The Design and Analysis of Mixture Experiments with Applications to Glazes

by Matthew Bursnall

A thesis submitted to the University of Sheffield for the degree of Doctor of Philosophy

Probability and Statistics Department

February, 2001

THESIS CONTAINS

VIDEO

CD

DVD

TAPE CASSETTE

TABLE OF CONTENTS

Acknowledgements

Firstly I would like to thank my supervisors: Dr Eleanor Stillman, Dr Richard Martin and Dr Angela Seddon, for support and guidance throughout the project. Particular thanks must go to Dr Stillman for her continued support and help, which was above and beyond the call of duty. I am also grateful to N. J. A. Sloane of AT&T Labs Research for supplying the gosset program and for help in its implementation.

I am extremely grateful to Adrian Asling and Dr Howard Winbow at Johnson Matthey Ceramics plc, for helping me to find my way in a completely new subject.

In the Statistics department I would like to thank the following postgraduates for making the University an enjoyable and stimulating place to work: Mark Pilling, Sofia Mucharreira, Jeremy Oakley, Julie Hopkins and Martin Day. I would also like to thank Maralyn Smith and Pauline Hollow in the Statistics Office.

I would also like to thank the non-statisticians who were 'involved' in the project. Firstly my house mates: Rob Tate, Nicole Roche and Patrick Kenny; who have had to carry me in terms of houshold chores and cooking as the final deadline approached. Secondly my parents for helping me collate and proof read the final drafts.

Finally I would like to thank Professor John Kent for understanding the stresses involved when one is both working full time and studying.

The Design and Analysis of Mixture Experiments with Applications to Glazes

by Matthew Bursnall

Summary

This thesis is concerned with *mixture experiments*. A mixture experiment is one in which certain properties of a product depend on the relative proportions of its ingredients. This thesis looks at existing and new methods for generating optimal mixture designs, as well as the form of the optimal design. A mixture experiment of industrial relevance is designed and analysed.

The example used is the production of suspensions of ceramic glaze. Johnson Matthey Ceramics plc (JM) are interested in producing a model for predicting the viscosity of these suspensions. This viscosity will depend on the proportions of each ingredient and certain processing variables.

In seeking optimal designs attention is restricted to designs over candidate sets consisting of points which are known to occur frequently in optimal designs. A variety of computer based methods for generating optimal designs were investigated by comparing the designs they produce for the regular simplex with known optima. It is shown that exchange algorithms using these candidate sets produced near efficient designs.

Little is known about the form of the optimal design for *constrained* mixture problems. An algorithm is developed to generate continuous D optimal designs over a given candidate set. The form of these D optimal designs is investigated for the quadratic and cubic models for three components (q=3), and the quadratic model for q=4. The designs for the quadratic models included all vertices plus selections from the mid points of long edges and the centroid. Designs for the cubic model also included selections from thirds of edges, Mikaeili points and axial check points. The weights of points were found to be particularly dependent on the those of their near neighbours. The form of the V optimal design for a quadratic model and q=3 was also investigated. The conclusions were similar to those for D optimality, although the centroid was often weighted considerably more highly.

Practical situations often require repeated trials. Various algorithms were developed for generating optimal designs with r trials included twice. For q > 3, the best results were obtained using an exchange algorithm which only considered designs with r trials repeated.

Following an extensive programme of experimental work, and detailed discussion with JM to identify important responses, models were produced to describe the relationship between viscosity and the explanatory variables. It was found that increasing the amount of clay and frit in the mixture increased the viscosity, while increasing water decreased it. As particle size was reduced and flocculant was added the effects which the mixture variables had on the top plateau and yield point was found to increase in magnitude.

1 Introduction

1.1 Aims

This thesis is concerned with *mixture experiments*. A mixture experiment is one in which certain properties of a product depend on the relative proportions of its q ingredients. For simplicity the ingredients of interest are usually scaled so that they sum to 1. The requirement that the ingredients sum to 1 is known as the fixed-sum constraint. If we change the proportion of one variable then the proportion of another, or a combination of others, must also be changed so that the sum remains equal to 1.

The initial motivation for this project came from two separate sources, one being statistical and the other being industrial. The statistical design and analysis of mixture experiments is a relatively new subject area with many unexplored aspects. The statistical aim of the project is to explore some of these aspects. The specific aspects to be explored are detailed in the objectives section (see §1.2)

The industrial motivation came from Johnson Matthey Ceramics plc. For the remainder of the thesis Johnson Matthey Ceramics plc will be referred to by the abbreviation JM. One of the products produced by JM is *liquid glaze*. Liquid glaze is a suspension of glass and clay particles in water, which can be applied to ceramic items, or *wares*, to provide a protective and decorative finish.

Glaze can be applied to a ware using a variety of processes. Each process requires a glaze with different *rheological* properties. Rheology is the study of *viscosity*. Viscosity is a measure of how well a liquid resists motion when a force is applied to it. The viscosity of a glaze is determined by the proportion of each ingredient and some qualitative factors which represent different production methods. These two types of variable are referred to as *mixture variables* and *process variables* respectively.

The mixture and process variables have been varied by JM, prior to the current project, through a process of trial and error, to provide a glaze of the required viscosity. The company now wish to control viscosity more accurately using a

mathematical model to describe how the relevant variables affect viscosity. The practical aim of the project is to construct this model.

1.2 Objectives

1.2.1 Experimental Design

When an experiment is performed both time and financial constraints can limit the number of trials which can be produced and analysed. The design of experiments deals with selecting the best trials to be performed given this restriction. Each of the trials selected is referred to as a design point. A discrete design is one where a finite number design points are specified. In a discrete design each design point can be included more than once. A *continuous design* is one where distinct designs points are specified together with a *weight*. The weights specify the proportion of trials which would ideally be placed at each point.

An *optimal design* is one which meets a given criterion. An *efficient design* is one which is good according to a given criterion. A variety of criteria can be used. The majority of these criteria involve minimising some function of the variance of the model parameters. There is a variety of software packages which can be used to find efficient designs for specific problems.

Previous work on efficient mixture designs has mainly been limited to experiments where each variable can range between 0 and 1 (*non-constrained problems*). The co-ordinate system used to represent all possible mixture formulations is known as the *simplex*. In real problems it is more common to have additional restrictions placed on the variables. These problems are known as *constrained problems*. These additional restrictions limit the possible mixture formulations to a subregion of the simplex known as the *constrained region*. Because much is known about optimal designs for non-constrained problems, but little is known about optimal designs for constrained problems, the thesis concentrates mainly on constrained problems.

The smallest non-trivial design problem is one where there are 3 mixture variables (i.e. 3 components). This type of problem also occurs most commonly in real situations, though problems with 4 components also occur relatively frequently.

The most commonly used optimality criterion is the D criterion, but the V criterion is becoming increasingly popular. The most commonly used starting model is the quadratic model. Other common models are the linear model, the special cubic model and the full cubic model.

All experiments are subject to a variety of sources of error. One way to estimate how much error an experiment is subject to, is to repeat (or replicate) some of the trials, and compare the results. There is only one method (Snee, 1979) which is currently available for finding optimal designs where a specific number of replicates is required.

In terms of experimental design, the main objectives of the thesis are as follows:

- Implement the KL algorithm (Atkinson & Donev, 1992) for the V optimality criterion.

- Assess the software which is available for producing D and V optimal designs.

- Investigate the general form of D and V optimal designs for unconstrained problems with 3 components. All four models are to be considered.

- Investigate the general form of designs which are optimal over a given set of points, for constrained regions. Such sets are known as the *candidate points* and will consist of points which occur frequently in known optimal designs. Optimal designs will be found for a range of constrained regions. We will start with the regular simplex and gradually modify it by using constraints of increasing complexity. The observed trends in the optimal designs will then be investigated. The following situations are to be considered. The first situation is to be investigated most thoroughly.

  - D optimal designs for 3 components and the quadratic model.

  - D optimal designs for 3 components and the full cubic model.

  - D optimal designs for 4 components and the quadratic model.

  - V optimal designs for 3 components and the quadratic model.

- Develop methods for finding optimal designs which have a specified number of points included twice. These methods will then be compared with the existing method.

1.2.2 Predicting Glaze Viscosity

JM is currently interested in how the viscosity of glaze is affected by a variety of predictor variables. The main variables of interest are the (mixture) components of the glaze formulation. Two process variables are also of interest. These process variables are both qualitative and are used to represent different production methods. To produce the glaze, three components; crushed glass, clay and water, are milled together until the solid particles reach a specified size. Because JM is only interested in the approximate particle size, the complete range of possibilities is partitioned into three size ranges and the level of the particle size process variable indicates which one of these ranges the glaze falls into. The second process variable is whether or not a *flocculant* has been added. Flocculant is the name given to any substance which is known to cause aggregation. If none of the aggregates sink to the bottom of the sample, the effect of adding flocculant will be an increase in the viscosity of the slip.

When the two process variables are combined there is a total of 6 possible settings. Due to practical considerations it is easier to split each batch of mixture into 6 separate samples for each of these settings. This means that the experiment has what is known as a *split plot structure* and means that samples from a particular batch are subject to similar error effects. That is, the measurements on each of the 6 samples in a batch are distorted to the same extent by external factors. To account for the expected similarities in the errors, we need to include a non-trivial covariance structure in the models. Numerical ML and REML are to be used to fit such models.

A key property of glaze is that its viscosity, i.e. the coefficient which determines its ability to resist flow, depends on the size of the *shear stress* trying to cause this flow. Shear stress is defined as the force trying to cause the flow divided by the area it is being applied to. JM is interested in the relationship between viscosity and shear stress.

In terms of viscosity prediction the objectives of the project are as follows.

- Mimic the industrial process used by JM in the Engineering Materials Department of the University of Sheffield.

- Perform a series of experiments where the viscosity is measured for different values of the predictor variables.

- Summarise the relationship between viscosity and shear stress by 3 suitably selected response variables.

- Fit a model for predicting the response variables (i.e. the 3 viscosity parameters) from the predictor variables.

- Analyse the model and determine how each variable affects viscosity. Recall that the variables are the mixture components; crushed glass, water and clay and the qualitative processing conditions; particle size and flocculant. Within the industry the crushed glass is referred to as *frit*.

1.3 Main Features

1.3.1 Experimental Design

The KL algorithm was successfully implemented in *matlab* for the V optimality criterion. All packages which were assessed for their experimental design capabilities were found to produce efficient designs successfully. Some packages regularly produced designs which were more efficient than those produced by other packages. However, in all but one case the differences in the criterion values were not considered to be large enough to provide any practical improvement.

The investigation of design optimality, for the unconstrained simplex, suggested a small number of minor new results.

An algorithm was implemented, to estimate the weights of the continuous D optimal design, over a given set of points, for constrained problems. Design problems outlined in the aims were all investigated. In all cases the optimal designs generated used points which were spread reasonably evenly throughout the feasible regions. The

weights of these points appeared to depend on the number, the weights, and the distance to, other points which were nearby.

Using the algorithm, the general form of designs which are optimal over specified candidate sets were thoroughly investigated for a quadratic model and q=3. The types of points used in these designs are now discussed. In every case, all *vertices*, i.e. points which occur where 2 edges of the constrained region meet, were included. When particular edges were relatively long, their *mid-points*, i.e. points which occur equi-distant from the 2 vertices, also tended to be included. The *centroid*, a point in the middle of the constrained region, was also usually included.

Less thorough investigations were carried out for the quadratic model and four components and for the cubic model and three components. For q=4 the conclusions were similar to those obtained for q=3. For the cubic model the D optimal design sometimes used non-central long-edge points instead of the mid-edge points, and other internal points instead of, or as well as, the centroid.

Using a less precise method, weights were estimated for a range of V optimal continuous designs for q=3 and the quadratic model. The continuous V optimal designs tended to use the same points as the continuous D optimal designs. However, the mid-points of long edges tended to be weighted higher than they were for the D criterion and the centroid tended to be weighted considerably higher than it was for the D criterion.

An algorithm was implemented to generate optimal designs, over a given set of points, which included a pre-specified number of points twice. For q > 3 this algorithm was found to improve on the existing method.

1.3.2 Predicting Glaze Viscosity

When the industrial process used by JM was mimicked at Sheffield University the results displayed some major errors, inconsistencies and counter-intuitive conclusions. As a result the experimental trials were performed in the JM laboratory. In these trials the rheological properties of a range of ceramic glaze were measured.

Three summary measures of glaze rheology were identified. These summary measures were estimated for each of the glaze slips produced, using the data collected in the trials. Models were then fitted to explain how the explanatory variables affected

the values of the three summary measures. It was found that in general, increasing the amount of clay and/or frit in the mixture increased the values of all three summary measures, while increasing water decreased them. The conclusion about the effect of frit was particularly interesting because it was not known before. It was also found that adding flocculant, and/or decreasing the size of the solid particles in the slip, also increased the value of all three summary measured.

Chapter 2

Background

2.1 Statistical Models

2.1.1 Fixed Effects

When a series of experiments is performed, each trial will usually yield a different value of a measured *response variable*. The response is usually represented by the random variable Y. The main aim of an analysis is usually to derive a model explaining how Y varies with the explanatory $\underline{x}$ variables.

A family of models which is commonly used is the polynomials. Let q denote the number of explanatory variables, and denote their values, for trial i, by $\underline{x}_i = (x_{1i}, x_{2i}, \ldots, x_{qi})'$. The polynomials of degree 1 to 3 are given in equations 2.1 to 2.3, for the special case of q=3.

Linear $\qquad E(y_i) = \mu + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i}.$ $\hfill (2.1)$

Quadratic $\qquad E(y_i) = \{\text{Linear model}\} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{33} x_{3i}^2$

$$+ \beta_{12} x_{1i} x_{2i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i}. \qquad (2.2)$$

Cubic $\qquad E(y_i) = \{\text{Quadratic model}\} + \beta_{112} x_{1i}^2 x_{2i} + \beta_{122} x_{1i} x_{2i}^2 + \beta_{113} x_{1i}^2 x_{3i}$

$$+ \beta_{133} x_{1i} x_{3i}^2 + \beta_{223} x_{2i}^2 x_{3i} + \beta_{233} x_{2i} x_{3i}^2 + \beta_{111} x_{1i}^3 + \beta_{222} x_{2i}^3 + \beta_{333} x_{3i}^3. \quad (2.3)$$

In addition to the quantitative explanatory variables, the response may also depend on some qualitative conditions. Factor variables are used to specify qualitative conditions and their values are often coded to be integers.

In general, it is usually preferable to express a model in terms of matrices. Imagine that n trials are to be produced and for each trial the value of the response variable Y is denoted by the corresponding element of a vector $\underline{y}$. The values of the explanatory variables

for each of the n trials are given by the appropriate row of the design matrix X. This row of the design matrix gives both the values of the explanatory variables themselves, and the appropriate functions of them which are specified in the chosen model. Each row of the design matrix thus specifies a *model augmented point*. The model augmented point for $\underline{x}_i$ is subsequently referred to as $\underline{x}_i$. The i'th row of the design matrix is therefore $\underline{x}_i'$. The model, in matrix form, is:

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}, \tag{2.4}$$

where $E(\underline{Y}) = X\underline{\beta}$.

Here $\underline{\beta}$ is the vector of the parameters in the chosen model and $\underline{\varepsilon}$ is a vector of errors. The error terms are present because in the majority of cases the observed values of $\underline{Y}$ are not determined precisely by X according to an exact relationship. If we assume that the errors are independent and identically distributed (iid) Normal variables, with variance σ^2, then $Y \sim N(X\underline{\beta}, \sigma^2 I)$.

2.1.2 Random Effects

In certain experimental situations it is more sensible to assume that the errors, or specific groups of errors, are correlated with each other. In such cases we can specify the covariance structure of the error vector , $\underline{\varepsilon}$, by $V\sigma^2$, where V is an n×n matrix. The diagonal elements of the matrix V give the scaled variance of the measured response for each trial. The off-diagonal elements give the scaled covariance between each pair of measured responses. For uncorrelated pairs of responses the appropriate off-diagonal element will be 0. If the error vector is also assumed to be multivariate Normal, the model is:

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}, \tag{2.5}$$

where $Y \sim N(X\underline{\beta}, V\sigma^2)$.

2.2 Analysis of Results

2.2.1 Outline

Once a statistical model has been chosen and the data has been collected, the model needs to be fitted. Fitting a model involves estimating the entries of $\underline{\beta}$, i.e. the model parameters. There are a variety of methods which can be used to fit a model. A commonly used approach is Least Squares Estimation (LSE). When some of the data points have considerably higher variances than others, or there is particular interest in certain regions of the explanatory variable space, Weighted Least Squares Estimation (WLSE) is a useful alternative to LSE.

When variance parameters are included in a model, the model parameters are often estimated using Maximum Likelihood estimation (ML). ML is a well known and generally applicable method which produces estimators with good asymptotic properties. However, in general ML estimators of variance parameters are biased downwards. Restricted Maximum Likelihood estimation (REML) is a modified form of ML estimation. In general REML estimators of variance parameters are less biased than ML estimators.

Once the model has been fitted, any assumptions which were made about the errors should be verified. If the main purpose of the model is for prediction, it is also sensible to investigate how well the model is likely to predict future response values. Predictive power can be investigated using the PRESS procedure.

2.2.2 Least Squares Estimation (LSE)

One way to estimate $\underline{\beta}$ is to use least squares estimation (LSE). The principle of LSE involves a quantity known as the squared error. To calculate the squared error for an individual trial we estimate its response using a model, find the difference between this estimated response and the observed response, and square this difference. We can then add up the squared errors for all trials to calculate the error sum of squares (SS_E). LSE involves choosing the model which has the lowest SS_E. Calculus shows that the LS estimate of $\underline{\beta}$ is:

$$(X'X)^{-1}X'\underline{y}. \qquad\qquad (2.6)$$

If the X′X matrix does not have full rank it can not be uniquely inverted and only a subset of the parameters can be estimated.

LS estimation is unbiased for $\underline{\beta}$ and, amongst the set of all other unbiased estimates of $\underline{\beta}$, it has the smallest variance. Additionally, if the individual entries of $\underline{\varepsilon}$ are independent, and $\underline{\varepsilon} \sim N(0, I\sigma^2)$, the LS estimate of $\underline{\beta}$ is also the ML estimate of $\underline{\beta}$, $\hat{\underline{\beta}}$.

2.2.3 Weighted Least Squares Estimation (WLSE)

When the response has a higher variance for some of the trials, for example the response for some trials may have been measured less accurately, Weighted Least Squares Estimation (WLSE) is a useful alternative to LSE.

WLS estimation involves a similar procedure to LSE but the squared error for each trial is multiplied by some constant known as its *weight*. In the example where the variance of the response is higher for some trials, the squared error for the trials with the higher response variance might be multiplied by k (k < 1) in the LSE process. If the weights are given by the vector $\underline{w}$ numerical minimisation of equation (2.7) over the range of $\underline{\beta}$ is needed.

$$\sum_i w_i (y_i - X_i \underline{\beta})^2 \tag{2.7}$$

2.2.4 Maximum Likelihood Estimation (ML)

If the covariance structure is specified by V≠I, as in 2.5, the parameters are often estimated using numerical maximisation of the appropriate likelihood function. The likelihood function for the multivariate Normal distribution with a covariance structure specified by $\sigma^2 V$ is given in equation 2.8. Maximising equation (2.8) is equivalent to minimising equation 2.10.

$$L(\underline{\beta}, \sigma^2, \rho) = (2\pi\sigma^2)^{-n/2} |V|^{-1/2} \exp\{(\underline{y}-X\underline{\beta})'V^{-1}(\underline{y}-X\underline{\beta})/2\sigma^2\} \tag{2.8}$$

The numerical minimisation involves using an iterative approach, with $\hat{\underline{\beta}}$ and $\hat{V}$ being given by solving equations 2.9 and 2.10 in turn. This iterative approach is based on a starting estimate of $\underline{\beta}$, obtained by ordinary least squares, i.e. using V=I in equation (2.9).

$$\hat{\underline{\beta}} = (X'V^{-1} X)^{-1}X' V^{-1} \underline{y} \qquad (2.9)$$

$$\min(|V|^{1/n} \underline{e}'V^{-1}\underline{e}) \text{ where } \underline{e} = \underline{y} - X\hat{\underline{\beta}} \qquad (2.10)$$

In addition to the iterative procedure involving equations 2.9 and 2.10, equation 2.10 needs to be solved by an iterative search over the range of the variance parameters.

2.2.5 Restricted Maximum Likelihood Estimation (REML)

ML estimation does not take into account the degrees of freedom associated with estimating the parameters in the model for E(Y). Consequently, in general, ML estimators of the variance parameters are biased downwards. Restricted Maximum Likelihood (REML) is a modified form of ML (see, e.g., Welham and Thompson, 1997). In general REML estimators of the variance components are less biased than ML estimators. The REML likelihood is given in equation 2.11, where p is the number of parameters in the model for E(Y). REML involves replacing equation 2.10 in the ML analysis with equation 2.12.

$$L_R(\underline{\beta}, \sigma^2, \rho)= (2\pi\sigma^2)^{-(n-p)/2}|V|^{-1/2} |X'X|^{1/2}|X' V^{-1} X|^{-1/2}\exp\{(\underline{y}-X\underline{\beta})' V^{-1} (\underline{y}-X\underline{\beta})/2\sigma^2\} \qquad (2.11)$$

$$\min(\{|V|.| X'V^{-1} X|\}^{1/n-p} \underline{e}'V^{-1}\underline{e}), \qquad (2.12)$$

2.2.6 Model Checking

Once a model has been fitted, the assumptions made about the error terms need to be verified. It is also useful to investigate how well the model is likely to predict future response values. This measure of a model is subsequently referred to as predictive power. A popular way of assessing predictive power is to use the PRESS procedure. This procedure involves

removing one of the data points, fitting the chosen model using the remaining points, estimating the response for the removed trial using the subsequent model and calculating the squared deviation between the predicted and observed responses. This is repeated for each data point individually, and the sum of the squared differences is calculated. If a number of models are being compared then the "best" model will have a comparatively small PRESS value but will not involve too many parameters.

2.3 Design Principles with a Focus on Mixture Experiments

2.3.1 Background to Mixture Experiments

A mixture experiment is carried out in order to assess how certain properties of a product depend on the relative proportions of its ingredients. The letter q will be used to denote the number of *components* or ingredients. If we change the proportion of one variable then the proportion of another, or a combination of others, is also changed so that their sum remains constant. The requirement that the sum of the ingredients remains constant is known as the *fixed-sum constraint*. For simplicity the ingredients of interest are usually scaled so that they sum to 1. If the proportions of the q components in a mixture are denoted by the variables $x_1, x_2, \ldots, x_q$ then the fixed sum constraint is expressed as

$$x_1 + x_2 + \ldots + x_q = 1.$$

Depending on whether or not any other constraints are imposed we can split mixture experiments into *simple* and *constrained* cases. The background to mixture experiments is best introduced using the q=3 component example.

A Simple case

A product is to be made out of the three ingredients A, B and C, the proportions of which are denoted by x_1, x_2 and x_3. The co-ordinate system used to represent all possible mixture formulations is known as the *simplex*. For the simple example where q=3 the simplex is illustrated in Figure 2.1. Each mixture formulation which is made and analysed can be referred to as a *trial*. Clearly there are an infinite number of possible mixture formulations. The diagram is annotated with 7 commonly used compositions.

Figure 2.1 Regular simplex The regular simplex when q=3. This simplex can be used to represent all possible formulations of 3 mixture components.
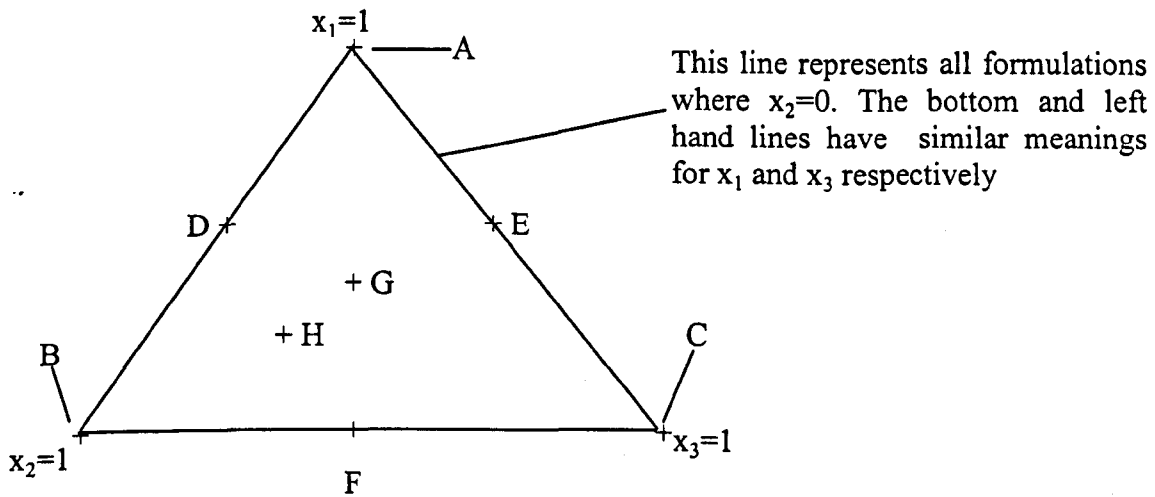
Point A on Figure 2.1 represents a *pure blend* including ingredient A only. In this case the proportions of the ingredients are $x_1=1$, $x_2=0$ and $x_3=0$. There are similar points for ingredients B and C, these are represented by points B and C.

Point D is the *binary blend* of A and B with proportions $x_1={}^1/_2$, $x_2={}^1/_2$ and $x_3=0$. For simplicity we denote the values of the variables using vectors, the vector which represents point D is $({}^1/_2, {}^1/_2, 0)$. Again E and F are similar binary blends.

Point G is called the *centroid* and is situated at the centre of gravity of the shape. In the simple case this is at the point where all proportions are equal, i.e. $({}^1/_3, {}^1/_3, {}^1/_3)$.

As mentioned above, any point on the simplex represents a possible mixture. Point H represents one of the infinite number of possibilities and has the proportions $({}^1/_4, {}^1/_2, {}^1/_4)$.

Diagrams similar to Figure 2.1 can be produced for q=2 and q=4. The 2 component case can be represented by the left hand line of the simplex triangle. For four components a triangular pyramid would be used where each face is similar to the three component simplex in Figure 2.1.

B Constrained case

In practice the proportions of components in a mixture experiments are often subject to lower and upper bounds in addition to the fixed sum constraint. In general these bounds take the form $a_i \leq x_i \leq b_i$. As an example imagine that the product is made out of three

ingredients, but it is known that it will not work unless the proportions of each ingredient are in the following ranges:

$$0.2 \leq x_1 \leq 0.8, \qquad 0.1 \leq x_2 \leq 0.7, \qquad 0.1 \leq x_3 \leq 0.7.$$

·· These ranges can be added to the simplex diagram to produce a *feasible region*. The feasible region for the above constraints is the hexagon shown in Figure 2.2.

Figure 2.2 Example of a feasible region This feasible region is generated by the lower and upper bounds given above.
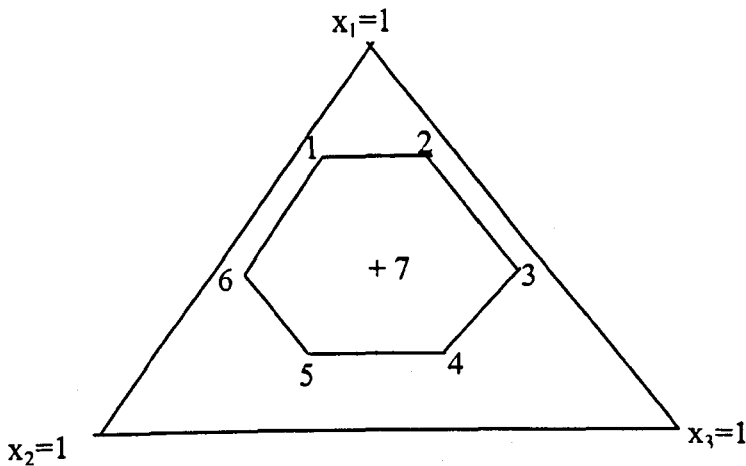
Points 1 to 6 are known as *extreme vertices* of the feasible region and point 7 is a central point. The central point is similar to the centroid in Figure 2.1 but due to the irregular nature of most feasible regions, the centre of gravity centroid can be hard to find. Instead of using the centre of gravity, the centroid is often estimated by averaging the co-ordinates of the extreme vertices to give the *average vertex (a.v) centroid*. The co-ordinates of the extreme vertices were averaged to locate the central point, i.e. point 7, in Figure 2.2.

15

Another form of constraint is the *multicomponent constraint*. Multicomponent constraints specify lower and upper bounds for linear combinations of components, in the following way

$$A_k \leq \sum_{j=1}^{q} c_{jk}x_j \leq B_k \ (k=1, \ldots, K).$$

To simplify constrained problems the feasible region can be transformed to fit tightly within the regular simplex using *pseudocomponents* (Cornell 1990). There are two types of transformation which lead to pseudocomponents, both transform the lower bounds to zero. If q=3 and there are no multicomponent constraints, the transformed regions are either hexagonal, pentagonal, quadrilateral or triangular (see §3.5.1). The transformations are performed using equations 2.13 and 2.14 below, and are known as the *L transformation* and the *U transformation* respectively. The L transformation pushes the lower bound for component i to the edge $x_i = 0$. The U transformation produces an enlarged mirror image of the feasible region by pushing the upper bound for component i to the line $x_i = 0$. Both transformations preserve the ratios of edge lengths and leave the type of shape unchanged. The transformation, which will lead to the transformed feasible region which occupies more of the regular simplex, varies from case to case. Crosier (1984) gives conditions for when each transformation leads to the smaller region. As the L transformation preserves the orientation of the feasible region it is in some sense the more natural of the two to use.

If the bounds are of the form $0 \leq a_i \leq x_i \leq b_i \leq 1$ (i=1, 2, . . . , q) then the L pseudocomponents (x_j') are defined by equation 1.2. and the U pseudocomponents (x_j'') are defined by equation 2.14.

$$x_j' = (x_j - a_j) / R_L, \qquad \text{where } R_L = 1 - \sum_{i=1}^{q} a_i \,. \qquad (2.13)$$

$$x_j'' = (b_j - x_j) / R_U, \qquad \text{where } R_U = \sum_{i=1}^{q} b_i - 1. \qquad (2.14)$$

The L transformation makes the lower bound equal zero for each x_j' with the upper bounds given by equation 2.15. The U transformation also makes the lower bound for each component equal to zero, the resulting upper bounds are given by equation 2.16.

$$b_j' = (b_j - a_j) / R_L.$$ (2.15)

$$b_j'' = (b_j - a_j) / R_u.$$ (2.16)

With the L transformation, if $b_j = 1 - \sum_{i \neq j}^{q} a_i$, then the upper bound for x_i' equals 1. If this happens for all components then the L transformed feasible region is the original simplex of Figure 2.1. The U transformed feasible region is the regular simplex if $1 - a_j = \sum_{i \neq j}^{q} b_i$ for all components. Whether or not each of the transformed upper bounds is 1 determines the shape of the transformed feasible region, and has consequences for the design of experiments (see §3.5.1).

2.3.2 Models for Mixture Experiments

There is a variety of models which have been presented for mixture experiments (Becker 1978, Scheffé 1958 and Cornell 1990 §6.3, §6.5, §6.9). In this thesis attention is restricted to the *Scheffé canonical polynomials*. The Scheffé equations of degree 1, 2 and 3 are given below for q=3.

1. Linear $\quad$ $E(Y)= \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$ (2.17)

2. Quadratic $\quad$ $E(Y)= \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3.$ (2.18)

3. Special Cubic

$\quad$ $E(Y)= \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3.$ (2.19)

4. Full Cubic

$$E(Y) = \{\text{Special Cubic}\} + \delta_{12}x_1x_2(x_1-x_2) + \delta_{13}x_1x_3(x_1-x_3) + \delta_{23}x_2x_3(x_2-x_3). \qquad (2.20)$$

The x_ix_j terms represent the interactions between mixtures i and j. The δ_{ij} are used to describe terms where the product of two variables is multiplied by their difference.

The two main departures from the polynomial models given above (see §2.1.1) are that there is no constant term and there are no pure powers in the quadratic or cubic forms. These terms are shown to be unnecessary in the following way. Suppose that constant term β_0 and pure quadratic term $\beta_{11}x_1^2$ are added to a quadratic model of the form given in equation 2.18. If one of the x_1's is replaced using the fixed sum constraint and the mean is multiplied by the fixed sum constraint, the model equation can be rearranged in the following way:

$$E(Y) = \beta_0(x_1+x_2+x_3) + \beta_1'x_1 + \beta_2'x_2 + \beta_3'x_3 + \beta_{11}x_1(1-x_2-x_3) + \beta_{12}'x_1x_2 + \beta_{13}'x_1x_3 + \beta_{23}'x_2x_3$$

$$= (\beta_0+\beta_1'+\beta_{11})x_1 + (\beta_0+\beta_2')x_2 + (\beta_0+\beta_3')x_3 + (\beta_{12}'-\beta_{11})x_1x_2 + (\beta_{13}'-\beta_{11})x_1x_3 + \beta_{23}'x_2x_3.$$

This yields the same model as given in equation 2.18,

i.e. $\beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3$, but with

$\beta_1 = \beta_1' + \beta_{11} + \beta_0$, $\beta_2 = \beta_2' + \beta_0$, $\beta_3 = \beta_3' + \beta_0$, $\beta_{12} = \beta_{12}' - \beta_{11}$, $\beta_{13} = \beta_{13}' - \beta_{11}$ and $\beta_{23} = \beta_{23}$.

$$\therefore \beta_1' = \beta_1 - \beta_{11}, \quad \beta_{12}' = \beta_{12} + \beta_{11}, \quad \beta_{13}' = \beta_{13} + \beta_{11}.$$

An alternative parameterisations of the Scheffé polynomials is sometimes used. The re-parameterisation is introduced here because it is used by the computer algorithm gosset (see §2.3.7D). The gosset algorithm is used in many parts of the thesis. The alternative parameterisations can be derived from the originals using the fixed sum constraint to replace one of the components. Unlike the original parameterisations, the alternatives include a constant term, and terms involving individual variables raised to powers. The equations for the alternative parameterisations are given below for the special case of $q=3$. Although these equations include the first two components and not the third, any of the q components could

be selected for removal without loss of generality. These re-parameterisations are subsequently referred to as the *intercept form* of the Scheffé polynomials.

1. Linear $\quad E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$

2. Quadratic $\quad E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{12} x_1 x_2 + \beta_{22} x_2^2.$

3. Special cubic $E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{12} x_1 x_2 + \beta_{22} x_2^2 + \beta_{(12)}(x_1^2 x_2 + x_1 x_2^2).$

4. Cubic $\quad E(Y) = \{\text{Quadratic model}\} + \beta_{111} x_1^3 + \beta_{112} x_1^2 x_2 + \beta_{122} x_1 x_2^2 + \beta_{222} x_2^3.$

It becomes apparent in the next section that the choice of model will influence the experimental design. Therefore the model that will be used initially to describe the data needs to be chosen in advance (though subsequent checking of quality of fit may cause one to revise the model used initially). However the parameterisation of the chosen model will not influence the design. The most common choice of model is the Scheffé quadratic because it is much simpler than the cubic model, but still allows the possibility of non linear relationships between a response variable and the mixture proportions. It should also be noted that the fixed sum constraint means that the interpretation of models which have been fitted to mixture data, will be different from the interpretation of the polynomial models (see §2.1.1).

Models combining mixture variables and qualitative variables

Due to the fixed sum constraint, when models are to include both mixture variables and qualitative variables, the interactions between mixture and qualitative variables can not all be used. If the qualitative variables were denoted by z_j, the interaction between a mixture variable and a qualitative variable would be denoted $x_i z_j$.

In this project all of the process variables are qualitative variables. Qualitative variables are usually specified in a mathematical model by an indicator vector of integers. Due to the fixed sum constraint it can be shown that the interactions between a qualitative variable and the q mixture variables will have a sum which is equal to the qualitative variable indicator vector. Imagine that a mixture has three components and that it is represented by the

design matrix X $\{X = [x_{ij}]\ i=1, .\ .,4 : j=1, 2, 3\}$. Imagine also that the process variable, z, has 2 levels and is represented by the (indicator) vector $I = (1, 0, 1, 0)'$. The three columns of the design matrix which would represent the three interactions between this process variable and each of the mixture components will have the following sum

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underline{x}_1 + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underline{x}_2 + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \underline{x}_3 = \begin{pmatrix} x_{11} + x_{12} + x_{13} \\ 0 \\ x_{31} + x_{32} + x_{33} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{By the fixed sum constraint}$$

This equality causes a linear dependence in the design matrix X which causes the information matrix to be rank deficient. Consequently the information matrix would not be uniquely inverted and only a subset of the parameters could be estimated. (see §2.2.2).

Cornell (1990, §7.5) suggests solving this problem by dropping one of the mixture variable interactions for each of the qualitative process variables, e.g. include x_1z and x_2z but omit x_3z. The choice of which interaction term to remove will not affect how well the model fits the data.

2.3.3 A Good Design

Usually, time or financial constraints mean that a maximum of n trials can be performed. The purpose of experimental design is to choose the best n trials to perform, in order to maximise the amount of information which will be obtained from the experiment. The set of trials which one chooses to perform is subsequently referred to as a *design* and represented by the symbol ξ. In this thesis attention is restricted to designs where the covariance of the response variable, Y, is specified by V=I.

The $\hat{\underline{\beta}}$ vector (see §2.2.2) is only an estimate of the true parameters. Because the majority of experiments are affected by random factors it is practically impossible to find an exact scientific relationship. These random factors are of two different types. The first type of random factor is *external variation*, which includes any variables that are not included in the model. Two common examples of external variation are wind resistance and temperature . If these variables are different for two different trials, then the differences in the measured

response might be artificially changed. A further example of this type of error could be that the measurement device may perform differently from one trial to the next. The second type of random factor is that the relationship itself may not be exact due to the random variation implicit in physical processes. For example in repeated tests, heat may not be passed through a body in the same way, even if every factor is identical, due to the chaotic nature of particulate behaviour.

A measure of how accurate the parameter estimates are, is:

$$Var(\hat{\beta}) = (X'X)^{-1}\sigma^2 \qquad\qquad (2.21)$$

Equation 2.21 specifies the variance of the sampling distribution for the parameter estimate $\hat{\underline{\beta}}$. The variance of a parameter estimate is a measure of how much the estimate would be likely to vary if the experiment, including the model fitting procedure, was performed many times.

The $100(1-\alpha)$ per cent confidence region for all p elements of β includes all values of $\underline{\beta}*$ which satisfy equation 2.22, where s^2 is an estimate of σ^2 based on a model which includes all p parameters, and $F_{p,\,n\text{-}p,\,\alpha}$ is the α percentage point of the F distribution with p and n-p degrees of freedom.

$$(\underline{\beta}* - \hat{\underline{\beta}})'X'X(\underline{\beta}* - \hat{\underline{\beta}}) \le ps^2 F_{p,\,n\text{-}p,\,\alpha}. \qquad\qquad (2.22)$$

If the experiment were performed many times and this confidence region were constructed each time, approximately $100(1-\alpha)$ per cent of the resulting confidence regions would contain the true $\underline{\beta}$. The volume of the $100(1-\alpha)$ percent *confidence region for the parameter estimates*, is inversely proportional to the square root of $|(X'X)|$ (Atkinson and Donev, 1992, §5.3). The $|(X'X)^{-1}|$ (see, e.g., equation 2.21) is subsequently referred to as the *generalised variance*.

For the model given in equation 2.4, and an estimate of β obtained using equation 2.6, the predicted value of y at $\underline{x}_i$ is given by $\underline{x}_i'\hat{\beta}$ (where $\underline{x}_i$ is the model augmented point for $\underline{x}_i$ (see §2.1.1)). This predicted value has variance given by equation 2.23.

$$\text{Var}(\hat{y}(\underline{x}_i)) = \sigma^2 \underline{x}_i'(X'X)^{-1} \underline{x}_i. \qquad (2.23)$$

For a given design, ξ, the *standardised variance* of $(\hat{y}(\underline{x}_i))$ is obtained by scaling the variance for σ^2 and n, i.e.

$$d(\underline{x}_i, \xi) = n\,\underline{x}_i'(X'X)^{-1}\underline{x}_i \qquad (2.24)$$

When $\underline{x}_i$ is not in the design, the standardised variance is sometimes referred to as the *scaled prediction variance.*

An ideal design would have two features. Firstly it would minimise σ^2. Secondly, it would minimise both the generalised variance of the parameter estimates, and $d(\underline{x}_i, \xi)$ over the design region.

Usually these criteria can not all be met. However, a well designed experiment will have the following features. Firstly σ^2 will have been made as small as possible. Two ways to reduce σ^2 are the use of covariates and the use of blocking. These two ideas are described in sections A and B below. Secondly we choose the $\underline{x}_i$ (for i=1, . . .n) to be 'favourable' in terms of the generalised or standardised variances described above. A discussion of what constitutes a 'favourable' choice is given below (see §C and §2.3.4).

A Covariates

The term σ^2 is an estimate of the variation in the data which has not been explained by the model. This is the failure of the model terms to explain the two types of variation discussed above. The second type of error, that due to the natural variability in physical processes, can not be explained by any measurements and will always contribute to σ^2. Recall that the first type of error which contributes to σ^2 is the external variation due to factors

which could be measured but have not yet been included in the model. By identifying these factors and including them in the model σ^2 can be reduced. These factors, or experimental conditions, are known as covariates. Covariates are usually of no real interest to the experimenter, but if included in the model, they will make the estimated values more accurate and improve the quality of any follow up analysis.

B Blocking

Another approach for reducing this type of error is to structure the data using *blocking*. This means that groups or blocks of trials are performed in because similar conditions as possible. All trials in a block are then subject to many of the same errors. Comparisons between trials within blocks are thus more accurate. A *factor variable* can be introduced to the model as a block identifier to describe the variability between blocks. This often results in more accurate estimates of the features of interest.

C The information matrix

The generalised variance and standardised variance are wholly dependent on the choice of experiments (or trials) which are performed, i.e. the design matrix, and the chosen model. Because there is a limit to the amount of external error which can be removed using blocking and covariates, the bulk of the work in experimental design involves selecting the design matrix, given that σ^2 is as small as possible. In general we try to identify a design which either minimises some summary of the standardised variance over the design region, the generalised variance or some other summary of the variance of the parameter estimates. Four appropriate *criteria of optimality* are discussed in the next section. Usually a choice is made between them. This choice is often influenced by the specific aim of the experiment.

2.3.4 Criteria of Optimality

Four commonly used *criteria of optimality* are now outlined:

a) The *A optimal design* minimises $\text{tr}(X'X)^{-1}$. This trace is the sum of the scaled sampling variances for the estimates of the parameters in the model for E(Y).

b) The *D optimal design* minimises the generalised variance $|(X'X)^{-1}|$.

An equivalent approach to minimising $|(X'X)^{-1}|$ is to maximise $|X'X|$. This alternative approach is used in the thesis. The quoted D-value elsewhere may be any of the following; $|X'X|$, $|X'X|^{1/p}$, $|n^{-1}X'X|$, $|n^{-1}X'X|^{1/p}$, $|(X'X)^{-1}| = |X'X|^{-1}$, $|X'X|^{-1/p}$, $|n^{-1}X'X|^{-1}$, $|n^{-1}X'X|^{-1/p}$ (the

latter 4 are minimised), or the logarithm of any of these. Although different packages use different choices from the above possibilities, the resulting optimisation procedures are similar (see §2.3.7).

c) The *G optimal design* minimises $\max\{\underline{x}_i{}'(X'X)^{-1}\underline{x}_i, \text{ for } x \in S\}$, where $\underline{x}_i$ is the design augmented point for $\underline{x}_i$, and S is the feasible region. That is the G optimal design minimises the maximum prediction variance over the feasible region. In some of the examples included later (see §3.5.4), we use a set of g points $\{x_1, \ldots, x_g\}$ instead of S.

d) The *V optimal design* minimises the average of $(x'(X'X)^{-1}x)$ over S, i.e. the average scaled prediction variance, over S. Minimising this criterion can be computer intensive because estimation of the V value may require Monte-Carlo integral estimation (see §3.3.1). It is not always an available option. This criterion is referred to as the *I-criterion* in gosset (see §2.3.7D).

In this thesis the efficiency of design A with respect to design B and a given criterion, is defined to be the criterion value for design A divided by the criterion value for design B. When B is not specified, the efficiency is with respect to the optimum design for the given problem. Elsewhere efficiency of a design has been defined as the criterion value for the design in question divided by the criterion value of the optimal design, all raised to the power 1/p (see e.g. Atkinson and Donev, 1992, §11.1).

The design of experiments is complicated by the fact that each criterion may lead to a different optimal design. However an optimal design for one criterion will often be nearly optimal for others.

In section 2.3.6 we outline some further aspects of well designed experiments. First, we look at some examples of optimal designs for mixture experiments.

2.3.5 Some Examples of Designs

The points in a design can be specified succinctly using expressions of the form, $x \leftrightarrow (1,1,0\ldots,0)$, as introduced in Chan (1995). This expression represents all points whose co-ordinates can be given by any vector containing a permutation of the elements in the vector $(1,1,0\ldots,0)$, scaled to sum to 1.

An important design which is used frequently in the thesis is the (q, m) simplex lattice design. The (q, m) simplex lattice design uses all feasible combinations of

$x_i = 0, ^1/_m, ^2/_m, .. ^{(m-1)}/_m, 1$ for i=1, ..., q. This includes all feasible combinations of the integer multiples of $^1/_m$. In the notation (q, m), m denotes how regularly points occur. If m=1 then only pure blends are used. Recall that pure blends are where the mixture is made entirely of one ingredient (see §2.3.1A). The m=2 case uses all feasible combinations of $x_i = 0, ^1/_2, 1$ for i=1, ..., q. The m=3 case uses all feasible combinations of $x_i = 0, ^1/_3, ^2/_3, 1$ for i =1, ..., q.). The (3, 2) and (3, 3) simplex lattice designs are given in Figure 2.3.

Figure 2.3 Simplex lattice designs
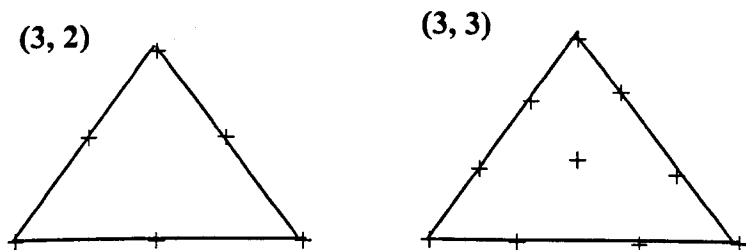
For some models and criteria, the optimal design on the full simplex is known. For the Scheffé polynomial models of degree 1 and 2 (see §2.1.1) the (q, 1) and (q, 2) simplex lattice designs are D optimal (Kiefer, 1961). For the special cubic model the (q, 3) simplex centroid design is D optimal (Uranisi, 1964). For the full cubic model the optimal design is the (q, 3) simplex centroid design, but with the midpoints of edges replaced by the q(q-1) edge points which are $(5-\sqrt{5})/10$ from each vertex (Mikaeili, 1993). A summary of known optimality results is available in Chan (1995).

The (q, m) simplex centroid design uses a similar idea to the (q, m) simplex lattice design but includes all feasible combinations of $x_i = 1, ^1/_2, .. ^1/_{m-1}, ^1/_m, 0$ for i=1, ..., q. In the notation of Chan (1999), the (q, m) simplex centroid design includes the following points: x $\leftrightarrow$ (1, 0, . . ., 0), which consists of $^qC_1$ points; x $\leftrightarrow$ (1, 1, 0, . . ., 0), which consists of $^qC_2$ points, ... , x $\leftrightarrow$ (1, . . , 1, 0, . . , 0) which consists of all ^qC_m points which have m ones and q–m zeros.

When the number of design points is equal to the number of terms in the chosen model, we say that the design is *saturated*. Because the number of design points in the (q, 1) and (q, 2) simplex lattice designs is equal to the number of terms in the Scheffé polynomial models of degree 1 and 2 respectively, these two designs have no degrees of freedom for

error. When a design leaves no degrees of freedom for error we say that it is saturated. The same is also true for the (q, 3) simplex centroid design for the special cubic model. An unfortunate consequence of these optimal designs being saturated, is that the PRESS approach to assessing the predictive power of models (see §2.2.6) can not be used.

In the majority of cases, theory can not be used to obtain an optimal design and are used to obtain near optimal designs. Algorithmic methods for generating experimental designs, often involve the selection of n points from a set of C possible points. The set of C points is known as the *candidate set.*

The simplex lattice designs are examples of *discrete* or *exact designs*, in which a finite number of design points are specified. A *continuous design* is one where *distinct support points* are specified and the design assigns each support point a *weight*. The weights specify the proportion of trials, which would ideally be placed at each support point.

The *equivalence theorem* states that the G optimal continuous design is also D optimal (see e.g. Atkinson & Donev, 1992 §9.2).

2.3.6 Further Aspects of Experimental Design
A Randomisation

If the trials are not performed in a random order there is likely to be some correlation between the error terms. This correlation will need to be accounted for in the model (see, e.g., §2.1.2).

Imagine that the trials were performed in some non-random order, in increasing order of the first variable for example. Given this ordering, any pattern, in an unidentified experimental condition which interacts with the first variable, might cause a pattern in the errors. In this way errors localised in time might be correlated with each other, and all errors would be correlated with the first variable. Additionally the interaction with the unidentified variable might be wrongly attributed to the first variable. Similar effects would probably occur for any non-random way of ordering the trials.

There is a variety of ways of putting trials into random order. The most common is to use computer generated random numbers in conjunction with a randomisation scheme. A randomisation scheme is a pre-specified way of translating a number into the name of one of the trials.

B Replication

All experiments are subject to a variety of sources of error. One way to estimate how much error an experiment is subject to, is to repeat some of the trials and compare the results. If a design includes a certain trial r times, then the design is said to include r *replicates* of that point.

A distinction must be made between *support points* and *design points*. The s support points are the set of all distinct points which are included in the design. The design points are the set of all points in the design, including support points and their replicates.

When replicates are present, the information matrix can be re-written in terms of a matrix of model augmented support points (denoted as Z below) and a diagonal matrix containing replicate numbers for each support point. This is shown below.

Information matrix $= X'X$.

$$\text{let } X = \begin{bmatrix} 1_{n_1} \underline{x}_1' \\ 1_{n_2} \underline{x}_2' \\ \vdots \\ 1_{n_s} \underline{x}_s' \end{bmatrix}, \qquad Z = \begin{bmatrix} \underline{z}_1' \\ \underline{z}_2' \\ \vdots \\ \underline{z}_s' \end{bmatrix}$$

where s denotes the number of support points, $\underline{z}_i' = \underline{x}_i'$ and n_i denotes the number replicates of suport point i.

$$[X'X]_{ij} = \sum_{m=1}^{s} \sum_{n=1}^{n_m} x_{mni} x_{mnj} \,,$$

where x_{mni} is the i'th term in the model augmented point for the n'th replicate of the m'th support point.

$$\therefore [X'X]_{ij} = \sum_{m=1}^{s} n_m z_{mi} z_{mj} \,,$$

as for all n_m replicates of support point m, the i'th term in its model augmented point is z_{mi}

Now

$$[Z'W^R Z]_{ij} = \sum_{m=1}^{s} n_m z_{mi} z_{mj} \quad \text{where } W^R = \text{diag}\,(n_1, n_2, \ldots, n_m),$$

i.e. a diagonal matrix of the replicate numbers

27

$$\therefore X'X = Z'W^R Z. \tag{2.25}$$

Note also that:

$$|X'X| = (\Pi n_i).|Z'Z|, \text{ if } s = p. \tag{2.26}$$

By dividing W by n, it can be seen that the determinant of the information matrix is directly proportional to $|Z'WZ|$, where W is a diagonal matrix of the weights.

In a similar way to equation 2.25, if Z specifies the model augmented support points of a continuous design Ψ, with weights given by the diagonal elements of W, the information matrix of the continuous design, i.e. $M(\Psi)$, is given by equation 2.27 (Atkinson and Donev, 1992, 9.1).

$$M(\Psi) = Z'WZ. \tag{2.27}$$

C Designs for split plot experiments

Split plot experiments occur when groups of experiments are performed in batches and each batch is subject to similar experimental conditions. This type of experiment was historically motivated by agricultural experiments where each field, or plot, was split into a variety of different sub-plots.

With split plot experiments each treatment is usually performed in more than one of the batches. In this way we can investigate the effect of each treatment under a range of experimental conditions. If possible, every treatment is performed at least once in each batch. Results from two treatments performed in the same batch, can be compared more accurately than if they had been performed in different batches, because the two treatments have been subject to similar experimental conditions (*covariates*). The experimental design for each batch is usually generated in conjunction with the experimental design for each of the other batches.

One example of a split plot experiment, which is relevant to our experiment, is where we have both qualitative and quantitative variables. If there is a total of *l* levels when the

individual levels of each qualitative variable are combined, we essentially have a split plot experiment with *l* main plots or batches. A *product design* is one where the same design over the quantitative variables is used at each combined level of the qualitative variables.

Kurotschka (1981, see, e.g., Atkinson & Donev, 1992, §13.2) considers optimal designs for experiments with both qualitative variables and quantitative variables. The following three situations are considered:

a) The chosen model includes all possible interactions between the qualitative and quantitative variables.

b) The chosen model includes no interactions between the qualitative and quantitative variables.

c) The chosen model includes some of the possible interactions between the qualitative and quantitative variables.

Kurotschka (1981) shows that for cases a and b, if the design region for the quantitative variables is the same for all l levels of the (combined) qualitative variables, and n is a multiple of l, the optimal continuous design will be a product design generated from the optimal continuous design for the quantitative design region. That is if the continuous optimal design for the quantitative variables is repeated at each level of the (combined) qualitative variables. One disadvantage in case b, is that the number of trials required will be large compared to the number of parameters in the model.

2.3.7 Relevant Computer Packages

Two packages and two algorithms which can be used for the design and analysis of mixture designs are looked at in this section. DX4 (Stat Ease), and minitab (Minitab Inc), are relatively simple to use packages which operate in the MS DOS and MS windows environments respectively. They can be used to deal with both the design and analysis of mixture experiments. KL (Atkinson and Donev 1992, §15), and gosset (Hardin and Sloane 1993), are algorithms for producing *efficient* designs. The term efficient design is used to describe a design which is near optimal. KL and gosset require more detailed inputs than DX4 or minitab but can handle constrained problems.

A Minitab (version 11)

Design facilities

The facilities in Minitab (Minitab Inc), for the design of mixture experiments, are extremely limited. The package has no facilities for choosing between different design types and can only generate the simplex lattice and simplex centroid designs (see §2.3.5). Minitab's facilities for these two designs types offer the following options.

a) Lower bounds can be imposed on the components and both types of design can still be generated. This is because regions defined by lower bounds only can be transformed to the regular simplex (see §2.3.1B). Designs with both upper and lower bounds can be handled if they can be U transformed to a regular simplex (see §2.3.1B). The majority of constrained regions can not be transformed to the regular simplex and can therefore not be designed over using Mintab.

b) The design points can be expressed in terms of proportions, original units or pseudocomponents (see §2.3.1B).

c) The two standard design types mentioned above can have *axial check blends* added. Axial check blends are midway between the centroid and each of the q vertices. If axial points are added, all q must be included, there is no option for selecting sub-sets of them

Analysis

Minitab's facilities for analysing results are again limited, but are superior to the mixture design facilities. This is because once the design matrix is known, analysis can be carried out using standard GLM techniques. Minitab has an extensive range of options for this type of analysis. However methods which are specific to mixture problems, such as trace plots (see §2.4), are not available.

B The KL algorithm

Overview

KL (Atkinson and Donev 1992, §15) is an algorithm which takes a set of C user specified *candidate points* and selects the best n points according to some optimality criterion. The candidate points usually consist of points which have previously been found frequently in good designs (see §3.5.3). A FORTRAN implementation of the algorithm for the D criterion is given in Atkinson and Donev (1992, §15). Its main problem is that it can

not generate its own candidate set for constrained design regions and needs to be used in conjunction with other algorithms such as CONSIM (Snee, 1979).

Basis of algorithm

A series of n point designs which are good according to a specified criterion is generated. The best of these designs is then chosen. Each design is generated using the following three stage process:

a) Addition of randomly selected points to the (possibly empty) set of previously analysed points

b) Sequential addition of points to make an n-trial starting design

c) Improvement of design by exchanging points.

In more detail these stages are:

a) In some cases experimental runs at q_1 points have already been analysed and must be included in the design. The starting set is then generated by the random selection of q_2 points from the candidate set, where q_2 itself is a randomly generated integer between 0 and int[p/2], where p is the number of parameters in the model. The q_2 points are then added to the initial q_1 points.

b) The n-trial starting design is then chosen by sequential addition of the $n-(q_1+q_2)$ points which give the biggest increase in criterion value.

c) The *leverage* and *prediction variance* of all design and candidate points are calculated and each set of points is ranked in increasing and decreasing order respectively. The leverage of a design point is given by the same equation as the prediction variance, i.e. equation 2.24, but with x being a point which is already included in the design. The design points with the k lowest leverages and the candidate points with the *l* highest prediction variances are then considered for exchange. These points are chosen because the update formulae (see §2.3.8C), show that the points with the biggest and smallest prediction variance and leverage respectively, give the biggest increase and smallest decrease in D criterion value when added to or removed from the design. The algorithm does not restrict attention to the single points giving the biggest and smallest change, because the best combined improvement will not necessarily be caused by the points giving the best individual improvement. This leads to k×*l* possible exchanges for which the change in criterion value is calculated using the updating formulae (see §2.3.8C). The exchange which gives the biggest improvement is

then carried out and the process repeated until there are no more exchanges which will improve the criterion value. The algorithm is then repeated t times.

The choice of exchange is limited to these k×l possibilities because it has been shown that they are likely to provide a bigger improvement than exchanges between other points. Unfortunately this restriction sometimes means that the best points for exchange are not considered. The best choice of k, l and t are problem dependent. However, as stated in Atkinson & Donev (1992), it was found that with a small amount of experimentation a rapid convergence to an optimal design can usually be achieved.

C Design Expert 4 (DX4)

Overview

DX4 (Stat Ease Inc) is a specialist package which deals with most types of experimental design and includes a range of features for the design and analysis of mixture experiments. It is extremely user-friendly and operates using dialogue boxes. DX5 has recently been made available, this is an updated version of DX4 which operates in the MS Windows environment. DX5 offers the same facilities as DX4 but some of the problems encountered with DX4 (see §3.2.2) are no longer encountered with DX5. The maximum design size has increased from 200 to 8 000, but the maximum number of candidate points has decreased from 8 000 to 3 001. It is suggested here that the maximum number of candidate points should have been increased, or at least left the same. This is suggested because the majority of experiments are likely to have less than 200 trials, and when q is large DX4 can often not store its full candidate set (Platts *et al* 1996).

Design facilities

DX4 offers a wide variety of design types including all of the ones which are available in Minitab. The simplex lattice and simplex centroid designs are available and D criterion values can easily be calculated so that the optimal choice between them can be made..

In addition to these designs, D optimal and distance-based designs can also be generated. Distance-based designs choose those n points which are most evenly spread over the simplex, according to a criterion based on the Euclidean distance between points. For both types of design, DX4 generates a set of candidate points using the CONSIM algorithm (Snee 1979) and searches through them using repeated applications of the Wheeler exchange algorithm (see §3.1.2). In the Wheeler algorithm f is repeatedly increased from 1 to 5, until no

further beneficial exchanges can be made. If at any point an improvement is found the process is restarted with f=1.

DX4 also has the following range of options:

a) The form of the displayed compositional values is easily switched between proportions, the variables in their original units and pseudocomponents. Minitab has to repeat its analysis separately for each of the three methods for specifying the points.

b) User specified designs can be entered into the package so that the highly automated analysis features can be used. A DX4 generated design can also be modified by the user.

c) Before the design is generated DX4 checks whether the user has requested enough design points to estimate the parameters. If a design has been modified or completely entered by a user and no response values are available yet, dummy values can be randomly generated or calculated from an existing model. This enables analysis to be carried out in order to check whether the specified design will allow model parameters to be estimated. Whether this can be estimated can easily be confirmed by checking whether the extended design matrix is of full rank. This is probably quicker than using dummy response values. However because the package is aimed at non-statisticians as well as experienced users this facility is useful.

d) The order of the design points can be randomised (see §2.3.6A).

e) A data entry form can be generated and printed. This form contains all design points in original units and in randomised run order. Spaces are left for response values to be entered as the experiments are conducted.

The above comments show that while not the ideal tool for an expert, DX4 is useful for a small company wishing to design their experimentation without the help of a statistician.

Analysis

For the chosen model, DX4 provides automated production of the following summary diagrams and values:

a) Taking any two components at a time, DX4 can produce a three dimensional response surface plot, which can be rotated to any viewing angle. A response surface plot shows the value of the response variable, for all feasible combinations of the three components chosen. With mixture problems, the other components which were not selected are kept constant at either their centroid values or alternative values selected by the user.

b) Trace plots can be produced (see §2.4). In problems with $q > 3$ these are useful for choosing three components which most affect the response, so that these 3 components can be used in plots of type (a) above.

Disadvantages of the package

Overall DX4 is a comprehensive package for most aspects of the design and analysis of mixture experiments. However it only offers a limited number of optimality criteria. DX4 also has problems storing its full set of candidate points for larger problems, and has no options for reducing the candidate set (Martin *et al.*, 1999). An example of a problem in which DX4 was unable to store its full candidate set is given in Platts (1996).

D Gosset

Overview

Gosset (Hardin and Sloane, 1993) is implemented using the computer language C and can be used to find the optimal design for a wide variety of variable types and a wide range of criteria. Gosset uses an algorithm which is based on a different approach from the candidate point algorithms discussed above. The gosset algorithm selects a random starting design and iteratively moves each point in the direction which gives the biggest improvement in criterion value. A similar approach has been used in a variety of other algorithms, including the adjustment algorithm of Atkinson & Donev (1992) and the perturbation algorithm of Martin *et al* (1999). In the Martin *et al*. algorithm, efficiency was increased by perturbating the points in starting design which was optimal over a set of candidate points and had been obtained using an exchange algorithm. The advantages of gosset are that it does not require a starting design, or even candidate points, and that it can be directed to work on a vast range of problems by simple user inputs.

One important improvement on the 3 packages discussed above, is that gosset will generate V-optimal designs. For mixture designs this involves Monte-Carlo estimation of an integral over the feasible region. This estimate is known as the moments matrix (M) and is described in more detail in Chapter 3 (see §3.3.1). The estimate of M can easily be exported to other packages so that algorithms for the V-criterion can be implemented in them.

The main drawback of gosset is that compared to the three methods discussed above it takes a long time to run, even on a large Unix machine.

2.3.8 Implementing KL for the V Criterion

The KL algorithm (Atkinson and Donev, 1992, §15.6, appendix A) was implemented in Matlab (Mathworks 1992) for the V and D criteria. The code for this algorithm, one version for each criterion, is given in electronic form on the disk which is included at the back of thesis. The names of the two files are klv.m and kld.m respectively. The FORTRAN code available in Atkinson & Donev (1992) was not modified for use with the V criterion, because writing matlab code from scratch proved to be easier. The FORTRAN code for the D criterion was also not used, because in chapter 3 a variety of modifications to KL are implemented. These modifications were again easier to implement on familiar code.

The option of V optimisation algorithm can only be used if the user can provide the moments matrix (M) (see §3.3.1), by some other means such as gosset (see §2.3.7D). The gosset M is based on the q-1 component model (see §2.3.2), and a transformation of the variables onto the scale (-1, 1), the matlab implementation of the KL algorithm for the V criterion also uses this approach since this proved to be far easier than converting the gosset M to correspond with the original variables. The aim of the KL algorithm is to generate an n point design from C candidate points. A flow diagram for the KL routine is given in Figure 2.4.

The first stage of each KL iteration is the sequential build up from a rank-deficient random start to an n point design. The choice of points should be performed in accordance with the relevant criterion. Because the random starting design is rank-deficient, the information matrix (see §2.2.2) can not be inverted, and the criterion value can not be calculated. To solve this problem Atkinson & Donev (1992) suggest modifying the information matrix in their D criterion version, to a form which can be inverted, by adding εI_p to it, where ε is a small number between 10^{-6} and 10^{-4}. Because all components of mixture designs are less than or equal to 1, and the V value is typically several orders of magnitude smaller than the D value, $\varepsilon = 10^{-15}$ was used..

Figure 2.4 Flow diagram for the KL exchange algorithm
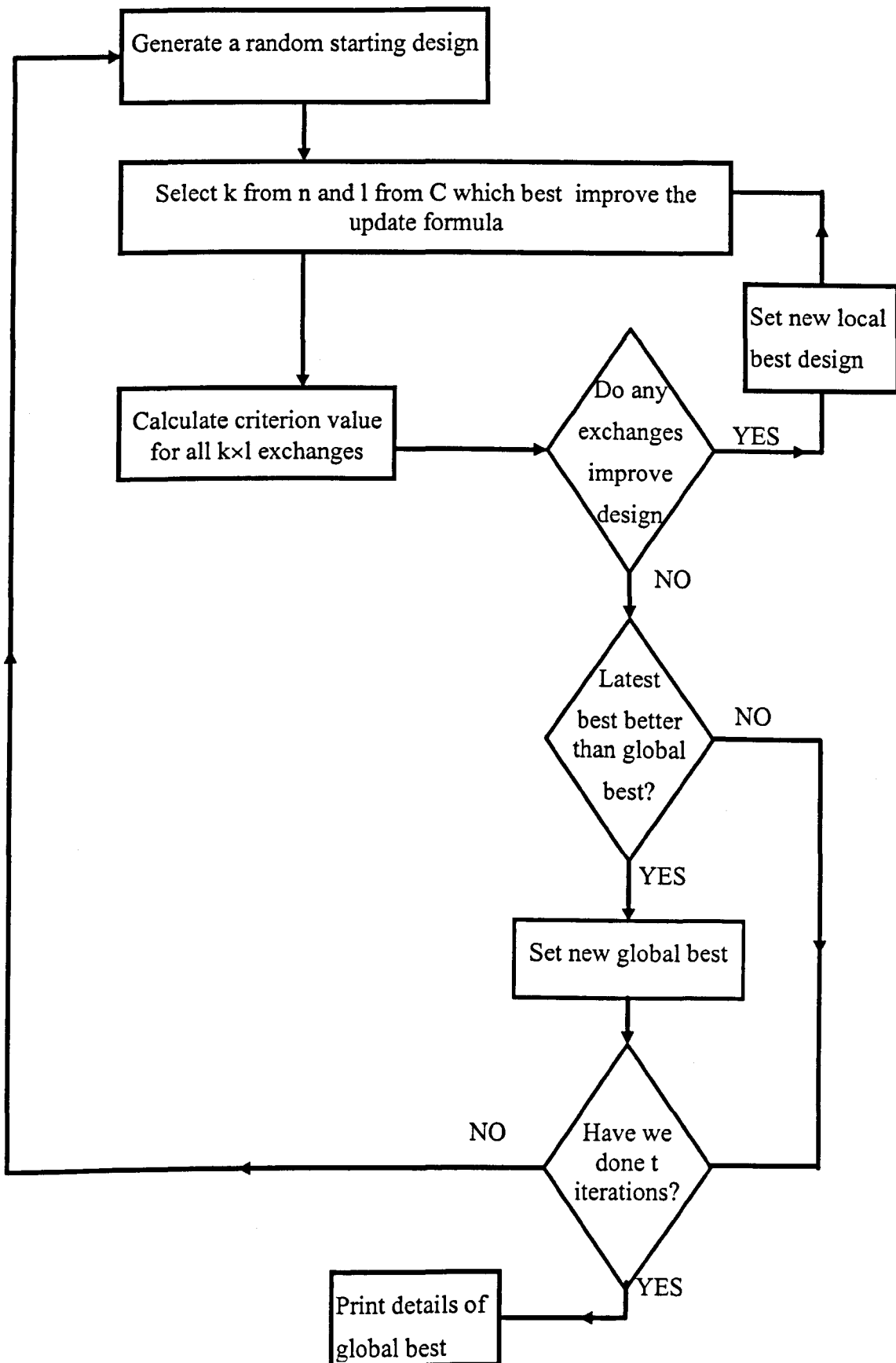
The rationale behind the method outlined in the flow diagram is that the best 1 candidates and the worst k design points are those most likely to improve the design when they are considered for exchange (see §2.3.7B). The algorithm is then repeated t times. When KL is used in the remainder of the thesis, the intensity of the search used will be specified using the notation KL(k, l, t).

C Programming aspects

To increase the efficiency of the algorithm the below updating formulae were used. These formulae update the information matrix, its determinant and its inverse, for a design specified by the design matrix X, when points have been added to, removed from or exchanged in, the given design. Let S denote the original X'X matrix, S_{new} the new X'X matrix, let $\underline{x}_k$ be a point to be removed from the design which is specified by X, and let $\underline{x}_l$ be a point to be added to the design specified by X. For ease of notation let the model augmented points (see §2.1.1) for the points to be removed be denoted by x_k and x_l respectively. Then, assuming all matrices are of full rank,

addition: $|S_{new}| = |S + x_l x_l'| = |S| \times (1 + x_l 'S^{-1} x_l).$ $\qquad$ (2.28)

removal: $|S_{new}| = |S - x_k x_k'| = |S| \times (1 - x_k 'S^{-1} x_k).$ $\qquad$ (2.29)

exchange: $|S_{new}| = |S - x_k x_k ' + x_l x_l '|$

$\qquad = |S| \times \{(1 - x_k 'S^{-1} x_k)(1 + x_l 'S^{-1} x_l) + (x_k 'S^{-1} x_l)^2\}.$ $\qquad$ (2.30)

The following formulae, give the inverse of the new information matrix S_{new}

addition: $(S_{new})^{-1} = (S + x_l x_l ')^{-1} = S^{-1} - (1 + x_l' S^{-1} x_l)^{-1} S^{-1} x_l x_l' S^{-1}$ $\qquad$ (2.31)

removal: $(S_{new})^{-1} = (S - x_k x_k ')^{-1} = S^{-1} + (1 - x_k' S^{-1} x_k)^{-1} S^{-1} x_k x_k' S^{-1}.$ $\qquad$ (2.32)

exchange:

$$(S_{new})^{-1} = (S - x_k^a\, x_k^{a\prime} + x_l\, x_l{}')^{-1} = S^{-1} + (1 - x_k{}' S^{-1} x_k)^{-1} S^{-1} x_k x_k{}' S^{-1}$$

$$- \{1 + x_l{}' S^{-1} x_l + (x_k{}' S^{-1} x_l)^2/(1 - x_k{}' S^{-1} x_k)\}^{-1}$$

$$\times \{ S^{-1} + (1 - x_k{}' S^{-1} x_k)^{-1} S^{-1} x_k x_k{}' S^{-1} \}$$

$$\times x_l x_l{}' \{ S^{-1} + (1 - x_k{}' S^{-1} x_k)^{-1} S^{-1} x_k x_k{}' S^{-1} \}. \qquad (2.33)$$

When the KL algorithm was implemented in matlab for the D criterion, with a small amount of experimentation with the values of k, *l* and t, the algorithm rapidly converged towards a solution and no problems were encountered. When the V criterion was used the following two problems occurred.

