International Journal of Development Mathematics Vol 1 Issue 2, Page No. 218 - 226

INTERNATIONAL JOURNAL OF DEVELOPMENT MATHEMATICS

ISSN: 3026-8656 (Print) | 3026-8699 (Online)

journal homepage: https://ijdm.org.ng/index.php/Journals

# A Modified Split-Plot Design Rice Yield Model and Analysis

**David I. John[a]* and Ali, I. Michael[a]**

[a]*Department of Mathematics and Statistics, Federal University Wukari, Taraba State, Nigeria*

**ARTICLE INFO**

***Article history:***
Received 10March 2024
Received in revised form 15 May 2024
Accepted 26 May 2024

**Keywords:**
Maximum likelihood estimation,
Model adequacy measures
Nonlinear function, Split-plot design,
Restricted maximum likelihood estimation

**MSC 2020 Subject classification:**
62M10, 91B84

**ABSTRACT**

In this research, a modified split-plot design model (SPDM) is introduced, where the mean part of the SPDM is tailored with a three-parameter Gompertz function, which modifies the SPDM to an intrinsically nonlinear SPDM. The model is applied to a balanced 2-replicated 3×42 mixed-level split-plot design experiment to determine the irrigation effect on four varieties of rice trial at four different rates of nitrogen fertilizer. The SPDM parameters were estimated through the methods of estimated generalized least squares (EGLS) and restricted maximum likelihood estimation (REML) for estimating the SPDM variance components. The parameter estimates are compared to estimates from ordinary least squares (OLS) and maximum likelihood estimation (MLE) estimates of the variance using four median adequacy measures and three information criteria for goodness of fit. The results obtained shows that the modified SPDM is a good fit, and its EGLS-REML estimates are of better reliability and adequacy compared to the OLS and EGLS-MLE techniques.

## 1. Introduction

Sir R. A. Fishers developed and introduced the split-plot design (SPD) of experiment in agricultural science in 1925 (David and Adehi, 2014; David and Ikwuoche, 2022), which has since been adopted in other sectors such as industrial experiments as a linear model (Wang, Kowalski and Vining, 2009; Myers, Montgomery, and Anderson-Cook 2009; Jones and Nachtcheim, 2009; Lu, Anderson-Cook, and Robinson, 2011; Lu and Anderson-Cook, 2012; Lu, Anderson-Cook and Robinson, 2012; Jones and Goos, 2012; Lu and Anderson-Cook, 2014; Anderson and Whitcomb, 2014; Lu, Robinson, and Anderson-Cook, 2014; Anderson, 2016; Kulahci and Menon, 2017; Huameng, Fan and Lei, 2017). However, little attention has been paid to the Intrinsically Nonlinear Split-Plot Design Model (INSPDM). The parameters in this class of model are not linearizable. The SPD model includes two sources of random variability, which are independently and identically distributed normally with zero mean and constant variance $\sigma^2$, that is, the Whole Plot Error (WPE) and Subplot Error (SPE). Therefore, traditional nonlinear regression is not a suitable choice because it cannot handle more than one random error variation. If used, the single Mean Square Error (MSE) produced will be a compromise between the WPE and SPE variances (Gumpertz and Rawlings, 1992; Knezevic *et al*., 2002; Blankenship *et al*., 2003). Gumpertz and Rawlings (1992) fitted a three-parameter Weibull function to the mean part of an unbalanced SPD experiment to study the effect of ozone (O3) exposure (WP treatment I) on soybean yield at two watering regimes (WP treatment II) on thirty chambers arranged in three randomized blocks (each block has 10 chambers). Two cultivars (SP treatments) are within each chamber where the soybeans are grown. Knezevic *et al*. (2002) and Blankenship *et al*. (2003) modelled the WP and SP effects of three nitrogen fertilizer rates on "Critical Period for Weed Control" (CPWC) in corn yield using Logistic and Gompertz functions. David *et al*. (2022a and 2022b) suggested the use of the estimated generalized least square technique through residual maximum likelihood estimation (EGLS-REML) technique on the variance components for estimating the parameters of a three-parameter Bertalanffy-Richards split-plot design model (SPDM) and a three-parameter Weibull SPDM, respectively, after comparing it to ordinary least squares (OLS) and EGLS-MLE. Also, the EGLS-REML, EGLS-MLE, and OLS estimators were compared by David *et al*. (2023a and 2023b) for Johnson-Schumacher and Chapman-Richards SPDMs, and the EGLS-REML was found to outperform the other two estimators based on its MAM, AIC, Corrected AIC, BIC, and SEE estimates.

