Thailand Statistician
April 2022; 20(2): 420-434
http://statassoc.or.th
Contributed paper

A Weibull Split-Plot Design Model and Analysis

David Ikwuoche John*[a][b], Asiribo Osebekwin Ebenezer [b] and Dikko Hussaini Garba [b]
[a] Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.
[b] Department of Mathematics and Statistics, Federal University, Wukari, Nigeria.
*Corresponding author; e-mail: davidij@fuwukari.edu.ng

Received: 4 August 2020
Revised: 8 November 2020
Accepted: 19 November 2020

Abstract

In this research, a class of nonlinear split plot design model where the mean function of the split-plot model is not linearizable is presented. This was done by fitting intrinsically nonlinear split-plot design (SPD) models using Weibull function. The fitted model parameters were estimated using ordinary least square (OLS) and estimated generalized least square (EGLS) techniques based on Gauss-Newton with Taylor series expansion by minimizing their respective objective functions. The variance components for the whole plot and subplot random effects are estimated using maximum likelihood estimation (MLE) and restricted maximum likelihood estimation (REML) techniques. The adequacy of the fitted intrinsically nonlinear SPD model was tested using four median adequacy measures namely resistant coefficient of determination, resistant prediction coefficient of determination, resistant modeling efficiency statistic and median square error prediction statistic based on the residuals of the fitted models which are influenced by the two parameter estimation techniques being applied, that is, the OLS and EGLS. Akaike's information criteria (AIC), Corrected Akaike's information criteria (AICC) and Bayesian information criteria (BIC) statistics were used to select the best parameter estimation technique. The results obtained showed that the Weibull SPD model is adequate and a good fit based on OLS but of less reliability and stability when the standard errors of the parameter estimates were compared to EGLS-MLE and EGLS-REML parameter estimates standard errors.

Keywords: Weibull function, intrinsically nonlinear, split-plot design, maximum likelihood estimation, restricted maximum likelihood estimation, median adequacy measures, information criteria.

1. Introduction

Split-plot design (SPD) of experiment has since been used in all aspect of agricultural experiments as introduced by Sir R.A. Fishers in 1925 and in the industry too as a linear model (Myers et al. 2009, Jones and Nachtcheim 2009, Lu et al. 2011, Lu et al. 2012, Lu et al. 2012, Jones and Goos 2012, Lu and Anderson-Cook 2014, Anderson and Whitcomb 2014, Lu et al. 2014, Anderson 2016, Kulahci and Menon 2017, Gao et al. 2017). However, intrinsically nonlinear SPD (NSPD) modeling has received little attention. This class of model has parameters that are not linearizable. Since the

SPD has two sources of random variations, that is, the WPE and SPE traditional nonlinear regression will not be suitable because it cannot handle more than one random error variation. If wrongly 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 and estimated the parameters of a Weibull unbalanced SPD of experiment for the effect of ozone (O_3) 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 were the soybean are grown. Knezevic et al. (2002) and Blankenship et al. (2003) modelled the WP and SP effect of three nitrogen rates on "critical period for weed control" (CPWC) in corn yield using logistic and Gompertz functions. Theoretical presentation on intrinsically nonlinear SPD modelling has been given by Gumpertz and Pantula (1992), David et al. (2018) and David et al. (2019).

In this research, an intrinsically nonlinear balanced SPD modeling is presented. The WP and SP are modeled using a three parameter Weibull function with fixed block effect. The variance covariance matrix, $\mathbf{V}$ is estimated using maximum likelihood estimation (MLE) technique and restricted maximum likelihood estimation (REML) technique for estimated generalized least square (EGLS) where their results will be compared to ordinary least square (OLS) estimates of the fitted model. All fitted models are assessed for goodness of fit using median adequacy measures (MAM) by David et al. (2016) and information criteria.

2. Methodology

In this section, we present the NSPD models and a theoretical frame work for estimating the parameters of the models using an iterative Gauss-Newton procedure with Taylor series expansion. The NSPD model which has WPE and SPE are special case of nonlinear model with random effects (also called nonlinear model with $\mathbf{V}$ that is, WPE and SPE). The formulated model and assumptions are given as follows. 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 $\mathbf{A}$ and $\mathbf{B}$. The corresponding NSPD model 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 WP error and ε_{ijk} is the SP error; $f(x_{ijk}, \theta)$ is the nonlinear function for the mean describing the relationship of fixed main and interaction effects to the response y_{ijk}. The parameters Reps, **A** and **B** are assumed fixed.