The first problem occurred when intermediate improvements included a replicate. With subsequent updates, considering further replication of this point produced a rank-deficient design matrix and a negative V value. Because the aim is to minimise the V criterion, this option would be wrongly selected and future updates would continue to get worse. Because the aim is to maximise the D criterion in its |X'X| form, negative determinant values do not cause the same problem. This problem was corrected by the replacement of negative values with unfavourably high ones. Unfortunately this slowed the algorithm down considerably, but no alternative solution could be found. The addition of εI_p to the information matrix for temporary updates was discounted because despite the fact that designs with further replication of the point never led to the best design, the latter method took much longer. The point which caused the problem was usually the centroid. The problem would probably not occur if the chosen form of the D criterion required minimisation (see §2.3.4), as the centroid is weighted considerably lower for the D criterion (see §3.9.5).

2.4 Trace Plots - An Analytical Method Specific to Mixture Experiments

A consequence of the fixed sum constraint is that it is not easy to interpret individual coefficients in isolation. This is because increasing one component will automatically decrease at least one of the other components, and so any effect which is then observed may have been due to the automatic decrease of the other components, rather than the increase in

the component under investigation. In order to rectify this situation, *trace plots* can be produced.

A plot of the response trace for component j shows how the predicted response varies as the value of component j is varied from a reference blend, usually the centroid (see §2.3.1), along a straight line to the edge of the feasible region. It is common to draw the response trace for all q components on the same set of axis to produce a *trace plot*. Two types of trace plot in common use are the *Cox direction plot* (see, e.g. Cornell, 19990, §5.9) and the *Piepel direction plot* (see, e.g. Cornell, 1990, §5.9.2). These two types of trace plot differ in the direction which is taken.

a) Cox direction plots show the effect of moving each component towards the vertex which represents its pure blend (see §2.3.1), while the other components maintain the same relative proportions as at the reference blend, usually the centroid. For component 1 this involves moving along arrow C in Figure 2.5.

Figure 2.5 Direction of trace for Cox (C) and Piepel (P) direction trace plots
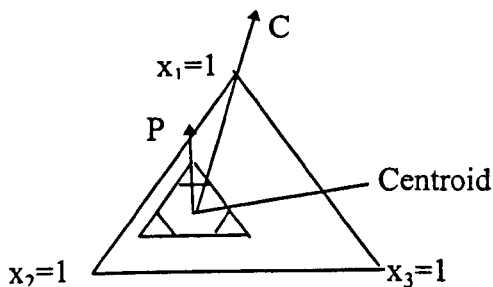
b) Piepel direction plots are a modification to the Cox direction plots which are often more useful for constrained examples. Piepel plots show the effect of moving toward the vertices of the smallest simplex which will bound the constrained region. This is illustrated by the arrow labelled P on Figure 2.5. The vertex for component i is calculated by setting the other q−1 components at their lower bounds and calculating component i to satisfy the fixed sum constraint.

An explicit formula for the value of x_j as component i moves in the Cox direction is now described. Consider starting from the reference blend, where component i takes the value s_i, and change x_i by an amount δ. The value of δ should be chosen small enough to produce a

smooth plot and should decrease as the degree of the model increases. The new value of x_i is therefore $s_i + \delta$. To satisfy the fixed sum constraint the changes in the other variables must sum to $-\delta$. To ensure that we are moving towards the pure blends, the relative proportions of the other components must remain the same as at the reference blend. To maintain these original proportions the amount of the $-\delta$ which component j must take is proportional to its relative proportion with respect to the other q-1 variables at the reference blend i.e. $s_j/(1-s_i)$. Therefore the new value of component j is

$$x_j* = \quad s_j - \delta.s_j /(1-s_i) \qquad = s_j.(1-s_i-\delta)/(1-s_i), \quad \forall j \ne i. \tag{2.34}$$

Both types of plot can be constructed using a simple computer routine, this routine takes each component in turn. For component i denote the reference blend and the appropriate vertex (see definitions above) by $\underline{s}_i$ and $\underline{u}$ respectively. Any position on a straight line between these two points can be represented by

$$\underline{x} = \underline{s} + \delta(\underline{u} - \underline{s}).$$

The trace plot is then constructed in the following way. The predicted response according to the chosen model is plotted against $\delta(u - s_i)$, the change in component i, for a range of δ which produce points on the line and within the constrained region. When using a computer to perform this process there is no need to calculate a range for δ explicitly. The value of δ can be varied by increments of $\pm\varepsilon$, for small ε, until the resulting point is found to occur outside the constrained region. When this occurs the appropriate point is discarded and no further points in the direction in question are considered. All q trace plots are then superimposed to form a single plot.

Figure 2.6 gives some trace plots for an example due to Cornell (1990). This example has q=3 with the following lower and upper bounds; (0.2, 0.6), (0.1, 0.6), (0.1, 0.5). For the data given in Cornell (1990), the following quadratic model was fitted

$$E(Y) = -21.2x_1 - 24.6x_2 + 20.6x_3 + 129.3x_1x_2 + 25.2\,x_1x_3 + 36.5\,x_2x_3. \tag{2.35}$$

The above algorithm was implemented with ε=0.01 using `matlab`. The code for this algorithm is given in electronic form on the attached disk as `trc.m`. Figure 2.6(a) and 2.6(b)

show the Cox and Piepel direction trace plots from the centroid (0.383, 0.334, 0.283). Figure 2.6(c) and 2.6(d) show similar plots but from the alternative reference blend (0.25, 0.4, 0.35)

Figure 2.6 Examples of trace plots. Component x_1 is denoted by the solid line, Component x_2 is denoted by the dashed line and component x_3 is denoted by the dotted line
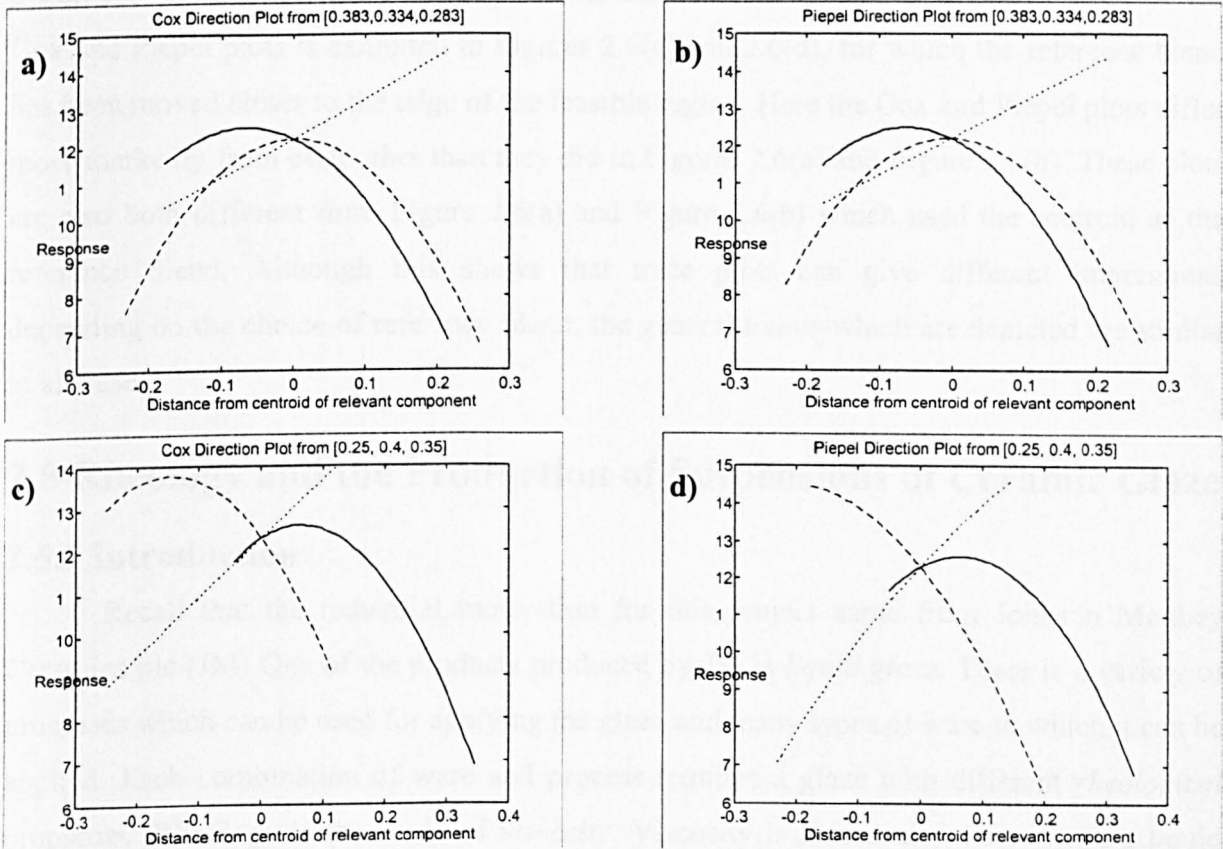
As an example of how trace plots can be interpreted, consider Figure 2.6(a). We can see that as x_3 increases from its minimum value the response increases linearly. As x_2 increases from its minimum value, up to the centroid, the response increases in a non-linear way. Then, as the value of x_2 increases between the centroid and its maximum value, the response decreases in a non-linear way. This decrease occurs at about the same rate the previous increase, and ends when the response is just lower than it was when x_2 was at a minimum. As x_3 increases from its minimum value up to the centroid, the response variable increases in a non-linear way from a comparatively high starting value. As it then increases between the centroid and its maximum value the response decreases in a similar non-linear way to that observed when x_2 was decreasing. When x_1 reaches its maximum value the response is slightly lower than it was when x_2 was at its maximum value.

41

From Figure 2.6(a) and 2.6(b) it can be seen that the Piepel plot is very similar to the Cox plot. This is because the constrained region is towards the centre of the regular simplex, causing the vertices of the regular simplex to be in a similar direction from the centroid as the vertices of the triangle containing the feasible region. As the constrained region, and subsequently its centroid, becomes less central with respect to the regular simplex, the difference between the two types of plot will increase. The increased difference between the Cox and Piepel plots is exhibited in Figures 2.6(c) and 2.6(d), for which the reference blend has been moved closer to the edge of the feasible region. Here the Cox and Piepel plots differ more markedly from each other than they did in Figures 2.6(a) and Figure 2.6(b). These plots are also both different from Figure 2.6(a) and Figure 2.6(b) which used the centroid as the reference blend. Although this shows that trace plots can give different impressions depending on the choice of reference blend, the general trends which are depicted are similar in all cases.

2.5 Rheology and the Production of Suspensions of Ceramic Glaze

2.5.1 Introduction

Recall that the industrial motivation for this project came from Johnson Matthey Ceramics plc (JM) One of the products produced by JM is *liquid glaze*. There is a variety of processes which can be used for applying the glaze and many types of ware to which it can be applied. Each combination of ware and process requires a glaze with different *rheological* properties. Rheology is the study of *viscosity*. Viscosity is a measure of how well a liquid resists motion when a force is applied to it. The viscosity of a glaze depends on a variety of variables which can be split into *process variables* and *mixture variables*. The process variables which are of interest in the project are all qualitative and are used to represent different production methods rather than the amount of each ingredient which is used. The amounts of each ingredient are specified by the mixture variables

The mixture variables in question are clay, water and crushed glass, the crushed glass is known as *frit*. It is widely accepted that increasing the proportion of clay will increase the viscosity, and increasing the proportion water will decrease the viscosity. The effect of frit is not known.

The first process variable is derived from a quantitative measurement of the particle size to which the three mixture variables have collectively been ground. As JM is only

interested in the approximate particle size, the complete range of possibilities is partitioned into three size ranges and the level of the process variable indicates which one of these ranges the glaze falls into. It is widely accepted that decreasing the particle size of a glaze will increase the viscosity. The second process variable is whether or not a *flocculant* has been added. Flocculant is the name given to any substance which is known to cause aggregation, in this case the clay particles, which will affect viscosity.

Some quantitative process variables also affect viscosity, for this project they were either kept constant (e.g. temperature, see §5.3.4) or accounted for in the viscosity model (e.g. the size of the force applied (see §2.5.8))

Some background to the production and application of glaze is given in this section. Knowledge about glaze is helped by an understanding of the basic properties of glass, therefore the basic properties of glass is the first area which is discussed in this section (see §2.5.2). The background given in sections 2.5.2 and 2.5.3 was obtained from Taylor and Bull (1984) and through personal communication with Dr Howard Winbow of JM.

2.5.2 Some Properties of Glass

Glass is a very complicated substance. Nearly any material may be made into a glass if it is possible to cool it down from the molten state quickly enough. Glass can be regarded as a new state of matter, in addition to solid, liquid and gas, rather than a particular substance. A material's state of matter can be identified by whether it can change its shape to fit into, or completely fill, a container. Another method of identification is to measure the rate at which its shape and volume are affected by temperature. As a glass-forming liquid is cooled its viscosity gradually increases. Just before it solidifies a point is reached where it has the energy, volume and atomic arrangement of a liquid, but the rate at which its energy and volume change with temperature is similar to that for a solid. When a substance displays these properties it has been permanently changed into a glass.

An x-ray diffraction pattern of a material in the glassy state shows its instantaneous molecular structure to be almost exactly the same as that of a liquid, despite the fact that the material appears to be solid to the human touch. This means that despite it appearing solid there is little long range ordering of its molecular structure.

2.5.3 Producing Glasses and Frits

Recall that a *frit* is a glass which has been crushed into small pieces. As mentioned above the production of glass requires a rapid reduction in temperature from the molten state. To perform this rapid reduction in temperature, soluble minerals and compounds are heated to 1400°C in a furnace and then poured into a bath of cold water known as a *'bosh'*. This process causes the material to shatter into small pieces of glass and is known as *quenching*. Alternative methods involve cooled steel rollers or blasts of air or water mists. The resulting glass is insoluble in water, but may be selectively leached, and can therefore be used in glaze. The frit can then be produced by grinding the quenched glass into smaller pieces.

The biggest problem in the production of glass frit is variation in the size and weight of the particles in the raw materials. If the particle sizes and weights vary substantially in any individual batch, a segregation of materials is likely to occur when they are mixed. This segregation is due to large particles with low specific gravity, moving around more slowly than the small particles with high specific gravity and also interparticle adhesion characteristics. To avoid this segregation raw materials with similar specific gravities are used and before they are mixed together the individual raw materials are processed so that all particles are of a similar size. The ideal conditions after mixing would be to have particles of each raw material adjacent to a particle of a different raw material. In reality we get apparently random mixing. This is sufficient in most cases and will help to produce a quenched product that is easy to grind.

2.5.4 Introduction to the Use and Application of Glazes

Ceramic items, which have been moulded from clay and left to dry, can be extremely weak and prone to degradation. Decay is caused by external sources such as everyday use and exposure to acids, alkalis and even water. To prevent degradation the majority of ceramic wares are protected. In most cases the first stage of the protection process is to bake or *biscuit fire* the ware in a special oven known as a kiln, at temperatures of around 1000°C. After an item has cooled down it is covered with *glaze*, a suspension of glass and clay, and *glost fired*, again in a kiln. During the glost firing, much of the water evaporates, and the glaze hardens to form a protective, and often decorative, coating. A glaze can therefore be considered as a layer of glass between 75μm and 500μm thick. It is widely believed that 'glazes' were discovered by accident due to the presence of the relevant substances in the air within the kiln or in certain types of clay. The ware is then safe to be used for a variety of different purposes.

Household crockery is perhaps the most common example of items produced in this way. With crockery, the main purpose of the glaze is to prevent food absorbing dangerous chemicals from the ware. Secondary reasons for glazing crockery are protection from scratching or corrosion and the provision of an attractive finish.

There is a variety of processes for applying glaze to a ware. Each application process requires glaze of a different viscosity. Recall that viscosity is a measure of how well a substance will resist flow when a force is applied to it. The basic ingredients of glaze are water, clay and frit (see §2.5.1). The practical aim of the thesis is to investigate how the proportions of the three basic ingredients of glaze affect its viscosity. Because it is the proportions which are of interest, this investigation defines a mixture experiment. There is a variety of processing conditions which are also believed to alter glaze viscosity. The processing conditions, or factors, which are of most interest, are the particle size to which the suspension has been milled, and whether a flocculant has been added. Storage time is also an important factor. It is believed that the viscosity of a glaze sample which has been taken from storage, will depend on how many of the glaze particles have settled on the bottom of the container during storage.

Viscosity is not always an absolute property, and for glazes, which are non-Newtonian, depends on the force or *shear stress* which is being applied. Viscosity is best illustrated on a graph displaying how viscosity varies with shear stress. To produce a viscosity graph, a force (or stress) is applied parallel (or shear) to a sample, in order to produce a set displacement. The size of the force which is required to produce this displacement is measured, and viscosity is proportional to the ratio of these two values (see §2.5.8). Because viscosity is best represented by a graph of values (rather than a single value) a further aim of the project is to choose suitable summary measures of these graphs.

2.5.5 Glaze Production and Modification

Recall that the three main ingredients of glaze are frit, clay and water (see §2.5.1). In addition to these main ingredients, comparatively small amounts of other chemicals are often added to modify glazes for specific uses. These additional chemicals can be added to the raw materials at the start of the process, or to the glaze slip towards the end of the process. An example of an additional chemical is the flocculant calcium chloride, which was used in the JM experiment. Calcium chloride was the only additional chemical which was used in the JM trials. The amount of calcium chloride which is usually added is approximately 0.05% of the

weight of solids in the glaze (see §5.2.6). This amount of flocculant is a very small proportion of the total mass, and was consequently treated as a process variable (i.e. a specified amount was either added or not added). The amounts of the three mixture ingredients were measured accurately by weight (g) to two decimal places using an accurate balance (Sartorius, maximum 2kg, tolerance ±0.01g, calibrated weekly).

The most common method for producing glaze involves wet grinding of the ingredients into a suspension of small particles. This suspension is known as a *glaze slip*. The process of wet grinding is described in separate section below (see §2.5.6). Methods for applying dry glaze powders do exist, but are rarely used. It is also possible to replace the water with other carrier media such as alcohol, but this too is rare. The purpose of the grinding procedure is to standardise the particle size of the solid ingredients and evenly distribute the mixture throughout the water. Once the slip particles have reached a specific size, if not before, the additional chemicals are added. The additional chemicals are then distributed evenly throughout the slip by shear mixing (see §5.2.6).

Glaze is normally sent to a JM customer in the slip form. In some cases the glaze is dried, so that the customer can add the required amount of water themselves. Ground frit can also be sold separately, so that the customer can produce a glaze to a variety of specifications.

2.5.6 Wet Grinding

The machine which is most commonly used for grinding the ingredients is a *tumbling mill*. The tumbling mill consists of a rotating ceramic cylinder filled with abrasive rocks, known as the *grinding* or *milling media*. As the mill rotates the media crash into the ingredients and gradually grind them. An alternative approach is to apply vibrations to the mill. Whichever method is used, the mill is lined with abrasion-resistant bricks to minimise the internal damage to the cylinder. The choice of material for these bricks depends on the type of frit being ground. In the majority of cases a form of silica rock known as silex is used.

The rate at which the glaze is ground depends on the following factors: lining type, frit type, ratio between the mass and volume of the materials, mill speed, particle size, and consistency and hardness of grinding media. The most important factor is the volume ratios between mill, media and ingredients. The grinding media and spaces caused by its irregular shape should fill about 55% by volume of the mill. The raw materials should then account for between 11% and 18% of the mill volume. The exact value within this range should increase

with the density of the grinding media. The amount of water to include in the mill also needs to be controlled. An acceptable range for the amount of water to use depends on the specific gravity of the raw materials.

A variety of shapes of grinding media can be used in milling. In wet milling spherical media are usually used. The rate at which the grinding occurs depends on the difference between the density of the grinding media and the density of the ingredients. The four most common choices of media are flint pebbles, porcelain, steatite and high density alumina.

The size of the grinding media is also an important factor. In most cases three different sizes of media are used. Once the biggest particles have ground the ingredients to a certain size they have a much diminished effect. When the big particles become less effective the second smallest size is needed to continue the grinding. After additional milling the second size of media become less effective and the smallest media complete the milling process. All three sizes of media are left in the mill jar throughout the process. Care is also needed when choosing which media to use with which mill lining. This is because certain combinations can increase unwanted contamination.

The final factor which needs consideration is the speed at which the mill rotates. If the mill rotates very slowly the main force present is gravity, consequently the contents remain at the bottom and little grinding occurs. At very high speeds centrifugal force pushes the contents towards the edges and again little grinding occurs. The speed at which this starts to occur is known as the *critical speed*. A balance between the centrifugal and gravitational forces is clearly required. In practice 60% of the critical speed is known to provide the best effect. The bigger the mill, the faster it will need to be rotated to get the same linear velocity around its edges.

The length of time for which the mill is rotated will affect the particle size distribution of the resulting glaze slip. The particle size distribution is usually summarised by measuring the percentage of particles which are below certain sizes. Because this is only a summary, two samples with the same percentage of particles less than a specified size can have different particle size distributions. When milling is performed, particle sizes within specified ranges are usually required rather than exact particle sizes. Because particle size can have a dramatic effect on the viscosity of a glaze slip, it is one of the processing conditions which are investigated in this thesis.

Because the internal volume of the mill increases through damage, the volume of material needs to be altered to retain the optimum effects. In a similar way, the size of the grinding media is also reduced. The range of media sizes needs to be checked approximately every 2 months (depending on usage) so that it can be restored to the optimum.

2.5.7 Glaze Application

Before any glaze is applied, it is sieved and then mixed, to remove any lumps. Lumps are removed because they may cause an uneven finish on the ware. This vigorous mixing is known as *shear mixing*. A glaze can be applied to a ceramic ware by a variety of different methods. The three most common application procedures are spraying, dipping and painting. A glaze will need a low viscosity under the high shear stress met in glaze application, to allow it to be evenly spread over the ware. The three methods of application use decreasing levels of shear stress and hence require a low viscosity at decreasing shear rates. Once the glaze has been applied it will be subject to the low shear stresses of gravity while it is left to harden. It is required that the glaze viscosity will be relatively high under the low shear stress of gravity, to prevent the glaze from running and causing an uneven finish.

The viscosity of a glaze can be modified until it is suitable for the application procedure in question. This modification can be done by the addition of dry glaze powder, water or other ingredients such as flocculant (see §2.5.1). Recall that the aim of the thesis is to produce a model for predicting glaze viscosity from a variety of mixture and process variables (see §2.5.1 and §5.2,1). This model is required so that glazes, with specific rheological properties for the different methods of application, can be produced with less need for trial and error.

It is known that the viscosity of a glaze can be increased by increasing the amount of clay and decreasing the amount of water. The effect of changing the amount of frit is less clear.

Recall that the viscosity will be affected by the size of the particles in the glaze slip. Because smaller particles tend to bond together more easily, decreasing the size of the solid particles is likely to increase the viscosity of a suspension.

An additional complication is that glaze is also known to be *thixotropic*. The viscosity of a thixotropic substance is non-permanently lowered by being exposed to large shear stresses. The level of shear stress which is needed to lower viscosity, depends on the type of

frit and clay used and the proportions of each. However, the larger shear stresses, which were used in the JM trials, were all high enough to apparently permanently lower the viscosity of the glazes investigated here.

A key property of glaze is that its viscosity, i.e. the coefficient which determines its ability to resist flow, depends on the size of the force trying to cause this flow (see §2.5.8). To ensure that a glaze has the potential for successful application to ceramic ware, it needs specific rheological properties (see §2.5.7). In summary these properties are a low viscosity under the high shear stresses involved in its application and a high viscosity under the low shear stresses as the glaze subsequently hardens.

Recall that the aim of the project was to produce a mathematical model for predicting viscosity, when small and large forces were applied to a glaze. Initially this aim was to be achieved by predicting the viscosity at shear rates of $150s^{-1}$ and $1050s^{-1}$ respectively. The trials discussed in Chapter 5 were designed to measure viscosity at these two shear rates. While these trials were being performed discussions with the new JM team, with responsibility for the project, led to a change of focus. It was now decided to investigate the relationship between glaze viscosity and shear stress, rather than the relationship between viscosity and shear rate. Additionally the scope of the project was increased since the low shear stresses of interest were considerably lower than those encountered at a shear rate of $150s^{-1}$. This led to the need for the additional trials, these trials are discussed in Chapter 6.

A better understanding of the theoretical shape of viscosity curves had also been obtained while the trials discussed in Chapters 4 and 5 were being performed. The aspects of rheology that were now of interest to JM are discussed in section 2.5.9. Before looking at these aspects of the curve we outline the procedure for measuring viscosity (see §2.5.10) and some standard rheology theory.

2.5.8 Basic Rheology Models

This section outlines four theoretical models for the relationship between shear rate and shear stress. The four models discussed below do not have specific names, but if one of them applies to a given material, the material is said to be; Newtonian, Bingham, Pseudoplastic or Pseudoplastic with a yield respectively. More complex models exist but are not considered here.

Before looking at the four relationships we define the following symbols:

Shear rate $\quad \dot{\gamma} \quad$ s^{-1}

Shear stress $\quad \tau \quad$ Pa

Viscosity $\quad \eta \quad$ Pa.s

A: Newtonian materials

In Newtonian materials the relationship between shear rate and stress is linear. This linear relationship is shown in Figure 2.8.

Figure 2.8 Rheology of Newtonian materials
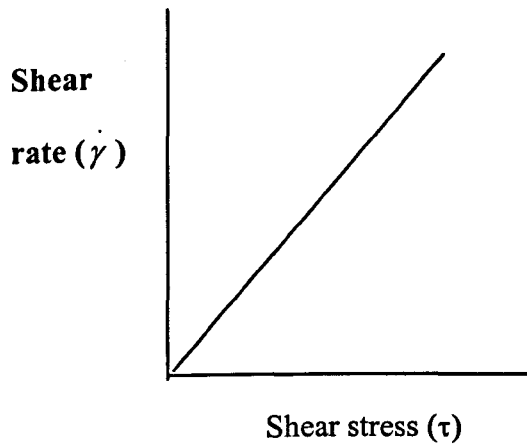
The nature of this linear relationship can be described by equation 2.36.

$$\tau = \eta \, \dot{\gamma} \tag{2.36}$$

To calculate the viscosity of a Newtonian material we simply calculate the gradient of the graph of shear rate against stress, and take the reciprocal of this gradient. To investigate rheology, the Haake viscometer (see §5.2.7) gradually increases the shear stress and measures the induced shear rate at various discrete time points. JM refer to this approach as a *shear stress ramp*. An alternative approach is to measure the shear stresses which are required to produce various (increasingly high) shear rates. JM refer to this approach as a *shear rate ramp*.

Irrespective of which variable is ramped, to calculate the *apparent viscosity*, we can then divide the measured shear stress by the measured shear rate. For Newtonian materials the apparent viscosity should be approximately the same at each point. The value of this constant is known as the *absolute viscosity*. To estimate the absolute viscosity we can calculate the mean of the measured apparent viscosities.

B: Bingham materials

Bingham materials exhibit the relationship shown in Figure 2.9. Up until a certain shear stress the shear rate is 0. The shear stress at which the shear rate first becomes non-zero is known as the *yield stress* or *yield point*. Yield stress is perhaps the most important measure of rheology. Once the yield point has been exceeded, the material displays Newtonian behaviour, i.e. the graph becomes linear increasing. Consequently Bingham materials are sometimes referred to as being Newtonian with a yield. The absolute viscosity of a Bingham material (η_{abs}) is the reciprocal gradient of the linear increasing section of the graph.

For Bingham materials, an applied shear stress will not cause any flow until the yield stress has been reached. For any given shear stress the viscometer has no way of knowing whether the yield point has been reached. Consequently the viscometer can only calculate apparent viscosities. Because the apparent viscosity is simply the instantaneous shear stress divided by the instantaneous shear rate, it is not equal to the absolute viscosity. Apparent and

Figure 2.9 Rheology of Bingham materials
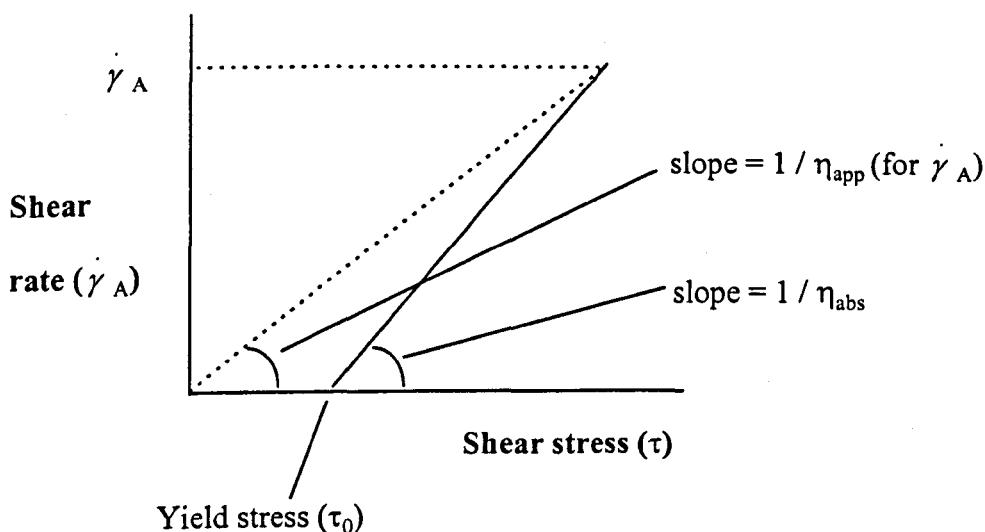
absolute viscosity are indicated on Figure 2.9 as η_{app} and η_{abs} respectively. As shear stress increases, the gradients of the two lines will converge, and apparent viscosity will tend towards absolute viscosity.

Many suspensions in water approximate to a Bingham liquid. With suspensions, as shear stress is increased towards the yield stress, its solid particles can clump together and force water towards the edges. This water can cause the viscometer bob to rotate before the yield point has been reached. This effect is known as *wall slip*. When wall slip occurs, the viscometer will measure the shear rate and calculate an apparent viscosity. This apparent viscosity is not the apparent viscosity of the glaze. The graph in Figure 2.10 illustrates the effect of wall slip on rheology measurement.

Figure 2.10 Bingham material with wall slip
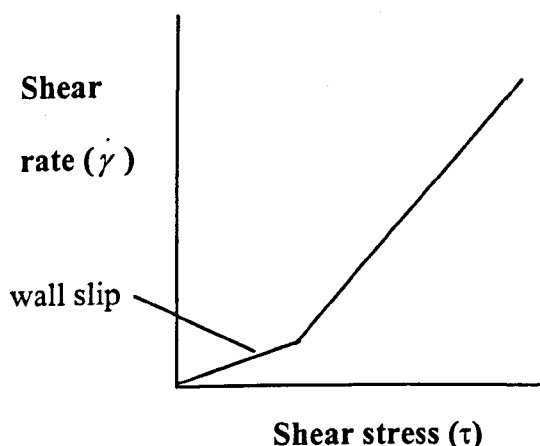
When wall slip occurs there are various ways to estimate yield stress. Some people estimate the yield stress as the shear stress at which the gradient changes in Figure 2.10. Other people extrapolate the Newtonian line back to the τ-axis and define the point where they intersect as the yield stress. It is suggested here that the former definition is intuitively more sensible.

The most informative summary measures of a suspensions rheology are the yield stress and the absolute viscosity.

C : Pseudoplastic material

With pseudoplastic materials the relationship between shear rate and shear stress is non-linear. Consequently the absolute viscosity at a given point is the reciprocal gradient of a tangent to the line. Figure 2.11 shows the relationship between shear rate and shear stress for pseudoplastic materials. As shear stress is increased the absolute viscosity decreases and the gradient increases. For pseudoplastic materials measurements of apparent viscosity have to be quoted together with the shear rate at which they were measured.

Figure 2.11 Rheology of Pseudoplastic material

Shear rate ($\dot{\gamma}$)

Shear stress (τ)
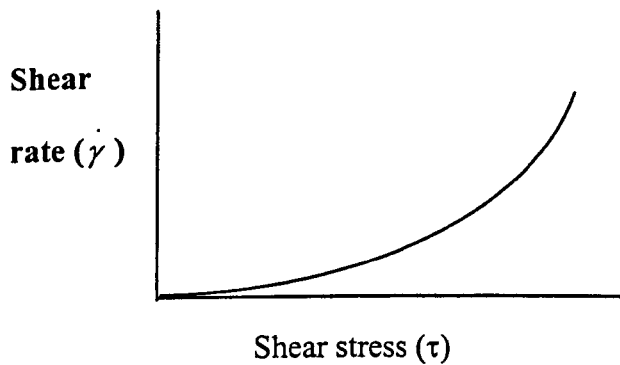
Pseudoplastic materials can also have a yield point, and consequently they can also experience wall slip. These two effects are illustrated in Figure 2.12.

Figure 2.12 Pseudoplastic material with yield stress and wall slip

Shear rate ($\dot{\gamma}$)
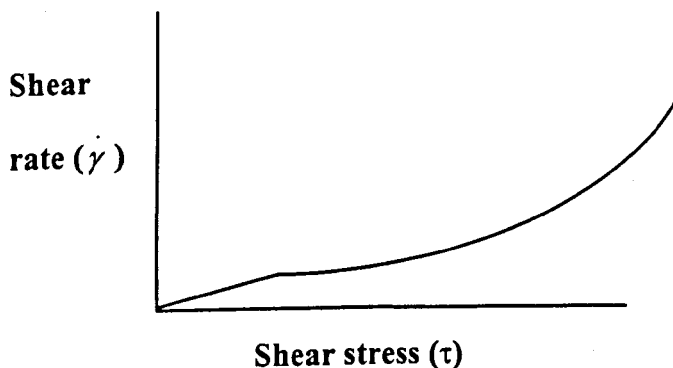
Shear stress (τ)

D Dilatant behaviour

In Bingham and pseudoplastic materials, when shear stress reaches very high levels, dilatant behaviour is sometimes observed. Dilatent behaviour means that the particles in the suspension can not move sufficiently fast relative to each other, as the shear rate increases, to slip past each other. Consequently they crowd together. This re-arrangement of the particles causes the viscosity to increase. Dilatent behaviour is illustrated in Figure 2.13.

Figure 2.13 Dilatant behaviour
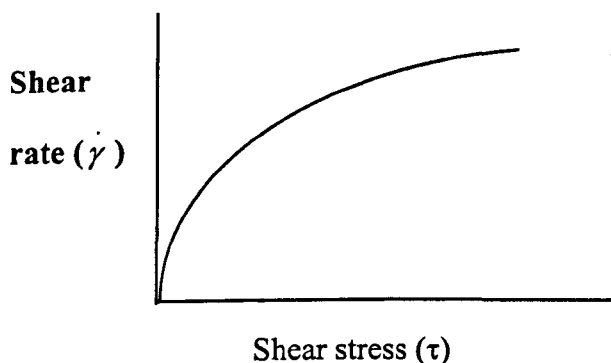
In extreme cases the viscosity becomes so high that it imparts a large resistive stress on the mechanism causing the shear stress. It is this effect which causes a spoon to fly across the room when cornflour is mixed vigorously with water.

2.5.9 The Rheological Properties of Interest to JM

Instead of the standard graph of shear rate against shear stress, JM requested an analysis based on the graph of apparent viscosity against shear stress. Aspects of the theoretical viscosity stress graph, which are of interest to JM, are discussed in section A. In section B these aspects of the viscosity stress graph are related to the rheology models discussed above (see §2.5.10).

A Glaze rheology as outlined by JM

JM are interested in the graph of apparent viscosity against shear stress. The theoretical form of this graph is given in Figure 2.14.

Figure 2.14 Theoretical viscosity curve used by JM
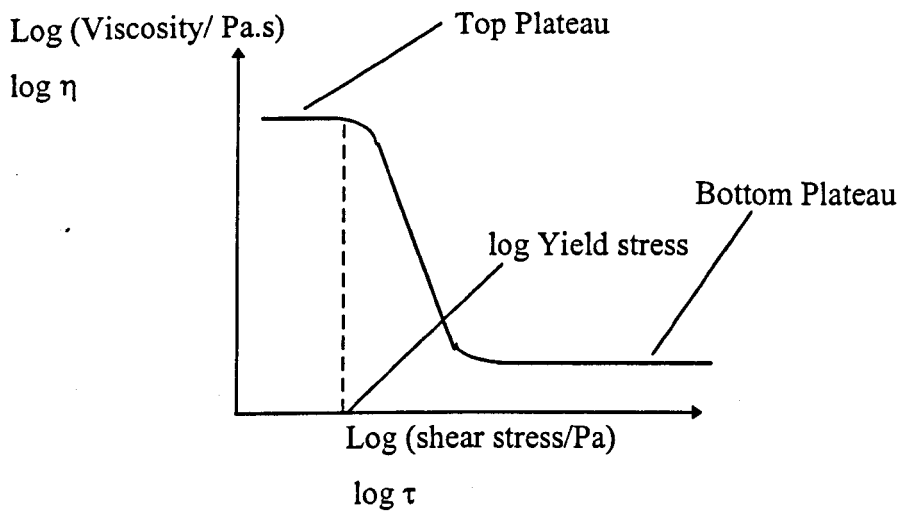
Viscosity in the low shear region has been observed to be several orders of magnitude greater than the viscosity in the high shear region. Within these regions the viscosity remains approximately constant and displays *Newtonian behaviour* (Herritage, 1994). Due to this behaviour, the relationship between viscosity and shear stress can be represented by two plateaux separated by a steep descent.

JM define the shear stress at which the top plateau ends as the yield stress. For any particular glaze, it is important that the yield point occurs between the high shear stress of the application method and the low shear stress experienced when the glaze is left to dry.

JM are interested in three main parameters: viscosity of the top plateau; the yield stress and the viscosity of bottom plateau. For glaze slip, the top plateau typically occurs between 1,000 Pa..s and 15,000 Pa.s, the yield point typically occurs between 1 Pa.s and 30 Pa.s and the bottom plateau typically occurs between 0.05 Pa.s and 10 Pa.s. JM specified that the only aspects of the curve which were of interest to them were the viscosities of the top and bottom plateau and the yield point. There was no interest in modelling shear stress as a continuous variable.

B Relating the rheology measures of interest to JM to the basic rheology models

Figure 2.10 shows that wall slip will cause the viscometer to record approximately constant values of apparent viscosity. It seems likely that the top plateau in Figure 2.13 is caused by wall slip. Although the level of wall slip will be related to the rheology of the glaze

in some way, it is a consequence of the viscosity measurement procedure and is unlikely to have any scientific significance .

It is likely that the yield stress defined by JM is the yield stress defined in the previous section (see §2.5.8).

As soon as the yield point has been reached the apparent viscosity begins to decrease. Because Figure 2.13 is a log-log graph, the short period over which the apparent viscosity decreases, actually represents a large increase in shear stress. Recall that when shear stress becomes reasonably high, apparent viscosity tends towards absolute viscosity. Therefore the rapid decrease in apparent viscosity, which is illustrated in Figure 2.13, is likely to be a consequence of the discrepancy between apparent viscosity and absolute viscosity. It is therefore likely that the height of the bottom plateau will be a good estimate of absolute viscosity.

2.5.10 Theory of Rheology Measurement

The physical aspects of the viscosity measurement are described using Figure 2.7. The large and small circles in the diagram represent the cup and sensor of the Haake viscometer respectively. The parallel plates are not part of the measurement system but are included to aid description.

In terms of Figure 2.15, one method of representing the measurement of viscosity involves moving the top plate along the block of sample at a speed of S ms^{-1}. This would cause all particles in the block of the sample to be displaced, with the displacement of particles decreasing with the distance, d (m), from the top plate. This reduction in displacement is indicated by the displaced sample being represented by the parallelogram to the right of the original block. Shear rate (γ (s^{-1})) is defined as the speed of the moving plate divided by the distance between the two parallel plates and given by equation 2.37.

$$\gamma = \frac{S}{d} \text{ (s}^{-1}\text{)} \qquad\qquad (2.37)$$

As the plate is moved the sample exerts a force F (N) in an attempt to resist the motion. Newton's third law of motion tells us that as the shear rate increases, the force induced to resist it also increases. The area of plate which is in contact with the sample is represented by A (m^2). The smaller this area is the larger the shear stress of the moving plate

on the sample, and the larger the resistant force. The resulting effect is a potential stress known as shear stress (τ (Pa)) and its magnitude is given by equation 2.38.

$$\tau = \frac{F}{A} \tag{2.38}$$

Figure 2.15 Parallel plate viscosity measurement: Theory of operation of Haake viscometer. The straight line plates are not part of the equipment but are drawn to aid the explanation in the text. S represents movement of the plate, A is the area of plate in contact with the sample, F represents the force with which the sample resists the movement.
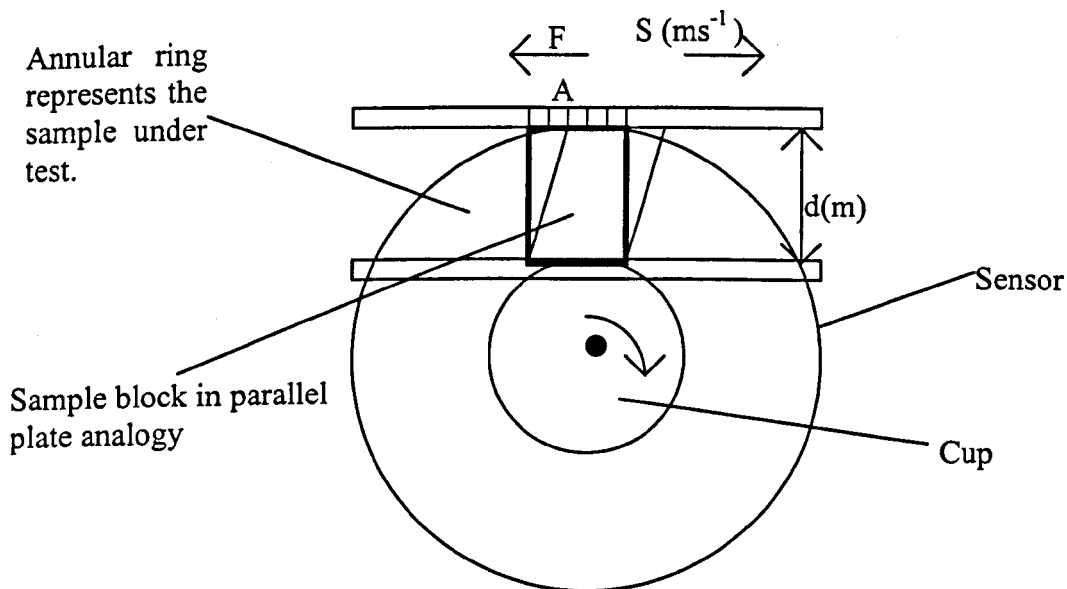
The size of the resisting force (F (N)) depends on the material's natural resistance to motion which is, as already discussed, known as viscosity (η (Pa.s)). This is defined by equation 2.39, as the ratio of the shear stress induced to the shear rate which causes it.

$$\eta = \frac{\tau}{\gamma} \tag{2.39}$$

The majority of viscometers, including the Haake, are based on the same principle, but instead of two plates make use of two cylindrical surfaces as shown in Figure 2.15 with rotation of the inner surface.

Viscosity can be measured by moving the plate in two different ways. Firstly, as was done in the trials of this chapter, the plate can be moved at a specified shear rate and the resulting resistance to motion can be measured. Secondly, as was done in the trials of chapter 6, the shear rate required to induce a specified shear stress can be measured. In both cases equation 2.39 is evaluated to calculate the apparent viscosity.

Tests as detailed as this are rarely used in ceramic ware production lines. A common replacement for the detailed test is a crude measure known as the *overswing test*. To measure the overswing value a body is immersed in the liquid sample and manually rotated through a specified angle. The rotation is opposed (rather crudely) by an elastic band in such a way that the kinetic energy of the applied rotation is converted into potential energy of the elastic band. The body is then released and allowed to rotate towards its starting point Any additional rotation past the starting point is known as the overswing and is measured in degrees. The larger this overswing value the lower the viscosity of the sample. According to one of the company's production employees, a satisfactory glaze is indicated by an overswing value of between 190° and 210°.

Chapter 3

Efficient Designs for Constrained Mixture Experiments

3.1 Introduction

3.1.1 Motivation

For the majority of mixture problems which are subject to constraints (see §2.3.1B), there are no theoretical results on the form of the optimal design. In these cases algorithmic methods are needed to locate near optimal designs (see §2.3.7B). Atkinson (1996, §3.5) notes "how difficult it is to guess the structure of good designs for constrained regions". This chapter compares the performance of the software DX4, KL and gosset, in the generation of good mixture designs, and investigates whether they can find the known D optimal designs for the unconstrained simplex, in a variety of examples for q=3. The unconstrained simplex is considered first because theoretical results can be used to asses the accuracy of any algorithmic results. This software is then used to help make conjectures on the form of D and V optimal designs (see §2.3.4) when q=3, and D optimal designs when q=4.

Attention is then turned the general form of designs which are optimal over a given set of points. Both the D and V optimality criteria will be considered for a range of constrained regions. We will start with the regular simplex and gradually modify it by using constraints of increasing complexity. The observed trends in the optimal designs will then be investigated.

3.1.2 Background

Some algorithms, for the location of near optimal designs for constrained mixture problems, were considered for a linear model by Snee & Marquardt (1974) and for a quadratic model by Snee (1975). In both papers the analysis is mainly based on the G criterion (see §2.3.4). Both of these papers stress that design algorithms should be used to aid the generation of designs and not to dictate them. It is advised that a variety of algorithms should be used and that the best design should then be selected. It is also suggested that if a design is greater than 50% G efficient, it can be considered as a good design and if a design is greater than 80% G efficient, any further improvements are of little practical benefit.

In Snee and Marquardt (1974), the XVERT algorithm is introduced and compared with the existing CADEX algorithm (Kennard and Stone, 1969). The CADEX algorithm selects designs which give an even spread of points over the simplex, and does not depend on the chosen model. Snee and Marquardt (1974) find that CADEX tends to be poor for highly constrained regions. The XVERT algorithm locates the extreme vertices of the constrained region and then, where appropriate, it will select the best subset of them with respect to the G criterion. In the majority of the cases considered in Snee and Marquardt (1974) XVERT generated better designs, in terms of the A criterion as well as the G, than CADEX. When large q leads to too many extreme points for the algorithm to run efficiently, not all of the extreme vertices need to be calculated. XVERT limits the number of candidate points, using the fact that G values tend to be higher when components with the smallest ranges are set at their extreme values. Only candidate points which have these components set at either of their bounds are considered. For the quadratic model Snee and Marquardt (1974) suggest that the centroids of each face should be added to the extreme vertices; these are estimated by averages of the relevant extreme vertices. In this chapter we restrict our attention to linear constraints and do not look at multiple constraints (see §2.3.1B).

In the Snee (1975) paper, the Wynn (1970) sequential algorithm and Wheeler (1972) exchange algorithm are used. The Wynn algorithm takes a starting design, it then sequentially adds the points with the highest prediction variance with respect to the latest design. As was discussed in the description of the KL algorithm (see §2.3.7B), the point with the highest prediction variance causes the biggest increase in $|X'X|$ (the value which is to be maximised under the D criterion) when added to a design. The Wheeler algorithm takes an n point starting design, where n is the size of the design required, and sequentially adds the f points with the highest prediction variance. It then sequentially removes the f points with the lowest prediction variance to return the design to the required size; this process is repeated until no further improvements in $|X'X|$ are possible. Using these two algorithms, Snee (1975) found that G and D efficient designs for a quadratic model could be generated by selecting points from amongst the following sets: extreme vertices, mid points of long edges, constraint plane centroids (i.e. face centroids when q > 3) and the overall centroid. The designs selected from these points tend to be more G efficient than both the extreme vertex design and the extreme vertex plus constraint plane centroid designs, which were advocated in Snee & Marquardt (1974).

The co-ordinates of the extreme vertices and their centroids (including mid points of edges) can be generated using the CONVRT and CONAEV algorithms (Piepel, 1988) for all regions constrained by both linear and multiple constraints. Alternatively, there is a variety of procedures for calculating the extreme vertices of linearly constrained regions (see, e.g., Cornell, 1990). The other candidate points which are listed in the previous paragraph, can then be calculated by averaging appropriate subsets of the extreme vertices.

· As mentioned above, the results in Snee & Marquardt (1974) and Snee (1975) are based mainly on the G criterion. It is now more usual to use the D criterion and increasingly the V criterion. However, for continuous designs G optimality implies D optimality (see §2.3.4).

It must be noted that the optimal designs will depend on the chosen model but not the parameterisation used (see §2.3.2). The D value itself will depend on the parameterisation.

3.2. The Regular Simplex

3.2.1 Choosing Candidate Points

When restricted to the full simplex D optimal continuous designs (see §2.3.4), for polynomial models up to the full cubic, use s equally weighted support points (Chan, 1995). The support points are the q vertices for the linear model, and these plus the $q(q-1)/2$ mid-edges for the quadratic model (Kiefer, 1961). These designs are otherwise known as the $(q, 1)$ and $(q, 2)$ simplex lattice designs respectively (see §2.3.5). The optimal design for the special cubic uses the $(q, 3)$ simplex centroid design (Uranisi, 1964). This design adds all $q(q-1)(q-2)/6$ averages of sets of three adjacent vertices, to the $(q, 2)$ simplex lattice design (see §2.3.5). For the full cubic model the support points are similar to those used for the special cubic, with the mid-edges replaced by the $q(q-1)$ edge points which are α units from each vertex, [where $\alpha = (5 - \sqrt{5})/10 \approx 0.2764$] (Mikaeili, 1993). These points are subsequently referred to as the *Mikaeili type points*. For the quartic model the D optimal design includes all vertices and edge midpoints plus two new types of point. These points are, using the notation of Chan (1995) (see §2.3.5) the $2.^qC_2$ points $x \leftrightarrow (\alpha_1, 1-\alpha_1, 0, \ldots, 0)$, and the $3.^qC_3$ points $x \leftrightarrow (\alpha_2, \alpha_2, 1-2\alpha_2, 0, \ldots, 0)$; where $\alpha_1 = (7 + \sqrt{21})/14 \approx 0.8273$ and $\alpha_2 = (7 - \sqrt{5})/22 \approx 0.2165$ (Kasatkin, 1974, see Chan, 1995).

When n is a multiple of s, discrete D optimal designs on the regular simplex have n/s replicates at each support point of the continuous D optimal designs. It seems likely that the

support points of continuous designs on the regular simplex, are also the support points in exact designs for all n. Recall that if support point i in an n point design is replicated n_i times then $|X'X| = (\Pi n_i).|X_s'X_s|$, where X_s is the design matrix which includes only one copy of each support point (see §2.3.5). This result for $|X'X|$ suggests that the discrete D optimal design on the regular simplex should have the s support points replicated as equally as possible. This situation is defined here as *even replication*. It is achieved for an n point design when there are $f = \text{int}[n/s]$ replicates at some points and f+1, such that the total number of points is n, replicates at the others. For the linear and quadratic model these even replication designs are known to be optimal (Gaffke, 1987).

From the designs generated later in this chapter it appears that the result is also true for the cubic and special cubic models (see §3.2.3). There are many ways in which the additional replicates can be allocated, each of the possible evenly replicated designs are defined here as being *equivalent*. Each support point is initially allocated int[n/s] replicates, n−s.int[n/s] of the support points can then be allocated an additional replicate each. The number of evenly replicated equivalent designs is therefore $^sC_{n-\text{int}[n/s].s}$.

The investigation of optimal designs on the regular simplex, in this chapter, is restricted to the special case of q=3. However many of the results and comments may also apply to the general q component case. Optimal designs are model dependent. However it is sometimes possible to find *robust designs* which are efficient for any model. (Kiefer, 1961) shows that for the linear and quadratic models, the optimal exact design will be chosen entirely from vertices, edge midpoints and the overall centroid. However as similar results are not known for the constrained case, and cubic models are also to be considered, the following five types of point, specified using the notation of Chan (1995), are used for all examples.

a) Vertices: $\qquad x \leftrightarrow (1, 0, 0)$

b) Centroid $\qquad x \leftrightarrow (1, 1, 1)$ i.e. $(^1/_3, ^1/_3, ^1/_3)$

c) Midpoints $\qquad x \leftrightarrow (1, 1, 0)$ i.e. $x \leftrightarrow (^1/_2, ^1/_2, 0)$

d) Thirds of edges $\qquad x \leftrightarrow (1, 2, 0)$

e) Axial check points $x \leftrightarrow (1, 1, 4)$

The actual points were calculated using DX4.

Less specific edge and axial points can be defined as:

$x \leftrightarrow (\lambda, 1-\lambda, 0)$ and $x \leftrightarrow (\psi, \psi, 1-2.\psi)$.

3.2.2 DX4 Using the D Criterion

The user can select any of the design point types which are given above, but subsets of a type are not allowed. Unless otherwise stated the searches below use all five categories. When inputting the number of points required, DX4 asks for the following three values: number of points to fit the model, number of points to detect lack of fit and number of points to calculate pure error. The points used to estimate lack of fit are always internal points which are not included in the known optimal designs (see §3.2.1). Pure error is calculated using replicated points (see §2.3.5). For n < s, where s denotes the number of support points (see §2.3.5), designs benefit most from the addition of one of the support points which have yet to be included in the design. If n > s then replicates are automatically included. Specifying separate numbers of points to be used for each of these three purposes, offers little benefit to the experienced user who is trying to locate the optimal design and investigate the efficiency of DX4 in doing this. Consequently in the analysis which follows, the required numbers of the latter two types of point are always set to zero.

DX4 offers optimisation according to the D criterion and also according to packing criteria which spread the points as evenly as possible. There is no option for the V criterion. There are many forms of the D-criterion (see §2.3.4), but DX4 gives the values of $|X'X|$ and $|(X'X)^{-1}|$

For the four Scheffé polynomials of degree up to 3 (see §2.3.2), the first designs generated had n = s = p design points. This is the minimum number of points which can be used to fit the models. The numbers of parameters, p, in the linear, quadratic, special cubic and full cubic models are 3, 6, 7 and 10 respectively.

A Linear model

Recall that for the linear model p = 3. The designs with between 3 and 8 points included all three vertices replicated evenly. In repeated application of the package for given n, exactly the same designs were generated every time. The D values for designs with between 3 and 8 points, were as follows (with the expressions in the brackets being obtained using equation 2.30): 1, 2 (= 1×2), 4 (= 1×2^2), 8 (= 1×2^3), 12 (= 1×2^2×3) and 18 (= 1×2×3^2). For a given n, the D values of any evenly replicated design will clearly be the same.

B Quadratic model

Recall that for the quadratic model $p = 6$. In accordance with Kiefer (1961) the optimal design with six points which was generated by DX4 was the (3, 2) simplex lattice (see §2.3.5). The D-value of this design is $(^1/_4)^6$ and was calculated in the following way.

$$D = |X'X| = |X|^2 \text{ as } n=p \text{ and } X \text{ (the model augmented design matrix) is square.}$$

The quadratic model is $\mu = E(y) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3$

$$\text{so } X = \begin{pmatrix} I_3 & 0_3 \\ L & (^1/_4).I_3 \end{pmatrix} \text{ where L is some block triangular matrix}$$

$$\therefore |X| = |I_3|.|^1/_4 .I_3| = 1.(^1/_4)^3$$

$$\therefore D = |X|^2 = (^1/_4)^6$$

Increasing the number of design points from 7 to 12 yielded designs which added an additional replicate of each support point. For each application of DX4 the same points were replicated for each size design. The order in which each point was replicated, for designs of increasing size, did not appear to exhibit any pattern. In accordance with equation 2.30 the D values for each of these designs were as follows: $2.(^1/_4)^6, 2^2.(^1/_4)^6, \ldots, 2^6.(^1/_4)^6$

For designs of size 13 to 16 the points were further replicated in the same order. For 17 points the design generated was not evenly replicated. This analysis shows that DX4 successfully locates the theoretical optimum design for $n=6$. It then locates designs which intuitively seem likely to be optimal for $n < 17$.

Identical results were observed in repeated applications of the package. This indicates a fixed, rather than random, selection of replicate order.

C Special cubic model

Recall that for the special cubic model $p = 7$. The optimal 7 point design is the (3, 2) simplex centroid design consisting of the simplex lattice generated for the quadratic model plus the centroid. DX4 was able to generate this design from a search of its entire candidate set. The D value of this design is 12^{-6}, as shown below.

Here $\mu = E(y) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$

$$\text{so } X = \begin{pmatrix} I_3 & 0_3 & 0_{3\times1} \\ L & {}^{1}/_{4}.I_3 & 0_{3\times1} \\ {}^{1}/_{3}.1_3 & {}^{1}/_{9}.1_3 & \left({}^{1}/_{3}\right)^3 \end{pmatrix}.$$

Proceeding as in §3.2.2B $|X| = \left({}^{1}/_{4}\right)^3.\left({}^{1}/_{3}\right)^3 = 12^{-3}$ and $|X'X| = 12^{-6}$.

As the number of points was increased from 8 to 13 the edge points were replicated evenly. The centroid was then replicated as the 14th point. In accordance with equation 2.30 the D values for each of these designs were as follows 2. $12^{-6}, 2^2.12^{-6}, \ldots, 2^7.12^{-6}$.

For designs with 15 and 16 points even replication was again observed, however for 17 points and above DX4 started to replicate the support points unevenly. A new notation is now introduced to specify the number of times each point was replicated. The expression *"Rcount = rep(a, b, c, ...)"* is used to indicate that 'a' points were replicated once, 'b' points were replicated twice, 'c' points were replicated three times and so on. Table 3.1 gives the Rcounts and D values of the designs which were generated by DX4, the R counts and D values which would have been obtained under even replication and the efficiency (see §2.3.4) of the DX4 designs with respect to the evenly replicated designs. The table includes these values for designs with between 15 and 19 points.

Table 3.1 Ratios of replication (Rcount), D values (obtained, and under even replication) and efficiencies for designs generated by DX4 with n = 15 to 19

n	15	16	17	18	19
Rcount	rep(0,6,1)	rep(0,5,2)	rep(0,5,1,1)	rep(0,5,0,2)	rep(0,4,1,2)
D value	$3.2^6.12^{-6}$	$3^2.2^5.12^{-6}$	$4.3.2^5.12^{-6}$	$4^2.2^5.12^{-6}$	$4^2.3.2^4.12^{-6}$
Rcount (even)	rep(0,6,1)	rep(0,5,2)	rep(0,4,3)	rep(0,3,4)	rep(0,2,5)
D Value (even)	$3.2^6.12^{-6}$	$3^2.2^5.12^{-6}$	$3^3.2^4.12^{-6}$	$3^4.2^3.12^{-6}$	$3^5.2^2.12^{-6}$
Efficiency	1	1	0.89	0.79	0.79

Table 3.1 shows that for n=17, 18 and 19 the designs generated by DX4 had lower D values than those designs in which the support points were replicated evenly. This reveals a failing of DX4. Larger designs were not considered because it seems likely that for n > 19 DX4 will continue to replicate unevenly.

D Full cubic model.

It is known that the continuous optimal design for this model includes the Mikaeili type edge points which are situated at a distance $[(5 - \sqrt{5}) / 10]$ from each vertex (Mikaeili, 1993). These points are not included in DX4's candidate set and alternative points can not be specified by the user. Because the Mikaeili type points could not be included it was not possible to generate the optimal design for the full cubic model. The ten point design generated by DX4 used the 6 thirds of edges instead of the Mikaeili type points and has a D value of $(^2/_9)^{18}$ ($\approx 1.75 \times 10^{-12}$). This value was calculated in the following way.

$D = |X'X| = |X|^2$ as n=p and X is square

The model is $\mu = E(y) = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3$

$$+ \delta_{12} x_1 x_2 (x_1 - x_2) + \delta_{13} x_1 x_3 (x_1 - x_3) + \delta_{23} x_2 x_3 (x_2 - x_3)$$

so $X = \begin{pmatrix} I_3 & 0_3 & 0_3 & 0_{3\times1} \\ A & r.s.I_3 & r.s.(r-s).I_3 & 0_{3\times1} \\ A^* & r.s.I_3 & r.s.(s-r).I_3 & 0_{3\times1} \\ ^1/_3.1_3 & ^1/_9 1_3 & 0_{1\times3} & (^1/_3)^3 \end{pmatrix}$

where $r = ^1/_3$, $s = ^2/_3$, $A = \begin{pmatrix} r & s & 0 \\ r & 0 & s \\ 0 & r & s \end{pmatrix}$ and A^* is A with r and s interchanged

$|X| = (1/3)^3 . \begin{vmatrix} rs & rs(r-s) \\ rs & rs(s-r) \end{vmatrix}^3$

$= (1/3)^3 . \{(rs)^2(s-r) - (rs)^2(r-s)\}^3$

$= (1/3)^3 . (rs)^6 . 2^3 . (s-r)^3$ (3.1)

$= (^1/_3)^3 . (^2/_9)^{2.3} . 2^3 . (^1/_3)^3 = (^2/_9)^9$ and so $|X'X| = (^2/_9)^{18}$.

The design based on the Mikaeili type points has a D value of $(^{2 \cdot \sqrt{5}}/_{375})^{18}$ ($\approx 2.88 \times 10^{-12}$). This D value was calculated using equation 3.1, with $r, s = (5 \pm \sqrt{5})/10$. The DX4 design is therefore only 61% D efficient compared to the known optimal design. The fact that alternative candidate points can not be added is perhaps the main failing of DX4.

As the number of design points was increased from 11 to 16, an additional replicate of each edge third was sequentially added. In accordance with equation 1.30 the D values for each of these designs were as follows; $2.(^{2}/_{9})^{18}$, $2^{2}.(^{2}/_{9})^{18}$, $2^{3}.(^{2}/_{9})^{18}$, $2^{4}.(^{2}/_{9})^{18}$, $2^{5}.(^{2}/_{9})^{18}$, $2^{6}.(^{2}/_{9})^{18}$.

General conclusions

All evenly replicated designs with the same number of design points, have the same D value and are equally efficient over the specified candidate sets. The same choice of evenly replicated design was generated every time the algorithm was run. This suggests that DX4 checks through the points in a deterministic order. If the algorithm only replaces points when an improvement in the D value is obtained, then the points included in the final design will be the first points in the list, for the type of candidate point in question. If it replaces for D values which are less than or equal to the previous best, then the points included will be the last ones in the list. The method which DX4 actually used was not investigated further as it is of little consequence.

It was observed for the special cubic model with all candidate points, that DX4 began to replicate unevenly when 17 points were asked for. This was remedied by removing the thirds of edges from the candidate set, as they have never been observed in optimal designs for the special cubic. Uneven replication re-occurred when a design of size 38 was searched for. In other cases non support points were found to be included in the designs. At least one of these errors was found to occur for designs with 40, 17 and 17 points for the linear, quadratic and special cubic models respectively. It was always possible to remedy this by reducing the number of candidate points. However in all cases a value of n was soon reached, at which the optimal designs were only found when the candidate sets were restricted to those points which were known to be in them. The same problems are likely to occur for the constrained examples, for which there are no theoretical results for reducing the candidate sets.

3.2.3 KL Using the D Criterion

For the KL algorithm (Atkinson & Donev, 1992) the candidate set must be specified by the user. All 22 points discussed above (see §3.2.1) were used in the searches described below. Unless otherwise specified KL(k=4, l=4, t=100) is used in all searches (see §2.3.7B).

Results for all models

The results in this section apply to all four Scheffé polynomial models used above, unless specified otherwise. In all cases with $p \leq n \leq 42$ KL produced the known optimal designs over the given candidate sets, with equal replication of support points. The partially random choice from the set of equivalent designs (see §3.2.1) depended on the search specifications used. This is because, unless modified by the user, the seed of the random number generator remains the same. If a given point was replicated f times in an n point design it was not always replicated f times in the n+1 point design. This is an obvious consequence of the stochastic element of the algorithm.

Additional results for the full cubic model

Specification of the canonical full cubic model, of the form given as equation 2.20, was found to be difficult using the KL input codes. Each of the difference terms (i.e. $\delta_{ij}x_ix_j(x_i - x_j)$) can be expanded into two terms with coefficients of the same magnitude but the opposite sign. However there exists no facility in the KL algorithm for constraining the coefficients in this manner. This problem was solved by using the intercept form of the Scheffé full cubic model (see §2.3.2).

The 10 point design generated by KL used five Mikaeili type points but one edge third. This design had a D value of 4.22×10^{-14} and was 94% efficient compared to the known optimal design which uses all six Mikaeili type points and has a D value of 4.49×10^{-14}. When the search was intensified by using KL(k=10, l=15, t=5 000) the same design was observed. It was not until the candidate set was reduced to contain the support points and three edge thirds (including the one incorrectly used in the design above) that the true optimal design over the candidate set was generated. However when the algorithm was used for designs of size 11 to 30 using the standard search, the ten points in the known optimal design (see §3.2.1), were replicated evenly every time.

3.2.4. Comparison of DX4 and KL Under the D Criterion

Flaws were identified in both DX4 and the KL algorithm in the above analysis. However DX4 failed to produce the evenly replicated design for a 17 point design based on the quadratic model whereas KL was successful for all designs up to 42 points. KL also performed adequately up to this size for the full cubic model, apart from for the 10 point design. In addition to this the ability to include user defined points, such as the six Mikaeili type points, in the KL search is a clear advantage.

Although we can prevent DX4 from going wrong by reducing the candidate set this approach is of little use when theoretical results are not known. One further advantage of DX4 however is that it will generate the candidate sets for irregular constrained regions.

3.2.5 Gosset Using the D Criterion

Gosset uses a completely different approach to the generation of optimal designs, based on a structured search of the entire constrained region which does not require candidate points. Although it takes considerably longer for gosset to run on a fast UNIX machine than it takes DX4 or KL to run using a pc with a Pentium 2 processor, gosset provided designs with equal or better D values than DX4 and KL, in every example where they were all used. Instead of the Scheffé model gosset uses the q-1 term parameterisation (see §2.3.2), and re-scales the q-1 components it chooses onto the range (-1, 1).

Specification of the special cubic was not straightforward. Each term had to be specified individually and forced in using the following model code (see gosset operating manual).

a( )+b( )+c( )+a( )*b( )+a( )*c( )+b( )*c( )+a( )*b( )*c( ).

When this approach was used the variables were not re-scaled onto the range (-1, 1), this had an effect on the accuracy of the designs generated (see §3.3.3C). No reason could be found in the gosset operating manual for why the transformation was not performed in this case.

The designs generated using 10 iterations of gosset, for n = p, . . , 2p, for the D criterion and each of the four canonical models used so far, replicated evenly the support points in the n=p continuous optimal designs. In many cases the non zero components in the support points were perturbated slightly (± 0.001 for each proportion). Different small perturbations of the same point were sometimes observed in the same design.

3.3 Gosset Using the V Criterion on the Regular Simplex

3.3.1 Introduction.

As the D criterion minimises the volume of the normal confidence region for the parameter estimates (see §2.3.4), it is an intuitively sensible choice of criterion when the aim of the experiment is model-based analysis. In many practical examples however the aim of the experiment is to produce a model for predicting future response values. In these cases the V value, which can be calculated using equation 3.2, is a more appropriate measure to minimise as it gives the average prediction variance over the entire feasible region R. The same model parameterisations as described above are again used (see §3.2.5).

$$V = \iiint_R \underline{x}' (X' X)^{-1} \underline{x} d\underline{x} \quad \text{where R is the region } \{x_i \geq 0 \text{ for } i=1,2,3 \text{ and } \Sigma x_i = 1 \tag{3.2}$$

The integrand can be rewritten in the following way

$$V = \iiint_R tr\{(X' X)^{-1} \underline{x}\underline{x}'\} d\underline{x}$$

$$= tr\{(X'X)^{-1} \iiint_R \underline{x}\underline{x}' \, d\underline{x}\}$$

$$= tr\{(X'X)^{-1} M\} \quad \text{where } M = \iiint_R \underline{x}\underline{x}' d\underline{x} \tag{3.3}$$

M is known as the *moments matrix*. Unless the integrand can be evaluated directly it is estimated by gosset using Monte Carlo estimation. This is done by defining a uniform pdf over R and then estimating $E(\underline{x}\underline{x}')$ for this uniform probability density function (pdf). This expected value for a uniform pdf is denoted here as $E_U(\underline{x}\underline{x}')$. The uniform pdf over the 3 dimensional simplex, R, is $= 1/ \iiint_R d\underline{x}$. Therefore $M = E_U(\underline{x}\underline{x}') / \iiint_R d\underline{x}$. An estimate of $E_U(\underline{x}\underline{x}')$ is obtained by surrounding the feasible region with a regular shape inside which it is easy to simulate random points. Many of these random points are then simulated and each

$f(\underline{x})$ is calculated and averaged over the points which lie in the constrained region. $\iiint_R d\underline{x}$ is estimated by multiplying the area of the surrounding shape by the proportion of points which fell inside the constrained region.

Every time gosset runs 4 000 Monte Carlo (MC) iterations are automatically performed to estimate M. If M has already been estimated then it is modified using the results of the new iterations. Every time the algorithm is run the V values and design points can change slightly. The number of iterations to be carried can also be specified. In the remainder of the thesis the intensity of the search used will be specified using the notation gosset (number of iterations, number of MC samples). One disadvantage of the V criterion is that it requires more computer time than the D criterion.

