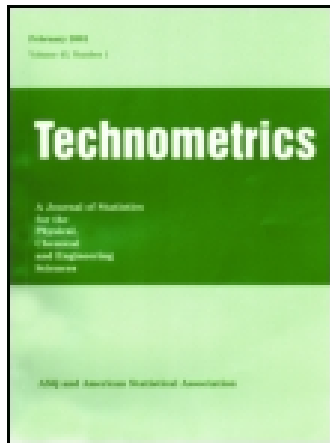


This article was downloaded by: [The University of Manchester Library]

On: 26 June 2015, At: 02:27

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Technometrics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/utch20>

### General Blending Models for Data From Mixture Experiments

L. Brown<sup>a</sup>, A. N. Donev<sup>a</sup> & A. C. Bissett<sup>b</sup>

<sup>a</sup> University of Manchester, Oxford Road, Manchester, M13 9PL, UK

<sup>b</sup> Federal-Mogul Friction Products Ltd, High Peak, Derbyshire, SK23 0JP, UK

Accepted author version posted online: 31 Jul 2014.



[Click for updates](#)

To cite this article: L. Brown, A. N. Donev & A. C. Bissett (2014): General Blending Models for Data From Mixture Experiments, *Technometrics*, DOI: [10.1080/00401706.2014.947003](https://doi.org/10.1080/00401706.2014.947003)

To link to this article: <http://dx.doi.org/10.1080/00401706.2014.947003>

Disclaimer: This is a version of an unedited manuscript that has been accepted for publication. As a service to authors and researchers we are providing this version of the accepted manuscript (AM). Copyediting, typesetting, and review of the resulting proof will be undertaken on this manuscript before final publication of the Version of Record (VoR). During production and pre-press, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal relate to this version also.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# General Blending Models for Data From Mixture Experiments

L. Brown<sup>1</sup>, A. N. Donev<sup>1</sup>, and A. C. Bissett<sup>2</sup>

<sup>1</sup>University of Manchester, Oxford Road, Manchester, M13 9PL, UK

<sup>2</sup>Federal-Mogul Friction Products Ltd, High Peak, Derbyshire, SK23 0JP, UK

## Abstract

We propose a new class of models providing a powerful unification and extension of existing statistical methodology for analysis of data obtained in mixture experiments. These models, which integrate models proposed by Scheffé (1958, 1963) and Becker (1968, 1978), extend considerably the range of mixture component effects that may be described. They become complex when the studied phenomena requires it, but remain simple whenever possible. This paper has supplementary material online.

KEYWORDS: Becker's models, Scheffé polynomials, model selection, nonlinear models

## 1 Introduction

We introduce a new class of statistical models for mixture experiments. In such experiments, the response depends on the proportions of the mixture components, but not on the amount of the mixture. For example, the strength of an alloy depends on the proportions of the metals of which it is comprised. Similarly, many features of friction materials (such as their friction coefficients and compressibilities) depend on the proportions of the chemicals from which they are made.

The common practice for analysing mixture experiments has evolved from the work of Scheffé (1958, 1963). Scheffé suggested the canonical polynomial models, which have provided the recourse for the majority of practitioners since, although alternatives have been proposed.

Quenouillé (1953, 1959, 1963) demonstrated that the Scheffé polynomials are incapable of describing common linear blending, a structure which he considers intuitively sensible. A component blends linearly when the effect of increasing its presence in the mixture, while keeping all other components in fixed relative proportions to each other, may be described by a linear relationship. For example, this is the effect of a diluent. In response, Becker (1968) suggested alternative models whose terms assume such effects. Where necessary, the veracity of this assumption may be judged by any practitioner who applies Becker's models, but this is equally true of the contrasting assumption in the Scheffé polynomials. Becker (1978) proposed related developments by introducing terms capable of describing a far broader range of effects than previously considered.

Prior to the work of Scheffé, ordinary polynomial models in mathematically independent variables (MIV) were applied to mixture experiments and such models endure, where they are deemed appropriate. Claringbold (1955), Draper and Lawrence (1965a,b), Thompson and Myers (1968), and Becker (1970) consider cases where the MIV are linear combinations of the component proportions, while Hackler et al. (1956) and Kenworthy (1963) take the MIV to be ratios of the component proportions. The utility of the Scheffé polynomials has been extended by Draper and John (1977a,b) and Chen et al. (1985), who propose the use of inverse terms and logarithmic terms, respectively, Gorman and Hinman (1962), who discuss a higher-order derivation, and Darroch and Waller (1985), Draper and Pukelsheim (1998), Cornell (2000) and Piepel et al. (2002), who each present useful reparameterisations. An overview of mixture experiment methodology is given by Cornell (2002).

The terms of the existing mixture experiments models do not allow sufficient flexibility to accommodate differences in the way the components affect the response. The joint effects, that is, those described by terms involving two or more components, are limited. For example, existing models have limited capability to represent rapid change in the response in certain areas of the experimental region. As a result, models may represent the response surface inaccurately or with a greater number of terms than necessary.

Models that are nonlinear in parameters have not been applied to data from mixture experiments, with the exception of those given by Focke et al. (2007) and Focke et al. (2012). The models proposed by these authors have application for only a small number of components. However, there are situations where more complex models would be preferable over models providing potentially crude polynomial approximations. The models proposed by Becker (1968, 1978) provide some of the required increased flexibility. However, these models assume linear blending, or alternatively inactivity, of one or more components.

The class of models we propose include many of the Scheffé (1958, 1963) and Becker (1968, 1978) models as special cases. However, this class of models also encompasses other ideas for modelling mixture experiments.

In Section 2 we summarise the main features of mixture experiments. We also discuss different effects that mixture components may have. In Section 3 we introduce a new general class of mixture models and discuss its relation to existing models. We focus on models with binary and ternary blending as they are useful in practice. In Section 4 we show that choosing the appropriate model for a specific study and its estimation can be combined, thus leading to a simple model selection procedure that can be implemented using many statistical packages. This is demonstrated with a simulated example, chosen to illustrate a situation when the new models provide excellent fit of the data, while the standard models do not. The dataset and computer code implementing the analysis are available as supplementary material on the journal's website. We conclude the paper with a discussion of the advantages and limitations of the new methodology.

## 2 Mixture experiments

In mixture experiments the response of interest,  $y$ , is dependent on the proportions of the  $q$  mixture components  $x_i, i = 1, \dots, q$ , such that

$$\sum_{i=1}^q x_i = 1, \quad x_i \geq 0. \quad (1)$$

The unconstrained composition space of the experiment is the  $(q-1)$ - dimensional simplex. However, individual component lower and upper bounds, linear multicomponent constraints, and non-linear constraints (Atkinson et al., 2007, p.230) often apply.

