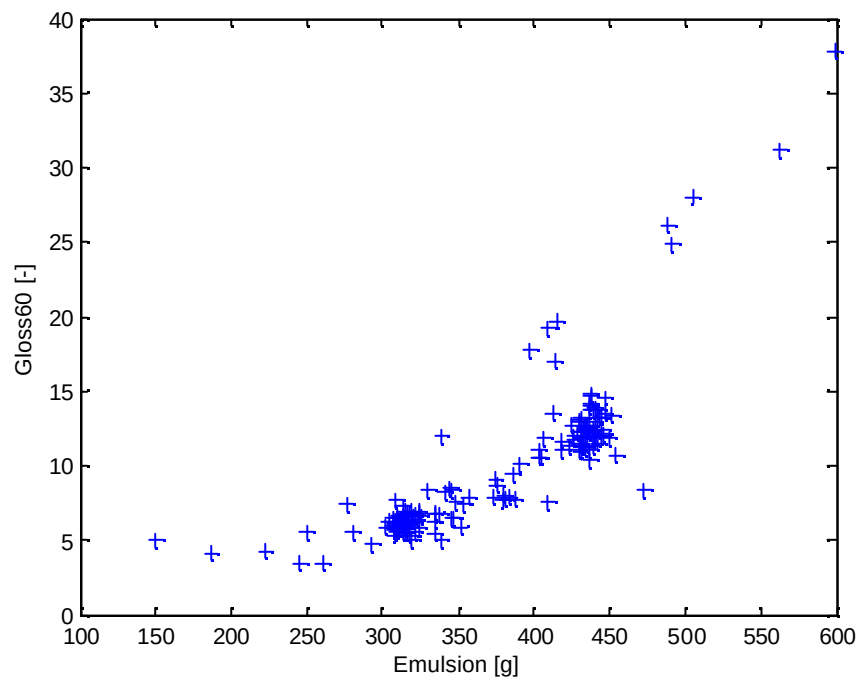
(b)



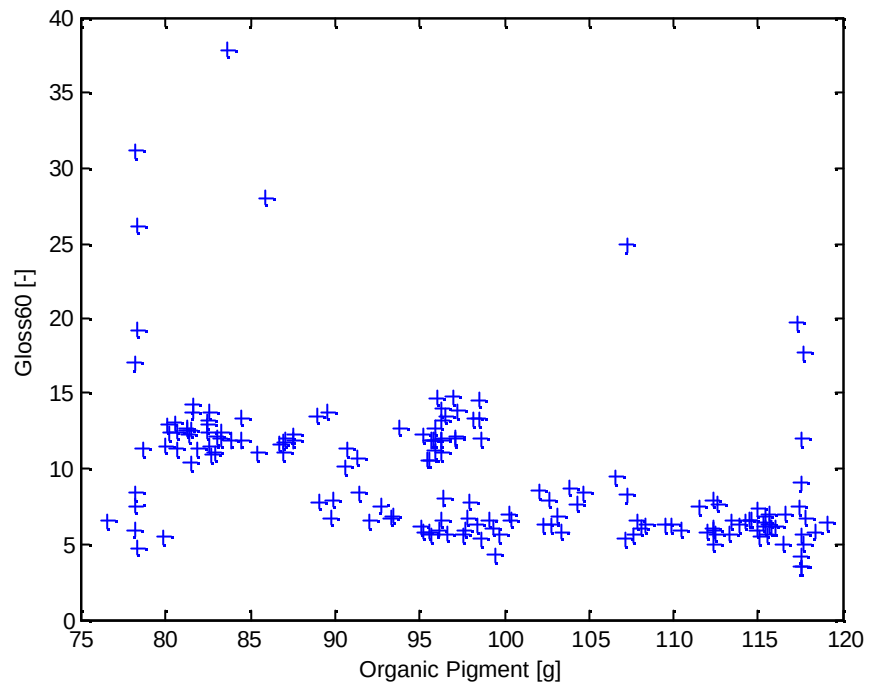
(c)



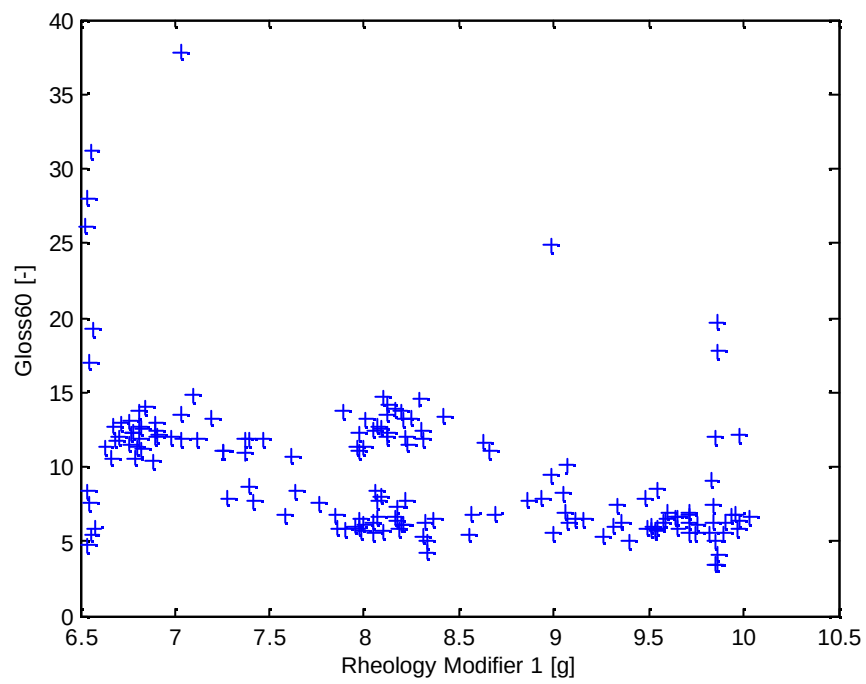
(d)



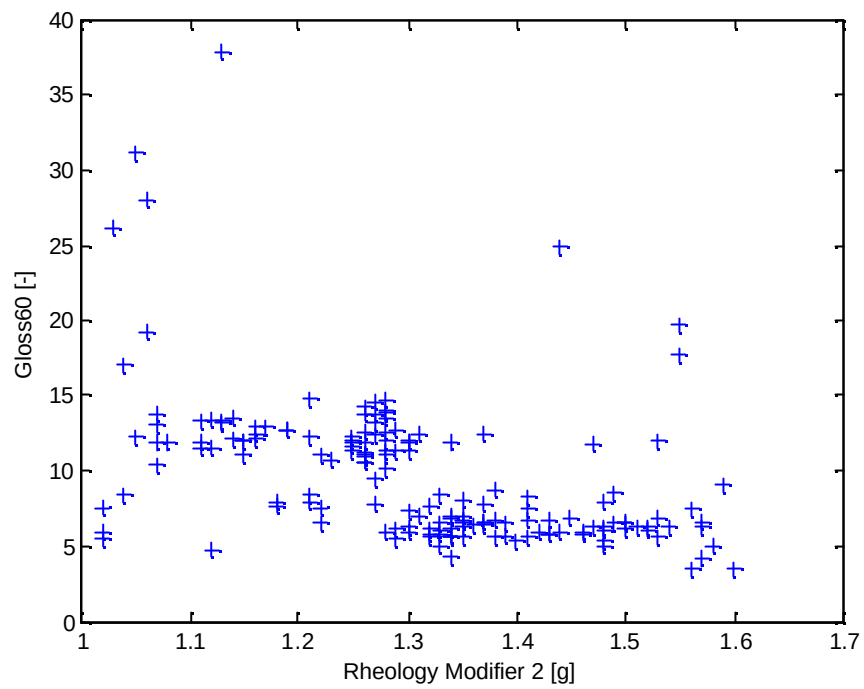
(e)



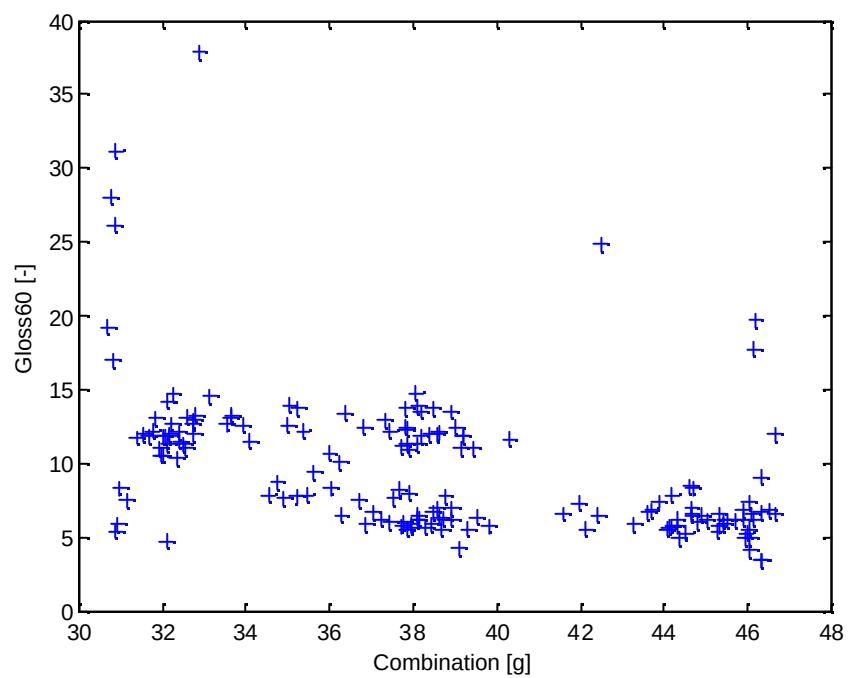
(f)



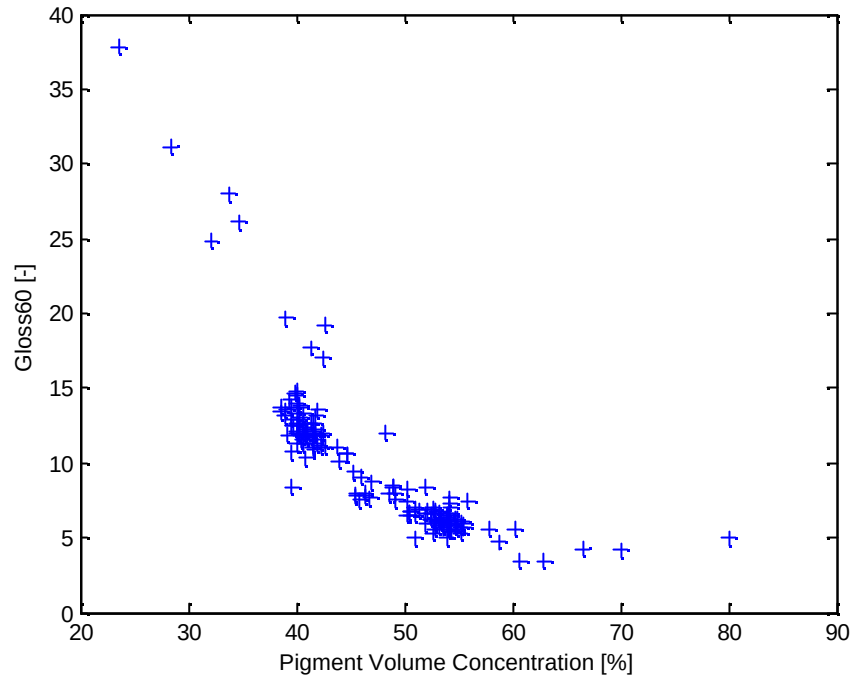
(g)



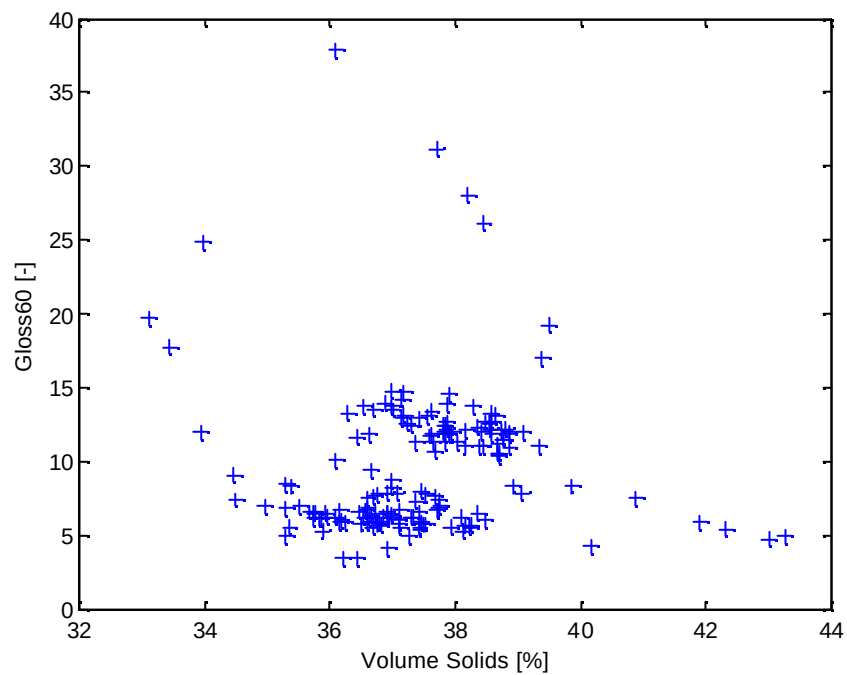
(h)



(i)



(j)



(k)

Figure 5.15 Scatterplot of Measured Gloss60 Values Against Predictor Variables

Linear relationships between *Gloss60* and the predictor variables are evident in most of the scatterplots of Figure 5.15 (excluding 'pigment' and 'volume solids'). The scatterplots of emulsion and pigment volume concentration (PVC) indicate some curvature, so do the other variables at

high and low concentrations, nonetheless, no severe curvature exists and linear assumptions seem to be a good approximation – at least for the largest part of the range (mostly the middle) of each predictor variable. The assumption of linearity is not grossly violated.

Another indication that at least one regressor variable contributes significantly in explaining the linear relationship between the response and the regressor variables, is by rejecting the null hypothesis, that not *one* of the regressor variables is significant in explaining the variance in the response. The p-values in Table 5.8 indicate with a high level of certainty that the variance in all the paint properties, except '*Open Time*' can be explained linearly by at least one of the predictor variables ($p < 0.05$). Therefore there exists at least one linear relationship between the paint property and the raw materials.

Table 5.8 Significance Test for Null Hypothesis

Paint Property	R^2	p-values
Dry Burnish (20)	0.3330	0.0001
Dry Burnish (60)	0.7021	0.0000
Dry Burnish (85)	0.5803	0.0000
Wet Burnish (20)	0.3189	0.0003
Wet Burnish (60)	0.4696	0.0000
Wet Burnish (85)	0.7022	0.0000
Dirt Pick-up (delta gloss 85)	0.5427	0.0000
Open Time - Glass	0.1595	0.1525
Dry Film Thickness	0.5506	0.0000
Gloss (20)	0.8482	0.0000
Gloss (60)	0.9266	0.0000
Gloss (85)	0.9526	0.0000
Hiding Power Roller - 1st coat	0.2662	0.0033
Krebs	0.8698	0.0000
Opacity	0.5753	0.0000
Anti-sag Index	0.1994	0.0440
Water Permeability	0.3096	0.0004

Variance (R^2) in some of the paint properties is explained well, e.g. $R^2_{Gloss85} = 0.95$, while variance in other properties is explained not that well, e.g. $R^2_{DryBurnish20} = 0.33$. High R^2 values do not necessarily mean good predictability and vice versa.

2. Mean Independence

The average value for the residuals of *Gloss60* is so small (5.33×10^{-6}) that the assumption of mean independence (that the independent variables are unrelated to the random error ε of

equation 5.6) can be considered as valid for the MLR model of *Gloss60*. Figure 5.16 also illustrates that the residuals of *Gloss60* are centred round zero. The assumption of mean independence is also true for all the other models (see Chapter 12 *Appendix E, Section 12.2 – Linear Modelling*).

3. Homoscedasticity

The scatterplot of residuals against case number shows no systematic pattern and therefore the errors in the model are independent of the run order.

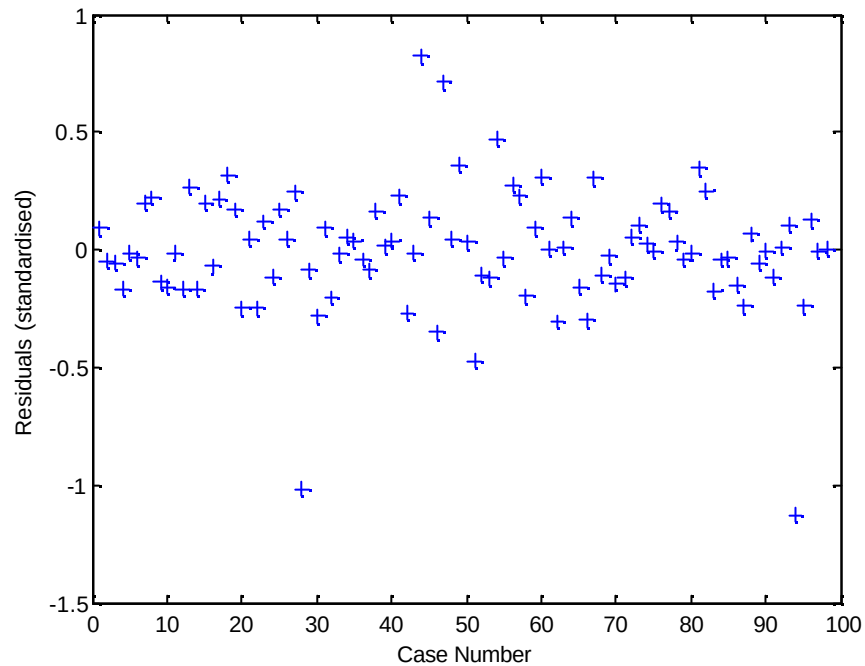


Figure 5.16 Scatterplot of Residuals (Gloss60) Against Case Number

4. Uncorrelated Errors

Once again, there is no easy way to determine whether contributions in the error term ε of equation 5.6 are correlated. However, it has been shown that the issue of correlated disturbances is strongly affected by sampling design [4] and that if samples are chosen at random order from an experimental space, that it is unlikely that correlated disturbances will be a problem. Furthermore, Design Expert[®] (the software package used for designing experiments) estimated the set of experiments so that multicollinearity between variables is kept at a minimum, reducing the possibility of correlation between error terms.

5. Normal Distribution

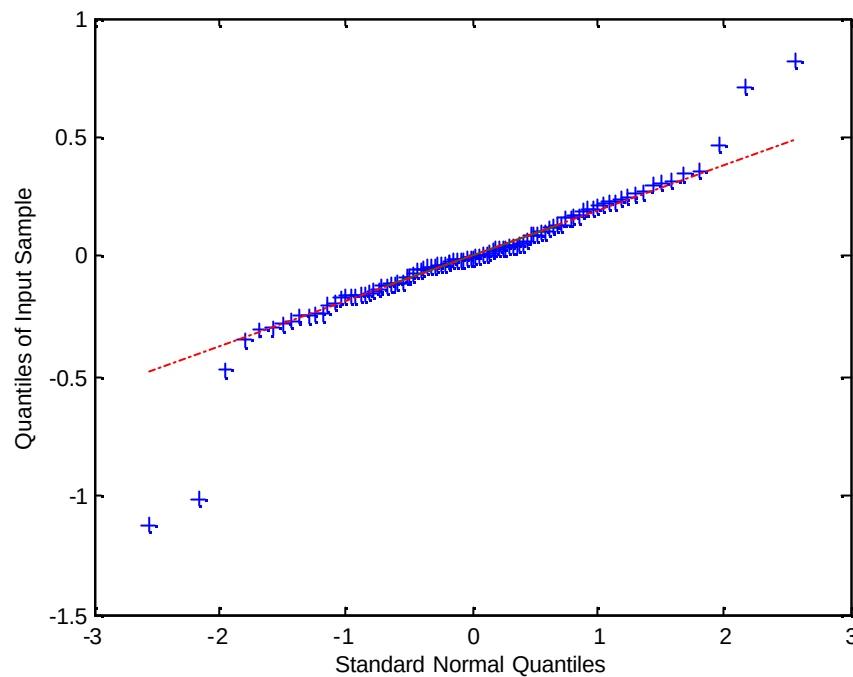


Figure 5.17 QQPLOT for Gloss60 Values

The quantile-quantile plot of residuals in Figure 5.17 follows a definite S pattern which is an indication of nonnormality. However, the assumption of normality is not grossly violated as most of the points follow a normal distribution. Multiple linear regression will be able to fit the normal distributed points accurately but might not be that accurate in modelling the data points that deviate from normality. It was observed in the previous section that paint samples with extreme PVC/VS are classified as influential data points; therefore it is not entirely surprising that the QQPLOT of Figure 5.17 also deviate from normality which is caused by non-linearities due to extreme PVC and VS.

6. Influential Data Points

Influential data points affect the results that are included in the analysis. The residual case order errorbar plot (RCOPlot) of confidence intervals on residuals from the MLR, indicates the residuals that are considered abnormally large according to the 95% confidence interval. These data points are not fitted accurately with the MLR model (indicated with  in Figure 5.18). These are either outliers or influential data points.

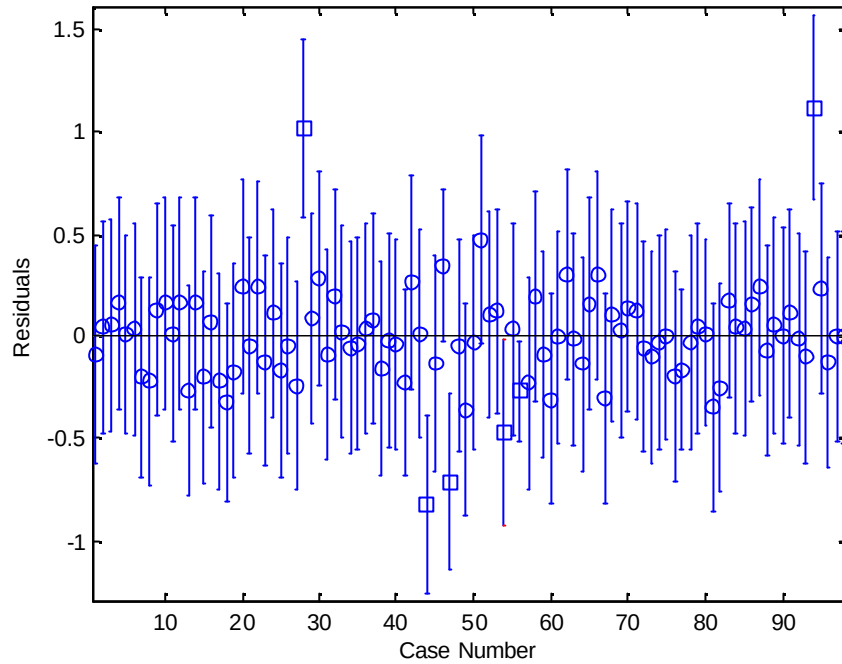


Figure 5.18 RCO PLOT of Gloss60

Closer inspection of Figure 5.18 reveals that 5 of the 6 'outliers' (83%) are paint samples with extreme PVC/VS. These data points can be considered as influential data points that are not fitted well by the MLR model. This is also an indication that the linear model is unlikely to predict future observations with high PVC/VS accurately and that other modelling techniques should also be investigated for prediction capabilities of these samples.

5.3.4 PAINT PROPERTY PREDICTIONS

Table 5.9 shows the regression coefficients used for paint property predictions. These regression coefficients are used in equation 5.6.

Table 5.9 Regression Coefficients for Model Parameters (MODEL 2)

Paint Property	β_{constant}	β_{pigment}	β_{extender}	β_{water}	β_{solvent}	β_{emulsion}	$\beta_{\text{org. pigment}}$	β_{RM1}	β_{RM2}	$\beta_{\text{combinatio}}$	β_{PVC}	β_{Vs}
Dry Burnish20	-0.1886	1.3641	1.5381	-2.8761	-0.9224	1.3205	0.5258	0.1385	0.1408	-0.9641	1.1511	-4.6903
Dry Burnish60	-0.1021	-0.6790	-0.1986	-2.3404	-0.5528	-0.3847	-0.3481	0.0552	0.0404	-0.6865	2.9014	-2.7524
Dry Burnish85	-0.0510	1.5127	2.1479	-2.3169	-0.5369	-2.1023	0.4650	0.0668	0.0294	-0.5995	-2.2855	-3.1479
Wet Burnish20	-0.1509	-3.1632	-2.7861	0.04147	0.0618	1.2819	-0.8515	-0.0361	0.0885	0.0914	4.8181	1.5900
Wet Burnish60	-0.0566	0.5149	0.6307	-1.0404	-0.3412	1.1707	0.2504	0.0267	0.0628	-0.5516	1.8681	-1.7323
Wet Burnish85	-0.0127	1.3170	1.7709	-1.1269	-0.2299	-0.3804	0.3626	0.0835	0.0016	-0.3874	-0.8532	-1.8126
Water Permeability	0.0280	-4.1105	-4.4080	2.8334	0.8135	3.3233	-1.6200	-0.0810	0.0367	0.8422	7.5934	4.1799
Krebs	-0.0023	0.1020	-0.3053	1.6895	0.4052	0.7302	-0.0363	0.2036	0.0963	0.6534	-0.1203	2.778
Sag	-0.1502	1.8322	1.6614	-1.4499	-0.4220	-0.1677	0.4555	0.1382	0.1470	-0.3395	-1.5640	-2.0098
Open Time	0.0899	-5.1919	-4.9833	4.9831	1.5487	-0.9808	-1.3039	-0.2783	-0.0005	1.7192	2.2522	8.1913
Gloss20	0.0022	-7.3902	-7.3519	3.7341	1.1221	5.0901	-2.7485	-0.0156	-0.0384	1.1738	12.4546	6.6933
Gloss60	0.0200	-6.4427	-6.6085	3.9873	-2.1874	3.3968	-2.1874	-0.0232	0.0103	1.2594	8.8941	-6.9358
Gloss85	-0.0147	-6.6797	-6.9707	3.1959	1.0031	2.2970	-2.1667	-0.0360	0.0268	0.9844	8.9128	6.2097
Dirt Pick-Up	-0.0371	0.6210	0.2786	-0.1389	-0.0114	-2.0610	0.1700	-0.2267	0.1503	-0.0277	-3.2483	-0.1293
Opacity	0.1341	5.9451	4.3084	-0.0047	-0.2401	1.7758	2.0301	0.1093	0.1268	-0.0297	-5.5124	1.8123
Hiding Power	-0.0204	-0.9265	-0.7332	-0.9704	-0.2646	-4.6708	-0.2787	-0.1008	-0.3318	-0.2679	-2.0018	-0.3163
Dry Film Thickness	0.2228	-15.7983	-15.2227	12.8085	3.8658	-6.2145	-3.9004	0.1074	0.1908	4.4121	4.8521	23.4258

5.3.4.1 MODEL PREDICTABILITY

The regression models with their regression coefficients and corresponding predictor variables, $X_{emulsion}$, $X_{pigment}$, $X_{extender}$, ..., X_{VS} are now used to predict future paint property data. Figure 5.19 gives the 95% prediction interval which indicates with 95% certainty where future observations for the *Gloss60* values are expected to be.

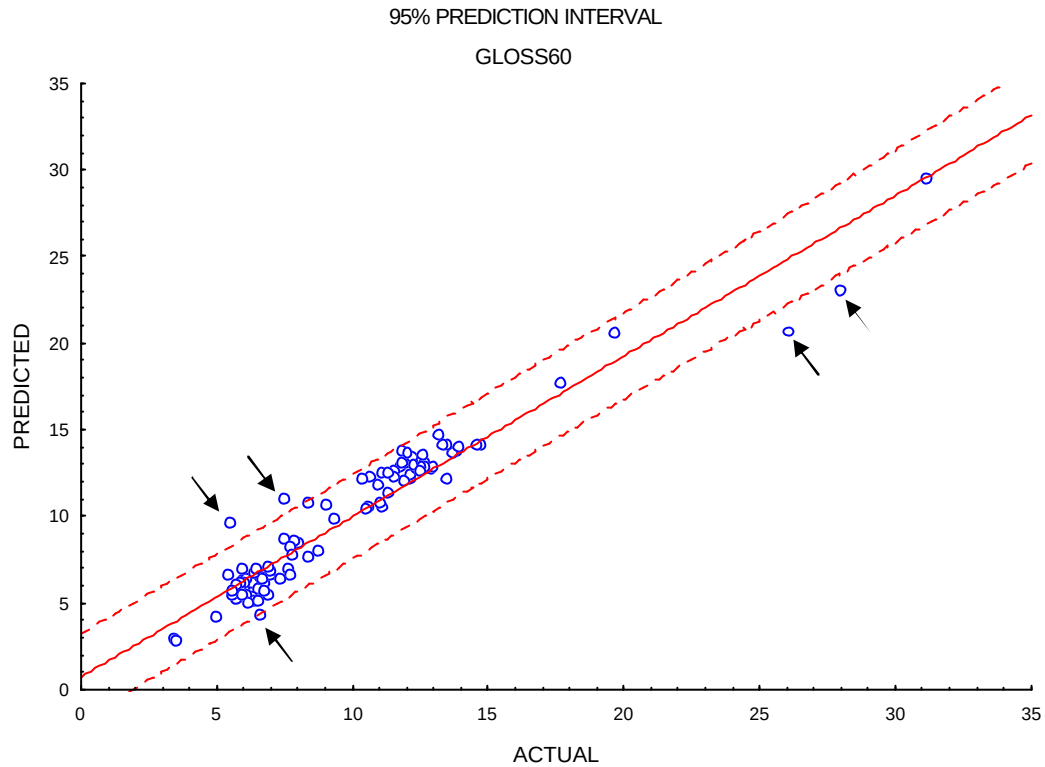


Figure 5.19 Actual and Model Gloss60 Values and 95% Prediction Interval

Closer inspection of the data points reveals that four of the five (80%) outliers are samples with extreme PVC/VS. The RCOPLLOT of Figure 5.18 has indicated that the MLR model does not fit the data points with extreme PVC/VS accurately and therefore it can also be expected that future predictions that fall outside the 95% prediction interval will mostly result in large residuals.

Once again, the model is tested with 'unseen' data (data that were not used to develop the model). Figure 5.20 indicates that three predictions of unseen data are predicted outside the 95% prediction interval. Closer inspection also reveals that all three points are samples with extreme PVC/VS.

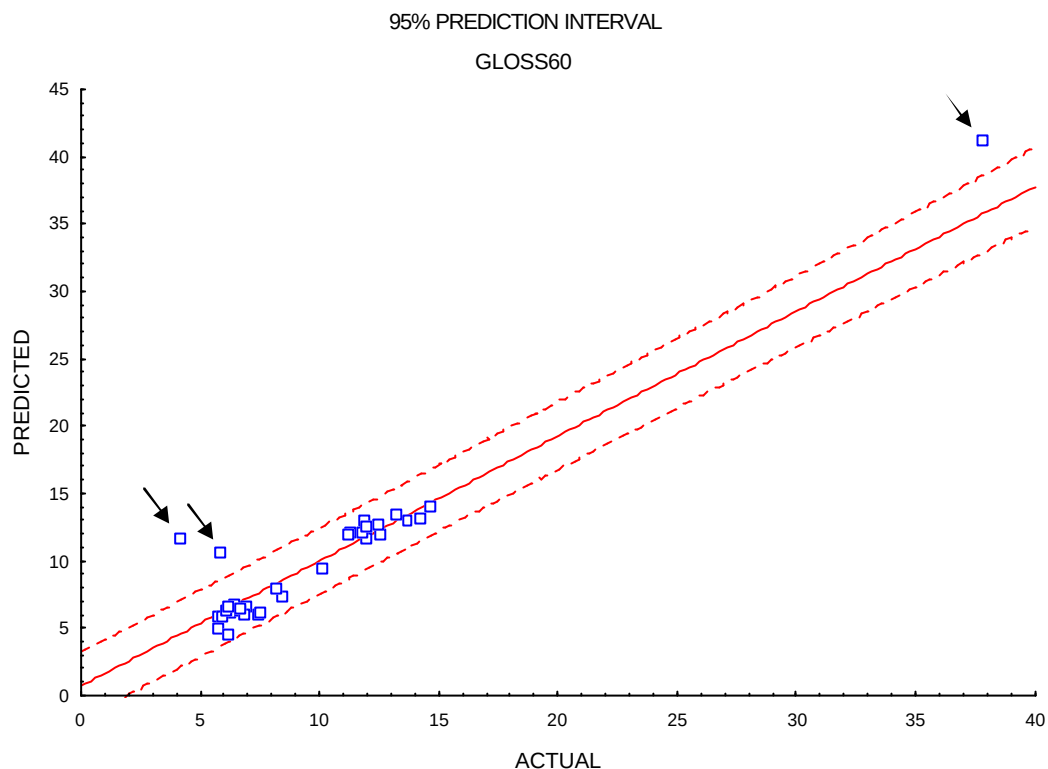


Figure 5.20 Predicted Gloss60 Values

It is evident that the MLR model is capable of predicting the gloss property accurately as far as 'normal' formulation levels of the paint is concerned. However samples with larger variations in the raw materials (larger than 20% on the standard formulation) are not predicted accurately with the linear models. Other models might perform better in predicting the gloss behaviour of paint samples with extreme variation in raw materials and it is worth some investigation.

5.3.4.2 MODEL ERRORS

The paint formulator would like to know what error margins could be expected in predicting paint properties with the models. Table 5.10 shows the relative error made by the models in predicting new paint properties.

Table 5.10 Relative Errors in Predicting Unseen Data

Paint Property	Relative Errors [%]	
	Including Extreme Data Points	Excluding Extreme Data Points
Dry Burnish (20)	20.99	14.82
Dry Burnish (60)	14.88	10.73
Dry Burnish (85)	34.45	22.71
Wet Burnish (20)	30.06	29.65
Wet Burnish (60)	15.47	14.78
Wet Burnish (85)	20.65	11.76
Dirt Pick-up (delta gloss 85)	46.02	43.57
Open Time - Glass	22.59	17.29
Dry Film Thickness	9.65	6.38
Gloss (20)	11.93	6.69
Gloss (60)	14.50	7.23
Gloss (85)	6.59	4.23
Hiding Power Roller - 1st coat	1.74	0.81
Krebs	1.67	1.57
Opacity	1.25	0.81
Anti-sag Index	42.16	43.15
Water Permeability	15.80	14.24

Note the difference between relative errors of paint properties with and without the influential data points. It is considerable less for predictions without extreme data points if one takes into account that influential data points from paint samples with extreme PVS/VS only make up 5% of the 'unseen' data and therefore it actually increases the error significantly.

5.3.4.3 PRACTICAL APPLICABILITY OF MODEL

Some paint properties are critical in determining the quality of the paint. These paint properties have definite specifications of the variation (error) that is allowed and are given in the technical data sheets. For the specific paint under investigation, the critical paint properties are:

i. Krebs viscosity at 23°

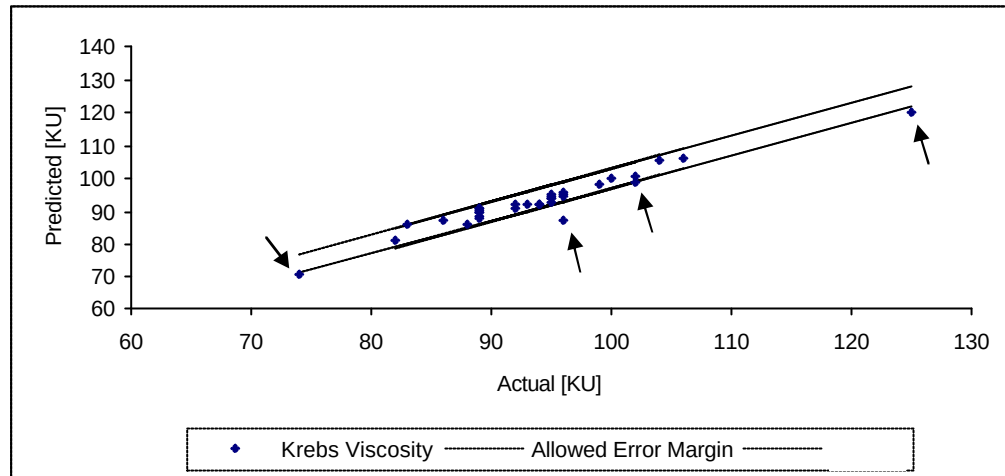


Figure 5.21 Predicted Krebs Viscosity with Allowed Error Margin

- $\pm 12\%$ of predicted *Krebs Viscosity* values lie outside the allowed error margin.
- 50% of the predicted *Krebs Viscosity* values outside of the allowed error margin have extreme PVC/VS.

ii. Gloss

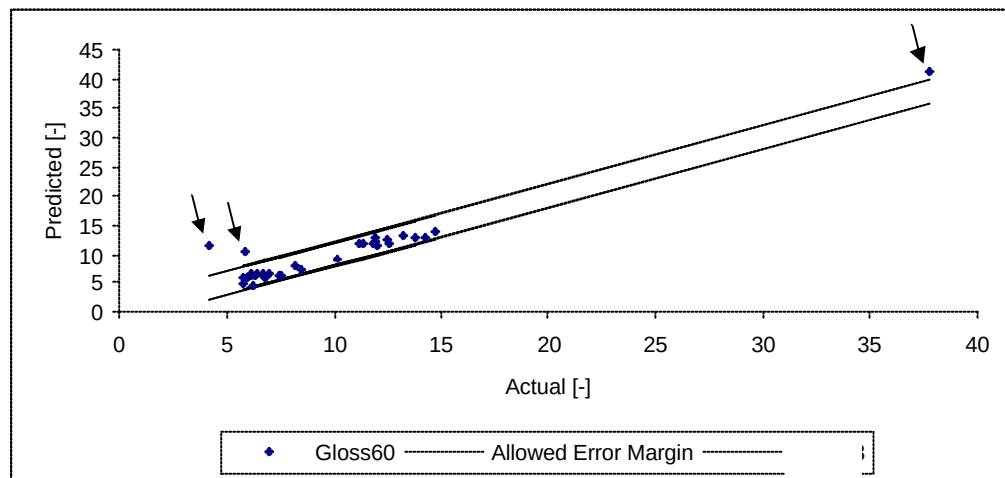


Figure 5.22 Predicted Gloss60 with Allowed Error Margin

- 10% of predicted *Gloss60* values lie outside the allowed error margin.
- All three (100%) predictions for *Gloss60* values outside the allowed error margin are from samples with extreme PVC/VS.

iii. Opacity

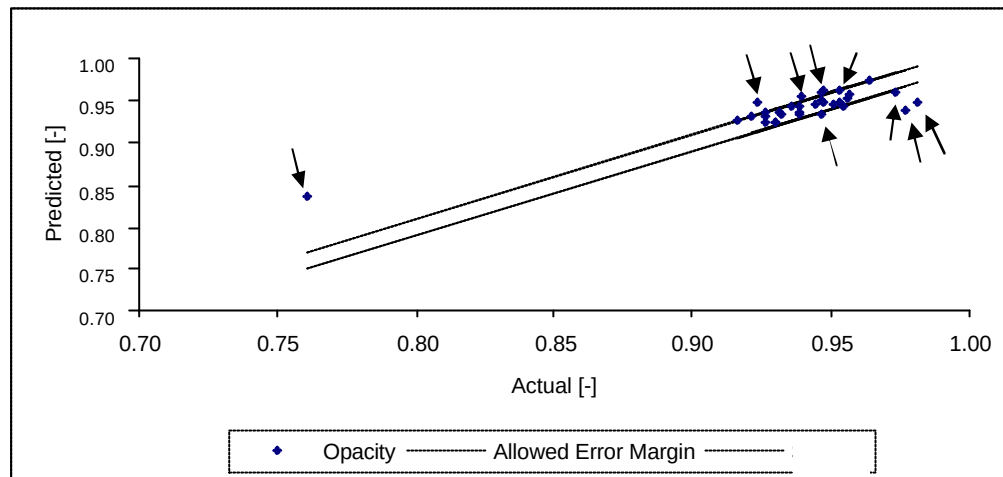


Figure 5.23 Predicted Opacity with Allowed Error Margin

- 27% of predictions for *Opacity* values lie outside the allowed error margin, although only 11% are clearly erroneous. *Opacity* is probably the most crucial paint property and has very tight error margins.
- 33% of the data points that lie outside the error margin have extreme PVC/VS and are also those points with the largest errors.

iv. Film thickness

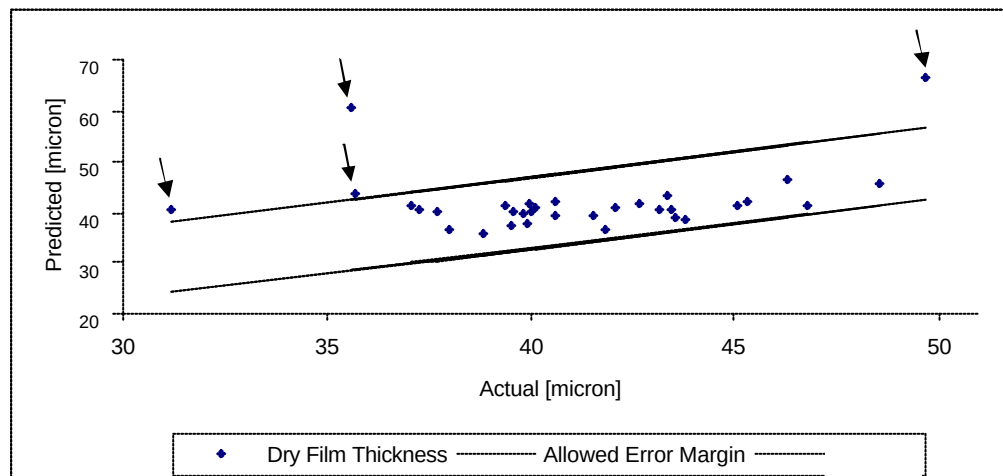


Figure 5.24 Predicted Dry Film Thickness with Allowed Error Margin

- ±10% of predicted *Dry Film Thicknesses* lie outside the allowed error margin.
- 75% of the data points outside the allowed error margin have extreme PVC/VS.

It is clear that MLR accurately predicts these four paint properties with only ±10% of the predictions for *Krebs Viscosity*, *Gloss60* and *Dry Film Thickness* that lie outside the allowed error

margin. Linear modelling is actually very accurate, taking into account that the incorrect predictions are mainly from samples with extreme PVC/VS.

5.3.5 MODEL 2 CONCLUSIONS

- Table 5.9 gives the matrix of regression coefficients that can be used in equation 5.6 to predict the paint properties of this specific paint.
- The linear models are accurate in predicting the paint properties as long as raw material levels fall within the normal 'plus or minus 20%' range of the standard formulation. Another model is recommended for paint samples whose variation in raw materials is larger than the 'plus or minus 20%' range (samples with extreme pigment volume concentrations (PVC) and volume solids (VS)).
- The paint property that is predicted most accurately by MLR is opacity. On average, an error of 0.81% can be expected when predicting the opacity value by using MLR when the raw materials vary within a range of 20% from the standard formulation. However, opacity is also predicted very accurately when the raw materials vary more than 20% from the standard formulation. On average, errors of 1.25% can then be expected.
- Another paint property that is also predicted accurately is the Krebs viscosity. On average, an error of 1.57% can be expected when predicting the Krebs viscosity by using MLR when the raw materials vary within a range of 20% from the standard formulation. Only slightly less accuracy is obtained when the predicted Krebs viscosity vary with more than 20% from the standard formulation, with expected errors of 1.67%.
- Other paint properties that are predicted with less than 10% error are gloss 60 (error = 7.23%) and dry film thickness (error = 6.38%)
- Paint properties that can be predicted with an expected error of between 10% and 20% are dry and wet burnish, open time and water permeability.
- Paint properties that are predicted the most inaccurate by MLR, that results in errors larger than 20% are dirt pick-up and sagging.
- Opacity, Krebs viscosity, gloss and dry film thickness are critical paint properties that specify the quality of this paint. All these paint properties are predicted with an error less than 10% by MLR. Furthermore, an average of 90% of predictions for these paint properties, fall within the allowed specifications range.

5.4 MODEL 3: RELATIONSHIP BETWEEN PAINT PROPERTIES AND PAINT RHEOLOGY

5.4.1 BACKGROUND

The understanding of coating performance in terms of rheology is far from complete. The reasons are (a) the complexity of coatings processes, which complicates the understanding of the role of rheology in process outcomes, and (b) the difficulty in linking measured fundamental rheological properties with real-world performance. This study attempts to succeed in the second point and

then in future work, these links can be used for a fundamental investigation of the effect of rheology on the paint properties. Certain paint properties, e.g. levelling, sag, spatter, rely heavily on the flow and deformation behaviour (rheology) of the paint and therefore the importance of understanding coating performance in terms of rheology is understood. Up to now, rheology has only been used qualitatively in explaining the coating performance, for example, *more* pseudoplasticity usually leads to *easier* application, but these type of correlations have never been explained quantitatively. One can also understand this, because qualitative behaviour is relative and can be used for comparing any two paints, while quantitative behaviour is case-specific. However, this does not take away the relevance of a quantitative model for a specific paint.

One solution to the lack of a quantitative model is computer simulation of coating processes, which utilises the fundamental rheological data as the required input. However, although computer modelling is perhaps the ideal approach, it has not always been readily accessible and many coating problems were attempted to be solved from shear viscosity and viscoelastic data from rheological measurements alone [19].

A major advantage of having a model that predicts the paint properties in terms of rheology, is that no explicit knowledge of the constitutive equations describing the chemical/physical interaction between the raw materials are required.

More specific, Leskovsek et al [11], [20] and Tusar et al [21], [22] have modelled some paint properties in terms of the rheological behaviour, but their work have the following weaknesses:

1. Single rheological measurements were used. This is an important weakness because a number of rheological measurements are available today to explain the coating performance.
2. Single points of measurements were used to model specific paint properties. This is also an important weakness, because the different parts of the rheology curves assist in explaining different coating behaviour (e.g. low frequency range of the frequency sweep can explain the sedimentation behaviour of pigments while spatter behaviour is explained at the high frequency range of the frequency sweep).

These existing weaknesses are overcome in this study by:

1. Using a set of rheological measurements and extracting the most important data from it with the use of principal component analysis.
2. Making use of preliminary modelling of the rheology curves (a method that has been used in the pharmaceutical industry by Gasperlin et al) to represent the whole rheology curve.

The theory behind both these improvements is discussed in *Chapter 4 Data Preparation*.

5.4.2 MULTIPLE LINEAR REGRESSION (MLR)

Linear modelling techniques are used as first approximations for models, although two other research groups (Leskovsek et al and Tusar et al) found that non-linear modelling techniques are

superior. The reason for first attempting linear modelling is that (a) each paint system is unique and therefore behaves differently and (b) linear models are good first approximations and give a lot of information about the level of nonlinearity.

5.4.2.1 MLR – CAUSAL MODEL

The rheological data is highly collinear (the condition number $C \gg 15$). Principal component analysis is used to form new linear independent variables (principal components) from the collinear rheological data. The drawback is that principal components lack simple interpretation since each is, in a sense, a mixture of the original variables. The use of MLR as a method to construct a causal model is therefore inappropriate, because principal components are used as regressor variables (there are complex statistical ways to reconstruct the original data from principal components but it is beyond the scope of this study).

5.4.2.2 MLR – PREDICTIVE MODEL

Multiple linear regression is used to model the relationship between the paint properties and the rheological behaviour. Three principal components explaining 95% of the variance in the rheological data (98 random samples) are used as predictor variables so that the models for the paint properties are expressed in the following way:

$$a_{\text{PaintProperty}} = b_0 + b_{PC1}x_{PC1} + b_{PC2}x_{PC2} + b_{PC3}x_{PC3} + e \quad (5.7)$$

Unlike the previous models, the independent variables that are used to predict the paint properties are stochastic and therefore the variance of the rheological variables should be substantially less than the variance of the measured paint properties for the predictions to be accurate. This is verified by repeating a single experiment 5 times to determine the variability of the paint properties and rheological variables. Table 5.11 gives the paint property and rheological variables with the largest variance.

Table 5.11: Largest Variance of Dependent/Independent Variables

	Variance
Open Time	28.8
Offset of Time Sweep (β_{TS})	7.8

The accuracy of MLR model predictions increase when the variance of the independent variables are small compared to the variance of the dependent variables. There exists no quantitative criterion for what the ratio of variances between the independent variable (β_{TS}) and dependent variable (open time) should be, but according to experts in the field of statistical modelling^{*****}, the values in Table 5.11 are acceptable for accurate predictions. It has also been verified that the

***** Personal correspondence with Prof C Aldrich (University of Stellenbosch)

variance of the remaining rheological variables are small compared to the variance of the paint properties.

The model fit for *Gloss60* for 98 samples can be seen in Figure 5.25

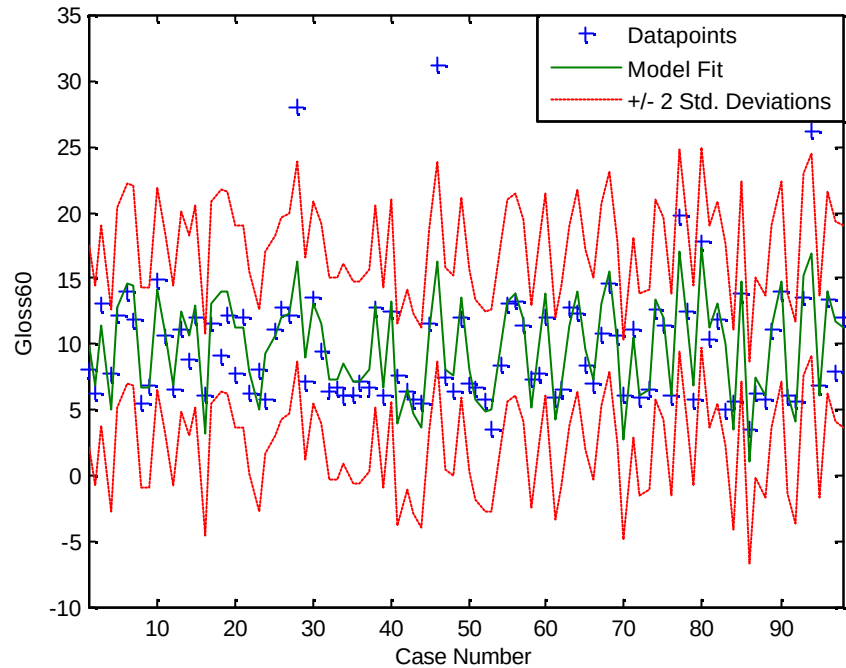


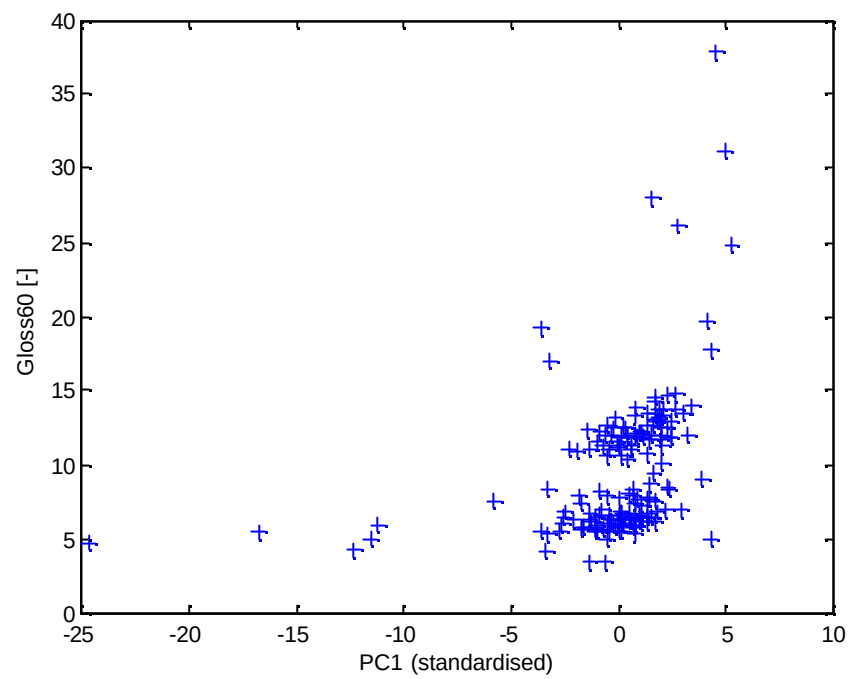
Figure 5.25 Model Fit (MLR) for Gloss 60 as Predicted by Principal Components of Rheology Data

Only three data points fall outside two standard deviations of the model and therefore, as a first approximation, it can be assumed that the model fits the data relatively well. Further analysis will indicate how accurate the model predicts Gloss60 of new observations.

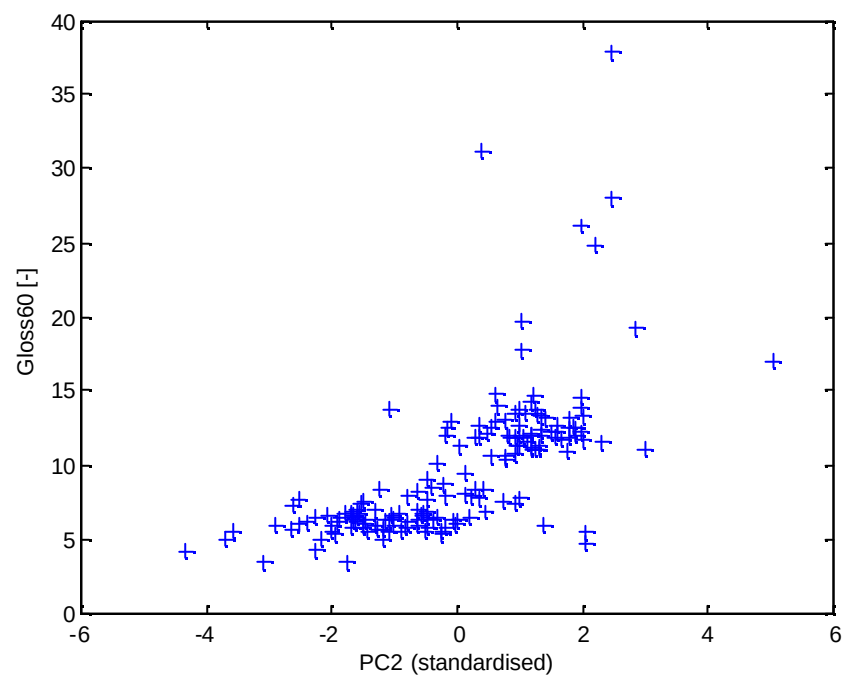
5.4.3 MODEL VALIDITY

The validity of the assumptions that are made during MLR is tested below.

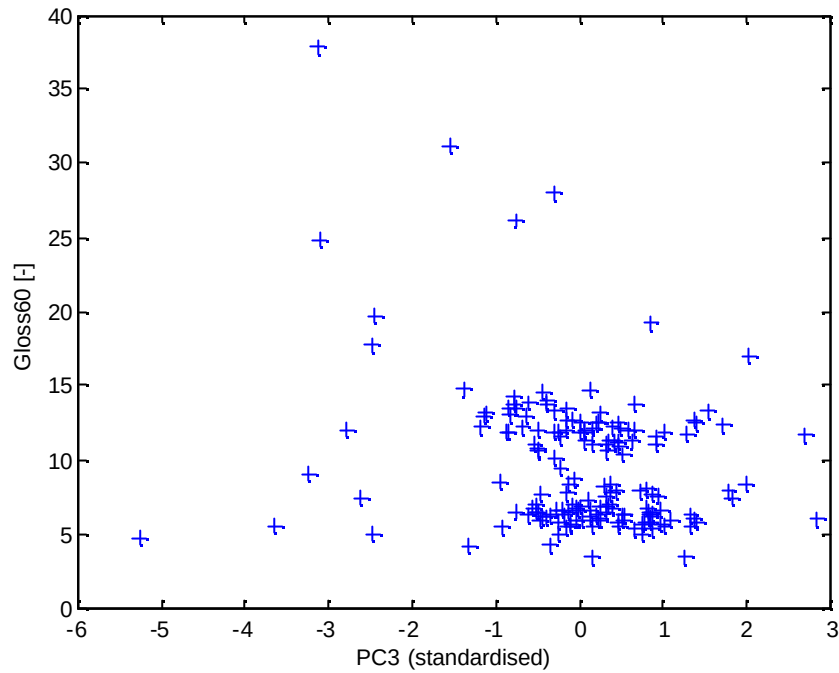
1. Linearity



(a)



(b)



(c)

Figure 5.26 Scatterplot of Gloss60 Against (a) PC1, (b) PC2 and (c) PC3

Some linearity seems to exist in PC1 and PC2, although curvature is also evident at the high and low ends of the curves. Still, the assumption of linearity is not grossly untrue or inappropriate keeping in mind that the linear model will only be an approximation.

2. Mean Independence

The average value for the residuals of *Gloss60* is so small (-5.66×10^{-18}) that the assumption of mean independence (that the independent variables are unrelated to the random error ε of equation 5.7) can be considered as valid for the MLR model of *Gloss60*. Figure 5.27 also illustrates that the residuals are randomly centered round zero.

3. Homoscedasticity

The scatterplot of residuals against case number (Figure 5.27) shows no systematic pattern and therefore the errors in the model are independent of the run order.

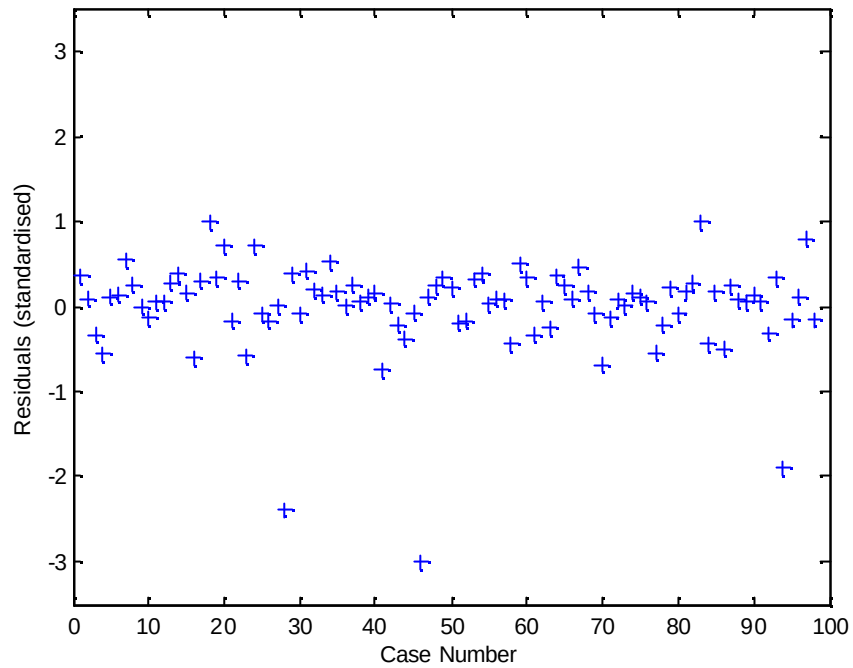


Figure 5.27 Scatterplot of Residuals for Gloss60 Model Fit

4. Uncorrelated Errors

It is highly unlikely that error contributions of the principal components are correlated. The principal components are a linear uncorrelated 'mixtures' of the original variables. Therefore one can also assume that the errors will be uncorrelated.

5. Normal Distribution

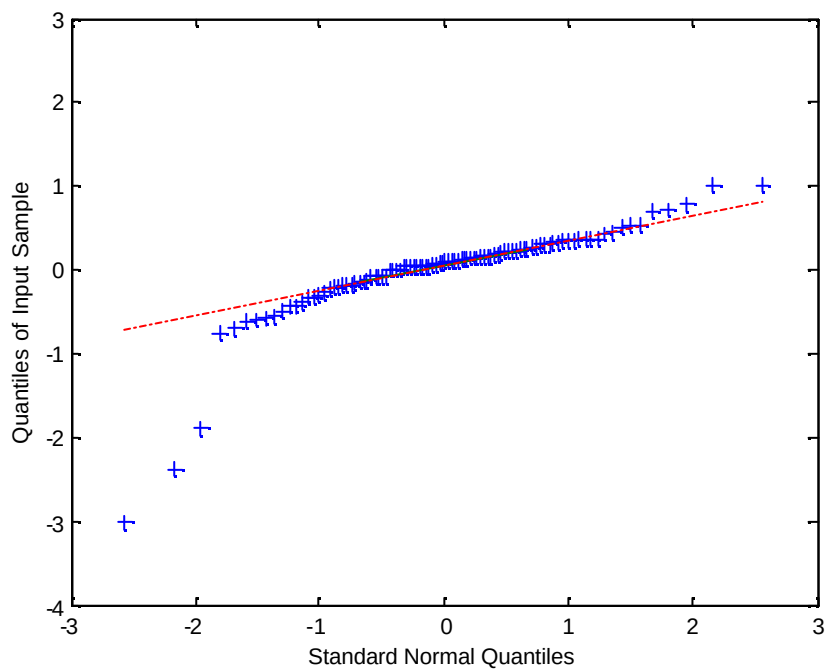


Figure 5.28 QQPLOT for Gloss60

The definite S pattern at the lower quantiles is an indication of nonnormality. This gives an indication that there are points which do not follow a normal distribution and that a linear model might not give the most accurate predictions for these points. Otherwise, the largest part of the curve follows a normal distribution and therefore the assumption of a normal distribution is not grossly untrue.

6. Influential Data points

Possible influential data points are already identified in the QQPLOT of Figure 5.28. One can already expect that most of these points that deviate from normality are samples with extreme PVC/VS. Figure 5.29 gives a clear indication of these possible 'outliers' or influential data points.

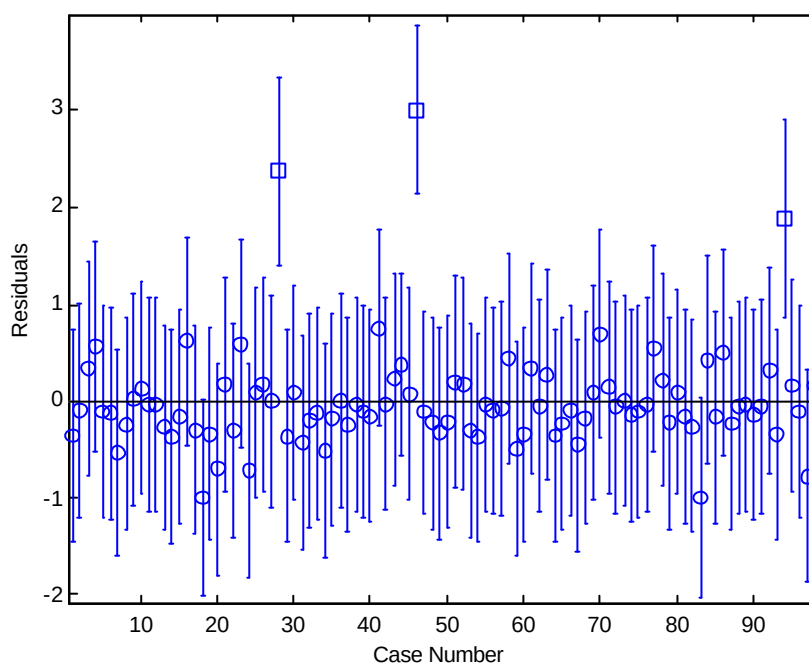


Figure 5.29 RCOPLLOT for Gloss60 Model Fit

All three 'outliers' (indicated with a $\square$) are samples with extreme PVC/VS and therefore they can be considered as influential data points in affecting *Gloss60*. The model does not fit these *Gloss60* values accurately and it can be expected that 'unseen' paint samples with extreme PVC/VS will not be accurately predicted either.

5.4.4 PAINT PROPERTY PREDICTIONS

Table 5.12 shows the regression coefficients used for paint property predictions. These regression coefficients are used in equation 5.7.

Table 5.12 Regression Coefficients for Model Parameters (MODEL 3)

Paint Property	b_{constant}	b_{PC1}	b_{PC2}	b_{PC3}
Dry Burnish20	-0.1198	0.1168	-0.0165	-0.3294
Dry Burnish60	0.0108	-0.1743	-0.4186	-0.3170
Dry Burnish85	0.0256	-0.1318	-0.3644	0.0759
Wet Burnish20	-0.1683	0.1738	0.2129	0.00873
Wet Burnish60	-0.0173	-0.1148	-0.3470	-0.0400
Wet Burnish85	0.0151	-0.1400	-0.4226	0.2465
Water Permeability	-0.0438	0.0077	-0.1588	-0.3423
Krebs	0.0276	-0.3314	-0.0251	0.0675
Sag	-0.0868	-0.1379	0.0526	0.0260
Open Time	-0.0135	0.0414	-0.0865	-0.0780
Gloss20	-0.0943	0.1903	0.3006	-0.0329
Gloss60	-0.0923	0.2305	0.4102	-0.0755
Gloss85	-0.0903	0.1807	0.3309	-0.3186
Dirt Pick-Up	-0.0642	0.1202	0.3327	-0.0571
Opacity	0.1548	-0.3199	-0.0602	-0.1878
Hiding Power	0.0155	-0.1034	-0.1790	-0.0219
Dry Film Thickness	-0.0093	0.0140	0.1281	-0.1575

5.4.4.1 MODEL PREDICTABILITY

The 95% prediction interval of Figure 5.30 indicates where future predictions are expected to be. Five of the 98 fitted data points lie outside the 95% prediction interval. Therefore it can be expected that $\pm 5\%$ of future predictions will lie outside the 95% prediction interval.

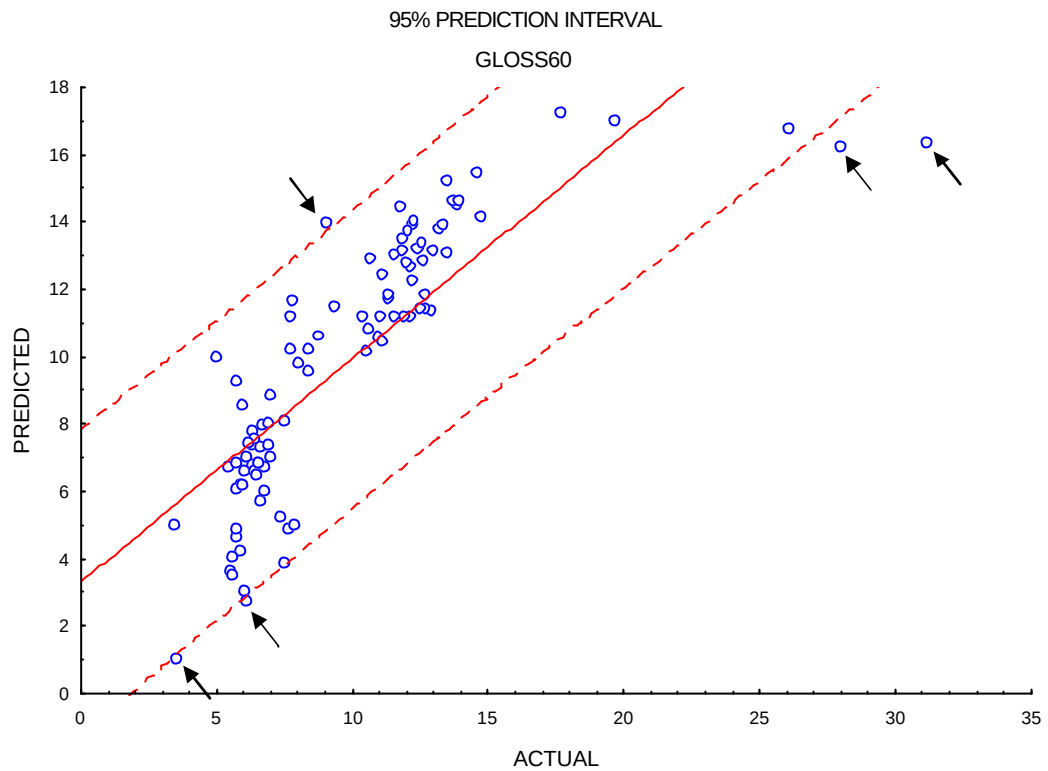


Figure 5.30 Fitted Gloss60 Values and 95% Prediction Interval

Four of the five data points (80%) that lie outside the allowed error margin have extreme PVC/VS and therefore it can be expected that $\pm 80\%$ of future 'unseen' data points that lie outside the 95% prediction interval are also samples with extreme PVC/VS. The RCOPLLOT of Figure 5.29 has indicated that these data points result in large residuals that affect the model fit and therefore the model predictions.

The model predictions for 33 'unseen' samples can be seen in Figure 5.31.

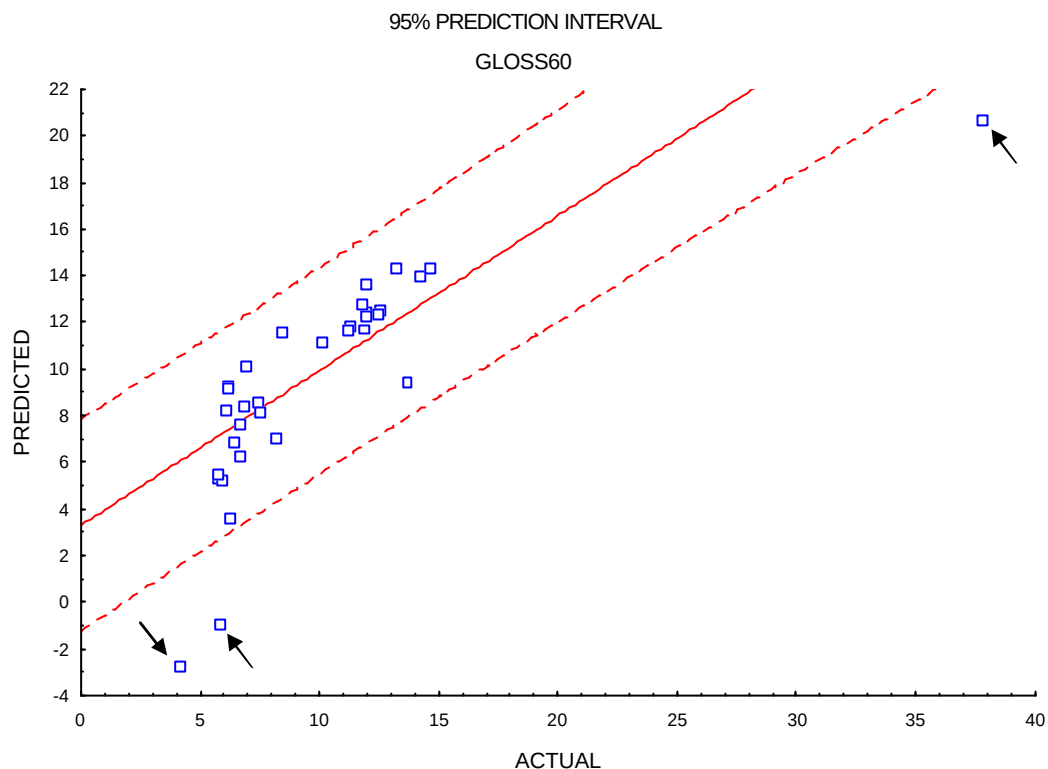


Figure 5.31 Predicted Gloss60 Values and 95% Prediction Interval

All three of the predictions outside the 95% prediction interval are samples with extreme PVC/VS. The RCOPLLOT of Figure 5.29 has indicated that large residuals (errors) can usually be expected with these predictions outside the 95% prediction interval.

5.4.4.2 MODEL ERRORS

The previous sections indicated that large residuals can usually be expected for paint samples with extreme PVC/VS. Table 5.13 gives the relative error⁺⁺⁺⁺⁺ in the predicted paint properties. The actual and predicted data is given in Chapter 12 – Appendix E.

$$+++++ \text{ error} = \frac{|actual - predicted|}{actual}$$

Table 5.13 Relative Errors (Residuals) for Paint Properties

Paint Property	Relative Errors [%]	
	Including Extreme Data Points	Excluding Extreme Data Points
Dry Burnish (20)	20.02	17.31
Dry Burnish (60)	28.19	19.98
Dry Burnish (85)	30.60	26.14
Wet Burnish (20)	31.73	30.11
Wet Burnish (60)	15.62	15.30
Wet Burnish (85)	18.43	16.33
Dirt Pick-up (delta gloss 85)	49.29	48.02
Open Time - Glass	19.61	17.36
Dry Film Thickness	8.87	7.11
Gloss (20)	14.14	8.70
Gloss (60)	23.89	15.33
Gloss (85)	11.54	9.94
Hiding Power Roller - 1st coat	1.81	1.53
Krebs	1.74	1.61
Opacity	1.82	1.05
Anti-sag Index	38.85	38.88
Water Permeability	13.89	12.51

Some of the errors in Table 5.13 might seem large. But it must be remembered, that except for the specifications of certain paint properties which are used for quality control, that error margins for many of these paint properties are very vague. An error of 20% in the *Burnish* value is already an 'accurate' prediction if there are no existing error margins. These error margins do not exist because these paint properties are not critical in quality control. Table 5.13 also indicates that the relative errors are smaller when extreme data points are not included in the prediction. In fact, they are considerably smaller if one considers the fact that samples with extreme PVC/VS only make up $\pm 10\%$ of the 'unseen' sample size and that errors are calculated on an average basis. Therefore, as with the other models, MLR accurately predicts the responses of the paint samples within the normal range of raw materials (plus/minus 20% of standard formulation), but fail to accurately predict the response for paint samples with raw materials that vary more than 20% from the standard formulation (paint samples with extreme PVC/VS). A different model is required for paint samples whose raw materials vary more than 20%.

5.4.4.3 PRACTICAL APPLICABILITY OF THE MODEL

The paint scientist will be very satisfied if the paint properties can be predicted accurately for all the different variations of raw materials. Certain error margins are allowed for the most critical paint properties that are used in quality control:

i. Krebs Viscosity

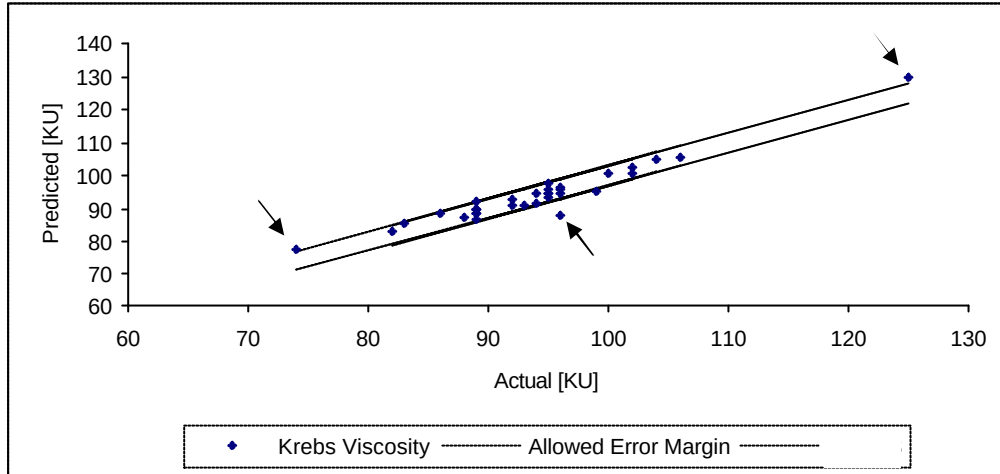


Figure 5.32 Error Margin for Krebs Viscosity

- 10% of the predicted Krebs Viscosity values lie outside the allowed error margin and can be seen as inaccurate predictions.
- 50% of the inaccurate predictions are paint samples with extreme PVC/VS.
- Four samples with extreme PVC/VS are in the 'unseen' sample size. The *Krebs Viscosity* for two of them (50%) is predicted inaccurately by the MLR model.

ii. Gloss

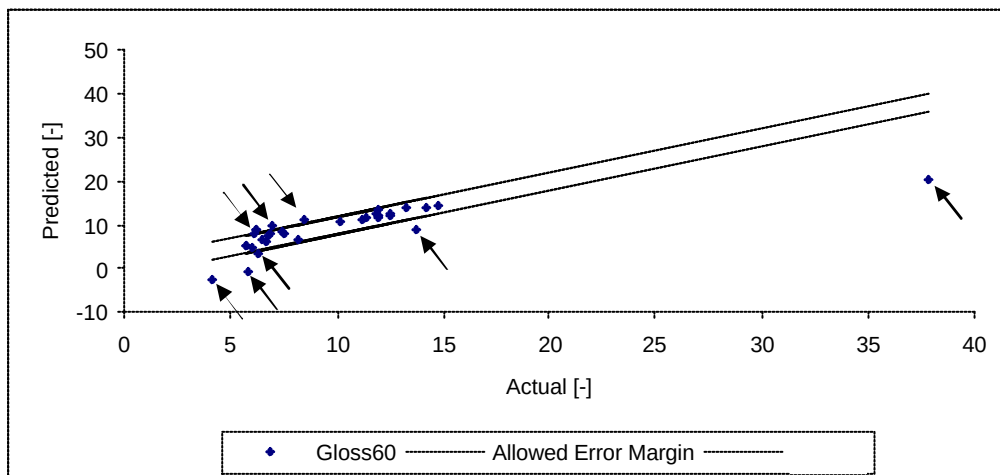


Figure 5.33 Error Margin for Gloss60

- 24% of predicted Gloss60 values lie outside the allowed error margin.
- 38% of the inaccurate predictions are paint samples with extreme PVC/VS.
- Four samples with extreme PVC/VS are in the 'unseen' sample size. *Gloss60* for three of them (75%) is predicted inaccurately by the MLR model.

iii. Opacity

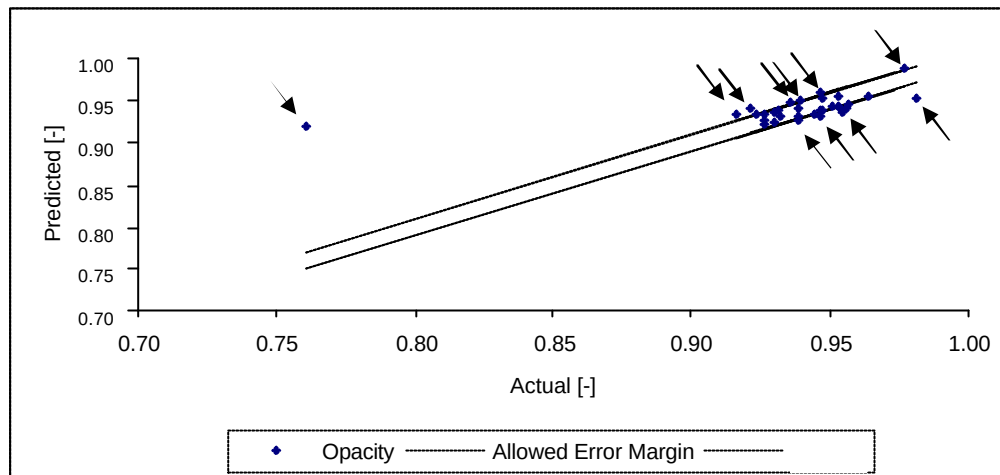


Figure 5.34 Error Margin for Opacity

- 33% of predicted Opacity values lie outside the allowed error margin.
- 38% of the inaccurate predictions are paint samples with extreme PVC/VS.
- Four samples with extreme PVC/VS are in the 'unseen' sample size. The *Opacity* for three of them (75%) is predicted inaccurately by the MLR model.

iv. Dry Film Thickness

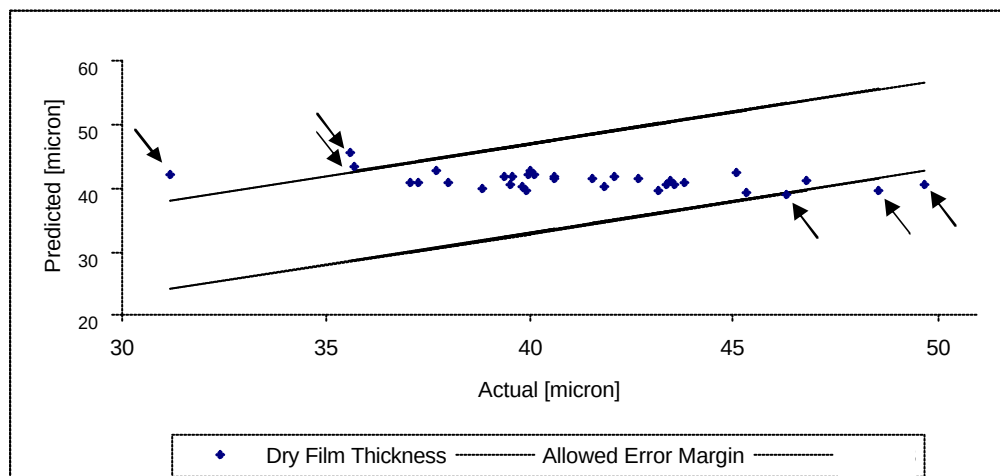


Figure 5.35 Error Margin for Dry Film Thickness

- 18% of predicted Dry Film Thickness values lie outside the allowed error margin.
- 67% of the inaccurate predictions are paint samples with extreme PVC/VS.
- Four samples with extreme PVC/VS are in the 'unseen' sample size. The *Dry Film Thickness* for all four of them (100%) is predicted inaccurately by the MLR model.

Therefore, MLR can be used confidently to predict paint properties from rheological data that originate from samples whose raw materials fall within the 'normal' range (plus/minus 20%

deviation from the standard formulation). However, inaccurate predictions can be expected for samples whose raw materials vary more than plus/minus 20% from the standard formulation.

It was mentioned in *Section 5.2.4.3* that the practical applicability of MODEL 1 is only useful when the quantitative relationship between the rheology and the paint properties are established (MODEL 3). In other words, MODEL 1 would only be useful when it can quantitatively explain whether the predicted differences in rheological behaviour would be significant in affecting a certain paint property. MODEL 3 has established this quantitative relationship between rheology and the paint properties. However, MODEL 3 uses the transformed rheological data, in the form of principal components to predict the paint properties. The principal components are basically a linear combination of all the rheological variables. Reconstruction of the rheological curves (working backwards from the principal components) is possible [5], but is an area of intense statistics that does not fall in the scope of this study, but is considered to be an area for future work.

5.4.5 MODEL 3 CONCLUSIONS

- Table 5.12 gives the matrix of regression coefficients that can be used in equation 5.7 to predict the paint properties from the rheological behaviour of this specific paint.
- The paint property that is predicted most accurately by MLR and the rheological data is opacity. On average, an error of 1.05% can be expected when predicting the opacity value by using MLR and rheology data. This is the case when the raw materials vary within a range of 20% from the standard formulation. However, opacity is also predicted very accurately when the raw materials vary more than 20% from the standard formulation. On average, errors of 1.82% can then be expected when predicting opacity values of these paints.
- Another paint property that is also predicted accurately is the Krebs viscosity. On average, an error of 1.61% can be expected when predicting the Krebs viscosity by using MLR with rheology data as predictor variables and when the raw materials vary within a range of 20% from the standard formulation. The Krebs viscosity of those paint samples that vary more than 20% from the standard formulation, can be expected to have relative errors of 1.74% in new predictions.
- Another paint property that is predicted with less than 10% error by using MLR and rheology data, is dry film thickness (error = 7.11%).
- Paint properties that can be predicted with an expected error of between 10% and 20% are dry and wet burnish, gloss 60, open time and water permeability.
- Paint properties that are predicted the most inaccurate by MLR and the rheology data, that result in errors larger than 20%, are dirt pick-up and sagging.
- Furthermore, an average of 90% of predictions for these paint properties, fall within the allowed specifications range.
- For new paint samples of which the raw materials vary within 20% from the standard formulation, only 5% of new paint samples are predicted outside the allowable Krebs

viscosity error margin. For gloss values, it is only 9%; opacity 13% and dry film thickness 12% of new paint samples that are predicted outside the allowable error margin. These low values indicate that rheology data and multiple linear regression can be used as reliable and accurate modelling technique to predict the critical paint properties of this specific paint.

5.5 MODEL INTERPRETATIONS⁺⁺⁺⁺⁺

All models are developed as tools to investigate the practical implications of a change in the paint formulation.

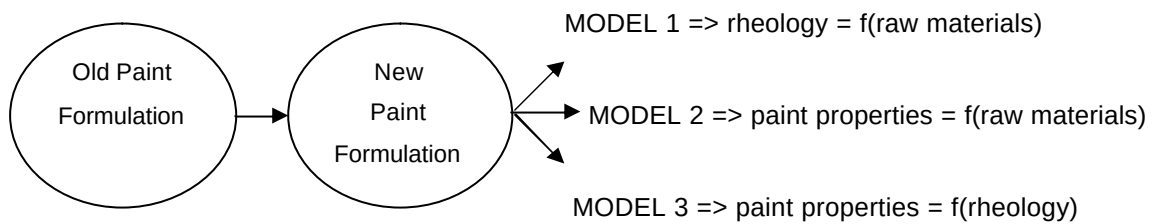


Figure 5.36: Models are Used to Investigate Practical Implications

The paint chemist wants to know what the practical implications are if the standard paint formulation changes in certain levels of raw materials. The models in Figure 5.36 can be used to investigate this behaviour. The following examples illustrate the use of these models to investigate the effect of:

- i) +2% change in pigment
- ii) +5% change in emulsion

The effect of these changes on i) Gloss 60, ii) Krebs Viscosity, iii) Opacity iv) Film Thickness and v) Frequency Sweep (as an example of rheology) is illustrated in Table 5.14

⁺⁺⁺⁺⁺ Model interpretations are not considered to be a fundamental part of this study. It is rather an indication that model predictions from the study can be used for interpretation of future phenomenological models.

Table 5.14 MLR Model Predictions of Change in Standard Paint Formulation ^{ssssss}

Paint Property	Standard Formulation	+2% Pigment	+5% Emulsion	+2% Pigment +5% Emulsion	Model
Gloss 60 [-]	7.73	5.21	21.5	22.56	2
Opacity [%]	98.33	98.90	93.97	93.83	2
Krebs Viscosity [KU]	89	93	91	89	2
Dry Film Thickness [μm]	37	51	49	45	2
Frequency Sweep	See Figure 5.37				1

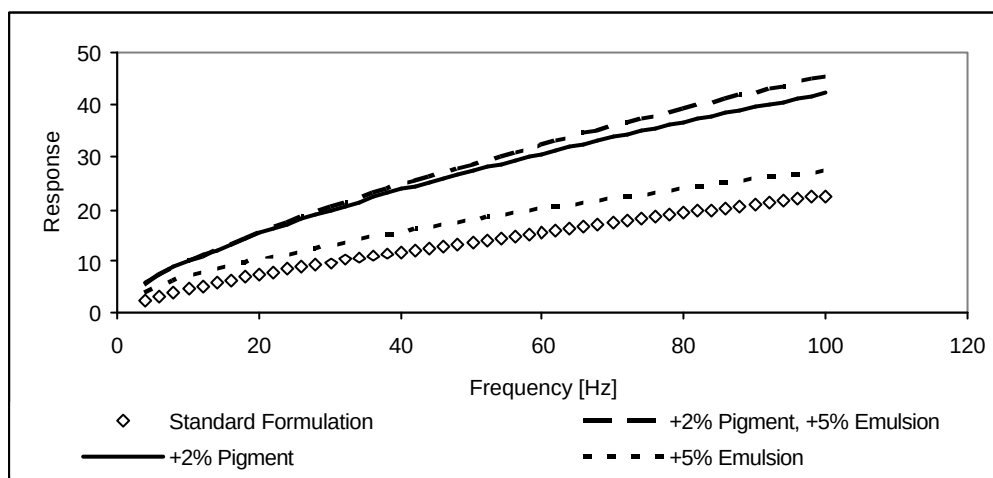


Figure 5.37 Frequency Sweep of Standard Paint Formulation and MLR Model Predictions for Formulations with Variation in Pigment and Emulsion Concentration

It is well known that the level of pigment and emulsion influence all the paint properties in Table 5.14 [17], [18]. It is interesting to interpret these model predictions and to see how they correlate with literature sources and basic paint chemistry.

Gloss 60:

The MLR model predicts a decrease in the gloss level with an increase in pigment concentration, while an increase in gloss is predicted with an increase in emulsion concentration. This prediction correlates with literature sources [17], [18]. According to these sources gloss of paints:

- i) decreases with an increase in pigment concentration [17], [18]
- ii) increases with an increase in emulsion levels [17]

^{ssssss} Model 2 predictions are based on equation 5.6 with model parameters obtained from Table 5.9. Model 1 predictions are based on equation 5.4 with model parameters obtained from Table 5.5.

According to the prediction of the MLR model, the combination of a 2% increase in pigment and 5% increase in emulsion also leads to an increase in gloss levels.

Opacity:

In common with much of paint technology, theory trails practical application and can only partially explain observed phenomena, but as predictive tool, it is overwhelmed by the complexities of real life and humbled by the simplicities of direct experimentation [17].

Opacity or hiding power can be related to the refractive index of the pigment. A widely accepted way of viewing pigment content is the pigment volume concentration (PVC) of the dry paint film.

$$PVC = \frac{\text{Volume of Pigment}}{\text{Volume of Binder} + \text{Volume of Pigment}} \times 100 \quad (5.8)$$

As levels of pigment are increased and the binder reduced, a pvc is reached at which there is just sufficient resin to coat the particles and fill the voids. This situation is obviously complicated by the ease of dispersion of the pigment, the ability of the resin to distribute evenly, the way the particles pack together, the presence of air bubbles and so on. The pigment crowding which causes deviations from 'ideal' behaviour, now causes the maximum opacity to occur at a particle diameter somewhat greater than the theoretical value of 220 nanometers.

The MLR model predictions correlate with these principles. The MLR model also predicts a decrease in opacity with a 2% increase and a 5% increase in emulsion. As mentioned, these are only possible explanations to explain what might happen with the opacity values when pigment and emulsion levels change.

Krebs Viscosity:

The Krebs viscosity is a measure of the paint's consistency and is generally affected by the volume solids (VS) of the wet paint [18]. Pigment and emulsion contribute in the following way to the volume solids (VS) of a paint:

$$VS = \frac{\text{Volume of Binder} + \text{Volume of Pigment}}{\text{Total Volume}} \times 100 \quad (5.9)$$

An increase in both of these raw materials (pigment and binder) will lead to an increase of the volume solids and therefore the Krebs viscosity can be expected to increase. This correlates with the model predictions. However, the model predicts no change in the Krebs viscosity if both pigment and emulsion are increased simultaneously. This is most likely because the resulting ratio of binder and pigment is similar to the original formulation and therefore the effect of increasing both the pigment and binder in such a ratio, is eliminated.

Dry Film Thickness:

Surface coatings are 'dried' polymeric films. For latex paints, the concept of being 'dry' is *mainly* related to the evaporation of water, but there are also other complex mechanisms involved [17]. In order to produce a continuous film it is necessary that the polymer particles begin to flatten and coalesce. Coalescence is the flowing together of polymer particles to form a continuous film. The driving force for coalescence arises from the surface tension and capillary forces that develop as the water evaporates and the polymer particles begin to pack together. It can be visualised that the polymer is being forced to fill the voids vacated by the vaporising water. The better the polymer particles can fill the vacated voids, the denser the film will be. With a 5% increase in polymer there is an excess of polymer particles, after all the voids have been filled. The excess polymer particles will result in a thicker film. The model predicts a 12 μm increase with a 5% increase in emulsion. The model therefore correlates with basic principles of paint chemistry.

Pigment particles provide another surface for the polymer to effectively coat. In general, the surface area of pigment particles will be greater than that offered by the substrate. To produce a solid, coherent film it is necessary that the polymer cover the total surface area of the substrate and pigments. A 2% increase of pigment on the standard formulation will result in a greater surface area but with the polymer at a constant level, the packing of the polymer and pigment particles are not efficient. This results in a lower density and greater film thickness. The MLR model predicts a increase of 14 μm in the dry film thickness with a 2% increase in pigment.

According to the model the film thickness is more sensitive for a change in pigment than it is for a change in emulsion. A 2% increase results in a greater film thickness than a 5% increase in emulsion. It is therefore also interesting to look at the model prediction for a combined effect of a 2% and 5% increase in pigment and emulsion. In this case the model predicts the smallest increase in dry film thickness (8 μm). As mentioned, the underlying film formation is complex and one can only speculate as to what is occurring. The volume solids increase with an increase in pigment and emulsion and therefore the resulting film is thicker because there are more solids. However it is not as thick as expected (combination of pigment and emulsion effect), possibly because the increase in pigment (surface area) is better covered with more available emulsion, resulting in a more dense film (lower film thickness).

Frequency Sweep (Rheology)

Rheology, a more comprehensive way to look at the flow behaviour of paint than Krebs Viscosity, is generally affected by the volume solids (VS) of the wet paint [18]. An increase in the frequency sweep indicates an increase in the structural strength of the paint [23]. The model predicts an increase in the frequency sweep, which can be related to the increase in volume solids of the paint. The model predictions correlate with basic rheology principles.

Table 5.14 only illustrates MODEL 1 and MODEL 2 predictions. MODEL 3 predictions are based on the measured rheology in the form of the principle components. The example given above is based on a variation in raw materials and therefore the rheology needs to be determined first and converted into principle components for MODEL 3 input. MODEL 3 input can be obtained without actually measuring the rheology (via MODEL 1 output) but this will lead to input that are already not entirely precise (because they are model predictions) and therefore the MODEL 3 predictions will have accumulated a lot of variance that might lead to inaccurate predictions. In other words, MODEL 3 needs actual rheological data for the most accurate predictions. However, the objective with MODEL 3 is to reduce the time spent on testing/evaluating a new paint formulation and not to produce a theoretical model.

5.6 GENERAL CONCLUSIONS

The following are the most important conclusions from this chapter:

1. Multiple linear regression is an accurate modelling technique for the prediction of the rheological behaviour of this specific paint. Model 1 predicts the rheological behaviour with raw materials as input. Table 5.5 gives the matrix of regression coefficients that can be used in equation 5.4 to predict the rheological behaviour of 9 rheology curves. 80% of new model predictions will be predicted with 95% confidence. The rheology curves that are predicted most accurately are the amplitude sweep (AS), 3-interval-thixotropy-test (3-ITT) and the flow curve (FC). Average model prediction errors are given in Table 5.6.
2. Multiple linear regression is an accurate modelling technique for the prediction of the paint properties of this specific paint. Model 2 predicts the paint properties with raw materials as input to the model. Table 5.9 gives the matrix of regression coefficients that can be used in equation 5.6 to predict the paint properties (Model 2). Figure 5.38 shows that critical paint properties (opacity, gloss, Krebs viscosity and dry film thickness) are predicted with relative errors smaller than 10%. An average of 90% of predictions for these paint properties fall within the allowed specifications range, set out by the paint manufacturer.
3. A novel model has been developed that predicts the paint properties with rheology data as input to the model (Model 3). Multiple linear regression has proven to be an accurate modelling technique for this purpose. Table 5.12 gives the matrix of regression coefficients that can be used in equation 5.7 to predict the paint properties of this specific paint. The strength of this model lies in the fact that a set of 9 rheology curves are combined by the use of principal component analysis to predict the paint properties. Figure 5.38 shows that accurate model predictions are obtained for the critical paint properties with relative errors smaller than 10% for opacity, Krebs viscosity and dry film thickness.
4. Figure 5.38 shows that the rheological model (Model 3) is almost just as good as the raw material model (Model 2) in prediction of paint properties. Model 2 can be used for further studies in identifying a possible phenomenological model describing paint properties. The advantage of Model 3 is that accurate predictions can be made by measurement of the

rheology only, without any knowledge of the raw material composition. In these terms, it can be classified as a “black box” model.

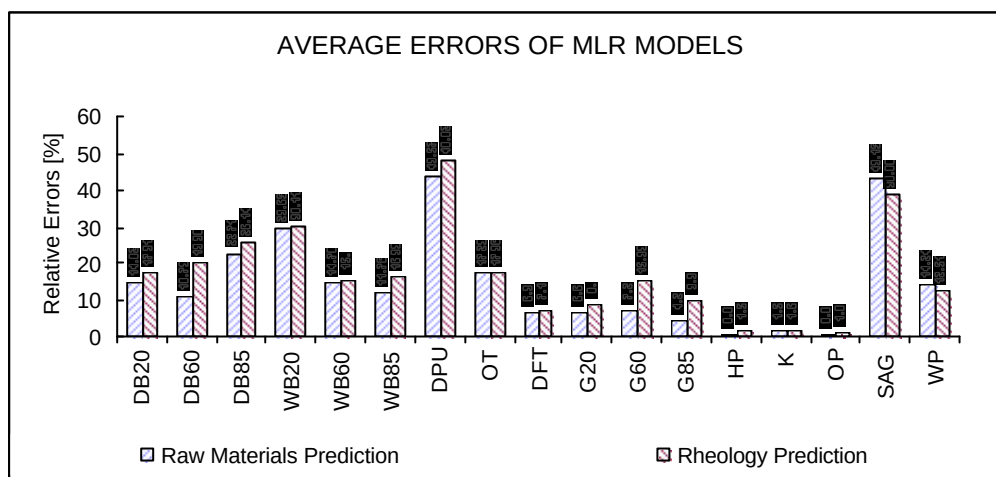


Figure 5.38 Relative errors in New Predictions

(The paint property nomenclature is as follows: DB20 – Dry Burnish20, DB60 – Dry Burnish60, DB85 – Dry Burnish85, WB20 – Wet Burnish20, WB60 – Wet Burnish60, WB85 – Wet Burnish85, DPU – Dirt Pick-Up, OT – Open Time, DFT – Dry Film Thickness, G20 – Gloss20, G60 – Gloss60, G85 – Gloss85, HP – Hiding Power, K - Krebs, OP - Opacity, SAG - Sag, WP – Water Permeability).