3.3.2. Exact Moments Matrix for the Regular q=3 Problem

For the regular simplex it is possible to calculate the exact moments matrix by evaluating equation 3.3. It is then possible to calculate exact V values. The integral of a matrix is defined as the matrix of each entry integrated individually. The evaluation of M is especially simple in this case as all model terms and therefore all cross products of model terms, which occur in $\underline{xx'}$ are of the form $x_1^a x_2^b x_3^c$. This yields integrals of the form

$$\iiint_R x_1^a x_2^b x_3^c \, dx_1 dx_2 dx_3$$

which, using the fixed sum constraint to replace x_3 become

$$\iint_S x_1^a x_2^b (1-x_1-x_2)^c \, dx_1 dx_2 \text{ where } S=\{x_i \geq 0 \text{ for } i=1,2 \text{ and } \Sigma x_i \leq 1\}$$

$$(\Sigma x_i \text{ is } \leq 1 \ i=1,2 \text{ as } \Sigma x_i = 1, \ i=1,2,3 \text{ and } x_3 \geq 0 \}.$$

This integral is equal to the expected value of the trivariate Dirichlet distribution, given by equation 3.4, scaled by the area of the feasible region

$$\Gamma(a+1)\Gamma(b+1)\Gamma(c+1) / \Gamma(a+b+c+3) = a!b!c! / (a+b+c+2)! \qquad (3.4)$$

The area of the feasible region can also be evaluated using the Dirichlet distribution in the following way.

$$\iiint_R 1 d\underline{x} = \iiint_R x_1^0 x_2^0 x_3^0 d\underline{x} = \Gamma(1)\Gamma(1)\Gamma(1) / \Gamma(3) = 1/2! = 1/2.$$

We can now write down equation 2.5 which can be used to evaluate exactly each term of the moment matrix.

$$2 \iiint_R x_1^a x_2^b x_3^c \, dx_1 dx_2 dx_3 = 2 \times a! b! c! / (a+b+c+2)!$$

(3.5)   From equation 3.5 it can be seen that a simplification of the calculation of M procedure is possible. The order of a, b and c makes no difference to either side of equation 3.5 and many of the entries of M will be the same (for a worked example see §3.3.3A).

All implementations of algorithms which use gosset estimates of M, must use the same model and transformation to the range (-1, 1), that was used by gosset to estimate M. The estimated M, $\hat{M}$, can be replaced by the exact M for the regular simplex. The transformation used by gosset for the regular simplex is given in equation 3.6 and the integral to be evaluated is given by equation 3.3 with x and X replaced by y and Y.

$$y_i = 2x_i - 1 \tag{3.6}$$

The following three stage procedure is then used to calculate M to replace $\hat{M}$

a) Write M in terms of the transformed y variables using the model in question.

b) Substitute equation 3.6 into M and expand so that it only includes x variables. in the form $f(x_1^a x_2^b x_3^c)$.

c) Assume a uniform pdf and calculate the values of $E_U(x_1^a x_2^b x_3^c)$, which make up each of the $f(x_1^a x_2^b x_3^c)$ in stage b, using equation 3.5.

3.3.3 Designs Generated

The performance of gosset will be investigated for the quadratic model as it was decided that the linear model is probably too small to provide interesting comparisons.

A) Linear model

The optimal designs generated for the D criterion, which consist of the vertices replicated evenly, were again generated for the V criterion. As mentioned above, the criterion values depend on the estimate of M, and different values can be given every time the program is run. To investigate the possible range of V values, gosset was run three separate times, using a new set of 4 000 MC samples, for designs with between three and seven points. Table 3.2 gives the observed ranges, and shows that V values are reasonably invariant to different estimates of M, though they do tend to be biased. The estimated V values are very close to the known V values of the true optimal design. This suggests that using MC estimates of M in the gosset algorithm, leads to efficient designs. The exact V values were calculated using the equation $\frac{1}{6}\sum n_i^{-1}$, which is now derived.

$$V = \text{trace}\{(X'X)^{-1}M]\}$$

Because the trace is required, only the diagonal elements are needed.

As $\mu = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$, $M = E[\text{diag}(x_i^2) + P + P']$ where P is an upper triangular matrix.

$$E_U(x_i^2) = {}^{2.2!}/_{4!} = {}^1/_6 \text{ by equation 3.5}$$

$$\therefore M = \text{diag}_3({}^1/_6) + E_U(P + P')$$

$$X = \begin{pmatrix} 1_{n_1 \times 1} & 0 & 0 \\ 0 & 1_{n_2 \times 1} & 0 \\ 0 & 0 & 1_{n_3 \times 1} \end{pmatrix} \therefore (X'X) = \text{diag}(n_i) \text{ and } (X'X)^{-1} = \text{diag}(n_i^{-1}), \text{ where } n_i \text{ is}$$

the number of times support point i is used in the design.

$\therefore (X'X)^{-1}M = \text{diag}_3(\frac{1}{6} n_i^{-1}) + E_U(P_2 + P_2')$, where P_2 is some upper triangular matrix

$\therefore V = \frac{1}{6} \sum n_i^{-1}$.

Table 3.2 Range of V values in repeated runs

V value	Number of points				
	3	4	5	6	7
Min	0.501	0.415	0.333	0.250	0.222
Max	0.503	0.417	0.333	0.251	0.224
Known	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{2}{9}$

The results in Table 3.2 were then compared to the V values of designs generated using the true M. Implementation of the three stage procedure for calculating M is now outlined for the linear model.

$$\underline{yy}' = \begin{pmatrix} 1 \\ y_1 \\ y_2 \end{pmatrix} (1 \quad y_1 \quad y_2) = \begin{pmatrix} 1 & y_1 & y_2 \\ y_1 & y_1^2 & y_1 y_2 \\ y_2 & y_1 y_2 & y_2^2 \end{pmatrix}$$

$E_U(y_1) = E_U(2x_1 - 1) = 2 E_U(x_1) - 1$. $E_U(x_1) = E_U(x_1^1 x_2^0 x_3^0) = 2.1! / 3! = 1/3$ by equation 3.5 Therefore $E_U(y_1) = 2 \times (1/3) - 1 = -1/3$.

$E_U(y_1^2) = E_U(y_2^2) = E_U[(2x_1 - 1)^2] = 4E_U(x_1^2) - 4E_U(x_1) + 1$. From equation 3.5 $E_U(x_1^2) = 1/6$ so $E_U(y_1^2) = 4 \times (1/6) + 4(1/3) + 1 = 1/3$

$E_U(y_1 y_2) = E_U[(2x_1 - 1)(2x_2 - 1)] = 4$. $E_U(x_1 x_2) - 4 E_U(x_1) + 1$. $E_U(x_1 x_2) = 1/12$ using equation 3.5 so $E_U(y_1 y_2) = 4 \times (12) - 4 \times (1/3) + 1 = 0$, which yields the following M^*.

$$
\begin{pmatrix}
1 & -1/3 & -1/3 \\
-1/3 & 1/3 & 0 \\
-1/3 & 0 & 1/3
\end{pmatrix}
$$

Gosset can be made to use the exact $M^{\bullet}$, once it has been saved in the appropriate form, by inserting the following line after the moments matrix generation stage of the gosset command sequence (see gosset operating manual).

!cp "saved name" moments.h.

When the true value of M was used, gosset generated evenly replicated designs, with V values equal to the known values in Table 3.2, for all cases other than for four and five design points. In these two cases completely different designs were produced. These consisted of the three vertices plus a seemingly random point on the edge of the simplex for n=4, and the vertices with rep(2,1) plus a seemingly random point, on the edge opposite the vertex which was replicated, for n=5. To ascertain whether this was a failing of gosset, V values for the designs where the vertices are replicated evenly, were calculated using the true M. These were found to be identical to the V values of the two alternative designs outlined above. Due to these seemingly random placements of the non-vertex points, further investigation was carried out. This revealed the new result that for n=4 a design with the 3 vertices appears to be V optimal if the fourth point is anywhere on the simplex edge. Similarly if n=5, a design with the 3 vertices replicated with the ratio rep(2,1) appears to be V optimal if the fifth point is anywhere on the edge opposite the replicated vertex.

It has been shown that D optimal designs for the linear model are also V optimal (Hardin & Sloane 1993). It has been conjectured that in general, when the two criteria produce similar designs, V optimality implies D optimality but the converse may not be true (Hardin & Sloane 1993). The two V optimal designs discussed above, i.e. those which use points which occur anywhere on a specified edge, give examples of designs which are V optimal but not D optimal.

B) Quadratic model

To evaluate gosset's performance for searches of varying intensity, the size of the required design was increased from 7 to 14, the number of iterations was varied between 1 and 100 and the number of MC samples used to estimate M was varied between 4 000 and 2

million. Because the number of MC samples was increased from 4 000 to 2 million, the generated designs got closer to the design generated using the exact M. However there was little difference (V values calculated using the exact M were all identical to 6 decimal places) between the designs generated using one million samples, two million samples and the exact M. From this we can conclude that the use of 1 million samples provides a highly accurate design. When the minimum number of 4 000 MC samples was used, the designs were often different from those generated using 1 million MC samples ($\pm$ 0.04 for at least one proportion), but were almost fully efficient (V values calculated using the exact M were all identical to 4 decimal places). It can be concluded that the use of 4 000 MC samples is likely to provide an efficient design, but if there is an interest in the exact form of the optimal design, one million samples should be used.

The number of iterations used tended to make little difference. In the majority of cases using the estimated M, 1, 10 and 100 iterations all produced the same best designs. In a limited number of cases increasing from 1 iteration to 10 iterations gave a better design, however increasing to 100 never gave additional improvement. As the use of 10 iterations does not increase running time dramatically, it is suggested here that this number should be used with an estimated M. When the exact M was used, increasing the number of iterations from 10 to 100 and in some cases even 300, was found to increase design efficiency. When the true design is required to a high level of accuracy ($\pm$ 0.001 for each proportion), considerably more than ten iterations should be used. The above analysis suggests that a search of intensity (10, 1 million) is sufficient for generating efficient designs. The required intensity is likely to be higher than this for large n.

Laake (1975) showed that for the quadratic model, the V optimal design over the (q, 2) simplex centroid, has r_1 (i.e. vertex weight) / r_2 (i.e. midpoint weight) = $\sqrt{\{(q^2+7q +18)/32\}}$. For q=3 this ratio is $\sqrt{3}/4$. This weights the vertices and midpoints as approximately 0.101 and 0.233 respectively. In Martin *et al* (1999) the optimal design over a grid of size 0.01 for n=6 was found to move the 3 midedges slightly inside the simplex towards the opposite vertex. If the moved mid-edges are written as (0.5 - ϵ, 0.5 - ϵ, 2ϵ), the most V efficient design which was found had $\epsilon \approx 0.00214$. In a later section the optimal weightings for the continuous optimal design, over the points of the (3, 3) simplex centroid lattice are estimated (see §3.9). This continuous design adds the centroid to the (3, 2) simplex

centroid design considered by Laake, and the vertices, midpoints and centroid appear to be weighted almost in the ratio (1, 2, 1).

Gosset was then used to generate V optimal designs for between 6 and 14 points using at least 10 iterations and at least 1 million samples to estimate M, i.e. at least gosset(10, 1 million). The best designs of size 6 to 14 consisted of various selections from the following four types of point: vertices, exact or slightly perturbated midpoints, axial (or near axial) points and the centroid.

The best designs of size 7, 10, 11 and 13 consisted purely of points from the (3, 3) simplex centroid design, to a high degree of accuracy (± 0.0002 for each proportion). In all other cases apart from n=14 the designs consisted of exact vertices (exactly once), perturbations of the midpoints (± 0.01 for each proportion) and up to two additional axial points (not the axial check blends (see §3.2.1)). For n=14 the design consisted of the exact vertices, with one of them replicated twice, one axial point near to the centroid and 3 replicates of perturbated midpoints (± 0.01 for each proportion). When different estimates of M were used, based on up to 8 million MC samples, the choice of vertex and position of the axial point varied, but the V value remained approximately the same (identical to 6 decimal places).

In conclusion it can be said that the designs can be split between two types; those which contain points of the (3, 3) simplex centroid design to a high level of accuracy (± 0.0002 for each proportion), and those which consist of vertices (exactly once), perturbations of the midpoints (± 0.01 for each proportion) and up to two additional axial (or near axial) points.

The improvements which were gained by not using exact candidate point designs in the appropriate cases given above were then investigated. The "candidate point designs" considered consisted of the points in the (3, 2) simplex centroid lattice which corresponded most closely to the perturbated points in the gosset(1, 1 000 000) designs. In the candidate point designs axial points were replaced by vertices or the centroid, depending on which one the axial point was closest to. This investigation was performed to investigate the drop in efficiency which would be caused by the use of an exchange algorithm based on candidate points. The V values of the candidate point designs and the gosset(1, 1 000 000) designs were then calculated using the exact M, so that the discrepancies could be investigated. As is

to be expected, the candidate point designs which were less efficient than the gosset designs were those where internal points had been used by gosset. These less efficient designs were for n = 6, 8, 9 and 12 and had efficiencies of 99.95%, 98.7%, 99.9% and 99.6% respectively. From these efficiencies we can conclude that in many cases, the points of the (3, 2) simplex centroid lattice need to be perturbated to find the optimal design. In terms of practical applications of design theory, only the 8 point design is inefficient enough to warrant any concern. The improvements in efficiency gained from the perturbations were not enough to reject the use of algorithms based on candidate points. Irrespective of whether perturbations were needed or not, there appeared to be a pattern in the way in which the different types of points were replicated. These patterns are investigated further using the KL algorithm (see §3.4).

By comparing these results with the designs generated for the D criterion (see §3.2.5), we see that the main difference in terms of the points used is that the V criterion uses the centroid.

C) Special cubic

To fit this model the canonical parameterisation had to be fully specified (see §3.2.5), and gosset did not convert the x variables onto the scale (-1, 1). This meant that only stage (c) of the M estimation procedure was needed (see §3.3.2). As mentioned above all searches conducted in this section have an intensity of at least (10, 1 million).

Laake (1973, 1975) showed that the V optimal design over the points of the (q, 3) simplex centroid lattice, replicates the vertices the midpoints and the centroid in the ratio $\sqrt{(q^4-10q^3+59q^2-218q+1608)}$: $16\sqrt{(2q^2-18q+49)}$: 108. When q=3 this is approximately 1 : 1.6 : 3 which gives the following weights: (0.093, 0.148, 0.278). It has since been shown that amongst all designs with s=p, the Laake weights are optimal over the whole simplex (Liu and Neudecker, 1995).

The designs generated for this model were more variable than for the quadratic. For n = 7 and 8 the designs generated used the points of the (3, 3) simplex centroid to a high degree of accuracy (±0.002 for each proportion). More specifically they used each of the vertices once, small perturbations of the midpoints once (±0.002) and one, then two, copies of the exact centroid. The 9 point design used the (3, 2) simplex lattice with small perturbations of the midpoints plus 3 near axial points of the following form (0.27, 0.36, 0.37). The 10 point

design uses the three vertices, three points close to the midedges (two being singletons and one replicated twice) and three near axial points comparatively close to the centroid. As the size of the design increases between 11 and 15 the points of the (3, 2) simplex centroid are again used to a high level of accuracy. The points chosen for replication in these designs tended to be the centroid followed by the midedges followed by the vertices. This is consistent with the Laake (1975) weightings given above.

The values of n times the Laake weights, i.e. n.(0.093, 0.148, 0.278), were then examined for n between 7 and 15. For n=9 and 10, i.e. those designs where the candidate points were deviated from, two of the resulting products were approximately 2.5 and 1.5. For all other values of n none of the products deviated from integers by more than 0.37. This suggests that when one of the products of n and the optimal weights deviates sufficiently from integers, efficient designs have some of the points of the (3, 2) centroid lattice replaced with axial points. This is considered further using KL (see §3.4).

V values based on the true M were calculated using matlab, and found to equal the V values estimated by gosset(10, 1 million) to 3 decimal places. For the other three models the V values estimated by gosset were equivalent to the values using the true M to 6 decimal places. The reduced accuracy when the re-scaling was not used suggests that re-scaling to the range (-1, 1) is done to improve the accuracy of gosset.

D) Full cubic model

Designs with between 10 and 21 points were generated using the true M and 100 iterations. The form of the designs was again complicated, but the designs generated for n between 10 and 16 followed a more definite pattern than for previous models. In all designs the three vertices were included exactly once. For designs with n between 11 and 16, edge points and axial points were used in addition to the centroid. All six edge points were always included. If (n–9)/3 = a *remainder* r then the design consisted of the vertices, all three axial points replicated a times, and r replicates of the centroid. The value of λ (which in the notation of 3.2.1 specified the edge points) was seen to increase monotonically with n from 0.2748 to 0.2943. The value of ψ (which in the notation of 3.2.1 specified the axial points) ranged non-monotonically between 0.177 and 0.2651. These two types of point are very similar to the Mikaeili type and axial check points respectively. If these two types of point were added to the DX4 candidate set and used in the KL algorithm, then it is hypothesised that similar designs could be generated by an exchange algorithm with little loss in

efficiency. For designs of size 19 and 20 a similar pattern was observed but with two replicates of the edge points and λ and ψ varying slightly on the different edges ($\pm 1 \times 10^{-4}$). The 17 and 18 point designs consisted of the edge points and points of the form $(5a, a, 14a)$ with $a \approx 0.05$. The value of λ varied for all points but no pattern was observed.

Using a similar approach to that used for the special cubic, the designs generated by gosset were compared to the rounded values of n times the weights in the hypothesised continuous optimal candidate point design. The candidate points were the vertices, the centroid and the Mikaeili type points. As the design generated in Laake (1975) does not use the Mikaeili type points, the weights were estimated by generating a 100 point design using KL, and dividing the number of replicates for each point by 100 (see §3.9.2). When these estimated weights were multiplied by values of n between 10 and 21, whenever the results were close to integers (± 0.4), the integers were equal to the numbers of replicates in the gosset designs. When the evaluation of n times the weights were not close to integers it was observed that internal points had been used by gosset. A similar procedure was used for the quadratic optimal design over the vertices, midedges and centroid, and the same conclusion was reached. These results for the quadratic, special cubic and cubic models suggest that if the values of n times the weights in the continuous optimal designs, are not excessively different from integers, then rounding these values can lead to a good n point discrete design.

3.4 KL Using the V Criterion on the Regular Simplex

The KL algorithm (Atkinson & Donev, 1992) implemented on matlab (Mathworks, 1992) for the V criterion (see §2.3.8) was used here. Recall that this algorithm requires the user to input a moments matrix. The moments matrix can be estimated by gosset (see §3.3.1) or, in the case of the regular simplex, it can be calculated directly (see §3.3.2).

In all cases the complete set of candidate points (see §3.2.1), will be used even though the above analysis suggests which ones are likely to be included. All candidate points were used to gain some insight on the performance of KL for future cases in which the best candidate points may not be known.

In the majority of cases for the linear, quadratic and special cubic models, the KL designs consisted of those exact points in the (3, 3) simplex centroid design, suggested by the perturbation designs of gosset (see §3.3.3). All designs were equivalent to or worse than

the designs generated by gosset. As is to be expected, the improvement gained by gosset depends on how much the perturbed points deviate from the points of the (3, 3) simplex centroid design. The improvement also depends on the chosen value of gosset's *step* parameter. The *step* parameter specifies how many criterion-improving iterations gosset will perform (see Hardin & Sloane, 1994, §4.4). The default step parameter can be modified by the user.

For the quadratic model for example, the designs generated by gosset, for n = 7, 10, 11 and 13 were found to consist of the simplex centroid points to a high level of accuracy. The KL V values were therefore identical to those from gosset with small rounding errors. Many of the points in designs of size 6, 9, 12 and 14 were perturbations of various sizes around midpoints, with the size of these perturbations affecting the efficiency of the KL designs. The best 9 point design generated by gosset contained three axial points in between the centroid and the axial check blends. Consequently there was no obvious candidate point design suggested by the perturbation design of gosset and the design which KL produced was nearly 1% less efficient.

For the special cubic model the designs generated by gosset for n = 7, 8 12, 13 and 14 were found to consist of the simplex centroid points to a high level of accuracy. KL easily located the relevant designs and calculated V values which were identical apart from rounding errors. The gosset designs of size 9 to 11 moved the centroid by varying degrees and the KL V values were thus less efficient.

In observing the more structured patterns which KL was forced to adopt, due to its use of exact candidate points, it appeared that the points (vertices, midpoints, centroid) were being replicated approximately in the proportions (1, 3, 1) and (1, 2, 3) for the quadratic and special cubic models respectively. These approximations are not hugely different from the algorithmic estimate (see §3.9) and the Laake (1975) result, which gave ratios of (1, 2, 1) and (2, 3, 6) respectively (see §3.3.3 B, C). As gosset was using similar patterns, but with the candidate points slightly perturbated, it may be possible to obtain the optimal designs using the KL algorithm followed by the perturbation algorithm of Martin *et al* (1999). However as the improvements for the designs considered here were all less than 1% it may not be worthwhile.

The patterns for the full cubic were more complicated. Again the reduction in efficiency from the gosset designs were greatest for those n where the Mikaeili type and

axial points had been moved by the greatest amounts. The observed patterns were much less consistent than for the other models but in general the points (vertices, Mikaeili type, axial, centroid) were replicated in the proportion (1, 2, 1, 4). Laake (1975) shows that the optimal design over the points of the (3, 3) simplex lattice, which replaces the Mikaeili type points with edge thirds, weights the points (vertices, edge thirds, centroid) according to the ratio $\sqrt{(2.(q^4 - 13q^3 + 98q^2 - 386q + 660)}) : 9.\sqrt{(q^2 - 9.q + 38)} : 54\sqrt{2}$. When q=3 this leads to a ratio of 15.1 : 40.25 : 76.37 which is approximately 1:3:5. If axial points are removed and the Mikaeili type points are given the weight of the thirds of edges, then the KL ratio is not dissimilar to this Laake result.

As gosset is not restricted to a list of pre-specified candidate points but KL is, it is not surprising that gosset produced better designs than KL. However, in the worst case the KL design was only 1% less efficient than the *gosset* design. Consequently for practical purposes it would seem that the KL algorithm is sufficiently accurate. However, this may not be the case when constrained regions are considered.

3.5 Constrained Problems on the Three Dimensional Simplex

3.5.1 Specifying Constrained Regions

In the majority of practical situations the mixture variables are constrained by upper and lower bounds of the form $a_i \le x_i \le b_i$ (see §2.3.1B). Multi-component constraints may also apply (see §2.3.1B). Both single and multi-component constraints were used for the initial trials of the JM experiment (see §5.3). In this investigation of the form of the optimal design, multi-component constraints are not considered.

The bounds are usually specified by scientists and are often imprecise. In an attempt to gain as much information as possible, scientists often choose unnecessarily wide constraints. These constraints often exceed the operating regions which are known to produce successful compositions. This, coupled with the fact that design algorithms favour extreme points of the feasible region, can lead to non-standard trials which fail to perform as expected. Analysis of the experimental results often leads to the constraints being narrowed or widened.

The constraints lead to a feasible region in which all compositions are expected to produce a satisfactory product. The feasible region is usually a convex polyhedron. An example of one is given in Figure 3.1a. The feasible region in Figure 3.1a is for an example taken from Atkinson and Donev (1992, §16.3), and has the following bounds: (0.2, 0.7),

82

(0.1, 0.6) (0.1, 0.6). The feasible regions can be expanded to fit tightly within the regular simplex by applying the L pseudocomponent or the U pseudocomponent transformation. Restricting attention to just one of the transformations leads to no loss of generality, as either transformation can be applied to any region which is constrained by single component upper and/or lower bounds. Because the L transformation preserves the orientation of the feasible region (see §2.3.1B), it is in some sense the more natural of the two to use. Here, attention is restricted to the L transformation and all regions considered have already been L transformed. If X is L transformed to T by $T = XA$, for some p×p matrix A, then $|T'T|$ can be written as $|A|^2.|X'X|$. This result shows that a D optimal design for the transformed region, will be D optimal for the original region when the inverse of the L transformation is applied to it. The condition for L transformed upper bounds being consistent is that $b_i + b_j \geq 1 \; \forall i,j$. The transformed region for the above example (Atkinson and Donev, 1992 §16.3) is shown in Figure 3.1b.

Apart from for the regular hexagon (see §3.6.6) and the parallelogram (see §3.6.2), there are no theoretical results specifying the form of the optimal design for constrained problems. Due to this lack of theoretical optima, near optimal designs need to be found using

Figure 3.1 Constrained region for example 16.3 of Atkinson and Donev (1992, §16.3)
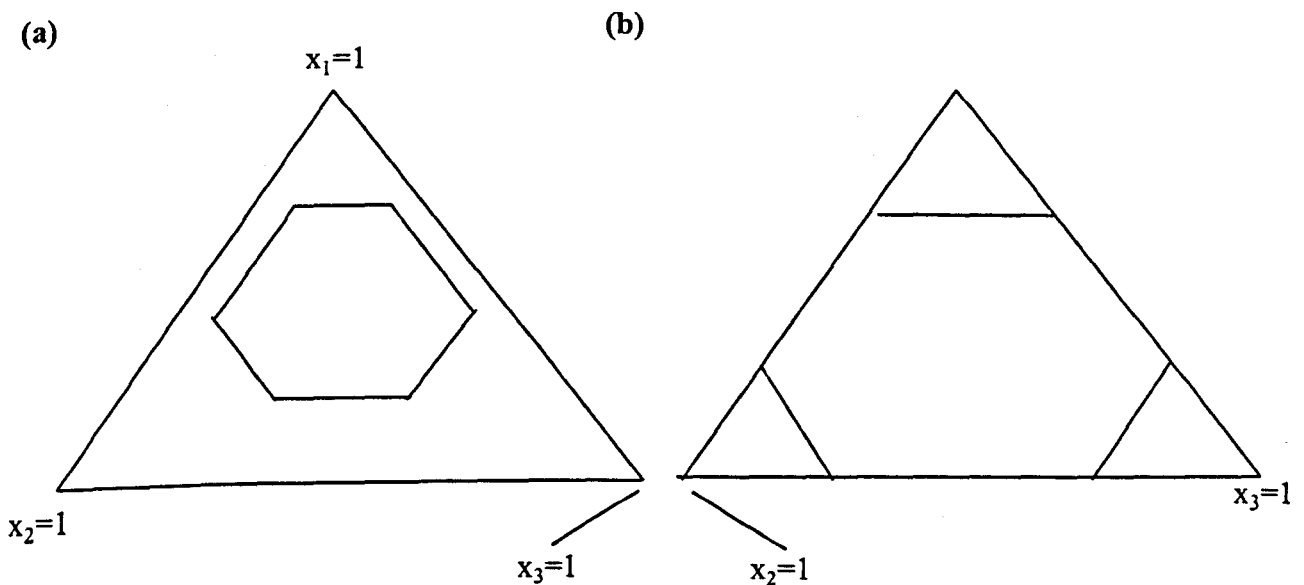
an algorithmic approach. The types of constrained region which can occur under single component upper and lower bounds can be subdivided into exhaustive categories by

considering the upper bound equation. Recall that the upper bound of each of the L

pseudocomponents is given by $(b_j - a_j) / (1 - \sum_{i=1}^{q} a_i)$. If $b_j = 1 - \sum_{i \neq j}^{q} a_i$ the L

pseudocomponent is defined on the full range between zero and 1. If this equation is satisfied for all j then the resulting region is the original simplex.

The shape of the L transformed constrained region depends on how many of the L transformed upper bounds equal 1. If the upper bound equals 1 in zero, one or two cases then the resulting shapes are the hexagon, quadrilateral, pentagon or parallelogram, and trapezium respectively. Each of these cases is discussed separately below (see §3.6).

3.5.2 Performance of Packages for a Constrained Example from Snee (1975)

Before investigating the form of the optimal design for constrained regions, the performance of the three computer based methods is considered. An example in Snee (1975) has the bounds (0.2, 0.8), (0.1, 0.8), (0, 0.5). After L transformation these bounds form an irregular pentagon with $b_1 = 6/7$, $b_2 = 1$ and $b_3 = 5/7$. Using KL for this example with low values of k and l, and the full candidate set of 31 points, illustrated some possible problems with the algorithm. In ten single trials the range of values obtained were as follows (the figure in brackets is the number of times in which the best design with a D value of 0.1239 and 8 support points was obtained): 0.1204-0.1239 ($\times$1) ($k = l = 2$), 0.1196-0.1239 ($\times$5) ($k = l = 3$). Over ten trials with 10 iterations each, the range of values obtained was 0.1230-0.1239 ($\times$8) ($k = l = 2$), 0.1233-0.1239 ($\times$9) ($k = l = 3$). Despite failing to locate the optimal design on a number of occasions these figures show that as long as the trials are not restricted to single iterations it is unlikely that the best design will be missed. DX4 also produced some strange results and failed to locate the optimal on a number of occasions.

Runs of gosset (360, 1 million) produced D values of 0.1218 and 0.1241. A variety of runs with fewer iterations were also performed and these obtained D values within the range given by KL, for designs with 10 or 11 support points. The best design had ten support points. It is likely that the optimal design with eight support points, which was located by KL, would be obtained if each of two pairs of close points in the ten point design were amalgamated into two new points and both replicated twice.

3.5.3 Candidate Points

Recall that in Snee (1975) it was found that, for constrained mixture problems, G and D efficient designs for a quadratic model could be generated by selecting points from amongst the following sets: extreme vertices, mid points of long edges, constraint plane centroids and the overall centroid. These four types of points can all be generated by DX4 for constrained regions and will be considered for inclusion in the designs generated in the remainder of this section.

It was found in Farrel *et. al.* (1967) that the continuous optimal design for the cubic model on a square, uses the vertices, plus points which are very close to the axial check blends and thirds of edges which are used by DX4. Partly due to this result and partly to provide a set of points which are well spread over the feasible region, points similar to these are added to the candidate sets for all of the models, regions and q, which are considered below. As DX4 quickly calculates the exact thirds of edges and axial check blends, the exact points rather than those used in Farrel *et. al.* (1967) are used here.

For constrained regions, Mikaeili type points are defined here, to be α of the way along each edge from each vertex, where $\alpha = (5 - \sqrt{5})/10 \times l \approx 0.2764 \times l$, and l is the length of the edge in question. This definition is based on the definition of Mikaeili points for the regular simplex. As the Mikaeili points were found to be optimal for the cubic model on the regular simplex, the Mikaeili points are also added to the above candidate sets when the full cubic model is considered.

For the remainder of this thesis attention is restricted to the optimal design over these candidate points. No further designs are generated by perturbing points. Due to this restriction some of the designs generated here, which are optimal for the given candidate sets, are different from the optimal designs generated in Martin *et al* (1999). The differences occur when the Martin *et al* (1999) designs were generated from a grid of points covering the region or by perturbating points. The drop in efficiency caused by these restriction is minimal (<1%) in all examples considered here, and will be of little consequence in a practical situation.

3.5.4 Location of Continuous D Optimal Designs for a Given Candidate Set Using the General Equivalence Theorem

One method for estimating the weights of the support points in a continuous optimal design over a given candidate set is to use algorithms to generate extremely large designs of

up to 1,000 points (due to Wynn, 1972, see Atkinson and Donev, 1992). The designs generated using an algorithmic approach often include single occurrences of points which are not in the optimal design. Any points which occur infrequently should therefore be discarded or investigated more rigorously, by using a variety of different search parameters or different algorithms. Once infrequently occurring points have been tested further, the weights can be estimated by dividing the number of times which each point occurs by the total number of points.

The Wynn procedure can only estimate the weights, and is likely to be highly computer intensive. Consequently an alternative approach was developed. This alternative approach begins by locating the most likely design points using the Wynn method with only 1 000 points. The alternative approach is described in full below.

To find D efficient designs for a range of constrained regions, and hence to identify the types of point which form the support for D efficient designs in general, the following approach was developed.

By the general equivalence theorem (see, e.g., Atkinson and Donev, 1992, §9.2) it is known that continuous G optimal designs are also D optimal.

Let Ψ represents a continuous design in which each of the support points in design matrix S have weights given by the diagonal elements of W (see §2.3.6B). The *S-prediction variance* of support point $\underline{s}_i$ is defined here as the scaled prediction variance at $\underline{s}_i$ of the continuous design Ψ. The S-prediction variance is denoted here as $pv_S(\underline{s}_i)$. Using the expression for prediction variance obtained by scaling equation 2.23 for σ^2, and the alternative expression for the information matrix given by equation 2.25, the S-prediction variance is given by equation 3.7.

$$pv_S(\underline{s}_i) = \underline{s}_i{}'(S'WS)^{-1}\underline{s}_i. \tag{3.7}$$

The S-prediction variance of each support point in a continuous G optimal design equals the number of parameters in the model (p) (see Atkinson and Donev, 1992, §10.8). Using this result, the weights of the continuous G optimal and hence continuous D optimal design over the given candidate set can be estimated by modifying the weights until the S prediction variances of all support points equal p. The following procedure for doing this was

implemented in `matlab`. The code for this algorithm is given in electronic form on the attached disk as wt.m.

1. A 100 point design was generated using KL with the full DX4 candidate set. For the full cubic model the Mikaeili points were calculated and included in the search. All points in the generated designs were used in the remaining stages.

The KL algorithm was used as DX4 was found to be less accurate, especially with large designs (see §3.2.2). The gosset output was harder to search because it is in random order and uses perturbations of the support points and internal points which are completely different from the candidate points.

2. The algorithm used the approach of assigning an associated value λ_i to every point, and dividing λ_i by $\Sigma\lambda_i$ to obtain the weights w_i.

3. An arbitrary starting value of 100 was associated with each point. Although better starting values could have been assigned using the number of times each point occurred in stage 1, this would have been unnecessarily time consuming because all that is needed is an arbitrary starting point.

4 The S-prediction variance of each point is calculated using equation 3.7 and is denoted by v_i.

5. By varying the weights and investigating the resulting change in prediction variance, it was observed that increasing the weight of a point would reduce S-prediction variance with respect to the subsequent continuous design. Similarly it was observed that decreasing the weight of a point would increase its S-prediction variance with respect to the subsequent continuous design. This means that if $v_i < p$, λ_i needs to be decreased, and if $v_i > p$, λ_i needs to be increased. It was hypothesised that the required change in λ_i would be proportional to the discrepancy between the S-prediction variance of support point i and the desired S-prediction variance (p). The design is therefore improved by calculating a new value of λ_i using equation 3.8 where L =10 000 000

$$\lambda_i' = \lambda_i + {}^1/_L \times \text{int}[L.(v_i - p)] \tag{3.8}$$

6. Stages 4 and 5 are repeated until the S-prediction variance of each point equals p to four decimal places. If this was not achieved after one million iterations then the process was re-started from stage 1, using KL to generate a design with 100 additional points.

7. If at any time it was found that the algorithm was trying to assign a weight of less than 0 to any of the points, it was realised that the point in question was not included in the optimal design. The point was therefore removed from the candidate set, which essentially sets its weight to 0, and stages 3 to 6 were repeated.

All continuous D optimal designs obtained in the remainder of the thesis are obtained using this approach based on the general equivalence theorem.

3.5.5 Estimating the Weights of Continuous V Optimal Designs for a Given Candidate Set

The above approach can not be used for the V criterion. Because the Wynn (1972) approach should be accurate enough to give a good picture of what constitutes an efficient design, it is used to find V efficient designs for a range of constrained regions (see §3.9). In finding V efficient designs for a range of regions the algorithm also identifies the types of point which will form the support for V efficient designs in general.

3.6 Designs for the Quadratic Model, the D Criterion and q=3

3.6.1 Outline

The aim for this section is to start with the regular simplex, for which many theoretical results are known, and to gradually modify it by removing vertices using constraints of increasing complexity. The regions to be looked at are listed below.

1) Parallelograms

2) Trapezia

3) Regular Pentagons

4) Irregular Pentagons

5) Hexagons which are symmetric under a third turn

6) Hexagons which are reflection symmetric in the line $x_i = x_j$

7) Hexagons which are symmetric under a half turn and in the line $x_i = x_j$

8) Hexagons with no symmetries

Recall that attention is restricted to designs using candidate points, and all regions have already been L transformed. The most general type of constrained region is the irregular hexagon which is shown in Figure 3.2. All other regions can be obtained from this one by removing some or all of its constraints. Expressions for its edge lengths are given below so that conditions for all other constrained regions can be derived using them.

Figure 3.2 Hexagonal constrained region with no symmetries

The numbers 1 to 6 are used to identify the 6 different edges of the feasible region.
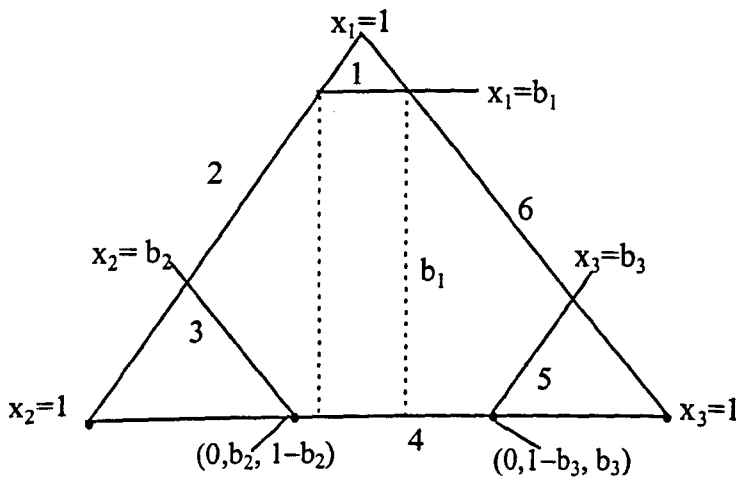
The edge length of an equilateral triangle with height 1 is $1 \times 2/\sqrt{3}$. Therefore the distance from the end of edge $x_i=0$ to any point on it, is $2/\sqrt{3} \times$[perpendicular distance from $x_i=0$] i.e. $2/\sqrt{3} \times$[component value of x_i].

$\therefore$ length of edge (4) $= 2/\sqrt{3}.b_2 - 2/\sqrt{3}.(1-b_3) = 2/\sqrt{3}[b_2+b_3-1]$

and length of edge (1) $= 2/\sqrt{3}.(1-b_1)$ as the triangle cut off by $x_1=b_1$ is also equilateral.

Similar results apply for the other four edges of the hexagonal constrained region.

3.6.2 Parallelogram

As the parallelogram can be linearly transformed to a square, it is a special case of constrained region where optimality results on the square can be applied. A parallelogram of the type shown in Figure 3.3(a) occurs if $b_1 = 1$ and the length of edge (4) is 0. The length of

edge (4) will be 0 if $b_2 + b_3 = 1$ (see §3.6.1). This condition is intuitively sensible because moving between the two ends of the line $x_1=0$, represents a total change of 1 in the values of x_2 and x_3. Due to the fixed sum constraint we can eliminate x_1 and represent the constrained region by the isosceles triangle of Figure 3.3(b). Any two constraints which satisfy the condition $b_2 + b_3 = 1$ can be represented on the isosceles triangle in a similar way to that in which they would be represented on the regular simplex. An example of two such constraints is given on Figure 3.3(b).

Figure 3.3 Parallelogram shaped constrained regions

(a) 3 Components   **(b) 2 Components**

NB the value of x_i for a composition represented by any particular point is not represented by the distance from the point to the line $x_i=0$
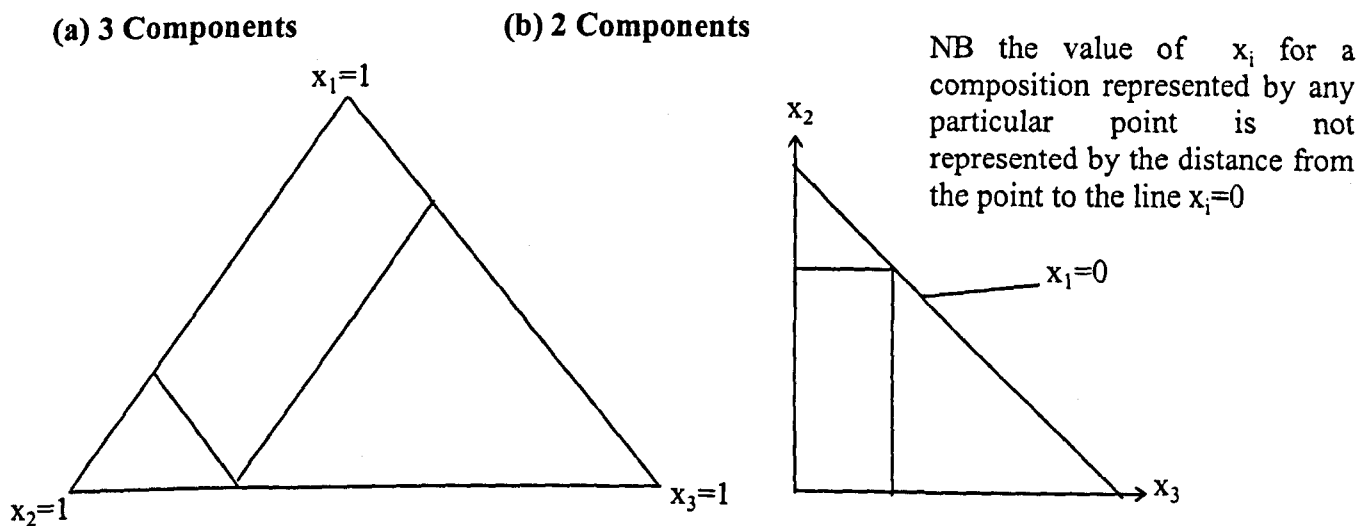
The rectangular constrained region of Figure 3.3(b) can be linearly transformed to a square. According to Atkinson and Donev (1992, §13), Kurotschka (1981) shows that optimality results for a quadratic model and two qualitative factors with three levels, can be applied to the same shaped regions for quantitative variables. The optimality results for the cube (Kiefer, 1961) can therefore be applied to the parallelogram. The continuous D optimal design on the cube, and hence the parallelogram, uses the vertices, midpoints and centroid with (approximate) weights 0.146, 0.080 and 0.096. In contrast with the constrained regions which follow, the midpoints of the long and short sides are equally weighted. The exact design for n=100, which was obtained by Snee (1975), is consistent with this because the points were replicated according to the rounded values of 100 times the weights.

3.6.3 Trapezia

The simplest modification to the regular simplex is when one vertex is removed. This occurs when only one of the L transformed upper bounds is less than one. For simplicity it is assumed that the $x_1=1$ vertex is removed and $b_1 = b < 1$. The resulting shape is the trapezium shown in Figure 3.4.

As b is reduced from 1 to zero the constrained region moves from the full simplex to an approximate parallelogram. The optimal designs at b=1 and in the limit as b approaches 0 are the (3, 2) simplex lattice (see §3.2.2B) and the known optimal design on the parallelogram (see §3.6.2). When b actually equals zero the design region is a two dimensional unconstrained simplex (i.e. a straight line) and the optimal design is the (2, 2) simplex lattice (see §2.3.5). As b moves from 1 the top vertex becomes an edge which clearly has two

Figure 3.4 A trapezium which is reflection symmetric in the line $x_2 = x_3$
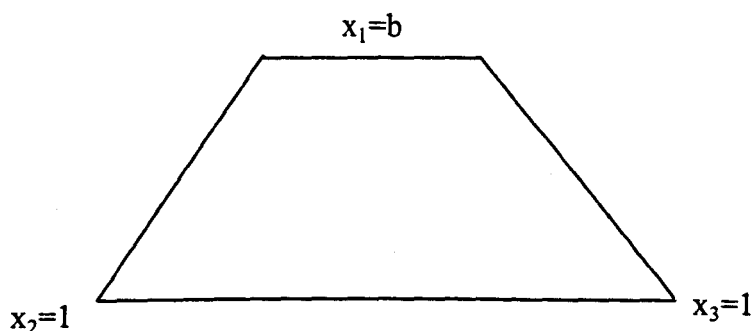
vertices and a midpoint. For b = 0.999 it was found that the optimal design occurred when the weight of the top vertex in the simplex lattice design was split equally between the two new vertices. For b=0.001 it was found that the vertices, midedges and centroid had approximately the same weights as the corresponding points of the parallelogram (see §3.6.2).

A variety of intermediate trapezia were then considered with b being reduced from 1 to 0 in steps of 0.2. The estimated weights for all six cases are given in Table 3.3 and plotted in Figure 3.5. In Table 3.3, and all subsequent tables of a similar type, the number in brackets next to the description of the type of point, indicates how many points of this type there are.

Table 3.3 Estimated weights, as $b_1=b$ is reduced from 1 to 0 in steps of 0.2, for the trapezium which is reflection symmetric about the line $x_2 = x_3$

b_1	$x_1=b_1$ vertices (2)	$x_1=0$ vertices (2)	$x_1=b_1$ midedge (1)	$x_2,x_3=0$ midedges (2)	$x_1=0$ midedge (1)	a.v. centroid (1)
~1	0.084	0.167	0	0.167	0.167	0
0.8	0.105	0.165	0	0.147	0.167	0
0.6	0.136	0.162	0	0.113	0.163	0.016
0.4	0.151	0.157	0	0.085	0.148	0.067
0.2	0.153	0.151	0.012	0.079	0.129	0.091
~0	0.146	0.146	0.080	0.080	0.080	0.096

Figure 3.5 Estimated weights, as $b_1=b$ is reduced from 1 to 0 in steps of 0.2, for the trapezium which is reflection symmetric about the line $x_2 = x_3$
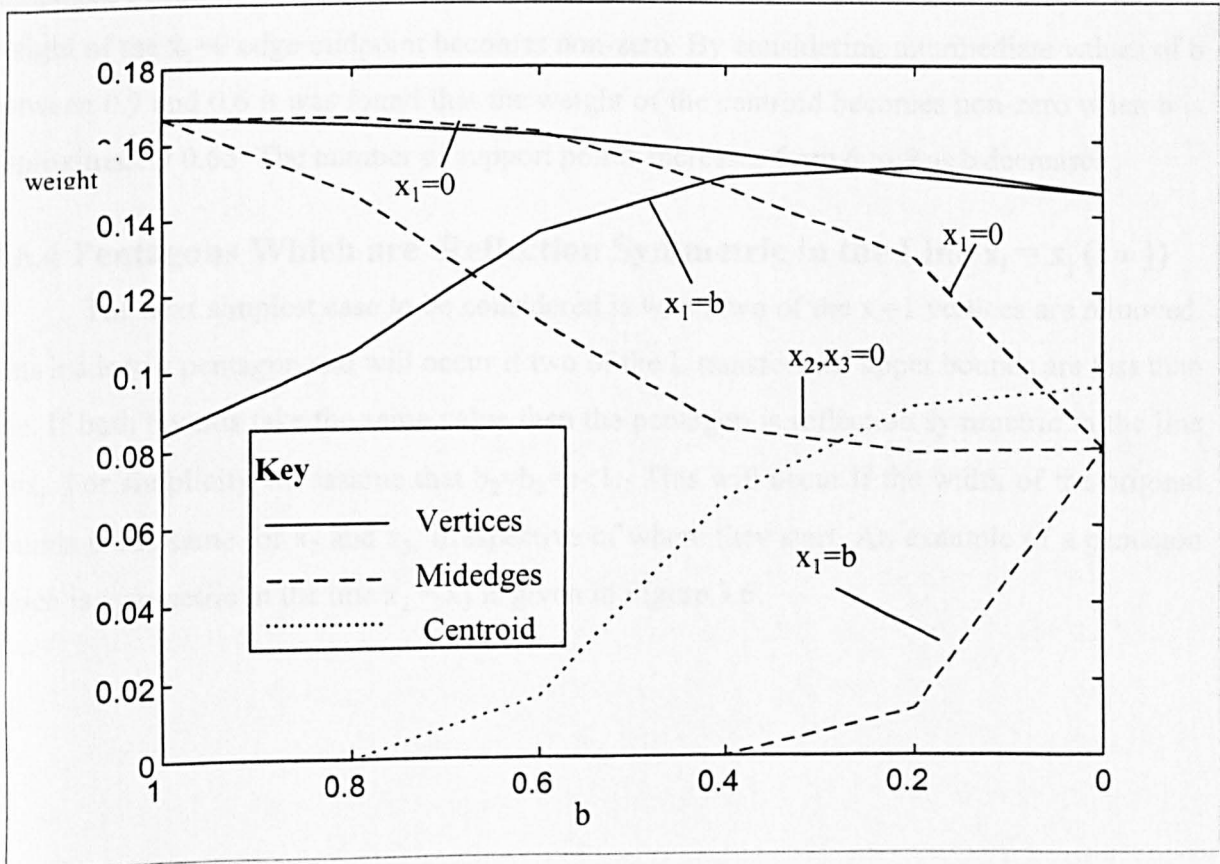
As the value of b decreases the weights exhibit clear patterns. From Figure 3.5 it can be seen that the $x_1=0$ vertices have comparatively high weighting throughout. The other points with comparatively high weighting are the $x_1=0$ midedge for b between 1 and 0.6, and the $x_1=b$ vertices for b close to 0.2.

As b decreases and the distance between the two $x_1=b$ vertices increases, their weight increases, but at a reducing rate. As b approaches zero and the region approaches a parallelogram, the weight of the $x_1=b$ vertices suddenly decrease slightly to the parallelogram weights. The distance between the $x_1=0$ vertices does not change with b, however a majority of the other points move towards them, and their weight decreases comparatively slowly.

The weight of the $x_1=b$ midedge remains zero until b_1 is approximately 0.2. By considering intermediate values of b between 0.4 and 0.2, it was found that its weight becomes non-zero very close to 0.2. The $x_2=0$ and $x_3=0$ simplex edges decrease in length with b, as this happens the weight of their midpoints decreases at a constantly reducing rate. The weight of the $x_1=0$ midedge decreases slowly at first but then at a faster rate as the weight of the centroid, and later the $x_1=b$ midpoint, become non-zero. The weight of the a.v. centroid becomes non-zero when b_1 is approximately 0.6, and then increases as the weight of the midpoints continue to decrease. This increase in the centroid weight appears to slow when the weight of the $x_1=0$ edge midpoint becomes non-zero. By considering intermediate values of b between 0.7 and 0.6 it was found that the weight of the centroid becomes non-zero when b is approximately 0.65. The number of support points increases from 6 to 9 as b decreases .

3.6.4 Pentagons Which are Reflection Symmetric in the Line $x_i = x_j$ (i ≠ j)

The next simplest case to be considered is when two of the $x_i=1$ vertices are removed. This leads to a pentagon and will occur if two of the L transformed upper bounds are less than one. If both bounds take the same value then the pentagon is reflection symmetric in the line $x_i=x_j$. For simplicity we assume that $b_2=b_3=b<1$. This will occur if the width of the original bounds is the same for x_2 and x_3, irrespective of where they start. An example of a pentagon which is symmetric in the line $x_2 = x_3$ is given in Figure 3.6.

Figure 3.6 Pentagon which is reflection symmetric in the line $x_i = x_j$
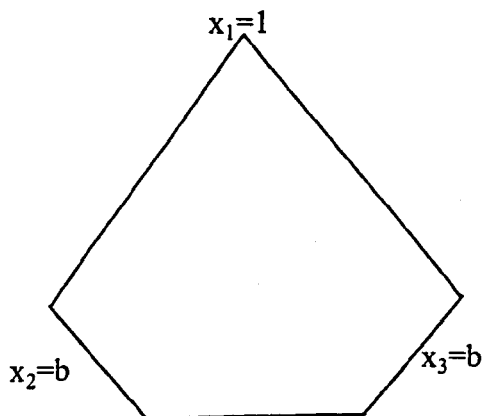
As b moves from 1 to 0.5 the constrained region moves from the regular simplex to a parallelogram. In the same way as for the trapezium, the optimal designs for the two extreme cases are the (3, 2) simplex lattice design (see §3.2.2B), and the optimal design on a parallelogram (see §3.6.2). This was confirmed by taking b = 0.999 and b= 0.5001.

To investigate the complete range of regular pentagons, b was gradually reduced from 1 to 0.5 in steps of 0.1. The estimated weights for these six cases are given in Table 3.4 and plotted in Figure 3.7. When b is close to 1 the weight which would originally have been assigned to the two bottom vertices, is split unevenly between the two new vertices on each of the two new edges. The x_2=0 and x_3=0 edges become longer than the x_1=0 edge, and the vertices on the x_1=0 edge get less of the weight. For this reason the tabulated weights for these vertices when b_1=1, have been split according to the same ratio as when b_1=0.9. In the case where b is increased from 0.5, the weight which was assigned to the vertex opposite the x_1=1 vertex, is split equally between the two new vertices.

In general the x_1=1 vertex has a comparatively high weighting throughout. Other points with comparatively high weighting are the midedges for b near to 0.9, and the other four vertices for b near to 0.6.

The weight of the x_1=1 vertex, which has no other points particularly close, continually reduces but by very small amounts. The weights of the other two vertices increase at steady rates as the length of the edge between them increases, then reduce again as the region approaches the parallelogram. For the x_2=0 and x_3=0 vertices this decrease is only slight. For the x_1=0 vertices the decrease is rapid and begins around b=0.6. By taking

intermediate values of b between 0.7 and 0.6, it was found that the weight of the $x_1=0$ vertices appears to reach a maximum at around b = 0.65.

Table 3.4 Estimated weights, as $b_2=b_3=b$ is reduced from 1 to 0.5 in steps of 0.1, for the pentagon which is reflection symmetric about the line $x_2 = x_3$

b	$x_1=1$ vertex (1)	x_2 and $x_3=0$ vertex (2)	$x_1=0$ vertices (2)	$x_2,x_3=0$ midedges (2)	$x_2,x_3=b$ midedges (2)	$x_1=0$ midedge (1)	a.v. centroid (1)
1	0.167	0.096	0.071	0.167	0	0.167	0
0.9	0.166	0.102	0.076	0.163	0	0.154	0
0.8	0.164	0.117	0.095	0.144	0	0.107	0.019
0.7	0.157	0.136	0.139	0.097	0	0	0.097
0.6	0.153	0.149	0.136	0.081	0	0	0.116
~0.5	0.146	0.146	0.073	0.080	0.080	0	0.096

As b decreases and the length of the $x_2=0$ and $x_3=0$ edges decrease, the weight of their midpoint decreases at an increasing and then decreasing rate. The weight of the $x_1=0$ edge midpoint reduces at a rapidly increasing rate and becomes zero when b is approximately 0.7. The weight of the $x_2=b$ and $x_3=b$ edge midpoints, appear to be zero until b is 0.5. Taking intermediate steps

between 0.6 and 0.5 it was seen to became non-zero when b is approximately 0.57. The weight of the centroid becomes non zero when b is approximately 0.8, it has rapidly increased by the time b=0.7, but has only increased by a small additional amount by the time b =0.6.

As the region approaches a parallelogram the trends which had previously been observed alter. As mentioned above the weight of the $x_2=0$ and $x_3=0$ vertices, which had previously increased, now decreased slightly. The two $x_1=0$ vertices combine to form a single vertex. Much of their weight appears to be transferred to the $x_2=b$ and $x_3=b$ edge midpoints,

whose weight suddenly becomes the same as the other non-zero-weighted midedges. A small amount of the centroid's previous weight also appears to be added to these midpoints.

The number of support points varies from 6 to 8 to 9 to 8 to 10 to 9 as b decreases.

Figure 3.7 Estimated weights, as $b_2 = b_3 = b$ is reduced from 1 to 0.5 in steps of 0.1, for the pentagon which is reflection symmetric about the line $x_2 = x_3$
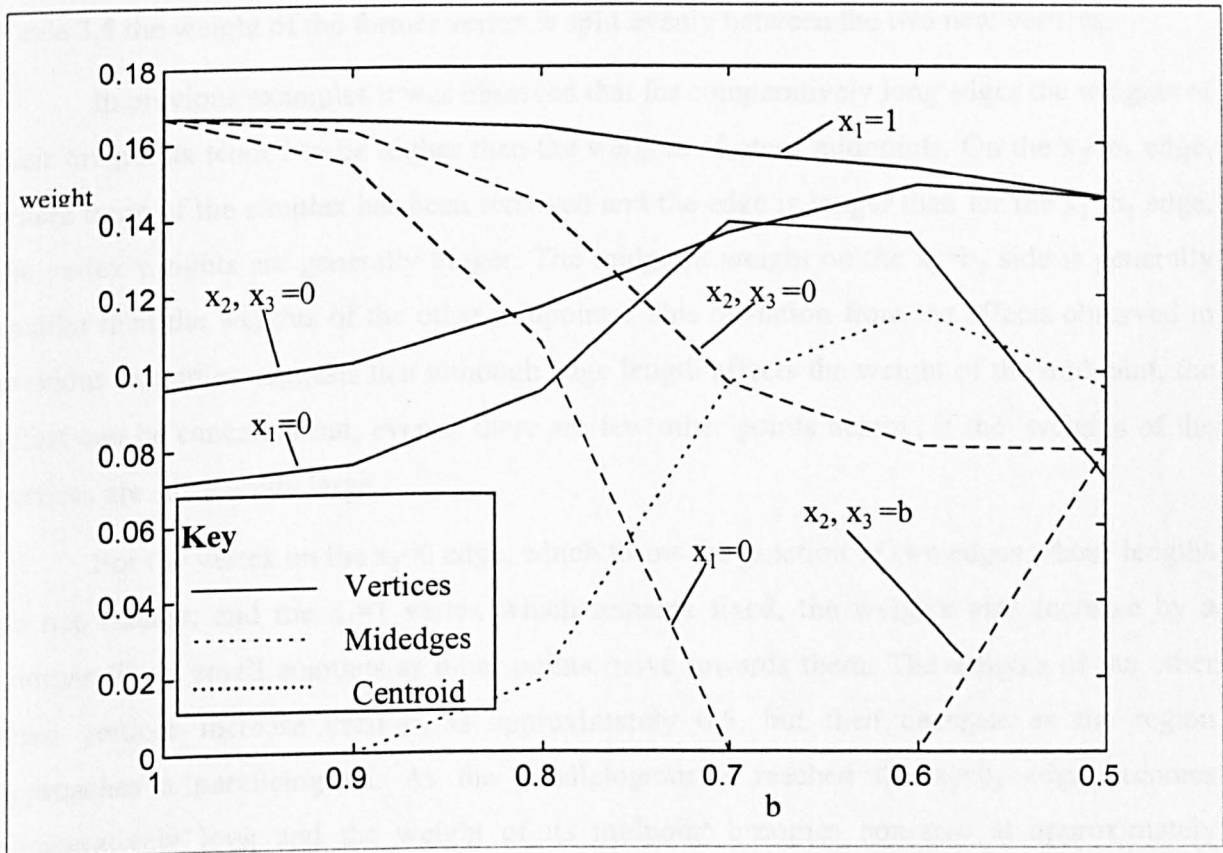
3.6.5 Non Symmetrical Pentagons

A non-symmetric pentagon will occur when one L transformed upper bound is 1 and the other two are less than 1 but different. For simplicity it was assumed that $b_1 = 1$ and b_2, b_3 < 1. To explore systematically a range of regions for constant b_2, b_3 was started close to b_2 but then reduced. It must be noted that keeping b_2 constant means that this example will only cover a small subset of the many non symmetric pentagons which may occur. It was decided that $b_2 = 0.9$ was too close to the regular simplex to provide an interesting analysis. The analysis of this section therefore takes $b_2 = 0.8$ and gradually reduces b_3 from 1 to 0.8 and then

between 0.7 and 0.3 in steps of 0.2. The estimated weights for these four cases are given in Table 3.5, along with the extreme cases of the trapezium (b_1, b_3 =1, b_2 =0.8), and parallelogram.

When the x_3=1 corner of the simplex is removed, the weighting of 0.165, which was previously assigned to the vertex, is split unevenly between the two new vertices. The listed weights for these vertices when b_3=1, have been split using the same ratio as when b_3=0.8. When b increases from 0.2 the single vertex on the line x_1=0 splits into two vertices, and in Table 3.5 the weight of the former vertex is split evenly between the two new vertices.

In previous examples it was observed that for comparatively long edges the weights of their midpoints tended to be higher than the weights of other midpoints. On the x_3=b_3 edge, where more of the simplex has been removed and the edge is longer than for the x_2=b_2 edge, the vertex weights are generally bigger. The midpoint weight on the x_3=b_3 side is generally smaller than the weights of the other midpoints. This deviation from the effects observed in previous examples suggests that although edge length affects the weight of the midpoint, the effect can be cancelled out, even if there are few other points nearby, if the weights of the vertices are sufficiently large.

For the vertex on the x_2=0 edge, which forms the junction of two edges whose lengths do not change, and the x_1=1 vertex which remains fixed, the weights also increase by a comparatively small amounts as other points move towards them. The weights of the other three vertices increase until b_3 is approximately 0.5, but then decrease as the region approaches a parallelogram. As the parallelogram is reached the x_3=b_3 edge becomes comparatively long and the weight of its midpoint becomes non-zero at approximately b_3=0.3.

The length of the x_3=0 edge remains unchanged but other points move towards it, and as this happens the weight of its midpoint gradually reduces. As the length of the x_2=0 edge decreases rapidly the weight of its midpoint also decreases rapidly. The weight of the x_2=0 midpoint then increases slightly at b=0.3 but attains the weight of the mid-edge point of a parallelogram at b_3=0.2. The weight of the x_3=b_3 midedge remains zero for the majority of the time but then rapidly increases to the weight of the mid-edge point of a parallelogram. As the length of the x_1=0 midedge rapidly decreases, so does the weight of its midpoint. It seems likely that the weight of this midpoint will become 0 for b slightly lower than 0.7. The weight

Table 3.5 Estimated weights for the non symmetrical pentagon with b_2=0.8 and b_3=0.8, then for b_3 reduced from 0.7 to 0.3 in steps of 0.2

As the complete table is too wide for the page it has been split into 2 parts.

Each point only occurs once in the region.

b_3	x_1=1 vertex	x_3=0 vertex	x_2=0 vertex	x_1=0, x_2=b_2 vertex	x_1=0, x_3=b_3 vertex
~1	0.165	0.105	0.091	0.105	0.074
0.8	0.164	0.117	0.117	0.095	0.095
0.7	0.160	0.119	0.133	0.107	0.128
0.5	0.156	0.132	0.150	0.122	0.153
0.3	0.148	0.151	0.146	0.097	0.129
~0.2	0.146	0.146	0.146	0.073	0.073

b_3	x_1=0 midedge	x_3=0 midedge	x_2=0 midedge	x_2=b_2 midedge	x_3=b_3 midedge	a.v. centroid
~1	0.147	0.147	0.167	0	0	0
0.8	0.105	0.144	0.144	0	0	0.019
0.7	0.044	0.128	0.112	0	0	0.069
0.5	0	0.109	0.074	0	0	0.105
0.3	0	0.095	0.083	0	0.045	0.107
~0.2	0	0.080	0.080	0.080	0.080	0.096

of the x_2=b_2 midpoint remains 0 until the region approaches a parallelogram. When a parallelogram is reached this midpoint receives the standard midpoint weight. The weight of the centroid increases rapidly at the start, it then slows down and begins to decrease slightly as the region approaches a parallelogram.

3.6.6 Hexagons Symmetric Under a Third Turn

This type of hexagon will occur when edges (1), (3) and (5) of the parallelogram in Figure 3.2 are equal and edges (2), (4) and (6) of Figure 3.2 are equal. These edge equalities will clearly occur when $b_j = b$ $\forall j$ (see §3.6.1), i.e. the original constraints are all the same width irrespective of their exact location (see §3.5.1). These hexagons are also symmetric about the lines $x_i = x_j$ $\forall i,j$. The constraints are consistent for b between 0.5 and 1, because the condition required is that $b_i + b_j \geq 1$ $\forall i,j$ (see §3.5.1).

The simplest case is the regular hexagon in which all lengths are equal. Therefore if the length of edge (1) is equal to the length of edge (2) the hexagon is regular.

$$\text{If}(1) = (2), \text{then}$$

$$1 - b = b + b - 1$$

$$\therefore b = \frac{2}{3}$$

If each hexagon in the range $\frac{1}{2} \leq b \leq \frac{2}{3}$ is rotated through 60° it is similar to one of the hexagons in the range $\frac{2}{3} \leq b \leq 1$. The smaller of any two similar hexagons, i.e. the ones which occur in the range $\frac{1}{2} \leq b \leq \frac{2}{3}$, can be expanded to fit the larger one, and the same optimal design applies. Consequently only hexagons in the latter range are considered here. The feasible region for $b = \frac{1}{2}$ is a triangle and corresponds to the regular simplex obtained when $b = 1$.

From runs of KL, the support points in the D optimal design for the regular hexagon appear to be the six vertices and the a.v. centroid. The weights for the optimal design over these seven points can be calculated by direct enumeration of the determinant followed by simple differentiation. Because the region is symmetric about all three axes, the six vertices, which are represented by the six permutations of the vector $(\frac{2}{3}, \frac{1}{3}, 0)$, will receive the same weight. Writing the vertex and centroid weights as w_v and w_c, the design matrix for these seven points as X, and the diagonal matrix containing their weights as W, `maple` was used to show that $|X'WX| = d.w_v^5 w_c$ where $d = \frac{2^5}{3^{15}}$. Based on the fixed sum constraint the substitution $w_c = 1 - 6w_v$ was made, and by differentiation, it was found that the determinant is

99

maximised if $w_v = \frac{5}{36}$ and hence $w_c = \frac{1}{6}$. Snee (1975) states that the optimal design, when n=36, has 5 replicates at each vertex and 6 at the centroid. This is consistent with the above result.

As b moves from $\frac{2}{3}$ towards 1, three of the sides get bigger and three of the sides get smaller. The weight of the long edge midpoint appears to gain increasingly non-zero weights until b=1 at which point the (3, 2) simplex lattice is optimal. To investigate the changes between these extreme cases, b was reduced from 1 to 0.7 in steps of 0.1 and the estimated weights are given in Table 3.6.

From Table 3.6 it appears that the centroid weight becomes non-zero between 0.9 and 0.8. By considering intermediate values of b it was found that the centroid weight actually becomes non-zero when b is approximately 0.88. By a similar approach it was found that the weights of the midpoints become zero at approximately 0.75. Between approximately b=0.75 and approximately $b=\frac{2}{3}$ the weight of the vertices and centroid each appear to take constant values which are approximately the values given in Table 3.6. This result can be proven and the exact values can be calculated by using a similar approach to that

Table 3.6 Estimated weights for the hexagon which is symmetric about a third turn with b reduced from 1 to 0.7 in steps of 0.1

b	Vertices (6)	Long edge Midpoints (3)	Centroid (1)
~1	0.083	0.167	0
0.9	0.090	0.154	0
0.8	0.111	0.074	0.111
0.7	0.139	0	0.167
$\frac{2}{3}$	0.139	0	0.167

used for the regular hexagon. If the weight of the mid-edge points is zero then the midedge points can be taken out of the design. Using the same design matrix as was used for the

regular hexagon, but with 2/3 and 1/3 replaced by b and 1-b respectively, maple was used to calculate that $|X'WX| = c'.w_v^5 w_c$ $\forall b$, where $c' = \frac{2}{3}b^4(1-3b+3b^2)^2(2b-1)^4(b-1)^4$. This result shows that the weights of the vertices and centroid will each take constant values ($w_v = \frac{5}{36}$ and $w_c = \frac{1}{6}$) for all hexagons which are symmetric about a third turn and have long mid-edge points with weight zero. Overall there are six support points at b=1, nine support points for $1 < b \leq 0.88$, ten support points for $0.75 \leq b \leq 0.88$ and six support points for $\frac{2}{3} \leq b \leq 0.75$.

As b moves between $\frac{2}{3}$ and 0.5 the patterns observed in the weights between 1 and $\frac{2}{3}$ are repeated in the opposite order but at a faster rate. The relationship between the weights of the optimal designs in the two ranges of b is non linear. Due to the width of the two ranges the rate of change for $\frac{1}{2} \leq b \leq \frac{2}{3}$ is on average twice the rate of change between $\frac{2}{3} \leq b \leq 1$. As b moves towards $\frac{1}{2}$, the differences in the rates of change becomes increasingly great. The rate is considerably more than three times faster for b near to $\frac{1}{2}$ than it is for b near to 1.

KL was used to generate efficient 20 and 30 point designs for hexagons which were reflection symmetric about a third turn with b=0.9, 0.8 and 0.7. In all six cases the best designs generated were all selected from the ten support points discussed above. In all cases the replicate numbers were close to the values of n times (w_v, w_m, w_c). For low n, and b slightly above 0.75, the lightly weighted mid-edges are unlikely to be included in the optimal design. For low n, and b slightly below 0.88, the lightly weighted centroid is unlikely to be included in the exact optimal design.

Example 16.3 in Atkinson & Donev (1992) has the following constraints; (0.2, 0.7), (0.1, 0.6), (0.1, 0.6). These constraints correspond to $b = \frac{5}{6}$ (= 0.833) after L transformation. It is noted here that this value of b is towards the middle of the range in which all ten support points are used. From the way in which the weights appear to be changing as b is decreased from 0.9 to 0.8, it looks likely that the weight of all three types of point will be reasonably close for this value of b. In Atkinson and Donev (1992), when KL was used to generate exact designs for n = 10, 20 and 30, the 10 support points were equally replicated in all cases. When DX4 was used to perform the same task it failed in some cases. For n=35, equal replication is not possible. The design for n=35 was slightly less efficient (compared with the continuous design) than the equally replicated designs. This design measure was based on the (normalised) D efficiency (see §2.3.4), of the continuous design with weights given by the proportions in which each support point was replicated. For n=997 an extremely good design was found. This design replicated the vertices, midedges and centroid; 100, 109 and 70 times

respectively. These replicate numbers suggest that the weights are not as close to each other as was suggested above. As discussed above however, the moderately similar weights are close enough to cause equal replication, for low n which are multiples of the number of support points. This was to be expected because of the large deviations from integers which were required before interior points were included in the gosset designs on the regular simplex (see §3.3.3 and §3.4).

· It is noted in Atkinson and Donev (1992) that the examples with n=35 and n=997 show that exact designs can sometimes depend in an unpredictable way on the value of n. The examples also show that exact designs for constrained examples, are likely to be more efficient, when n allows the proportion of the design points, at each support point, to be close to the weights in the continuous optimal design. This is similar conclusion to that obtained using gosset for the regular simplex (see §3.3.3).

3.6.7 Hexagons Which are Reflection Symmetric in the Line $x_i = x_j$

This occurs when all three L transformed constraints are less than 1, and two of them share the same value. For simplicity it is assumed that $b_2=b_3=b$. To investigate this type of region it was decided that b_1 would be kept constant and b would start at a value close to b_1 and gradually be reduced. Values of $b_1=0.8$ and b=0.9 were considered to be too close to the regular simplex to provide an interesting analysis, so b_1 was taken to be 0.7 and b was started at 0.9 and reduced to 0.5 in steps of 0.2. The estimated weights for these three cases are given in Table 3.7, together with the weights for the extreme cases of the trapezium and a pentagon. It must be noted that keeping b_1 constant means that this example will only cover a small subset of the many non reflection symmetric hexagons which may occur.

