Journal of Scientific Research and Reports

Volume 29, Issue 11, Page 24-33, 2023; Article no.JSRR.108686
ISSN: 2320-0227

Modeling the Yield of *Glycine max* (L.) Merrill Using Mixture Process Variable Model within an Optimal Split-Plot Design

Samson W. Wanyonyi [a*] and Julius K. Koech [b]

[a] Department of Mathematics and Computer Science, Pwani University, Kilifi, Kenya.
[b] Department of Mathematics and Computer Science, University of Eldoret, Eldoret, Kenya.

Authors' contributions

This work was carried out in collaboration between both authors. Author SWW designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author JKK managed the analyses of the study and the literature searches. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JSRR/2023/v29i111805

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here:
https://www.sdiarticle5.com/review-history/108686

Received: 07/09/2023
Accepted: 13/11/2023
Published: 01/12/2023

Original Research Article

ABSTRACT

A mixture design has become famous in statistical modeling in a mixture process variable experiment owing to its usefulness in modeling the blending surface that predicts the response of any mixture empirical. The mixture blends included manure from cows, chickens, goats, and sheep while the process variable was seeding rate of *Glycine max* seeds and the pH of the soil. The effect of variety of the seed used was established through variation of seeds per acre with uniform application of organic and inorganic fertilizer. This study's main aim was to determine the best desirable split-plot design for performing the *Glycine max* experiment with the settings mixture-process variables. The split-plot design (SPD) was used to solve the problem of restricted randomization. It constituted a simplex centroid design (SCD) of four design points of mixture

---

**Corresponding author: E-mail: samsonwangila@gmail.com, s.wangila@pu.ac.ke;*

components and 2^2 factorial design with a central composite design (CCD) of the process variable. We formulated a new Scheffe model and the proposed design for SPD for the combined second-order mixture process variable model with CCD. We used the restricted maximum likelihood method to approximate values for P parameter models within the SPD. We also found the effect of mixture component at vertices of components of the mixture plus with interaction effect between mixture and process variable to have the highest impact on the growth and pod development of *Glycine max*. The optimum total yield of *Glycine max* for variety R184 and Blyvoor used in Bushel per acre was 180.53 and 219. 217, respectively on the Whole Plot with a pH of soil being 5.4. The mean response maximum optimum yield for the total number of pods per plant and seeds per pod of *Glycine max* were found to be 32.30 and 2.331, respectively. We recommend using SPDs in experiments involving mixture settings formulations to measure the interaction effects of both the mixture components and the processing conditions like a pH of the soil and seeding rate.

Keywords: Process variable; mixture design; simplex centroid design; split plot design; soil pH.

1. INTRODUCTION

A mixture design has become popular in statistical modeling in a mixture process variable (MPV) experiment owing to its usefulness in modeling the blending surface that predicts the response of any mixture empirical [1-4]. In MPV, the response is a function of the mixture part proportion and the process variable. The explanatory variable and response in mixture experiments are dependent only on the relative proportion of the mixture ingredient, not on the mixture's volume. [5,6]. Process variables are variables that do not make up a portion of the mixture in an experiment but influence the ingredients' blending properties when their levels are modified [1,7].

Glycine max is known as a leguminous seed. It contains a high percentage of high-quality protein (40-42) and oil (18-20%), and other nutrients such as calcium, iron, and glycine which helps to prevent diseases like heart disease, cancer, and a slew of others, according to Jackson (2016). It also improves soil fertility by fixing significant atmospheric nitrogen levels by root nodules and leaf fall on the ground at maturity.

Glycine max production in Kenya is still inadequate, according to Mahasi et al. [8], averaging 2000-5000 metric tons per year. However, due to functional and economic considerations, some process variables (noise variables) are challenging to modify in some cases. According to Goldfarb et al. [9], these limitations prevent complete randomization of the experimental runs. Incomplete randomization of experimental runs has become one of the leading causes of cereal crop yield declines. Many researchers [10-13] believe that other variables are also to blame if they are not well considered, such as seed row spacing, seeding rates, soil nutrient management strategies, soil pH. Some factors mentioned above, such as soil pH, row spacing, and seeding rate, are examples of the process variables affecting the optimum yield of the crop if not well managed [14].