Scheffé (1958) proposes the use of  $\{q, m\}$  symmetric canonical polynomial models obtained by reparameterisation of standard polynomials of degree  $m$  for  $q$  components by using (1). The quadratic (S2), cubic (S3) and special cubic (SSC3) Scheffé polynomials for mixtures are

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i \neq j}^q \beta_{ij} x_i x_j, \quad (2)$$

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i < j}^q \beta_{ij} x_i x_j + \sum_{i < j}^q \gamma_{ij} x_i x_j (x_i - x_j) + \sum_{i < j < k}^q \beta_{ijk} x_i x_j x_k, \quad (3)$$

and

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i < j}^q \beta_{ij} x_i x_j + \sum_{i < j < k}^q \beta_{ijk} x_i x_j x_k, \quad (4)$$

respectively, where  $\beta_1, \beta_2, \dots, \gamma_{12}, \gamma_{13}, \dots$  are the parameters that must be estimated using the data.

As an alternative to the quadratic Scheffé polynomial, Piepel et al. (2002) suggest partial quadratic mixture (PQM) models, which are reduced forms of the model

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i < j}^q \beta_{ij} x_i x_j + \sum_{i=1}^q \beta_{ii} x_i^2, \quad (5)$$

where up to  $q(q-1)/2$  terms of binary joint effects  $x_i x_j$  ( $i \neq j$ ) or square terms  $x_i^2$  are included in the model in addition to the linear terms.

A full PQM model provides a fit equivalent to the full quadratic Scheffé polynomial. However, a reduced PQM model, containing squared terms, may prove more parsimonious than reduced quadratic Scheffé polynomials (e.g., when one or more components have strong quadratic curvature effects). This may equivalently be said to be the case for the models proposed by Draper and Pukelsheim (1998).

Becker (1968) introduces models that allow for describing linear blending:

$$H1 : E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i<j}^q \beta_{ij} \min(x_i, x_j) + \sum_{i<j<k}^q \beta_{ijk} \min(x_i, x_j, x_k) + \dots, \quad (6)$$

$$H2 : E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i<j}^q \beta_{ij} \frac{x_i x_j}{x_i + x_j} + \sum_{i<j<k}^q \beta_{ijk} \frac{x_i x_j x_k}{x_i + x_j + x_k} + \dots, \quad (7)$$

and

$$H3 : E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i<j}^q \beta_{ij} \sqrt{x_i x_j} + \sum_{i<j<k}^q \beta_{ijk} \sqrt[3]{x_i x_j x_k} + \dots \quad (8)$$

Reports of applications of these models include those of Becker (1968), Snee (1973), Johnson and Zabik (1981), Chen et al. (1996) and Cornell (2002), among others, most of whom demonstrate them to be advantageously used in comparison to Scheffé polynomials.

Becker (1978) progressed to propose the general model form,

$$E[y] = \sum_{i=1}^q \beta_i x_i + \sum_{i<j}^q h(x_i, x_j)(x_i + x_j) + \sum_{i<j<k}^q h(x_i, x_j, x_k)(x_i + x_j + x_k) + \dots \quad (9)$$

of which the H2 and H3 models are each a special case, where  $h()$  are each homogenous of degree zero, that is, their effect remains consistent for all values  $x_i + x_j$ , where  $x_i$  and  $x_j$  remain in fixed relative proportion, and similarly  $x_i + x_j + x_k$ , where  $x_i$ ,  $x_j$  and  $x_k$  remain in fixed relative proportions. Reduced forms of this model form allow the linear blending effect of one (or more) components on the response to be described. This is useful when one or more components has an additive effect on the response, such as a diluent. Becker suggested the nonlinear terms

$$h(x_j, \dots, x_k) = \prod_{i<j}^k \left( \frac{x_i}{x_j + \dots + x_k} \right)^{r_i} \quad (10)$$

which potentially provide greater flexibility in joint effects of the components. However, he gives little guidance on how these nonlinear parameters could be utilised, nor does he consider their estimation.

As described so far, the full statistical model chosen for a particular study may have too many terms describing joint effects, particularly if the number of components is large. Often only a subset of them will be needed. Stepwise regression (Efroymson, 1965) can be applied to achieve this.

### 3 Model generalisation

The linear (in the parameters) models discussed thus far will describe well situations where their terms accommodate the specific joint effects of the mixture components. However, their terms do not accommodate particular non-additive effects and this could lead them to perform poorly. They may not adequately represent the response or do so in a manner detrimental to model parsimony. This section proposes a general class of models which can represent responses of mixtures whose components have a wide range of different effects. We first discuss joint effects of two components and then extend the presented ideas to three components. The joint effects of more than three components are rarely considered when modelling mixture experiments using existing methodology and therefore are not considered here. We start by describing an idea for combining and generalising standard binary blending models.

#### 3.1 Motivation

The models

$$E[y] = \sum_{i=1}^q \beta_i x_i + \beta_{ij} x_i x_j \quad (11)$$

and

$$E[y] = \sum_{i=1}^q \beta_i x_i + \beta_{ij} \frac{x_i x_j}{x_i + x_j}, \quad (12)$$

where  $1 \leq i, j \leq q$ ,  $i \neq j$ , characterise the response surface in contrasting ways with respect to the joint effect of  $x_i$  and  $x_j$ . While (12) allows an additive blending effect through the linear blending of Becker's (1968) H2 model, (11) utilises the quadratic blending effect of the Scheffé polynomial. This contrast can be seen along any ray where  $x_i$  and  $x_j$  remain in a fixed relative proportion. Models (11) and (12) differ by the form of their last term. Where more than one pair of mixture components demonstrate joint effects, the best model fit may be achieved where the blending effect of the term in (11) is used for one pair of components and that of the term in (12) for another (Johnson and Zabik, 1981).

Firstly, a generalised binary blending effect is defined by introducing the parameter  $s_{ij}$  in the model

$$E[y] = \sum_{i=1}^q \beta_i x_i + \beta_{ij} \left( \frac{x_i}{x_i + x_j} \right) \left( \frac{x_j}{x_i + x_j} \right) (x_i + x_j)^{s_{ij}}. \quad (13)$$

The generalised binary blending term in (13) could be mathematically reduced to the form  $x_i x_j (x_i + x_j)^{s_{ij}^*}$ , but it is written that way to more easily see subsequently that the Scheffé and Becker H2 models are special cases.

The blending effects corresponding to five different values of  $s_{ij}$  ( $s_{ij} = 0.2, 0.5, 1, 2, 5$ ) are shown in Figure 1. Note that increasing  $s_{ij}$  above 1 results in a term whose effect is very small as  $x_i + x_j$  approaches zero, while reducing  $s_{ij}$  toward 0 results in a term whose impact decreases rapidly as  $x_i + x_j$  approaches zero.

Further flexibility can be added by introducing  $r_{ij}$  and  $r_{ji}$  to the model, which gives

$$E[y] = \sum_{i=1}^q \beta_i x_i + \beta_{ij} \left( \frac{x_i}{x_i + x_j} \right)^{r_{ij}} \left( \frac{x_j}{x_i + x_j} \right)^{r_{ji}} (x_i + x_j)^{s_{ij}}, \quad (14)$$



where, if  $s_{ij} = 1$ , this is a reduced form of (9) with only one term of joint effect. Model (14) is linear in the parameters  $\beta_{ij}$  for any values of the parameters  $s_{ij}$ ,  $r_{ij}$ ,  $r_{ji}$  that define the form of the terms.