When b is 0.9 and the two corners of the simplex are removed the weight of 0.164 assigned to the $x_2=1$ and $x_3=1$ vertices was split unevenly between the two sets of two new vertices. The $x_2=0$ and $x_3=0$ edges are shorter than the $x_1=0$ edge and the vertices on the $x_1=0$ edge get more of the weight. The weights for these vertices when b=1 have been split in the same ratio as when b=0.9. When b increases from 0.5 the weight of the $x_1=0$ vertex is split evenly between the two new vertices.

The weight of the $x_1=b_1$ vertices increases between b=1 and b=0.8 but then remains approximately the same. As b decreases the $x_2=0$ and $x_3=0$ vertices move further away from the $x_1=0$ edge, their weight increases at a gradually decreasing rate but drops as the parallelogram is reached. The $x_1=0$ vertices increase to start with but begin to drop very

Table 3.7 Estimated weights for hexagons which are reflection symmetric in the line $x_i = x_j$ with $b_1 = 0.7$ and $b_2 = b_3 = b$ reduced from 1 to 0.5 in steps of 0.1

b	$x_1=b$ vertices (2)	$x_2=x_3=0$ vertices (2)	$x_1=0$ vertices (2)	$x_2=x_3=0$ midedges (2)	$x_1=0$ midedge (1)	$x_2=0, x_3=0$ midedge (2)	a.v. centroid (1)
~1	0.122	0.068	0.096	0.131	0.167	0	0
0.9	0.126	0.074	0.104	0.098	0.132	0	0.065
0.8	0.142	0.121	0.115	0.019	0.060	0	0.146
0.7	0.139	0.139	0.139	0	0	0	0.167
0.6	0.140	0.147	0.130	0	0	0	0.167
~0.5	0.139	0.136	0.079	0	0	0.097	0.097

rapidly when b is approximately 0.6. As b decreases the lengths of edges with weighted mid points also decrease, the weights of these midpoints consequently decrease and are zero when b is approximately 0.7. The weight of the centroid increases, at a gradually reducing rate, to a maximum at around b=0.7. This maximum weight is maintained until b=0.6 but by b=0.5 an additional support point has been added and the weight has dropped to 0.097.

3.6.8 Hexagons Which are Symmetric Under a Half Turn and in the Line $x_i = x_j$

For simplicity it is assumed that the hexagon is symmetric about the line $x_2 = x_3$, this will occur if $b_2 = b_3$ (see §3.6.1). For rotational symmetry sides (1) and (4) must be equal, i.e. $(1 - b_1) = (b_2 + b_3) - 1$ (see §3.6.1).

$$\therefore 1 - b_1 = b_2 + b_3 - 1$$

$$\therefore \sum b_i = 2$$

Four examples of this type of region were looked at. The first example is due to Snee (1975) and has L transformed constraints of $(^2/_{13}, \, ^{12}/_{13}, \, ^{12}/_{13})$. The other three examples have $b_1 = b$

decreasing from 0.9 to 0.3 in steps of 0.3, and $b_2=b_3=$ $^1/_2(2-b)$. The estimated weights for these four cases are given in Table 3.8 together with the extreme cases of the parallelograms which occur at b=1 and b=0. The constrained region for the example due to Snee (1975) is given in Figure 3.8.

Table 3.8 Estimated weights for hexagons which are symmetric about a half term in the line $x_i = x_j$ with $b_1=0.7$ and $b_2=b$ reduced from 1 to $^2/_{13}$

b_1	$x_1=b_1$ or 0 vertices (4)	$x_2=x_3=0$ vertices (2)	$x_1=0$, $b_1=0$ midedges (2)	Other midedges (4)	a.v. centroid (1)
~1	0.073	0.146	0	0.080	0.096
0.9	0.100	0.146	0	0.040	0.141
0.6	0.142	0.132	0	0	0.167
0.3	0.143	0.102	0.044	0	0.137
$^2/_{13}$	0.144	0.089	0.067	0	0.113
0	0.146	0.080	0.080	0	0.096

Figure 3.8 Hexagons which are symmetric under a half turn
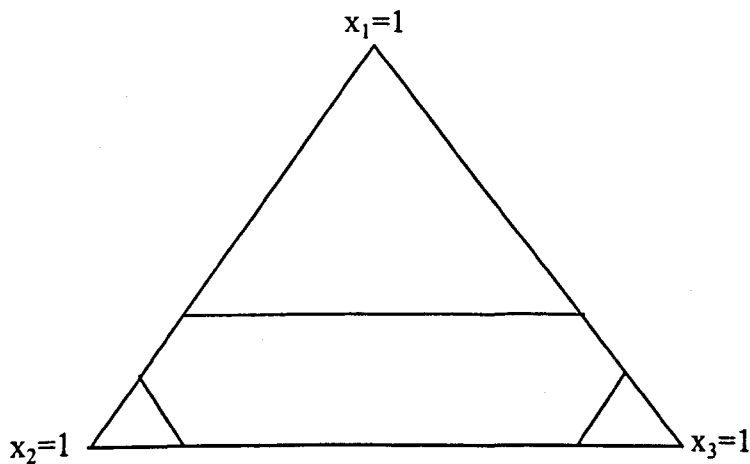
As b_1 decreases and the length of the $x_1=b_1$ and $x_1=0$ edges increases, the weight of their vertices increases considerably between 0.9 and 0.6 but remains approximately constant

after this. The weights of their (short) midedge points remain 0 until $b_1=0.3$ after which they rapidly increase. The other two vertices gradually become closer to the other points and their weights decrease. Between 1 and 0.6 the weights of the long edge midpoints decrease and the weight of the centroid increases. At $b=0.3$ the weights of the (now) long edge midpoints become non-zero and begin to increase, and the weight of the centroid decreases. For the parallelograms at the two extremes the patterns which were observed for the other values of b continue to be observed. This is unusual because for all previous shapes sudden deviations in the patterns were observed.

3.6.9 Hexagons With No Symmetries

There is no obvious approach for systematically modifying the hexagonal regions with no symmetries. Instead, the constrained regions were summarised using edge lengths, excluding the $2/\sqrt{3}$ multiplier (see §3.6.1). The following five sets of upper bounds were used: (0.9, 0.8, 0.7), (0.8, 0.7, 0.6), (0.7, 0.6, 0.5), (0.9, 0.7, 0.5) (0.8, 0.7, 0.4). So that the different constrained regions could be used in the same analysis, the summary lengths were scaled by the total edge length of their region. This scaling was required because if the total edge length were low, some midedges would get a high weighting because the points on the short edges were not weighted.

Correlation coefficients between certain summary lengths (defined below) and the weights were calculated. The designs generated for the above upper bounds collectively included 30 vertices, 8 midedges and 5 centroids as their support points. Though five and eight values are quite small samples on which to base a correlation coefficient, there should be enough to indicate whether a relationship exists. In terms of individual points and in general, the vertices tended to have the highest weights, closely followed by the centroid, and the midpoints tended to have considerably lower weights.

For the vertices, the summary values used were the sum of the lengths of the two edges forming the vertex and the total length of the next edges on either side of the first two. The correlation coefficients were calculated as 0.551 and −0.501. For the midpoint the values used were the length of the edge in question and the total length of the two adjacent edges. The correlation coefficients were calculated as 0.714 and −0.906 The first values in each pair, which relate to the edges a point is on, confirm that weight increases with edge length, but that the relationship is far stronger for midpoints than vertices. The second value in each pair is the total length of the two adjacent edges. From previous results in this section it would

appear that the weight increases as other points move away. This would imply a positive values for the second coefficient in each pair. The negative values observed are probably because long adjacent edges are likely to include midpoints which will have a stronger negative effect on the weight.

Using a 2 sample t-test, the length of the 22 edges whose midpoints were not included, were compared with the lengths of the 8 edges whose midpoints were included, to investigate whether in general the former are likely to be shorter than the latter. A similar test was performed to investigate whether, in general, the total lengths of the 2 edges adjacent to those edges whose midpoints were used, are likely to be higher than the total lengths of the 2 edges adjacent to those edges whose midpoints were not used. For a null hypothesis of no difference against the appropriate one sided alternative hypotheses, p-values of 0, to three and four decimal places respectively, were obtained. These p-values are consistent with the conclusions from the correlation coefficients.

For the centroid, the summary value was the total length of all edges. The correlation coefficient between the weights and this summary measure was -0.632. This coefficient suggests that for regions which use more of the original simplex, and have boundaries which are further from the centroid, the weight of the centroid tends to be lower. This is the opposite effect from what was observed for the other types of point because for the vertices and midedges the weights were found to decrease as other points got closer to them.

3.6.10 Conclusions from the Systematic Variation of Constrained Regions when q=3

From the systematic variation of the constrained region, which is discussed above, the following conclusions can be drawn:

(1) All vertices are always included.

(2) The biggest influence on weight appears to be edge length. As the length of an edge increases the weights of its vertices increase, but at a decreasing rate. When a particular length is reached the weight of the midpoint becomes non zero, but low compared to the vertices, and the weights of the vertices tend to drop.

(3) Another factor which influences weight is the number of weighted points which are nearby. If a particular point has few weighted points nearby, its weight will tend to be

higher. There is a trade off between this effect and edge length, but edge length appears to have considerably more influence.

(4) In contrast to the general relationship between the weight of a point and the distance to nearby points, the weight of the centroid appeared to increase as other points moved towards it.

(5) In most cases conclusions 2 and 3 did not apply as the regions approached parallelograms.

(6) For the trapezium and pentagon the vertices can be split into two types. Those caused by an upper bound removing a vertex from the regular simplex, and those vertices of the regular simplex which remained. These vertices are subsequently referred to as removal and intact vertices. The weight of the removal vertices increases rapidly as b is reduced but the weights of the intact vertices decrease slowly as b is reduced.

(7) The slow change in weight which was observed for the intact vertices appears to be caused by the other vertices moving towards it (see conclusion 3). The rapid increase in the weight observed for the removal vertices appears to be due to the increase in length of the edge they are on (see conclusion 2). This difference between the two types of vertex strengthens the conclusion that the weight change associated with length is greater than the weight change associated with nearby points.

(8) Although the above conclusions show that the weights assigned are highly dependent on the constrained region, some general trends were observed. In terms of individual points the vertices tended to have the highest weight, but were closely followed by the centroid. The weights of the edge midpoints tended to be considerably lower. Two exceptions to this general trend were the trapezium and parallelogram for which the midpoints tended to have higher weights than the centroid.

Many of these conclusions are connected with the distance between points. Similar conclusions would most likely be obtained if a packing criterion was used.

3.7 Designs for the Full Cubic Model, the D Criterion and q=3

3.7.1 Outline

Initially both the special cubic and full cubic were considered. The first two examples considered were the hexagon which is symmetric under a third turn with b=0.8, and the hexagon which is symmetric under a half turn with b=0.3. In the case of the special cubic

model the same support points as the quadratic model were used but the mid edges appeared to be weighted slightly higher. For the full cubic model, Mikaeili type points (see §3.5.3), and axial check points were used. Due to the fact that for these examples there appeared to be little difference between the continuous optimal designs for the quadratic and special cubic models, but considerable difference between the continuous optimal designs for the quadratic and full cubic models, attention was subsequently restricted to the full cubic model.

· The investigation of the full cubic model was considerably smaller than the investigation of the quadratic model because the latter is more commonly used in initial investigations of feasible regions. A subset of examples will be used to investigate how optimal designs for the full cubic model deviate from optimal designs for the quadratic model. The constrained regions which are used were the special cases of the parallelogram, the trapezium, the pentagon and the hexagon which is symmetric about a third turn. Now that the full cubic model was being considered the initial search in stage one of the weight estimation algorithm used the DX4 candidate set plus the Mikaeili type points.

3.7.2 Parallelogram

It was found in Farrel et al (1967) that the continuous optimal design for the cubic model on a square uses the vertices, together with points which are very close to the axial check blends and thirds of edges used by DX4. The weightings of the vertices, near thirds of edges and axial points are 0.092, 0.058 and 0.043. When the exact thirds of edges and axial check blends were used, the continuous optimal design weighted the vertices, thirds of edges and axial check points as (0.0929, 0.0560, 0.0450). Since the optimal design located in Farrel et al (1967) did not use Mikaeili points, they are not investigated here. As this thesis is interested in the standard candidate points rather than perturbations of them, the Farrel et al points are not considered further.

3.7.3 Trapezia

In the same way as was done for the quadratic model (see §3.6.3), the value of $b=b_1$ was reduced from 1 to 0 in steps of 0.2. The weights for the resulting six regions are given in Table 3.9. The following expressions were used to calculate the Mikaeili type points on the left hand side of the feasible region; (b, (1-b).r, 1-b-(1-b).r), where $r = 0.2764$ for the $x_1=b_1$ edge, (0, r, 1-r) for the $x_1=0$ edge, (b.r, 1-b.r, 0) for the lower point (i.e. the one near $x_1=0$) on the $x_3=0$ edge and the same expression with 1-r replacing r for the higher point (i.e. the one

near $x_1=b$) on the $x_3=0$ edge. The Mikaeili type points on the right hand edge were the same as those on the left hand edge but with the x_2 and x_3 components interchanged.

Table 3.9 Estimated weights in the D optimal design for the full cubic model and q=3 for the trapezium with b_1 reduced from 1 to 0 in steps of 0.2 Unless an alternative number is given, each point occurs only once in the constrained region. For b~0 the thirds of edges on the $x_1=b$ and $x_1=0$ edges (2 on each) have weight 0.056. The weights for the $x_2=0$ edge also apply for the $x_3=0$ edge.

b	$x_1=b$ vert's (2)	$x_1=0$ vert's (2)	$x_1=b$ mid'	$x_2=0$ Mik' near $x_1=b$ (2)	$x_2=0$ third near $x_1=b$ (2)	$x_2=0$ Mik' near $x_1=0$ (2)	$x_2=0$ third near $x_1=0$ (2)	$x_1=0$ Mik' (2)	$x_1=b$ axial (2)	$x_1=0$ axial (2)	a.v. cen'd
~1	0.05	0.100	0	0.100	0	0.100	0	0.100	0	0	0.100
0.8	0.08	0.099	0	0.076	0	0.077	0.017	0.100	0	0	0.100
0.6	0.085	0.098	0.027	0.061	0	0	0.081	0.095	0.017	0.019	0.061
0.4	0.090	0.096	0.068	0.035	0.019	0	0.067	0.086	0.034	0.039	0
0.2	0.094	0.094	0.080	0.007	0.049	0	0.059	0.077	0.036	0.044	0
~0	0.093	0.093	0	0	0.056	0	0.056	0	0.045	0.045	0

As b decreases and the distance between the $x_1=b$ vertices increases, their weight increases steadily. The weight of the midpoint on this edge becomes non-zero when b is approximately 0.6.

At b~1 the only edge points used are the Mikaeili type points. As b decreases the length of $x_2=0$ edge and the weight of the Mikaeili type points also decreases. As b continues to reduce, the thirds of edges begin to share the weight originally assigned to the Mikaeili type points and eventually take all of it. The weight of the edge third near to $x_1=0$, i.e. the vertex which remains intact, is the first to become non-zero at approximately b=0.8. The weight is then shared with the Mikaeili type points until the weight of the Mikaeili type point

becomes zero at around b=0.6. A similar thing happens with the Mikaeili type point near to x_1=b, i.e. where the vertex of the regular simplex has been removed, but not until b reaches a lower value. The x_1=b edge is shorter than the x_1=0 edge and the combined weights of the Mikaeili type and edge thirds are slightly lower. As the reductions in b lead to a shortening of the x_2=0 edge, the combined weight of all four edge points decreases but the combined weight of the points near x_1=b remain comparatively lower. In summary it can be said that for the points near to the intact vertex, the combined weight is slightly higher, and the sharing of weight begins earlier than for the points near the removed vertex. In all cases the support points on the x_1=b and x_1=0 edges are either the midpoints or Mikaeili type points. As b changes the weights of these points gradually increase and decrease respectively.

A similar sharing effect occurs with the centroid and axial points. The weight of the centroid gradually reduces with b. The weight of axial points become non-zero around b=0.6, the weight assigned to the internal points is then split between them and the centroid, until the weight of the centroid becomes zero at around b=0.4. Again the weights of the axial points near to the x=b edge are lower than the weights of those near to the x_1=b edge.

3.7.4 Pentagon which is Symmetric in the line x_i = x_j

In a similar way to the analysis of the pentagon for the quadratic model (see §3.6.4), the value of b_2=b_3=b was reduced from 1 to 0.6 in steps of 0.2. The weights for these three pentagons are given in Table 3.10 along with the weight for b=0.5 where the results for the parallelogram apply (see §3.7.1). The two lower Mikaeili type points on each of the right hand edges were calculated using the following expressions ([1-b].r_1, [1-b].r_2, b), ([1-b]+b.r_1, 0, b.r_2) where r_1 = 0.2764 and r_2 =1− r_1. The other two points on the right hand edges were obtained using the same expressions with r_1 and r_2 interchanged. The four left hand points were obtained by taking the equations for the first four points and interchanging the x_2 and x_3 components.

The weight of the x_1=1 vertex appears to decrease very slowly as the other points move towards it. The weight of the x_2=b vertices (referred to as x_2=0 and x_1=0 in Table 3.10) generally increase as b decreases and the distance between them increases .

In a similar way to the trapezia the weights of edge points on the longer edges gradually shift from the Mikaeili type points to the thirds of edges as the edges become shorter. Again, with the trapezium the switch to thirds of edges occurs first for the points near the intact vertices.

Table 3.10 Estimated weights in the D optimal design for the full cubic model and q=3 for the pentagon with $b_2 = b_3 = b$ reduced from 1 to 0.5

In this table all values quoted for edge points on $x_2=0$, are identical to the values for similar edge points on $x_3=0$. Unless an alternative number is given, each point occur once in the constrained region. At $b \sim 1$ the 2 Mikaeili type points on the line $x_1=0$ receive a weight of 0.1. At $b \sim 0.5$ the Mikaeili type points on $x_2=0$ and $x_3=0$ (2 on each) receive a weight of 0.056 each.

b	$x_1=1$ vert	$x_2=0$, $x_3=0$ vert's (2)	$x_1=0$ vert's (2)	$x_2=0$ Mik' nr $x_1=1$ (2)	$x_2=0$ third nr $x_1=1$ (2)	$x_2=0$ Mik' nr $x_2=b$, $x_3=b$ (2)	$x_2=0$ third nr $x_2=b$, $x_3=b$ (2)	$x_1=0$ mid $x_2=b$ mid (2)	$x_1=1$ axial	$x_2=0$ $x_3=b$ axial (2)	$x_1=0$ $x_3=b$ axial (2)	a.v cen'd
1	0.100	0.050	0.050	0.100	0	0.100	0	0	0	0	0	0.100
0.8	0.099	0.086	0.090	0	0.087	0.050	0	0.083	0.027	0.069	0	0.028
0.6	0.096	0.092	0.089	0	0.06	0.009	0.056	**0.038**	0.050	0.054	0.029	0
0.5	0.093	0.093	0.047	0	0.056	0	0.056	0	0.045	0.045	0.022	0

The combined weights of the edge points again decrease, but at b=0.8 the points near to $x_1=1$ are weighted more highly and at b=0.6 the points near to $x_3=b_3$ are higher. The weight of the $x_1=0$ midpoint briefly becomes non-zero when b is approximately 0.8, this weight is then shifted to the $x_2=b$ midpoints at b=0.6 and finally to the $x_2=b$ thirds of edges as the region approaches a parallelogram.

The weights of the top and middle axial points become non-zero when b is approximately 0.8. At b=0.6 the weight of the centroid has become zero and the weight of the top axial point has increased. The distance between the bottom two axial points has increased, and their weight become non-zero, but the weight of the middle axial point has reduced. When the region becomes a parallelogram the weight of the axial point near the $x_1=1$ vertex goes against these trends and suddenly decreases, but the weight of the other two axial points continues to decrease.

The weight of the centroid rapidly decreases, and from its weight at b=0.8 it seems likely to have become zero shortly afterwards.

3.7.5 Hexagons Symmetric Under a Third Turn

In the same way as for the quadratic model (see §3.6.6), $b_i=b$ was gradually reduced from 1 to $^2/_3$. The estimated weights of the optimal design over the usual candidate set are given in Table 3.11. Recall that all regions with $b<^2/_3$ are similar to one of the regions with $b>^2/_3$ and the same optimal design will apply. The Mikaeili type points are all three permutations of each of the following two points: $([1-b]+[2b-1].r, 0, [1-b]+[2b-1].[1-r])$ and $([1-b].r, [1-b].[1-r], b)$.

As b is varied, in a similar way to the trapezium and pentagon, the weight of the edge points shifts away from the Mikaeili type points as the edges get shorter.

Table 3.11 Estimated weights in D optimal designs for the full cubic model and q=3 for the hexagon symmetric about a third turn, with b reduced from 1 to $^2/_3$.

b	Vertices (6)	long mid edges (3)	short mid edges (3)	Mikaeili type (6)	thirds (6)	axial (6)	a.v. centroid (1)
~1	0.05	0	0	0.100	0	0	0.100
0.9	0.064	0	0	0.083	0	0.0047	0.09
0.85	0.079	0	0	0.011	0.043	0.027	0.035
0.8	0.091	0.073	0	0	0	0.039	0
0.7	0.089	0.055	0.014	0	0	0.043	0
$^2/_3$	0.087	0.036	0.036	0	0	0.043	0

The weights of the vertices do not appear to follow any definite trend, apart from being greater than all of the other weights in all cases except for the regular simplex (where all points have an equal weighting) and b=0.9. By considering intermediate values of b it was found that the weight of the axial points becomes non-zero at approximately 0.91. The weight of the axial points then increases at a gradually reducing rate.

From the table it appears that the weight assigned to the edge points, shifts from long edge Mikaeili type points to long edge thirds, to long edge midpoints. By considering intermediate values of b it was found that the progression was slightly more complicated than this. Thirds of long edges obtain non-zero weighting at approximately 0.87, the weight is then shared with the Mikaeili type points until approximately 0.83. At 0.83 the weight of the Mikaeili type points becomes zero but the weight of the long edge mid points becomes non-zero. Between 0.83 and 0.81 the weight is shared between the thirds and midpoints of long edges, at b=0.8 the weight of the thirds becomes zero. The long mid edges are then the only edge points to be used until the weight of the short edge mid points becomes non zero at approximately 0.7. After this all midedges are weighted until some point for $b < {}^2/_3$.

The weight of the centroid gradually decreases until it becomes zero at approximately 0.82. From b=0.7, where the shape is rapidly approaching a regular hexagon, the weights of the vertices and axial points remain fixed while the weights of the long and short edge mid points equalise.

3.7.6 Conclusions for the Full Cubic Model and Comparison with the Optimal Design for the Quadratic Model

A Conclusions

The support points used in the optimal designs for the full cubic model are the vertices, Mikaeili type edge points, edge thirds, axial check points and centroid. As the edge lengths decrease the weight shifts from the Mikaeili type points to the thirds of edges for the trapezium and pentagon, and from Mikaeili type points to thirds of edges then to edge mid points with hexagons which are symmetric about a third turn. In the intermediate stages the weight is sometimes shared between the points involved in the shift. It is hypothesised here that if the points were allowed to range freely over an edge they would gradually move inwards as the edge length decreases. As b decreases and less of the regular simplex is used the weight shifts from the centroid to the axial points. In a similar way to the regular simplex (see §3.2.3), any exact designs generated by DX4 will use thirds of edges instead of Mikaeili type points and thus be inefficient.

The first seven conclusions obtained for the quadratic model also apply to the four examples considered here, with the edge mid points complemented with thirds of edges and Mikaeili type points and the centroid complemented by axial points. However different general trends were observed in terms of the weight assigned to individual points. In all cases

the vertices tended to have the highest weights. For the parallelogram and the hexagon the edge points tended to have the next highest weights and the internal points tended to have the lowest weights. For the trapezium and pentagon the internal points tended to have higher weights than the edge points. This shows that the weights for the trapezium and pentagon again behave differently from the weights for other shapes (see §3.6.10).

B Comparison with the quadratic model

We first consider the parallelogram. The vertex weights used by the full cubic model are considerably lower than those used by the quadratic model. The centroid and axial check blends have a greater combined weight than the centroid in the quadratic model. Recall that the axial check blends are half way between the centroid and the vertices. It seems likely that some of the weight assigned to the vertices under the quadratic model has been assigned to these axial check blends under the full cubic model. Each pair of adjacent edge thirds and Mikaeili type points have a greater combined weight than the midpoints used for the quadratic model.

For the hexagon which is reflection symmetric about a third turn, the weight of the vertices is considerably lower than it was for the quadratic model, but reduces at a similar rate. For the full cubic model the centroid weight starts high but then decreases rapidly when the axial points attain non-zero weights. For the quadratic model, the centroid weight starts low and then increases. The two models therefore induce completely different trends in the weight of the centroid. The points on the long edges start lower and become non-zero quicker than the corresponding midpoints under the quadratic model. However, the general trend in edge point weights is in the same direction. For the full cubic model the short edge mid points become non-zero as b approaches $^2/_3$. These points never attained non-zero weight for the quadratic model.

The comparisons for the trapezium and pentagon are very similar. For both shapes the vertex weights are again lower for the full cubic model, but the proportional change in weight is then similar to the trends observed for the quadratic model. The trends in the weight of the centroid are again completely opposite for the two models. For the full cubic model the centroid weight starts high and gradually decreases to 0 with b. For the quadratic model the centroid weight starts low and gradually increases to the weight which it took on the parallelogram.

114

Additionally for the trapezium, the decrease in the weight of the centroid does not commence until the weights of the axial points become non-zero. This delay may also occur for the pentagon, but b was changed too rapidly to tell (i.e. from 1 to 0.8 in 1 step). The inclusion of axial points is hence a possible reason for the radical difference in centroid weight.

We now consider the edge points. For the trapezium, the weight of the mid point on the x_i=b edge becomes non-zero earlier than it did for the quadratic model. For the trapezium, the combined weight of all points on the x_1=0 edge starts at a slightly higher value than it did for the quadratic model, but then decreases more slowly to a lower weight than that of the single edge point observed under the quadratic model. For the pentagon the weights of the points on x_2=0 and x_3=0 edges, start slightly higher and finish slightly lower than the values observed for the quadratic.

In general the trends for the vertices and edge points are similar to the trends observed for the quadratic model. The only differences are that the weights change at slightly different rates and between slightly different values. For the internal points however the trends in the weights as b is reduced are completely opposite to those observed for the quadratic model. In the quadratic model the weight gradually increases from zero but for the for full cubic model the weights rapidly decrease to zero.

3.8 Designs for the Quadratic Model, the D Criterion and q=4

3.8.1 Introduction

When q=4 the regular simplex is a triangular based pyramid which is fairly easy to represent on a three dimensional plot. The feasible regions considered were limited to the four simple cases which were considered for the full cubic model with q=3 (see §3.7). For a four dimensional L transformed constrained region to be consistent, the sum of each combination of three upper bounds must be at least 1. The four dimensional extensions of the trapezium, pentagon, and hexagon symmetric about a third turn, occur when apices of equal size are removed from one, two or four vertices respectively. Each of these cases is studied.

3.8.2 One Apex Removed

For simplicity it is assumed that b_1=b≤1 and b_2=b_3=b_4=1. In a similar way to the previous examples b is reduced from 1 to 0 in steps of 0.2. An example of a constrained

region of this type is given in Figure 3.9 and estimated weights for the six regions are given in Table 3.12.

Figure 3.9 Constrained region for q=4 with one apex removed (i.e. only one component bounded)
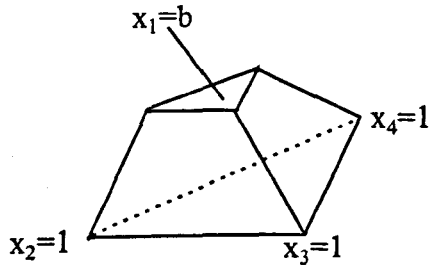
Table 3.12 Estimated weights in the D optimal design for the full cubic model and q=4 with one component having an upper bound of b, with b being reduced from 1 to 0 in steps of 0.2

b	$x_1=b$ vertices (3)	$x_1=0$ vertices (3)	top edge midpoints (3)	side edge midpoints (3)	bottom edge (3) midpoints	a.v. centroid (1)
~1	$^1/_{30}$	0.100	0	0.100	0.100	0
0.8	0.047	0.099	0	0.087	0.100	0
0.6	0.076	0.097	0	0.061	0.100	0
0.4	0.089	0.095	0	0.045	0.099	0.020
0.2	0.086	0.090	0.014	0.039	0.091	0.038
~0	0.080	0.080	0.060	0.033	0.061	0.059

As b decreases the length of the top edges increases, the length of the bottom edges remains the same and the top gets closer to the bottom. As these changes occur the weights of the vertices and edge mid points on the bottom reduce by a comparatively small amount, the weights of the side edge mid points reduce by considerably more but at a decreasing rate and

116

the weights of the top vertices increase at a decreasing rate until approximately b=0.2. At b=0.2 the weight of the midpoints on the to edges becomes non-zero and the weight of the top vertices dips slightly. The weight of the centroid remains zero until b=0.4 but then increases. As b approaches zero the changes in all weights, apart from those for the vertices on the top, continue to change in the same direction, but at an accelerated rate. The weights appear to converge towards a point where they are symmetrical about a plane drawn through the midpoints of the three side edges.

3.8.3 Two Apices Removed

For simplicity it was assumed that $b_1=b_3=b \leq 1$ and $b_2=b_4=1$. When b reaches 1/2 a major change in shape occurs. For $\frac{1}{2} < b < 1$ the faces of the region include two triangles, two trapezia and two pentagons. When $0 < b < \frac{1}{2}$ the faces of the region are all quadrilaterals (4 of which being trapezia). When $b=\frac{1}{2}$ two of the faces which were quadrilaterals when $0 < b < \frac{1}{2}$ become triangles and there is only one B type vertex (see Figure 3.10). The two shapes displayed by the region in these two ranges are given in Figure 3.10. When the shape of the region changes at $b= \frac{1}{2}$, the vertices on $x_2=x_4=0$ are replaced by vertices on $x_1=x_3$ and the midpoint, $(\frac{1}{2}, 0, \frac{1}{2}, 0)$, on the former edge is replaced by a midpoint on the new edge at $(b, \frac{1}{2}-b, b, \frac{1}{2}-b)$.

Figure 3.10 Constrained region for q=4 with two components bounded

a) $\frac{1}{2} < b < 1$

b) $0 < b < \frac{1}{2}$
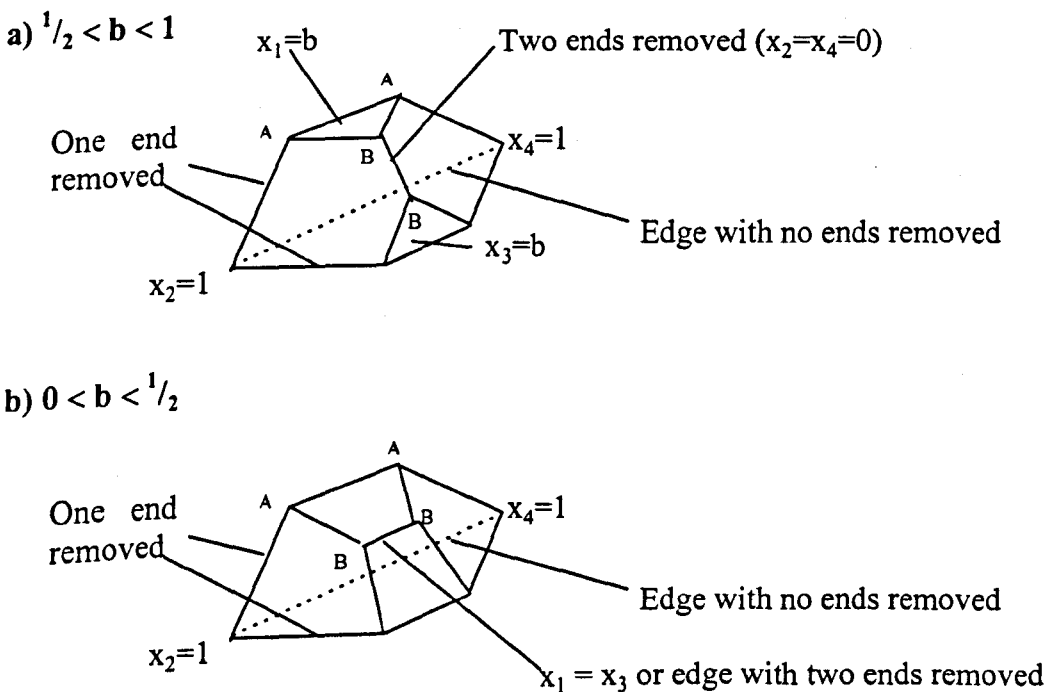
As with previous examples b was reduced from 1 to approximately zero in steps of size 0.2 or 0.1. The estimated weights for these regions are given in Table 3.13.

Table 3.13 Estimated weights in the D optimal design for the quadratic model and q=4 with two components with upper bounds of b, where b is reduced from 1 to 0 in small steps of 0.2 or 0.1.

Bold type is used to indicate that for one region there was only one B type vertex.

b	Intact vertices (2)	A. vertex where 1 end removed (4)	B. vertices where 2 ends removed (2) or (1)	midpoint of intact edge (1)	Mid bet' intact and A type (4)	Mid bet' B type (1)	mid bet' A type vertices on new face (2)	mid bet' A/B type vertices on new face (4)	a.v. centroid (1)
~1	0.100	0.05	0	0.100	0.100	0.100	0	0	0
0.8	0.098	0.051	0.039	0.100	0.087	0.074	0	0	0
0.6	0.092	0.080	0.072	0.099	0.052	0	0	0	0.048
0.5	0.088	0.082	**0.085**	0.097	0.043	0	0	0.024	0.047
0.4	0.085	0.080	0.047	0.097	0.030	0	0	0.038	0.050
0.2	0.070	0.068	0.066	0.078	0.038	0	0.019	0.031	0.061
~0	0.064	0.064	0.064	0.035	0.035	0.035	0.035	0.035	0.066

From Table 3.13 it can be seen that in all cases for the midpoints, and in the majority of cases for the vertices, the weights are lower for those edges which have had more of their ends removed. As b decreases the length of the intact edge remains the same but the other points move towards it. As points move towards the intact edge the weights of its vertices and midpoint on this edge reduce. This reduction is slow at first then becomes more rapid. The patterns observed for the remaining points are far more complicated. The weight of the A

type vertices increases as b decreases until it appears to reaches a maximum when b=1/2, and then begins to decrease again

The first edge considered is the one which has had two ends removed. As b decreases it appears that the weight of the mid point on this edge decreases so rapidly that, despite the fact that the length of the edge decreases, the weight of its vertices increases. At around the same time as the weight of this midpoint has become zero, the weight of the mid edge points between vertices A and B becomes non-zero and the weight of the A and B vertices decreases as expected. After this the weights of the midpoints between the A and B type vertices fluctuate as other factors intervene. When b=0.2 and the sides created by the apices are very big, the weight of the midpoint of the edges between the A vertices becomes non-zero.

As b decreases the two faces caused by the apices get closer together and the region becomes more squashed. As this happens the length of the edges between the intact and type A vertices decreases, and the weights of their midpoints decreases until b=0.4. By b=0.2 the weight of the A type vertices suddenly drops and the weights of the midpoints on the line between these and the intact vertices increase again. The length of the edge between the two type B vertices decreases at twice the rate of the edges with one end removed (i.e. the edge between the B type and intact vertices) and as mentioned above, the weight of its midpoint quickly becomes zero at b=0.6.

The weight of the centroid becomes non-zero at b=0.6 and then increases slowly as the shape becomes more squashed. When b approaches zero, the region is approximately symmetrical about a line through the two intact vertices and the midpoint of the edge with two ends missing. At this point the weights shift in accordance with this symmetry and some of the previous weight patterns are no longer observed.

3.8.4 The Octahedron with Two Sets of Four Identical Sides

The q=4 extension of the hexagon which is symmetric about a third turn, occurs when four apices of equal size are removed from the four $x_i=1$ vertices. The removal of these apices results in an octahedron with two sets of four identical sides. Four of these sides are triangles and four are hexagons. When b=1/2 the region is the regular octahedron in which all the faces are triangles and there are only six vertices. When b reaches its lower limit of $1/3$, caused by the requirement for consistency (see §3.8.1), the region becomes the regular simplex again and the known optimal design applies. To investigate a selection of regions within these

bounds the value of b was reduced from 1 to 0.4 in steps of 0.1, and the weights of the optimal designs were estimated.

When the region is very close to the regular octahedron the weights are likely to take unusual values. Weights were calculated for four additional values of b in this region. The weights for all types of point are given in Table 3.14 and are plotted in Figure 3.11. Because the region only has six vertices when b=$\frac{1}{2}$, the weights of the design for this region do not fit in with the general trends. The weights for this region were omitted from the plot so that the general trend could be compared more easily. The weight of the centroid is zero at this point and this information is also missing from the plot

Table 3.14 Weights of the D optimal designs for the octahedron with two sets of four identical sides as b is reduced from 1 to $\frac{1}{3}$

Bold type is used to indicate that certain points occurred less frequently for certain values of b.

b	Vertices (12) or (6)	Long edge midpoints (12) or (6)	Centroid (1)	b	Vertices (12) or (6)	Long edge midpoints (12) or (6)	Centroid (1)
1	0.0333	**0.100**	0	**0.5**	**0.081**	0.043	0
0.9	0.037	**0.093**	0	0.49	0.040	0.042	0.009
0.8	0.047	**0.065**	0.045	0.48	0.041	0.040	0.022
0.7	0.075	0	0.100	0.47	0.043	0.037	0.036
0.6	0.075	0	0.100	0.45	0.053	0.026	0.065
0.575	0.059	0.018	0.077	0.425	0.075	0	0.100
0.55	0.048	0.031	0.052	0.4	0.075	0	0.100
0.533	0.043	0.037	0.035	0.367	0.050	**0.057**	0.056
0.5	**0.081**	0.043	0	1/3	0.0333	**0.100**	0

From Figure 3.11 it can be seen that the patterns in the weights for $\frac{1}{2} \leq b \leq 1$ is repeated at a faster rate for $1/3 \leq b \leq 1/2$. The patterns are repeated because as b moves between 0.5 and 1/3, the regions which occur, and hence the optimal designs, are the same as those which occur as b is increased from 0.5 and 1. A similar result was observed for the hexagon (see §3.6.6). The relationship between the weights of the optimal designs in the two ranges of b is non linear. Due to the width of the two ranges the rate of change for $\frac{1}{2} \leq b \leq \frac{1}{3}$ is on average three times the rate of change between $1 \leq b \leq \frac{1}{2}$. As b moves towards $\frac{1}{3}$ the differences in the rates of change become increasingly great. For b near to 1/3, the rate of change is considerably more than three times the rate that it is for b near to 1.

Between 1 and 0.7 the weight of the vertices increases at an increasing rate and the weights of the long edge midpoints decrease at an increasing rate. At 0.7 the weight of the long edge midpoints become zero and the weight of the vertices reaches a maximum. Recall that when b is $\frac{1}{2}$ the region is the regular octahedron. As the regular octahedron is approached some interesting behaviour is observed. When b is approximately 0.6 the weight of the midedges becomes non-zero and begins to increase and the weight of the vertices begins to decrease. The weights of the midedges and vertices continue to increase and decrease, at gradually reducing but approximately equal rates, until the regular octahedron is reached.

By considering intermediate values of b it was found that the weight of the centroid first becomes non zero at approximately b=0.85. The weight of the centroid appears to increase as the weight of the (long) midedges decrease and then decrease as the weight of the (now long) midedges become non-negative and increase again. The weight of the centroid varies between 0 and 0.1 and has a higher maximum weight than the vertices. For b approximately in the range (0.6, 0.75) the weight of the midedges is zero and the weights of the vertices and centroid remain constant. This is a similar effect to that which was observed for q=3 (see §3.6.6). Another interesting effect which is observed is that for all values of b < 0.9 but not close to 1/2 or 1/3, the changes in the weight of the centroid appears to mirror the changes in the weight of the midedges.

Figure 3.11 Weights of the D optimal designs for the vertices and mid edge points of the octahedron with two sets of four identical sides as b is reduced from 1 to $^1/_3$.
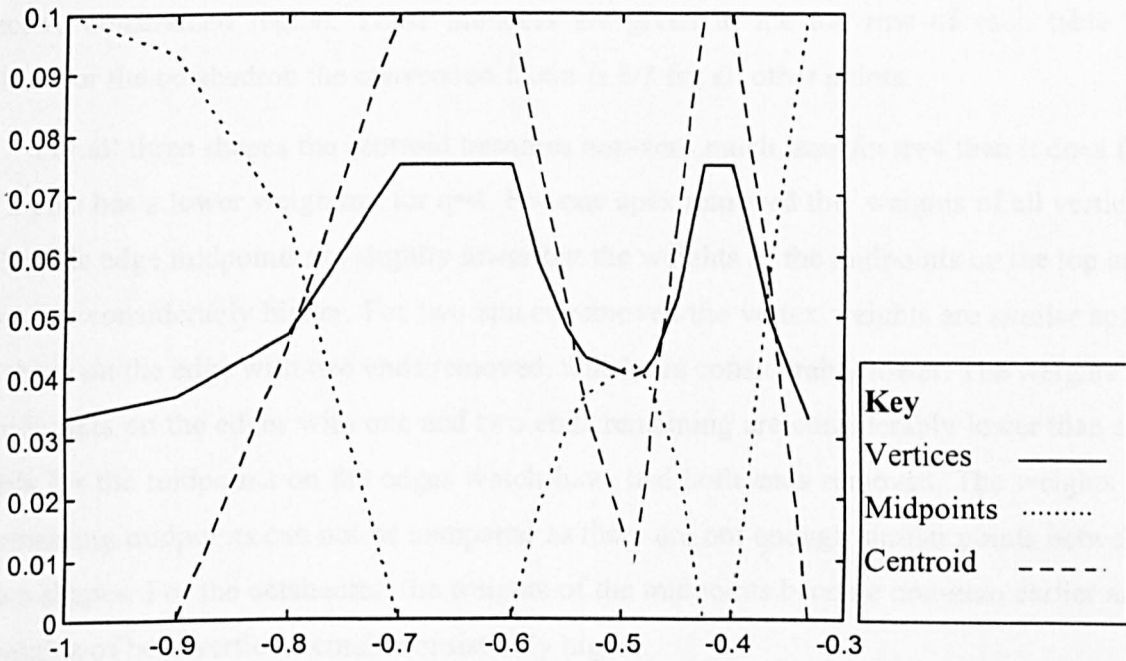
3.8.5 Conclusions on Optimal Designs for the Quadratic Model and q=4 and Comparison with the Same Model when q=3

For the cases where one or two apices are removed, the first seven conclusions obtained for the q=3 quadratic model were again observed here (see §3.6.10). For the octahedron the behaviour of the weights was considerably more complicated and few conclusions can be drawn. General trends over all b were considerably less clear but the vertices again tended to have higher weights in all cases. For the regions which are generated by the removal of two or four apices the centroid tended to have higher weights than the mid edge points but the reverse was true for the regions generated when only one apex was removed.

The weights for q=4 are now compared to those for q=3. As there are more of each type of point when q=4 the weights of individual points will automatically be lower. To account for this a conversion factor is used to convert the weights for q=4, into a value which can be compared with the weights for q=3. Where applicable the factor used was [# occurrences for q=4]/ [# occurrences for q=3].

122

No conversion factor is required for the centroid as in all cases it only occurs once. For one and two apices removed the conversion factors for those points which can be compared are easily calculated from the numbers of each type of point which occurs in the respective constrained region. These numbers are given in the top row of each table of weights. For the octahedron the conversion factor is 8/3 for all other points.

For all three shapes the centroid becomes non-zero much later for q=4 than it does for q=3, it also has a lower weighting for q=4. For one apex removed the weights of all vertices and the side edge midpoints are slightly lower but the weights of the midpoints on the top and bottom are considerably higher. For two apices removed the vertex weights are similar apart from those on the edge with two ends removed, which are considerably lower. The weights of the midpoints on the edges with one and two ends remaining are considerably lower than the weights for the midpoints on the edges which have had both ends removed. The weights of the remaining midpoints can not be compared as there are not enough similar points between the two shapes. For the octahedron the weights of the midpoints become non-zero earlier and the weights of both vertices remain consistently higher.

3.9 V Optimal Designs for the Quadratic Model and q=3

The result that G optimality implies D optimality and the algorithmic method which was derived from it (see §3.5.2) can clearly not be used here. Instead of using this result the weights for the V optimal designs were estimated using a more rigorous application of stage 1 of the algorithmic procedure (see §3.5.2). This use of stage 1 involves generating a 100 point design and estimating the weights by dividing the replicate numbers by 100. In some cases designs with n= 101 or n=102 were generated, to ensure that two points which are each other's reflection in a line of symmetry had the same weights (see §3.9.1). In these cases the numbers of replicates were divided by the appropriate n As with the previous two sections the analysis of this section is limited to the special cases of the parallelogram, trapezium, regular pentagon and hexagon which is symmetric about a third turn. The weight estimates obtained in this section will clearly not be as accurate as those for the D optimal designs but will be accurate enough to indicate trends and allow comparison with D optimal designs.

3.9.1 V Optimal Design for the Parallelogram and q=3

In the same way as for the previous two q=3 examples, the V optimal design on the square also applies for the parallelogram. The initial results, based on a 100 point design,

suggested that the midpoints and vertices had equal weighting. An attempt to verify this was made by generating a 200 point design. This was shown to be false when a 200 point design was used. In the 200 point design three of the midpoints were replicated 19 times and one was only replicated 18 times. If the weights were estimated from these figures then different weights would be estimated for those two points which are each other's reflection in a line of symmetry. In an attempt to remedy this uneven replication a 201 point design was generated. Using the 201 point design the weights of the vertices, midpoints and centroid were estimated to be (0.089, 0.094, 0.271).

3.9.2 V Optimal Design for the Trapezium and q=3

As was the case when the D optimality and q=3, it was assumed that $b_2 = b_3 = 1$ and $b_2 = b$ was gradually reduced from 1 to 0 in steps of 0.2. The estimated weights for these six cases are given in Table 3.15.

Table 3.15 shows that as b decreases and the length of the $x_1 = b$ edge increases, the weight of its vertices increases at a gradually reducing rate. When b reaches 0.2 and the length of the $x_1 = b$ edge becomes comparatively large, the weight of its midpoint becomes non-zero and the weight of its vertices decreases slightly. The decrease in b also leads to a decrease in the length of the $x_2 = 0$ and $x_3 = 0$ edges. As this happens the weight of their midpoints decreases rapidly at first but then at a rapidly decreasing rate.

As b decreases the length of the $x_1 = 0$ edge remains the same but the other points move towards it. As this happens the weight of its vertices initially reduces slightly, then remains approximately constant. However, unlike for D optimality the weight of the $x_1 = 0$ midpoint reduces rapidly and at an increasing rate. The weight of the centroid increases at a rapid rate as b decreases and then drops slightly as the parallelogram is reached. When b approaches zero the weights of the parallelogram apply and many of the previously exhibited trends no longer apply.

3.9.3 V Optimal Design for the Pentagon and q=3

As was the case with D optimality for q=3 it was assumed that $b_1 = 1$ and $b_2 = b_3 = b$ was gradually reduced from 1 to 0.5 in steps of 0.1. The estimated weights for these six cases are given in Table 3.16.

Table 3.15 Estimated weights in the V optimal design for trapezium using the quadratic model and q=3, with b_1 reduced from 1 to 0 in steps of 0.2

b_1	$x_1=b_1$ vertices (2)	$x_1=0$ vertices (2)	$x_1=b_1$ midedge (1)	$x_2=x_3=0$ midedges (2)	$x_1=0$ midedge (1)	a.v. centroid (1)
~1	0.050	0.100	0	0.202	0.202	0.095
0.8	0.078	0.088	0	0.157	0.196	0.157
0.6	0.098	0.088	0	0.118	0.167	0.225
0.4	0.109	0.089	0	0.099	0.149	0.257
0.2	0.099	0.089	0.059	0.089	0.119	0.277
~0	0.089	0.089	0.094	0.094	0.094	0.271

Table 3.16 shows that as b decreases and the lengths of the $x_i=b$ edges increase, the weight of their vertices increased with no obvious pattern. Although many similarities were observed the $x_2=0$ and $x_3=0$ edge vertices attain greater weights than the $x_1=0$ vertices. As b decreases the majority of the other points move towards the $x_1=1$ vertex. As this happens its weight reduces slightly initially and then remains approximately constant. When b reaches 0.6 and the length of the $x_i=b$ edges becomes comparatively large, the weight of its midpoint becomes non-zero and the weight of the vertices decreases again slightly. The decrease in b also leads to a decrease in the lengths of the $x_2=0$ and $x_3=0$ edges, as this happens the weight of their midpoint decreases with no obvious pattern.

The length of the $x_1=0$ edge decreases faster than any of the other edges, the weight of its midpoint rapidly becomes zero by b=0.7. The weight of the centroid increases at a rapid rate with non-uniform trend, but slows down towards the end. When b approaches zero the weights of the parallelogram apply and many of the previously exhibited trends no longer apply.

Table 3.16 Estimated weights in the V optimal design for the pentagon using the quadratic model and q=3, with b=b$_2$, b$_3$ reduced from 1 to 0 in steps of 0.2 or 0.1

b	$x_1=1$ vertex (1)	$x_2=x_3=0$ vertices (2)	$x_1=0$ vertices (2)	$x_2=x_3=0$ midedges (2)	$x_2=x_3=b$ midedges (2)	$x_1=0$ midedge (1)	a.v. centroid (1)
~1	0.100	0.050	0.050	0.202	0	0.202	0.095
0.9	0.089	0.059	0.059	0.188	0	0.139	0.158
0.8	0.089	0.069	0.079	0.149	0	0.099	0.208
0.7	0.089	0.109	0.089	0.119	0	0	0.277
0.6	0.089	0.099	0.099	0.099	0.009	0	0.297
~0.5	0.089	0.089	0.045	0.094	0.094	0	0.271

3.9.4 V Optimal Design for the Hexagon Symmetric About a Third Turn when q=3

In the same way as for D optimality, the weight changes for values of b between 1 and $^2/_3$ will be repeated in the opposite order at twice the rate for b between $^2/_3$ and 0.5. Table 3.17 gives the estimated weights for b=$^2/_3$ and b decreasing from 1 to 0.7 in steps of various sizes.

Table 3.17 shows that as b decreases and the lengths of the long edges decrease, the midpoint weight gradually reduces to zero. By considering intermediate values of b it was found that the midpoint weight becomes zero for b very close to 0.7. As b decreases and the distance between the vertices increases, the weights of the vertices and centroid increase steadily and reach a plateau when b is approximately 0.7. This plateau appears to start when the weight of the midpoints becomes zero. This behaviour is similar to that which was observed for the D criterion in this region.

Table 3.17 Estimated weights in the V optimal design for the hexagon symmetric about a third turn using the quadratic model and q=3, with b reduced from 1 to $^2/_3$

b	Vertices (6)	Midpoint (long edges) (3)	Centroid (1)
~1	0.050	0.202	0.095
0.9	0.061	0.151	0.182
0.8	0.071	0.102	0.265
0.7	0.110	0	0.340
$^2/_3$	0.110	0	0.340

3.9.5 Conclusions on V Optimal Designs for the Quadratic Model and q=3 and Comparisons with the D Optimal Designs.

In general the first seven conclusions on D optimality were also observed here (see §3.6.9). Contrary to conclusion eight however the centroid tended to have a considerably higher weight than the vertices and midpoints and the midpoints tended to have slightly higher weight than the vertices. This was the main difference between V and D optimal designs. In some cases the steps, by which the weights varied as b decreased, displayed less regular patterns than for the D criterion.

Comparing individual points for the parallelogram with the D criterion, the vertex weights tended to be approximately one third lower and the midpoint weights tended to be slightly higher than for the D criterion. For the trapezium and pentagon the weights of the intact vertices were halved and the weights of the other vertices were reduced by a third. The weight of the midpoint on the edges replacing the vertices became non-zero earlier and finished slightly higher. The weights of the other midpoints varied at similar rates but between slightly higher values. For the first three shapes the centroid was weighted from the start, rather than starting at zero as for the D criterion, and then rose to be almost three times higher than the maximum value attained under the D criterion.

For the hexagon which is symmetric about a third turn the comparison for the vertices was the same as for the parallelogram, the centroid weight only finished at about twice the size as for the D criterion and the weight of the midpoints did not become zero until later.

3.10 General Conclusions for Constrained Regions

For all examples considered every vertex was included in the optimal design. As a wide range of constrained regions were systematically investigated it is conjectured that optimal designs of the four types considered will always include every vertex. The weight assigned to the vertices varied between the quadratic model design types in the following descending order: D optimal q=3, D optimal q=4, and V optimal q=3.

For the quadratic model the remaining points in the optimal designs were chosen from the edge midpoints and the centroid, at least one additional point was chosen from one of these categories in each optimal design. The weight assigned to midpoints varied between the design types in the following descending order; D optimal q=4, V optimal q=3, and D optimal q=3. The longer an edge was the more likely it was that its midpoint would be included in the optimal design. The edge length required for the inclusion of midpoints appeared to decrease between the design types in the same order as the tendency to assign weight. The weight assigned to the vertices and midpoints appears to decrease with the number of support points near by.

In many cases, but by no means more the further a constrained region was from the full simplex, i.e. the smaller the values used for the upper bounds, the higher the weight assigned to the centroid. In general the weight assigned to the centroid varied between the quadratic model design types in the following descending order; V optimal q=3, D optimal q=3 and D optimal q=4.

For D optimality and q=3, the points used in descending weight order, were the vertices, centroid and midpoint, though the midpoint and centroid were in the opposite order for the pentagon and hexagon. For V optimality and q=3 the points in descending weight order were the centroid, midpoints and vertices in all cases. This is the only case where the centroid had the highest weight. The fact that the centroid weight was considerably higher for V optimality is the main difference from D optimality. For D optimality and q=4 the vertices tended to have the higher weights in all three cases. For the pentagon and hexagon the

centroid tended to have higher weights than the mid edge points but the reverse was true for the trapezium.

For the full cubic model the optimal design uses the vertices, edge thirds, Mikaeili type edge points, axial check points and centroid. This shows that the Mikaeili type points are used in constrained mixture examples as well as for the regular simplex. This was the main conclusion obtained for the full cubic model. If additional time were available it would be interesting to investigate whether further improvements in efficiency could be obtained by using alternative edge points. As b decreases and less of the regular simplex is used the weight shifts from the centroid to the axial points. As the edge lengths decrease, the weight shifts from the Mikaeili type points to the thirds of edges for all shapes. For hexagons which are symmetric about a third turn a final shift to midpoints only was observed. During the intermediate stages the weight is often shared between the points involved in the shift.

In all cases the vertices tended to have the highest weights. For the parallelogram and the hexagon the edge points tended to have the next highest weights and the internal points tended to have the lowest weights. For the trapezium and pentagon the internal points tended to have higher weights than the edge points. Again the weights for the trapezium and pentagon are different from the weights for the parallelogram and hexagon, the shapes can therefore be partitioned into two sets. The weight order for the centroid and midpoints under the full cubic model is exactly opposite to the weight order for the quadratic model for both sets of shapes in the partition discussed above.

None of the optimal designs found for constrained regions were saturated. Recall also that for the regular simplexes discussed at the start of this chapter the optimal designs were saturated.

When a constrained region is similar to one where the optimal design over the standard candidate set is known, a similar design is often efficient or even optimal for the new region. When the only difference is that the vertices of the known region are expanded into edges, in it was found that the weight of the appropriate vertex in the original shape was split evenly between the new vertices. It is likely that this generalisation will be applicable to further examples, including ones where the new edges are reasonably large, with little loss in efficiency.

It is conjectured here that in a similar way to the regular simplex, optimal exact designs for constrained regions will tend to use the same points as the optimal continuous

designs. It is also conjectured that the efficiency of exact designs for constrained regions will be higher if n is such that the proportion of design points at each support point can be close to the weights in the continuous optimal design.

Chapter 4

Techniques for the Design of Mixture Experiments Which Require at Least r Support Points Included at Least Twice

4.1 Preliminaries

4.1.1 Introduction

In many of the examples considered in chapters 2 and 3, the points in the optimal discrete designs were all distinct from each other. In chapter 2, some of the qualities of well designed experiments were discussed (see §2.3.3). A well designed experiment should also allow an internal estimate of error to be calculated from replication (Atkinson and Donev, 1992). That is, it is sensible to repeat some of the trials in order to investigate the amount of random variation which is inherent in the experiment. The question of how many points to replicate is a complicated one. This question is not considered here. Instead we look at a more specific problem. We look at ways to generate D or V efficient n point (discrete) designs, which include at least r of their points at least twice.

Recall that p is the number of parameters in the model, s is the number of support points in the continuous optimal design and n is the number of points in the exact design. In some cases the n point optimal design with r points included at least twice is simply the n point optimal design. Subsequently, when this is the case, we say that the appropriate design problem is *trivial*. For the unconstrained full simplex, the exact optimal n point design with at least r points included at least twice is known in some cases. Vuchkov (1982) presents results for the canonical polynomials on the unconstrained simplex. Vuchkov shows that if the support points of the continuous optimal design (see §2.3.5) are used to provide support for a discrete design with n > p, the variance of the predicted response at support point $\underline{x}_i$ is

$$\sigma^2 \, \underline{x}_i'(X'X)^{-1} \, \underline{x}_i = \frac{\sigma^2}{n(\underline{x}_i)}, \text{ where } n(\underline{x}_i) \text{ is the number of replicates at the support point } \underline{x}_i.$$

From this result we can deduce that for canonical polynomials on the unconstrained simplex, the exact G (and hence D) optimal design, over the support points of the continuous D optimal design, replicate the support points evenly (see §3.2.1). For the linear and quadratic model, when p < n < 2p, the discrete D optimal design over the whole simplex also replicates

these support points evenly. This result is due to Gaffke (1987). In both of these situations, if n is sufficiently greater than r, such optimal designs will automatically have at least r points replicated at least twice.

For constrained problems the situation can be more complicated, but in some situations, the problem of finding the optimal design with at least r points included at least twice is again trivial. The various situations which can arise are discussed fully in §3.2.2. The cases in which the optimisation problems are not trivial are when s ≥ n > p and when n > s > n-r. In this chapter we derive a variety of methods for addressing the above optimisation problem in these non-trivial cases.

One method for addressing the above design problem is presented in Snee (1979). All of the new methods which are developed in this chapter are compared to the Snee approach. Using Snee's method one generates the optimal n-r point design, and then adds those r of these points which lead to the biggest increase in criterion value when added sequentially. In this chapter we find that the best method for the specific problem outlined above is to use a modified form of the KL algorithm (Atkinson & Donev, 1992) which only considers designs with the required number of replicates.

4.1.2 Some Definitions

If a design includes a certain trial t times, then the design is said to include t *replicates* of that point. In some cases below, points which are included once are referred to as *singletons* and points which are included twice are referred to as *doubletons*.

The *pure error sum of squares*, an internal estimate of error, is:

$$\frac{1}{\sum_{j=1}^{r}(r_j-1)} \sum_{j=1}^{r}\sum_{k=1}^{r_j}(y_{jk}-\overline{y_{j.}})^2 \, ,$$

where there are r replicates each produced r_j times and y_{jk} is the k'th observation of the j'th replicate, i.e. k=1, . . ., r_j and j=1, . . . r. The pure error sum of squares has $\sum_{j=1}^{r}(r_j-1)$ degrees of freedom.

In finite designs the number of distinct points will depend on the weights in the continuous optimal design. The term *loadings* is defined here to be the rounded (to integers) values of n times the weights.

4.2 Outline of Chapter

4.2.1 Situations that can Occur

Firstly, it is noted that the variance of a measured response may vary over the design region. In most situations the variance is either constant over the design region or the fact that it is not constant is not known. We restrict attention to an analysis based on the assumption of constant variance.

It was observed in chapter 3, that in the majority of cases the support points in the continuous optimal design were sufficient for the exact optimal designs. In general, increasing the number of design points when $n < s$, is likely to improve efficiency more than increasing the number of design points when $n > s$. This is likely to be because additional support points can be added when $n < s$. If n is increased when all the support points are already included (i.e. $n > s$), the improvement is likely to be maximised by replicating the support points, rather than adding new distinct points.

The numbers of design points required in all cases considered in this chapter are not high enough to warrant any of the points being replicated more than twice. For the purposes of this chapter, r is defined as the number of points which are to be included in the design twice. The number of distinct points depends on the values of n and r. The number of distinct points is defined here as d and equals $n-r$. A *singleton* is defined here as a point which is used in the design exactly once. The number of singletons also depends on the values of n and r and equals $n-2 \times r$.

Recall that if the values of n times the weights are appreciably different from integers then additional distinct points, often approximately equi-distant from other support points, may be included in the exact design. If $n > s$ and the values of n times the weights are not appreciably different to integers, the exact optimal design will tend to include a maximum of s distinct points, with the remaining points being replicates of the s support points.

4.2.2 Situations to be Considered

Depending on the values of n, s, p, r and d, there is a variety of situations which can arise. The different situations can be subdivided into the following categories.

1) **n=p** All parameters in the model for E(Y) can be estimated if the information matrix can be uniquely inverted (see §2.2.2). In well designed experiments, this will occur if the number

of distinct points in the design is at least p, the number of parameters in the model. If n=p then all n points will therefore have to be distinct so that d=p; hence r=0.

2) $s \geq n > p$ In this case the design will be more efficient if all of the points were distinct and no replicates were included. It is suggested here, however, that the advantages of including some replicates (see §4.1.1) outweigh the drop in efficiency which would occur if support points were left out of the design and so it is suggested here that rather than having no replicates at all, a small number of them should be included.

3) $n > s$ In this case, the recommended approach for selecting optimal designs with r replicates will depend on the number of replicates required and the loadings (see §4.1.2). The situations which arise can again be split into three cases (see below). The action to take in each of these three cases will depend on whether there are enough replicates in the optimal design and whether even replication occurs. Recall that even replication means that if any points are replicated f+1 times then all other points are replicated at least f times (see §3.2.1).

Whether or not even replication occurs when n > s, will depend on the loading of the support points. If the r support points with the highest weights all have loading close to 2, then each of these will be replicated twice and even replication will have occurred. If m of the support points have loading close to 3 while others have loadings which are considerably lower than 2, some of the m may be replicated three times, before others are replicated twice. If m < r then the n point optimal design would probably not include r replicates, but, a design which does could be obtained by a simple modification to the n point optimal design. If some of the points are replicated three times before others are replicated twice then uneven replication will have occurred. Uneven replication is more likely to occur for the V criterion because the centroid was found to be weighted considerably more highly than the other support points.

There are three situations which can arise if n > s

(a) $d = n-r < s$ In this case, depending on the loadings, the optimal design will tend to include the s support points at least once. The remaining n−s points will tend to be replicates of the support points. If even replication occurs, the n−s replicates will all be different. If so they can be considered as the first n−s replicates. An additional $a_1 = r-(n-s)$,

replicates can then be chosen to bring the total number of replicates to r. Selection of the a additional replicates can be done by removing a of the distinct points from the design, and including a of the remaining distinct points twice. If a is small then a simple exchange algorithm can be used to choose the a points to remove and the a points to include twice. The simple exchange algorithm will consider all ${}^sC_a \times {}^{(s-a)}C_a$ possibilities. If a is large then an exchange algorithm of this type may take too long. In this case it will probably be better to use the alternative methods discussed in this chapter (see §4.2.3).

This exchange algorithm is unlikely to yield the optimal design with r replicates, because the continuous optimal design will include points which are good for n point optimal designs, and these points may be different from those which are good for designs which include r replicates. If uneven replication occurs or the optimal design with r replicates is required (rather than a good design with r replicates), then the alternative methods discussed in this chapter can be used. If alternative methods are sought and even replication occurs then, depending on the loading, it may be possible to locate the optimal design with r replicates very easily. Method R5 (see §4.2.3), involves selecting a number of good n point designs and modifying them to include r replicates. If some of the support points have low loading, efficient n point designs may automatically have r replicates. If the search for all good n point designs (see §4.2.3) is constructed well, then some of the good designs which are chosen may already have r replicates. If this is the case the best of these designs will probably be the optimal n point design with r replicates.

(b) $d = n-r = s$ Here, depending on the loadings, the optimal n point design will tend to include the s support points with the remaining n-s points being repeated support points (replicates). If the support points are evenly replicated then the optimal design will include the required number of replicates. Again if uneven replication occurs then the alternative methods discussed in this chapter can be used.

(c) $d = n-r > s$ Unless there is a specific reason for wanting more distinct points than support points, it should be avoided, because it will lead to an inefficient design. A possible reason for having more than s distinct points would be a special interest in a certain area of the feasible region. If there is a special interest in a certain area of the feasible region, it may be necessary to have more design points in this region than would be selected under optimality.

Uneven replication

In the event of uneven replication, selection of the replicates needs some care. If the loadings of the support points suggest that a small number of points should be replicated a large number of times, the reason for including the replicates needs to be considered. Although cases of this type are not considered here, some V optimal designs may include the centroid in a design many times, before any other points are replicated. It seems sensible that the estimate of internal error should be applicable to all parts of the feasible region (S). To satisfy this requirement, one should replicate a range of points which are spread reasonably evenly throughout S. Once replicates have been placed in all parts of the feasible region, it is suggested that the remaining replicates should be chosen to maximise the criterion value, irrespective of whether uneven replication occurs. Although the replication of a range of points seems sensible, the degrees of freedom associated with the error sum of squares will increase with the number of times each point is included in the design. In addition to this, the accuracy of the internal error estimate is likely to improve if each repeated trial is repeated more than twice. Because there is a variety of conflicting factors, the design should be chosen so that a balance between the above three factors and D efficiency is obtained.

In some cases restricting designs to have exactly r points replicated exactly r times would lead to inefficient designs. It is for this reason that the algorithms discussed below will search for designs with at least r of their points included at least twice. Designs which have this quality are subsequently referred to as designs with r replicates.

4.2.3 Outline of Methods to be Considered

In cases 2 and 3a above, methods for generating optimal designs with r replicates are needed. A variety of algorithms is investigated.

Obtaining a good design with r replicates can be approached in one of three ways. The first possible approach is to construct the optimal design with no restrictions on the number of replicates, using the theory of chapter 3, and then modifying this design to have r replicates in some way (e.g., see below, and §4.2.2, 3a above). This first approach is subsequent referred to as the modification approach. The second possible approach is to use the weights of the continuous optimal design to choose which points should be replicated. This approach based on weights is used in method R6 (see below). The third possible approach is to use a KL type algorithm which only considers designs with r points replicated twice. This KL type approach is implemented in the R7 algorithm (see below).

We now consider the modification approach. Two different types of good starting designs can be modified. The first approach, which is used in Snee (1979), is to start with a d=n−r point design and build up to an n point design. Three variants to the Snee method are also investigated below. The second approach is to take a good n point design and amalgamate some of the points. This second approach is used in the R5 algorithm below.

We concentrate first on the approach which involves generating a good design with d distinct points, and then choosing r of these points to replicate. The good designs with d distinct points will be generated using KL with the same candidate set as was used in chapter 3 (see §3.5.3). It is conjectured that the optimal d point design will not necessarily form the basis of an optimal n point design with r replicates. All methods which involve modifying d or n point efficient designs to include r replicates will be used on a variety of efficient designs. The d or n point optimal design will also be included in the selection of efficient designs. These unrestricted starting designs are subsequently referred to as *bases*. A range of bases is used to investigate whether the relationship between bases and final designs is monotonic. Because the relationship was found to be strong but not monotonic, the improvements which could be made by using a range of good bases, instead of the optimal basis, were then investigated. To generate these multiple bases, mkl, a modified version of KL, was implemented. mkl involved running KL t times and taking every distinct final design produced as one of the bases.

The only method for selecting r replicates from d distinct points which could be found in the literature is given in Snee (1979). Two alternatives to the Snee approach are introduced below as R1 and R4. These two new methods are then compared with the method of Snee (1979), and are also compared with a modified version of the Snee method. The original and modified methods of Snee are referred to as R2 and R3 respectively.

Because R1 to R4 all involve extending good d point starting designs, they are subsequently referred to collectively as *extension* methods. Three alternative methods, which do not involve extending good d point design, are also considered. The three alternative methods are given as R5 to R7 below.

Extension methods

R1. Consider all ^dC_r selections of r replicates and select the one which gives the best value for the chosen criterion.

R2 Add to the basis, one point at a time, those r distinct points which give the biggest improvement in criterion value, for designs of size d+1, d+2, . . . , n (=d+r). The points giving the biggest improvement can be selected using the updating formulae (see §2.3.8C). Each point is replicated a maximum of once. The method of Snee (1979) involves a similar procedure but does not insist that all points chosen for replication are distinct from each.

This algorithm was later combined with R3 (see §4.7.4).

R3 Add those r points from the d, which give the r biggest improvements in criterion value when added individually to the d points, to make a design of size d+1.

R4. Use KL (see §2.3.7B), to locate the optimal n point design, using the d distinct points as its candidate set. If, using the d points given, the loading of the support points leads to uneven replication (see §3.2.1), R4 will not produce a design with r points replicated exactly twice. R4 will never improve upon R1 because R1 considers all possible choices of r replicates. If $^d/_2 < r \leq d$, then as the number of replicates required increases, R1 will rapidly become computationally intensive and R4 can be used as an alternative.

Alternative methods

R5 An n point design will be generated using KL and the same candidate set as was used in chapter 3 (see §3.5.3). Based on Euclidean distance, the two points which are closest to each other will be identified. Ignoring this first pair, the next pair of closest points is identified. Once the r pairs of closest points have been identified, r sets of three points will be formed, using each original pair and their midpoint. From each of the r sets, one point will be replicated in the final design as a replacement for the original pair. All 3^r possible selections will be considered and the one which produces the design with the best criterion value will be selected as the final design. As r increases the number of possible selections increases rapidly but with $r \leq 8$ this method will be viable if a reasonably powerful computer is available.

R6 The motivation for this method was to use knowledge about optimal designs for constrained regions, i.e. the results obtained in chapter 2. This is an entirely different approach from the 5 methods discussed above. R6 is based on the weights of the support points in the continuous optimal design. If a design with r replicates is constructed by taking the d points with the highest weights, and then replicating the r of these, which have the highest weights, it seems likely that the design will be efficient. Although there is no

guarantee that the support of the continuous and discrete designs will coincide, this approach has been shown to work in a number of situations. For example, the support for the exact optimal design, when a quadratic model is used on the full simplex, coincides with the support for the continuous optimal design (i.e. the (3, 2) simplex lattice) (see §4.1.1). It was concluded that R6 may lead to a good optimality solution in cases where the support does coincide.