Assumption 1: It is presumed that the WPE and SPE are random effects. Also, it is assumed that $w_{ij} \overset{i.i.d.}{\sim} N(0, \sigma^2_{WP})$ and $\varepsilon_{ijk} \overset{i.i.d.}{\sim} N(0, \sigma^2_{SP})$.

Assumption 2: Let $\hat{\theta}$ be the model parameter estimate of θ which follows an asymptotic normal distribution with mean θ and variance $\sigma^2(\mathbf{F}'\mathbf{F})^{-1}$, where $\mathbf{F}$ is the $n \times u$ matrix with elements $\left(\partial f\left(x_{ijk}, \theta\right) / \partial\theta'\right)$ where the columns, u of the matrix is a full rank.

Assumption 3: If the number of parameters in the mean function, $f(x_{ijk},\theta)$ is p and the number of random effects is r, then the number of measurements in the data set, n, must be at least $p+r+1$ in order to estimate all of the parameters. This implies that $n \geq p+r+1$.

3. Split-Plot Model with Weibull Function as the Mean Curve

The mean curve, $f(x_{ijk},\theta)$ in (2) is substituted with the Weibull function. The Weibull function used for this research is a three-parameter function. Let $f(x_{ijk},\theta)$ be a Weibull function. Therefore,

$$f(x_{ijk},\theta) = \alpha_{ijk} \times \exp\left[-\left(x_{ijk}\omega^{-1}\right)^{\lambda}\right], \tag{3}$$

where α_{ijk} is the asymptote and it is tailored as $\alpha_{ijk} = \alpha + Rep_i + A_j + B_k + (AB)_{jk}$. Hence, (3) can be rewritten as follows

$$f(x_{ijk},\theta) = \left[\alpha + Rep_i + A_j + B_k + (AB)_{jk}\right] exp\left[-\left(x_{ijk}\omega^{-1}\right)^{\lambda}\right]. \tag{4}$$

The SPD model with the Weibull function as the mean curve is therefore given as follows

$$y_{ijk} = \left[\alpha + Rep_i + A_j + B_k + (AB)_{jk}\right]\exp\left[-\left(x_{ijk}\omega^{-1}\right)^{\lambda}\right] + w_{ij} + \varepsilon_{ijk}, \tag{5}$$

where α is the average yield at zero rate or dose, Rep_i is the i^{th} replicate or block, A_j is the effect of the j^{th} levels of factor A, B_k is the effect of the k^{th} levels of factor B, $(AB)_{jk}$ is the j^{th} and k^{th} levels interaction effect of the factors A and B, respectively, x_{ijk} is the mean covariate effect in the i^{th} replicate at the j^{th} factor A and k^{th} factor B, ω and λ are the Weibull scale and shape parameters respectively, w_{ij} is the WP error and ε_{ijk} is SP error.

4. Method of 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)

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

with respect to θ, where $\mathbf{V}$ is a known positive definite (non-singular) covariance matrix which arises from the model

$$y_{ijk} = f(x_{ijk},\theta) + w_{ij} + \varepsilon_{ijk}, \tag{7}$$

where $E(w_{ij}) = 0$, $Cov(w_{ij}) = \sigma_w^2\mathbf{I}_{\text{N}}$, $E(\varepsilon_{ijk}) = 0$, and $Cov(\varepsilon_{ijk}) = \sigma_\varepsilon^2\mathbf{I}_{\text{N}}$. Let the $\mathbf{V}$ matrix of the observations $\text{var}(y)$ be written as

$$\mathbf{V} = \sigma_w^2\text{I}_{\text{N}} + \sigma_\varepsilon^2\text{I}_{\text{N}} = \sigma^2\mathbf{I}.$$

By Cholesky decomposition, multiply model (7) by J^{-1} on both sides yield that

$$\text{J}^{-1}y_{ijk} = \text{J}^{-1}f(x_{ijk},\theta) + \text{J}^{-1}(w_{ij}) + \text{J}^{-1}(\varepsilon_{ijk}). \tag{8}$$