- Linear models are inaccurate in prediction of rheology and paint properties of which the raw materials vary more than 20% from the standard formulation. Alternative models are required for these paint samples.

5.7 NOMENCLATURE

Abbreviations

3-ITT	Three-interval-thixotropy test
3-ITT (osc)	Three-interval-thixotropy test in oscillatory mode
3-ITT (rot)	Three-interval-thixotropy test in rotational mode
ANTF	Antifoamer
AS	Amplitude sweep
CLNT	Coalescent
CNST	Constant
COMB	Combination
C _{XLF}	Constant of extra low frequency
DISP	Dispersant
EMLS	Emulsion
FC	Flow curve
FNGD	Fungicide
FS	Frequency sweep

HS	High shear
H_0	Null hypothesis
LS	Low shear
MLR	Multiple linear regression
OPIG	Organic pigment
PC	Principal component
PIGM	Pigment
PVC	Pigment volume concentration
RM1	Rheology modifier 1
RM2	Rheology modifier 2
RMC	Raw material contribution
SA	Soda ash
SOLV	Solvent
SURF	Surfactant
TS	Time sweep
VS	Volume solids
WTR	Water
XLF	Extra low frequency
XTND	Extender

Symbol	Description
C	Condition number (-)
F	Frequency (s^{-1})
G'	Storage modulus (Pa)
G''	Loss modulus (Pa)
G^*	Complex modulus (Pa)
t	Time (s)
x	Independent variable
x_{avg}	Average value of variable
x_i	Value of variable
$x_{standard deviation}$	Standard deviation of variable
$x_{standardised}$	Standardised value of variable
y	Dependent variable

Greek Symbol	Description
a	Preliminary model parameter (gradient)
a_i	Gradient of preliminary model where i=HS, LS, 3-ITT _(rot) , etc

$a_{oscillatoy}$	Gradient of oscillatory measurement
$a_{Gloss60}$	Predicted Gloss60 value
$a_{PaintProperty}$	Value of predicted paint property
$a_{rotational}$	Gradient of rotational measurement
b	Preliminary model parameter (offset)
b_i	Preliminary model parameter (offset) where $i=HS, LS, 3-ITT_{(rot)}, etc$
$b_{oscillatoy}$	Intersect of oscillatory measurement
$b_{emulsion}$	Regression coefficient associated with the emulsion
$b_{rotational}$	Intersect of rotational measurement
b_0	Constant associated with regression analysis
b_k	Regression coefficient associated with variable $k=1,2,3,...$
b_{PC1}	Regression coefficient associated with principal component 1
g	Deformation [%] or strain [-]
g_A	Deformation [%] or strain [-] amplitude
$\dot{g}$	Shear rate [1/s]
d	Phase shift, loss angle [°]
e	Random error
h	Shear viscosity [Pa.s]
h^*	Complex viscosity [Pa.s]
t	Shear stress [Pa]
t_A	Shear stress amplitude [Pa]
w	Angular frequency [rad/s]

5.8 REFERENCES

1. Neter, J., Kutner, M. H., Nachtsheim, C. J., Wasserman, W. (1996). Applied Linear Statistical Models, Chapter 8. Tom Casson.
2. Noveon, Statistical Methods for Optimising Coating Formulations (2006). The Lubrizol Corporation©
3. ASK Chemicals, House Paints with Improved Levelling and Excellent Brushability (2006). Ashland-Sudchemie-Kernfest GmbH.
4. Allison, P. D. (1999). Multiple Regression, A Primer, Chapter 1. California, SAGE Publications Ltd.
5. Chatterjee, S., Price, B. (1991). Regression Analysis by Example, Second Edition. Chapter 1,3,7. New York, John Wiley & Sons, Inc.

6. Engineering statistics handbook, NIST. 2006.
(<http://www.itl.nist.gov/div898/handbook/pmd/section1/pmd142.htm>)
7. Montgomery, D. C. (2005). Design and Analysis of Experiments, Chapter 1-7. John Wiley and Sons Inc.
8. Demuth, H., Beale, M., (2000). Neural network toolbox user's guide, for use with MATLAB, Mathworks.
9. StatSoft (2006). Statistica 7.1.
10. Gasperlin, M. (1995). "Empirical modelling of rheological and mechanical properties of paint." Rheology: 140.
11. Leskovsek, N., Tusar, M., Tusar, L. "Study of rheological properties of paints with designing of experiments: a neural network approach." p. 709. Supplied by The British Library.
12. Hester, R. D., Squire, D. R. (1997). 'Rheology of waterborne coatings." Journal of Coatings Technology **69**(864): 109.
13. Franck, A., Technical Report: Rheological characterisation of paints and coatings, p.1-6. TA Instruments.
14. Osterhold, M. (2000). "Rheological methods for characterising modern paint systems." Progress in Organic Coatings **40**: 131-137.
15. Gasperlin, M., Tusar, L., Tusar, M., Smid-Korbar, J., Zupan, J., Kristl, J. (2000). 'Viscosity prediction of lipophilic semisolid emulsion systems by neural network modelling." International Journal of Pharmaceutics **196**: 37-50.
16. Aldrich, C. (1997). Neural Networks for the Process Industries. Unpublished Manuscript, Stellenbosch.
17. Parsons, P. (1993). Surface Coatings: Raw materials and Their Usage, Chapter 17, 24, 26, 28, 31. London, Chapman and Hall.
18. (2003). Introduction to Paint Technology. Johannesburg, SA Paint Industry Training Institute: 1-48.
19. Eley, R. R. (2001). Applied Rheology, ICI Paints. Strongsville, OH 44136
20. Leskovsek, N., Tusar, M., Tusar, L. (1995). "Optimisation of rheological and mechanical properties of paint using methods of experimental design and neural network modelling (Polish)." p.258-259. Medvode, Ljubljana.
21. Tusar, L., Tusar, M., Leskovsek, N. (1995). "A comparative study of polynomial and neural network modelling for the optimisation of clear coat formulations." Surface Coatings International **10**: 427-434.
22. Tusar, L., Tusar, M., Barle, N., Leskovsek, N., Kunaver, M. (1995). "A study of the influence on hiding power of the composition of a paint and its film thickness." Surface Coatings International **11**: 473-481.
23. Mezger, T. (2002). The Rheology Handbook, Chapter 8. Hannover, Vincent Verlag.

6 NON-LINEAR MODELLING

6.1 INTRODUCTION

Traditional empirical modelling is based on linear statistical techniques. Nothing in nature is absolutely linear, so it helps to take non-linearities into account rather than to ignore them. Non-linear modelling is empirical or semi-empirical modelling which takes non-linearities into account. There are many ways of doing that, including linear regression with non-linear terms, polynomial regression, non-linear regression, splines, etc. Some of the new techniques based on artificial neural networks have advantages over the traditional methods mentioned above. Multilayer perceptrons, a kind of feed-forward neural network, are capable of approximating any continuous, first order differentiable function to any desired degree of accuracy with a single series (a single layer) of activation functions.

Neural networks are useful mathematical techniques inspired by the study of the human brain. Although the brain is a very complex organ that is still largely an enigma despite considerable advances in neurosciences, it is clear that it operates in a massively parallel model. Unlike most other mathematical techniques, where knowledge is stored explicitly as rules or heuristics, neural networks generate their own implicit rules by learning from examples and therefore reasonable responses are obtained when the network is presented with incomplete, noisy or previously unseen inputs.

Neural networks are computational structures consisting of large numbers of primitive process units connected on a massively parallel scale. These units (nodes or neurons) are relatively simple devices by themselves, and it is only through the collective behaviour of these nodes that neural networks can realise their powerful ability to form generalised representations of complex relationships and data structures. A basic understanding of the structure and functioning of a typical neural network node is therefore necessary for a better understanding of the capabilities and limitations of neural networks.

The main advantage of modelling techniques based on the use of neural networks, is that no *a priori* assumptions with regard to the functional relationship between the inputs and the targets are required. This is especially applicable in the coatings industry where many of the paint properties (including rheology) are dependent not only on the main effects of each of the raw materials but also on the interactions between them. However, *a priori* knowledge can also be incorporated into the neural network model. In this study, *a priori* knowledge about the non-linear behaviour of coatings is obtained from:

1. Literature sources (Chapter 2)

Other researchers ([1], [2], [3], [4]) also noted the non-linear behaviour of paint properties. Leskovsek et al who mostly studied the interaction between rheological behaviour and the paint components (similar to that of MODEL 1), modelled the rheological behaviour with both linear and non-linear models. It was found that non-linear techniques express the relationship between paint rheology and the raw materials more accurately than linear techniques.

Gasperlin et al [5], [6] also found that the relationship between rheology and raw materials (MODEL 1), which was modelled with the complex viscosity, were expressed more accurately with non-linear techniques. Non-linear techniques included both quadratic models and neural networks. It was found that neural networks are superior to both the linear and quadratic models. They also modelled the phase behaviour of pharmaceutical creams (similar to that of MODEL 3) in terms of the rheological parameter $\tan \delta$ more accurately with neural networks than with linear or quadratic models [7].

Tusar et al [8], [9] who modelled the relationship between the paint properties and the raw materials (similar to that of MODEL 2) also found that non-linear techniques such as neural networks were superior to both quadratic and linear models.

2. Data preparation (Chapter 4)

It was already noted in the early stages of data preparation (removing outliers) that possible influential data points exist which do not fall within 2 standard deviations from the average of the dataset. These data points were identified as samples with extreme PVC/solids and were not removed from the data set but were kept for the specific reason to add to the information content of the dataset.

3. Linear modelling (Chapter 5).

Evaluation of the linear models (scatterplots, QQPLOTS, RCOPLLOTS and the prediction intervals) revealed that both the rheology and paint properties behave in non-linear ways for paint samples with higher/lower levels of PVC and solids.

The three points discussed above indicate that paint properties (including rheology) behave in non-linear ways. In Chapter 5 it was concluded that linear models are sufficient to explain the paint behaviour (paint properties and paint rheology) when the raw materials are varied within the 'normal' levels (plus and minus 20% from the standard formulation). However, the paint chemist must often extrapolate beyond the 'normal' levels of raw materials. The linear models produce inaccurate predictions (extrapolations) for the paint samples with higher/lower levels of PVC and VS. It should be noted that neural networks are not nearly as sensitive to ill-distributed data as linear modelling techniques. Neural nets can accommodate variables with distributions that standard linear models simply cannot cope with [10].

Therefore, the aim of this section of work is to implement neural networks that accurately express the non-linear behaviour of the paint. There are various ways to optimise neural networks (neural network architecture, training, etc.). Although the neural networks in this study have undergone

some optimisation, it is not considered to be the major part of this study and therefore further improvement in terms of optimisation of the neural networks are possible (see Chapter 7 – Future Work).

6.1.1 THE SINGLE NODE (NEURON)

Each node consists of a processing element with a set of input connections, as well as a single output connection, as illustrated in Figure 6.1. The interaction between the nodes is characterised by weight values (w_i) associated with the strength of the connections between the nodes.

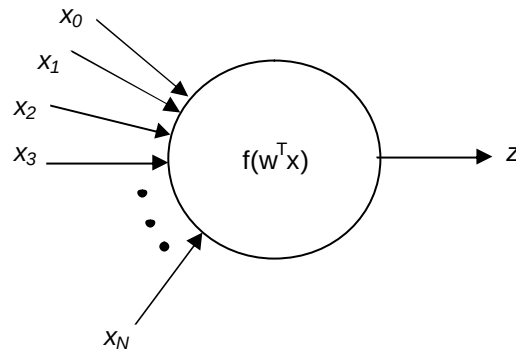


Figure 6.1 Model of a Single Node

The output of the neuron can be expressed as follows:

$$z = f(\sum_{i=1}^N w_i x_i) \text{ or } z = f(w^T x) \quad (6.1)$$

where w is the weight vector of the neural node, defined as:

$$w = [w_1, w_2, w_3, \dots, w_N] \quad (6.2)$$

and x is the input vector, defined as:

$$x = [x_1, x_2, x_3, \dots, x_N]. \quad (6.3)$$

The function $f(w^T x)$ is referred to as the activation function of the node, defined on the set of activation values, which are the scalar product of the weight and input vectors (equations 6.2 and 6.3 respectively) of the node. Sigmoidal activation functions (Figure 6.2) are used widely in neural network applications.

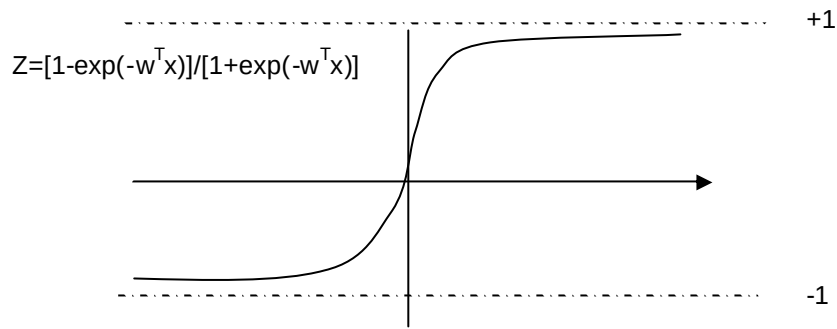


Figure 6.2 Sigmoidal Activation Function

These non-linear functions are popular for a variety of reasons. First of all they are continuously differentiable which enabled the implementation of gradient search learning algorithms. Moreover, in many applications continuous-valued outputs are required.

Additional input can be defined for some nodes, i.e. x_0 , with associated weight w_0 . This input is referred to as a bias and has a fixed value of -1. Like the other weights $w_1, w_2, w_3, \dots, w_N$, the bias weight is also adaptable. The use of the bias input value is sometimes necessary to enable the neural network to form accurate representations of process trends, by offsetting the output of the neural network.

The nodes of a neural network are in reality simple processors that receive input data, operate on the data, and pass them on to successive nodes in the network.

6.1.2 DEVELOPING THE NEURAL NETWORK ARCHITECTURE

A neural network consists of interconnections of nodes. The nodes are usually divided into different subsets, in which all the nodes have similar computational characteristics. A distinction is made between input, hidden and output nodes, depending on their relation to the information environment of the neural network (Figure 6.3). The nodes in a particular subset (layer) are linked to other nodes in successive layers by means of the weighted connections discussed above. Multilayer neural networks can be formed by simply cascading the single-layer neural networks. Theoretical results indicate that a back propagation neural network with more than one hidden layer can approximate any continuous non-linear relationship to an arbitrary degree of accuracy [10].

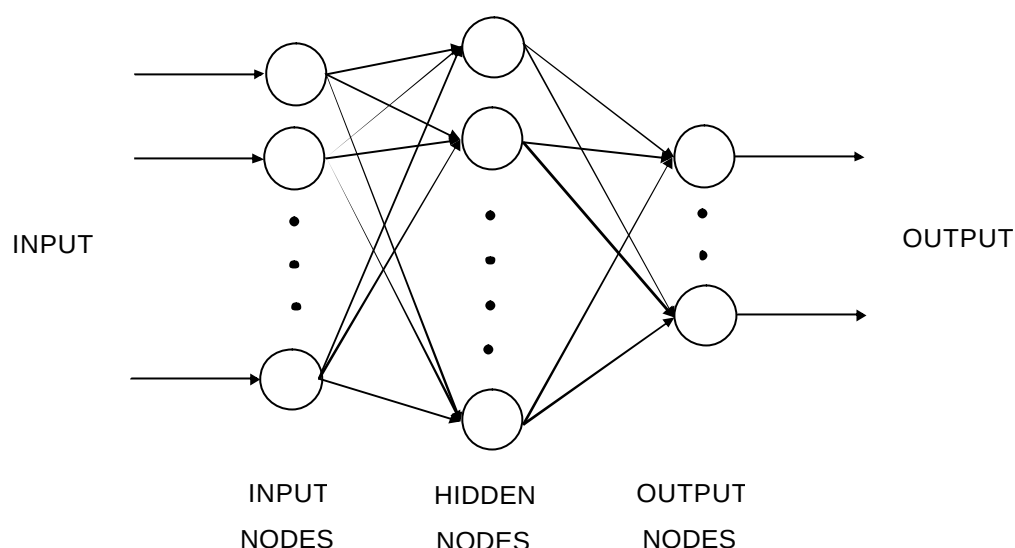


Figure 6.3 Architecture of Multilayer Neural Network

The development of neural network models thus consists of first determining the overall structure of the neural network (number of layers, number of nodes per layer, types of nodes, etc.). One strategy of developing the architecture of a neural net model is known as the 'top-down' strategy. In this, a large complex structure is initially specified and this structure is gradually pruned down as indicated by the ability of the neural net to generalise on the test data. Therefore, the architecture of the neural network is dependent on the ability of the model to generalise. Also, the more complicated the underlying relationships the network has to identify, the more data would be needed for proper generalisation. Too small a network would prevent the net from adequate generalisation, while too large a net might again be unable to generalise the trends in the data, owing to overfitting of the training data.

One method of reducing the complexity of the model is by reducing the number of input variables [10]. This is called principal component analysis, where the set of input variables is replaced by its corresponding set of principal components as discussed in *Chapter 4, Section 4.3.3 Redundant Rheological Data*.

Once the structure of the neural network is fixed, the parameters (weights) of the neural network have to be determined. This is done by training of the weight matrix of the neural network.

6.1.3 NEURAL NETWORK TRAINING

Feed forward neural networks (FFNN) learn by repeatedly attempting to match sets of input data to corresponding sets of output data or target values (a process called supervised learning). The optimised weights constitute a distributed internal representation of the relationships between the inputs and the outputs of the neural network. Learning typically occurs by means of algorithms designed to minimise the mean square error (MSE) between the desired and the actual output of the network through incremental modification of the weight matrix of the network. In feed forward neural networks (FFNN), information is propagated back through the network during the learning

process, in order to update the weights incrementally. As a result, these neural networks are also known as back propagation neural networks. Different training algorithms are possible for back propagation neural networks. The faster algorithms fall into two main categories.

1. Algorithms using heuristic techniques:
 - Variable learning rate back propagation
 - Resilient back propagation.
2. Algorithms using numerical optimisation techniques:
 - Conjugate gradient
 - Quasi-Newton
 - Levenberg-Marquardt.

All of these algorithms use the gradient of the performance function (MSE) to determine how to adjust the weights to minimise the performance. These implementations generally have the following characteristics, i.e. they are computationally intensive, they are massively parallel, and the neural networks require large computer memory. These can put serious constraints on the size of the network. Some algorithms are better than others to ensure that the global minimum in the performance is reached in the shortest time. Often, neural networks get 'stuck' in an area of a local minima, resulting in a weight matrix that does not reflect the true global minima.

Training of the neural network is terminated when the network has learnt to generalise the underlying trends or relationships exemplified by the data. Generalisation implies that the neural networks can interpolate sensibly at points not contained in its training set. A method for improving generalisation is called *early stopping*.

6.1.4 NEURAL NETWORK VALIDATION

With *early stopping*, the available data is divided into three subsets. The first subset is the training set, which is used for computing the gradient of the performance and updating the network weights and biases. The ability of a neural net to generalise is influenced by three factors, i.e. physical complexity of the problem, the architecture of the network and the size and efficiency of the training set. Of these, only the latter two can be controlled. That is why so much attention is given to an appropriate design of experiments (see Chapter 3). The second subset is the validation set. The ability of the neural network to generalise is assessed by means of cross-validation of the validation set, where the performance of the network is evaluated against a novel set of test data, not used during training. The error on the validation set is monitored during the training process. The validation error will normally decrease during the initial phase of training, as does the training set error. However, when the network begins to overfit the data, the error on the validation set will typically begin to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned and used as model parameters for further predictions.

6.1.5 NEURAL NETWORK PREDICTION

The third subset is the test set. The test error is not used during the training, but is only used once the net has been developed, to assess the accuracy of the model. It is also useful to plot the test set error during the training process. If the error in the test set reaches a minimum at a significantly different iteration number than the validation set error, this may indicate poor division of the data set. The presence of random noise can also have a significant effect on the ability of the neural network to discriminate between cases near decision boundaries which will lead to inaccuracy and poor predictions.

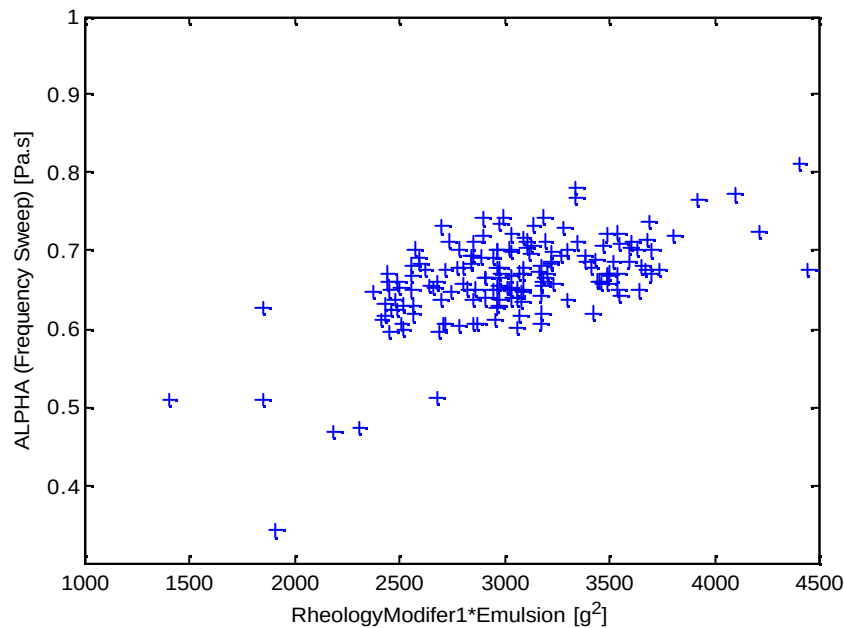
6.2 MODEL 1: RELATIONSHIP BETWEEN PAINT RHEOLOGY AND RAW MATERIALS

6.2.1 BACKGROUND

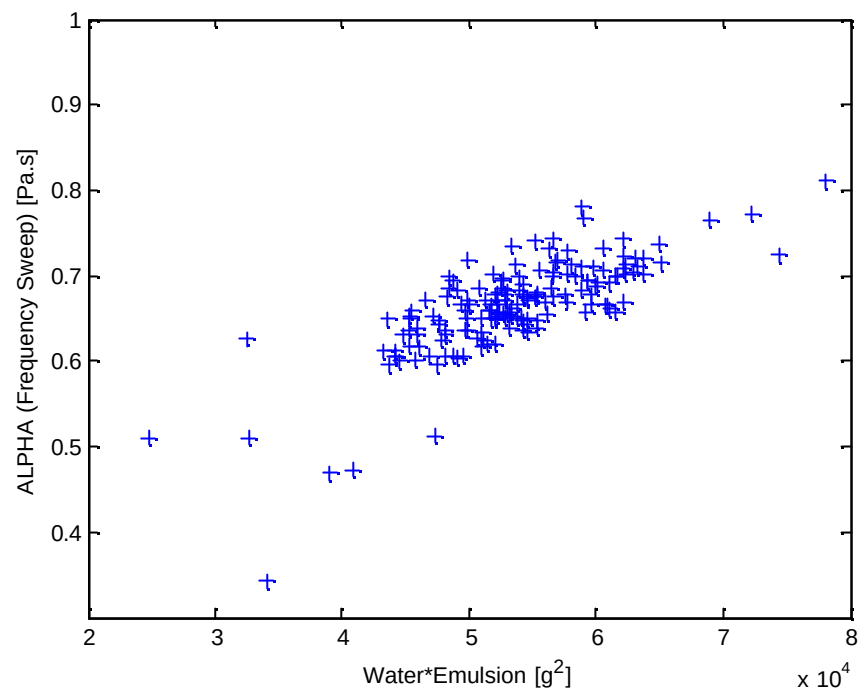
Rheological parameters of the frequency sweep, a_{FS} (gradient) and b_{FS} (offset), are used once again as an example to illustrate the non-linear modelling process. The rheology curve of the frequency sweep is expressed as:

$$\ln y_{FS} = a_{FS} \ln x_{FS} + b_{FS} \quad (6.4)$$

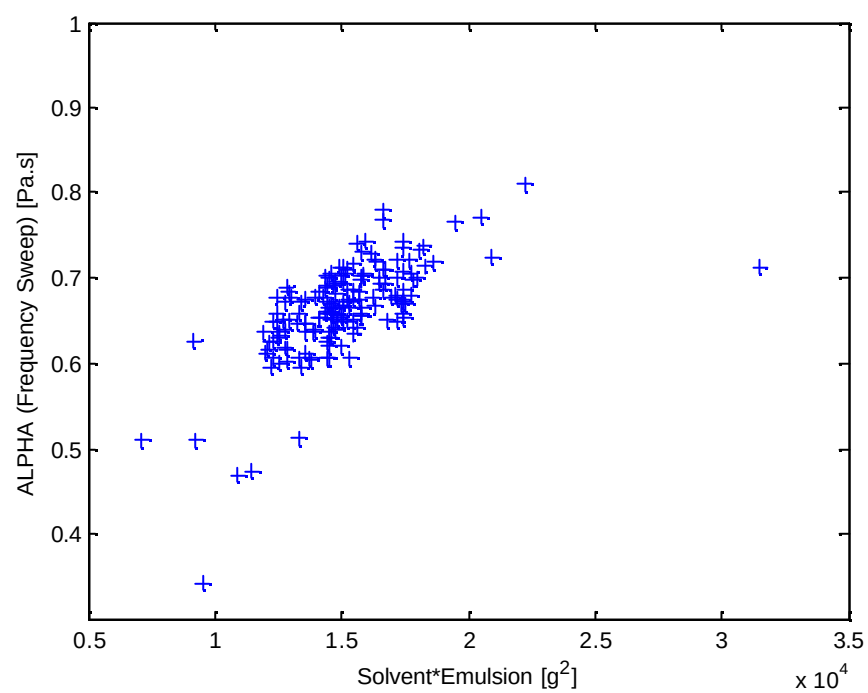
The scatterplot of the response of the frequency sweep gradient (a_{FS}) against some of the interaction terms (Figure 6.4), gives a first (although only graphical) indication of the non-linear behaviour. A pattern (i.e. linear) in these figures will indicate that two terms interact in a certain way so that the response is no longer a result of a single factor.



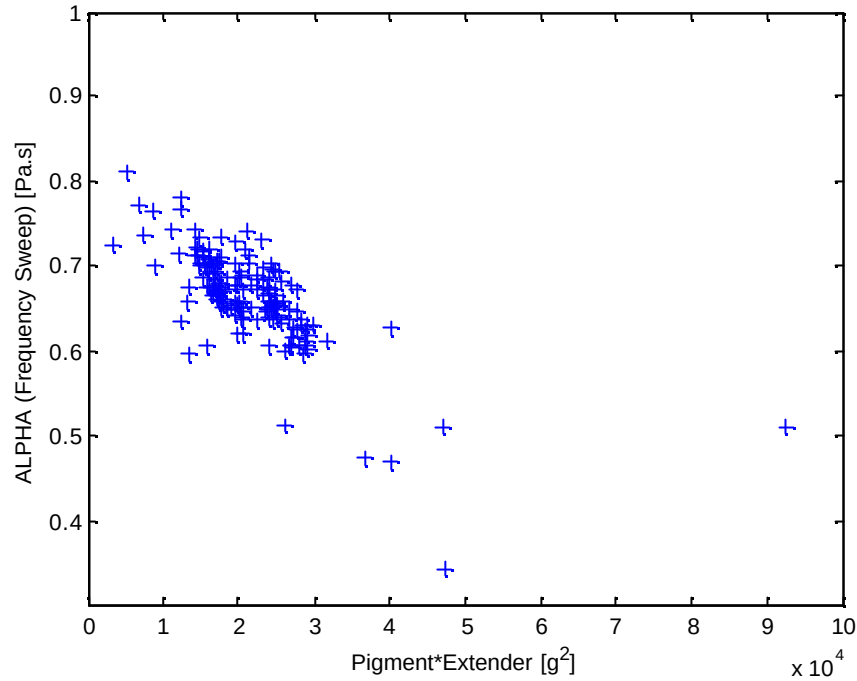
(a)



(b)



(c)



(d)

Figure 6.4 Interaction Terms Affecting ALPHA (Frequency Sweep)

Very definite patterns exist in Figure 6.4 (a) to (d) indicating pairwise interaction between some of the raw materials, noticeably the emulsion with other raw materials. The number of pairwise interactions in the 9 component system amounts to 36. Developing polynomial models with 45 (36 pairwise interaction terms + 9 linear terms) terms make the modelling process very difficult, remembering that a 45-term polynomial does not even take the higher order interactions (3, 4, etc.) into account. Neural networks are ideal for these situations where n types of interactions are possible.

6.2.2 NEURAL NETWORK MODELLING

Neural networks are developed for prediction of the rheological parameters. Linear models predicted the rheological behaviour of the 'normal' samples relatively accurately. *The aim of the neural networks is to predict accurately the rheology of the paint samples whose behaviour is non-linear.* The non-linear behaviour is mostly as a result of large variations in the raw materials (larger than the plus/minus 20% from the standard formulation). These samples were identified in Chapter 5 as samples with residuals that are outside the linear regression confidence interval, as indicated by the RCOLOTS of a_{FS} and b_{FS} , Figure 6.5 and Figure 6.6 respectively.

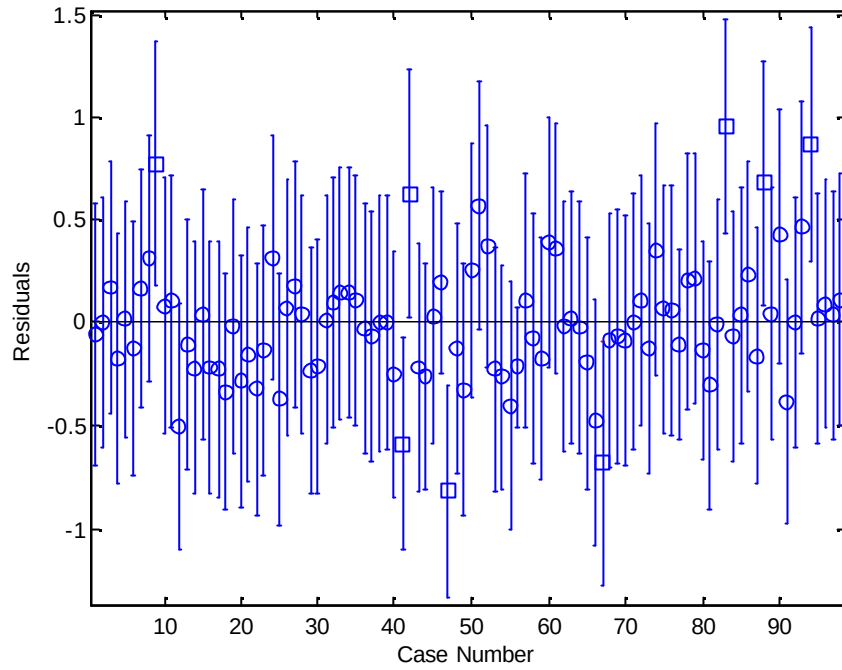


Figure 6.5 RCOplot for ALPHA (Frequency Sweep)

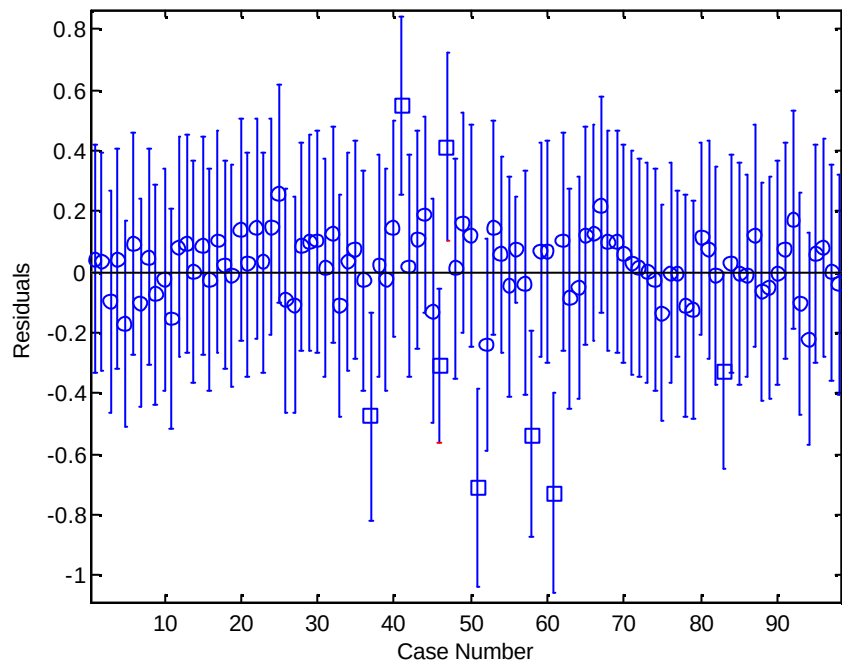


Figure 6.6 RCOplot for BETA (Frequency Sweep)

The residuals that are considered as outliers are indicated in Figure 6.5 and Figure 6.6 as those with square markers (). Their error bars do not cross the zero mean line. These paint samples vary with the following percentages from the standard formulation (Table 6.1).

Table 6.1 Composition of Samples with Outlier Residuals

Case Number	Paint Composition								
	Pigment	Extender	Water	Solvent	Emulsion	Organic Pigment	RM1	RM2	Combination
#9 a_{FS}	20.31	9.83	3.43	15.38	15.03	0.16	18.23	5.98	8.07
#37 b_{FS}	5.50	1.82	9.63	1.95	13.97	20.31	21.21	3.84	10.81
#41 a_{FS} and b_{FS}	16.60	49.95	18.23	20.40	9.31	20.12	20.15	21.57	25.53
#42 a_{FS}	2.90	18.87	16.09	11.50	15.81	17.55	0.54	5.20	7.72
#46 b_{FS}	50.01	28.37	18.17	20.39	50.08	20.11	20.01	19.25	26.14
#47 a_{FS} and b_{FS}	59.84	59.96	23.48	20.03	26.21	19.74	19.91	19.91	10.03
#51 b_{FS}	0.06	21.66	5.79	13.47	13.06	2.46	22.26	5.33	11.55
#58 b_{FS}	20.38	0.95	2.31	16.09	16.31	17.30	0.44	0.07	0.33
#61 b_{FS}	15.89	14.07	11.06	10.01	19.16	12.83	4.14	12.32	11.76
#67 a_{FS}	49.46	23.77	4.52	6.27	21.18	6.90	7.27	5.58	14.03
#83 a_{FS} and b_{FS}	63.24	60.85	23.74	20.76	8.77	20.89	20.81	22.23	10.82
#88 a_{FS}	19.22	7.78	18.42	1.74	17.14	1.37	16.22	2.97	9.33
#94 a_{FS}	7.54	50.08	18.16	20.07	30.38	19.96	20.30	20.79	26.24

Table 6.1 indicates that 6 of the 13 paint samples are samples with variations in raw materials that are larger than the 20% on the standard formulation. It is therefore a clear indication that the linear models do not model these type of paints accurately and that a non-linear model is required to model the samples with large residuals (based on the linear regression model).

6.2.2.1 NEURAL NETWORK ARCHITECTURE AND TRAINING

A neural network is developed and trained so that the residuals of the outliers in Table 6.1 are minimised. The errors involved with these paint samples, based on multiple linear regression are given in Table 6.2.

Table 6.2 Outlier Errors for ALPHA (Frequency Sweep) and BETA (Frequency Sweep)

Case Number	Outlier Error of Frequency Sweep Parameters (Based on multiple linear regression models)	
	a_{FS} (MLR) [%]	b_{FS} (MLR) [%]
#9	6.45	NA
#37	NA	280.52
#41	6.47	24.51
#42	5.04	NA
#46	NA	1239.02
#47	6.77	67.14
#51	NA	1118.14
#58	NA	306.13
#61	NA	381.87
#67	5.78	NA
#83	6.96	1378.32
#88	5.61	NA
#94	6.53	NA
Average:	6.20	599.46

The average errors of the outliers in Table 6.2 (6.20% for a_{FS} and 599.46% for b_{FS}) are large in comparison with the average errors of the values that are not considered as outliers (1.43% for a_{FS} and 11.15% for b_{FS}). Another way to get an idea of the severity of the errors^{*****} associated with these points, can be seen in Figure 6.7 and Figure 6.8.

$$***** \text{ error} = \frac{|actual - predicted|}{actual} \times 100$$

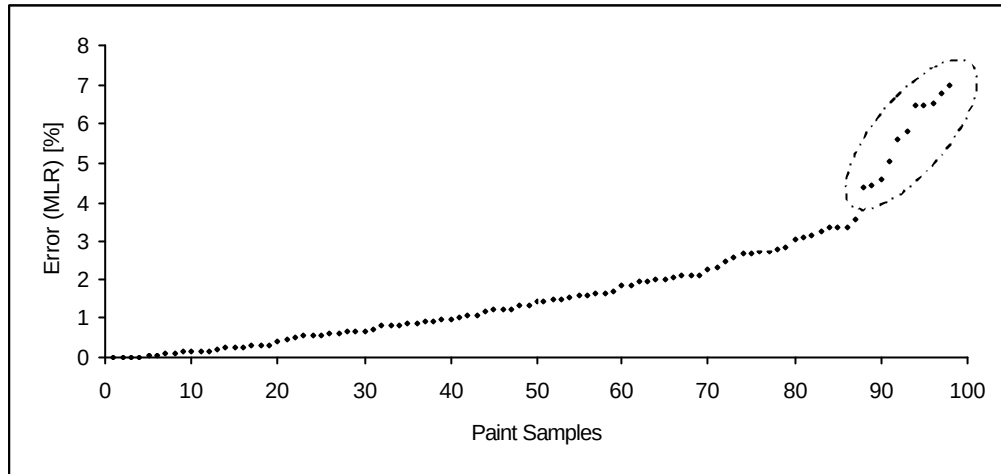


Figure 6.7 MLR Error for ALPHA (Frequency Sweep) (in increasing order)

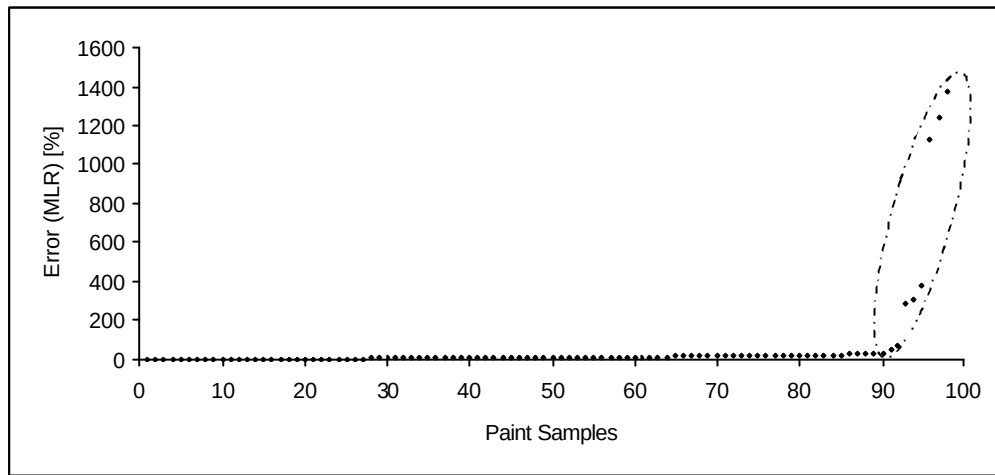


Figure 6.8 MLR Error for BETA (Frequency Sweep) (in increasing order)

The aim of the neural network modelling is to minimise the large MLR errors as marked in Figure 6.7 and Figure 6.8.

There is not a fixed set of rules for developing the architecture of the neural network as every problem is unique. However, selection of specific network architecture (number of layers, number of nodes, etc) also depends on previous experience and knowledge of the problem. It has been shown that standard back propagation neural nets with one hidden layer are very versatile and robust in prediction of some rheological properties (Leskovsek et al ([1], [2], [3], [4]) and Gasperlin et al [5], [6]). Also, a 'top-down' strategy for developing the architecture of the model is used to find the optimum architecture for the neural network model. In other words, a large complex structure is initially specified and this structure is gradually pruned down as indicated by the ability of the neural net to generalise. The effect of the number of output nodes (rheological parameters) of the neural net can be seen in Figure 6.9 and Figure 6.10 for a_{FS} and b_{FS} respectively. The aim is to determine the number of rheological outputs that will result in the smallest error. In Figure 6.9 and Figure 6.10 the last ten paint samples with the largest errors are plotted for each model.

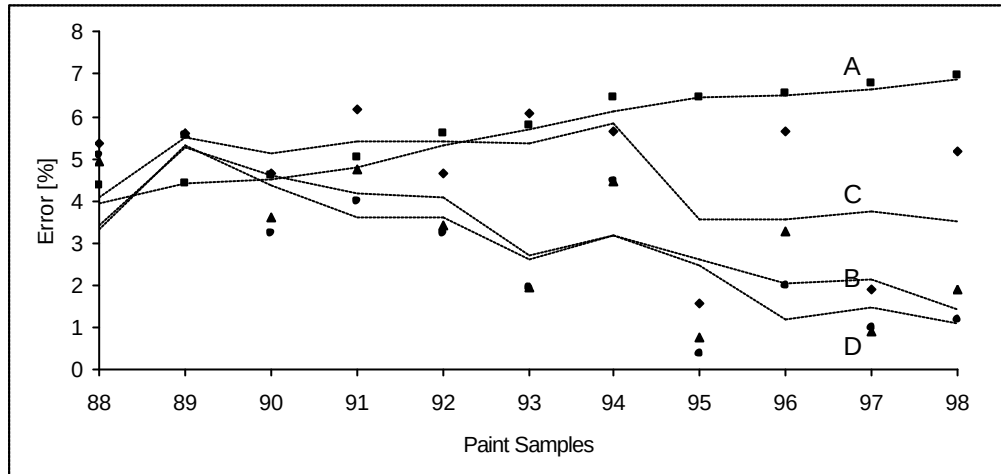


Figure 6.9 ALPHA (Frequency Sweep) Errors for Different Neural Net Architectures(A – MLR, B – NN with 18 outputs; C – NN with 2 outputs; D – NN with 1 output)

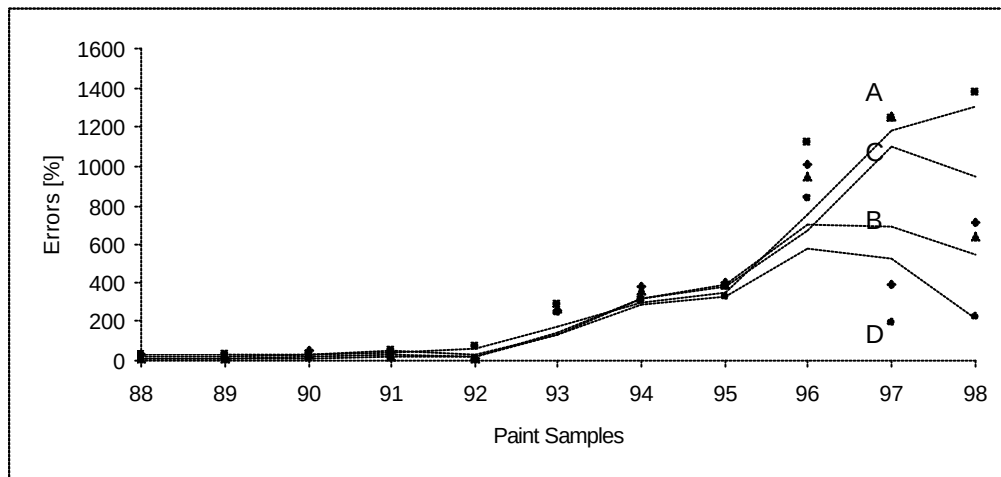


Figure 6.10 BETA (Frequency Sweep) Errors for Different Neural Net Architectures (A – MLR, B – NN with 18 outputs; C – NN with 2 outputs; D – NN with 1 output)

In both cases the linear model (■), with moving average⁺⁺⁺⁺⁺ A, is used as reference. It is clear for both cases that the neural network with 18 rheological outputs (▲), with moving average B, performs better than the linear model. Pruning the number of outputs from 18 to a neural net with 2 (a_{FS} and b_{FS} , ◆), with moving average C, does not result in minimising the error any further and actually increases the error. However, pruning the net further to just one output, a_{FS} in Figure 6.9 and b_{FS} in Figure 6.10, results in the smallest errors (●), with moving average D. The network architecture that produces the minimum error therefore looks as follows (Figure 6.11).

⁺⁺⁺⁺⁺ The average calculated by taking each consecutive point into account.

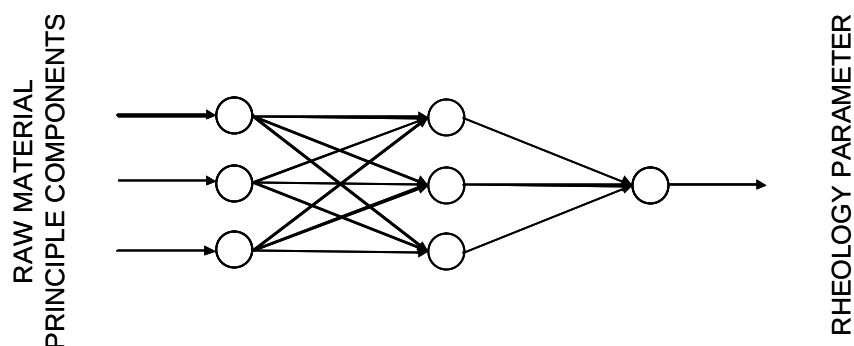


Figure 6.11 Network Architecture for ALPHA and BETA (Frequency Sweep)

Figure 6.11 shows that the neural network architecture used for the determination of a rheological parameter is very basic. The input to the net consists of the three principle components that explain 80% of the variance in the raw materials. The hidden layer consists of 3 hidden nodes each using the tan-sigmoidal activation function to send its summed output to a single output node is used for estimation of the rheological parameter. The advantage of using only one output node is that the specific rheological parameter can be estimated very accurately but a disadvantage lies in the fact that the optimum paint formulation for a combination of rheological parameters can not be estimated (see Chapter 7 – Future Work).

It is clear from Figure 6.9 and Figure 6.10 that the neural net drastically reduces the errors of the paint samples with large linear (outlier) errors and that the architecture of the network has a significant effect in minimising the error as well. Table 6.2 can now be modified to include the outlier errors for a_{FS} and b_{FS} as predicted by the single output neural network (Table 6.3).

Table 6.3 Comparison Between Outlier Error of MLR and NN Models

Case Number	Outlier Error of Frequency Sweep Parameters (Based on MLR and NN models)			
	a_{FS} (MLR) [%]	a_{FS} (NN) [%]	b_{FS} (MLR) [%]	b_{FS} (NN) [%]
#9	6.45	4.48	NA	NA
#37	NA	NA	280.52	245.73
#41	6.47	0.38	24.51	0.38
#42	5.04	3.99	NA	NA
#46	NA	NA	1239.02	198.68
#47	6.77	0.98	67.14	9.82
#51	NA	NA	1118.14	830.97
#58	NA	NA	306.13	317.65
#61	NA	NA	381.87	330.69
#67	5.78	1.94	NA	NA
#83	6.96	1.19	1378.32	226.83
#88	5.61	3.23	NA	NA
#94	6.53	1.97	NA	NA
Average:	6.20	2.27	599.46	270.09

The average errors of Table 6.3 indicate that the neural net (NN) performs significantly better than the linear model (MLR) in fitting the outlier values, although the errors are still large. Appendix E, Section 12.4 – Non-Linear Modelling (MODEL1) gives the 'ACTUAL vs MODEL' graphs for all the rheology parameters modelled by neural networks. Figure 6.12 illustrates the average errors of the outliers for each of the rheological parameters (r1 – r18).

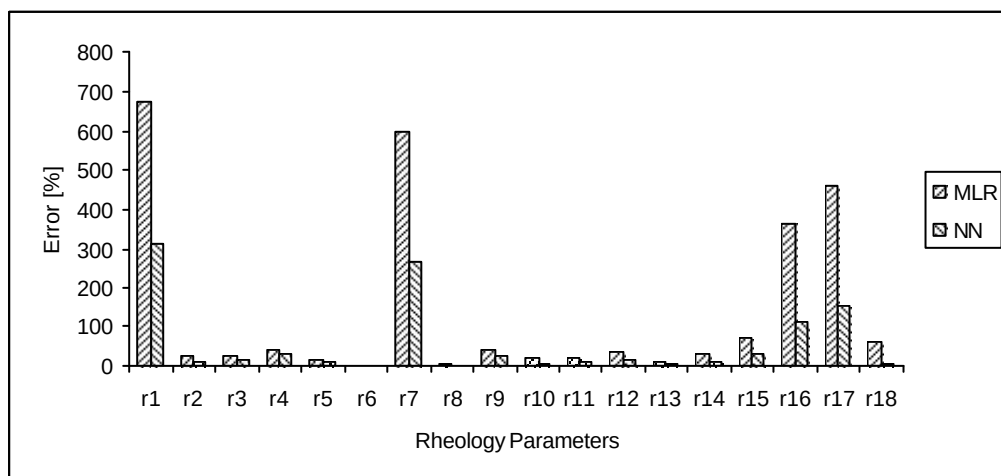


Figure 6.12 Neural Net (NN) and Multiple Linear Regression (MLR) Model Errors

These error values of the outliers are also presented in Table 6.4 for more accurate inspection.

Table 6.4 Model Errors

Rheology Curve	Rheology Parameter	ID	RELATIVE ERROR [%]	
			MLR	NN
3-ITT (Oscillatory)	$b_{3-ITT(Osc)}$	r1	675.61	312.09
	$a_{3-ITT(Osc)}$	r2	24.52	9.27
3-ITT (Rotational)	$b_{3-ITT(Rot)}$	r3	27.36	14.67
	$a_{3-ITT(Rot)}$	r4	41.81	32.31
Amplitude Sweep	b_{AS}	r5	15.56	11.17
	a_{AS}	r6	0.94	0.59
Frequency Sweep	b_{FS}	r7	599.46	270.09
	a_{FS}	r8	6.2	2.27
Flow Curve	b_{FC}	r9	38.22	24.70
	a_{FC}	r10	19.61	7.47
High Shear Curve	a_{HS}	r11	22.03	12.20
	b_{HS}	r12	34.42	14.04
	g_{HS}	r13	12.48	6.54
Low Shear Curve	b_{LS}	r14	29.70	9.29
	a_{LS}	r15	70.94	30.77
Time Sweep Curve	b_{TS}	r16	367.06	113.71
	a_{TS}	r17	462.94	151.9
Extra Low Frequency	C_{XLF}	r18	59.71	5.01

It is clear from figure Figure 6.12 and Table 6.4 that the neural networks (NN) perform significantly better than the multiple linear regression (MLR) models in modelling the outlier values. However, although the neural network models perform significantly better, some of the neural network errors are still relatively large, e.g. $b_{3-ITT(Osc)} = 312\%$, $b_{FS} = 270\%$ and $a_{TS} = 151\%$. This also illustrates the major drawback of neural networks – interpretation of results is almost impossible due to the complex mathematical techniques (rules) that neural network predictions are based on. It must be remembered that these networks are not fully optimised yet in terms of network architecture (number of layers, learning algorithm, etc). It has been shown that optimising the number of

output nodes already has a significant effect in reducing the model errors. Optimisation of the neural network is a study in itself and is not included in the scope of this work (see Chapter 7 - Future Work).

Therefore, although the neural networks are not fully optimised yet, the models should also perform better in predicting unseen data than the MLR model could not predict accurately.

6.2.2.2 MODEL IMPLEMENTATION

The neural networks are now implemented to predict the rheological parameters of 33 unseen paint samples. In Figure 6.13 and Figure 6.14 the error in predictions of the 33 unseen data points are given for a_{FS} and b_{FS} respectively. A comparison is made between multiple linear regression (MLR, ■) and neural network (NN, ●) predictions. The paint samples are placed in increasing order according to the MLR model error values.

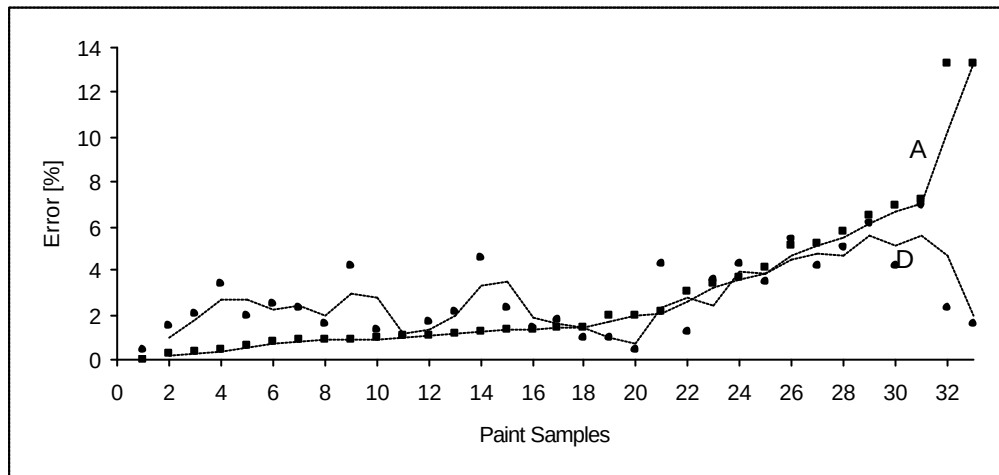


Figure 6.13 Model Errors In Prediction of Unseen Data points for ALPHA (Frequency Sweep): A – MLR; D – NN with 1 output

Although the moving average in error of the NN (D) is higher than that of the MLR model (A) in the first half of Figure 6.13 (up to Paint Sample 17), the much more important part of the figure is the latter part where the moving average D is smaller than the moving average A. It is in this part of the graph that the MLR model performs poorly and predicts the last seven paint samples as outliers (Chapter 5). It can clearly be seen that, although the neural net is not that accurate in predicting every single data point (first half of the figure), the NN generalises better than the MLR model (last half of figure) and predicts the last seven data points more accurately. The same trend is observed in the error values of b_{FS} .

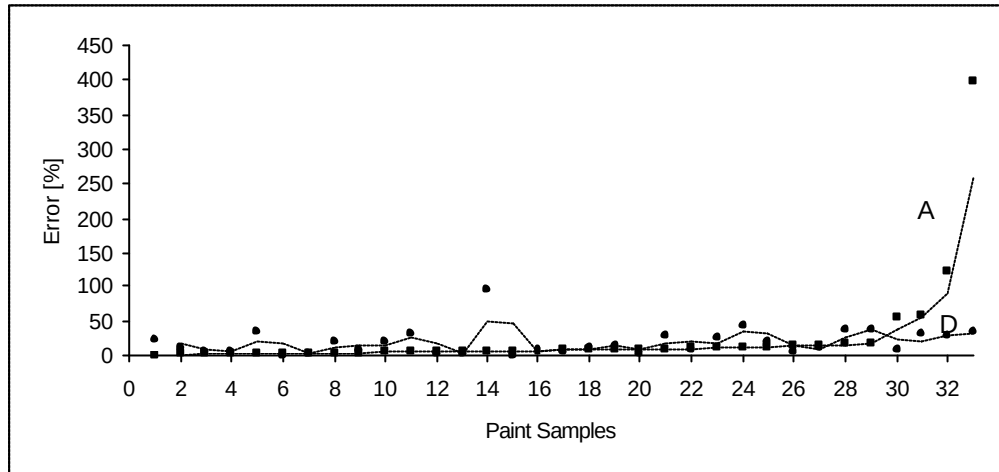


Figure 6.14 Model Errors In Prediction of Unseen Data points for BETA (Frequency Sweep)
(Frequency Sweep): A – MLR; D – NN with 1 output

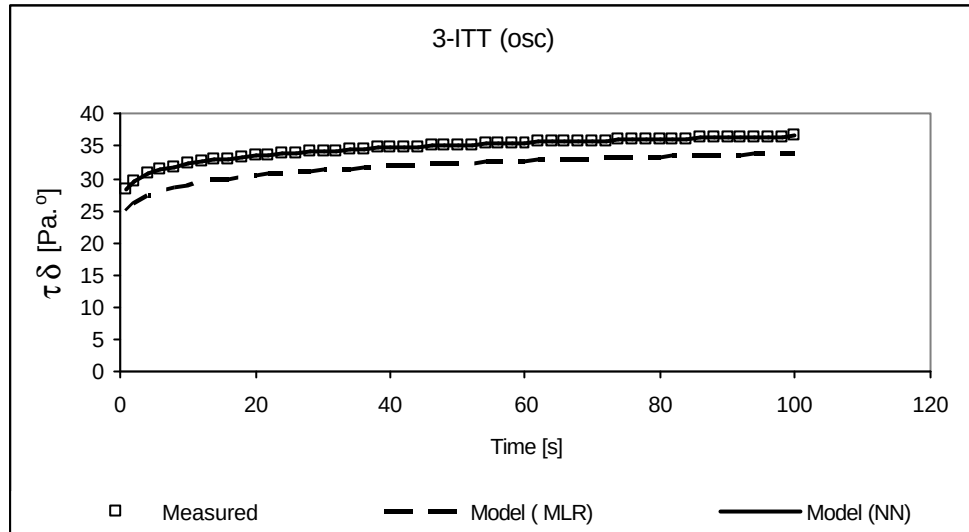
The same trend is followed in the error values of b_{FS} . The MLR model (■) predicts the last four paint samples of Figure 6.14 as outliers (Chapter 5). The neural net is specifically trained so that it will be more accurate in predicting these points. It is in this area of the graph where the NN (●) performs better and it is also in this area where the neural net will be used to make predictions about the rheology of the paint samples with extreme PVC/VS. Table 6.5 illustrates the average improvement in accuracy of predictions of samples that the MLR model could not predict accurately (as indicated as outliers in Chapter 5).

Table 6.5 Outlier Errors of Unseen Predicted Rheology Parameters ALPHA and BETA
(Frequency Sweep)

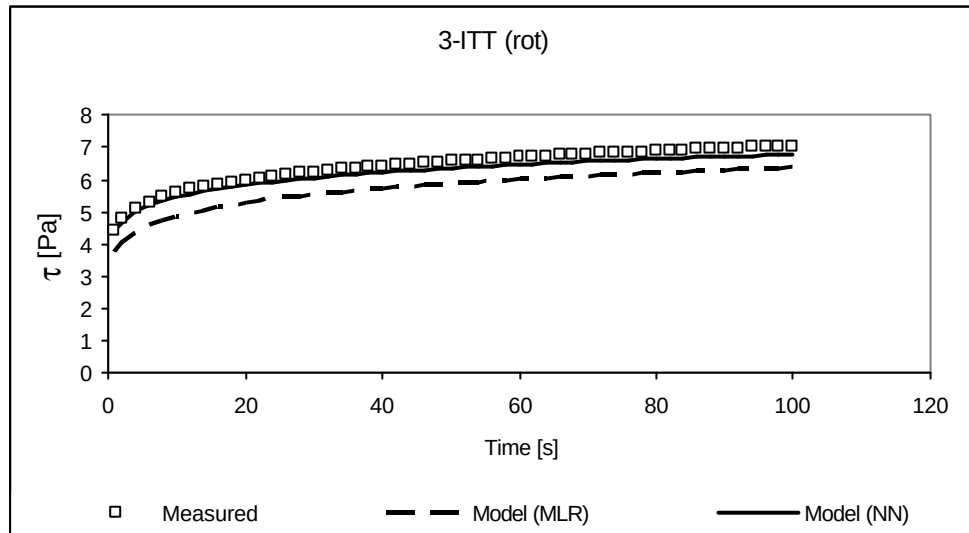
Rheology Parameter	Number of Outliers	Outlier Error [%]	
		MLR	NN
a_{FS}	7	8.32	4.36
b_{FS}	4	157.82	25.42

Clearly, the neural net performs better than multiple linear regression in predicting the rheology parameters of the frequency sweep for samples with extreme variations in raw materials levels.

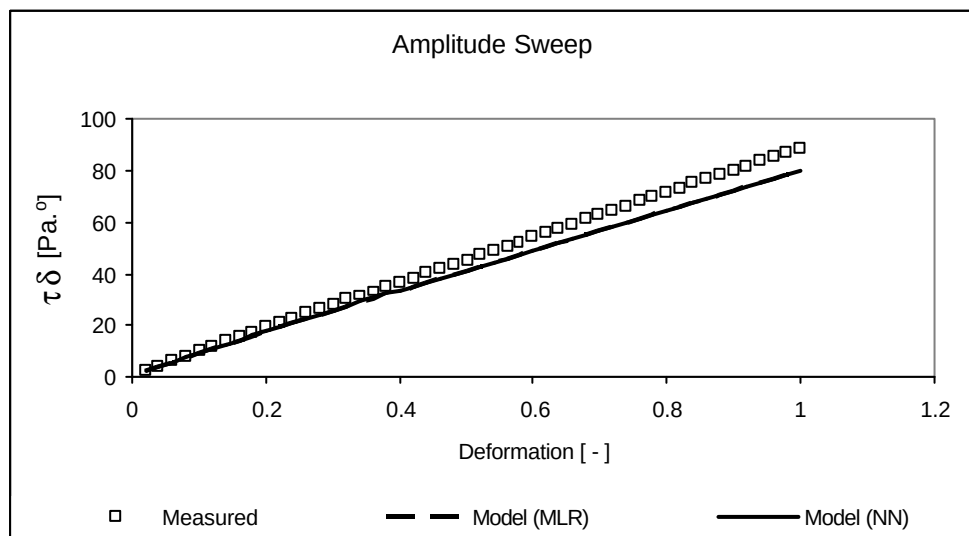
Figure 6.15 now shows an example of the entire set of rheology curves ($\dot{\gamma}$) for a specific paint sample with extreme variations in raw material levels. The model predictions obtained from multiple linear regression (MLR) and neural network (NN) modelling are also shown.



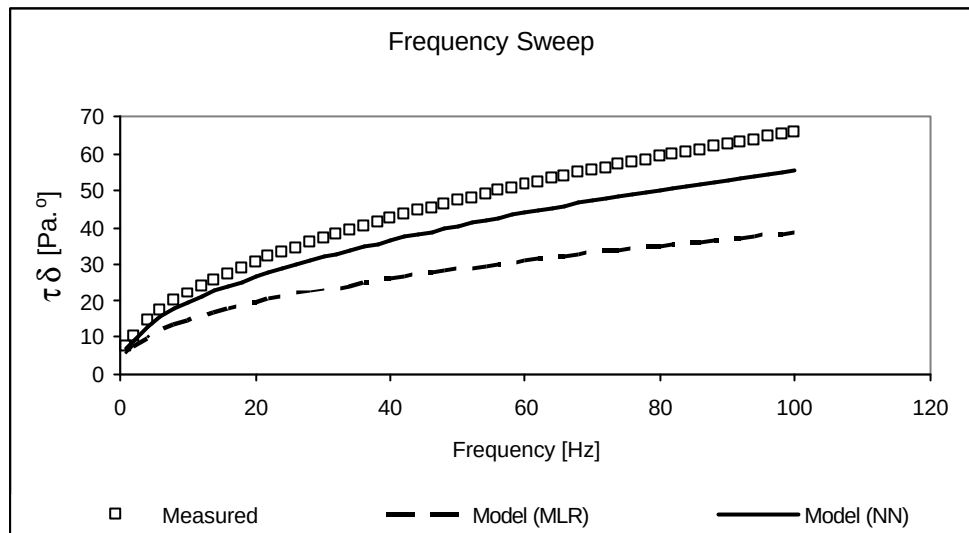
a)



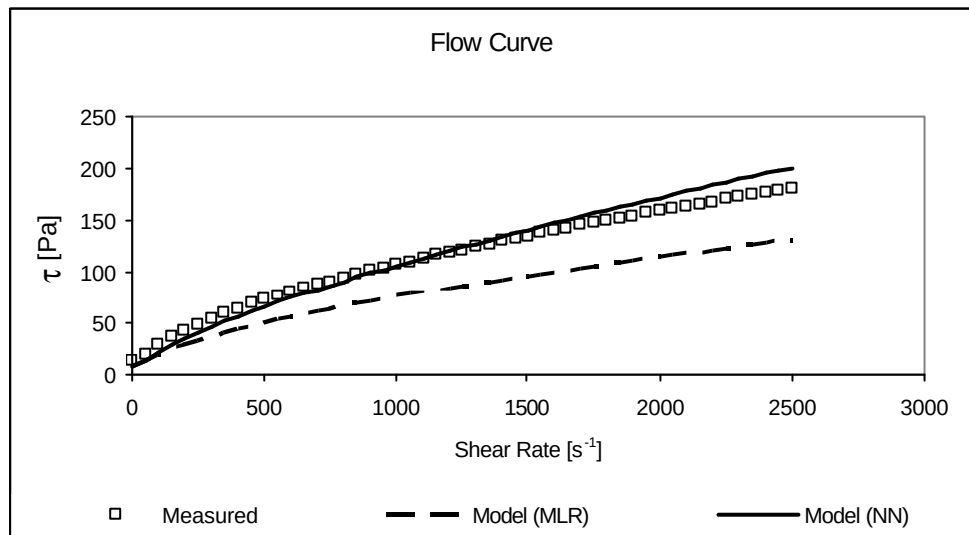
b)



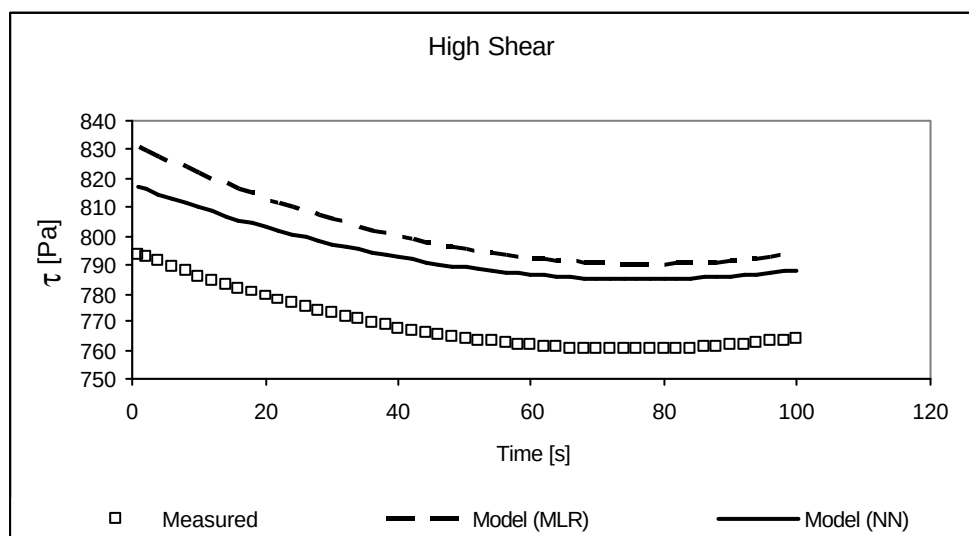
c)



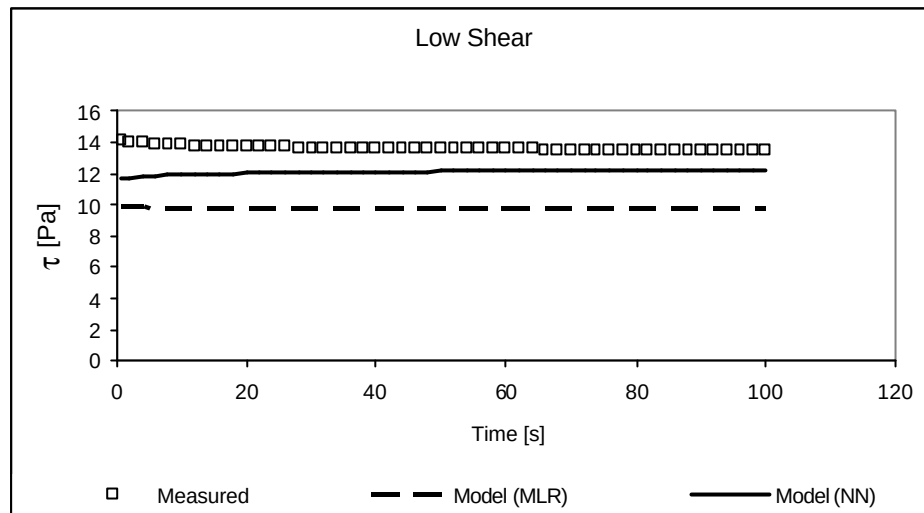
d)



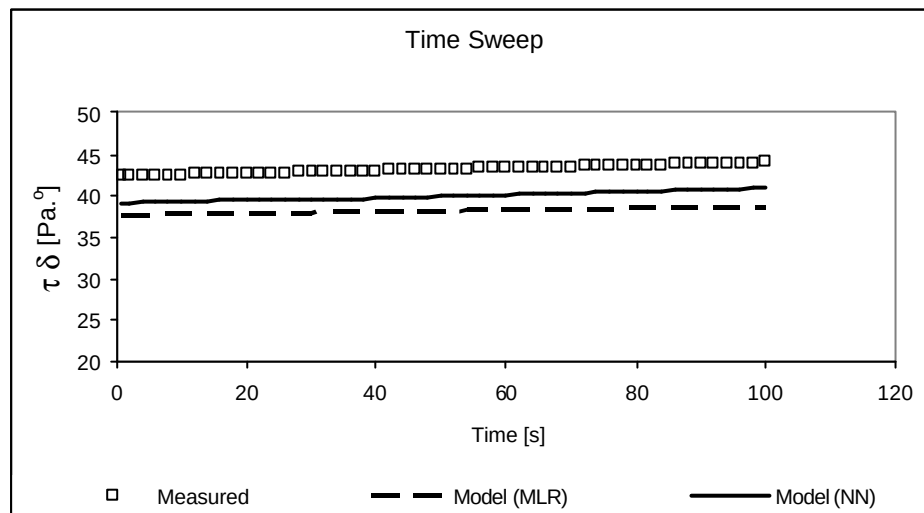
e)



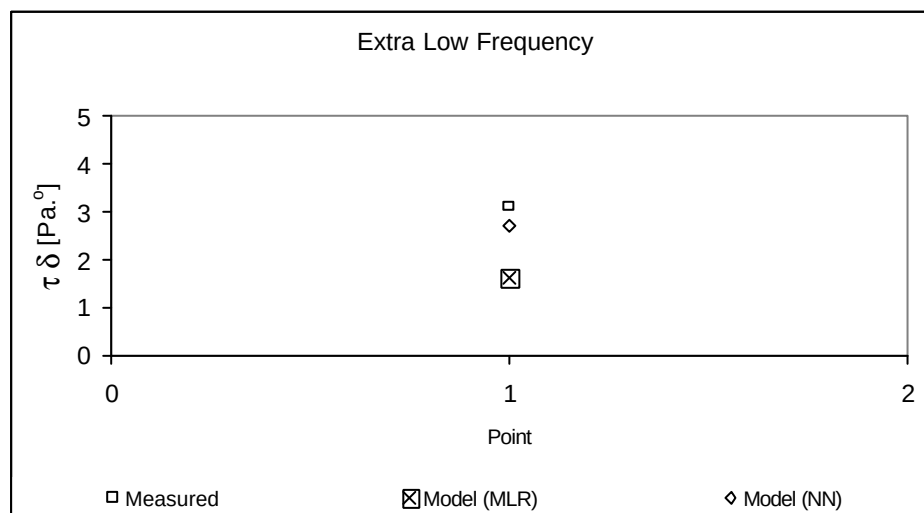
f)



g)



h)



i)

Figure 6.15 Model Predictions (MLR vs NN) for Rheology Curves of a Paint Sample With Extreme Variations in Raw Materials

The neural network predicts the rheological response of all the rheology curves Figure 6.15 (a) to (i) more accurately than the multiple linear regression model. Although the neural networks do not predict the response entirely accurately, it is a major improvement on multiple linear regression models.

Repeatability of the rheological measurements introduces a certain degree of uncertainty around a measurement. Model predictions can at most be as accurate as the band of uncertainty around the true value. Repeatability is within 2% for oscillatory tests and within 5% for rotational tests [12].

6.2.3 MODEL 1 CONCLUSIONS

It became clear after data inspection, the literature study and the linear modelling, that a certain level of non-linearity exists in the rheological behaviour of the paint. In Chapter 5 it was shown that the linear models (MLR models) could not model and predict the non-linear rheological behaviour accurately for paint samples of which the raw materials vary by more than 20% from the standard formulation. Non-linear modelling techniques such as neural networks were developed to overcome these inadequacies. The following are the most important conclusions:

- Neural networks perform better than multiple linear regression in predicting the rheological behaviour of paint samples for which the raw materials vary by more than 20% from the standard formulation.
- Table 6.4 shows the improvement that neural networks make on the prediction of the rheological variables. Although some of the predictions made by the neural networks still result in large relative errors, the neural networks show a major improvement on the predictability of the rheology of those paint samples that vary more than 20% from the standard formulation. On average, neural networks improve predictability of the rheological parameters by 54%. The largest improvement in predictability is made on the rheological variable C_{XLF} , where multiple linear regression resulted in relative errors of 59%, while neural networks resulted in errors of only 5%. Other predictions of rheology curves where neural networks have shown a major improvement on predictability are the: Time Sweep Curve (TS) – 68% increase in accuracy; Low Shear Curve (LS) – 63% increase in accuracy. The smallest increase that the neural network had on the predictability of a rheology curve, was a 33% increase in accuracy of the amplitude sweep (AS) predictions.
- Although it is clear that neural networks result in significantly more accurate predictions for the rheological behaviour of those paint samples that vary more than 20% from the standard formulation, the relative errors are still large, e.g. 312% for $b_{3-ITT(Osc)}$ and 270% for b_{FS} . Optimisation in terms of the neural net architecture is not complete and

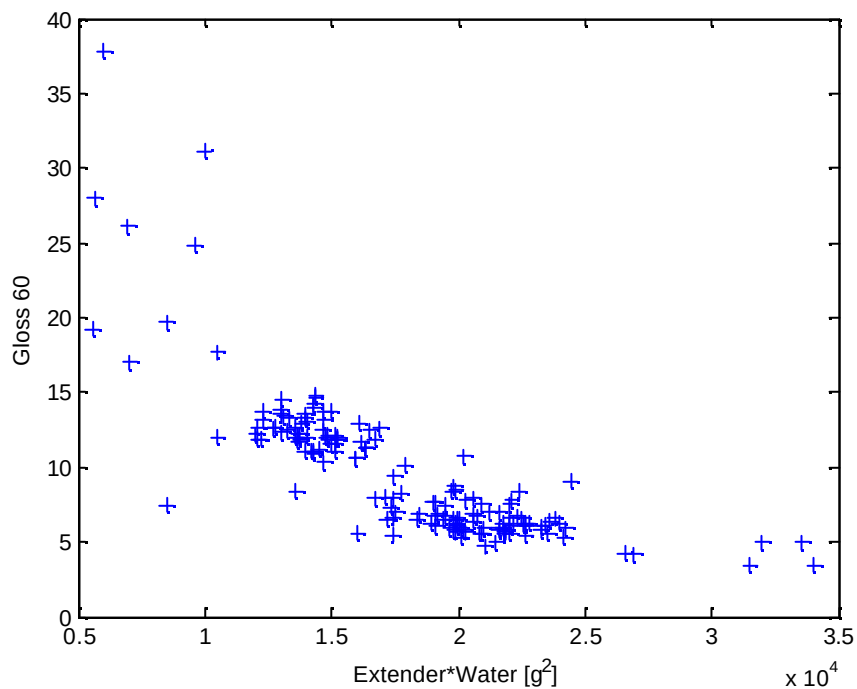
various options are available for further network optimisation (e.g. training algorithm, activation function, number of hidden layers, etc.) that will result in smaller relative errors.

6.3 MODEL 2: RELATIONSHIP BETWEEN PAINT PROPERTIES AND RAW MATERIALS

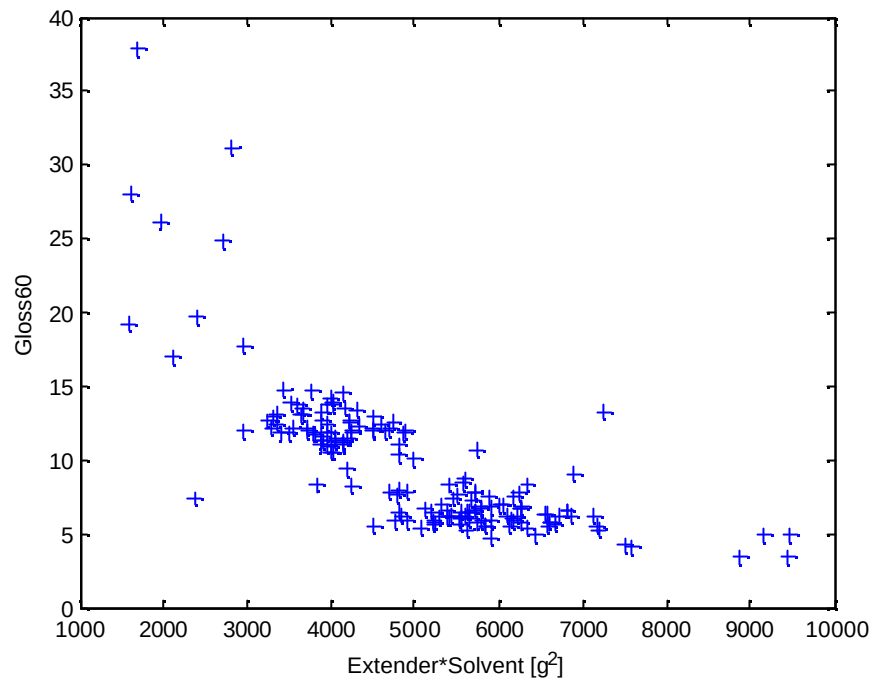
6.3.1 BACKGROUND

In Chapter 5 the linear relationship between the paint properties and the raw materials was evaluated. In the same way that larger variation of raw materials resulted in a certain level of non-linearities in the rheological behaviour, so it is expected that the paint properties will also show some non-linear behaviour. Preliminary indications of non-linearity (literature sources (Chapter 2), data preparation (Chapter 4) and linear modelling (Chapter 5)) were also identified and the need for other non-linear modelling techniques arose.

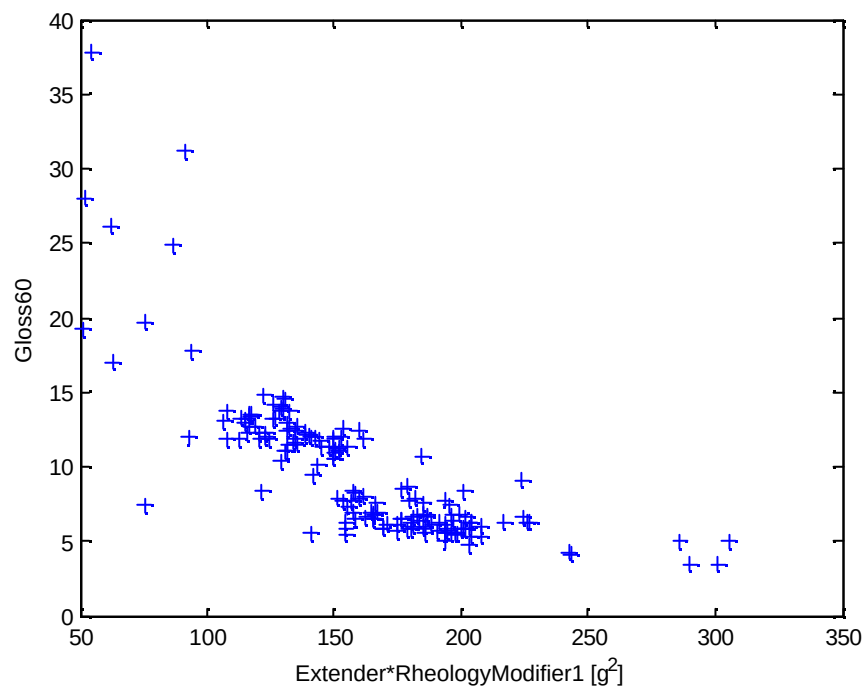
The paint property, *Gloss60*, is used once again as an example to illustrate the non-linear modelling process. The scatterplots of the *Gloss60* values against some of the interaction terms (Figure 6.16), give a first (although only graphical) indication of the non-linear behaviour. Any structure (e.g. linear) in these graphs will be an indication of non-linear behaviour.



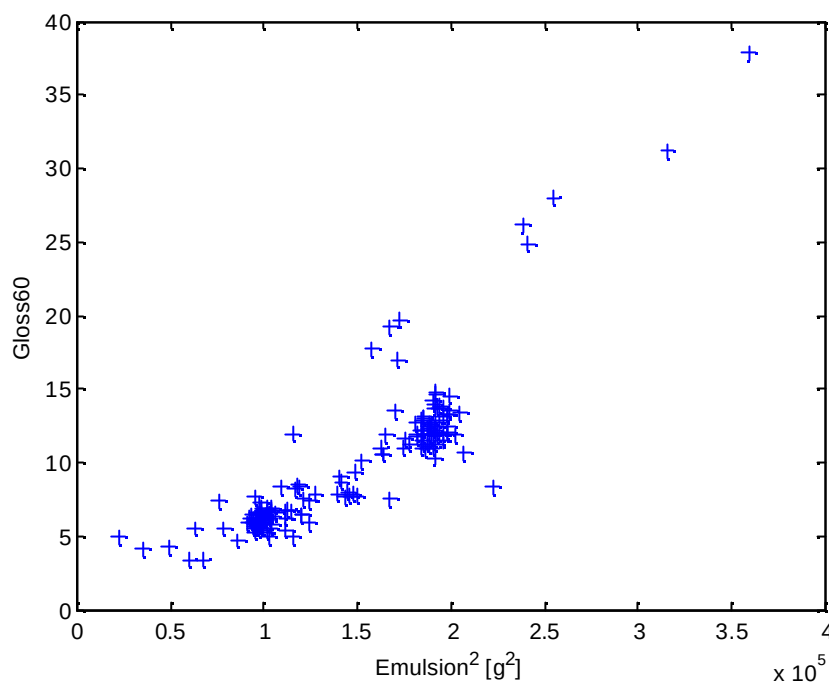
(a)



(b)



(c)



(d)

Figure 6.16 Interaction Terms Affecting Gloss60

From Figure 6.16 (a) to (c) it can be concluded that the *extender* interactions with some of the other raw materials (water, solvent and rheology modifier1) play an important role in affecting the Gloss60 values. Figure 6.16 (d) indicates that the non-linear behaviour of the *emulsion* also play an important role. All of the above are examples of possible interactions and many more (pairwise and higher order) can be expected.

Linear modelling of the Gloss60 values (Chapter 5) also revealed that the linear model is not entirely appropriate for prediction of all of the Gloss60 values. The linear model fit is illustrated in Figure 6.17.

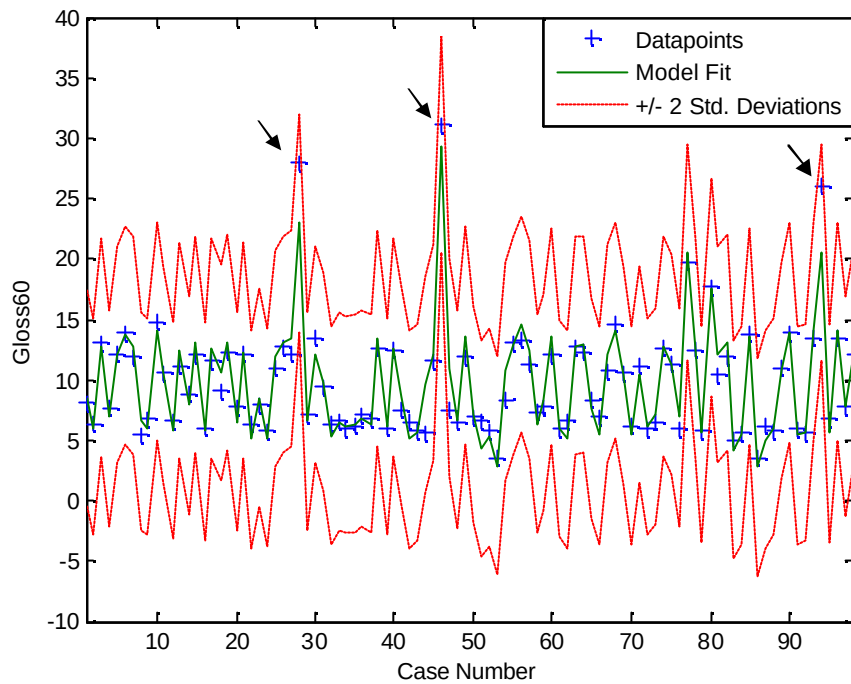


Figure 6.17 Linear Model Fit for Gloss60 Values

Closer inspection of Figure 6.17 reveals that the linear model fits the data points, in general relatively well; however, some extreme data points (indicated with the arrows) are not modelled accurately by the linear model. The linear model is not flexible enough to fit these data points accurately. The effect of the extreme data points on the model predictions can be seen in Table 6.6.

Table 6.6 Effect of Extreme Gloss60 Values on Model (MLR) Predictions

	Average Relative Error [%]	
	Including Extreme Data Points	Excluding Extreme Data Points
Model Errors	14.50	7.23

The linear model is therefore accurate in predicting Gloss60 values as long as it includes no extreme data points. The error of the linear model fit, more or less doubles when the extreme data points are included. However, the need exists for prediction of the extreme data points and therefore another model (non-linear) is developed for prediction of these values. A neural network model is developed after other researchers have indicated [8], [9], [11] that these modelling techniques can be used accurately to predict coating properties.

6.3.2 NEURAL NETWORK MODELLING

6.3.2.1 NEURAL NETWORK ARCHITECTURE AND TRAINING

The neural network is developed to minimise those large errors in predicted values caused by the linear model. These data points, considered as outliers by the MLR model can be seen in Figure 5.18 (RCOPlot of GLOSS60 values) of Chapter 5. Figure 6.18 also illustrates the sharp increase in error of the last four model predictions.

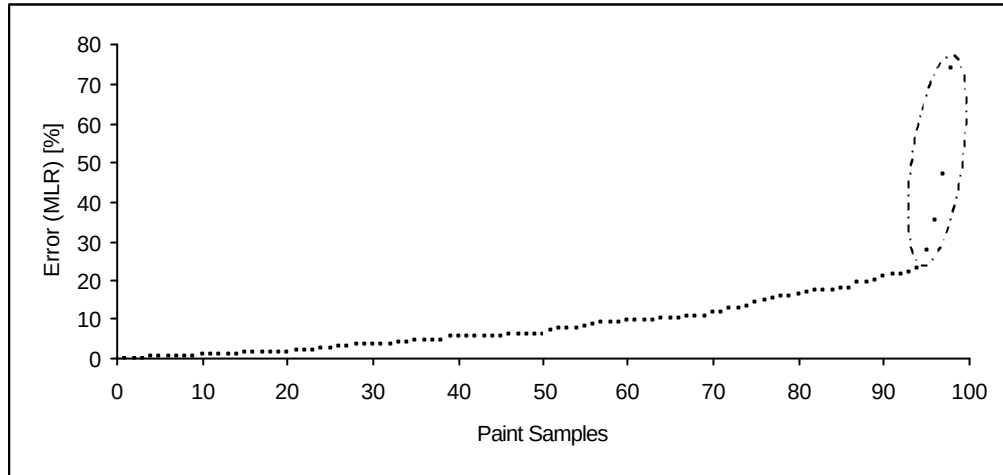


Figure 6.18 MLR Error for Gloss60 Values (In Increasing Order)

The neural network, if developed adequately, should mainly minimise the errors that are encircled in Figure 6.18.

Once again, no set of rules exists for developing the neural network architecture and initial selection of a specific network architecture (number of layers, number of nodes, etc) depends on previous experience and knowledge of the problem. To begin with, the same architecture and training algorithm is used to model the Gloss60 values, as that of what other researchers [8], [9] have used successfully in predicting certain paint properties. This is a back propagation feed forward neural network model with the following architecture (Figure 6.19)

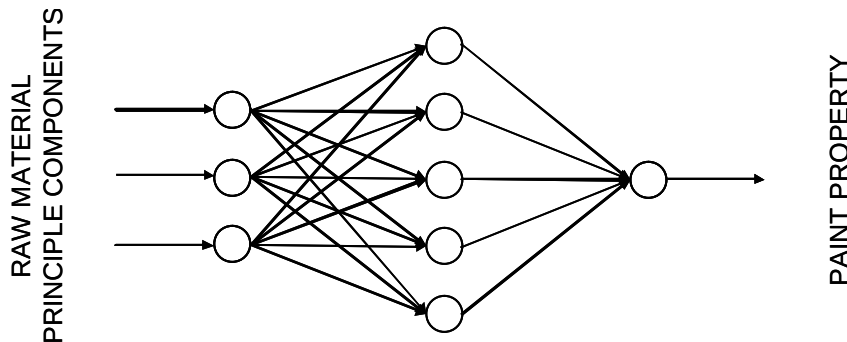
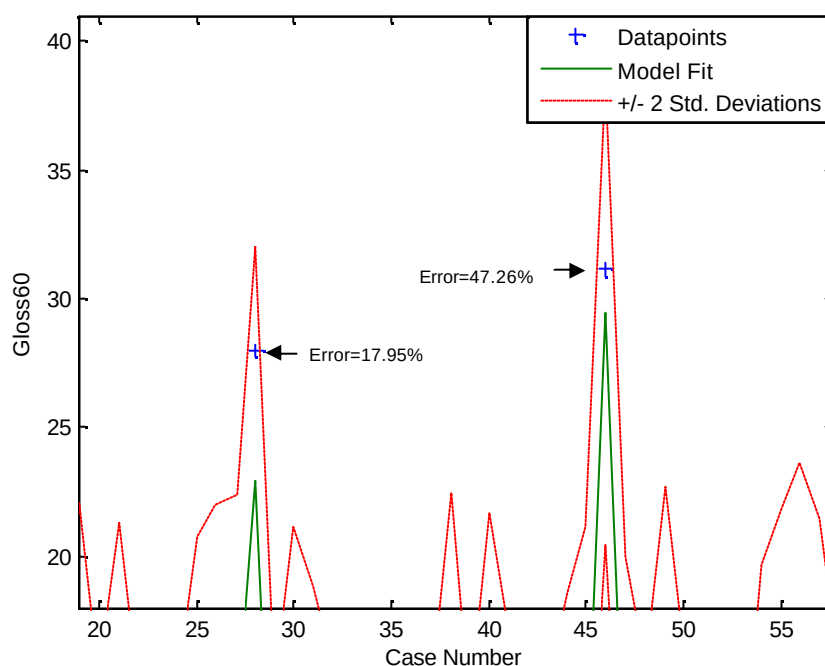


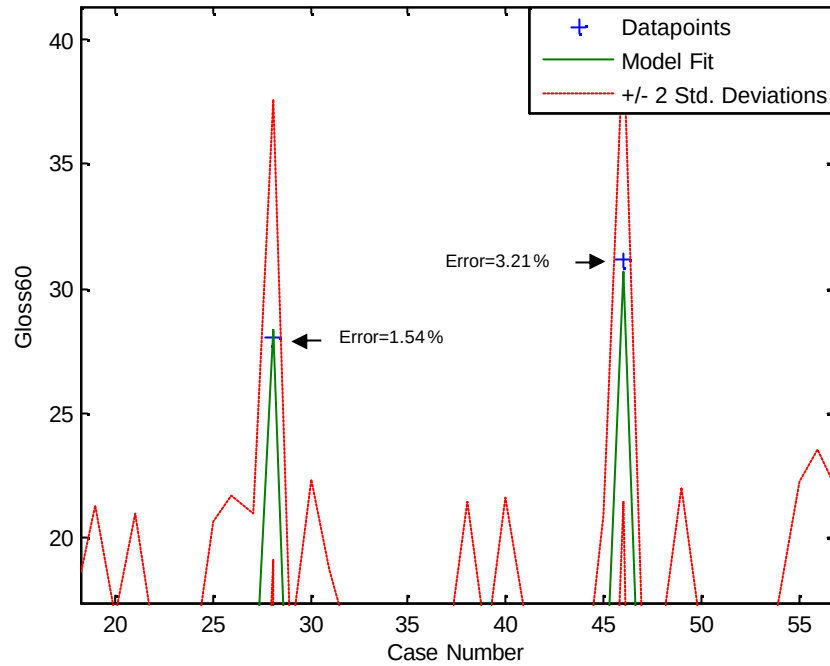
Figure 6.19 Neural Network Architecture for Gloss60

Figure 6.19 shows that a relative basic neural network architecture is used for the determination of the paint properties. The input to the net consists of the three principle components that explain 80% of the variance in the raw materials. The hidden layer consists of 5 hidden nodes each using the tan-sigmoidal activation function to send its summed output to a single output node that is used for estimation of the rheological parameter. A single output node corresponding to a single paint property will result in the most accurate prediction, but the disadvantage is once again that an optimum paint formulation is not found which generalises a set of paint properties the best. However, neural network architecture is very flexible and a set of output nodes corresponding to more than one paint property is possible but will need further network optimisation (see Chapter 7 – Future Work).

The first two data points that are indicated with arrows in Figure 6.17 are zoomed into in Figure 6.20. In Figure 6.20 (b) it can be seen that the neural net is more flexible than the MLR model in Figure 6.20 (a). (Similar model fit graphs for all the paint properties are given in Chapter 12 - Appendix E (Section 12.5))



a)



b)

Figure 6.20 Gloss60 Model (MODEL2) Fit for (a) MLR Model and (b) NN Model

In Figure 6.21 a comparison is made between the error values of the linear (MLR) model and non-linear (NN) model for the Gloss60 values. Only the last 10 data points (with increasing error values) are illustrated, because they are the points with the largest errors. These data points are not modelled accurately by the linear model and the last 6 are classified as outliers by the linear model. It is clear that the neural network (NN) predicts the Gloss60 values more accurately than the linear model (MLR), resulting in smaller error values.

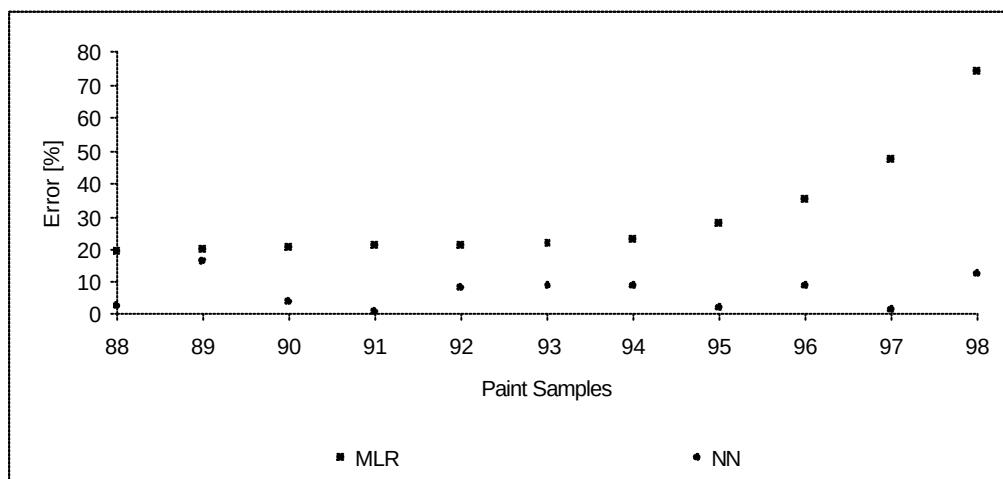


Figure 6.21 Gloss60 Errors (MODEL2) for MLR and NN Models (last 10 data points)

The neural network performs particularly better in the areas where the MLR model performs the worst (last 6 data points). In other words, the neural network generalises better than the MLR

model and is more accurate in predicting data points with extreme Gloss60 values. The MLR classifies the last 6 data points as 'outliers' with an average error of 24.85% for them while the NN reduces that average error of these data points to 4.87%, which is a clear improvement by the NN.

Chapter 5 indicated that a similar pattern is followed with all the other paint properties when MLR models are used for modelling. That is, the MLR models do not perform well in prediction of paint properties that are included in the design to act as influential data points whose values are extreme. Figure 6.22 illustrates the relative errors associated with these extreme/influential data points as a result of MLR and NN modelling for all the paint properties (PP1-PP17 of Figure 6.22 is identified in Table 6.7).

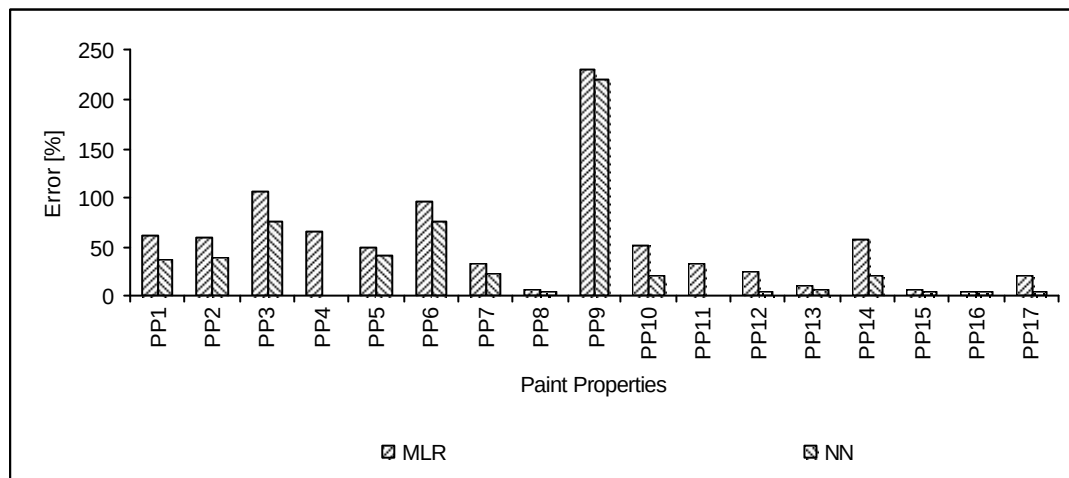


Figure 6.22 Comparison Between Outlier Errors

The neural net reduces the error of all the paint properties and is therefore better in generalising and prediction. Table 6.7 gives the relative error values for the outliers of the paint properties as modelled by the MLR and NN models.

