Using geometric mean to compute robust mixture designs

Wanida Limmun, Boonorm Chomtee, John J. Borkowski

Quality and Reliability
Engineering International

Using Geometric Mean to Compute Robust Mixture Designs

Journal:	*Quality and Reliability Engineering International*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	LIMMUN, WANIDA; Walailak University School of Science, Mathematics and Statistics Chomtee, Boonorm; Kasetsart University Faculty of Science, Statistics Borkowski, John; Montana State University Bozeman, Mathematical Sciences
Keywords:	Experimental design, Robust design, Reliability
Additional Keywords:	mixture experiments, optimal design, genetic algorithm, D-efficiency, geometric mean

Using Geometric Mean to Compute Robust Mixture Designs

Wanida Limmun[a*], Boonorm Chomtee[b], John J. Borkowski[c]

[a] Department of Mathematics and Statistics, Walailak University, Thasala, Nakhon Si Thammarat, 80160, Thailand

[b] Department of Statistics, Kasetsart University, Chatuchak, Bangkok, 10903, Thailand

[c] Department of Mathematical Sciences, Montana State University, Bozeman, MT, 59717, USA

Abstract

Mixture experiments involve developing a dedicated formulation for specific applications. We propose the weighted D-optimality criterion using the geometric mean as the objective function for the genetic algorithms. We generate a robust mixture design using genetic algorithms (GAs) of which the region of interest is an irregularly shaped polyhedral region formed by constraints on proportions of the mixture component. This approach finds the design that minimizes the weighted average of the volume of the confidence hyperllipsoids across a set of reduced models. When specific terms in the initial model display insignificant effects, it is assumed that they are removed. The design generation objective requires model robustness across the set of the reduced models of the design. Proposing an alternative way to tackle the problem, we find that the proposed GA designs with an interior point are robust to model misspecification.

Keywords: genetic algorithm; mixture experiments; D-efficiency; geometric mean; optimal design

***Correspondence :** Wanida Limmun, Department of Mathematics and Statistics, Walailak University, Thasala, Nakhon Si Thammarat, 80160, Thailand.

Email : lwanida@mail.wu.ac.th

1. Introduction

Mixture experiments are a special class of response surface experiments that involve mixing various proportions of two or more components within an experimental region and then measure the response characteristics of the end product. Mixture experiments occur commonly in chemical and food industries, and pharmaceutical drug formulations, as well as other industrial processes. The goal of the mixture experiment is to calculate the proportion of the mixture components that lead to desirable response characteristics of the end product. In a typical mixture experiment, it is believed that the response is only a function of the proportions of the components in the mixture. The response is independent of the total amount of the mixture. Suppose that the experiment involves q mixture components with proportions given by x_i, for $i = 1, 2,..., q$. The mixture components proportions (x_i) must satisfy:

$$0 \leq x_i \leq 1 \;\; ;i = 1,2,...,q \text{ and } \sum_{i=1}^{q} x_i = 1 \qquad (1)$$

These restrictions are the primary limitations assigned to the mixture experiments. Due to these constraints the mixture components cannot be varied independently, and the mixture design space is reduced to a regular $(q - 1)$-dimensional simplex. In addition to the natural constraints on the component proportions, the components are often further confined by lower (L_i) and upper (U_i) bounds, i.e.,

$$0 \leq L_i \leq x_i \leq U_i \leq 1 \quad ;i = 1, 2,...,q. \qquad (2)$$

If the constraints in (2) are imposed on some or all of the component proportions, the mixture design space becomes an irregularly shaped polyhedron subspace within the simplex. (See Cornell[1], Myers et al.[2] and Smith[3] for more information on examples and applications of mixture experiments.)

When the design space is an irregularly shaped polyhedron within the simplex, the standard mixture designs in the full simplex are not implementable, and optimal designs are commonly used. It is common to generate near-optimal designs using a point-exchange algorithm for selecting the design points from a set of candidate points with the goal of optimizing a particular optimality criterion for a specified regression model. Optimality criteria are single-valued measures based on different desirable variance properties that characterize how good a design is. Details of optimal designs and optimality criteria are well-documented by Atkinson et al.[4], Fedorov[5], and Kiefer[6]. The most commonly-used criterion has been D-optimality which is focused on parameter estimation; specifically, on minimizing the volume of the joint conference region on the model coefficients. In this research, we consider only D-optimality.

Genetic algorithms (GAs) are classified as an evolutionary algorithm (EA) motivated by natural selection that yields robust search capabilities in complex spaces. They offer a logical approach to solve sophisticated mathematical modeling and optimization problems in physical science, engineering, and industry during the last thirty years. This approach considers a set of solutions where the population size is the number of potential solutions. The result yields the best, the desired optimal, or acceptable solution. Holland and his collaborators developed the genetic algorithms based on the concept of Darwin's theory of evolution in the 1960s and 1970s. This algorithm was further extended by Goldberg[7] in 1989 to cope with a difficulty in the control of the gas-pipeline transmission. GAs have also been successfully used to generate optimal experimental designs. Recent applications of genetic algorithms provided an alternative approach to classic exchange-point algorithms for constructing optimal designs. Uses of GAs to find optimal designs can be found in Heredia-Langner et al.[8,9], Borkowski[10], Park et al.[11], Drain et al.[12], Rodriguez et al.[13], and Limmun et al.[14,15,16].

In general, the true relationship between the response of interest and the number of associated control variables is unknown but it can be estimated by using a good approximating model. Frequently, the final model in the analysis phase differs from the initial model when the experimenter found that some effect terms in the initial model are not statistically significant and are subsequently removed from the model. A design is considered to be non-robust if the design does not have desirable prediction variance properties with respect to any potential final model. Thus, the problem of robust designs can overcome this problem. In this research, we zero in on the problem of robustness of the design against model change, the so-called model robustness. We extend the work by Limmun et al.[14,15,16]. The key difference is that in Limmun et al.[14], a single model was used to generate an optimal design, while in this research, all possible reduced models are taken parts in the generation of a robust optimal design. Another distinction is that in Limmun et al.[15,16], the weighted arithmetic mean was employed to construct the objective function of the genetic algorithm, whereas in this research the geometric mean is employed to formulate the objective function of the genetic algorithm because of desirable properties in comparison to the arithmetic mean.

In Section 2, the Scheffé model and some background theory concerning the mixture experiments are introduced. The weighted D-efficiency based on the geometric mean and its benefits as the objective function of the genetic algorithm is described in Section 3. The genetic algorithm organization is presented in Section 4, whereas in Section 5 the proposed scheme is demonstrated with an example to compare and to illustrate the productivity of the scheme. Conclusions are in Section 6.