Let $\mathbf{I} = \text{T}^{-1} = \text{JJ}^{\text{t}}$ then the Cholesky factorization of the error variance is as follows:

$$\text{J}^{-1}\left[Cov\left(\varepsilon_{ijk}\right) + Cov\left(w_{ij}\right)\right]\text{J}^{-\text{t}} = \text{J}^{-1}Cov\left(\varepsilon_{ijk}\right)\text{J}^{-\text{t}} + \text{J}^{-1}Cov\left(w_{ij}\right)\text{J}^{-\text{t}} = \text{J}^{-1}\text{J}^{-\text{t}}\left[Cov\left(\varepsilon_{ijk}\right) + Cov\left(w_{ij}\right)\right]$$

$$= \text{J}^{-1}\left(\sigma^2\mathbf{I}\right)\text{J}^{-\text{t}} \quad = \sigma^2\text{J}^{-1}\text{JJ}^{\text{t}}\text{J}^{-\text{t}} \quad = \sigma^2\mathbf{I}.$$

Define $\mathbf{T}_{ijk} = \mathrm{J}^{-1} y_{ijk}$, $\mathbf{M}(x_{ijk}, \theta*) = \mathrm{J}^{-1} f(x_{ijk}, \theta)$ and $\Omega_{ijk} = \mathrm{J}^{-1}(w_{ij}) + \mathrm{J}^{-1}(\varepsilon_{ijk})$. Then (8) becomes

$$\mathbf{T}_{ijk} = \mathbf{M}(x_{ijk}, \theta^*) + \Omega_{ijk}, \tag{9}$$

where $E(\mathbf{\Omega}_{ijk}) = 0$ and $\mathbf{V}(\mathbf{\Omega}_{ijk}) = \sigma^2 \mathbf{I}$. Thus, the GLS model has been transformed to an OLS model. Hence, model (9) is to be solved using the OLS technique as follows. Taking the summation of both sides of (9) and squaring we have

$$\sum_{i}^{s}\sum_{j}^{a}\sum_{k}^{b} \Omega_{ijk}^2 = \sum_{i}^{s}\sum_{j}^{a}\sum_{k}^{b} \left[\mathbf{T}_{ijk} - \mathbf{M}(x_{ijk}, \theta*)\right]^2, \tag{10}$$

Let $L(\theta*) = \sum_{i}^{s}\sum_{j}^{a}\sum_{k}^{b} \Omega_{ijk}^2 = \sum_{i}^{s}\sum_{j}^{a}\sum_{k}^{b} \left[\mathbf{T}_{ijk} - \mathbf{M}(x_{ijk}, \theta*)\right]^2$, minimize $L(\theta*)$ w.r.t. $\theta*$, equate to zero and divide through by –2, we have

$$\frac{\partial L(\theta^*)}{\partial \theta_h^*} = \sum_{i}^{s}\sum_{j}^{a}\sum_{k}^{b} \left[\mathbf{T}_{ijk} - \mathbf{M}(x_{ijk}, \theta^*)\right] \times \left[\frac{\partial \mathbf{M}(x_{ijk}, \theta^*)}{\partial \theta_h^*}\right]_{\theta^* = \hat{\theta}^*} = 0. \tag{11}$$

At this point, (11) has no closed form hence will be solved iteratively using the Gauss-Newton method with Taylor series expansion of $\mathbf{M}(x_{ijk}, \theta^*)$ at first order. Therefore, we have

$$\mathbf{M}(x_{ijk}, \theta^*) = \mathbf{M}(x_{ijk}, \theta_0^*) + (\theta_1^* - \theta_{10}^*) \left.\frac{\partial \mathbf{M}(x_{ijk}, \theta^*)}{\partial \theta_1^*}\right|_{\theta^* = \theta_0^*} + (\theta_2^* - \theta_{20}^*) \left.\frac{\partial \mathbf{M}(x_{ijk}, \theta^*)}{\partial \theta_2^*}\right|_{\theta^* = \theta_0^*} + \ldots$$

$$+ (\theta_h^* - \theta_{h0}^*) \left.\frac{\partial \mathbf{M}(x_{ijk}, \theta^*)}{\partial \theta_h^*}\right|_{\theta^* = \theta_0^*}. \tag{12}$$

Let $\mathbf{M}(x_{ijk}, \theta^*) = \eta(\theta^*)$ and $d_{ijkl} = \left.\frac{\partial \mathbf{M}(x_{ijk}, \theta^*)}{\partial \theta_h^*}\right|_{\theta^* = \theta_0^*}$ for all N cases and $\delta = \theta^* - \theta_0^*$ then (12) becomes

$$\eta(\theta^*) = \eta(\theta_0^*) + D_0 \delta, \tag{13}$$

where D_0 is the $N \times H$ derivative matrix with elements $\{d_{ijk}\}$ at h iterations and this is equivalent to approximating the residuals for the model, that is, $\mathbf{\Omega}(\theta^*) = \mathbf{T} - \eta(\theta^*)$ by

$$\Omega(\theta^*) = \mathbf{T} - \left[\eta(\theta_0^*) + D_0 \delta\right] = \mathbf{T} - \eta(\theta_0^*) - D_0 \delta = z_0 - D_0 \delta, \tag{14}$$

where $z_0 = \mathbf{T} - \eta(\theta_0^*)$ and $\delta = \theta^* - \theta_0^*$.