Table 6.7 Error Values for Outliers as Modelled by MLR and NN

Paint Property	ID	Number of Outliers	Relative Error [%]	
			MLR	NN
'Dry Burnish 20'	PP1	7	63.31	30.37
'Dry Burnish 60'	PP2	6	60.28	38.14
'Dry Burnish 85'	PP3	6	107.74	76.13
'Wet Burnish 20'	PP4	1	66.53	0.61
'Wet Burnish 60'	PP5	3	50.50	42.35
'Wet Burnish 85'	PP6	3	97.74	75.71
'Water Permeability'	PP7	6	33.06	22.32
'Krebs'	PP8	5	6.62	4.83
'Sag'	PP9	8	229.08	219.17
'Open time'	PP10	4	51.66	20.93
'Gloss 20'	PP11	6	32.01	0.95
'Gloss 60'	PP12	6	24.85	4.87
'Gloss 85'	PP13	6	10.45	6.51
'Dirt Pick-Up'	PP14	5	58.67	19.55
'Opacity'	PP15	4	6.36	3.83
'Hiding Power'	PP16	5	5.04	4.33
'Dry Film Thickness'	PP17	5	20.34	3.92

The neural network makes clear improvements on the error values, however some errors are still unacceptably large. It must be remembered that the neural network architecture is not fully optimised yet and that there is still room for improvement (Chapter 7 – Future Work). However, the important thing to note is that a preliminary neural network architecture already predicts the most important paint properties more accurately.

6.3.2.2 MODEL IMPLEMENTATION

The neural networks are now implemented to predict the paint properties of 33 unseen paint samples. The error in predictions of the 33 unseen data points is given for Gloss60. In Figure 6.23 a comparison is made between the relative error in prediction resulting from multiple linear regression (MLR, ■) and neural network (NN, ●) predictions. The paint samples are placed in increasing order according to the MLR model error values.

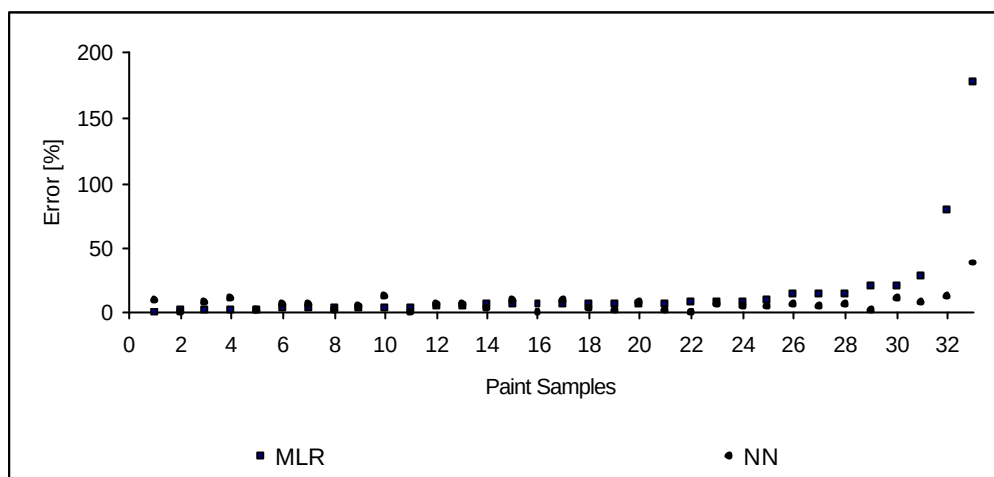


Figure 6.23 Model Errors in Predictions of Unseen Data points for Gloss60

The important part of the graph is once again there where the MLR (■) model is unable to predict the Gloss60 values accurately, resulting in large errors (end part of the graph - paint samples 22 to 33). In this part of the graph it is clear how the NN (●) predicts the Gloss60 values more accurate, resulting in smaller errors. In Figure 6.24 a comparison is made between the predicted values for Gloss60 of the 33 unseen paint samples obtained from MLR and NN modelling.

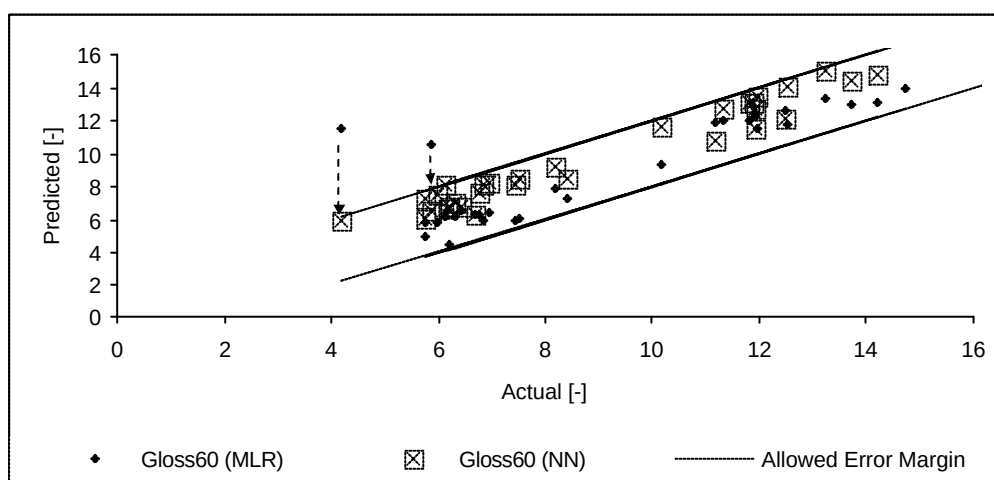


Figure 6.24 Model Predictions for Gloss60

Figure 6.24 shows what improvement the neural has on prediction of the Gloss60 values. MLR failed to predict all the values within the allowed error margin. This is especially obvious for the two outliers at the low range of the Gloss60 values. The neural net is much more accurate in prediction of these outliers. Although the neural net results in more scatter of the data outisded the allowed error margin, all the predicted data points fall within the allowed error margin. It is therefore clear that the neural net is the appropriate model to use when there are obvious outliers (e.g. extreme PVC/solids) in the data. This might result in a larger relative error in prediction, but at least there should not be any predictions that fall outside the allowed error margin.

The same trends can be observed for the other important paint properties (Krebs, Opacity and Dry Film Thickness) as can be seen in Figure 6.25 to Figure 6.27.

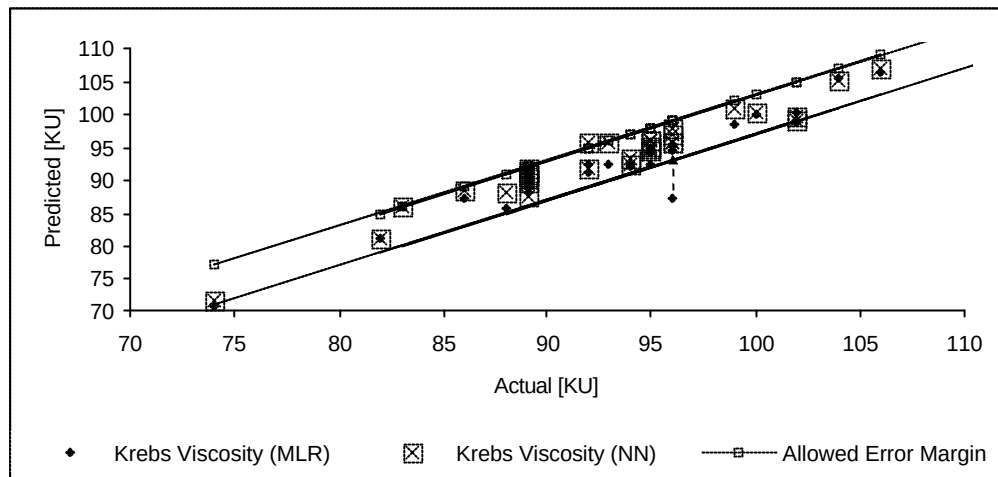


Figure 6.25 Model Predictions for Krebs Viscosity

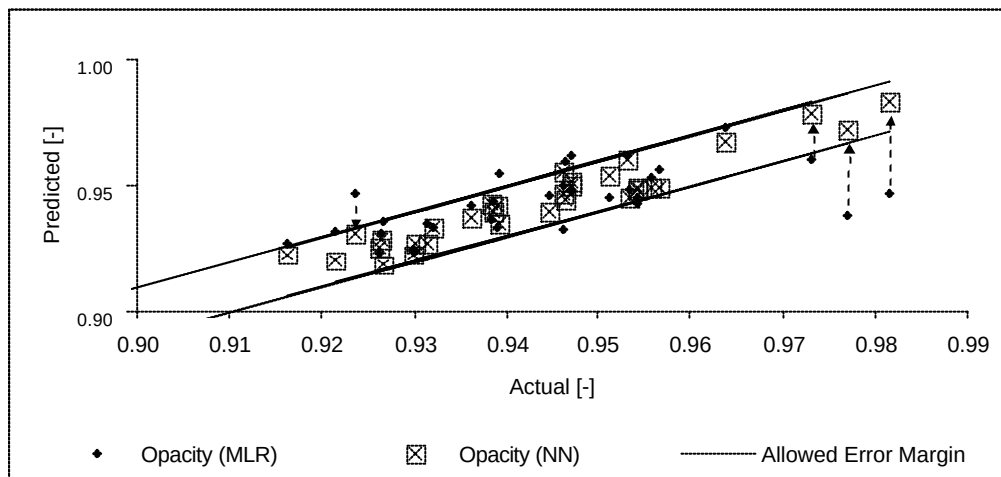


Figure 6.26 Model Predictions for Opacity

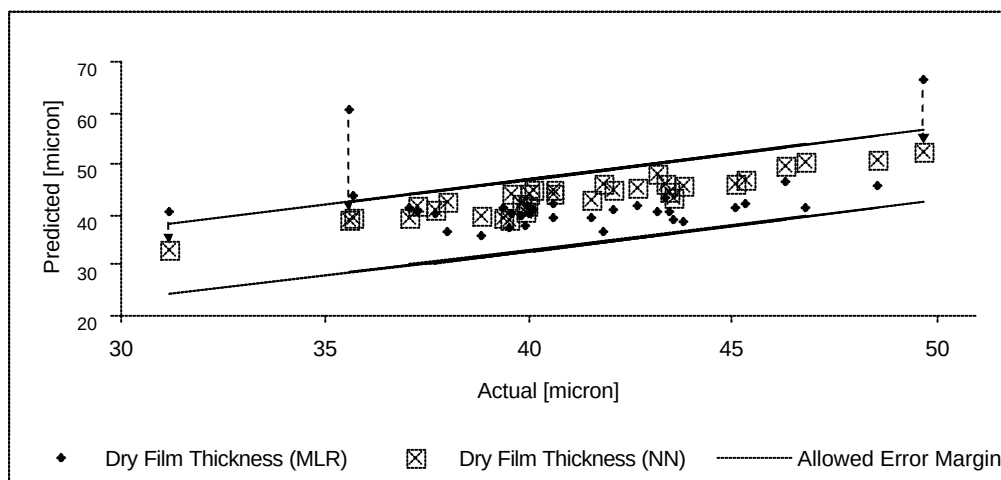


Figure 6.27 Model Predictions for Dry Film Thickness

Some of the major improvements on prediction for these paint properties has been illustrated with arrows in Figure 6.24 to Figure 6.27.

6.3.3 MODEL 2 CONCLUSIONS

As expected, non-linearity exists in the paint properties of the specific paint. In Chapter 5 it was shown that the linear models (MLR models) could not model and predict the non-linear paint properties accurately for paint samples of which the raw materials vary by more than 20% from the standard formulation. Non-linear modelling techniques such as neural networks were developed to overcome these inadequacies. The following are the most important conclusions:

- Neural networks perform better than multiple linear regression in predicting the paint properties of paint samples for which the raw materials vary by more than 20% from the standard formulation.
- Table 6.7 shows the relative errors that can be expected when neural networks are used to predict paint properties for which the raw materials vary more than 20% from the standard formulation. Although some of the predictions made by the neural networks still result in large relative errors, the neural networks show a major improvement on the predictability of the paint properties of those paint samples that vary more than 20% from the standard formulation. On average, neural networks improve predictability of the paint properties by 47%. The largest improvement in predictability is made on the wet burnish 20 prediction, where multiple linear regression resulted in relative errors of 66%, while neural networks resulted in errors of only 0.6%. Other paint property predictions where neural networks have shown a major improvement on predictability of more than 80% in accuracy, are: Gloss – 80% increase in accuracy and Dry Film Thickness – 81% increase in accuracy. The smallest increase that the neural network had on the predictability of a paint property, was a 33% increase in accuracy of the sag predictions.

6.4 MODEL 3: RELATIONSHIP BETWEEN PAINT RHEOLOGY AND PAINT PROPERTIES

6.4.1 BACKGROUND

Another way to model the paint properties, is in terms of the rheological behaviour. The importance of understanding coating performance in terms of rheology is clear. Certain paint properties, e.g. levelling, sag and spatter rely heavily on the flow and deformation behaviour (rheology) of the paint.

In the previous sections (*Section 6.2 and 6.3*), the importance of the interactions between the raw materials have been illustrated. These interactions are numerous and very complex and almost impossible to model with constitutive equations. The advantages of having a model that predicts the paint properties in terms of rheology is as follows:

- i. No explicit knowledge of the constitutive equations describing the chemical/physical interactions between the raw materials are required. The rheological behaviour of paint represents the flow and deformation behaviour of all the possible interactions.
- ii. Rheology tests are relative fast and inexpensive to do and are less operator dependable compared to the testing procedure for many of the paint properties. Therefore if good correlation can be found between the rheology and certain paint properties, many of the time/cost consuming paint property measurements can be replaced with fast and cost effective rheology measurements.

Although the rheology represents all the possible interactions between the raw materials, non-linear behaviour is still present as was illustrated by the lack of the linear model to predict certain paint properties (Chapter 5) of influential data points. The inability of the linear MLR model to predict some of the extreme Gloss60 values is illustrated with arrows in Figure 6.28.

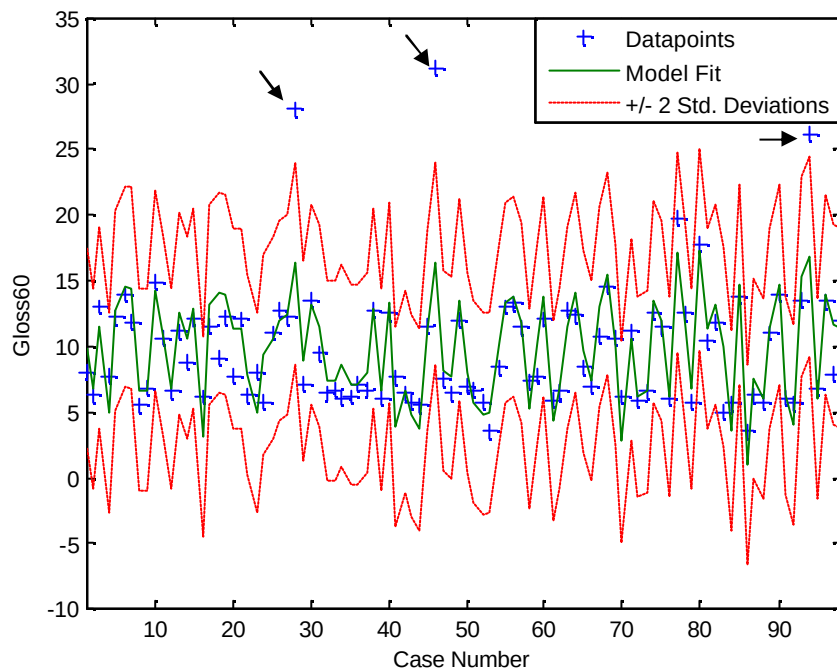


Figure 6.28 Linear Model Fit for Gloss60 Values (MODEL 3)

These three points, indicated with arrows, are the same data points that the linear model in the previous section could not predict accurately. However, in this case, the linear model is even more inaccurate in modelling these three data points and the model does not even predict the Gloss60 values within two standard deviations of the model.

Non-linear techniques are therefore developed to predict these extreme/influential data points more accurately. Neural networks are chosen as the preferred method for non-linear modelling, because of the proven success of these type of non-linear models to accurately predict the rheology and paint properties in terms of raw materials. Also, neural networks have been used

successfully in pharmaceutical research to predict the performance of creams in terms of its rheological behaviour [11].

6.4.2 NEURAL NETWORK MODELLING

6.4.2.1 NEURAL NETWORK ARCHITECTURE AND TRAINING

The neural network is therefore developed to minimise the large error values that the linear model could not predict accurately. Figure 6.29 illustrates the increase in error values as predicted by the MLR model. Also note the sharp increase in error of the last three model predictions.

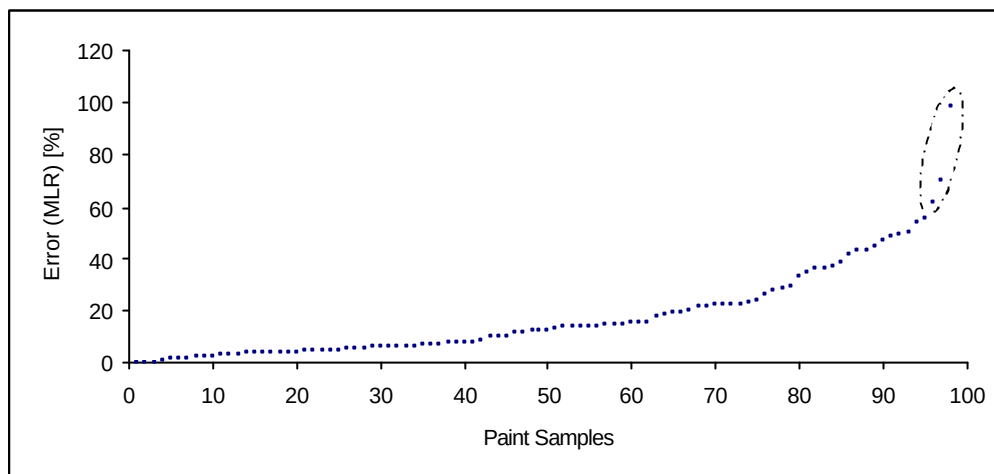


Figure 6.29 Error (MLR Model) for Gloss60 Values (MODEL 3)

A neural network is developed to predict the paint properties of the samples with large MLR errors more accurately, thereby reducing large error values at the end part of the graph. The neural network architecture is given in Figure 6.30.

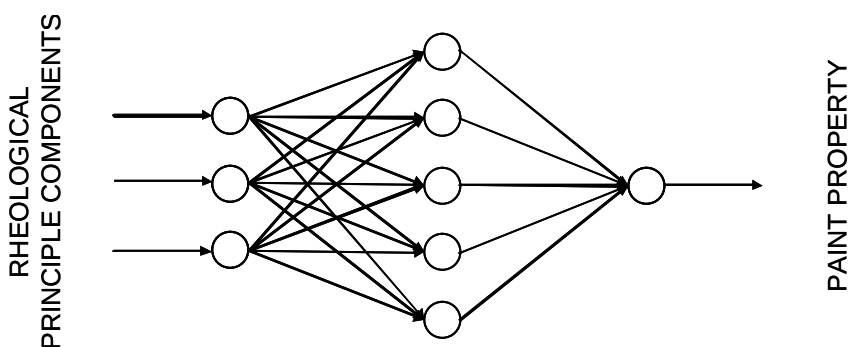
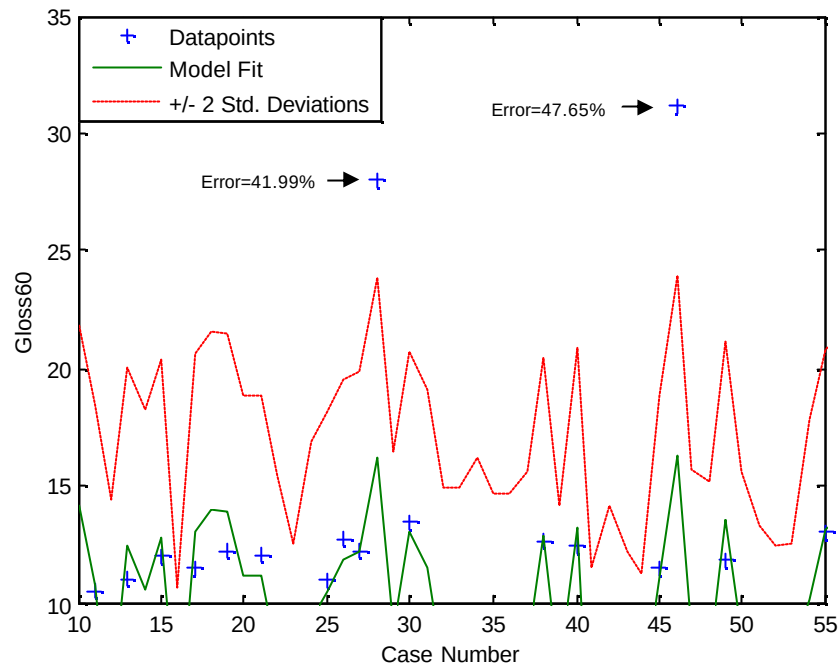


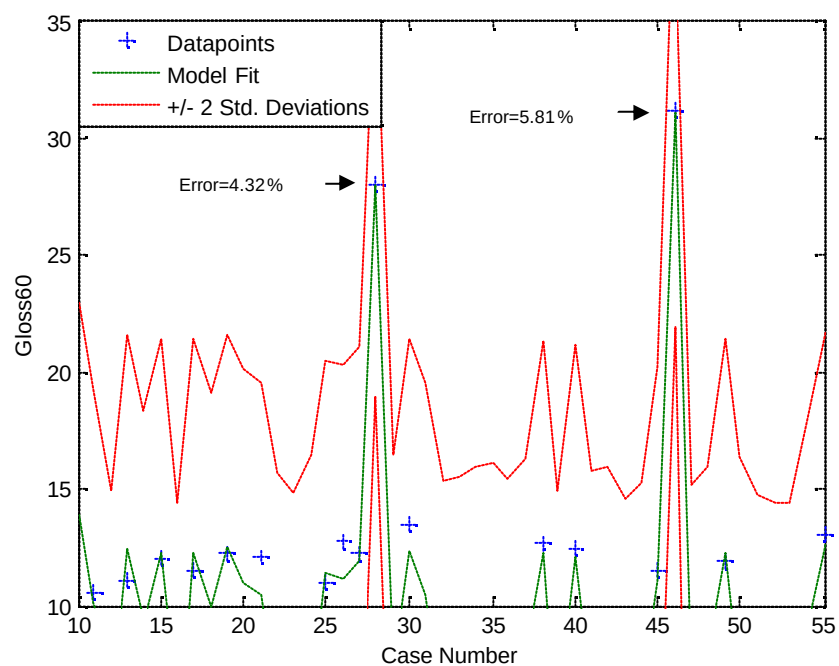
Figure 6.30 Neural Network Architecture for Gloss60 (MODEL3)

The basic neural network architecture in Figure 6.30 consists of three principle components describing 80% of the variance in rheological behaviour as input nodes, 5 nodes with tan-sigmoidal activation functions in the hidden layer and a single output node for estimation of the paint properties. It is possible to modify the neural network architecture so that various paint properties can be estimated simultaneously and therefore result in a more generalised model (see Chapter 7 - Future Work).

Although the network architecture is not fully optimised yet, it already gives improvements on the predictions compared to that of the MLR model (Figure 6.31 (a) and (b)). The flexibility of the neural network is illustrated in Figure 6.31 (b). It can clearly be seen how the neural network model fits two outlier points more accurately than the MLR model (Figure 6.31 (a)). The neural network is more flexible in predicting extreme data points and drastically reduces the error associated with these two data points.



(a)



(b)

Figure 6.31 Gloss60 Model Fit (MODEL3) for (a) MLR and (b) NN Model

The neural network therefore reduces large errors in Gloss60 prediction significantly. In Figure 6.32 a comparison is made between the errors of the last 10 data points of the multiple linear regression (MLR) model fit and the neural network (NN) model fit. It is clear that the neural network significantly reduces the errors in the part of the graph where the largest MLR errors occur.

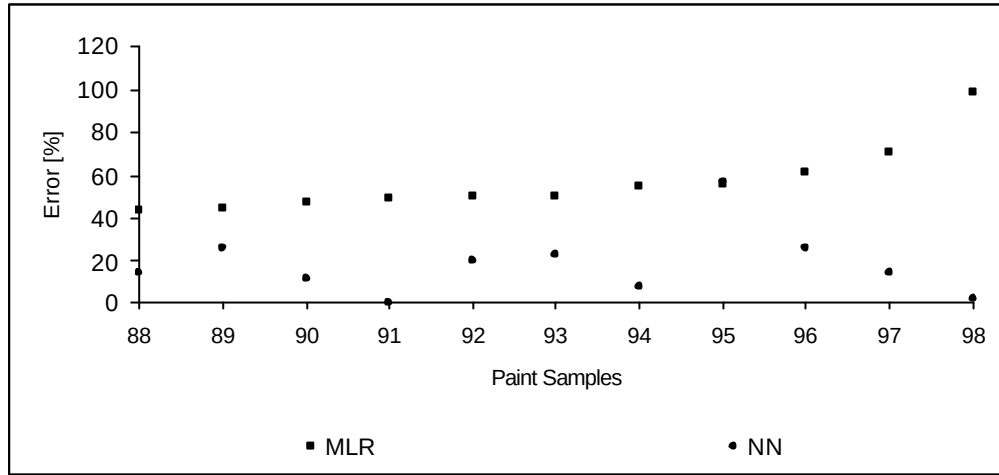


Figure 6.32 Gloss60 Errors (MODEL3) of the MLR and NN Models (last 10 data points)

The aim of the neural network is not only to reduce the large errors predicted by the MLR model (Figure 6.32), but also to predict the outliers from the MLR more accurately. Figure 6.33 illustrates the flexibility of the neural network which leads to a reduction in outlier error. The neural network results in a drop in error of the relative outlier values for all the paint properties. Even larger improvements in outlier error reduction can be expected when the neural network architecture is fully optimised (see Chapter 7 - Future Work). Values and identification of the bar graphs in Figure 6.33 are given in Table 6.8.

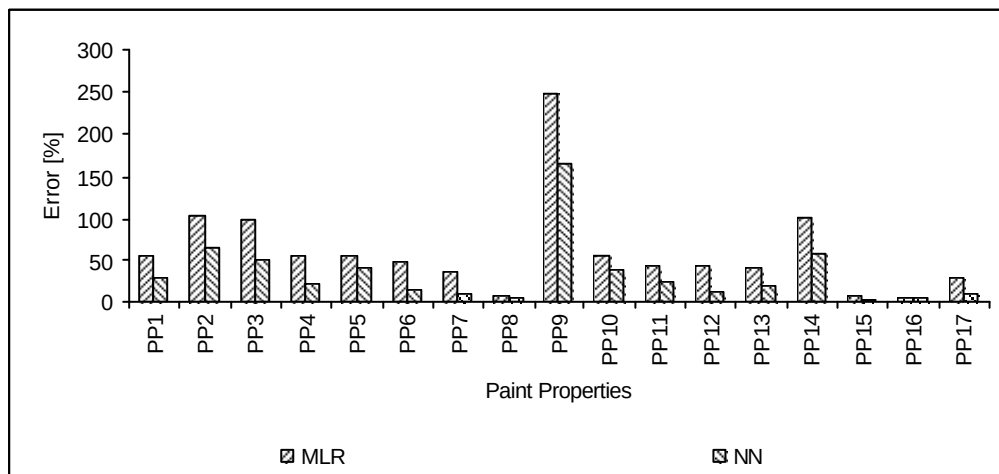


Figure 6.33 Comparison of Outlier Errors (MODEL3)

Table 6.8 Error Values for Outliers as Modelled by MLR and NN

Paint Property	ID	Number of Outliers	RELATIVE ERROR [%]	
			MLR	NN
'Dry Burnish 20'	PP1	5	53.93	27.61
'Dry Burnish 60'	PP2	7	104.04	65.81
'Dry Burnish 85'	PP3	6	98.30	48.60
'Wet Burnish 20'	PP4	2	54.33	21.30
'Wet Burnish 60'	PP5	3	55.09	40.37
'Wet Burnish 85'	PP6	4	46.88	14.75
'Water Permeability'	PP7	5	34.30	9.31
'Krebs'	PP8	6	6.41	4.00
'Sag'	PP9	8	247.17	165.66
'Open time'	PP10	3	57.11	37.67
'Gloss 20'	PP11	3	43.45	22.55
'Gloss 60'	PP12	3	41.79	12.74
'Gloss 85'	PP13	8	39.01	19.45
'Dirt Pick-Up'	PP14	4	100.83	59.30
'Opacity'	PP15	6	6.60	2.37
'Hiding Power'	PP16	7	5.45	4.34
'Dry Film Thickness'	PP17	2	27.58	8.72

The error values for sag are unexpectedly high. The prescribed ASTM technique only allows for a specified range of measurement values. Therefore, sag measurements above the maximum of this range have been rounded to the ceiling value.

Other error values in Table 6.8 are surprisingly low, e.g. that of opacity which one does not directly relate to the rheology of the paint. This is most likely because the opacity of this specific paint is already high so that a change in the raw materials results in a relative small experimental range for the opacity.

6.4.2.2 MODEL IMPLEMENTATION

The neural networks are now implemented to predict the paint properties of 33 unseen paint samples. The Gloss60 model predictions for these paint samples are shown in Figure 6.34.

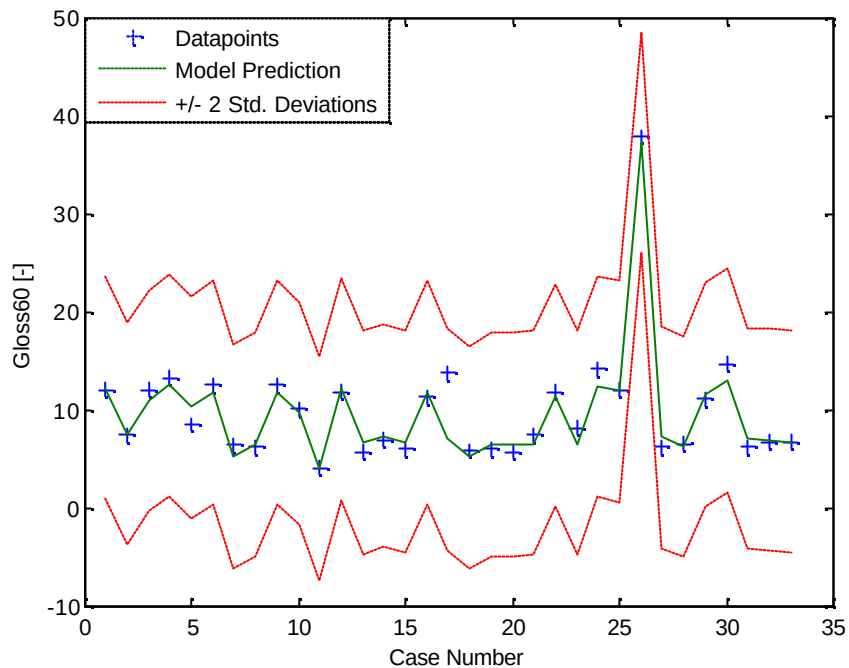


Figure 6.34 Gloss60 Model Predictions for Unseen Paint Samples

What is clear from Figure 6.34 is how accurate the neural net is in prediction of the extreme data point. The neural net reduces the error of the extreme data points significantly. Figure 6.35 shows the error in predictions of the 33 unseen data points for their Gloss60 values. A comparison is made between multiple linear regression (MLR, ■) and neural network (NN, ●) predictions. (The paint samples are placed in increasing order according to the MLR model error values and do not follow the same order as in Figure 6.34).

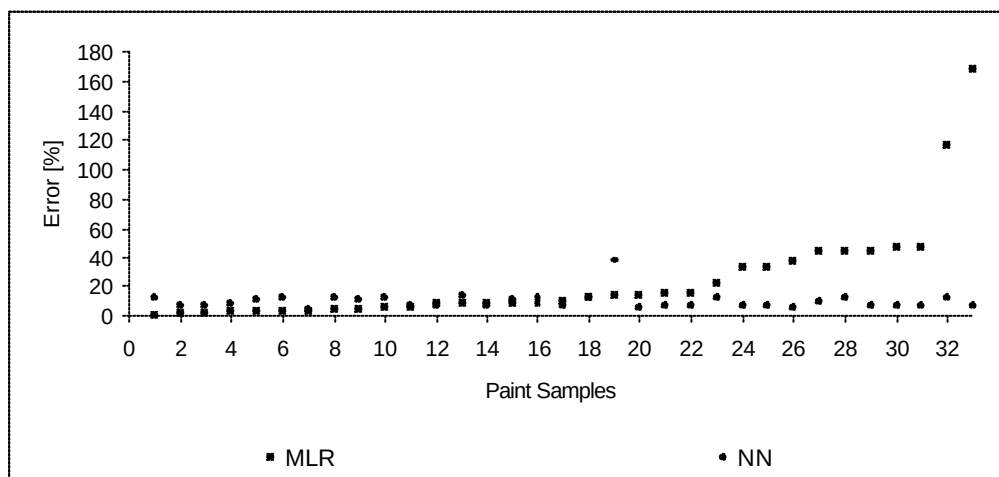


Figure 6.35 Model Predictions in Errors of Unseen Data points for Gloss60 values

As expected, after a more accurate model fit for the neural network than the MLR model, the neural network is more accurate in prediction of the unseen data points. This is especially the case where the MLR model predicts large errors (end part of the graph).

In Figure 6.36 the arrows give an indication of the improvement the neural net makes on prediction of Gloss60 values. The neural net is able to predict all but one of the paint samples within the allowed error margin – a clear improvement on the MLR model.

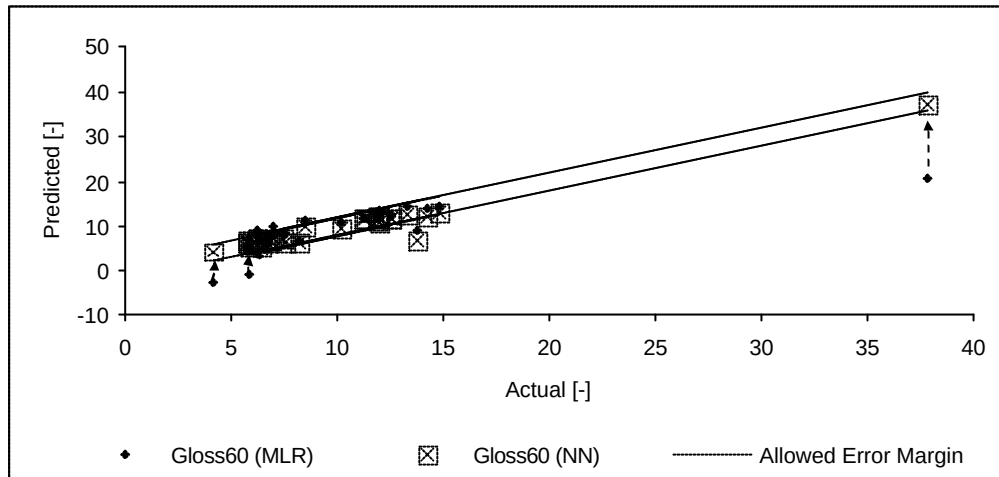


Figure 6.36 Model Predictions for Gloss60

Figure 6.37 shows that the neural is even more accurate than the multiple linear regression model in predicting the Krebs viscosity of the paint samples, even though the MLR model already predicts the Krebs Viscosity of the paint samples quite accurately. Once again, the neural net predicts the Krebs Viscosity of only one sample outside the allowed error margin.

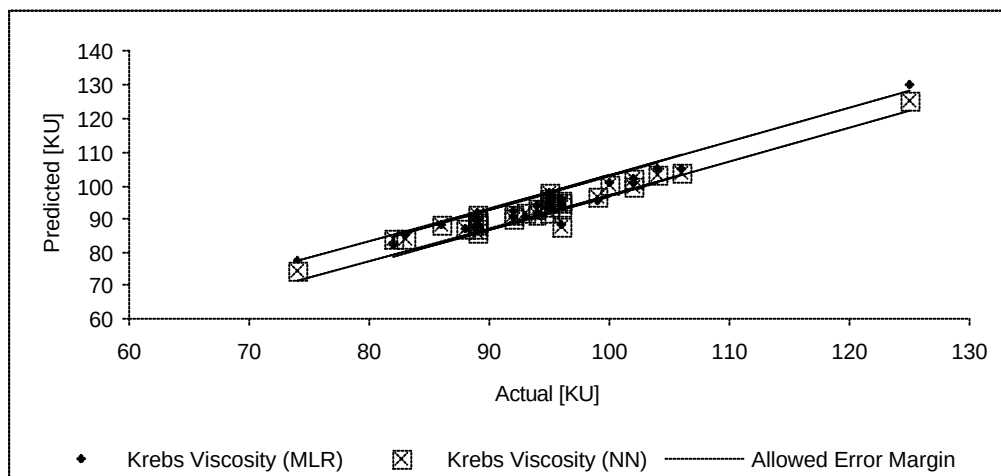


Figure 6.37 Model Predictions for Krebs Viscosity

In Figure 6.38 the neural net illustrates that it is most efficient in reducing the error between the actual and the predicted Opacity values of the extreme opacity values (indicated with an arrow). It is clear that in this case the neural net should only be used when paint samples are manufactured with extreme levels of raw materials that might result in extreme Opacity values.

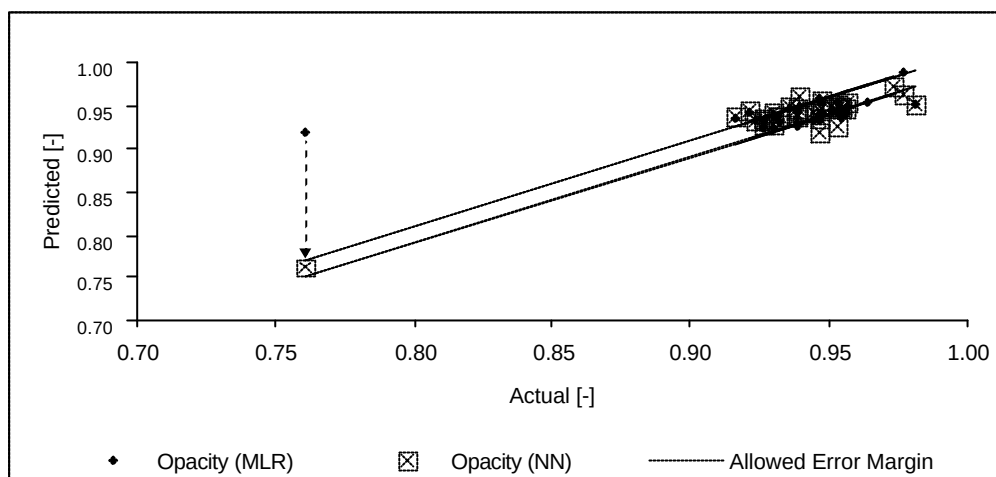


Figure 6.38 Model Predictions for Opacity

The neural net is also more accurate than the multiple linear regression model in predicting Dry Film Thickness, with fewer predictions falling outside the allowed error margin compared to the MLR model.

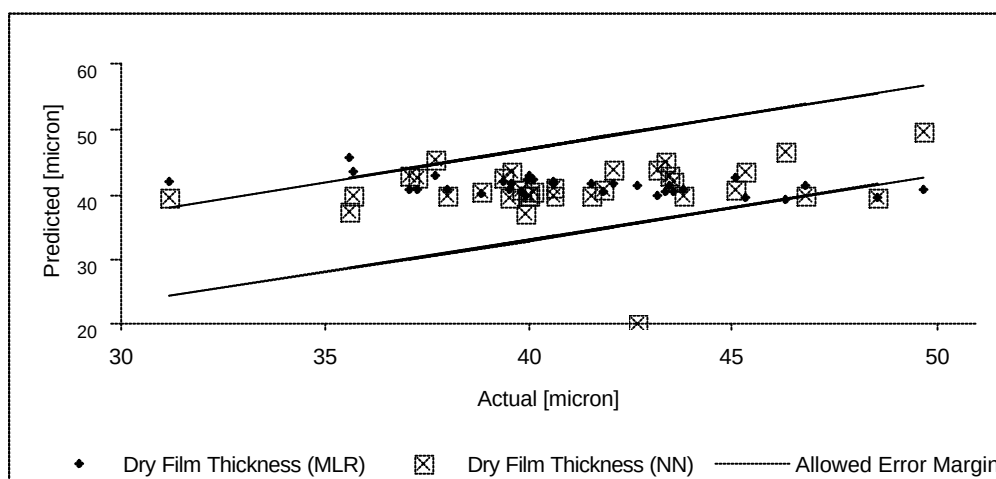


Figure 6.39 Model Predictions for Dry Film Thickness

6.4.3 MODEL 3 CONCLUSIONS

As expected, non-linearity also exists in the paint properties of the specific paint. In Chapter 5 it was shown that the linear models (MLR models) could not model and predict the non-linear paint properties accurately for paint samples of which the raw materials vary by more than 20% from the standard formulation. Non-linear modelling techniques such as neural networks were developed to overcome these inadequacies. The following are the most important conclusions:

- Neural networks perform better than multiple linear regression in predicting the paint properties of paint samples for which the raw materials vary by more than 20% from the standard formulation.

- Table 6.8 shows the relative errors that can be expected when neural networks are used to predict paint properties for which the raw materials vary more than 20% from the standard formulation. On average, neural networks that use rheology data as input, predict the paint properties for those paint samples with raw materials that vary more than 20% from the standard formulation, 49% more accurate than equivalent multiple linear regression models. The greatest improvement in model predictability is for water permeability (73% increase in accuracy) and gloss (70% increase in accuracy).
- Two models have been developed that predicts the paint properties of those samples that vary more than 20% from the standard formulation. The first model makes use of raw materials as input data (MODEL 2) and the second model makes use of rheological data as input to the model (MODEL 3). Table 6.9 show a comparison of the two models. In certain cases MODEL 3 is more accurate in predicting certain paint properties such as dry/wet burnish, water permeability, Krebs viscosity, sag, dry film thickness and open time. In other cases, MODEL 2 performs better in predicting certain paint properties such as open time, gloss and dirt pick-up. Table 6.9 can be used as reference to make a decision between which model to use for the most accurate prediction of paint properties for those paint samples that vary more than 20% from the standard formulation.

Table 6.9 Comparison Between MODEL 2 and MODEL 3 Paint Property Predictions

Paint Property	MODEL 2 Error (NN) [%]	MODEL 3 Error (NN) [%]
'Dry Burnish 20'	30.37	27.61
'Dry Burnish 60'	38.14	65.81
'Dry Burnish 85'	76.13	48.60
'Wet Burnish 20'	0.61	21.30
'Wet Burnish 60'	42.35	40.37
'Wet Burnish 85'	75.71	14.75
'Water Permeability'	22.32	9.31
'Krebs'	4.83	4.00
'Sag'	219.17	165.66
'Open time'	20.93	37.67
'Gloss 20'	0.95	22.55
'Gloss 60'	4.87	12.74
'Gloss 85'	6.51	19.45
'Dirt Pick-Up'	19.55	59.30
'Opacity'	3.83	2.37
'Hiding Power'	4.33	4.34
'Dry Film Thickness'	3.92	8.72

6.5 MODEL INTERPRETATIONS*****

Non-linear models have been developed to predict paint properties (rheological and mechanical) for paint samples with extreme variations in some of the raw materials. The predictions of the neural networks for a paint with a 25% increase in emulsion is illustrated in Table 6.10.

Table 6.10 Neural Network Model Predictions of Change in Standard Paint Formulation

Paint Property	Standard Formulation	+25% Emulsion	Model
Gloss 60 [-]	7.73	29.4	2
Opacity [%]	98.33	90.21	2
Krebs Viscosity [KU]	89	115	2
Dry Film Thickness [μm]	37	40	2
Frequency Sweep	See Figure 6.40		1

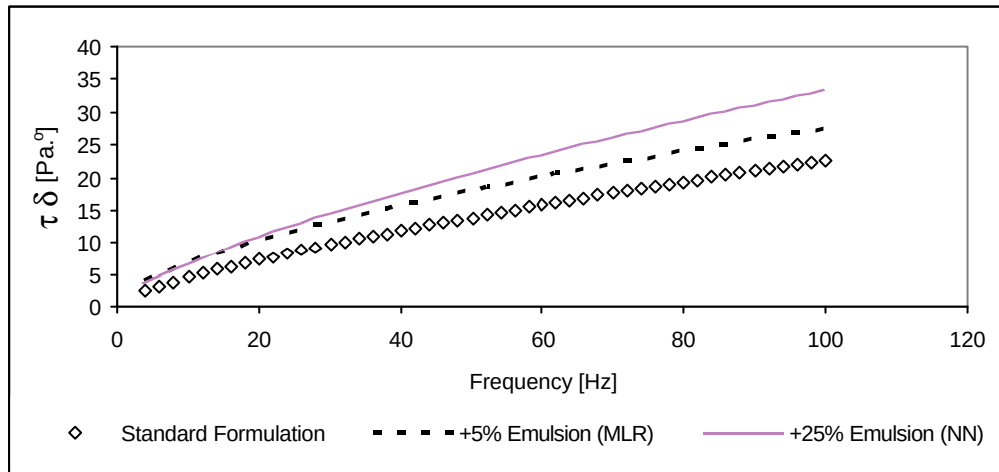


Figure 6.40 Frequency Sweep of Standard Paint Formulation, MLR Model Prediction for 5% Increase in Emulsion and NN Model Prediction for 25% Increase in Emulsion

All of the neural network predictions in Table 6.10, except for the Dry Film Thickness, correlate with what can be expected for a large increase in emulsion. The increase in emulsion results in the following predictions:

- a large increase in gloss
- a large decrease in opacity
- a large increase in Krebs Viscosity
- a small increase in dry film thickness

***** Model interpretations are not considered to be a fundamental part of this study. It is rather an indication that model predictions from the study can be used for interpretation of future phenomenological models.

- v) an increase in the response ($\tau \delta$) of the frequency sweep

All of the predictions above, except dry film thickness can be explained by following the basic principles of paint chemistry as discussed in Chapter 5, Section 5.5. The dry film thickness is expected to be substantially more than what the neural network predicts and the small increase in the predicted value of the dry film thickness can not be explained. The prediction could possibly be a model error.

Only MODEL 1 and MODEL 2 predictions are included in the interpretation because actual rheology measurements are required for MODEL 3 predictions.

6.6 MODEL DISCRIMINATION

Akaike's information criterion (AIC) developed by Hirotugu Akaike in 1971 [13] is a measure of goodness of fit of an estimated statistical model and is given as:

$$AIC = 2k - 2\ln(L) \quad (6.4)$$

where k is the number of parameters in the statistical model and L is the likelihood function. AIC can also be written as follows [14]:

$$AIC = 2k + n \left[\ln \left(\frac{2pRSS}{n} \right) + 1 \right] \quad (6.5)$$

where n is the number of observations and RSS is the residual sum of squares. AIC takes the model complexity into account when the goodness of fit is evaluated. A lower value for AIC indicates better goodness of fit at the specific model complexity.

The number of parameters k for MLR are taken as the number of input variables of the regression model – see equation 5.4 (MODEL 1 $\rightarrow k=11$), equation 5.6 (MODEL 2 $\rightarrow k=11$) and equation 5.7 (MODEL 3 $\rightarrow k=3$). The number of parameters k for the NN's are taken as the number of weights between neurons in an entire network – see Figure 6.11 (MODEL 1 $\rightarrow k=12$), Figure 6.19 (MODEL 2 $\rightarrow k=20$) and Figure 6.30 $\rightarrow k=20$). Therefore an increase in the complexity of the NN (number of connections between neurons), can drastically increase the value of the AIC if the goodness of fit does not increase significantly.

The values for AIC of the MLR and NN of MODEL 1 are given in Table 6.11.

Table 6.11: Akaike Information Criterion for Model 1

Rheology Curve	Rheology Parameter	AIC	
		MLR	NN
3-ITT (Oscillatory)	$b_{3-ITT (Osc)}$	505	526
	$a_{3-ITT (Osc)}$	120	122
3-ITT (Rotational)	$b_{3-ITT (Rot)}$	139	147
	$a_{3-ITT (Rot)}$	119	122
Amplitude Sweep	b_{AS}	527	608
	a_{AS}	122	124
Frequency Sweep	b_{FS}	179	175
	a_{FS}	121	123
Flow Curve	b_{FC}	217	217
	a_{FC}	110	112
High Shear Curve	a_{HS}	112	114
	b_{HS}	121	123
	g_{HS}	910	1051
Low Shear Curve	b_{LS}	192	210
	a_{LS}	130	132
Time Sweep Curve	b_{TS}	125	127
	a_{TS}	533	547
Extra Low Frequency	C_{XLF}	135	129

Table 6.11 indicates that, in general, the added complexity of the neural networks does not result in a relative increase in goodness of fit. It is only the b_{FS} and C_{XLF} parameters for which the goodness of fit, relative to its model complexity, increases, or $AIC_{NN} < AIC_{MLR}$.

The values for AIC of the MLR and NN of MODEL 2 and MODEL 3 are given in Table 6.12.

Table 6.12: Akaike Information Criterion for Models 2 and 3

Paint Property	AIC			
	MODEL 2		MODEL 3	
	MLR	NN	MLR	NN
Dry Burnish 20	123	140	108	141
Dry Burnish 60	139	149	131	152
Dry Burnish 85	125	141	115	145
Wet Burnish 20	119	136	105	139
Wet Burnish 60	128	144	113	144
Wet Burnish 85	125	141	111	142
Water Permeability	581	570	582	554
Krebs	462	530	444	432
Sag	468	484	456	450
Open time	695	693	693	687
Gloss 20	145	140	174	180
Gloss 60	355	248	481	392
Gloss 85	429	456	623	574
Dirt Pick-Up	123	141	105	139
Opacity	118	136	101	135
Hiding Power	120	138	104	138
Dry Film Thickness	541	470	597	496

Table 6.12 indicates that, in general, the MLR models of MODEL 2 have lower values for the AIC (except for Water Permeability, Open Time, Gloss20/60 and Dry Film Thickness). Therefore, according to AIC the increase in model complexity of the neural networks relative to the complexity of the MLR models, does not contribute in increasing the goodness of fit (except for the abovementioned paint poroperties for which $AIC_{NN} < AIC_{MLR}$).

Table 6.12 also indicates that the increase in model complexity of the neural networks of MODEL 3, only results in 7 models (out of 17) for which the goodness of fit, relative to its model complexity, increases. The paint properties for which $AIC_{NN} < AIC_{MLR}$ are Water Permeability, Krebs Viscosity, Sag, Open Time, Gloss60/85, Dry Film Thickness.

6.7 GENERAL CONLCLUSIONS

Multiple linear regression models are accurate in predicting the rheology and paint properties of paint samples for which the raw materials do not vary more than 20% from the standard formulation. However, it has been shown that other modelling techniques were required for those paint samples that do vary more than 20% from the standard formulation. Three neural network

models were developed for the purpose of predicting the rheology and the paint properties of these paint samples more accurately.

- 1) The neural network of MODEL 1 was developed to predict the *rheology* of those paint samples that vary more than 20% from the standard formulation, by using *raw materials* as input to the model. It has been shown that neural networks perform better than multiple linear regression in predicting the rheological behaviour of paint samples for which the raw materials vary by more than 20% from the standard formulation. On average, neural networks improve predictability of the rheological parameters by 54%. The largest improvement in predictability is made on the rheological variable C_{XLF} , where multiple linear regression resulted in relative errors of 59%, while neural networks resulted in errors of only 5%. Other predictions of rheology curves where neural networks have shown a major improvement on predictability are the Time Sweep Curve (TS) – 68% increase in accuracy and Low Shear Curve (LS) – 63% increase in accuracy. The smallest increase that the neural network had on the predictability of a rheology curve, was a 33% increase in accuracy of the Amplitude Sweep (AS) predictions.

The Akaike information criterion (AIC) shows that multiple linear regression models are the preferred models (in terms of goodness of fit versus model complexity) when predicting all of the rheological parameters, except b_{FS} and C_{XLF} .

- 2) The neural network of MODEL 2 was developed to predict the *paint properties* of those paint samples that vary more than 20% from the standard formulation, by using *raw materials* as input to the model. The neural networks show a major improvement on the predictability of the paint properties for those paint samples that vary more than 20% from the standard formulation. On average, neural networks improve predictability of the paint properties by 47%. The largest improvement in predictability is made on the wet burnish₂₀ prediction, where multiple linear regression resulted in relative errors of 66%, while neural networks resulted in errors of only 0.6%. Other paint property predictions where neural networks have shown a major improvement on predictability of more than 80% in accuracy, are Gloss – 80% increase in accuracy and Dry Film Thickness – 81% increase in accuracy. The smallest increase that the neural network had on the predictability of a paint property, was a 33% increase in accuracy of the sag predictions.

The Akaike information criterion (AIC) shows that neural networks are the preferred models when predicting Water Permeability, Open Time, Gloss_{20/60} and Dry Film Thickness.

- 3) A neural network of MODEL 3 was developed to predict the *paint properties* of those paint samples that vary more than 20% from the standard formulation, by using *rheology data* as input to the model. The neural networks perform better than multiple linear regression in predicting the paint properties of paint samples for which the raw materials vary by more than 20% from the standard formulation. On average, neural networks that use rheology data as input, predict the paint properties 49% more accurate than equivalent multiple linear

regression models. The greatest improvement in model predictability is for water permeability - 73% increase in accuracy and gloss - 70% increase in accuracy.

The Akaike information criterion (AIC) shows that neural networks are the preferred models when predicting Water Permeability, Krebs Viscosity, Sag, Open Time, Gloss60/85 and Dry Film Thickness.

- 4) Two models have been developed that predicts the paint properties of those samples for which the raw material composition varies more than 20% from the standard formulation. The first model makes use of raw materials as input data (MODEL 2) and the second model makes use of rheology data as input to the model (MODEL 3). In terms of the relative error between the actual and the predicted value of a paint property, MODEL 3 is more accurate in predicting paint properties such as dry/wet burnish, water permeability, Krebs viscosity, sag, dry film thickness and open time. For paint properties such as open time, gloss and dirt pick-up, MODEL 2 performs better.
- 5) Although it is clear that neural networks result in significantly more accurate predictions for the rheological behaviour of those paint samples that vary more than 20% from the standard formulation, some of the relative errors are still large, e.g. 312% for $b_{3-ITT(Osc)}$ and 270% for b_{FS} (MODEL 1); 76% for dry burnish85 and 75% for wet burnish85 (MODEL 2); 65% for dry burnish60 and 59% for dirt pick-up (MODEL 3). Optimisation in terms of the neural net architecture is not complete and further work in terms of network optimisation (e.g. training algorithm, activation function, number of hidden layers, etc.) is required that will aim at minimising these errors.

6.8 NOMENCLATURE

Abbreviations

3-ITT	Three-interval-thixotropy test
3-ITT (osc)	Three-interval-thixotropy test in oscillatory mode
3-ITT (rot)	Three-interval-thixotropy test in rotational mode
ANTF	Antifoamer
AS	Amplitude sweep
CLNT	Coalescent
CNST	Constant
COMB	Combination
C _{XL} F	Constant of extra low frequency
DISP	Dispersant
EMLS	Emulsion
FC	Flow curve
FNGD	Fungicide

FS	Frequency sweep
HS	High shear
H_0	Null hypothesis
LS	Low shear
MLR	Multiple linear regression
OPIG	Organic pigment
PC	Principal component
PIGM	Pigment
PVC	Pigment volume concentration
RM1	Rheology modifier 1
RM2	Rheology modifier 2
RMC	Raw material contribution
SA	Soda ash
SOLV	Solvent
SURF	Surfactant
TS	Time sweep
VS	Volume solids
WTR	Water
XLF	Extra low frequency
XTND	Extender

Symbol	Description
C	Condition number (-)
G'	Storage modulus (Pa)
G''	Loss modulus (Pa)
G^*	Complex modulus (Pa)
L	Maximum likelihood
t	Time (s)
w	Neural network weight
x	Independent variable
x_{avg}	Average value of variable
x_i	Neural network input
$x_{standard deviation}$	Standard deviation of variable
$x_{standardised}$	Standardised value of variable
y	Dependent variable
Z	Sigmoidal activation function

Greek Symbol	Description
a	Preliminary model parameter (gradient)
a_i	Gradient of preliminary model where i=HS, LS, 3-ITT _(rot) , etc
$a_{oscillatory}$	Gradient of oscillatory measurement
$a_{Gloss60}$	Predicted Gloss60 value
$a_{PaintProperty}$	Value of predicted paint property
$a_{rotational}$	Gradient of rotational measurement
b	Preliminary model parameter (offset)
b_i	Preliminary model parameter (offset) where i=HS, LS, 3-ITT _(rot) , etc
$b_{oscillatory}$	Offset of oscillatory measurement
$b_{emulsion}$	Regression coefficient associated with the emulsion
$b_{rotational}$	Offset of rotational measurement
b_0	Constant associated with regression analysis
b_k	Regression coefficient associated with variable $k=1,2,3,\dots$
b_{PC1}	Regression coefficient associated with principal component 1
g	Deformation [%] or strain [-]
g_A	Deformation [%] or strain [-] amplitude
$\dot{g}$	Shear rate [1/s]
d	Phase shift, loss angle [°]
e	Random error
h	Shear viscosity [Pa.s]
h^*	Complex viscosity [Pa.s]
t	Shear stress [Pa]
t_A	Shear stress amplitude [Pa]
w	Angular frequency [rad/s]

6.9 REFERENCES

1. Leskovsek, N., Tusar, M., Tusar, L. (1995). "Optimisation of rheological and mechanical properties of paint using methods of experimental design and neural network modelling (Polish)." p.258-259. Medvode, Ljubljana.
2. Leskovsek, N., Tusar, M., Tusar, L.; Zupan, J. (1994). "Surface coating formulation using neural networks and statistical methods (Polish)." **28**(1). Kemijski Institut, University of Ljubljana, Slovenia.
3. Gasperlin, M. (1995). "Empirical modelling of rheological and mechanical properties of paint." Rheology: 140.
4. Leskovsek, N., Tusar, M., Tusar, L. "Study of rheological properties of paints with designing of experiments: a neural network approach." p. 709. Supplied by The British Library.
5. Gasperlin, M. Tusar, L., Tusar, M., Smid-Korbar, J., Zupan, J.; Kristl, J. (2000). "Viscosity prediction of lipophilic semisolid emulsion systems by neural network modelling." International Journal of Pharmaceutics **196**: 37-50.
6. Gasperlin, M., Tusar, L., Smid-Korbar, J., Kristl, J. "Viscous behaviour of pharmaceutical creams: characterisation by neural network and polynomial modelling." p. 14-20. University of Ljubljana, Faculty of Pharmacy, Slovenia.
7. Gasperlin, M. Tusar, L. Tusar, M. Smid-Korbar, J. Kristl, J. (1998). "Lipophilic semisolid emulsion systems: viscoelastic behaviour and prediction of physical stability by neural network modelling." International Journal of Pharmaceutics **168**: 243-254.
8. Tusar, L., Tusar, M., Leskovsek, N. (1995). "A comparative study of polynomial and neural network modelling for the optimisation of clear coat formulations." Surface Coatings International **10**: 427-434.
9. Tusar, L. Tusar, M., Barle, N., Leskovsek, N., Kunaver, M. (1995). "A study of the influence on hiding power of the composition of a paint and its film thickness." Surface Coatings International **11**: 473-481.
10. Aldrich, C. (1997). Neural Networks for the Process Industries. Unpublished Manuscript, Stellenbosch.
11. Hussain, A. S., Yu, X., Johnson, R. D. (1991). "Application of neural computing in pharmaceutical product development." Pharmaceutical Research **8**(10): 1248-1252.
12. Moolman, P. L. (2003). Rheology of Coating Systems. Department Chemical Engineering, University of Stellenbosch.
13. Akaike, H (1974). "A new look at the statistical model identification." IEEE Transactions on Automatic Control **19**(6): 716-723.
14. http://en.wikipedia.org/wiki/Akaike_information_criterion

7 CONCLUSIONS AND FUTURE WORK

Understanding the effect that raw materials have on paint properties is the ultimate goal of the paint chemist. This will enable the him to develop new paint technology that is ultimately less expensive, less time consuming and less labour intensive; while still retaining those fundamental characteristics of the paint that determines its quality. Empirical models that accurately predict paint properties are useful tools in obtaining this goal. The work presented in this study is an attempt in obtaining accurate paint property predictions, based on empirical models.

Rheology plays a fundamental role in explaining certain paint properties. The relationship between certain paint properties and rheology is clear, but there is also underlying relationships between certain paint properties and the rheology that is not so clear. Advances in rheological measurement techniques, combined with powerful modelling techniques, make more accurate paint property predictions a possibility. Rheological measurement in the paint industry has evolved from a single viscosity measurement (1930's) to measurements of the viscous (1970's) and elastic properties (1990's) of paint. Combining these advanced rheological measurement techniques with powerful modelling techniques have resulted in novel contributions in this field of study

7.1 CONTRIBUTIONS

The novel and most important contributions resulting from this study are as follows:

- Historically it was easy to characterise the flow behaviour of paint with a single viscosity measurement and therefore this single rheological measurement has been the subject of numerous rheological modelling. With the advent of new rheological measurement techniques other modelling possibilities also surfaced and scientists began to model paint properties in terms of the viscoelastic behaviour. But paint properties were still being modelled by a single point from a single rheological measurement. This study not only uses a single rheological measurement to make a prediction on a certain paint property, but combines a set of nine rheological measurements by means of preliminary modelling techniques to make those predictions. This makes more sense because paint is time and shear dependent. Certain paint properties (e.g. levelling) develop during long time scale (low shear) conditions while other (e.g. spatter) during short time scale (high shear) conditions (spatter). Preliminary modelling has made it possible to capture the rheological response of the paint over the widest range of time scale and shear conditions possible. Harmful effects that are caused by redundant data

have been minimised by making use of principal component analysis for obtaining the principal components in the rheological data. The principal components were then used for paint property predictions. **It has been shown that these preliminary modelling techniques are useful and contributes to accurate paint property predictions.** A major drawback of principal component analysis is that the principal components lack simple interpretation since each is, in a sense, a mixture of the original variables and although these preliminary modelling techniques make model predictions more accurate, it makes the task of developing a phenomenological model more challenging.

- **A model has been developed for the prediction of the rheological behaviour of a specific paint (MODEL1).** Quantitative prediction of the rheological response of new paint formulations are accurate, resulting in small errors between actual and predicted rheological responses. The predicted rheological parameters enables the paint chemist to accurately construct rheology curves for new paint formulations, based only on raw material composition. The model consists out of two parts:
 - i) A linear model that enables the paint chemist to predict the linear correlations very accurately. The linear model is based on multiple linear regression and is useful to predict the rheology of a new paint formulation, for which the raw materials do not vary by more than 20% from the standard formulation.
 - ii) A non-linear model that enables the paint chemist to predict the non-linear correlations with higher accuracy than the linear model. This model is based on neural network modelling techniques and is useful to predict the rheology of a new paint formulation, for which the raw materials vary by more than 20% from the standard formulation. The non-linear model significantly reduces the error between the actual and predicted rheological value, but still results in relative large inaccuracies for certain rheological parameters. Improving the accuracy of these models requires more work in network optimisation.
- **A model has been developed for the prediction of paint properties, based on the rheological behaviour of the specific paint (MODEL 3).** Historically, other researchers have already succeeded in predicting paint properties based on the rheological data, but the current model is based on techniques that surpass those of previous models in numerous ways:
 - i) The predictability of an entire set of paint properties have been evaluated. This resulted not only in obtaining accurate predictions for certain paint properties that intuitively depend on the rheology of the paint (e.g. Krebs viscosity), but also those paint properties which do not intuitively relate to the rheology (e.g. opacity).
 - ii) Rheological parameters of an entire set of rheological measurements (e.g. flow curve, amplitude sweep, 3-ITT, etc) are used to predict paint properties. Previously, only single rheological parameters were used.

- iii) The input to the model is the rheological response to a range of time-scale/shear conditions and not only a single point as previously used.
- iv) Paint properties are predicted by the principal components of the rheological data.
- v) All of these techniques have been combined in the development of a model that accurately predicts paint properties from rheological data. The model also consists out of two parts that are run independently. A linear model accurately predicts critical paint properties such as opacity, gloss, dry film thickness and Krebs viscosity, within the allowable error margin set out by the paint manufacturer. Nowhere in the literature has such a model been developed that predicts paint properties so accurately by means of a set of rheological parameters. This is considered to be the most important contribution of this work. The model predictions have the same level of accuracy as predictions that originate from other models that use raw material data as input. The current rheological model only makes use of rheology measurements as input to the model and no information about the paint composition is required to make these predictions. This is very useful when the paint composition is not known. The linear model is inaccurate in predicting paint properties with extreme values and a non-linear model has been developed to predict these paint properties more accurately. Although the non-linear model still results in large errors between the actual and predicted value, the model shows a significant increase in model accuracy compared to the linear model. The non-linear model is not fully optimised yet, in terms of network architecture, and requires more work.
- vi) The **correlation between fundamental rheological properties and real world performance** has been quantified for a specific paint. There has always been an understanding that rheology can explain certain paint properties, but the problem was that there were no real-world correlations between rheological behaviour and paint properties. Interpretation of paint properties in terms of rheological behaviour has been quantified in this study and it is possible now to estimate the value of a specific paint property, by evaluating the rheology.

Other smaller, yet significant contributions include:

- Many paint properties can be explained by rheological parameters that are well defined, e.g. shear viscosity (η), shear modulus (G^*), complex viscosity (η^*), etc. These parameters are mere manipulations of fundamental flow and deformation behaviour. It has been shown that the fundamental flow and deformation behaviour can be used as accurate model input to predict paint properties. These fundamental rheological parameters that were used as input to the rheological model are the shear stress (τ) response from the rotational tests (flow behaviour) and a combined parameter ($t_A \cdot d$) from the oscillatory tests (viscoelastic behaviour). **These parameters have never been used before in a rheological model for prediction of paint properties.**

- A **screening process** is successfully used for identification of the most important raw materials that affect the paint properties. These raw materials are then used for further detailed models.
- Models have been developed for the **prediction of paint properties** in terms of the raw materials for the specific paint under investigation (MODEL 2). These type of models are not new in the coatings industry and many models exist that accurately model the relationship between paint properties and raw materials. However, although fundamental characteristics can be linked to certain raw materials, each paint property is a direct result of the paint's raw material composition. A model that accurately predicts the paint properties of this specific paint is therefore also unique. It has been shown that the developed model is accurate in predicting critical paint properties such as opacity, gloss, film thickness and Krebs viscosity. Two independent models have been developed that deal with linear and non-linear correlations between the raw materials and the paint properties. The linear model makes it possible for the paint chemist to identify the range of levels of raw materials that result in acceptable tolerances around a certain paint property. Very little work has been performed on non-linear modelling of paint properties and even less on neural networks as tools for non-linear modelling techniques. The neural networks are not fully optimised yet and require further work in terms of network optimisation.
- An **extensive database of rheological/mechanical and performance properties** of a certain paint family has been developed. This will be extremely useful for future investigations on this specific paint family.

The contributions emanating from this study assist in a better understanding of the relationship between raw materials and paint properties. The models developed in this study are not able to explain the paint properties and rheological behaviour in a fundamental way as a phenomenological model would have been able to. Such phenomenological models are considered to be extremely complex due to all the possible interactions between raw materials. However, the models developed in this study form the basis of such phenomenological models – the relationship between raw materials, paint rheology and paint properties for this specific paint is now established and endless virtual correlations between these can now be accurately modelled, forming a data base for a future phenomenological model that can one day be tested with actual data.

7.2 CONCLUSIONS

The following conclusions regarding this specific paint can be made:

7.2.1 CONCLUSIONS BASED ON LINEAR MODELLING

- 1) A high level of multicollinearity exists between the measured rheological parameters. This is proved by a condition number of $C = 2.67 \times 10^6$. Condition numbers > 15 can lead to inaccurate predictions.
- 2) 80% of the variance in the rheological data is explained by three principal components. Principal components are linear transformations of the original data.
- 3) A screening process based on multiple linear regression has identified the most important raw materials that affect the paint properties. In decreasing order of importance, the raw materials that affect the paint properties the most are: emulsion, pigment, extender, organic pigment, water and solvent
- 4) A total of 43 paint properties have been evaluated. A screening process based on multiple linear regression identified the paint properties that are best predicted by the 6 most important raw materials mentioned in point 3 above. These raw materials explain more than 80% of the variance in the following paint properties: gloss, Krebs viscosity, water permeability and ease of stir. Paint properties that can not be modelled accurately by means of MLR by at least one of these raw materials are: chemical resistance (NH_3 , H_2O), wet film thickness, hiding power (brush/roller), spreading rate (1st and 2nd coat), stain removal, adhesion, fibre levelling roller drag (1st coat). The variance of these paint properties are too small for accurate predictions.
- 5) Table 5.5 gives the matrix of regression coefficients that can be used in equation 5.4 to predict the rheological behaviour of this specific paint.
- 6) Rheology curves that are predicted most accurately by means of multiple linear regression and the six most important raw materials (emulsion, pigment, extender, organic pigment, water and solvent) are: the amplitude sweep (AS), the three-interval-thixotropy-test in the rotational mode (3-ITT-rot) and the flow curve (FC). The rheology curves that are predicted least accurately are: low shear curve (LS), time sweep (TS) and extra low frequency point (XLF).
- 7) Table 5.9 gives the matrix of regression coefficients that can be used in equation 5.6 to predict the paint properties of this specific paint, based on raw material data as input to the model.
- 8) The linear models of equation 5.6 predict the critical paint properties of the paint (gloss, opacity, dry film thickness and Krebs viscosity) within the specified error margin set out by the paint manufacturer. The expected average relative error on prediction of these paint properties by means of multiple linear regression for new paint samples are: gloss – 7.23%; opacity – 0.81%; dry film thickness – 6.38% and Krebs viscosity – 1.67%.

- 9) Paint properties that are predicted the most inaccurate by means of the linear models of equation 5.6 are dirt pick-up and sagging.
- 10) Table 5.12 gives the matrix of regression coefficients that can be used in equation 5.7 to predict the paint properties from the rheological behaviour of this specific paint.
- 11) Paint properties that are predicted most accurately by equation 5.7 are gloss, dry film thickness, opacity and Krebs viscosity. The following errors can be expected when using equation 5.7 to predict these properties for new paint samples: gloss – 15.33%; opacity – 1.05%; dry film thickness – 7.11% and Krebs viscosity – 1.61%.
- 12) Figure 7.1 shows that the rheological model (equation 5.7) is comparable to the raw material model (equation 5.6) when predicting paint properties of new paint samples.

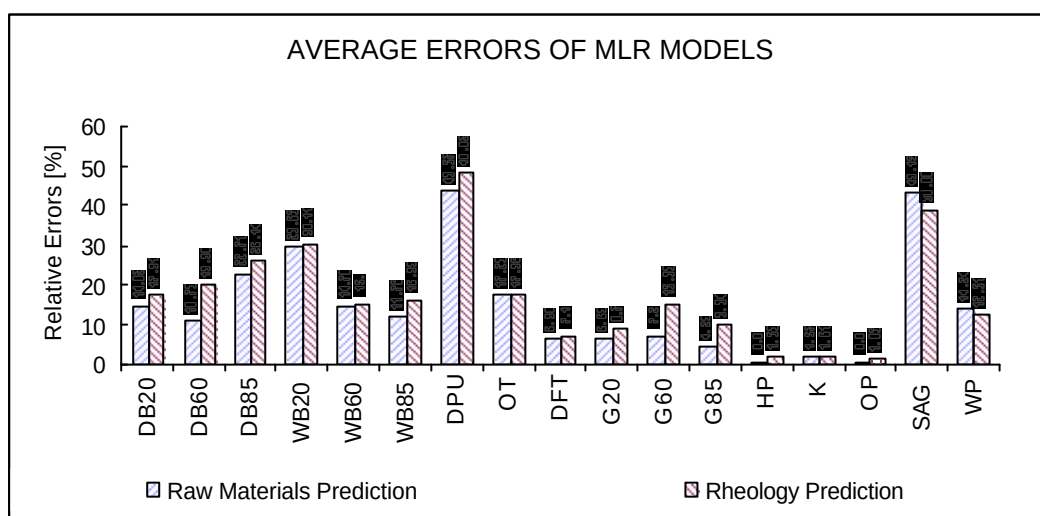


Figure 7.1 Relative errors in New Predictions

(The paint property nomenclature is as follows: DB20 – Dry Burnish20, DB60 – Dry Burnish60, DB85 – Dry Burnish85, WB20 – Wet Burnish20, WB60 – Wet Burnish60, WB85 – Wet Burnish85, DPU – Dirt Pick-Up, OT – Open Time, DFT – Dry Film Thickness, G20 – Gloss20, G60 – Gloss60, G85 – Gloss85, HP – Hiding Power, K - Krebs, OP - Opacity, SAG - Sag, WP – Water Permeability).

- 13) Paint property predictions based on equation 5.6 and equation 5.7 are only accurate when the raw materials do not vary by more than 20% from the standard formulation.

7.2.2 CONCLUSIONS BASED ON NON-LINEAR MODELLING

- 1) It has been shown that neural networks perform better than multiple linear regression in predicting the rheological behaviour of paint samples for which the raw materials vary by more than 20% from the standard formulation (MODEL 1). On average, neural networks improve predictability of the rheological parameters by 54%. The largest improvement in predictability is made on the rheological variable C_{XLF} , where multiple linear regression resulted in relative errors of 59%, while neural networks resulted in errors of only 5%.

Other predictions of rheology curves where neural networks have shown a major improvement on predictability are the Time Sweep Curve (TS) – 68% increase in accuracy and Low Shear Curve (LS) – 63% increase in accuracy. The smallest increase that the neural network had on the predictability of a rheology curve, was a 33% increase in accuracy of the Amplitude Sweep (AS) predictions. The Akaike information criterion (AIC) shows that multiple linear regression models are the preferred models when predicting all rheological parameters, except b_{FS} and C_{XLF} .

- 2) The neural networks that predict paint properties from the raw material composition (MODEL 2) show a major improvement on the predictability of the paint properties for those paint samples that vary more than 20% from the standard formulation. On average, neural networks improve predictability of the paint properties by 47%. The largest improvement in predictability is made on the wet burnish20 prediction, where multiple linear regression resulted in relative errors of 66%, while neural networks resulted in errors of only 0.6%. Other paint property predictions where neural networks have shown a major improvement on predictability of more than 80% in accuracy, are Gloss – 80% increase in accuracy and Dry Film Thickness – 81% increase in accuracy. The smallest increase that the neural network had on the predictability of a paint property, was a 33% increase in accuracy of the sag predictions. The Akaike information criterion (AIC) shows that multiple linear regression models are the preferred models when predicting any of the paint properties except Water Permeability, Open Time, Gloss20/60 and Dry Film Thickness.
- 3) The neural networks that predict paint properties from the rheological data (MODEL 3) perform better than multiple linear regression in predicting the paint properties of paint samples for which the raw materials vary by more than 20% from the standard formulation. On average, neural networks that use rheology data as input, predict the paint properties 49% more accurate than equivalent multiple linear regression models. The greatest improvement in model predictability is for water permeability - 73% increase in accuracy and gloss - 70% increase in accuracy. The Akaike information criterion (AIC) shows that multiple linear regression models are the preferred models when predicting any of the paint properties except Water Permeability, Krebs Viscosity, Sag, Open Time, Gloss60/85 and Dry Film Thickness.
- 4) Two models have been developed that predict the paint properties of those samples for which the raw material composition varies more than 20% from the standard formulation. The first model makes use of raw materials as input data (MODEL 2 - raw material models) and the second model (MODEL 3 - rheological model) makes use of rheology data as input to the model. In terms of the relative error between the actual and the predicted value of a paint property, the rheological model is more accurate in predicting paint properties such as dry/wet burnish, water permeability, Krebs viscosity, sag, dry film

thickness and open time. For paint properties such as open time, gloss and dirt pick-up, the raw material model performs better. In terms of the Akaike information criterion (AIC), which not only rewards goodness of fit, but also includes a penalty that is an increasing function of the number of estimated parameters, the neural network of the raw material model outperforms the neural network of the rheological model in predicting all of the paint properties except dry burnish85, water permeability, Krebs viscosity and sag.

- 5) Although it is clear that neural networks result in significantly more accurate predictions for the rheological behaviour of those paint samples that vary more than 20% from the standard formulation, some of the relative errors are still large, e.g. 312% for $b_{3-ITT(Osc)}$ and 270% for b_{FS} ; 76% for dry burnish85 and 75% for wet burnish85 (raw material model); 65% for dry burnish60 and 59% for dirt pick-up (rheological model). Optimisation in terms of the neural net architecture is not complete and further work in terms of network optimisation (e.g. training algorithm, activation function, number of hidden layers, etc.) is required that will aim at minimising these errors.

7.3 FUTURE WORK

This study identified the following areas that need further work:

- Empirical models that predict paint properties have been developed in this study. These models result in accurate predictions but do not lend themselves for accurate model interpretations. At the same time, a large data base of raw materials, rheology and paint property data have been obtained. A **phenomenological model** that makes use of this data and model predictions can be developed to establish the required fundamental behaviour that dictates the paint properties.
- Paint property predictions based on rheological behaviour is currently based on the principal components describing most of the variance in the rheological behaviour. However, a major drawback of principal component analysis is that the principal components lack simple interpretation since each is, in a sense, a mixture of the original variables. Future work can be done on these principal components, in order to break them down to a more understandable relationship between rheology and paint properties.
- Non-linear models show a major improvement in predictive capabilities. However, the neural network models are not fully optimised yet in terms of network architecture. The major purpose of this study, in terms of non-linear modelling, was to illustrate the principle that neural networks can be used for the prediction of paint properties. Predictions that are even more accurate are possible with fully optimised neural network models.
- This study has proved that both paint composition and rheological data can be used to predict paint properties accurately. A combined model (raw materials and rheology as input) can be

even more powerful in the accuracy of its predictions and the present dataset can be used to investigate this.

- Optimisation techniques to manufacture paints of high quality and low cost need to be developed. These techniques can be based on the model predictions obtained from this study. Paint quality is dependent on various paint properties which all have to be optimised with the same paint formulation. Therefore a single model to predict a range of properties is sought after. An option is a model that incorporates all the models developed in this study as subroutines, which can then be optimised simultaneously. A universal optimum in the quality of a paint can then be obtained.
- Models can be extended to include a variety of different paint families. This will require that similar studies need to be done on other paint families and then a comprehensive study which draws correlations between different paint families (e.g. solvent based and water based) can be undertaken.
- Cost factors for the raw materials can be included to optimise paint composition for the lowest cost, but with acceptable properties. This is especially true for cheaper paint formulas where the profit margins are small.

7.4 NOMENCLATURE

Abbreviations

3-ITT	Three-interval-thixotropy test
3-ITT (osc)	Three-interval-thixotropy test in oscillatory mode
3-ITT (rot)	Three-interval-thixotropy test in rotational mode
ANTF	Antifoamer
AS	Amplitude sweep
CLNT	Coalescent
CNST	Constant
COMB	Combination
C_{XLF}	Constant of extra low frequency
DISP	Dispersant
EMLS	Emulsion
FC	Flow curve
FNGD	Fungicide
FS	Frequency sweep
HS	High shear
H_0	Null hypothesis
LS	Low shear
MLR	Multiple linear regression
OPIG	Organic pigment
PC	Principal component
PIGM	Pigment

PVC	Pigment volume concentration
RM1	Rheology modifier 1
RM2	Rheology modifier 2
RMC	Raw material contribution
SA	Soda ash
SOLV	Solvent
SURF	Surfactant
TS	Time sweep
VS	Volume solids
WTR	Water
XLF	Extra low frequency
XTND	Extender

Symbol	Description
C	Condition number (-)
G'	Storage modulus (Pa)
G''	Loss modulus (Pa)
G*	Complex modulus (Pa)
t	Time (s)
x	Independent variable
x_{avg}	Average value of variable
x_i	Value of variable
$x_{standard deviation}$	Standard deviation of variable
$x_{standardised}$	Standardised value of variable
y	Dependent variable

Greek Symbol	Description
a	Preliminary model parameter (gradient)
a_i	Gradient of preliminary model where i=HS, LS, 3-ITT _(rot) , etc
$a_{oscillatory}$	Gradient of oscillatory measurement
$a_{Gloss60}$	Predicted Gloss60 value
$a_{PaintProperty}$	Value of predicted paint property
$a_{rotational}$	Gradient of rotational measurement
b	Preliminary model parameter (offset)
b_i	Preliminary model parameter (offset) where i=HS, LS, 3-ITT _(rot) , etc

$b_{oscillatory}$	Intersect of oscillatory measurement
$b_{emulsion}$	Regression coefficient associated with the emulsion
$b_{rotational}$	Intersect of rotational measurement
b_0	Constant associated with regression analysis
b_k	Regression coefficient associated with variable $k=1,2,3,\dots$
b_{PC1}	Regression coefficient associated with principal component 1
g	Deformation [%] or strain [-]
g_A	Deformation [%] or strain [-] amplitude
$\dot{g}$	Shear rate [1/s]
d	Phase shift, loss angle [°]
e	Random error
h	Shear viscosity [Pa.s]
h^*	Complex viscosity [Pa.s]
t	Shear stress [Pa]
t_A	Shear stress amplitude [Pa]
w	Angular frequency [rad/s]

8 APPENDIX A: GLOSSARY

AND TERMS

8.1 ACRONYMS AND ABBREVIATIONS

3-ITT	Three-interval-thixotropy-test
3-ITT (osc)	Three-interval-thixotropy test in oscillatory mode
3-ITT (rot)	Three-interval-thixotropy test in rotational mode
ANTF	Antifoamer
AS	Amplitude sweep
ASTM	American standard test methods
CLNT	Coalescent
CNST	Constant
CPVC	Critical pigment volume concentration
CRGI	Coatings research group international
CSR	Controlled shear rate
CSD	Controlled shear deformation
CSS	Controlled shear stress
C_{XLF}	Constant of extra low frequency
DISP	Dispersant
EMLS	Emulsion
FC	Flow curve
FNGD	Fungicide
FS	Frequency sweep
HEUR	Hydrophobically modified ethylene urethane oxide rheology modifier
HS	High shear
H_0	Null hypothesis
LS	Low shear
LVER	Linear viscoelastic range
MLR	Multiple linear regression
OPIG	Organic pigment
PC	Principal component
PIGM	Pigment
PVC	Pigment volume concentration
RM1	Rheology modifier 1
RM2	Rheology modifier 2
RMC	Raw material contribution
RMSE	Root mean square error

SA	Soda ash
SOLV	Solvent
SURF	Surfactant
TS	Time sweep
VOC	Volatile organic compounds
VS	Volume solids
WTR	Water
XLF	Extra low frequency
XTND	Extender

8.2 PAINT TERMS

Adhesion: The degree of attachment between a coating film and the underlying paint or other material. (www.generalpaint.com/glossary)

Brush Drag: Resistance encountered when applying a coating by brush (ASTM)

Burnish Resistance: The ability of a coating to resist an increase in its gloss (sheen) value after polishing or rubbing. (ASTM)

Chemical resistance: The ability of a coating to resist damage by chemicals. (www.homegarden.move.com)

Dirt pick-up: The degree of dirt collection on a coating. (ASTM)

Ease of Stir: The ability of a coating to resist stirring. (Plascon)

Film thickness (wet and dry): The thickness of a coating applied to a substrate. (ASTM)

Gloss: The relative luminous reflectance factor of a specimen in the mirror direction. (ASTM)

Hiding power: The ratio of the reflectance of a film on a black substrate to that of an identical film on a white substrate. (ASTM)

Krebs units: Values of a scale commonly used to express the consistency of paints generally applied by brush or roller. (ASTM)

Levelling: A measure of the ability of a paint to flow out after application so as to obliterate any surface irregularities such as brush marks, orange peel, peaks, or craters that have been produced by the mechanical process of application. (ASTM)

Microbial Attack: Degree of the presence of fungal or algal growth on a coatings. (ASTM)

Opacity: The ratio of the reflectance of a film on a black substrate to that of an identical film on a white substrate. (ASTM)

Open Time: Length of time a coating remains wet enough to allow for brushing-in at the laps; also called wet edge time. (ASTM)

Roller Drag: Resistance encountered when applying a coating by roller. (ASTM)

Sagging: Non-uniform downward flow of a wet paint film that occurs between the times of application and setting, resulting in an uneven coating having a thick lower edge. (ASTM)

Scrub resistance: Resistance of paints to erosion caused by scrubbing. (ASTM)

Settling: Degree of pigment suspension of a shelf-aged sample of paint under homogeneous conditions. (ASTM)

Spatter: Droplets of paint that spin or mist off the roller as paint is being applied. (www.flexbon.com)

Spreading Rate: The area covered by a unit volume of coating material. (ASTM)

Stain Removal: The relative ease of removal of common stains from interior coatings by manual or mechanical washing with a sponge and a liquid cleanser. (ASTM)

Syneresis: Degree of phase separation of a shelf-aged liquid paint that is stored under homogeneous conditions. (ASTM)

Water Permeability (vapour): The rate at which water vapour passes through films of paint. (ASTM)

Water Resistance: The ability of a coating to resist water (liquid) permeating through the film coating, causing effects such as colour change, blistering, loss of adhesion, softening or embrittlement. (ASTM)

8.3 STATISTICAL TERMS

8.3.1 THE NULL HYPOTHESIS

The following description is taken from [Montgomery*]

$$\begin{aligned}H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0 \\ H_1 : \beta_j \neq 0 \quad \text{for at least one } j\end{aligned}$$

The statement $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0$ is called the null hypothesis and $H_1 : \beta_j \neq 0$ is called the alternative hypothesis. One way to report the results of a hypothesis test is to state that the null hypothesis was or was not rejected at a specific α -level or level of significance. Rejection of H_0 implies that at least one of the regressor variables contributes significantly to the model and therefore it is customary to call the test statistic (and the data) significant when the null hypothesis H_0 is rejected. The p-value approach has been adopted widely in practice to evaluate the null hypothesis (Montgomery). The p-value is defined as the smallest level of significance that would lead to rejection of the null hypothesis H_0 . The p-value conveys much information about the weight of evidence against H_0 , and so the decision maker can draw a conclusion at a specified level of significance. H_0 is rejected if the p-value is less than α .

8.3.2 MULTICOLLINEARITY

The following definition is taken from [Applied linear statistical models]

'When the predictor variables are correlated among themselves, *intercorrelation* or *multicollinearity* among them is said to exist.'

The effect of multicollinearity is that estimated regression coefficients tend to vary widely from one sample to the next when the predictor variables are highly correlated. As a result, only imprecise information may be available about the individual true regression coefficients.

The fact that some or all predictor variables are correlated among themselves does not, in general, inhibit our ability to obtain a good fit nor does it tend to affect inferences about mean responses or predictions of new observations, provided these inferences are made within the region of observations. Also, multicollinearity is not so serious when the main goal of the regression analysis is to predict the dependent variable.

1. * Montgomery, D. C. (2005). Design and Analysis of Experiments, John Wiley and Sons Inc.

9 APPENDIX B: RHEOLOGY

FUNDAMENTALS

The following description on rheology fundamentals is compiled from various sources [1], [2], [3], [4].

9.1 FLOW BEHAVIOUR AND VISCOSITY

The ‘Two-Plate-Model’ (Figure 9.1) is used to define some of the fundamental rheological parameters encountered during flow behaviour. The upper plate, with shear area A , is moved by the shear force F and the resulting velocity is measured. The lower plate is immovable. The distance h is the distance between the plates. The liquid sample is sheared in this gap.

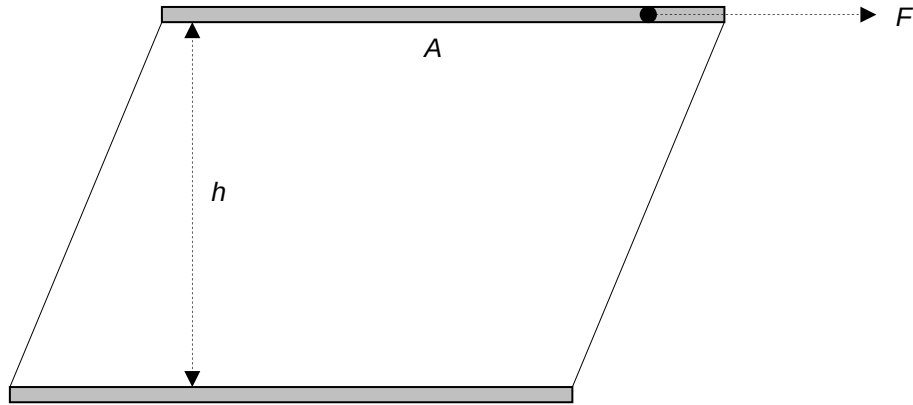


Figure 9.1 The Two-Plate Model

In all fluids there are frictional forces between the molecules and, therefore, they display a certain flow resistance which can be measured as viscosity. For ideal viscous (Newtonian) fluids at a constant temperature, the value of the ratio of the shear stress (τ) to the corresponding shear rate ($\dot{\gamma}$) is a material constant. The definition of shear viscosity is:

$$\eta = \frac{\tau}{\dot{\gamma}} \quad (9.1)$$

Where the shear stress (τ) is defined as:

$$\tau = \frac{F}{A} \quad (9.2)$$

Where F is the shear force that is applied over a shear area A . The definition of the corresponding shear rate ($\dot{\gamma}$) is:

$$\dot{\gamma} = \frac{dv}{dh} \quad (9.3)$$

Where dv is the differential velocity between two neighbouring flowing layers and dh the differential thickness of one of the individual flowing layers.

There are two common approaches for rheological measurements; namely controlled-rate and controlled-stress measurements. In controlled-rate tests, the speed or the shear rate is set and controlled and the torsional force produced is measured. This test method is called a “controlled shear rate test” or “CSR” test. This test method is usually selected when specific flow velocities of technical processes have to be simulated (e.g. flow in pipes or particle sedimentation). In controlled stress tests, a stress is applied and the resulting displacement or rotational speed is measured.

ROTATIONAL TESTS: CONTROLLED SHEAR RATE (CSR) TESTS

In shear rate tests the speed or the shear rate is set and controlled. The test method with controlled shear rate is usually selected when specific flow velocities of technical processes have to be simulated. The following tests are performed in the CSR mode:

- a) Flow and viscosity curves

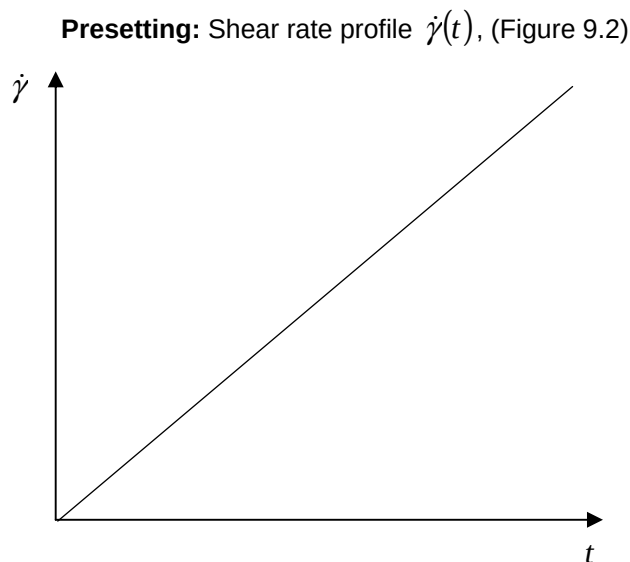


Figure 9.2 Preset Profile: Shear Rate as a Function of Time

Measuring result: Flow curve $\tau(\dot{\gamma})$, (Figure 9.3)

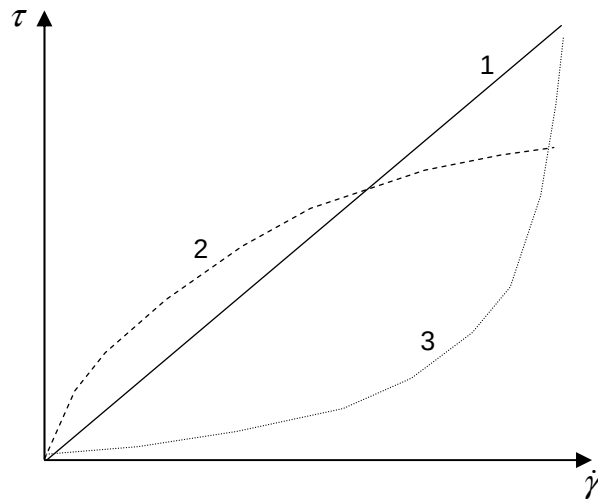


Figure 9.3 Measuring Result: (1) Ideal viscous[†] (Newtonian) (2) Shear-thinning[‡] (Pseudoplastic) (3) Shear-thickening[§] (Dilatant)

Result: Viscosity curve $\eta(\dot{\gamma})$, (Figure 9.4)

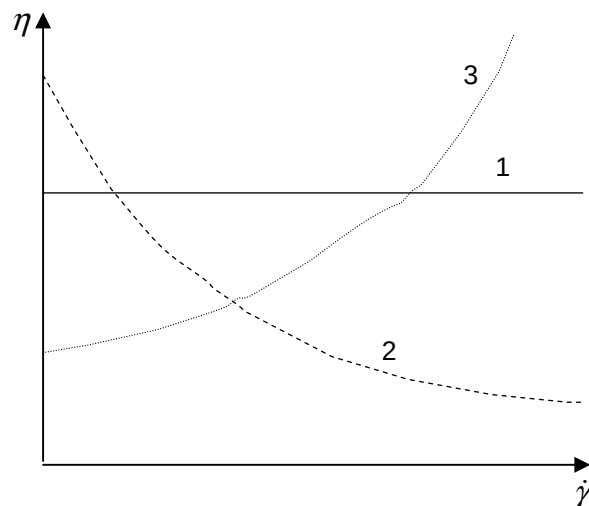


Figure 9.4 Result: (1) Ideal viscous (2) Shear-thinning (3) Shear-thickening

When measuring at shear rates of $\dot{\gamma} < 1 \text{ s}^{-1}$, it is important to ensure that the duration of each measured point is long enough. This is especially true for high-viscosity samples when tested at

[†] The shear viscosity of an ideal viscous (Newtonian) fluid is independent of the degree or duration of the shear load.

[‡] The shear viscosity is dependent on the degree of shear load. The flow curve shows a decreasing curve slope.