R7 It seems likely that the best method for generating an n point design with r replicates would be to construct an algorithm which only considers good n point designs that have r replicates. This method will certainly provide the best design if it is possible to include all possible designs with r replicates. The description of this algorithm is given in the next section (see §4.3).

4.3 Exchange Algorithm which only Considers Designs with r Replicates

In the previous section it was suggested that the best method for generating an efficient design with r replicates would be to use an algorithm which only considered designs which had r replicates. Construction of this algorithm is discussed in this section. The code itself is given in electronic form on the attached disk as klr.m.

With many practical design problems, some runs may already have been performed. In cases where runs have already been performed, the design should be constrained to include these points. In other cases there may be a special interest in certain parts of the feasible region (it may be known that the error variance is higher in some regions) and a user may insist that the design includes specific points in this area. The exchange algorithm constructed here will have an option for finding the best additional points, given that the user specified points must be included in the design. These user specified points will subsequently be referred to as *fixed* points.

It was decided that in the R7 algorithm, singletons and doubletons would be stored separately. The singletons and replicates would then be combined for calculation of the criterion value. The algorithm will also ensure that the r points selected for replication are distinct from each other. The option of fixed replicates as well as fixed singletons was also included. Because the optimal design might include one of the fixed singletons as a replicate, a third group of points was identified. The third group of points are subsequently referred to as 'fixed singletons as replicates'. If more exchanges are still to be considered, points in this

new category are different from the normal fixed replicates, because they can revert back to being fixed singletons. Figure 4.1 shows which types of point exchange are permissible. To ensure that at least r replicates are included in the design, all points in the set of replicates must be distinct. To ensure that all points specifically selected for replication are distinct from each other, the exchanges illustrated by bold lines in Figure 4.1 must be restricted to those points which are not already included in the set of replicates.

Figure 4.1 Possible exchanges in the R7 algorithm
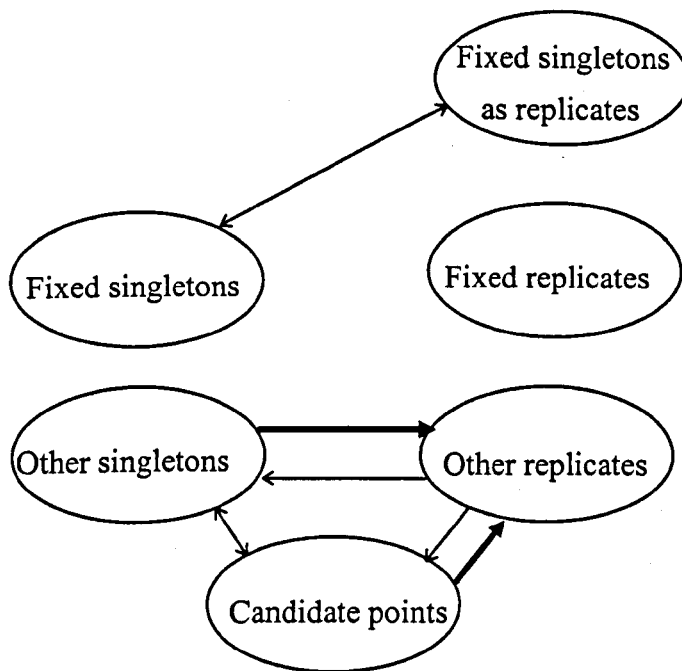
Each iteration of the KL algorithm (Atkinson and Donev, 1992), is initiated by the random selection of $q < n$ points. The value of q is also chosen randomly. In the R7 algorithm the random points were selected to be singletons and replicates in the same ratio (rounded to the nearest integers) as the number of singletons and replicates required in the final design. If any repeats were observed in the set of random points selected as replicates, the additional point was not added to the set and a new random point was generated. The starting set of random points was then built up to an n point starting design by adding, one point at a time, those points which led to the biggest improvement in criterion value at each stage. In the R7 algorithm the points for addition are alternately assigned to the sets of replicates and singletons until each set contains the required number of points. To ensure that the r points

selected for replication are distinct from each other, if any of the points selected for the starting design is already in the starting set of replicates, it is automatically assigned to the starting set of singletons.

Once the starting design has been selected there are three types of exchange which can be made. Exchanges between candidate points and singletons, between singletons and replicates and between candidate points and replicates. Two separate implementations of the R7 algorithm were produced. The first implementation performed each type of exchange sequentially. The algorithm first considered exchanges between the ls best candidate points and the ks worst singletons which were currently selected in the design. The second stage of the algorithm considered exchanges between all singletons and all replicates and the final stage considered exchanges between the kr best candidate points and the lr worst replicates currently included in the design. A flow diagram summarising this first implementation is given in Figure 4.2.

It was hypothesised that the R7 algorithm would be likely to produce better designs if all three types of exchange were considered simultaneously. If each type of change were considered sequentially then beneficial updates which involved 2 or more of the three types of exchange would never be found. In other words, the exchange of singleton A for replicate B, with temporary loss of efficiency, would be ignored, even though a subsequent exchange between point B and the candidate set may have led to an overall improvement. Consequently a second implementation, which considered exchanges involving all three types of points at the same time, was investigated. A flow diagram summarising this second implementation of the algorithm is given in Figure 4.3. To investigate which approach was best both algorithms were implemented.

In Figures 4.2 and 4.3 the terms *best point* and *worst point* are used to mean the points which lead to the biggest and smallest improvements in the chosen criterion, when they are added to the design which is currently selected. The best and worst points are selected using the updating formulae (see §2.3.8C). The parameters of the R7 algorithm are k, l and t. The values of k and l represent the number of points which will be considered for removal from and addition to the currently chosen design, for each exchange. The value of t is the number of iterations which are to be performed

In both implementations of the R7 algorithm the values of ks and kr are selected so that they sum to k and are in the same proportion as the number of singletons and replicates required in the design. The values of lr and ls are calculated in a similar way.

In the formula for the number of new designs to be considered in the second implementation (see Figure 4.3 stage C), one is added to the number of each type of point. The addition of one is necessary, because the designs may be improved if only one or two of the exchange types are made. Because three types of exchange are to be considered the values of k and l do not need to be very big before the algorithm becomes computationally intensive. By keeping k and l low and increasing t, the algorithm can be made to be relatively efficient.

Figure 4.2 Flow diagram for the first implementation of the R7 algorithm
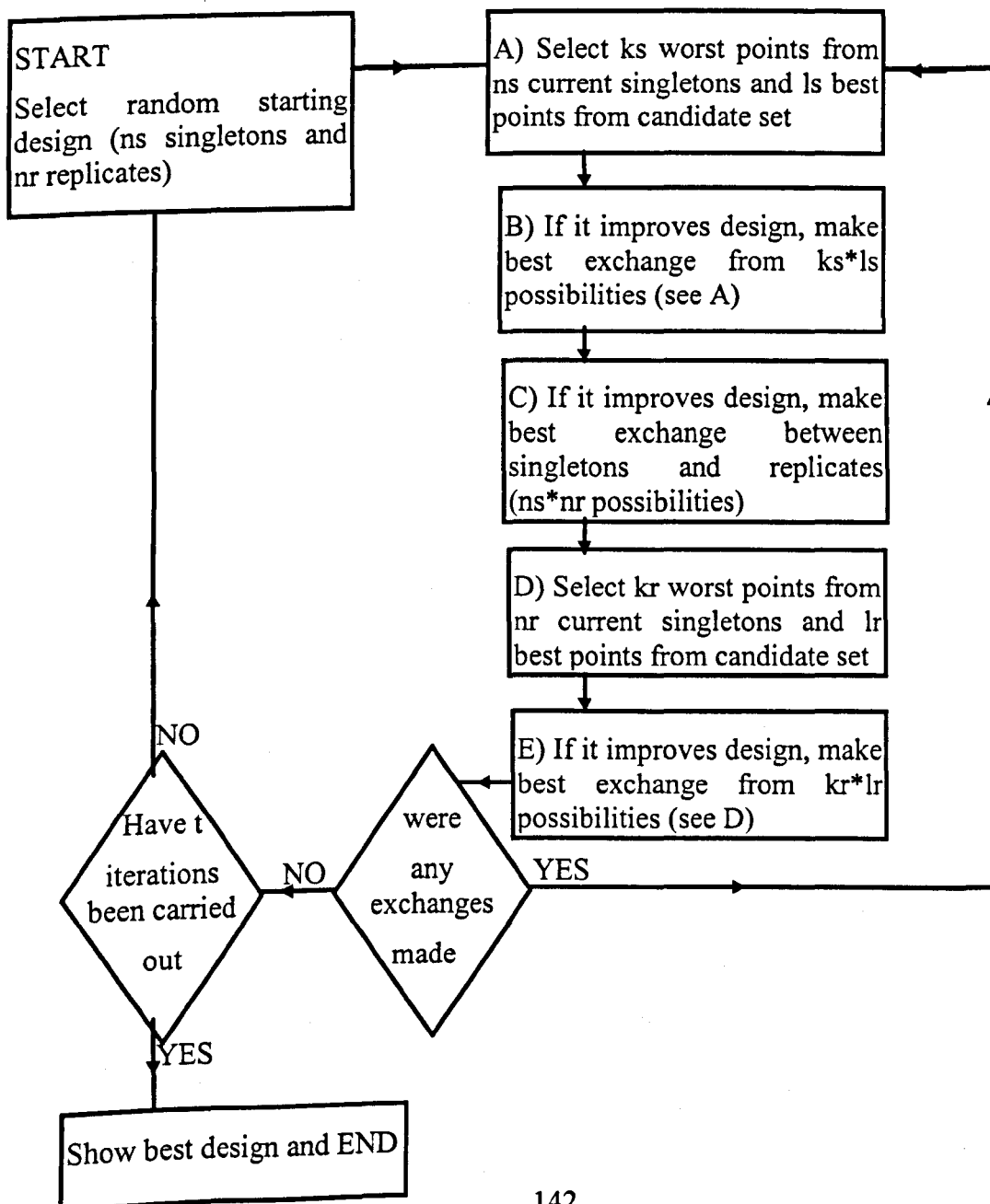
Figure 4.3 Flow diagram for the second Implementation of the R7 algorithm

```
┌──────────────────────────┐      ┌──────────────────────────┐
│ START                    │      │ A) Select ks worst points│◄─────────┐
│ Select random starting   │─────▶│ from ns current singletons,│         │
│ design (ns singletons and│      │ and ls best points from  │         │
│ nr replicates)           │      │ candidate set            │         │
└──────────────────────────┘      └──────────────────────────┘         │
                                             │                          │
                                             ▼                          │
                                  ┌──────────────────────────┐         │
                                  │ B) Select kr worst points│         │
                                  │ from nr current singletons│        │
                                  │ and lr best from          │        │
                                  │ candidate set             │        │
                                  └──────────────────────────┘         │
                                             │                          │
                                             ▼                          │
                              ┌──────────────────────────────┐         │
                              │ C) Consider all              │         │
                              │ (ks×ls+1)×(kr×lr+1)×(ns×nr+1) new│     │
                              │ designs (see stages A and B).│         │
                              └──────────────────────────────┘         │
                                             │                          │
                                             ▼                          │
                              ┌──────────────────────────────┐         │
                              │ D) Make the exchange (see C), which│   │
                              │ most improves the design      │         │
                              └──────────────────────────────┘         │
                                             │                          │
              NO                             ▼                          │
          ┌─────────┐                 ┌─────────────┐                   │
          │ Have t  │                 │   were      │      YES          │
          │iterations│◄─── NO ────────│    any      │──────────────────▶┘
          │been carried│              │  exchanges  │
          │   out    │                │   made      │
          └─────────┘                 └─────────────┘
              │ YES
              ▼
    ┌──────────────────────────┐
    │ Show best design and END │
    └──────────────────────────┘
```

Both versions of the R7 algorithm were used for the five component example of Snee (1979) (see §4.6.1). Each implementation was run 10 times with the following parameters: $k=10$, $l=6$, $t=100$. The version which considered the three exchange types simultaneously, produced designs with D values in the following range: $(4.1299 \times 10^{-46}, 6.4162 \times 10^{-46})$, the version which considered the three exchange types separately, produced designs with D values in the following range: $(2.4687 \times 10^{-46}, 5.6361 \times 10^{-46})$. These results give some evidence that the second implementation is more effective. Due to this result and the fact that

it is intuitively likely to be more efficient, the second implementation is used for all examples considered.

It is expected that, because the R7 algorithm only considers designs with r replicates, the designs it generates will be better than the designs generated using the extension methods. In addition to the expected improvements in criterion values, in order to use R7 the number of support points and their respective weights are not needed. Irrespective of whether there are r replicates in the optimal n point design, or uneven replication occurs, the R7 algorithm will generate a design in which the r points selected for replication are distinct from each other (see §4.3).

It is ensured that the r points specifically selected for replication are distinct from each other, because no repeats are allowed in the set of replicates. Unless there is some prior notation that at most r replicates are needed, it would be unwise to impose this restriction as it may reduce efficiency. The number of replicates is not restricted to r, i.e. points may be repeated more than twice, because there are no constraints on the set of singletons being distinct from each other or the set of replicates.

If the continuous optimal design were not considered when the extension methods were being used, the points in each new best design would need to be monitored carefully to ensure that the r points selected for replication were distinct from each other.

4.4 Selecting an Appropriate Problem

4.4.1 Examples where the Location of the Optimal n Point Design with r Replicates does not Require any of the Algorithms Developed in this Chapter

The first example to be considered was the glaze slip problem (see §5.4). The continuous D optimal design for this problem and a quadratic model has 7 support points. It was decided that a design with 3 replicates would be used. This replication structure means that the problem is of type 3(b) (see §4.2.2). As none of the weights were particularly low the D optimal 10 point design included the required number of replicates.

The next problem to be considered was a 4 component example from Snee (1975). This four component example had constraints with the following lower and upper bounds: $x_1 \in (0.07, 0.18)$, $x_2 \in (0, 0.3)$, $x_3 \in (0.37, 0.7)$ and $x_4 \in (0, 0.15)$. The continuous D optimal

design for this region and a quadratic model has 19 support points with the following weights.

(0.0555 0.0842 0.0791 0.0851 0.0537 0.0515 0.0807 0.0762 0.0408 0.0227

0.0298 0.0288 0.0369 0.0401 0.0923 0.0139 0.0339 0.0358 0.0590)

An example of type 2 can be created from this constrained region if the design is required to have 18 points with 4 replicates. As some of the support points have comparatively low weights, one of the good n point designs, which were generated as bases for R5, automatically included 4 replicates.

4.4.2 The Problem Selected

It was decided that a problem of type 3(a) should be considered first as in problems of type 2 there is a good argument for keeping the number of replicates to a minimum (see §4.2.2). Recall that problems of type 3(a) have n > s > d.

The new replication algorithms were not needed for the two examples in the previous section, as some of the good n point designs for these examples automatically included r replicates. To provide a good test of the algorithms we need an example for which the good n point designs will only include a small number of replicates.

If we choose an example in which the support point weights (in the continuous optimal design) are all similar, then the good n point designs should include most, if not all, of the s support points. If we then choose a value of n which is only slightly bigger than s, there will only be a few spaces for replicates (n-s if all of the support points are used). If we then choose r to be greater than n-s by 3 or 4 points, the replication algorithms will be needed.

In chapter 2 we looked at the continuous D optimal design for a quadratic model and a hexagon which is symmetric under a third turn (see §3.6.6). It was noted that for b=5/6 the 10 support points all had similar weights. It was decided that a good example of type 3(a) would be created if an 11 point design with four replicates were required, i.e. n=11, r=4, d=7, s=10, p=6. To create a slightly different example from the one considered in chapter 2, b=0.83 was chosen.

4.4.3 Selecting Starting Designs

Two approaches for selecting bases are investigated. The bases can either be D efficient designs or designs consisting of randomly generated points from the candidate set. The choice between these two options will be based partly on how much the D value of the basis influences the D value of the final design. The relationship between the criterion values of bases and final designs is therefore investigated in this chapter.

· If the best basis always leads to the best final design, multiple bases will not be needed. If only efficient bases lead to efficient final designs, it will be sufficient to consider a small number of efficient bases. If very inefficient bases can lead to efficient final designs, the best approach will probably be to consider a large number of random bases. The designs generated from random bases will therefore be compared to the designs generated by efficient bases. As all five modification methods are different, especially R5, the use of random bases may work better for some of them than others. Consequently the use of random bases will be investigated as well as the use of efficient bases, for all 5 methods.

The following sets of bases were considered:

1. N good designs generated using KL

2. 49 designs made up of distinct random points from the candidate set.

The KL algorithm was modified to perform a large number of small individual searches, and to save each of the distinct designs which was generated. This modified algorithm is subsequently called mkl. The code for the algorithm is given in electronic form on the attached disk as mklv.m and mkld.m, where the fourth letter indicates the optimality criterion. The parameters of mkl were chosen so that approximately N=10 efficient designs were generated and that these 10 designs included the optimal design. The parameters required to produce these bases were located by a process of trial and error.

It was decided that the number of random bases should be relatively large (i.e. greater than 10) because similar algorithmic approaches generally need a large number of different starts to find the optimal solution. 49 were used as any more would be unnecessarily time consuming unless this method was found to be useful.

For R1 to R4, the same starting designs were used so that reasonable comparisons could be made between the results. A different set of bases was clearly needed for R5 as an additional three points are required.

To investigate the relationship between the criterion values of bases and final designs and whether more efficient designs can be generated by the use of multiple bases, the following summary statistics were calculated for each method: a) the mean, standard deviation, maximum and minimum criterion values for the final designs; b) the correlation coefficient, r_c, between the criterion values of the final designs and bases they were generated from; c) p-vectors, where the n'th value in the p-vector is the rank of the basis which yielded the n'th best final design. The p-vectors given below are restricted to, at most, the best 12 final designs.

4.5 Comparisons of the Seven Methods for the D Criterion, a Quadratic Model and q=3

4.5.1 Results

A Generation of bases

To generate approximately 10 bases to be used with R1 to R5 the modified version of KL, mkl (see §4.2.3), was used with the following parameters.

Table 4.2 Parameters used to generate multiple bases in mkl

Design Points	Number of Runs	k	l	Random starts (t)
7	300	1	2	1
10	200	1	2	2

All D values which are quoted for this example are scaled up from the calculated value by a factor of 10^3. In all subsequent tables of a similar form to Table 4.3, r_c denotes the correlation coefficient between the D value of the basis and the D value of the final design from that basis.

B Results for R1

Table 4.3 Summary statistics for the designs generated by R1

Generation of bases	N	Mean	St-dev	Max	Min	r_c
mkl	9	0.2182	0.0562	0.2831	0.1126	0.9844
Random	49	0.0224	0.0506	0.2647	0.0001	0.9994

Bases	p-vector
mkl generated	2 1 3 4 5 6 7 8 9
Random	1 2 3 4 5 6 7 8 9 10 11 12

From Table 4.3 and the p-vector for bases generated by mkl it can be seen that the best final design, which had a D value of 0.2831, was generated by the second best starting design. The design generated from the best basis had a D value of 0.2819 which is 99.6% efficient compared to the best design generated. This small increase in efficiency suggests that the improvement gained by using multiple bases is unlikely to be of much practical benefit when q=3. The improvement does suggest that multiple bases will be more effective in locating optimal designs.

From Table 4.3 it can be seen that the bases which were generated randomly, and were therefore inefficient, gave by far the worst designs. This suggests that the D value of the starting design has a strong influence on the D values of final design. The possibility of improving the designs generated using a large number of random starts is investigated below (see §4.5.4).

From the above results it can be seen that better designs can be generated if a range of efficient bases are used. It also suggested that for R1, as the bases get better, the relationship with the final designs becomes weaker. This can be seen by the fact that the ranks in both p-vectors are in numerical order until the best bases are considered. When good bases are considered, the best designs were not necessarily constructed from the best bases. This suggests that the relationship is not monotonic but that, in order to get good final designs, only highly efficient bases need be used.

C Results for R2

Exactly the same results were obtained in all cases for R2 where the best r points are replicated sequentially. The fact that the results were identical suggest that R2 should be used because it takes considerably less time to run. Because R1 considers all possible replication strategies, when the size of the problem increases and/or the number replicates required increases, it is hypothesised that R1 will lead to a better final design. This possibility is considered later for a bigger example (see §4.6).

D Results for R3

Table 4.4 Summary statistics for the designs generated by R3

Generation of bases	No	Mean	St-dev	Max	Min	r_c
mkl	9	0.1866	0.0421	0.2608	0.1126	0.8455
Random	49	0.0171	0.0382	0.1946	0.0000	0.9971

Bases	p-vector
mkl generated	2 3 1 6 4 7 5 8 9
Random	1 2 3 4 5 6 7 9 13 8 11 12

From Table 4.3 it can be seen that as expected, R3 produced considerably worse designs than R1 or R2. The best design generated by using good bases for R3 was worse than the best design generated using random bases for R2. It is not surprising that R2 produced better design than R3 because each additional point is based on all previous updates. It is however possible that R3 will produce a better design than R2. This is discussed in a later example (see §4.7.4).

Again it can be seen that better designs can be generated if a range of efficient bases is used. From the correlation coefficients and p-vectors it can be seen that the relationship between bases and final designs is weaker than for R1 and R2.

E Results for R4

KL was modified to use more than one candidate set and to record the results from each candidate set used. When the algorithm was run repeatedly with k=3, l=4, t=100, for the same multiple starts, the summary statistics were different every time. The results in Table 4.5 are all based on 100 iterations of the algorithm.

Table 4.5 Summary statistics for the designs generated by R4

Generation of bases	N	Mean	St-dev	Max	Min	r_c
mkl	9	0.2730	0.0305	0.2831	0.1918	0.2252
Random	49	0.1954	0.0908	0.2647	0.0003	0.1285

Bases	p-vector
mkl generated	2 8 4 5 1 6 9 3 7
Random	1 38 42 28 49 35 30 26 44 10 12 25

The best design generated by R4, for this example, replicated the chosen points evenly. The best design was identical to the one generated by R1 and R2. From the p-vector for the good bases generated by KL, it can be seen that the best basis produced the 5th best final design. This fifth best final design, which was produced from the best basis, was the same design as was produced using the best basis by R1. The fact that with R4 the best basis only produced the 5th best final design suggests that R4 makes better use of inefficient bases than the other extension methods and does not suggest that k, l and t in the modified version of KL were too low.

F Results for R5

From Table 4.6 it can be seen that R5 failed to generate the best design which had been generated by R1, R2 and R4. It can also be seen that the relationship between the criterion values of bases and final design is even weaker, for the mkl generated bases, than for the previous methods. Given this weaker relationship it is not surprising that the use of multiple bases led to a better design than if the optimal basis had been used. From the table it can be seen that the best basis led to a design which was inefficient compared to the best

Table 4.6 Summary statistics for the designs generated by R5

Generation of bases	No	Mean	St-dev	Max	Min	r_c
mkl	10	0.2316	0.0293	0.2819	0.1899	0.0116
Random	49	0.0820	0.0764	0.2096	0.0000	0.9308

Bases	p-vector
mkl generated	7 4 8 2 3 5 10 9 6 1
Random	7 13 8 12 5 1 3 6 4 10 15 14

design. This low efficiency of the design generated from the best basis suggests that multiple bases are probably needed for R5.

In general, the relationship between bases and final designs appears to be considerably weaker for R4 and R5 than for R1 to R3. This weaker relationship suggests that an increase in the number of good bases may be more beneficial for methods R4 and R5 than for methods R1 and R2. The use of R4 and R5 with an increased number of starting designs may lead to designs which are better than those generated by R1 and R2, especially if the bases ranged from being moderately efficient to being highly efficient. From the results for the randomly generated bases, it is clear that the range of bases should not be extended to include them. The possibility that a larger range of bases will lead to better design is investigated for a bigger example below (see §4.6).

G Results for R6

The weights for this example were calculated using the method described in chapter 3 (see §3.5.4). The weights for the 6 vertices, 3 long edge midpoints and centroid were found to be 0.1016, 0.1053 and 0.0742 respectively. The weights suggest that the centroid and any two vertices should be removed and three midedges and one of the remaining vertices should be replicated.

The criterion value for the design with r replicates will vary depending on which of the vertices are chosen. In chapter 2 it was concluded that D efficient design can be selected by attempting to spread the chosen points evenly over the constrained region. For a given set of points, it was concluded in chapter 3 (see §3.6.10), that the weights depended on how

many points were near by. Based on this conclusion, it seems likely that the best design will be obtained if any vertex is replicated, as long as the vertices removed are the two which are adjacent to it. The resulting design had a D value of 0.2440, and this design is 86% efficient compared to the best design produced by methods R1 to R5. This lack of efficiency suggests that algorithmic methods may lead to better designs than those constructed using the weights of the support points in the continuous optimal design.

H Results for R7

When the R7 algorithm was run with k=4, l=3 and t=100, the same best design, as was generated by R1, R2, and R4, was again generated.

4.5.2 Discussion

The p-vectors for methods R1 to R5 for the example considered above, show that the best final design was never generated from the best basis. This suggests that the use of multiple near optimal starting designs, leads to better final designs than if only the optimal basis were used. However, the differences in D values between the best design generated overall and the design generated by the best basis, were relatively small for all approaches apart from R5. This suggests that the use of multiple bases only leads to a real practical improvement for the R5 algorithm. However, if the aim of the study is to obtain the optimal design, rather than a design which is highly efficient, then the use of multiple bases is advantageous.

Despite the fact that the best basis never led to the best design, the results for R1 to R5 all suggest a positive correlation between the D values of bases and final designs. Because the random bases led to poor final designs it can also be concluded that efficient bases are more likely to produce efficient design with r replicates. The relationship between bases and final designs is not monotonic and also varies considerably between the 5 methods. With R1 and R2 the relationship is relatively strong, though the use of multiple bases still leads to a better final design than the use of the optimal basis. With R4 and R5 the relationship is weaker and it is hypothesised that the use of some less efficient bases in addition to the efficient ones may lead to better designs than R1 and R2 in bigger examples.

Methods R1, R2, R4 and R7 all generated the same best design for this example. However this example is very small and a better comparison would probably be obtained from a larger example. A 5 component example is considered below for R1, R2, R4 and R7

(see §4.6.2). Because methods R3, R5 and R6 produced comparatively poor designs for this small example, they are not considered for the 5 component example.

For all extensions methods apart from R4, the correlation between the D values of the bases and the D values of the final designs they produced, was stronger for the randomly generated bases. This was to be expected because any useful points in the random bases would be chosen for replication; therefore the efficiency of the final designs would depend on the number of useful points in the basis and hence the efficiency of the basis. As all points in the good bases should be "useful" anyway, the efficiency of the final designs they produce will depend on whether the points they contain are "useful" for replication. Whether or not a point is useful for replication does not necessarily depend on whether they are useful for unrestricted designs. The fact that the correlation is so low for the efficient bases actually suggests that points which are very useful in unrestricted designs are not the best points for replication.

4.5.3 Comments on the Best Design Generated

The best design produced is given in Figure 4.4. The exact points are given in Appendix 1.1. The best design is not unique as at least 12 designs share the same D value. Three of these equivalent designs are: the design in Figure 4.4 and the designs produced by rotating this design through one or two rotations of 120°. Each of these 3 equivalent designs can then be reflected in each of the lines $x_i=x_j$ to produce the other 9 equivalent designs.

Figure 4.4 Points chosen in the D optimal design with 11 points and 4 replicates for a quadratic model and the hexagon which is symmetric about a third turn with b=0.83
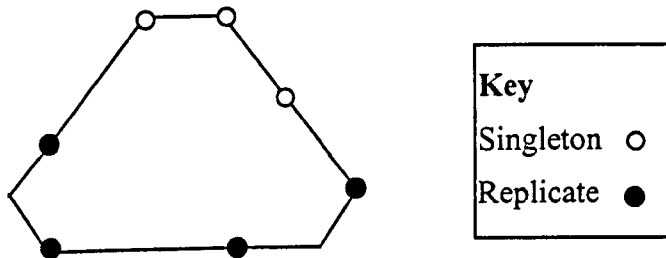
The 3 singletons in the best design are two vertices and the midpoint of a long edge. The 4 replicates are 2 vertices and the thirds of the other two long edges. As was observed for the continuous designs in chapter 3, this design spreads the points reasonably evenly around the feasible region. The main difference is that here the overall centroid is not used, whereas for the corresponding continuous design in chapter 3 (see §3.6.6) the overall centroid was weighted reasonably highly for b <0.88. The other main difference was that thirds of edges were not used in any of the quadratic designs in chapter 3. A more detailed comparison with the designs of chapter 3, which looks at all of the designs generated in this chapter, is given below (see §4.9.2).

4.5.4 Further Investigations into the use of Random Bases

From the tables of summary statistics for R1 to R5, it can be seen that the results from the random bases are reasonably competitive with those from the good bases. As R3 was the least computationally intensive, it was run with increasingly large numbers of bases to see how rapidly an improvement could be made. The results of R3 for increasingly large numbers of bases are given in Table 4.7. Recall that the best design with r replicates which was generated from good bases had a D value of 0.2831.

Table 4.7 Maximum D values using R3 with an increasing number of random bases

Number of Bases	D value
49	0.2647
999	0.2754
3333	0.2791

As is to be expected the table shows that we can improve our results by increasing the number of random starts. This is intuitively sensible because if the number of random starts was increased by a sufficient amount, the optimal design would be obtained because all possibilities would eventually be considered. For a small problem, this method would probably be feasible in terms of computing time. As the number of components and candidate points increases, the feasible region rapidly becomes much larger. To search the feasible

region sufficiently to obtain the optimal design would probably be extremely inefficient. This approach is not considered further.

4.6 A Five Component Example for the D Criterion

4.6.1 Introduction

In the previous section the relationship, between D values of bases and final designs, appeared strong enough to warrant the use of good bases rather than randomly generated bases. It was also concluded that using a range of good bases was likely to lead to a better design than if only the optimal basis was used. Because methods R1, R2, R4 and R7 all located the best design they are all considered here for an example with q=5.

A 5 component example was taken from Snee (1979). The Snee (1979) example has the following lower and upper bounds; $x_1 \in (0.5, 0.7)$, $x_2 \in (0.05, 0.15)$, $x_3 \in (0.05, 0.15)$, $x_4 \in (0.1, 0.25)$, $x_5 \in (0, 0.15)$. In Snee (1979) additional constraints were imposed and the fixed sum constraint only summed to 0.997. These complications are of no interest here and are ignored. Using the algorithmic method described in chapter 3 (see §3.5.4) it was found that the continuous optimal design for this problem and a quadratic model had 23 support points. To create a replication problem of type 2 (see §4.2.2), it was decided that a 21 point design with 4 replicates would be searched for, i.e. n=21, r=4, d=17, s=23, p=15.

It was also hypothesised in the previous section that increasing the number of good bases above 10, may improve the results for R4. The number of bases is increased to 28 for this example. The parameters used in mkl were again selected by trial and error so that approximately 30 bases would be generated. Using 100 searches with k=5, l=7 and t=40 produced 28 bases.

The full DX4 candidate set includes 542 points. Because each run of the algorithm took over an hour, the use of this many candidate points was not considered to be practical. Because triple blends and axial check blends have never been included in any of the optimal designs for the quadratic model, these points were removed and the reduced candidate set included 196 points.

4.6.2 Comparison of R1, R2, R4 and R7

The same design, with a D value of 5.578×10^{-46}, was generated by R1, R2 and R4. This best design was obtained using the second best basis, and the best basis led to the second

best final design. The final design which was located using the best basis had a D value of 5.51×10^{-46} which is 98.8% efficient compared to the best final design. The same conclusion as was reached for q=3, on the relationship between the D values of bases and final designs, was again obtained here. The relationship was extremely strong for R1 and R2 and comparatively weak for R4.

As was the case for the q=3 example (see §4.5), R1 and R2 located exactly the same designs from each basis. It is hypothesised that for the D criterion and q ≤ 5, the method of Snee (R2) is unlikely to be improved upon by any of the alternative seven methods. Because R1 is computationally more intensive than R2 it is suggested that R2 should always be used instead of R1 when q ≤ 5.

The overall best design was located by R7 with k=6, l=4 and t=300. This best design had a D value of 6.42×10^{-46} The best design which was generated using extension methods was only 86% efficient compared to the design generated by R7. It can be concluded that R7 can provide substantial improvements over the extension methods (see §4.2.3). The only problem with the use of R7, though not a serious one, was that it took 3 hours to run using a Pentium 333 processor whereas R2 only took approximately one hour.

4.6.3 Comments on the Best Design Generated

The points in the best design generated here are given in Appendix 1.2. The 13 singletons are comprised of 9 vertices, 2 face centroids, one edge midpoint and one edge third. The 4 replicates were all vertices. As with the q=3 design discussed above, the points used in this design are slightly different from the points which tended to be used in the designs of chapter 3. In particular, this design does not use the overall centroid but includes an edge third. Again the points are spread reasonably evenly over the feasible region. Whether or not the design is unique was not investigated. Unlike the q=3 design above, the region is relatively complicated. There are no designs which are obviously equivalent due to symmetry. A more detailed comparison with the designs of chapter 3, which looks at all of the designs generated in this chapter, is given below (see §4.9.2).

4.7 Comparisons of Methods for the JM Design (q=3), Under the V Criterion

4.7.1 Introduction

It has been observed that the weights in continuous V optimal design tended to be less evenly spread among the support points than for continuous D optimal design (see §3.9.2). Due to this difference in weightings it is likely that methods for finding designs with r replicates (see §4.2.3), are likely to perform differently for the V criterion. Some of the 7 methods will be considered for two examples using the V criterion. The first example considered will be the design for the JM glaze problem (see §5.4).

The aim of the JM experiment is to produce a good model for predicting the viscosity of future mixtures (see §1.1). The V criterion minimises the average scaled prediction variance over the feasible region (see §2.3.4) and was consequently chosen as the most appropriate criterion.

4.7.2 Deciding the Number of Design Points for the JM Problem

The constraints for this problem are given in chapter 5 (see table 5.4 in §5.4.2). We were informed by JM that as slip analysis was a lengthy procedure, we should keep the number of slips analysed to a minimum. As a quadratic model was being used, at least 6 distinct design points were needed to fit the model, and at least 7 distinct design points were needed to perform the model checking procedure discussed in chapter 2 (see §2.2.6). Consequently it was decided that we

Figure 4.5 Plot of V Values as n is increased for the JM example

157
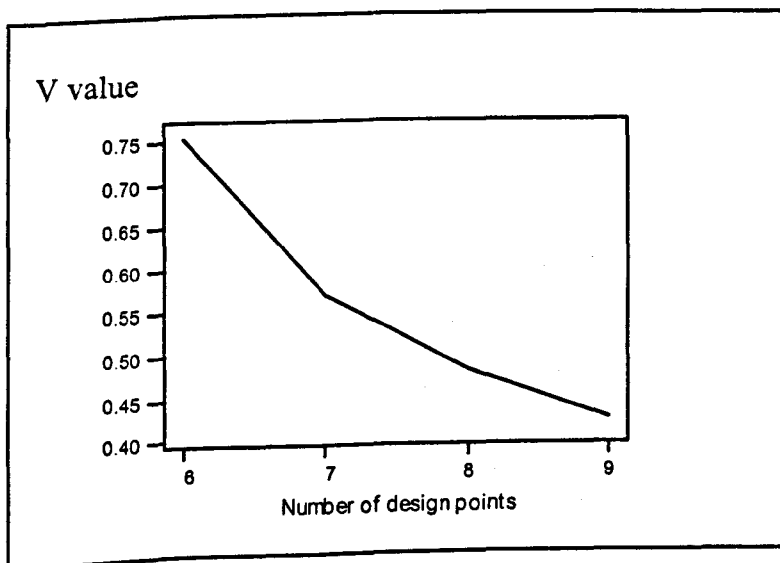
would perform at least 7 distinct trials. To investigate the benefits of performing more than 7 distinct trials, a scree plot of V values, for V optimal designs with between 6 and 9 points, is given in Figure 4.5. The V optimal designs were generated by gosset (Hardin & Sloane, 1993). The scree plot shows that the improvement achieved by adding an additional point, decreases sharply after 7 points. Consequently it was decided that a design with 7 distinct points would be used. As time was limited we decided to produce 3 of these 7 trials once. This led to a replication problem with d=7, r= 3, n=10.

4.7.3 Results

Using the method described in chapter 3 (see §3.5.4), the continuous V optimal design for the JM example was estimated. The continuous optimal design had 8 support points and the optimal 10 point design is likely to have two replicates. Because the optimal 10 point design is likely to have 2 replicates this example is likely to allow a restricted comparison of the methods being considered (recall that R3 was no longer part of the investigation).

The full candidate set which can be generated by DX4 was used here. The candidate set has 31 points. To generate the bases, mkl was used with the following parameters. Correlation coefficients and p-vectors were obtained using the remaining 5 modification methods with both efficient and random bases. The p vectors and correlation coefficients led to similar conclusions on the relationship between V values of bases and final designs, as were obtained for D values. Due to the similarity in conclusions no results are given for the random bases.

To generate the good starting bases mkl was used (see §4.2.3), with the parameters given in Table 4.9. For the V criterion the number of bases was increased from 10 to 28. The parameters of mkl were selected by trial and error so that approximately 40 bases were generated.

Table 4.9 Parameters used to generate multiple bases in mkl using the V criterion

Design Points	Number of Runs	k	l	Random starts (t)	Number of Bases
7	20	4	6	2	39
10	20	5	7	3	36

Table 4.10 Summary statistics for the JM example extension methods for the V criterion

Method	Mean	St-dev	Min	Max	r_c
R1	0.3912	0.0417	0.3466	0.5281	0.970
R2	0.3936	0.0409	0.3492	0.5385	0.966
R4	0.3907	0.0409	0.3466	0.5281	0.959
R5	0.3695	0.0225	0.3466	0.4337	0.511

p-vectors

R1   1 2 3 4 5 6 7 8 9 11 10 12 13 14 15 17 23 16 26 18

R2   1 2 3 4 5 6 8 7 9 10 11 12 13 14 15 18 24 16 25 17

R4   1 2 4 3 5 6 7 8 9 10 11 13 15 12 16 14 18 21 22 17

R5   3 1 27 2 23 4 16 31 5 7 6 28 9 19 14 21 32 33 30 25

Recall that the aim is to find the design with the minimum V value. The best design generated by R7 had a V value of 0.3469. The first thing to be noticed from the table is that the method of Snee (1979), i.e. R2, failed to locate the best design. The design located by R2 was 99.3% efficient compared to the best design. Although this drop in efficiency is negligible it is likely that as the number of components increases this discrepancy could also increase. The best design was located by R1, R4 and R5. Even though R5 had been found to perform poorly for the D criterion examples, it is not surprising that it performed well here because the third best basis, which led to the best final design, included three replicates. The best design located by R7 was 99.9% efficient, which is also a negligible difference. Again R7 was used on designs with more components before conclusions about its performance were made.

As with the D criterion the relationship between the V values of bases and final designs was strong for R1, R2 and weak for R5. Using R4 the relationship was weak for the D criterion but can be seen to be strong for the V criterion. For all methods using very good bases the relationship is even stronger for the V criterion than it is for the D criterion. For the V criterion the best basis led to the best design when methods R1, R2 and R4 were used. It is

suggested that the use of multiple bases may be more appropriate for the D criterion than for the V criterion.

R7 was used with the following parameters (k=5, l = 6, t=50). The design produced by R7 was identical to the best design located by the extension methods. Out of the points which were replicated in the best design, two were midedges and one was a third of edge. The singletons in the best design consisted of three vertices and one triple blend. This selection of points suggests that a larger variety of points were required for V optimal designs with r replicates than for unrestricted V optimal designs (see 3.9.5).

4.7.4 Comparison of R2 and R3

R2 is likely to produce better designs than R3 because each additional point is based on all previous updates. It is possible for R2 to produce a worse design than R3. This will happen when the addition of the a'th best point for the b'th update stage (any a, b >1) would ultimately have led to a better final design than the addition of the best point at the b'th update stage. If the best sequentially updated design uses the r points which would be the r best selections at the first updating stage, both methods will produce identical designs.

An initial aim for this section is to show that R2 can produce a worse design than R3 in practice. A comparison of the two methods is of interest, because if we find that R3 is appreciably better than R2, in some situations, an amalgamation of the two methods should be used. If we find that R2 is always better than R3, then R3 does not need to be considered further. An algorithm was written to use both methods for each start, and then calculate the number of times where the sequential method led to a worse design. As neither the good bases or the 49 random bases provided any examples of R3 improving on R2, the combined algorithm was used with 999 random starts. For 10 out of the 999 random starts R3 produced a better design than R2. The final V values and percentage improvements are given in the Table 4.11 below.

Table 4.11 V values for final designs and improvements gained by using R3

V value	2.82	3.01	3.63	14.69	23.11	24.27	32.69
%	1.5659	2.1990	4.9162	6.3077	3.4355	4.2298	1.0426
V value	51.86	73.23	584.72				
%	1.3471	1.0691	1.0875				

Table 4.11 shows that R3 can sometimes improve on the result obtained by R2, but that the improvements appear to be restricted to poor designs. This restriction to poor designs may be because random designs are more likely to produce poor designs. Because the additional computer time required to run both methods is insubstantial, it is suggested that the two methods should be amalgamated if this approach were to be used. The amalgamation involved using both approaches to generate a design separately, and then choosing the best of them. In the algorithm used for the five component example below (see §4.8), the two methods were amalgamated.

4.7.5 Comments on the Best Design Generated

The points in the best design generated for this example are given in Appendix 1.3. The 4 singletons are comprised of three vertices and one triple blend. The 3 replicates are two edge midpoints and one third of edge. This design compares to those in chapter 3 in a similar way to how the D optimal designs generated above compared to the trends observed in chapter 3. Again the points are spread reasonably evenly around the feasible region and an edge-third is included. However, an additional difference is that this design also includes a triple blend (see §3.2.1). It is even more surprising that the centroid is not included here, as in chapter 3 we found that the centroid was weighted more highly in V optimal designs than in D optimal designs. A more detailed comparison with the designs of chapter 3, which looks at all of the designs generated in this chapter, is given below (see §4.9.2).

4.8 A 5 Dimensional Example for the V Criterion

4.8.1 The Design Problem and Algorithms Used

The JM design problem, which was discussed in the previous section (see §4.4.1), was not large enough to indicate any difference between the seven methods. In this section a five component examples is considered for R1, an updated version of R2/R3, R4, R5 and R7.

The five component example of Snee (1979), which was used for the D criterion (see §4.6), was again used here. However the following constraints, which were originally used in the Snee paper, were also used.

$$x_4 + x_5 \in (0.18, 0.26),$$

$$x_3 + x_4 + x_5 < 0.35.$$

Note also that as $\Sigma x_j = 0.997$, not 1, x_j was scaled to $x_j/0.997$, so that $\Sigma x_j = 1$. A reduced candidate set with 128 points was used. This candidate set can be found in the code for the klrv algorithm. Using the method described in chapter 3 (see §3.5.4), there were found to be approximately 40 support points with the centroid receiving a weight of approximately $\frac{1}{6}$. A 26 point design with 6 replicates was searched for. This is a replication problem of type 2 (see §4.2.2), with n=26, r=6, d=20, s≈40, p=15. Bases were generated using mkl with the following parameters: runs =100, k=7, l = 10, t=5.

4.8.2 Comparison of the Algorithms

Table 4.12 shows that R2 again produced a design which was as good as the best design produced by R1. It can also be seen that the use of multiple bases led to a better design for all of the extension methods. The design generated by the best basis was 99.9% efficient compared to the design generated by multiple bases. As with the D criterion the drop in

Table 4.12 Results for the Snee (1979) example with q=5.

Method	Bases	V
R1, R2, R4, R5	1 (best)	0.3199
	35	0.3185
R7	n/a	0.3134

efficiency is unlikely to have practical consequences but does suggest that multiple bases are likely to be required to find the optimal design.

The design generated by R2 was as good as the design generated by R4. In the cases considered R2 took considerably less time to run than R4 and appears to be the better choice.

From Table 4.12 it can be seen that R7 produced the overall best design. The best design which was generated using extension methods was only 98.4% efficient compared to the design generated by R7.

4.8.3 Comments on the Best Design Generated

The points in the best design generated for this example are given in Appendix 1.4. The 16 singletons are 9 vertices, 6 mid points of edges and one face centroid. The 5 replicates are all face centroids. This design is more similar to the n point optimal designs in chapter 3 than the previous 3 examples of designs which include r replicates. Again the points are spread reasonably evenly around the feasible region. However, in this design the only edge points are mid edges. The other points, including 5 replicates, are face centroids. The only real difference between this design and the designs generated in chapter 3, is that here, face centroids are used instead of the overall centroid. Although some of the points used here were face centroids rather than the overall centroid, the use of face centroids suggests a tendency to use points which are not at the extremities of the region. Therefore, the fact that there are 6 face centroids, five of them being replicated, is in line with the high weightings observed for the centroid in the V optimal designs of chapter 3.

4.9 Conclusions

4.9.1 Performance of the Algorithms

Depending on the number of design points, the number of replicates required and the weights of the support points in the continuous optimal design, many discrete optimal designs include a satisfactory number of replicates. In some cases, especially if the number of support points is large, methods for generating good designs with r replicates may be needed. It is recommended that the points selected for replication are distinct from each other, so that the estimate of internal error (see §4.1.2) can be applied to more than one area of the feasible region.

For the examples considered where $q > 3$, the best results were obtained using an exchange algorithm which only considered designs with r replicates which are distinct from each other, i.e. the R7 algorithm. In one case the best extension method was only 86% efficient compared to the design generated by this specially constructed algorithm. This

approach, of only considering designs with r replicates, has the advantage that the continuous optimal design does not need to be known. Unlike the extension methods (see §4.2.3), in the design generated by R7, the points specifically chosen for replication will be distinct from each other, irrespective of the support point weights. For the algorithm which only considers design with r replicates, it is suggested that all three types of exchange should be considered simultaneously. This is suggested so that good exchanges which involve all three types of point can be identified.

If an implementation of the specially constructed exchange algorithm were not available, good designs with no constraints on the number of replicates can be extended into good designs with r replicates (see §4.2.3). The remainder of this section is concerned with these extension methods. With the extension methods, the continuous optimal design needs to be generated, to verify whether the r points selected for replication are all distinct support points. If the continuous optimal design were not known, the best design at each intermediate stage would need to be checked, so that if the points selected for replication were not distinct from each other, a less efficient design in which they were distinct could be selected from that stage instead.

The starting design which is to be extended is referred to as a basis (see §4.2.3). It was found that the relationship, between the criterion values of bases and final designs, was relatively strong but not monotonic. The relationship was found to be stronger for the V criterion than for the D criterion. For the majority of examples, the use of multiple good bases improved on the design which was generated from the optimal basis. This improvement is likely to be because the relationship between the D values of the bases and final designs was strong but non-monotonic. In all but one example the Snee (1979) method produced designs which were identical to those produced by the more extensive extension methods. In the one example where the Snee (1979) method was bettered, the design generated by the Snee method was more than 99% efficient.

If a design which is highly efficient is adequate, and the values of d and r are such that R1 is too computationally intensive (see §4.5), it is recommended that the Snee (1979) approach should be used with multiple bases. In a few isolated cases a modified version of the Snee (1979) approach, which selected the points for addition non sequentially (see §4.2.3) improved upon the design which was generated by the sequential method. Because the non-

sequential method is not computationally intensive, it is advised that the sequential and non sequential approaches be combined if the idea used in the Snee method is used.

If we want to locate an efficient design for a real experiment then the Snee (1979) method is appropriate. However, if the optimal design is required it is suggested that the Snee method will not always be sufficient. If r is small then the recommended alternative is the R1 algorithm. The R1 algorithm takes a (n-2×r) point design and considers all possible choices of r replicates. If r is large then R1 is computationally intensive and the R4 algorithm is the recommended alternative. The R4 algorithm uses efficient (n-2×r) point designs as the candidate set for KL.

4.9.2 The Form of Optimal Replicate Designs and Comparisons with the Designs of Chapter 3

As chapter 3 restricted attention to q < 5, comparisons between the q=5 examples of this chapter, and the designs in chapter 3, may be flawed. However, in chapter 3 the same types of points were used for q=3 and q=4. The below comparison assumes that the same types of points would have been used for q=5.

In chapter 3 the continuous optimal designs for the quadratic model had their points spread reasonably well throughout the feasible region. The points used were vertices, midedges and the overall centroid. The weight of the centroid tended to be much higher for the V criterion than it was for the D criterion. In the replicate designs of this chapter the points were again spread evenly around the feasible region. However, many of the other trends were different.

The most notable difference is that in the replicated designs the overall centroid was never used. In the q=5 designs the overall centroid appears to have been replaced by face centroids. As was observed with the centroid weights in chapter 3, here, the most V efficient design has considerably more face centroids than the most D efficient design (11 compared to 2).

Another difference was that the points used in the replicate designs were more varied. Thirds of edges were used in three of the designs and a triple blend occurred in one. However, in line with the designs of chapter 3, the majority of the points were still vertices or midedges. As this chapter relied on new and unproved methods, it is possible that the occurrence of these new points was because the best designs generated were not optimal.

Finally it must be noted that considerably more examples would need to be looked at before we could make any conjectures on the overall form of replicated designs.

4.9.3 Further Work

If the optimisation problem discussed in this chapter were investigated further, it is suggested that as well as considering many more examples, the following approach, which is similar to the one used in chapter 3, should be taken. For the constrained case, the D optimal design with r replicates replicates each of the points evenly. With this result in mind, the design region could gradually be changed in a similar way to that used in chapter 3, and the way in which the optimal replicate design changed, could be investigated.

Chapter 5

Experimental Methods and Preliminary Trials

5.1 Introduction

The previous chapters have been concerned with mixture experiments. A mixture experiment is one in which certain properties of a product depend on the relative proportions of its ingredients (see §1.1). Chapter 1 also discussed the use of glaze slip for forming a protective and decorative coat on ceramic ware. Glaze slip is a suspension of frit and clay particles in water. There are various methods for applying glaze to ceramic ware (see §2.5.7). Different application methods require glazes with different viscosities, i.e. differing capabilities for resisting motion when a force is applied. The viscosity of a glaze depends on the relative proportion of each ingredient which it contains. The analysis of glaze viscosity therefore defines a mixture experiment.

The aim of chapters 5 to 7 is to provide a model for predicting glaze viscosity using a variety of mixture and process variables (see §1.2.2). The model which is to be fitted is for use by JM. This chapter discusses some preliminary trials which mimic the industrial process used by JM. These trials are split into two stages. Stage 1 included just 2 trials and was conducted at the JM plant (see §5.2). Stage 2 included 11 trials and was conducted in the Department of Engineering Materials at the University of Sheffield. Some modifications to the JM process were made to ensure accurate experimentation using the University equipment (see §5.3). The aims of stage 1 were to become familiar with the production and measurement process and to investigate the effects on viscosity of storing glaze samples for short periods of time. The aim of stage 2 was to provide data for modelling. Unfortunately the response data from stage 2 was unsuitable for modelling and further trials were required (see chapter 6) for modelling in chapter 7. Both stages of the preliminary trials helped to clarify a suitable experimental protocol and to identify possible causes of experimental error (see §5.2.9). For this reason, they are described here, though the final models were developed based on the data in chapter 6 only.

Sections 2 and 3 of this chapter discuss the methodology used in the preliminary trials. Much of this methodology was also used in the main trials (see chapter 6). Section 4 of this chapter looks at two experimental designs generated for stage 2 of the preliminary trials. Section 5 then gives a basic analysis of the viscosity measurements made in the preliminary trials.

When the preliminary trials were being analysed, it was believed, by the supervising employee at JM, that the best measures of glaze rheology were viscosity at the low and high shear rates of $150s^{-1}$ and $1050s^{-1}$ (see §2.5.7). By the time the main trials of chapter 6 were being planned, it was realised that other rheological properties were more important. These alternative properties are discussed in chapter 2 (see §2.5.9).

5.2 Stage 1 Experimental Protocol used in the JM Laboratories

5.2.1 Materials and Processing Conditions and Outline of the Experimental Protocol

As discussed earlier, the core ingredients of a glaze slip are frit, water and clay. These ingredients are referred to as mixture variables, as it is their proportions which are of interest not their individual masses. Frit was itself produced from a mixture of oxides and has mixture properties of its own. As the same frit is used in all of the experiments discussed in the thesis, it is treated as a single mixture variable. Many other materials can be added to a glaze but here attention is restricted to these core ingredients. Ranges for these core ingredients, within which JM is interested, are given in Table 5.4 (see §5.4.2) where their effect on the design is discussed.

In addition to the mixture variables there are also two processing conditions which are of interest. Each processing condition will be represented by a *factor variable*. Factor variables are used to represent qualitative variables rather than quantitative variables. The first factor variable is the particle size of the glaze slip. Although particle size is a continuous variable, the aim was to take one sample from anywhere within each of the ranges of interest. The particle 'size' of a sample is specified by the proportions of particles below 14 μm maximum dimensions. The ranges used were 65% to 70%, 70% to 75%, or 75% to 80% of particles with this maximum size. These ranges will subsequently be referred to as the 'coarse', 'medium' and 'small' particle size respectively. Particle size measurements of this

type can be obtained from the majority of laser particle sizers (see §5.2.5). The particle size is represented by a factor variable as JM typically specify glazes by the range in which its particles' size should be, rather than by giving a precise value. The possibility of aiming for narrower ranges was considered, but was rejected because it is not standard procedure in the industry.

The second factor variable of interest is whether a *flocculant* had been added to the slip. Flocculants are substances which promote attraction between particles. A 10% by mass aqueous calcium chloride solution was used as a flocculant in the trials discussed below (see §5.2.6). As flocculants are widely used to modify glazes after production, their effect is of interest.

A possible third non-mixture explanatory variable is how long a sample has been stored for. Measurement of viscosity was therefore repeated after intervals of 1, 24 and 48 hours. This factor is of interest as the majority of glazes sold by JM are aqueous suspensions whose particles tend to settle during storage. Settling causes samples which are taken from the top of the jars to have a higher percentage of water, and hence have a lower viscosity.

The seven stages of the experimental procedure are summarised in Table 5.1. Each stage is explained in detail in sections 5.2.2 to 5.2.7.

5.2.2 Stage 1. Weighing the Ingredients

To calculate the required mass of each core ingredient the total mass was multiplied by the relevant proportions. The total mass should be approximately the maximum mass which will still leave enough room in the mill for efficient grinding. This is often determined by a process of trial and error. The total mass used here was 1.5kg, this having been suggested by a JM employee as being an effective amount for the mills being used. The appropriate mass of each ingredient was measured using a balance (Sartorius, maximum 2kg tolerance 0.01g, calibrated weekly). It should be noted that water was measured by mass not volume. This balance is believed to be highly accurate because it is calibrated weekly. Consequently the proportions which were actually milled, are likely to be extremely close to those specified in the design.

Table 5.1 Summary of the seven steps of the experimental procedure used by JM

```
┌─────────────────────────────────────────────────────────────────┐
│ 1) Weigh the required amounts of frit, water and clay (see 5.2.2)│
└─────────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│ 2) Ball mill ingredients (see 5.2.3)                             │
└─────────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│ 3) Sieve and take samples for particle size measurement (see 5.2.4)│
└─────────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│ 4) Ensure particle size is in the largest required range and take two│
│    samples (see 5.2.5)                                           │
└─────────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│ 5) Add flocculant to one and shear mix (see 5.2.6)               │
└─────────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│ 6) Measure rheological properties (see 5.2.7)                    │
└─────────────────────────────────────────────────────────────────┘
                              │
                              ▼
┌─────────────────────────────────────────────────────────────────┐
│ 7) Repeat stages 2 to 6 for finer particle sizes                │
└─────────────────────────────────────────────────────────────────┘
```

5.2.3 Stage 2. Ball Milling

Specially hardened ceramic balls known as *milling media* were placed along with the ingredients in a ceramic ball mill jar (Pascal model 5612), which had a capacity of 1 gallon. For efficient grinding a range of different sized milling media are required in the mill (see §2.5.6). The jar was then rotated for an initial period of 16 hours at 65 revolutions per minute using a ball mill rotator (no make or model number available). This is the standard milling time which is used by JM for ball mills of this size. As the jar rotates the ingredients are gradually ground into smaller particles through collisions with the media (see §2.5.6).

5.2.4 Stage 3. Sieving and Taking Samples

After 16 hours of ball milling all lumps of glass frit which had not been ground sufficiently were removed from the mixture using a 120 μm mesh sieve (Endecott test sieve BS410 HSSS125 nominal aperture size 125μm, mesh size 120μm). The glaze was poured

into a plastic bucket (Fisons 10litre YRW 250 050K), through the sieve, aided by a lightly vibrating sieve shaker (Gosling & Gatensbury). The sieved mixture is known as a glaze slip. Sieving was done at this stage to ensure that the sample was representative of the mixture which would be used in rheology measurement because sieving is always performed before rheology is measured (see §2.5.8). A small sample of the sieved slip, approximately 2ml, was then transferred into a sample jar (Fisherbrand 15ml glass jar BTF600010M), and taken for particle size measurement (see §5.2.5). A gap of 10 minutes was left between use of the sieve shaker and measurement of viscosity, this was to encourage the structure of the sample to return to equilibrium.

5.2.5 Stage 4. Measuring Particle Size

The machine used by JM for measuring particle size is the *Malvern Mastersizer*. The Malvern is used for sizing in both laboratory trials and large scale industrial production of glaze. The Malvern consists of an electronic measurement unit, a small well of water into which a small amount of the sample is dropped and a computer link. It must be noted that this water well is separate to the water bath used in stage 2 of the preliminary trials (see §5.3.4) A simple diagram of it is given in Figure 5.1.

Figure 5.1 The Malvern Mastersizer
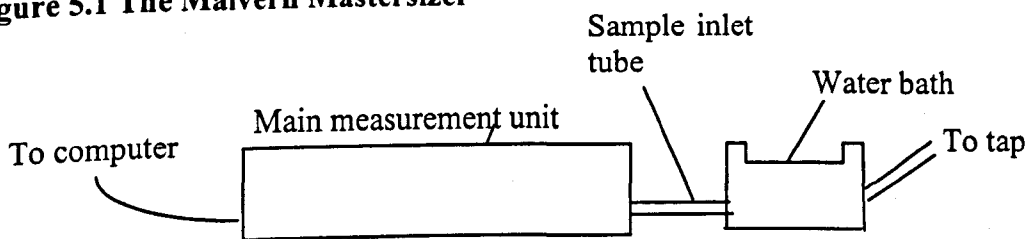
The operating procedure is summarised in Table 5.2. Each stage is described in detail below.

Notes relating to Table 5.2

Obscuration This is a measure of the (volume) percentage of solid particles in the water well. It must be noted that the quoted value is the not the percentage itself but a value proportional to it.

Agglomerates These are formed when the original particles lump together to produce bigger particles.

Table 5.2 Summary of the procedure for particle size analysis of glaze slip samples

a) Automatically rinse the water bath

b) Automatically align particle detectors

c) Measure particle size distribution of impurities in well (tap) water

d) Ensure zero obscuration

e) Add small amounts of glaze slip sample until obscuration $\in$ (0.15, 0.17)

f) Perform test

g) If agglomerates are detected repeat process
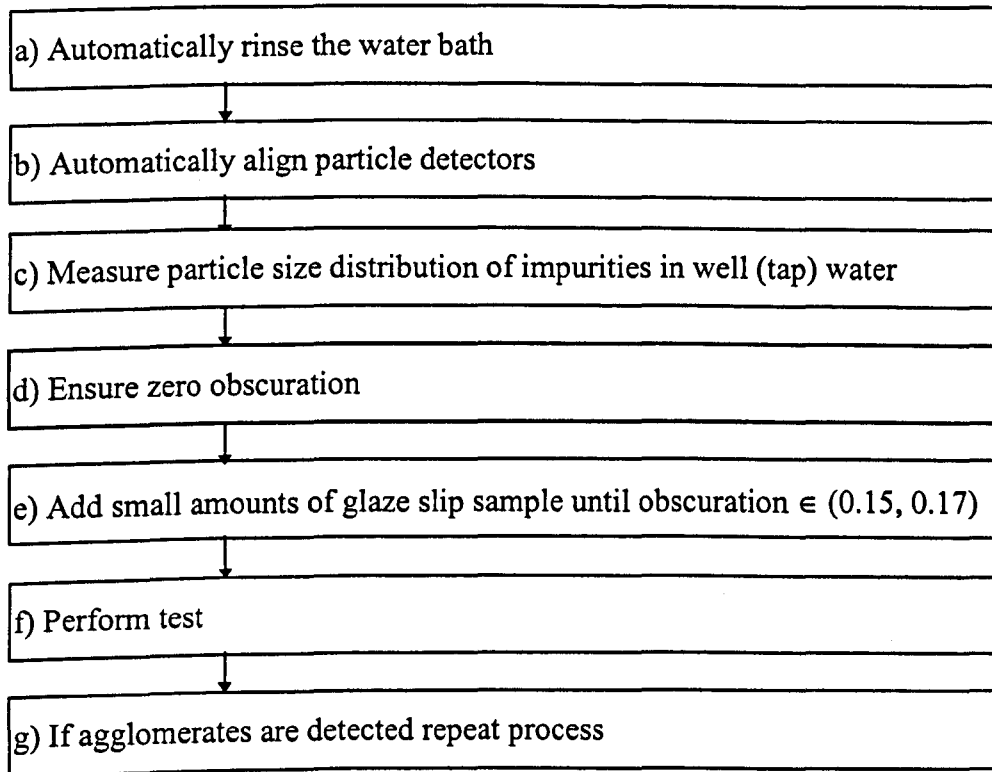
Details of individual stages

a) Rinsing of water bath

This was performed to ensure that no particles from previous samples remained in the bath. Any remaining particles would have distorted the measurement of the new sample. This was an automated procedure which was initiated using the computer link. The procedure involved filling the bath with tap water and then slowly emptying it. This was performed three times and took approximately 5 minutes.

b) Align particle detectors

The measurement process involved passing a low energy laser beam through the sample and detecting it on the opposite side. If the beam generator and detector were not perfectly in line then the reading would have been inaccurate. The detectors were easily disturbed while the machine was not in use. Two possible causes of disturbance were slight knocks to the machine and changes in temperature. The alignment was an automated procedure and was initiated using the computer link.

c) Measure the particle size distribution of impurities in well water, which is taken from the tap

To measure the particle size of a sample it was dropped into a well of water so that the particles were spread out and distinguished from each other (see stage e below). Because the water in the well came from the tap it contained various impurities in the form of solid particles and dissolved ions which would distort measurement of the particle size of a sample. If the background measurements were not removed from the sample data, the results would be non-representative of the glaze slip.

The particulate impurities and ions (but not ultimately the interaction of the tap water ions with the glaze slip) in tap water were accounted for by measuring the (background) particle size for the well water before the sample was added. The result of this measurement is known as the *background particle size*. The measurement of the well water was performed using the same procedure as for the measurement of the sample (see stage f below). Together with the obscuration reading, the background particle size was used by the computer to remove the unwanted measurements for the impurity particles in the water. This correction procedure involved splitting the range of particle sizes into classes. For each class the counts for the background water supply were subtracted from the counts for the glaze sample.

d) Ensure zero obscuration

The procedure of accounting for the background particle size (see stage c above), was based on a small sample from the well and was not exact. Consequently the more impurities there were in the bath, the less accurately their measurements could be removed from the sample measurements. A negligible amount of impurities had to be observed before sufficiently accurate measurements could be performed. When carrying out the experiments if

the 'obscuration value' for the background water was greater than zero the experiment was restarted.

e) Add small amounts of glaze slip sample until obscuration is in the range (0.15, 0.17)

Once an obscuration reading of zero had been observed, a small amount of the sample taken from the mill, approximately 0.25ml, was added to the well using a spatula. The aim was to add enough sample to bring the obscuration reading within the range 0.15 and 0.17. As only a small amount was needed and it was very easy to exceed this range accidentally. When an obscuration of between 0.15 and 0.17 was observed, the measurement process was initiated using the computer link. If the obscuration reading had been outside this range then the measurement would have been inaccurate by an amount which is proportional to the discrepancy. If this range was exceeded then the entire process was repeated.

f) Perform test

This stage was entirely automated and was initiated using the computer link. The measurement procedure involved passing a low energy laser beam through the sample and then absorbing it into a detector on the other side. The reduction in beam energy depends on the size of the particles encountered in passage through the sample tube. Using a measure of the reduction in beam energy, an estimate of the particle size distribution is derived by the computer linked to the apparatus.

g) If agglomerates detected repeat process

In some cases the experiment had to be repeated due to the detection of agglomerates (large particles caused by smaller particles sticking together). The Malvern output included a particle size frequency plot. Agglomerates were indicated on these plots by additional peaks which occurred beyond the main peak of the distribution. If agglomerates were detected then the measurement had to be repeated. This is because the agglomerates would have contained many of the particles which were initially less than 14μm. As all agglomerates are removed by shear mixing before a glaze is used (see §2.5.7), the measured percentage of particles less than 14μm would not be representative of the glaze slip. This would have led to an increase in experimental error. If agglomerates were detected during the experiments then the bucket of slip was stirred manually with a standard household spatula and a new sample was taken for measurement.

h) Interpreting the particle size measurement

In addition to the particle size frequency plot, output from the Malvern includes the percentage of particles which were less than 14μm in diameter. If the percentage of particles were in the required range (see §5.2.1) then two samples of 200g were placed in one litre polythene sample jars (Nalgene 660 110F), and taken for the addition of flocculant and a viscosity measurement (see §5.2.6 to §5.2.7). To weigh the samples the sample jars (see stage h), were placed on the balance (see §5.2.2) and the reading was set to zero. Glaze slip was then gradually poured into the jars until the required mass was reached (±1g).

If the percentage of particles less than 14μm was outside the required range, the slip was returned to the ball mill and milled for an additional hour. After this hour it was re-sieved and re-measured. If the result was still outside the required range then the amount of additional milling time required was estimated using a linear interpolation based on the first two readings. This interpolation was based on aiming for 67.5% of particles less than 14μm. Techniques for linear interpolation can be found in most standard texts which deal with the mathematical analysis of scientific data. Using these techniques the additional milling time was calculated in the following way, where p_1 and p_2 represent the first and second readings respectively.

Time needed = required decrease in particle size / rate of decrease

= required decrease in particle size / (previous decrease / time taken)

= $(67.5 - p_2) / [(p_2 - p_1) / 1 \text{ hour}]$

= $1 \text{ hour} \times (67.5 - p_2) / (p_2 - p_1),$

If the particle size was below this range already, then measurement at the coarse particle size was not possible and analysis continued at the medium particle size.

The above procedure for milling and measuring (see §5.2.2 to §5.2.5), was repeated to produce samples from the other two ranges of particle size (see §5.2.1). Any slip which remained after the six samples had been taken was stored in case further analysis was required.

5.2.6 Stage 5. Adding Flocculant and Shear Mixing

Flocculant was added to one of the 200g samples from each particle size. The flocculant used in this experiment was a 10% by mass calcium chloride aqueous solution., and the amount of solution added was 0.05% by mass of the glaze's solid mass. Because this was such a small amount the procedure, which was recommended by JM, and was used for addition to the sample was to place the sample on the balance and add drops from a disposable dropping pipette (Fisherbrand 1ml PMK355 030X) until the mass had increased by an appropriate amount. The mass of flocculant was not included in the subsequent modelling as it was believed to be negligible compared to the mass of the sample. Instead flocculant was treated as a processing condition and specified using a qualitative factor variable. Immediately after addition, each flocculated sample was mixed using a free standing shear mixer (Parvalux). The shear mixing was performed for 10 seconds at the mixer's maximum speed of 4 000 revolutions per second. The non-flocculated sample was not shear mixed. The samples were then taken for viscosity measurement (see §5.2.7).

5.2.7 Stage 6. Measuring Viscosity of Glaze Slip Samples

Viscosity was measured using the *Haake M10 Viscometer with MV IV Sensor* shown schematically in Figure 5.2. Measurement was performed immediately after glaze production and then again after storage periods at ambient of 1, 24 and 48 hours (see §5.2.1). After each storage period shear mixing was performed to redistribute some of the settled particles. This did not contravene the investigation of settling because all slips are shear mixed prior to application (see §2.5.7), but not all of the settled particles are redistributed.

The viscometer consisted of a temperature control unit, a sample cup, a sensor which was suspended from a frame, a control panel and a computer link. A diagram of this apparatus is given in Figure 5.2.

It is known that viscosity can vary with temperature. Unless there is interest in quantifying the effect of temperature, all samples should be tested at the same temperature, to allow effective comparison. A temperature of 25°C is usually used by employees of JM and this temperature was used in the trials which are discussed in chapters 5 and 6.

The fact that future glazes will not necessarily be used at 25°C is not necessarily problematic. The purpose of the model is to predict standard (25°C) viscosity from the

ingredients, to investigate whether the rheological properties, of a glaze with a given formulation, are likely to be appropriate for the chosen application procedure.

Figure 5.2 Schematic diagram of the Haake viscometer
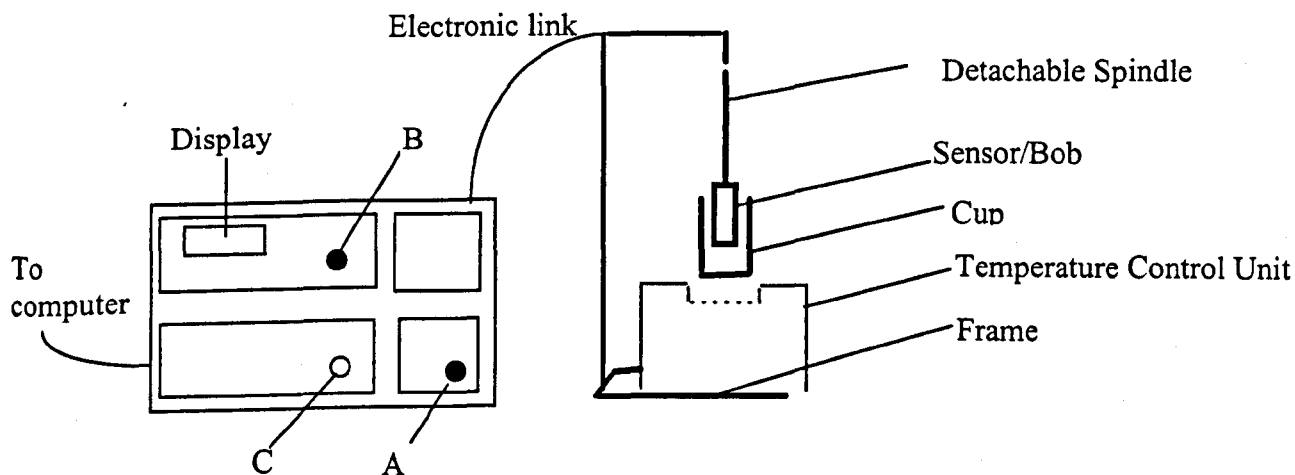
Table 5.3 Parameters used for the Haake viscometer measurement of glaze slip viscosity

Measurement type	From (s^{-1})	To (s^{-1})	Time (minutes)	# readings
$\tau / \dot{\gamma}$	0	1100	0.65	50
τ / t	1100	1100	1	10
$\tau / \dot{\gamma}$	1100	0	0.65	50

Outline of viscosity measurement procedure

1) The measurement parameters shown in Table 5.3 were inputted to the computer. The symbols in this table have the following meanings; τ (shear stress), $\dot{\gamma}$ (shear rate), t (time). The entry after the "/" in the "measurement type" column indicates which variable was to be varied, the entry before the "/" indicates which variable was to be measured while this variation was taking place. The columns labelled 'from' and 'to' indicate the range of values between which the shear rate (γ) was varied. The time column indicates the length of

time over which the changes should be spread. The final column indicates that 110 viscosity readings were taken for each sample measured. These measurement parameters mean that the viscosity is measured 50 times while the shear rate is being increased, ten times while it is kept constant and then another 50 times while the shear rate is being decreased.