In Africa, natural soil fertility is addressed by applying nitrogen, phosphorus, and potassium fertilizers at low rates. There is always a generally expected response of Cereals to NPK fertilizer application at current recommendations. However, the response remains far below the potential level, particularly on-farm due to nutrient deficiencies and imbalances. The predicted responses of *Glycine max* to N, P, and K, as well as the scale of macronutrient (N, P, and K) and micronutrient (Zinc) deficiencies, have been studied and reported in Kenyan Soybean (*Glycine max*) growing areas. However, there has been little investment in research to determine the best method of combining mixture components with simplex centroid design (SCD) of organic fertilizers derived from livestock manure within split-plot design (SPD) using a 2^k factorial configuration with a central composite design (CCD) of the process variable. As a result, this study evaluates the impact of MVP design on *Glycine max* production using farm trials in a SPD. The SPD is used to solve the problem of restricted randomization on mixture-process variable layouts in this case.

2. METHODOLOGY

2.1 Data Source

The data was primarily collected from the field of experiment. The data consists of two response measurements obtained from *Glycine max* (L.) Merrill. The two responses measured include the number of entire pods per plant (η_1) and seeds per pod(η_2). The mixture settings included four components x_1, x_2, x_3, and x_4 , derived from different organic matter varieties, which represent goat manure, cow manure, chicken manure, and sheep manure, respectively. The mixtures were the subplots and process variables the whole plots. The model for the fixed part of this experiment is represented in model (1).

2.2 Description of Experimental Sites

The study was conducted in Spande and Munge's villages in Kakamega County, Mautuma Ward, Lugari Sub-District, and Western Kenya. Both sites are about 8 km apart. The two regions lie between $(0.706373^{0}\ N, 35.0722^{0}\ E)$ and $(0.695366^{0}\ N, 35.028022^{0}\ E)$, with an elevation of between 1800 and 1900 m above sea level, respectively. The region receives bimodal rains with an annual mean precipitation of about 1971 mm, and an annual mean temperature of about $20.4^{0}c$, as reported by Althof [15]; Mbau et al. [16]. Additionally, prolonged rain usually occurs between April and July, while short precipitation occurs between August and December, as described by Mbau et al. [16]. Further, the reliability growth period for *Glycine max* (L.) Merrill lies between 75 and 140 days [17].

Further, as Isaev et al. [18], the best period for sowing Glycine max is when the temperature in the $0-10\ cm$ layer of soil is about $12-14^{0}\ c$. According to Tsikhungu et al. [19], the Lugari sub-county grounds are predominantly well-drained deep red to dark, sandy loams to sandy clays that are not very fertile. Still, well-drained soils with moderately to slight condition with soil pH, lie between 5.3 to 5.9. However, some part of this region contains low inherent fertility as evidenced by low amounts of Nitrogen, soil organic carbon and exchangeable base as described by Ayuke et al. [20]. The experimental site encompasses farmlands adjacent to the Lugari forest. The area was initially inhabited by a sparse population of former forest residence communities who practiced shifting cultivation, hunting, and gathering. The study sites have a settlement history dating more than a hundred years with relatively intensive sedentary mixed subsistence agriculture as reported by Kimetu et al. [21] for over the last sixty years. Pender et al. [22] found that landholding per household has reduced drastically because of the high population growth rate and immigration into the area. Currently, most agricultural land is characterized by low soil fertility, low crop yields, and low farm income [20]. However, cereals (maize), legumes (beans), and sugarcane have become the primary crops, with most fields described by Mbau et al. (2006).

2.3 Method of Analysis

We formulated a second order Scheffe polynomial model within SPD as shown in model (1) and Fig. 1 by extending the model proposed by Kowalski et al. [23], Cho [2] and Njoroge et al. [3] from 3 to 4 mixture blends in the presences of two process factors as in Equation (1). The process variable in the model were Z_1 and Z_2 where Z_1 the control is process variable (seeding rate (seeds per acre) at a constant row spacing of *Glycine max* seed) and Z_2 is the noise process variable (soil pH).

$$\begin{aligned} Y(x,z) = {} & \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{14} X_1 X_4 + \beta_{23} X_2 X_3 \\ & + \beta_{24} X_2 X_4 + \beta_{34} X_3 X_4 + \vartheta_{11} X_1 Z_1 + \vartheta_{12} X_1 Z_2 + \vartheta_{21} X_2 Z_1 + \vartheta_{22} X_2 Z_2 \\ & + \vartheta_{31} X_3 Z_1 + \vartheta_{32} X_3 Z_2 + \vartheta_{41} X_4 Z_1 + \vartheta_{42} X_4 Z_2 + \varrho + \varepsilon \end{aligned} \quad (1)$$

Where β_i is the vector of fixed effect resulting from mixture blend of the vertices of component X_i, β_{ij} is the vector of random effect resulting from the interaction between mixture components, ϑ_{ij} is the vector of random effect resulting from the interaction between mixture components and process factors, $\varrho \sim N(0, \sigma_\delta^2)$, represent the random error associated with the whole-plot factor by itself during the randomization level, and $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ indicate the random error that is associated with sub plot randomization level. However, σ_δ^2 and σ_ε^2 are assumed to be statistically independent and distributed. The model 1 was formulated based on the design shown in Fig. 1. Fig. 1 depicts the newly generated design for SPD to Support Fitting the Combined Second-Order MPV model where the center point $[z_1, z_2] = \{0,0\}$, and v, k is the number of times that treatment combination is replicated.