* Corresponding author. Tel.: +23421484696
E-mail address: davidij@fuwukari.edu.ng (David, I. J.)
https://doi.org/10.62054/ijdm/0102.18

Many methods for estimating the variance covariance matrix exist and have been applied to various research projects (Ikeda *et al*., 2014; Hasegawa *et al*., 2010; Weerakkody and Johnson, 1992; Gumpertz and Rawlings, 1992; Liao & Lipsitz, 2002). The theoretical presentation of INSPDM has been given by Gumpertz and Pantula (1992), David *et al*. (2018), and David *et al*. (2019). In this research, a balanced INSPDM is presented. The WP and SP are modeled using a three-parameter Gompertz function (3PGF) with a fixed block effect. The variance covariance matrix (VC), is estimated using the restricted maximum likelihood estimation (REML) technique for estimating generalized least squares (EGLS), where the results obtained are compared to ordinary least squares (OLS) estimates and maximum likelihood estimation (MLE) technique for the EGLS of the fitted model. All fitted models are assessed for goodness of fit using median adequacy measures (MAM) by David *et al*. (2016 and 2020) and information criteria.

## 2 Methodology

The formulated model for this research is given below.

Let

$$Y_{ijk} = \mu + \gamma_i + \alpha_j + w_{ij} + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk} \tag{1}$$

be the linear SPD model with two factors **A** and **B**. The corresponding INSPDM is given as follows.

$$y_{ijk} = f(x_{ijk},\ \theta) + w_{ij} + \varepsilon_{ijk} \tag{2}$$

where, $y_{ijk}$ is the response variable; $i = 1, ..., s$ replicates (***Reps***) or block; $j = 1, ..., a$ levels of the *WP* factor **A**; $k = 1, ..., b$ levels of the *SP* factor **B**; $w_{ij}$ is the *WPE* and $\varepsilon_{ijk}$ is the *SPE*; $f(x_{ijk},\ \theta)$ is the nonlinear function for the mean describing the relationship of the fixed main and interaction effects to $y_{ijk}$. The parameters ***Reps***, **A** and **B** are assumed to be fixed. The WPE and SPE are random effects and are independent and identically distributed normally with zero mean and constant variance, that is, $w_{ij} \sim N(0,\sigma^2_{WP})$ and $\varepsilon_{ijk} \sim N(0,\sigma^2_{SP})$. To be able to estimate all parameters of the INSPDM, the number of parameters, $p$ in $f(x_{ijk},\ \theta)$ and the number of random effects $r$, the number of measurements in the data set, $n$, must be at least $p + r + 1$ and this implies $n \geq p + r + 1$.

2.1 Modified Split-Plot Design Model for a Balanced $2\times3\times4^2$ SPD Rice Trial

Let the mean curve, $f(x_{ijk},\ \theta)$ in equation (2) be substituted with a 3PGF for a three factor SPD to become

$$f(x_{ijkl},\ \theta) = \alpha_{ijkl} \times exp\left[-\lambda\, exp\left(-\omega\, x_{ijkl}\right)\right] \tag{3}$$

where, $\alpha_{ijkl}$ is the asymptote and it is tailored as $\alpha_{ijkl} = \alpha + Rep_i + V_j + N_l + (VN)_{jl} + (VI)_{jk} + (IN)_{kl}$. Hence, equation (3) can be rewritten as,

$$f(x_{ijkl},\ \theta) = \left[\alpha + Rep_i + V_j + N_l + (VN)_{jl} + (VI)_{jk} + (IN)_{kl}\right] exp\left[-\lambda\, exp\left(-\omega I_{ijk}\right)\right]. \tag{4}$$

The modified SPD model with 3PGF as the mean curve is therefore given as,

$$y_{ijkl} = \left[\alpha + Rep_i + V_j + N_l + (VN)_{jl} + (VI)_{jk} + (IN)_{kl}\right] exp\left[-\lambda\, exp\left(-\omega I_{ijk}\right)\right] + w_{ijk} + \varepsilon_{ijkl} \tag{5}$$

where, $i = 1, 2$; $j = 1, 2, 3, 4$; $k = 1, 2, 3$; $l = 1, 2, 3, 4$, $\alpha$ is the average yield at zero rate or dose, $Rep_i$ is the $i$th replicate or block, $V_j$ is the effect of the $j$th levels of rice variety, $I_k$ is the effect of the $k$th levels of Irrigation, $N_l$ is the $l$th levels of nitrogen fertilizer effect, $(VN)_{jl}$ is the jth and lth levels interaction effect of rice varieties and nitrogen fertilizer, $(VI)_{jk}$ is the $j$th and $k$th levels interaction effect of the rice varieties and irrigation, and $(IN)_{kl}$ is the $k$th and $l$th levels interaction effect of irrigation and nitrogen fertilizer, $\omega$ and $\lambda$ are the Gompertz function scale and shape parameters respectively, $w_{ij}$ and $\varepsilon_{ijk}$ are the WPE and SPE respectively.

