Paper 209-30

Mixture Experiments and Their applications in Agricultural Research

K. Bondari
Experimental Statistics, Coastal Plain Station, University of Georgia, Tifton, GA 31793-0748

ABSTRACT

Statistical techniques used to analyze data from Mixture Experiments involve fitting Multiple Regression models with the intercept set to zero, Canonical Analysis to determine the shape of the fitted response, and the Response Surface System. One can use Mixture Design Macros (e.g., ADXMIX File) in SAS or the point-and-click ADX Interface of SAS to analyze data and obtain graphs. Contour plot, the residual plot showing the deviations of the observed values from their predictions, and Box-Cox transform plot, etc. would also be provided by the ADX Interface of SAS. Two experiments with 3-component (X1, X2, and X3) Mixture Design consisting of equally spaced proportions of 0, 0.5, and 1 (designated as {3, 2} Design) and 0, 1/3, 2/3, and 1 (designated as {3, 3} Design), and a 4-component mixture design (X1, X2, X3, and X4) will be considered to show analyses and interpretations of the results. The Simplex Lattice Design using all possible blends of these 3 components will create 6 different mixtures for {3, 2} Design and 10 different mixtures for {3, 3} Design. These design points are the blending coordinates arranged such that all combinations of factor levels are tested. The Simplex-Centroid Design will be applied to the 4-component design and to the augmented 3-component design. One purpose of statistical modeling in a Mixture Experiment is to model the blending surface such that predictions of the response for any mixture component, singly or in combination, can be made empirically. Testing of the model adequacy will also be an important part of the statistical procedure in a Mixture Experiment.

Keywords: ADX Interface, Simplex Lattice, Mixture experiment, Simplex Centroid.

INTRODUCTION

Research in many disciplines frequently involves blending two or more ingredients together. The design factors in a mixture experiment are the proportions of the components of a blend, and the response variables vary as a function of these proportions making the total and not the actual quantity of each component. The total amount of the mixture is normally fixed in a mixture experiment and the component settings are proportions of the total amount. The component proportions in a mixture experiment cannot vary independently as in factorial experiments since they are constrained to sum to a constant (1 or 100% for standard designs). Imposing such constraints on the component proportions complicates the design and the analysis of mixture experiments. Although the best known constraint in a mixture experiment is to set the sum of the components to one (100%), additional constraints such as imposing a maximum or minimum value on each mixture component may also apply.

Mixture experiments are commonly encountered in industrial product formulations (e.g., food processing, chemical formulations, textile fibers, pharmaceutical drugs) but the application in many fields of agricultural research is limited. Similar to industrial settings, a mixture experiment in agriculture may involve yield measurements of a crop due to applications of various mixtures of fertilizers or pesticides when the same amounts are applied making the yield response a function of the ingredient proportions and not the amounts included in the mixture. Many situations may exist in agriculture where an overall mixture response is more useful than the traditional individual responses (e.g., monoculture vs. multicultural cultivars, seed mixtures, soil mixtures, feed mixtures for animals), and mixture analysis techniques may have practical significance.

A researcher may conduct preliminary investigations to reduce the number of components to be included in a Mixture Experiment or to screen the mixture components to determine those components which are most important or effective in blending. The objective is to fit an appropriate mathematical model to express response variables as functions of the proportions of the mixture components. One purpose of statistical modeling in a Mixture Experiment is to model the blending surface such that predictions of the response for any mixture component, singly or in combination, can be made empirically. Testing of the model adequacy will also be an important part of the statistical procedure. The main objective of this study is to demonstrate the applicability of the mixture experiment in agricultural research. The use of the mixture methodology will be illustrated with a hypothetical example.