This concept can be extended to introduce a general ternary joint effect in the model:

$$E[y] = \sum_{i=1}^q \beta_i x_i + \beta_{ijk} \left( \frac{x_i}{x_i + x_j + x_k} \right)^{r_{ijk}} \left( \frac{x_j}{x_i + x_j + x_k} \right)^{r_{jki}} \times \left( \frac{x_k}{x_i + x_j + x_k} \right)^{r_{kij}} (x_i + x_j + x_k)^{s_{ijk}}. \quad (15)$$

Here, the joint effect of the  $x_i$ ,  $x_j$  and  $x_k$  is governed by  $s_{ijk}$ ,  $r_{ijk}$ ,  $r_{jki}$ ,  $r_{kij}$  and the corresponding  $\beta_{ijk}$ . In particular,  $s_{ijk}$  governs the blending effect between  $x_i$ ,  $x_j$  and  $x_k$  and the remainder of the mixture, in an analogous manner to  $s_{ij}$  above. Thus, contrasting effects may be seen along any ray where  $x_i$ ,  $x_j$  and  $x_k$  remain in fixed relative proportions. The new terms for the binary and ternary cases are referred to as generalised terms of binary and ternary joint effects.

Model (14) may alternatively be expressed as

$$E[y] = \sum_{i=1}^q \beta_i x_i + \beta_{ij} \left( \frac{x_i}{x_i + x_j} \right)^{g_{ij} h_{ij}} \left( \frac{x_j}{x_i + x_j} \right)^{g_{ji} (1-h_{ij})} (x_i + x_j)^{s_{ij}}, \quad (16)$$

where  $g_{ij} h_{ij} = r_{ij}$  and  $g_{ji} (1 - h_{ij}) = r_{ji}$ , so that  $g_{ij} = r_{ij} + r_{ji}$  and  $h_{ij} = r_{ij} / g_{ij}$ . This allows us to better interpret the effects  $r_{ij}$  and  $r_{ji}$  through constrained values of  $h_{ij}$ ,  $g_{ij}$  and  $g_{ji}$ , where  $0 \leq h_{ij} \leq 1$ ,  $g_{ij} > 0$  and  $g_{ji} > 0$ .

The interpretation of  $g_{ij}$ ,  $g_{ji}$ , and  $h_{ij}$  may be understood, without loss of generality, along the edge where  $x_i + x_j = 1$ . First  $h_{ij}$  describes location, that is the point of greatest departure from linearity, with  $h_{ij} = 0.5$  indicating a symmetrical effect. Meanwhile,  $g_{ij}$  defines the localisation of the effect, where  $g_{ij} \gg 1$  indicates a joint effect contained to a region localised about the point of greatest departure from linearity. In contrast, where  $g_{ij} \ll 1$ , the effect of the term changes little around this point but instead causes rapid change as  $x_i / (x_i + x_j)$  approaches 0 and 1. The effect of the term is proportional to  $x_i^{g_{ij} h_{ij}} x_j^{g_{ij} (1-h_{ij})}$ . To illustrate the way the shape changes with  $x_i$  (or conversely with  $1 - x_j$ ), the effects for  $g_{ij} = 20$  and  $g_{ij} = 0.2$  are shown in Figure 2 for  $h_{ij} = 0.75$ .

A general binary term has no effect when  $x_i / (x_i + x_j)$  or  $x_j / (x_i + x_j)$  approaches zero. When  $h_{ij} = 1$  (or conversely,  $h_{ij} = 0$ ) the general binary blending term has greater significance as  $x_i / (x_i + x_j)$

approaches 1, and  $x_i$  dominates the effect. The manner in which this occurs is governed by  $g_{ij}$ . For values of  $g_{ij} > 1$  the term becomes increasingly influential at an increasingly rapid rate as  $x_i/(x_i + x_j)$  approaches 1. For values of  $g_{ij} < 1$  the term becomes increasingly influential at a decreasingly rapid rate. The impact of the term diminishes for larger values of  $g_{ij}$ . The parameter  $\beta_{ij}$  sets the magnitude of the effect given its specification by  $s_{ij}$ ,  $g_{ij}$  and  $h_{ij}$ .

Similar analysis may be made of the general term for a ternary joint effect. Model (15) may alternatively be expressed

$$E[y] = \sum_{i=1}^q \beta_{ijk} x_i x_j x_k \left( \frac{x_i}{x_i + x_j + x_k} \right)^{g_{ijk} h_{ijk}} \left( \frac{x_j}{x_i + x_j + x_k} \right)^{g_{ijk} h_{jki}} \times \left( \frac{x_k}{x_i + x_j + x_k} \right)^{g_{ijk} (1 - h_{ijk} - h_{jki})} (x_i + x_j + x_k)^{s_{ijk}},$$

where  $g_{ijk} = r_{ijk} + r_{jki} + r_{kij}$ ,  $h_{ijk} = r_{ijk}/g_{ijk}$ ,  $h_{jki} = r_{jki}/g_{ijk}$ ,  $g_{ijk} h_{ijk} = r_{ijk}$ ,  $g_{ijk} h_{jki} = r_{jki}$  and  $g_{ijk} (1 - h_{ijk} - h_{jki}) = r_{kij}$ . The new terms for the generalised ternary effects can be interpreted in a similar way as the generalised binary effects discussed earlier, without loss of the generality, as the effect of the term across the two-dimensional simplex where  $x_i + x_j + x_k = 1$ . The parameters  $h_{ijk}$  and  $h_{jki}$  describe the location of the point of greatest departure from linearity, with  $h_{ijk} = h_{jki} = 1/3$  indicating a rotationally symmetrical effect, while  $g_{ijk}$  once again describes the localisation of the effect about that point.

### 3.2 General Blending Mixture Models

We propose a class of Generalized Blending Mixture (GBM) models, for  $q$  components, of the form

$$\begin{aligned}
 E[y] &= \sum_{i=1}^q \beta_i x_i \\
 &+ \sum_{i<j} \beta_{ij} \left( \frac{x_i}{x_i + x_j} \right)^{r_{ij}} \left( \frac{x_j}{x_i + x_j} \right)^{r_{ji}} (x_i + x_j)^{s_{ij}} \\
 &+ \sum_{i<j<k} \beta_{ijk} \left( \frac{x_i}{x_i + x_j + x_k} \right)^{r_{ijk}} \left( \frac{x_j}{x_i + x_j + x_k} \right)^{r_{jki}} \times \\
 &\quad \left( \frac{x_k}{x_i + x_j + x_k} \right)^{r_{kij}} (x_i + x_j + x_k)^{s_{ijk}} .
 \end{aligned} \tag{17}$$

In the second and third sums, we have  $\binom{q}{2}$  and  $\binom{q}{3}$  terms, respectively. Although one could have multiple terms involving the same variables with different powers, we do not consider this.