[§] The shear viscosity is dependent on the degree of shear load. The flow curve shows an increasing curve slope.

very low shear rates. Otherwise start effects or time-dependent transition effects are measured, this means that transient viscosity instead of the desired steady-state viscosity is measured. Steady-state viscosity is only dependent on the applied shear rate. When $\dot{\gamma} > 1 \text{ s}^{-1}$, transient effects only influence samples with pronounced viscoelastic properties. Therefore, for liquids with low or medium viscosities the duration of $t = 5 \text{ s}$ is sufficient in most cases for each measuring point. The preset duration of each measured point should decrease logarithmically as the shear rate increases. This method ensures that longer measuring point durations are set for lower shear rates than for higher ones so that time-dependant transition effects are not measured which usually occur at lower shear rates. Consequently, hardly any significant transient effects can be expected over the entire shear rate range.

b) Viscosity / time curve

Presetting: Constant shear rate $\dot{\gamma}$, (Figure 9.5)

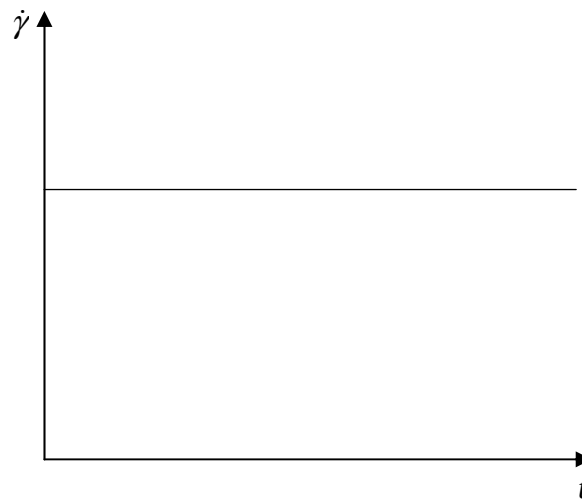


Figure 9.5 Preset Profile: Shear Rate/Time Function with Constant Shear Rate

Measuring result: Shear stress / time curves (Figure 9.6)

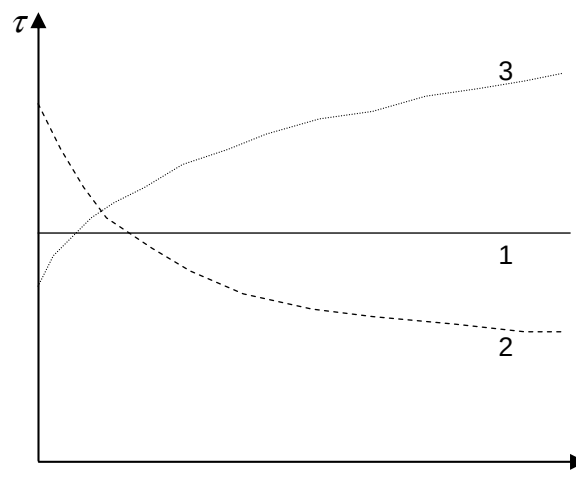


Figure 9.6 Measuring Result: (1) With No Change in Viscosity (2) Decrease in Viscosity Over Time (3) Increase in viscosity Over Time

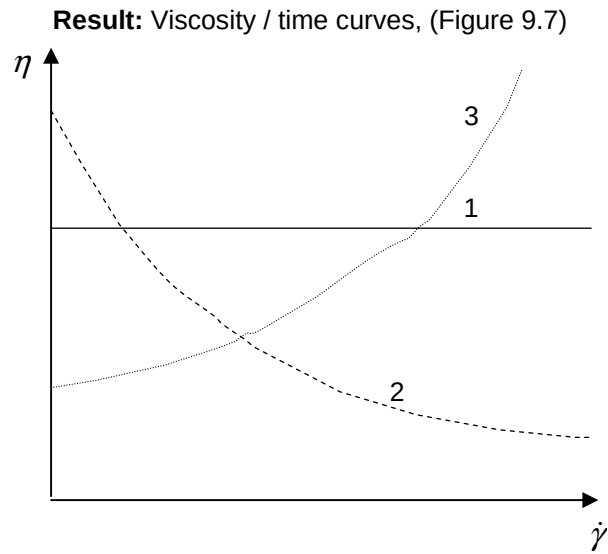


Figure 9.7 Measuring Result: (1) No Change in Viscosity (2) Decrease in Viscosity (3) Increase in Viscosity

When presetting at low shear rates, the selected measuring point duration must be long enough to overcome transient effects.

c) Structural decomposition and regeneration (thixotropy)

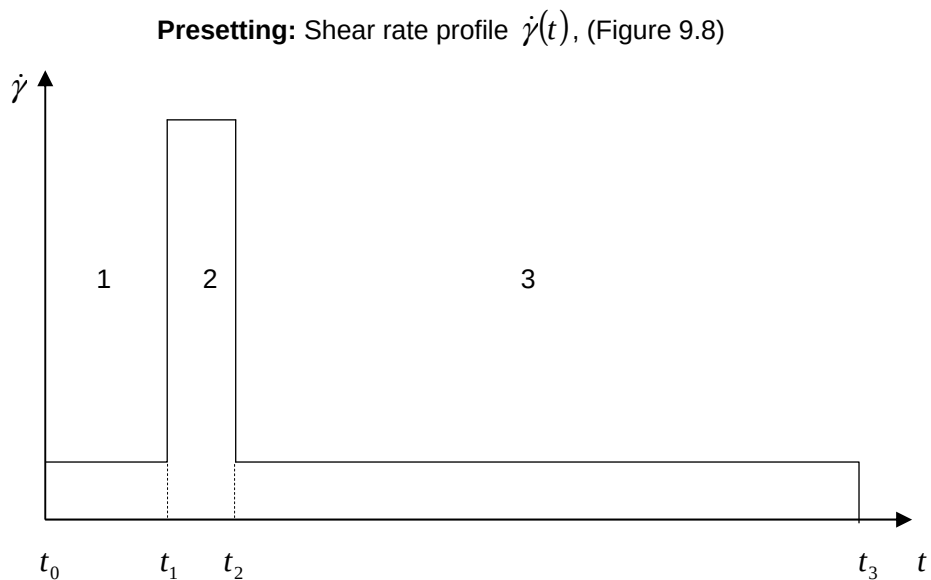


Figure 9.8 Preset Profile: Shear Rate / Time Function with Three Test Intervals

For measurements like this, three test intervals are preset:

- 1) Rest phase under low-shear conditions during the time period between t_0 and t_1 . The aim is to achieve a fairly constant viscosity (η) value for the whole first interval, since it is then used as a reference value for the third interval
- 2) Load phase under high-shear conditions during the time period between t_1 and t_2 in order to decompose the structure of the sample
- 3) Phase after removing the load under low-shear conditions during the time period between t_2 and t_3 , under the same shear conditions as in the first interval to facilitate regeneration of the structure.

Result: Viscosity as a function of time $\eta(t)$, (Figure 9.9)

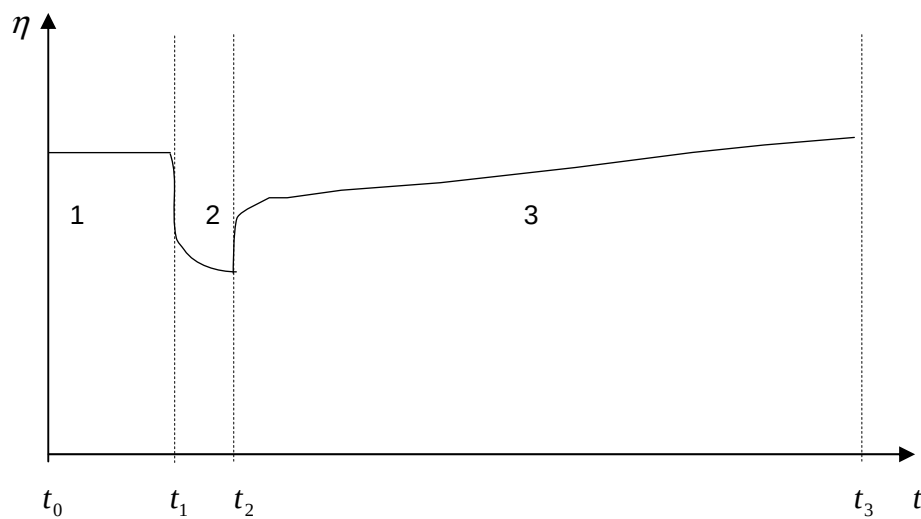


Figure 9.9 $\eta(t)$ Curve of a Thixotropic Material Under (1) Low-shear Conditions (2) During Structural Decomposition and (3) During Structural Regeneration

There are various methods for analysing structural regeneration, all of them which are based either on i) the time it takes for the structure to reach a certain degree of regeneration based on the 1st interval or ii) the degree of structural regeneration achieved (as a percentage) after a specified period of time, based on the 1st interval.

9.2 ELASTIC BEHAVIOUR AND SHEAR MODULUS

For elastic solids at a constant temperature, the value of the ratio of the shear stress (τ) to the corresponding deformation (γ) is a material constant if measured in the reversible deformation range (the so-called “linear-elastic range”). It is referred to as the shear modulus (G) and reveals information about the rigidity of a material. Oscillatory tests are performed to characterise elastic behaviour, which can be expressed in terms of the shear modulus. Materials with stronger intermolecular cohesive forces show a higher internal rigidity and therefore also a higher value of G . The definition of the shear modulus is:

$$G = \frac{\tau}{\gamma} \quad (9.4)$$

Where the shear stress (τ) is defined in equation (9.2) and the deformation (γ) is defined as:

$$\gamma = \frac{s}{h} \quad (9.5)$$

With s the deflection and h the distance between the plates.

9.3 VISCOELASTIC BEHAVIOUR

The 'Two-Plate-Model' (Figure 9.10) is also used to explain some of the fundamental rheological parameters encountered during viscoelastic behaviour. The upper plate with shear area A is moved back and forth with shear force $\pm F$ in an oscillatory way. This causes shearing of the sample between the two plates.

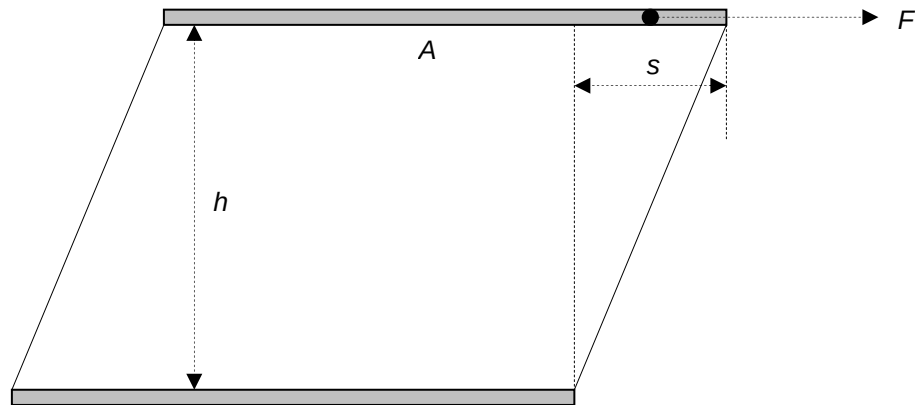


Figure 9.10 'Two-Plate-Model' for Viscoelastic Materials

A viscoelastic material always shows viscous and elastic behaviour simultaneously and it is therefore important to understand the ideal elastic and ideal viscous responses to the oscillatory deformation. Ideal elastic behaviour of a solid (Figure 9.11) is characterised by immediate response (or no delay) in the deformation behaviour $\gamma(t)$ when subjected to a shear stress $\tau(t)$, resulting in a phase shift angle (δ) of 0° .

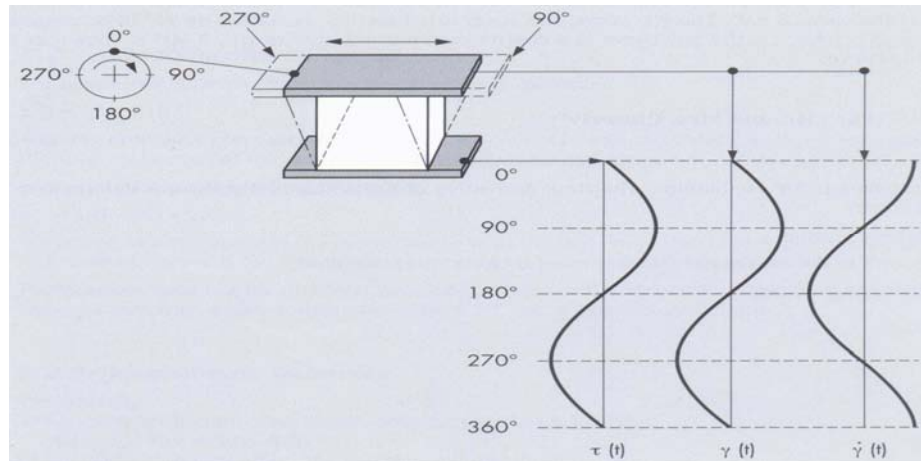


Figure 9.11 Ideal Elastic Behaviour of a Solid [1]

Ideal viscous behaviour of a liquid (Figure 9.12) is characterised by a delay in the deformation behaviour $\gamma(t)$ when subjected to shear stress $\tau(t)$, resulting in a phase shift angle (δ) of 90° .

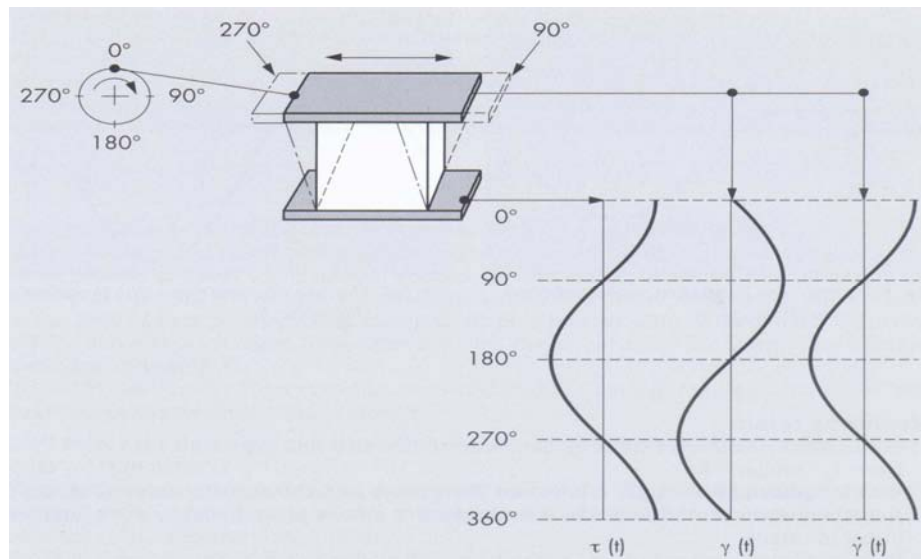


Figure 9.12 Ideal Viscous Behaviour of a Liquid [1]

Therefore, when real materials are subjected to oscillatory shear stress deformation, the phase shift angle will show a delay of $0^\circ < \delta < 90^\circ$. The closer δ is to 0° the more elastic the material is, while the closer δ is to 90° the more viscous the material is.

OSCILLATORY TESTS: CONTROLLED SHEAR DEFORMATION (CSD) TESTS

The following tests are performed in the CSD mode:

a) Amplitude sweep

An amplitude sweep is an oscillatory test with variable amplitude and constant frequency values. Amplitude sweeps are mostly carried out for the sole purpose of determining the limit of the linear

viscoelastic range (LVER, γ_{LVER}) but also gives information about the internal structure of the sample.

Presetting: Controlled shear deformation: $\gamma(t) = \gamma_A \cdot \sin(\omega t)$, with γ_A the variable strain amplitude and $\omega = \text{constant}$. (Figure 9.13)

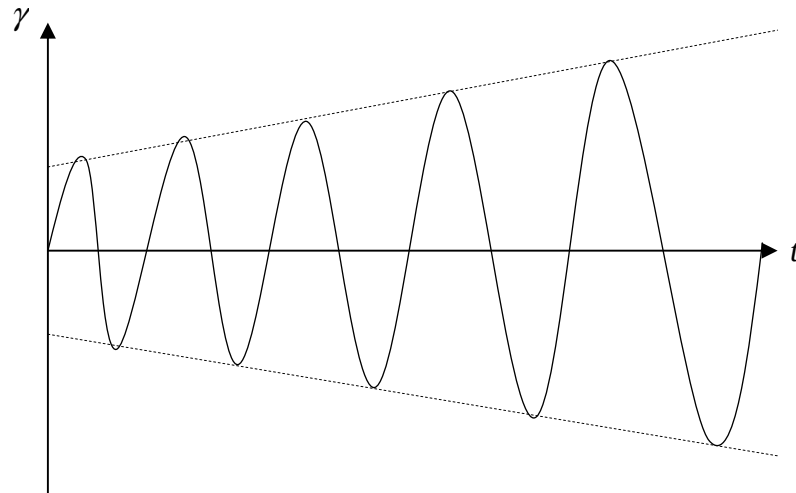


Figure 9.13 Presetting: Amplitude Sweep

Measuring result: Both G' and G'' are represented on the y-axis (log scale) with γ on the x-axis (log scale), (Figure 9.14)

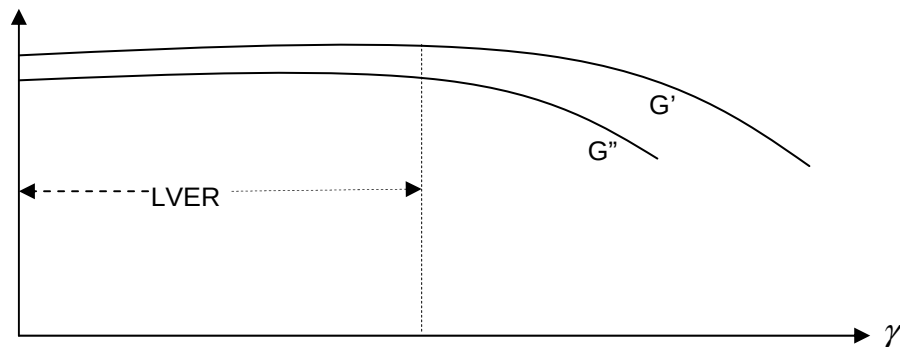


Figure 9.14 Measuring Result: Amplitude Sweep

Amplitude sweeps are used to investigate the character of the sample. In the linear viscoelastic range (LVER) each of the $G'(\gamma)$ and $G''(\gamma)$ moduli shows a constant plateau value and the structure of the sample is stable under this low-deformation condition. The term linear viscoelastic range (LVER) takes its name from the proportionality of the preset and measured parameters, and as a consequence, the resulting curve is a linear straight line in the diagram. It is within the LVER that the structure of the sample deforms reversibly, however at deformations larger than the limiting

deformation (γ_{LVER}), the structure deforms irreversibly. The limit of the LVER is also an indication of the yield point of the sample, or the stress that needs to be overcome for the sample to flow. Therefore, in consequent oscillatory tests, if any information regarding the structure of the sample needs to be determined, then care must be taken to perform these tests within the LVER ($\gamma < \gamma_{LVER}$).

The following relations concerning the structure of the sample are visible in the LVER:

1) Gel character if $G' > G''$

Here, the elastic behaviour dominates over the viscous behaviour. The structure shows certain rigidity. Dispersions such as coatings that show $G' > G''$ in the LVER have a gel structure and therefore a certain form of stability.

2) Liquid character if $G'' > G'$

In this case, the viscous behaviour dominates over the elastic behaviour. The sample shows the character of a liquid in the LVER. This type of material does not remain in the stable form; they run or creep, although sometimes with a very low viscosity.

It must also be remembered that amplitude sweeps do not examine the sample at rest, because the samples are often tested at angular frequencies that do not represent the condition at rest.

b) Frequency sweep

A frequency sweep is an oscillatory test with variable frequency and constant amplitude values. The importance of frequency sweeps is that the time-dependent shear behaviour is examined. Short-term behaviour is simulated by rapid oscillations (at high frequencies) and long-term behaviour is simulated by slow oscillations (at low frequencies)

Presetting: Controlled shear deformation: $\gamma(t) = \gamma_A \cdot \sin(\omega t)$, with variable angular frequency $\omega = \omega(t)$ and with $\gamma_A = \text{constant}$. (Figure 9.15)

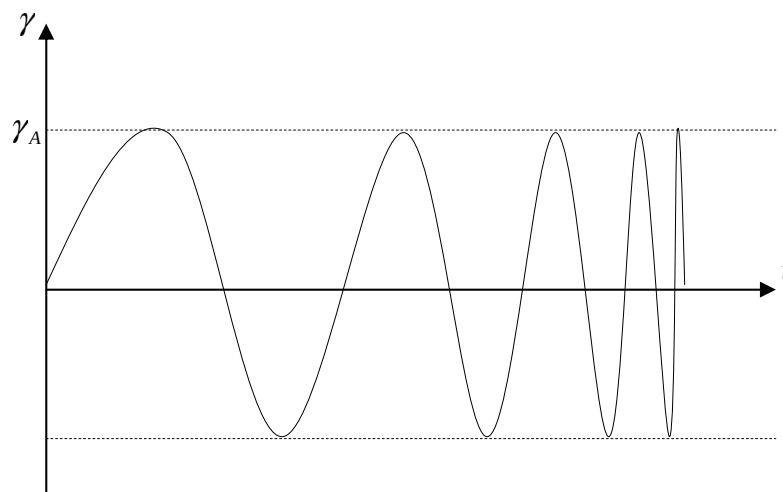


Figure 9.15 Presetting: Frequency Sweep

A frequency sweep must be performed below the limit of the LVER (γ_{LVER}) and therefore an amplitude sweep must always be carried out first to determine this limit.

Measuring result: Both G' and G'' are represented on the y-axis (log scale) with the frequency f on the x-axis (log scale), (Figure 9.16)

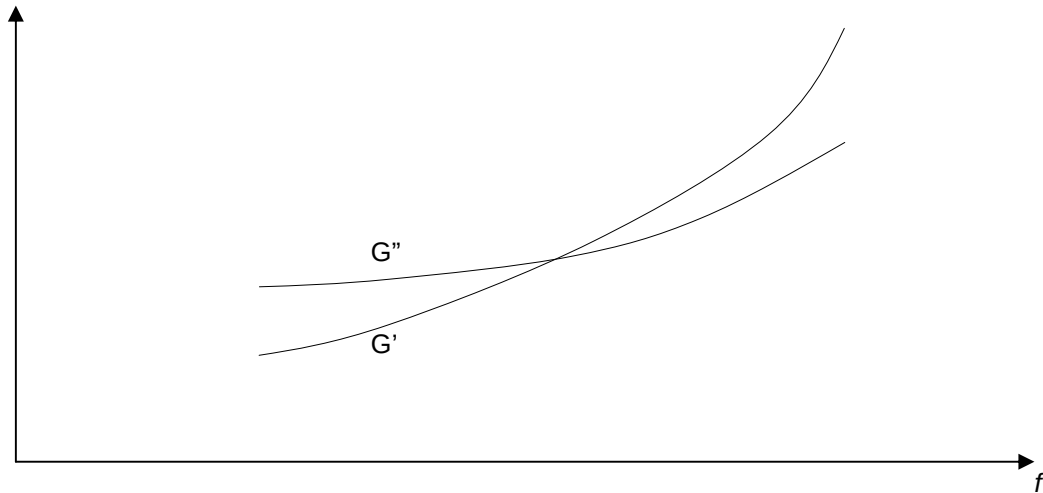


Figure 9.16 Measuring Result: Frequency Sweep

Frequency sweeps are time tests because they simulate time-dependent behaviour. Short and long-term behaviour can be evaluated by investigating the high and low frequency regions respectively.

A sample with a stable structure usually displays $G' > G''$; the elastic behaviour dominates over the viscous behaviour and the sample shows gel character. For a stable structure, the G' and G'' curves are almost parallel straight lines throughout the entire frequency range with a shallow slope, often with $G':G'' = 10:1$ up to $100:1$.

The analysis of the sample's consistency is undertaken in the LVER at low frequencies and has many advantages over the conventional rotational tests where the structure of the sample is evaluated at low shear conditions:

- 1) The measurements actually represent the sample's structure at rest and make the evaluation of the results even more precise.
- 2) Two rheological parameters (stress (τ) and phase shift (δ)) are measured instead of a single parameter (only stress (τ) as in the case of rotational tests).

c) Time sweep

In this type of oscillatory test, both the frequency and the amplitude are set at a constant value. Therefore constant mechanical shear conditions are given.

Presetting: Controlled shear deformation: $\gamma(t) = \gamma_A \cdot \sin(\omega t)$, with constant angular frequency $\omega = \omega(t)$ and with constant amplitude strain γ_A . (Figure 9.17)

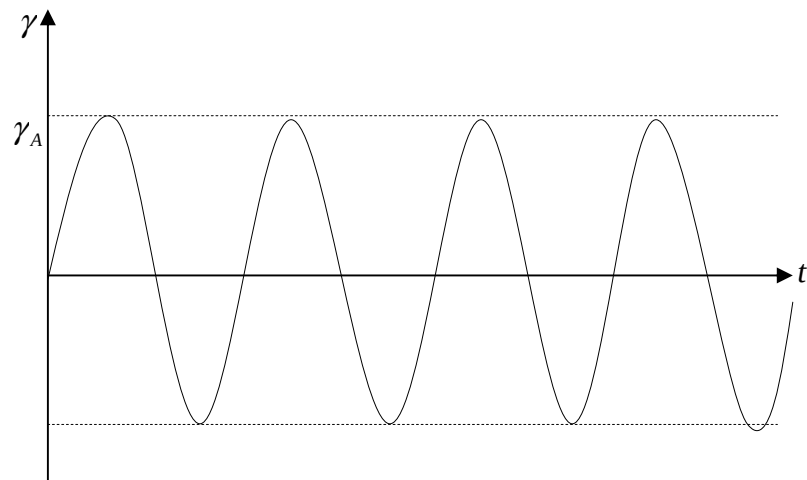


Figure 9.17 Presetting: Time Sweep

A time sweep must be performed below the limit of the LVER (γ_{LVER}) as determined by the amplitude sweep.

Measuring result: Both G' and G'' are represented on the y-axis (log scale) with the time t on the x-axis (linear scale), (Figure 9.18)

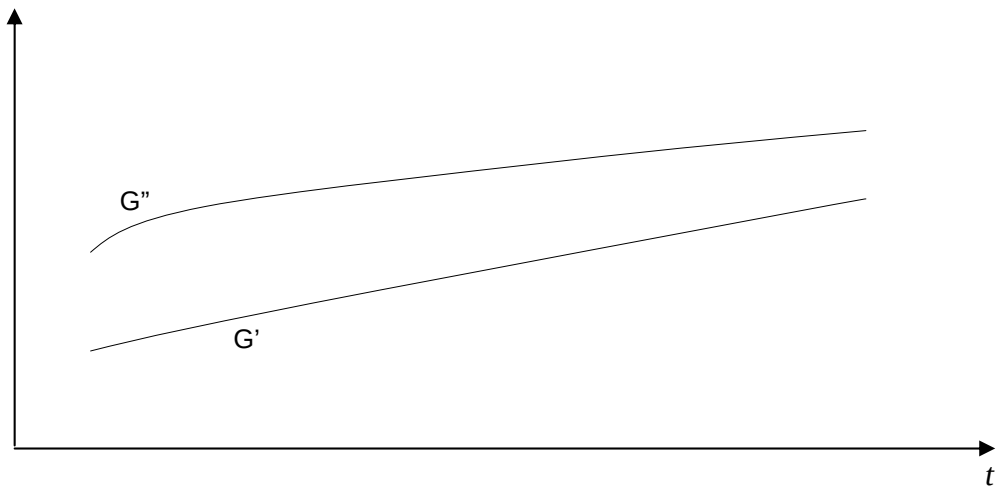


Figure 9.18 Measuring Result: Time Sweep

The test aims to determine the influence of the shear load on the physical properties, e.g. showing the change in structural strength over time.

d) Structural decomposition and regeneration (thixotropy)

Thixotropic behaviour means that a material shows a decrease in the structural strength during the load phase and shows regeneration during the rest phase. In other words, during the rest phase, the original structure regenerates completely in a certain period of time.

Presetting: Oscillatory test with three intervals, each under constant dynamic-mechanical conditions. (Figure 9.19)

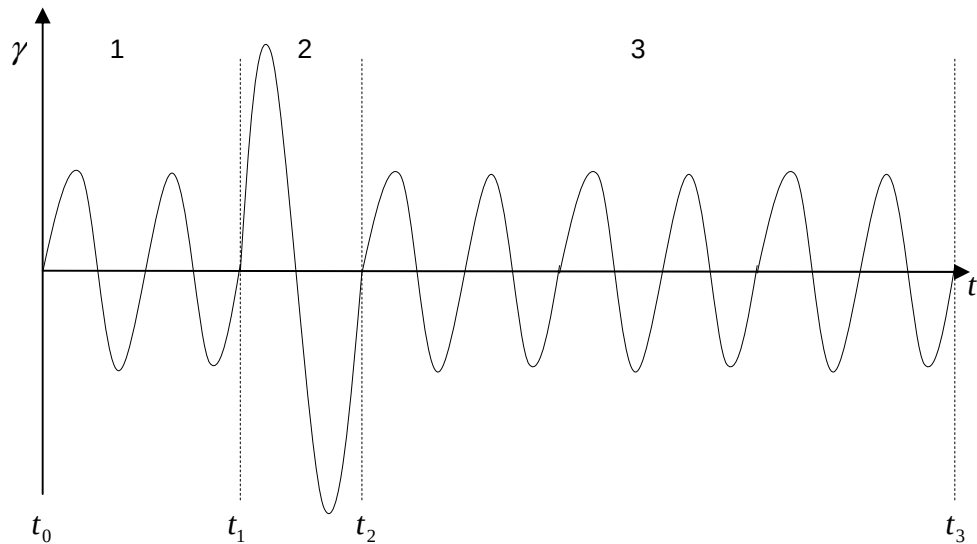


Figure 9.19 Presetting: Structural Breakdown and Regeneration

The shear load is set at a constant value in each of the individual test intervals and this is achieved by keeping both the amplitude and the frequency constant in each interval. A measurement like this consists of three test intervals:

- 1) Rest phase under low-shear conditions (within the LVER) during the time period t_0 and t_1 . The first interval is used as reference for the third interval.
- 2) Load phase under high-shear conditions (outside the LVER) during the time period t_1 and t_2 in order to decompose the sample's structure.
- 3) Phase after removing the load under low-shear conditions (within the LVER) during the time period t_2 and t_3 , under the same conditions as in the first interval, to facilitate regeneration of the structure.

Measuring result: Both G' and G'' are represented on the y-axis (log scale) with the time t on the x-axis (linear scale), (Figure 9.20)

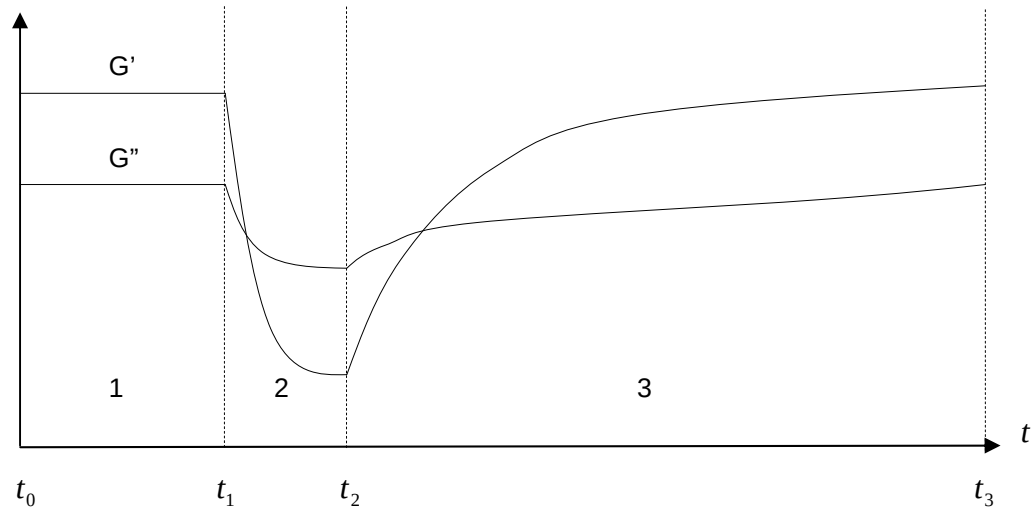


Figure 9.20 Measuring Result: Structural Breakdown and Regeneration

There are various options for the analysis (as is in the case for the rotational tests) of thixotropic behaviour, once again based on i) degree of structural recovery based on the 1st interval and ii) the time it takes for a certain structural recovery to occur.

9.4 NOMENCLATURE

Symbol

Symbol	Description
A	Shear area of Two-Plate Model (m^2)
h	Distance between plates of Two-Plate Model (m)
F	Shear force (N)
G	Shear modulus (Pa)
G'	Storage modulus (Pa)
G''	Loss modulus (Pa)
s	Deflection (m)
t	Time (s)
t_0 to t_3	Different points in time (s)

Greek Symbol

Greek Symbol	Description
γ	Deformation or strain (% or 1)
γ_A	For oscillatory tests: deformation or strain amplitude (% or 1)
γ_{LVER}	For amplitude tests: limiting deformation value of the LVER (% or 1)
$\dot{\gamma}$	Shear rate (1/s)
δ	Phase shift angle, loss angle ($^\circ$)

η	Shear viscosity (Pa.s)
τ	Shear stress (Pa)
ω	Angular frequency (rad/s)

9.5 REFERENCES

1. Mezger, T.G. 2002, *The Rheology Handbook: For users of rotational and oscillation rheometers*. Hannover: Vincentz Verlag
2. Schramm, G. *A Practical Approach to Rheology and Rheometry*. HAAKE GmbH, Karlsruhe, Germany
3. Slatter, P. 2001, *Practical Rheology of Polymer Systems*. Cape Technikon Rheology Centre, Cape Town, South Africa
4. Barnes, H.A. 1989, *An Introduction to Rheology*. Amsterdam: Elsevier

10 APPENDIX C: RAW DATA

10.1 PAINT FORMULATION (RAW MATERIALS)

10.1.1 SCREENING

Paint ID	Water	Soda Ash	Fungicide	Surfactant	Dispersant	Antifoamer	Solvent	Pigment	Extender	Emulsion	Organic Pigment	Coalescent	RM1	RM2
8	15.73	0.12	0.16	0.18	1.42	0.39	3.26	13.87	14.50	39.68	7.84	1.91	0.82	0.13
4	14.43	0.15	0.20	0.23	1.18	0.49	4.09	17.34	12.12	37.46	9.77	1.59	0.82	0.13
7	15.73	0.12	0.16	0.18	0.94	0.39	3.26	20.81	14.50	29.95	11.74	1.27	0.82	0.13
12	15.73	0.18	0.16	0.18	1.42	0.59	3.26	13.87	9.66	40.97	11.76	1.27	0.82	0.13
10	15.73	0.12	0.24	0.28	1.42	0.59	3.26	13.87	14.50	39.94	7.83	1.27	0.82	0.13
6	13.23	0.12	0.16	0.18	0.94	0.39	3.26	17.07	9.66	44.94	7.83	1.27	0.82	0.13
15	15.73	0.18	0.24	0.18	0.94	0.39	4.90	13.87	14.50	39.01	7.83	1.27	0.82	0.13
26	15.73	0.12	0.24	0.28	1.42	0.39	3.26	20.81	9.66	33.48	11.75	1.91	0.82	0.13
22	13.24	0.18	0.16	0.28	0.94	0.59	4.90	20.81	14.50	30.44	11.75	1.27	0.82	0.13
25	13.23	0.12	0.24	0.18	1.42	0.59	3.26	20.81	14.50	31.68	11.75	1.27	0.82	0.13
2	14.46	0.15	0.20	0.23	1.18	0.49	4.09	17.34	12.11	37.45	9.77	1.59	0.82	0.13
16	13.23	0.18	0.16	0.18	1.42	0.59	4.90	20.81	14.50	33.34	7.83	1.91	0.82	0.13
5	14.42	0.15	0.20	0.23	1.18	0.49	4.09	17.34	12.12	37.46	9.77	1.59	0.82	0.13
18	13.23	0.18	0.16	0.28	0.94	0.39	3.26	20.81	9.66	40.39	7.83	1.91	0.82	0.13
28	15.73	0.18	0.24	0.18	1.42	0.59	4.90	20.81	9.66	36.24	7.83	1.27	0.82	0.13
17	13.23	0.12	0.24	0.28	1.42	0.39	4.90	20.81	9.66	38.90	7.83	1.27	0.82	0.13
21	13.23	0.12	0.16	0.18	1.42	0.59	4.90	14.12	9.66	44.93	7.83	1.91	0.82	0.13
3	14.44	0.15	0.20	0.23	1.18	0.49	4.01	17.35	12.13	37.49	9.78	1.59	0.82	0.13
13	15.73	0.18	0.16	0.28	1.42	0.39	3.26	13.87	9.66	41.08	11.75	1.27	0.82	0.13
11	15.73	0.18	0.24	0.28	0.94	0.59	3.26	13.87	14.50	39.72	7.83	1.91	0.82	0.13
9	13.23	0.18	0.24	0.28	1.43	0.59	3.26	20.80	14.50	30.89	11.74	1.91	0.82	0.13
27	13.24	0.18	0.16	0.28	1.42	0.39	4.90	13.87	14.49	37.10	11.75	1.27	0.82	0.13
19	13.36	0.18	0.25	0.18	0.95	0.60	3.29	14.00	9.75	42.72	11.86	1.93	0.82	0.13
1	14.42	0.15	0.20	0.23	1.18	0.49	4.09	17.34	12.12	37.46	9.77	1.59	0.82	0.13
29	13.23	0.13	0.24	0.18	0.94	0.39	4.90	13.87	14.50	37.01	11.75	1.91	0.82	0.13
24	13.23	0.12	0.24	0.28	0.94	0.59	4.90	15.06	9.66	44.92	7.84	1.27	0.82	0.13
20	15.74	0.18	0.24	0.18	0.94	0.39	4.90	20.81	9.66	32.36	11.75	1.91	0.82	0.13
23	15.73	0.12	0.16	0.28	0.94	0.59	4.90	20.81	14.50	31.28	7.83	1.91	0.82	0.13
14	15.73	0.12	0.16	0.28	0.94	0.59	4.90	13.87	9.67	39.14	11.75	1.91	0.82	0.13

10.1.2 MODELLING

Paint ID	Water	Soda Ash	Fungicide	Surfactant	Dispersant	Antifoamer	Solvent	Pigment	Extender	Emulsion	Organic Pigment	Coalescent	RM1	RM2
1	18.96	0.12	0.16	0.18	0.94	0.39	3.11	16.07	9.21	40.00	8.82	1.27	0.63	0.11
2	23.28	0.14	0.18	0.21	1.09	0.45	3.75	15.54	13.22	29.03	10.62	1.47	0.89	0.13
3	21.50	0.14	0.18	0.21	1.07	0.44	3.53	18.22	10.44	31.85	10.00	1.44	0.84	0.13
4	22.06	0.13	0.18	0.20	1.04	0.43	4.21	15.81	12.68	32.40	8.61	1.41	0.72	0.12
5	22.91	0.14	0.18	0.21	1.08	0.45	4.36	18.25	11.00	28.81	10.13	1.46	0.89	0.14
6	22.52	0.17	0.22	0.25	1.30	0.54	4.30	15.80	13.45	29.51	9.13	1.76	0.91	0.14
7	19.43	0.17	0.22	0.25	1.30	0.54	4.45	16.38	11.23	32.63	10.75	1.75	0.79	0.12
8	18.76	0.15	0.21	0.24	1.22	0.50	4.50	19.84	11.37	29.78	10.89	1.64	0.77	0.12
9	21.36	0.15	0.19	0.22	1.15	0.48	3.35	13.90	9.92	38.78	8.03	1.55	0.80	0.11
10	17.70	0.12	0.16	0.18	0.93	0.39	3.20	16.53	11.30	38.41	9.07	1.25	0.64	0.12
11	17.23	0.13	0.18	0.21	1.06	0.44	3.91	13.83	9.87	42.84	7.99	1.42	0.80	0.11
12	20.22	0.14	0.18	0.21	1.09	0.45	3.17	13.16	9.39	40.76	8.99	1.46	0.66	0.11
13	22.44	0.14	0.18	0.21	1.08	0.45	3.73	19.28	12.12	28.94	8.95	1.46	0.89	0.12
14	23.85	0.17	0.22	0.25	1.30	0.54	3.81	15.78	12.71	29.49	9.12	1.76	0.88	0.12
15	20.09	0.14	0.19	0.21	1.10	0.46	4.42	19.47	13.31	29.22	9.04	1.48	0.76	0.12
16	19.79	0.17	0.23	0.26	1.35	0.56	4.50	18.34	11.65	30.53	9.73	1.82	0.94	0.13
17	18.49	0.16	0.22	0.25	1.29	0.54	3.88	19.35	13.68	30.04	9.29	1.74	0.93	0.13
18	19.15	0.14	0.18	0.21	1.08	0.45	3.14	16.23	9.30	40.39	7.54	1.45	0.63	0.11
19	17.70	0.12	0.16	0.18	0.93	0.39	3.20	16.53	11.30	38.41	9.07	1.25	0.64	0.12
20	16.25	0.13	0.17	0.19	1.00	0.41	3.29	14.49	11.62	42.30	7.89	1.35	0.79	0.11
21	20.17	0.13	0.17	0.20	1.00	0.42	3.71	13.13	9.36	40.65	8.97	1.35	0.64	0.12
22	22.78	0.14	0.18	0.21	1.09	0.45	4.38	19.31	11.07	28.99	8.97	1.46	0.85	0.12
23	15.81	0.12	0.16	0.18	0.94	0.39	3.24	16.72	9.58	41.60	9.18	1.27	0.71	0.10
24	18.86	0.14	0.19	0.22	1.12	0.47	4.20	19.94	11.43	29.93	10.95	1.51	0.92	0.12
25	20.98	0.17	0.23	0.26	1.33	0.55	3.89	17.11	11.50	30.12	11.01	1.79	0.93	0.12
26	22.84	0.14	0.18	0.21	1.08	0.45	3.72	18.58	11.01	28.83	10.54	1.46	0.84	0.12
27	18.92	0.16	0.22	0.25	1.29	0.54	4.43	19.52	12.27	29.30	10.44	1.74	0.76	0.14
28	21.80	0.14	0.18	0.21	1.07	0.44	3.68	18.97	12.97	28.48	9.61	1.44	0.88	0.14
29	16.58	0.14	0.19	0.22	1.12	0.46	3.82	13.52	10.57	41.87	9.24	1.50	0.66	0.12
30	17.93	0.17	0.22	0.25	1.30	0.54	3.80	19.59	13.40	29.41	10.75	1.75	0.76	0.13
31	18.37	0.14	0.18	0.21	1.06	0.44	3.66	18.87	12.91	32.99	8.76	1.43	0.87	0.12
32	23.77	0.16	0.22	0.25	1.29	0.53	3.76	15.60	11.84	29.14	10.66	1.73	0.90	0.14
33	22.13	0.16	0.21	0.24	1.24	0.51	4.24	16.22	12.76	31.09	8.67	1.67	0.73	0.13
34	23.65	0.16	0.22	0.25	1.29	0.53	3.76	15.58	12.19	29.11	10.65	1.73	0.76	0.12
35	21.28	0.13	0.17	0.19	0.99	0.41	3.43	14.20	12.09	35.26	9.70	1.34	0.69	0.11
36	18.79	0.17	0.22	0.26	1.32	0.55	4.01	19.86	11.39	29.82	10.90	1.77	0.81	0.12
37	24.39	0.16	0.21	0.25	1.26	0.52	4.49	15.87	11.32	29.65	9.17	1.70	0.88	0.12
38	21.44	0.13	0.17	0.20	1.00	0.42	4.04	14.31	12.18	35.53	8.27	1.35	0.82	0.13
39	20.41	0.14	0.19	0.21	1.10	0.46	3.62	13.28	9.47	41.13	7.67	1.48	0.73	0.12
40	19.68	0.14	0.18	0.21	1.08	0.45	3.15	13.08	10.83	40.51	8.48	1.45	0.64	0.12
41	21.57	0.13	0.17	0.20	1.01	0.42	3.47	14.40	12.26	35.75	8.32	1.36	0.83	0.11
42	21.25	0.16	0.22	0.25	1.30	0.54	4.44	15.71	13.37	29.36	10.74	1.75	0.76	0.14
43	19.99	0.15	0.19	0.22	1.14	0.47	4.61	16.30	12.75	30.45	11.13	1.54	0.94	0.13
44	21.73	0.14	0.19	0.22	1.13	0.47	4.54	16.06	13.67	30.01	9.28	1.52	0.92	0.12
45	22.11	0.13	0.17	0.20	1.01	0.42	3.47	14.39	10.26	35.72	9.83	1.36	0.83	0.11
46	16.89	0.15	0.20	0.23	1.15	0.48	3.46	17.86	10.24	38.56	8.29	1.56	0.83	0.11
47	18.68	0.14	0.19	0.21	1.10	0.46	3.21	13.33	11.35	41.28	7.70	1.48	0.75	0.12
48	22.15	0.14	0.18	0.21	1.07	0.45	3.69	18.12	13.03	28.60	10.06	1.45	0.74	0.12
49	19.10	0.12	0.15	0.18	0.91	0.38	3.14	16.19	9.28	40.28	8.20	1.23	0.75	0.12
50	22.33	0.14	0.19	0.21	1.10	0.46	4.26	15.66	13.33	29.26	10.70	1.48	0.76	0.12
51	21.46	0.16	0.22	0.25	1.29	0.54	4.42	19.47	11.16	29.23	9.04	1.74	0.88	0.14
52	23.65	0.16	0.22	0.25	1.29	0.53	3.76	15.58	12.19	29.11	10.65	1.73	0.76	0.12
53	22.44	0.14	0.18	0.21	1.08	0.45	3.73	19.28	12.12	28.94	8.95	1.46	0.89	0.12
54	17.51	0.14	0.19	0.22	1.11	0.46	3.24	14.28	9.60	41.67	9.19	1.50	0.77	0.12
55	16.84	0.12	0.16	0.18	0.94	0.39	3.79	13.41	11.41	41.53	9.16	1.27	0.69	0.12
56	19.80	0.12	0.16	0.18	0.92	0.38	3.72	14.79	9.40	40.80	7.61	1.24	0.76	0.12
57	15.22	0.12	0.16	0.18	0.94	0.39	3.36	16.63	11.38	41.40	8.08	1.26	0.77	0.12
58	16.32	0.13	0.17	0.19	0.99	0.41	3.21	16.57	9.49	41.22	9.09	1.33	0.77	0.12
59	19.56	0.12	0.15	0.18	0.91	0.38	3.15	13.05	11.11	40.43	8.92	1.23	0.69	0.12
60	20.23	0.12	0.16	0.18	0.93	0.39	3.35	14.12	9.49	41.22	7.69	1.26	0.77	0.10
61	22.44	0.16	0.21	0.24	1.23	0.51	3.79	19.02	10.90	28.55	10.44	1.66	0.74	0.13
62	19.02	0.12	0.15	0.18	0.91	0.38	3.66	16.12	9.24	40.11	8.17	1.22	0.63	0.12
63	19.68	0.13	0.17	0.20	1.00	0.42	3.71	13.14	11.18	40.68	7.59	1.35	0.64	0.12
64	22.14	0.16	0.21	0.25	1.27	0.53	3.70	18.50	13.08	28.70	8.88	1.71	0.75	0.13
65	21.15	0.13	0.17	0.19	0.99	0.41	3.99	14.12	12.02	35.05	9.64	1.33	0.69	0.13
66	21.36	0.14	0.19	0.22	1.11	0.46	4.46	15.79	13.44	29.50	10.79	1.49	0.91	0.14
67	18.82	0.14	0.19	0.22	1.12	0.47	3.87	18.01	13.66	29.98	10.96	1.51	0.92	0.12
68	15.40	0.14	0.19	0.22	1.10	0.46	3.23	16.12	11.40	41.46	7.97	1.49	0.71	0.12
69	15.12	0.12	0.16	0.18	0.93	0.39	3.75	16.53	11.30	41.13	8.37	1.25	0.64	0.12
70	22.70	0.16	0.22	0.25	1.28	0.53	4.21	19.24	11.02	28.88	8.93	1.72	0.75	0.12
71	16.02	0.12	0.16	0.18	0.94	0.39	3.39	14.29	11.46	41.69	9.20	1.27	0.77	0.12
72	20.85	0.14	0.18	0.21	1.07	0.44	4.12	18.15	12.41	31.73	8.43	1.44	0.71	0.13
73	23.45	0.16	0.21	0.25	1.26	0.52	4.31	15.26	10.88	31.63	9.35	1.70	0.88	0.14
74	18.13	0.12	0.16	0.18	0.95	0.39	3.82	13.51	9.63	41.83	9.23	1.27	0.68	0.10
75	18.83	0.14	0.19	0.22	1.12	0.46	3.83	13.54	9.66	41.93	7.82	1.51	0.66	0.10

76	16.95	0.14	0.19	0.22	1.11	0.46	3.81	13.50	11.49	41.80	8.04	1.50	0.68	0.12
77	19.20	0.12	0.16	0.18	0.92	0.38	3.42	16.27	9.33	40.50	7.56	1.23	0.63	0.10
78	19.80	0.12	0.16	0.18	0.92	0.38	3.16	13.92	9.81	40.62	8.96	1.24	0.64	0.10
79	17.78	0.16	0.22	0.25	1.28	0.53	4.37	18.61	13.16	32.05	8.93	1.72	0.82	0.14
80	15.66	0.12	0.16	0.18	0.95	0.39	3.26	16.83	10.58	41.89	7.82	1.28	0.78	0.10
81	19.31	0.14	0.19	0.21	1.10	0.46	3.78	16.20	11.29	35.84	9.10	1.48	0.77	0.12
82	18.59	0.16	0.22	0.25	1.27	0.53	3.71	15.40	13.11	33.51	10.52	1.71	0.89	0.13
83	16.69	0.14	0.19	0.22	1.13	0.47	3.77	13.68	11.01	42.36	7.90	1.52	0.79	0.12
84	18.73	0.15	0.21	0.24	1.21	0.50	4.50	19.81	11.35	29.73	10.87	1.64	0.92	0.14
85	20.06	0.14	0.19	0.21	1.09	0.45	4.41	19.44	13.29	29.18	9.02	1.47	0.90	0.14
86	21.56	0.14	0.18	0.21	1.08	0.45	4.37	19.23	11.02	28.87	10.56	1.46	0.75	0.12
87	24.11	0.14	0.19	0.21	1.10	0.46	4.43	15.69	11.19	29.32	10.72	1.48	0.83	0.12
88	19.55	0.17	0.23	0.26	1.34	0.56	4.58	16.20	13.78	30.26	10.21	1.80	0.93	0.15
89	22.42	0.15	0.21	0.24	1.21	0.50	4.31	19.00	10.89	28.52	10.03	1.63	0.74	0.14
90	16.75	0.12	0.16	0.19	0.96	0.40	3.29	13.66	10.68	42.29	9.33	1.29	0.79	0.11
91	15.73	0.13	0.17	0.19	0.99	0.41	3.81	16.77	10.07	41.73	7.79	1.33	0.78	0.12
92	16.44	0.13	0.17	0.20	1.01	0.42	3.48	17.97	12.29	35.85	9.86	1.36	0.70	0.11
93	22.32	0.13	0.17	0.20	1.02	0.42	4.10	14.52	10.36	36.05	8.39	1.37	0.84	0.11
94	15.88	0.12	0.16	0.18	0.95	0.39	3.83	15.19	11.52	41.92	7.82	1.28	0.66	0.10
95	20.23	0.14	0.18	0.21	1.06	0.44	3.19	13.22	9.73	40.95	8.34	1.43	0.76	0.12
96	17.56	0.14	0.19	0.21	1.10	0.46	3.20	16.52	9.47	41.12	7.67	1.48	0.76	0.12
97	20.68	0.14	0.18	0.21	1.07	0.44	4.15	18.97	12.97	28.47	10.41	1.44	0.74	0.14
98	18.43	0.17	0.22	0.26	1.32	0.55	4.41	19.92	12.89	29.90	9.25	1.78	0.78	0.12
99	18.00	0.17	0.22	0.25	1.30	0.54	3.81	19.68	13.46	29.54	10.25	1.76	0.89	0.14
100	15.71	0.14	0.19	0.21	1.10	0.46	3.36	16.61	9.52	41.33	9.12	1.48	0.65	0.12
101	15.73	0.13	0.17	0.19	0.99	0.41	3.81	16.77	10.07	41.73	7.79	1.33	0.78	0.12
102	16.35	0.12	0.16	0.18	0.95	0.39	3.55	13.54	11.53	41.94	9.25	1.28	0.66	0.10
103	19.96	0.13	0.18	0.20	1.05	0.44	3.19	13.24	10.81	40.99	7.65	1.42	0.64	0.10
104	15.97	0.14	0.19	0.22	1.11	0.46	3.24	14.25	11.43	41.57	9.17	1.49	0.65	0.12
105	19.13	0.16	0.22	0.25	1.28	0.53	4.37	15.48	11.57	33.68	10.58	1.72	0.89	0.14
106	19.48	0.17	0.23	0.26	1.33	0.55	4.23	16.14	13.74	30.15	11.03	1.79	0.78	0.12
107	19.14	0.12	0.16	0.18	0.92	0.38	3.15	14.67	11.13	40.49	7.56	1.23	0.75	0.12
108	16.53	0.12	0.16	0.19	0.96	0.40	3.78	13.69	11.66	42.41	7.91	1.29	0.79	0.11
109	20.26	0.12	0.16	0.18	0.92	0.38	3.45	13.18	9.40	40.82	9.01	1.24	0.76	0.12
110	18.66	0.14	0.19	0.22	1.13	0.47	4.55	20.05	12.60	30.10	9.31	1.52	0.93	0.12
111	19.76	0.17	0.23	0.26	1.35	0.56	4.40	16.37	13.94	30.59	9.46	1.82	0.94	0.13
112	19.99	0.15	0.19	0.22	1.14	0.47	4.61	16.30	12.75	30.45	11.13	1.54	0.94	0.13
113	16.03	0.15	0.19	0.22	1.14	0.47	3.34	17.23	10.83	40.05	8.00	1.54	0.69	0.11
114	17.02	0.14	0.19	0.22	1.14	0.47	3.32	13.77	10.30	42.66	8.32	1.53	0.79	0.11
115	15.91	0.13	0.17	0.20	1.03	0.43	3.82	16.82	9.64	41.87	7.81	1.39	0.66	0.10
116	22.33	0.16	0.21	0.25	1.26	0.52	3.68	17.12	12.98	28.50	10.42	1.70	0.74	0.14
117	18.46	0.12	0.16	0.18	0.94	0.39	3.25	16.06	10.99	39.97	7.46	1.27	0.63	0.12
118	17.95	0.14	0.19	0.22	1.10	0.46	4.13	19.62	13.42	29.46	10.77	1.49	0.91	0.14
119	22.48	0.16	0.22	0.25	1.30	0.54	3.79	18.90	11.20	29.34	9.07	1.75	0.88	0.12
120	21.56	0.16	0.22	0.25	1.28	0.53	3.73	19.25	12.45	28.89	8.94	1.72	0.89	0.14
121	17.03	0.12	0.16	0.19	0.96	0.40	3.86	13.67	9.75	42.33	9.34	1.29	0.79	0.11
122	20.19	0.14	0.19	0.22	1.12	0.47	3.84	16.08	9.69	37.91	7.85	1.51	0.66	0.12
123	16.39	0.12	0.16	0.18	0.93	0.39	3.29	16.51	11.29	41.08	7.67	1.25	0.64	0.10
124	20.12	0.12	0.16	0.18	0.93	0.39	3.75	13.26	10.37	41.07	7.66	1.25	0.64	0.10
125	18.32	0.14	0.19	0.22	1.11	0.46	4.48	18.76	13.49	29.61	10.83	1.50	0.77	0.13
126	20.64	0.16	0.22	0.25	1.29	0.53	3.76	19.42	11.13	29.15	10.66	1.74	0.90	0.14
127	15.72	0.12	0.16	0.18	0.94	0.39	3.77	16.62	9.53	41.36	9.12	1.26	0.71	0.12
128	16.47	0.14	0.19	0.22	1.13	0.47	3.67	13.65	11.62	42.27	7.89	1.52	0.66	0.11
129	17.68	0.14	0.19	0.22	1.11	0.46	3.26	13.51	10.10	41.84	9.23	1.50	0.66	0.10
130	15.96	0.12	0.16	0.18	0.94	0.39	3.70	16.17	9.58	41.59	9.18	1.27	0.65	0.11
131	23.45	0.16	0.21	0.25	1.26	0.52	4.21	15.26	10.88	31.63	9.62	1.70	0.74	0.12
132	16.27	0.14	0.19	0.21	1.10	0.46	3.77	16.08	10.45	41.36	7.72	1.49	0.65	0.12
133	22.22	0.17	0.22	0.26	1.31	0.54	4.48	15.86	13.50	29.63	9.16	1.76	0.77	0.13
134	23.68	0.16	0.22	0.25	1.28	0.53	4.39	16.15	11.07	29.00	10.60	1.73	0.80	0.14

EXTREME POINTS

Paint ID	Water	Soda Ash	Fungicide	Surfactant	Dispersant	Antifoamer	Solvent	Pigment	Extender	Emulsion	Organic Pigment	Coalescent	RM1	RM2
1	21.88	0.17	0.23	0.26	1.34	0.56	4.63	24.59	14.66	17.70	11.11	1.80	0.93	0.15
2	23.37	0.17	0.22	0.26	1.31	0.55	4.54	14.87	16.81	24.19	10.90	1.77	0.91	0.14
3	16.63	0.15	0.20	0.23	1.17	0.49	4.04	25.42	17.71	21.76	9.70	1.58	0.81	0.13
4	16.38	0.14	0.18	0.21	1.06	0.44	3.68	25.15	9.73	31.90	8.83	1.43	0.74	0.12
5	19.98	0.14	0.18	0.21	1.09	0.45	3.76	15.63	8.46	38.72	9.03	1.47	0.76	0.12
6	24.91	0.16	0.22	0.25	1.29	0.53	4.45	18.39	5.49	30.87	10.68	1.73	0.89	0.14
7	24.65	0.15	0.20	0.23	1.15	0.48	3.99	7.76	5.41	43.92	9.58	1.56	0.80	0.13
8	19.04	0.14	0.19	0.22	1.13	0.47	3.92	17.02	14.32	31.70	9.40	1.53	0.79	0.12
9	23.16	0.17	0.22	0.26	1.32	0.55	4.55	24.19	8.55	23.28	10.93	1.77	0.92	0.15
10	12.33	0.12	0.16	0.18	0.94	0.39	3.24	20.06	17.97	34.84	7.77	1.26	0.65	0.10
11	15.53	0.11	0.15	0.18	0.90	0.37	3.12	17.82	5.77	46.62	7.48	1.22	0.63	0.10
12	13.72	0.12	0.16	0.18	0.92	0.38	3.19	14.10	17.68	39.93	7.64	1.24	0.64	0.10
13	19.10	0.13	0.17	0.20	1.03	0.43	3.56	8.11	14.00	42.50	8.55	1.39	0.72	0.11
14	10.91	0.12	0.16	0.19	0.95	0.40	3.29	26.21	18.26	29.58	7.89	1.28	0.66	0.10
15	17.16	0.11	0.15	0.17	0.88	0.37	3.06	8.12	8.11	52.62	7.34	1.19	0.61	0.10
16	25.57	0.16	0.22	0.25	1.28	0.53	4.41	7.81	12.70	33.75	10.58	1.72	0.89	0.14
17	19.79	0.14	0.19	0.22	1.11	0.46	3.83	12.20	14.16	36.35	9.18	1.49	0.77	0.12
18	20.69	0.16	0.21	0.24	1.22	0.51	4.23	20.45	13.01	26.49	10.16	1.65	0.85	0.13
19	13.84	0.12	0.16	0.18	0.92	0.38	3.18	25.35	5.89	40.38	7.63	1.24	0.64	0.10
20	25.99	0.16	0.21	0.25	1.27	0.53	4.39	12.94	5.41	35.60	10.52	1.71	0.88	0.14
21	15.18	0.12	0.15	0.18	0.91	0.38	3.13	6.65	18.54	45.30	7.51	1.22	0.63	0.10
22	26.32	0.16	0.21	0.25	1.26	0.52	4.37	12.37	4.31	37.03	10.48	1.70	0.88	0.14
23	11.90	0.12	0.16	0.18	0.94	0.39	3.25	20.73	19.26	33.23	7.80	1.27	0.65	0.10
24	15.90	0.11	0.15	0.18	0.90	0.37	3.10	17.36	4.60	47.95	7.45	1.21	0.62	0.10
25	13.72	0.12	0.16	0.18	0.92	0.38	3.19	27.07	4.72	39.93	7.64	1.24	0.64	0.10
26	23.05	0.17	0.22	0.26	1.32	0.55	4.56	15.30	18.00	22.79	10.94	1.78	0.92	0.15
27	19.68	0.12	0.16	0.18	0.92	0.38	3.20	6.36	4.43	54.91	7.67	1.25	0.64	0.10
28	23.68	0.17	0.22	0.25	1.31	0.54	4.52	25.63	4.46	25.55	10.85	1.76	0.91	0.14
29	19.34	0.17	0.22	0.25	1.31	0.54	4.52	26.83	18.69	14.48	10.84	1.76	0.91	0.14
30	24.91	0.16	0.22	0.25	1.29	0.53	4.45	6.30	17.57	30.87	10.68	1.73	0.89	0.14

10.2 PAINT PROPERTIES

Paint properties are given on the CD at the end of the thesis Appendix C – Raw Data/Paint Properties.

10.3 RHEOLOGY DATA

Paint properties are given on the CD at the end of the thesis Appendix C – Raw Data/Rheology.

11 APPENDIX D: RESULTS -

DATA PREPARATION

(SCREENING)

11.1 INDIRECT MODELLING OF RHEOLOGICAL DATA

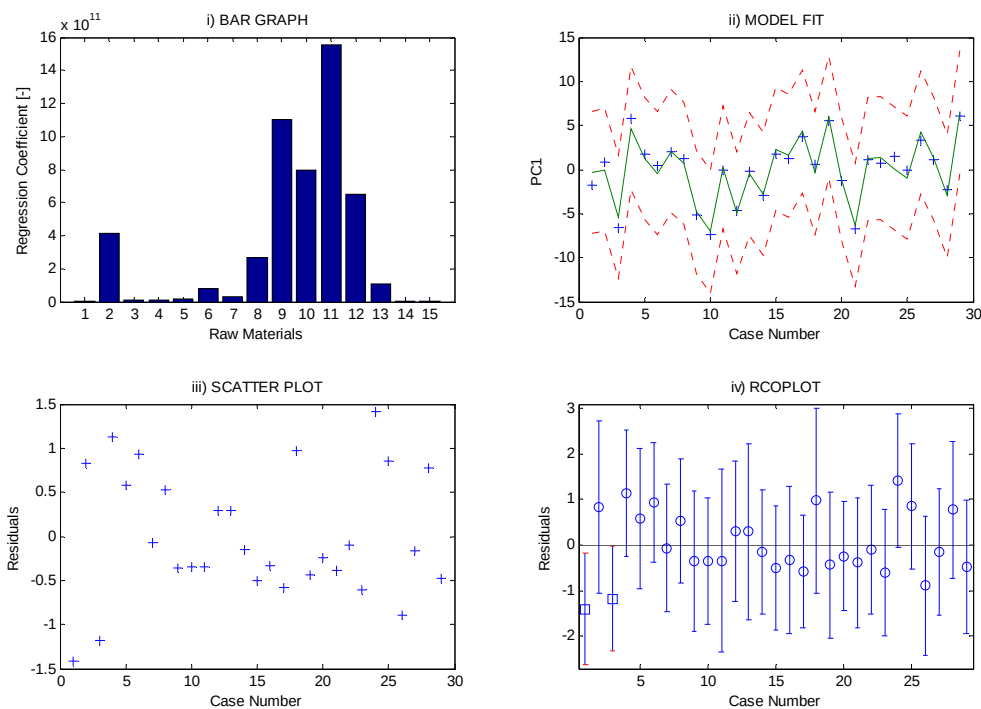
Indirect model parameters (α , β and γ) for the rheology curves of the 29 paint samples of the screening data is presented in electronic format on the CD at the end of the thesis (Indirect Modelling/Screening)

11.2 RAW MATERIAL SCREENING

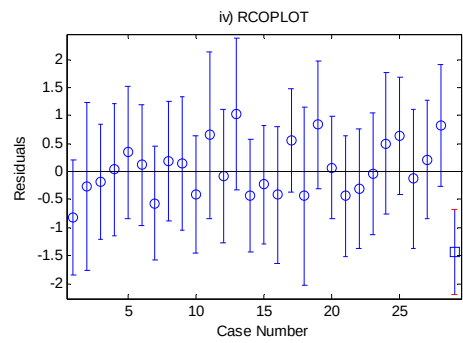
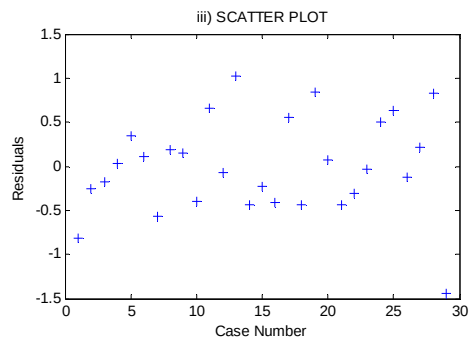
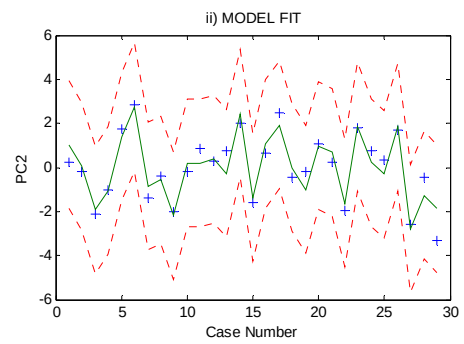
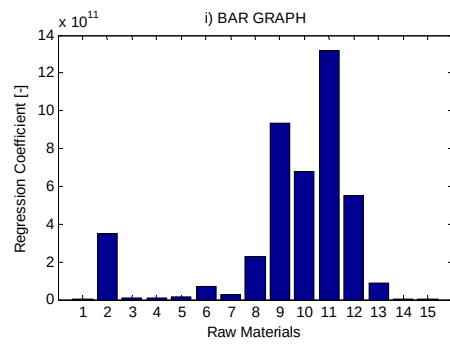
The x-axis numbering corresponds to the raw materials as follows:

1 – constant, 2 – water, 3 – soda ash, 4 – fungicide, 5 – surfactant, 6 – dispersant, 7 – antifoamer, 8 – solvent, 9 – pigment, 10 – extender, 11 – emulsion, 12 – organic pigment, 13 – coalescent, 14 – rheology modifier 1, 15 – rheology modifier 2

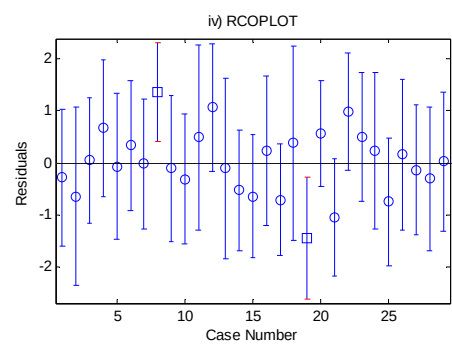
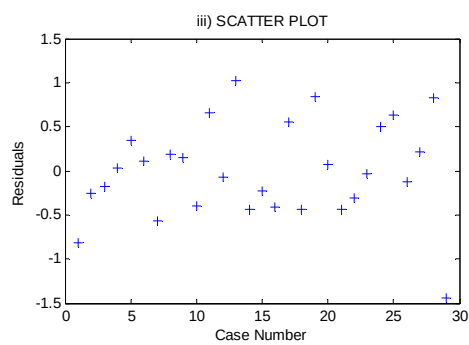
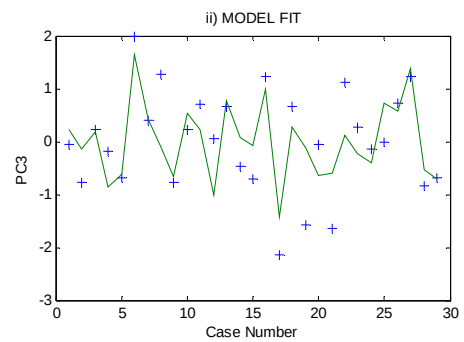
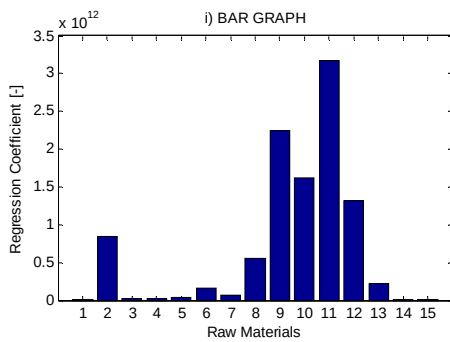
PC1



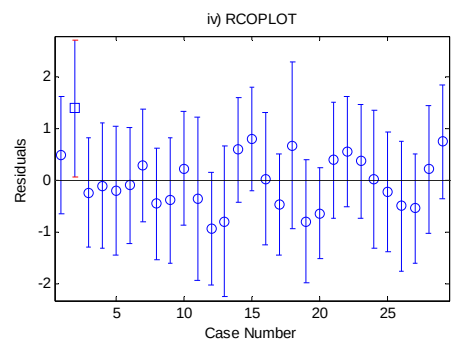
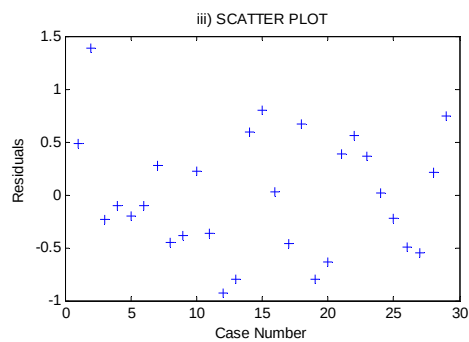
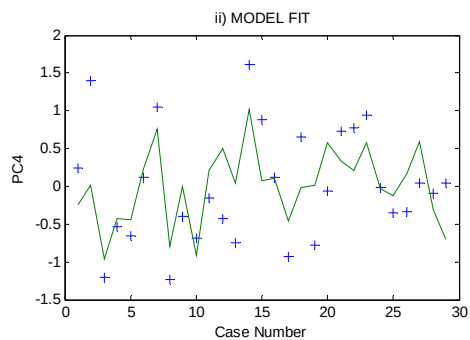
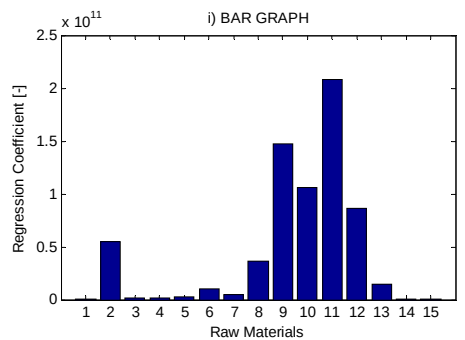
PC2



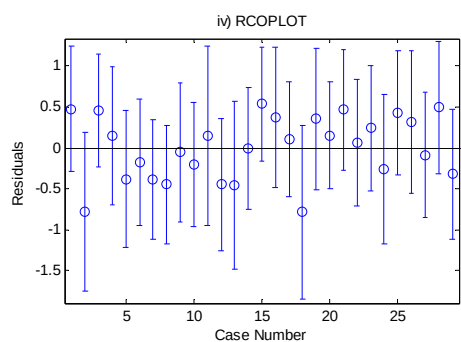
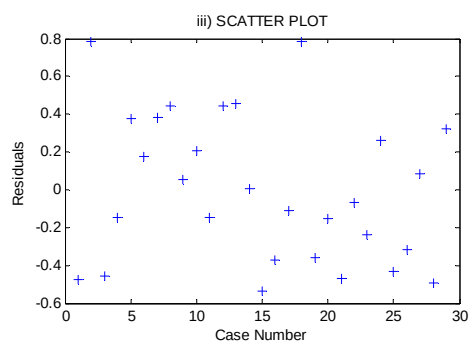
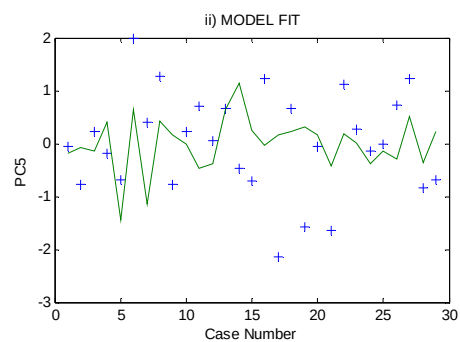
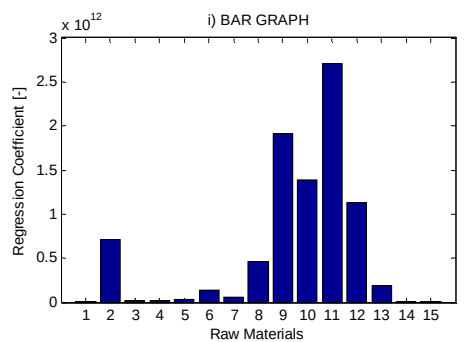
PC3



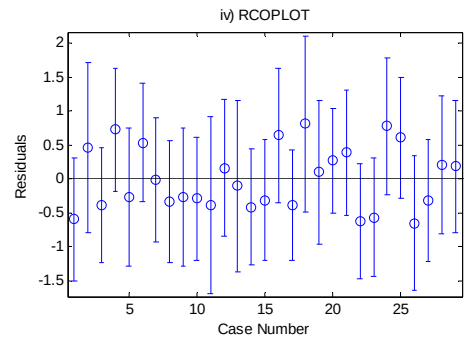
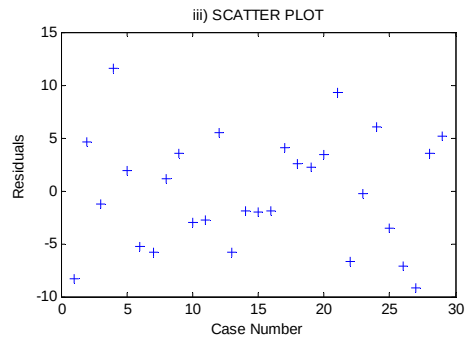
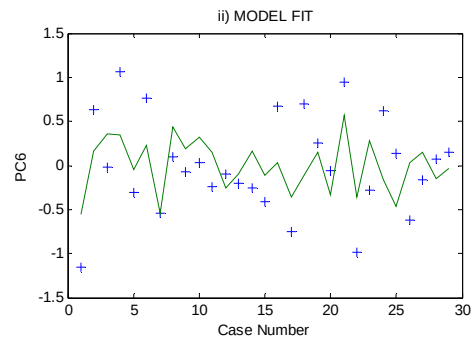
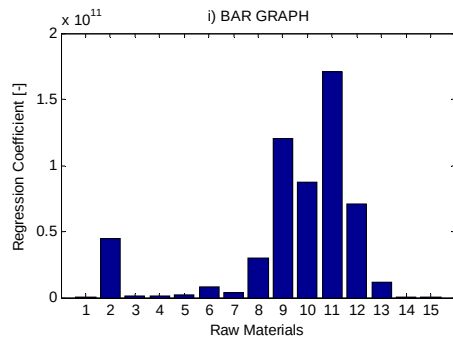
PC4



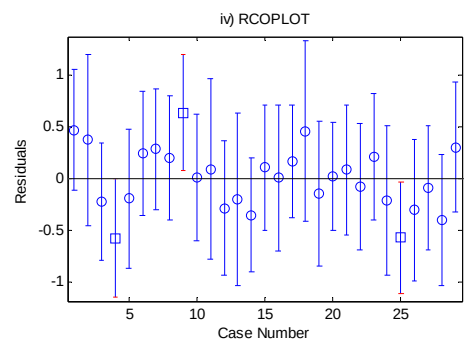
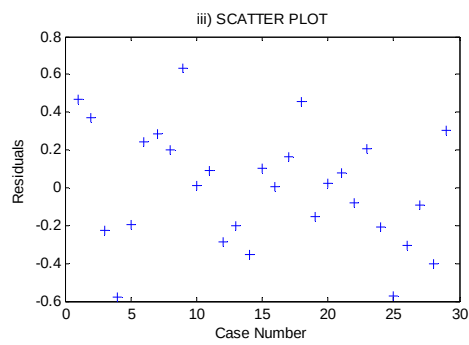
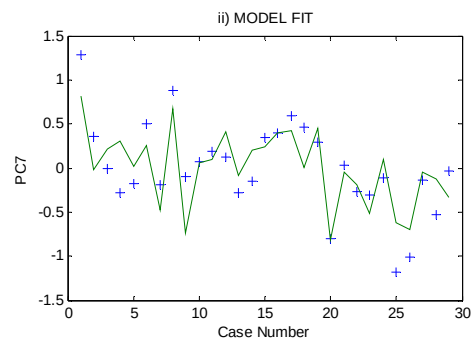
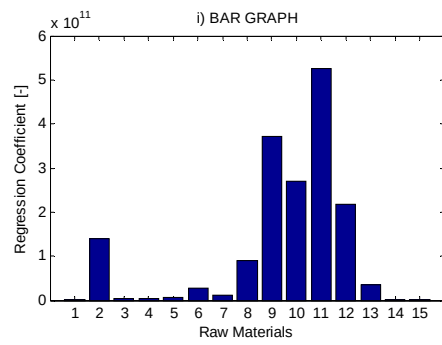
PC5



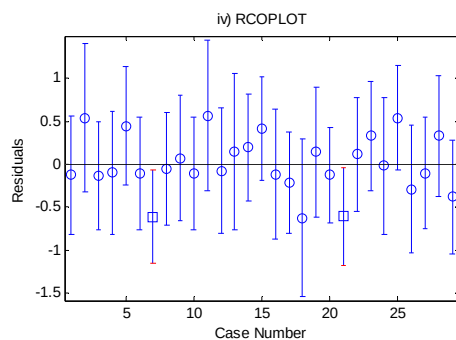
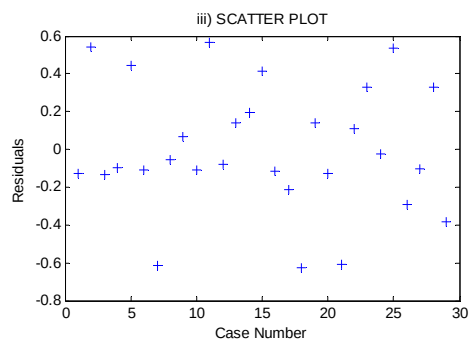
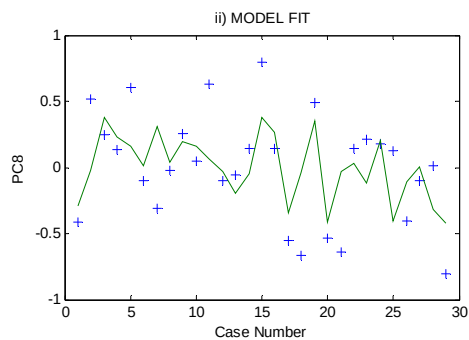
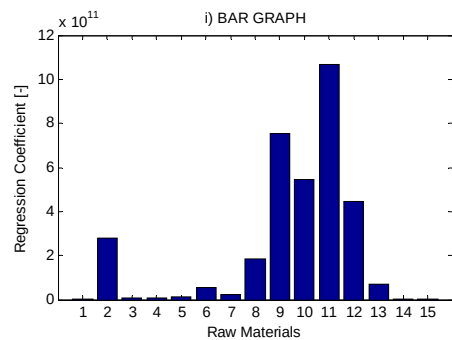
PC6



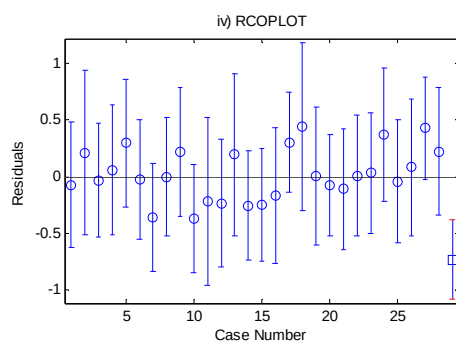
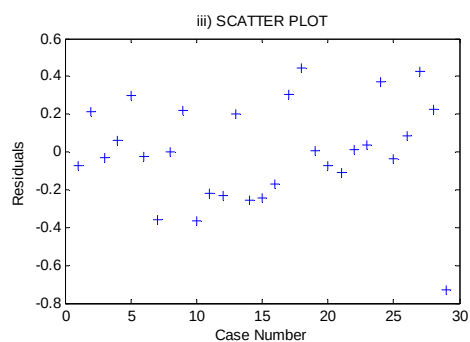
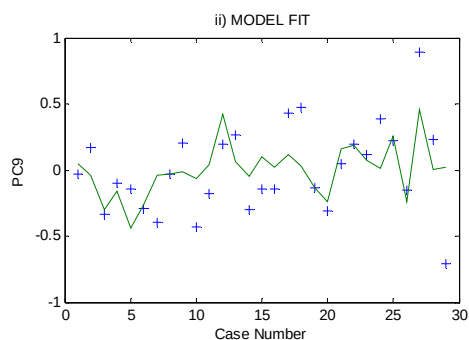
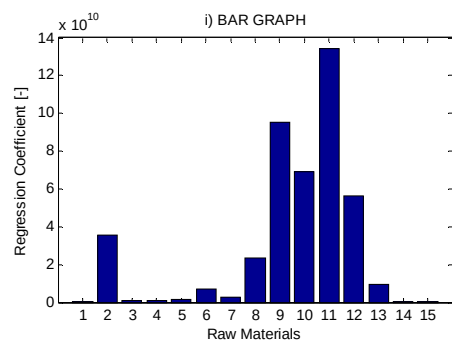
PC7



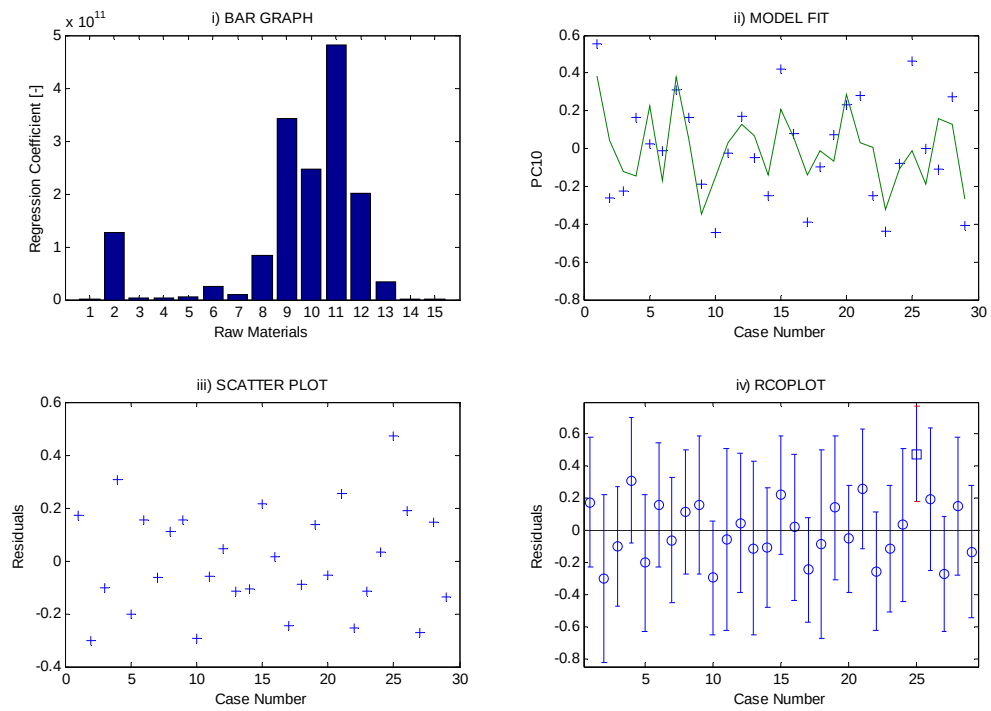
PC8



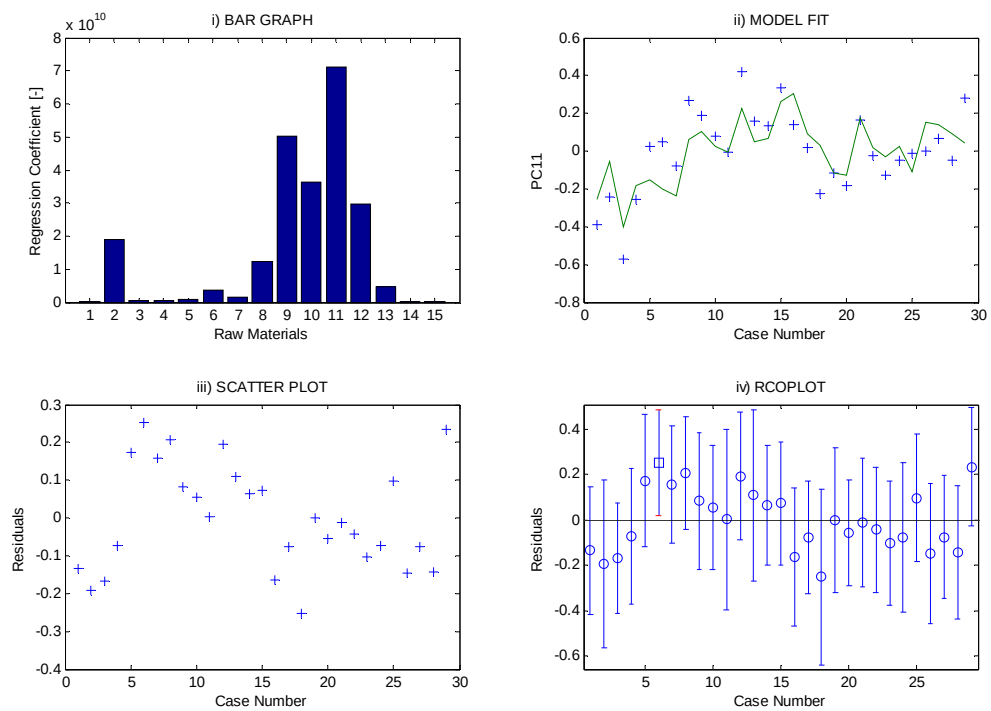
PC9



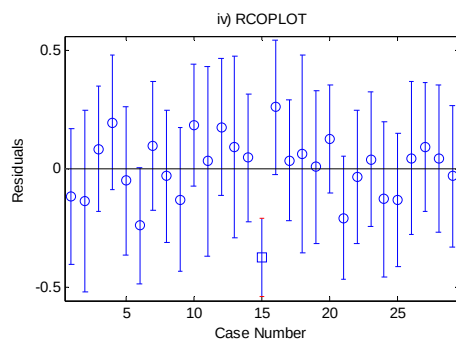
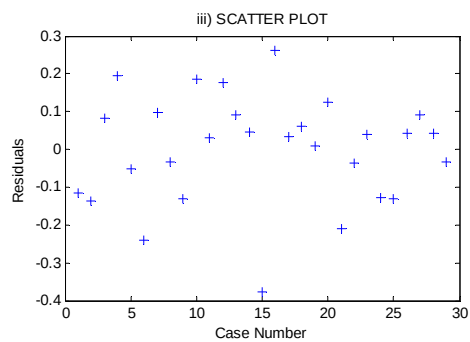
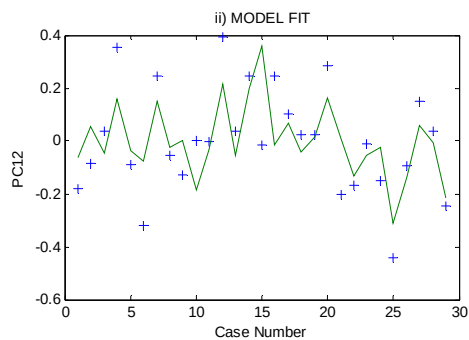
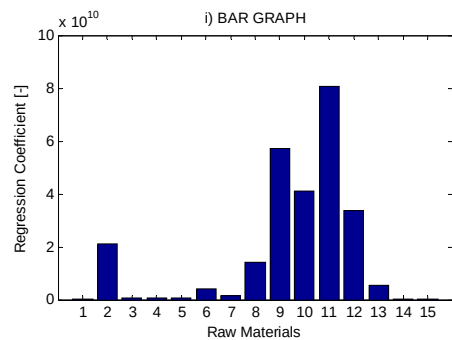
PC10



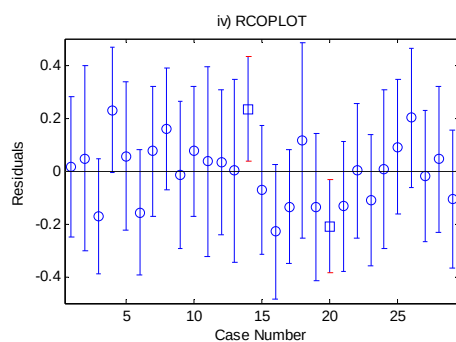
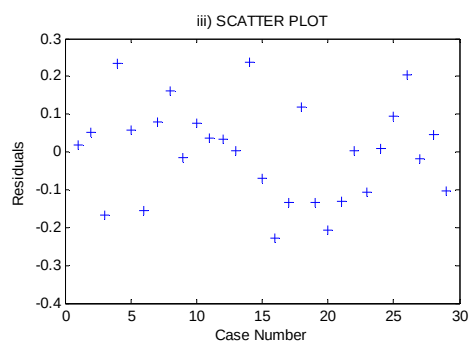
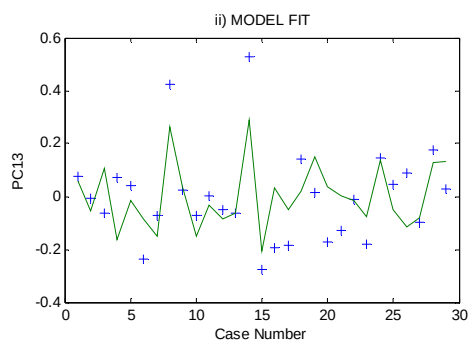
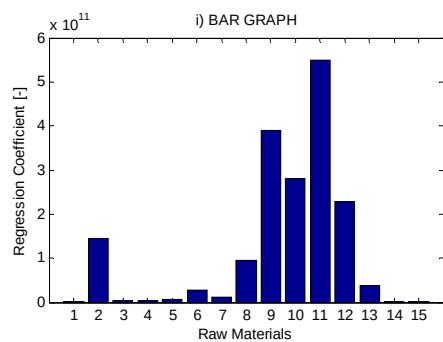
PC11



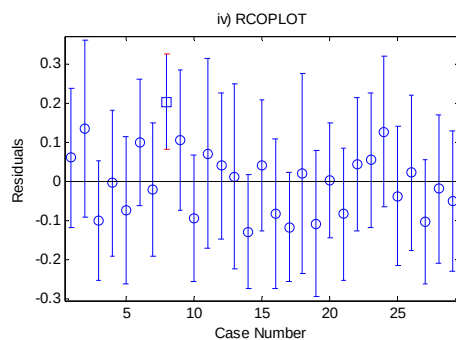
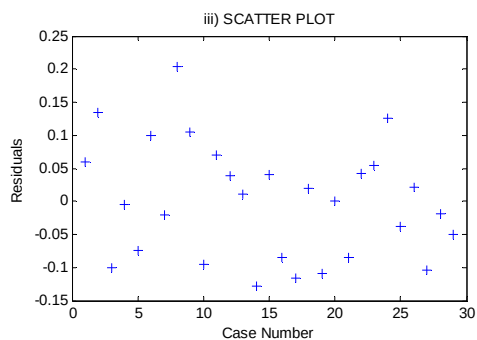
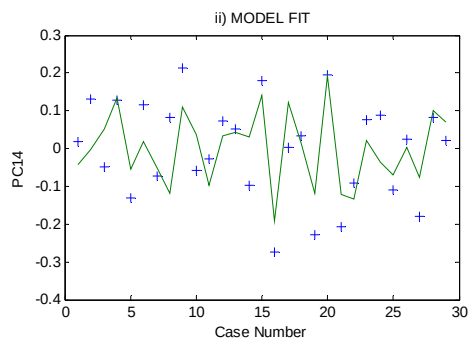
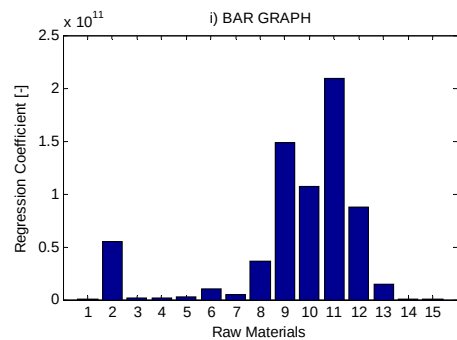
PC12



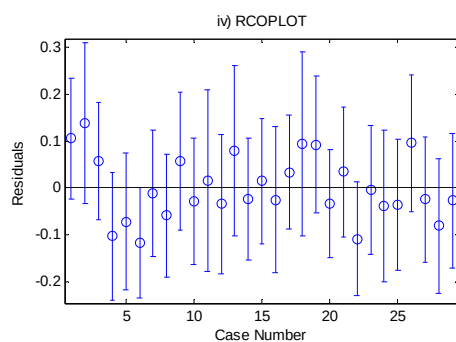
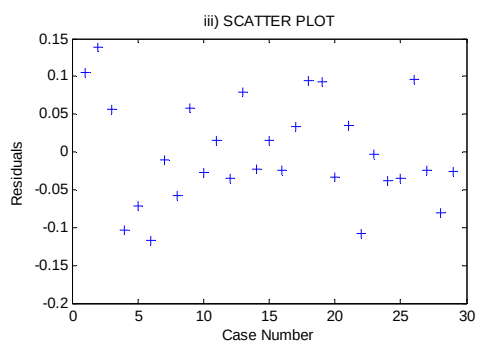
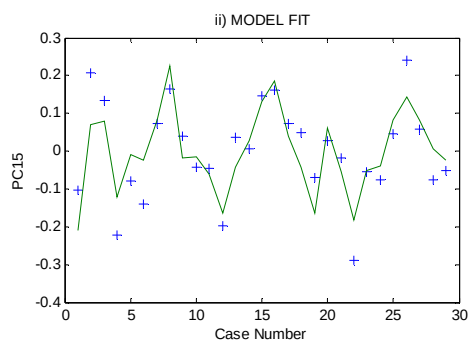
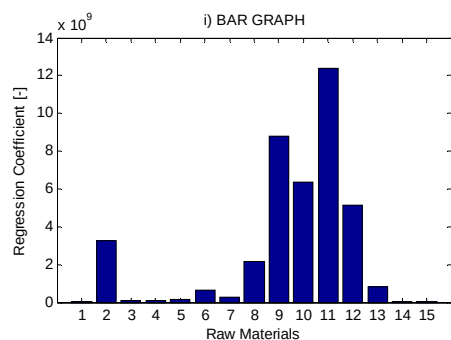
PC13



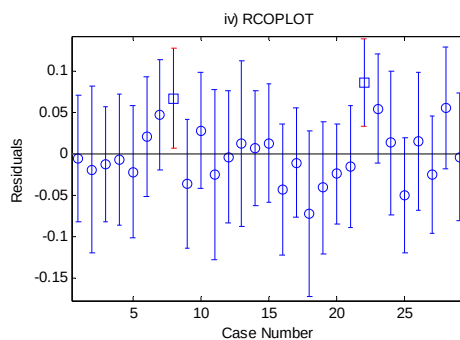
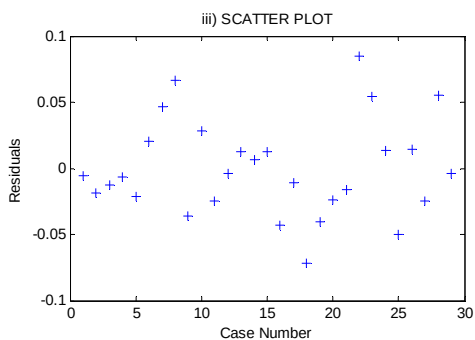
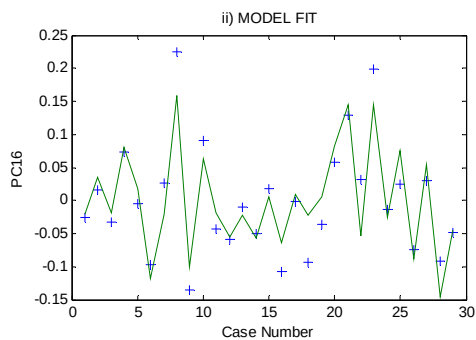
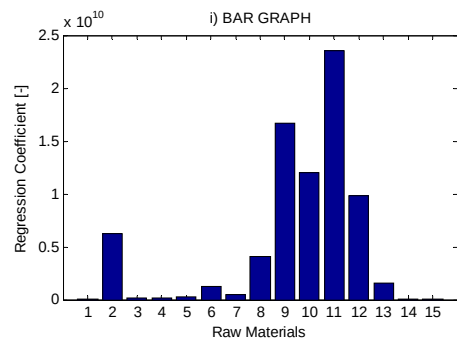
PC14



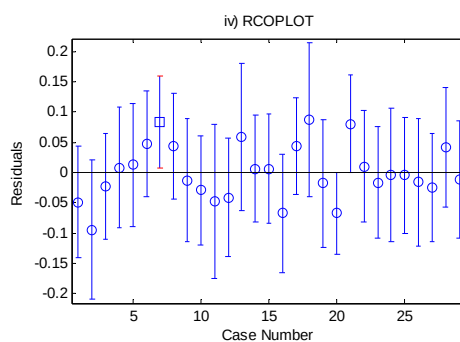
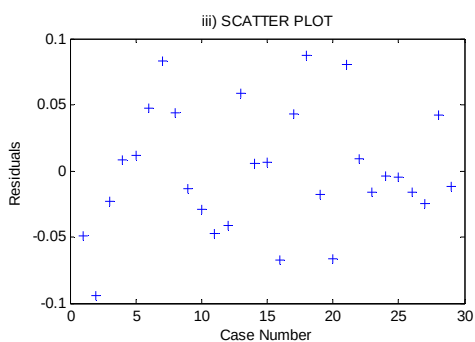
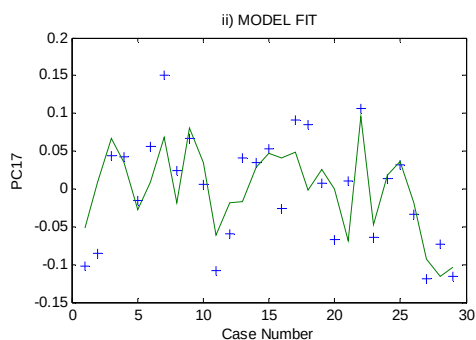
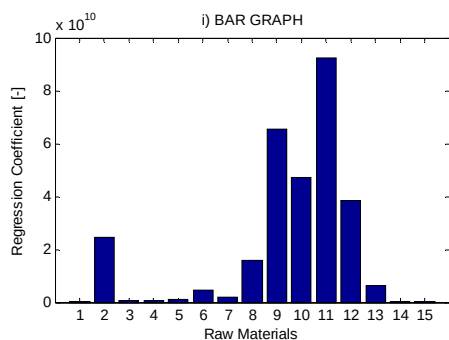
PC15



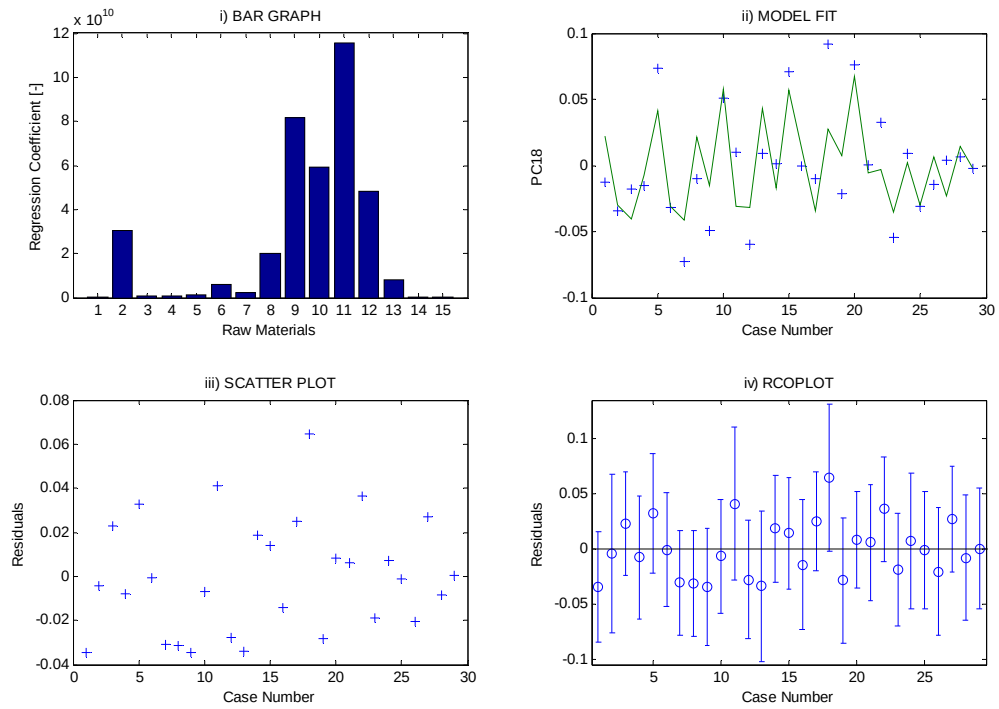
PC16



PC17



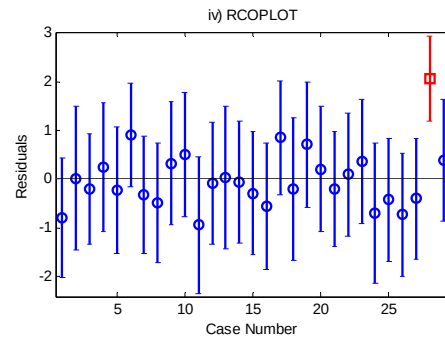
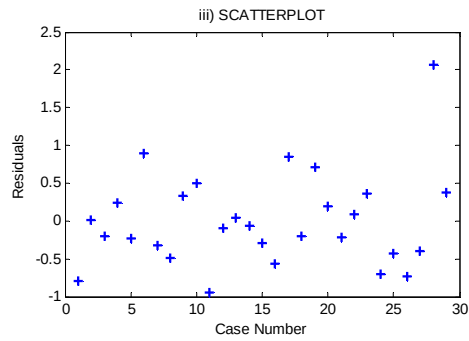
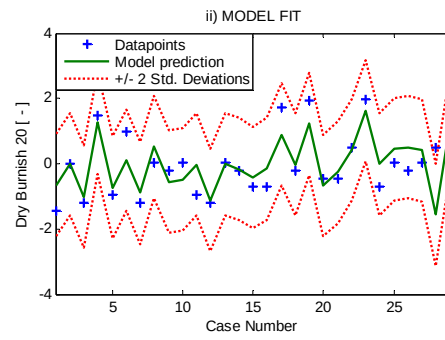
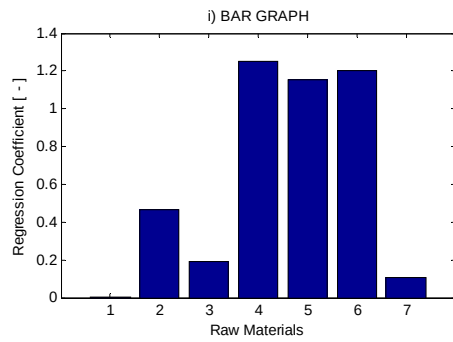
PC18



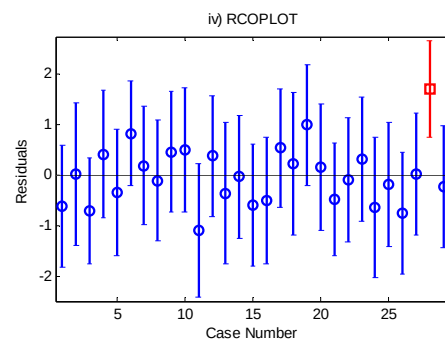
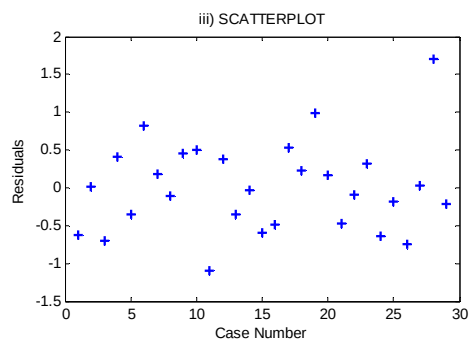
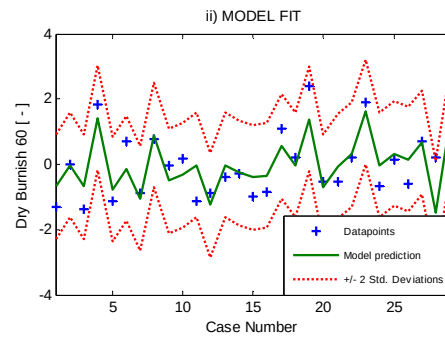
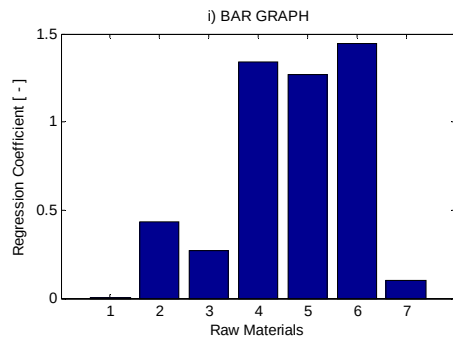
11.3 PAINT PROPERTIES SCREENING

The following are results obtained from the MLR process for the paint property screening. The graphs occur in the following order: i) Bar plot of regression coefficient values of raw materials (1=Constant, 2=Water, 3=Solvent, 4=Pigment, 5=Extender, 6=Emulsion, 7=Organic Pigment) ii) Model fit of data points iii) Scatterplot of residuals (standardised) and iv) RCO PLOT (standardised).