The model (1) is an empirical model that corresponds well with the experience and plots of the data. The random component effect of the model has a whole plot and split-plot contribution. The whole plot error is nested under x_1, x_2, x_3, and x_4, while the subplot error is the standard residual error term. The model was analyzed using restricted maximum likelihood described in Njoroge et al. [3] to account for the split-plot random structure.

However, the model (1) under split plot design was further simplified to

$$Y_{jk} = X_{jk}\beta + d_{jk}\delta_j + \varepsilon_{jk} \qquad (2)$$

here Y_{jk} represents whole plot j at k^{th} measurement response variable subject to split-plot factors and process variable. n_w denotes the number whole plot while n_j number of measurements in whole plot j. d_{jk} indicates a covariate vector of j^{th} whole plot at k^{th} measurement for random effects $\delta_j \in \mathbb{R}^q$ associated with whole plot effect where q is the number of factor components applied in split plot layout experiment.

Fig. 1 shows the same design produced in table form shown Table 1. The design in Table 1 was generated using the candidate set free algorithm described in Jones and Goos [24] and implemented in JMP version 15.1 [25] based on the design shown in Fig. 1. However, the split plot structure constituted a simplex centroid design (SCD) of four design points of mixture components and 2^2 factorial design with a central composite design (CCD) of the process variable.

2.4 Composting Farmyard manure through the framework of Mixture design

Four agro-organic wastes commonly found in the test sites were selected for our study. The selection criteria for the required farm manure (FYM) obtained from livestock were based on the region's availability of material. FYM includes goat manure, poultry manure, sheep manure, and animal manure derived from the Spande farm. Compound composting was done using the pit method using the structure of the mixture design under control

$$x_1 + x_2 + x_3 + x_4 = 1,$$

in reference to the design in Table 1. Composting was done in line to the literature using Pit method where each pit was measuring $3 \times 3 \times 2\ m$ [26], Mbau et al. 2008.

Table 1. Shows the MPV settings in the context of SPD

Run	Whole plot	X_1	X_2	X_3	X_4	Z_1	Z_2
1	1	0.25	0.25	0.25	0.25	-1	1
2	1	0	0	0	1	-1	1
3	1	0	1	0	0	-1	1
4	1	0.25	0.25	0.25	0.25	-1	1
5	1	0	0	1	0	-1	1
6	1	1	0	0	0	-1	1
7	2	0.25	0.25	0.25	0.25	1	-1
8	2	0.25	0.25	0.25	0.25	1	-1
9	2	0	0	1	0	1	-1
10	2	0	1	0	0	1	-1
11	2	0	0	0	1	1	-1
12	2	1	0	0	0	1	-1
13	3	0.5	0.5	0	0	1	1
14	3	0.5	0	0.5	0	1	1
15	3	0.5	0	0	0.5	1	1
16	3	0	0.5	0	0.5	1	1
17	3	0	0.5	0.5	0	1	1
18	3	0	0	0.5	0.5	1	1
19	4	0	0.5	0.5	0	-1	-1
20	4	0.5	0.5	0	0	-1	-1
21	4	0	0	0.5	0.5	-1	-1
22	4	0.5	0	0.5		-1	-1
23	4	0.5	0	0	0.5	-1	-1
24	4	0.25	0.25	0.25	0.25	-1	-1
25	5	1	0	0	0	0	0
26	5	0	1	0	0	0	0
27	5	0	0	1	0	0	0
28	5	0	0	0	1	0	0
29	5	0.25	0.25	0.25	0.25	0	0
30	5	0.25	0.25	0.25	0.25	0	0
31	6	0.5	0	0.5	0	0	1
32	6	0.5	0.5	0	0	0	1
33	6	0.5	0	0	0.5	0	1
34	6	0	0.5	0.5	0	0	1
35	6	0	0.5	0	0.5	0	1
36	6	0	0	0.5	0.5	0	1
37	7	0.5	0	0.5	0	0	-1
38	7	0.5	0.5	0	0	0	-1
39	7	0.5	0	0	0.5	0	-1
40	7	0	0.5	0.5	0	0	-1
41	7	0	0.5	0	0.5	0	-1
42	7	0	0	0.5	0.5	0	-1
43	8	0.25	0.25	0.25	0.25	-1.414	0
44	8	0.25	0.25	0.25	0.25	-1.414	0
45	8	0.25	0.25	0.25	0.25	-1.414	0
46	8	0.25	0.25	0.25	0.25	-1.414	0
47	8	0.25	0.25	0.25	0.25	-1.414	0
48	8	0.25	0.25	0.25	0.25	-1.414	0
49	9	0.25	0.25	0.25	0.25	1.414	0
50	9	0.25	0.25	0.25	0.25	1.414	0
51	9	0.25	0.25	0.25	0.25	1.414	0
52	9	0.25	0.25	0.25	0.25	1.414	0
53	9	0.25	0.25	0.25	0.25	1.414	0
54	9	0.25	0.25	0.25	0.25	1.414	0