### 2.2 Parameter Estimation: Estimated Generalized Least Square (EGLS)

When the covariance matrix of $y$ is known then the GLS estimator, $\hat{\theta}_{\text{GLS}}$ is found by minimizing the objective function (Gumpertz and Rawlings, 1992; David *et al*., 2019)

$$\left(y - f(X,\ \theta)\right)^{t} \boldsymbol{Q}^{-1}\left(y - f(X,\ \theta)\right) \tag{6}$$

where, $\boldsymbol{Q}$ is a known positive definite (non-singular) covariance matrix which arises from equation (2).

Since $\boldsymbol{Q}$ is unknown it needs to be estimated using the methods of REML and MLE to estimate the variance components. According to David *et al*. (2018 and 2019) the derived estimates are given as

$$\left\langle tr\left(\hat{\mathbf{T}}_{(h)}\hat{\mathbf{Q}}_i\ \hat{\mathbf{T}}_{(h)}\hat{\mathbf{Q}}_j\right)\right\rangle \times \left\langle\left(\hat{\sigma}^2_{j(h+1)}\right)\right\rangle = \left\langle\left(y^{t}\hat{\mathbf{T}}_{(h)}\hat{\mathbf{Q}}_i\ \hat{\mathbf{T}}_{(h)}y\right)\right\rangle \tag{7}$$

$$\left\langle\left(\hat{\sigma}^2_{j(h+1)}\right)\right\rangle = \left\langle tr\left(\hat{\mathbf{T}}_{(h)}\hat{\mathbf{Q}}_i\ \hat{\mathbf{T}}_{(h)}\hat{\mathbf{Q}}_j\right)\right\rangle^{-1} \times \left\langle\left(y^{t}\hat{\mathbf{T}}_{(h)}\hat{\mathbf{Q}}_i\ \hat{\mathbf{T}}_{(h)}y\right)\right\rangle \tag{8}$$

and

$$\begin{aligned}\hat{\sigma}^2_{(h+1)} = &\left\langle tr\left(\mathbf{K}_j\mathbf{K}'_j\hat{\mathbf{Q}}^{-1}_{(\mathrm{h})}\mathbf{K}_i\mathbf{K}'_i\hat{\mathbf{Q}}^{-1}_{(\mathrm{h})}\right)\right\rangle^{-1} \\ &\times\left\langle\left(z_0 - \hat{D}_0(\hat{\theta}^*_{(h+1)} - \ \hat{\theta}^*_0)\right)' \hat{\mathbf{Q}}^{-1}_{(\mathrm{h})}\mathbf{K}_j\mathbf{K}'_j\hat{\mathbf{Q}}^{-1}_{(\mathrm{h})} \right. \\ &\left.\times\left(z_0 - \hat{D}_0(\hat{\theta}^*_{(h+1)} - \ \hat{\theta}^*_0)\right)\right\rangle\end{aligned} \tag{9}$$

respectively.

The solutions to the equations may turn out to be negative when further iteration does not improve the log-likelihood. In such a case, the negative value is changed to zero before the next iteration.

## 2.3 Median Adequacy Measure (MAM) Statistics and Information Criteria (IC)

Four proposed Median Adequacy Measure (MAM) statistics for assessing the adequacy of linear SPD models and regression models (David *et al*. 2016, 2020) and three IC are used for this research to assess the adequacy and goodness of fit of the fitted MSPDM. The four MAM statistics used are the resistant coefficient of determination ( $r_r^2$ ) proposed by Kvalseth (1985), resistant prediction coefficient of determination (*Pred*-$r_r^2$), resistant modeling efficiency (*RMEF*) and median square error prediction (*MedSEP*). These statistics are called resistant due to their ability to withstand outliers or extreme values and not increase or decrease unnecessarily when a variable is added or removed from the original model. The ICs used are the Akaike IC (AIC), Corrected AIC (CAIC), and Bayesian IC (BIC). All the MAM and IC formulas used can be found in David *et al*. (2022a; 2022b; 2023a; 2023b). Below are the statistics as presented in subsection 2.5.1 to 2.5.5.