EXPERIMENT BACKGROUND

Consider two situations with a 3-component mixture experiment (X1, X2, and X3 are assumed to be 3 chemical compounds to be used for insect control): (1) X1, X2, and X3 consist of equally spaced proportions of 0, 0.5, and 1 designated as {3, 2} Design and (2) X1, X2, and X3 consist of equal proportions of 0, 1/3, 2/3, and 1 for each factor designated as {3, 3} Design. The Simplex Lattice using all possible blends of these 3 components (Table 1) shows 6 different mixtures for {3, 2} Design and 10 different mixtures for {3, 3} Design. These design points are the blending coordinates arranged such that all combinations of factor levels are tested. The sum for each run of the mixture is 1.0 and the component values are interpreted as proportions. The statistical design is referred to as {q, m} Simplex Lattice Design, where q represents the number of factors involved with m+1 equally spaced proportions from 0 to 1 for each component. All possible mixtures for {q=3, m=2} and {q=3, m=3} situations are shown in Table 1.

Table 1. Three-component (X1, X2, and X3) Simplex Lattice Design for 0, 0.5, and 1 designated as {3, 2} Design and for 0, 1/3, 2/3, and 1 designated as {3, 3} Design situations.

	{3, 2} Design			{3, 3} Design		
	X1	X2	X3	X1	X2	X3
Run 1	1	0	0	1	0	0
Run 2	0	1	0	0	1	0
Run 3	0	0	1	0	0	1
Run 4	0.5	0.5	0	1/3	2/3	0
Run 5	0	0.5	0.5	1/3	0	2/3
Run 6	0.5	0	0.5	0	1/3	2/3
Run 7				2/3	1/3	0
Run 8				2/3	0	1/3
Run 9				0	2/3	1/3
Run 10				1/3	1/3	1/3

Note that the design points in Runs 1-3 in Table 1 are various permutations of the pure blends consisting of only one component (single-component blends). Runs 4-9 are the permutations of the binary blends involving a mixture of 2 components and Run 10 is the multi-component mixture involving equal proportions of all 3 components (assuming equal importance). The 3 coordinates (0.5, 0.5, 0), (0, 0.5, 0.5), and (0.5, 0, 0.5) in Table 1 represent 3 simplex midpoints. The “Run” in Table 1 refers to the basic experimental unit with one setting for each of the factors of the experiment and for which a single value for the response will be observed. A Run (used in industrial experimentation) in agriculture may refer to a plot of land assigned to a particular treatment. It is recommended that the experimental Runs be replicated for the mixture experiments but it will not be necessary to collect replicate observations at all of the design points. One interesting aspect of these simplex-lattice designs is to determine how the 3 selected chemical components behaved singly and in combination in their effectiveness for insect control.

Analysis and modeling of mixtures of components that must sum to a constant (e.g., 1, 100%, or however else the mixtures are scaled) differ from the traditional approach and have been addressed by several investigators (see Cornell, 2002 for details). Triangular graphs are commonly used to summarize mixture proportions since the data points for valid mixtures frequently form a triangle in the 3D space in the scatter plot (with 2 components the simplex is a straight line). The simplex coordinates of a point in a triangular layout are determined by a line extended from each vertex to the side of the triangle (Fig. 1). Simplex coordinates of the 3 vertex points (corners of the triangle) representing a pure blend (consisting of only one component as in Runs 1-3 in Table 1) is 1 for a particular component and 0 for all other components (Fig. 1). However, including pure blends may not be practical in many applications of mixture experiments in agriculture. Fig. 1 also shows simplex coordinates for the center point (centroid) containing equal proportions of all 3 components (Run 10 in Table 1) and midpoints of the sides representing binary points (Runs 4-9 in Table 1). The interior points of the triangular layout represent all 3 components and are used by experimenters to augment the design.

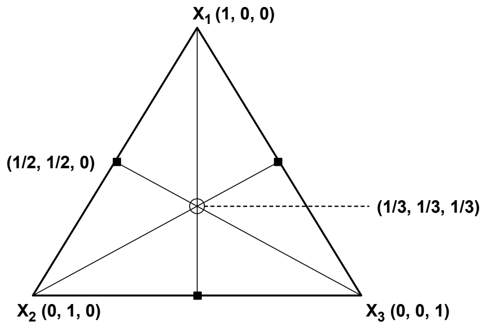


Fig. 1. Triangular layout for 3-component mixture

Standard designs for mixture experiments include Simplex-Lattice designs and Simplex-Centroid designs (introduced by Scheffe, 1963). A Simplex-Centroid Design is composed of mixtures with one component, two components in equal proportions, three components in equal proportions, to mixtures with all components present in equal proportions. Both design types evaluate points at vertices (corners of the triangle) and centroid (center of the triangle), and they are sometimes augmented by including additional interior points.