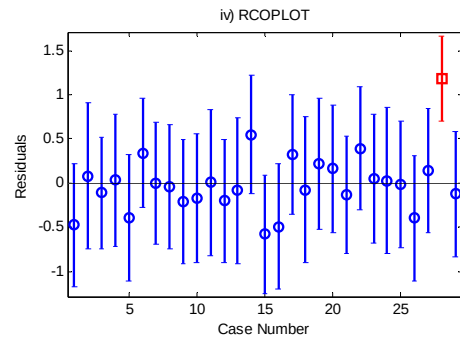
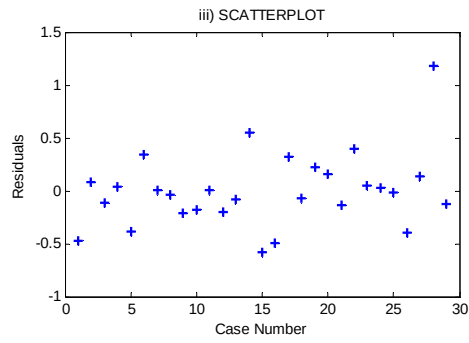
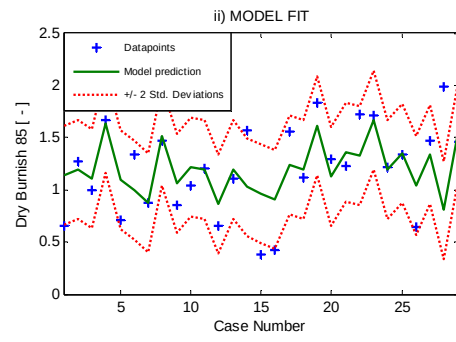
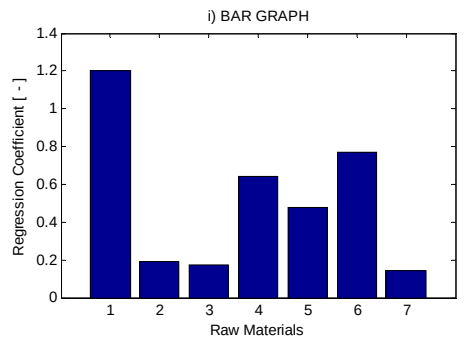
Dry Burnish (20)



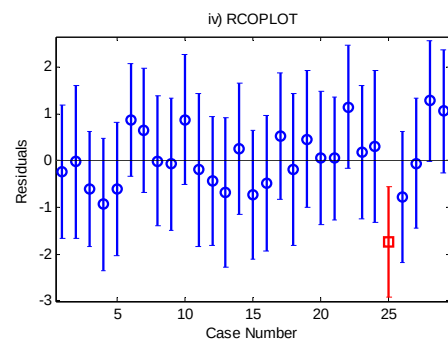
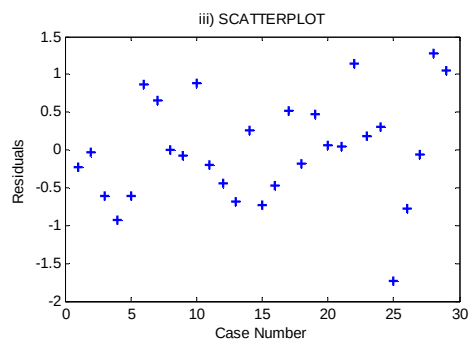
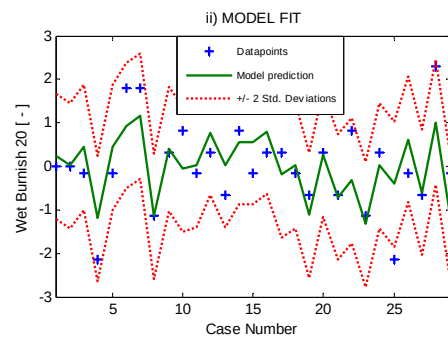
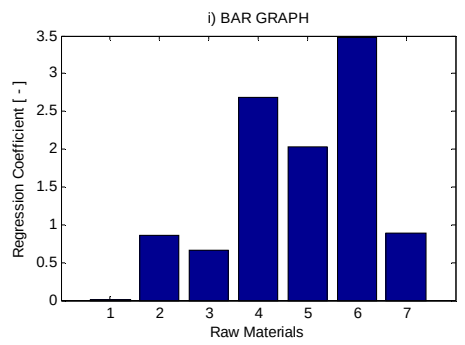
Dry Burnish (60)



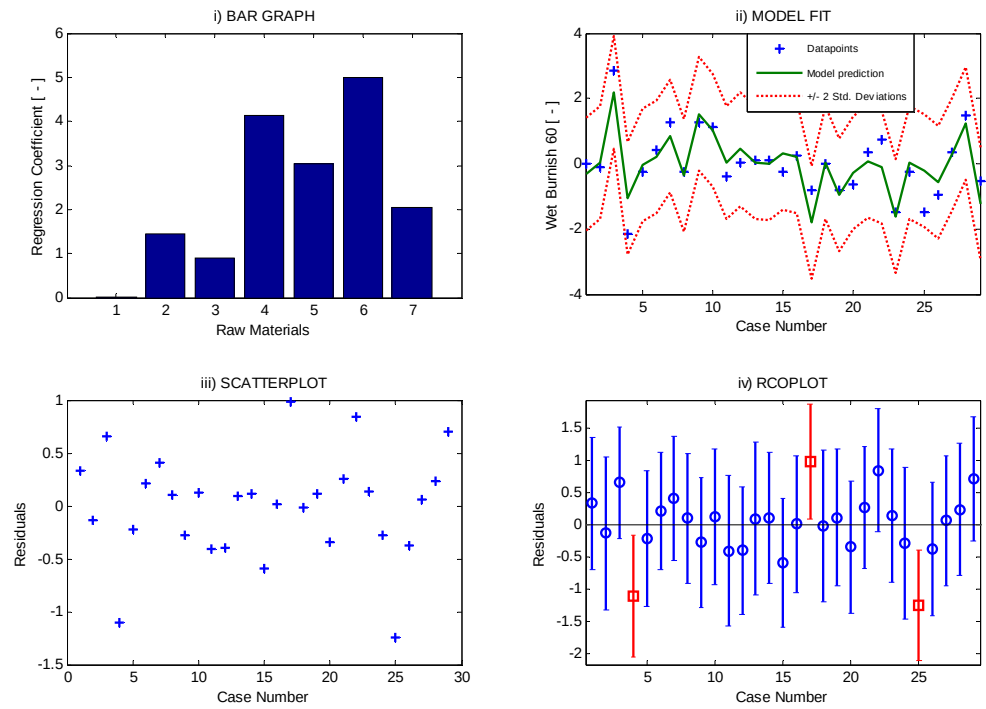
Dry Burnish (85)



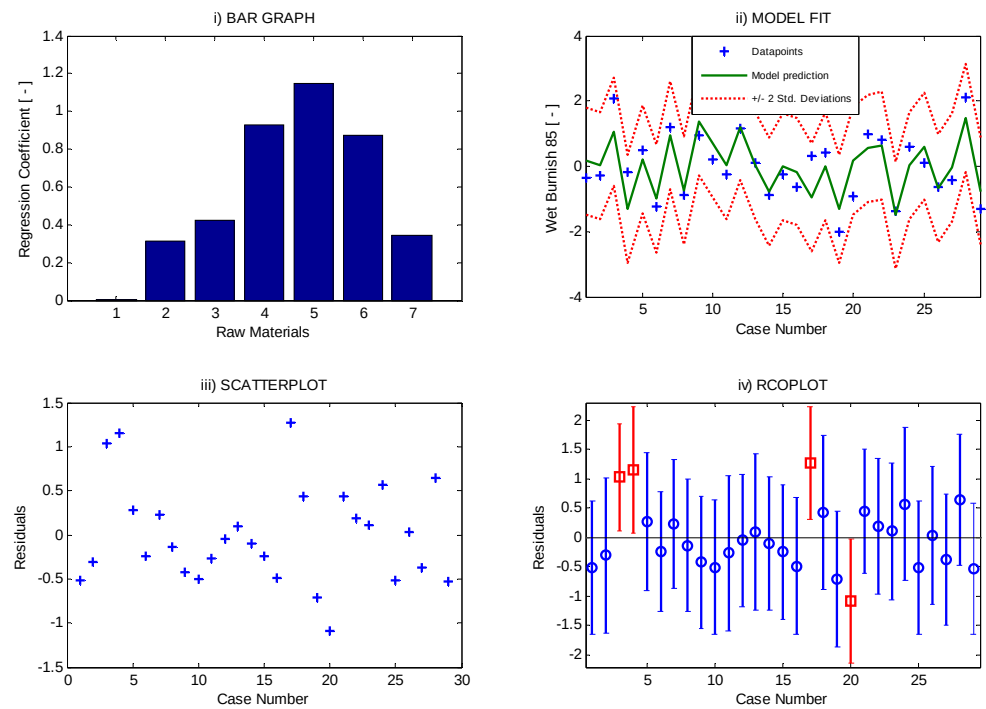
Wet Burnish (20)



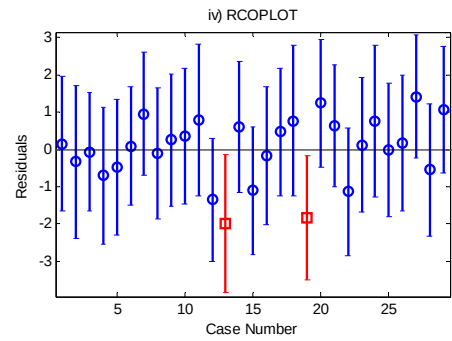
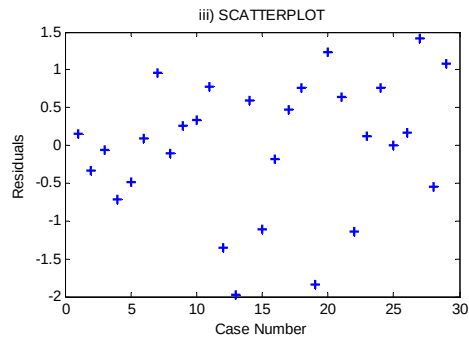
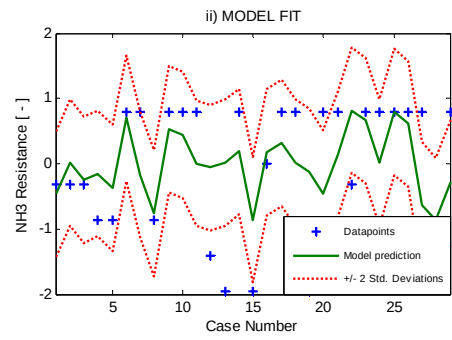
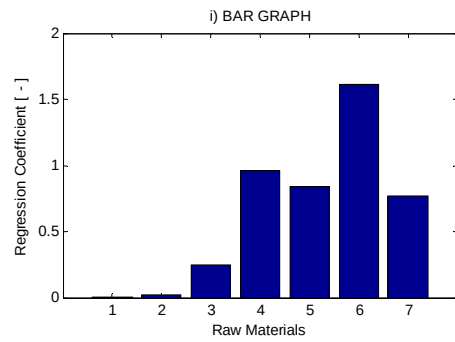
Wet Burnish (60)



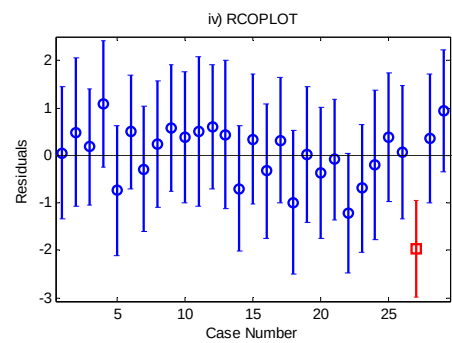
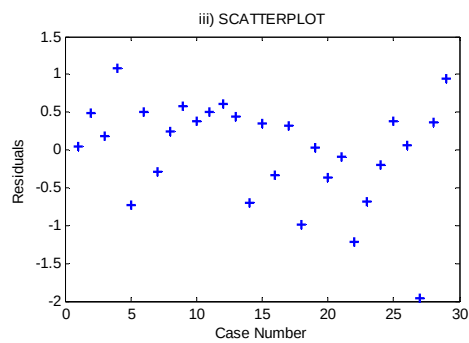
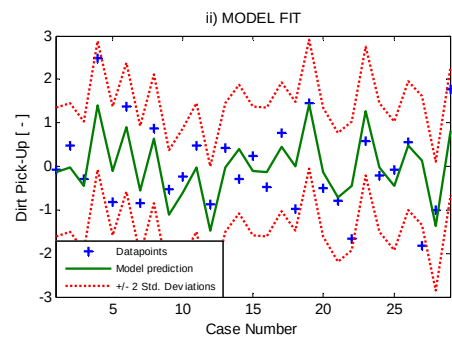
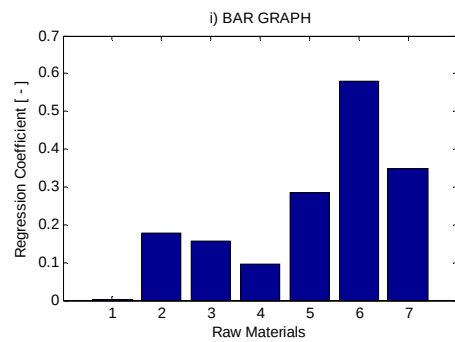
Wet Burnish (85)



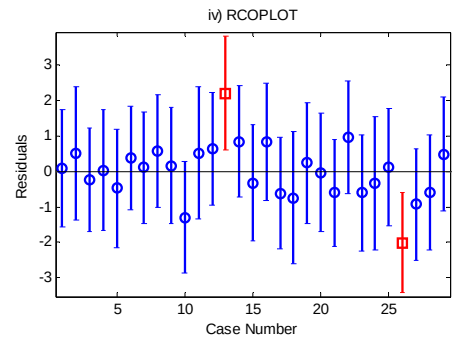
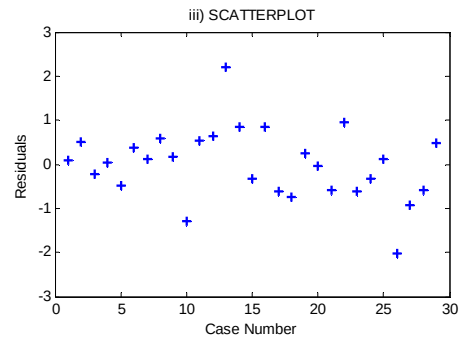
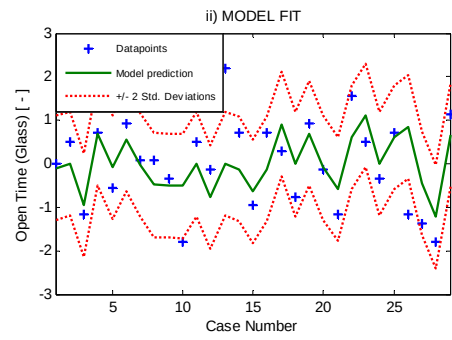
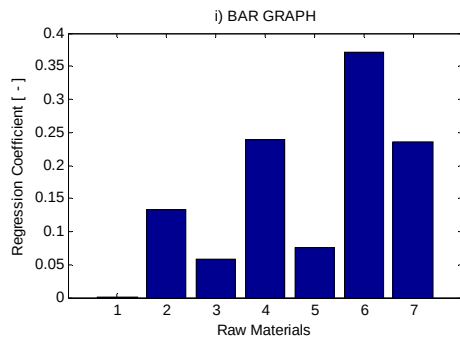
Chemical Resistance Blistering NH3



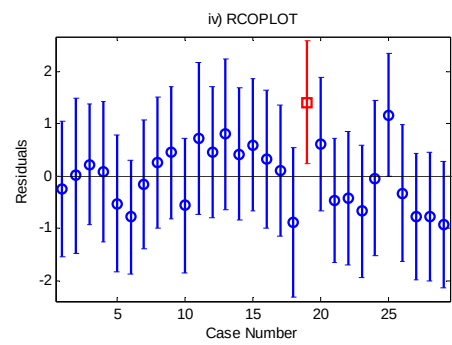
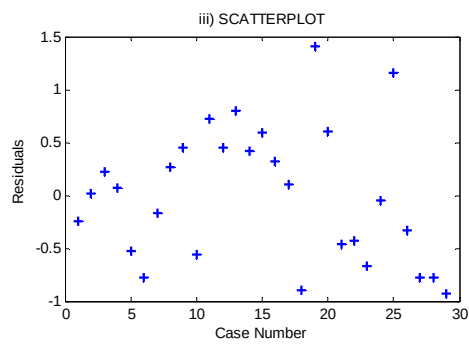
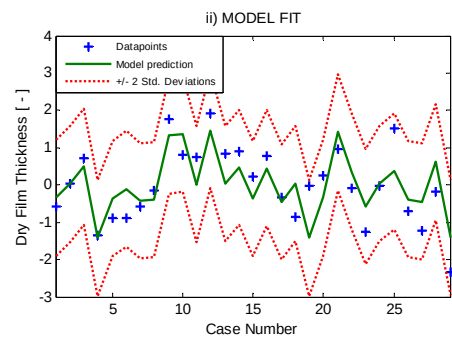
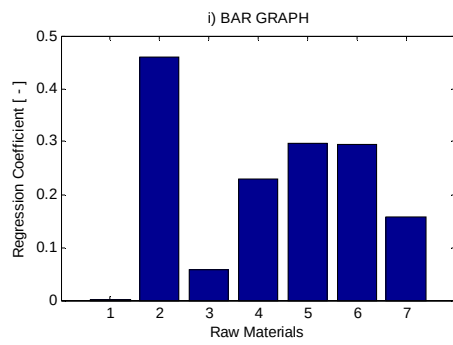
Dirt Pick-up (delta gloss 85)



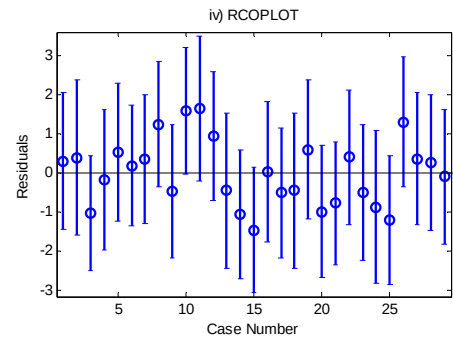
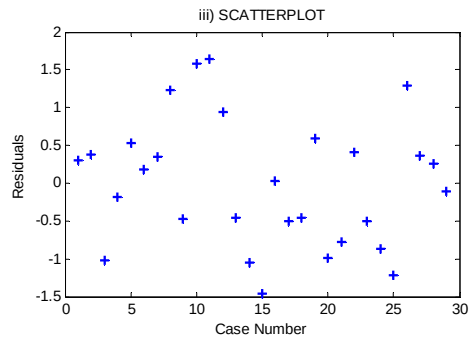
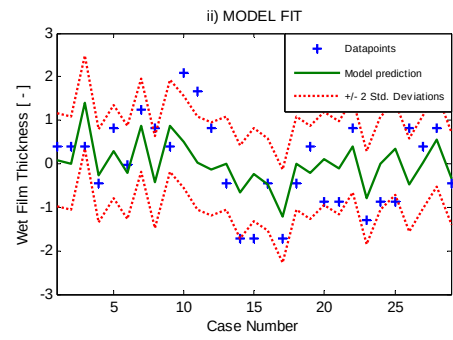
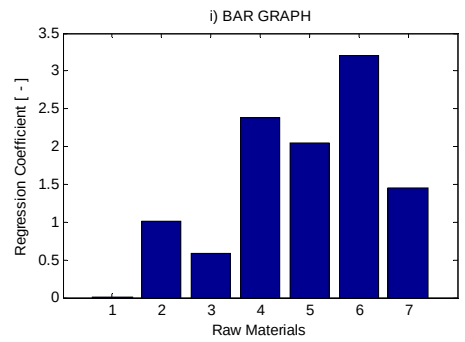
Open Time – Glass



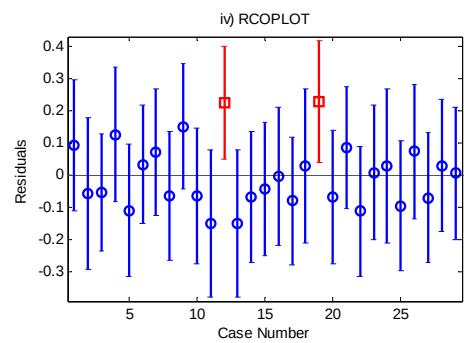
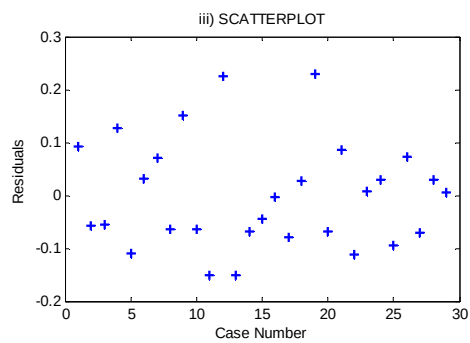
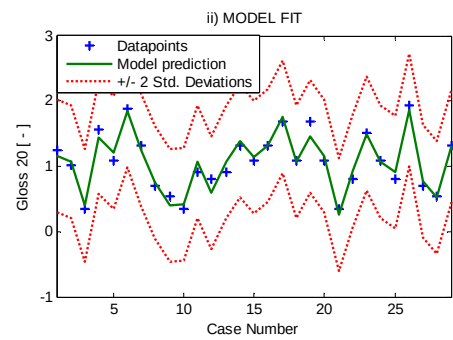
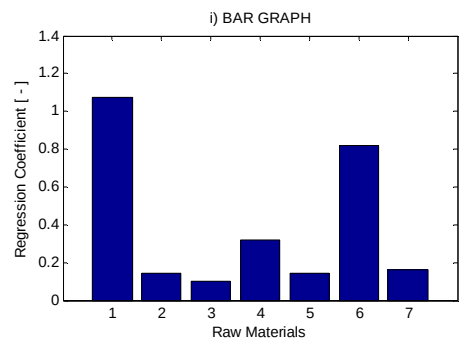
Dry Film Thickness



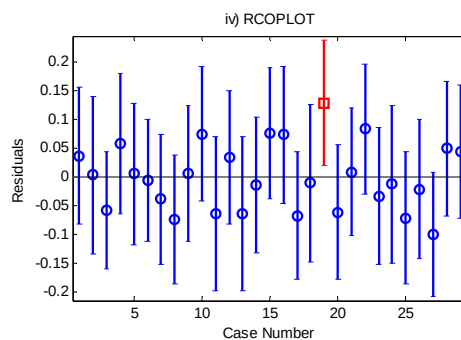
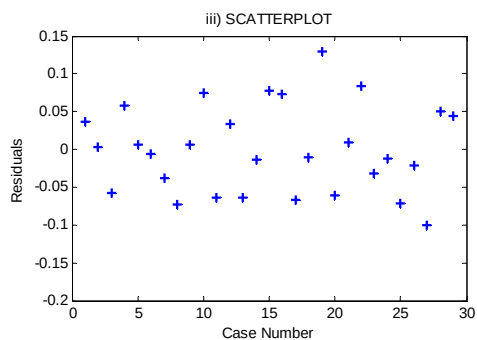
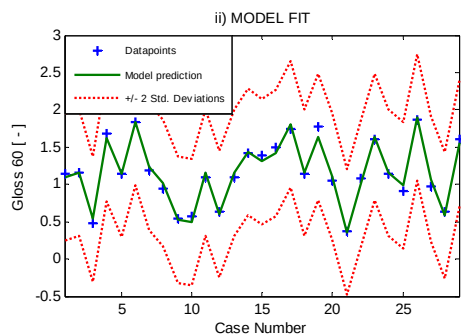
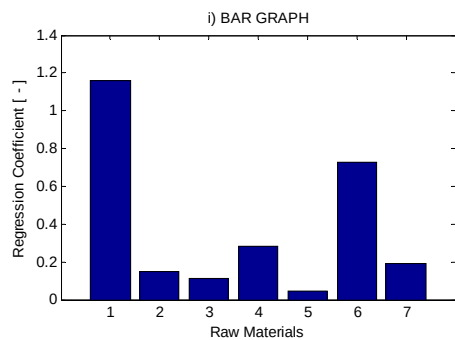
Wet Film Thickness



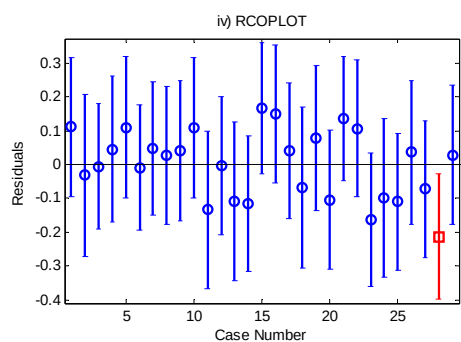
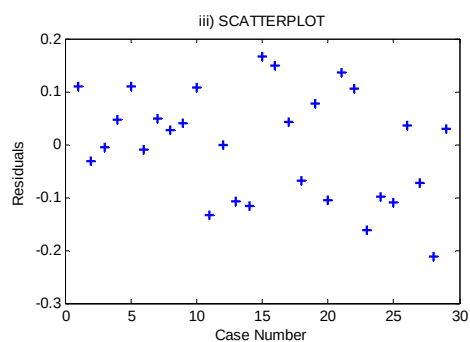
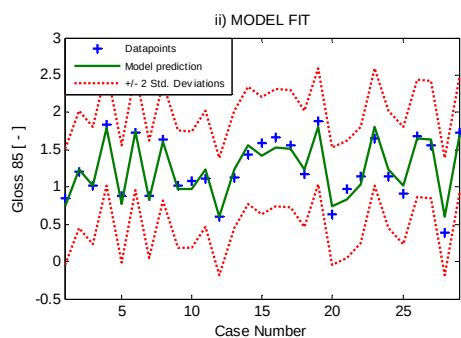
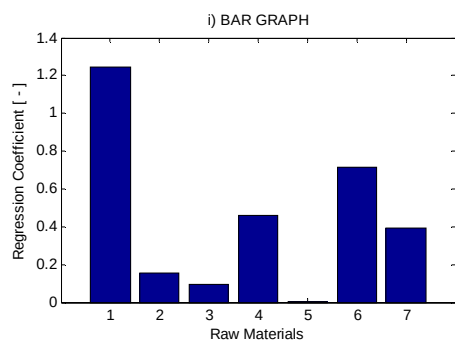
Gloss (20)



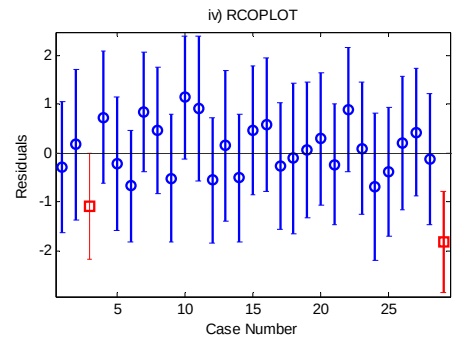
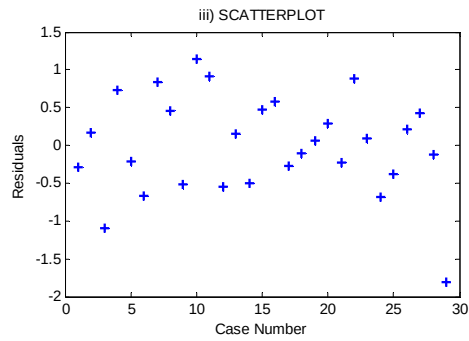
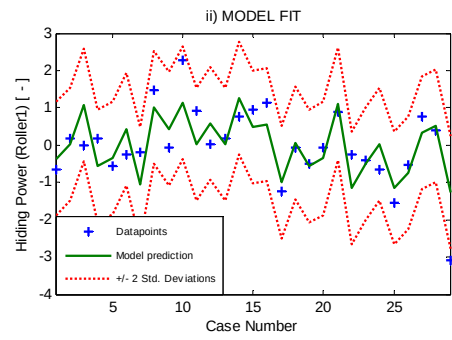
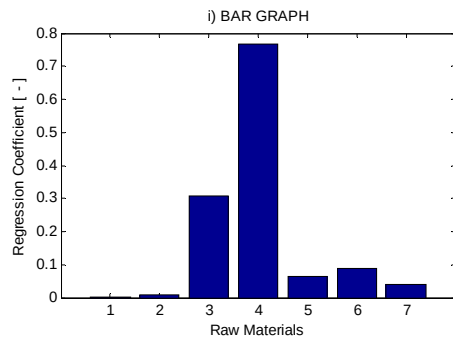
Gloss (60)



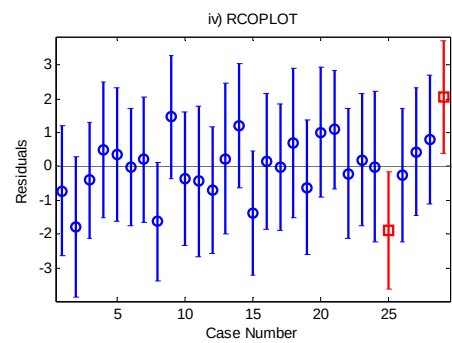
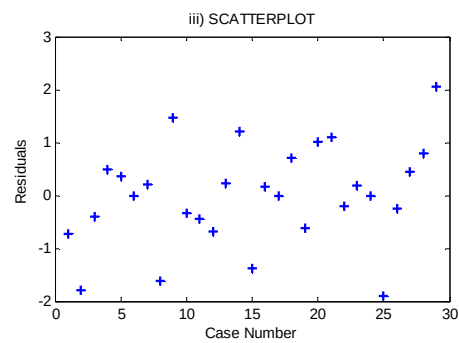
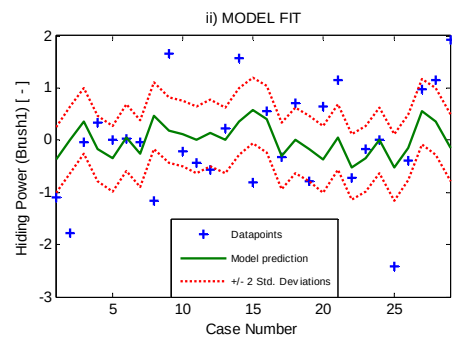
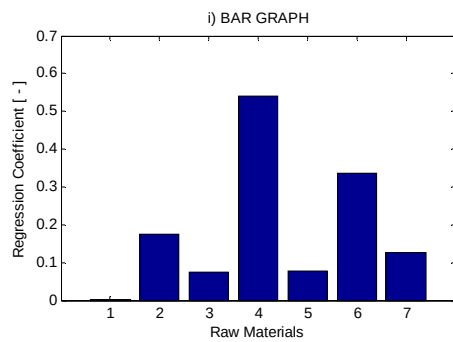
Gloss (85)



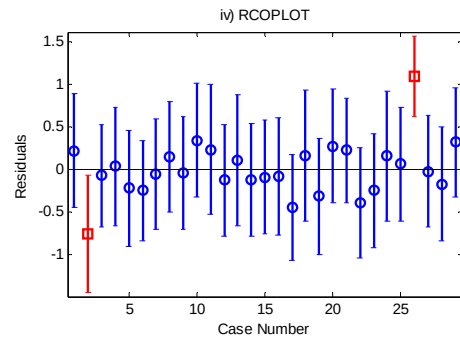
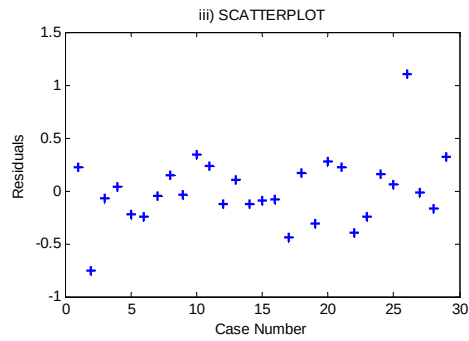
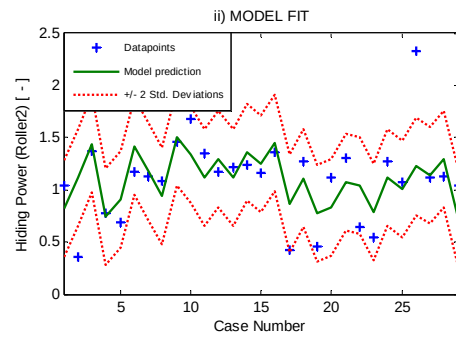
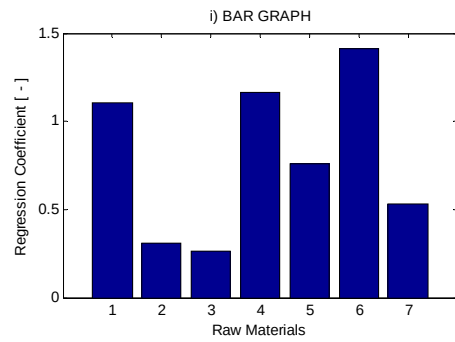
Hiding Power Roller - 1st coat



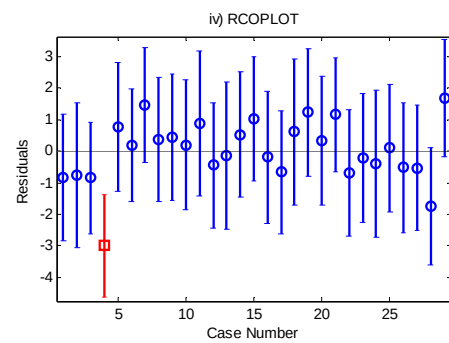
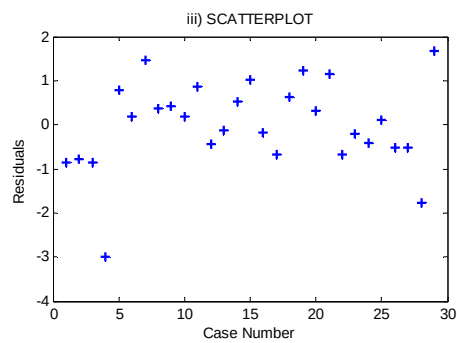
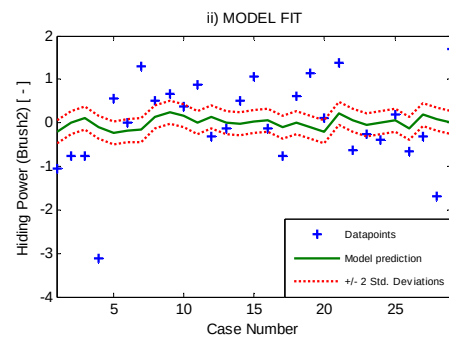
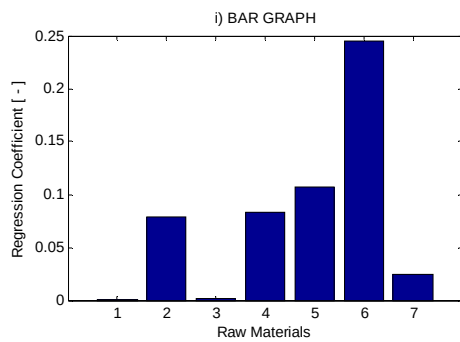
Hiding Power Brush - 1st coat



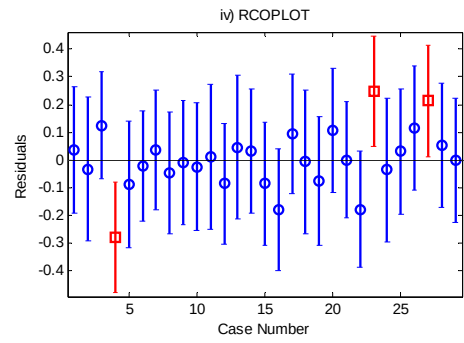
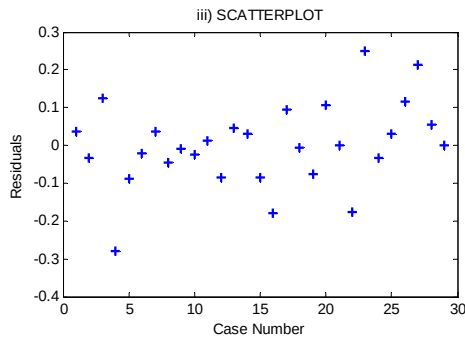
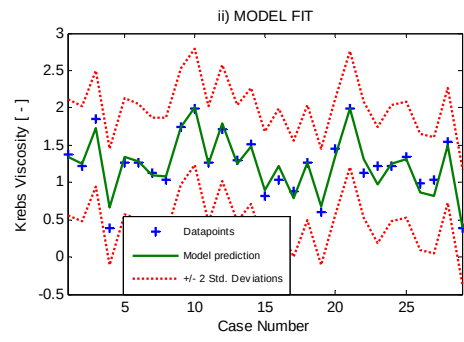
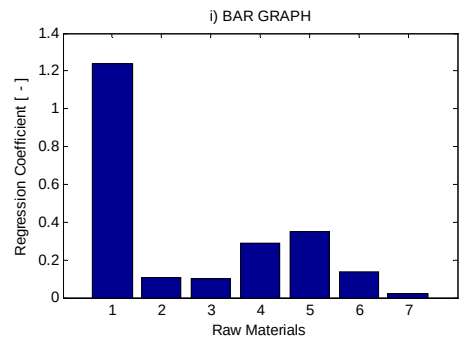
Hiding Power Roller - 2nd coat



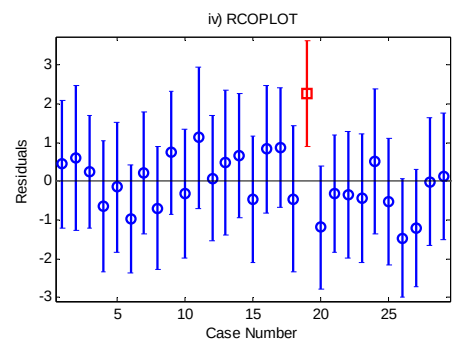
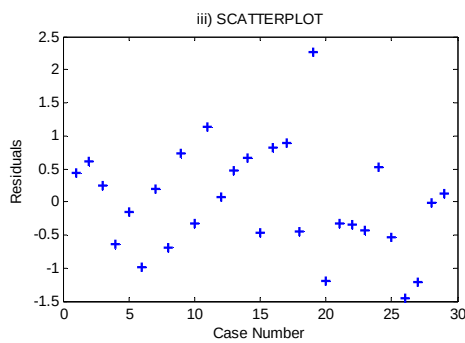
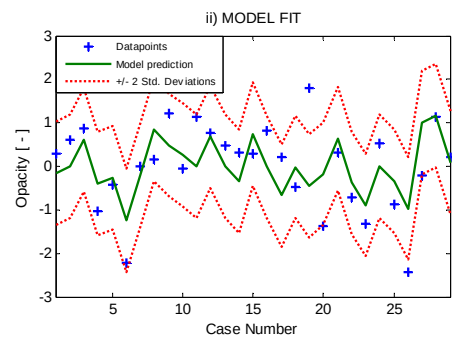
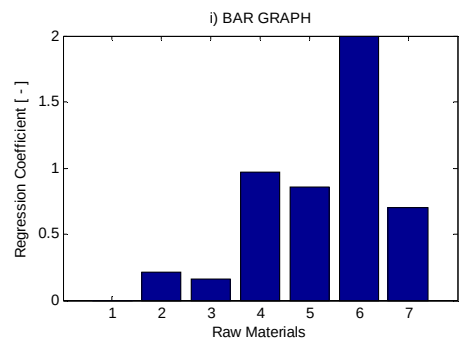
Hiding Power Brush - 2nd coat



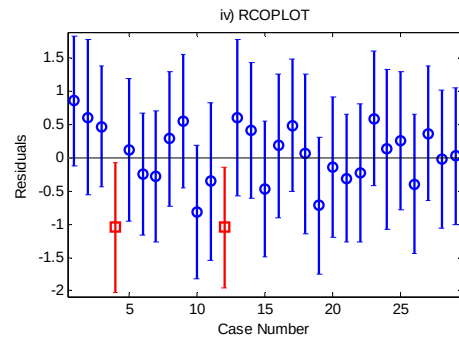
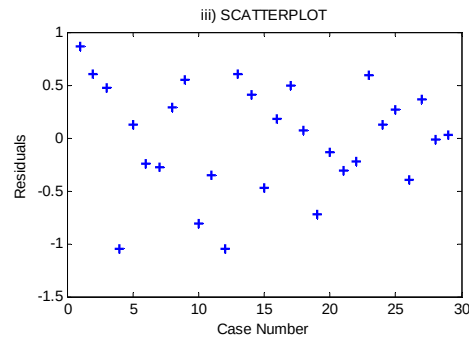
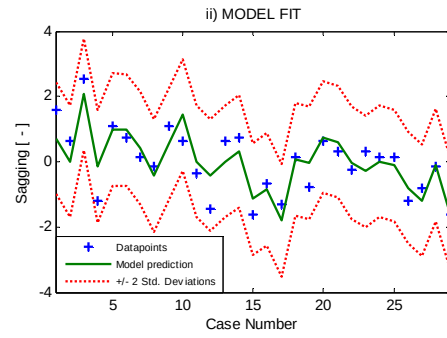
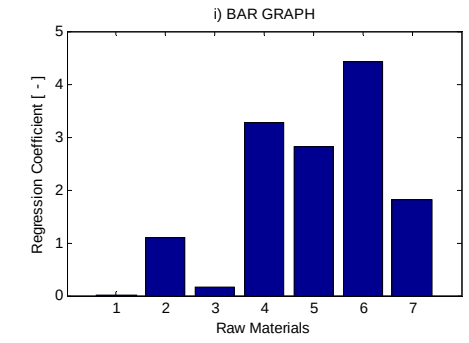
Krebs Viscosity



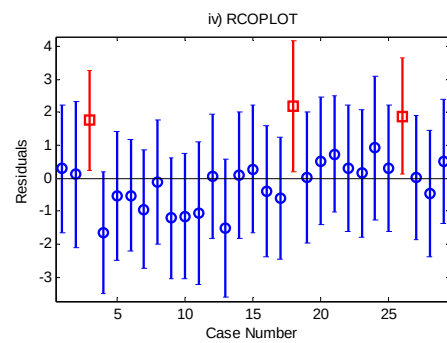
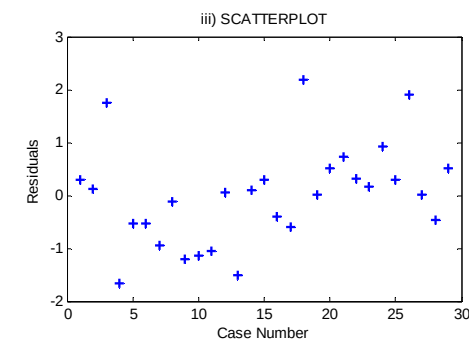
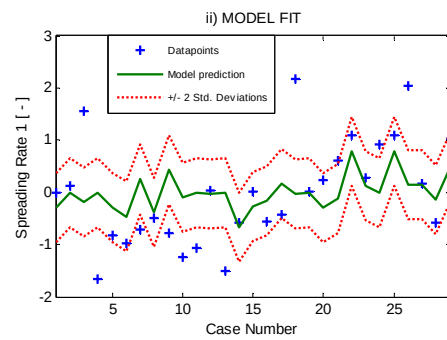
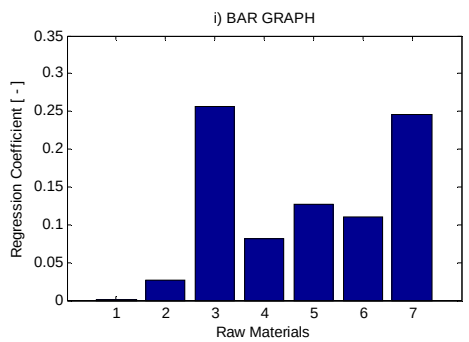
Opacity



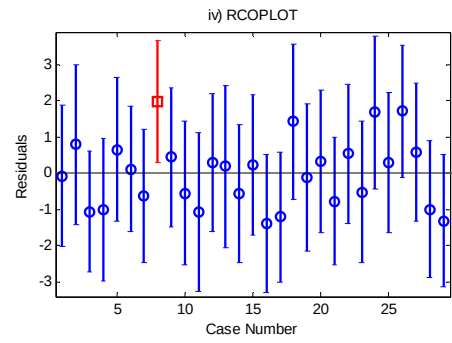
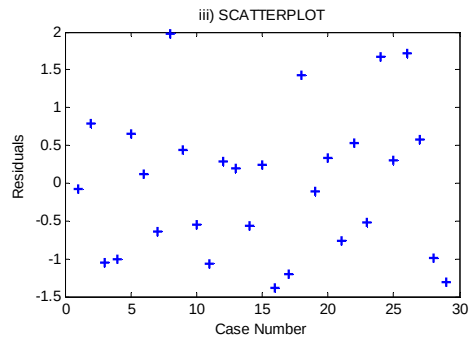
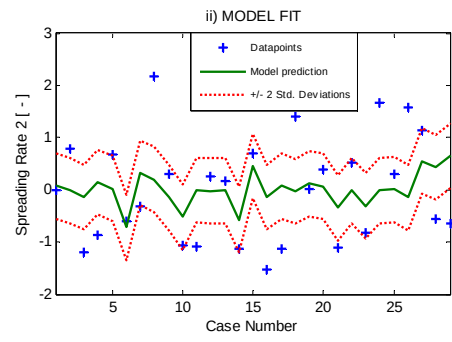
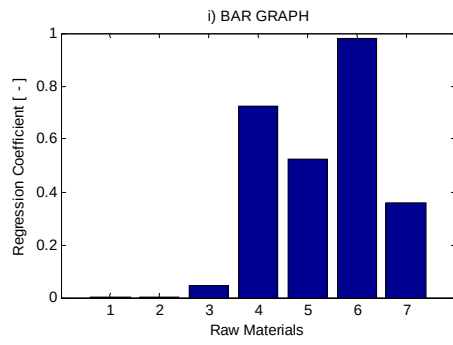
Anti-sag Index



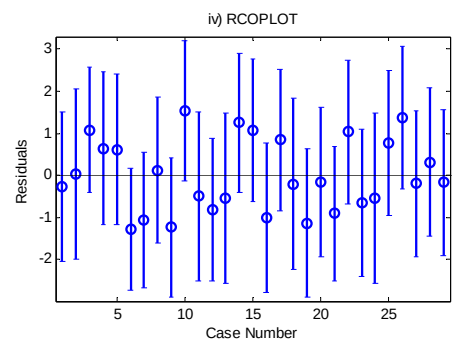
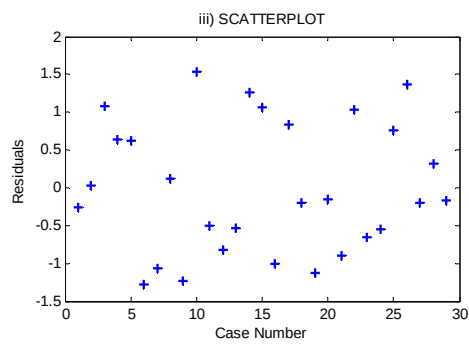
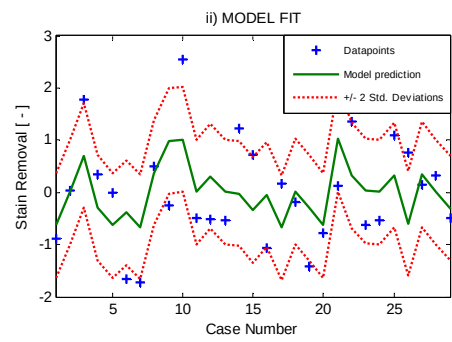
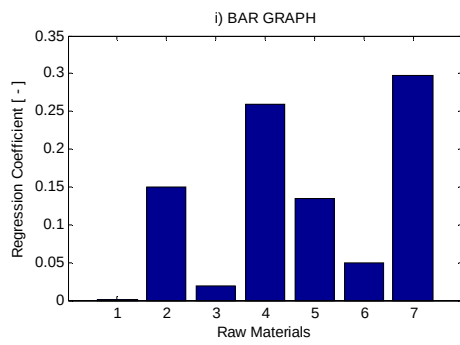
Spreading Rate 1st coat



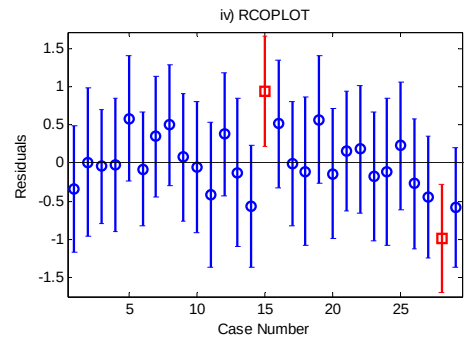
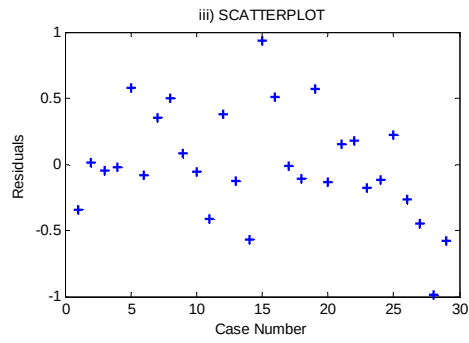
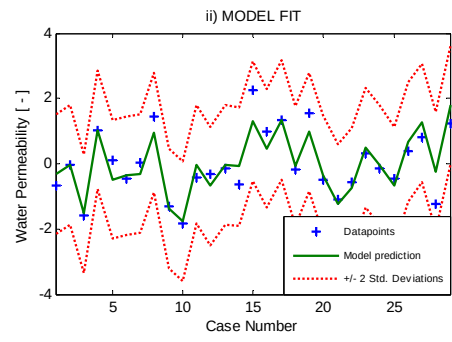
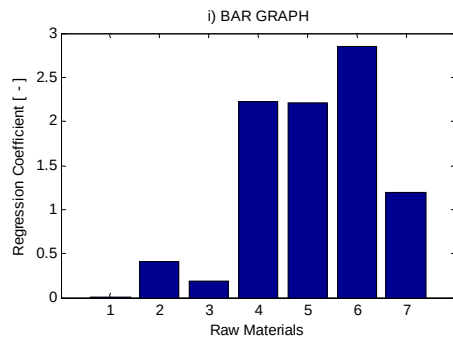
Spreading Rate 2nd coat



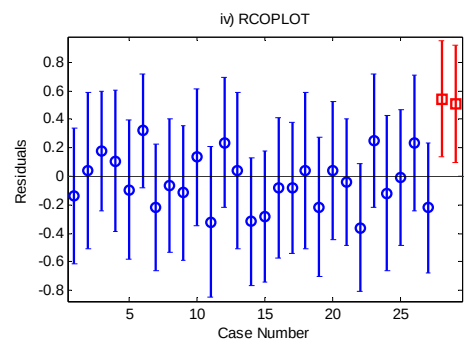
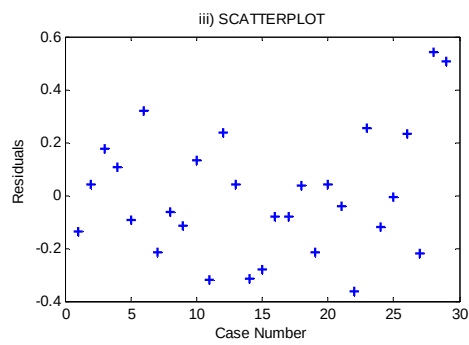
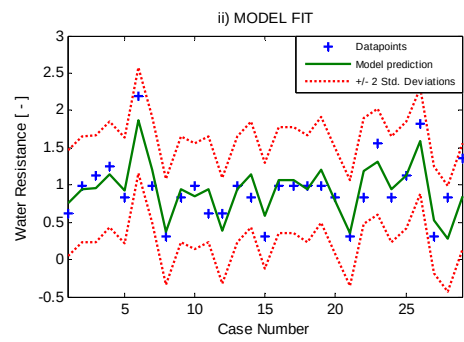
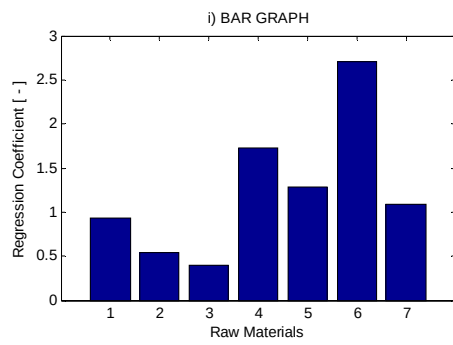
Stain Removal



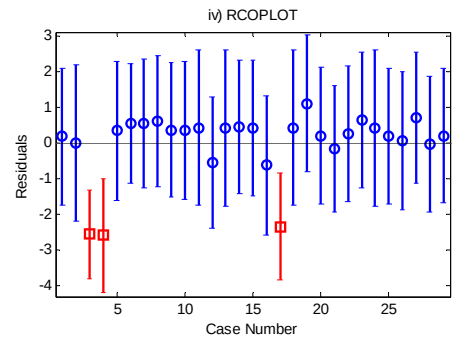
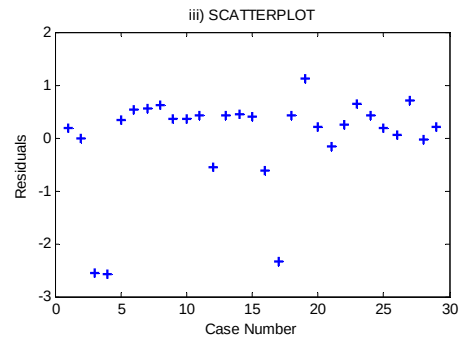
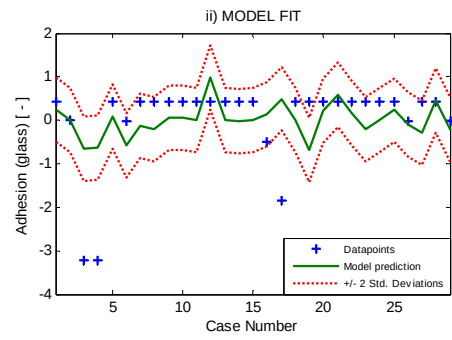
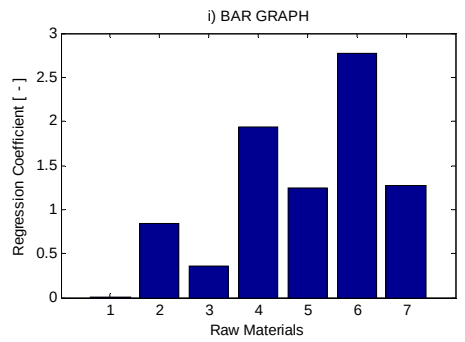
Water Permeability



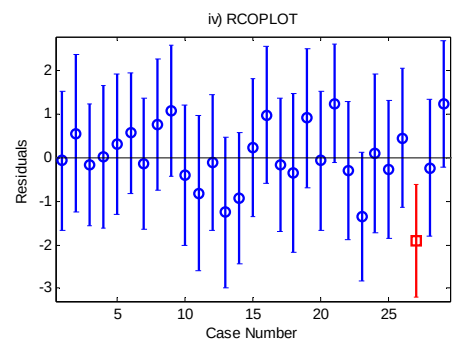
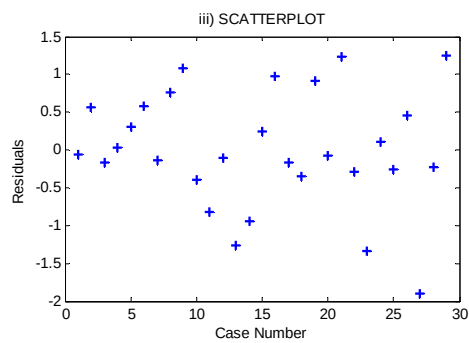
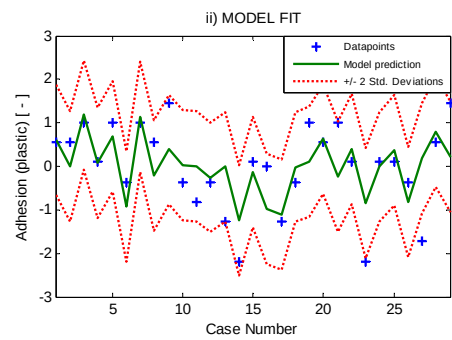
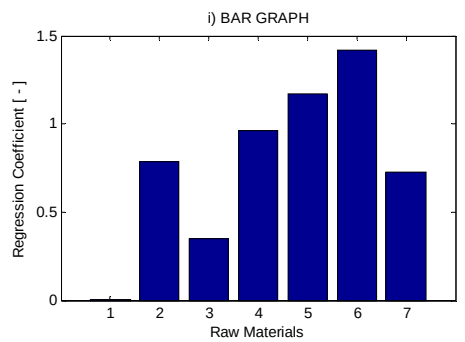
Water Resistance



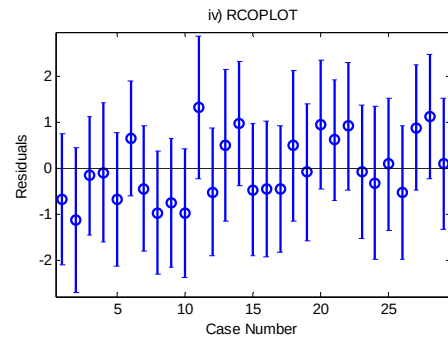
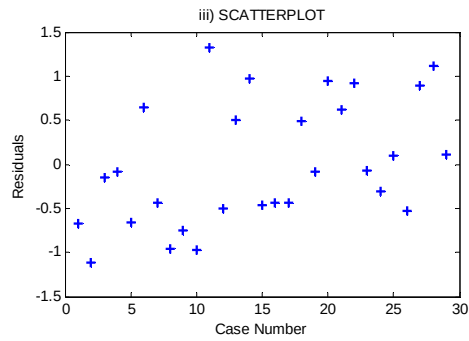
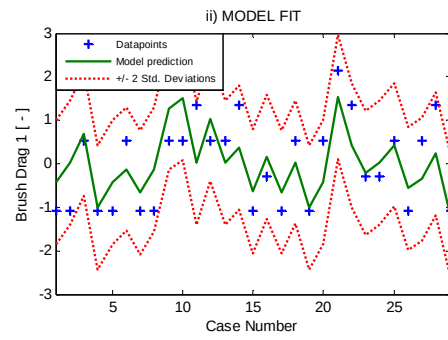
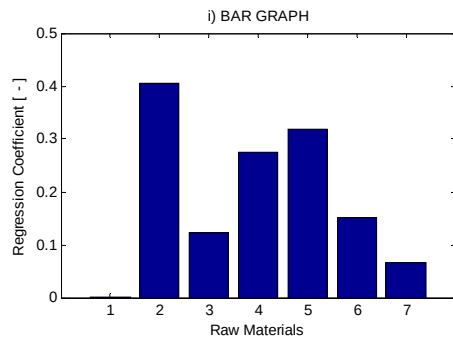
X-hatch Adhesion(glass)



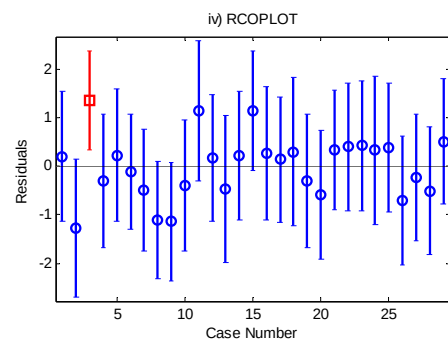
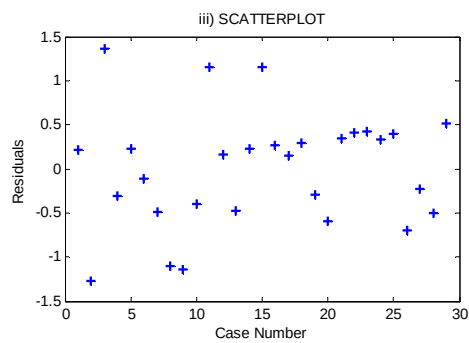
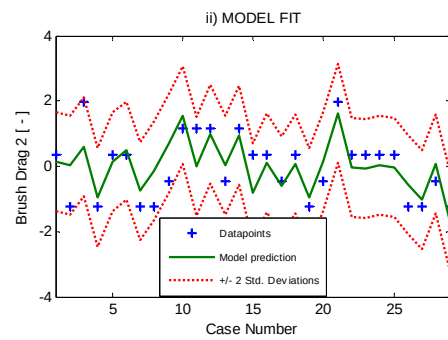
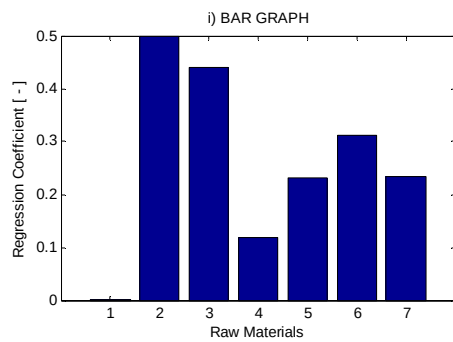
X-hatch Adhesion(panel 2A)



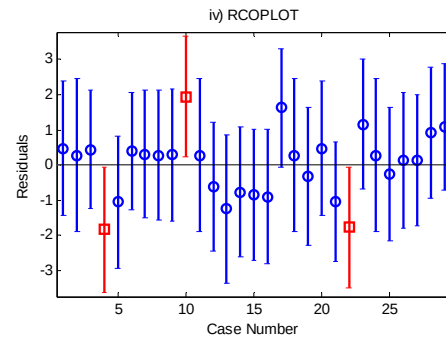
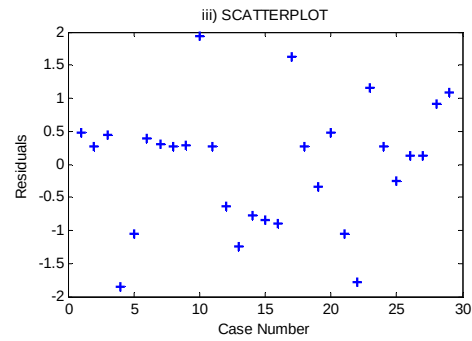
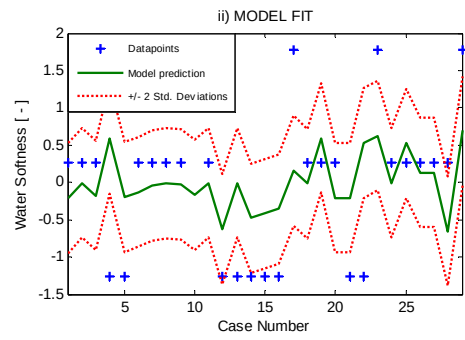
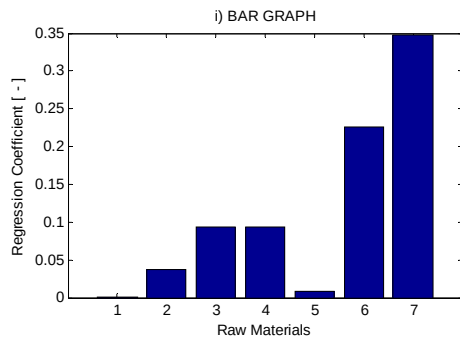
Brush Drag1st coat



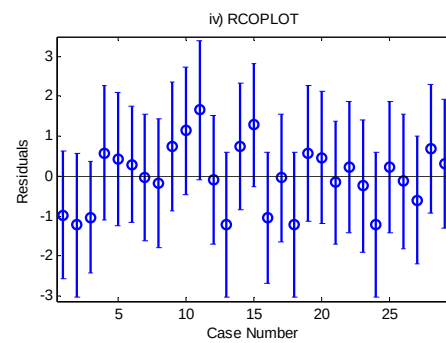
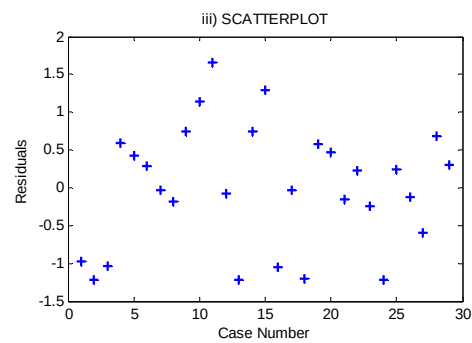
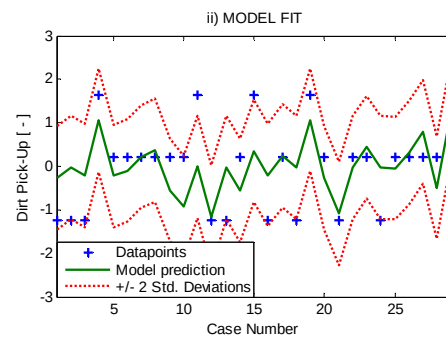
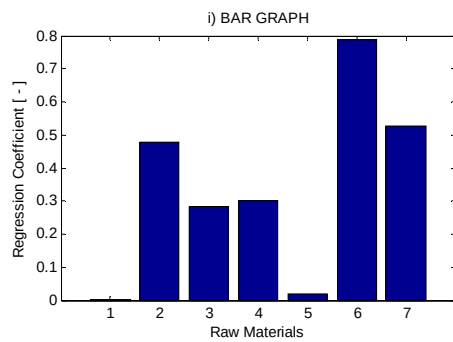
Brush Drag 2nd coat



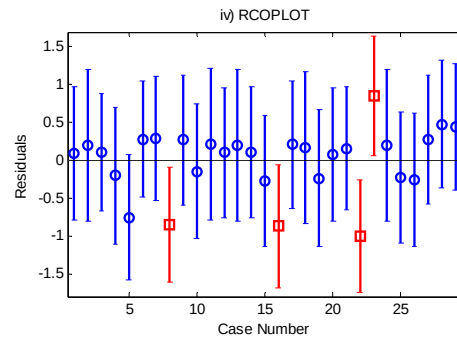
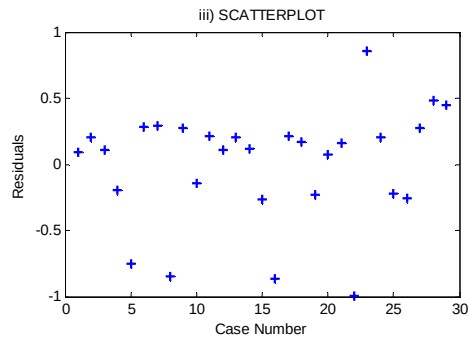
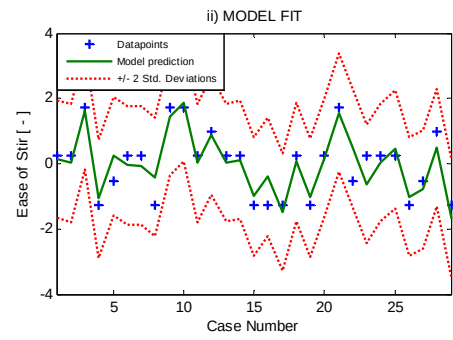
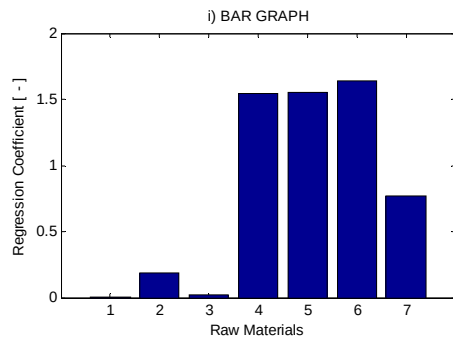
Chemical ResistanceWater Softness



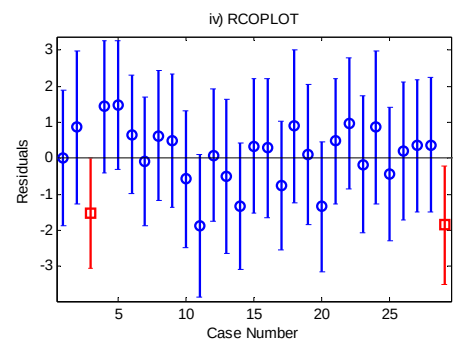
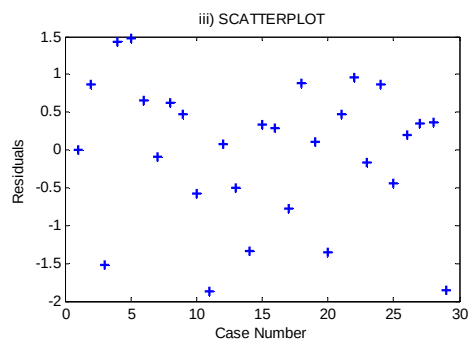
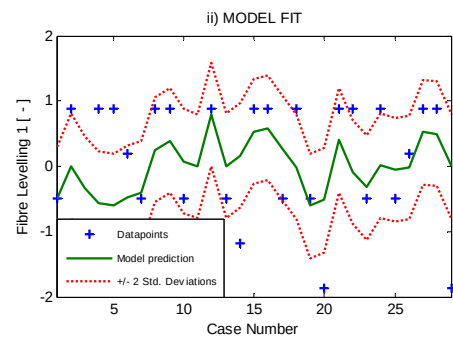
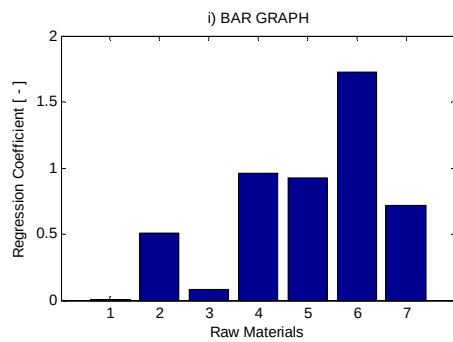
Dirt Pick-up



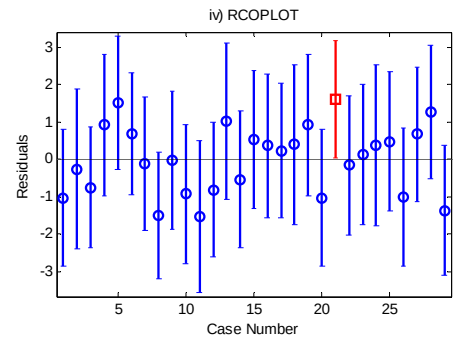
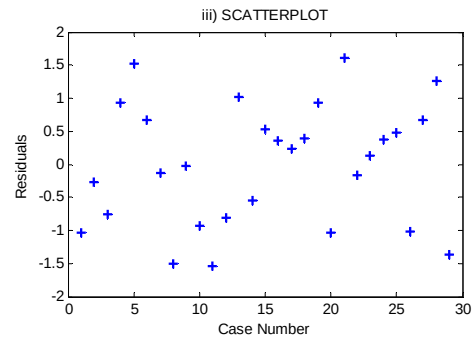
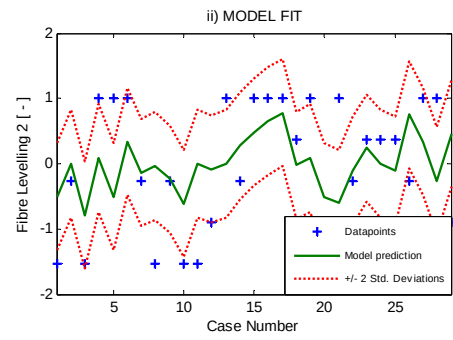
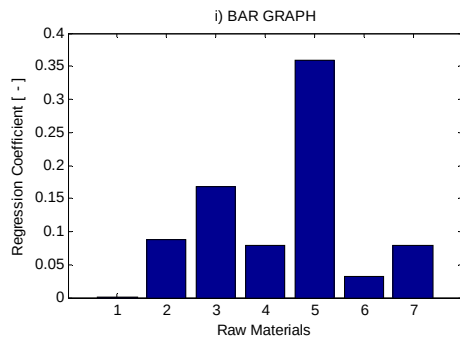
Ease of Stir



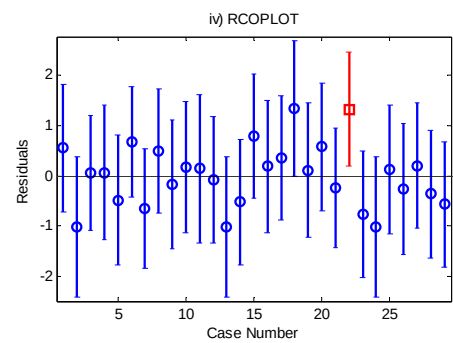
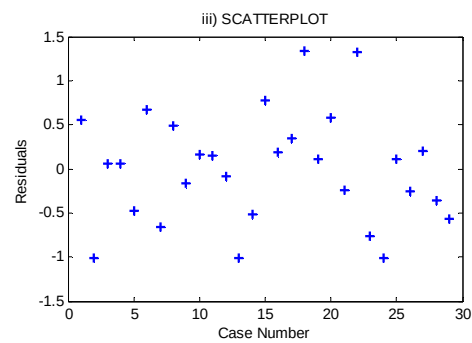
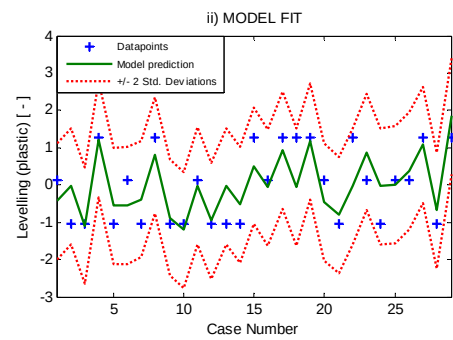
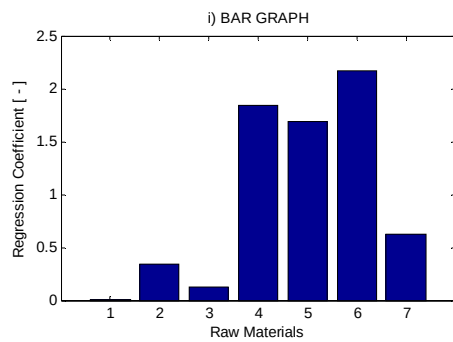
Fibre Levelling 1st coat



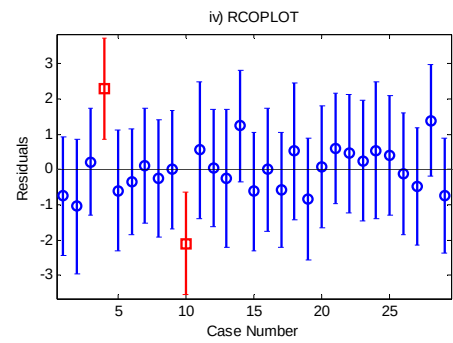
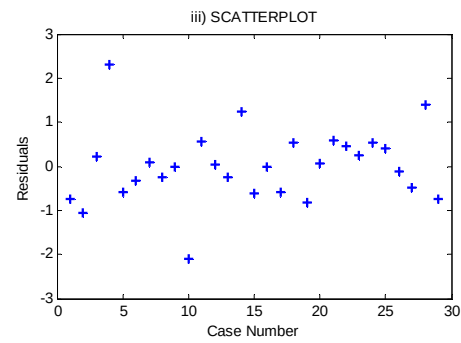
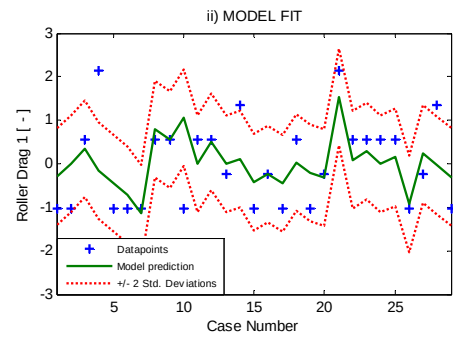
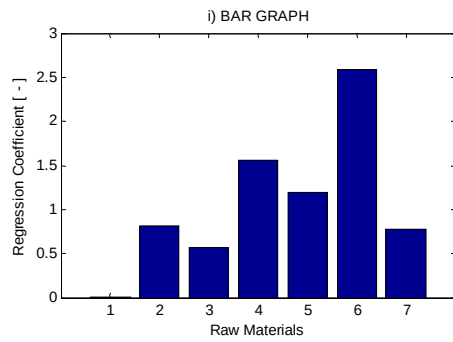
Fibre Levelling 2nd coat



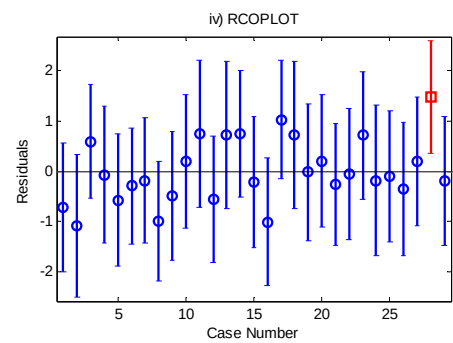
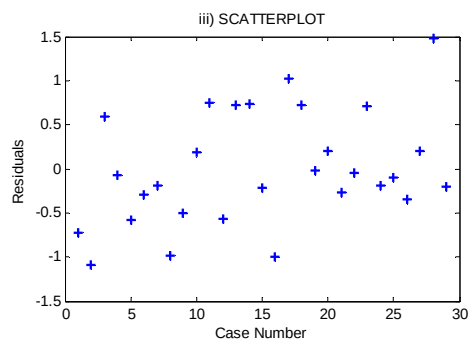
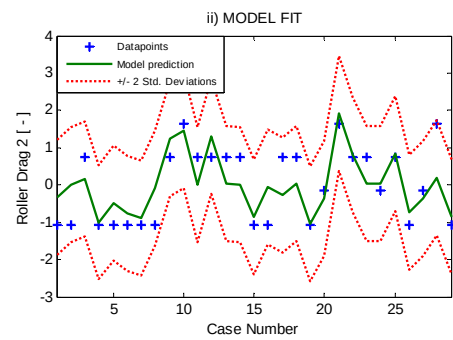
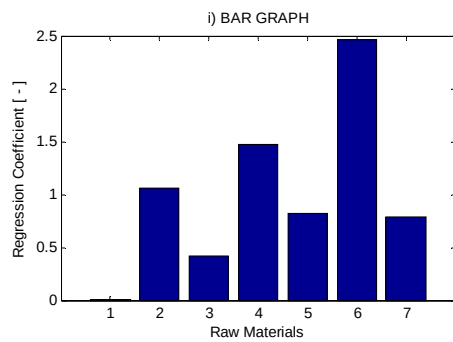
Levelling (Leneta)



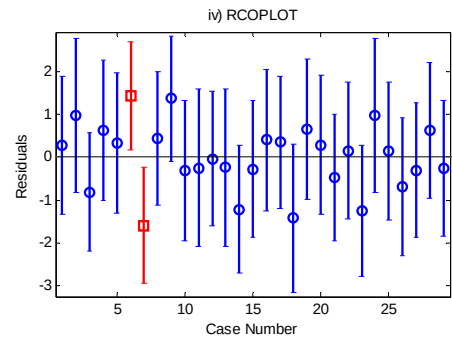
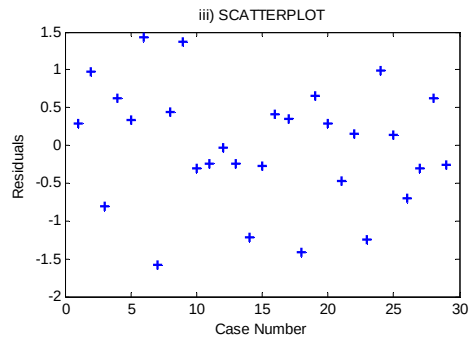
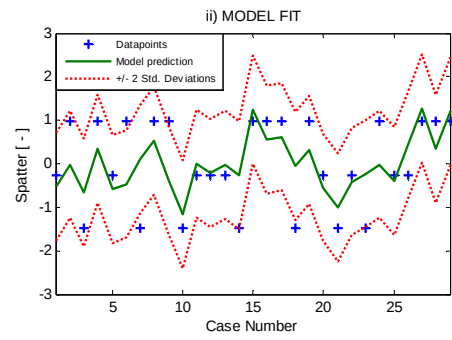
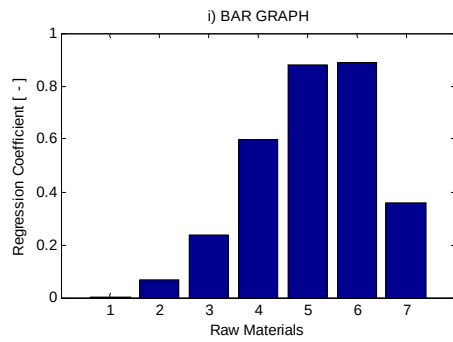
Roller Drag 1st coat



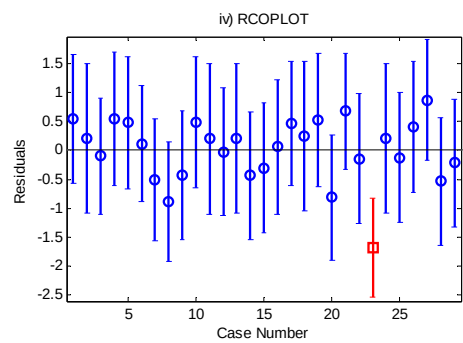
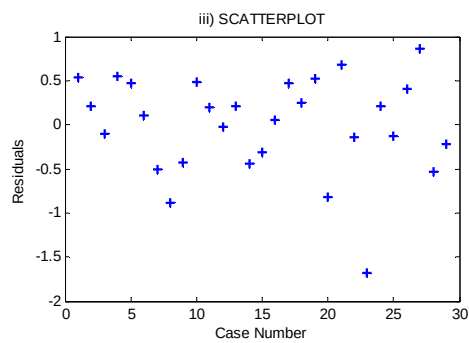
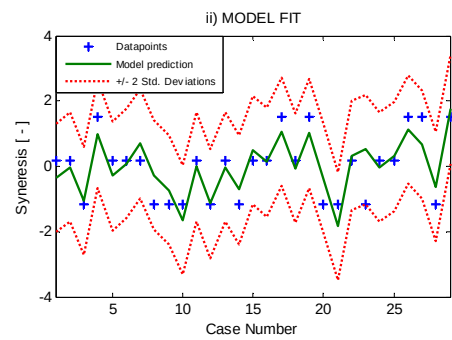
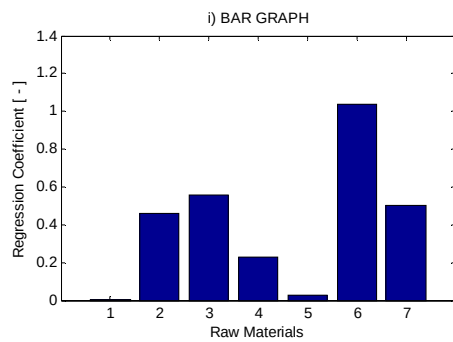
Roller Drag 2nd coat



Spatter



Synerisis



12 APPENDIX E: RESULTS - MODELLING

12.1 INDIRECT MODELLING OF RHEOLOGICAL DATA

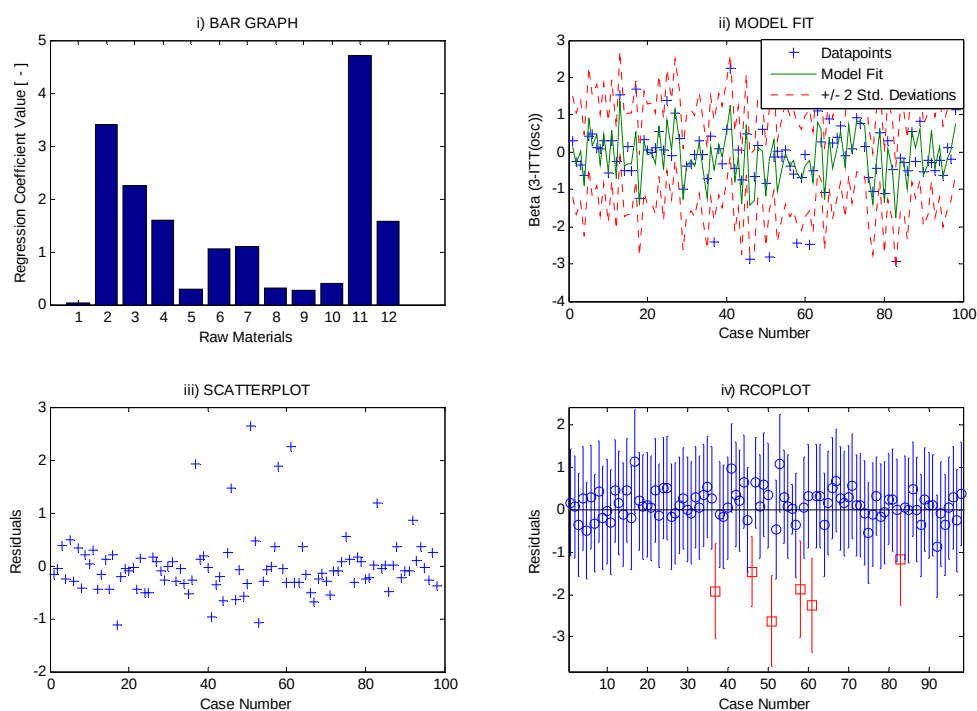
Indirect model parameters (α , β and γ) for the rheology curves of the 164 paint samples of the modelling data is presented in electronic format on the CD at the end of the thesis (Indirect Modelling/Modelling)

12.2 LINEAR MODELLING (MODEL 1)

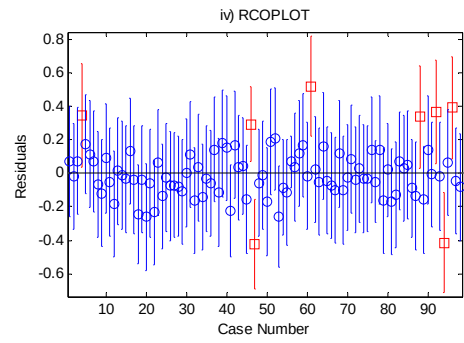
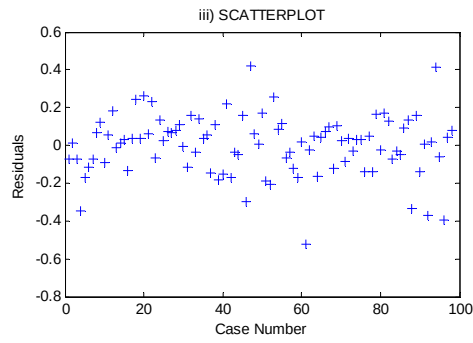
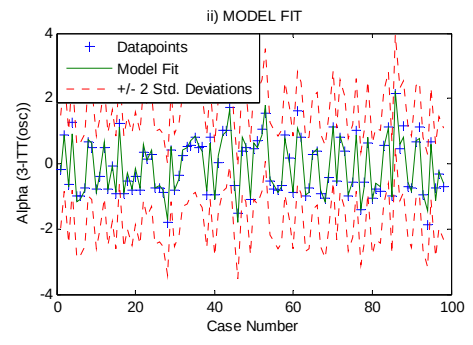
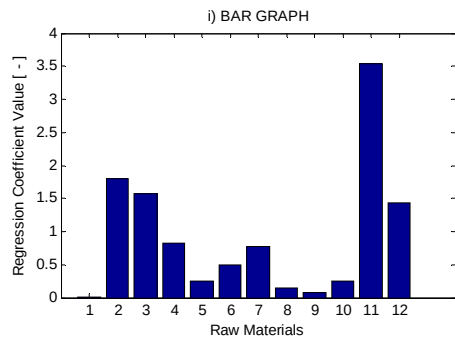
The x-axis numbering correspond to the raw materials as follows:

1 – constant, 2 – pigment, 3 – extender, 4 – water, 5 – solvent, 6 – emulsion, 7 – organic pigment, 8 – rheology modifier1, 9 – rheology modifier2, 10 – combination, 11 – PVC, 12 - SOLIDS'

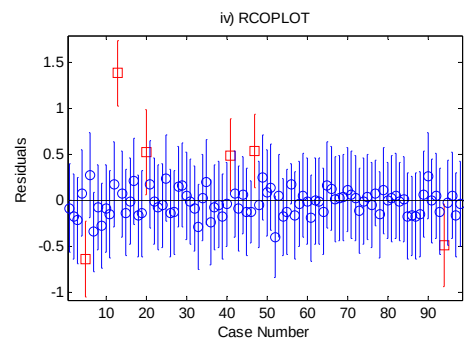
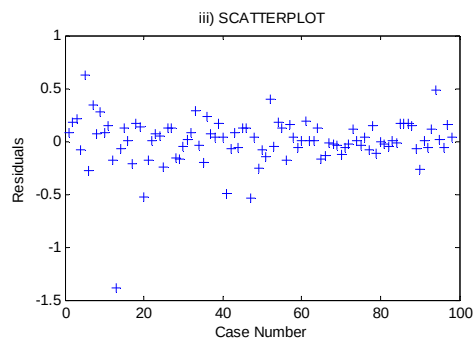
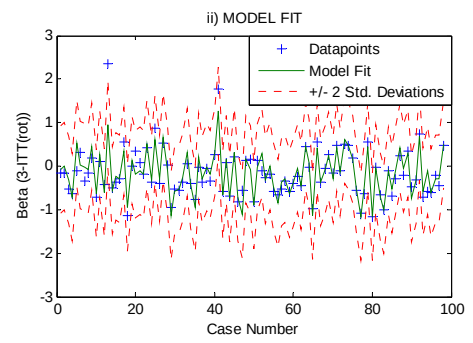
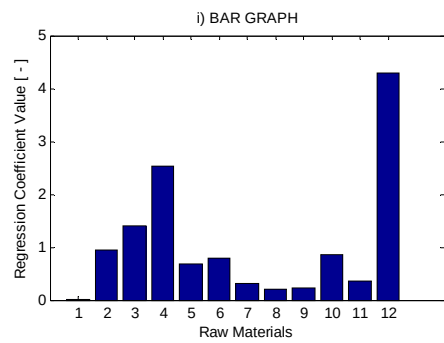
$\beta_{3-ITT(osc)}$