### 2.3.1 Resistant Coefficient of Determination ( $r_r^2$ )

The statistic to calculate the WP and SP $r_r^2$ values are as follows:

$$r^2_{r(wp)} = 1 - \left(\frac{\underset{i=1}{\overset{n}{M}}\left(\left|\varepsilon_i\right|\right)_{WP}}{\underset{i=1}{\overset{n}{M}}\left(\left|Y_i - \bar{Y}\right|\right)_{WP}}\right)^2 \tag{10}$$

$$r^2_{r(sp)} = 1 - \left(\frac{\underset{i=1}{\overset{n}{M}}\left(\left|\varepsilon_i\right|\right)_{SP}}{\underset{i=1}{\overset{n}{M}}\left(\left|Y_i - \bar{Y}\right|\right)_{SP}}\right)^2 \tag{11}$$

where, $M$ is the median of the absolute values from $i = 1$ to $n$ and $\varepsilon_i$ is the fitted models residuals. The statistics (10 and 11) above uses the median instead of the mean in obtaining a coefficient of determination value that is highly resistant to outliers as proposed by Kvalseth (1985), $0 \le r_r^2 \le 1$. However, for nonlinear models the coefficient of determination value can be negative when the fit is worse, that is, $-1 \le r_r^2 \le 1$.

**2.3.2 Resistant Prediction Coefficient of Determination (Pred- $r_r^2$ )**

The statistic to calculate the WP and SP Pred- $r_r^2$ values are as follows:

$$_{WP}\,Pred-r_r^2=1-\left(M_{i=1}^{n}\left[\left(|e_i|\right)\Big/\left(1-h_{ii}\right)\right]_{WP}^{2}M_{i=1}^{n}\left[\left|Y_{i(WP)}-\bar{Y}_{(WP)}\right|\right]^{-2}\right) \tag{12}$$

$$_{SP}\,Pred-r_r^2=1-\left(M_{i=1}^{n}\left[\left(|e_i|\right)\Big/\left(1-h_{ii}\right)\right]_{SP}^{2}M_{i=1}^{n}\left[\left|Y_{i(SP)}-\bar{Y}_{(SP)}\right|\right]^{-2}\right) \tag{13}$$

where, $M$ is the median of the squared values from $i$ = 1 to $n$, $e_i$ is the residual, $h_{ii}$ is the hat matrix and -1 ≤ Pred- $r_r^2$ ≤ 1. However, for nonlinear models the prediction coefficient of determination value can be negative when the fit is worse, that is, -1 ≤ Pred- $r_r^2$ ≤ 1.

**2.3.3 Resistant Modeling Efficiency (RMEF)**

The statistic to calculate the WP and SP RMEF values are as follows:

$$\text{RMEF}_{WP}=1-\left(\frac{M_{i=1}^{n}\left[\left|Y_i-f\left(X_i\ldots,X_p\right)_i\right|\right]_{WP}}{M_{i=1}^{n}\left(\left|Y_i-\bar{Y}\right|\right)_{WP}}\right)^2 \tag{14}$$

$$\text{RMEF}_{SP}=1-\left(\frac{M_{i=1}^{n}\left[\left|Y_i-f\left(X_i\ldots,X_p\right)_i\right|\right]_{SP}}{M_{i=1}^{n}\left(\left|Y_i-\bar{Y}\right|\right)_{SP}}\right)^2 \tag{15}$$

where, $M$ is the median of the absolute values from $i$ = 1 to $n$ and $f(X_i\ldots,X_p)_i$ is the model-predicted values. In a perfect fit RMEF would result in a value equal to one. The upper bound is one and the (theoretical) lower bound is negative infinity (– ∞ < RMEF ≤ 1).

**2.3.4 Median Square Error Prediction (MedSEP)**

The statistic to calculate the WP and SP *MedSEP* values are as follows:

$$_{WP}\,MedSEP=(n)_{WP}^{-1}\left(M_{i=1}^{n}\left[\left|Y_i-f\left(X_i\ldots,X_p\right)_i\right|\right]_{WP}\right)^2 \tag{16}$$

$$_{SP}\,MedSEP=(n)_{SP}^{-1}\left(M_{i=1}^{n}\left[\left|Y_i-f\left(X_i\ldots,X_p\right)_i\right|\right]_{SP}\right)^2 \tag{17}$$

where, $M$ is the median of the absolute values from $i$ = 1 to $n$ and $f(X_i\ldots,X_p)_i$ is the model-predicted values. A model with the smallest *MedSEP* value is termed as more adequate.

**2.3.5 Information Criteria (IC) Statistics**

In this research, Akaike's Information Criteria (AIC), Corrected AIC (AICC) and Bayesian Information Criteria (BIC) are used for testing the goodness of fit of the models and to complement the results obtained from MAM. The statistic for each criterion is given as follows.

$$\text{AIC}=2f\left(\hat{\theta}\right)+2p \tag{18}$$

$$\text{AICC}=2f\left(\hat{\theta}\right)+\frac{2np}{n-p-1} \tag{19}$$

$$\text{BIC}=2f\left(\hat{\theta}\right)+p\log(s) \tag{20}$$

where, $f$ () is the negative of the marginal log-likelihood function, $\hat{\theta}$ is the vector of parameter estimates, $p$ is the number of parameters, $n$ is the number of observations and $s$ is the number of subjects.