STATISTICS AND DATA ANALYSIS

Analysis and modeling of Mixture Experiments have been discussed by Cornell (2002) and SAS (2003). Statistical techniques used to analyze data from Mixture Experiments involve fitting Multiple Regression models with the intercept set to zero, Canonical Analysis to determine the shape of the fitted response, and the Response Surface System. One can use Mixture Design Macros (e.g., ADXMIX File) in SAS using SAS (2002-2003) or the point-and-click ADX Interface of SAS (see Fossceco, 1999; Ramirez and Ramirez, 2001; and SAS Publishing, 2004 for instructions) to analyze data and obtain graphs. The ADX Interface can be accessed from your desktop SAS System by selecting “Solutions”, “Analysis”, and “Design of Experiments” from the main SAS menu. Use the “Fit” button of the main ADX menu to fit the selected model and analyze the data. Several other software packages are available for analyzing data from mixture experiments which will not be discussed here.

Consider the following hypothetical study generating a set of data from 3 replications of the {3, 2} Design (Table 1) involving 3 chemical compounds (18 observations). The responses measured will be designated by R which is assumed to be the % of dead insects from applying these chemical treatments (Table 2). Each chemical component was applied individually and in combination with each of the other two chemical components. Further, assume that the 3 single-component blends (pure) and 3 binary blends were applied in 3 replications.

Table 2. Response (% dead insect) data from applications of 3 chemical compounds (X1, X2, and X3) in a mixture experiment

Run	Blend	Components			Observed Response (R)
		X1	X2	X3	
1	Pure	0.0	0.0	1.0	43.6
2	Pure	0.0	0.0	1.0	42.0
3	Pure	0.0	0.0	1.0	43.0
4	Binary	0.0	0.5	0.5	30.0
5	Binary	0.0	0.5	0.5	29.4
6	Binary	0.0	0.5	0.5	33.6
7	Pure	0.0	1.0	0.0	17.6
8	Pure	0.0	1.0	0.0	30.0
9	Pure	0.0	1.0	0.0	28.0
10	Binary	0.5	0.0	0.5	45.4
11	Binary	0.5	0.0	0.5	42.8
12	Binary	0.5	0.0	0.5	43.2
13	Binary	0.5	0.5	0.0	40.0
14	Binary	0.5	0.5	0.0	39.6
15	Binary	0.5	0.5	0.0	42.2
16	Pure	1.0	0.0	0.0	32.0
17	Pure	1.0	0.0	0.0	34.8
18	Pure	1.0	0.0	0.0	34.0

In this experiment, the researcher’s interests are assumed to be on pure blends (design points of the vertices which allow estimation of main effects) and binary blends (midpoints of the edges which allow estimation of 2-way interactions) of these 3 chemical compounds for insect control. In other words, the researcher wishes to investigate the insect control effectiveness of these 3 chemical compounds singly and in combination. The model of choice to represent the response as a function of the mixture variables is the low degree polynomials such as linear or quadratic polynomial equations.

RESULTS AND DISCUSSION

The canonical form of the 2nd degree polynomial fit to the data of Table 2 using the ADX Interface feature of the Mixture Experiment (SAS V9.1, 2002-2003) shows the following analyses of variance results (Table 3) and reveals the following predictive estimates for a quadratic polynomial fit to the data (Table 4). The 2 cross product terms in Table 3 (X1*X2 and X1*X3) are significant showing an adequate quadratic fit (R^2 = 86.12% and adjusted R^2 = 81.85%) for the predictive model. The interaction effect X2*X3 is not significant and thus was not included in the prediction equation. The ANOVA table shows sums of squares (SS), mean squares (MS), F values, and the significance probabilities for both master and predictive models (Table 3).