As discussed earlier, the GBM models can also be reparametrised as

$$\begin{aligned}
 E[y] &= \sum_{i=1}^q \beta_i x_i \\
 &+ \sum_{i<j} \beta_{ij} \left( \frac{x_i}{x_i + x_j} \right)^{g_{ij}h_{ij}} \left( \frac{x_j}{x_i + x_j} \right)^{g_{ji}(1-h_{ij})} (x_i + x_j)^{s_{ij}} \\
 &+ \sum_{i<j<k} \beta_{ijk} \left( \frac{x_i}{x_i + x_j + x_k} \right)^{g_{ijk}h_{ijk}} \left( \frac{x_j}{x_i + x_j + x_k} \right)^{g_{jki}h_{jki}} \times \\
 &\quad \left( \frac{x_k}{x_i + x_j + x_k} \right)^{g_{kij}(1-h_{ijk}-h_{jki})} (x_i + x_j + x_k)^{s_{ijk}} .
 \end{aligned} \tag{18}$$

Models (17) and (18) may be used to establish a broad range of joint effects. In fact, many models presented in the literature are special cases of our class of models. For example, the quadratic crossproduct terms in the Scheffé polynomial or the PQM model, occur when  $h_{ij} = 0.5$ ,  $g_{ij} = 2$  and  $s_{ij} = 2$ . The squared terms in the PQM model occur when  $h_{ij} = 0$ ,  $g_{ij} = 2$  and  $s_{ij} = 2$ . The binary

blending terms of Becker's H2 and H3 models occur when  $h_{ij} = 0.5$ ,  $s_{ij} = 1$  and  $g_{ij} = 2$  or  $1$ , respectively. Furthermore the ternary term of the special cubic model occurs when  $h_{ijk} = h_{jim} = \frac{1}{3}$ ,  $g_{ijk} = 3$  and  $s_{ijk} = 3$ . Thus, the GBM model allows us to consider commonly used terms, as well as new terms, with considerable flexibility.

## 4 Model selection

Model (17) is complex, being a nonlinear function of some of its parameters. Its estimation is difficult but possible. However, in most cases it is unnecessary to estimate all its parameters simultaneously. When the parameters  $r_{ij}$ ,  $r_{ji}$ ,  $s_{ij}$ ,  $r_{ijk}$ ,  $r_{jik}$  and  $s_{ijk}$  are specified, the estimation of the remaining parameters of (17) becomes trivial as the resulting models are linear in the parameters. Therefore a sensible alternative to estimating (17) is to choose a model from a list of models that includes traditional models as well as new GBM models obtained for a grid of values for  $r_{ij}$ ,  $r_{ji}$ ,  $s_{ij}$ ,  $r_{ijk}$ ,  $r_{jik}$  and  $s_{ijk}$ . The model selection criterion

$$AIC_c = 2pn/(n - p + 1) - 2\log(L) \quad (19)$$

is used, where  $L$  is the likelihood function and  $p$  is the number of the parameters of the estimated model.

There are various ways of implementing such a comparison. The results presented here were obtained by a forward selection stepwise regression procedure, i.e. by fitting first the model including the effects of the individual mixture components and then sequentially adding the best possible term representing joint action of two or three components, defined by all possible values of  $r_{ij}$ ,  $r_{ji}$ ,  $s_{ij}$ ,  $r_{ijk}$ ,  $r_{jik}$  and  $s_{ijk}$  and judged by the  $AIC_c$  criterion. The fitting was terminated when the models became unnecessarily complicated. The computer implementation was done with the free software package R using the  $AIC_c$  function of the library *AICcmodavg* (Mazerolle, 2013). The computer program is provided as supplementary material.

Most published datasets obtained in mixture experiments are provided with satisfactory statistical analyses using standard models. Fitting GBM models to such data therefore usually brings modest benefits, which is not surprising. Furthermore, the experimental designs used in such studies often do not allow for fitting the GBM models, as they either have too few observations or their location in the design region does not allow the estimation of some of the model parameters. This is why in order to illustrate the features of the new models, data were simulated for a scenario where their advantages in comparison with standard models could be seen. Certainly, this is not a typical case and in most practical situations the differences seen in this example are likely to be considerably smaller.

**Example.** This example involves three mixture components with their proportions varying from 0 to 1. The response surface for this example was chosen to be asymmetric, but ordinary, see Figure 3. The maximum of the response was attained by a combination of a large proportion of  $x_1$  and similar but small proportions of the remaining components  $x_2$  and  $x_3$ . However the joint effect of the mixture components was strong. This was achieved by using the model

$$E[y] = 3x_1 + 4x_2 + 5x_3 + 20x_1x_2^3 + 80\frac{x_1^{2.5}x_2^{0.5}x_3^{0.5}}{x_1 + x_2 + x_3}. \quad (20)$$

The data were generated for the 22-trial, 3-component simplex lattice design with an additional centroid point shown in Figure 4. Independent and normally distributed random errors with homogeneous variance equal to  $0.25^2$  were added to the model-calculated values to yield the simulated data.

The best GBM model with four terms was chosen by comparing all possible models obtained by adding to the model having just three terms, i.e.

$$E[y] = \sum_{i=1}^3 \beta_i x_i,$$

a single term of joint action. There were four types of terms to consider adding:  $\beta_{ij} \left(\frac{x_i}{x_i+x_j}\right)^{r_{ij}} \left(\frac{x_j}{x_i+x_j}\right)^{r_{ji}} (x_i + x_j)^{s_{ij}}$ , for  $i = 1, j = 2$ ;  $i = 1, j = 3$ ; and  $i = 2, j = 3$ ; and  $\beta_{123} x_1^{r_{123}} x_2^{r_{231}} x_3^{r_{312}}$ , where the last term simplified from that in (17) as  $x_1 + x_2 + x_3 = 1$ . Each of these terms was considered for all possible combinations of the values 0.5, 1, 1.5, 2, 2.5 or 3 for  $r_{12}, r_{13}, r_{23}, r_{123}, r_{231}$  and  $r_{312}$  and the values 0, 1, 2 or 3 for  $s_{12}, s_{13}$  and  $s_{23}$ . Hence,  $3(4 * 6^2) + 6^3 = 648$  models were considered. The model with smallest  $AIC_c$  was chosen. It included a term representing binary blending for the components  $x_1$  and  $x_2$ .

The best GBM model with five terms was chosen by comparing all models obtained by adding one more term to the best GBM model with four terms. The list of models to compare was obtained in the same way as that used to obtain the best GBM model with four terms. At this stage a term representing ternary blending for the components was included.

The same approach was used again to obtain the best GBM model with six and then, with seven terms. The terms that were added represented binary blending for components  $x_1$  and  $x_3$  and for components  $x_2$  and  $x_3$ , respectively. As expected, the model with seven terms had the same structure as (17).

It may be beneficial to use different blending terms for the same components only in very rare situations. If this possibility is excluded in the model selection, it becomes faster as the number of models to consider becomes smaller as more terms are added to the model.

The full Scheffé cubic polynomial (3), Becker's models H2 (7) and H3 (8) were also fitted to the data. Their  $AIC_c$  statistics, as well as those for the best GBM models with three, four, five, six and seven terms, are given in Table 1.