### 2.4 Experimental Data and Analysis Procedure

The data used for this research is balanced $3^1$ x $4^2$ replicated mixed-level SP experimental design data. The WP has two factors, which are irrigation and rice varieties. The irrigation was administered three different times: 7 days, 14 days, and 21 days, on four different rice varieties: NERICA 2, NERICA 3, NERICA 4, and NERICA 14. The SP factor is nitrogen fertilizer, and it was administered at four different rates: 30kg N ha-1, 60kg N ha-1, 90kg N ha-1, and 120kg N ha-1 on each of the four varieties of rice. The aim of the field trial was to determine the irrigation effect on the yield of four rice varieties. The research was conducted by the Institute of Agricultural Research, Ahmadu Bello University, Zaria, at their experimental field station in Kano State, Nigeria. The procedures for analysis are as follows:

1. Perform a traditional SP experimental design analysis. This is done to see which of the effects are significant because only the significant effects will be included in the main nonlinear model. Another reason is to avoid the unnecessary inclusion of factors in the model and to decrease the number of parameter estimates. To achieve this step using SAS 9.4 M5 software, the **Proc Mixed** code is used.
2. After identifying the significant effects, a reanalysis is performed to obtain the parameter estimates in terms of the regression model. The reason is that the size of the parameters to be estimated will be too large for meaningful nonlinear modeling as well as interpretation of results. At this stage, the main effects and their significant interaction effects, the WP and SP variance components, are estimated using the MLE and REML methods as implemented in SAS software through **Proc Mixed**. A total of 11 parameters are estimated, including the asymptote, scale, and shape parameters. These parameter estimates are used as initial values for the MSPD models under study.
3. The asymptote, shape, and scale parameters for each of the nonlinear functions used for remodeling the traditional SPD model were estimated using the **Proc Nlin** code in SAS.
4. At this step, the SAS **Proc Nlmixed** code is used to obtain EGLS results. The OLS results are obtained using the **Proc Nlin** code.

## 3. Results and Discussion

Table IV presents the Gompertz SPD model parameter estimates, standard errors and P-values from the OLS and EGLS via MLE and REML.

Table I Gompertz Split-Plot Design Model Parameter Estimates

| Parameter | OLS | EGLS (MLE) | EGLS (REML) | Std. Error a | Std. Error b | Std. Error c | P-value a | P-value b | P-value c |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha_0$ | 30.3170 | 36.7667 | 28.1836 | 76.3604 | 11.5378 | 2.6551 | 0.6922 | **0.0019** | **<.0001** |
| $\alpha_1$ | 1.4363 | 0.004338 | 0.58 | 2.1245 | 0.9007 | 0.615 | 0.5006 | 0.9962 | 0.348 |
| $\alpha_2$ | -1.7969 | 1.4347 | -0.08488 | 2.7323 | 4.3501 | 1.442 | 0.5123 | 0.7423 | 0.9532 |
| $\alpha_3$ | 0.06209 | 0.3893 | 0.2391 | 0.07429 | 0.2735 | 0.07494 | 0.4054 | 0.1579 | **0.0019** |
| $\alpha_4$ | -0.00428 | -0.2186 | -0.06172 | 0.07494 | 0.2710 | 0.07331 | 0.9546 | 0.4219 | 0.4019 |
| $\alpha_5$ | -0.00418 | -0.02331 | -0.01368 | 0.004774 | 0.01515 | 0.003848 | 0.3830 | 0.1272 | **0.0006** |
| $\alpha_6$ | 0.02488 | 0.03395 | 0.02042 | 0.03095 | 0.01659 | 0.008173 | 0.4234 | **0.0435** | **0.0142** |
| $\omega$ | -0.02954 | 0.1959 | 0.2651 | 0.08841 | 0.1052 | 0.05441 | 0.7390 | 0.0657 | **<.0001** |
| $\lambda$ | 0.3051 | 7.5504 | 8.9347 | 0.9188 | 4.9718 | 3.8233 | 0.7405 | 0.1321 | **0.0215** |
| $\hat{\sigma}_{\delta}^{2}$ | 6.9440 | 0.7373 | 266.62 | 56.0591 | 0.2588 | 70.7317 | 0.9017 | **0.0054** | **0.0003** |
| $\hat{\sigma}_{\varepsilon}^{2}$ | 1.4432 | 1.9870 | 4.8669 | 0.05006 | 0.1739 | 0.6533 | **<.0001** | **<.0001** | **<.0001** |

*Letters a, b and c represents OLS, EGLS (MLE) and EGLS (REML) respectively. Bold values imply significance at 5%.*