To achieve numerical stability of the parameter estimates D_0 is decomposed using QR decomposition into the product of an orthogonal matrix and an inverted matrix (Klotz 2006, David et al. 2019). A point $\hat{\eta}_1 = \eta(\theta_1^*) = \eta(\theta_0^* + \delta_0)$ should now be closer to y than $\eta(\theta_0^*)$, and then move to a better parameter value $\theta_1^* = \theta_0^* + \delta_0$ and perform another iteration by calculating new residuals $z_1 = \mathbf{T} - \eta(\theta_1^*)$, a new derivative matrix D_0, and a new increase. This process is reiterated until convergence is achieved, that is, until the increment is so small that there is no useful change in the elements of the parameter vector (Bates and Watts 1988). A small step in the direction δ_0 is introduced if the new value is not small as expected. A step factor λ is introduced such that $\theta_1^* = \theta_0^* + \lambda \delta_0$ where λ is chosen to ensure that the new residual sum of squares is less than the initial estimate. As suggested by Bates and Watts (1988) it is to start with $\lambda = 1$ and halve it until it is satisfied that the new residual sum of squares is less than the initial estimate.

5. Variance Component Estimation via MLE

To estimate $\mathbf{V}$, the mean function $\mathbf{M}(x_{ijk}, \theta^*)$, is first approximated by a Taylor series at first order centered at $\hat{\theta}_0^*$. Therefore, the log-likelihood function is given as

$$\ln L(\mathbf{\Delta}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{V}| - \frac{1}{2}\left(y - f(X, \theta)\right)^t \mathbf{V}^{-1}\left(y - f(X,\theta)\right), \tag{15}$$

where $\mathbf{\Delta}^t = \left[\theta^t, \left(\sigma_{WP}^2, \sigma_{SP}^2\right)^t\right]$ is then approximated by the surface and (15) becomes

$$\ln L(\mathbf{\Delta}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{V}| - \frac{1}{2}\left(z_0 - D_0\delta\right)^t \mathbf{V}^{-1}\left(z_0 - D_0\delta\right), \tag{16}$$

where $z_0 = T - \eta(\theta_0^*)$, $\delta = \theta^* - \theta_0^*$, $\mathbf{V} = \sigma^2\mathbf{I} = \sum_{i=1}^{2}\sigma_i^2 I_i I_i^t$, $\mathbf{V}\mathbf{V}^{-1} = \sum_{i=1}^{2}\sigma_i^2\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t$ then (16) becomes

$$\ln L(\mathbf{\Delta}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right| - \frac{1}{2}\left(z_0 - D_0\left[\theta^* - \theta_0^*\right]\right)^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1} \times \left(z_0 - D_0\left[\theta^* - \theta_0^*\right]\right) \tag{17}$$

and the gradient is given by partially differentiating (17) w.r.t. θ^* and σ_i^2 we have

$$\frac{\partial \ln L(\mathbf{\Delta})}{\partial(\theta^*)^t} = -\frac{1}{2}\left(z_0 - D_0\left[\theta^* - \theta_0^*\right]\right)^t \mathbf{V}^{-1}\left[-D_0\right] = -\left(z_0 - D_0\delta\right)^t \mathbf{V}^{-1}\left[-D_0\right] \tag{18}$$

and

$$\begin{aligned}
\frac{\partial \ln L(\mathbf{\Delta})}{\partial \sigma_i^2} &= -\frac{1}{2}\frac{1}{\left|\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right|}(\mathbf{I}_i\mathbf{I}_i^t) + \frac{1}{2}(z_0 - D_0\delta)^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}\mathbf{I}_i\mathbf{I}_i^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}(z_0 - D_0\delta) \\
&= -\frac{1}{2}\left(\frac{1}{\left|\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right|}(\mathbf{I}_i\mathbf{I}_i^t) - (z_0 - D_0\delta)^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}\mathbf{I}_i\mathbf{I}_i^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}(z_0 - D_0\delta)\right) \\
&= -\frac{1}{2}\left(tr\left[\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}(\mathbf{I}_i\mathbf{I}_i^t)\right] - (z_0 - D_0\delta)^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}\mathbf{I}_i\mathbf{I}_i^t\left(\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\right)^{-1}(z_0 - D_0\delta)\right) \\
&= -\frac{1}{2}\left(tr\left(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}\right) - (z_0 - D_0\delta)^t\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0 - D_0\delta)\right).
\end{aligned} \tag{19}$$