The GBM model with five terms had a smaller  $AIC_c$  statistic than those for the GBM models with three, four, six or seven terms, and was overall the best model. This model is

(21)

where the figures in the brackets are the standard deviations of the estimates of the parameters that precede them. The  $AIC_c$  value for this model is 9.91 and compares rather favourably with the corresponding values for the fitted full Scheffé cubic polynomial and Becker's H2 and H3 models for which it is 95.85, 80.53 and 78.72, respectively. While we emphasise that such advantageous differences in favour of the GBM models are not typical, this example shows that for certain studied phenomena they can be achieved.

The contour plots of the predicted surfaces with the full Scheffé cubic polynomial and the selected GBM model are shown in Figures 5 and 6. It can be seen that the estimated response surfaces of the GBM model was very similar to the true response surface shown in Figure 3. This cannot be said for the estimated full Scheffé cubic polynomial as its estimated response surface is notably different from the true surface.

Using a grid with a larger number of possible values for  $r_{12}$ ,  $r_{13}$ ,  $r_{23}$ ,  $r_{123}$ ,  $r_{231}$ ,  $r_{312}$ ,  $s_{12}$ ,  $s_{13}$  and  $s_{23}$  was attempted but did not bring any benefits. It was felt that the reason for that was that the amount of simulated data was not sufficiently large to allow to distinguish between models with such small differences of the values of the nonlinear parameters. However, as different models were found, some with structures somewhat different to that of the true model, they all produced predictions which would be considered indistinguishable in a practical application and well representing the underlying relationship.

## 5 Discussion

The general class of models that we propose provide a powerful unification and extension of the existing statistical methodology for analysis of data obtained in mixture experiments. The complexity of the models fitted to the data will closely match the complexity of the studied phenomena: they will be models equivalent to those proposed by Scheffé (1958, 1963), Becker (1968) and (Piepel et al., 2002) when possible, but more complex when needed. The main benefits of using GBM

models are that they are parsimonious and can accurately describe response surfaces in situations where sometimes standard models will offer only a crude and possibly even misleading approximation.

Estimating simultaneously all parameters of the GBM models (17) would require a substantial computational effort. The authors have made considerable progress in developing a computational tool capable of doing this, though its discussion remains outside the scope of the presented research. However, the method of choosing a GBM model proposed in this paper is effective, simple and computationally stable, thus it is good for its purpose.

The development of the general class of mixture models naturally creates the need to re-evaluate the usefulness of the standard and computer generated experimental designs for mixture experiments. It is clear that fitting GBM models requires more data than fitting any of the standard models. It is possible to use space-filling experimental designs, collecting as much data as the available resources allow for. Such designs have been explored, for example, by Fang and Wang (1994); Borkowski and Piepel (2009) and Ning et al. (2011). Further work aiming to formulate a better experimental design strategy for estimating the class of general blending mixture models that takes into account their structure is underway.

## Supplementary material

Data and Code: The simulated dataset used in the example in Section 4 of the paper, along with R code to perform the analysis (zip folder).

## 6 Acknowledgement

We were fortunate to have our paper been refereed by an Editor, Associate Editor and the two referees who understand well the theory and practice of mixture experiments. Their extensive and constructive suggestions were invaluable to us in the revisions of this paper.



## References

- Atkinson, A. C., A. N. Donev, and R. Tobias (2007). *Optimum Experimental Designs, with SAS*. New York: Oxford University Press.
- Becker, N. G. (1968). Models for the response of a mixture. *Journal of the Royal Statistical Society, Series B* 30, 349–358.
- Becker, N. G. (1970). Mixture designs for a model linear in the proportions. *Biometrika* 57, 329–338.
- Becker, N. G. (1978). Models and designs for experiments with mixtures. *Australian Journal of Statistics* 20, 195–208.
- Borkowski, J. J. and G. F. Piepel (2009). Uniform designs for highly constrained mixtures experiments. *Journal of Quality Technology* 41, 35–47.
- Chen, J. J., L. A. Li, and C. D. Jackson (1996). Analysis of quantal response data from mixture experiments. *Environmetrics* 7, 503–512.
- Chen, W. S., W. Y. Zhu, and C. J. Hu (1985). A new mixture model with three or four components and their quasi D-optimality. Technical report, Northern Illinois University.
- Claringbold, P. J. (1955). Use of the simplex design in the study of joint action of related hormones. *Biometrics* 11, 174–185.
- Cornell, J. A. (2000). Fitting a slack-variable model to mixture data: Some questions raised. *Journal of Quality Technology* 32, 133–147.
- Cornell, J. A. (2002). *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data (3rd edition)*. New York: Wiley.

Darroch, J. N. and J. Waller (1985). Additivity and interaction in three-component experiments with mixtures. *Biometrika* 72, 153–163.

Draper, N. R. and R. C. S. John (1977a). Designs in three and four components for mixtures models with inverse terms. *Technometrics* 19, 117–130.

Draper, N. R. and R. C. S. John (1977b). A mixtures model with inverse terms. *Technometrics* 19, 37–46.

Draper, N. R. and W. Lawrence (1965a). Mixture designs for four factors. *Journal of the Royal Statistical Society, Series B* 27, 473–478.

Draper, N. R. and W. Lawrence (1965b). Mixture designs for three factors. *Journal of the Royal Statistical Society, Series B* 27, 450–465.

Draper, N. R. and F. Pukelsheim (1998). Mixture models based on homogeneous polynomials. *Journal of Statistical Planning and Inference* 71, 303–311.

Efroymson, M. A. (1965). Multiple regression analysis. In A. Ralston and H. S. Wilf (Eds.), *Mathematical Methods for Digital Computers*. New York: Wiley.

Fang, K.-T. and Y. Wang (1994). *Number Theoretic Methods in Statistics*. London: Chapman and Hall.

Focke, W. W., M. H. Ackermann, and R. L. J. Coetzer (2012). Mixing rules based on power means and generalized q-fractions. *Journal of Chemical Engineering of Japan* 45, 1–8.

Focke, W. W., C. Sandrock, and S. Kok (2007). Weighted-power-mean mixture model: Application to multicomponent liquid viscosity. *Industrial Engineering Chemistry Research* 46, 4660–4666.

Gorman, J. W. and J. E. Hinman (1962). Simplex lattice designs for multicomponent systems. *Technometrics* 4, 463–487.

Hackler, W. C., W. W. Kriegel, and R. J. Hader (1956). Effect of raw-material ratios on absorption of whiteware compositions. *Journal of the American Ceramic Society* 39, 20–25.

Johnson, T. M. and M. E. Zabik (1981). Response surface methodology for analysis of protein interactions in angel food cakes. *Journal of Food Science* 46, 1226–1230.

Kenworthy, O. O. (1963). Fractional experiments with mixtures using ratios. *Industrial Quality Control* 19, 24–26.

Mazerolle, M. J. (2013). *AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c)*. R package version 1.35.

Ning, J.-H., K.-T. Fang, and Y.-D. Zhou (2011). Uniform design for experiments with mixtures. *Communications in Statistics - Theory and Methods* 40, 1734–1742.

Piepel, G. F., J. M. Szychowski, and J. L. Loepky (2002). Augmenting scheffe linear mixture models with squared and/or crossproduct terms. *Journal of Quality Technology* 34, 297–314.

Quenouillé, M. H. (1953). *The Design and Analysis of Experiments*. New York: Hafner.