2. Notation and Model development

2.1 Notation and model

Due to the restrictions on component proportions, the intercept term is not included in a Scheffé mixture model which can be written in the matrix form as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{4}$$

where $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)'$ is the vector of n observed responses; $\boldsymbol{X}$ is the $n \times p$ matrix with columns associated with the p model effects; $\boldsymbol{\beta}$ is the $p \times 1$ vector of model parameters; and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)'$ is the vector of random errors associated with natural variation of $\boldsymbol{y}$ around the underlying surface and are assumed to be independent and identically normally distributed with zero mean and variance $\sigma^2 \boldsymbol{I}_n$. Although there are many possible mixture experiment models, the Scheffé models most commonly used in applications. For example, the Scheffé linear mixture model is:

$$y = \sum_{i=1}^{q} \beta_i x_i + \varepsilon, \tag{5}$$

and the Scheffé quadratic mixture model is expressed as

$$y = \sum_{i=1}^{q} \beta_i x_i + \sum_{i} \sum_{i<j}^{q} \beta_{ij} x_i x_j + \varepsilon, \tag{6}$$

where y is the response variable, the coefficient β_i is the expected response when the i-th component equals unity ($x_i = 1$), the coefficient β_{ij} repesents the nonlinear blending effect of the i-th and the j-th components, and ε indicates the random error. There are also higher-order mixture models that may be considered. The model under consideration in this research is the Scheffé quadratic model mixture model. For a review of Scheffé mixture models, see Cornell[1] and Smith[3].

2.2 Optimality criteria

Most design optimality criteria are single-valued and are used to assess how good a design is for certain design properties. Therefore, most optimality criteria can used as a basis for generating, as well as, comparing designs. They are mainly concerned with optimal properties of the information matrix, $\boldsymbol{X}'\boldsymbol{X}$, where $\boldsymbol{X}$ is the model matrix, i.e., the design matrix with columns expanded to the model form. Four commonly used "alphabetic" optimality criteria for evaluating design are the D, A, G and IV criteria. If the goal of the research is focused on the parameter estimation, the D- and A-optimality criteria are often used. However, the G- and IV-optimality criteria are recommended if the goal of the research is focused on prediction variance. The D-optimality criterion is the most widely-used design criteria and the most well-studied problem as seen in the statistical literature by Atkinson et al.[4], Kiefer[6], and Fedorov[5]. This optimality criterion aims to minimize the volume of the joint conference region on the model coefficients, or equivalently, maximize the determinant of the information matrix, $|\boldsymbol{X}'\boldsymbol{X}|$. Therefore, the larger the determinant of the information matrix, the better the estimation of the model parameters. The D-efficiency measure used in this research is

$$D_{eff} = \frac{100|\boldsymbol{X}'\boldsymbol{X}|^{1/p}}{n}, \tag{7}$$

where p and n are the number of model parameters and design points, respectively. In the model robustness problem, the experimenter prefers a design having good properties concerning any potential final model. In this research, we focus on D-optimality because the D-optimality is essentially a parameter estimation criterion. The comprehensive review and uses of the optimality criteria can be found in Atkinson et al.[4], Kiefer[6], Fedorov[5], and Pukelsheim[17].

2.3 The scaled prediction variance

Box and Hunter[18] recommended that the precision of the coefficient and the variances of the individual model coefficient estimates should be studied simultaneously. The scaled prediction variance (SPV) of the expected response can be expressed as

$$\mathrm{SPV} = \frac{n\mathrm{Var}(\hat{y}(\boldsymbol{x}))}{\sigma^2} = n\boldsymbol{x}'^{(m)}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}^{(m)}, \tag{8}$$

where $\mathrm{Var}(\hat{y}(\boldsymbol{x}))$ is the variance of the estimated response at model vector $\boldsymbol{x}^{(m)}$ which represents design point $\boldsymbol{x}$ expanded to model form, $\boldsymbol{X}$ is the design matrix expanded to model form; and σ^2 is the variance. Thus, SPV depends on the selected model, and is penalized by for larger design size N. For example, for the Scheffé quadratic model with q mixture components, then $\boldsymbol{x}'^{(m)} = (x_1,x_2,\ldots,x_q,x_1x_2,x_1x_3,\ldots,x_{q-1}x_q)$ and matrix $\boldsymbol{X}$ has dimension $n \times \left(\frac{q(q+1)}{2}\right)$. Thus, a design having a small maximum SPV over the design space is a desirable design. The SPV is broadly used to determine the precision of the predicted response on a per observation basis. That is, it has capacity for comparing design with different sizes of runs.

3. Weighted optimality criterion on Geometric mean

A design optimality criterion can be applied to evaluate a proposed design for a specified model. Suppose, however, that after the data have been collected and the model parameters are estimated, the experimenter found that some model terms are insignificant, and a reduced model is adopted. That is, the final reduced model only retains terms for further analysis. The design optimality criterion value for the reduced model is equally, if not more important than that for the initial or "full" model. Consequently, the design should have robust properties over the class of possible reduced models given the final model is unknown a priori.

Li and Nachtsheim[19] considered a weighted criterion based on a design's estimation capacity and the D-criterion. Butler[20] presented the Gaussian weighted function on the optimality criteria for polynomial responses surfaces. The authors found that the weighted ASV-optimality is equivalent to weighted A-optimality based on the Gaussian weighted function. Heredia-Langner et al.[9] employed a genetic algorithm to produce a robust design based on optimizing a desirability function. Borkowski et al.[21] presented the prior probabilities to the set of all possible reduced models based on model heredity principles for creating the weighted D, G, A, and IV-efficiencies. They found that the design optimality criteria can be affected by deviations from the full model. Stallings and Morgan[22] proposed a weighted information matrix for any positive definite weight matrix. Limmun et al.[15,16] developed the weighted A-optimality criterion and the weighted G-efficiency based on an arithmetic mean of efficiencies as the objective function of a genetic algorithm. Rosa[23] extended work of Stallings and Morgan[22] in 2015 and proposed a weighted eigenvalue-based criterion.

Weighted optimality criteria form a methodology for illustrating a variety of properties of model parameter estimates and the prediction variance. The weighted D-optimality (WD-optimality) criterion is developed in this research. The goal of the WD-optimality criterion is to search for a set of points minimizing the weighted average of the efficiencies across a set of reduced models. For the concept of weighted optimality criterion, we will assume that models having the same number of parameters have equal weight. An additional assumption is that all linear $\beta_i x_i$ effects must occur in each reduced model. The WD-optimality criterion is calculated based on the geometric mean because one of advantages of the geometric mean is that it is very sensitive to the existence of one or more small efficiencies. The complete enumeration of the D-efficiencies across all possible reduced models is necessary for calculating the WD-efficiency. We define the WD-efficiency based on the geometric mean as

$$WD = \prod_{i=1}^{r} D_i^{w_i}, \tag{9}$$

where $D_i = \frac{100|\boldsymbol{X}'\boldsymbol{X}|^{\frac{1}{p_i}}}{n}$ is the D-efficiency for model i, r is the number of reduced models for a given full model, w_i is the weight for model i, p_i is the number of parameters for model i, $\boldsymbol{X}$ is a mixture model matrix for model i, and n is the number of design points.