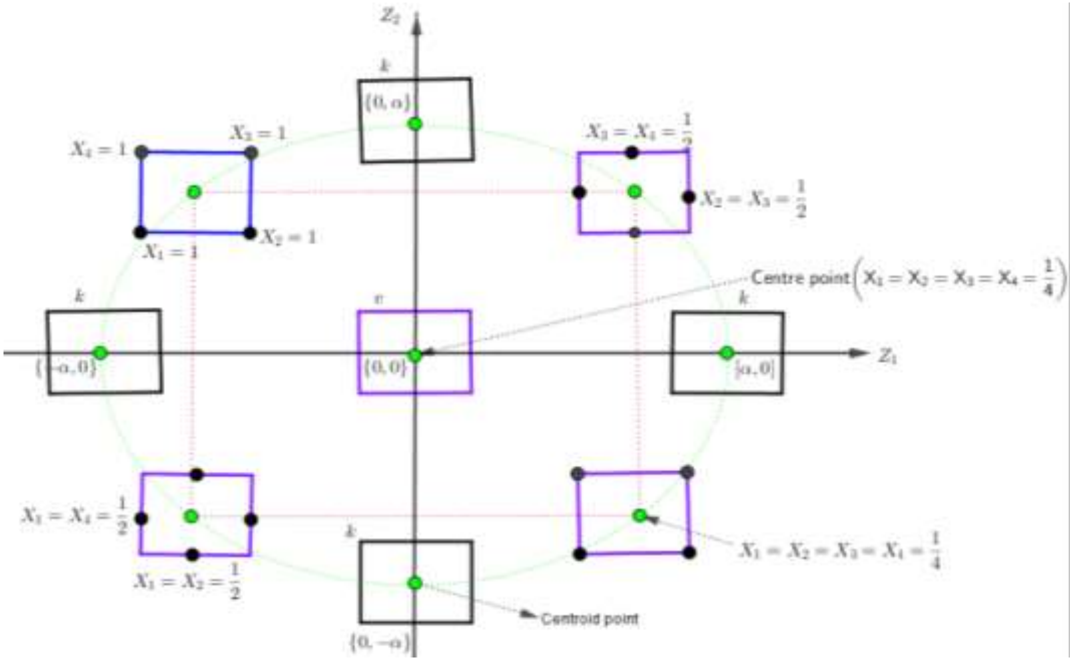

Fig. 1. Shows a newly developed design for split-plot layout for combined 2nd order MPV with CCD

2.5 Treatment combinations Manure in the context of MPV within SPD

Field trials were performed on two farms. The experiment was carried out using a well randomized complete block in a split-plot arrangement with replication, as shown in Table 1. The split-plot structure comprised nine whole plots, with each field having six sub-plot treatments. Each plot's plot size was $95.5\ ft \times 170\ ft$, while each experimental subplot unit was $15.5\ ft \times 50\ ft$. Split plot treatments were applied based on the proposed design in Fig. 1 using composite compost manure from the 11 compost pits. There were four lime treatments ($0, 1.7, 5, \text{and } 15 - \text{ton } aglime\ /acre$) being applied to 9 main plots with at least twice at axial part as shown below with correspondence of soil pH obtained after the application.