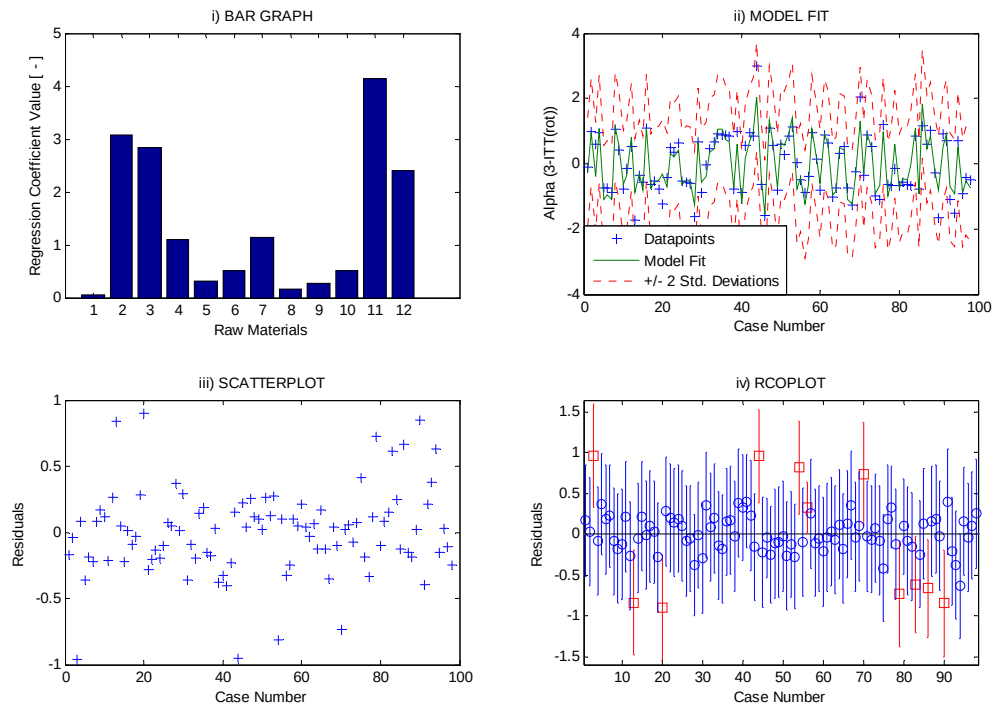
$\alpha_{3-ITT(osc)}$



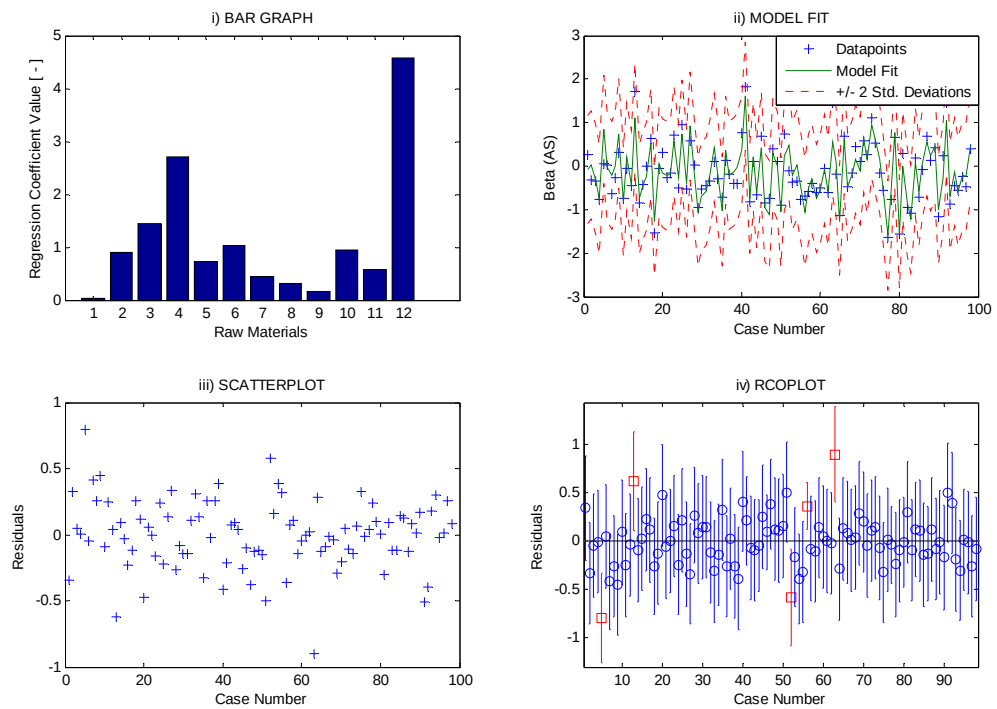
$\beta_{3-ITT(rot)}$



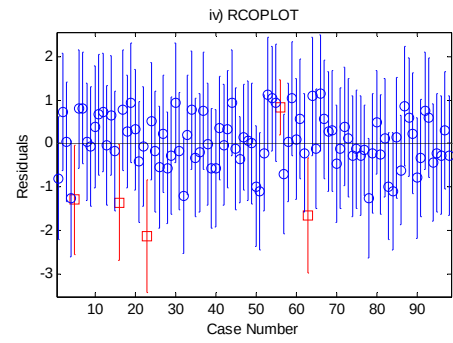
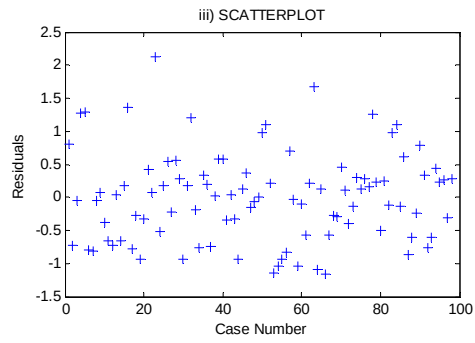
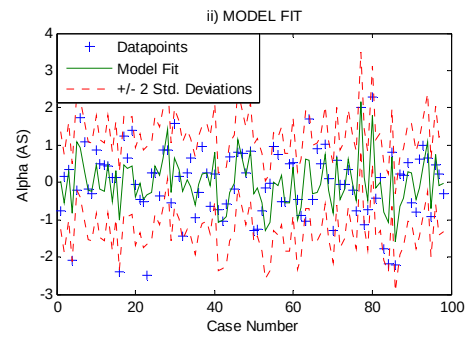
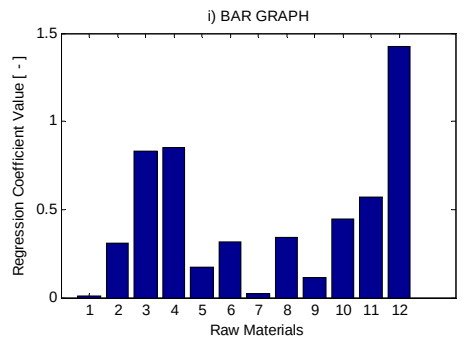
$\alpha_{3-ITT(rot)}$



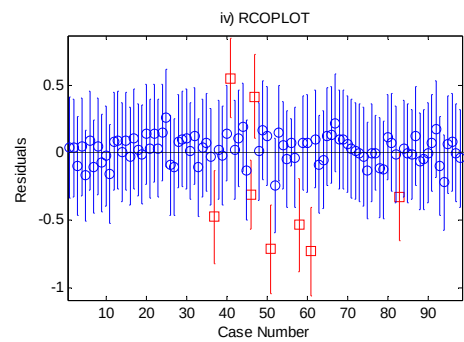
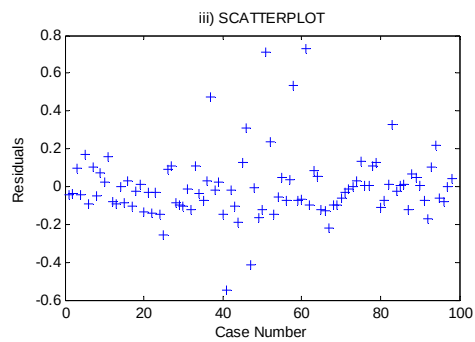
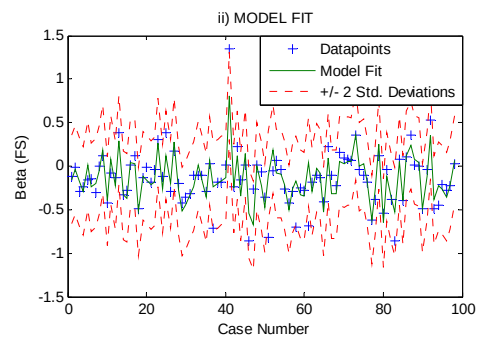
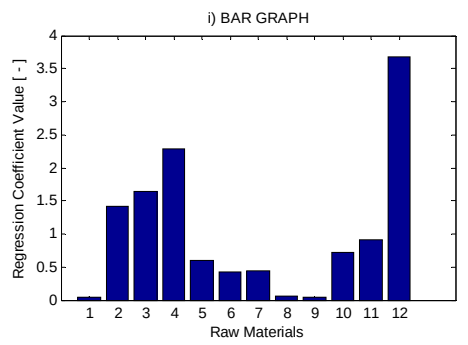
β_{AS}



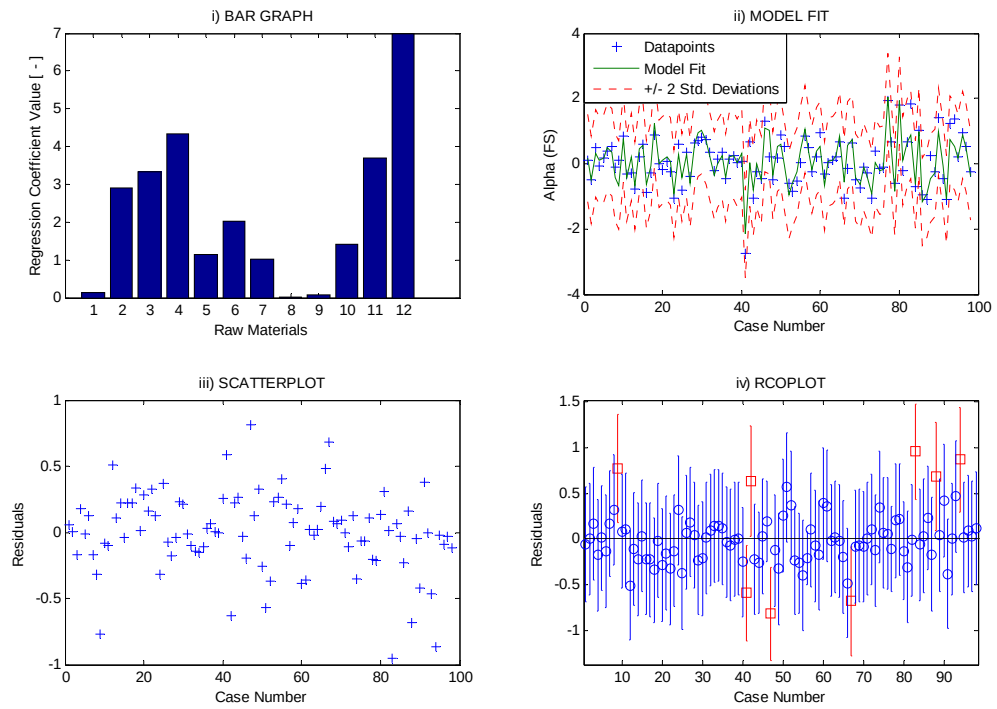
α_{AS}



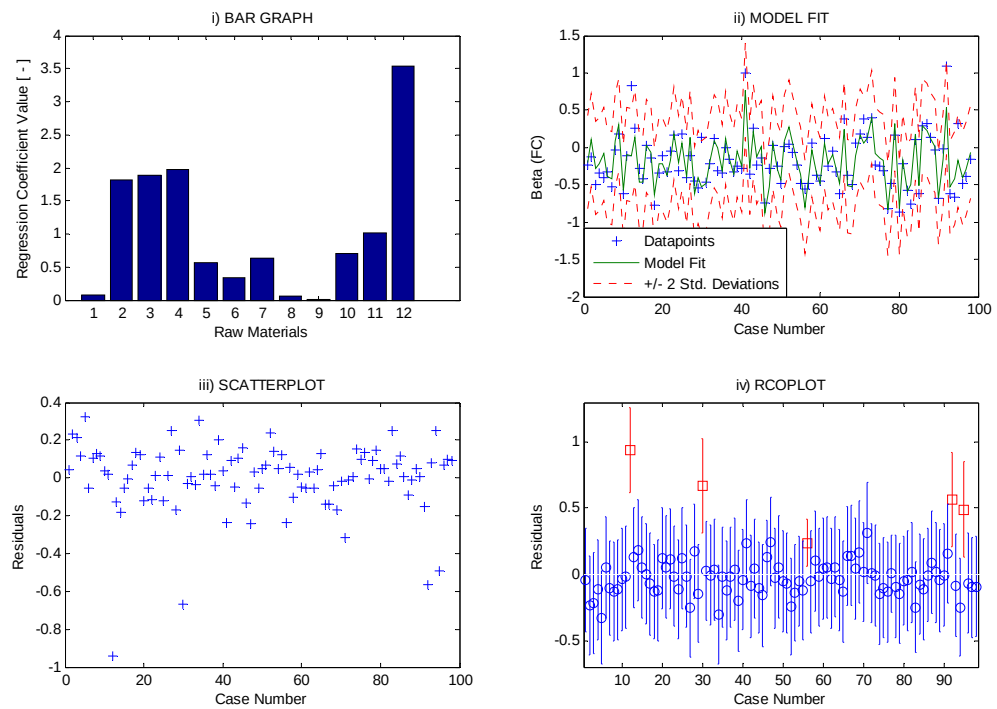
β_{FS}



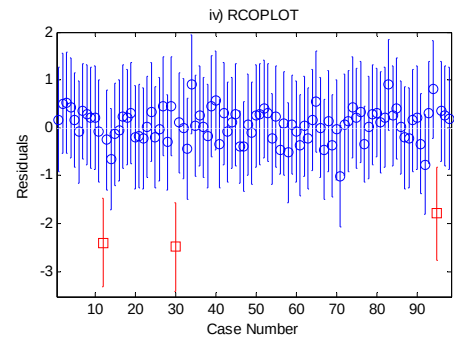
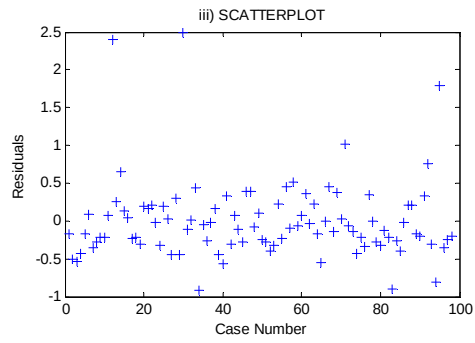
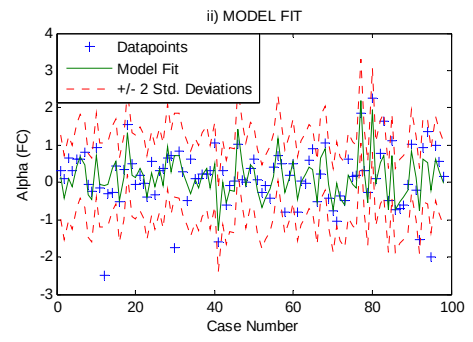
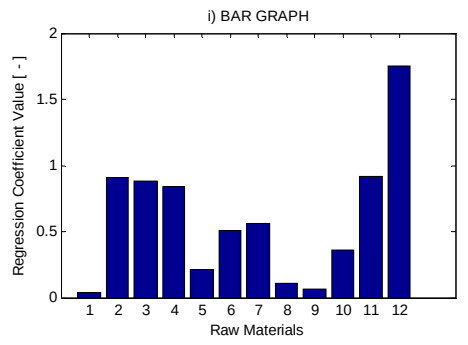
α_{FS}



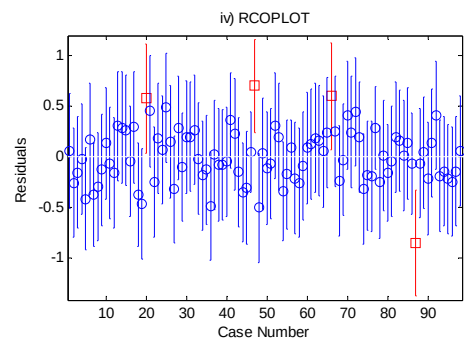
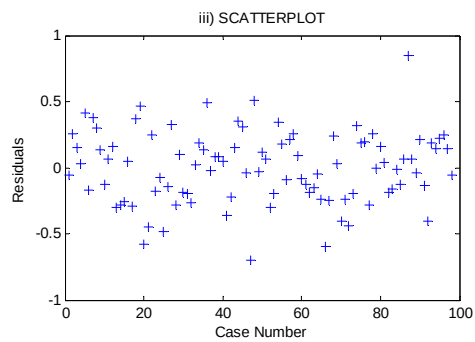
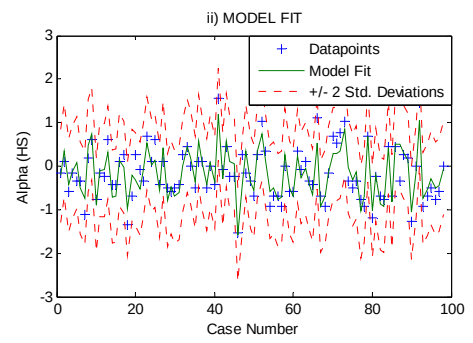
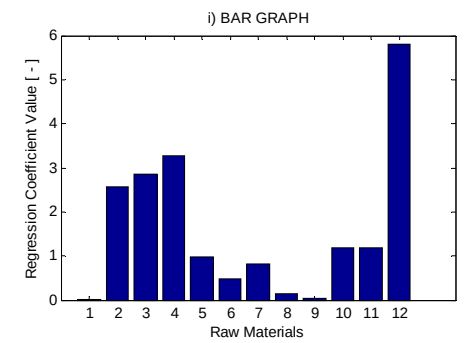
β_{FC}



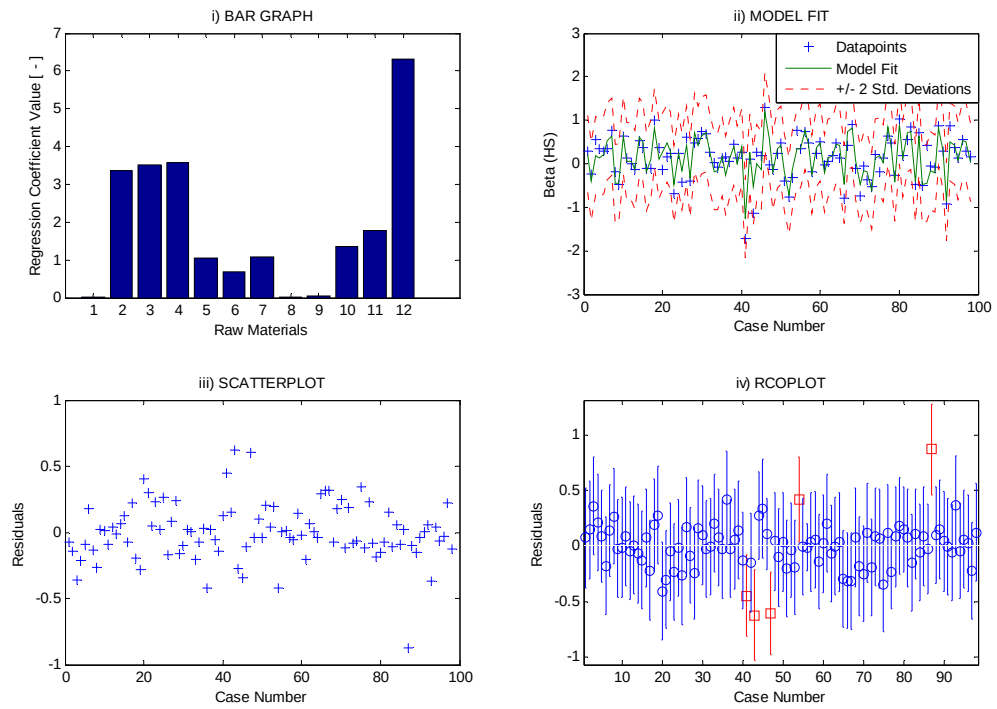
α_{FC}



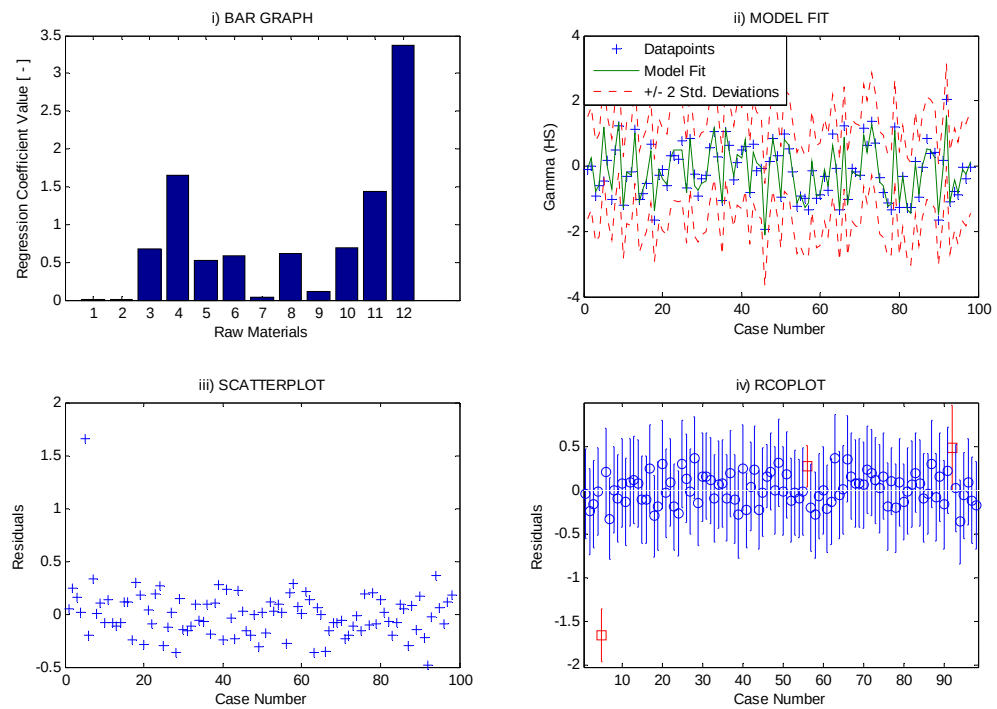
α_{HS}



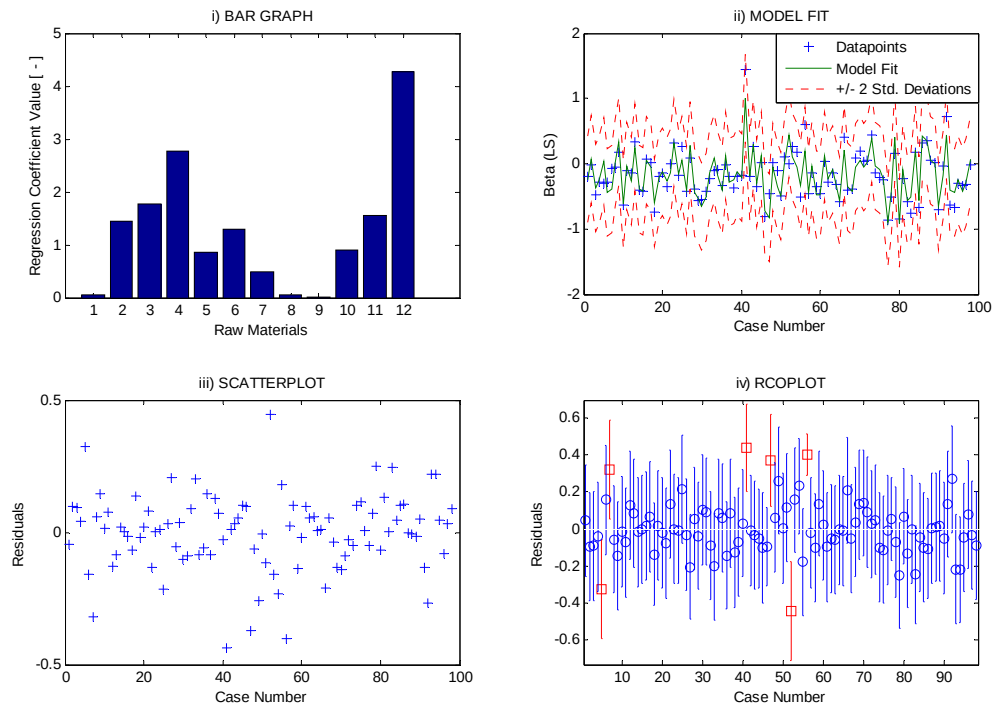
β_{HS}



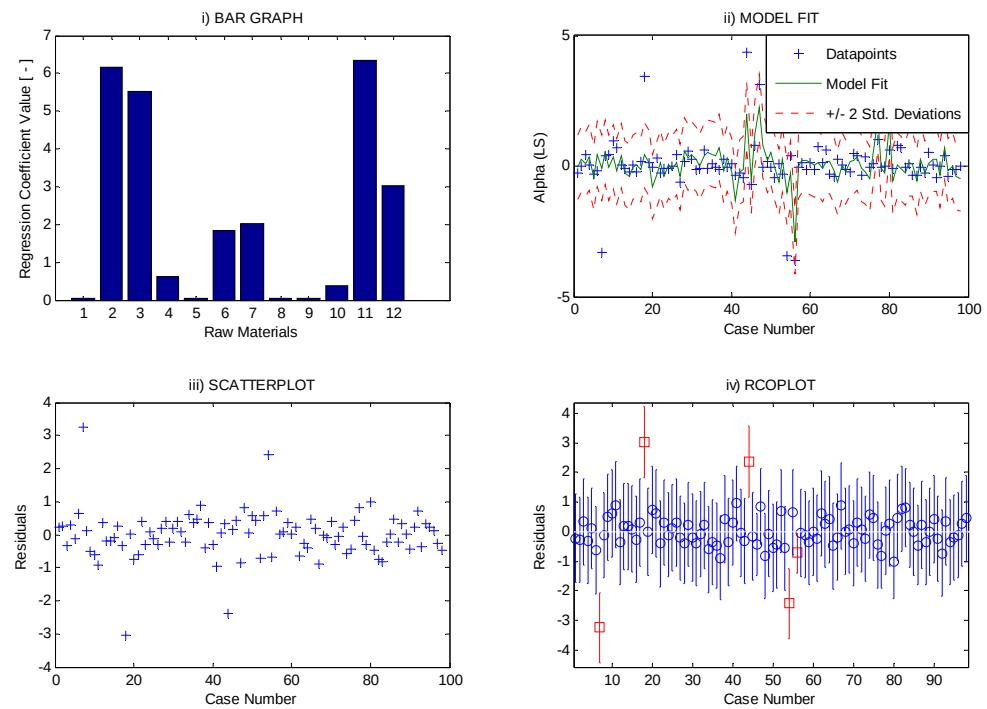
γ_{HS}



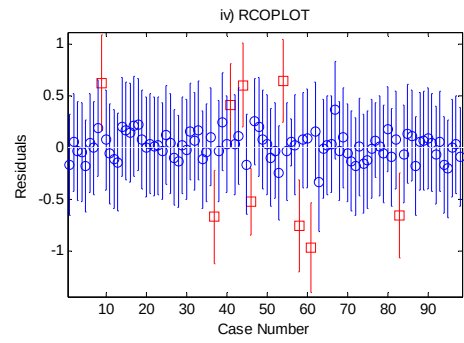
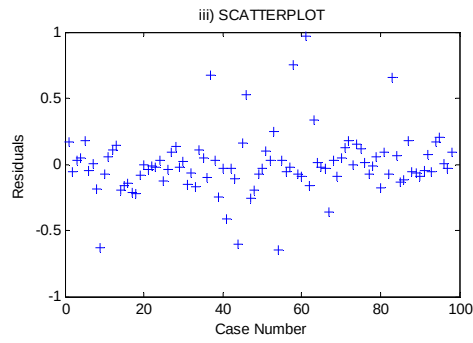
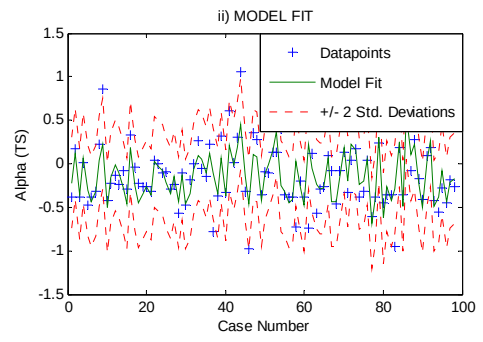
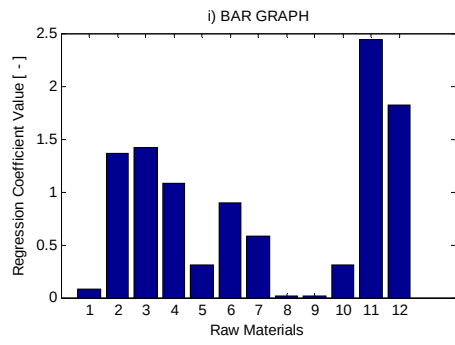
β_{LS}



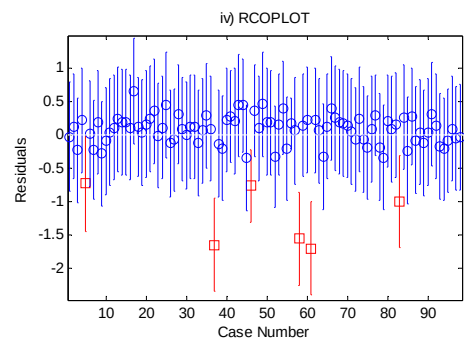
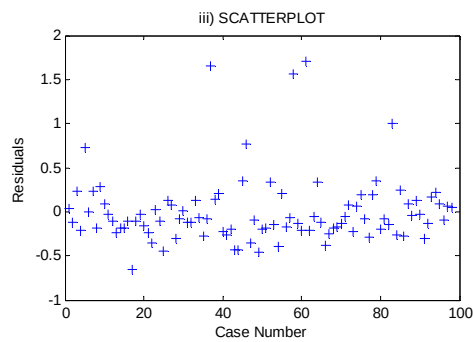
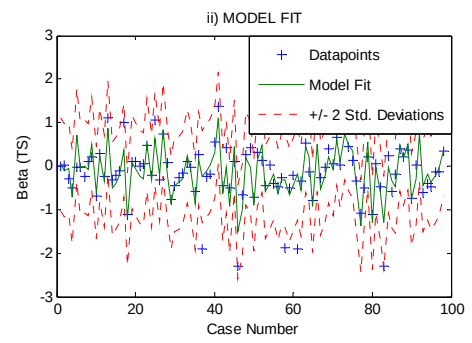
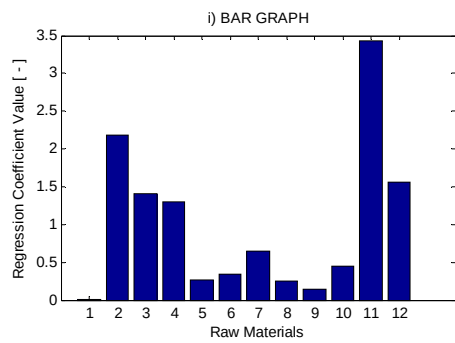
α_{LS}



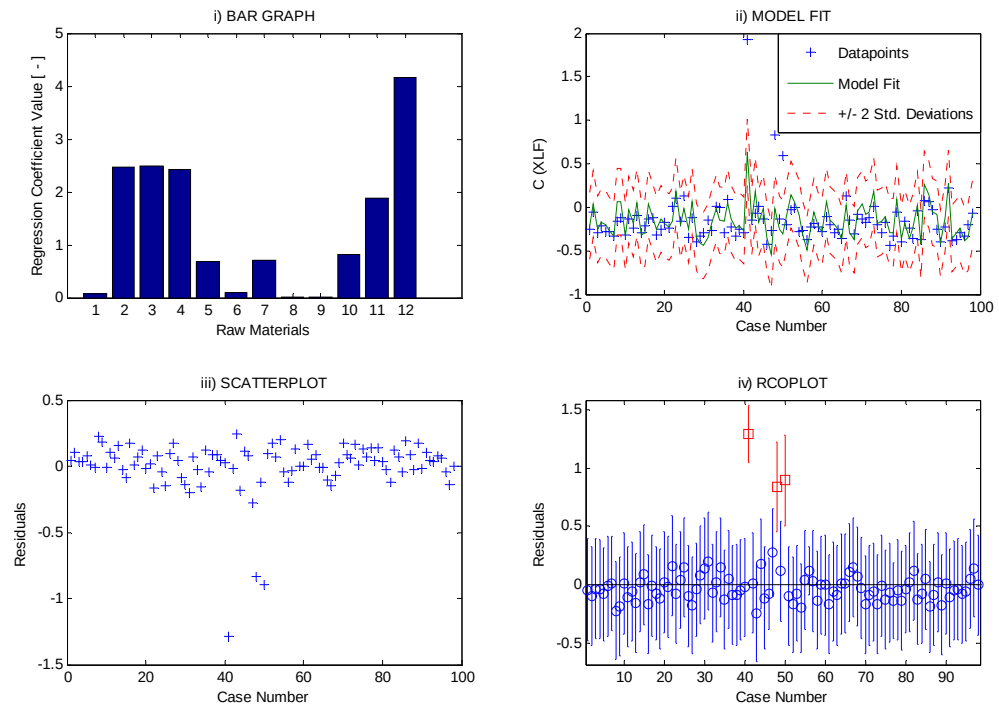
α_{TS}



β_{TS}



C_{XLF}



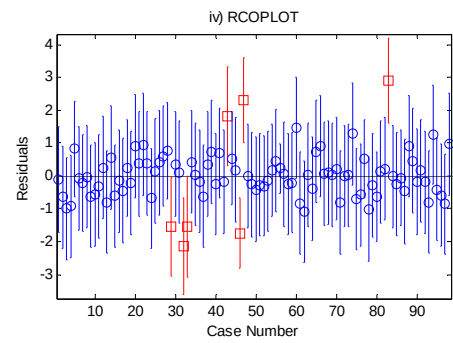
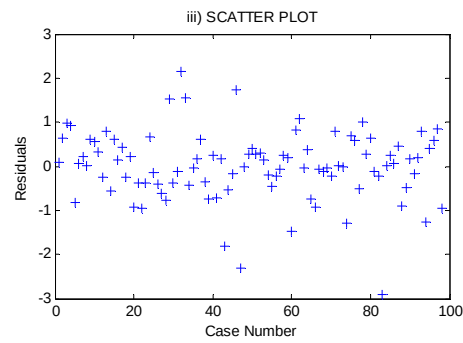
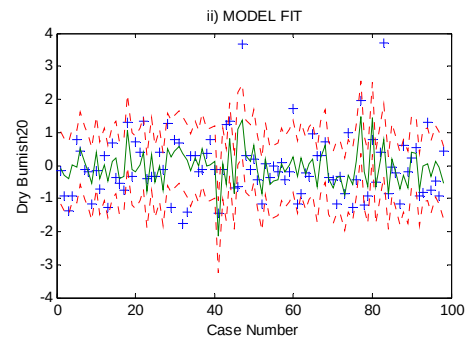
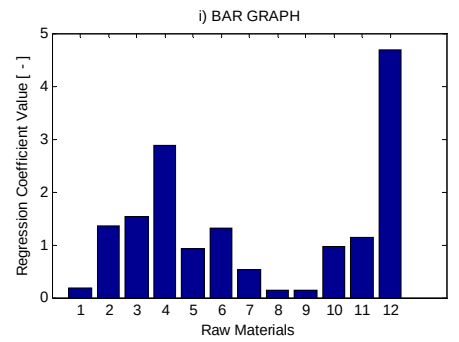
12.3 LINEAR MODELLING (MODEL 2)

The x-axis numbering of the bar graph correspond to the raw materials as follows:

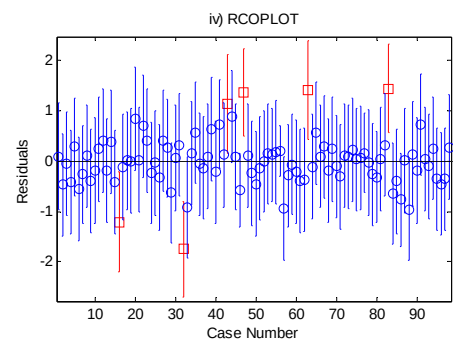
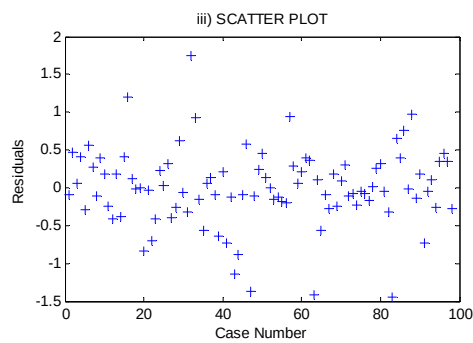
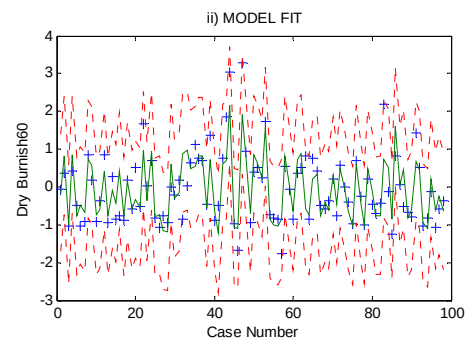
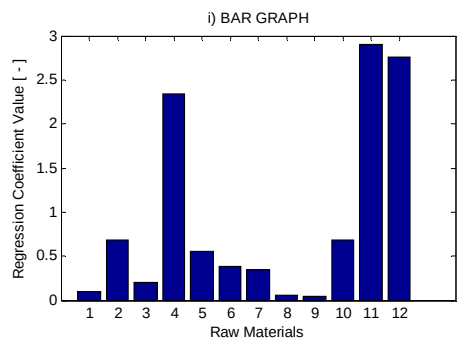
1 – constant, 2 – pigment, 3 – extender, 4 – water, 5 – solvent, 6 – emulsion, 7 – organic pigment, 8 – rheology modifier1, 9 – rheology modifier2, 10 – combination, 11 – PVC, 12 - VS

(The model fit for tha paint properties are given as standardised values).

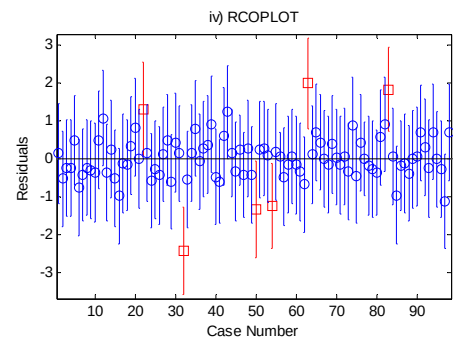
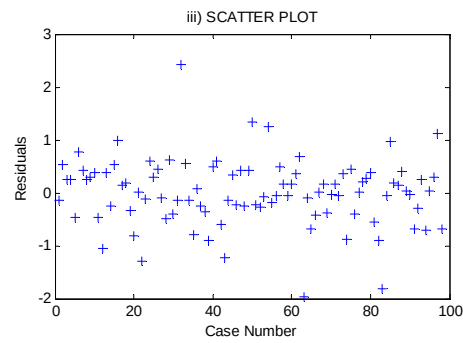
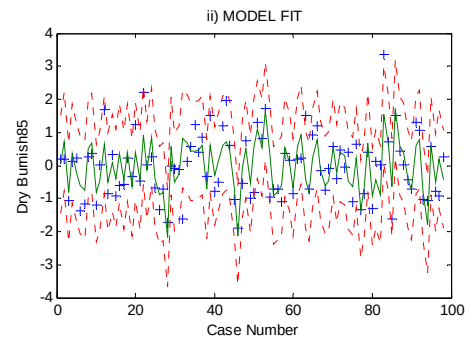
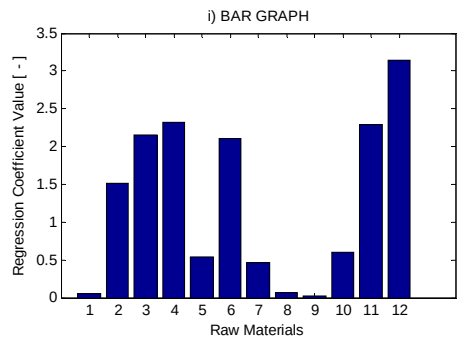
Dry Burnish 20



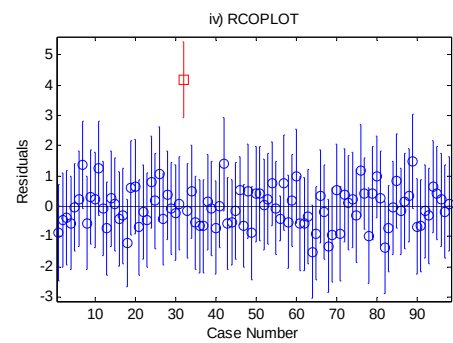
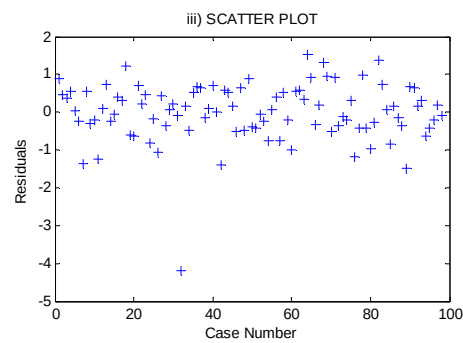
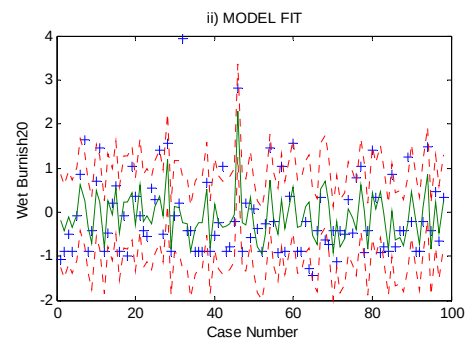
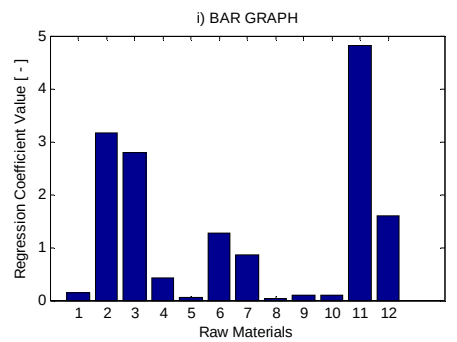
Dry Burnish 60



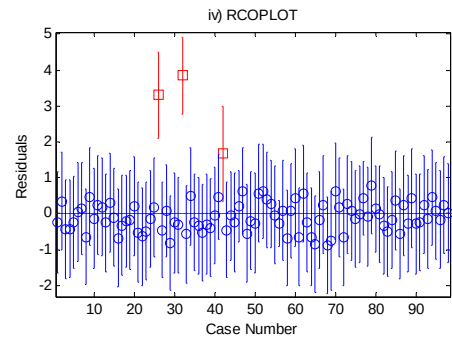
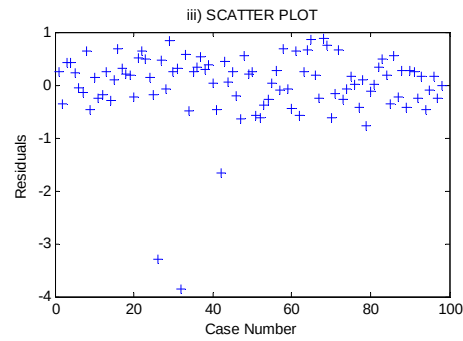
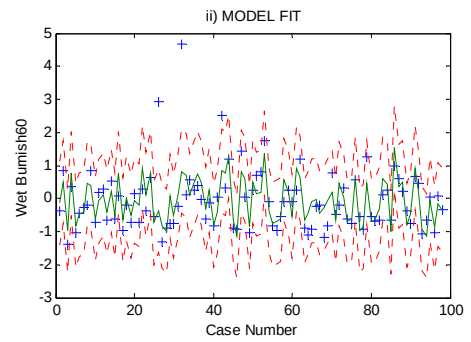
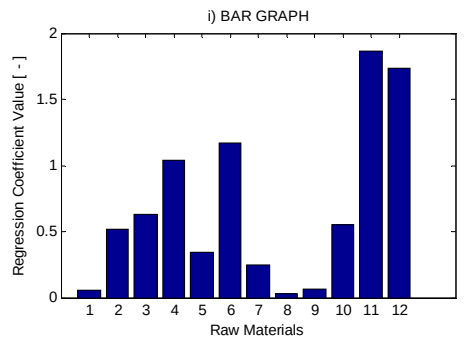
Dry Burnish 85



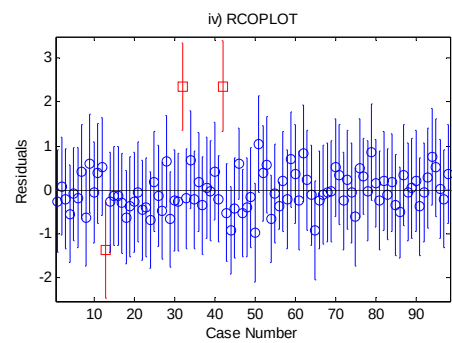
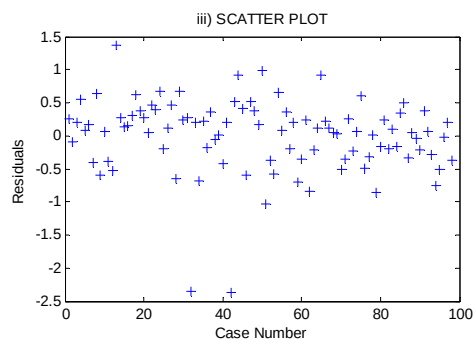
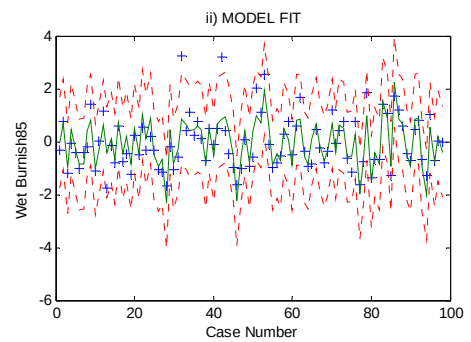
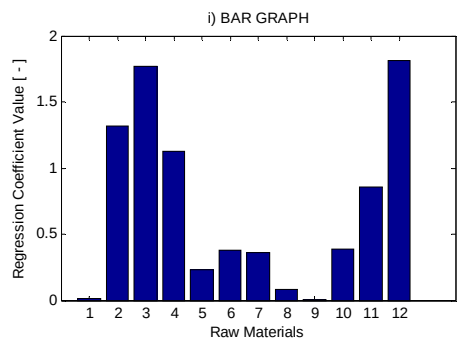
Wet Burnish 20



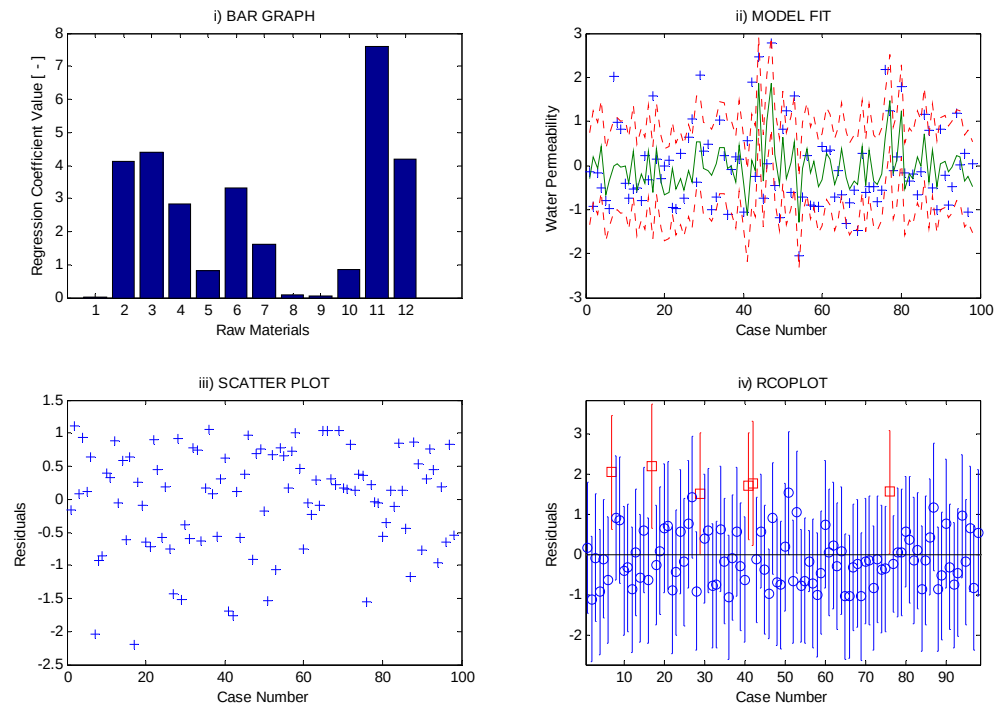
Wet Burnish 60



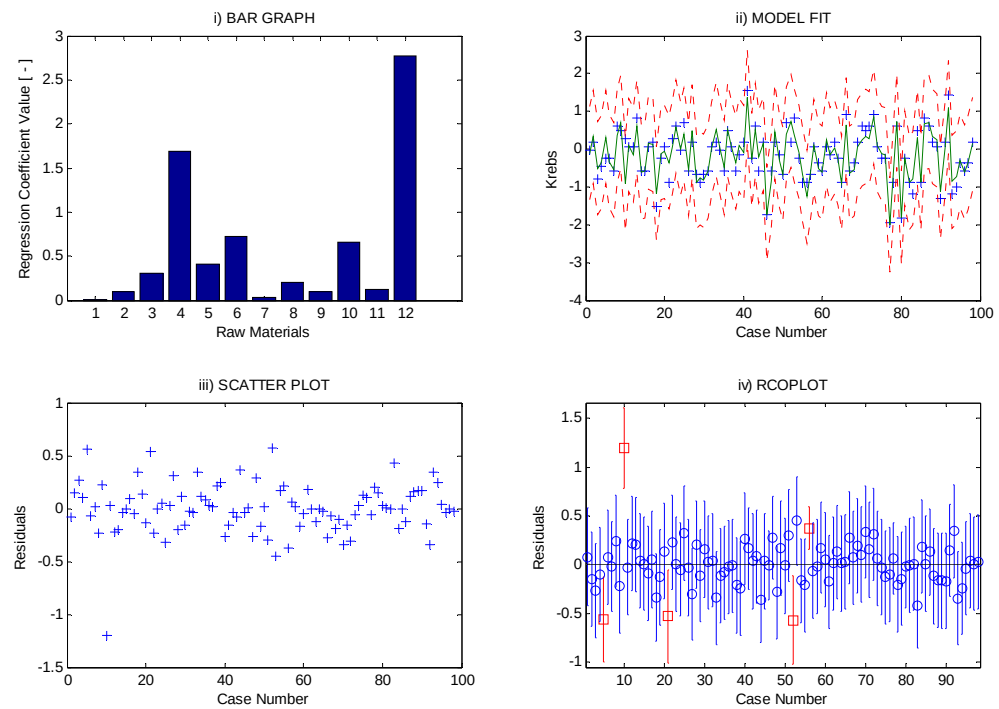
Wet Burnish 85



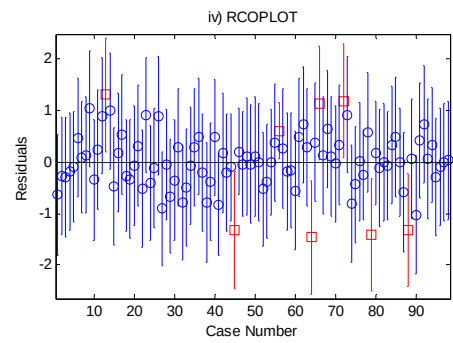
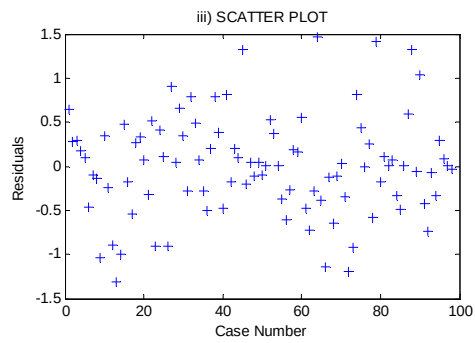
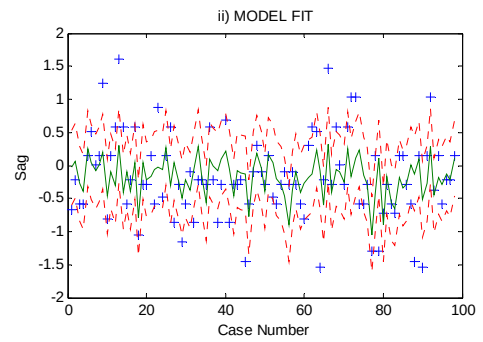
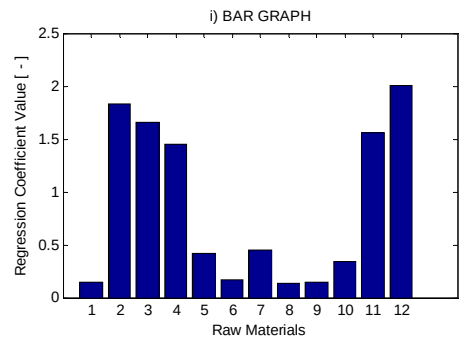
Water Permeability



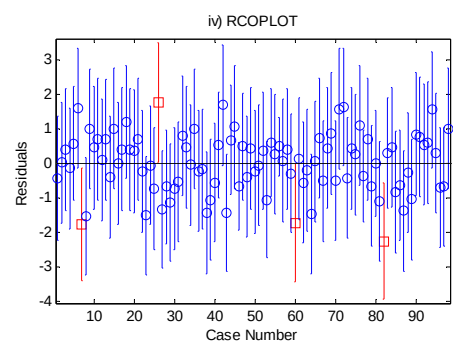
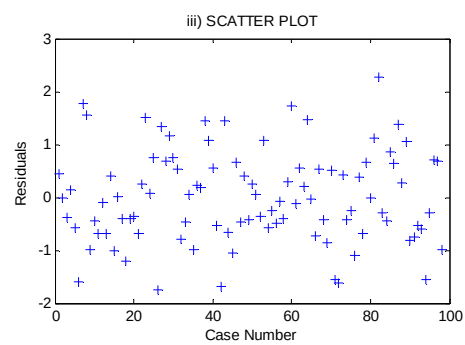
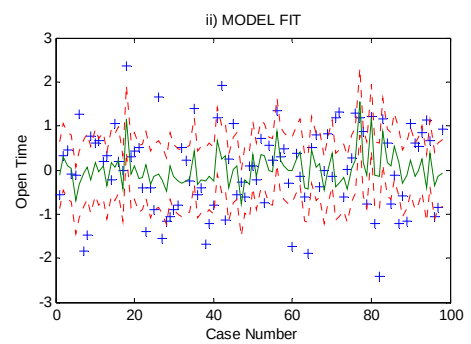
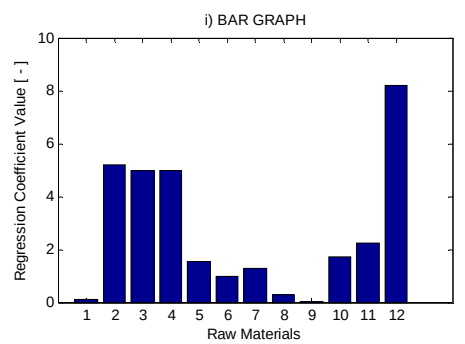
Krebs



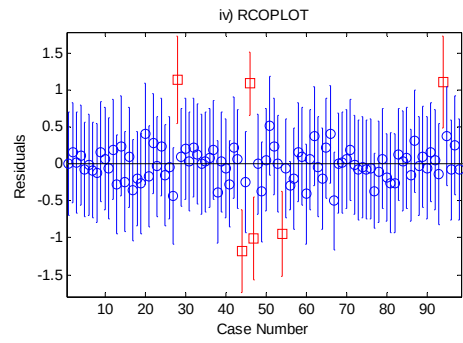
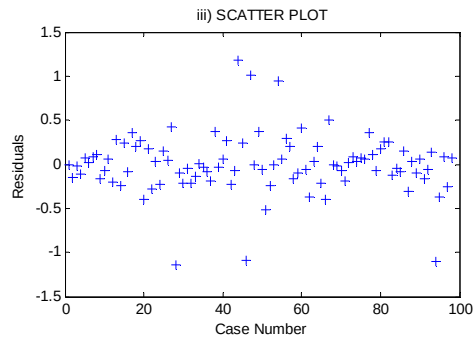
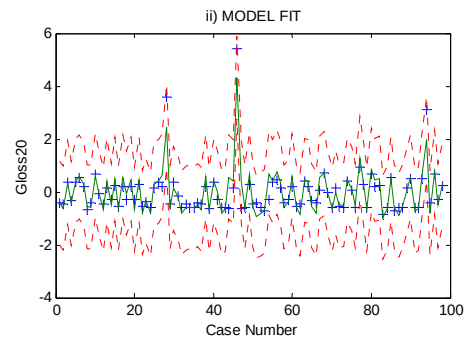
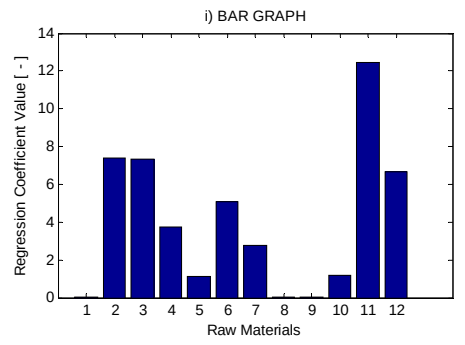
Sag



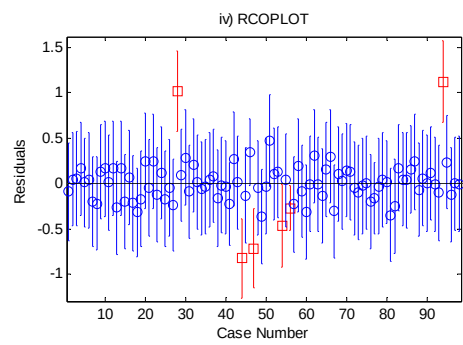
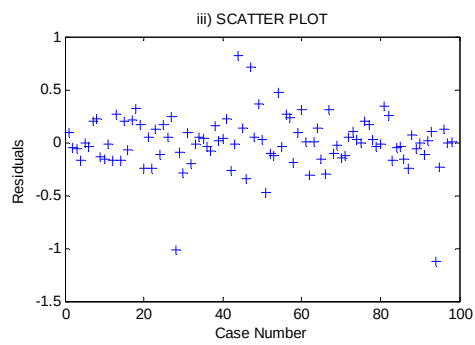
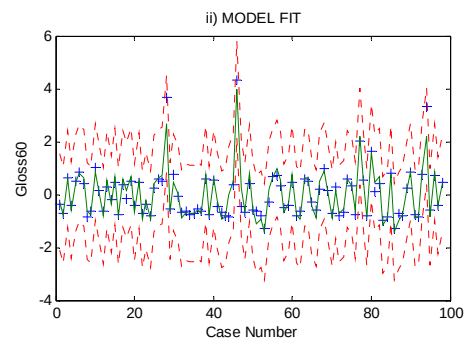
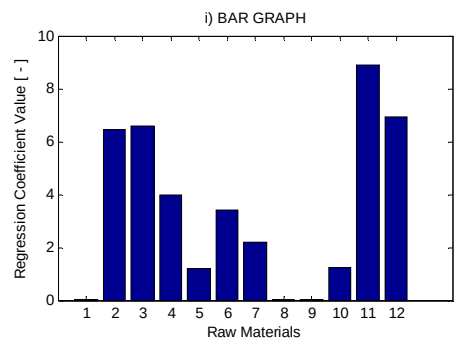
Open time



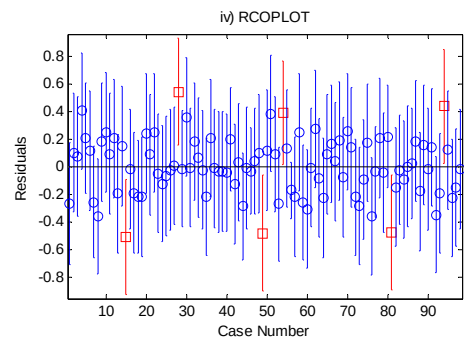
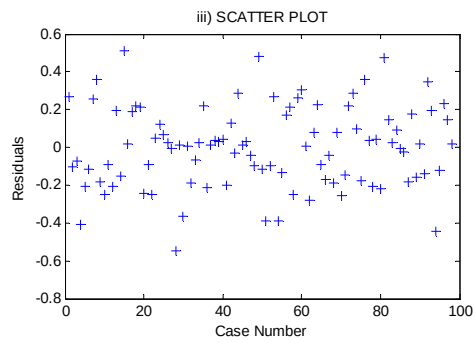
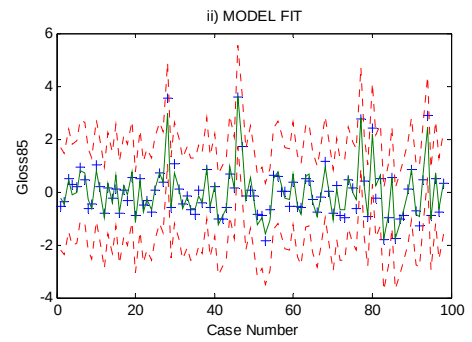
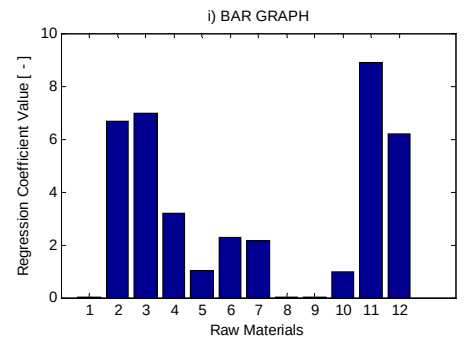
Gloss 20



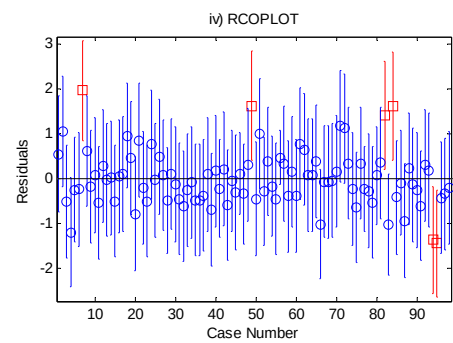
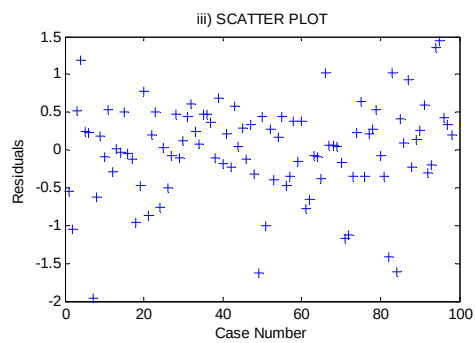
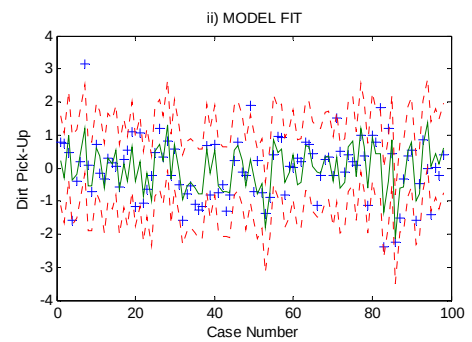
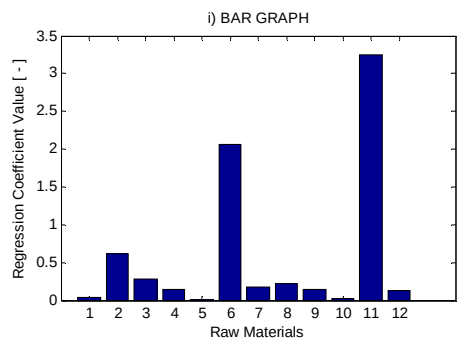
Gloss 60



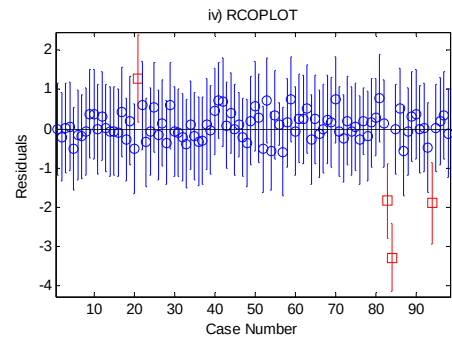
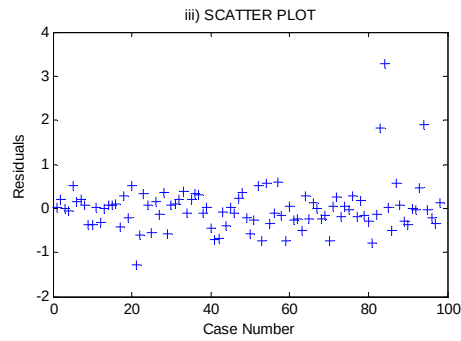
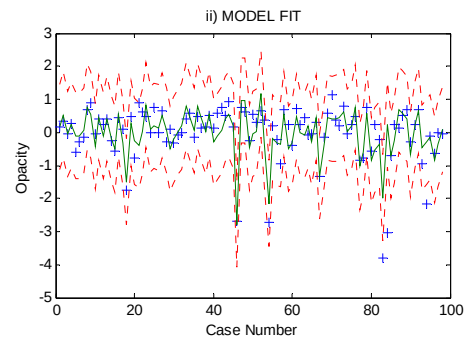
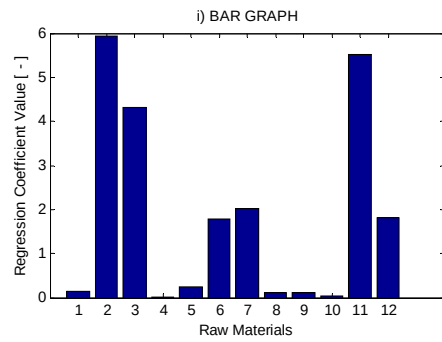
Gloss 85



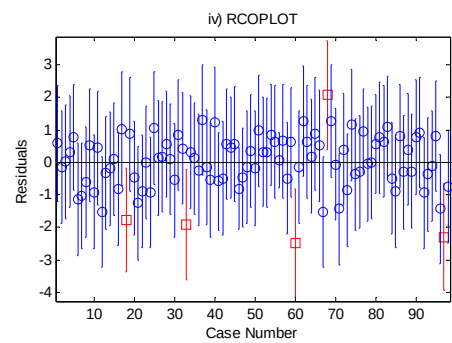
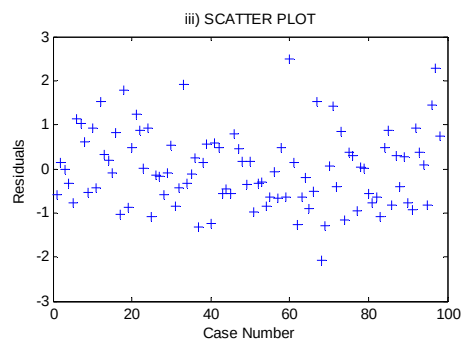
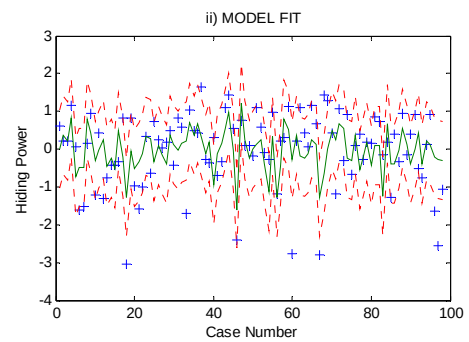
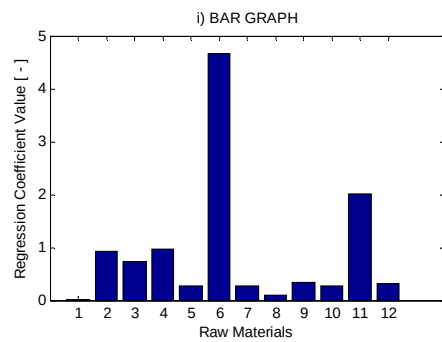
Dirt Pick-Up



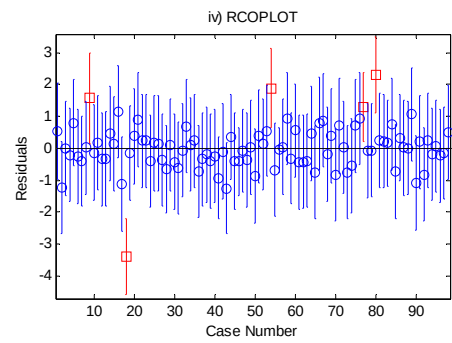
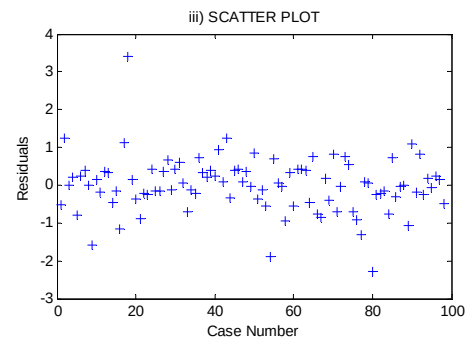
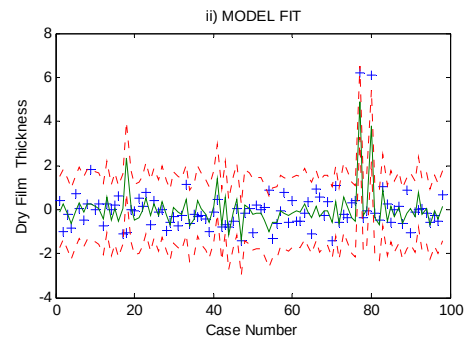
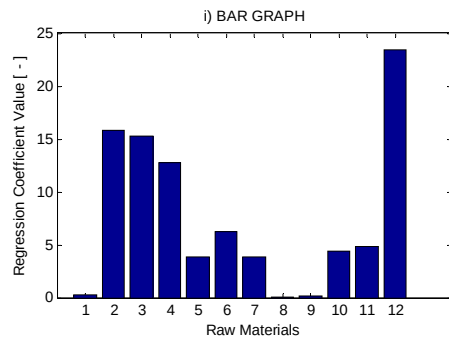
Opacity



Hiding Power

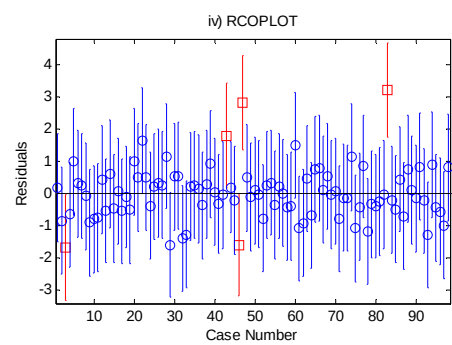
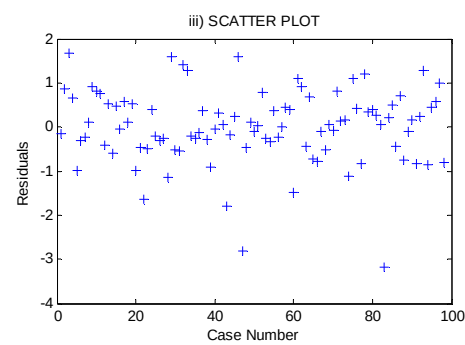
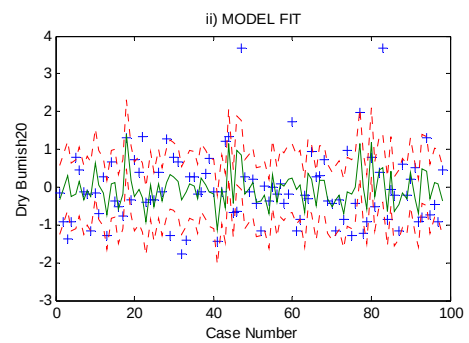
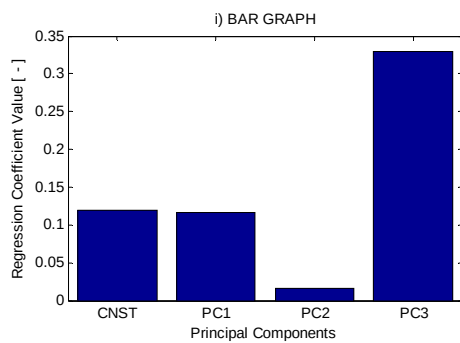


Dry Film Thickness

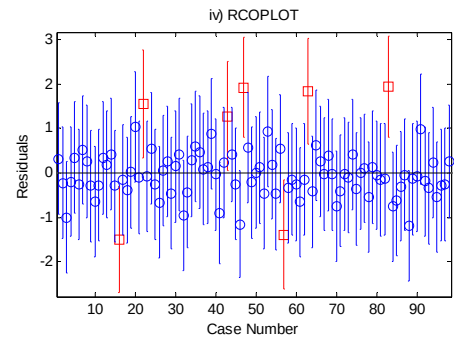
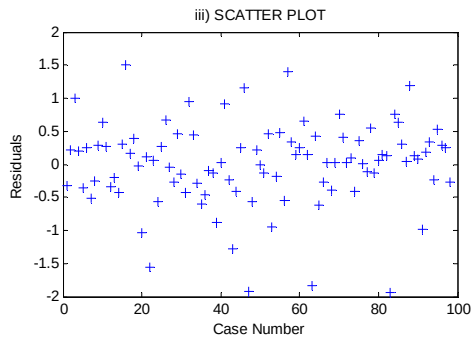
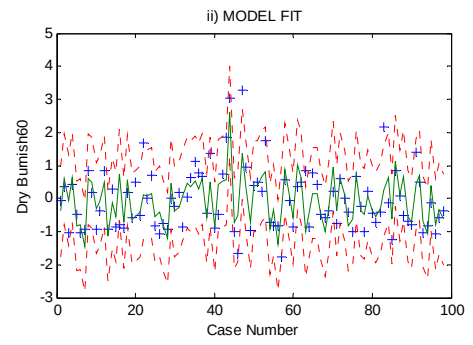
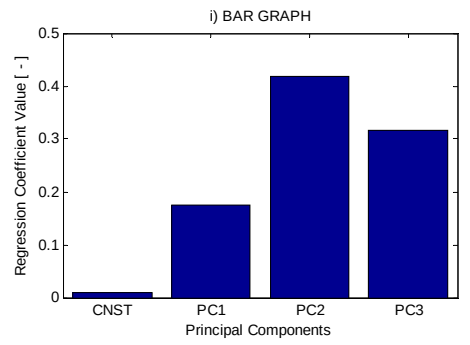


12.4 LINEAR MODELLING (MODEL 3)

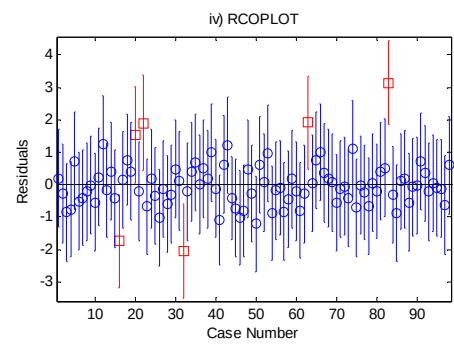
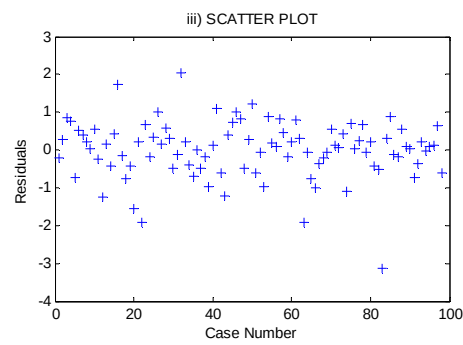
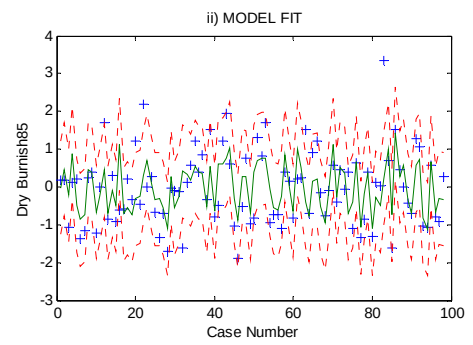
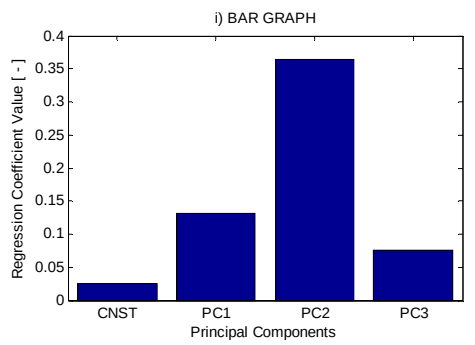
Dry Burnish 20



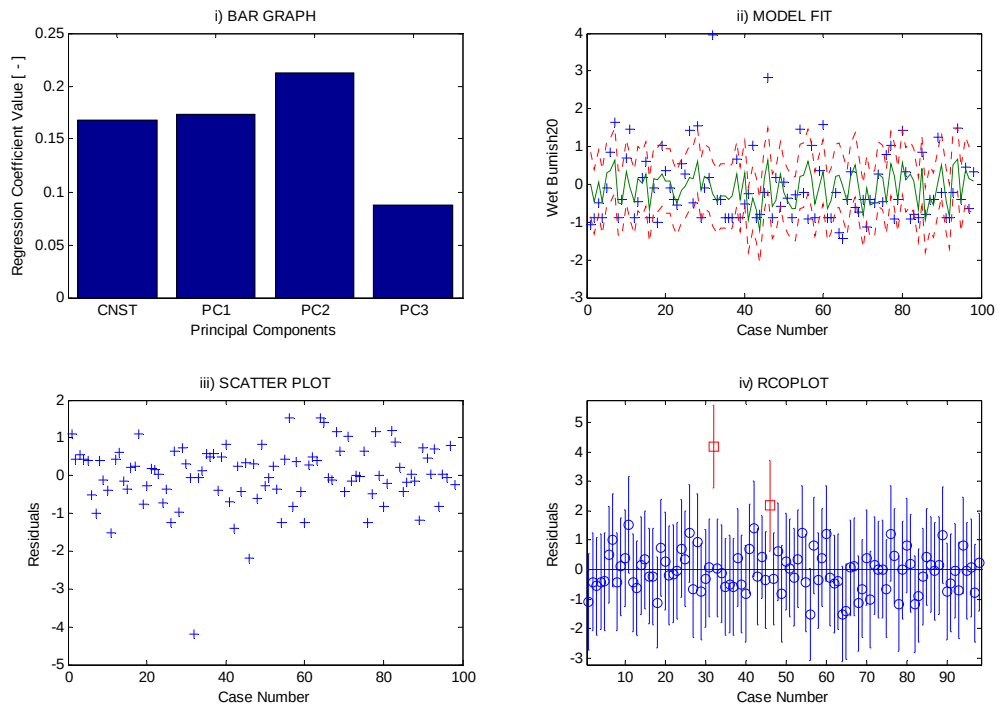
Dry Burnish 60



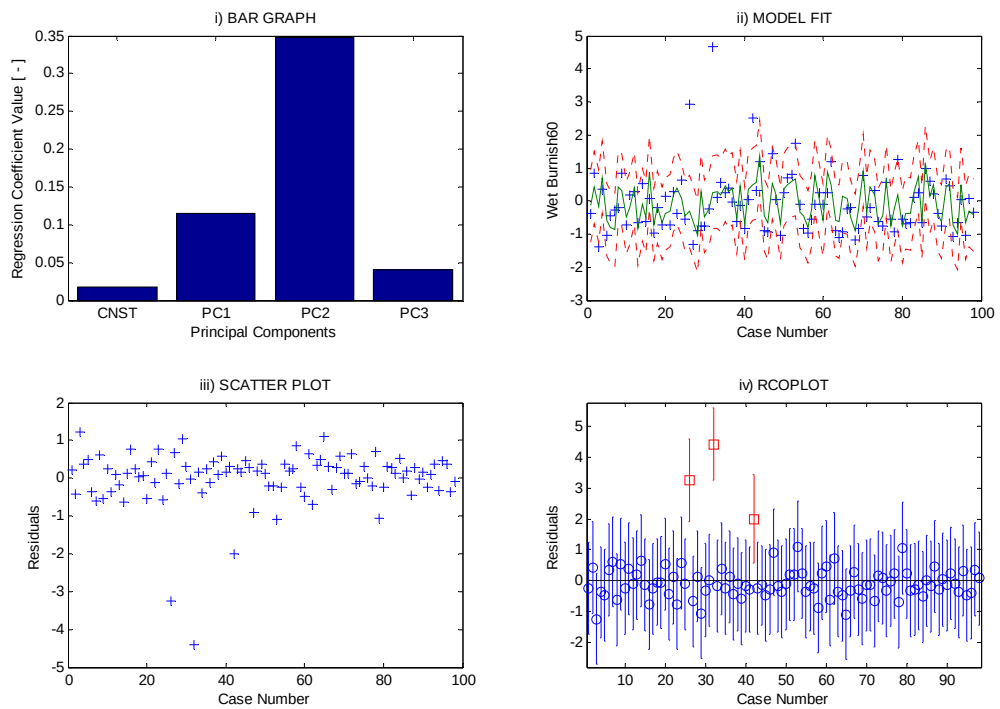
Dry Burnish 85



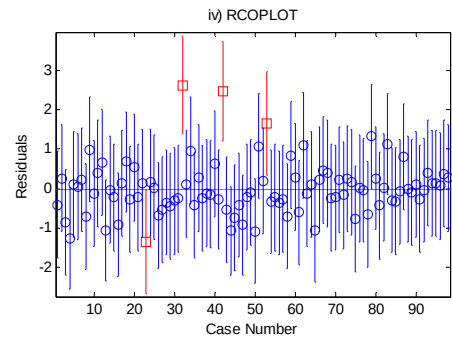
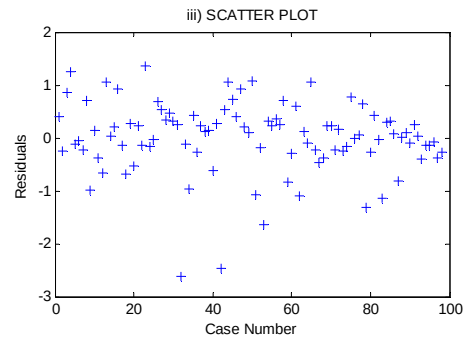
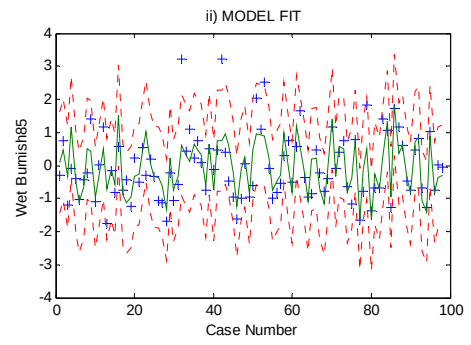
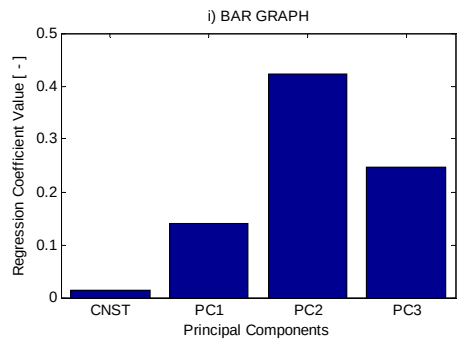
Wet Burnish 20



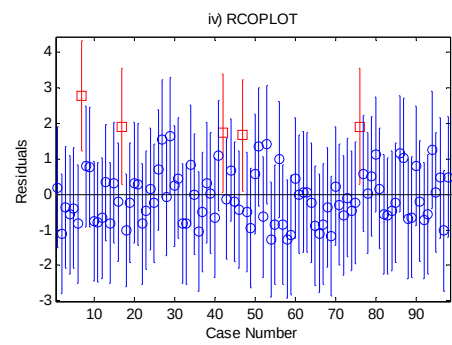
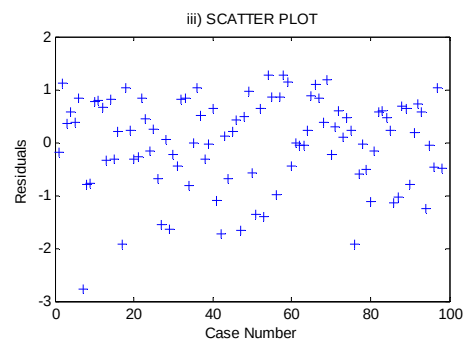
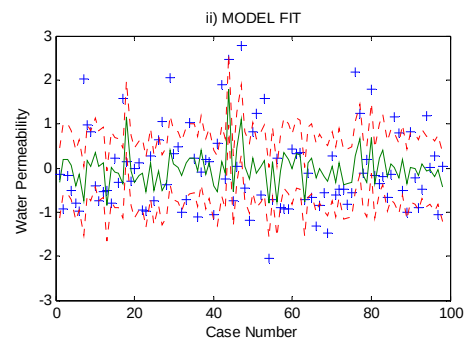
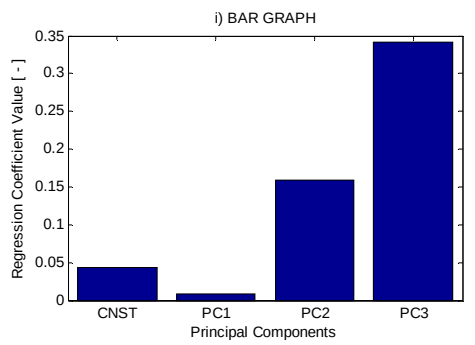
Wet Burnish 60



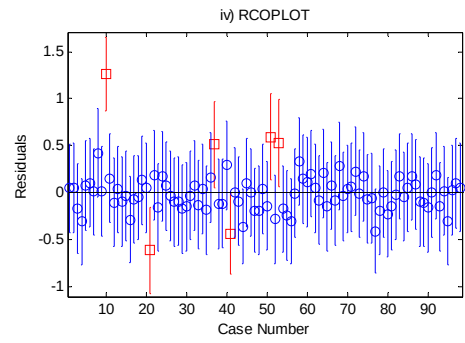
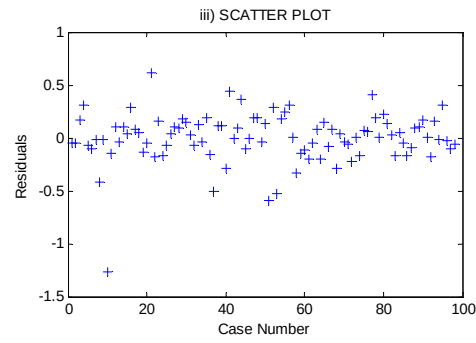
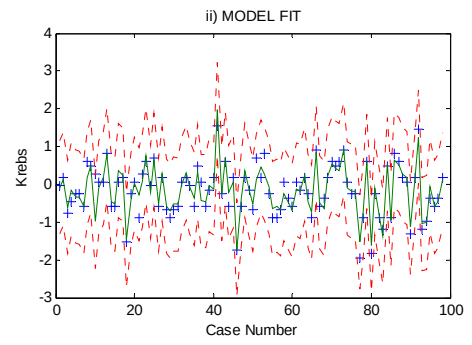
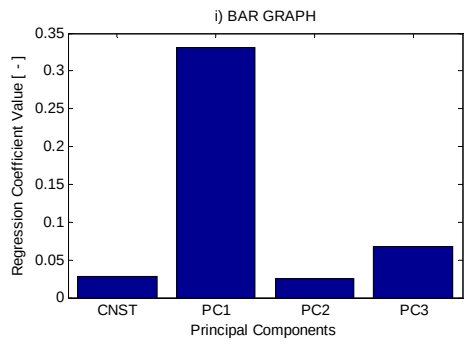
Wet Burnish 85



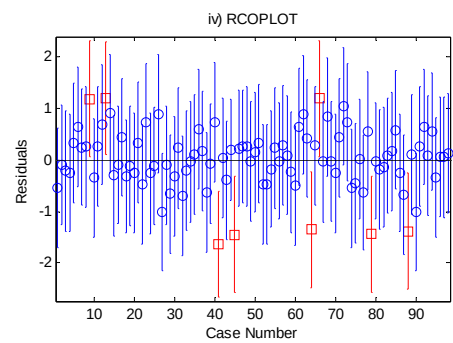
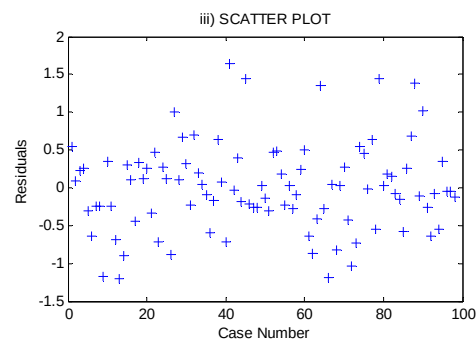
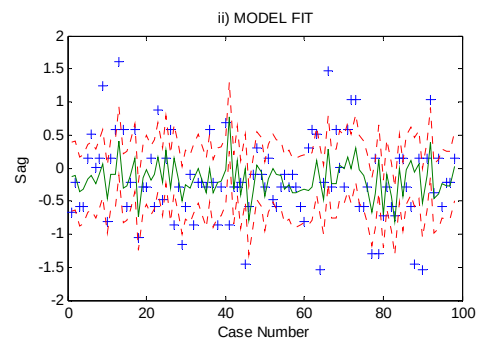
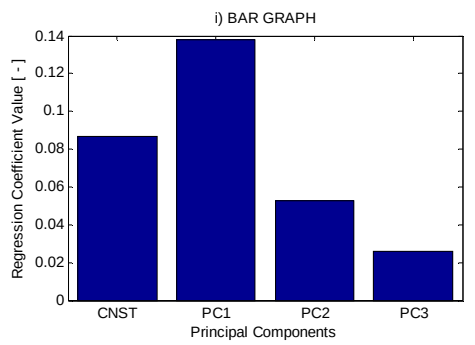
Water Permeability



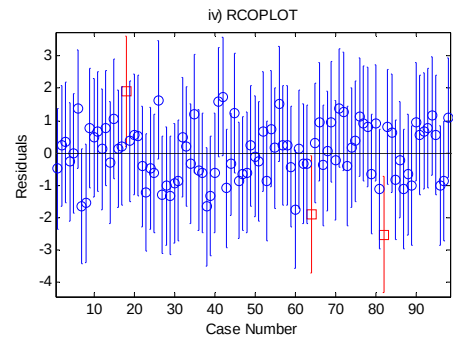
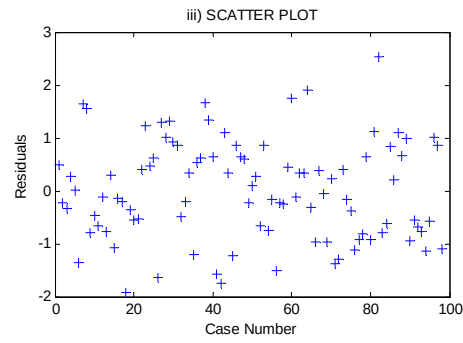
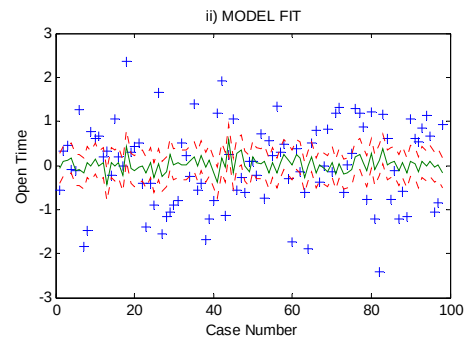
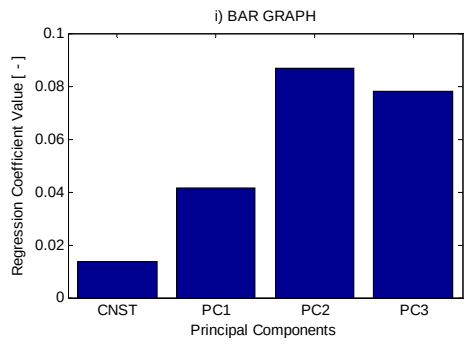
Krebs



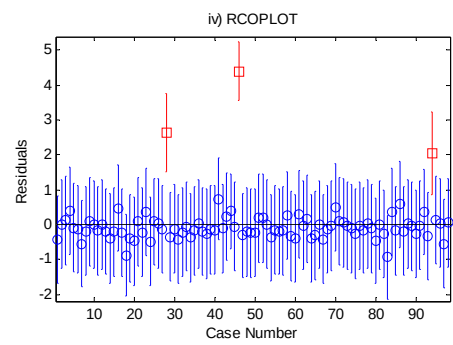
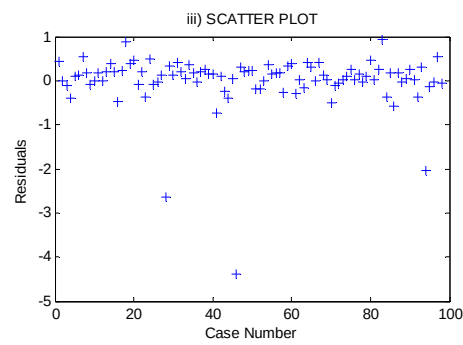
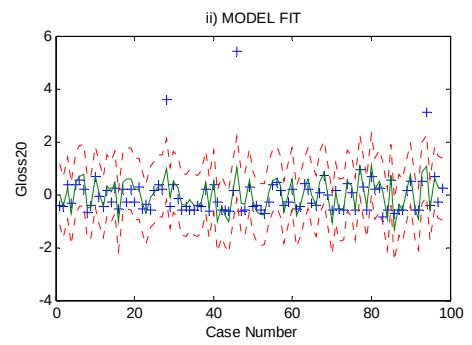
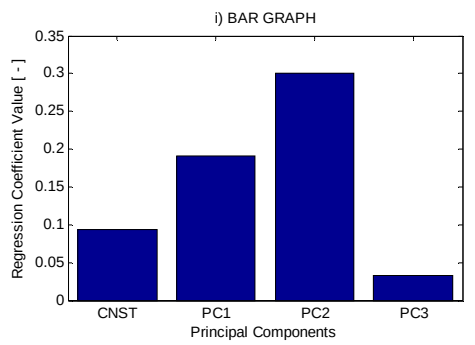
Sag



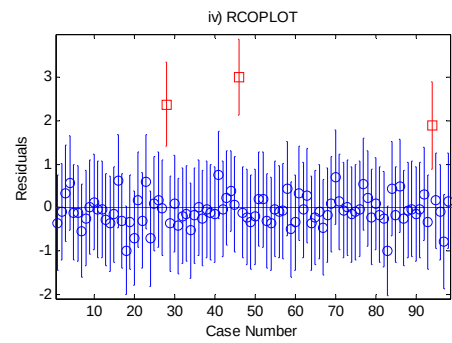
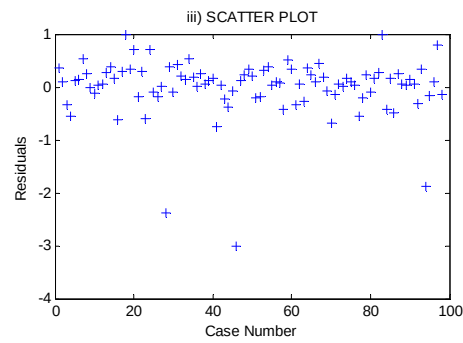
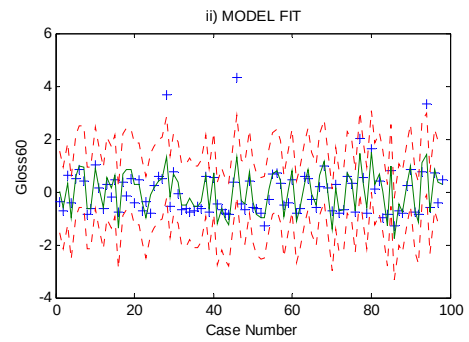
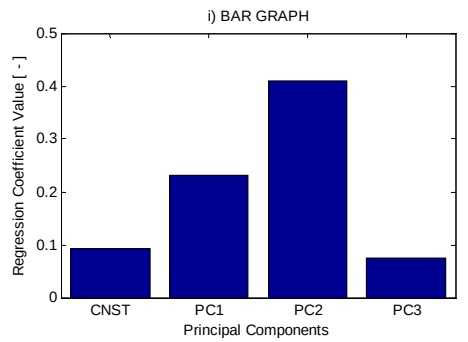
Open time



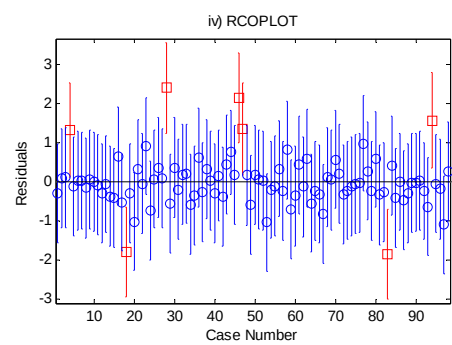
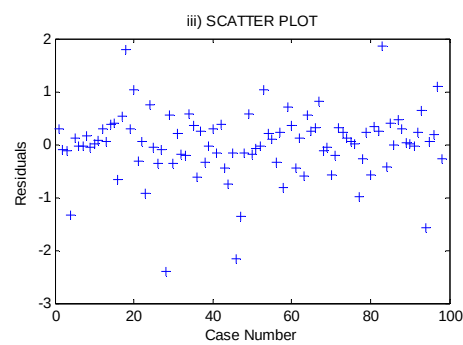
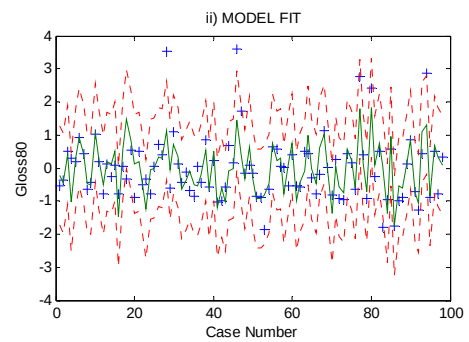
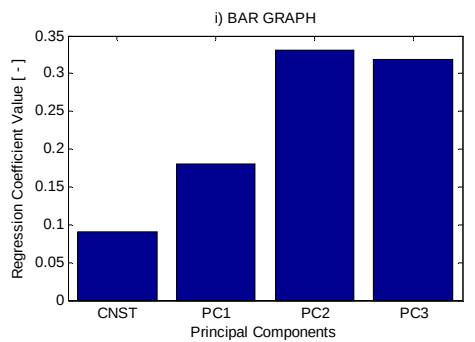
Gloss 20