Prior to generating a design, we must assign weights to each reduced model. When calculating *WD*. Suppose that each parameter is equally important. The weight of each reduced model reflects the importance of the model with respect to the number of parameters. We now define the weight associated with each reduced model. Suppose model i has p_i parameters and $m = \sum_{i=1}^{r} p_i$ is the sum of the number of parameters of all possible reduced models. We define a set of model weights $(w_i, i = 1,2,\ldots,r)$ that is positive weights satisfying $\sum_{i=1}^{r} w_i = 1$. Then, the weight for model i is defined as

$$w_i = \frac{p_i}{m}. \tag{10}$$

Recall that the considered full model is the Scheffé quadratic model. To calculate the WD-efficiency, the Scheffé linear model must receive the smallest weight, whereas the Scheffé quadratic model receive the largest weight. Recall that models containing an equal number of model parameters have equal weight.

For three components, suppose the full model is the Scheffé quadratic model and can be expressed as

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

The full model and all possible reduced models are shown in Table 1. In Table 1, a 1 and 0 in any terms in the model columns represents the presence and absence of terms in the model parameters, respectively. The column p_i represents the number of model parameters and the column w_i represents the weight of each model. Consider the model having five parameters, the 2nd , 3rd, and 4th model, we find that the weight of each model is 5/36 ≈ 0.1389. Note that *WD* will be large when each D_i is large, and will be small if any D_i is close to zero. This is not necessarily the case with an arithmetic mean. For example, consider the design with mixtures (1,0,0), (0,1,0), (0,0,1), (1/3,2/3,0), (2/3,1/3,0), (1/3,0,2/3), (2/3,0,1/3), (0,1/3,2/3), (0,2/3,1/3). This contains 9 points equi-spaced on the boundary of the simplex. Table 2 contains the D-efficiencies (D_i), model weights (w_i), arithmetically and geometrically weighted efficiencies ($D_i w_i$ and $D_i^{w_i}$) across the 8 models. The weighted geometric $\mathrm{WD_G}$-eff and weighted arithmetic $\mathrm{WD_A}$-eff efficiencies are 7.7954 and 6.7948, respectively, where $\mathrm{WD_A}$-eff

$= \sum_{i=1}^{8} D_i w_i$. As seen in this example, the variability in the $D_i^{w_i}$ values is small indicating no model is dominating the efficiency calculation of WD and keeping WD from going to 0. In contrast, however, the models having the fewest number of parameters contribute the most to WD_A-eff thereby putting greater importance on those models in the efficiency calculation. This leads to a weighted efficiency that is inflated by models with relatively few model parameters, which is not the case using WD efficiency. It is important to note that our assignment of weights is just one possible assignment with the goal of demonstrating how WD can be used. For example, if an experimenter has some prior knowledge about model terms such as being certain of an x_1x_2 blending effect, a different set of model weights can be used that assign 0 weight to models without an x_1x_2 blending effect.

4. A genetic algorithm for constructing optimal designs

Genetic algorithms (GAs) are inspired by natural selection where the strongest individuals are chosen for the reproductive process in order to produce offspring. A chromosome is a potential solution to an optimization problem and a gene is an assigned location within a chromosome. The GA selection process ensures fittest individuals from a population are retained for breeding in future generations. The crossover and mutation processes alter the population of chromosomes temporally. The fitness function gives a fitness score that is used to determine the fitness of population chromosome. Based on the idea of survival, the genetic algorithm selects the chromosomes with the highest fitness values to be parents for the next generation. If the best chromosome is not improved for a specified number of generations, the algorithm terminates. For an introduction and review of genetic algorithms, see Goldberg[7], Michalewicz[24], Kinnear[25], and Haupt and Haupt[26].

In this research, a chromosome represents the experimental design matrix, while a gene represents a row in a chromosome or design matrix. Thus, a potential design is characterized by a chromosome C having dimension $n \times q$ where n and q are, respectively, the number of the design points and components. Each row in C is a gene $\boldsymbol{x_i} = [x_{i1} \;\; x_{i2} \;\; \ldots \;\; x_{iq}]$ for $i = 1,2,\ldots n$. For example, if there are six design points, a chromosome C is comprised of $(\boldsymbol{x_1},\boldsymbol{x_2},\boldsymbol{x_3},\boldsymbol{x_4},\boldsymbol{x_5},\boldsymbol{x_6})$.

We propose WD-efficiency as the objective function in order to determine the fitness of a chromosome as a model robust design. Recall that the aim of the WD-optimality criterion is to obtain a design that minimizes the weighted geometric average of *D*-

efficiencies across a set of reduced models representing possible misspecification of the full model. Again, this extends the genetic algorithm developed by Limmun et al.[14, 15, 16].

The following summarizes the steps in our genetic algorithm:

Step 1 Determine the genetic parameters, such as the initial population size (M), the number of iterations, the selection method, the blending rate (α_b), the crossover rate $(\alpha_{cb}, \alpha_{cw})$, the extreme rate (α_e), the midpoint rate (α_{mi}) and the mutation rate (α_{mu}). We use elitism with random parent pairing for selection process.

Step 2 Generate an odd number M chromosomes (designs) as the initial population. We encode each chromosome using the real-valued encoding with four decimal places. We use the function in Borkowski and Piepel[27] for generating the initial population which takes randomly sampled points in a hypercube and then maps them into a constrained mixture space.

Step 3 Calculate the WD-efficiency as the objective function for each chromosome in the initial population.

Step 4 Identify the elite chromosome having the highest WD-efficiency. The remaining $M-1$ chromosomes are randomly paired for the reproduction process.

Step 5 Generate offspring by using reproduction operators: *blending*, *between-parent crossover*, *within-parent crossover*, *extreme gene*, *midpoint gene* and *mutation*.

Higher rates for genetic parameter operations are employed in the early iterations while smaller rates are used for the later. A genetic operator will only be performed if a probability test is passed (PTIP). Let U be a uniform [0,1] random variable with parameter α. If $U \leq \alpha$, then a PTIP. Thus, when a PTIP, a reproduction process is imposed on a parent chromosome to create an offspring chromosome. For blending and crossover, let A and B be two parents paired in the selection process. Let gene A_a be the a-*th* row of A and gene B_b be the b-*th* row of B.

Blending: A_a is combined with a random row of B, say B_b, to form two new rows, say A_a^* and B_b^*, using the linear combinations:

$$A_a^* = \beta A_a + (1-\beta)B_b \quad \text{and} \quad B_b^* = (1-\beta)A_a + \beta B_b \tag{12}$$

respectively, where β is a $[0,1]$ uniform random variable.

Between parent crossover: An offspring is obtained by swapping the tail fractions of two genes of two paired parents.

Within parent crossover: An offspring is obtained by swapping the tail fractions within a gene.

For *extreme, midpoint,* and *mutation* operators, let A_i be the i-*th* row or gene $x_i$of chromosome A. If a PTIP, one of the q components, say h, is randomly chosen and the remaining component proportions in gene x_i must be adjusted to ensure that the sum of the component proportions equals one.

Extreme gene: If the extreme point of the h-*th* component $(\mathrm{EP}_h = 0.5 - \beta) < 0$ then $\varepsilon_h = x_{ih} - L_h$ and the new gene $x_{ih}^* = L_h + \varepsilon_{ih}$. If $\mathrm{EP}_h > 0$ then $\varepsilon_{ih} = U_h - x_{ih}$ and the new gene $x_{ih}^* = U_h + \varepsilon_{ih}$.