After initial testing of soil pH at the farm was 5.4, we prepared five different soils for selected plant growth at optimal pH values, as shown in Table 2. pH chosen deals from the initial 5.4 pH of the soil using a control method according to the literature [27]:

Table 2. Showing scaled seeding rate $(\mathbf{Z_1})$ and soil pH $(\mathbf{Z_2})$ according to 2^2 factorial design with CCD

Whole plot	Lime application (tons/ acre)	Un coded Z_1	Un coded Z_2	Coded Z_1	Coded Z_2
1	15	125000	7.0	-1	1
2	1.7	225000	6.0	1	-1
3	15	225000	7.0	1	1
4	1.7	125000	6.0	-1	-1
5	5	175000	6.5	0	0
6	15	175000	7.0	0	1
7	1.7	175000	6.0	0	-1
8	0	100000	5.4	-1.414	0
9	0	275000	5.4	1.414	0

The whole plots consisted of five primary seeding rates (100000, 125000, 175000, 225000, and 275 000 seeds per acre) applied to sub-plot experimental units taking into account row spacing whole plot as shown in Table 2. According to the literature review, we used the seeding rate and row spacing [27]. The *Glycine max* varieties were R 184, and Blyvoor were planted on April 24th, 2020. The seeds were first inoculated with *Bradyrhizobium Japonicum*, and each subplot was grown using row spacing specified in Table 2 and 1- inch depth. Plots were harvested on August 27th, 2020. Grain yields obtained were then adjusted to 13 % moisture.

3. RESULTS AND DISCUSSION

The estimate, standard errors, t values and p values of the fitted Scheffe model for the total number of pods of Glycine max per plant stem observed. The fitted Scheffe model is therefore,

$$\widehat{Y} = 21.8741\, x_1 + 23.3074 x_2 + 23.5741 x_3 + 21.54074\, x_4 + 39.36272\, x_1x_2 + 40.8294\, x_1x_3 + 47.5961\, x_1x_4 + 38.9627\, x_2x_3 + 41.6515\, x_2x_4 + 38.7299\, x_3x_4 + 2.9268\, x_1z_1 + 1.4991\, x_2z_1 + 1.3146 x_3z_1 + 2.5141 x_4z_1 + 1.7823\, x_1z_2 + 2.8601\, x_2z_2 + 2.1625\, x_3z_2 + 2.1221\, x_4z_2 \quad (3)$$

The significant factors were x_1, x_2, x_3, x_4, $x_1\, x_2$, x_1x_3, x_1x_4, x_2x_3, x_3x_4 and x_1z_1 as shown in Table (3). Thus, the final model is

$$\hat{Y} = 21.8741\, x_1 + 23.3074 x_2 + 23.5741 x_3 + 21.54074\, x_4 + 39.36272\, x_1x_2 + 40.8294\, x_1x_3 + 47.5961\, x_1x_4 + 38.9627\, x_2x_3 + 41.6515\, x_2x_4 + 38.7299\, x_3x_4 + 2.9268\, x_1z_1 \quad (4)$$

Table 3. Shows the t student test for the fitted scheffe model for the total number of pods of glycine max per plant stem using REML analysis

Term	Estimate	Standard Error	t value	p value
x_1	21.8741	1.2191	17.94	0.0001
x_2	23.3074	1.2191	19.12	0.0001
x_3	23.5741	1.2191	19.34	0.0001
x_4	21.54074	1.2191	17.67	0.0001
x_1x_2	39.36272	2.7087	14.53	0.0001
x_1x_3	40.8294	2.7087	15.07	0.0001
x_1x_4	47.5961	2.7087	17.57	0.0001
x_2x_3	38.9627	2.7087	14.38	0.0001
x_2x_4	41.6515	3.0267	13.76	0.0001
x_3x_4	38.7299	2.6341	14.70	0.0001
x_1z_1	2.9268	1.2446	2.35	0.0420
x_2z_1	1.4991	1.2580	1.19	0.2615
x_3z_1	1.3146	1.2454	1.06	0.3175
x_4z_1	2.5141	1.2503	2.01	0.0433
x_1z_2	1.7823	1.3687	1.30	0.2299
x_2z_2	2.8601	1.3801	2.07	0.0516
x_3z_2	2.1625	1.3679	1.58	0.1535
x_4z_2	2.1221	1.3743	1.54	0.1614

The estimate, standard errors, t values and p values of the fitted Scheffe model for the total number of seeds per pod of Glycine max per plant stem observed. The fitted Scheffe model is therefore,

$$\hat{Y} = 2.0597\, x_1 + 2.0597 x_2 + 2.0597 x_3 + 2.0597\, x_4 + 1.2119\, x_1 x_2 + 1.1119\, x_1 x_3 + 1.1119\, x_1 x_4 + 1.2119\, x_2 x_3 + 1.2207\, x_2 x_4 + 1.0695\, x_3 x_4 + 0.0719\, x_1 z_1 + 0.0426\, x_2 z_1 + 0.0634 x_3 z_1 + 0.0795 x_4 z_1 + 0.0431\, x_1 z_2 + 0.0097\, x_2 z_2 + 0.0462\, x_3 z_2 + 0.0492\, x_4 z_2 \tag{5}$$