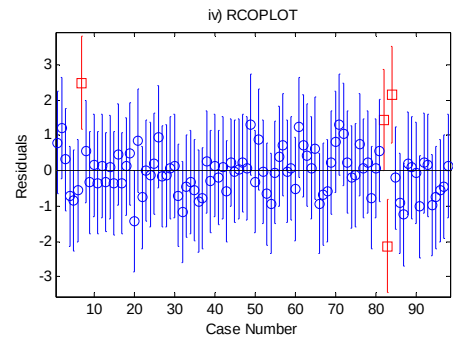
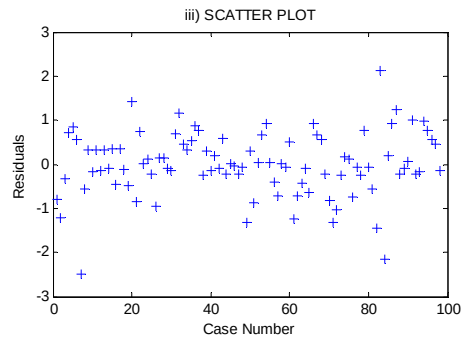
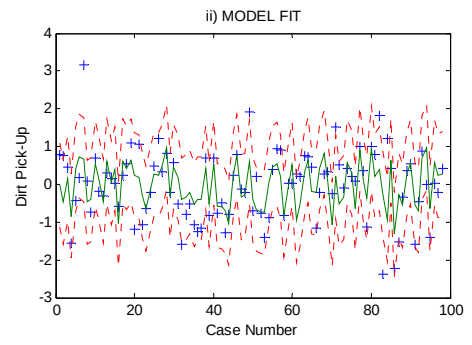
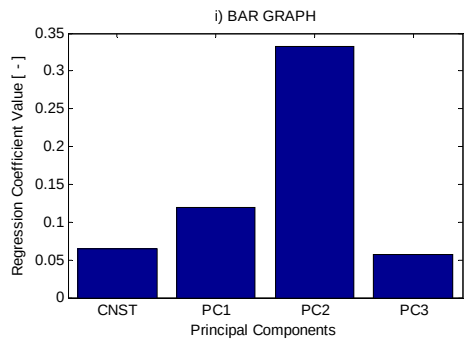
Gloss 60



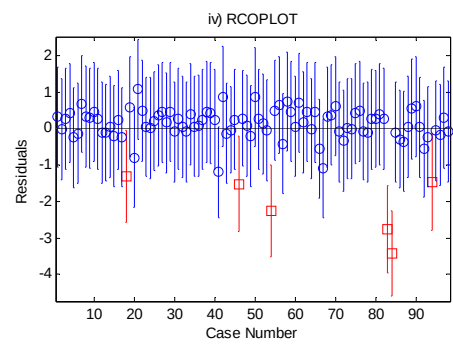
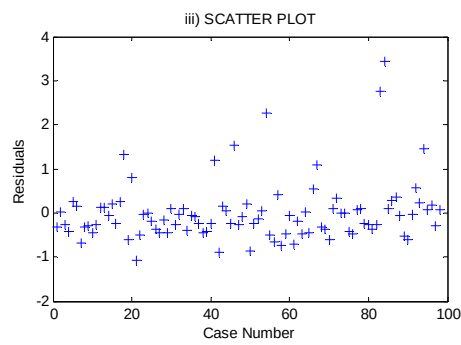
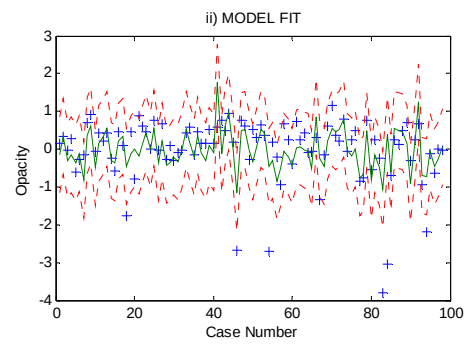
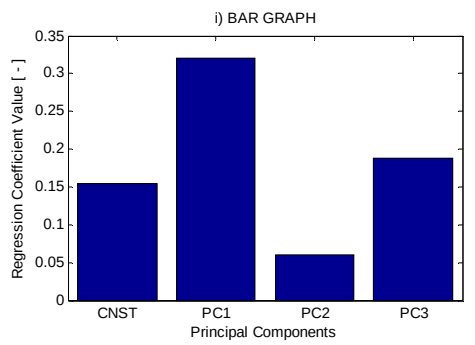
Gloss 85



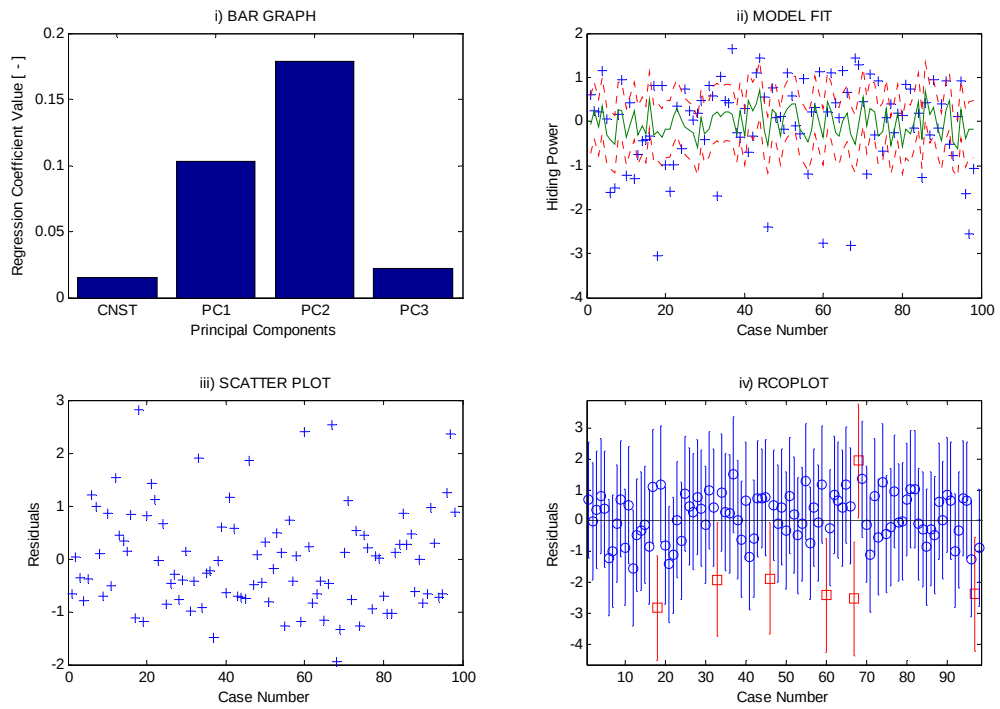
Dirt Pick-Up



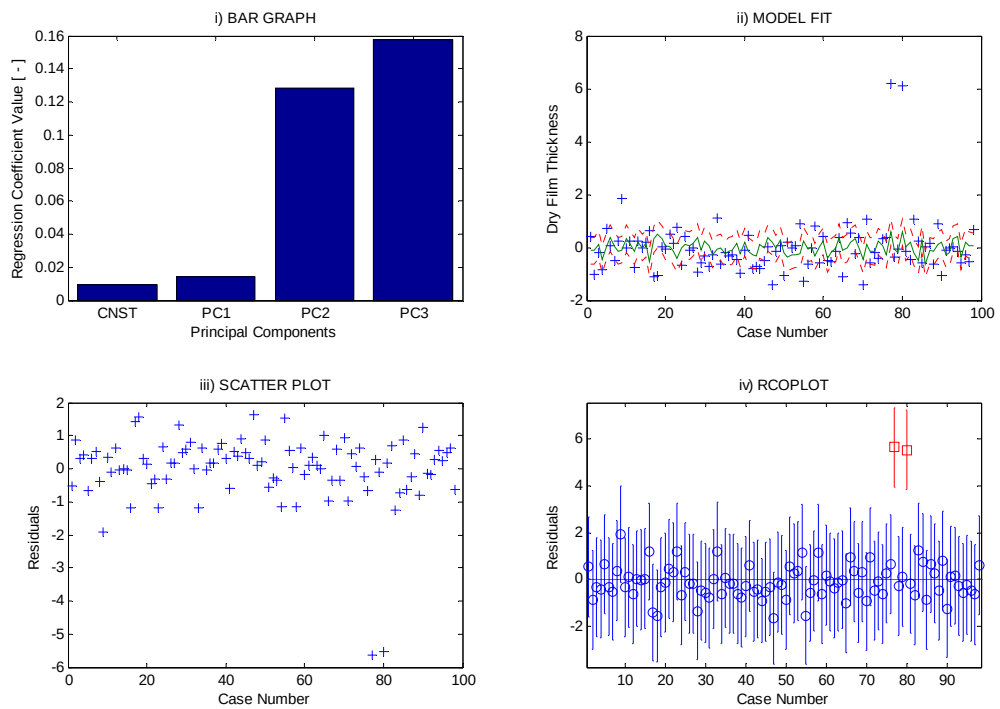
Opacity



Hiding Power



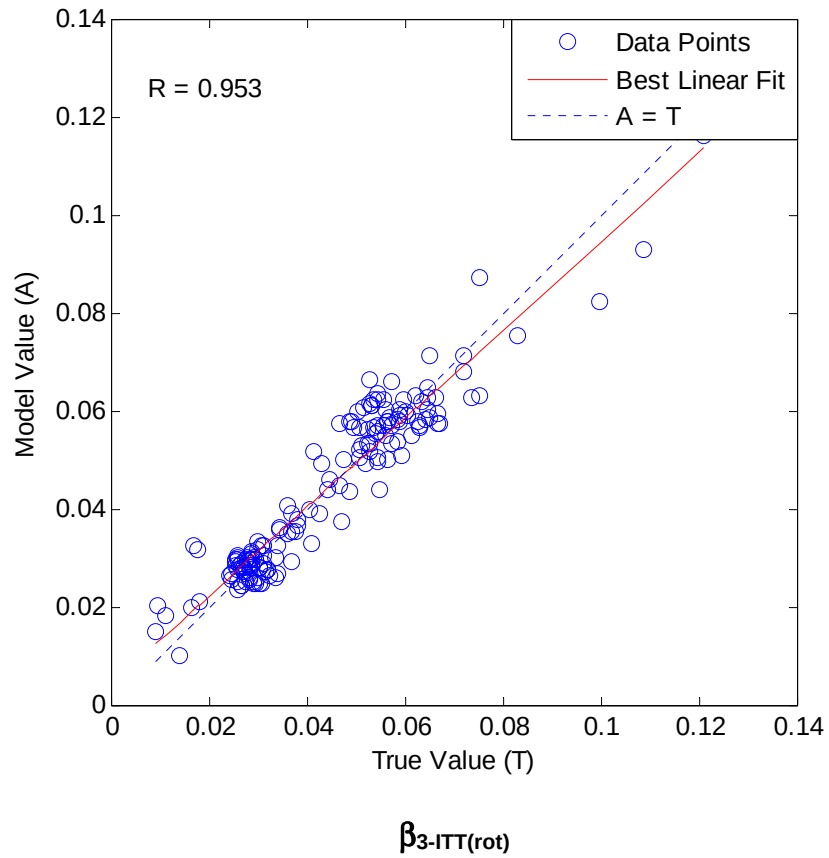
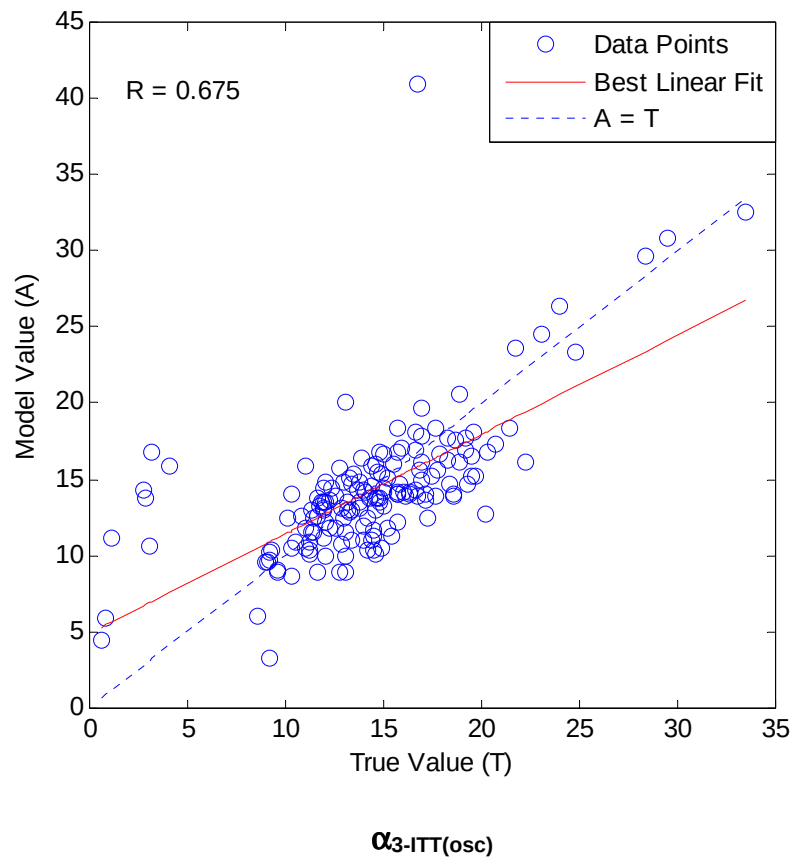
Dry Film Thickness

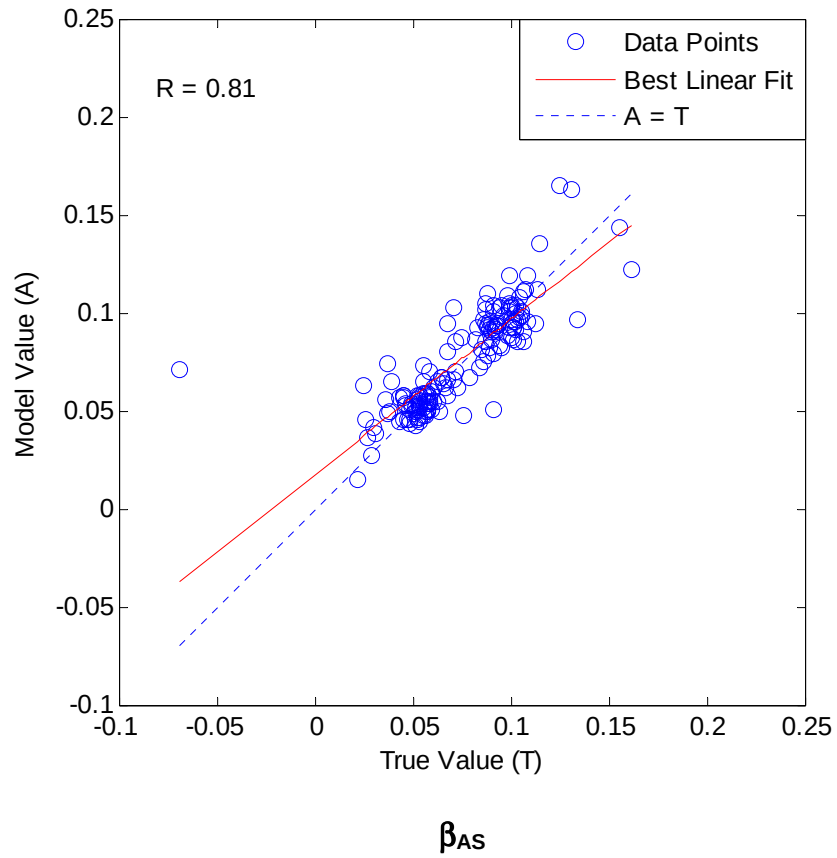
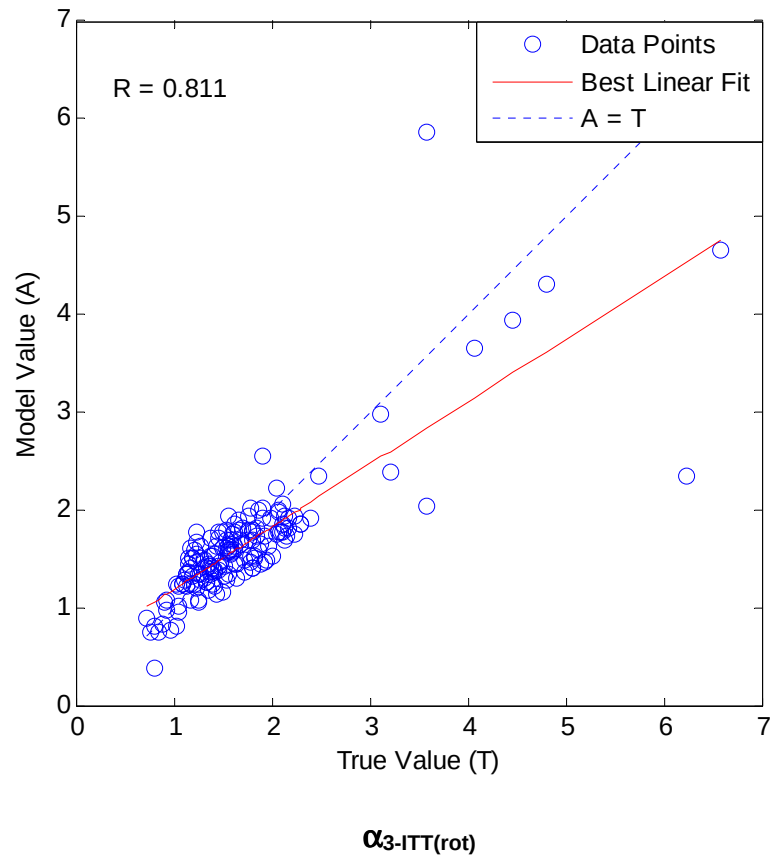


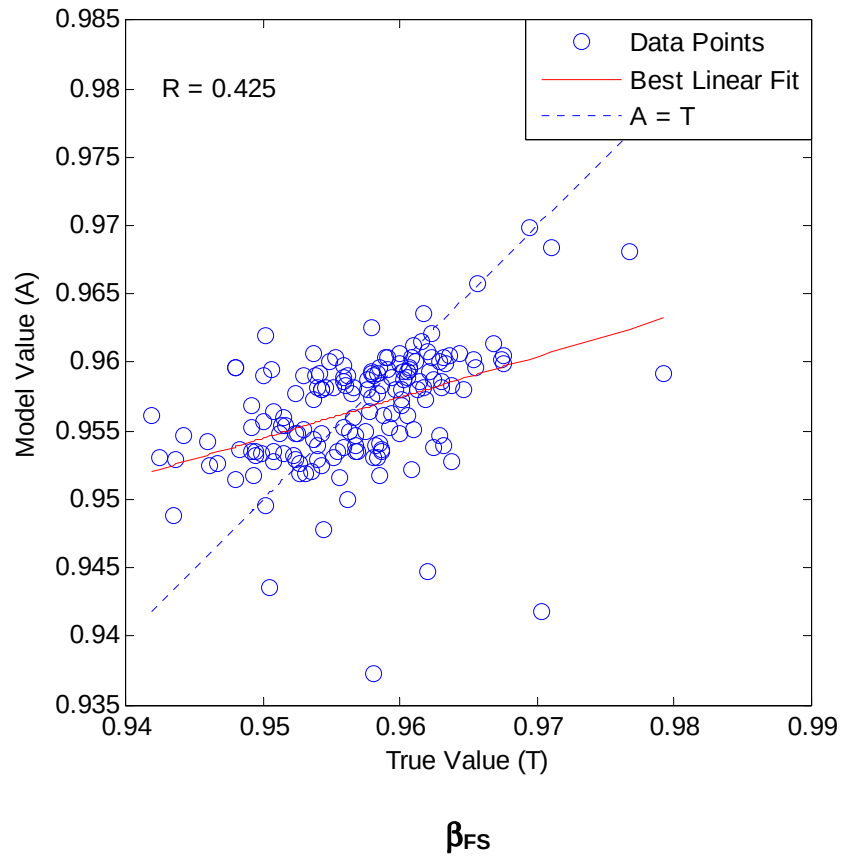
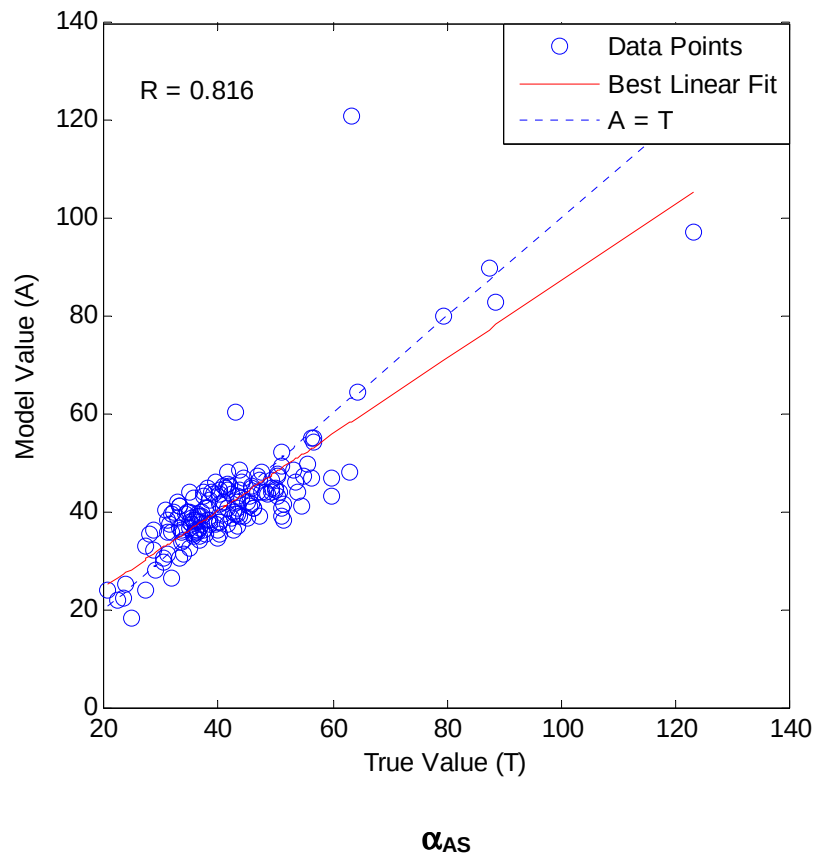
12.5 NON-LINEAR MODELLING (MODEL 1)

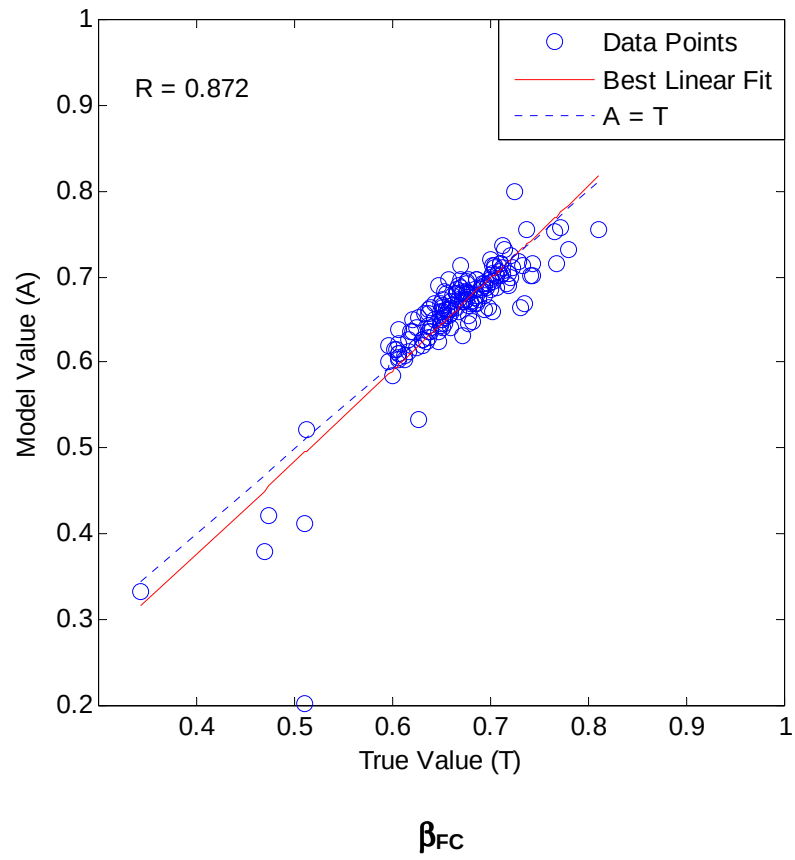
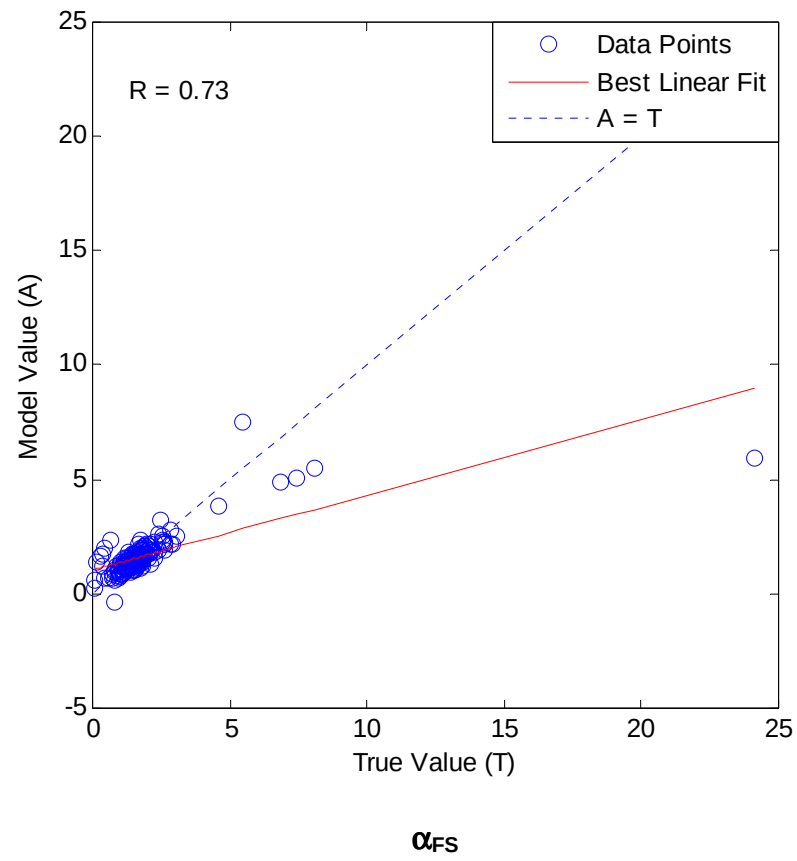
'ACTUAL vs MODEL' graphs are given below for each of the rheological parameters.

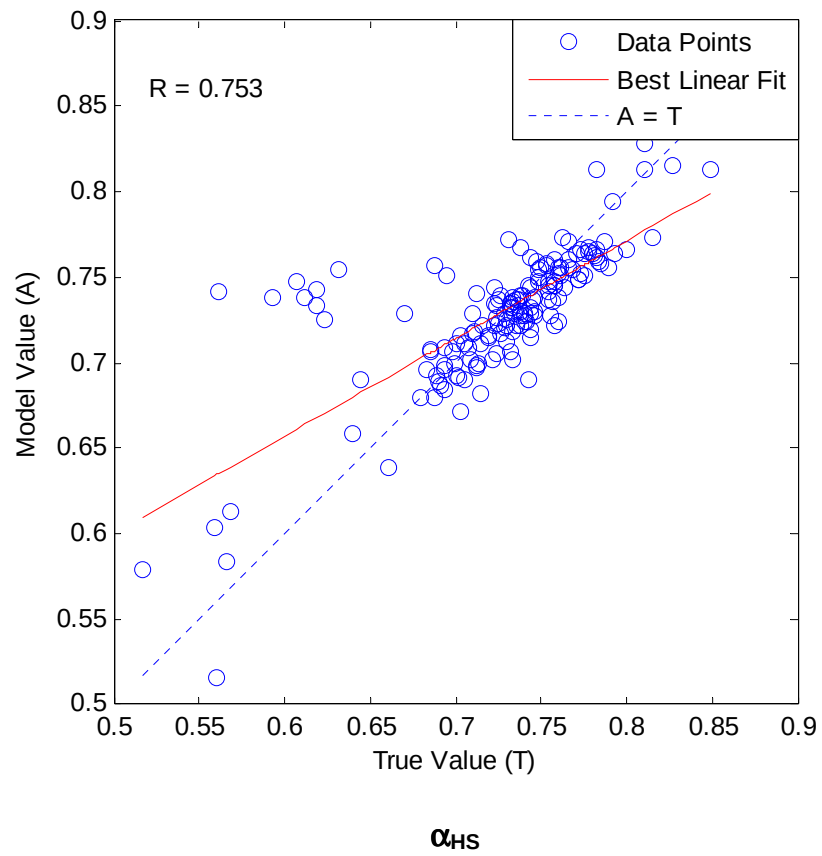
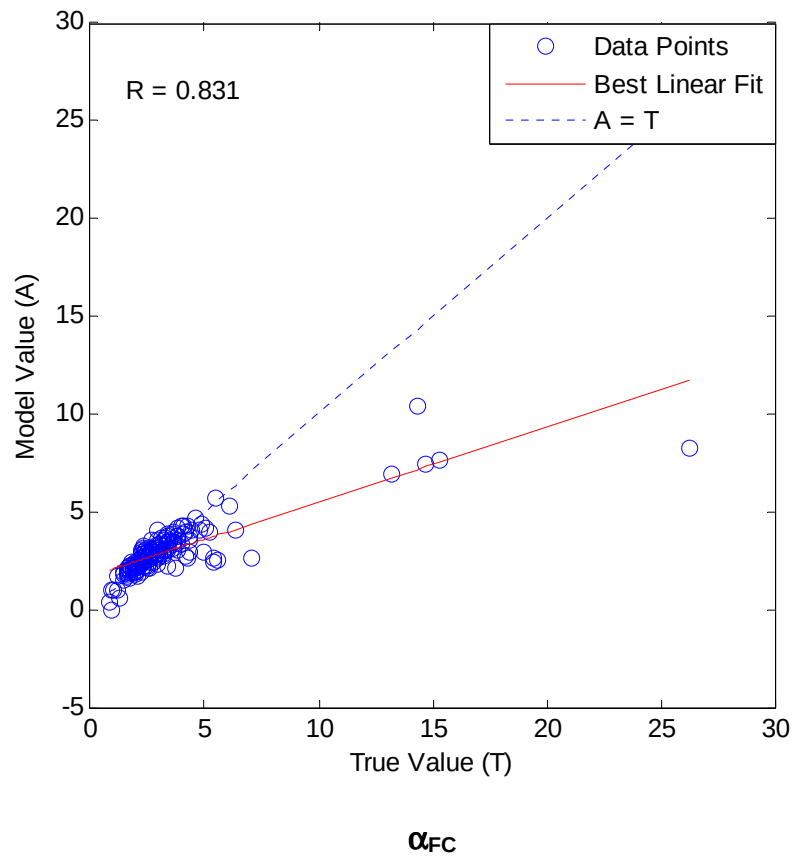
$$\beta_{3-ITT(osc)}$$

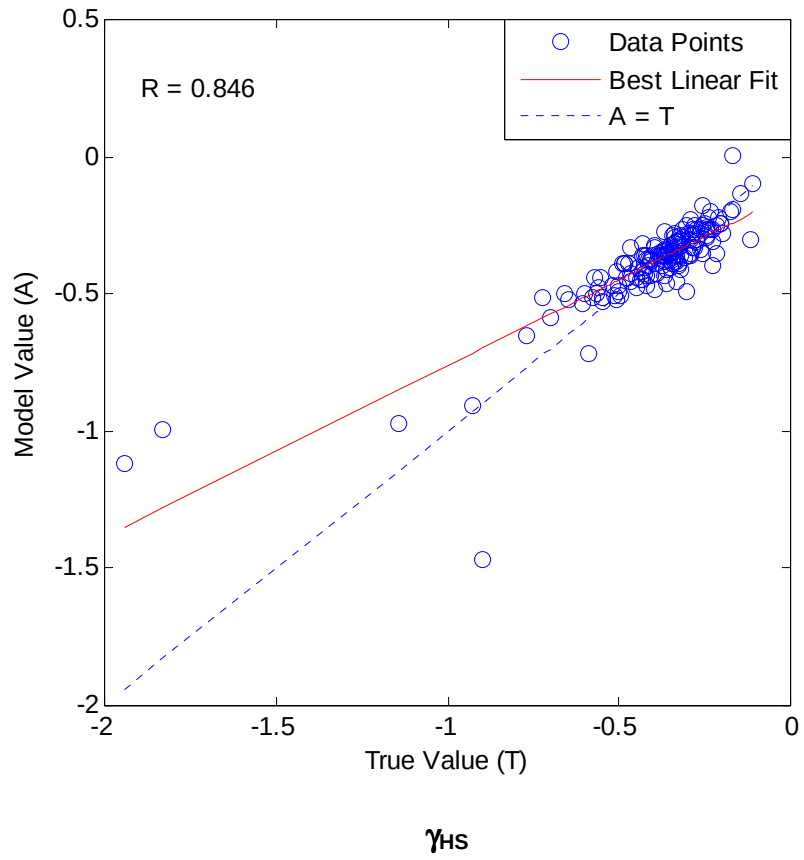
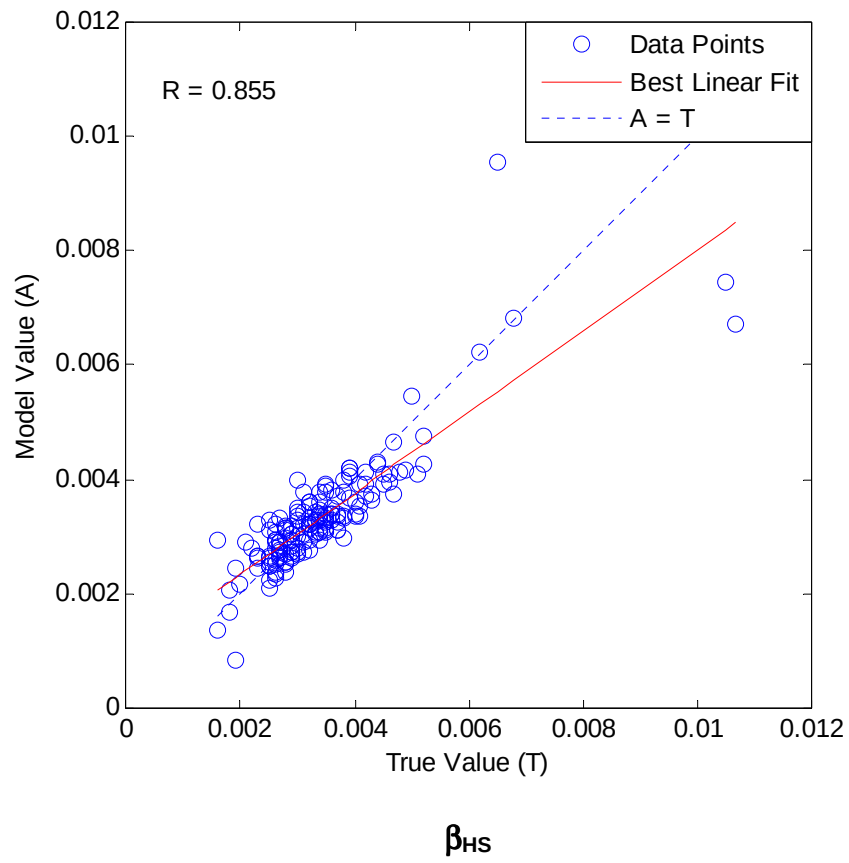


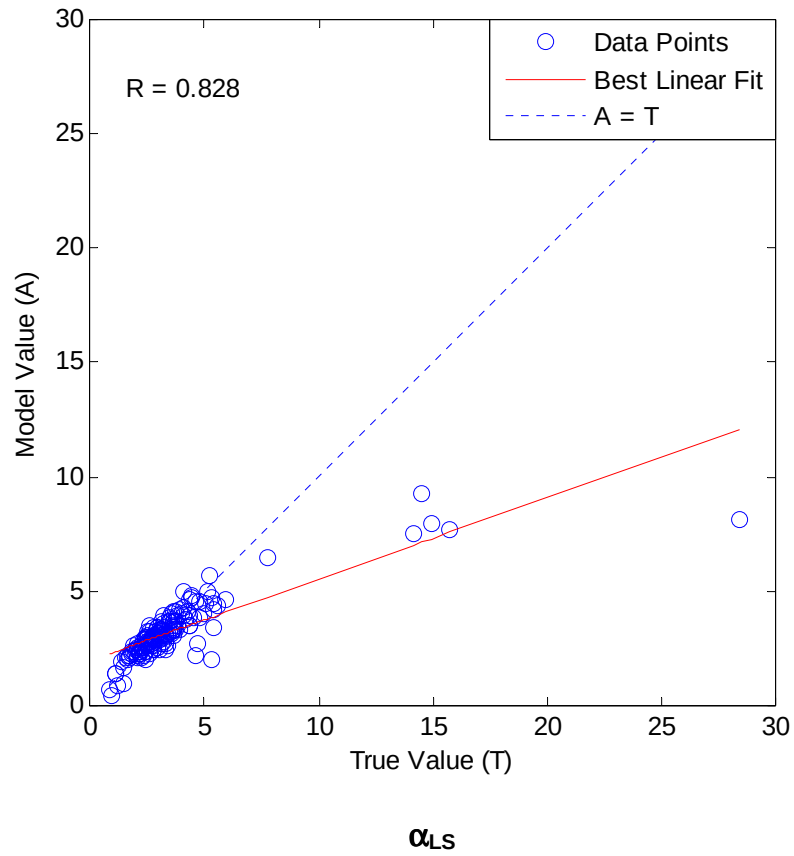
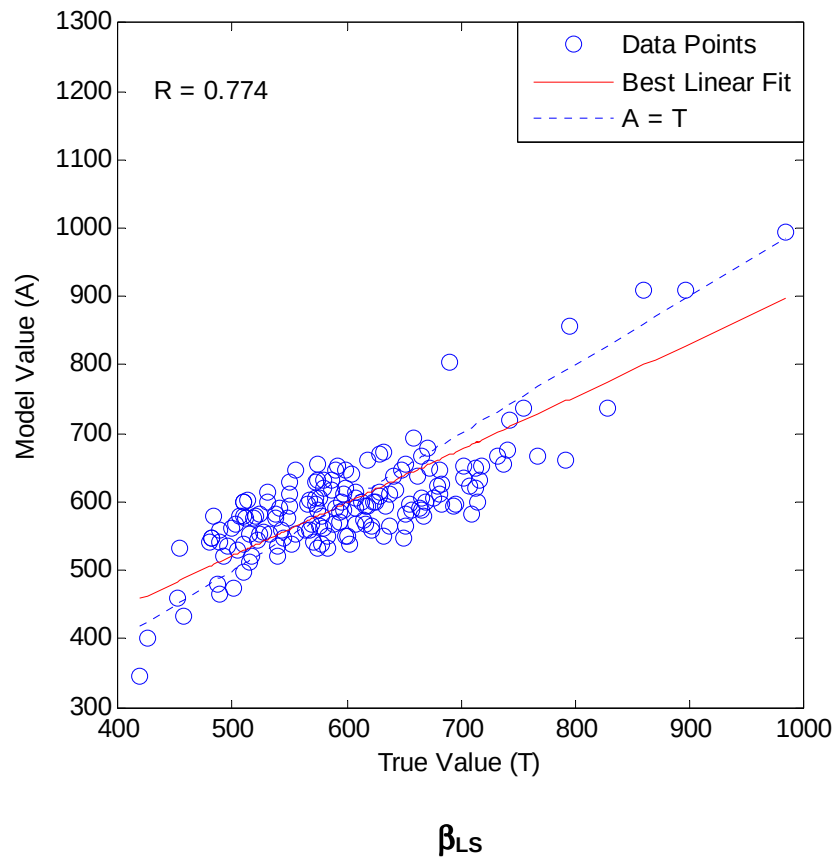


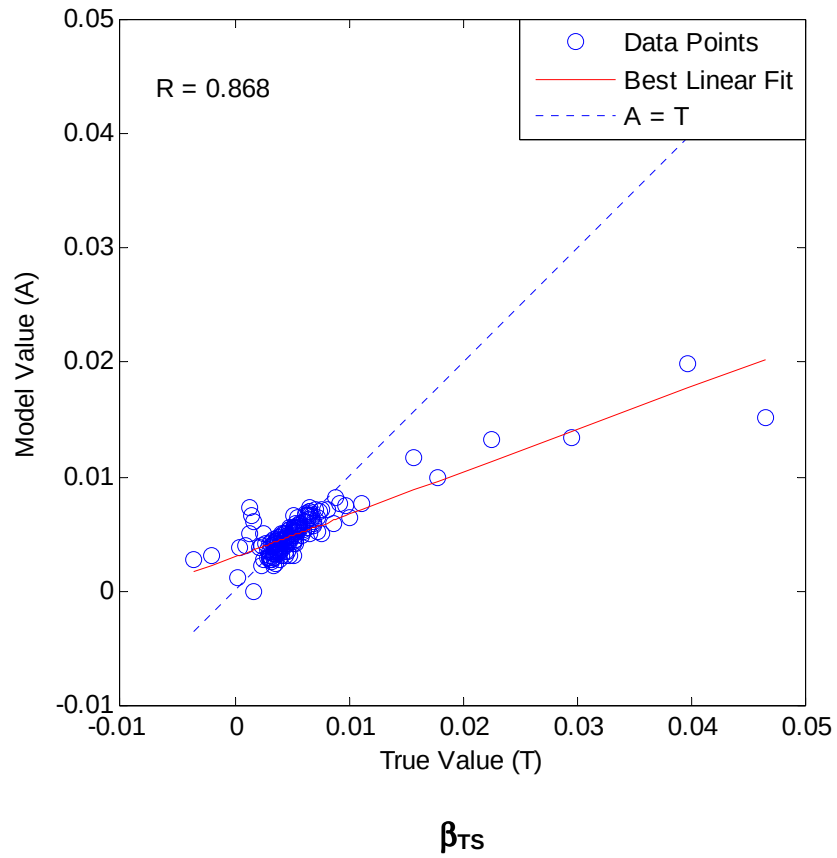
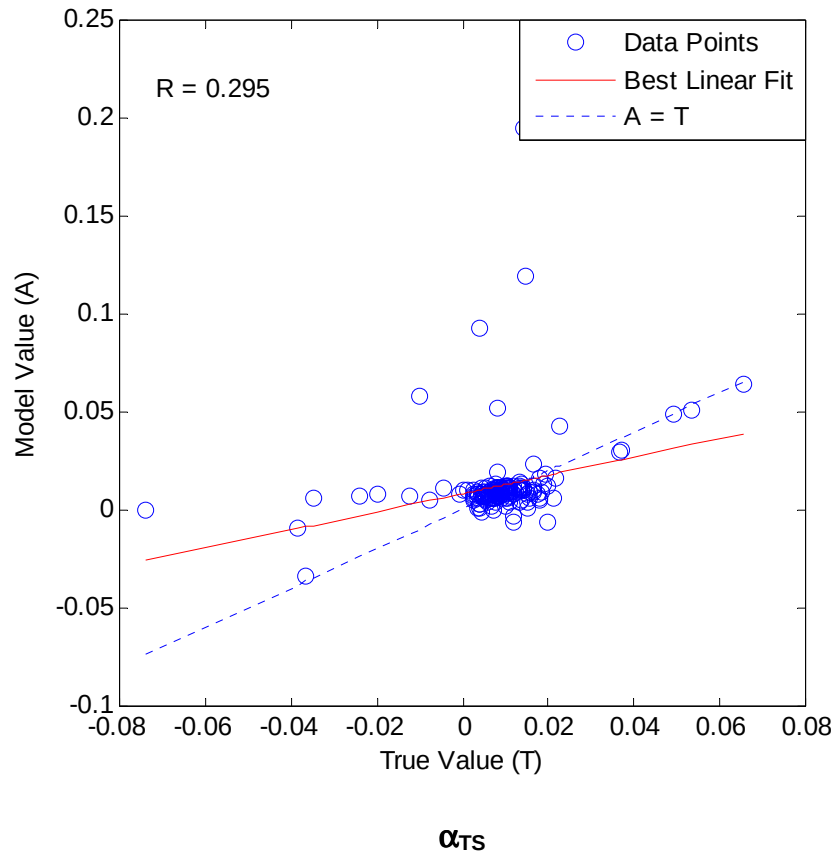


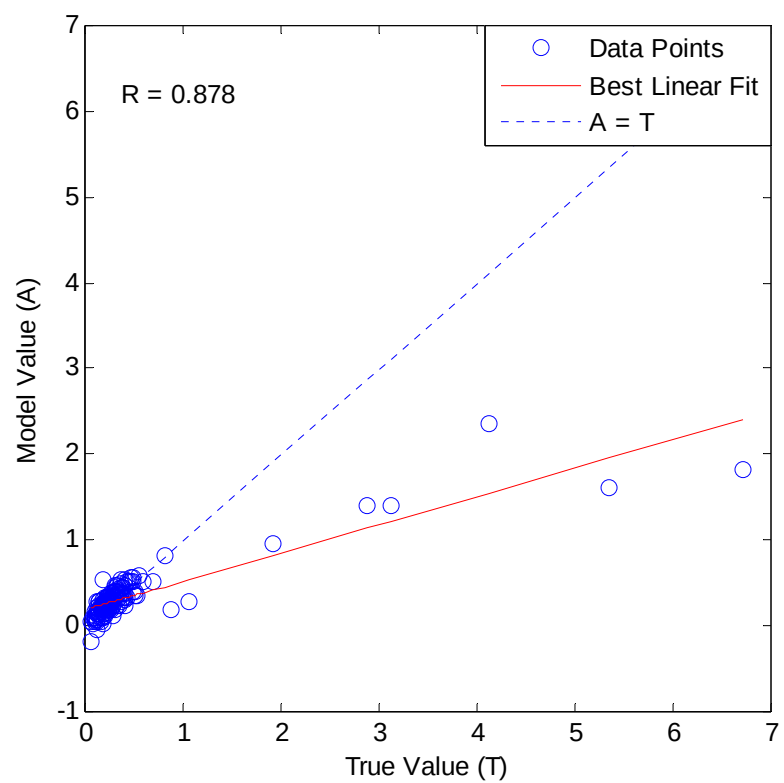
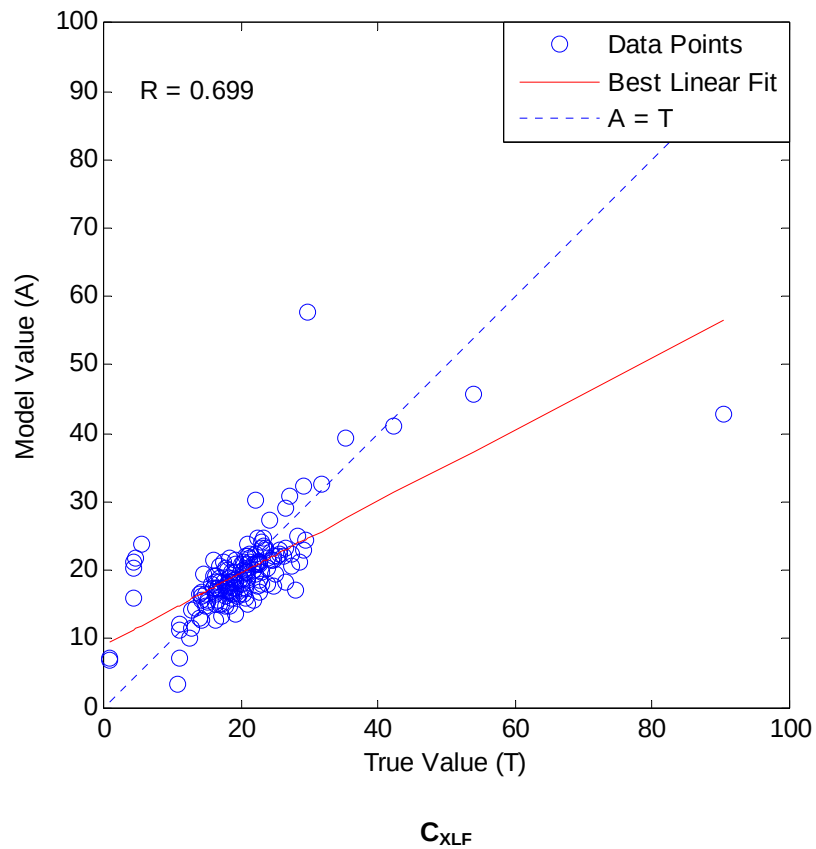






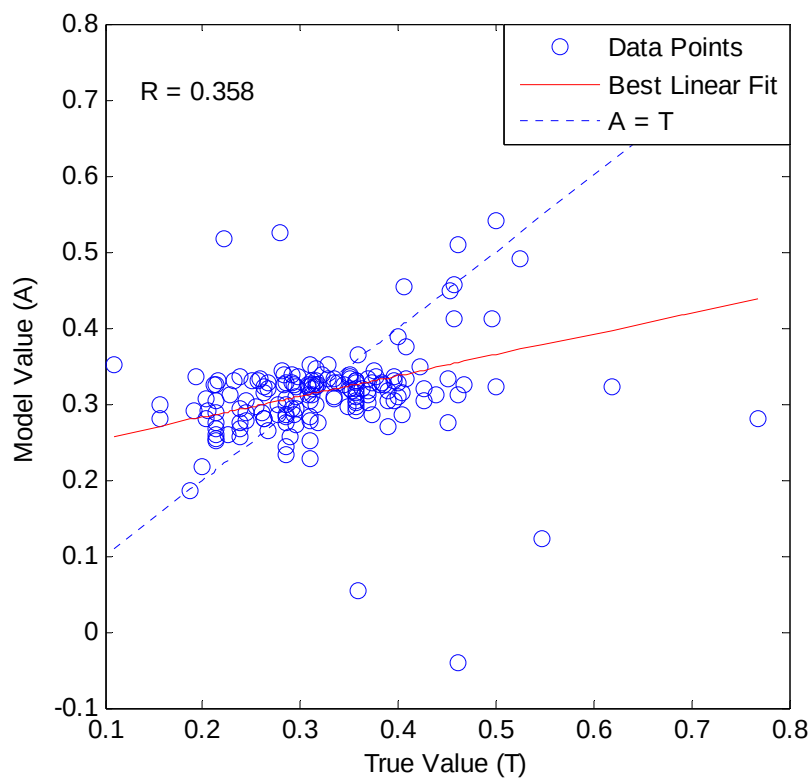
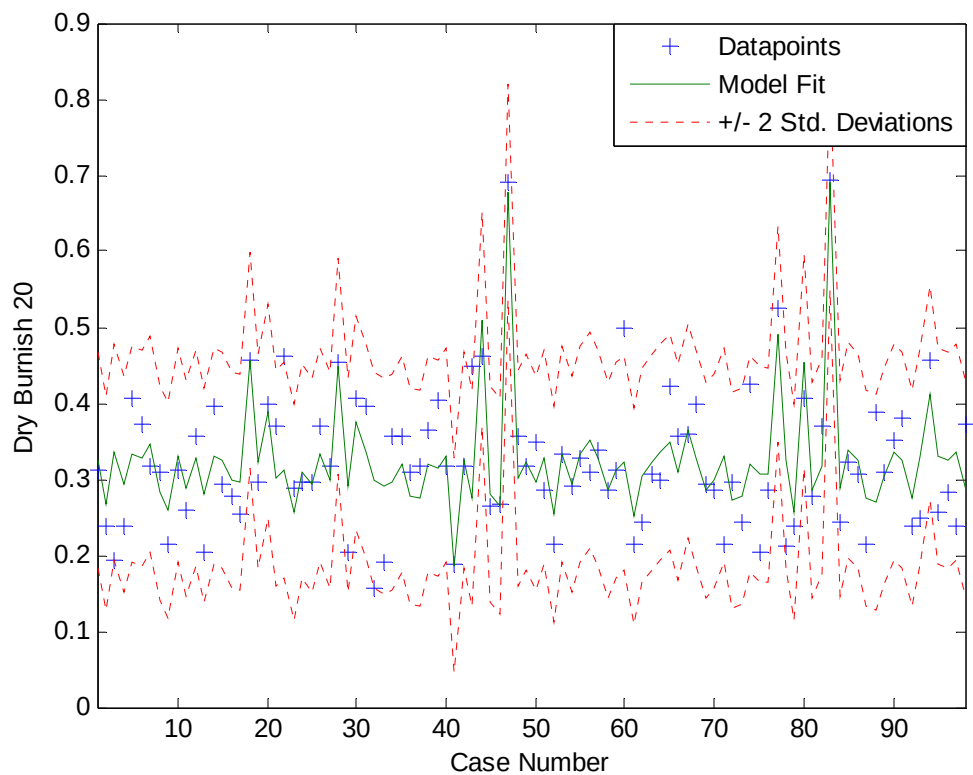




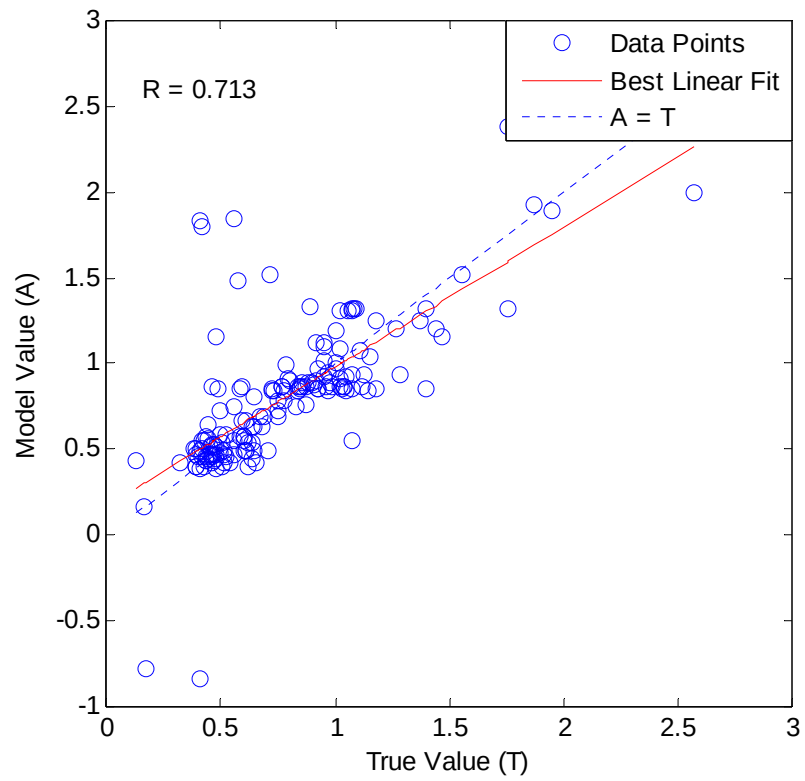
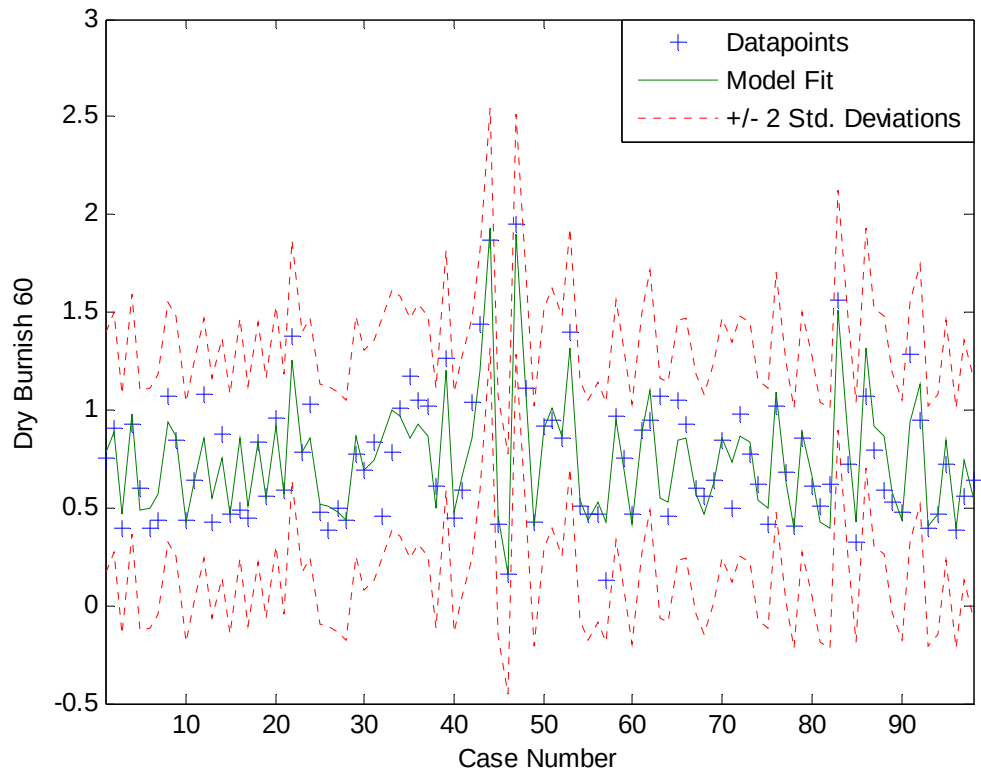


12.6 NON-LINEAR MODELLING (MODEL 2)

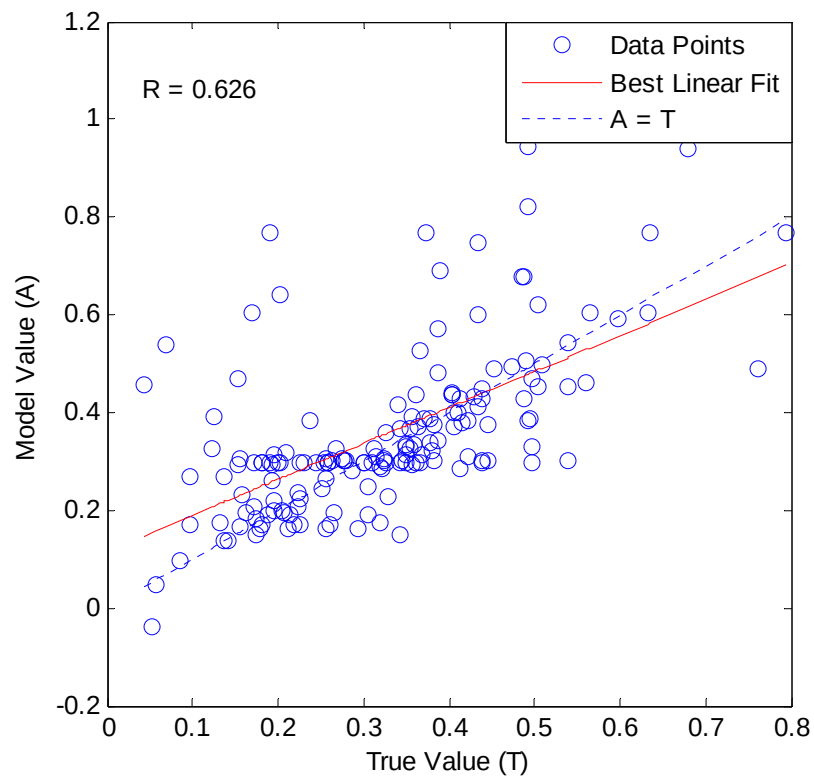
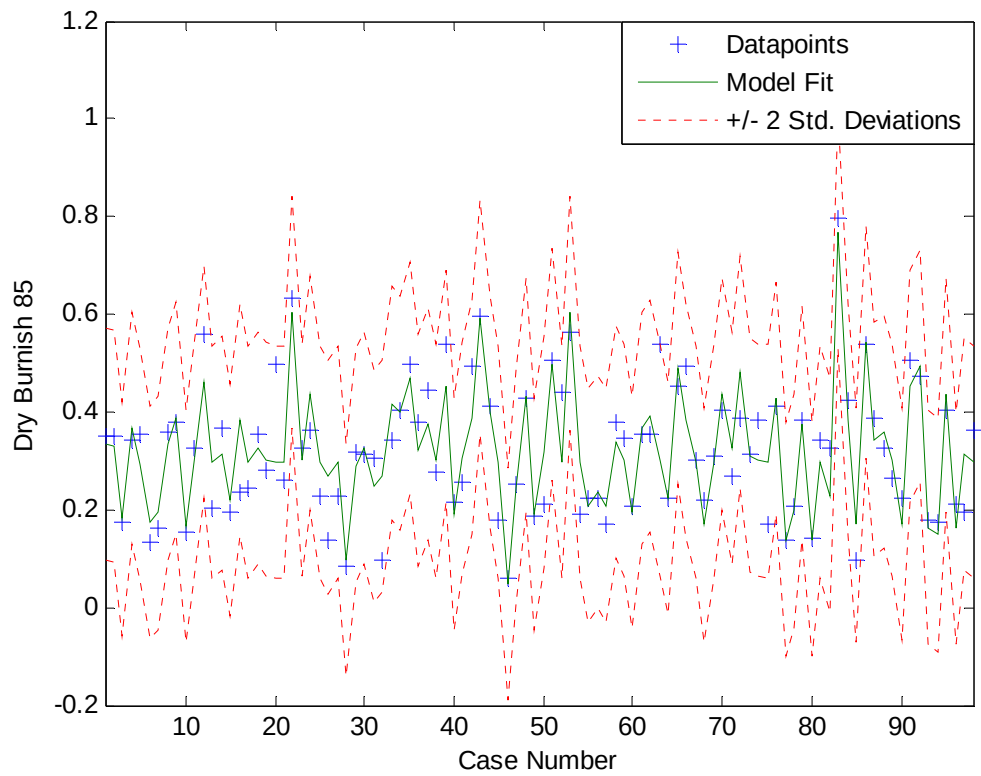
Dry Burnish 20



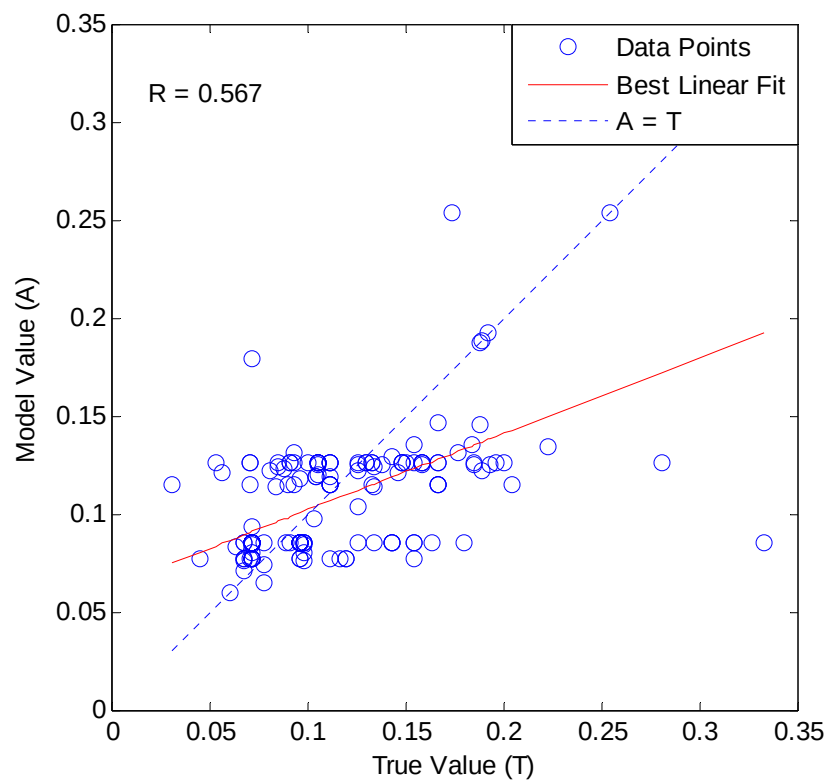
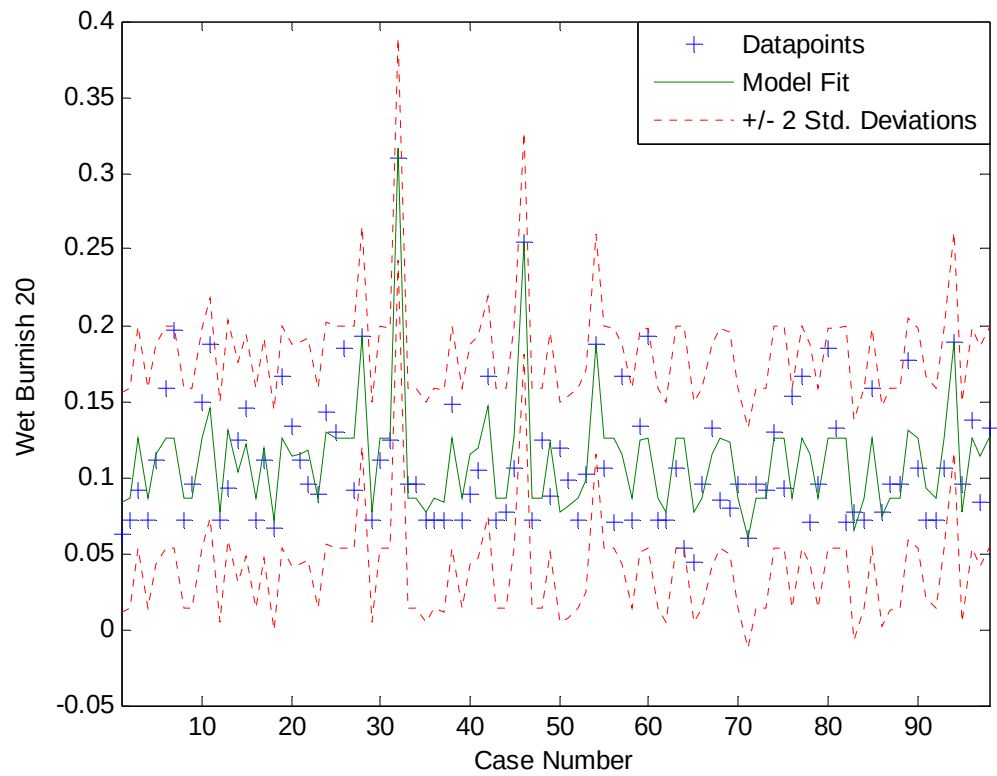
Dry Burnish 60



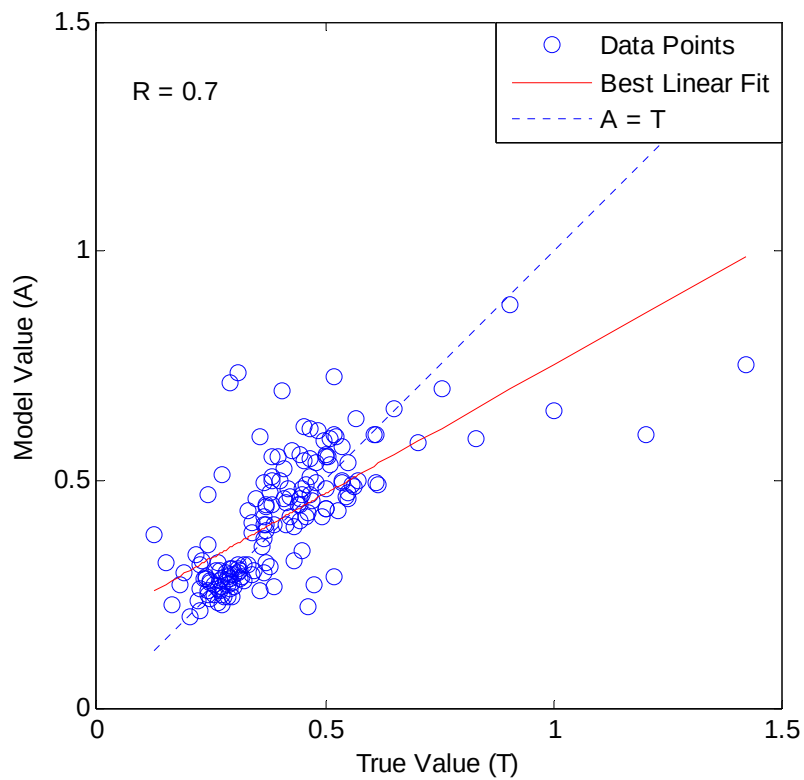
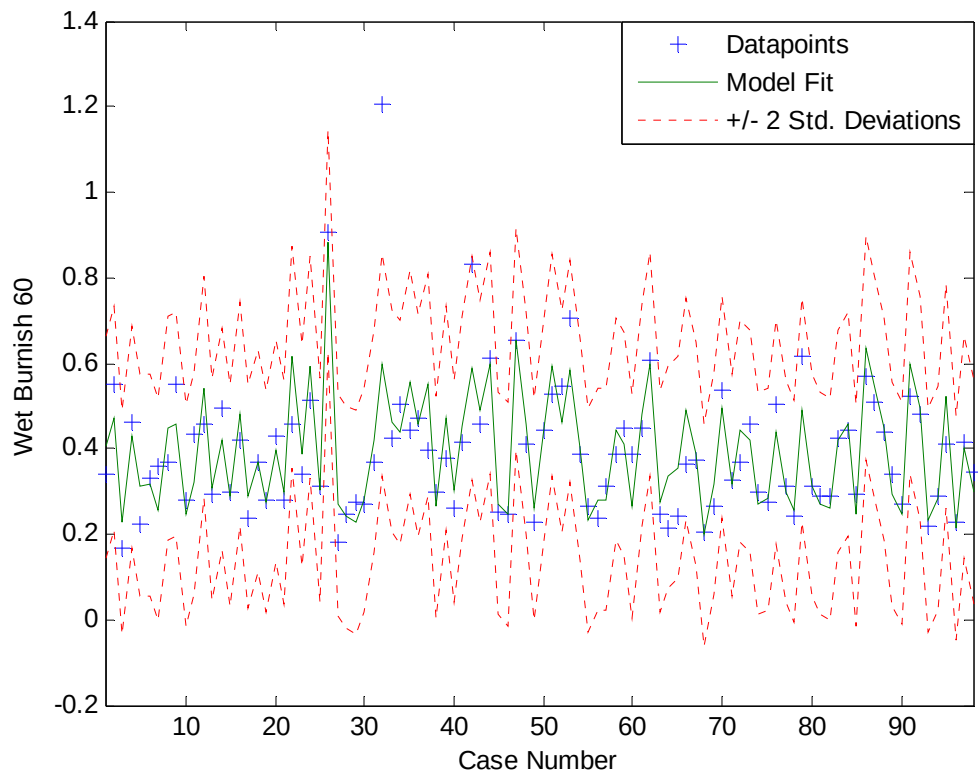
Dry Burnish 85



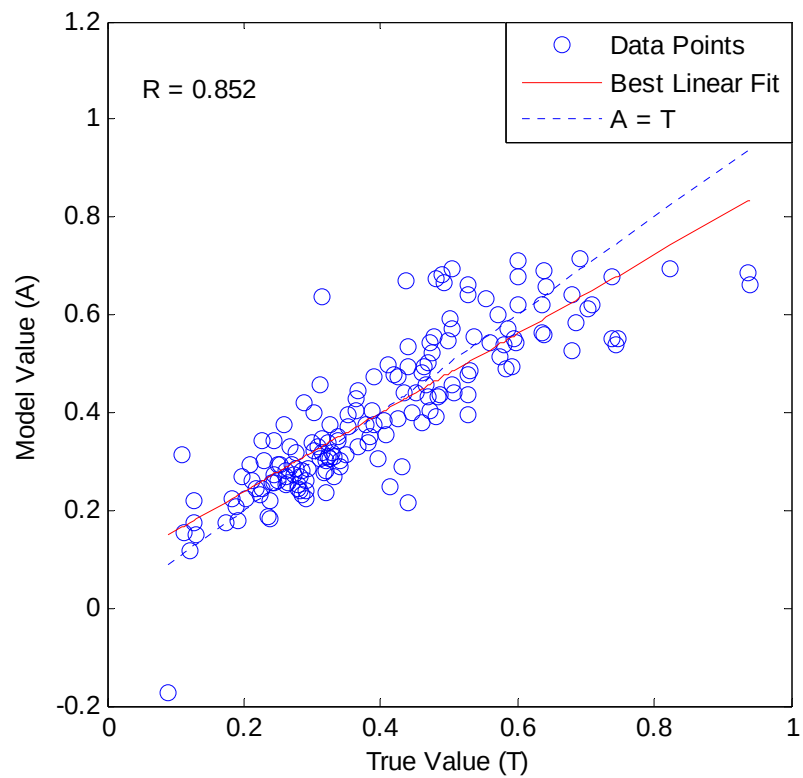
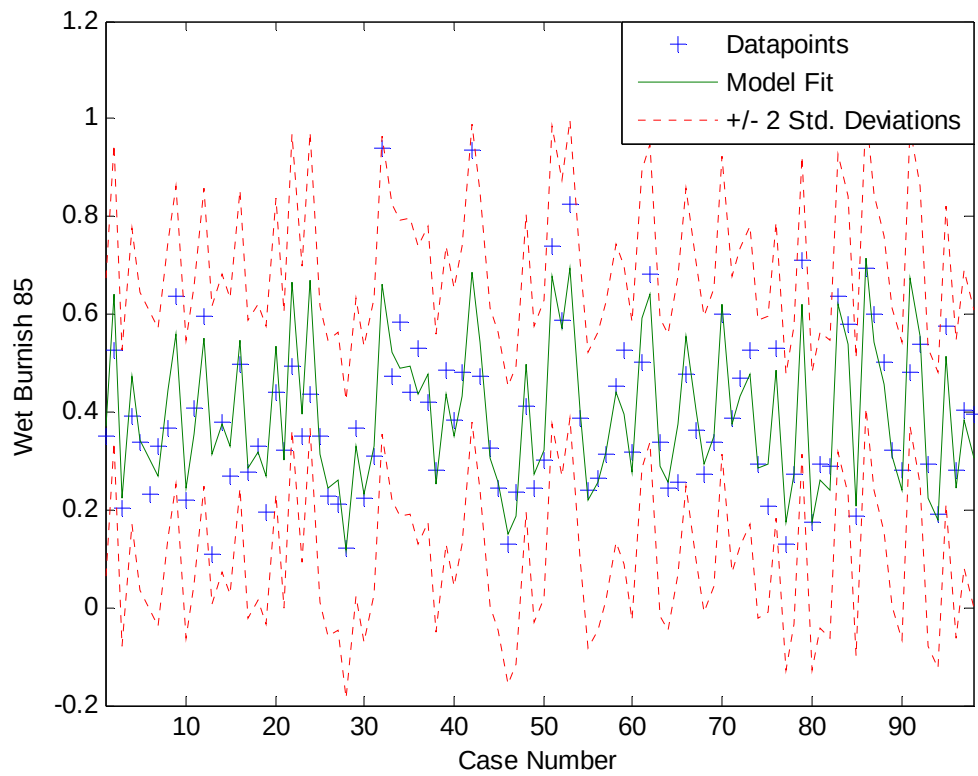
Wet Burnish 20



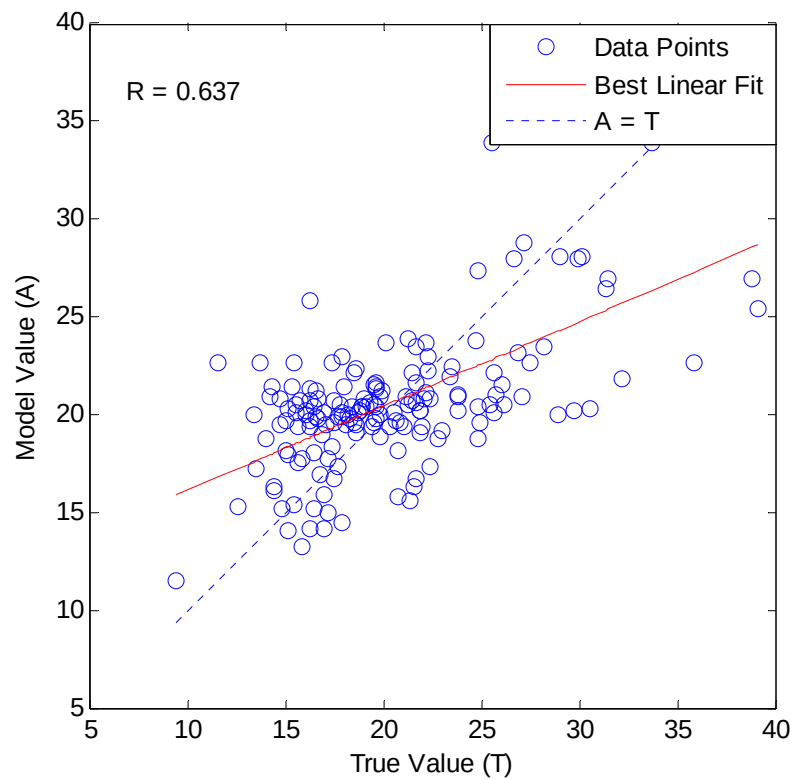
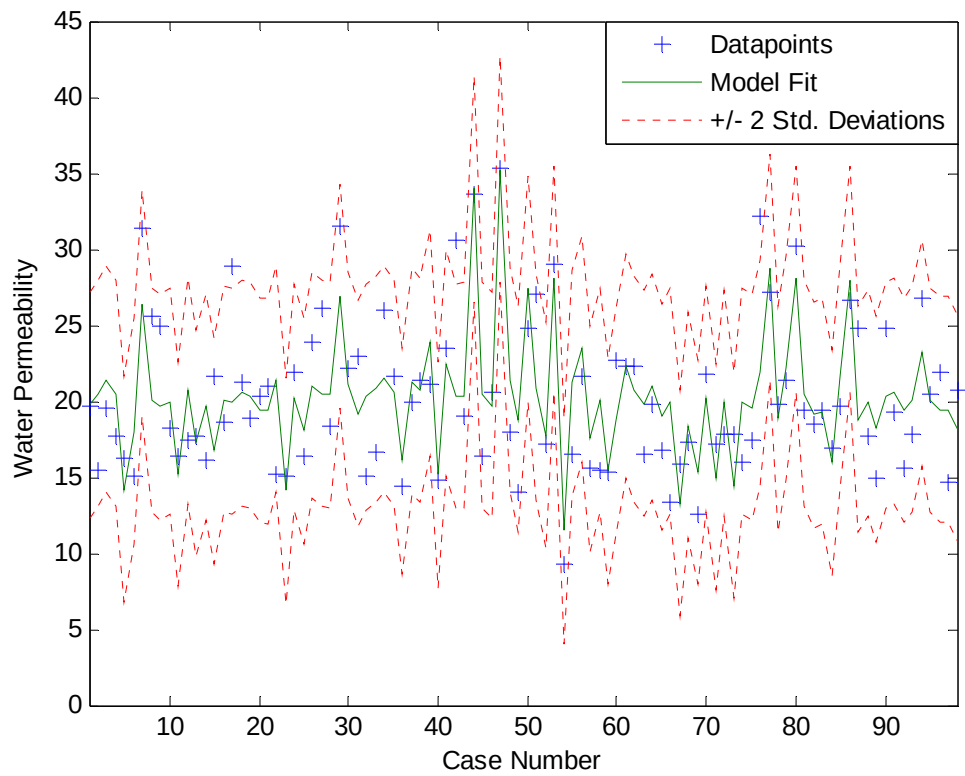
Wet Burnish 60



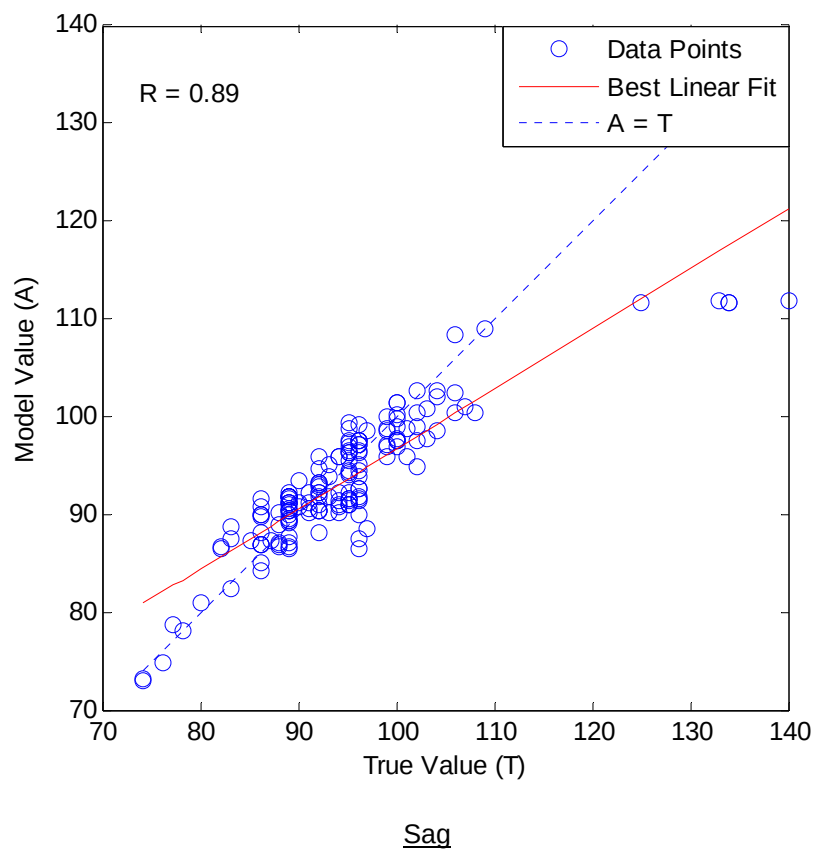
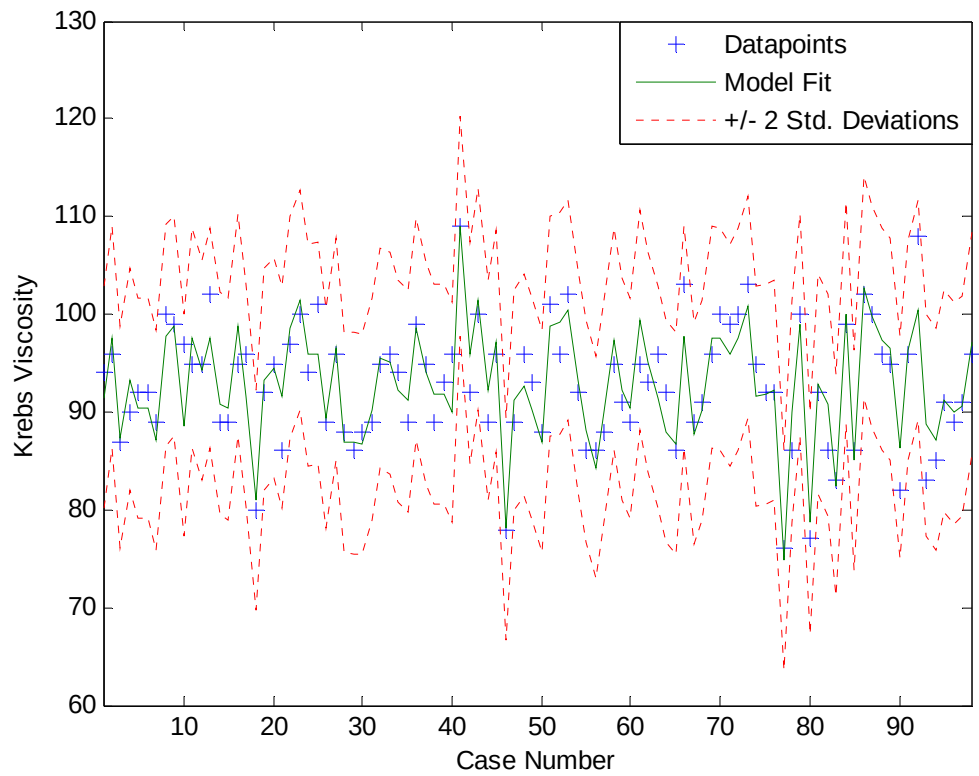
Wet Burnish 85

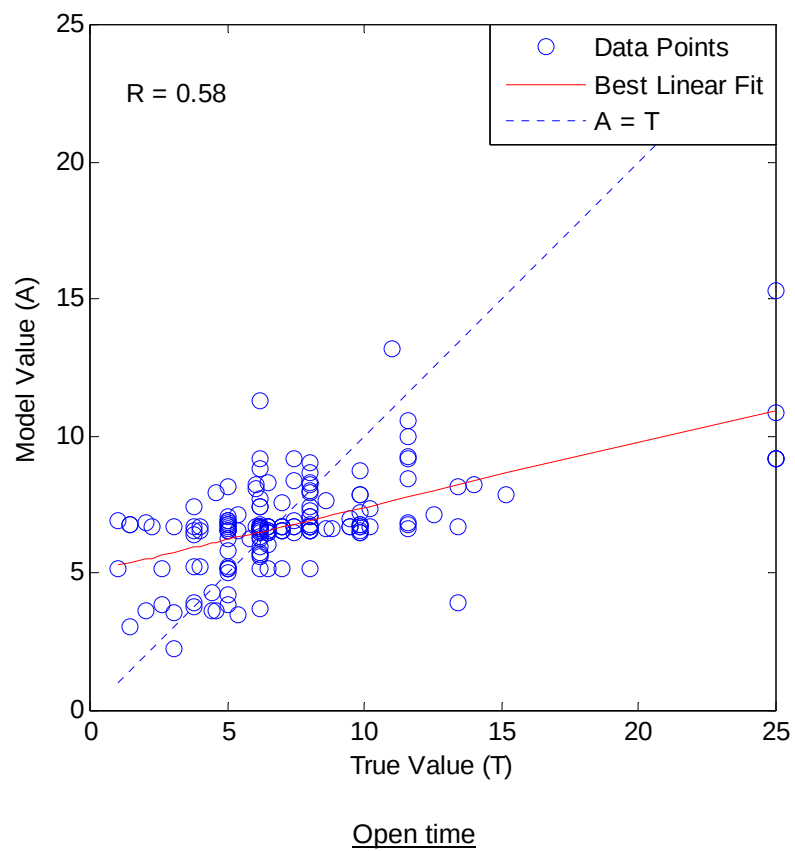
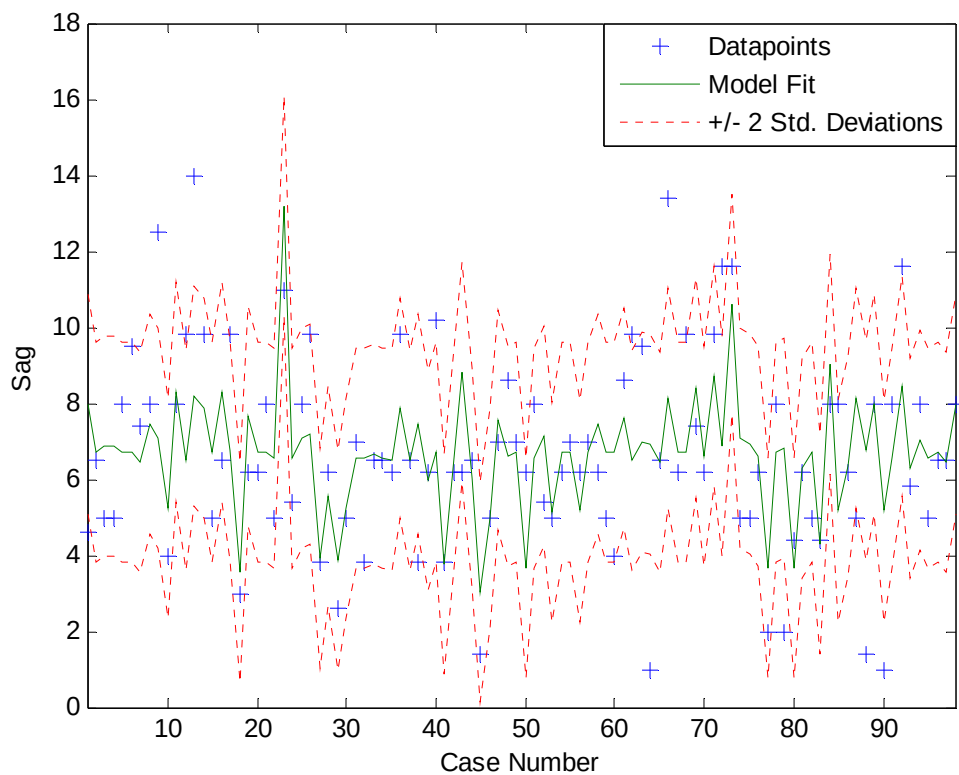


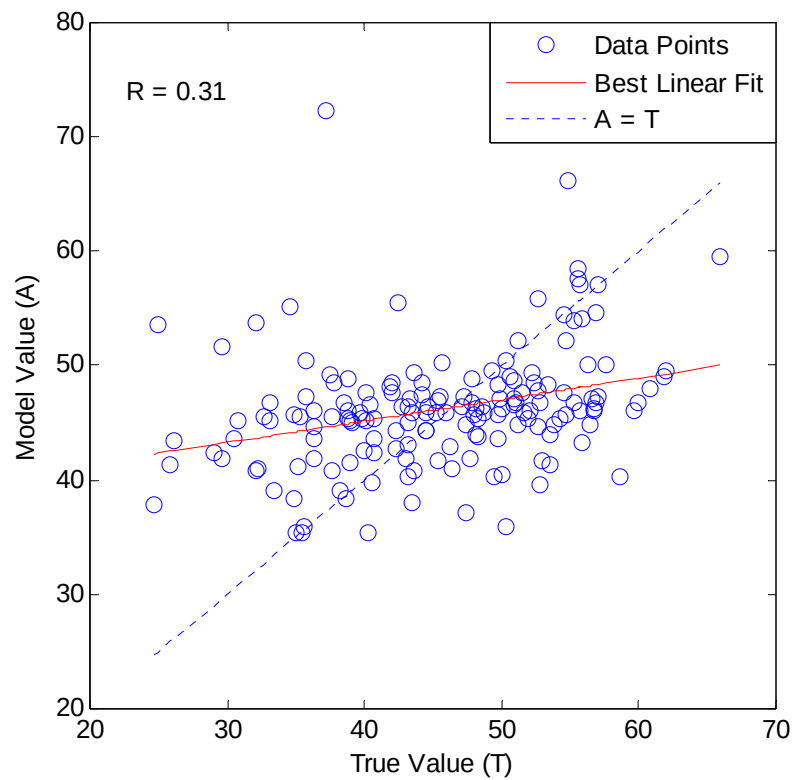
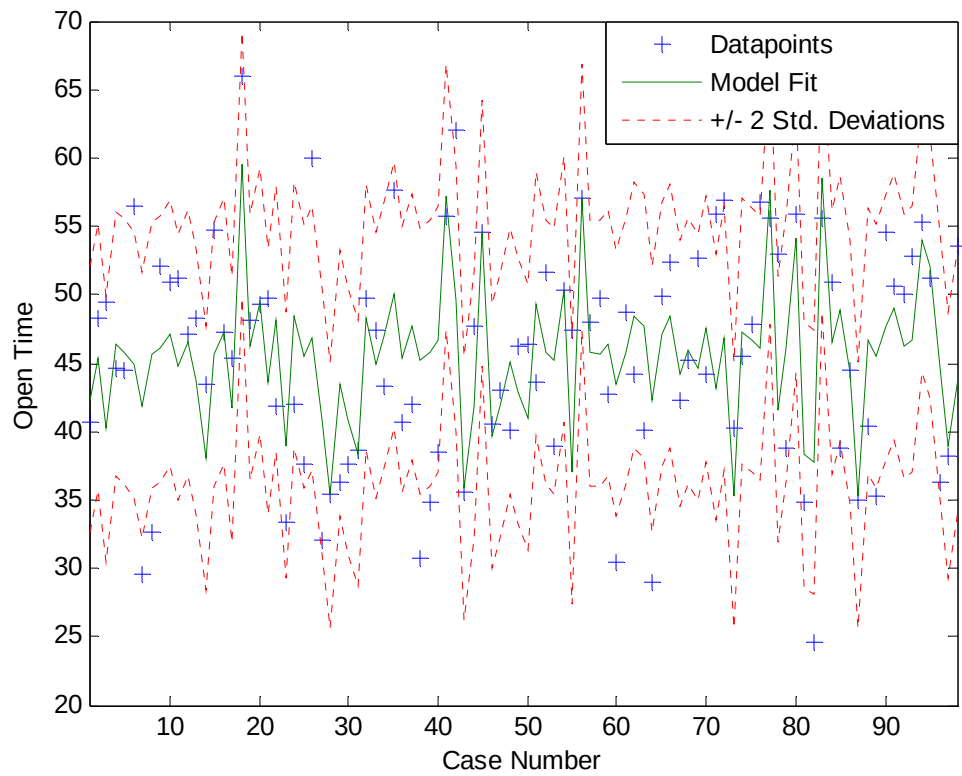
Water Permeability



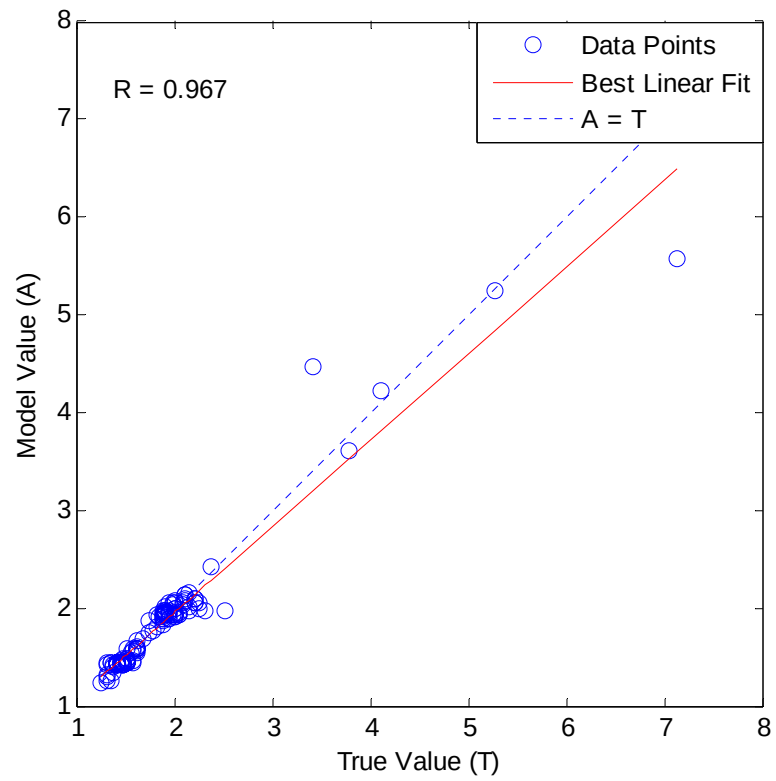
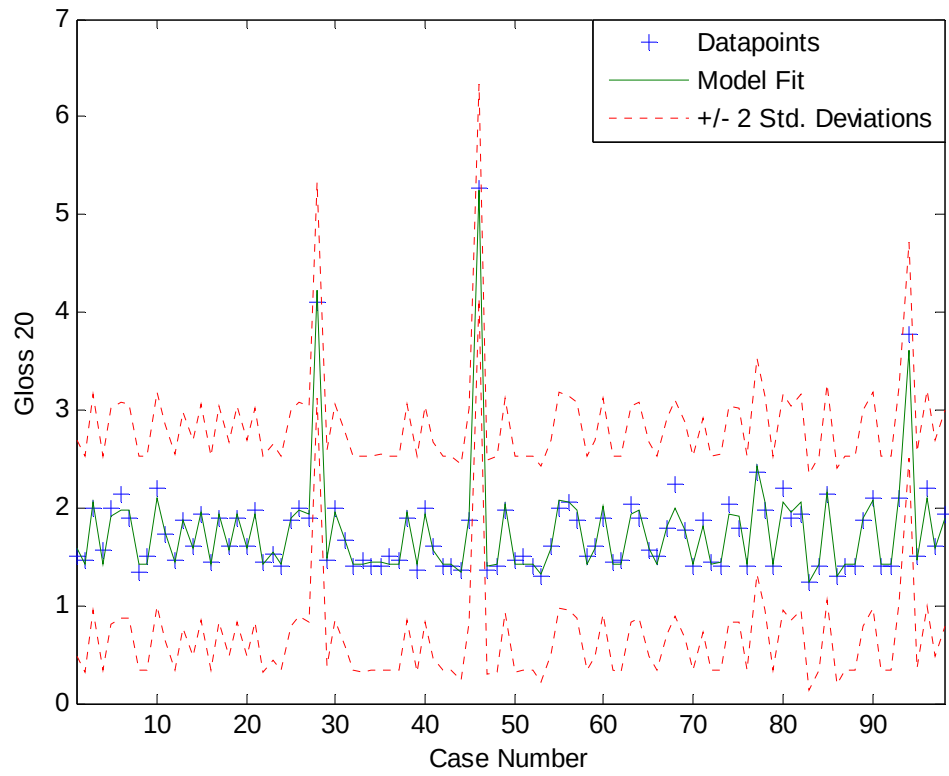
Krebs