Table 3. Analysis of variance results using ADX Interface of SAS

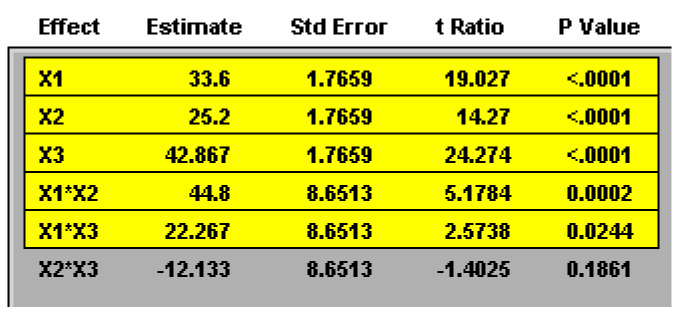
```
______________________________________________________________________________

                   Master Model                      Predictive Model
       ____________________________________    _____________________________

Source  DF    SS         MS        F     Pr > F    DF  SS      MS      F    Pr > F
______________________________________________________________________________

X1      1  3386.88   3386.88   362.02  0.0001    1 3386.88 3386.88 336.95 0.0001
X2      1  1905.12   1905.12   203.64  0.0001    1 2106.37 2106.37 209.56 0.0001
X3      1  5512.65   5512.65   589.24  0.0001    1 6306.79 6306.79 627.45 0.0001
X1*X2   1   250.88    250.88    26.82  0.0002    1  281.87  281.87  28.04 0.0001
X1*X3   1    61.98     61.98     6.62  0.0244    1   75.85   75.85   7.55 0.0166
X2*X3   1    18.40     18.40     1.97  0.1861

Model   5   829.08    165.82    17.72  0.0001    4  810.68  202.67  20.16 0.0001
Error  12   112.27      9.36                    13  130.67   10.05
(Lack of fit)                                    1   18.40   18.40   1.97 0.1861
(Pure Error)                                    12  112.27    9.36
Total  17   941.35                              17  941.35
______________________________________________________________________________
```

The prediction equation for the model is:

$$R=33.6\,X1 + 25.2\,X2 + 42.9X3 + 44.8X1X2 + 22.3X1X3.$$

Since the parameter estimate for X3 is greater than X1 and X2 (Table 4), it can be concluded that the X3 component will be the most effective single chemical in insect control. Additionally, since the estimated coefficients for X1*X2 and X1*X3 cross products are positive, blending chemicals X1 and X2 or X1 and X3 will be more effective in insect control than if used alone. Although the X2*X3 coefficient is not significant, indications are (a negative estimate in Table 4) that blending chemicals X2 and X3 may produce antagonistic effects for insect control. In general, positive interaction coefficients indicate synergism if higher responses are better and negative coefficients indicate antagonism. The inverse is true if lower responses are favored (see Bondari, 1999 for the discussion of interactions).

Table 4. Results of the quadratic fit using ADX Interface of SAS



Effect	Estimate	Std Error	t Ratio	P Value
X1	33.6	1.7659	19.027	<.0001
X2	25.2	1.7659	14.27	<.0001
X3	42.867	1.7659	24.274	<.0001
X1*X2	44.8	8.6513	5.1784	0.0002
X1*X3	22.267	8.6513	2.5738	0.0244
X2*X3	-12.133	8.6513	-1.4025	0.1861

Table 3 and Table 4 results can be verified by using SS III of PROC GLM of SAS with the intercept set to zero. For the data in Table 2, the following SAS statements should be included in the SAS program to analyze the data:

```
data chemical;
input run blend X1 X2 X3 R;
proc GLM;
model R=X1 X2 X3 X1*X2 X1*X3 X2*X3/noint solution;
run;
```

Data could be further explored by the Response Surface analysis of SAS:

```
proc rsreg;
model R=X1 X2 X3/lackfit;
run;
```

In this example, the analysis of variance has detected no significant difference between pure (Mean=33.9, SE of mean=2.56) and binary blends (Mean=38.5, SE of mean=2.56) in their effectiveness to control insects.

GRAPHIC PRESENTATION

The ADX Interface in SAS provides various plot options (located in the “Explore” button of ADX) including contour plots (Fig. 2) over the mixture design space. The contours of constant height are plotted on the 2D triangle in a contour plot to visualize the fitted model in triangular surface plots. Residual plots (the plots of observed vs. predicted values) are also provided by ADX and should be examined for outliers (Fig. 3).