The significant factors were x_1, x_2, x_3, x_4, $x_1 x_2$, $x_1\, x_3$, $x_1\, x_4$, $x_2\, x_3$, $a_3\, x_4$, $x_1 z_1$ and $x_4 z_1$ as shown in Table (4). Thus, the final model is

$$\hat{Y} = 2.0597\, x_1 + 2.0597 x_2 + 2.0597 x_3 + 2.0597\, x_4 + 1.2119\, x_1 x_2 + 1.1119\, x_1 x_3 + 1.1119\, x_1 x_4 + 1.2119\, x_2 x_3 + 1.2207\, x_2 x_4 + 1.0695\, x_3 x_4 + 0.0719\, x_1 z_1 + 0.0795 x_4 z_1. \tag{6}$$

Table 4. shows the t student test for the fitted Scheffe model for the Number of seeds per pod of Glycine max per plant stem using REML analysis

Term	Estimate	Standard Error	t value	p value
x_1	2.0597	0.03460	59.52	0.0001
x_2	2.0597	0.03460	59.52	0.0001
x_3	2.0597	0.03460	59.52	0.0001
x_4	2.0597	0.03460	59.52	0.0001
$x_1 x_2$	1.2119	0.10790	11.23	0.0001
$x_1 x_3$	1.1119	0.10790	10.30	0.0001
$x_1 x_4$	1.1119	0.10790	10.30	0.0001
$x_2 x_3$	1.2119	0.10790	11.23	0.0001
$x_2 x_4$	1.2207	0.12058	10.12	0.0001
$x_3 x_4$	1.0695	0.10491	10.19	0.0001
$x_1 z_1$	0.0719	0.03396	2.12	0.0491
$x_2 z_1$	0.0426	0.03473	1.23	0.2352
$x_3 z_1$	0.7634	0.03400	1.86	0.0546
$x_4 z_1$	0.0795	0.03429	2.32	0.0326
$x_1 z_2$	0.0431	0.03511	1.23	0.2434
$x_2 z_2$	0.0097	0.03581	0.27	0.7916
$x_3 z_2$	0.0462	0.03505	1.32	0.2120
$x_4 z_2$	0.0492	0.03545	1.39	0.1898

The results shown in Table 3 and 4 were obtained by using a REML. The results in both tables clearly shows that x_1, x_2, x_3, x_4, and all the interaction x_1, x_2, x_3 and x_4 are all significant and have a great impact on the number of pods on the main stem per plant. Also, the interaction between the process variable and mixture component factor $x_1 z_1$ and $x_4 z_1$ are significant at 5%. However, we can also observe that the interaction between $x_3 z_1$, and $x_2 z_2$ are almost significant at the same level with $p = 0.0546$ and 0.0516, respectively. Further, this is indicating a possible effects of the mixture process variable interaction resulting from soil pH (z_2) and the number of seeds used per acre (z_1). The whole plot error variance and sub-plot error variance was found to be 0.0048252 and 0.0018023, respectively corresponding to 2.677 Variance ratio $\left(\boldsymbol{d} = \frac{\sigma_\delta^2}{\sigma_\varepsilon^2}\right)$ with Wald $\boldsymbol{P}-$ value= 0.109. This shows that random effect resulting from MPV interaction was not significant at 5% level and therefore, restricted randomization was completely solved with SPD.

The averagely adjusted R^2 from Table 5 shows that 96.83% of the variation in the response was explained by the model. The result shows clearly that the model fits the data well for the two responses. Also, the results indicate that the second-order MPV model (1) formulated adequately represents the growth and pod development of *Glycine max*. We can also observe that the model has a reliability of 96.83% on averagely which can also provide some vital information regarding germination of *Glycine max*. The result also shows that average number of pods per plant and number of seeds per pod is 32.30 and 2.331, respectively.

The Table 6 shows the predicted Glycine max yield per acre in Bushels for each variety basing on the yield obtained from each whole plot in terms $\boldsymbol{\eta}_0, \boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ that represents plant per

Table 5. Shows the summary fit of the two responses obtained using MPV setting model structure

Summary of fit	Pods per plant	Seeds per pod
Multiple R^2	0.9857	0.9711
Adjusted R^2	0.9790	0.9575
Mean response	32.30	2.331

Table 6. Estimation of Glycine max yield of variety two (Blyvoor and R 184) in Bushel per acre max