Quenouillé, M. H. (1959). Experiments with mixtures. *Journal of the Royal Statistical Society, Series B* 21, 201–202.

Quenouillé, M. H. (1963). Discussion of ‘the simplex-centroid design for experiments with mixtures’ by H. Scheffé. *Journal of the Royal Statistical Society, Series B* 25, 235–263.

Scheffé, H. (1958). Experiments with mixtures. *Journal of the Royal Statistical Society, Series B* 20, 344–360.

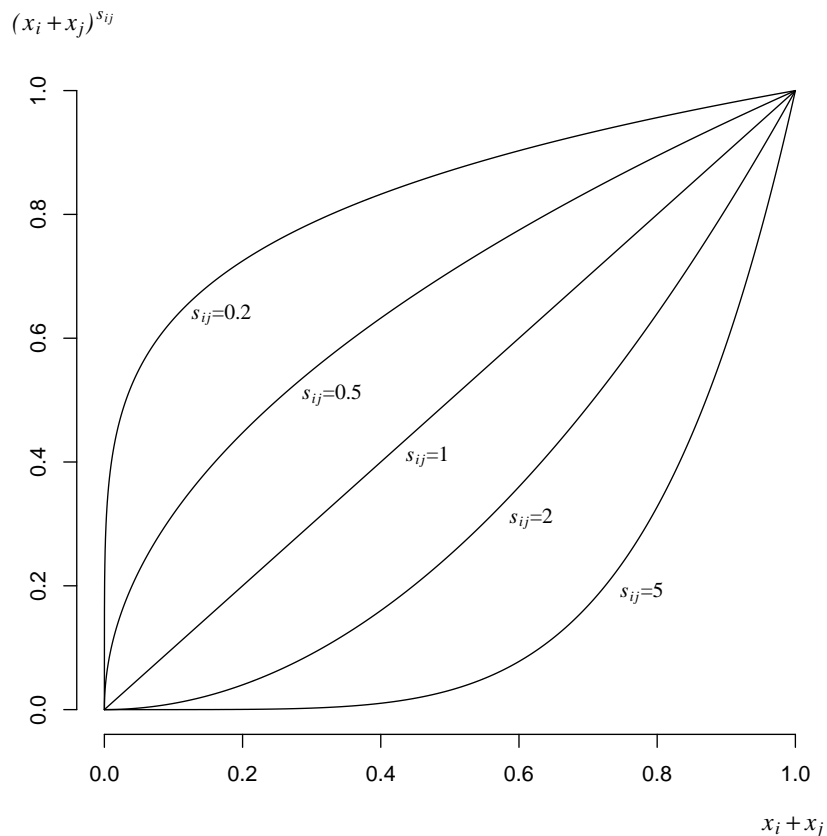
Scheffé, H. (1963). The simplex-centroid design for experiments with mixtures. *Journal of the Royal Statistical Society, Series B* 25, 235–263.

Snee, R. D. (1973). Techniques for the analysis of mixture data. *Technometrics* 15, 517–528.

# ACCEPTED MANUSCRIPT

Thompson, W. and R. Myers (1968). Response surface designs for experiments with mixtures.  
*Technometrics* 10, 739–756.

Model	Scheffé	H2	H3	GBM	GBM	GBM	GBM	GBM
Terms	10	7	7	3	4	5	6	7
$AIC_c$	95.85	80.53	78.72	80.60	45.53	9.91	10.49	11.81

Table 1:  $AIC_c$  statistics for models fitting the simulated data.Figure 1: Blending effects for  $s_{ij} = 0.2, 0.5, 1, 2$  and  $5$ .

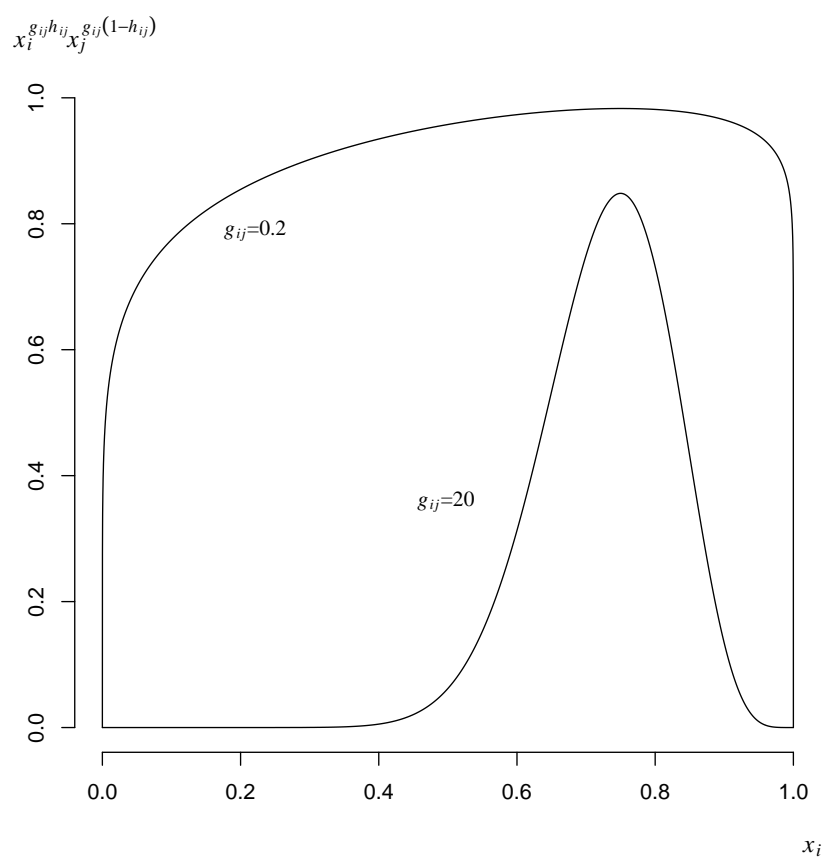


Figure 2: Binary blending effects for  $g_{ij} = 0.2$  and 20.