It should be noted that these were the initial input parameters suggested by JM. It eventually transpired that these settings did not yield all of the information required (see §6.2.1), and they were radically modified for the trials discussed in chapter 6 (see §6.2.8)

2) 35ml of sample was poured into the viscometer cup and the sensor was lowered in. Additional sample was added until a small amount overflowed onto the top of the sensor. The sensor was smaller than the cup and none of the sample was lost due to this small overflow.

3) The shear stress setting (τ) was selected using the display selection control (A on Figure 5.2). The sensitivity control was then set to 10% (B on Figure 5.2), and the displayed viscosity reading was adjusted to zero using the fine tuning control (C on Figure 5.2). The sensitivity was then returned to the 1% level.

4) The measurement process was initiated using the computer link.

5) Glaze is known to be *thixotropic*. This means that when it has been exposed to a sufficiently high shear stress, its viscosity is decreased in the long term. The high shear stresses involved in the measurement process are known to alter virtually permanently and radically the rheological properties of the sample. Because of this the sample could not be used for further measurement and so it was discarded.

6) To prepare the apparatus for the next sample the cup and sensor were washed using distilled water and dried thoroughly using a cotton cloth.

Output from the Haake viscometer included a table of the following values: shear rate, shear stress, viscosity and time.

5.2.8 Possible Causes of Bias and Experimental Error, and Proposed Solutions.

It is important that all variables and processing conditions which are not being investigated are kept constant during the experiment. These variables and processing conditions are subsequently referred to as *external conditions* or *covariates*. If the external conditions vary bias can occur and the standard deviation of measurements can be high. If

bias and standard deviation are high, any conclusions obtained are less likely to reflect the true situation accurately. In general the procedures used by the company do little to standardise external conditions. In this section we outline five possible causes of bias or experimental error which were identified in the stage 1 trials. In stage 2 and the main trials of chapter 6 every effort was made to eliminate these problems.

A) When a glaze sample is left at rest, solid particles slowly sink to the bottom. This sinking is referred to in the industry as *settling*. The presence of clay and flocculant tend to reduce settling but do not prevent it entirely. In all subsequent trials any sample which was left at rest for longer than 5 minutes was shear mixed, prior to measurement, in an attempt to homogenise the distribution of the solid particles.

In stage 1 the non-flocculated samples were measured before the corresponding flocculated samples. This order was chosen due to the higher risk of settling in the non-flocculated samples. As the gap between the sieving and the measurement of the flocculated samples was still only 5 minutes, the flocculated samples were not shear mixed prior to the initial measurement. The fact that the non flocculated slip was not shear mixed might have caused the sample to have a higher proportion of water than specified and a lower viscosity. This discrepancy might result in downwards bias in the coefficient of the flocculant parameter. This bias problem was subsequently tackled by shear mixing all samples prior to measurement (see §5.2.6).

B) Between measurements dust can settle in the viscometer cup and cause an artificial increase in the viscosity of the next measurement. If the non-flocculated sample is always measured first the dust will be a possible cause of upwards experimental bias in the coefficient of the flocculant parameter. To prevent this bias in all subsequent experiments a coin was tossed to decide which sample to measure first. This randomisation was performed until one type of sample had been measured first, for half of the proposed pairs. After this point all remaining pairs were measured in the other order.

C) Every time that slip is moved there is a possibility that air bubbles can enter it. The majority of these air bubbles can be removed by stirring the sample with a standard household spatula for approximately one minute prior to measurement. This procedure was used in all subsequent trials.

D) Viscosity can vary considerably with temperature. JM test all glazes at 25°C. This temperature is maintained using the temperature control unit attached to the Haake

viscometer. As this unit operates to a high degree of accuracy, changes in room temperature did not affect the measurement procedure. The stage 1 trials were conducted at the JM plant and made use of the temperature control unit. The stage 2 trials, which are described in the next section, were performed in the Department of Engineering Materials at the University of Sheffield. Because the University viscometer did not have a temperature control unit, temperature was maintained at 25°C using a water bath (see §5.3.3).

E) If residues of a previous sample are left on the viscometer cup or bob, future viscosity measurements can be distorted. If the residue is still aqueous it might blend into the new sample and alter its viscosity. If the residue had solidified, lumps could have broken away and impeded rotation of the bob because there was only a 3mm gap (see §5.2.7). This impedance would have been interpreted by the viscometer as the sample's resistance to flow, thus increasing the measured viscosity (see §5.2.8). To reduce the likelihood of this happening the cup and bob are washed stringently after each test. The approach used by scientists at JM was to rinse them with water and then dry them thoroughly with a cloth (see §5.2.7). It was realised that there was a possibility that cloth fibres could be left on the equipment. These fibres could act in a similar way to the break away lumps and again impede rotation. In addition to this, if the drying were not done thoroughly, any remaining water would mix into the next sample and reduce its viscosity. To remedy the problem of residual fibres in all subsequent experiments, after being rinsed with water, the cup and bob were sprayed with acetone (Fisher Chemicals CAS 67-64-1). Acetone was used to solubilise and wash away the water. The acetone was then evaporated using the cool air setting of a conventional household hairdryer, which was only likely to cause a negligible change in cup temperature.

5.3 Modifications to the Experimental Protocol for the Stage 2 Preliminary Trials Performed at The University of Sheffield

5.3.1 Introduction

The seven step experimental structure, which is outlined in Table 5.1, also formed the basis of stage 2 of the preliminary trials. However, a number of changes were made to the following 4 steps:

1 Weighing the mixture ingredients (see §5.3.2).

2 Ball milling (see §5.3.2).

5 Adding flocculant and shear mixing (see §5.3.3).

6 Viscosity Measurement (see §5.3.4).

Other procedures were unaltered.

5.3.2 Revised Steps 1 and 2. Weighing the Mixture Ingredients and Ball Milling

It is known that if approximately 70% of the mill is filled, there will be enough room to allow successful grinding (see §2.5.6). To estimate the weight of ingredients which would fill 70% of the mill, the frit, water and clay were gradually added to the mill in the required ratio (see Table 5.4) 200g at a time. Together with the milling media, 2.2kg of the ingredients was found to fill approximately 70% of the mill. This weight was different to the one used at JM (see §5.2.2) because the total volume of the milling media was different.

Once filled appropriately, the mill jar was rotated for an initial period of 16 hours at 65 revolutions per minute, using a ball mill rotating device (Crypto G(R)/3AT2). All other details about the weighing and ball milling of the ingredients were unaltered from sections 5.2.2 and 5.2.3

5.3.3 Revised Step 5. Adding Flocculant and Shear Mixing

The protocol for adding flocculant that was used in the JM laboratories (see §5.2.6) was also used in Sheffield. The shear mixing procedure varied and is described below.

Shear mixing was performed to redistribute any glaze slip particles which may have settled on the bottom of the sample jar during storage (see §5.2.1). In industry glazes are always shear mixed before use to ensure that all particles are evenly distributed (see §2.5.7). Because the aim of the experiment was to mimic the industrial process the shear mixer was used prior to each viscosity measurement taken after each period of storage. The mixer was also used to distribute flocculant throughout the mixture as soon as the deflocculant was added. To standardise the pre-measurement conditions of each sample and to decrease the chances of experimental bias (see §5.2.8A) shear mixing was performed before the measurement of both flocculated and non-flocculated samples. In the protocol of section 5.2

the non-flocculated sample was always measured first. It was discussed in section 5.2 that this may have led to experimental bias (see §5.2.8B). To prevent this bias the first sample to be measured was selected using the toss of a coin (see §5.2.8B).

It is desirable to measure all viscosities at a constant temperature (see §5.2.7). The University viscometer did not have a temperature control unit. Consequently all samples were stored in a water bath at 25°C between measurements (see §5.2.6). If the samples to be shear mixed had been removed from the water bath and taken to the mixer, the temperature of the mixer blade may have altered the temperature of the sample. In this case the sample would have needed time to regain the required temperature when placed in the viscometer sample cup. During this time additional settling would have occurred and might have distorted the viscometer measurements. Consequently the blade and spindle of the mixer were detached and placed inside the sample cup to attain a temperature of 25°C for half an hour prior to use. The sample jar, spindle and blade were then moved over to the mixer frame so that the mixing could be performed. The shear mixing was performed for 10 seconds at the mixer's maximum speed of 2,000 revolutions per minute.

5.3.4 Revised Viscosity Measurement

It is known that even small changes in temperature can have a noticeable effect on viscosity. There was no interest in quantifying this effect within this project (see §5.2.1), therefore all samples were tested at the same temperature to allow reliable comparisons (see §5.2.7).

Between measurements the samples were stored in plastic jars (Nalgene 660 110F). Each jar was labelled according to how the glaze it contained was produced. Each label contained a code which consisted of the run numbers given in Table 5.9 and the process variable codes given in Table 5.10 (see §5.5.2). Trials are subsequently referred to by these labels.

It was originally suggested by an employee of JM that the viscosity be tested at 25°C and that the samples be stored at 20°C ±5°C. As there was no temperature control unit on the Haake viscometer this was facilitated by storing the viscometer cup and sample jars in a water bath. Storing the samples at one temperature and then measuring their viscosity at a different temperature would have been problematic. When shear mixing had been performed and the samples were poured from the jars into the viscometer cup up to an hour would be needed for

them to adjust between the two temperatures. This waiting time would have allowed settling to occur and would probably have led to inaccurate measurement. To solve this problem without buying a new viscometer with a temperature control the samples were stored at the measurement temperature of 25°C. In stage 1 and in the main trials discussed in chapter 6, the samples did not need to be stored at a constant temperature because the temperature control unit acts almost instantaneously.

The water bath contained an internal heater which maintained temperatures above room temperature. During the summer the room temperature exceeded 25°C. Consequently a cooling coil was submerged in the bath to reduce temperature below 25°C, and allow the heater to operate. The coil required a constant flow of water which was provided from a laboratory tap via high strength rubber tubing. The temperature of the bath was checked every half an hour at with a thermometer (Fisons THL 210-051R 76mm −10°C-100°C). No deviation from 25°C was ever detected.

The sample jars and viscometer cup were kept stationary in the water bath by a specially constructed aluminium rack which fitted into the bath. The rack was constructed from an aluminium offcut with a diameter of 4mm. A schematic drawing of the rack constructed is given in Figure 5.4. Each sample jar was stored in the water bath for 3 days and the slip viscosity was re-measured; 1, 24 and 48 hours after the initial measurement. Each slip was shear mixed prior to each additional measurements.

Monitoring Sample Temperature

Recall that the viscosity measurement involved rotating the bob at high speeds (see §5.2.7). These speeds can generate heat and increase the sample's temperature. Immediately after each test the sample's temperature was measured and found to equal 25°C ±0.3°. This indicates that the temperature stabilisation procedure was successful and it provided a relatively high level of accuracy considering the *adhoc* method used (see §5.2.6). A possible explanation for this success is that the cup was made from stainless steel which is 8mm thick. Because the cup was constantly stored in the water bath and stainless steel has relatively high thermal conductivity, its temperature was likely to be kept constant and it would have quickly absorbed any heat generated.

183

Figure 5.4 Diagram of metal rack to secure sample jars

NB. Not to scale.

a) Plan view.

Lip to fit over side of water bath (internal width 14mm)

Holes for sample jars (radius 35 mm)

120mm

160mm

b) Side elevation.

120mm

Lip to fit over side of water bath (internal height/width 8/10mm)

Holes for sample jars
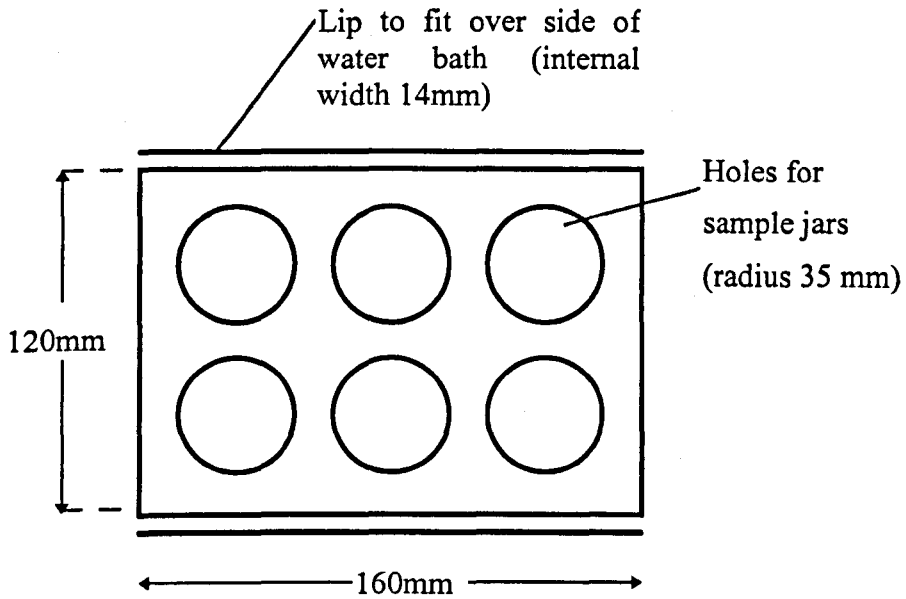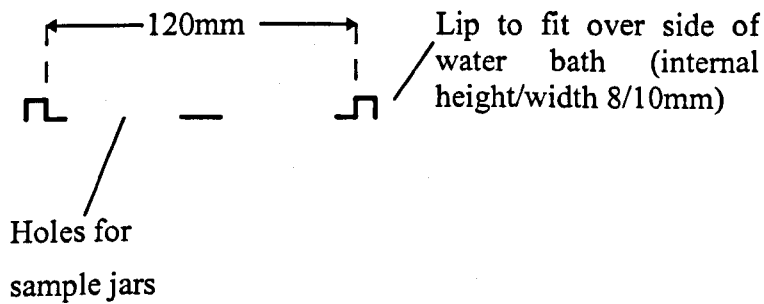
5.4 Mixture Constraints and Designs Generated

In this section we consider the designs generated for the preliminary trials. Recall that design optimality is model dependent. The starting model which was originally to be used in the was the quadratic. All designs discussed in this section are generated with respect to the quadratic model.

In stage 1 of the preliminary trials (see §5.2) two mixtures were analysed in the JM laboratories. The mixture formulations (see Table 5.6 in §5.5.1A) were suggested by

employees of JM as standard formulations used at the plant. No experimental design theory was used at this stage.

Stage 2 of the preliminary trials was performed in the Department of Engineering Materials at the University of Sheffield. It was decided in chapter 4 that we would analyse 7 distinct compositions; 4 singletons and 3 doubletons. An appropriate design was generated. This design was not constrained to include the two mixtures from stage 1, because they would only have provided two extra trials, and would require a blocking variable and a non-constant variance structure, to explain the differences between the two sites.

During the production and measurement of the compositions in the initial design, it was discovered that many of them were too viscous to be measured by the Haake (see §5.4.3A). The failure of the Haake occurred after the first two trials had been analysed successfully. Consequently the feasible region was modified and a new design, constrained to include these two points, was generated.

5.4.1 Using a Product Design

The explanatory variables of interest are three mixture variables and two process variables (particle size and flocculant). As the two process variables have 3 and 2 levels respectively, there are 6 different processing conditions. The processing conditions do not affect the constrained region for the mixture variables.

The most time consuming aspect of the experiments was milling each glaze formulation to the first particle size. The time required to mill the slips to the other two particle sizes, and add flocculant, were relatively small. Consequently it was decided to produce a sample for each of the six processing conditions for each of the mixture compositions. This structure means that we are using a product design for a split plot experiment (see §2.3.6C). As the starting model in the analysis was to include no interactions between the mixture and process variables, the product design will be optimal, if the mixture design is optimal (see §2.3.6C).

The remainder of this section is only concerned with mixture designs. When the trials were performed, the mixture designs generated were crossed with the 6 processing conditions to form product designs. The drawback of taking six samples from each glaze slip milled, is that within-batch errors are likely to be correlated. Appropriate analysis is used in chapter 7.

5.4.2 Design Generated for the Original Constrained Region

The original bounds on the mixture components which were suggested by JM are given in Table 5.4. Using the R7 algorithm, with ks = ls = 3, kr = lr = 2 and t = 100 (see §4.3) the design given in Figure 5.5 was generated. Doubletons are indicated by larger circles. It must be noted that Figure 5.5 depicts an L transformed feasible region (see §2.3.1B). The L transformed regions were used so that the designs generated could be compared with the general trends observed for continuous optimal designs in chapter 3. The real and L transformed co-ordinates are given in Table 5.5.

Table 5.4 Weight constraints originally suggested by JM. Material codes used by JM are given in brackets.

Name of component	Lower Wt %	Upper Wt %	Symbol
1 Frit (261887)	51	66.5	x_1
2 Water (1000)	30	40	x_2
3 Clay (2622)	3	10.5	x_3

Figure 5.5 Design generated for the original constrained region

The dotted line indicates the multi-component constraint imposed later (see §5.4.3A)
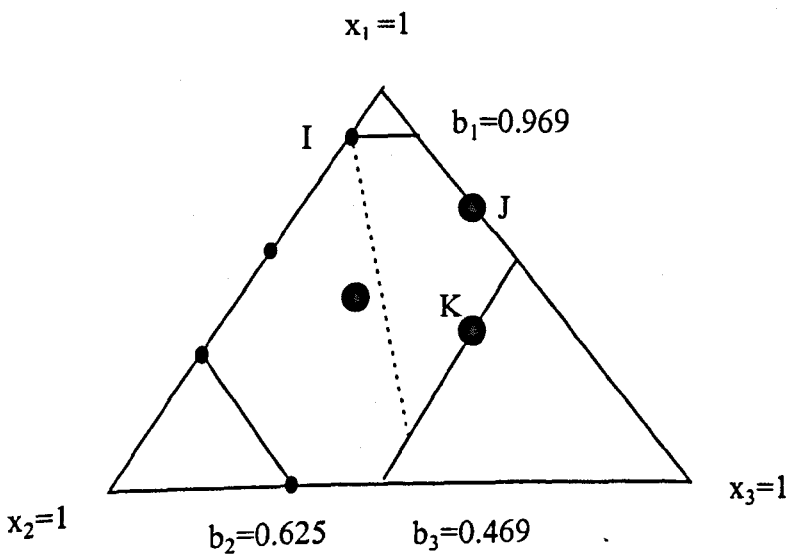
Table 5.5 Design generated for the original constrained region

Singletons	Actual co-ordinates			L transformed			Type
	x_1	x_2	x_3	x_1	x_2	x_3	
	0.5700	0.4000	0.0300	0.3750	0.6250	0.0000	vertex
	0.5100	0.4000	0.0900	0.0000	0.6250	0.3750	vertex
I	0.6650	0.3050	0.0300	0.9688	0.0312	0.0000	vertex
	0.6175	0.3525	0.0300	0.6719	0.3281	0.0000	mid-edge
Doubletons							
J	0.6300	0.3000	0.0700	0.7500	0.0000	0.2500	mid-edge
K	0.5667	0.3283	0.1050	0.3542	0.1771	0.4688	edge-third
	0.5858	0.3484	0.0658	0.4270	0.3021	0.2240	centroid

Using the algorithm written in chapter 3 (see §3.5.4) it was determined that the continuous optimal design for this region has 11 support points: all six vertices; the four mid points of the long edges and the overall centroid. The design in Figure 5.5 is now compared to the general trends observed for continuous V optimal designs in chapter 3 (see §3.9).

In accordance with these trends the points are reasonably well spread over the feasible region; edge points and an internal point were chosen as doubletons, and the edge points are on the three long edges. However, the design differs from the general trends because 3 of the 6 vertices are not included. These discrepancies probably occurred because it is desirable to spread the points as evenly as possible over the feasible region, and because, while there are 11 support points in the continuous optimal design there are only 7 distinct support points in the exact design. Points on three of the four long edges were probably included because they tend to have higher weights than the vertices. It is probable that the edge chosen to include no internal points was selected because both of its vertices were included.

5.4.3 The Design Generated for the Modified Constrained Region

A Modifications to the constrained region

For mixtures with a high clay content and a low water content the viscometer tended to cut out when the induced resistive force (see §5.2.8) exceeded a certain magnitude. This meant that no more measurements could be taken. This failure occurred for points J and K in the design given in Figure 5.5. The Haake indicated that a maximum shear stress had been reached by a τ (shear stress) reading of 100%.

It was discussed in chapter 1 that for a glaze to be of use to a ceramics manufacturer, each ingredient must be within certain bounds. There are no strict rules about these bounds but approximate ones which the company usually work within are given in Table 5.4. When the Haake failure occurred, additional discussion with JM experts was initiated, to identify whether the original feasible region correctly represented all known constraints. At this stage (and in accordance with the mixtures for which failure had occurred) it was suggested that in addition to the individual bounds, the lower bound for water should depend on the proportion of clay. High proportions of clay should lead to an increase in the lower bound for water, to prevent a glaze being too viscous for measurement. Although mixtures which include clay at its upper bound could be measured, and mixtures which included water at its lower bound could be measured, a mixture which attained both of these bounds, or values which were both near to them, was likely to be too thick for measurement by the Haake viscometer. Consequently it was suggested by JM that the experimental region should be restricted further. An appropriate restriction on the relative proportions can be described by a multi-component constraint of the general form

$$l \leq a.x_1 + b.x_2 + c.x_3 \leq u \qquad (5.1)$$

If the coefficients of the multi-component constraint were selected properly it was envisaged that the problem with the viscometer cutting out would be solved and the constrained region would still contain all mixtures of real interest.

As failure occurs when the proportion of clay is too high, compared to the proportion of water, the constraint should be of the form:

$$b.x_2 - c.x_3 \geq l \qquad\qquad (5.2)$$

To maximise the area of the constrained region, the value of l should be chosen to be as small as possible, while still being high enough to prevent the viscometer from cutting out. Prior to the failure of the Haake, mixture I of Figure 5.5 had been measured successfully; with a τ reading of 99% being observed. The difference between the proportion of water and the proportion of clay for mixture I, was 0.275 (see Table 5.5). Because the analysis of mixture I produced what appears to be the highest permissible value of τ, the following values were chosen: $l = 0.275$, b=c=1. The resulting multi-component constraint is drawn on the diagram of the original constrained region (see Figure 5.5).

B The design generated

Because the multi-component constraint changed the feasible region quite radically (see figure 5.5 above) a new design was needed. Two compositions had already been analysed successfully so the new design was constrained to include them. The R7 algorithm, with ks = ls = 3, kr = lr = 2 and t = 100 (see §4.3) was used to generate an appropriate design with 8 distinct points; 5 singletons and 3 doubletons. The number of distinct points was increased to 8 because there was sufficient time to analyse 9 further trials rather than 8. The design generated is given in Figure 5.6. It must be noted that Figure 5.6 depicts the L transformed feasible region. The real and L transformed co-ordinates are given in Appendix 2.1.

Using the algorithm which was written to locate continuous optimal designs over a given candidate set (see §3.5.4) it was determined that the continuous optimal design for this region has 9 support points: all five vertices; the three mid points of the long edges and the overall centroid. The discrete design is now compared to the general trends observed for continuous V optimal designs in chapter 3 (see §3.6).

As was observed for the stage 2 design, before the multi-component constraint was added, in general, the design concurs with the trends observed in chapter 3. Although the same deviations and similarities were again observed here, two additional deviation were also

observed. Firstly, the centroid carried over from the previous design was very close to the mid point of the longest edge. Consequently none of the points on this edge were included in the design. Secondly, the centroid of this modified region was not included. Again, this centroid was probably omitted because it was so close to the original centroid. Additionally, point L was probably selected to provide a good spread of points over the region. The mid point of this edge was probably too close to the two internal doubletons.

Figure 5.6 Design generated for the modified constrained region
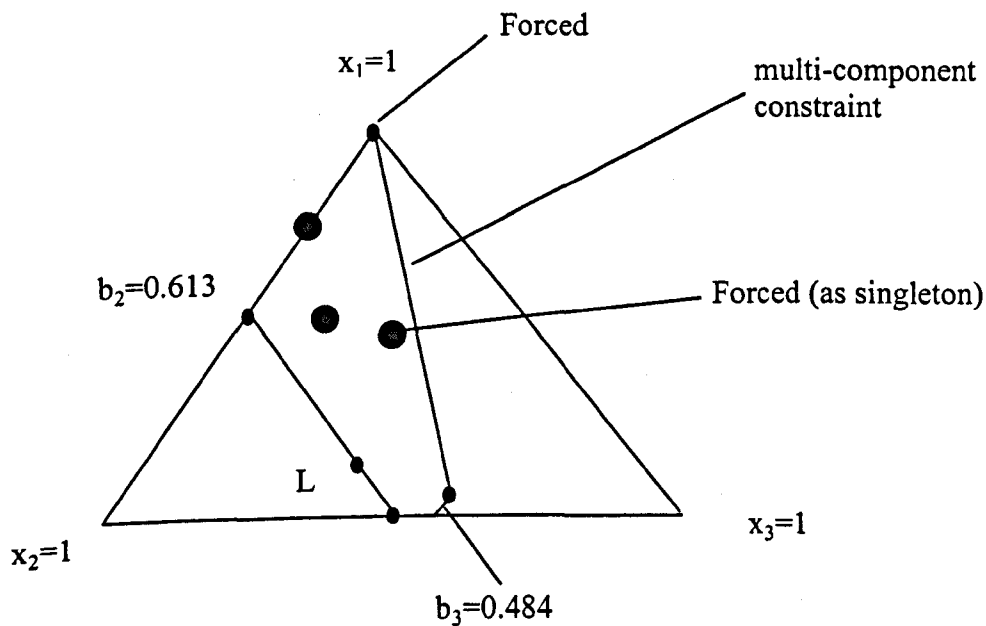
After the 11 trials in the design for the modified region had been performed, it was decided that one of the singletons would be used as a further experimental trial, in an attempt to further understand the experimental errors which had been observed. It was found that making point L a doubleton led to the design with the lowest V value. This choice is not surprising as edge points tend to have higher weights than the vertices, and the other 2 long edges already had doubletons on them or near them.

5.5 Analysis of the Preliminary Trials

5.5.1 The Two Trials of Stage 1

A Trials analysed

Some rheological properties of two different glaze slips were measured. The two mixture specifications are given in Table 5.6 and are referred to as JM1 and JM2 respectively. Both mixtures were milled to the coarse particle size range. These exploratory trials were limited to this size because the experimental time available was limited. Both batches of mixture were split into two samples and flocculant was added to one of them. A total of four different slip formulations was therefore analysed.

Table 5.6 **Mixture specifications for exploratory trials.** The JM product code is given in brackets

	Frit (261887)	Clay (2622)	Water(1000)
Mixture JM1	0.607	0.036	0.357
Mixture JM2	0.568	0.09	0.342

The additional measurements could not be taken after 1, 24 and 48 hours as was the original intention (see §5.2.1) because the Haake viscometer was not always available. After 3 hours of storage all samples were shear mixed and measured. The samples were then stored for 4 days, shear mixed again, and re-measured. Measurement without shear mixing was then repeated 30 and 90 minutes later. This gave five data sets for each of the four slip formulations.

B Graphs of viscosity against shear rate at various time points

Graphs of viscosity against shear rate were plotted to investigate the effect of leaving samples in storage:

a) the two readings from day 1 only

b) the three readings from day 4 only and

c) graphs to compare the first measurement on days 1 and 4.

Three separate graphs were plotted for each of the 4 experimental conditions.

The sets of graphs for each of the four slip formulations exhibited similar patterns. Only the curves for mixture JM2 with flocculant are included here (see Figure 5.8). This formulation was chosen arbitrarily. All graphs exhibited an extremely high viscosity, approximately 500 Pa.s, for the lowest shear rate. This high viscosity was always followed by a rapid reduction followed by a flat viscosity plateau as the shear rate was increased. The graphs only show the viscosity measurements which were taken while the shear rate were being increased. The measurements taken while the shear rate was kept constant, and then while it was being decreased (see §5.2.7), are not plotted.

Figure 5.8 Viscosity curves for mixture JM1 with flocculant

a) Day 1 Immediately (solid), +3hrs (dashed)

Viscosity (microPa.s)

Shear rate / s^{-1}
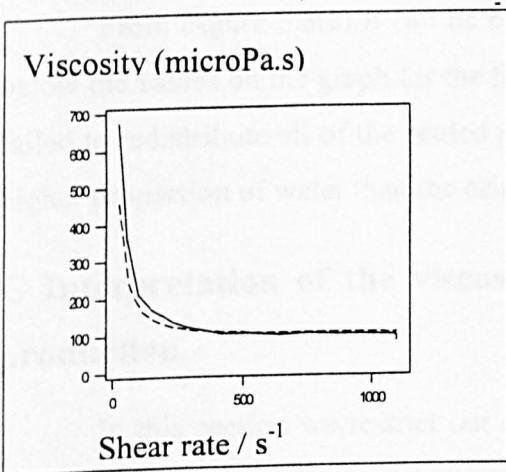
b) Day 4 Immediately (solid) +30mins (dashed), +90 mins (dotted)

Viscosity (microPa.s)

Shear rate / s^{-1}
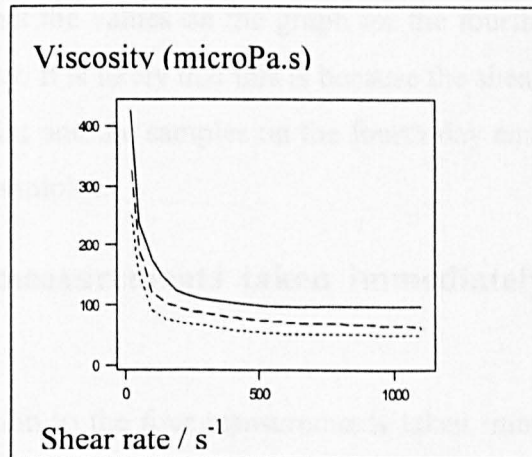
c) Immediately after shear mixing. Day 1 (solid), Day 4 (dashed)

Viscosity (microPa.s)

Shear rate / s^{-1}
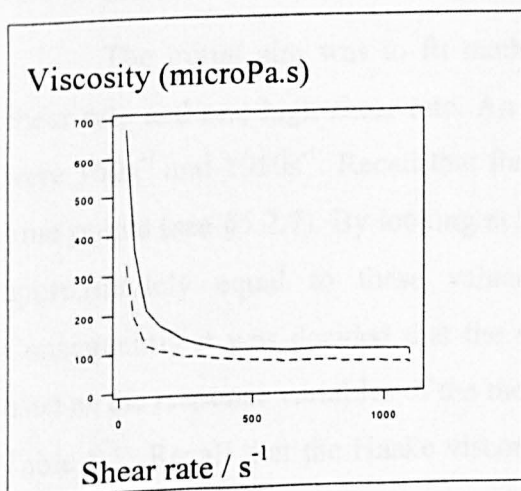
Recall that on day 4 the samples were only shear mixed prior to the first measurement. From Figure 5.8(b) it can be seen that on day 4 the curves get lower the longer they are left to stand. The most likely explanation for the curves getting lower is that progressively more solid particles were sinking to the bottom during storage, causing an increase in the proportion of water in the samples taken for measurement. The lowering of the curves can then be attributed to the low viscosity of water. By looking more closely at Figure 5.8(b) it appears that the effect of the additional settling decreases with time.

From Figure 5.8(a), which shows the measurements for day 1, a similar effect is observed but the drop in viscosity appears to be less. In the later stages of the graph when the flat plateau is reached, the curve for the second measurement rises above the first. The smaller reduction in viscosity on day 1 is likely to be because shear mixing was performed to redistribute the settled particles. The fact that a decrease still occurred in the initial stages suggests that the shear mixing failed to redistribute all of the settled particles. The only explanation of why the second curve rose above the first on day 1 is experimental error.

From Figure 5.8(c) it can be seen that the values on the graph for the fourth day are below the values on the graph for the first day. It is likely that this is because the shear mixing failed to redistribute all of the settled particles and the samples on the fourth day contained a higher proportion of water than the original samples.

C Interpretation of the viscosity measurements taken immediately after production

In this section we restrict our attention to the four measurements taken immediately after production, because the main aim is to investigate how the mixture formulation affects viscosity, not to investigate the effect of settling.

The initial aim was to fit mathematical models for predicting viscosity at one low shear rate and one high shear rate. An employee of JM decided that appropriate shear rates were $150s^{-1}$ and $1050s^{-1}$. Recall that for each sample of slip the viscosity is measured at 110 time points (see §5.2.7). By looking at the full set of readings it was observed that shear rates approximately equal to these values were measured at the 8th and 45th readings. Consequently it was decided that the viscosity at the 8th and 45th measurement would be used as the response variables of the model. The relevant viscosity measurements are given in Table 5.7. Recall that the Haake viscometer measures the shear rate for a given shear stress

193

(see §2.5.10). This means that the 8th and 45th viscosity values were not measured at the same shear rate for all four samples. This is not a measurement problem but a consequence of the Haake control system. As the 8th and 45th values were at similar shear rates ($\pm 2s^{-1}$), and this is only an exploratory model, they were considered to be sufficiently accurate by employees of JM.

As only two distinct mixtures were analysed it was not possible to fit models. Therefore, the analysis in this section is based on simple description of the data.

Table 5.7 gives the viscosity readings for the samples with and without flocculant for both mixture compositions at the specified low and high shear rates. Table 5.8 gives the observed increase due to flocculant for each mixture composition at both shear rates.

Table 5.7 Response data (viscosity) from stage 1 of the preliminary trials

Experimental conditions	Viscosity at low shear rate (i.e. 150 s^{-1}) (micro Pa.s)	Viscosity at high shear rate (i.e. 1050s^{-1}) (microPa.s)
JM1 no floc	53.74	46.55
JM1 floc	62.38	60.57
JM2 no floc	93.38	88.73
JM2 floc	169.60	96.46

Table 5.8 Difference in viscosity (Pas) between samples with and without flocculant

	Difference for Low shear rate	Difference for High shear rate
Mixture JM1	8.64	14.02
Mixture JM2	76.22	7.73

Figure 5.9 Results of stage 1 of the preliminary trials

Low shear rate (- -), High shear rate (—)
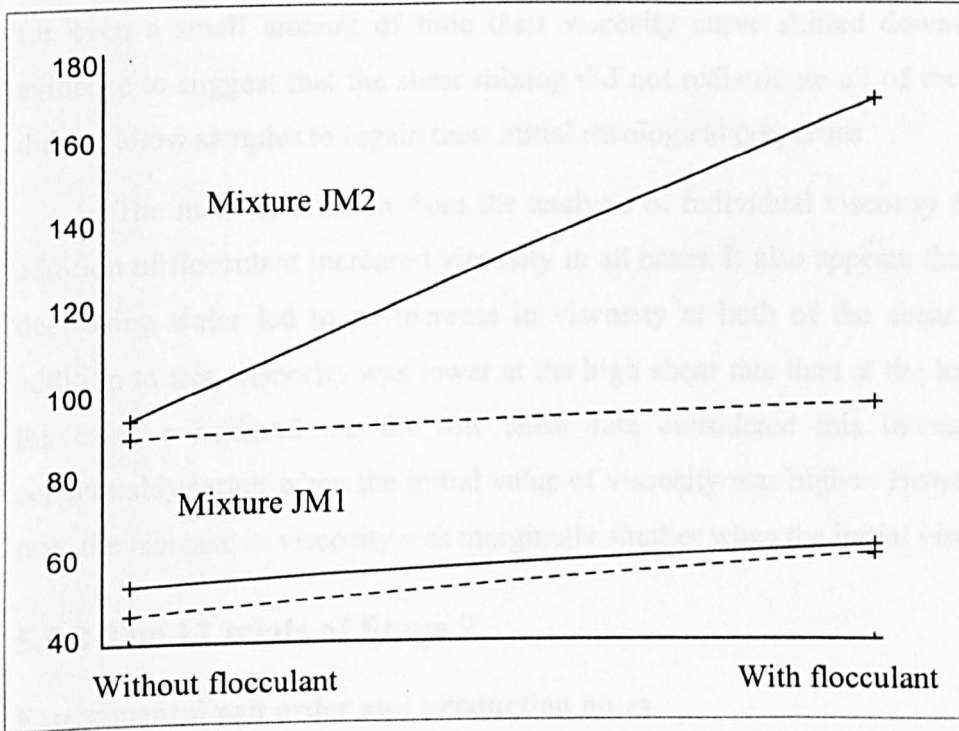
Interpretation of tables and graphs

The first thing to be noted is that in accordance with known results on glaze rheology (see §2.5.8) the viscosity was lower at the high shear rate than at the low shear rate.

From Table 5.6 it can be seen that mixture B had more clay but less frit and less water than mixture JM1. The tables and graphs show that changing the composition in this way led to a higher viscosity for both levels of shear rate. It can be said that mixture JM2 had a comparatively high viscosity compared to mixture JM1. This is consistent with theory because it is known that increasing the amount of clay and/or decreasing the amount of water in a glaze slip will increase its viscosity (see §2.5.1).

It can be seen that adding flocculant increased the viscosity for both mixture compositions and at both levels of shear rate. At low shear rates flocculant appears to increase viscosity considerably more, when the initial viscosity of a glaze slip is higher.

D Conclusions

The main conclusion from the above graphs is that if samples were allowed to stand for even a small amount of time their viscosity curve shifted downwards. There is also evidence to suggest that the shear mixing did not redistribute all of the settled particles and did not allow samples to regain their initial rheological properties.

The main conclusion from the analysis of individual viscosity readings was that the addition of flocculant increased viscosity in all cases. It also appears that increasing clay and decreasing water led to an increase in viscosity at both of the shear rates considered. In addition to this, viscosity was lower at the high shear rate than at the low shear rate in all of the cases considered. At the low shear rate considered this increase in viscosity was considerably larger when the initial value of viscosity was higher. However at the high shear rate, the increase in viscosity was marginally smaller when the initial viscosity was higher.

5.5.2 The 11 trials of Stage 2

Experimental run order and production notes

It is standard practice when performing an experiment, to conduct the trials in a random order (see §2.3.6A). Table 5.9 specifies the design which was generated for the modified feasible region (see §5.4.3B). The numbers in the last column, headed 'rand', were assigned for the purposes of run order randomisation. A dash in this column indicates that the point was a successfully analysed trial from the original design (see §5.4.2). To prevent repetition, the run order randomisation numbers were included in the same column as the run order they led to. The 10 trials in the initial design had been placed in a random order using a similar process to the one described below. The points with 2 numbers in the first column, or 1 number and a dash, are the doubletons. The run order given in the first column was selected using random numbers generated by Matlab:

6, 3, 4, 2, 9, 1, 7, 5, 8.

The last run to be analysed (run 12 in Table 5.9) was not in the original design and was added after the original 11 trials had been completed. Run 10 was initially over milled and the only samples which could be analysed were those with fine particles. Run 16 was also over milled and the only samples which could be analysed were those with medium and fine particles The remaining runs were produced and analysed successfully. When the data from this analysis

was interpreted it was realised that there were a lot of inconsistencies and experimental errors (see below). In order to understand these experimental errors better, it was decided to perform an additional replicate. Consequently run 12 was added.

Summary of graphs produced

In the analysis which follows, each trial performed in Sheffield is referred to using the run numbers of Table 5.9, and the process variable codes of Table 5.10. The two trials which were performed in the JM laboratories (see §5.5.1) are referred to by the names in Table 5.6.

Table 5.9 Trials performed

Run	Frit	Water	Clay	Rand
1, 8	0.5858	0.3484	0.0658	- , 6
2	0.665	0.305	0.03	-
5, (12)	0.53	0.4	0.07	3
6	0.5157	0.38	0.105	4
4	0.57	0.4	0.03	2
11, 3	0.6175	0.3535	0.03	9, 1
9	0.51	0.4	0.09	7
7, 10	0.5833	0.3617	0.0550	5, 8

For each of the trials a graph of viscosity (η), against shear rate (γ) was produced. As an example the graph for run 8A is given in Figure 5.10.

The graphs in figures 5.10 and 5.11 do not look the same as the graphs plotted for stage 1 (see Figure 5.8 in §5.5.1). Firstly, each of the curves plotted here has two sections. The top and bottom sections show the measured viscosity as the shear rate was increased and then decreased respectively (see §5.2.7). The graphs plotted for stage 1 do not show the bottom section. Secondly, the graphs plotted for stage 1 show rapid decrease before the eventual plateau. This rapid decrease was observed because a very high viscosity measurement was observed for the first shear rate. These single high viscosity measurements

were also observed for each of the trials in stage 2. These high viscosity readings are not plotted here, so that the extent of the random fluctuation can be seen.

Table 5.10 Processing conditions codes

Code	Particle size	Flocculant
A	(Coarse) 65%-70%	No
B	(Coarse) 65%-70%	Yes
C	(Medium) 70%-75%	No
D	(Medium) 70%-75%	Yes
E	(Small) 75%-80%	No
F	(Small) 75%-80%	Yes

The top and bottom sections, in figures 5.10 and 5.11, should not be confused with the top and bottom plateaux in the S shaped theoretical viscosity curve. (see §2.5.9). The two sections are essentially both estimates of the bottom plateau in the theoretical curve (see §2.5.9). The bottom section shows the viscosity as the shear rate is gradually reduced (see §5.2.7). The bottom section is lower because glaze slip is thixotropic (see §2.5.7).

Figure 5.10 A graph of viscosity against shear rate for run 8A
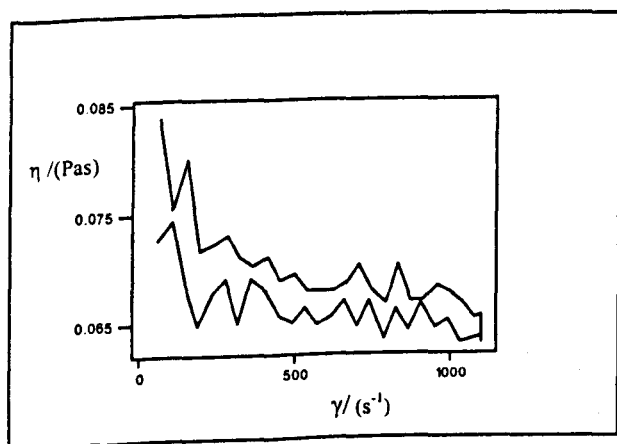
The graphs show the measured viscosity as the shear rate was increased gradually, kept constant for 60 seconds and then decreased gradually (see §5.2.7A). The viscosity is lower during the decrease than during the increase. This is because glaze is thixotropic (see §2.5.7). A typical viscosity curve for glaze slip consists of two plateau separated by a steep descent (see §2.5.9).

In each set of viscosity measurements taken in the preliminary trials, the first measurement for each sample had a very high viscosity compared to the subsequent points. This point is likely to be on the top plateau of the S shaped theoretical curve (see §2.5.9), with the remaining points being on, or near to, the bottom plateau. As the data also appeared to exhibit large amounts of experimental error (see below), estimation of the top plateau from this single value was considered unwise. These initial points are omitted from the subsequent graphs. The graphs are discussed in two sections:

A) Comparing the stage 1 trials performed at JM (see §5.5.1), with the stage 2 trials performed at Sheffield. (See §5.5.2).

B) Investigating inconsistencies in the effects of reducing particle size and adding flocculant.

A Comparing the trials performed at the JM plant with the trials performed at Sheffield

In this section graphs are produced to compare the results obtained when identical formulations were produced at the two different sites. The compositions analysed in stage 2 at Sheffield are given in Table 5.9. The compositions analysed in stage 1 at JM are given in Table 5.6. From these tables it can be seen that run JM1 corresponds to Sheffield run 8. The results from both sites, for trials with coarse particles but without flocculant, and for trials with coarse particles with flocculant, are given in figures 5.11 (a) and 5.11 (b) respectively. The flocculated sample from run 8 is represented by both solid and dotted lines. Different lines were used for this run because the stages of increase and decrease were difficult to distinguish. The solid line represents the viscosity increase stage.

From Figure 5.11 it appears that the Sheffield trials were subject to more random variation than the JM trials. The Sheffield trials also appear to exhibit higher viscosities. The between site differences in viscosity appears to be considerably bigger for the flocculated sample than for the non-flocculated sample.

Both measurements taken at the JM plant appear to exhibit reasonably flat bottom plateaux which could be approximated using straight lines. For the Sheffield sample with

199

flocculant, the extreme variation in the measurements would make it difficult to estimate a reasonable plateau. This excessive variation indicates that there is either a major flaw in the experimental protocol, or that considerably greater random errors were experienced in the Sheffield trials.

Figure 5.11 A graph of viscosity against shear rate for run 8 of the stage 2 trials and run JM1 of the stage 1 trials. The Sheffield trials had identical formulations to the JM trials.

a) No Floc' Sheffield (8A) (–), JM1 (- -)

b) With Floc' Sheffield (8B) (–, ..), JM1 (- -)
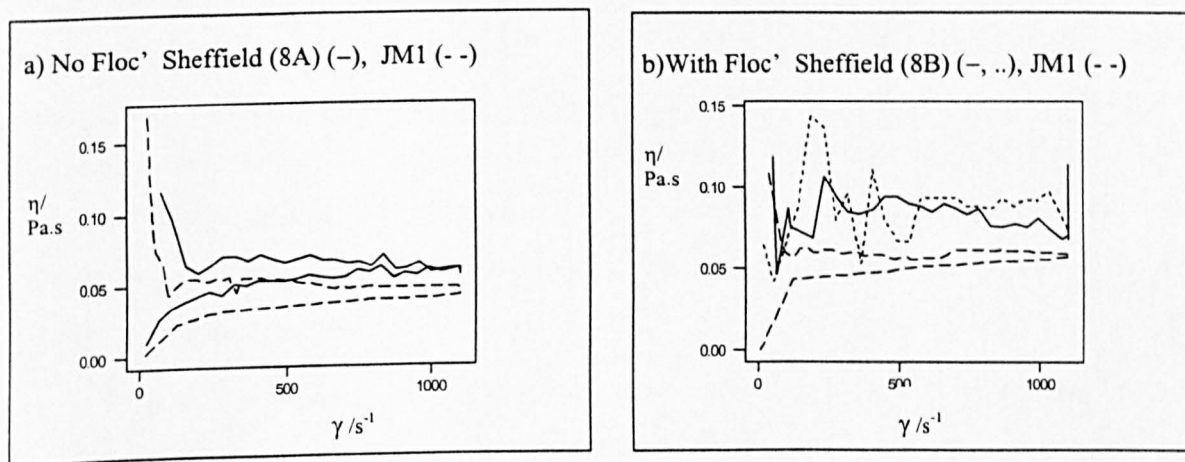
B Investigating the effects of particle size and flocculant

The analysis in this section is based entirely on the results from the 12 Sheffield trials. The three particle sizes to be considered were coarse, medium and fine. To analyse the effect of varying particle size, plots of all three particle sizes were produced for each composition and level of flocculant. This led to the construction of 24 graphs. The graphs are not included here as the effect of increasing particle size which each graph suggested was clear. The majority of the graphs showed that decreasing particle size increased the viscosity. Approximately 10% of the lines showed the unacceptable and apparently random variation observed in Figure 5.11 for run 8 with flocculant.

It is known that decreasing particle size increases viscosity. However 11 of the 24 graphs suggested the counter intuitive conclusion that decreasing particle size actually decreased viscosity.

The effect of adding flocculant was investigated using a similar approach to the investigation of varying particle size. This led to the construction of 36 graphs. These graphs are again omitted as the effects of flocculant were clear. Of the 31 graphs produced, 21

showed that flocculant increased viscosity. However 10 graphs suggested that flocculant decreased viscosity which contradicts theoretical results (see §2.5.1).

The above analysis throws doubt on the validity of the results and suggests that there was inadequate protection against experimental error. Consequently it was decided that the results of these trials were unsuitable for modelling. To provide data for modelling, further trials were conducted at JM. The further trials are outlined in chapter 6, and their results are analysed in chapter 7.

Chapter 6

Full Trials performed at Johnson Matthey

6.1 Introduction

. Chapter 5 discussed experimental trials which were mainly conducted at the University of Sheffield. The purpose of these trials was to produce a mathematical model for predicting glaze viscosity from compositional and process variables (see §1.1) but by the time these trials had been performed the company's understanding of glaze slip rheology had increased and the data collected in these trials were no longer comprehensive enough to model all of the aspects of interest (see §6.2.1). In addition, the information on the bottom plateau was subject to such large experimental error (see §5.5.2), that the validity of any fitted models would be in doubt.

Due to the problems detailed above it was concluded that further trials would be needed. It was decided that these trials would be performed at the JM production plant rather than at the University, due to the experimental error encountered in the Sheffield trials (see §5.5.2) and because at the JM plant multiple runs could be performed concurrently. Chapter 6 deals with the new trials performed at JM. The data from these trials are then analysed in chapter 7. The initial plan for the trials discussed in this chapter was to perform ten runs in two batches of five. These ten runs only included seven distinct formulations and one of them failed to produce usable data (see §6.3.2). With only six distinct runs there would have been little scope for checking the adequacy of models fitted to the data (see §2.2.6). Consequently an additional block of three trials was analysed.

Some of the apparatus available at JM was different from that available at the University. The JM equipment had also changed since the preliminary trials discussed in chapter 5. As additional sources of experimental error were identified in chapter 5 (see §5.2.8 and §5.5.2) the majority of the protocol from these chapters is modified even if the same equipment was available at both sites. Those aspects of protocol which remain unchanged from the previous chapter are omitted here.

6.2 Protocol Modification for Full Trials

6.2.1 Aspects of Glaze Rheology of Interest to JM

Recall that the key property of glaze is that its viscosity, i.e. the coefficient which determines its ability to resist flow (see §1.2.2), depends on the size of the force trying to cause this flow (see §2.5.8). Recall also that a typical graph of viscosity against shear stress has an inverted S shape. This S shape consists of top and bottom plateaux separated by a steep descent (see §2.5.9). The revised opinion of JM was that the most important aspects of the graph are the viscosity of the two plateaux and the value of shear stress at which the top plateau ends. The value at which the top plateau ends is known as the yield point.

6.2.2 Outline of Modifications

A Three problem areas

Recall the three main problems with the results from the trials performed in Sheffield:

(1) They were designed to investigate the relationship between viscosity and shear rate. This is a problem because the relationship between viscosity and shear stress is of more importance (see §2.5.9). It is not a major problem because shear stress can be calculated by multiplying shear rate by viscosity (see §2.5.10).

(2) Not enough measurements were taken at low shear rates and the data sets include only one point on the top plateau (see fig 6.1). This means that the height of the top plateau and the value of the yield point cannot be estimated accurately.

(3) Results for the bottom plateau were subject to such large experimental errors, that the validity of any conclusions drawn from them would be in doubt (see §5.5.2).

B General modifications

Since the stage 1 trials, JM had purchased a new *Bohlin* viscometer. The first problem was easily remedied by using the Bohlin operating software to select shear stress as the input variable.

The second problem occurred because the shear rate was increased linearly from $0s^{-1}$ to $1100s^{-1}$ and the yield stress usually occurs for very low values of shear rate and shear stress. This problem was solved by using the Bohlin operating software to select a logarithmic increase in shear stress causing the majority of readings to occur for low shear stresses.

6.2.3 Protocol Modifications to Correct Experimental Error

To remedy the excessive experimental error, the protocol was radically modified. The modifications are discussed in the relevant sections below but a list summarising them is given here.

1) Use of the more accurate Bohlin viscometer (see §6.2.8).

2) Use of a weaker flocculant solution so that a higher mass would be needed and could be measured more accurately (see §6.2.7).

3) More effective shear mixing (see §6.2.7).

4) Samples shaken prior to measurement to remove residues on sides of jar (see §6.2.8).

5) Use of syringe to measure samples for viscometer (see §6.2.8). Although the syringe will impart a small shear stress on the sample, it was envisaged that the resulting error would be less serious than the error caused by each sample having a different volume.

6) Pre-shearing prior to test to standardise shear history (see §6.2.8).

6.2.4 Outline of Experimental Procedure

The experimental stages which were outlined in Table 5.1, also apply to the new protocol. The table is reproduced here.

Table 6.2 The 7 stages of the experimental procedure

1) Weigh the required amounts of frit, water and clay

2) Ball mill the ingredients

3) Sieve

4) Ensure particle size is in the required range and take 2 samples

5) Add flocculant to one sample and then shear mix

6) Measure rheological properties

7) Repeat stages 2 to 6 for remaining particle sizes
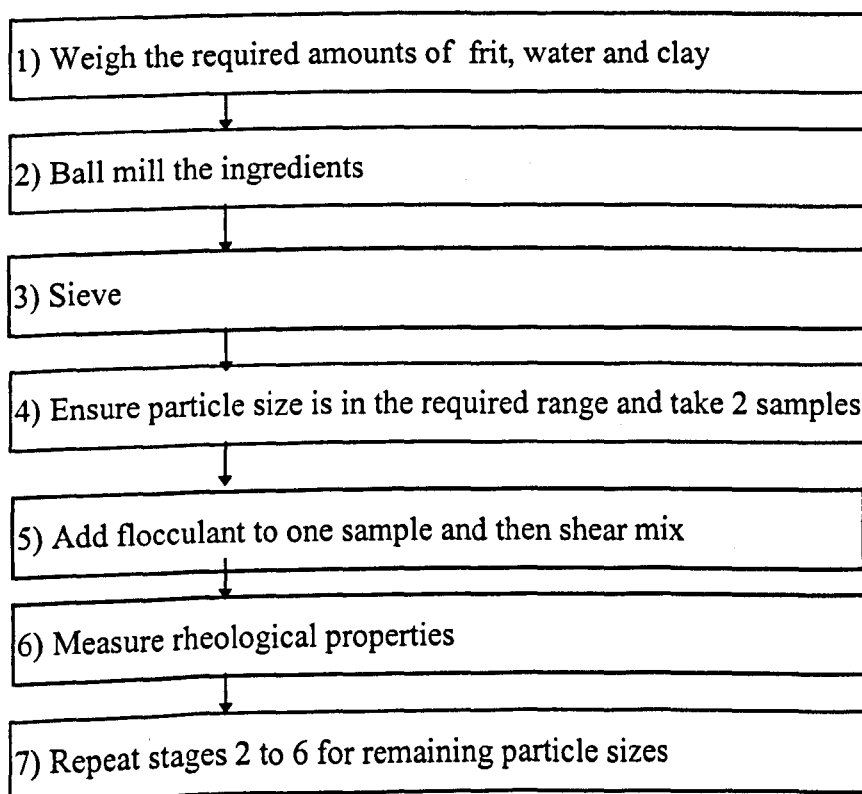
Only modified aspects of the protocol are discussed here.

6.2.5 Stage 1 Weighing the Ingredients

This was performed in a similar way to the method described in chapter 5 (see §5.2.2). The only difference was that half gallon mill jars were used instead of one gallon mill jars (see §6.2.6). The total weight of ingredients was therefore changed to ensure efficient milling (see §2.5.6). Mill jars of this size operate efficiently with 1kg of solids. Because water accounts for approximately 33% of each composition (see Table 6.5) a total of 1.5kg was used in each mill.

6.2.6 Stage 2 Ball Milling

This was performed in a similar way to the method described in chapter 5 (see §5.2.3). The only difference was that the ten runs (see §6.3) were performed in two blocks of five. This was to reduce the time of the experiment. In order to minimise experimental error all five mills were placed on the same rollers. Due to the size of the rollers half gallon ball mill jars (Pascal model 5160) were used instead of the one gallon ball mill jars used for the trials in chapter 5 (see §5.2.3) In contrast to the circumstances in chapter 5 the same media could not be used in each trial (see §5.2.2). In an attempt to standardise the milling as much as possible (see §2.5.6) each mill was loaded with a similar selection of media. To facilitate this the milling media were split into 3 size ranges; large, medium and small. Each mill was then loaded with 675g of the large media, 1250g of the medium media and 675g of the small media.

Sieving and particle size measurement were performed in the same way as was described in chapter 5 (see §5.2.4 and §5.2.5)

6.2.7 Stage 5 Adding Flocculant and Shear Mixing

Flocculant was added using a similar procedure to that described in chapter 5 (see §5.2.5). The mills which were used (see §6.2.6) yield approximately 800ml of glaze slip. Once the correct particle sizes had been confirmed (see §5.2.1 and §5.2.5), a sample of 125g was taken for each of the 6 process variable settings. The recommended flocculant was a 10% by mass calcium chloride solution. The recommended weight to use was 0.05% of the total weight of solids (see §5.2.6). Because each sample weighed 125g and approximately 66% was solids (see Table 6.5) approximately 0.04g ±0.005g of the 10% flocculant would have

been needed for each sample. This would have been impossible to measure accurately using the balance available (Sartorius; maximum 2kg tolerance 0.01g; calibrated weekly) because it is only accurate to two decimal places. Consequently 0.5% of a 1% weight solution was added instead. The resulting increase in the proportion of water was considered negligible by employees of JM. The weight of flocculant added for each sample is given in Table 6.5 to 2 decimal places. The procedure for adding the correct weight of flocculant was the same as that described in chapter 5 (see §5.2.5).

A completely new procedure was used for shear mixing the samples. Sedimentation is an inevitable process when glazes are stored for any period of time (see §5.2.5). This includes the short periods of time while the particle size is being measured or the viscosity of other samples is being measured. To redistribute the sedimentation which occurred in the mill jars while the particle size was being measured (see §5.2.3), the mills were rotated for a further 2 minutes before the 125g samples were taken. The additional grinding of the particles was likely to be negligible given the width of the particle size range (see §5.2.1). Once the samples had been transferred to the one litre sample jars (Nalgene 660 110F) an attempt was made to remove the sedimentation by shear mixing. The same shear mixer that was used in the trials of chapter 5 (see §5.2.6), was also used here. This mixer failed to work because the samples only filled the jars to a depth of 2cm, and its blade could not be fully immersed. This was solved in the following *adhoc* but effective manner. Four pieces of small milling media (see §6.2.6) were placed in each sample jar and the jars were rotated on the milling rollers for 2 minutes. This process was sufficient to dissipate sediment and evenly redistributed the particles. Once again the additional grinding was likely to be negligible given the width of the particle size range (see §5.2.1).

6.2.8 Stage 6 Measuring Rheological Properties

Since stage 1 of the preliminary trials (see §5.2) JM had purchased a new Bohlin viscometer. All viscosity measurements were taken using a Bohlin Viscometer. Visually the Bohlin was similar to the Haake Viscometer described in chapter 5 (see §5.2.7). The only difference between the two viscometers is that the Bohlin did not have a control unit because it was completely controlled by the computer link. The Bohlin consisted of a temperature control unit housing a removable sample cup, a sensor suspended from a frame and a computer link. The temperature control unit was programmed to keep the sample at 25°C

during measurement. This is the same temperature which was used for the trials in chapter 5 (see §5.2.7).

To minimise the experimental error due to sedimentation (see §5.2.5) a two stage procedure was performed prior to each viscosity measurement. The first stage of this procedure was the shear mixing described above (see §6.2.7). The sample jars were then shaken and banged on the table to remove any residue which may have formed on the sides of the jar during shear mixing. The measurement procedure was physically the same as that described for the Haake (see §5.2.7). Recall that this involved rotating a sensor inside a measurement cup filled with sample. Because the measurement can be affected by the volume of sample used, 11.5 ml was placed in the Bohlin measurement cup using a syringe (Plastipak disposable syringe 20ml Eccentric luer slip SZR-150-051E).

Before the measurement phase was begun a constant shear stress, known as a pre-shear, was imparted on the sample (Bohlin, 1998). This initial shear stress is always followed by a period of recovery so that the sample could regain its original structure. The pre-shear was performed so that all samples started the measurement procedure with the same shear history. This was done to validate comparisons between trials because experimental error can occur if glazes experience different levels of shear stress. These shear stresses can occur during storage, transport or shear mixing. The previous approach to standardising shear history was the prevention of unnecessary external forces. This had been abandoned in favour of the pre-shearing approach which was found to be more effective (Bohlin, 1998).

In the trials performed at Sheffield the viscosity was measured while the shear rate was increased, then while the shear rate was kept constant and finally while it was reduced to zero (see §5.2.7). The only purpose of the latter two stages is to investigate whether a liquid is thixotropic (see §2.5.7). If a suspension is thixotropic high shear stresses will cause permanent reduction in its viscosity. Because it is already accepted that this is the case for glazes these parts of the measurement procedure were not performed in the trials discussed in this chapter.

The viscosity measurement was performed using the Bohlin settings given in Table 6.3. The term *stress ramp* means that specified shear stresses of gradually increasing magnitude were induced and the shear rate needed to produce them was measured. The apparent viscosity was then calculated as the ratio of shear stress over shear rate (see §5.2.8).

These settings were easy to modify using the computer link. The measurement was an entirely automated process which was initiated using the computer link.

Table 6.3 Bohlin viscometer measurement parameters

Procedure	Shear stress/Pa	Time/s	Number of Readings
Pre-shear	50	60	--
Recovery	--	60	--
Stress ramp	logarithmic increase from 0.4 to 100	120	50

When the measurement procedure had been performed the cup and bob were thoroughly washed with water and dried using highly absorbent cloth. Employees of JM believed that this process was sufficiently free of experimental error and the complex acetone procedure described in chapter 5 was abandoned (see §5.2.7).

To allow direct comparison between trials it is desirable to use the same measurement parameters throughout the experiment (Herritage, 1994, §3D). When some compositions were analysed, a yield point was not detected using the viscometer parameters given in Table 6.3. This problem has occurred in previous JM trials. In order to identify a yield point in these cases the measurement parameters have been changed. For those trials which did not exhibit a yield point a new sample was placed in the viscometer cup. The measurement was then repeated with a pre-shear stress of 100 Pa and the stress ramp commencing at 0.1 Pa. Because compositions 2, 6 and 1 were in the second block of trials to be performed (see §6.4.1) it was too late to repeat the whole experiment using the new measurement conditions. The effect of the alternative measurement conditions are included in the model using an additional factor variable (see §7.4).

6.3 Mixture Constraints and the Design Generated

In this section we consider the designs generated for the main trials. As with the preliminary trials (see §5.4) all designs discussed in this section are generated with respect to the quadratic model.

The main trials were designed in two stages. The first stage included 10 design points. After performing the trials in the initial design it was decided that a second stage of trials would be performed. A new design, which added three new points to the original design, was then generated.

As with the designs generated for the preliminary trials, this section is only concerned with mixture designs. When the trials were performed, these mixture designs were crossed with the 6 processing conditions to form product designs (see §5.4.1).

A The constrained region

It was suggested in chapter 2 that scientists who have no knowledge of experimental design often choose unnecessarily wide constraints for completeness. This, coupled with the fact that design algorithms favour extreme points of the feasible region, can lead to the analysis of non-standard trials. When this feature of optimal designs was discussed with experts at JM it was decided that the initial constraints given in Table 5.4 should be narrowed. The constraints for frit and water were narrowed by one percent either side and the constraint for clay was narrowed by half of one percent either side. The resulting constraints are given in Table 6.4.

Table 6.4 Revised percentage bounds on the glaze ingredients

Name	Lower %	Upper %	Component
1 Frit (261887)	52	65.5	x_1
2. Water (1000)	31	39	x_2
3. Clay (2622)	3.5	10	x_3

B The design generated

Using the R7 algorithm, with ks = ls = 3, kr = lr = 2 and t = 100 (see §4.3) the design specified by the circles in Figure 6.1 was generated. Doubletons are indicated by larger circles. The decision to perform 7 distinct trials, four singletons and three doubletons, is discussed in chapter 4 (see §4.7.2). When these trials were performed, point M failed (see §6.3.2). This meant that only six distinct points were analysed and there would be few

degrees of freedom for hypothesis tests about the model. Consequently it was decided to analyse 3 further compositions. The R7 algorithm, with the same parameters as above, was used to add three more points to the set of 9 trials which were analysed successfully. These new points are depicted by the squares in Figure 6.1. Once again, note that Figure 6.1 depicts an L transformed feasible region. The real and L transformed co-ordinates are given in Appendix 2.2.

Figure 6.1 Design generated for the main trials
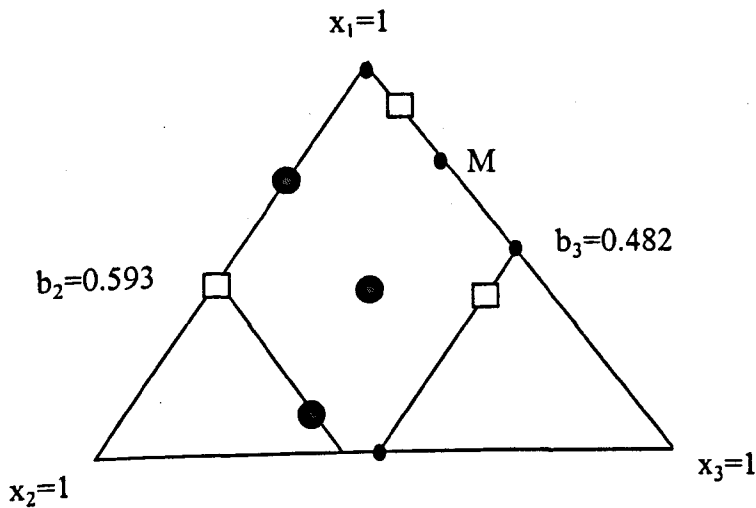
Using the algorithm to estimate the weights of continuous optimal designs (see §3.5.4) it was determined that the continuous optimal design for this region has 10 support points: all five vertices; the four mid points of the long edges and the overall centroid. The initial design in Figure 6.1, i.e. the design depicted by circles, is now compared to the general trends observed for continuous V optimal designs in chapter 3 (see §3.9). Once again, in general the design concurs with the general trends. However, similar discrepancies occurred here, because, while there are 10 support points in the continuous optimal design, there are only 7 distinct points in the exact design. The design region here is a pentagon. The initial design region for stage 2 of the preliminary trials (see Figure 5.5) is a hexagon. However, because the constraints are similar and the $x_1 = b_1$ edge in Figure 5.5 is so short, the designs are very similar.

The three additional points were one vertex, one point on the fourth long edge and an edge third on the edge which contained the failed trial. It must be noted that the resulting set of 9 singletons and 3 doubletons is unlikely to be the optimal replicate design over the candidate points considered. This is because it was generated in two stages, with the second stage constrained by the first. Again it is not surprising that the new points comprised one vertex and one point on each of the empty long edges. Note that two edges were empty because no results were obtained for point M. The fifth and final vertex was probably omitted because there are two points near to it already, and edge points tend to have higher weights. The only surprising aspect of the choice of points was that edge thirds were chosen over mid-edges.

6.4 Performing the Experiment

6.4.1 Allocation of Runs to Blocks and Mills

To maximise the amount of data which could be collected it was decided that more than one batch of sample would be milled at any one time. Because the time available for performing these trials was limited the samples were not re-tested after periods of storage .

The 10 runs of Table 6.5 were performed in two blocks of five, each block containing 5 batches of mixture (see §5.4.1). The term block is used to specify that sets of five samples are milled and their viscosities measured on the same day. The runs were allocated to blocks using the following.

Firstly each composition was given a number. The singletons were numbered 1 to 4 arbitrarily. The doubletons, or replicate pairs, were numbered 5 to 7 arbitrarily. To distinguish between each trial in a replicate pair, each of them was specified using two digits. For example, the two runs of composition 5 were referred to as runs 51 and 52.

Allocation to blocks

The singletons and doubletons were allocated to blocks separately. To gain additional information on whether the blocks affected internal error (see §4.1.2) allocation of the replicate pairs was not done completely at random. One replicate pair was performed entirely within the first block, another pair was performed entirely within the second block and the final pair was split between the two blocks. To decide which pairs would be assigned to each of these options the following randomisation scheme was used.

1, 2, 3	Replicate pair 51 and 52	7, 8, 9	Replicate pair 71 and 72
4, 5, 6	Replicate pair 61 and 62	0	Ignore

Only two random numbers were needed. The third replicate pair, i.e. the one to be split between the blocks, was chosen by default. The random numbers generated using matlab were 0, 3, 1, 0, 6. Any numbers which specified a replicate pair which had already been allocated were ignored. The relevant numbers were 3 and 6 so pair 5 was allocated to block 1 and pair 6 was allocated to block 2. For pair 7 it was decided arbitrarily that 71 would be analysed in block 1.

Once the replicate pairs, or doubletons, had been allocated to blocks, the singletons were allocated to blocks at random, in an attempt to prevent day bias. The following randomisation scheme was used:

1, 2	Composition 1	5, 6	Composition 3	9, 0	Ignore
3, 4	Composition 2	7, 8	Composition 4		

Only two numbers were required as the composition to be allocated to the second block were to be chosen by default. Numbers 9, 2, 0, 1 and 7 were generated using matlab. When relevant numbers had been ignored this yielded numbers 2 and 7. Compositions 1 and 4 were therefore allocated to blocks 1 and 2.

Allocation to mills

In an attempt to remove any mill bias the compositions were allocated to mills randomly. The compositions in block 1 were allocated to mills using the following randomisation scheme.

1, 2	Composition 51	5, 6	Composition 71	9, 0	Composition 4
3, 4	Composition 52	7, 8	Composition 1		

The mills were arbitrarily numbered 1 to 5 and the following random numbers were generated using matlab: 9, 2, 0, 4, 1, 9, 6, 5, 0, 3, 8. As each random number was generated, the appropriate composition was allocated to mills 1 to 5 respectively. That is mill 1 was allocated the first composition chosen, mill 2 was allocated the second composition chosen, etc. Mills 1 to 5 were therefore allocated compositions 4, 51, 52, 71 and 1 respectively.