Midpoint operator gene: If the midpoint of the h-*th* component (mid_h) is less than x_{ih}, then $\varepsilon_h = x_{ih} - \mathrm{mid}_h$, otherwise $\varepsilon_h = \mathrm{mid}_h - x_{ih}$. This variate ε_h is added to a gene x_{ih} to form a new gene $x_{ih}^* = x_{ih} + \varepsilon_h$.

Mutation: a random variate ε_h from a normal $N(0,\sigma^2)$ distribution is added to a gene x_{ih} to form a new gene $x_{ih}^* = x_{ih} + \varepsilon_h$.

Step 6 Calculate the objective function for each parent/offspring pair. The fitter chromosome in the pair is kept for the next generation.

Step 7 Repeat Steps 4 to 6 until the given stopping criterion is satisfied indicating convergence to a near optimal design.

Step 8 After the last iteration, a local grid search is applied to the best design (near optimal design) for finding the optimal design.

5. Illustrative example

To illustrate our process, a mixture design problem in an irregularly-shaped polyhedral region is considered. We compare the designs produced by our genetic algorithm based on the geometric mean and arithmetic mean with the designs given by the Design-Expert[28] software. The candidate set of the Design-Expert 11 is comprised of edge centroids, axial blends, interior blends, vertices, and an overall centroid. We used the D-criterion with Coordinate Exchange searches for all designs generated by the Design-Expert 11. The GA designs are generated using an author-written program in MATLAB. For the genetic algorithm, α_b,α_{cb},α_{cw},α_e,α_{mi} and α_{mu} represent the success probabilities of blending, between-parent crossover, within-parent crossover, extreme gene, midpoint gene, and mutation, respectively. σ is the standard deviation of the mutation operator. The values of genetic parameters are investigated before running the genetic algorithm. An efficient set of α_i and σ values was selected to produce the highest objective function values which then stabilize throughout the later iterations.

Two commonly used methods for comparing the properties of competing designs are design efficiencies and graphical displays. The design efficiency represents a percentage calculated relative to some theoretically best value according to the respective criterion. We consider the fraction of design space (FDS) plot for assessing the design robustness to model reduction. An FDS plot displays the prediction variance over the entire fraction of the design space and indicates the minimum and maximum prediction variance, as well as any percentile of the prediction variance. We limit our reduced model comparison in the FDS plots to the Scheffé quadratic and linear models because the curves for these models give a manageable summary that allows for comparison of prediction variance performance. For additional information of the FDS curves for examining model robustness, see Goldfarb et al.[29] and Ozol-Godfrey et al.[30].

Example: This thread elongation experiment with three mixture components is taken from Cornell[1]. The objective of the experiment is to explore the effects of the component proportions on the thread elongation of yarn. There are three components including polyethylene (x_1), polystyrene (x_2), and polypropylene (x_3). These components have the following lower and upper proportion constraints:

$$0.27 \le x_1 \le 0.59; \quad 0.15 \le x_2 \le 0.45; \quad 0.20 \le x_3 \le 0.34.$$

This boundary is formed by six vertices. The full model is the Scheffé quadratic model:

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

The following summarizes the ranges of genetic parameter values:

$0.03 \le \alpha_b, \alpha_{cb}, \alpha_{cw} \le 0.25$; $0.02 \le \alpha_e, \alpha_{mi} \le 0.15$; $0.03 \le \alpha_{mu} \le 0.25$;$0.01 \le \sigma \le 0.10$.

Our genetic algorithm performs for 8,000 iterations and the success probabilities are reset to new values after 2,000 iterations for every case.

We consider designs with 7 to 10 design points for testing the performance of the GA designs based on the geometric mean (GA-G), the GA designs based on the arithmetic mean (GA-A), and designs generated by the Design-Expert (DX). The DX designs are generated using the D-optimality criterion. In Tables 3 and 4, the digit attached to GA and DX represents the number of design points. The list of design points of the GA-G and GA-A designs are summarized in Table 3 and displayed in Figure 1 which highlights the differences in the distributional point patterns of these competing designs. Note that all DX designs have a near overall centroid point and no replicated points, while all GA-A designs have no near overall centroid point but have some replicated points that vary for different vertices, whereas all GA-G designs contain all six vertices. The GA7-G and GA10-G designs have an overall centroid point, and a near- overall centroid point, respectively, while the GA9-G and GA10-G designs contain some replicated vertices.

Table 4 presents weighted D-efficiencies based on the geometric mean (WD_G-eff), weighted D-efficiency based on the arithmetic mean (WD_A-eff), D-efficiencies for the quadratic model only (D-eff), and maximum SPV values for all competing designs. Recall that larger is better for efficiencies, while smaller is better for the maximum SPV. As expected, the largest WD_G-eff and WD_A-eff are associated with GAn-G and GAn-A designs, respectively, given these designs were optimized for those criteria. What is interesting is that for 2 of the 4 cases, the GAn-G designs have the largest D-efficiencies for the quadratic model and very close in the other 2 cases, despite the Design-Expert objective to maximize this efficiency. This can occur because Design-Expert generates designs from a finite candidate set while the GA searches the continuous space which can therefore produce better D-efficient designs. Also, not surprising are the worst WD_G-eff and WD_A-eff values for the Design-Expert designs because of the different objective function (i.e., D-efficiency).

In Table 4, the GAn-A designs have demonstrably larger maximum SPV values in comparison to the DXn and GAn-G designs for the quadratic model. In addition, for the linear model, the GAn-G designs have the smallest maximum SPV values in comparison to the other designs. Thus, the GAn-G designs have desirable properties based on this supplemental maximum SPV criterion. These results support the assertion of the model robustness properties of a GAn-G optimal design. We will now compare FDS plots across the 8 models in Table 1 to further assess the robustness of the GA-G designs. Because The GA-A designs are no longer competitive with respect to the larger maximum SPV values for both the Scheffé quadratic and linear models, we did not include their FDS plots.

Figure 2 shows the FDS plots for the GA-G and DX designs for the full quadratic and linear models (models 1 and 8, respectively). For $n = 8$ and 9, the DXn designs perform better than the GAn-G designs throughout the design space for the quadratic model which is not surprising given that the quadratic model is the criterion for generating DX designs. As seen in Figure 1, GA8-G and GA9-G have no interior points while DX-8 and DX-9 do. Therefore, if the experimenter decides to use GA8-G and GA9-G, there is a risk of having poor prediction variance properties in a potentially important region of the design space for the full quadratic model. This, however, will not be the case for the linear model in which case the FDS plots are very close but with GA8-G and GA9-G slightly better.

For $n = 7$ and 10, however, the quadratic model FDS plots of the GAn-G and the DXn designs very similar except for about 10% of the design space in which the DXn designs have the largest SPV values. For the linear model, the FDS plots are very close but with GA designs being slightly better for all n. The GA7-G and GA10-G designs display good model robustness properties since their FDS curves are much flatter and lower for both the Scheffé quadratic and linear models, indicating superior SPV distributions. We will now consider the FDS plots for the other six models. Figure 3 and 4 contain three FDS plots for the GA-G and DX designs for models having 5 parameters (models 2,3,4) and 4 parameters (models 5,6,7), respectively.