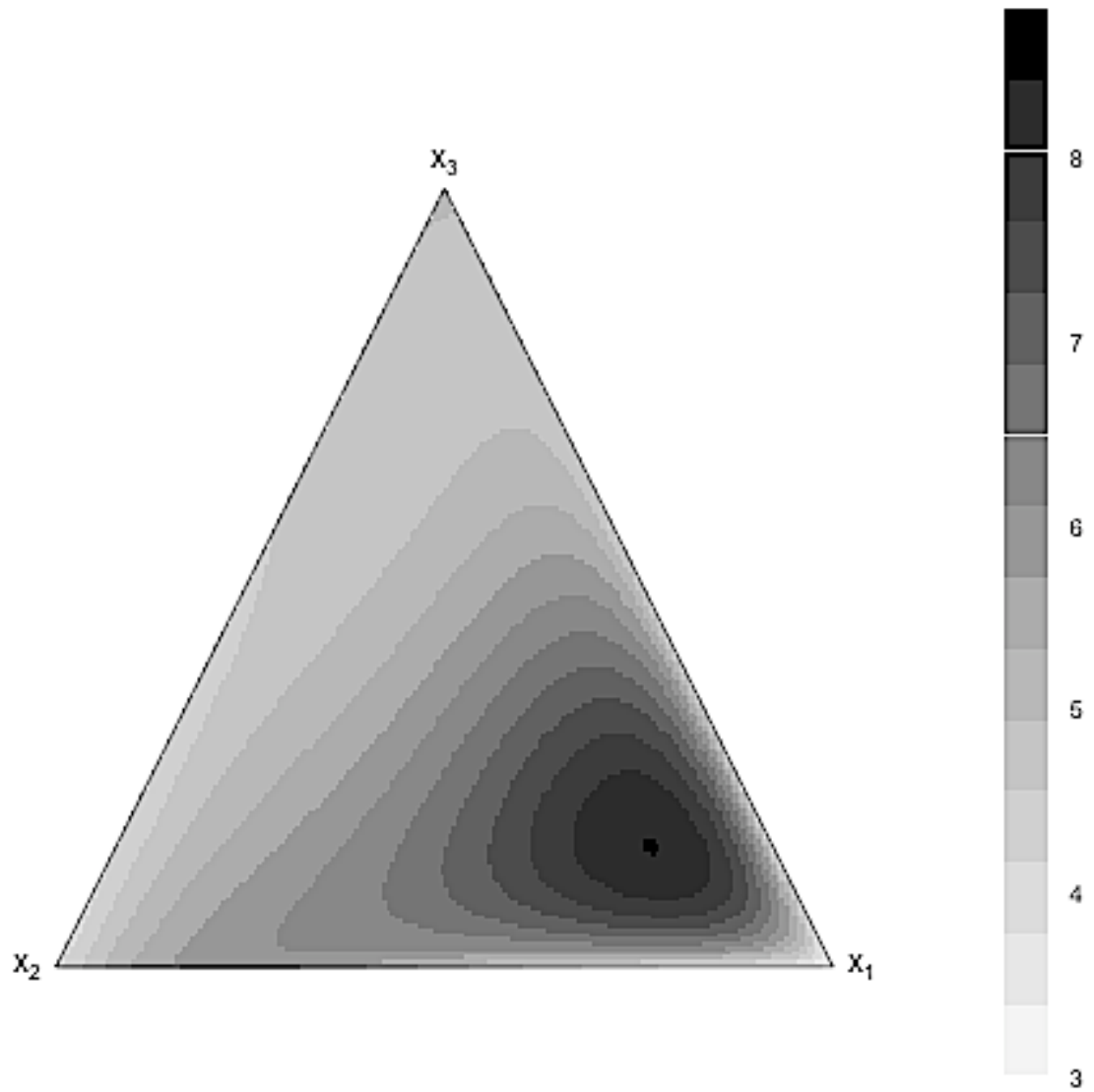


Figure 3: Contour plot for the underlying model.

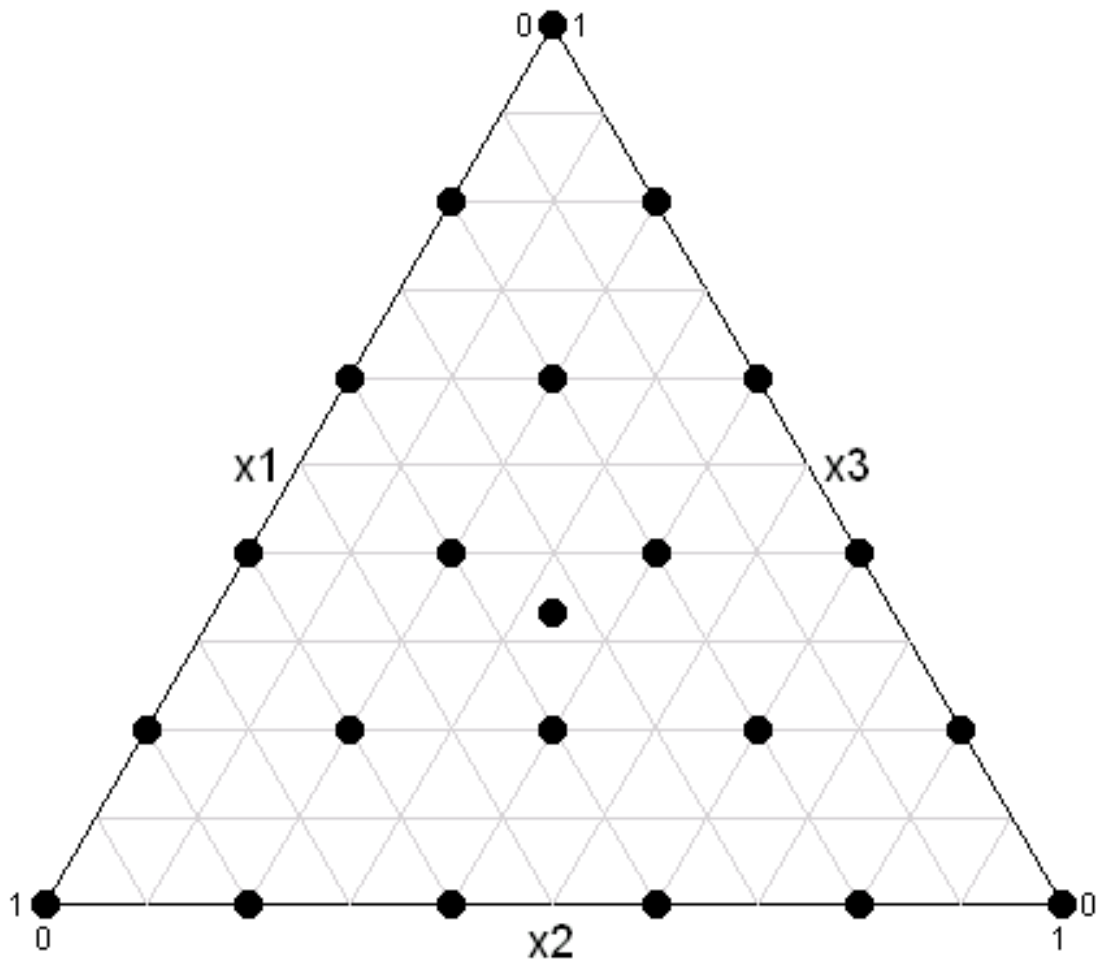


Figure 4: Plot of design for simulated data.