It can be observed from Table I that the parameter estimates obtained from the OLS estimation technique and the EGLS estimation technique via MLE and REML are a bit different. Also, the EGLS estimates via MLE are quite different from the REML estimates. The EGLS via REML mean estimate of 28.1836 is smaller compared to the OLS and EGLS via MLE mean estimates of 30.3170 and 36.7667, respectively. However, their respective p-values of 0.6922, 0.0019, and 0.0001 shows that the EGLS estimates via MLE and REML are significant at the 5% significance level but not significant at 5% for the OLS estimate. However, the replicate parameter estimates of 1.4363, 0.004338, and 0.58 with p-values of 0.5006, 0.9962, and 0.348 for OLS and EGLS via MLE and REML, respectively, are not significant at the 5% significance level. Also, the variety parameter estimates of -1.7969, 1.4347, and -0.08488 with p-values of 0.5123, 0.7423, and 0.9532 for OLS and EGLS via MLE and REML, respectively, are not significant at the 5% significance level. However, for nitrogen fertilizer, estimates of 0.06209, 0.3893, and 0.2391 with p-values of 0.4054, 0.1579, and 0.0019 for OLS and EGLS via MLE and REML, respectively, are not significant at the 5% significance level, except for the EGLS via REML estimate, whose p-value is less than 5%.

Table I shows that I*V interaction parameter estimates from OLS (-0.00428) and EGLS via MLE (-0.2186) and REML (-0.06172) are not significant because their p-values of 0.9546, 0.4219, and 0.4019 are all greater than 5% significance level. However, I*N interaction parameter estimates of -0.00418, -0.02331, and -0.01368 with p-values of 0.383, 0.1272, and 0.0006 from OLS and EGLS via MLE and REML, respectively, are not significant except for EGLS via REML, whose p-value is less than 5% significance level. Similarly, for V*N interaction effect parameter estimates of 0.02488, 0.03395, and 0.02042 with p-values of 0.4234, 0.0435, and 0.0142 from OLS and EGLS via MLE and REML, they are significant at 5%, except for the OLS estimate, whose p-value is greater than 5% significance level.

The OLS estimate for scale parameter ($\omega$) of -2.681 is smaller compared to the EGLS estimates via MLE (0.1959) and REML (0.2651). Their respective p-values of 0.7390, 0.0657, and 0.0001 indicates a significant scale parameter estimate from EGLS via REML, but the OLS and EGLS via REML estimates are not significant at the 5% significance level. Similarly, the shape parameter ($\lambda$) estimate from OLS (0.3051) is smaller than those of the estimates from EGLS via MLE (7.5504) and REML (8.9347), but their respective p-values of 0.7405, 0.1321, and 0.0215 indicate that only the EGLS via REML estimate is significant because its p-value is less than 5% significance level.

The final WP variance component ( $\hat{\sigma}_{\delta}^2$ ) parameter estimate of 266.62 from EGLS via REML is larger than the OLS estimate of 6.944 and the EGLS via MLE estimate of 0.7373. However, the OLS estimate is not significant because its p-value of 0.9017 is greater than 5%, but the EGLS via MLE and REML p-values of 0.0054 and 0.0003, respectively, are significant at the 5% significance level. While the SP variance component ( $\hat{\sigma}_{\varepsilon}^2$ ) estimate from OLS (1.4332) is also smaller compared to the EGLS via MLE (1.9870) and REML (4.8669) estimates, their p-values of 0.0001, respectively, indicate a significant covariance parameter estimate for the subplot at the 5% significance level. Table I reveals clearly that the covariance estimates from EGLS via REML for the WP and SP are larger compared to the OLS and EGLS via MLE estimates for the whole plot and split-plot.

Generally, the standard errors for each of the estimates from the OLS and EGLS via MLE and REML in Table I show that the EGLS via REML produced standard errors that are smaller compared to the OLS and EGLS via MLE parameter estimates standard errors. This gives pre-confirmation that the EGLS via REML estimates for the Gompertz SPD model are estimated adequately with better stability. Hence, the technique is more proficient than OLS and EGLS via MLE. The OLS, EGLS-MLE, and EGLS-REML estimated fitted models for the MSPDM are presented as follows:

$$y_{ijkl} = \left[30.317 + 1.4363Rep - 1.7969V + 0.06209N - 0.00428IV - 0.00418IN + 0.02488VN\right] \times exp\left[-0.3051 exp\left(-0.02954 \times I_{ijk}\right)\right] \tag{21}$$

$$y_{ijkl} = \left[36.7667 + 0.004338Rep + 1.4347V + 0.3893N - 0.2186IV - 0.02331IN + 0.03395VN\right] \times exp\left[-7.5504 exp\left(0.1959 \times I_{ijk}\right)\right] \tag{22}$$

$$y_{ijkl} = \left[28.1836 + 0.58Rep - 0.08488V + 0.2391N - 0.06172IV - 0.01368IN + 0.02042VN\right] \times exp\left[-8.9347 exp\left(0.2651 \times I_{ijk}\right)\right] \tag{23}$$