For the 5-parameter models in Figures 3(a) to 3(d), the FDS plots of DX designs are worse than those for the GA designs for model 2. For models 3 and 4, the DX7 and DX10 designs are superior to the GA designs except for the largest 10% to 20% (right tails) of the FDS, and for $n = 9$ and 10, the GA and DX have similar FDS plots for 80% of the distribution. Note that for all models, the right-tail extremes for the DX designs are noticeably larger than those for the GA designs.

For the 4-parameter models in Figures 4(a) to 4(d), the FDS plots of DX designs are worse than those for the GA designs for models 5 and 6 (with the exception in the largest SPV values for $n = 10$ which are close but just slightly larger). For model 7, the DXn designs are superior to the GA7-n designs for $n = 7, 9$ and 10 except for the largest 10% to 20% of the FDS. For model 8, the DX8 and GA8-G FDS plots are very close for the first 70% but separate in the right tail with the GA8-G design being superior.

Studying the geometry of the design points in Figure 1 reveals several patterns: (i) All GAn-G designs contain all 6 vertices which is never the case for any of the other designs (ii) GAn-A designs contain more replicated points at vertices, (iii) no GAn-A design contains an interior point, and (iv) all DXn designs contain an overall centroid. Pattern (ii) is consistent with GAn-A designs assigning greater weight to certain reduced models limiting the need to spread the 9 points on the boundary.

In summary, this example highlights the potential of using the *WD* optimality criterion based on a geometric mean for finding a model robust design as in the cases with $n = 7$ and 10. Except for two cases for the full quadratic model ($n = 8$, and 9), the GAn-G designs have the best right-tail FDS plot properties, and in that sense, the GAn-G designs are model robust. It is important to acknowledge that there can be infrequent drawback cases using any single value optimality criterion, as in the aforementioned two cases for the quadratic model. Thus, before choosing a model-robust design it is important to use other tools, such as FDS plots, to aid in design selection.

6. Conclusions

Our basic assumption in this model-robustness problem is that an experimenter will decide to remove insignificant terms from the initial full mixture model leading to a reduced final mixture model. In this case, it is important for the experimenter to exercise caution when selecting an experimental design. The design having robust properties over classes of reduced models should be considered in order to addressing this problem.

To address the design selection problem, we introduced the weighted D-optimality criterion *WD* based on the geometric mean. *WD* assures that no reduced model efficiency will be too close to zero. Although any user-supplied model weights can be used, for demonstration purposes, we assume that the models possessing an equal number of model parameters receive equal weight for calculation of *WD*. The problem of generating *DW*-optimal designs is studied for and irregularly-shaped polyhedron mixture space.

A GA was used to generate *WD*-optimal designs because GAs are structurally different from the traditional exchange algorithms in that they are very effective in performing an intense search inside the continuous design space, and require no selection points from a user-defined candidate set of mixtures.

According to the example results, the GA-G designs having all vertices and an interior point appear to be model-robust because these designs start with small values of the SPV and remain relatively flat throughout the design space across a set of possible reduced models. More research is needed to see if this happens in general. Therefore, if the experimenter is concerned with protection against the worst-case for the prediction variance, the addition interior point in GA-G design may be warranted. The computing time for generating GA*n*-G and GA*n*-A designs ranged from 45 to 55 seconds. In practice, although there is no single approach that can perform well for all situations, the experimenter may consider using GA with *WD*-optimality as an alternative approach for generating model-robust mixture designs.

Acknowledgement

This work was financially supported by Walailak University under the grant number WU62211.

References

1. Cornell JA. *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*. 3rd ed. New York, NY: John Wiley & Sons; 2002.

2. Myers RH, Montgomery DC. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 3rd ed. New York, NY: John Wiley & Sons; 2009.
3. Smith WF. *Experimental Design for Formulation*. Philadelphia, Pennsylvania: The American Statistical Association and the Society for Industrial and Applied Mathematics; 2005.
4. Atkinson AC, Donev AN, Tobias RD. *Optimum Experimental Designs, with SAS*. Oxford: Oxford University Press Inc.; 2007.
5. Fedorov VV. *Theory of Optimal Experiments*. New York: Academic Press; 1972.
6. Kiefer J. Optimum experimental designs. *J. Royal Stat. Soc B.* 1959; **21**(2): 272-304.
7. Goldberg DE. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Boston: Addison-Wesley; 1989.
8. Heredia-Langner A, Carlyle WM, Montgomery DC, Borror CM, Runger GC. Genetic algorithms for the construction of D-optimal designs. *Journal of Quality Technology.* 2003; **35**(1): 28-46.
9. Heredia-Langner A, Montgomery DC, Carlyle WM, Borror CM. (2004). Model-robust optimal designs: a genetic algorithm approach. *Journal of Quality Technology.* 2004; **36**(3): 263-279.
10. Borkowski JJ. Using genetic algorithm to generate small exact response surface designs. *Journal of Probability and Statistical Science.* 2003; **1**: 65-88.
11. Park Y, Montgomery DC, Fowler JW, Borror CM. Cost-constrained G-efficient response surface designs for cuboidal regions. *Quality and Reliability Engineering International.* 2006; **22**(2): 121-139.
12. Drain D, Carlyle WM, Montgomery DC, Borror C, Anderson-Cook C. A genetic algorithm hybrid for constructing optimal response surface designs. *Quality and Reliability Engineering International.* 2004; **20**(7): 637-650.
13. Rodriguez M, Montgomery DC, Borror CM. Generating Experimental Designs Involving Control and Noise Variables using Genetic Algorithms. *Quality and Reliability Engineering International.* 2009; **25**(8): 1045-1065.
14. Limmun W, Borkowski JJ, Chomtee B. Using a Genetic Algorithm to Generate D-optimal Designs for Mixture Experiments. *Quality and Reliability Engineering International.* 2013; **29**(7): 1055–1068.
15. Limmun W, Borkowski JJ, Chomtee B. Weighted A-optimality criterion for generating robust mixture designs. *Computers and Industrial Engineering*. 2018; **125**: 348-356.
16. Limmun W, Chomtee B, Borkowski JJ. Constructing model robust mixture designs via weighted G-optimality criterion. *International. Journal of Industrial Engineering Computations.* 2019; **10**(4): 473-490.
17. Pukelsheim F. *Optimal Design of Experiments*. New York: Wiley; 1993.
18. Box GEP, Hunter JS. Multi-Factor Experimental Designs for Exploring Response Surfaces. *Annals of Mathematical Statistics* 1957; **28**(1): 195-241.