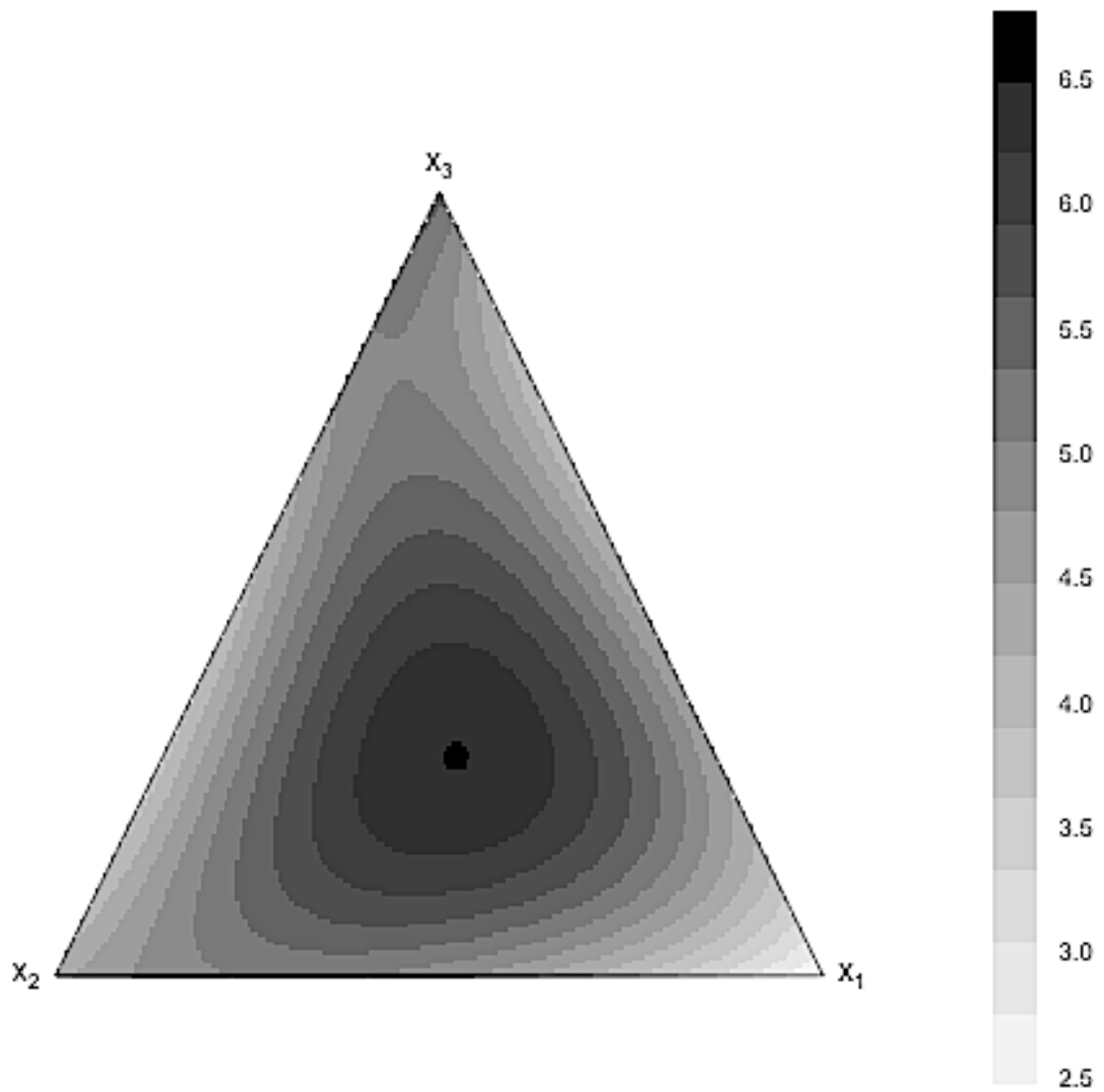


Figure 5: Prediction contour plot for Scheffé special cubic model fit to the simulated data.

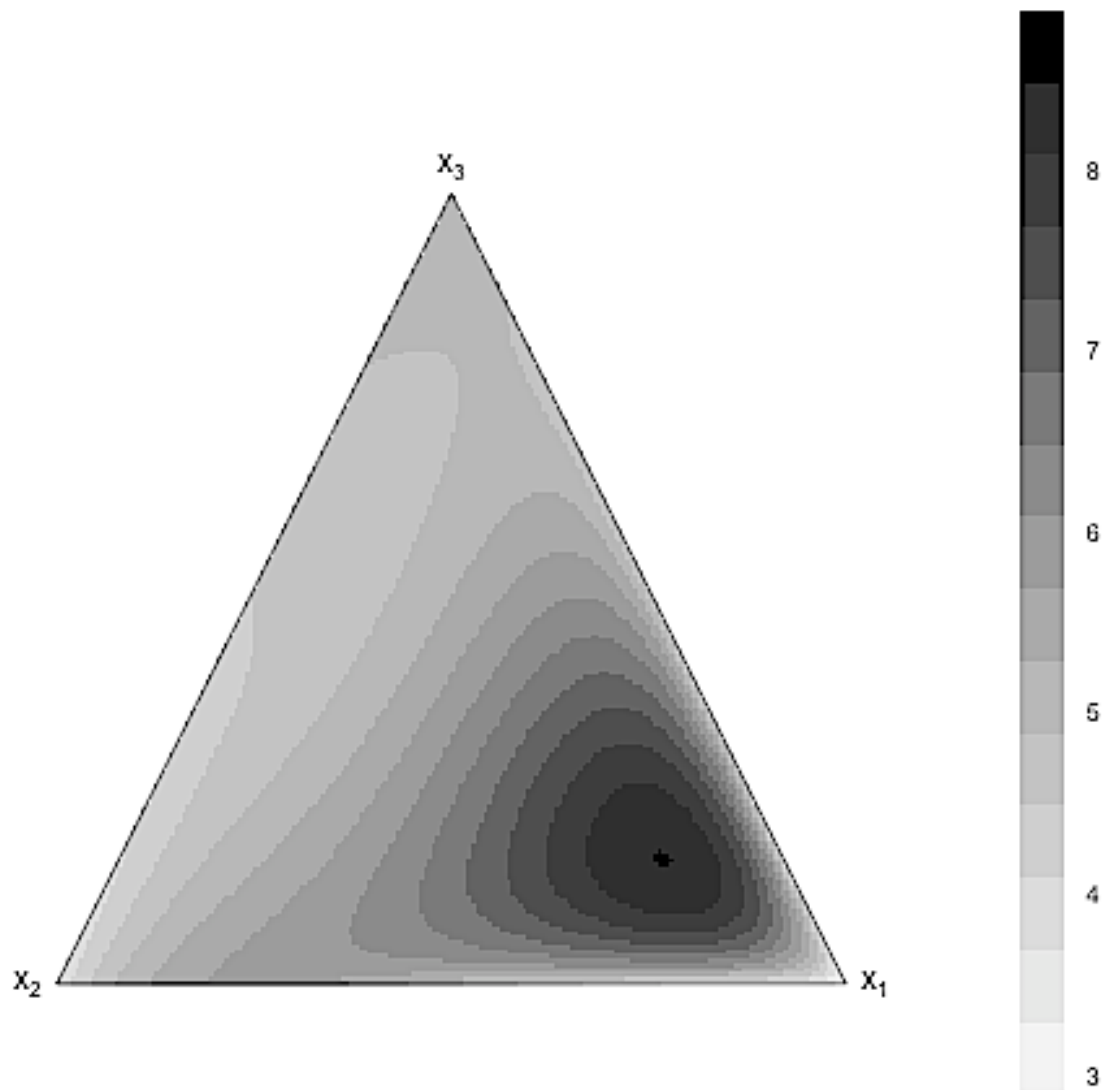


Figure 6: Prediction contour plot for GBM model fit to the simulated data.