Mill 5, containing composition 1, fell off the rollers and spilt its contents during the processing of batch 1. Composition 1 was consequently added to the second block and was again processed in mill 5. This led to two blocks of sizes four and six. An additional mill (labelled 6) was obtained. To allocate the remaining compositions in the second block to mills 1 to 4 and 6, the following randomisation scheme was used.

1, 2	Composition 2	5, 6	Composition 61	9, 0	Composition 72
3, 4	Composition 3	7, 8	Composition 62		

The following random numbers were generated using matlab 8, 2, 0, 8, 1, 6, 1, 9, 5, 3. Using the same approach as described for block 1, mills 1 to 4 and 6 were allocated compositions 62, 2, 72, 61 and 3 respectively.

Together with the glaze formulations, the blocks and mills to which the glazes were assigned are given in Table 6.5.

Table 6.5 Mixture formulations and block / mill allocations of the 10 trials

Composition Number	Frit %	Water %	Clay %	Frit : Clay %	Mass of flocculant/g	Block	Mill
1	52	38	10	84	0.38	2 (orig 1)	5
2	65.5	31	3.5	95	0.43	2	2
3	62.25	31	6.75	90	Trial Failed	2	6
4	59	31	10	86	0.43	1	1
51,52	61.5	35	3.5	95	0.41	1, 1	2, 3
61, 62	55.67	39	5.33	91	0.38	2, 2	4, 1
71, 72	57.2	35.6	7.2	89	0.43	1, 2	4, 3

This is a very small experiment. During the milling process mill 6 containing composition 3 fell off the rollers and the glaze formulation it contained could not be analysed. This left only 6 distinct compositions and would provide no degrees of freedom for evaluating a quadratic mixture model. Consequently some additional compositions were analysed in a further block of experiments. As another experiment was being conducted at the JM plant, only 3 compositions could be analysed in the third block. As these additional trials will require an additional block parameter they will only add a maximum of 2 to the degrees of freedom at the mixture level.

6.4.2 Particle Sizes Reached

The process of attaining the specified particle sizes (see §5.2.1) involved measuring the particle sizes at a number of time points (see §5.2.1) and then predicting how much additional time was needed. Tables 6.6 and 6.7 give the percentage of particles which were smaller than 14μm at each time point, for blocks 1 and 2 respectively. The tables are included to illustrate the milling rates and to allow particle size to be treated as a continuous variable rather than a factor variable. The highlighted values in the tables indicate particle sizes within the relevant ranges (see §5.2.1). Every time a sample was removed the time counter was reset to 0. This is denoted in the tables by a double line separating the relevant time points and by a plus sign preceding all times after the counter had been reset. The mill

containing run 3 fell off the rollers and spilt its contents. It was excluded from the experiment and is omitted from Table 6.7. With run 4, the samples for processing conditions E and F were extremely thick. Consequently they were lost as residues when being transferred between containers, and no measurements could be made.

Table 6.6 Milling times (hours) and particle sizes reached for block 1

Run / time	12	15:15	18:45	21:15	26:15	27.15	+1:45	+1:10	+2:00	+3:00
4	46.1	50.1	55.6		67.7		71.2	72.0		76.56
51	50.0	57.2	61.9	65.5			70.2	75.7		
52	43.1	46.5	53.2		64.8	66.1	71.1	73.6	76.16	
71	45.0	51.3	56.9		69.8		74.5	77.7		

Table 6.7 Milling times and particle sizes for block 2

Run / time	17:30	18:30	23:00	25:30	+2:00	+1:45	+2:15	+1:15	+1:30
1	64.4		66.0		72.3			77.2	
2	56.8	56.4	62.1	65.4		69.1	70.9		75.6
61	60.3	62.3	62.5	67.7		70.23			77.4
62	63.53		65.1		70.7			76.5	
72	55.7	57.6	65.7		69.6		71.5		77.0

With compositions 1, 2, 61 and 62 a yield point was not detected initially and the alternative measurement parameters had to be used (see §6.2.8). With composition 1, this only occurred for processing conditions E and F.

6.4.3 A Third Block of Experiments

Because the data collected in the trials discussed above only contained a small number of distinct trials it would have been difficult to asses fitted models. A third block of experiments with three new compositions was therefore performed. A 12 point design, conditioned to include the nine completed runs (see §6.3.1), was generated using the R7

algorithm (see §4.3). To use the R7 algorithm it was assumed that all data from run 4 had been obtained. This was because major modifications to the algorithm would have been required if the design were to consider the process variables as well as the mixture variables.

Two of the compositions selected were very similar to run 4 which could only be measured at 4 levels of the process variables (see §6.4.2). To avoid losing data from two of the process variable levels for the new trials the percentages were modified slightly. The original formulations are given in brackets in Table 6.8. The amount of clay in run 9 was decreased and the amount of water in run 10 was increased. These changes were performed by rounding the relevant percentages to the nearest integer in the appropriate direction. The resulting proportions are given in Table 6.8. It is conjectured that the resulting loss in efficiency would have little practical effect on the experiment (see §3.1.2).

The mills and media used in blocks 1 and 2 had been redistributed and the same mills that were used in batch 1 and 2 could not be used (see §6.2.6). The mills in this third batch are numbered 7 to 9. The same approach of allocating specified masses from three different sizes of media was again used (see §6.2.6). All other aspects of the experiment remained the same. Table 6.9 gives the milling times and particle sizes reached, specified in the same way as for the first two batches in Table 6.6 and Table 6.7 (see §6.4.2).

Table 6.8 Additional runs for block 3 The values in brackets indicate the percentages specified by the design. The values in not in brackets are modified values which were used instead.

Run/Design Point	Frit %	Water %	Clay %	Frit : Clay %	Mass of flocculant (g)	Block
8	57.5	39	3.5	94.3	0.38	3
9	62(61.2)	31(31)	7(7.83)	89.9	0.39	3
10	56(56.7)	34(33.3)	10(10)	84.8	0.41	3

Table 6.9 Milling times and particle sizes for batch 3

Run / time	22:30	23:30	+1:00	+2:00	+2:00
8	74.52		76.29		
9	64.73	66.17		71.1	77.33
10	66.07			71.02	77.1

From Table 6.9 it can be seen that run 8 had been over-milled by the first reading. Consequently only process variable levels C to F were analysed.

6.5 Summary of Trials Performed and Suggestions for Analysis

The three batches of experiments performed in the main trials yielded a total of nine distinct compositions and 50 distinct combinations of the compositional and process variables. Because 18 of these 50 combinations had been replicated the rheology of 68 trials had been investigated. In chapter 7 the rheology data is analysed. Firstly the three aspects of viscosity which are of interest were summarised for each of the 68 trials. After simple exploratory analyses of these values a variety of models are fitted for each of the three summary values. All trials which were performed in the same block are likely to experience some of the same error factors. In an attempt to quantify this dependency between the errors some of the models fitted in chapter 7 will include a covariance structure (see §7.6).

Chapter 7

Analysis of Experimental Results

7.1 Introduction

7.1.1 Outline

The main aim of this chapter is to identify models for predicting some rheological properties of glaze slip, using data from the 68 experimental trials discussed in chapter 6. Before any analysis can be performed the first stage is to extract from the data some summary measures of the rheological properties of interest. The estimation of three response variables of interest is discussed in section 7.2. Before the modelling is performed, section 7.3 gives a simple exploratory analysis of the data. In sections 7.4 to 7.7 a variety of models for predicting the three summary variables are fitted. Two different approaches to modelling are used. The first assumes independence between the observations, and uses standard General Linear Modelling (GLM) techniques (see §7.5). The second approach takes into account the split plot structure (see §7.6), and uses both maximum likelihood (ML) and restricted maximum likelihood (REML) techniques (see §7.6). Once the models have been selected, the predictive power of a range of models is assessed using a form of the PRESS statistic (see §2.2.6) which has been modified to account for replication. This PRESS type analysis is performed in section 7.8. The fitted models are then interpreted in section 7.9.

7.1.2 Variables Used

The response variables to be modelled describe the relationship between viscosity and shear stress (see §2.5.9). The three summary variables of interest to JM are the viscosity of the top and bottom plateaux and the shear stress at which the top plateau ends (see §2.5.9).

The explanatory variables for the rheology modelling will be the proportions of each ingredient used and some qualitative information on the processing conditions. Experiments in which the variables are proportions are known as mixture experiments (see §1.1). The mixture variables used here are frit, water and clay. The process variables are: whether flocculant has been added; and the particle size range to which the solid particles in the glaze have been milled.

7.2 Summarising the Viscosity Graphs

7.2.1 Viscosity Graphs for Chapter 6 Data and Theoretical Shape

A Theoretical viscosity curves

The inverted S shaped curve which JM use to describe glaze rheology is given in Figure 7.1. Recall that the aspects of the curve which are of interest to JM are the viscosity of the top and bottom plateaux and the shear stress at which the top plateau ends (the *yield point*). As the top plateau is usually several orders of magnitude greater than the bottom plateau, log-log (base e) graphs are usually used to plot viscosity against shear stress data.

Figure 7.1 Typical viscosity curve displaying an inverted S shape.
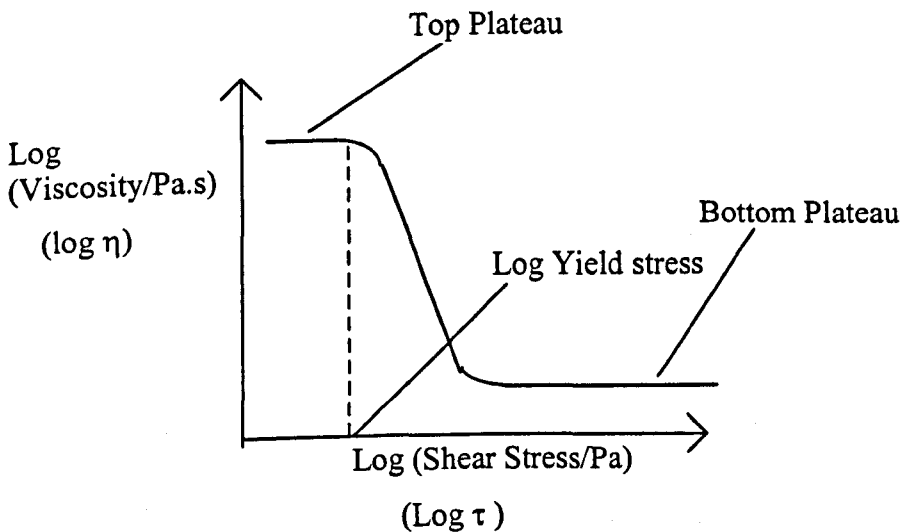
When graphs of this type were plotted for the experiments discussed in chapter 6, most of them could be well approximated by the graph shown in Figure 7.1. This shape suggests that the glazes studied approximate to Bingham liquids (as discussed in chapter 2). The rheology of a Bingham liquid can be summarised by two parameters; the lowest shear stress which when applied to the liquid will cause flow, and the subsequent absolute viscosity of the liquid. The former is known as the *yield stress*.

In chapter 2 the S-shaped curve was compared to the Bingham model. It was suggested that the yield point defined by JM is likely to be the same measure as the yield stress, and that the bottom plateau is likely to be a good estimate of the absolute viscosity. However, it was suggested that the top plateau is probably caused by wall slip (see §2.5.9). Although the level of wall slip

will be related to the rheology of the glaze in some way, it is a consequence of the viscosity measurement procedure and is unlikely to have any scientific significance (Messer, pers.comm).

Because it was requested by JM, we investigate all three summary measures here. Based on the Bingham model, the analysis of the yield point and bottom plateau will have the most scientific significance.

B Overview of the graphs plotted for the trials discussed in chapter 6

Graphs of log viscosity against log shear stress were constructed for each of the 68 trials. For each sample of glaze slip, the Bohlin viscometer measures viscosity at 60 different settings of shear stress (see §6.2.8). Each of the 68 graphs was constructed using all 60 of these viscosity measurements. It is often necessary to use several different sensor systems to investigate rheological properties over a large range of shear stresses (Herritage, 1994). The fact that both plateaux were exhibited in all 68 graphs, shows that the experiment was relatively successful. Apart from some minor deviations, the majority of these graphs have the inverted S shape of Figure 7.1. There were six ways in which the graphs were seen to display slight deviations from the S shape. The four plots in Figure 7.2, exhibit between them the six ways in which the observed curves deviated from the theoretical S shape. Possible explanations for these deviations are given below, where the deviations are described. The six types of deviation are as follows:

Top Plateau

1) **Random fluctuation**. This is shown by the seemingly random sawtooth type variation on the top plateaux in Figures 7.2 (a) and 7.2 (b). The variation is not a major problem because the general trends could easily be estimated by straight lines.

2) **Gradual increase in viscosity**. A gradual increase in viscosity occurs at the start of each of the top plateaux in Figure 7.2. This increase was observed in the majority of the 68 plots and was expected due to a consequence of the measurement procedure known as *thixotropic build-up* (Herritage, 1994, §3D). This effect is a separate phenomena from the wall slip discussed above. Figures 7.2 (a), (b) and (d) exhibit linear increases. The increase in Figure 7.2 (c) appears to be curved.

3) **Eventual plateau**. For all 68 graphs the gradual increase was followed by a flat region prior to the yield point.

Bottom Plateau

4) Approximate plateau. In many cases an actual minimum was never reached. In Figure 7.2(b) it is clear that no minimum has been reached. In Figure 7.2(a) it appears that the viscosity is still decreasing at the end of the curve. It is hypothesised that in cases like this the shear stress would need to be increased further in order to identify the bottom plateau.

Figure 7.2 Four examples of the graphs of log viscosity against shear stress

5) Gradual increase. In the majority of cases where a minimum was observed, it was followed by a gradual increase in viscosity. An example of this increase can be seen in Figure

7.2(c). This increase in viscosity is probably due to dilatant behaviour (see §2.5.9) and is a consequence of the measurement procedure.

6) **Intermediate plateau** In some cases the drop in viscosity slowed down, an intermediate plateau was then observed for a while until the viscosity decrease suddenly increased in rate. In the cases where this intermediate plateau was observed a bottom plateau was never reached. The failure to reach a bottom plateau is discussed in point 4 above. An example of an intermediate plateau can be seen in Figure 7.2(b).

Finally, run 51B showed extreme deviation from the theoretical S-shape. The graph for run 51B is given in Figure 7.3. Due to this extreme deviation, location of a top plateau and a yield point would be extremely subjective. Consequently run 51B is not used in the analysis of these two responses but is still used in the analysis of the bottom plateau.

Figure 7.3. Run 51B
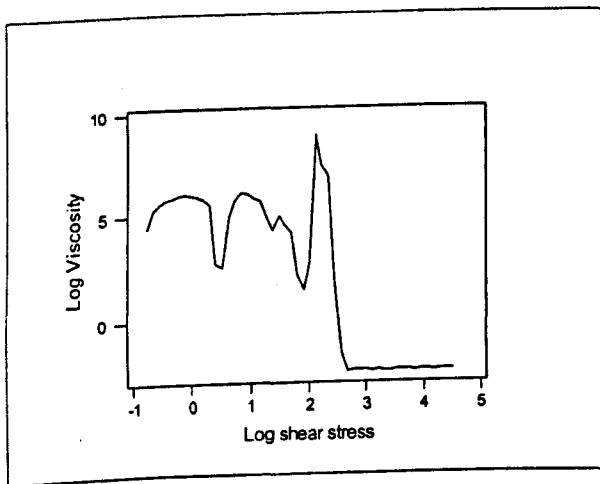
7.2.2 Potential Methods for Estimating the Responses of Interest

Recall that the aspects of the rheology graphs which are of interest to JM, are the viscosity of the top and bottom plateaux and the yield point The following four options were considered for estimating these three summary values.

Method A) Estimate the top and bottom plateaux by the maximum and minimum viscosity respectively. The yield point would be estimated by the shear stress which produced the maximum viscosity.

Method B) Estimate the top and bottom plateaux by the average viscosity of the points before and after the steep descent respectively. The yield point would then be estimated by the shear stress of the last point which is considered to be before the steep descent.

Method C) Estimate the summary variables using the same procedure as method B after removing clearly atypical points from the data sets. The first points to be removed would be those which deviate from the theoretical S shape due to thixotropic build-up and dilatancy. Any remaining points which were atypical of the region they were in, and likely to be due to experimental error, would also be removed (see §7.2.5).

Method D) Estimate the summary variables from parametric curves fitted to each of the 68 data sets.

All four methods, or combinations of them, were considered viable by JM scientists. The feasibility of each approach is considered below.

7.2.3 Method A Using Maximum and Minimum values

Method A involves estimating the top and bottom plateaux by the observed maximum and minimum viscosity measurements. The yield point would be estimated by the shear stress which produced the maximum viscosity.

Maximum value in the top plateau region

Figure 7.4 shows the viscosity data for two trials exhibiting peaks which were considerably higher than the majority of points on the top plateau. The highest of these peaks would clearly yield estimates which are considerably higher and lower than the top plateau and the yield point respectively.

Errors of this type could be reduced if non-typical peaks and troughs were removed. The removal of non-typical points is a subjective procedure and could lead to the loss of information on important effects. An example of how subjective the removal of points could be is given in Figure 7.4(b). In Figure 7.4(b) there are no clear plateaux and the selection of points to delete would be difficult. A more conservative approach to the removal of points from the graphs in Figure 7.4, would be to retain all points except the clear peaks near the beginning, i.e. all points labelled 'O'. However, consider graph 7.4(b). If only the clear peak 'O' was removed, the maximum would be either of the points labelled M and N. If point M were the maximum, the yield point estimate would be considerably lower than it should be. However, due to the

inaccuracies in estimating the top plateau described above, method A is rejected as a method for estimating (at least) the top plateau.

Figure 7.4 Examples of data sets with non typical peaks and troughs
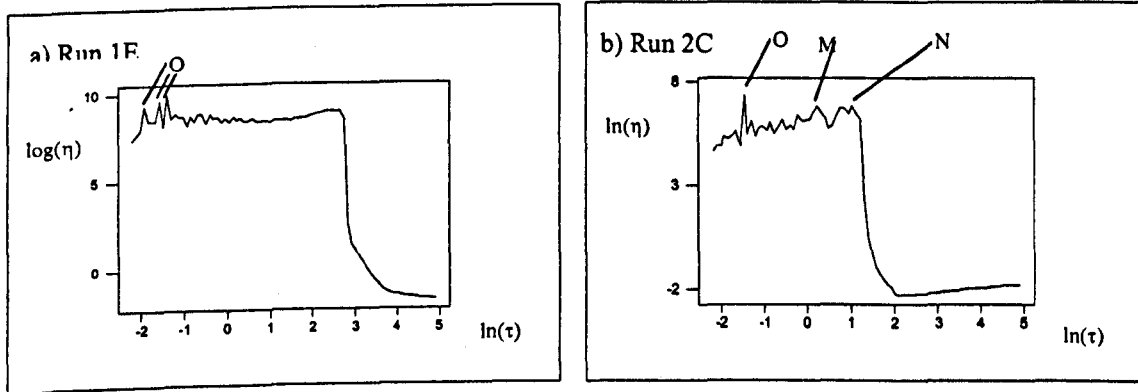
Minimum value in the bottom plateau region

The graphs in Figure 7.2 exhibited three different effects in the bottom plateau region. However, in terms of calculating a minimum value, there were essentially only 2 effects observed (the frequencies with which each effect was observed is highlighted next to its description):

(a) minimum point followed by gradual increase. (§7.2.1B5) **44 occurrences**

(b) no minimum reached (§7.2.1B4,6) **24 occurrences**

In case (a) the increase was probably caused by dilatancy and is therefore of no interest. In case (b), where a minimum viscosity was not reached, the best estimate of the minimum which would eventually be reached is probably the minimum observed so far. Therefore, in both cases the minimum viscosity is likely to provide a good estimate of the true bottom plateau and hence the absolute viscosity. As no spurious values were observed in the region of the bottom plateau, the minimum viscosity can be calculated without removing any values from the data sets.

7.2.4 Method B Using Mean Values of the Points on the two Plateaux and Taking the Yield Point from the Raw Data

Method B involves splitting the data into groups of points which occur before and after the rapid descent in the middle. If it was not clear whether the points at the beginning and end of the steep descent were on the plateaux they were excluded from the groups. The mean of each set of

points is used to estimate the top and bottom plateaux respectively. The method B estimates for run 2A are illustrated in Figure 7.5.

Figure 7.5 Plateaux estimates using average values for run 2A
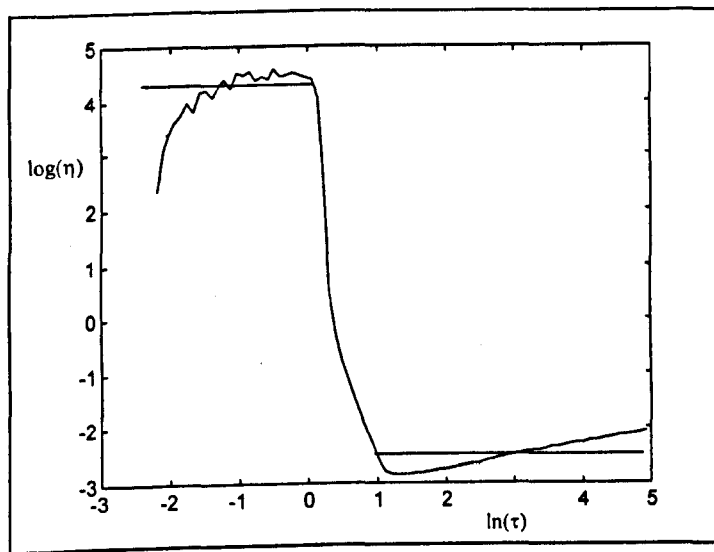
The estimates illustrated in Figure 7.5 appear to be below and above the top and bottom plateaux respectively. Consequently it was decided that this would be a poor method of estimation. These discrepancies are caused by thixotropic build-up and dilatancy (see §7.2.1B). Both of these effects are of no interest to the investigation and ways to remove the appropriate data points are discussed in the next section.

7.2.5 Method C Removing Atypical Points and Calculating the Mean of Those Which Remain

The approximately flat region prior to the steep descent (the top plateau) is a period of high limit Newtonian behaviour (Herritage, 1994, §3D). The gradual increase in viscosity, which occurs before this flat region, is due to the sample recovering from the pre-shear and is known as thixotropic build-up (Herritage, 1994, §3D) (see §6.2.8). The thixotropic build-up is a consequence of the measurement procedure and is an effect which is not relevant to the investigation. From Figure 7.5 it can be seen that the data in the region of the thixotropic build-up causes the method B estimate of the top plateau to be below the region of high limit Newtonian behaviour. Though run 2A was an extreme case, thixotropic build-up was observed, to varying extents, in all trials.

Although the Bohlin manual (Herritage, 1994) highlighted the problem of thixotropic build-up, no protocol was given for removing the appropriate points. The removal of the appropriate points was performed in the following *adhoc* manner, which was approved by scientists at JM. Starting from a visual estimate of the yield point, and working backwards along the top plateau, the first point at which a clear visual decrease was observed was identified. The mean and standard deviation of the set of points which occurred between a visual estimate of the yield point and this clear decrease were then calculated. Those points near to the visual estimate of the yield point, whose status as a top plateau point was unclear, were excluded from this calculation. A 95% confidence interval for the mean of this subset was calculated. The confidence interval was based on the t_{n-1} distribution, where n is the number of points in the subset. If any points at either end of the subset were within the 95% confidence interval they were added to the subset. Using the new set of points an updated confidence interval was calculated and the process repeated until the points considered to be on the plateau did not change between iterations. Once the data which corresponded to thixotropic build-up had been removed, the top plateau was estimated by the mean of the remaining points before the steep descent.

The yield can then be estimated by the shear stress of the last point considered to be on the top plateau. In the majority of cases the last point was easily selected, as the viscosity was seen to suddenly drop by a factor of approximately 10. In other cases this large drop was preceded by a smaller one and identification of the yield stress was less easy. If the point which represented the initial drop was considered to be part of the top plateau the yield point was taken as the shear stress of the next point.

In 44 cases a bottom plateau was reached. In these cases taking the mean viscosity, once atypical points had been removed, would probably provide a good estimate of the bottom plateau. However, in 24 cases a bottom plateau was not reached. Therefore, in these cases, none of the points are typical of a bottom plateau. Consequently the calculation of an average would be a poor estimate of the bottom plateau which would eventually be reached.

7.2.6 Method D Fitting a Parametric Curve to the Data

The final method to be considered involved fitting parametric curves to the data. As the other methods were considered to be more appropriate by JM, the results from this estimation method were never used. An outline of the method is included for possible consideration in future analysis.

Curves which approximate the inverted S shape of the theoretical viscosity curve (see §2.5.9) are given as equations 7.1 and 7.2. Equation 7.2 is known in the literature as the *Cross model* or the *extended Williamson model* (Caufin *et al*, 1992). Equation 7.1 is subsequently referred to as the *arctan model*. In equation 7.2, η_0 and η_∞ represent the viscosity of the top and bottom plateaux respectively. In both models, η and τ represent viscosity and shear stress, respectively.

$$\eta = a + b.\mathrm{Tan}^{-1}(c\tau + d) \tag{7.1}$$

$$\eta = \eta_\infty + (\eta_0 - \eta_\infty) / (1 + \alpha\tau^n) \qquad (n, \alpha > 0) \tag{7.2}$$

Increasing the value of n increases the rate of descent between the two plateaux but decreases the yield point. Decreasing α increases the yield point but makes little change to the rate of descent. Because the data plots all exhibit similar rates of descent, but different yield points, it was decided that changing the value of α would produce sensible example plots and the value of n would be a good indication of relative magnitudes of each yield point.

Examples of each model are given in Figure 7.6. Figure 7.6(a) shows the Cross model for three different values of α, with the other parameters remaining constant. The full set of parameter values are as follows: $\eta_\infty = 2$, $\eta_0 = 98$, n=10, $\alpha = 10^{-10}$ (single line); 10^{-5} (dashed); and 1 (dotted). Figure 7.6(b) shows an example of the arctan model with the following parameters: a = 75; b = −40; c = $^\pi/_{18}$; d = $−^\pi/_2$. As is the case in the data sets, the shear stress varies from 0.5Pa to 90Pa in both graphs.

Figure 7.6 Two possible models for approximating the data

(a)

(b)

227
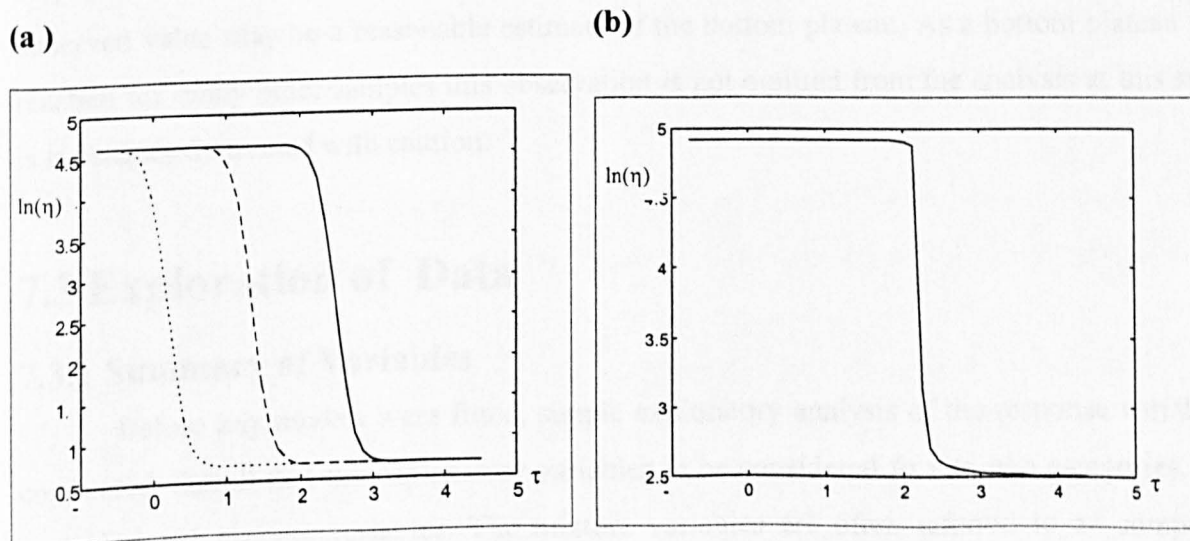
Both models can be fitted using an iterative least squares method. However, a weighted least squares, with the weights given by the observed viscosity, would probably be more appropriate because the observed viscosity varies so much between the top and bottom plateaux regions.

7.2.7 Conclusion on the Choice of Summary Variables

Because methods A and B gave such poor estimates of the top plateau and the yield point, method C will be used to calculate these two summary measures. The analysis of the top plateau will be based on the mean viscosity of all typical points in the region of high limit Newtonian behaviour (see §2.5.8). The analysis of the yield point will be based on the shear stress of the last point on the top plateau (see §7.2.5).

Recall that 2 different patterns were observed in the bottom plateau region (see §7.2.1). The only approach which provides realistic estimates of the bottom plateau for both patterns is method A. The bottom plateau will therefore be estimated using the minimum viscosity observed.

It is hypothesised that the yield point data will be the most reliable, as from Figure 7.2 it appears that the points at which the top plateaux end are clear, but the viscosity of points on it varies considerably. The data for the bottom plateau is likely to be less reliable than the yield point because in many cases a minimum was not reached (see §7.2.1B).

The 67 values obtained for the top plateau and yield point and the 68 values obtained for the bottom plateau are given in Appendix 3. From the data for the bottom plateau it is clear that run 4F is an extreme outlier. Run 4F was one of the samples for which a bottom plateau was not reached. Sample 4F was the flocculated fine particle sample for a glaze with low water and high clay (see Table 7.2 and Table 7.3). As such samples are expected to have high viscosity, the observed value may be a reasonable estimate of the bottom plateau. As a bottom plateau was not reached for many other samples this observation is not omitted from the analysis at this stage but is subsequently treated with caution.

7.3 Exploration of Data

7.3.1 Summary of Variables

Before any models were fitted, simple exploratory analysis of the response variables was conducted. Recall that the explanatory variables to be considered fit into two categories, mixture variables and process variables. The mixture variables are often referred to as compositional

variables and are the proportions of frit, water and clay in the glaze. Nine different glaze formulations were produced, each of these formulations is referred to as mixture. Recall that three of the mixtures were produced twice.

The processing conditions (or factor variables) are expressed using *indicator variables*. Indicator variables take the values 1 and 0, depending on whether the process in question was used or not. In our case the production processes were the addition of flocculant and the particle size to which the glaze had been ground (see §5.2.1). A single indicator variable was used to represent flocculant. Two indicator variables were used to represent particle size. An alternative to using two indicator variables would have been to use a single variable taking the values −1, 0 and 1. This single variable would force the effect of moving from medium to coarse particles to be equal, but opposite to, the effect of moving from medium to fine particles. Using the single variable would therefore reduce the generality of the model. Interaction terms between each of the variables will also be considered (see §7.3.5).

For the remainder of this chapter the variables are referred to by the following letters; top plateau (y_1), yield point (y_2), bottom plateau (y_3), proportion of frit (x_1), proportion of water (x_2), proportion of clay (x_3), flocculant indicator (z_1), medium particle indicator (z_2), fine particle indicator (z_3) The original symbols (η and τ) are not used in the analysis of this section to distinguish between measured viscosity over a range of shear stresses and the estimates of the plateaux and yield point which will be used in modelling.

7.3.2 Correlation Between Response Variables

Two variables are said to be *correlated* if inferences on the likely value of one of them can be made from the known value of the other. The relationship between two variables can be investigated by taking measurements of them both from the same item and calculating the correlation coefficient between the two resulting data sets. A correlation coefficient with a high magnitude suggests a strong relationship between the variables.

Because all three of the response variables used here (see §7.2) are essentially summaries of the same property, they are likely to be correlated with each other. The correlation coefficient is likely to be especially high between the viscosity of the top plateau and the yield point. The viscosity of the top plateau (y_1) is the viscosity under low shear stresses. If the viscosity under low shear stresses is high, it is likely to be because the bonds between molecules are very strong. To break these strong bonds and cause the glaze to have the viscosity of the bottom plateau, a high

shear stress is likely to be required. Because this high shear stress is the yield point, if y_1 is high then y_2 is also likely to be high.

To investigate whether the three response variables are correlated, the scatter plots given in Figure 7.7 were constructed. Pearson's correlation coefficients are also given in Table 7.1. Because the top plateau and yield point for run 51B have already been removed, its bottom plateau can not be used in the analysis of this section. Additionally, the scatter plots were constructed after run 4F had been omitted. Run 4F was removed for the purposes of these plots because its bottom plateau is at least 4 times bigger than all other bottom plateaux. Inclusion of this point would have caused the other points to appear more clustered, thus making any relationship appear stronger. Run 4F was not omitted when the correlation coefficients were calculated.

Figure 7.7 Plots to compare relationship between response variables

a) Top Plateau against Yield Point

b) Top Plateau against bottom plateau

c) Yield Point against Bottom Plateau
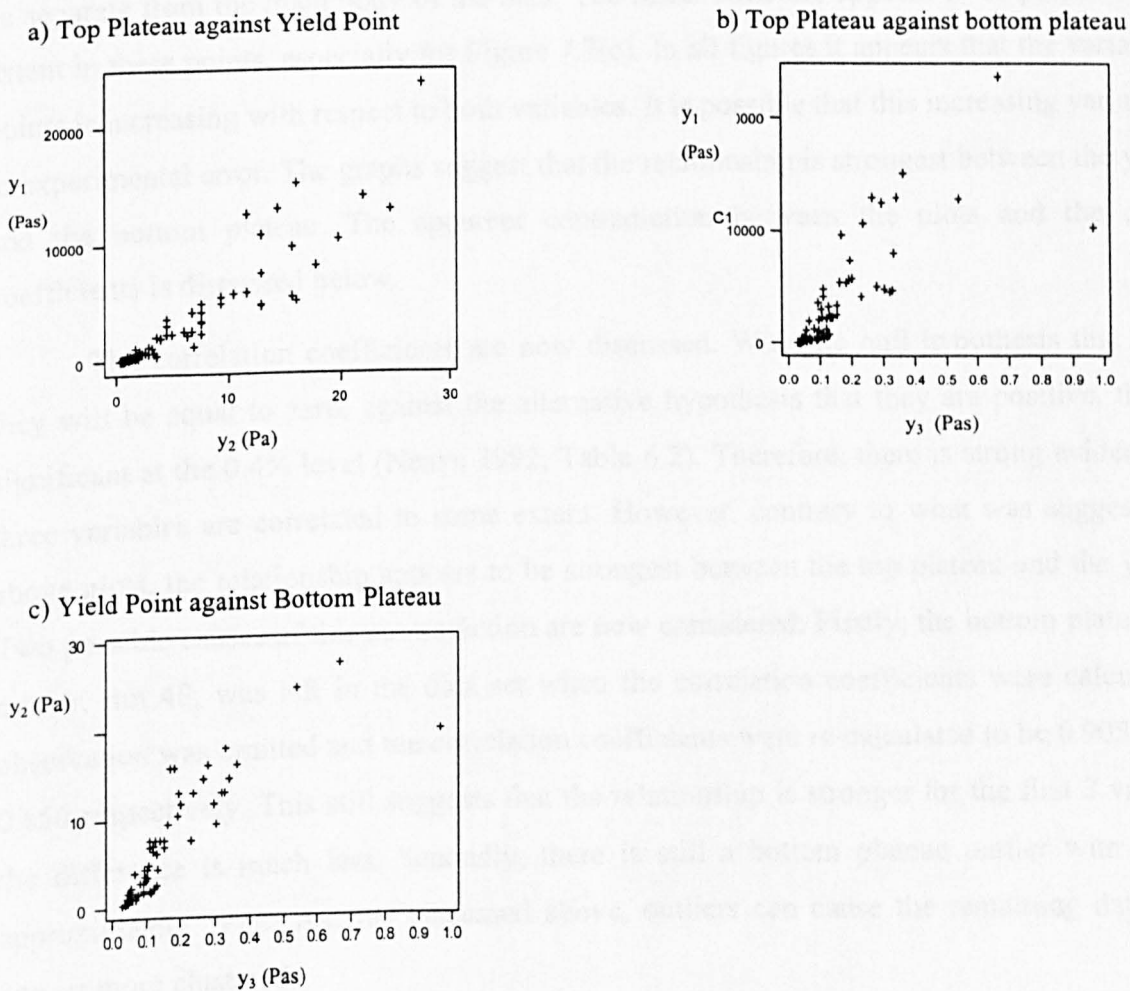
In all three plots many of the points are clustered in the same region. In all three cases a relationship between the variables could be reasonably well approximated using a straight line. Consequently, each variable is probably correlated to some extent with the other two variables.

Table 7.1 Pearson's correlation coefficients

Variables	Pearsons r
Top / Yield	0.913
Top / Bottom	0.493
Bottom / Yield	0.534

Despite the clustering of points, approximately one quarter of the points in each graph appear to be separate from the main body of the data. The linear structure appears to be preserved to some extent in these points, especially for Figure 7.7(c). In all figures it appears that the variance of the points is increasing with respect to both variables. It is possible that this increasing variance is due to experimental error. The graphs suggest that the relationship is strongest between the yield point and the bottom plateau. The apparent contradiction between the plots and the correlation coefficients is discussed below.

The correlation coefficients are now discussed. With the null hypothesis that in general they will be equal to zero, against the alternative hypothesis that they are positive, they are all significant at the 0.4% level (Neave 1992, Table 6.2). Therefore, there is strong evidence that all three variables are correlated to some extent. However, contrary to what was suggested by the above plots, the relationship appears to be strongest between the top plateau and the yield point. Two possible causes of this contradiction are now considered. Firstly, the bottom plateau extreme outlier, run 4F, was left in the data set when the correlation coefficients were calculated. This observation was omitted and the correlation coefficients were re-calculated to be 0.905, 0.789 and 0.856 respectively. This still suggests that the relationship is stronger for the first 2 variables but the difference is much less. Secondly, there is still a bottom plateau outlier with a value of approximately 1Pa.s. As was discussed above, outliers can cause the remaining data points to appear more clustered.

The fact that the r values changed so much when run 4F was omitted, suggests that the bottom plateau value for run 4F may need to be analysed separately from the other 67 trials.

7.3.3 Marginal Analyses

A Outline

This section takes each response variable separately, and subdivides the 68 values in a number of different ways. Each subdivision/grouping of the data is according to a different explanatory variable. Recall that there are five explanatory variables of explicit interest. The two processing conditions can also be combined with each other to give a sixth explanatory variable with six levels. Together with the five explanatory variables, this combination of the process variables means that there are six different ways of subdividing/grouping the data for each of the response variables. For each subdivision, the mean and standard error of the mean were calculated for each response. Groupings for the process variables were made according to their respective levels. For the mixture variables the data was put into groups depending on whether 'low' (0), 'medium' (1) or 'high' (2) proportions of the ingredient were used. The proportions which constituted each of these categories were chosen so that in each case the 12 different mixture compositions were split into three groups of four. The groupings and the mixture compositions are given in Table 7.3. The frit to clay ratio is given as it is commonly used by scientists at JM, it is calculated as $100F/(F+C)$. The mean and standard error of the mean are given in Table 7.4 for each grouping.

Table 7.2 Processing condition codes

Code	Particle size	Flocculant
A	(Coarse) 65%-70%	No
B	(Coarse) 65%-70%	Yes
C	(Medium) 70%-75%	No
D	(Medium) 70%-75%	Yes
E	(Small) 75%-80%	No
F	(Small) 75%-80%	Yes

The codes used in Table 7.4(d), to specify the combined levels of the processing conditions, are defined in Table 7.2.

B The marginal effects of the explanatory variables on the response variables

The comparisons performed in this section are only used to give an indication of what may be observed when models are fitted later, not to produce definitive conclusions. By grouping the trials in the way described above, we have probably lost a great deal of information. Any hypothesis tests based on Table 7.4 are likely to be extremely conservative compared to more advanced tests using all the information contained in the data.. Consequently attention is restricted to simple visual comparisons and very approximate z-tests with no adjustment for multiplicity (see §7.5.1B).

Table 7.3 Mixture proportions and appropriate grouping for each trial The numbers in the groupings column have the following interpretations; 'low' (0), 'medium' (1) or 'high' (2)

Run	Percentages				Groupings			
	Frit	Water	Clay	Frit:Clay	Frit	Water	Clay	Frit:Clay
1	52	38	10	84	0	2	2	0
2	65.5	31	3.5	95	2	0	0	2
4	59	31	10	86	1	0	2	0
51, 52	61.5	35	3.5	95	2,2	1,1	0,0	2,2
61, 62	55.67	39	5.33	91	0,0	2,2	1,1	1,1
71, 72)	57.2	35.6	7.2	89	1,1	1,1	1,2	0,1
8	57.5	39	3.5	94.3	1	2	0	2
9	62	31	7	89.9	2	0	1	1
10	56	34	10	84.8	0	0	2	0

A simple visual analysis is performed first. Because the data was radically simplified to produce Table 7.4, the effects identified by visual comparisons may prove to be significant when modelling is performed.

The main effects of each mixture variable are clear, and appear to be similar for all three response variables. This similarity in the effect of each mixture variable is further evidence that the three response variables are correlated (see §7.3.2). It appears that the addition of flocculant increases all three response variables. Flocculant was expected to increase viscosity because this is its purpose (see §2.5.1). Decreasing particle size can also be seen to increase both the viscosity of the top plateau and the yield point. This increase in viscosity, when particle size is decreased, was also expected (see §2.5.1). For the bottom plateau the effect of decreasing the particle size appears to be non-monotonic. The bottom plateau increased when the particles were milled from coarse to medium but then decreased again when the particles were further milled to the fine particle size. When the combined levels of the process variables are considered for the top plateau it can be seen that the addition of flocculant had a bigger effect than decreasing particle size. This can be seen by comparing level B with level C, and comparing level D with level E. For the other two response variables comparisons of this kind are less clear.

Increasing clay can be seen to be associated with an increase in viscosity in all three cases. In general an increase in water can be seen to decrease the viscosity. The increase and decrease in viscosity which was observed for increasing clay and water respectively, was to be expected from what is already known about glaze (see §5.2.1). Less is known about the effect of frit (see §5.2.1), and the effects observed in the data are unclear. For the top plateau an increase in frit appears to decrease the viscosity. For the yield point the effect of frit is small relative to the other mixture variables and is again non-monotonic. For the bottom plateau the effect of frit is again non-monotonic but is of a similar size to the effect of the other variables. In general, however, the marginal data suggests that changing frit will have a smaller effect on viscosity than changing any of the other variables. Because little is known about frit, its interaction with the processing conditions should probably be chosen as one of the interactions to be estimated explicitly (see §2.3.2).

Pairwise comparisons of appropriate group means are now performed using very approximate z-tests. If two group means differ by more than 2 times the sum of their respective standard errors it is concluded that the group means are likely to be different in general. Based on this approach only three pairwise comparisons are significant. However, 3 further pairwise

comparisons were borderline. These 6 comparisons are those between low and medium clay, and low and medium water, for all three response variables. Increasing water from a low level to a medium level appears to decrease the mean for all three responses. Increasing clay from a low level to a medium level appears to increase the mean for all three responses. The fact that these 6 pairwise comparisons were significant, despite being based on drastically simplified data, suggests that the effects of clay and water are extremely strong.

7.3.5 Interactions

A Terminology

The analysis above considers the effect of each variable in isolation. This is known as considering the *main effects*. The effect of one variable often depends on the values of other relevant variables due to an effect known as 'between variable interaction'. An analysis which allows for 'between variable interaction' can often give conclusions which are completely different from those for the main effects.

Interactions are included in a model if the effects of one variable are different at the various settings of another variable. Interaction terms are best described in terms of process variables and this is done here using the processing condition codes of Table 7.2. Recall that A, B and C are used for the three particle sizes when no flocculant has been added and D, E and F are used for the particle sizes when flocculant has been added. Using these codes, Figure 7.8(a) shows an example in which the effect of particle size is the same at both levels of flocculant. Figure 7.8(b), however, shows an example where the effect of particle size is different when flocculant has been added. In the latter case we say that there is an interaction between flocculant and particle size.

Table 7.4 Means and standard errors of groups (For groups see table 7.3)

	Top Plateau	Yield Point	Bottom Plateau
a) Overall	3694 (584)	6.77 (0.80)	0.213 (0.060)
b) By Flocculant Status			
Without Floc'	2831 (589)	6.32 (1.08)	0.163 (0.031)
With Floc'	4582 (1006)	7.24 (1.20)	0.263 (0.116)
c) By Particle Size Status			
Coarse(Cse)	2462 (652)	4.97 (1.07)	0.133 (0.020)
Medium (Md)	3651 (971)	6.17 (1.30)	0.320 (0.166)
Fine	4916 (1270)	9.14 (1.64)	0.177 (0.035)
d) By Combined Processing Conditions			
A Cse	1792 (615)	4.22 (1.31)	0.125 (0.027)
B Cse/Floc	3199 (1184)	5.79 (1.77)	0.141 (0.032)
C Md	3034 (1124)	5.82 (1.74)	0.192 (0.074)
D Md/Floc	4268 (1617)	6.53 (2.00)	0.448 (0.326)
E Fine	3650 (1211)	8.95 (2.33)	0.170 (0.045)
F Fine/Floc	6182 (2233)	9.34 (2.42)	0.184 (0.055)
e) By Mixture Proportion			
Frit			
Low	4039 (998)	6.42 (1.28)	0.129 (0.021)
Medium	3677 (796)	7.45 (1.44)	0.359 (0.198)
High	3348 (1187)	6.55 (1.51)	0.176 (0.032)
Water			
Low	7545 (1327)	12.37 (1.57)	0.466 (0.175)
Medium	2076 (396)	4.62 (0.68)	0.112 (0.007)
High	1534 (524)	3.43 (0.98)	0.072 (0.010)
Clay			
Low	753 (156)	2.80 (0.48)	0.090 (0.007)
Medium	3768 (1152)	6.49 (1.54)	0.159 (0.035)
High	6419 (955)	10.87 (1.27)	0.395 (0.177)

Figure 7.8 Simple plots showing the effect of an interaction
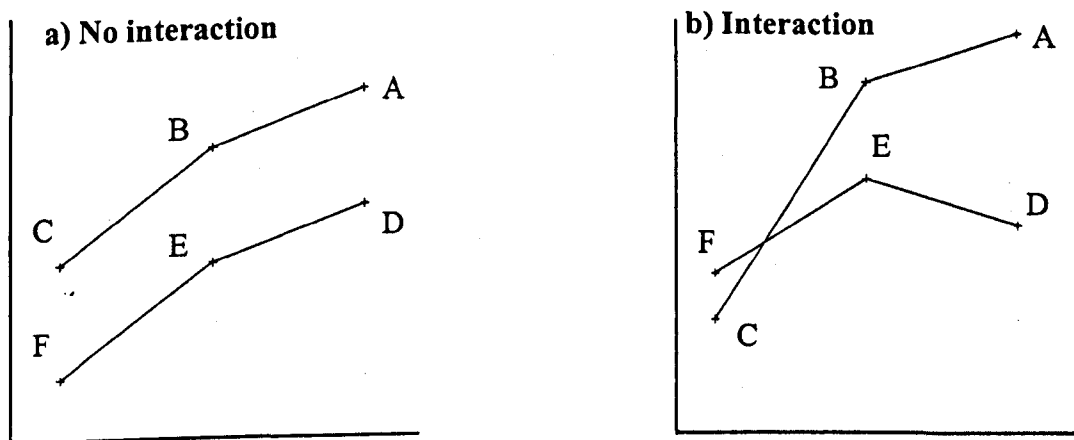
Interactions can also occur between process variables and quantitative variables, or between two quantitative variables. Here the quantitative variables are the mixture proportions. If the low medium and high groupings for the mixture variables are considered as qualitative conditions, interactions with and between mixture variables can be described in a similar way to the interactions between process variables.

B Consequences of the fixed sum constraint on interactions

The number of ingredients in the mixture is denoted by q. Recall that if the main effect of a process variable is included in the model, it is not possible to include all q interaction terms between the mixture components and the process variable (see §2.3.2). Cornell (1992, §7.5) suggests solving this problem by dropping one of the main effects from the mixture model and using the q-1 term parameterisation (see §2.3.2). However inclusion of the main effects normally leads to a model which is easy to interpret. Cornell's approach is therefore rejected here, in favour of the full canonical model (see §2.3.2), and one interaction with a mixture variable will be excluded for each process variable.

When q–1 of the interactions are used, the model will fit the data equally well irrespective of which q-1 are chosen (see §2.3.2). Although the choice of interactions will not affect the accuracy of the model, there may be more of an interest in estimating the effect of certain interactions explicitly.

237

7.4 Modelling Preliminaries

7.4.1 Explanatory Variables of Interest to JM

Recall that the aim of this chapter is to produce a model for predicting the rheological properties of glaze slip (see §1.1). The explanatory variables to be used in the model are the proportions of the three mixture components and some qualitative information on the processing conditions. Recall that the design used for the experiment was optimal (over a set of candidate points) for the quadratic model. The quadratic model for q=3 has 6 parameters and is given as equation 7.4, where η_q represents the mixture variable effect on viscosity. For a description of what each of the model terms represent see section 7.3.1.

$$\eta_q = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 \quad (x_{1,2,3} \in [0, 1]) \qquad (7.4)$$

To model the effect of the processing conditions it was decided that all main effects and interactions would be considered for inclusion in the model. The appropriate full model is given as equation 7.5, where η_{pv} represents the process variable effect on viscosity. For a description of what each of the model terms represent see section 7.3.1. It must be noted that as terms z_2 and z_3 represent three levels of the same variable, there is clearly no interaction between z_2 and z_3. It must also be noted that because equation 7.5 does not include a mean it will only provide an adequate model when used in conjunction with the mixture model.

$$\eta_{pv} = \alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3 + \alpha_{12} z_1 z_2 + \alpha_{13} z_1 z_3 \qquad [z_{1,2,3} = 0 \text{ or } 1] \qquad (7.5)$$

To produce a full model, which includes all mixture terms, all process variable terms and all interactions between them, equations 7.4 and 7.5 are combined to give a 42 term model. The 42 term model is given in equation 7.6. Equation 7.6 introduces the new notation γ_a^b. Here a and b denote the mixture variable and processing condition indicator variable respectively (Cornell 1992). For example, γ_{12}^{13} represents the coefficient of the interaction between frit, water, flocculant and fine particles. Although it is not possible to include the main effect of a process variable as well as its interactions with all of the mixture variables,

all interactions are given in equation 7.6 to illustrate the full range of terms which could be selected.

$$\eta \quad = \sum_{i=1}^{3} \gamma_i^{0} x_i \quad + \sum_{i<j=1}^{3}\sum \gamma_{ij}^{0} x_i x_j \tag{P}$$

$$+ \sum_{m=1}^{3} \gamma_0^{m} z_m \quad + \sum_{m=2}^{3} \gamma_0^{1m} z_1 z_m \tag{R}$$

$$+ \sum_{m=1}^{3} [\sum_{i=1}^{3} \gamma_i^{m} x_i + \sum_{i<j=1}^{3}\sum \gamma_{ijm} x_i x_j].z_m \tag{T}$$

$$+ \sum_{m=2}^{3} [\sum_{i=1}^{3} \gamma_i^{12} x_i + \sum_{i<j=1}^{3}\sum \gamma_{ij}^{1m} x_i x_j].z_1 z_m \tag{V} \tag{7.6}$$

The P and R type terms describe the effects of the mixture and process variables independently. The first order interactions (T) represent modification to the pure mixture parameters when the levels of the process variables are changed. The V type terms represent three or four way interactions between process variable interactions and the mixture variables, or their interactions.

Equation 7.6 is given for completeness. Because of a lack of degrees of freedom for error (see §7.4.4) there is no intention of including all of its terms in either the starting or the final model. However, depending on how the model selection procedure progresses, all terms in equation 7.6 may be considered for inclusion.

7.4.2 Consequences of the Experimental Structure

The standard approach when conducting an experiment is to perform all trials in a random order. A random order will often increase the chance that the assumptions required for linear modelling are met (see §2.3.6A) If these linear modelling assumptions have been verified, standard General Linear Modelling (GLM) analysis can be used. A discussion of GLM analysis can be found in most standard statistical text books.

Due to practical considerations it was much easier to embed the processing variables within the mixture blends (see §5.4.1). This meant that for each glaze slip milled, a separate sample was taken for each of the six settings of the process variables. This is known as a *split plot structure*. Each set of six samples will subsequently be referred to as a *batch*. They are

239

not referred to as mixtures, because 3 of the mixtures were replicated. Note that a batch is different from a block. Each of the three blocks contained 4, 5 and 3 batches respectively (see §6.4) The six samples in a batch are likely to be subject to similar error effects (see §2.3.6C). Consequently the error terms within a batch are likely to be correlated. Models which include a correlation structure are discussed in section 7.6.

7.4.3 Nuisance Parameters

In addition to the mixture and process variables, model terms are required to describe the variations in experimental conditions which are not of explicit interest. Recall that the variables which are not of explicit interest are the alternative pre-shear conditions (see §6.2.8), the block effects (see §6.4.1) and the mill effects (see §6.4.1).

To allow direct comparison between all samples it would have been desirable to use the same measurement settings throughout the experiment (Herritage 1994 §3D). Unfortunately the pre-shear stress for 20 of the 68 experiments had to be modified so that a yield stress could be estimated (see §6.2.8). The indicator variable *psh* is also added to the model for specifying the trials which required these alternative conditions. This indicator variable was not tested for removal from the model, because it is known that changing the pre-shear history will have an effect on the test (Herritage, 1994), whether it is statistically significant or not.

Another experimental condition which needs to be taken into account is the fact that the experiments were performed in three blocks (see §6.4.1). Two indicator variables will be used to describe the block effects. The first variable will describe the difference between block 1 and block 2, and the second variable will describe the difference between block 1 and 3. The indicator variables used to represent the block effects are subsequently referred to as bl_1 and bl_2.

The final experimental condition which needs to be taken into account is the mill jars used (see §6.4.1). The four mills which were used on the first day were also used on the second day with exactly the same media. The remaining mills on the second day, and all mills on the third day, were all different. Four indicator variables were considered for addition to the model, to explain how each mill may have affected the measurements. The indicator variables used to represent the mill effects are subsequently referred to as ml_1 to ml_4.

The terms described in this section are subsequently referred to as nuisance parameters, because their effect is of no explicit interest.

7.4.4 Restrictions on the Size of Model

Due to experimental time constraints, only nine distinct mixtures were analysed. This lack of distinct mixtures meant that although 68 samples were measured, there were extreme limits on the size of model which could be fitted. This was not considered to be a major problem because it is often more sensible to have a final model which is small and easily interpreted, rather than a complex model.

If the quadratic model is used there will be 12 columns of the design matrix which have the same entry for all trials in a batch. These 12 columns are the six which represent the quadratic mixture model (see §7.4), the two which represent the block effects and the four which represent the mill effects. The 12 parameters which are represented by these columns are subsequently referred to as mixture level parameters. Twelve batches of mixture were produced with nine of them being distinct and three being replicates (see §6.3B). Because none of the replicates were ground in the same mill, the mill effects distinguish each of the trials in a replicate pair from each other, and there are essentially 12 batches which are distinct at the mixture level. If we concentrated on a particular level of the processing conditions, and fitted a model including all 12 mixture level parameters, there would be no degrees of freedom for error, i.e. the design would be saturated. If one of the mixture level parameters were omitted from the model, it would be possible to perform hypothesis tests but the degrees of freedom would be low and the tests might be less reliable (that is unnecessary terms are more likely to be significant and/or necessary terms are more likely to be non significant). The more mixture level terms which are left out the higher the degrees of freedom would be, and the reliability of the hypothesis tests would increase. Because three of the trials only differ in the factor variables used to describe the mill effects, the problem with having few degrees of freedom is compounded.

Each batch was then split into separate samples for the 6 levels of the process variables. Because four of the resulting 72 samples could not be analysed, there are 68 distinct trials for fitting the models. Recall that there are only 5 parameters in the process variable model. If a model were fitted, including these 5 parameters and the 12 mixture level parameters, the design would still be saturated at the mixture level. If the model were reduced so that it only contained 11 mixture level parameters there would still only be 1 degree of

freedom for error. Consequently, hypothesis testing would still be unreliable. Additionally, the large amount of repetition within the 68 trials is likely to produce a high degree of collinearity between the parameter estimates. These collinearities mean that the problems with hypothesis testing which were discussed above are not alleviated.

The fact that the mill parameters only apply to twelve trials each worsens the above problem. It was therefore decided that the model selection procedure would start with the simplest sensible model and gradually add significant terms. The simplest sensible model will include the three main effects of the mixtures, the three main effects of the process variables, and the indicator variable to describe the different measurement conditions (see §7.4). It was also decided that a maximum of 8 mixture level parameters would be included in the final model, to allow at least 4 degrees of freedom for error at the mixture level.

7.4.5 Outline of the Modelling

Despite the split plot structure of the data, and the correlated errors which this is likely to produce, the first stage of the analysis will be to fit a series of models under the assumption that the errors are uncorrelated with equal variance. The models based on uncorrelated errors will give an indication of what conclusions might be obtained from the modelling which assumes a correlated error structure. The procedure for fitting the models is given below (see §7.5.3)

Once the models which assume independent errors have been fitted, numerical maximum likelihood will be used to fit a model which takes the correlation structure (see §7.4.2) into account. As well as the standard maximum likelihood approach (i.e. ML, see §2.2.4), models fitted using the restricted maximum likelihood approach (i.e. REML, see §2.2.5), will also be considered. In both cases the adequacy of the fitted models will be assessed using the multivariate Normal profile log likelihood. A Matlab routine was written to fit the models and to evaluate the log likelihood. The code for this routine is contained on the disk at the end of the thesis and is called reml.m.

7.5 Modelling Assuming Uncorrelated Errors with Equal Variance

7.5.1 Outline

A Introduction

There are two ways of approaching model selection, removing terms from the biggest possible model or adding terms to the simplest sensible model. These two approaches are known as backward and forward selection respectively. It is usually best to fit the fullest model possible and then remove the terms which are not needed, i.e. backward selection. Because there were so few distinct trials a combined approach was used. This combined approach involved extending the simplest sensible model and then testing some of the original terms for deletion.

B Dealing with multiplicity

The repeated testing of many parameters results in an effect known as *multiplicity*. Every time a test is performed there is a chance of finding terms to be significant when they are not really needed. In multiple tests the probabilities of this for each individual stage combine to produce a strong possibility that terms which are not needed have been included. Because there are so many possible parameters in this analysis (see §7.4), testing all of them is likely to result in multiplicity. To reduce the risk of multiplicity, groups of *similar terms* will be considered instead. These groupings can be done in many ways. Two examples of the types of group used are all interactions between frit and the three process variables, or all interactions between flocculant and two of the mixture variables (see §7.4). Additionally, given that they fit well, models which include sets of parameters are advantageous because they are simpler to interpret. Models consisting of terms with no clear patterns are likely to be confusing.

Another way of tackling multiplicity is to use the Bonferroni correction. In the simple case, Bonferroni's correction involves dividing the chosen significance level by the number of tests being performed. When forward selection is being used we do not know in advance the number terms that we will eventually test. Additionally, there are so many terms for possible inclusion that Bonferroni's adjustment is likely to result in highly conservative tests. As a compromise we will initially use a 5% confidence level. We will then consider the whole process again and re-consider the inclusion of terms if many of them are only just significant at the 5% level.

C Simultaneous testing for non-nested terms

In all cases, when there is no natural nesting of individual terms, or groups of terms, the relevant groups for possible inclusion were all tested before any were added. If one or more of the terms, or groups of terms, are significant, the one which gives the biggest improvement in log likelihood will be added. The same procedure will then be repeated for the remaining terms for possible inclusion. This procedure will be repeated until no more terms are significant, or all terms have been added. This approach will subsequently be referred to as *simultaneous* testing.

D Order in which terms will be considered

After the residuals for the smallest sensible model have been checked (see §7.5.2), individual terms, and groups of terms, will be tested in the order given below. No additional nesting is assumed and the groups within each group type will be tested simultaneously. Testing will be based on the hypothesis tests described below (see §7.5.3). The first six of the group types are terms for possible addition to the model. The last two are terms for possible removal.

1) Batch effects.

2) Mill effects.

3) Interactions between just mixture variable terms. Recall that the number of terms at the purely mixture level will be restricted to 8 (see §7.4.4). As some mixture level nuisance parameters may already have been added (see §7.4) it may not be possible to consider all three mixture terms as a group.

4) Interactions between just process variables terms.

5) Groups of terms which include interactions with the same process variable. An example of this type of group would include the interaction between flocculant and frit, the interaction between flocculant and clay and the interaction between flocculant and all mixture interactions selected in stage 2. Note that the interaction between flocculant and water is omitted in the above example. This is a consequence of the fixed sum constraint (see §7.3.5B)

6) Interactions between mixture variables and any process variable interactions which might have been added in stage 4.

7) Removal of groups of points which are similar at the mixture variable level, i.e. all interactions between the process variables and frit which were added in stage 2.

Removal of main effects. The only main effect to be tested for removal was medium particle size. The main effects of the mixture variables were not considered for omission. If a main effect were to be removed the usual interpretation of a mixture model would not apply. The process variable main effects were assumed to be necessary without testing their statistical significance, since for the majority of the 34 pairs the samples with flocculant were visibly less easy to pour. Also for the majority of the 12 batches the samples which had been milled the longest were again visibly less easy to pour.

7.5.2 Residual Checking (Diagnostics)

The hypothesis tests discussed in section 7.5.3, and the model based predictions generated in section 7.8 are based on the assumptions about the residuals (see §2.2). If these assumptions are not met then the validity of tests and predictions are put in doubt. The assumptions were tested regularly throughout the modelling procedure. In all cases these tests are performed on the standardised residuals. Here the standardised residuals are defined to be the raw residuals divided by σ. The value of σ is estimated by the standard deviation, s, of the set of all residuals. The following three procedures were used to check whether the three assumptions about the residuals were met.

1) A test was performed to establish if there was any dependence between the residual errors of trials performed at similar times. To do this each residual was paired with the residual from the trial performed at the next time point, and the correlation coefficient between the two sets of values was calculated. These correlation coefficients were then checked against the 5% critical values for a one sided test (Neave, 1992). The residual values were assumed to be time independent if the correlation coefficient was lower than the critical value.

2) Each residual was plotted against its corresponding fitted value. These plots were then checked visually to see if they exhibited a random spread of points with constant variance, equally distributed around the fitted value axis. Any point which had a standardised residuals greater than 4 was considered to be a potential problem, and the appropriate graph of viscosity against shear stress was consulted.

3) To test the assumption that the residual values were normally distributed the Anderson-Darling test was used. Linear modelling techniques are more robust to non-normality of residuals than they are to non randomness of residuals. Consequently very strong evidence ($p < 0.02$) was required before non-Normality was deemed to be a possible problem.

7.5.3 Hypothesis Test Used in the Model Selection Procedure Discussed in Section 7.6.1

To determine whether a (set of) parameter(s) is needed, the log likelihood of the model excluding them is compared with the log likelihood of the model including them. The log likelihood is calculated using equation 7.8 which was derived from the likelihood in equation 7.7.

$$L(\underline{\beta}, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\{-\frac{1}{2} \sigma^{-2}(\underline{y}\text{-}X\underline{\beta})' \ (\underline{y}\text{-}X\underline{\beta})\} \qquad (7.7)$$

$$l(\underline{\hat{\beta}}, \hat{\sigma}^2) \quad = -\frac{n}{2} \ \{1+\ln(2\pi\hat{\sigma}^2)\} \qquad (7.8)$$

The test involves comparing twice the difference between the log likelihoods for the two models with the $\chi^2_{df,1\text{-}p}$ critical value. The term df represents the number of parameters which are being tested for and is known as the change in the degrees of freedom. The term p represents the level of the test. Initially a p value of 0.05 was used here.

Every time the fitted model changes by two parameters the three linear modelling assumptions were re-tested (see §7.5.2).

7.6 Modelling Assuming a Non-Identity Covariance Structure

7.6.1 Rationale for Including A Covariance Structure

Recall that due to practical considerations the processing variables were embedded within the mixture blends (see §5.4.1). The embedding of the process variables meant that separate samples were taken from a single batch of mixture for each of the six settings of the process variables, rather than milling a separate mixture for each of the six settings. All six samples from a particular batch were therefore subject to similar error effects. This effect can be described by assuming a correlated error structure.

Here we assume that the errors from trials in the same batch will be correlated with each other (see §5.4.1). The rationale for this is that covariates not included in the model are likely have a similar effect on all measurements within the same batch, causing the errors in

the batch to be similar and be positively correlated. For example, imagine the temperature is high when a particular batch is being prepared and a lot of the water contained in the batch is evaporated off. The resulting reduction in water is likely to cause the viscosity of each sample taken from the batch to be higher than would be expected for the given composition, and higher than would be predicted by the model. All errors for this batch would therefore be relatively high. In another batch of mixture, where evaporation did not occur, all errors might be relatively low. Any individual error is therefore likely to be similar to the error for another trial in its batch. That is the errors in a batch are correlated.

7.6.2 Modelling the Covariance Structure

Three different covariance structures are considered (see §7.6.3). In the simplest case it was assumed that all pairs of correlated trials had the same correlation coefficient. This correlation coefficient is denoted by ρ. The covariances between the errors are then specified by $V\sigma^2$, where V is an n×n matrix with entries of the form $f(\rho)$ as shown below (see §7.6.3). The diagonal elements of the matrix give the scaled variance of the measured response for each trial. The off-diagonal elements give the scaled covariance between each pair of measured responses. It is then assumed that the data vary as $N(X\beta, \sigma^2 V)$ and numerical maximisation of the likelihood function can be used to estimate V and β. The likelihood function of the multivariate normal distribution is given in equation 7.9.

$$L(\beta, \sigma^2, \rho) = (2\pi\sigma^2)^{-n/2}|V|^{-1/2}\exp\{(\underline{y}-X\beta)'V^{-1}(\underline{y}-X\beta)/2\sigma^2\} \qquad (7.9)$$

When both β and V are unknown they can be estimated by an iterative approach in which equations 7.10 and 7.11 are solved in turn. This approach is based on a starting estimate of β obtained by using V=I in equation 7.10.

$$\hat{\underline{\beta}} = (X'V^{-1}X)^{-1}X'V^{-1}\underline{y}. \qquad (7.10)$$

$$Min(|V|^{1/n}\underline{e}'V^{-1}\underline{e}), \text{ where } \underline{e} = \underline{y} - X\hat{\underline{\beta}}. \qquad (7.11)$$

In general, ML estimators of variance parameters are biased downwards. An alternative approach is to use Restricted Maximum Likelihood (REML) (see, e.g., Welham and Thompson, 1997). n general REML estimators of variance components are less biased than ML estimators. The REML likelihood is given in equation 7.12, where p is the number of parameters in the model for E(Y).

$$L_R(\beta, \sigma^2, \rho) = (2\pi\sigma^2)^{-(n-p)/2} |V|^{-1/2} |X'X|^{1/2} |X' V^{-1} X|^{-1/2} \exp\{(y-X\beta)' V^{-1} (y-X\beta)/2\sigma^2\} \qquad (7.12)$$

To find the REML estimates of β and V involves replacing equation 7.11, in the ML analysis, with equation 7.13.

$$\min(\{|V|.|X'V^{-1}X|\}^{1/(n-p)} \underline{e}'V^{-1}\underline{e}) \qquad (7.13)$$

The ML and REML procedures were implemented in matlab. In addition to the iterative procedure described above, equation 7.11 and 7.13 were minimised by an iterative search over the range of ρ. The matlab code is provided on the disk at the end of the thesis and is called reml.m.

7.6.3 Different Forms of the Covariance Matrix V

As discussed above, the main features of the results from the JM experiment can be summarised in several ways (see §7.2). The first stage of the analysis for each of these possibilities is to identify the most appropriate form for the covariance structure V. The covariance matrix can take many forms and in each case ρ could take the value zero. If we assume for illustration that there are only three batches of size 6, then V can take one of the following three forms.

Covariance structure 1 (Simplest)

The simplest approach is to assume that the correlation between each pair of trials in the same batch is ρ and the correlation between any 2 trials in different batches is 0.

$$V = \begin{bmatrix} R & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & R \end{bmatrix} \text{ where } R = (1-\rho).I_6 + \rho.1_{6\times6}$$

Covariance structure 2 (Toeplitz)

When the experiment was performed, the six trials in a batch were measured sequentially, according to the experimental protocol (see table 5.1 in §5.2.2). In the Toeplitz structure we assume that there is a geometric decay in the correlation for each stage of this sequence. This approach is subsequently referred to as covariance structure 2, and is implemented using the above formula for V, with R given by:

Covariance structure 2 (Toeplitz)

$$R = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \rho^4 & \rho^5 \\ \rho & 1 & \rho & \rho^2 & \rho^3 & \rho^4 \\ \rho^2 & \rho & 1 & \rho & \rho^2 & \rho^3 \\ \rho^3 & \rho^2 & \rho & 1 & \rho & \rho^2 \\ \rho^4 & \rho^3 & \rho^2 & \rho & 1 & \rho \\ \rho^5 & \rho^4 & \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Covariance structure 3 (Modified)

Recall that each batch of slip was analysed in three stages. When the coarse particle size had been reached, some mixture was removed form the mill and split into two samples. Flocculant was added to one of these samples and their viscosity was measured. The slip which was left in the mill was then milled further, and the same procedure was performed for the medium and fine particles. Because the milling process and particle size was identical for the two samples, one with flocculant and one without, taken at each particle size, each pair of samples of this form were likely to exhibit similar errors. This could be modelled using ρ_1 and ρ_2 for correlations within and between particle sizes. Here we modify the geometric decay structure discussed above and use ρ and ρ^2 for ρ_1 and ρ_2 respectively.