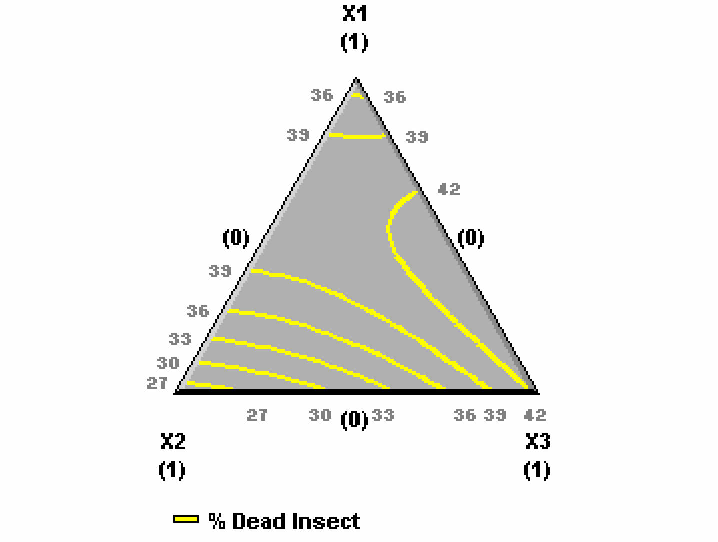


Fig. 2. Contour plot from ADX

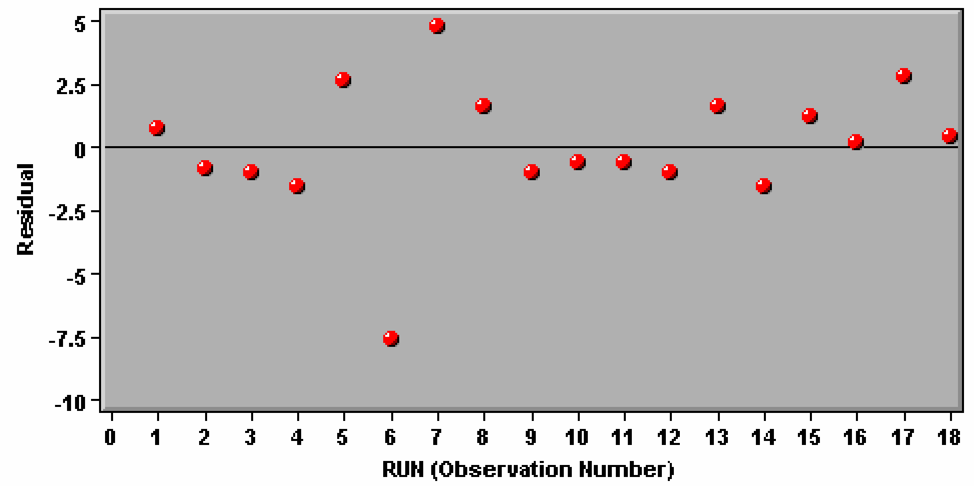
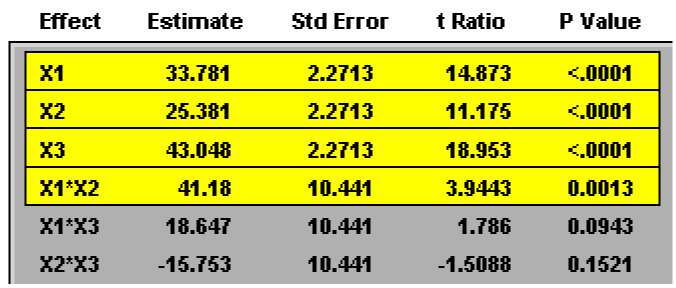


Fig. 3. Residual plot from ADX Interface of SAS

AUGMENTING THE DESIGN

Mixture designs are sometimes augmented by adding interior points. A center point (X1=0.333, X2=0.333, and X3=0.333) will be added to the design data (Table 2) with 3 replicate runs making 21 runs total) to augment the 3-component design selected for this study. This addition will change the design from simplex-lattice to simplex-centroid design which was created using the “create new design” option of ADX Interface of SAS. The triangular surface showing the coordinate points for this design resembles those of Fig. 1. Results of the analysis of variance (quadratic fit) using ADX Interface of SAS V9 indicates that effects of X1, X2, X3, and the X1*X2 interaction are significant and will be used in optimizing the response (Table 5). The X1*X3 cross product previously significant (Table 4) is no longer statistically significant (P=0.0932) in this new design.