19. Li W, Nachtsheim CJ. Model-robust factorial designs. *Technometrics*. 2000; **42**(4): 345-352.
20. Butler NA. On weighted design optimality criteria and response surface parameterizations. *Statistics and Probability Letters* 2001; **51**(1): 41-46.
21. Borkowski JJ, Chomtee B, Turk P. Using Weak and Strong Heredity to Generate Weighted Design Optimality Criteria for Response Surface Designs. *Journal of Statistical Theory and Applications*. 2011; **10**: 163-192.
22. Stallings JW, Morgan JP. General weighted optimality of designed experiments. *Biometrika.* 2015; **102**(4): 925 -935.
23. Rosa, S. Equivalence of weighted and partial optimality of experimental designs. *Metrika*. 2019; **82**(6): 719-732.
24. Michalewicz Z. *Genetic Algorithms + Data Structures = Evolution Programs*. New York: Springer-Verlag; 1992.
25. Kinnear KE. *Advances in Genetic Programming*. Cambridge, MA: MIT Press; 1994.
26. Haupt RL, Haupt SE. *Practical Genetic Algorithms*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2004.
27. Borkowski JJ, Piepel GF. Uniform designs for highly constrained mixture experiments. *Journal of Quality Technology.* 2009; **41**(1): 1-13.
28. Stat-Ease, Inc. *Design-Expert Version 11*. Minneapolis, MN; 2018.
29. Goldfarb HB, Anderson-Cook CM, Borror CM, Montgomery DC. Fraction of design space plots for assessing mixture and mixture-process design. *Journal of Quality Technology.* 2004; **36**(2): 169-179.
30. Ozol-Godfrey A, Anderson-Cook CM, Montgomery DC. Fraction of design space plots for examining model robustness. *Journal of Quality Technology.* 2005; **37**(3): 223-235.

Authors' biographies

Dr. Wanida Limmun is an Assistant Professor in the Department of Mathematics and Statistics, Walailak University, Thailand. She received her PhD in Statistics from Kasetsart University, Thailand. Her area of interests include model robust design, experimental design, mixture experiments, and heuristic algorithms. Her email address is lwanida@mail.wu.ac.th.

Dr. Boonorm Chomtee is an Associate Professor in the Department of Statistics, Kasetsart University, Thailand. She received her PhD in Statistics from Montana State University, USA. Her research interests include experimental design, response surface methodology and multivariate analysis. Her email address is fsciboc@ku.ac.th.

Dr. John J. Borkowski is a Professor in the Department of Mathematical Sciences, Montana State University, USA. He received his PhD in Statistics from University of Delaware, USA. His research interests include experimental design, response surface methodology, statistical quality control methods, sampling, and applications to industry, ecology, and environmental science. His email address is john.borkowski@montana.edu.

Table 1

The set of reduced models for the full quadratic mixture model with three components.
For a table value =1, the term is included in the model. For a table value = 0, it is not.

p_i is the number of model terms and w_i is the model weight for model i.

model	Terms in model						p_i	$w_i = p_i/36$
	$\beta_1 x_1$	$\beta_2 x_2$	$\beta_3 x_3$	$\beta_{12} x_1 x_2$	$\beta_{13} x_1 x_3$	$\beta_{23} x_2 x_3$		
1	1	1	1	1	1	1	6	0.1667
2	1	1	1	0	1	1	5	0.1389
3	1	1	1	1	0	1	5	0.1389
4	1	1	1	1	1	0	5	0.1389
5	1	1	1	0	0	1	4	0.1111
6	1	1	1	0	1	0	4	0.1111
7	1	1	1	1	0	0	4	0.1111
8	1	1	1	0	0	0	3	0.0833

Table 2

Consider the design with mixtures (1,0,0), (0,1,0), (0,0,1), (1/3,2/3,0), (2/3,1/3,0), (1/3,0,2/3), (2/3,0,1/3), (0,1/3,2/3), (0,2/3,1/3). D-efficiencies (D_i), model weights (w_i), arithmetically and geometrically weighted efficiencies ($D_i w_i$ and $D_i^{w_i}$), respectively, across the 8 models are calculated. The WD_G and WD_A efficiencies are 7.7954 and 6.7948, respectively.

Model i	1	2	3	4	5	6	7	8
D_i	3.6760	5.3517	5.3517	5.3517	9.2269	9.2269	9.2269	22.5267
w_i	6/36	5/36	5/36	5/36	4/36	4/36	4/36	3/36
$D_i w_i$	0.6127	0.7433	0.7433	0.7433	1.0252	1.0252	1.0252	1.8772
$D_i^{w_i}$	1.2423	1.2623	1.2623	1.2623	1.2801	1.2801	1.2801	1.2964

Table 3

The mixtures in the GA-G and GA-A designs for the thread elongation experiment.

Common design points for all GA designs		
x_1	x_2	x_3
0.27	0.39	0.34
0.35	0.45	0.20
0.51	0.15	0.34
0.59	0.21	0.20

$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	**Additional design points for**
0.27	0.45	0.28	GA7-G to GA10-G, GA8-A, GA10-A
0.59	0.15	0.26	GA7-G to GA10-G, GA7-A, GA9-A, GA10-A
0.35	0.45	0.20	GA9-G, GA10-G, GA7-A, GA9-A, GA10-A
0.51	0.15	0.34	GA10-G, GA8-A to GA10-A
0.59	0.21	0.20	GA10-G, GA10-A
0.43	0.30	0.27	GA7-G
0.3973	0.2627	0.34	GA8-G
0.4627	0.3373	0.20	
0.39	0.27	0.34	GA9-G
0.4951	0.3049	0.20	
0.4202	0.2976	0.2822	GA10-G
0.388	0.272	0.34	GA7-A
0.4276	0.3724	0.20	GA8-A
0.59	0.21	0.20	
0.27	0.39	0.34	GA9-A
0.4959	0.3041	0.20	
0.3536	0.3064	0.34	GA10-A

Table 4

The weighted D-efficiency based on geometric mean (WD_G-eff), the weighted D-efficiency based on arithmetic mean (WD_A-eff), D-efficiency (D-eff), and Max SPV values for the thread elongation experiment. Bold-faced values indicate the best design for that criterion.

Runs	Design	WD_G-eff	WD_A-eff	D-eff	model	Max SPV
7	GA7-G	**0.3418**	0.6193	**0.0899**	Quadratic	**7.3895**
					Linear	3.8778
	GA7-A	0.3269	**0.6523**	0.0685	Quadratic	39.3470
					Linear	3.6900
	DX7	0.3261	0.5949	0.0840	Quadratic	13.9341
					Linear	4.7077
8	GA8-G	**0.3445**	0.6451	0.0796	Quadratic	21.4423
					Linear	3.5106
	GA8-A	0.3251	**0.6606**	0.0643	Quadratic	60.0870
					Linear	3.6150
	DX8	0.3189	0.5932	**0.0804**	Quadratic	**18.1227**
					Linear	3.9314
9	GA9-G	**0.3431**	0.6471	0.0778	Quadratic	24.7607
					Linear	3.7679
	GA9-A	0.3315	**0.6678**	0.0661	Quadratic	59.4376
					Linear	3.4974
	DX9	0.3104	0.5725	**0.0792**	Quadratic	**15.6679**
					Linear	5.1683
10	GA10-G	**0.3454**	0.6436	**0.0861**	Quadratic	**10.4975**
					Linear	4.1110
	GA10-A	0.3416	**0.6679**	0.0706	Quadratic	56.1382
					Linear	3.4370
	DX10	0.3042	0.5669	0.0766	Quadratic	17.3307
					Linear	4.4032