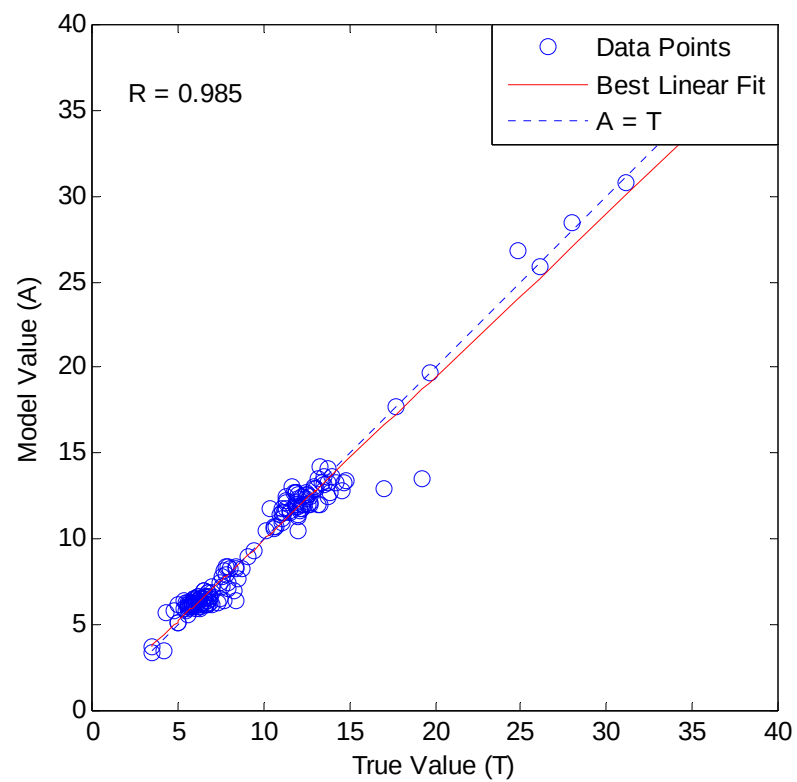
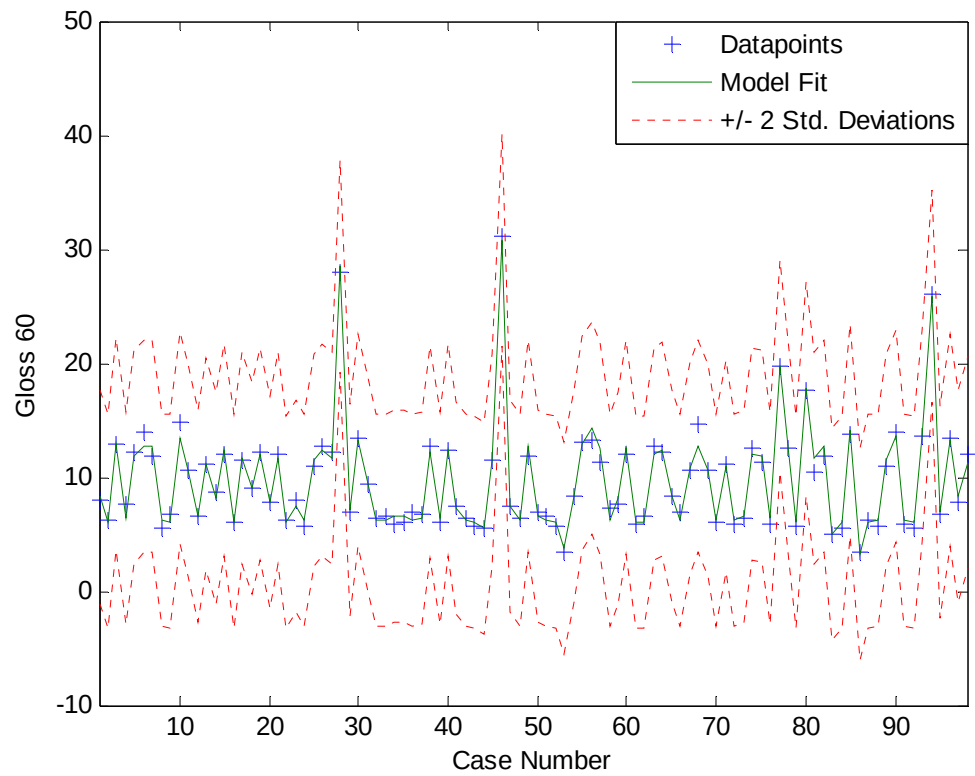




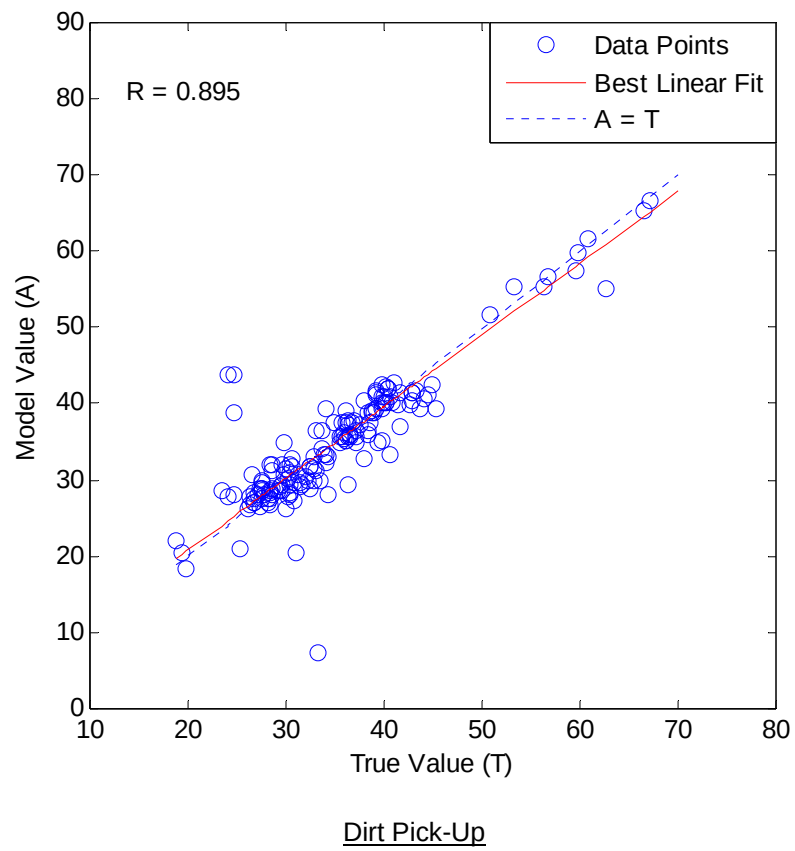
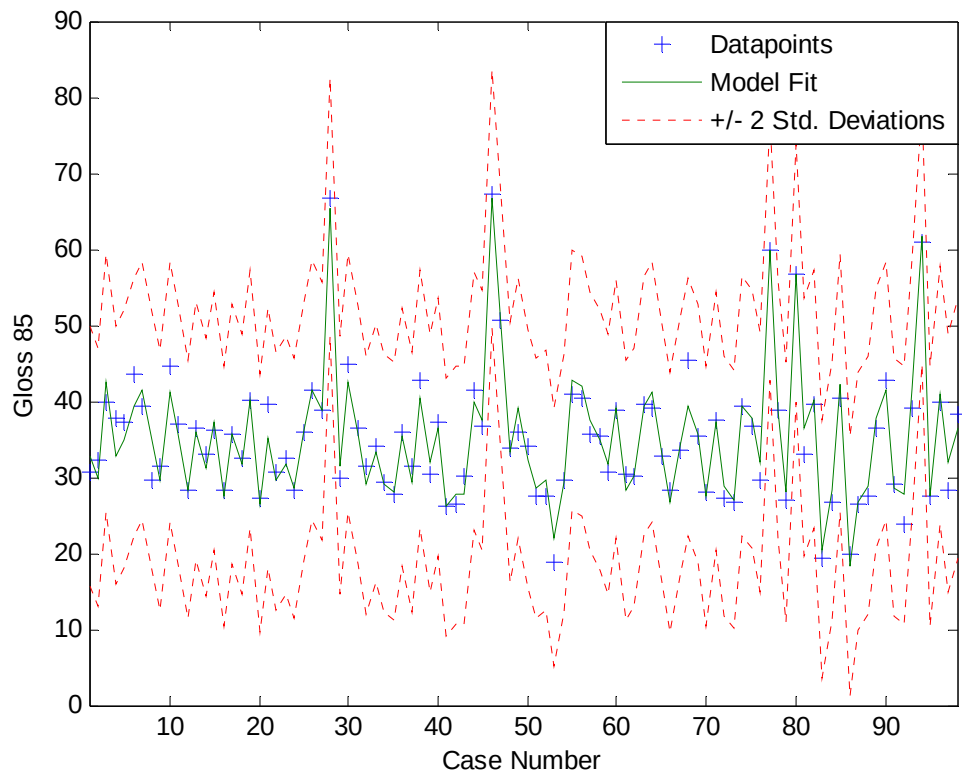
Gloss 20

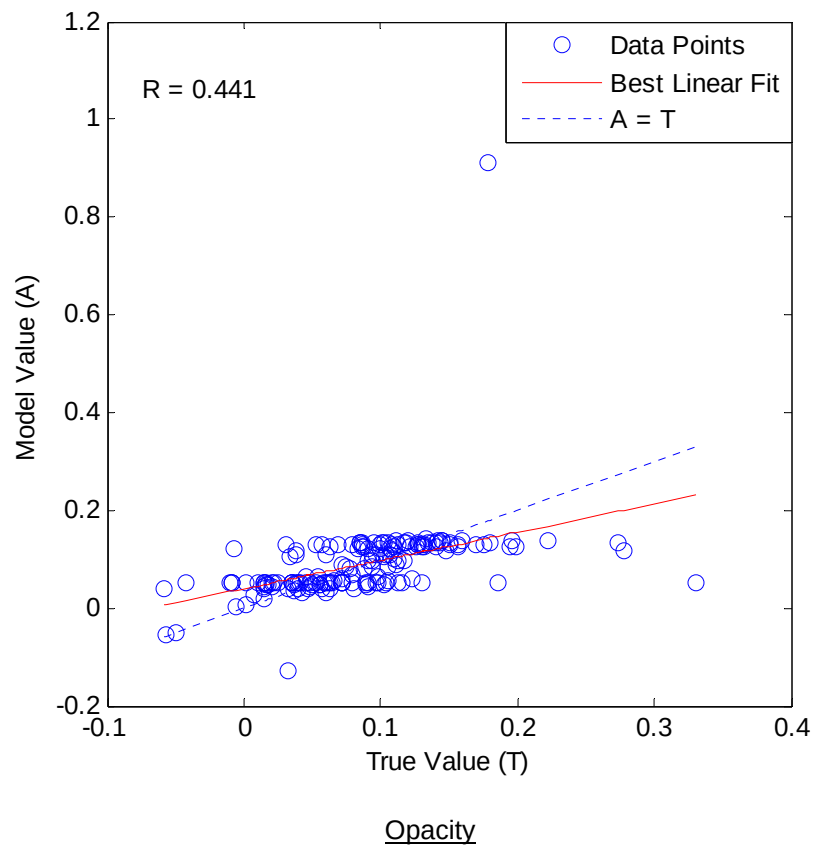
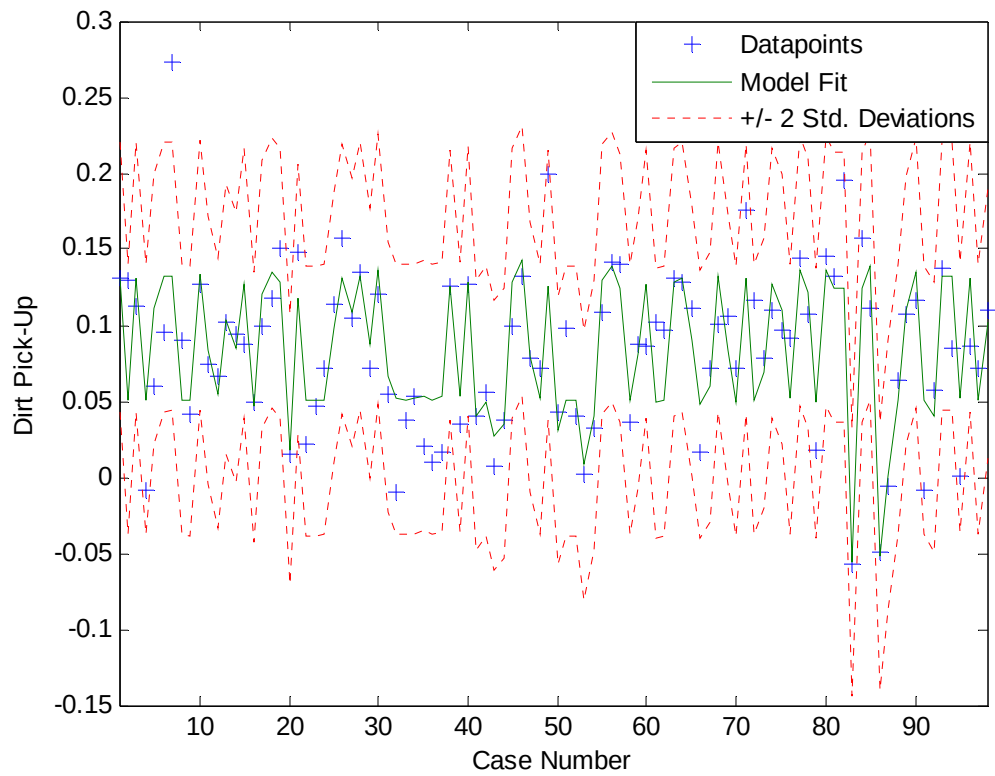


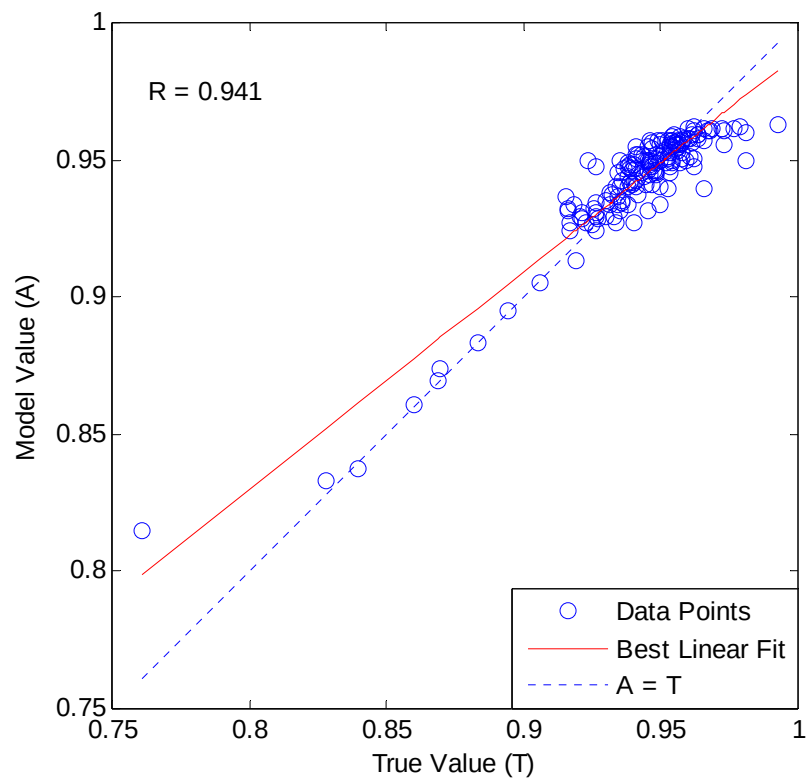
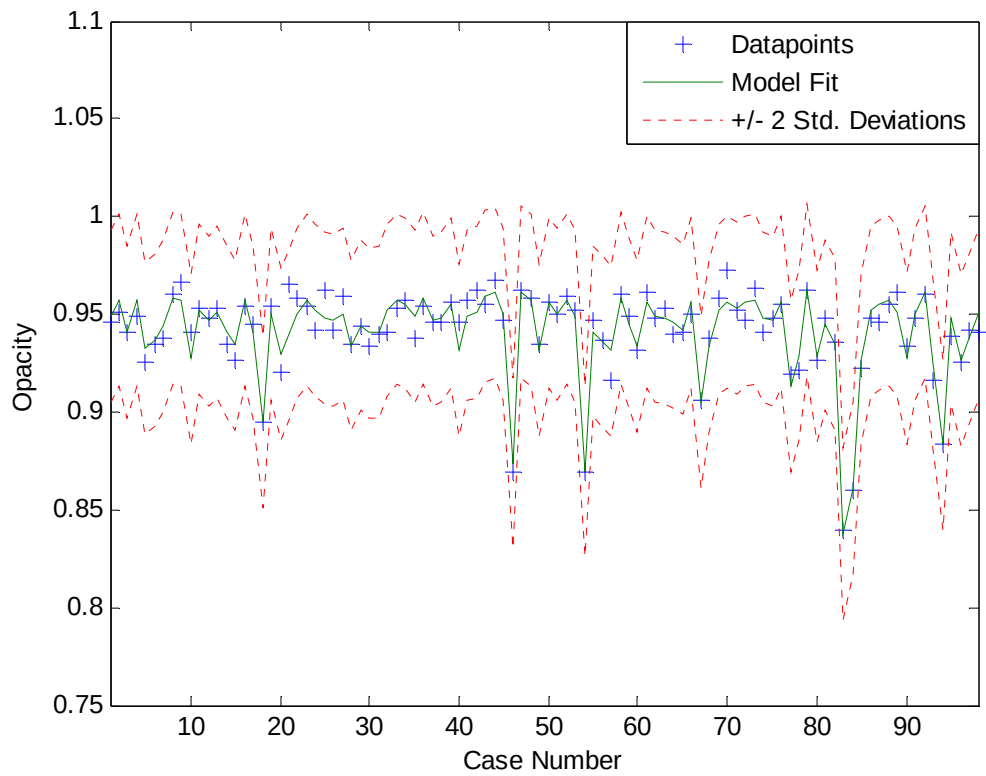
Gloss 60



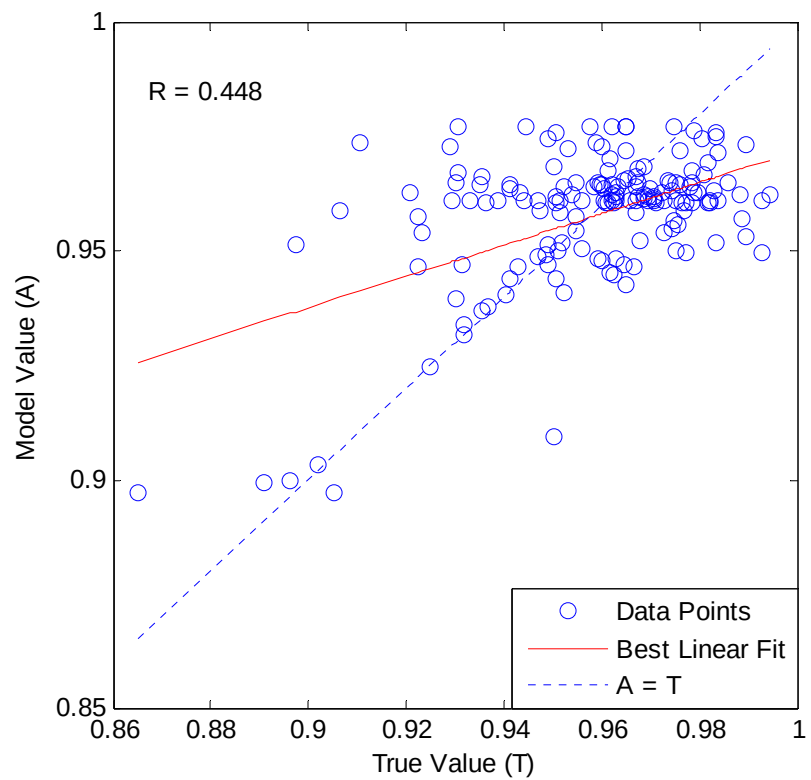
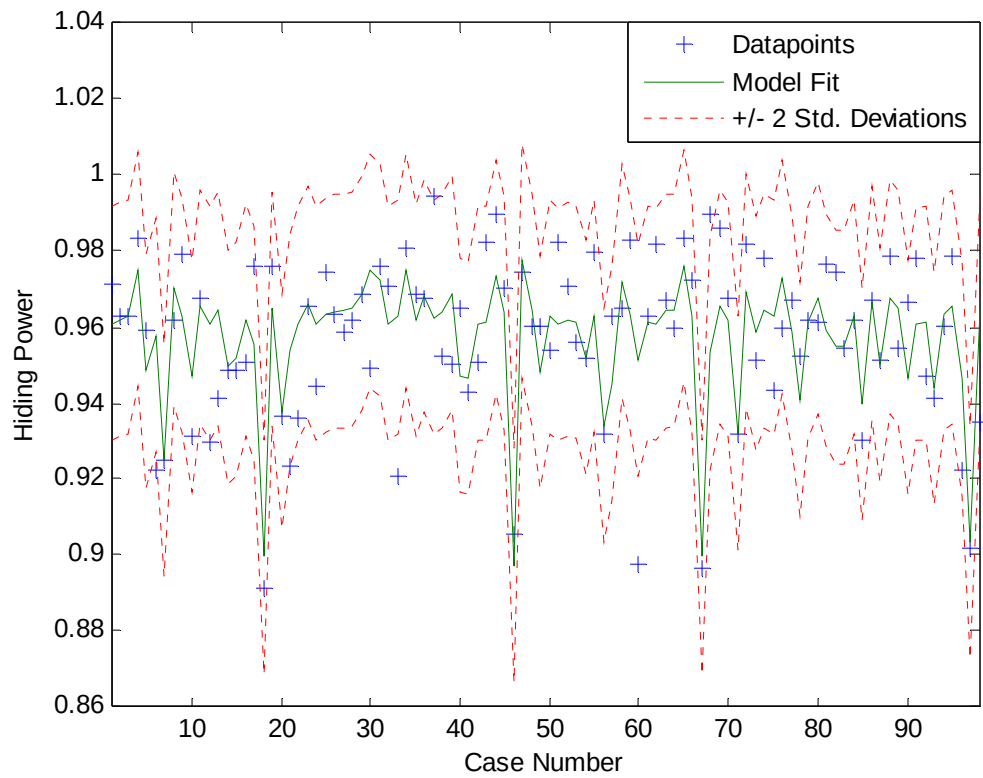
Gloss 85



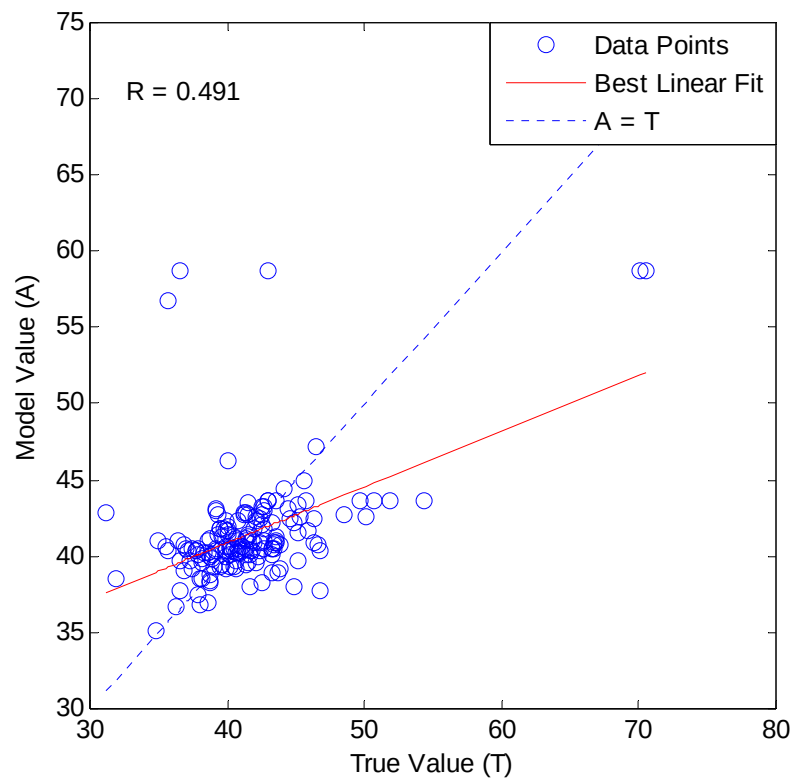
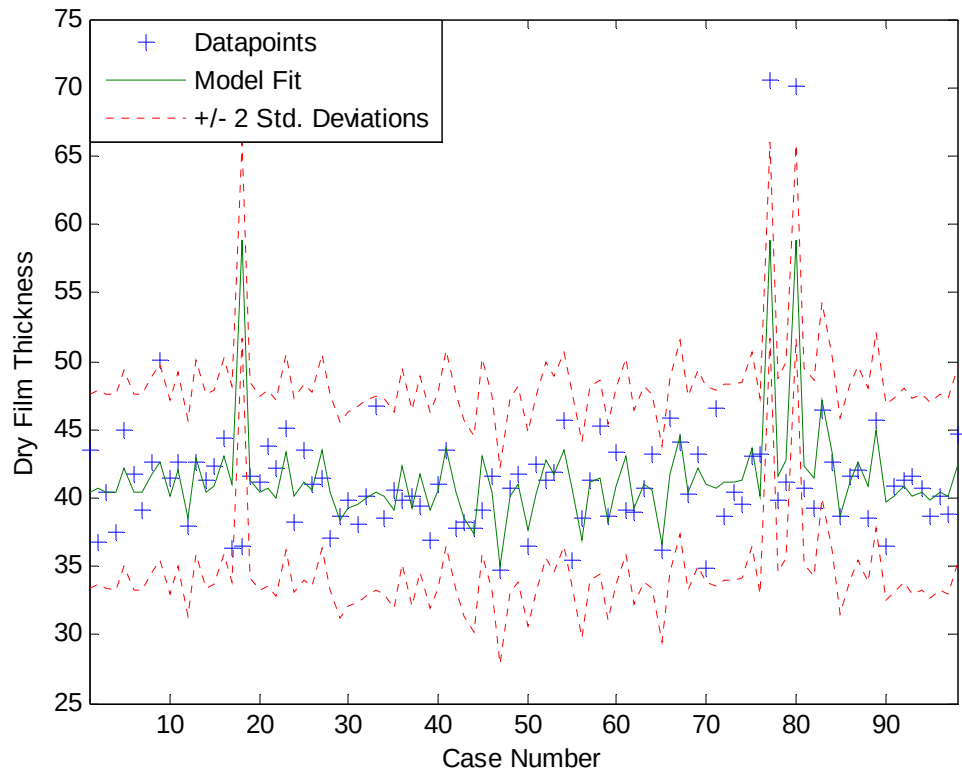




Hiding Power

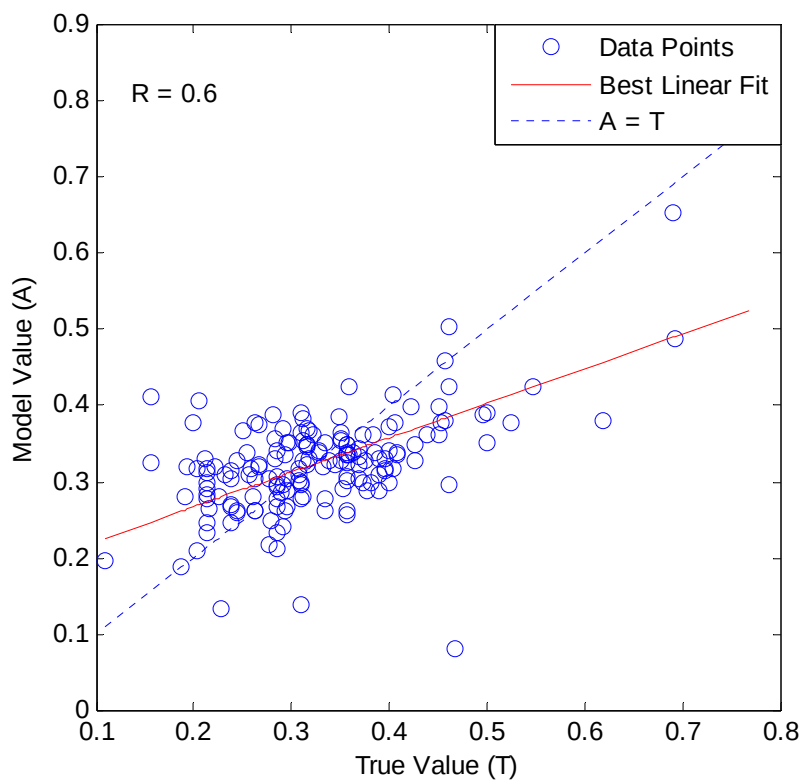
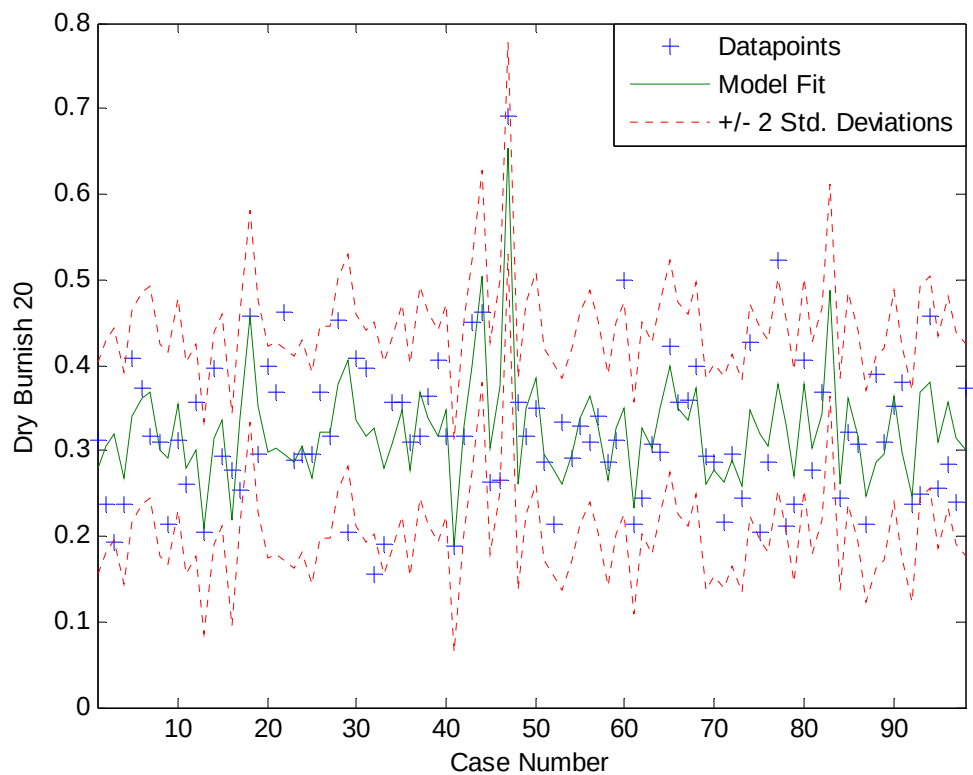


Dry Film Thickness

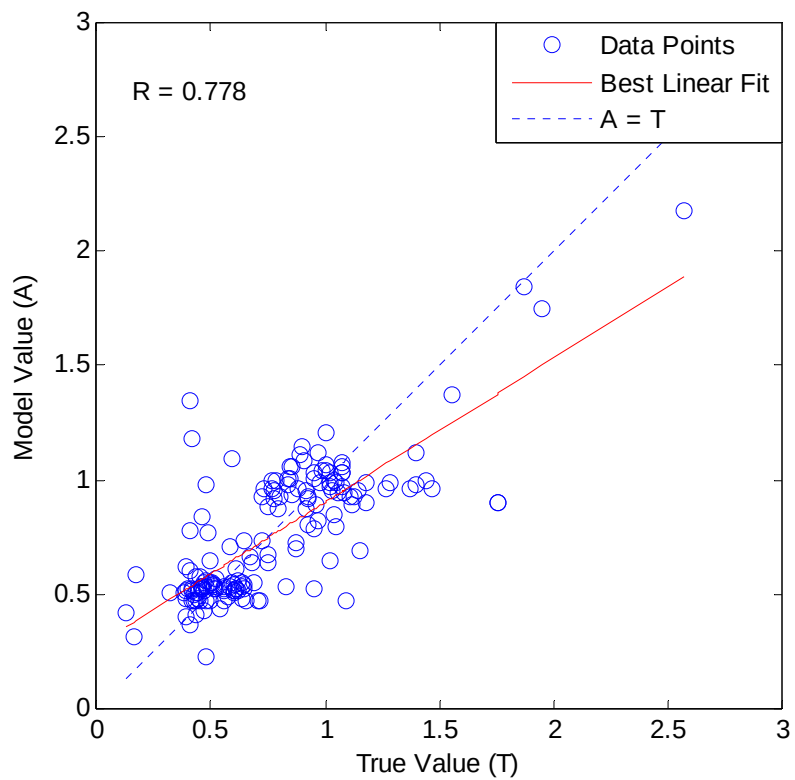
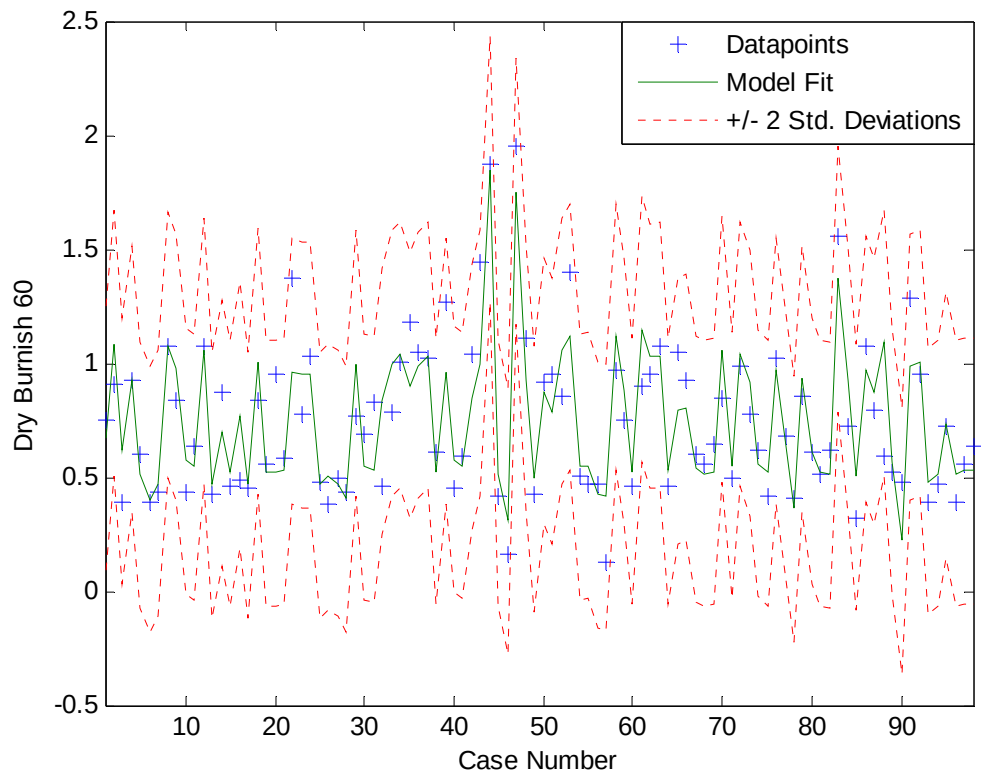


12.7 NON-LINEAR MODELLING (MODEL 3)

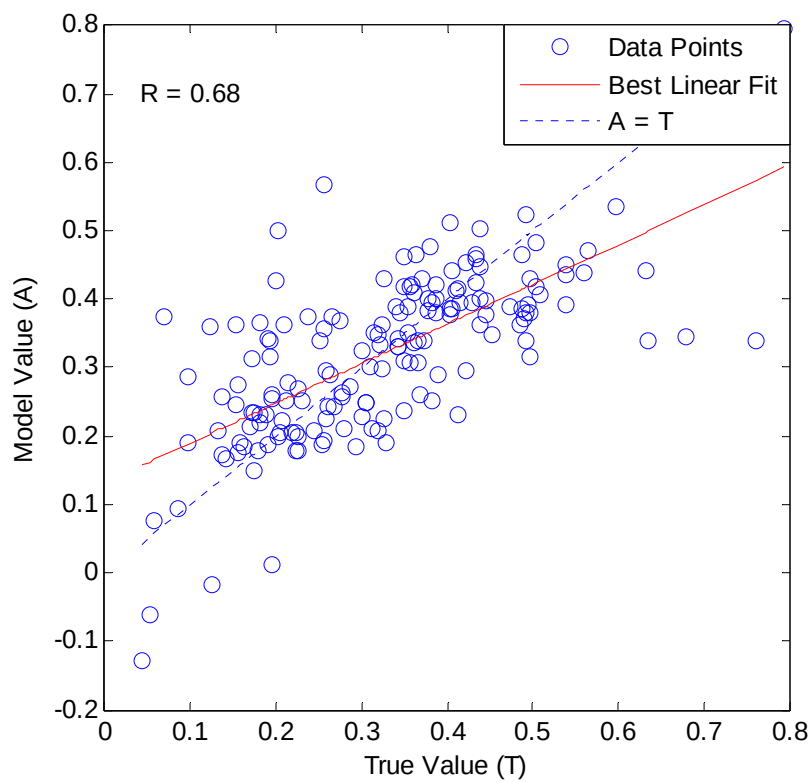
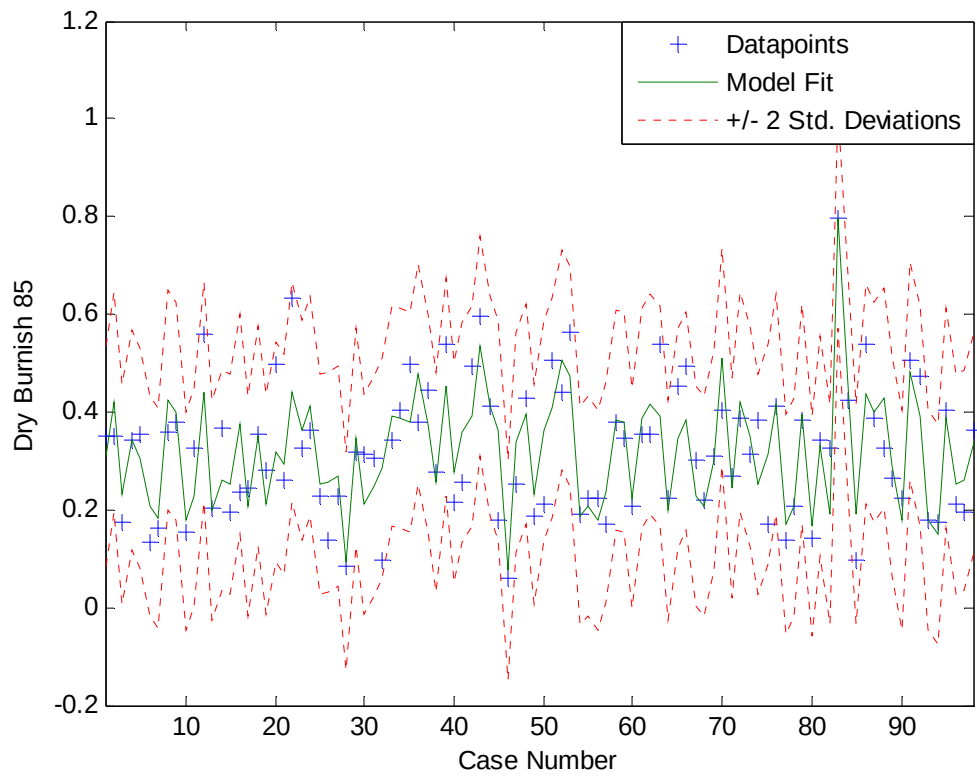
Dry Burnish 20



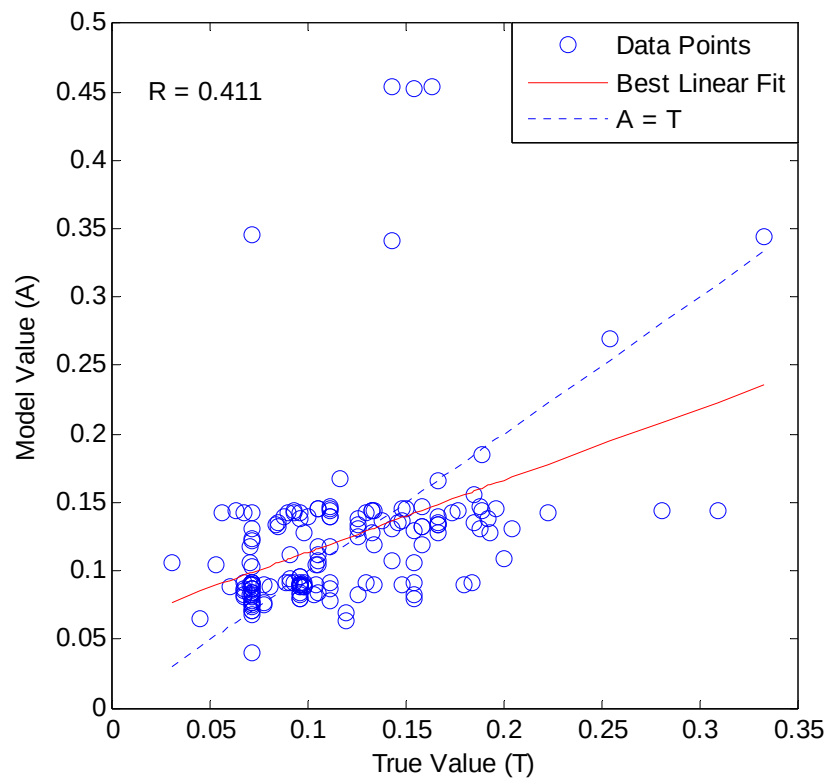
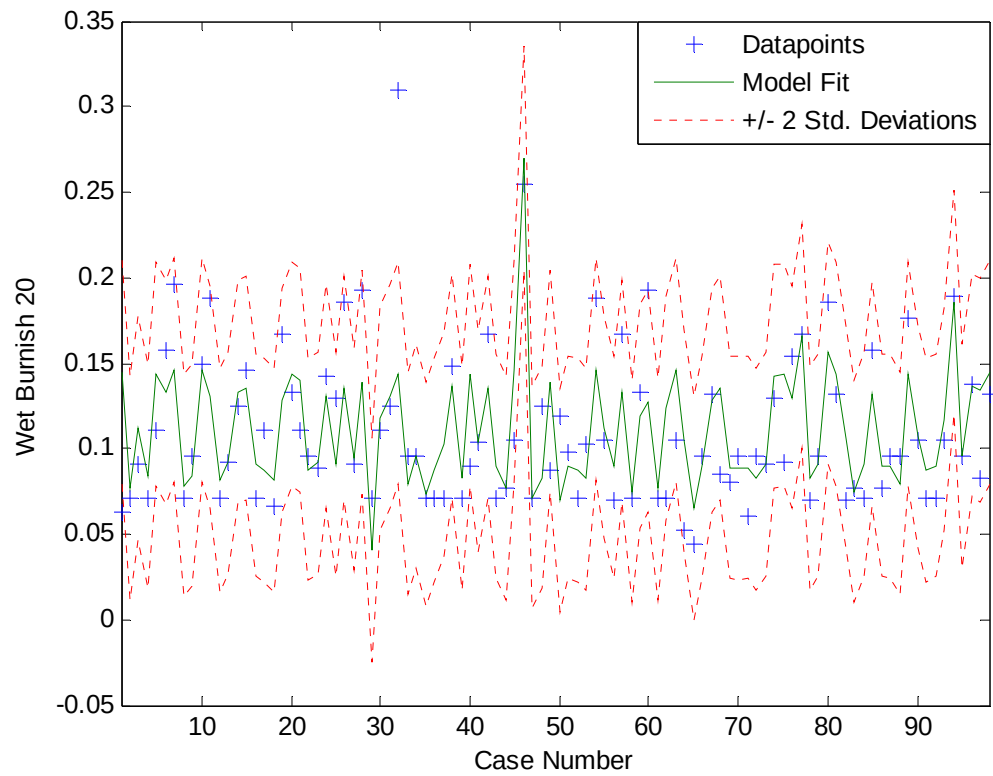
Dry Burnish 60



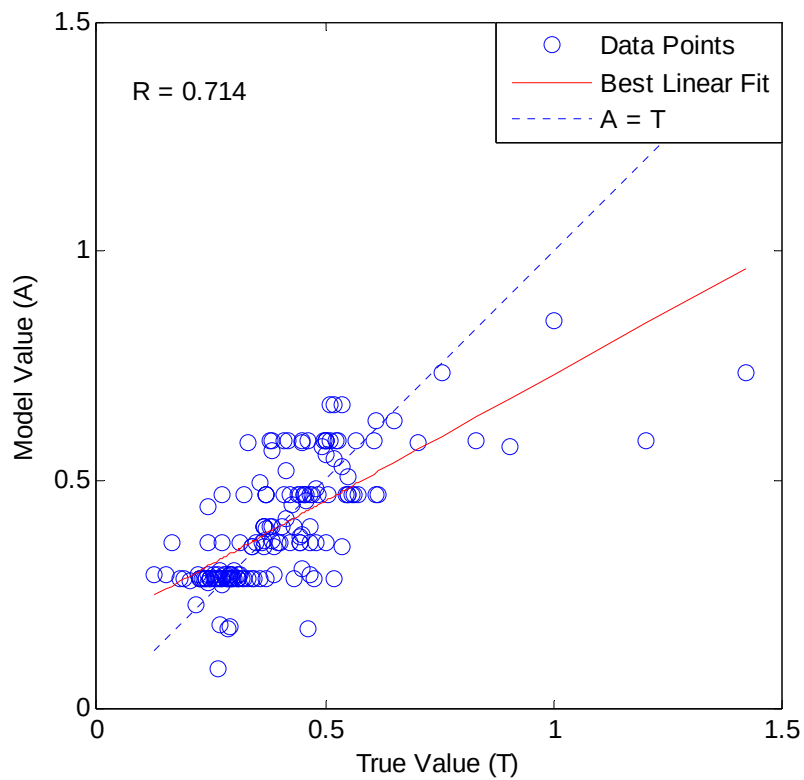
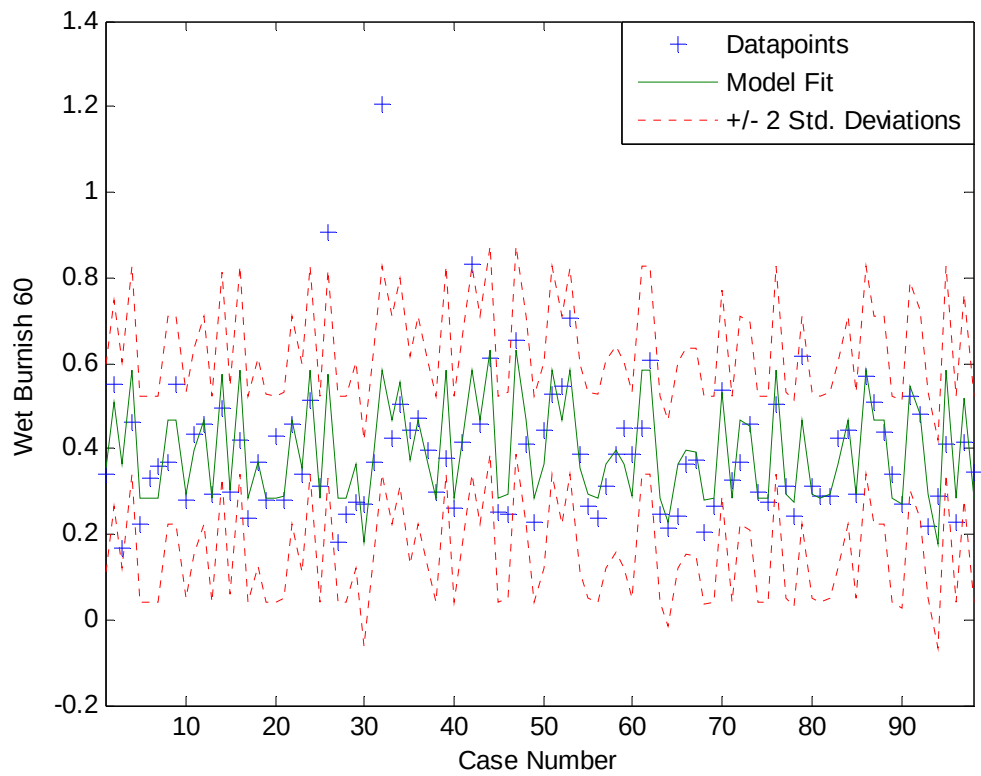
Dry Burnish 85



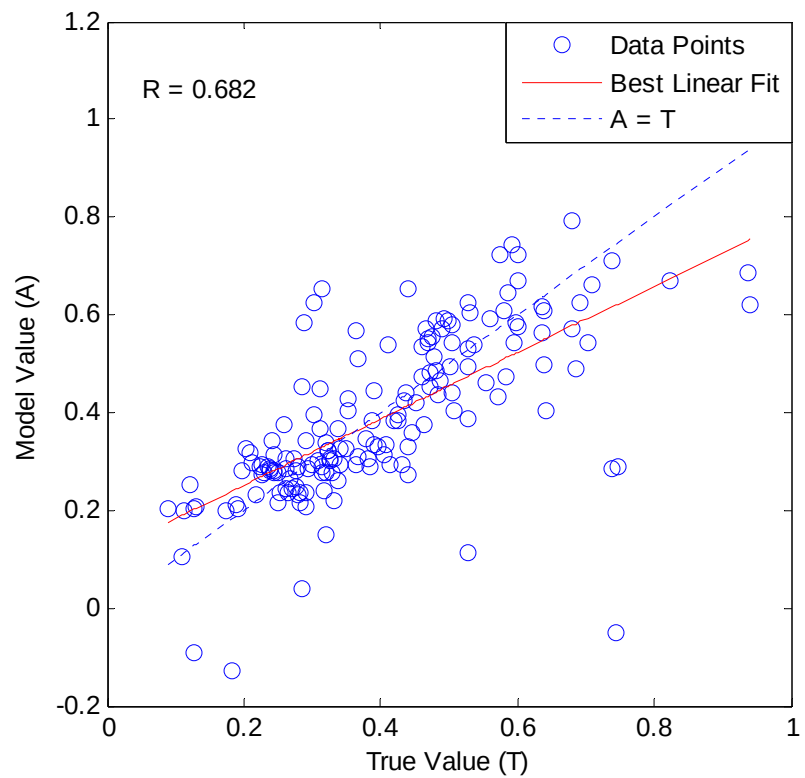
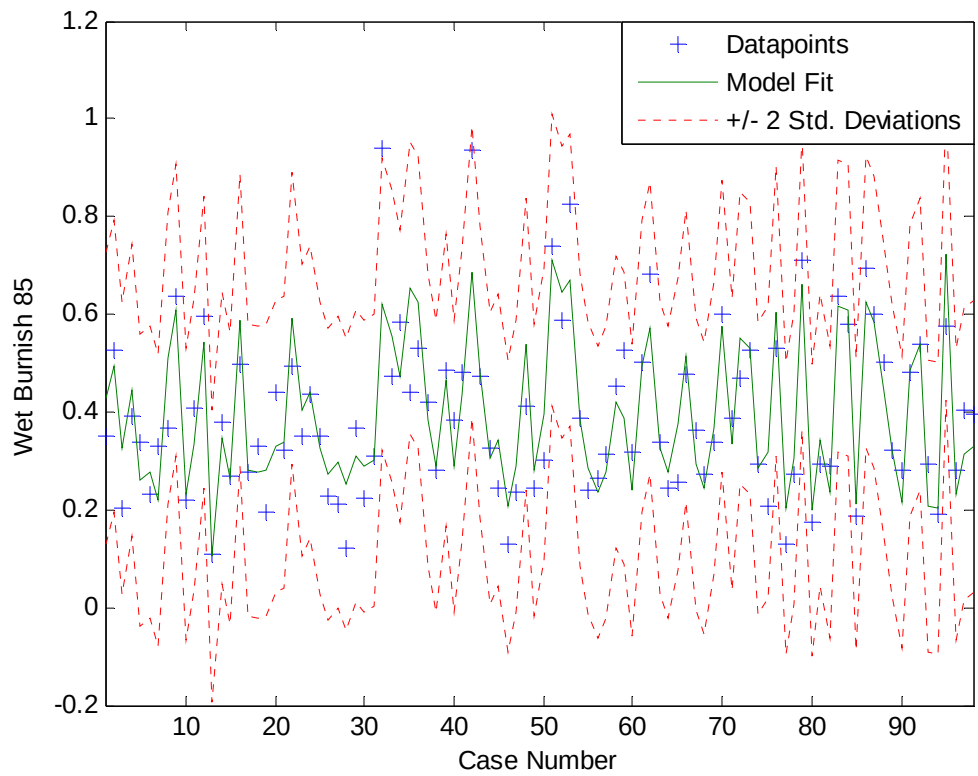
Wet Burnish 20



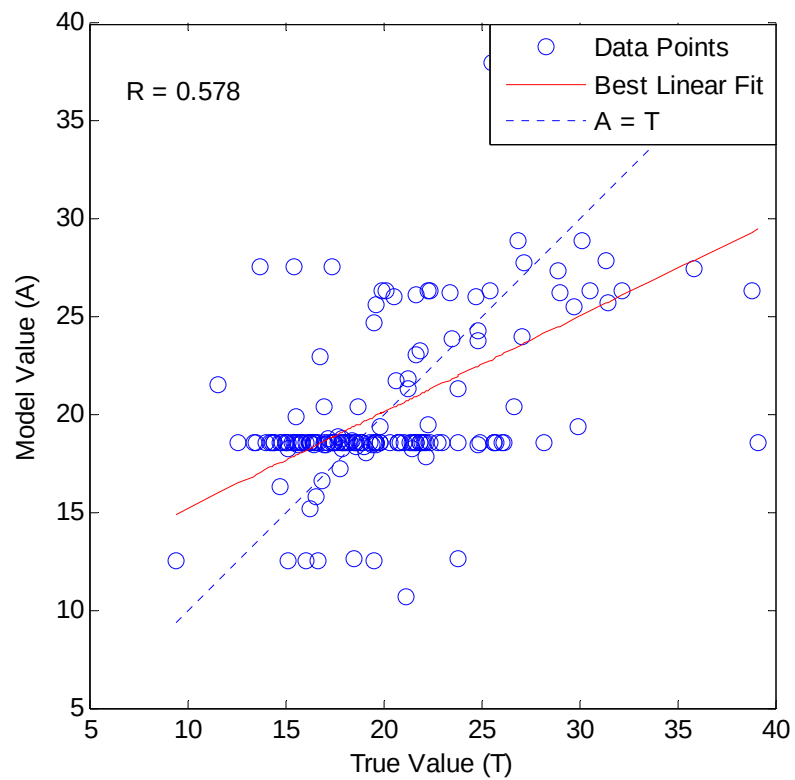
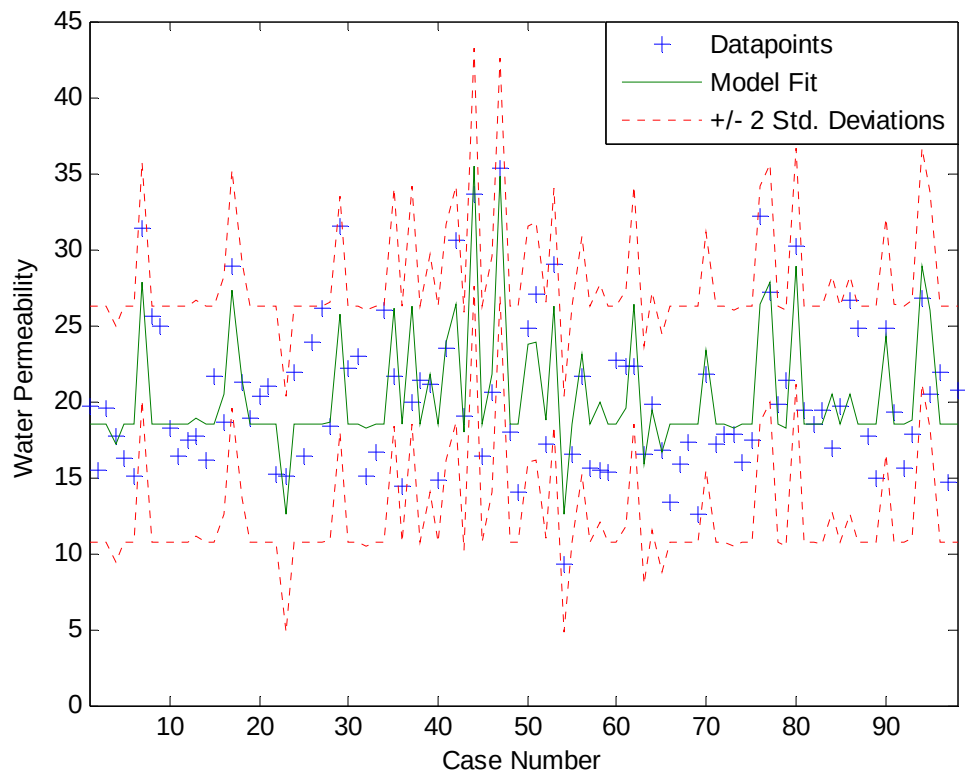
Wet Burnish 60



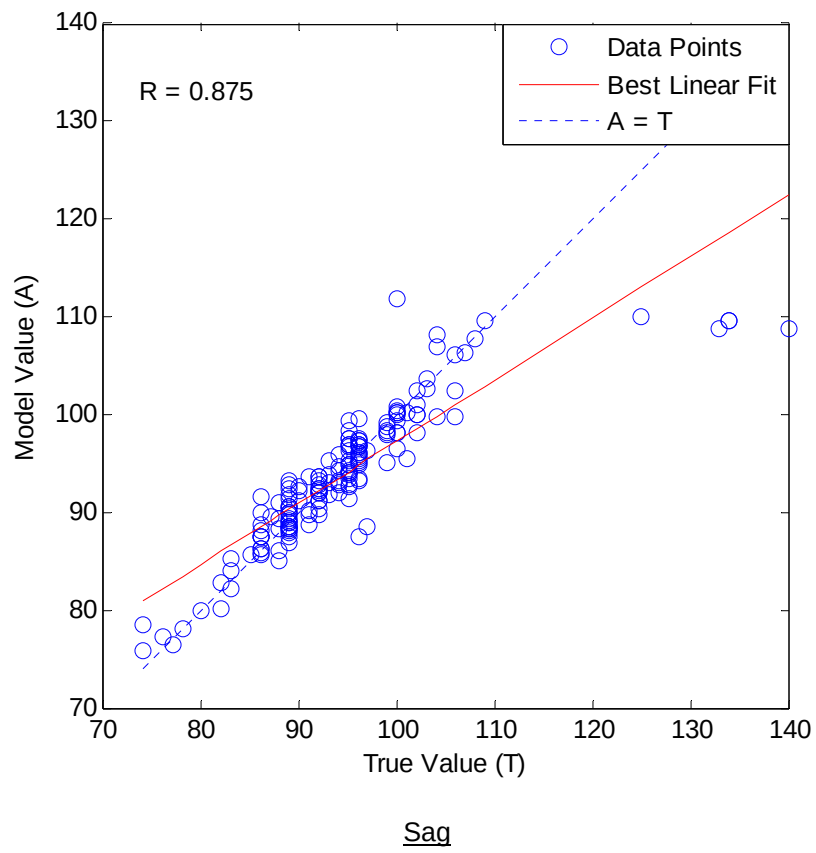
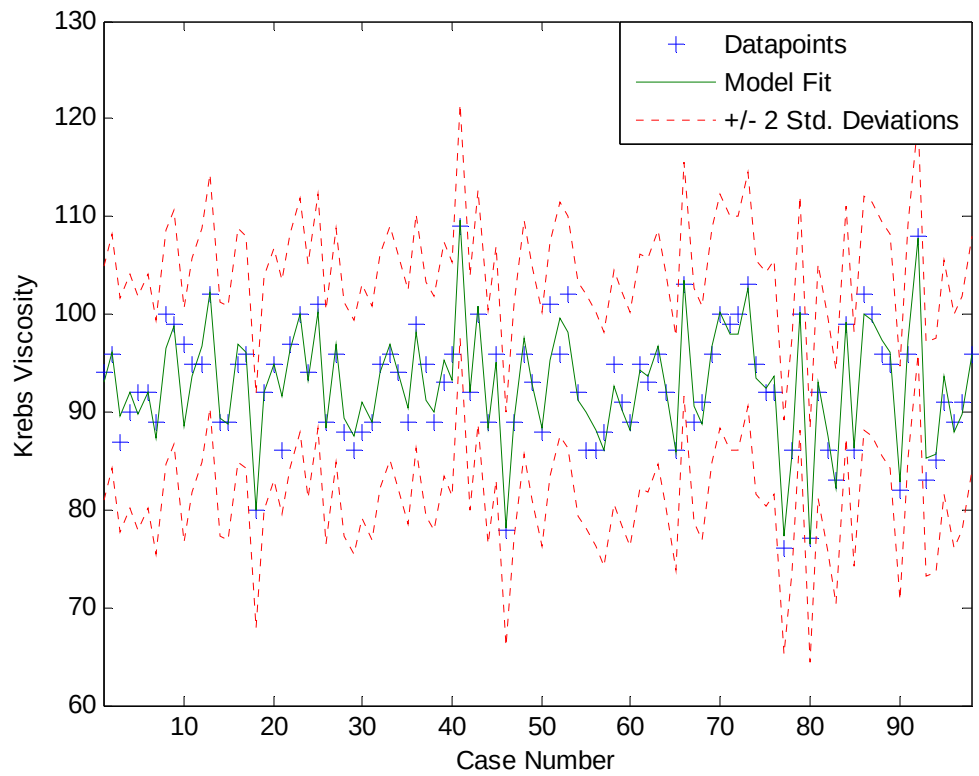
Wet Burnish 85

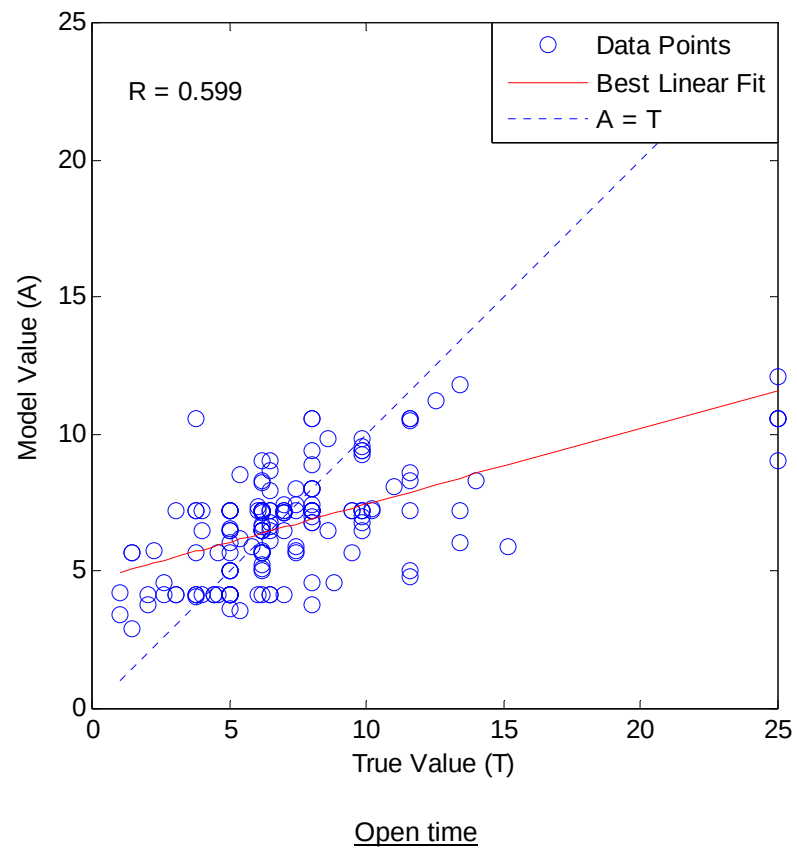
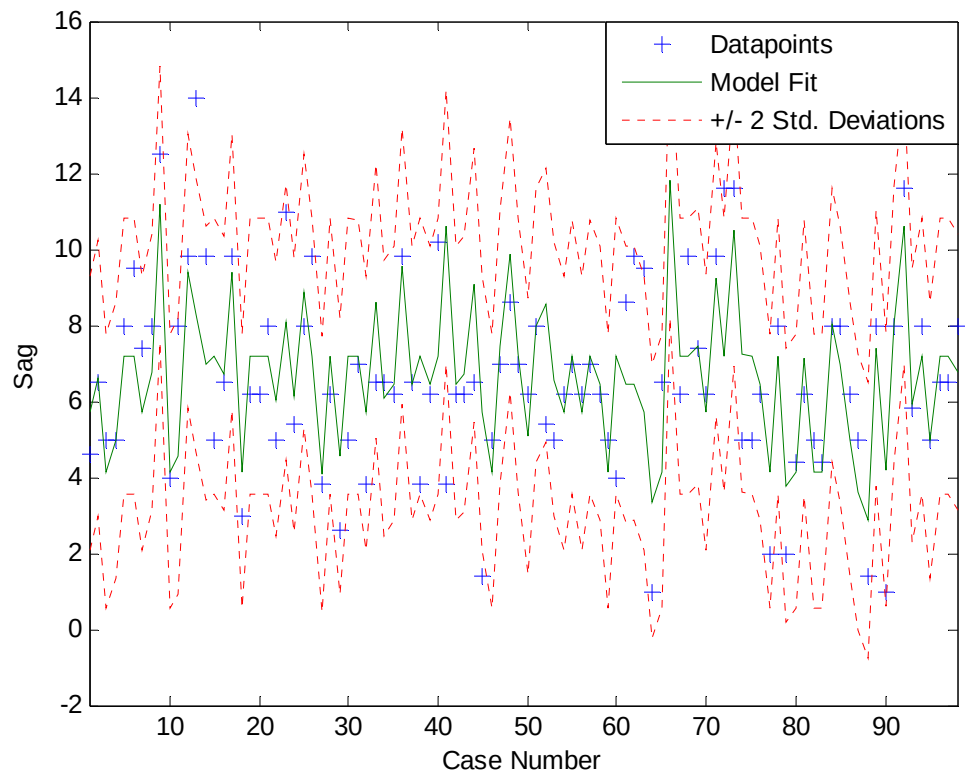


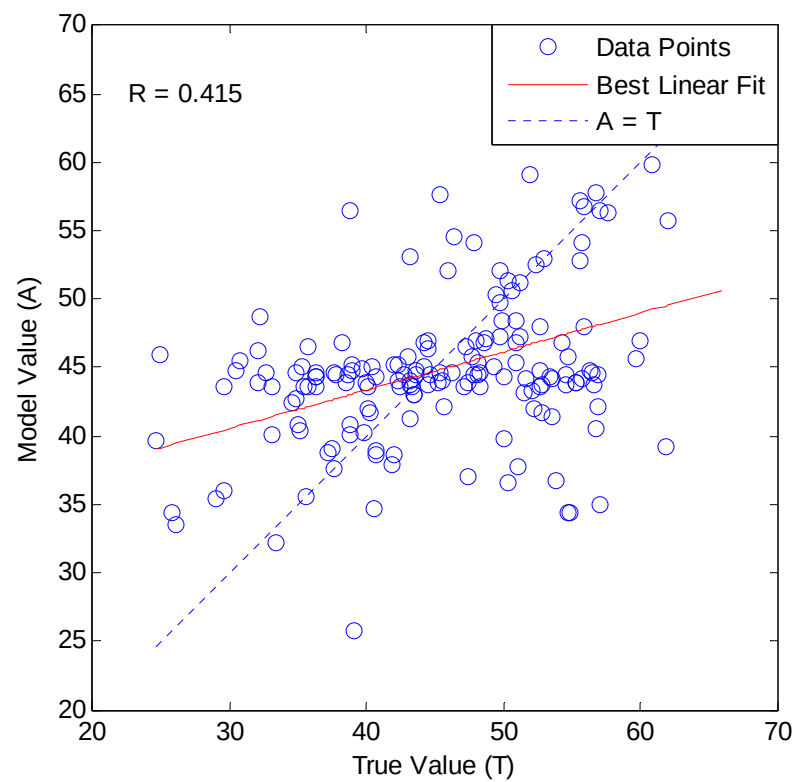
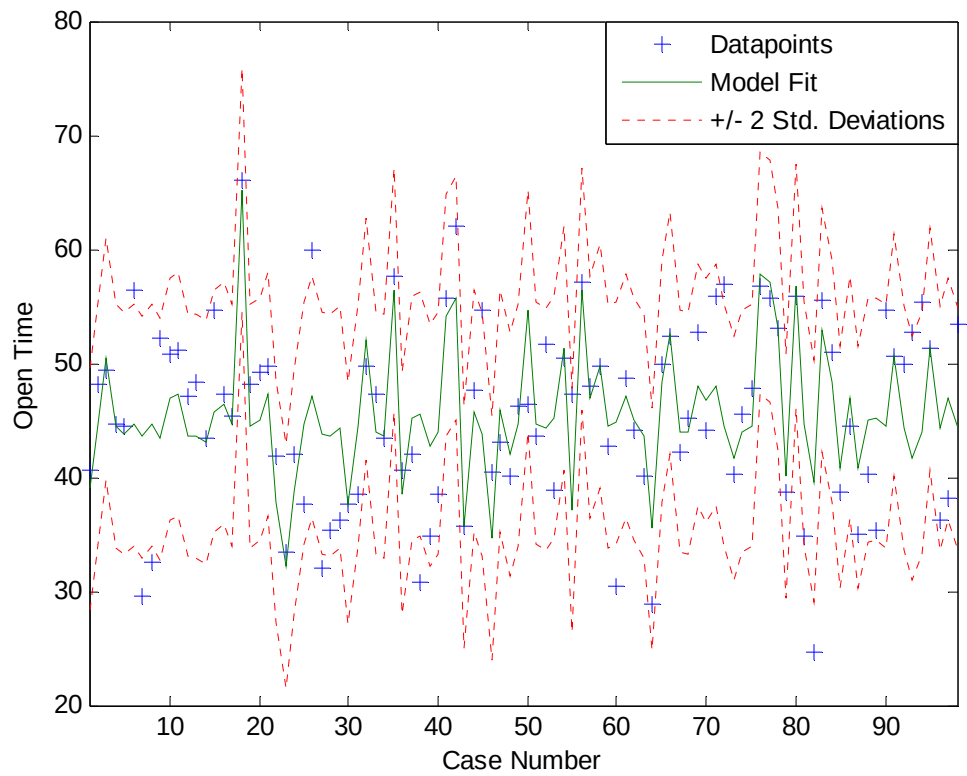
Water Permeability



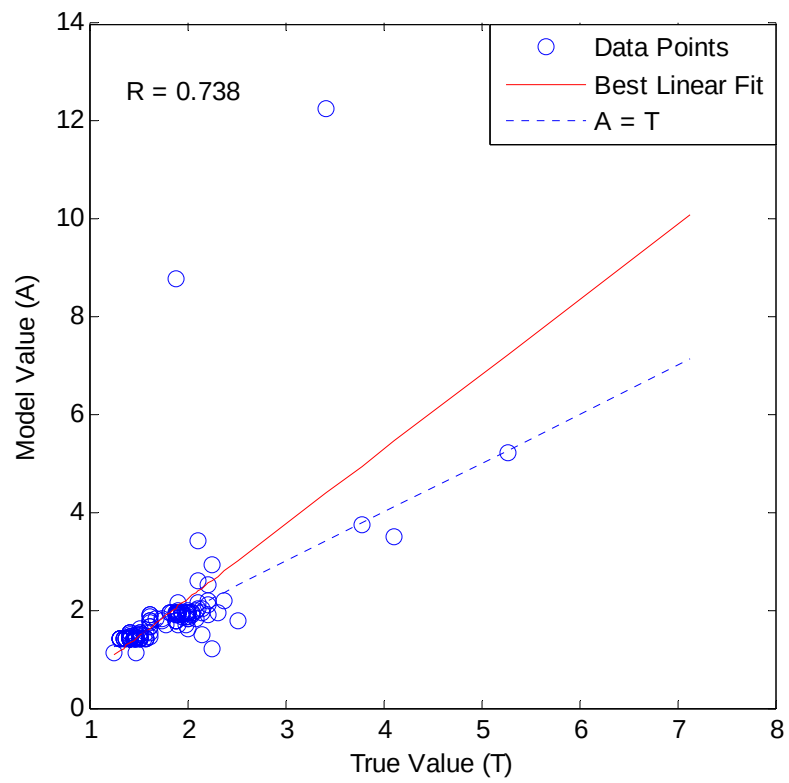
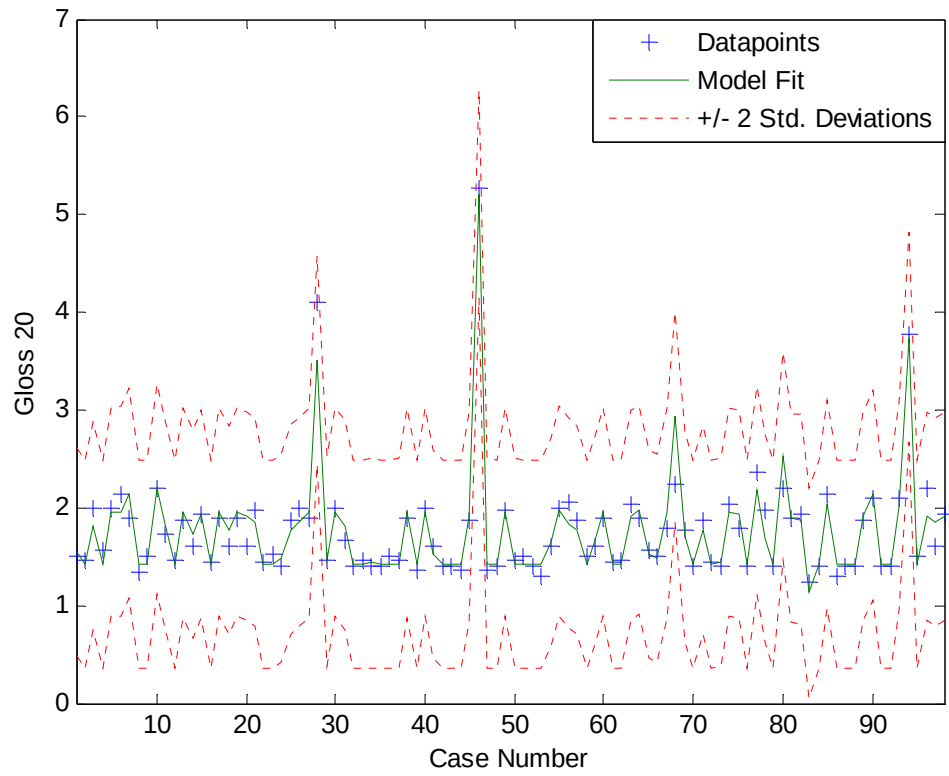
Krebs



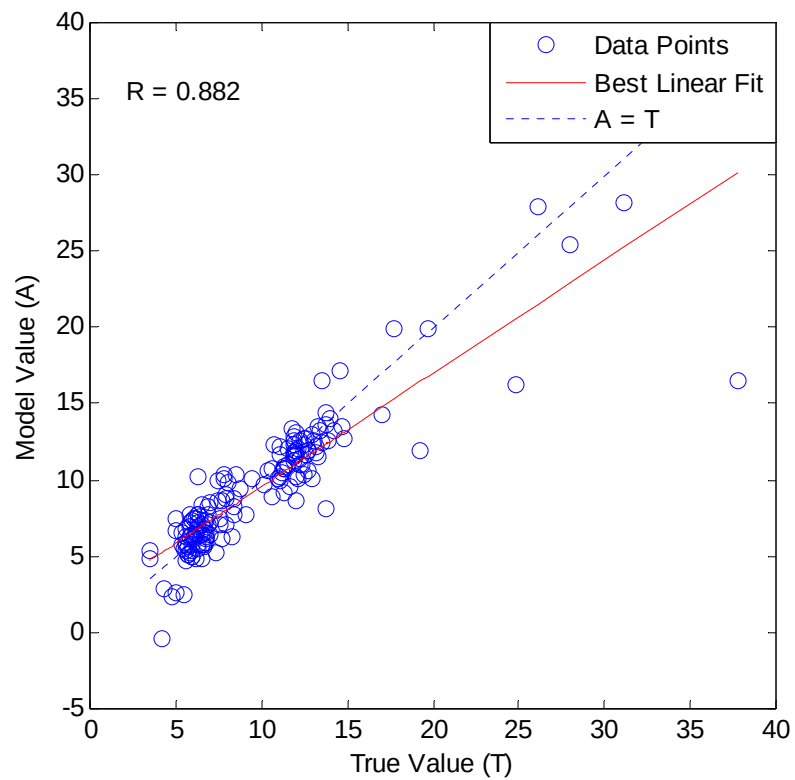
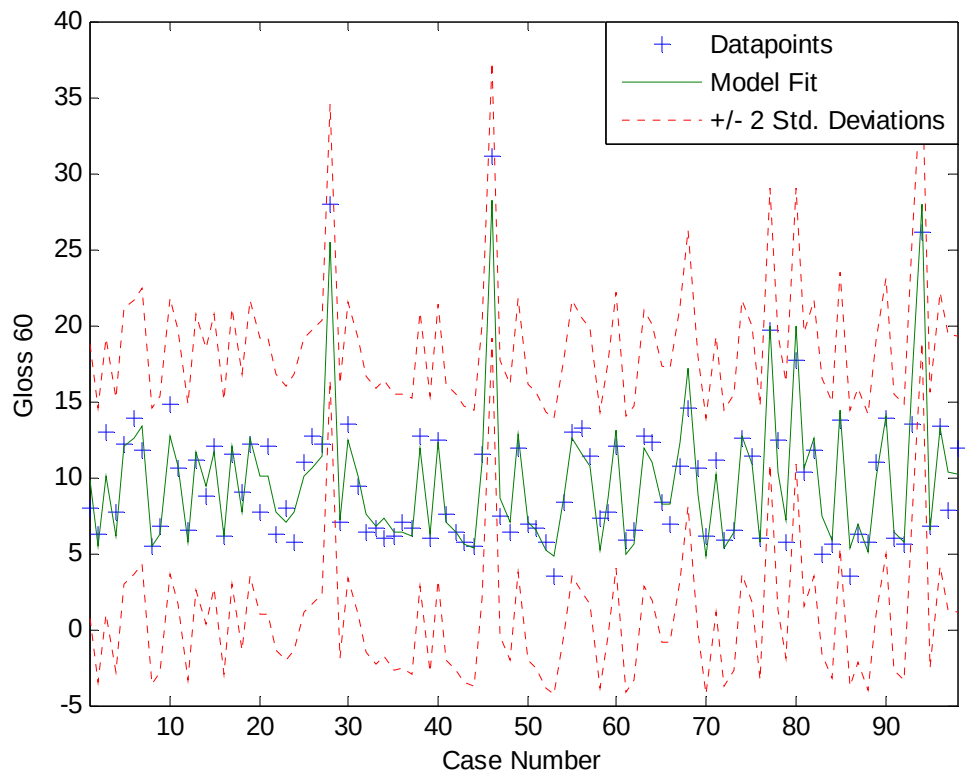




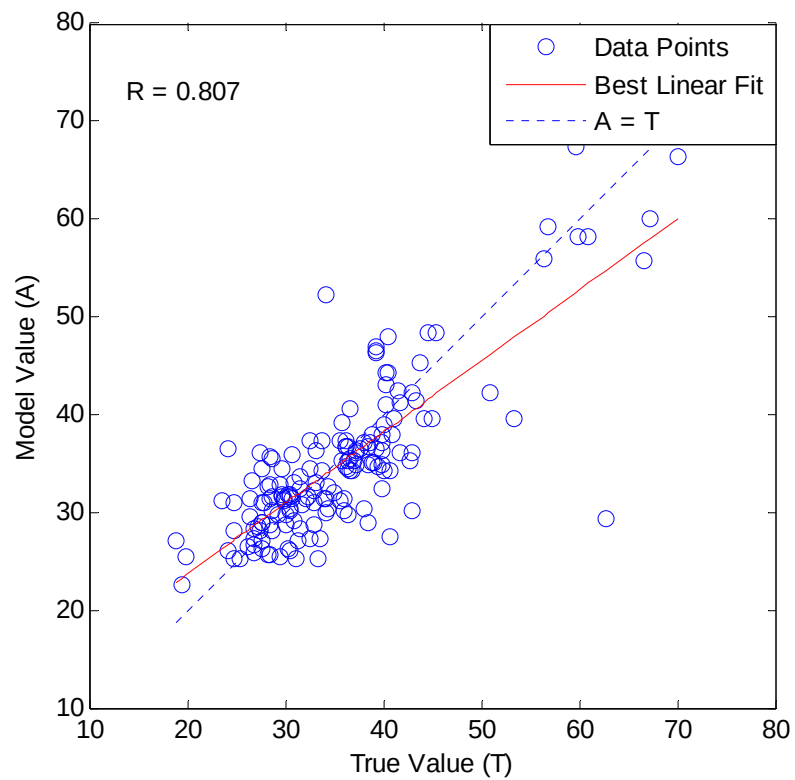
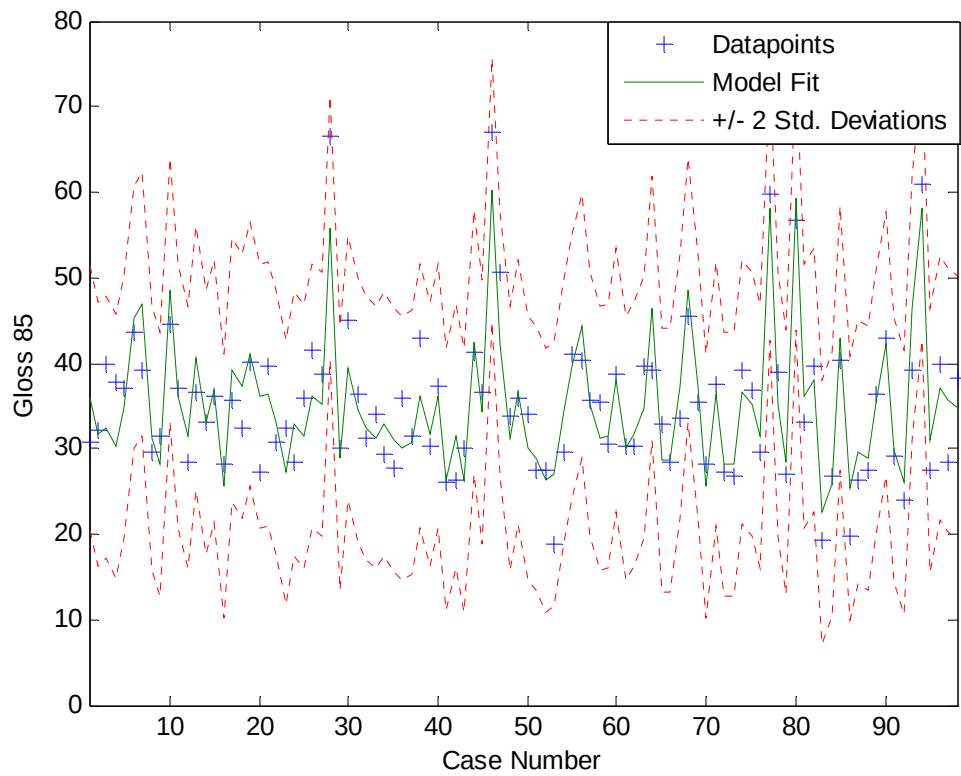
Gloss 20



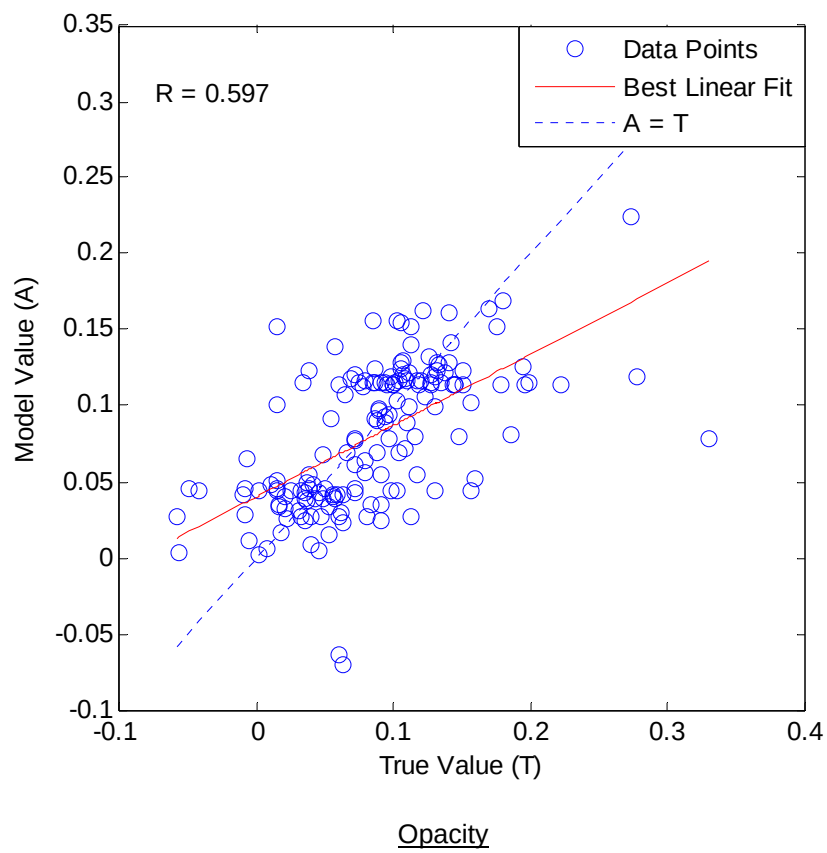
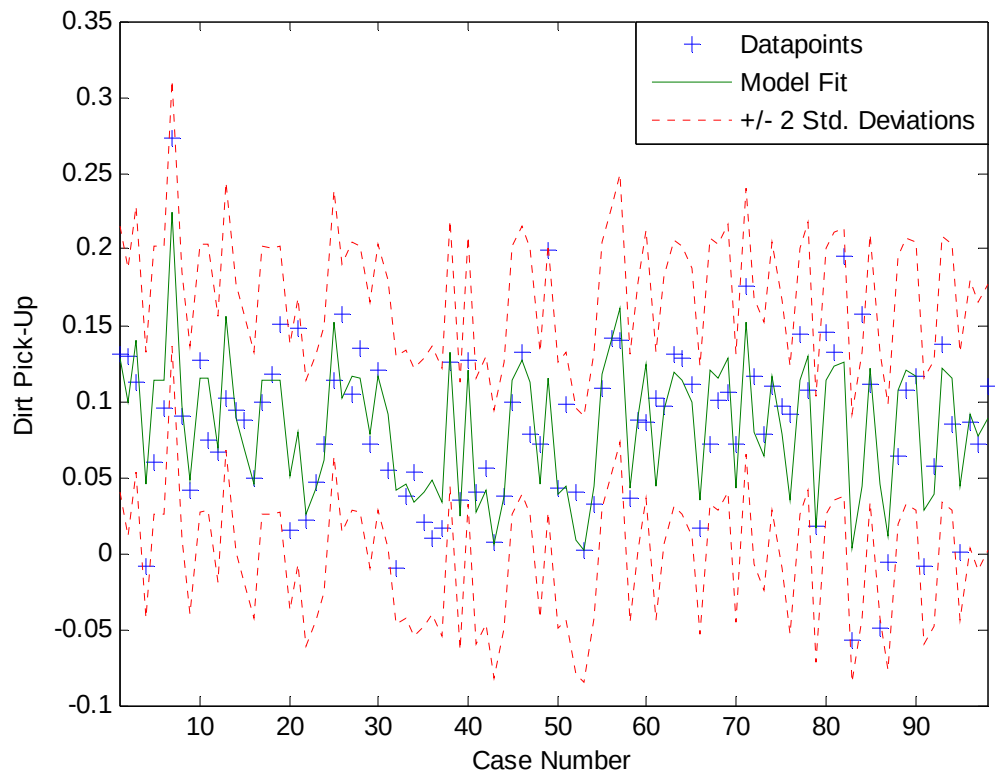
Gloss 60

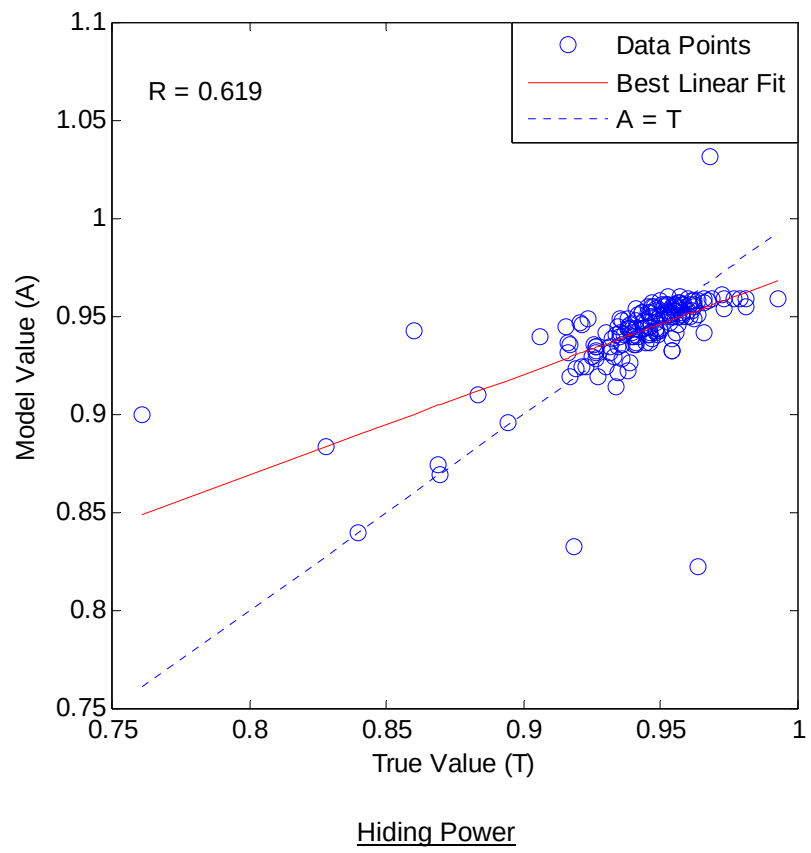
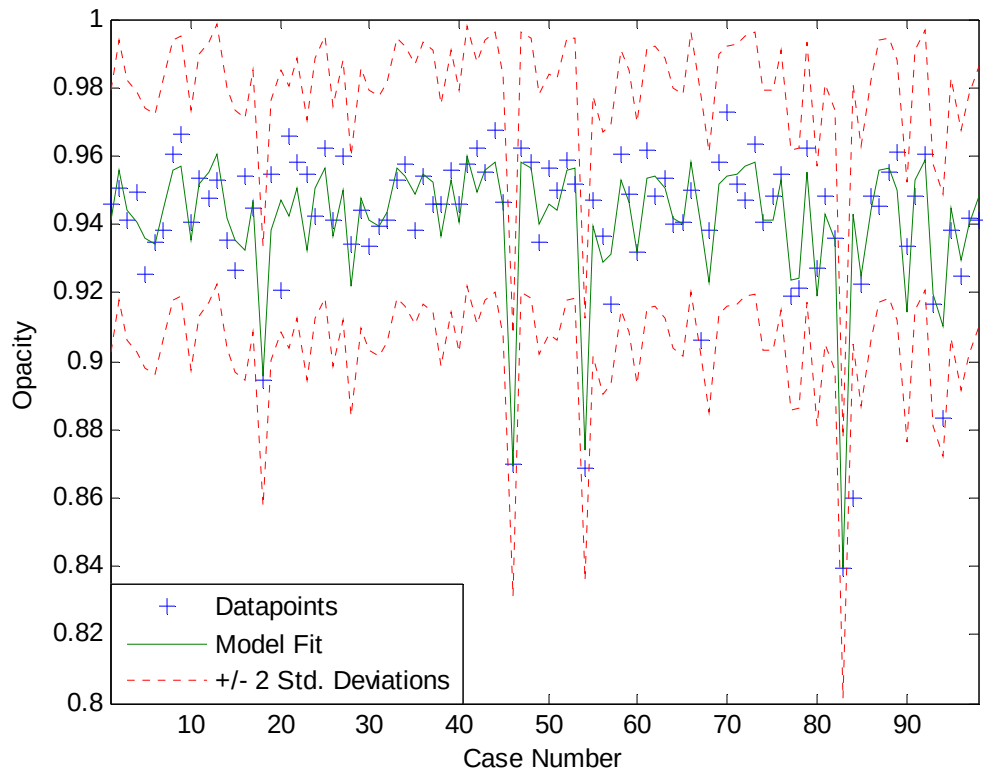


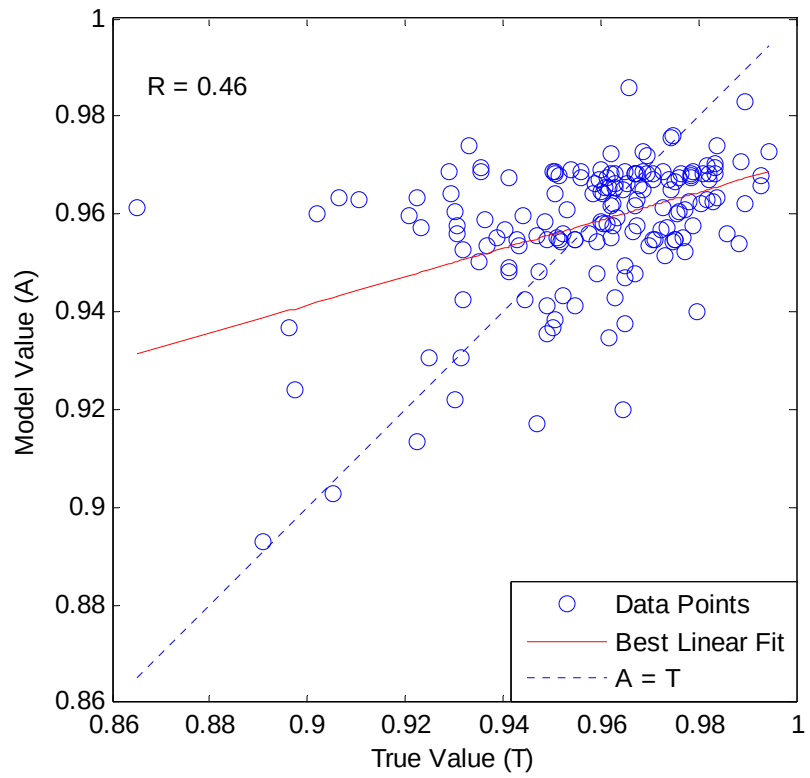
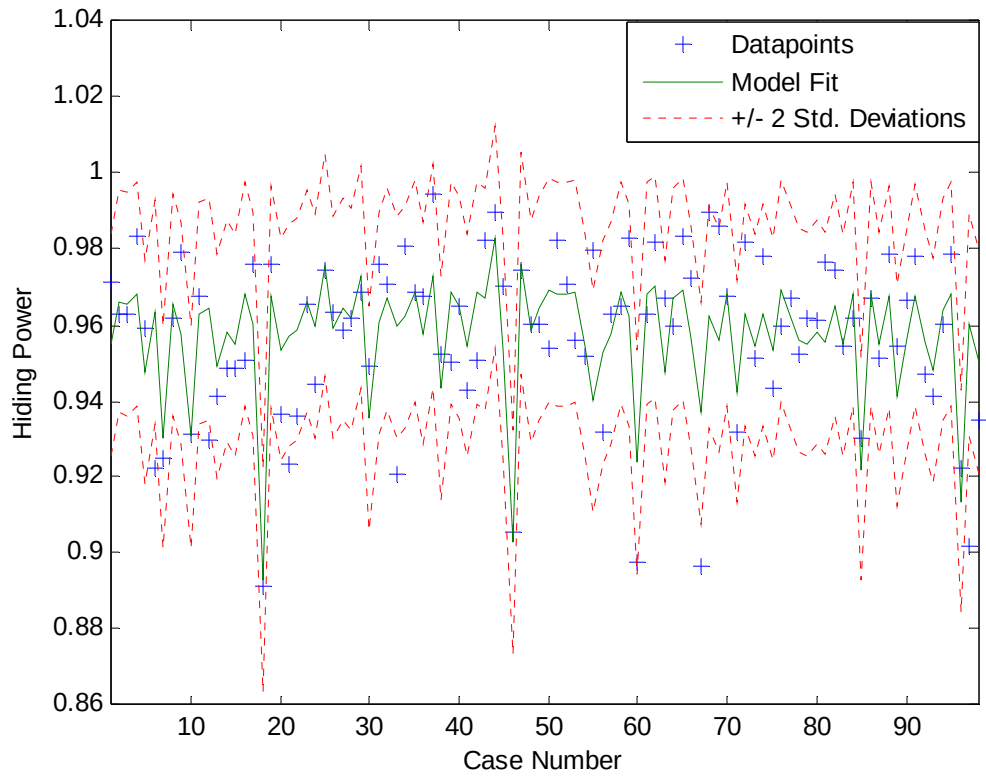
Gloss 85



Dirt Pick-Up







Dry Film Thickness