Whole plot	η_0	η_{1B}	η_{2B}	η_{1R}	η_{2B}	Bushels per acre for Blyvoor	Bushels ($1\ bushel = 25.4\ Kg$) per acre for R 184
1	125000	32.2	2.4	31.2	2.3	64.4	59.8
2	225000	32.6	2.4	31.6	2.3	117.36	109.02
3	225000	37.9	2.5	36.9	2.4	142.13	132.84
4	125000	29	2.3	28	2.3	55.58	53.67
5	175000	33.2	2.4	32.2	2.3	92.96	86.4
6	175000	35.1	2.4	33.9	2.3	98.28	90.97
7	175000	30.5	2.3	29.5	2.3	81.84	79.16
8	100000	40.4	2.5	39.4	2.5	67.33	65.67
9	275000	46	2.6	45	2.7	219.27	180.53

acre, Pods per plant and seed per pod, respectively. On the other hand, the subscript B and R denotes the variety Blyvoor and R 184. We used the formula described in Chad Lee and Jim (2005) to estimate anticipated total yield per acre in Bushel where one bushel of *Glycine max* (L.) Merrill weighs 60 pounds. The result shown in Table 6 indicates the *Glycine max* growth and pod development increase with the application of MPV settings used. Averagely, the variety Blyvoor does well as compared to variety R 184. The result also shows that the maximum *Glycine max* yield of the two varieties is directly proportional to the number of plant per acre. In addition, the result also indicates that the variety of the seed used has also the impact on the optimum yield obtained.

4. SUMMARY, CONCLUSION AND RECOMMENDATIONS

We formulated the proposed design for a split-plot layout structure for the combined second-order mixture process variable model with CCD. We used the restricted maximum likelihood method to approximate values for P parameter models within the SPD. We also found the effect of mixture component at vertices of component x_1, x_2, x_3 and x_4 to have the highest impact on the growth and pod development of *Glycine max* together with permutation interaction of these mixture components at 5 % significance level. The two-process variable used included various seeding rate of two varieties of R 184 and Blyvoor and soil pH. The variety Blyvoor was found to perform better than variety R 184 in terms of the yield of seeds harvested and the same condition mixture setting and pH of soil as evidenced in Tables 5. The optimum total yield of *Glycine max* for variety R184 and Blyvoor in Bushel per acre was 180.53 and 219. 217, respectively on the 9^{th} Whole Plot with a pH of soil being 5.4. The two source errors that result from main treatments (Whole plots) and a subplot treatment for each of the two responses were measured and fitted using the same model formulated. We found the average whole plot error variance (σ_δ^2) and average subplot (split-plot) error variance (σ_ε^2) to be 0.0048252 and 0.0018023, respectively corresponding to 2.677 Variance ratio $\left(\boldsymbol{d} = \frac{\sigma_\delta^2}{\sigma_\varepsilon^2}\right)$ with Wald $\boldsymbol{P}-$ value $=$ 0.109. Normally, the Whole Plot error variance larger than the split-plot (Residual) error variance, as shown by Box and Jones (1992). Therefore, this implies that the model adequately represented the mixture data collected from the field, and also, restricted randomization was completely solved with the SPD layout. The least square mean response maximum optimum yield for the total number of pods per plant stem of *Glycine max* was 41.083. We thoroughly investigated the interpretation and estimation of parameters from MPV settings in conjunction with CCD within SPD in the context of the Scheffe polynomial. We recommend using SPDs in experiments involving mixture settings formulations to measure the interaction effects of both the mixture components and the processing conditions like a pH of the soil and seeding rate.

DISCLAIMER

This paper is an extended version of a Thesis document of the same author.

The Thesis document is available in this link: http://41.89.164.27/handle/123456789/1037

[As per journal policy, preprint /repository article can be published as a journal article, provided it is not published in any other journal]