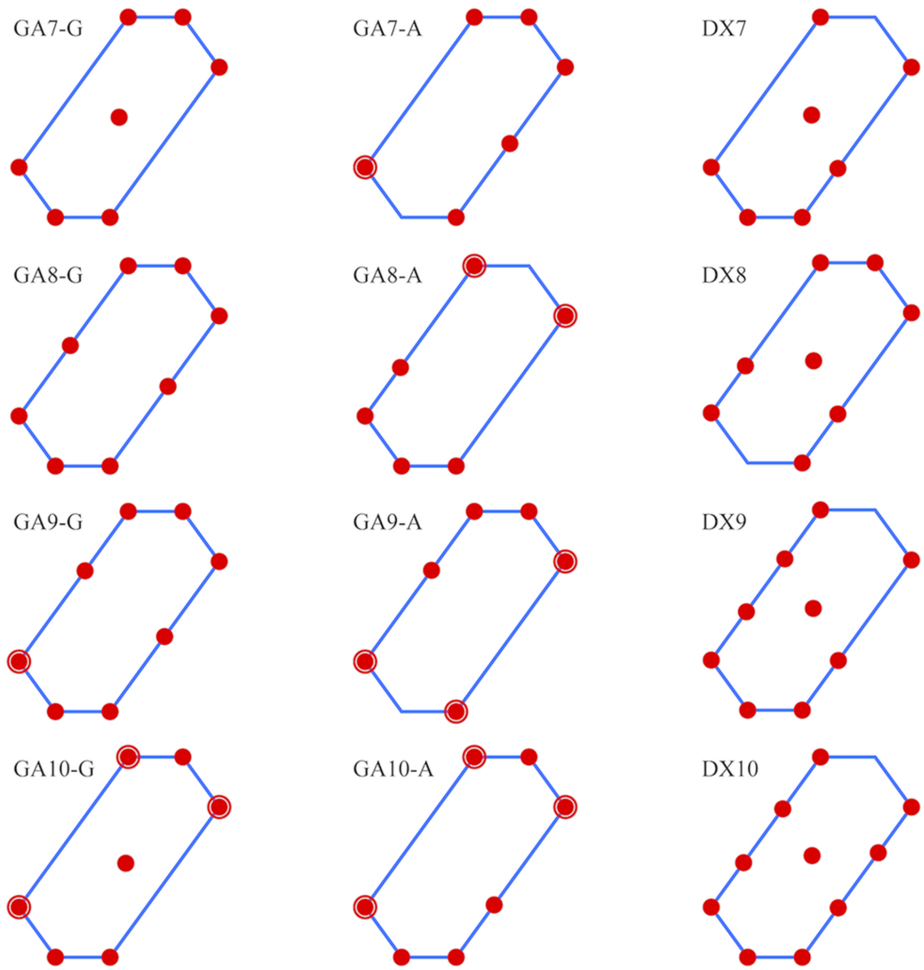

Figure 1. The mixtures in the robust GA and DX designs having 7 to 10 points for the thread elongation experiment. Concentric circles indicate two replications of that mixture.

90x88mm (300 x 300 DPI)

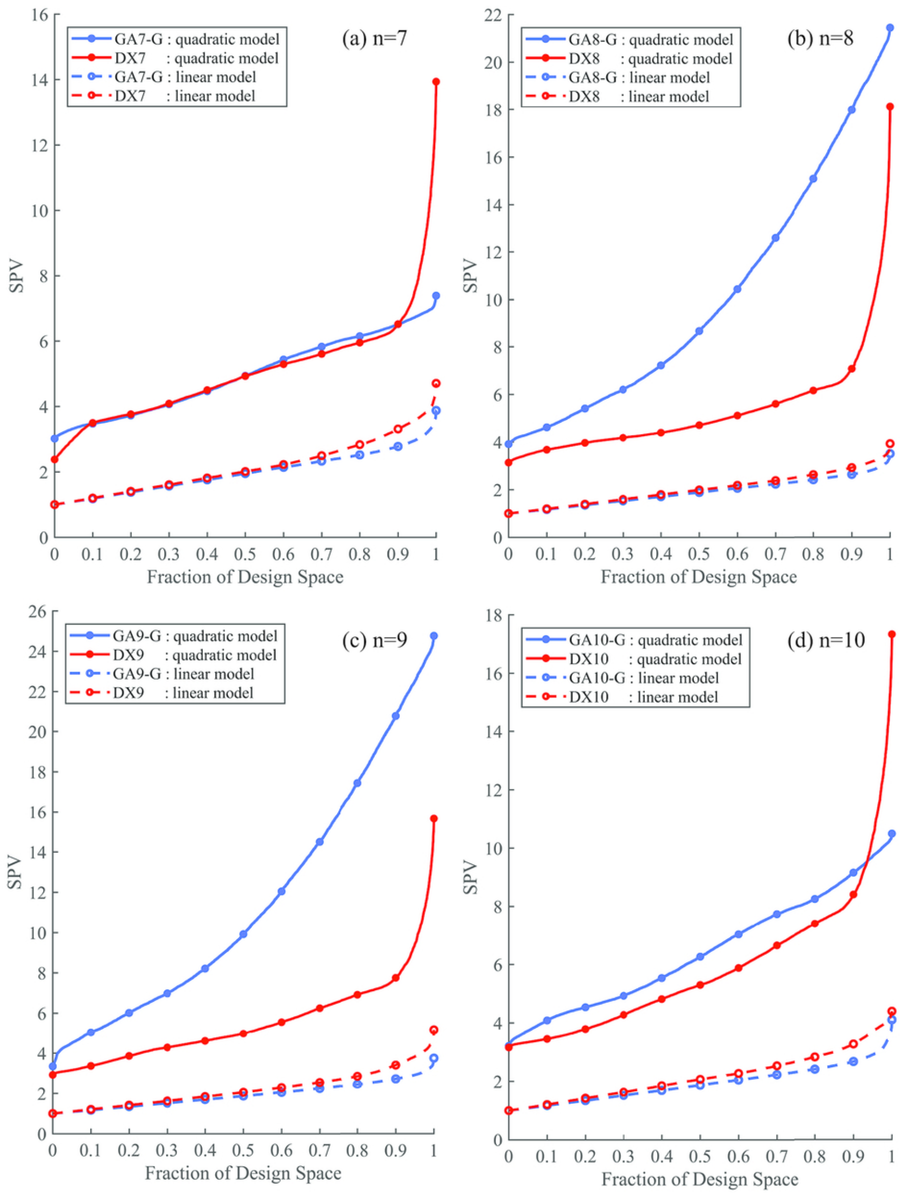

Figure 2. FDS plots for comparing GA designs (blue) and DX designs (red) for the Scheffé quadratic mixture model (model 1) and the Scheffé linear mixture model (model 8) for the thread elongation experiment.