249

7.6.4 Modifications to the Model Selection Procedure When a Non-Identity Covariance Structure is Assumed.

The procedure used here is similar to the one described for independent errors (see §7.5.3) but involves numerical ML to estimate V and $\underline{\beta}$. Numerical REML (restricted maximum likelihood) is also considered. Similar hypothesis tests to the ones used for uncorrelated errors will then be used. When the errors were uncorrelated, i.e. V=I, the |V| term disappeared because |I| =1. Now that correlated errors are being assumed, the term does disappear, and is included in the profile log likelihood (see equation 7.14). The hypothesis

Covariance structure 3 (Modified)

$$
R= \begin{bmatrix}
1 & \rho & \rho^2 & \rho^2 & \rho^2 & \rho^2 \\
\rho & 1 & \rho^2 & \rho^2 & \rho^2 & \rho^2 \\
\rho^2 & \rho^2 & 1 & \rho & \rho^2 & \rho^2 \\
\rho^2 & \rho^2 & \rho & 1 & \rho^2 & \rho^2 \\
\rho^2 & \rho^2 & \rho^2 & \rho^2 & 1 & \rho \\
\rho^2 & \rho^2 & \rho^2 & \rho^2 & \rho & 1
\end{bmatrix}
$$

tests are therefore performed using the same approach (see §7.5.3), but with equation 7.8 replaced by equation 7.14.

$$
l(\hat{\underline{\beta}}, \hat{\sigma}^2, \hat{\rho}) = -\frac{n}{2}(1+\ln(2\pi\hat{\sigma}^2))-\frac{1}{2}\ln|\hat{V}| \tag{7.14}
$$

There are no accepted methods for using the REML likelihood in hypothesis testing, though some have been proposed (see, e.g., Welham and Thompson, 1997). It is usually suggested that parameter estimation should be based on ML during the model selection phase, but that the parameters in the final model should be estimated using REML techniques. This approach to model selection is labelled (2) below. It is suggested here, that because REML is of benefit when estimating the variance parameters in the final model, it should also be of benefit when estimating the parameters in each of the intermediate models considered.

To investigate the effect of using REML parameter estimation for intermediate models, both approaches are used and the results compared. As stated in section 7.6.3, three forms of covariance structure are considered.

The covariance structure used throughout will be the one which had the highest profile log likelihood when used for the starting model. In each cases the starting model will be the simplest sensible model (see §7.4.4) plus any nuisance parameters which were found to be significant for when the uncorrelated errors were assumed. Because there was no prior notion that the covariance structure should be the same for all three responses, the models were not constrained to have identical covariance structures.

To verify the assumptions made about the errors (see §2.2), the standardised errors were plotted against the fitted values. Here the error vector was standardised by being pre-multiplied by the Cholesky factorisation (of V) available in `matlab`. The Cholesky factorisation of V returns R, where $R' \times R = V$. The Anderson-Darling test for normality (see §7.5.2) was also used on the standardised errors.

The next stage of the model fitting procedure when a non-identity covariance approaches (2) and (3) was to investigate, first if any of the nuisance parameters included in stage (1) could be removed, and secondly if any of the other nuisance parameters were needed. The explanatory variables of interest to JM were then investigated using the same groupings and ordering that had been used for approach (1) (see §7.5.1D).

7.7 Models Fitted

Three methods for estimating the model parameters will be considered. The first method of parameter estimation assumes uncorrelated errors. The other two methods assume correlated errors and estimate the parameters using numerical ML or numerical REML. These three approaches are subsequently referred to as approaches 1, 2 and 3 as follows:

1) $V = I$

2) V estimated by ML

3) V estimated by REML

In the description of the modelling below, the test statistic and significance level are given in brackets next to the significant terms. No details of the non-significant terms are given. The first term in the brackets is the test statistic; twice the change in the appropriate log likelihood. The second term is the degrees of freedom. If no degrees of freedom are stated then the value is 1. The third term is the significance level. The significance levels were obtained using Table 3.2 of Neave, 1992.

7.7.1 Logging the Response to Ensure Meaningful Predictions

One of the most important criteria when selecting a model, especially when the model is to be used for prediction, is that the predicted values are realistic. When the smallest sensible models were fitted, some of the fitted values obtained for each of the response variables were found to be negative. Even when a number of significant terms had been added, negative values were still predicted. The three aspects of rheology which are of interest (see §7.2.1) are only defined for positive values. To solve the problem of negative fitted values, the fitted/predicted values were constrained to be positive, by log transforming the response variables. The predicted response will be the (positive) exponential of the value predicted by the fitted model. The log transformation was used because other transformations which are only defined for a positive domain are more complex to interpret. Unless it is stated otherwise, the response variables in the subsequent analysis are these log transformed responses. The responses are not exponentially transformed back to the original scale until the trace plots of section 7.9.

7.7.2 Diagnostics, Nuisance Parameter Selection and Choice of Covariance Structure

A Diagnostics

Recall that the starting models included seven terms: the three mixture term main effects, the three process variable main effects and the effect of the alternative measurement conditions. For all three fitting methods the starting models for the top plateau and the yield point produced standardised errors which satisfied the appropriate assumptions. For all three methods of parameter estimation the starting model for the bottom plateau yielded a standardised error of approximately 5 for run 4D. It was discussed above that this run had a bottom plateau viscosity which was at least 4 times bigger than the corresponding response for all other runs (see §7.7.2A). It was also mentioned that this run was likely to have a high viscosity because of its composition. However, because this was one of the 24 runs for which a yield point was not reached, it is possible that the true bottom plateau is considerably lower than the observed one. It was decided, therefore, to omit this run from all further model fitting. For all three fitting methods the starting model was fitted to the remaining 67 points and all errors satisfied the assumptions.

B Nuisance parameters needed in models which assume uncorrelated errors

When uncorrelated errors were assumed the log likelihoods of the starting model for the top plateau, the yield point and the bottom plateau were -68.4, -25.2 and -18 respectively. The following nuisance parameters were then found to be significant:

Top Plateau: bl_1 (20.2, 0.0005), ml_1 (17, 0.0005), ml_2 (9.6, 0.005).

Yield Point: bl_1 (17, 0.0005), ml_1 (12.4, 0.0005), ml_2 (7, 0.01).

Bottom Plateau: bl_1 (34, 0.0005), ml_1 (8.8, 0.005).

C Covariance structure and nuisance parameters needed in the model when approaches (2) and (3) were used

To decide which covariance structure to use, a model was fitted which added the above nuisance parameters to the simplest sensible model. The evaluated log likelihood for each covariance structure and both ML and REML estimation are given below. Because the different covariance structures are not nested, the hypothesis tests discussed above can not be used. However it seems intuitively sensible that if the log likelihoods differ by an amount which would provide evidence of a difference if the models were nested, then the correlation structure with the highest likelihood would be likely to give a better fit. In the tables below the *difference* column gives twice the difference between the log likelihood for the modified structure and the log likelihood for the standard structures, and the appropriate p-value for a test of nested models.

The structure selected is highlighted. For the top plateau the difference was too small to suggest that any of the structures were likely to be better than the others. If there had been some prior notion that the covariance structure should be the same for all three responses, the modified structure would have been used. Because there was no such notion the standard structure was chosen because it is the simplest.

This analysis shows that the error structures are of the same form for the bottom plateau and the yield point but of a simpler form for the top plateau. It was suggested earlier that the top plateau is likely to be caused by wall slip. Wall slip is a consequence of the measurement procedure and is unlikely to be scientifically significant in terms of rheology. The yield point and bottom plateau, however, are scientifically significant. It is therefore not surprising that the errors structure used for the top plateau is different to the error structure used for the other two responses.

Table 7.5 Starting model log likelihoods, for all three covariance structures, and model fitting methods (2) and (3)

Top Plateau

	Modified	*difference*	Standard	Toeplitz
(2) ML	-40.4	0.4, 0.6	-40.2	-40.8
(3) REML	-41.4	0.2, 0.7	-41.3	-41.4

Yield Point

	Modified	*difference*	Standard	Toeplitz
(2) ML	**1.3**	2.2, 0.15	0.2	-3
(3) REML	**0.3**	2, 0.2	-0.7	-3.7

Bottom Plateau

	Modified	*difference*	Standard	Toeplitz
(2) ML	**13.5**	8.6, 0.005	9.2	8.4
(3) REML	**12.3**	8.4, 0.005	8.1	7.4

D Nuisance parameters needed in the model when the chosen covariance structure had been chosen

When the nuisance parameters, which had been added using method 1, were tested for removal some of them were significant. The change in likelihood for both ML and REML parameter estimation, i.e. methods (2) and (3) are given below for the terms which were now found to be non-significant. The criteria for the terms to be non-significant was that if they had been added using forward selection, then would not have been chosen for inclusion. The appropriate significance level is also given.

Table 7.6 Non-significant nuisance parameters using methods (2) and (3)

	(2) ML	(3) REML
Top Plateau	No terms	No terms
Yield Point	ml_2 (2.6, 0.15), ml_1 (4, 0.05).	ml_2 (1.8, 0.2), ml_1 (3.6, 0.075).
Bottom Plateau	ml_2 (0.2, 1), ml_1 (2.4, 0.15).	ml_2 (-0.8, 1), ml_1 (1.8, 0.2).

Note that when ML estimation was used for the yield point, mill one was significant. Note that the bottom plateau and the yield point have the same error structure and mill two was not significant for the bottom plateau or when REML was used for the yield point. Therefore it seems intuitively sensible that the error structure is homogenous. Consequently mill two was removed when ML was used for the yield point.

Note also that when mill two was removed from the bottom plateau the REML likelihood increased. The REML procedure involves projecting the data into the vector space where all parameter contrasts have zero expectation (see, e.g., Welham and Thompson, 1997). Therefore it is feasible that the REML likelihood could increase.

7.7.3 Significant Explanatory Parameters to be Included in the 'Best' Models

In this section individual terms and groups of terms are tabulated in the order in which they were found to be significant. Details of the ordering are given above (see §7.5.1D). The notation $A \times z_i$ is used to indicate the set of all A interactions between the process variable z_i and the mixture terms included in the model. For example, in Table 7.7 the term $3 \times z_1$ indicates the interactions between x_2, x_3 and the $x_2 x_3$ term which was found to be significant at the first stage. Recall that the absence of the x_1 interaction is a consequence of the fixed sum constraint (see §2.3.2).

Table 7.7

A Top Plateau

Term	Method used to fit models		
	(1) V=I	(2) ML	(3) REML
x_2x_3	(19.2, 0.0005)	(24.2, 0.0005)	(22, 1, 0.0005)
$3 \times z_3$	(13.8, 3, 0.005)	(9, 3, 0.05)	(9, 3, 0.05)
$3 \times z_1$	(12.4, 3 0.01)	(11.2, 3 0.025)	(22, 2, 0.005)

B Yield Point

Term	Method used to fit models		
	(1) V=I	(2) ML	(3) REML
x_2x_3	(10.3, 0.0005)	(8.4, 0.005)	(7.8 0.01)
$3 \times z_3$	(24.4, 3, 0.0005)	(10.2, 3, 0.025)	(10.8, 3, 0.025)
$3 \times z_1$	(10.8, 3, 0.005)	(11.8, 3, 0.01)	(11.8, 3, 0.01)

Note that the 'best' model for the top plateau and the yield point contain the same parameters of explicit interest. However, while the 'best' model for the top plateau contains bl_1, ml_1 and ml_2, the 'best' model for the yield point only contains bl_1.

C Bottom Plateau

No additional terms were found to be significant at the 5% level for the bottom plateau.

Note that the models selected above are referred to as the 'best' models. This is to distinguish them from the final models, which will be selected once the $PRESS_{mod}$ statistic (see §7.8) has been taken into account.

Recall that a consequence of the fixed sum constraint is that the model can not contain the q interactions between a process variable and each of the q mixture variable main effects.

Models which contain any q-1 of them fit the data equally well. It was decided above that the interactions with x_2 would be the ones omitted. When the parameter to describe the affect of $x_2 x_3$ was found to be significant, the interactions between the process variables and x_1 were omitted and were replaced by the appropriate interactions with x_2. This was done because it is not intuitively sensible to have interactions with $x_2 x_3$, but not interactions with x_2, unless there is some prior notion that this might reflect reality better.

7.7.4 Re-assessing the Hypothesis Tests

Because most of the significant terms were significant at the 0.5% level, it was decided that there was no need to reconsider the testing using the Bonferroni correction (see §7.5.1B). The final stage of the model fitting procedure is to test whether the 'best' models were significantly different from the starting models and whether models which contain many additional parameters were significantly different from the 'best' models. The model which contains many additional parameters is given in Figure 7.9. This model is subsequently referred to as the *moderately full model*. The only nuisance parameters to be included are those which were found to be significant by the hypothesis tests. This is because their inclusion would cause the model to be saturated at the mixture level (see §7.4.4). Given the high levels of significance which were observed for the majority of terms it is unlikely that these final tests will suggest any models other than the ones selected. Details of the test statistics and the subsequent significance levels are given in Table 7.8.

Table 7.8 suggests that in all cases the 'best' models are significantly better than the simple models. The table also shows that in all cases the full models are also significantly better than the 'best' models. This is not a surprising result, because the full models leave few degrees of freedom for error at the mixture level, and hypothesis tests are unlikely to be reliable. The full models are considered further in the PRESS type analysis of the next section. However, given the low degrees of freedom at the mixture level, and the expected collinearities between the parameter estimates (see §7.4.4) the full models will not be selected as the 'best' models. For the same reasons, the full models are unlikely to perform well according to any criteria. It is envisaged that the PRESS type analysis will give evidence to suggest that the full models should not be used.

Table 7.8 Testing the significance of the 'best' model

The bracketed terms in the 'best' model and full model columns are details of the hypothesis tests in the usual format. The bracketed terms in the starting model column indicate the number of parameters in the starting models for E(Y).

(a) Top Plateau

	Starting	'best'	Moderately full
(1) V=I	-64.8(8)	-18.7 (92.2, 9, 0.0005)	-5.7(26, 13, 0.025)
(2) ML	-40.2(10)	-18.2 (44, 7, 0.0005)	-5.2(26, 13, 0.025)
(3) REML	-41.3	-20.4 (41.8, 7, 0.0005)	-8.4(24, 13, 0.05)

(b) Yield Point

	Starting	'best'	Moderately full
(1) V=I	-25.2(7)	10.8(28.6, 10, 0.005)	45.2(68.8, 13, 0.0005)
(2) ML	1.3(10)	13.2(23.8, 5, 0.0005)	29.8(33.2, 13, 0.005)
(3) REML	0.3(10)	12.8(25, 5, 0.0005)	28.9(32.2, 13, 0.005)

(c) Bottom Plateau

	Starting	'best'	Moderately full
(1) V=I	-18(7)	7.4(50.8, 3, 0.0005)	24.9(35, 19, 0.025)
(2) ML	13.5(9)	No terms added	33.2(39.4, 20, 0.01)
(3) REML	12.2(9)	No terms added	32.3(41, 20, 0.005)

Figure 7.9 The explanatory variables of explicit interest in the 'moderately full model'

Terms which apply to the mixture variables only are represented by the left hand column. Terms which apply to the process variables only are represented by the top row. A process variable in the middle denotes an interaction between that process variable and the mixture term to its left.

	z_1	z_2	z_3	$z_1 z_2$	$z_1 z_3$
x_1					
x_2	z_1		z_3		
x_3	z_1		z_3		
$x_1 x_2$	z_1		z_3		
$x_1 x_3$	z_1		z_3		
$x_2 x_3$	z_1		z_3		

7.8 Predictive Power of Models

7.8.1 Outline

The main aim for the final models is to use them for predicting the viscosity of future glazes. One way to investigate how well a model is likely to predict future responses is to fit it using a reduced data set and see how well it predicts the deleted responses (see §2.2.6). The discrepancy between the observed and predicted values are subsequently referred to as *prediction residuals*. If this is repeated for all data points then the sum of the squared prediction residuals is known as the PRESS statistic. Here we use a modified form of the PRESS statistic which accounts for the replicated trials. This modified form of the PRESS statistic is subsequently referred to as the $PRESS_{mod}$ statistic.

259

A selection of the following models will be investigated for each response variable. Model types 1, 2, 5 and 6 will be investigated for all three responses. Types 3 and 4 only apply to the bottom plateau. The appropriate model types will be investigated for all three fitting procedures (see §7.7).

1. Simplest sensible model plus significant nuisance parameters.

2. Best model suggested by hypothesis tests

3. Any models which were suggested by one of the other 2 fitting procedures.

4. Models including all parameters for which the hypothesis tests showed weak evidence of an improvement at the 15% level (see. §7.5.3). The models which include all terms for which there was weak evidence of an improvement are referred to as 'weak evidence' models below.

5. Moderately full model. This is the model given in Figure 7.9 plus all nuisance parameters which were found to be significant.

6. Full Model. This is the moderately full model plus all interactions between the quadratic process variable model and the processing condition interactions.

7.8.2 Methodology

The individual prediction residuals and the $PRESS_{mod}$ statistic were calculated using the following procedure.

1. Remove the data for one of the trials. For the replicate pairs this will involve the removal of the data for two trials. Both trials in a replicate pair were removed because if one of them remained we would be calculating a fitted value for the other one rather than a predicted value.

2. Fit the model in question to the remaining data points.

3. Using the fitted model, estimate the response for the glaze slip whose data was removed in stage 1.

4 Calculate the difference between the observed and predicted response (and the square of this difference) for the glaze removed in stage one.

5. Replace the data which was removed in stage 1.

6 Repeat stages 1 to 5 for each distinct glaze.

7. Add up the squared differences to calculate the $PRESS_{mod}$ statistic. This name is used because we removed the data for 2 trials when we consider a point which was replicated.

Using `matlab` an algorithm was constructed for the analysis of predictive power as described above (see §7.8.1). The code for this algorithm is included on the disk at the end of the thesis and is called `pres.m`.

7.8.3 How to Interpret the $PRESS_{mod}$ Statistic.

Although it may not be the best measure for correlated data, the performance of each model was assessed by the sum of the squared differences calculated in stage 4. This statistic is similar to the PRESS statistic. Because the standard PRESS statistic has no known distribution (Draper and Smith, 1998) there is no accepted hypothesis test for investigating whether differences between models are significant. Given that the standard PRESS statistic has no known distribution, the $PRESS_{mod}$ statistic is unlikely to have a known distribution.

Models selected for their predictive power are potentially different from models selected for explanation. However, as the predictive power is likely to improve with the understanding of the system, the models produced for the two purposes are likely to be similar. In the case of independent errors the chi squared values used in the hypothesis tests are related to the predictive power. For correlated errors the relationship between the PRESS and chi squared statistic are less clear. It is likely that if terms were found to be significant in the hypothesis tests, then any improvement they cause in the PRESS or $PRESS_{mod}$ statistic is also likely to be a real effect. The method which was used for investigating predictive power is extremely computer intensive. As models produced for the two different purposes are likely to be similar, the hypothesis tests discussed above (see §7.7) can be used to indicate which terms are likely to improve predictive power.

Models which do not give the lowest $PRESS_{mod}$ value might actually be the best models for other reasons. For example, a high $PRESS_{mod}$ value might be caused by a small number of poor predictions in specific areas of the explanatory variable space. Predicted values in other areas might be considerably better than the predicted values from the model which did have the lowest $PRESS_{mod}$ value. If this were found to be the case then it would be sensible to choose the model which did not have the lowest $PRESS_{mod}$ value but to restrict its area of application. A separate model could then be fitted for glaze slips which fell outside

this area. To investigate the possibility of the PRESS statistic being misleading in ways such as this, box plots of the prediction residuals for each model will be compared.

7.8.4 PRESS$_{mod}$ Analysis

All models analysed in this section assumed the covariance structures identified as best in the hypothesis tests (see §7.6.4). In the tables below the highlighted model is the one which was found to be the best and p represents the number of terms in the model.

For the top plateau all three approaches to parameter estimation suggested the same 'best' model and there were no terms for which weak evidence was found.

For the yield point the assumption of uncorrelated errors with equal variance led to a slightly different model than the other two approaches. When uncorrelated errors were assumed, the hypothesis tests gave evidence to suggest that the effects of block 1 and mills 1 and 2 should be included. When a correlated error structure was used, the tests gave evidence to suggest that only block 1 was required. In Table 7.9 the notation (1)V=I is used to indicate the model with the additional nuisance parameters which was suggested by the fitting method (1). Consequently two tables are given below for the yield point.

For the bottom plateau, weak evidence was obtained for the inclusion of a variety of additional terms. The details of the appropriate hypothesis tests are detailed below:

(1)V=I: x_3x_3 (8, 0.005)

(2)ML/ (3) REML: z_1z_2 (2.2, 0.15) 2×z_3 (4, 0.15)

Recall that the 'best' model selected for the top plateau and the yield point contained the same parameters of explicit interest. In the analysis below we investigated whether these parameters would lead to a good predictive model for the bottom plateau. This model is referred to as 'Top / Yield' model in Table 7.9(c)

Additionally the nuisance parameters which were significant were different when uncorrelated errors were assumed to when correlated errors were assumed. When uncorrelated errors were assumed the hypothesis tests suggested that the effects of the block 1 and mill 1 should be included. When correlated errors were assumed the tests gave evidence to suggest that only block 1 was needed. Consequently two tables are given below for the bottom plateau.

Table 7.9 PRESS$_{mod}$ statistics

(a) Top Plateau

Model	p	PRESS$_{mod}$		
		(1)V=I	(2)ML	(3)REML
Start	10	19.2	18.9	18.6
'Best'	17	13.8	14.6	13.9
Moderately Full	30	19.3	26.9	19.6
Full	40	51.9	68.9	49.9

When the box plots were produced they all suggested that the models with the lowest PRESS$_{mod}$ values did in fact give the best fit in terms of individual prediction residuals. Two example of these box plots are given in Figure 7.10. Both figures relate to the models fitted for the top plateau. Figure 7.10(a) shows the prediction residuals for the four models fitted using REML. Figure 7.10(b) compares the prediction residuals for the 'best' model when it was fitted using REML with those obtained when it was fitted using ML.

For the top plateau, the PRESS$_{mod}$ analysis suggests that the most successful model in terms of prediction was the 'best' model suggested by the hypothesis tests when the 'modified' correlated error structure was used.

For the yield point, the PRESS$_{mod}$ analysis suggests that the most successful model in terms of prediction is actually the 'best' model which was suggested by the hypothesis tests when an uncorrelated error structure was used. When the effects of mills 1 and 2 were added to the model, the PRESS$_{mod}$ value reduced by approximately 35%. When there is no hypothesis test to investigate whether terms are needed, parsimony should be a consideration. Here it was decided that a decrease of 35% is enough to warrant the inclusion of 2 additional terms. The fact that these 2 terms have now been put back in the model highlights the fact that hypothesis tests can suggest the removal of necessary terms, especially when a large number of tests are being performed.

Table 7.9(b) Yield Point

Model	p	PRESS$_{mod}$	
		(2)ML	(3)REML
Start	8	7.5	7.5
'Best'	15	4.9	
V=I	17	3.2	3.1
Moderately full	28	4.5	4.6
Full	38	5.6	5.7

Model	p	PRESS$_{mod}$
		(1)V=I
Start	10	6.64
'Best'	15	5.6
V=I	17	3.12
Moderately full	30	3.5
Full	40	7.4

For the bottom plateau there was little difference between the starting model and the two models suggested by the hypothesis tests. For the sake of parsimony, and because it was suggested by the REML analysis, the starting model is chosen.

The comparatively poor performance of the 'moderately full' models, and extremely poor performance of the 'full' models, suggests that the decision to limit the number of parameters, due to saturation at the mixture level and collinearities between the parameters estimates (see §7.4.4) was well founded.

Table 7.9(c) Bottom Plateau

Model	p	PRESS$_{mod}$	
		(2)ML	(3)REML
Start	8	22.1	22.2
+x$_1$ x$_3$	9	22.3	22.5
+z$_1$z$_2$, 2×z$_3$, -ml$_1$	11	22.4	22.6
Yield / Top	17	24.8	26.6
Moderately Full	28	24.8	24.5
Full	38	34.9	34.9

Model	p	PRESS$_{mod}$
		(1)V=I
Start	9	23.2
+x$_1$ x$_3$	10	23
+z$_1$z$_2$, 2×z$_3$, -ml$_1$	12	23.6
Yield / Top	16	24.5
Moderately Full	29	26.4
Full	39	48.9

It has been suggested that REML is the best method for estimating the parameters of a final model . In terms of the predictive power of the models, REML only led to a noticeable improvement for the top plateau. It is likely that the ML values were higher because of the additional outlier in Figure 7.10(b). For the other responses REML led to a worse PRESS$_{mod}$ value as often as it led to a better one.

Figure 7.10

(a) Prediction residuals for the top plateau models fitted by REML
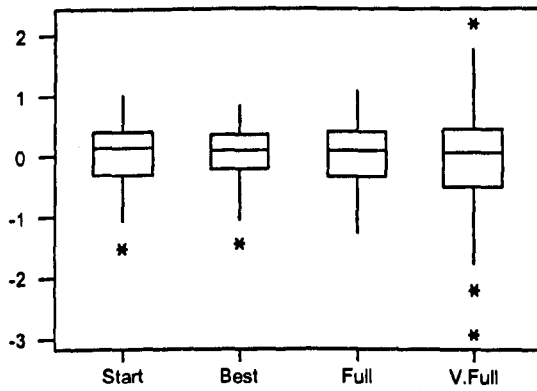
(b) Prediction residuals fitted for the 'best' model fitted by ML and REML
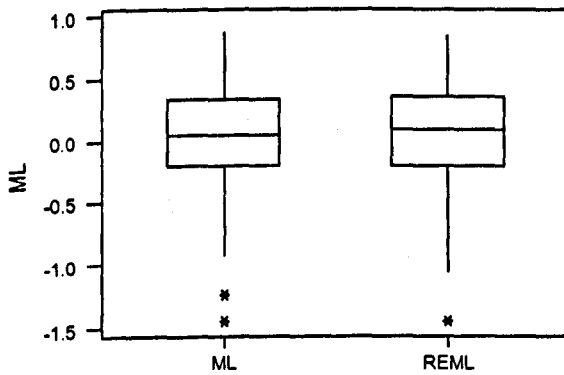
7.8.5 Summary of Final Models

Figure 7.11 indicates the form of the final models using a similar format that was used for the full model above (see §7.7.4). The difference here is that the nuisance parameters are included on the right hand side of the top line.

Figure 7.11 Terms in the final models

(a) Top Plateau / Yield Point

	z_1	z_2	z_3
x_1			
x_2	z_1		z_3
x_3	z_1		z_3
$x_2 x_3$	z_1		z_3

psh, bl_1, ml_1, ml_2 .

(b) Bottom Plateau

	z_1	z_2	z_3
x_1			
x_2			
x_3			

psh, bl_1

7.9 Interpreting the Final Models

7.9.1 Introduction

The models selected in the previous section were fitted using REML estimation. The parameter estimates are given in tables 7.10 to 7.12 for the three response variables respectively. It must be noted that in each case the models are for estimating the log of the response. Separate Piepel direction response trace plots (see §2.4) are then given for each of the six levels of the process variables (see §7.3.3) for each of three responses. To produce the Piepel plots the log response variables were estimated using the final models. The plots are then constructed using the exponential of the predicted response, so that they are on the same scale as the original measurements. There is no sensible way of converting the parameter estimates onto the original scale.

It must be noted that interpreting the individual parameter estimates, can be misleading if higher level interactions are present. The highest level interactions can be interpreted to some extent. However, this approach is not used here, as high level interactions are inherently complicated, and the trace plots should provide more reliable conclusions.

Additionally, interpreting mixture parameter estimates can be misleading due to the fixed sum constraint. The misleading nature of mixture parameter estimate interpretation is explained here by way of an example. Imagine that ingredients A and B both lead to an increase in the response but ingredient B has a much stronger effect. An increase in A may cause a decrease in B due to the fixed sum constraint. Because the effect of B is stronger than the effect of A this decrease in B will lead to an overall reduction in the response. Due to this overall reduction it will appear that increasing A decreases the response even though it is known that it actually increases the response.

In the majority of cases the effects of the explanatory variables on the response variables will be interpreted using the trace plots. In a few exceptions appropriate interpretations are given after the tables below.

7.9.2 Estimated Parameters for the Final Models

Table 7.10 Estimated parameters for the yield point (modified covariance structure)

Note that the individual parameter estimates can not be interpreted individually because of the interactions and the fixed sum constraint.

x_1	17.449	z_1	-4.047	$x_2 z_1$	12.296	$x_2 x_3 z_1$	-168.695
x_2	-33.999	z_2	0.436	$x_2 z_3$	-17.309	$x_2 x_3 z_3$	320.140
x_3	-111.631	z_3	7.662	$x_3 z_1$	57.478		
$x_2 x_3$	452.002			$x_3 z_3$	-119.668		

ρ	0.700
bl_1	-0.607
ml_1	-0.700
ml_2	-0.084
psh	-0.158

It is not surprising that the interaction between water and clay (i.e. x_2x_3) is important. These are the two variables which had the most dramatic visual effect on the glazes produced. They are also the variables which are usually modified when a glaze with specific properties is required. The coefficient estimates for x_2 and x_3 are both negative. However it is widely accepted that increasing clay actually increases the viscosity. The counter intuitive coefficients are a consequence of the fixed sum constraint. Given that they actually have the opposite effect it is not surprising that they interact.

B Top Plateau

Table 7.11 Estimated parameters for the top plateau (modified covariance structure)

Note that the individual parameters can not be interpreted individually because of the interactions and the fixed sum constraint.

x_1	28.950	z_1	-7.700	$x_2 z_1$	24.093	$x_2x_3 z_3$	-282.829
x_2	-41.837	z_2	0.531	$x_2 z_1$	-12.353	$x_2x_3 z_3$	203.350
x_3	-170.957	z_3	6.704	$x_3 z_3$	95.512		
x_2x_3	703.215			$x_3 z_3$	-87.643		

ρ	0.154
bl_1	-0.987
ml_1	-1.164
ml_2	-0.343
psh	-0.058

Again it is not surprising that the interaction between water and clay is significant.

C Bottom Plateau

Table 7.12 Estimated parameters for the bottom plateau (Toeplitz Covariance structure)

Note that the individual mixture parameter estimates can not be interpreted individually because of the fixed sum constraint. However, because there are no interactions, the process variable parameter estimates can be interpreted individually.

x_1	1.137		z_1	0.054		ρ	0.785
x_2	-11.623		z_2	0.188		bl_1	-0.469
x_3	18.566		z_3	0.413		psh	0.278

From Table 7.12 it appears that adding flocculant and milling to the next particle size both increase viscosity of the bottom plateau. This conclusion is consistent with what is widely accepted in the industry. The fact that the model contains so few terms indicates that the effects of the explanatory variables are far simpler for the bottom plateau than they are for the other responses.

7.9.3 Producing Trace Plots

Figures 7.13 to 7.15 show the trace plots for the yield point, the top plateau and the bottom plateau respectively. In each of the three figures, the trace plots are positioned according to the levels of the two process variables which they represent. A key for the positioning of the trace plots in given in Figure 7.12.

To produce the trace plots, pure mixture models were produced from the full models by setting the process variables to the appropriate levels. The nuisance parameters were not included in these models because their effects will not be present in general.

So that the trace plots could be compared with each other, the six trace plots for each response were plotted over the same range. Appropriate ranges were from 0, to 180 000Pa.s and from 0 to $28s^{-1}$, and from 0.04 to 0.33Pa.s for the top plateau and yield point and bottom plateau respectively. A routine was implemented on `matlab` to produce these plots from the

parameter vector returned by reml.m. This routine is included on the disk at the end of the thesis and is called jmtrc.m.

It must be noted that the bottom plateau data for run 4D was removed from the investigation. This point was removed because, being an unusual formulation, the model did not predict a high enough value for it. Recall also that trials 4E and 4F could not be analysed, because they were so thick that the sample was lost when it was being transferred between containers (see §6.4.2). Table 7.3 shows that for run 4, water was at its lower limit and clay was at its upper limit. The fact that the above three trials were not used for fitting the model, indicates that the model is not applicable for processing conditions D to F, when water and clay are near to their respective lower and upper bounds. For glazes of this type it is likely that the viscosity of the bottom plateau would increase at a faster rate than suggested by the appropriate model. Clearly the model should not be used to estimate viscosity of glazes of this type. To account for this limitation of the model, when trace plots of the bottom plateau are produced for process conditions D to F, the upper bound for clay is reduced to 0.09, and the lower bound for water is increased to 0.32. The range for frit is still consistent with these modified ranges.

Figure 7.12 Key to Figures 7.13 to 7.15, i.e. positions of the six trace plots with respect to the levels of the six projected levels of the two processing conditions they represent
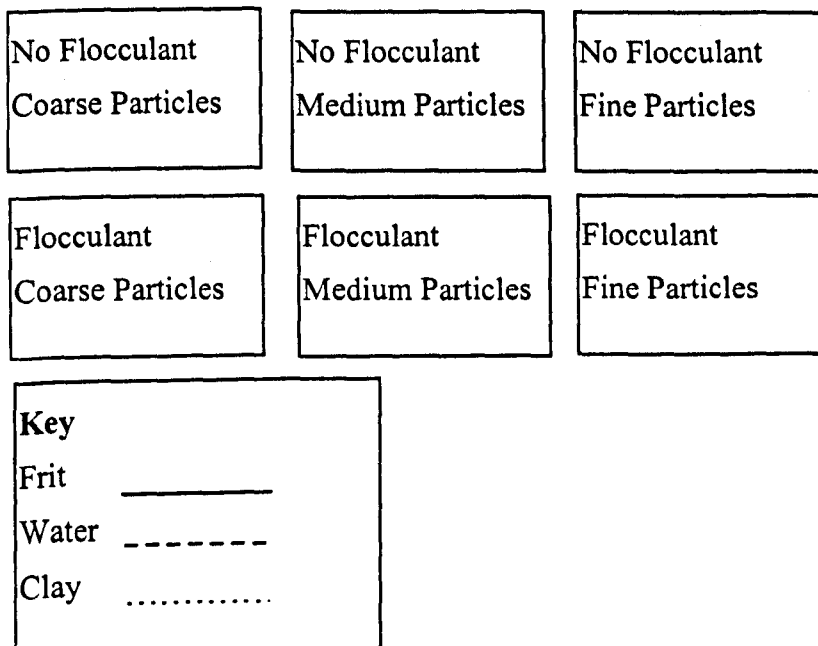
No Flocculant Coarse Particles	No Flocculant Medium Particles	No Flocculant Fine Particles
Flocculant Coarse Particles	Flocculant Medium Particles	Flocculant Fine Particles

Key
Frit _____
Water _ _ _ _ _ _ _
Clay ....

Figure 7.13 Trace plots for predicted top plateau viscosity (t.p) at the six levels of the process variables
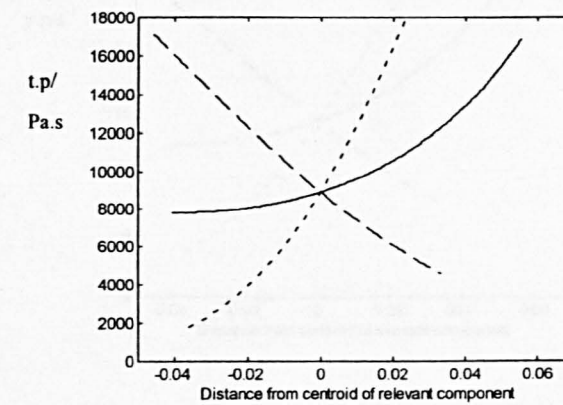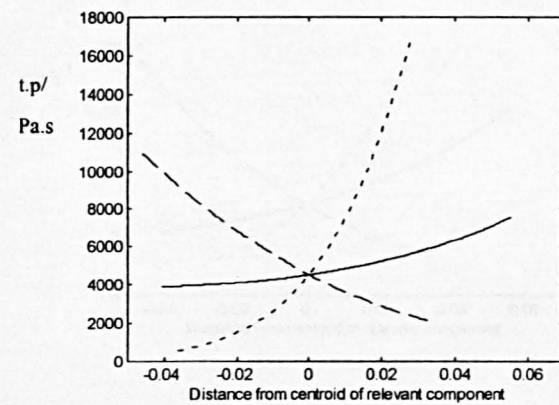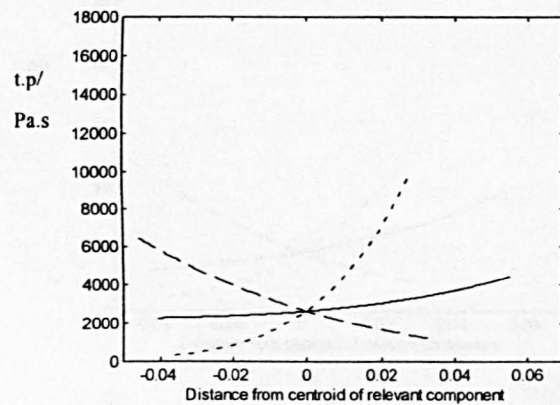
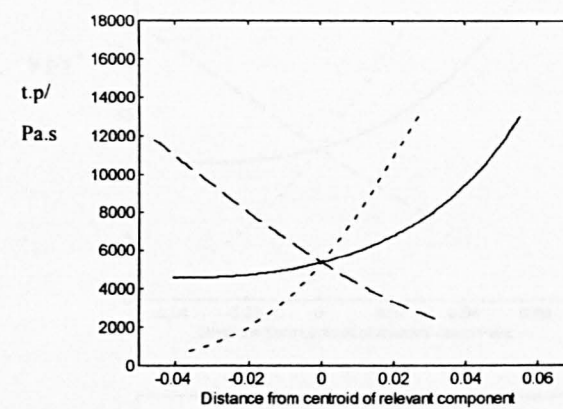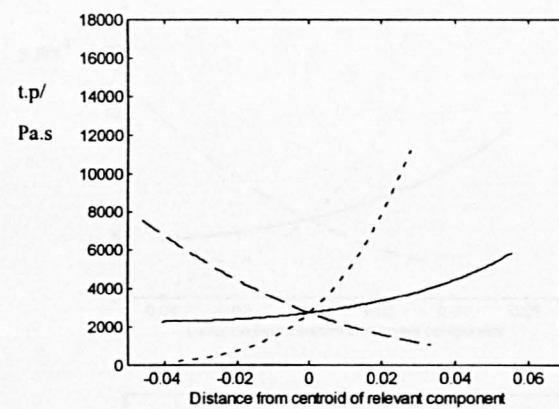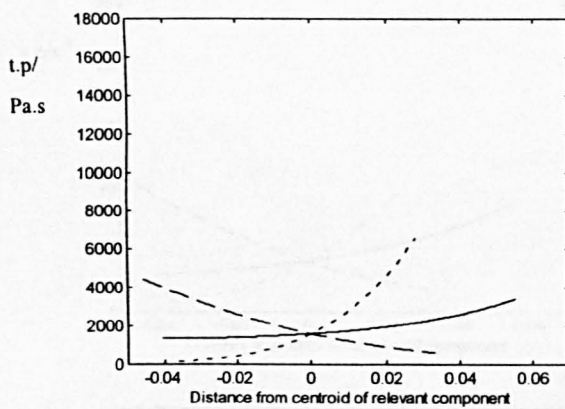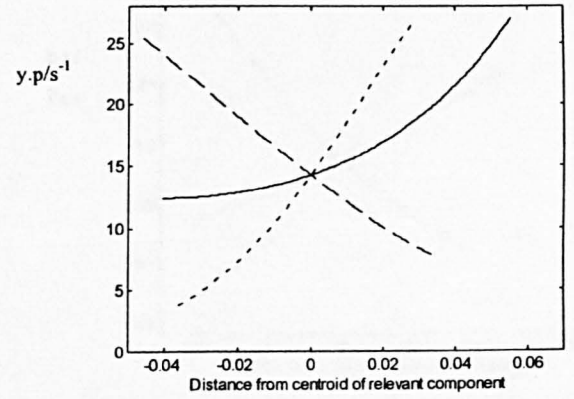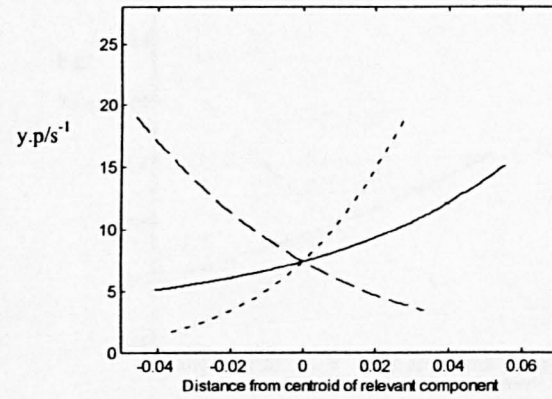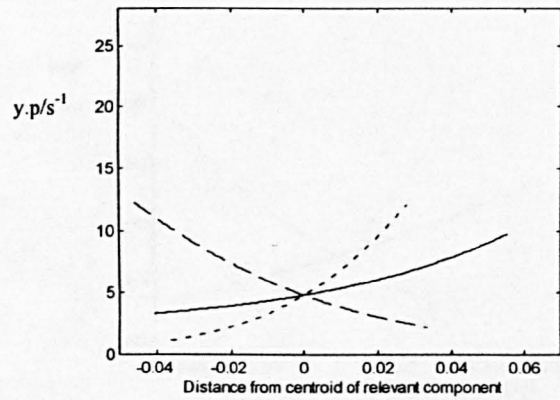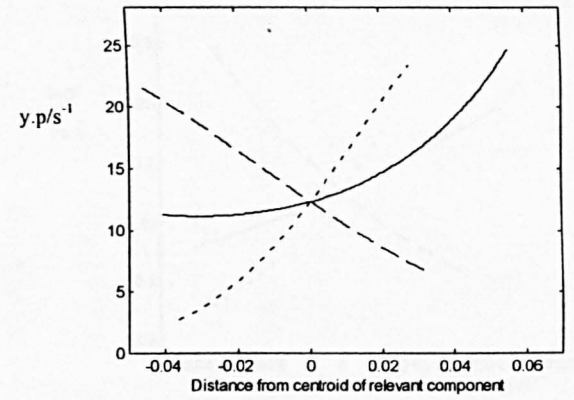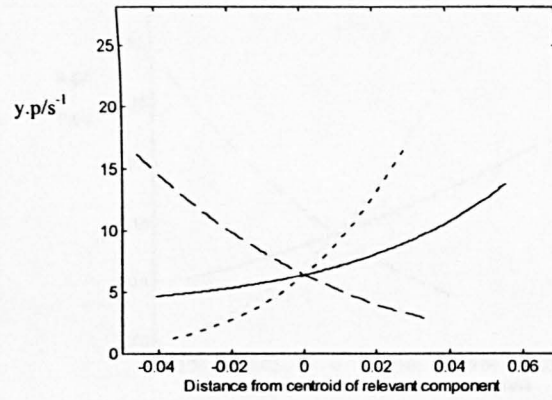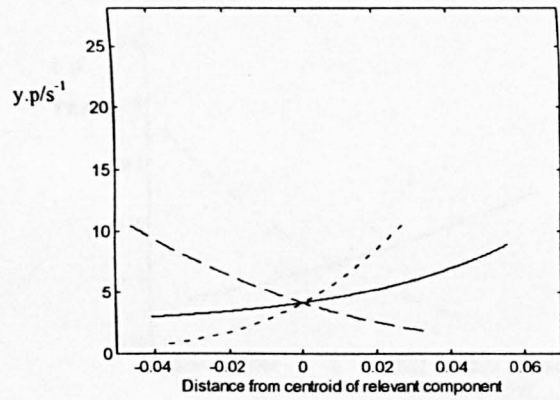Figure 7.15 Trace plots for the bottom plateau (b.p) at the six levels of the process variables
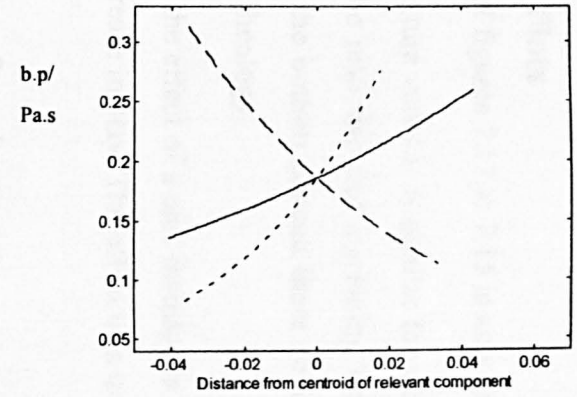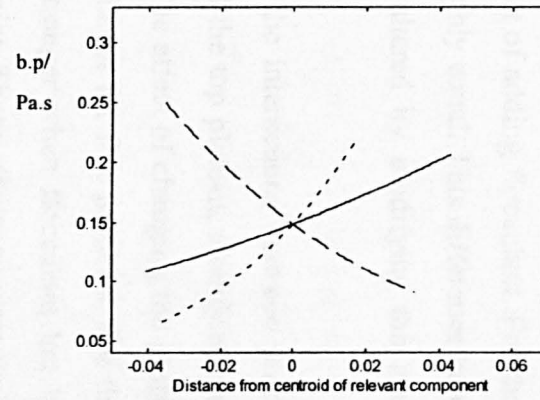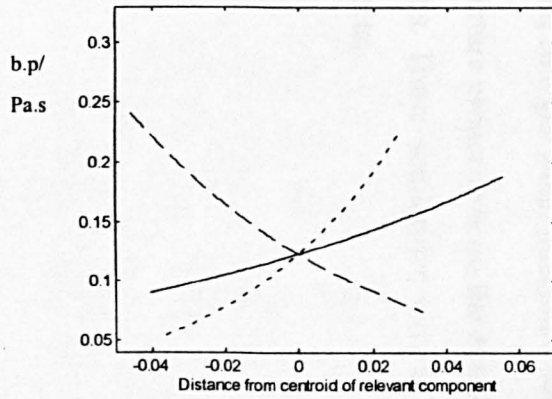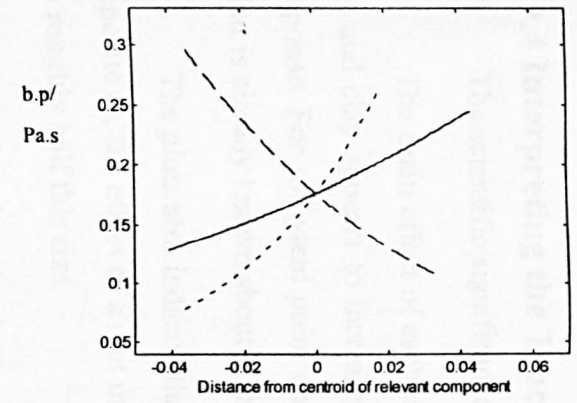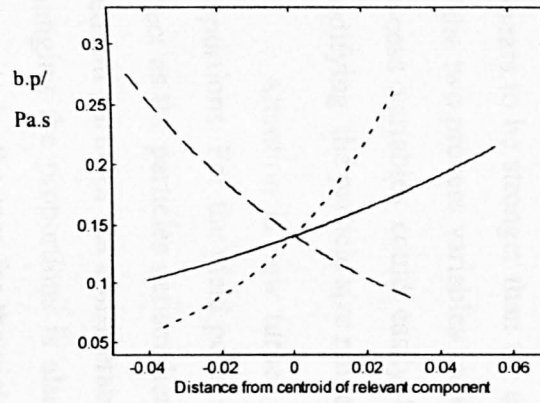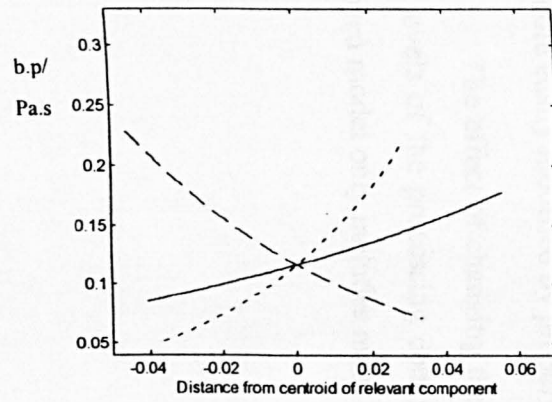
7.9.4 Interpreting the Trace Plots

The scientific significance of figures 7.13 to 7.15 is now discussed.

The main effect of each mixture variable is similar for all three responses. Increasing frit and clay appears to increase the response and increasing water appears to decrease the response. For the yield point and the bottom plateau these conclusions are consistent with what is already known about glaze rheology.

The plots also indicate that the effect of a unit increase in water is roughly equal, but opposite to, the effect of a unit increase in clay. The effect of a unit increase in frit appears to be roughly half this size.

It is also clear that both adding flocculant and decreasing particle size will increase all responses. For the top plateau and the yield point, the effect of milling to the next particle size appears to be stronger than the effect of adding flocculant. For the bottom plateau the effect of the two process variables are roughly equal. This difference between the effects of the two process variables could easily be altered by modifying the amount of flocculant and/or modifying the particle size ranges.

Attention is now turned to the interactions between process variables and mixture proportions. For the yield point and the top plateau, modifying the proportions has a bigger effect as the particles get smaller. The effect of changing the proportion is slightly larger for medium particles but considerably larger for fine particles. For the top plateau the effect of changing the proportions is also stronger when flocculant has been added. This does not appear to be the case for the yield point. These effects suggest the top plateau and yield point are more easily modified by proportional changes, for fine particles; and that the top plateau is more easily modified by proportional changes when flocculant has been added.

The effect of changing the mixture proportions on the bottom plateau is the same for all levels of the processing conditions. These similarities were to be expected because the chosen model only includes main effects.

Chapter 8

Conclusions and Suggestions for Future Work

8.1 Introduction

·   This thesis was concerned with *mixture experiments* A mixture experiment is one in which certain properties of a product depend on the relative proportions of its ingredients. Existing and new methods for generating optimal mixture designs were investigated, together with the form of the optimal design. One example of a mixture experiment is the production of ceramic glaze slip. An experiment to investigate how the proportions of each ingredient affect the viscosity of glaze slip was designed and analysed. Some models were then fitted to the results of this experiment.

8.2 The Form of the Optimal Design

Recall that an evenly replicated design includes each of the s support points either r or r+1 times. For the regular simplex, q=3, and all Scheffé polynomial models of degree up to three, all exact designs, which evenly replicated the continuous D optimal design support points, appeared to be D optimal. For the linear and quadratic models the support points are the vertices and the vertices plus midedges respectively. For the special cubic model the support points are the support points of the quadratic model plus the centroid. For the full cubic model the midedges are replaced by Mikaeili points.

The gosset algorithm was used to investigate V optimal designs over the regular simplex. For the linear model, D optimal designs were also found to be V optimal. Some interesting results were obtained for the linear model with n=4 and n=5. In many cases for the full simplex and a quadratic model, V optimal exact designs used the support points of the D optimal designs plus the centroid.

For the special cubic model, Laake (1973, 1975) calculated the V optimal weights over the (3, 3) simplex centroid. It was found here that in many cases, V optimal n point exact designs consisted of the points in the (3, 3) simplex centroid, replicated according to the rounded values of n times the weights. When some of the values of n times the weights deviated considerably from integers, the V optimal n point designs included near axial points. For the full cubic model the V optimal exact designs also included edge points and axial

points. KL was able to generate fairly V efficient designs despite being restricted to exact candidate points. The efficiency of these exact designs was found to decrease as the values of n times the Laake weights became increasingly different from integers.

An algorithm was implemented, to estimate the weights of the continuous D optimal design, over a given set of points which were in the optimal design for n=100 generated by KL. The algorithm was used to conduct a thorough investigation of the form of the optimal design for a quadratic model and three components. Less thorough investigations were carried out for the quadratic model and four components and for the cubic model and three components. In future it would it will be interesting to use the algorithm to conduct more thorough investigations for these two cases. It will also be of interest to investigate the special cubic for q=3, the special and full cubic for q=4 and all three models for q=5.

Using a less precise method, weights were estimated for a range of V optimal continuous designs for q=3 and the quadratic model. The same support points as used in the D optimal designs were again used here though the centroid tended to have an unusually high weight. In future work it would be especially interesting to conduct as thorough an investigation of the form of the V optimal continuous design as that suggested for D optimality. Even less is known about V optimal designs than is known about D optimal designs.

It was conjectured that in a similar way to the regular simplex, optimal exact designs for constrained regions will tend to use the same points as the optimal continuous designs. It is also conjectured that the efficiency of exact designs for constrained regions will be higher if n and s are such that the proportion of design points, at each support point, can be close to the weights in the continuous optimal design. In future work it will be interesting to investigate whether these conjectures are indeed correct for constrained problems. It will also be interesting to investigate the relationship between the efficiency of exact designs, and how much the values of n times the weights differ from integers. Whether this relationship varies for different constrained regions is also of interest.

Recall that in the majority of cases attention was restricted to designs which consisted of commonly used candidate points. It is known that the efficiency of designs can be increased if some of these candidate points are perturbated slightly. In future work it would be interesting to investigate by how much designs of the type generated in this thesis can be improved by perturbations.

8.3 Designs with Repeated Trials

Because many experimental processes are subject to random effects, it is often suggested that some of the trials should be replicated. Algorithms may be required to select the best points for replication depending on the number of design points, the number of replicates required and the weights of the support points in the continuous optimal design. A variety of appropriate algorithms were investigated for the simplified case where a set number of points are to be replicated twice. Here, the points included twice are referred as doubletons and the points included once are referred to as singletons. For $q > 3$, the best results were obtained using an exchange algorithm, which only considered designs with r doubletons which were distinct from each other. In this exchange algorithm there are three types of exchanges; those where a doubleton is exchanged with a singleton, those where a doubleton is exchanged for a point from the candidate set, and those where a singleton is exchanged for a point from the candidate set. For this algorithm it is suggested that all three types of exchange should be performed simultaneously, rather than sequentially, so that good exchanges which involve all three types of point can be identified.

An alternative method takes good designs with no constraints on the number of replicates and extends them into good designs with r points repeated twice. The starting design which is to be extended is referred to as a basis. It was found that the relationship between the optimality criterion values of bases and final designs was relatively strong but not monotonic. For the majority of examples the use of multiple good bases improved on the design which was generated from the optimal basis.

It is noted in the thesis that it may be better to replicate a smaller number of trials more times each. It was also recommended that the trials selected should be evenly spread around the feasible region. In future work it would be interesting to develop methods which strike a balance between having a good spread of points, having trials repeated more than twice and obtaining criterion efficiency. Future work of this type should consider constrained problems and both D and V optimality.

8.4 Experimental Work

A variety of glazes was produced and tested to investigate the relationship between glaze viscosity and glaze formulation. The effects of three mixture variables (frit, water and clay), and two processing conditions (particle size and whether flocculant had been added) were to be investigated. A number of problems were encountered with the experimental work

and the entire procedure had to be modified. An extensive programme of additional experimental work was then carried out. It was found that the use of a split plot structure greatly increased the number of samples that could be analysed in a given time but that the data had to be analysed using numerical ML estimation, under the assumption of correlated errors. REML is a modified form of ML which in general produces estimators of the variance parameters which are less biased than those produced by ML. For the bottom plateau there was evidence to suggest that REML produced a better final model in terms of predicting future responses.

The data produced included many points which were not representative of the true relationship between viscosity and shear stress. Many of these points were by-products of the measurement procedure. An ad hoc iterative procedure was used to remove atypical points, whereby confidence intervals were constructed for acceptable points and only points within these intervals were added to the data set.

After detailed discussion with JM three important response variables were identified and models to explain how they were affected by the explanatory variables were fitted. To fit sensible models for the three viscosity parameters the response data needed to be log transformed. The models produced suggested that the relationship between the explanatory variables and the viscosity of the bottom plateau was very simple, no interactions were necessary. The relationship for the other two responses was slightly more complicated but still parsimonious. It was found that increasing the amount of clay and frit in the mixture increased all three responses, while increasing water decreased them. This conclusion about the effect of frit was particularly interesting because it was not known before. The results also indicated that the effect of a unit increase in water, is roughly equal, but opposite to, the effect of a unit increase in clay. The effect of a unit increase in frit appears to be roughly half this size.

Decreasing the particle size and adding flocculant was also found to increase the viscosity of the bottom plateau. For the top plateau and the yield point the difference between fine particles and medium particles was far greater than the difference between medium particle and coarse particles. For the top plateau and yield point it was found that as particle size was reduced and flocculant was added the effects of the mixture variables increased in magnitude.

In terms of producing glazes for specific application procedures the above analysis suggests that the bottom plateau is less susceptible to proportional changes. It is therefore suggested that when designing a glaze, an appropriate bottom plateau should first be obtained by modifying the process variables. Once this has been achieved, changing the proportions should quickly lead to an appropriate yield point, with little change to the bottom plateau.

In future work the use of a split plot structure would mean that a similar procedure to that described above could be used for a large variety of different glazes. JM has been given information on how to design the appropriate experiments and how to analyse the results using numerical ML estimation. Details of this procedure, and other recommendations about how JM could use the methods discussed and the conclusions reached, were given in a report which was submitted to the company.

Appendix 1 Designs with r Replicates

A1.1 D optimal design, 3 compenent hexagon with b=0.83, 3 points replicated twice. (see §4.5.3)

D value 0.2831

Singletons

0.8300	0.1700	0.0000	vertex
0.8300	0.0000	0.1700	vertex
0.5000	0.0000	0.5000	long edge midpoint

Replicates

0.0000	0.8300	0.1700	vertex
0.1700	0.0000	0.8300	vertex
0.0000	0.3900	0.6100	third of long edge
0.3900	0.6100	0.0000	third of long edge

A1.2 D optimal design, 5 component irregular shape, 4 points replicated twice
Snee(1979) design with single component constraints and sum x_i =0.997 (see §6.4.3)

D value 6.4162e-046

Singletons

vertices

2	0.5000	0.1500	0.0500	0.1500	0.1500
3	0.7000	0.1500	0.0500	0.1000	0.0000
6	0.5000	0.1500	0.1500	0.2000	0.0000
9	0.5000	0.1500	0.1000	0.1000	0.1500
10	0.7000	0.0500	0.0500	0.2000	0.0000
17	0.5500	0.1500	0.0500	0.2500	0.0000
18	0.5000	0.1500	0.1000	0.2500	0.0000
19	0.5000	0.1000	0.1500	0.1000	0.1500

20 0.5000   0.0500   0.1500   0.1500   0.1500

38 0.5500   0.1500   0.1500   0.1000   0.0500 midpoint of edge
89 0.7000   0.0833   0.0500   0.1000   0.0667 third of edge

164 0.5950   0.1050   0.1050   0.1950   0.0000 face centroid
167 0.6000   0.0500   0.1000   0.1750   0.0750 face centroid

Replicates

(all vertices)

1   0.6500   0.0500   0.0500   0.1000   0.1500
5   0.5500   0.0500   0.1500   0.2500   0.0000
13 0.7000   0.0500   0.1500   0.1000   0.0000
14 0.5000   0.0500   0.0500   0.2500   0.1500

A1.3 V optimal design, 3 compenent JM problem, 3 points replicated twice.
(see §4.7.5) D value 0.3466

NB KLR did not provide the best design.

Singletons
1   0.5700   0.4000   0.0300 vertex
2   0.6650   0.3050   0.0300 vertex
4   0.5100   0.3850   0.1050 vertex
22 0.5300   0.3950   0.0750 triple blend

Replicates
6   0.6175   0.3525   0.0300 midpoint of edge
8   0.5900   0.3425   0.0675 midpoint of edge created by multi component constraint
14 0.5300   0.4000   0.0700 third of edge

A1.4 V optimal design, 5 component irregular shape, 5 points replicated twice (see 4.8.3) Snee(1979) design with single and multi component constraints and resacled so sum to 1

V value 0.3134

Singletons

vertices

1	0.5015	0.1474	0.0903	0.2508	0.0100
2	0.7021	0.0502	0.0502	0.1003	0.0972
6	0.7021	0.0502	0.0502	0.1003	0.0972
8	0.7021	0.0672	0.0502	0.1805	0.0000
9	0.5015	0.1474	0.1505	0.1003	0.1003
17	0.6488	0.0502	0.0502	0.2508	0.0000
23	0.5987	0.0502	0.0903	0.1103	0.1505
27	0.6188	0.0502	0.1505	0.1003	0.0802
34	0.5987	0.0502	0.1505	0.2006	0.0000

mid points of edges

45	0.5686	0.1505	0.1003	0.1003	0.0802
48	0.5837	0.1505	0.0502	0.2156	0.0000
75	0.5100	0.1505	0.1505	0.1890	0.0000
84	0.5385	0.1505	0.0502	0.1805	0.0802
101	0.5015	0.1505	0.0922	0.1053	0.1505
108	0.5887	0.1003	0.0502	0.1103	0.1505

115	0.5902	0.1011	0.0978	0.2109	0.0000	face centroid

Replicates

face centroids

118	0.6383	0.0502	0.0887	0.1615	0.0612
120	0.6282	0.0956	0.0502	0.1631	0.0628
122	0.5902	0.1011	0.0978	0.1003	0.1106
125	0.6321	0.0937	0.0937	0.1404	0.0401
127	0.5501	0.0988	0.1137	0.1689	0.0685

Appendix 2 Designs for Experimental Trials

A2.1 Preliminary trials, Stage 2 Modified Constrained Region (see §5.4.3)

V value 0.3242

Singletons

5	0.6650	0.3050	0.0300	1.0000	0.0000	0.0000	vertex
12	0.5300	0.4000	0.0700	0.1290	0.6129	0.2581	edge third
3	0.5100	0.4000	0.0900	0.0000	0.6129	0.3871	vertex
2	0.5150	0.3800	0.1050	0.0323	0.4839	0.4839	vertex
1	0.5700	0.4000	0.0300	0.3871	0.6129	0.0000	vertex

Replicates

	0.5858	0.3484	0.0658	0.4890	0.2800	0.2310	previous design
21	0.5833	0.3617	0.0550	0.4731	0.3656	0.1613	triple blend
7	0.6175	0.3525	0.0300	0.6935	0.3065	0.0000	mid-edge

A2.2 Main JM trials (see §6.3B)

Main Trials

4	0.5200	0.3800	0.1000	0.0000	0.5185	0.4815	vertex
2	0.6550	0.3100	0.0350	1.0000	0.0000	0.0000	vertex
8	0.6225	0.3100	0.0675	0.7593	0.0000	0.2407	mid-edge
5	0.5900	0.3100	0.1000	0.5185	0.0000	0.4815	vertex

Replicates

6	0.6150	0.3500	0.0350	0.7037	0.2963	0.0000	mid-edge
13	0.5567	0.3900	0.0533	0.2716	0.5926	0.1358	edge-third
31	0.5720	0.3560	0.0720	0.3852	0.3407	0.2741	centroid

Additional 3 trials

1	0.5750	0.3900	0.0350	0.4074	0.5926	0.0000	vertex
16	0.6117	0.3100	0.0783	0.6790	0.0000	0.3210	edge-third
20	0.5667	0.3333	0.1000	0.3457	0.1728	0.4815	edge-third

Appendix 3 Values of the Three Responses from the Main Trials

(see §7.2.7)

Run	Mixture Batch	Process Variable	Top Platea (pa.s)	Yield Stress (Pa)	Top Plateau (Pa.s)
1	1	A	2162	4.91	0.106
2	1	B	3513	4.41	0.094
3	1	C	5280	9.36	0.162
4	1	D	3292	7.55	0.151
5	1	E	5354	15.70	0.180
6	1	F	9597	15.70	0.168
7	2	A	84	1.05	0.059
8	2	B	267	1.64	0.074
9	2	C	646	3.38	0.102
10	2	D	274	1.97	0.090
11	2	E	1214	6.96	0.119
12	2	F	2127	6.36	0.119
13	4	A	4472	12.93	0.318
14	4	B	4977	16.04	0.278
15	4	C	10252	19.90	0.968
16	4	D	13995	22.16	4.020
17	51	A	193	0.98	0.066
18	51	B	n/a	n/a	0.093
19	51	C	245	1.35	0.066
20	51	D	305	1.50	0.069
21	51	E	616	2.87	0.089
22	51	F	1081	3.19	0.097
23	52	A	373	1.67	0.113
24	52	B	800	2.07	0.111
25	52	C	655	2.57	0.129
26	52	D	592	2.31	0.122
27	52	E	2461	6.78	0.151
28	52	F	2227	7.55	0.139

Run	Mixture Batch	Process Variable	Top Platea (pa.s)	Yield Stress (Pa)	Top Plateau (Pa.s)
29	61	A	64	0.51	0.037
30	61	B	179	0.76	0.044
31	61	C	132	0.67	0.044
32	61	D	259	1.15	0.050
33	61	E	320	1.72	0.045
34	61	F	478	1.97	0.051
35	62	A	28	0.30	0.027
36	62	B	141	0.51	0.037
37	62	C	112	0.76	0.044
38	62	D	323	1.15	0.054
39	62	E	309	1.50	0.052
40	62	F	542	2.25	0.055
41	71	A	2219	4.41	0.132
42	71	B	2957	4.41	0.102
43	71	C	2455	6.09	0.128
44	71	D	4034	6.78	0.108
45	71	E	5702	11.61	0.200
46	71	F	7338	12.93	0.194
47	72	A	550	1.67	0.090
48	72	B	1974	3.55	0.095
49	72	C	1224	2.87	0.081
50	72	D	1840	3.96	0.064
51	72	E	3086	7.55	0.125
52	72	F	4808	7.55	0.110
53	8	C	44	0.64	0.036
54	8	D	363	0.98	0.041
55	8	E	245	1.21	0.041
56	8	F	1006	1.67	0.055

Run	Mixture Batch	Process Variable	Top Platea (pa.s)	Yield Stress (Pa)	Top Plateau (Pa.s)
57	9	A	4053	7.55	0.229
58	9	B	4686	12.93	0.329
59	9	C	4703	9.36	0.302
60	9	D	12957	14.40	0.342
61	9	E	12846	24.67	0.541
62	9	F	23591	27.48	0.668
63	10	A	5518	10.43	0.192
64	10	B	12494	11.61	0.296
65	10	C	10655	12.93	0.237
66	10	D	12983	14.40	0.267
67	10	E	7996	17.86	0.330
68	10	F	15200	16.04	0.364

References

Atkinson, A.C. (1996). The usefulness of optimum experimental designs. J. Roy. Statist. Soc. B 57,59-76 (discussion 95-111)

Atkinson, A.C. and Donev, A.N. (1992). Optimum experimental designs. Clarendon: Oxford

Becker, N.G. (1978). Models and designs for experiments with mixtures. Ausytral. J. Stat,. 20, 195-208.

Bohlin Instruments Limited, Unit 6, The Corinium Centre, Gloucestershire GL7 1YJ.

Caufin, B., de Pretis, A. and Papo, A. (1992). Rheological Models for Ceramic Powder suspensions Sprechsaal. Vol 125, no. 10.

Chan, L. -Y. (1995). A review on optimal design for mixture models. In 'Five decades as a mathematician and educator - on the 80th birthday of Professor Wong', Eds. L. -Y. Chan & M. C. Liu, 45-88. World Scientific: Singapore.

Cornell, J. A. (1990) Experiments with mixtures. (Wiley).

Crosier, R. B. (1984). Mixture experiments: geometry and pseudocomponents. Technometrics 26, 209-16.

Draper, N. R. and Smith, H. (1998). Applied regression analysis, 3rd edition. Wiley.

Dekker (1990) Design and analysis of experiments with applications to engineering and physical sciences, Eds. Ghosh, S.

Elfving, G. (1952). Optimum allocation in linear regression theory, Ann. Math.. Statist., 23, 255-262.

Farrel, R. H. Kiefer, J. and Walbran, A. (1967). Optimum Multivariate Designs. Proceedings of the 5th Berkeley symposium, January 1967.

Gaffke, N. (1987). On D-optimality of exact linear regression designs with minimum support. J. Statist. Plann. Inference 15, 189-204.

Hardin and Sloane (1993) A new approach to the construction of Optimal Designs, Journal of Statistical Planning and Inference Vol 37 pp339-369.

Hardin and Sloane (1994) gosset Operating manual. Mathematical Sciences research Centre, AT&T Bell Laboratories, Murray Hill, New Jersey 07974 USA.

Herritage, A. (1994). A Basic Introduction to Rheology, Bohlin Instruments Ltd, Love Lane, Cirencester, Glos, UK.

Kasatkin, O.G. (1974, see Chan, 1995) On the construction of D optimal design on a simplex (in Russian), Application of Mathematical Methods for Multi-component systems investigation, Metallurgia, Moscow, 43-51.

Kennard, R.W. and Stone, L (1969). Computer aided design of experiments, Technometric, 11, 137-148.

Kiefer, J. (1961). Optimum designs in regression problems, II. Ann. Math. Statist. 32, 298-325.

Kurotschka, V.D. (1981). A general approach to the optimum design of experiments with qualitative and quantitative factors. In Proc. Indian Statistical Institute Golden Jubilee Int.

Conf. on Statistics: Applications and new directions (ed. J. K. Ghosh and J. Roy), pp. 353-68. Indian Statistical Institute, Calcutta.

Laake, P. (1973) Some optimal properties for experiments with mixtures (In Norwegian), Institute of Mathematics University of Oslo.

Laake, P. (1975) On the optimal allocation of observations in experiments with mixtures, Scand, J. Statist.,. 2, 153-157.

Liu, S. and Neudecker, H. (1995). A V Optimal Design for Scheffé's Polynomial model. Statistics and Probability Letters 23 (1995) 253-258.

Maple V. Computation Centre, University of Texas at Austin USA.

Martin, R.J., Bursnall, M.C., Stillman, E.C. Efficient Designs for Constrained Mixture Experiments. Statistics and Computing 9, 229-237.

Maths Works (1992). Matlab version 3.5. The Maths Works Inc.: Natick, MA.

Mikaeili, F. (1993). D-optimum design for full cubic on the q-simplex. J. Statist. Plann. Inference 35, 121-30.

Minitab Inc. 3081 Enterprise Drive, State College, PA 16801-3008 USA

Neave, H.R. (1992) Statistics Tables. Routledge: London.

Piepel, G. F. (1988) Programs for generating extreme vertices and centroids of linearly constrained experimental regions. Journal of Quality Technology, Vol 20, pp 125-139

Platts, L. (1996) The chemical Durability of Lead Free Glass Enamel Flux Compositions. Submitted for degree of Ph.D University of Sheffield.

Snee, R. D. (1975). Experimental designs for quadratic models in constrained mixture spaces.Technometrics 17, 149-59.

Snee, R. D. (1979). Experimental designs for mixture systems with multicomponent constraints. Commun. Statist. -Theor. Meth. 8, 303-26.

Snee, R. D. and Marquardt, D. W. (1974). Extreme vertices designs for linear mixture models. Technometrics 16, 399-408.

Stat-Ease (1993/7). Design-Expert version 4.0, 5.0: User's guide. Stat-Ease Inc.: Minneapolis.

Taylor, J.R. and Bull, A.C (1986) Ceramics glaze technology. Pergamon Press

Uranisi, H. (1964). Optimum design for the special cubic regression on the q-simplex. Mathematical Reports, General Education Dept., Kyushu University, No. 1, 7-12.

Vuchkov, I. N. (1982). Sequentially generated designs. Biometric J. 24, 751-63.

Wheeler, R.E. (1972) Efficient Experimental Design. Presented at the annual meeting of the Am. Stat. Assn., Montreal, August 1972.

Wynn, H. P. (1970). The sequential generation of D optimum experimental designs. Ann. Math. Statist., 41, 1655-64