Table 5. Results of the quadratic fit using ADX Interface of SAS



Effect	Estimate	Std Error	t Ratio	P Value
X1	33.781	2.2713	14.873	<.0001
X2	25.381	2.2713	11.175	<.0001
X3	43.048	2.2713	18.953	<.0001
X1*X2	41.18	10.441	3.9443	0.0013
X1*X3	18.647	10.441	1.786	0.0943
X2*X3	-15.753	10.441	-1.5088	0.1521

The fitted predictive model is:

$$R = 35.614\ X1 + 23.773\ X2 + 43.291\ X3 + 41.052\ X1{*}X2$$

and R^2 = 98.88% (adjusted R^2 = 98.68%) for the predictive quadratic model showing adequate fit. These estimates differ from those of Table 5 because the cross product terms X1*X3 and X2*X3 were not included in the prediction equation.

EXTENSION TO 4 COMPONENTS

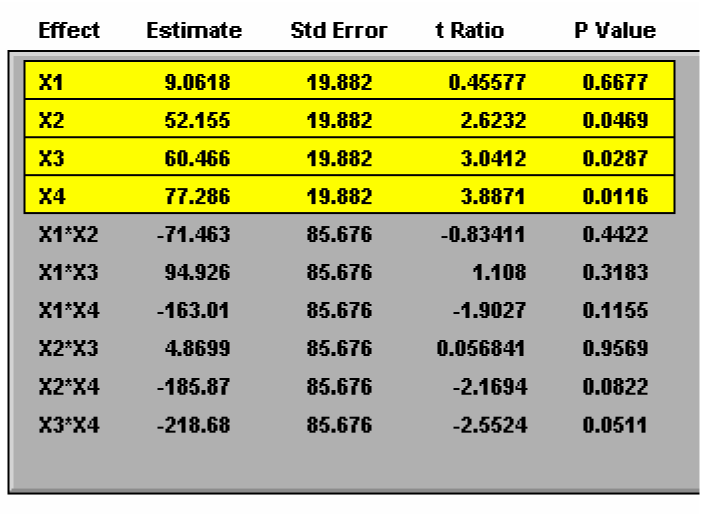
A new design (Simplex-Centroid from the Mixture Design) will be created for this experiment by selecting “File, Creating New Design” options of the ADX Interface. Again, the assumption is that a researcher wishes to examine 4 chemical compounds (X1, X2, X3, and X4) for their effectiveness (independently or in combinations) for insect control (Table 6). The % dead insect is determined as a response to these chemicals in this hypothetical experiment. Extension to 4 or more components can be accomplished in the ADX Interface by selecting the “Factor” tab of the “Define Variables” option of the ADX menu.

Table 6. Response (Resp) data from applications of 4 chemical compounds (X1, X2, X3, and X4) in a mixture experiment.

			Components				
Run	Blend	Type	X1	X2	X3	X4	Resp
1	Pure	Vertex	1	0	0	0	4.6
2	Pure	Vertex	0	1	0	0	51.8
3	Pure	Vertex	0	0	1	0	58.2
4	Pure	Vertex	0	0	0	1	78.0
5	Binary	Edge Centroid	0	0	0.5	0.5	10.8
6	Binary	Edge Centroid	0	0.5	0	0.5	7.2
7	Binary	Edge Centroid	0	0.5	0.5	0	58.4
8	Binary	Edge Centroid	0.5	0	0	0.5	7.8
9	Binary	Edge Centroid	0.5	0	0.5	0	75.8
10	Binary	Edge Centroid	0.5	0.5	0	0	22.4
11	Ternary	Face Centroid	0	1/3	1/3	1/3	45.0
12	Ternary	Face Centroid	1/3	0	1/3	1/3	6.2
13	Ternary	Face Centroid	1/3	1/3	0	1/3	5.8
14	Ternary	Face Centroid	1/3	1/3	1/3	0	23.2
15	All	Overall Centroid	1/4	1/4	1/4	1/4	2.6

The results of the analysis of variance (Table 7) indicate no significant interaction effects showing that the 4 components are most effective when used alone. The results further indicate that the possibility of antagonistic effects of X4 with the other 3 components may exist which requires further investigations by replicating the entire design or replicating some of the 15 runs of the experiment.