80x105mm (300 x 300 DPI)

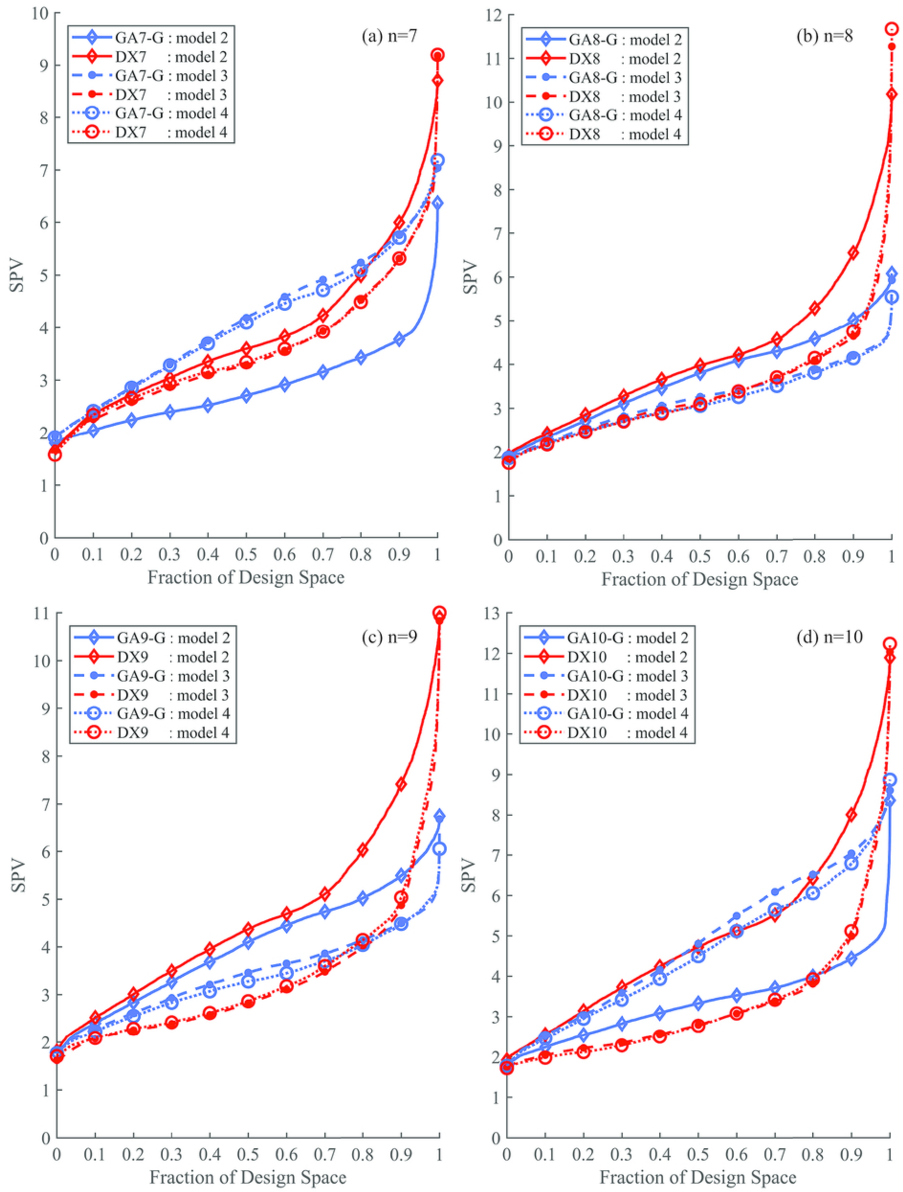

Figure 3. FDS plots for comparing GA designs (blue) and DX designs (red) for reduced Scheffé mixture models 2, 3, and 4 for the thread elongation experiment.

80x106mm (300 x 300 DPI)

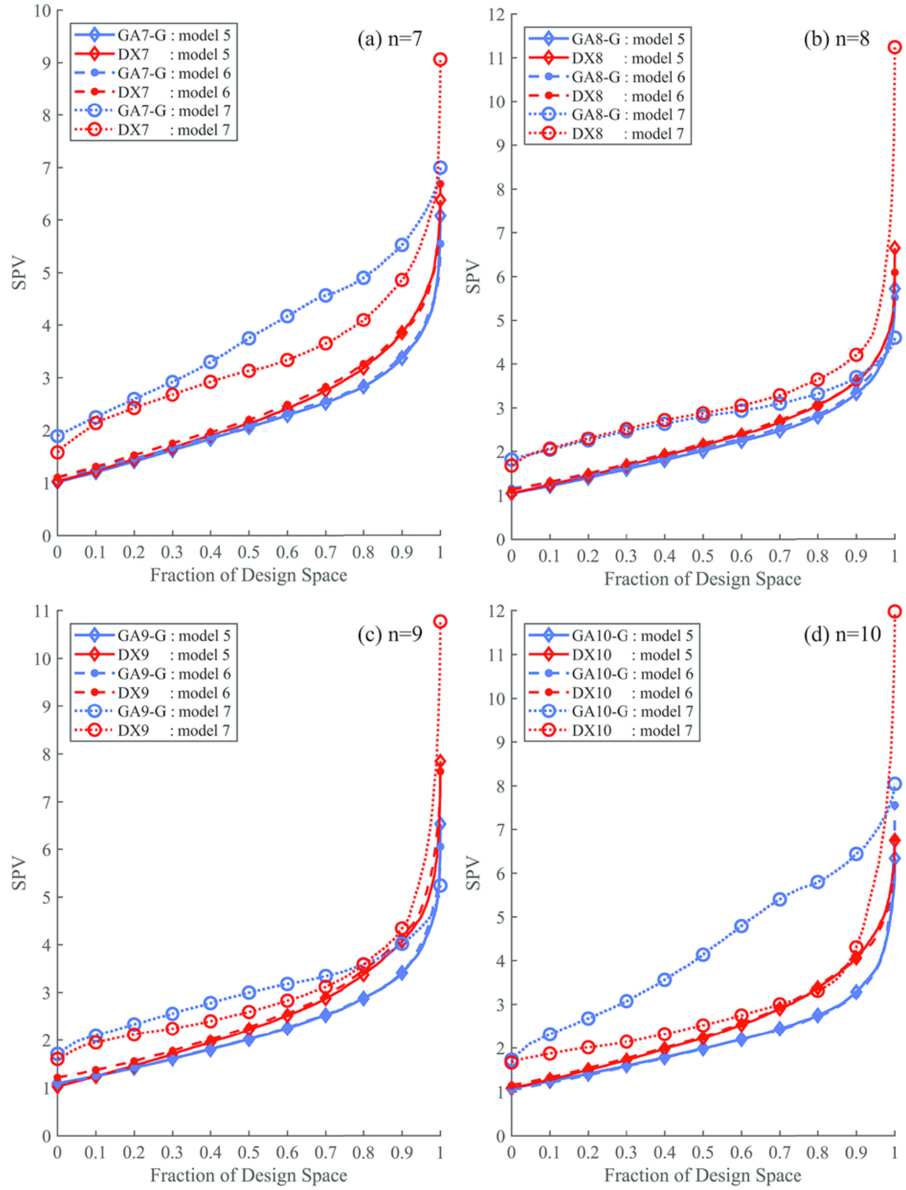

Figure 4. FDS plots for comparing GA designs (blue) and DX designs (red) for reduced Scheffé mixture models 5, 6, and 7 for the thread elongation experiment.

80x106mm (300 x 300 DPI)