Note that $\sigma_i^2 = \sigma_{WP}^2, \sigma_{SP}^2$ in (19) hence $i = 2$ for the two error variance. Multiplying the partial derivative first terms of (18) and (19) by the identity, $\mathbf{V}\mathbf{V}^{-1}$ and equate to zero, gives the estimating equations

$$\begin{aligned}
-(z_0 - D_0\delta)^t\mathbf{V}^{-1}\left[-D_0\right] &= 0 \\
(z_0 - D_0\delta)^t\mathbf{V}^{-1}\left[-D_0\right] &= 0 \\
\mathbf{V}^{-1}D_0^t z_0 - D_0^t\mathbf{V}^{-1}D_0\delta &= 0 \\
D_0^t\mathbf{V}^{-1}D_0\delta &= D_0^t\mathbf{V}^{-1}z_0,
\end{aligned} \tag{20}$$

and

$$-\frac{1}{2}\Big(tr(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1})-(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)\Big)=0$$

$$tr(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1})-(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)=0$$

$$\mathbf{V}\mathbf{V}^{-1}\Big[tr(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1})\Big]-(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)=0$$

$$\mathbf{V}\mathbf{V}^{-1}\Big[tr(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1})\Big]=(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)$$

$$\sum_{i=0}^{j}\sigma_i^2\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\Big[tr(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1})\Big]=(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)$$

$$\sigma^2\sum_{i=0}^{j}\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\Big[tr(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1})\Big]=(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)$$

$$\Big\langle tr\Big(\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}\mathbf{I}_j\mathbf{I}_j^t\mathbf{V}^{-1}\Big)\Big\rangle\sigma^2=\Big\langle(z_0-D_0\delta)^t\,\mathbf{V}^{-1}\mathbf{I}_i\mathbf{I}_i^t\mathbf{V}^{-1}(z_0-D_0\delta)\Big\rangle. \tag{21}$$

The estimates of $\hat{\theta}^*$ and $\sigma^2=\sigma_{WP}^2,\sigma_{SP}^2$ are iteratively obtained at $(h+1)^{\text{st}}$ iteration by substituting a prior estimate of σ^2 into (20) to get an updated estimate of $\hat{\theta}^*$, then the updated $\hat{\theta}^*$ and prior estimate of σ^2 are substituted into (21) to obtain updated estimates of the $\mathbf{V}$. These two steps are iterated till convergence is achieved. Therefore, (20) and (21) becomes

$$\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}\hat{D}_0(\hat{\theta}^*-\hat{\theta}_0^*)=\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}z_0,$$

$$\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}\hat{D}_0\hat{\theta}^*-\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}\hat{D}_0\hat{\theta}_0^*=\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}z_0,$$

$$\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}\hat{D}_0\hat{\theta}^*=\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}\hat{D}_0\hat{\theta}_0^*+\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}z_0,$$

$$\hat{\theta}_{(h+1)}^*=\hat{\theta}_0^*+\Big[\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}\hat{D}_0\Big]^{-1}\hat{D}_0^t\hat{\mathbf{V}}_{(h)}^{-1}z_0, \tag{22}$$

and

$$\hat{\sigma}_{(h+1)}^2=\Big\langle tr\Big(\mathbf{I}_j\mathbf{I}_j^t\hat{\mathbf{V}}_{(h)}^{-1}\mathbf{I}_i\mathbf{I}_i^t\hat{\mathbf{V}}_{(h)}^{-1}\Big)\Big\rangle^{-1}\times\Big\langle\Big(z_0-\hat{D}_0(\hat{\theta}_{(h+1)}^*-\hat{\theta}_0^*)\Big)^t\,\hat{\mathbf{V}}_{(h)}^{-1}\mathbf{I}_j\mathbf{I}_j^t\hat{\mathbf{V}}_{(h)}^{-1}\times\Big(z_0-\hat{D}_0(\hat{\theta}_{(h+1)}^*-\hat{\theta}_0^*)\Big)\Big\rangle. \tag{23}$$

When further iteration does not improve the log-likelihood, the solutions to the equations may turn out to be negative. In such scenario, the negative value is retuned to zero before the next iteration.