Table 7. Results of the quadratic fit using ADX Interface of SAS



Effect	Estimate	Std Error	t Ratio	P Value
X1	9.0618	19.882	0.45577	0.6677
X2	52.155	19.882	2.6232	0.0469
X3	60.466	19.882	3.0412	0.0287
X4	77.286	19.882	3.8871	0.0116
X1*X2	-71.463	85.676	-0.83411	0.4422
X1*X3	94.926	85.676	1.108	0.3183
X1*X4	-163.01	85.676	-1.9027	0.1155
X2*X3	4.8699	85.676	0.056841	0.9569
X2*X4	-185.87	85.676	-2.1694	0.0822
X3*X4	-218.68	85.676	-2.5524	0.0511

CONCLUSION

Mixture experiments are conducted to determine if mixed blends are more desirable than monocultures. Mixture experiments are appropriate to use when a researcher wishes to determine if synergism exists in mixing components which increases productivity or desirability of a product. A more general goal of a researcher might be to use mixture designs to determine which mixture or monoculture is most desirable, productive, or effective. An important property of the Mixture Experiment is that the change in the response depends on proportionality of the individual components present in the mixture and not on the amount of the mixture. Mixture experiments are commonly used by industrial experimenters (e.g., food industry, textile industry) and less in agriculture where factorial experiments are more common. However, many situations in agricultural research exist where the response varies as a function of the relative proportions and not the total amount of the components (e.g., research on formulations of pesticide and applications of fertilizers mixes) and thus mixture designs are more appropriate to use than factorials. Three-component and 4-component designs presented in this study illustrate how to apply mixture designs in agricultural research. The ADX interface in SAS was used to design experiments, fit statistical models, and explore responses using response surface designs. The ADX Interface of SAS is a point-and-click guided interface for the design, response optimization, graphical exploration, and statistical analysis of the mixture experiments in a PC environment. The ADX Interface of SAS is user friendly and reliable and could be very time saving for agricultural researchers.

REFERENCES

Bondari, K. 1999. Interactions in entomology: Multiple comparisons and statistical interactions in entomological experimentation. *Journal of Entomological Science* 34: 57-71.

Cornell, J. A. 2002. *Experiment with Mixtures: Designs, Models, and the Analysis of Mixture Data, 3rd Edition,* 680 pages, John Wiley & Sons, Inc., USA.

Fossceco, S. L. 1999. The SAS ADX Interface – designing the future. *Proceedings of the 24th Annual SAS Users Group International Conference (SUGI 24), Information Visualization Section:* 1-9.

Ramirez, B. and J. Ramirez. 2001. The ABCs of ADX: A solution for design and analysis of experiments. *Proceedings of the 26th Annual SAS Users Group International Conference (SUGI 26), Statistics, Data Analysis, and Data Mining Section:* 1-11.

SAS Institute Inc. 2002-2003. *SAS/STAT User's Guide, Version 9.1.* Cary, NC: SAS Institute Inc.

SAS Publishing. 2004. *Getting Started with the SAS 9.1 ADX Interface for Design of Experiments.* SAS Publishing, 320 pages.

Scheffe, H. 1963. Simplex-centroid designs for experiments with mixtures. *Journal of Royal Statistical Society, Series B,* 25: 235-263.

CONTACT INFORMATION

Contact the author at:

K. Bondari
University of Georgia
Experimental Statistics, Tifton Campus
Tifton, GA 31793-0748
Work Phone: (229) 386-3385
Fax: (229) 391-2501
Email: Bondari2@uga.edu
Web: http://myprofile.cos.com/bondarik85