The estimated (OLS, EGLS-MLE, and EGLS-REML) fitted MSPDM adequacy measures for the WP and SP sub-design models are presented in Table II. The results revealed that all models produced similar values for both the WP and SP sub-design models. However, the WP sub-design models have larger adequacy measure values compared to the SP sub-design models for $r_r^2$ , Pred-$r_r^2$ , and *RMEF*. The *MedSEP* values for the WP sub-design models are lower than those for the SP sub-design models. Despite the fact that all values for the fitted Gompertz SPD models showed that a large proportion of variability in the data is explained, high prediction power, better model efficiency, and better error prediction strength are obtained.

Table II Median Adequacy Measures Results

| | $r_r^2$ | | **Pred-** $r_r^2$ | | **RMEF** | | ***MedSEP*** | |
|---|---|---|---|---|---|---|---|---|
| | WP | SP | WP | SP | WP | SP | WP | SP |
| **OLS** | 0.9991759 | 0.8191939 | 0.99906732 | 0.724423 | 0.999079 | 0.816553 | 4.11E-07 | 0.034905 |
| **MLE** | 0.99898 | 0.8174575 | 0.99884561 | 0.721776 | 0.998806 | 0.820929 | 6.3E-07 | 0.035579 |
| **REML** | 0.9999999 | 0.8192343 | 0.99999992 | 0.724485 | 1 | 0.8185 | 2.69E-15 | 0.034889 |

However, Table III below presents the goodness of fit results for the fitted models, and it showed that EGLS via REML produced the lowest AIC, AICC, and BIC values of 475.7, 478.9, and 465.8, respectively. This implies that the EGLS-REML estimation technique produces reliable and stable estimates compared to OLS and EGLS-MLE parameter estimates.

Table III Model Goodness of Fit Test Results

| **Method** | **AIC** | **AICC** | **BIC** |
|---|---|---|---|
| **OLS** | 494.4 | 497.5 | 522.6 |
| **MLE** | 504.9 | 508.1 | 495 |
| **REML** | **475.7** | **478.9** | **465.8** |

## 4. Conclusion

In this study, balanced 2-replicated 3×42 mixed-level SPD experiment data was analyzed through a MSPDM, where the mean curve of the traditional linear SPDM was tailored to accommodate a 3-parameter Gompertz function. The method of EGLS was performed for estimating the parameters of the model, and the estimated variance components were estimated through REML and MLE. All results obtained were compared with OLS in terms of adequacy and goodness of fit. The MAM and IC revealed that the EGLS-REML performed better than the EGLD-MLE and OLS. Also, the EGLS-REML parameter standard errors for the fitted MSPDM are smaller compared to the OLS and EGLS-MLE, which suggests that the EGLS-REML produced adequate, stable, efficient, and better estimates for the MSPDM parameters.

**Competing Interest:** The authors declare that there is no any competing interest.

## References

Anderson, M. J. (2016). Design of Experiments (DoE): How to handle hard-to-change factors using a split plot. *Chemical Engineering, 123*(9), 12-1-12-5.

Anderson, M. J. and Whitcomb, P. J. (2014). Employing power to ‘right-size’ design of experiments. *Journal of Test Evaluation, 35*, 40-44.

Blankenship, E. E., Stroup, W. W., Evans, S. P. and Knezevic, S. Z. (2003). Statistical inference for Calibration Points in Nonlinear Mixed Effects Models. *Journal of Agriculture, Biology, and Environmental Statistics, 8*(4), 455–468.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2016). Assessing the Adequacy of Split-plot Design Models. *International Journal of Scientific and Technology Research, 2*(6), 197-203.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2016). Resistant Measures in Assessing the Adequacy of Split-plot

Design Models. *International Journal of Data Science*, *1*(4),382–396.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2018). Nonlinear Split-Plot Design Model in Parameters Estimation using EGLS-MLE. ComTech: *Computer, Mathematics and Engineering Applications, 9*(2), 65-71.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2019). Parameter Estimation of Nonlinear Split-Plot Design models: A Theoretical Framework. *Journal of Reliability and Statistical Studies*, *12*(1), 117-129.

David, I. J., Adubisi, O. D., Ogbaji, O. E., Eghwerido, J. T. and Umar, Z. A. (2020). Resistant measures in assessing the adequacy of regression models. *Scientific African*, *8*, e00437. https://doi.org/10.1016/j.sciaf.2020.e00437

David, I. J. and Adehi, M. U. (2014). Effectiveness of Split-Plot Design over Randomized Complete Block Design in Some Experiments. *Journal of Biology, Agriculture and Healthcare*, *4*(19), 75-80.