ACKNOWLEDGEMENT

I thank all my co-authors for having made this research successful. I also take this opportunity to thank Sevu Mueni from Kenya Agriculture and Livestock Organization, Food crop Research institute KALRO Alupe, Busia, Kenya for her great effort contribution in soil sample analysis and interpretation. Further, I give our deepest gratefulness to almighty God for giving us good health during this entire period of research.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Anderson-Cook CM, Goldfarb HB, Borror CM, Montgomery DC, Canter KG, Twist JN. Mixture and mixture–process variable experiments for pharmaceutical applications. Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry. 2004;3(4):247-260.
2. Cho TY. Mixture-process variable design experiments with control and noise variables within a split-plot structure. Arizona State University; 2010.
3. Njoroge GG, Simbauni JA, Koske JA. An Optimal Split-Plot Design for Performing a Mixture-Process Experiment. Science Journal of Applied Mathematics and Statistics. 2017;5(1):15.
4. Sitinjak MA, Syafitri UD. A Split Plot Design for an Optimal Mixture Process Variable Design of a Baking Experiment. In Journal of PHysics: Conference Series. IOP Publishing. 2019;1417(1):012018.
5. Cornell JA. Analyzing data from mixture experiments containing process variables: a split-plot approach. Journal of Quality Technology. 1988;20(1):2-23.
6. Goldfarb HB, Borror CM, Mongomery DC. Mixture-process variable experiments with noise variables. Journal of Quality Technology. 2003;35(4):393-405.
7. Goos P, Donev AN. Tailor-made split-plot designs for mixture and process variables. Journal of Quality Technology. 2007;39(4):326-339.
8. Mahasi JM, Mukalama J, Mursoy RC, Mbehero P, Vanlauwe B. A Sustainable approach to increased soybean production in western Kenya. In 10th African Crop Science Conference Proceedings, Maputo, Mozambique. African Crop Science Society; 2011.
9. Goldfarb HB, Borror CM, Montgomery DC, Anderson-Cook CM. Three-dimensional variance dispersion graphs for mixture-process experiments. Journal of Quality Technology. 2004;36(1):109-124.
10. Chad Lee, Jim Herbek. Kentucky estimating soybean yield, grain crops extension, plant and soil sciences. University of Kentucky, Kentucky Agricultural Experiment Station; 2005.
11. Lee CD, Egli DB, TeKrony DM. Soybean response to plant population at early and late planting dates in the Mid-South. Agronomy Journal. 2008;100(4):971-976.
12. Tittonell P, Vanlauwe B, Corbeels M, Giller KE. Yield gaps, nutrient use efficiencies and response to fertilisers by maize across heterogeneous smallholder farms of western Kenya. Plant and Soil. 2008;313(1-2):19-37.
13. Wanyama WE. Response of soybean *(Glycine max l.)* to application of inorganic fertilizers, cattle manure and lime in western Kenya (Doctoral dissertation); 2013.
14. Pedersen P. Soil pH and plant population effects on soybean yield. Iowa State University Research and Demonstration Farms Progress Reports. 2004;(1).
15. Althof AJ. Human impact on flora and vegetation of Kakamega Forest, Kenya; 2005.
16. Mbau SK, Karanja N, Ayuke F. Short-term influence of compost application on maize yield, soil macrofauna diversity and abundance in nutrient deficient soils of

Kakamega County, Kenya. Plant and soil. 2015;387(1-2):379-394.
17. Jaetzold R, Schmidt H, Hornetz B, Shisanya C. Farm management handbook of Kenya Vol. II: Natural conditions and farm management information Part C East Kenya Subpart C1 Eastern Province. Cooperation with the German Agency for Technical Cooperation (GTZ); 2006.
18. Isaev S Kh, Kadirov ZZ, Khamraev K Sh, Atamuradov BN, Sanoev Kh A. Scientific basis for soybean planting in the condition of grassy alluvial soil prone to salinization. JCR. 2020;7(4):354-360. DOI: 10.31838/jcr.07.04.68
19. Tsikhungu PW. Analysis of micronutrients and heavy metals of indigenous reed salts and soils from selected areas in Western Kenya (Doctoral dissertation, Egerton University); 2016.
20. Ayuke FO, Karanja NK, Wickama J, Awiti A, Hella J. Soil fauna community structure across land management systems of Kenya and Tanzania; 2007.
21. Kimetu JM, Lehmann J, Ngoze SO, Mugendi DN, Kinyangi JM, Riha S, Pell AN. Reversibility of soil productivity decline with organic matter of differing quality along a degradation gradient. Ecosystems. 2008;11(5):726.
22. Pender J, Place F, Ehui S. (Eds.). Strategies for sustainable land management in the East African highlands. Intl Food Policy Res Inst; 2006.
23. Kowalski SM, Cornell JA, Vining GG. Split-plot designs and estimation methods for mixture experiments with process variables. Technometrics. 2002;44(1):72-79.
24. Jones B, Goos P. A candidate-set-free algorithm for generating D-optimal split-plot *designs.* Journal of the Royal Statistical Society: Series C (Applied Statistics). 2007;56(3):347-364.
25. Jones B, Sall J. JMP statistical discovery software. Wiley Interdisciplinary Reviews: Computational Statistics. 2011;3(3):188-194.
26. Agromisa Inckel M, Tersmette T. The preparation and use of compost. 2005;27.
27. De Bruin JL, Pedersen P. Effect of row spacing and seeding rate on soybean yield. Agronomy journal. 2008;100(3):704-710.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/108686