David, I. J. and Ikwuoche, P. O. (2022). Analysis of Asparagus Africanus Vigor from Split-plot Experimental Design. *Open Access Journal of Biomedical Science*, *4*(6), 2201-2204.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2022a). A Bertalanffy-Richards Split-Plot Design Model and Analysis. *Journal of Statistical Modeling and Analytics*, *4*(1), 56-71.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2022b). A Weibull Split-Plot Design and Analysis. *Thailand Statistician*, *20*(2), 420-434.

David, I. J., Asiribo, O. E. and Dikko, H. G. (2023a). Nonlinear Split-Plot Design Modeling and Analysis of Rice Varieties Yield. *Scientific African*, *19*, e01444

David, I. J., Asiribo, O. E. and Dikko, H. G. (2023b). Johnson-Schumacher Split-Plot Design Modelling of Rice Yield. *Biometrical Letters*, *60*(1), 37-52.

Gumpertz, M. L. and Rawlings, J. O. (1992). Nonlinear Regression with Variance Components: Modeling Effects of Ozone on Crop Yield. *Crop Science*, *32*, 219–224.

Gumpertz, M. L. and Pantula, S. G. (1992). Nonlinear Regression with Variance Components. *Journal of the American Statistical Association*, *87*(417), 201–209.

Hasegawa, Y., Ikeda, S., Matsuura, S. and Suzuki, H. (2010). A study on methodology for total design management (the 4th report): A study on the response surface method for split-plot designs using the generalized least squares. In: *Proceedings of the 92nd JSQC Technical Conference*, Tokyo: The Japanese Society for Quality Control, pp. 235–238 (in Japanese).

Huameng, G., Fan, Y. and Lei, S. (2017). Split Plot and Data Analysis in SAS. American Institute of Physics Conference Proceedings 1834, 030024 (2017). https://doi:10.1063/1.4981589.

Ikeda, S., Matsuura, S. and Suzuki, H. (2014). Two-Step Residual-Based Estimation of Error Variances for Generalized Least Squares in Split-Plot Experiments. *Communications in Statistics-Simulation and Computation*, *43*(2), 342-358.

Jones, B. and Nachtsheim, C. J. (2009). Split-plot Designs: What, why, and How. *Journal of Quality Technology*, *41*(4), 340-361.

Jones, B. and Goos, P. (2012). I-optimal versus D-optimal split-plot response surface designs. *Journal of Quality Technology*, *44*(2), 85-101.

Knezevic, S. Z., Evans, S. P., Blankenship, E. E., Van Acker, R. C. and Lindquist, J. L. (2002). Critical period for weed control: the concept and data analysis. *Weed Science*, *50*, 773–786.

Kulachi, M. and Menon, A. (2017). Trellis plots as visual aids for analyzing split plot experiments. *Quality Engineering*, 29(2), 211–225. https://doi.org/10.1080/08982112.2016.1243248

Kvalseth, T. O. (1985). Cautionary note about $R^2$. *American Statistician*, *39*, 279–285.

Liao, J. and Lipsitz, S. R. (2002). A type of restricted maximum likelihood estimator of variance components in generalised linear mixed models. *Biometrika*, *89*(2), 401-409.

Lu, L., Anderson-Cook, C. M. and Robinson, T. J. (2011). Optimization of designed experiments based on multiple criteria utilizing a Pareto frontier. *Technometrics*, *53*(4), 353-365.

Lu, L. and Anderson-Cook, C. M. (2012). Rethinking the optimal response surface design for a first-order model with two-factor interactions, when protecting against curvature. *Quality Engineering*, *24*(3), 404- 422.

Lu, L., Anderson-Cook, C. M. and Robinson, T. J. (2012). A case study to demonstrate a Pareto Frontier for selecting a best response surface design while simultaneously optimizing multiple criteria. *Applied Stochastic Models in Business and Industry*, *28*(3), 206-221.

Lu, L., Robinson, T. J. and Anderson-Cook, C. M. (2014). A case study to select an optimal split-plot design for a mixture process experiment based on multiple objectives. *Quality Engineering*, *26*(4), 424-439.

Lu, L. and Anderson-Cook, C. M. (2014). Balancing multiple criteria incorporating cost using Pareto front optimization for split-plot designed experiments. *Quality and Reliability Engineering International*, *30*(1), 37-55.

Myers, R. H., Montgomery, D. C. and Anderson-Cook, C. M. (2009). Response surface methodology: Process and product optimization using designed experiments (3rd ed.). USA: John Wiley and Sons.

Weerakkody, G. J. and Johnson, D. E. (1992). Estimation of within model parameters in regression models with a nested error structure. *Journal of American Statistical Association*, *87*, 708–713.