6. Variance Component Estimation via REML

The REML system does not include $\hat{\theta}^*$ in the estimation of $\mathbf{V}$. The log-likelihood function is based on vectors in the error space, that is, on linear combinations of y which have expectation to be zero rather than y itself. To obtain these vectors in the error space the linear approximation of the residuals is used $z_0=D_0\delta+\varepsilon$. To estimate the $\mathbf{V}$ from the nonlinear functions of y that will not involve $\hat{\theta}^*$, vectors of the form K^ty are formed whereby K is chosen so that $K^tD_0=0$ which falls in the linear approximation to the error space. Elements of K^ty are called error contrasts (Harville 1977), that is, the part of the data that is orthogonal to the fixed effects (not dependent on the values of the fixed effect estimates), K is a full rank matrix satisfying $K^tD_0=0$ and applying maximum likelihood to K^ty, the log likelihood function of K^ty is

$$\ln L(\boldsymbol{\Theta}) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left|K^t\mathbf{V}K\right| - \frac{1}{2}\left(K^t y - K^t f(X,\theta)\right)^t \times \left(K^t\mathbf{V}K\right)^{-1}\left(K^t y - K^t f(X,\theta)\right), \tag{24}$$

where $\boldsymbol{\Theta}^t = \left(\sigma^2_{WP}, \sigma^2_{SP}\right)^t$ is then approximated by the surface and (24) becomes

$$\begin{aligned}\ln L(\boldsymbol{\Theta}) &= -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln\left|K^t\mathbf{V}K\right| - \frac{1}{2}\left(K^t y - K^t f(x,\theta)\right)^t\left(K^t\mathbf{V}K\right)^{-1}\left(K^t y - K^t f(x,\theta)\right)\\ &= C - \frac{1}{2}\ln\left|K^t\mathbf{V}K\right| - \frac{1}{2}\left(K^t y - K^t f(x,\theta)\right)^t\left(K^t y\left(K^t\mathbf{V}K\right)^{-1} - K^t f(x,\theta)\left(K^t\mathbf{V}K\right)^{-1}\right)\\ &= C - \frac{1}{2}\ln\left|K^t\mathbf{V}K\right| - \frac{1}{2}\left(\left(K^t y - K^t f(x,\theta)\right)^t K^t y\left(K^t\mathbf{V}K\right)^{-1}\right)\\ &\quad + \frac{1}{2}\left(\left(K^t y - K^t f(x,\theta)\right)^t K^t f(x,\theta)\left(K^t\mathbf{V}K\right)^{-1}\right).\end{aligned} \tag{25}$$

By matrix algebra on the third and fourth terms of (25) respectively and inserting

$\mathbf{V} = \sigma^2\mathbf{I} = K\sum_{i=1}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K^t$ and

$\mathbf{V}\mathbf{V}^{-1} = \left(K\mathbf{V}K^t\right)\left(K\mathbf{V}K^t\right)^{-1} = \mathbf{V}\left(K\left(K\mathbf{V}K^t\right)^{-1}K^t\right) = \sigma^2\mathbf{I}\left(\mathrm{Q}_h\right) = \sigma^2\left(\mathrm{Q}_h\mathbf{V}_j\right),$ (25) becomes

$$\begin{aligned}\text{In}L(\boldsymbol{\Theta}) &= C - \frac{1}{2}\ln\left|K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right|\\ &\quad -\frac{1}{2}\left(y^t K^t\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1} yK - f(x,\theta)^t K^t\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1} yK\right)\\ &\quad +\frac{1}{2}\left(y^t K^t\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1} f(x,\theta)K - f(x,\theta)^t K^t\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1} f(x,\theta)K\right).\end{aligned} \tag{26}$$

Differentiate partially (26) w.r.t. σ_i^2 and equate to zero. By transformation all other terms in the equation becomes zero since $K^t D_0 = K^t f(x,\theta) = Kf(x,\theta)^t = 0.$ Hence we have

$$\frac{\partial\ln L(\boldsymbol{\Theta})}{\partial\sigma_i^2} = -\frac{1}{2}\frac{1}{\left|\mathrm{K}^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\mathrm{K}\right|}\left(\mathrm{K}^t\mathbf{I}_i\mathbf{I}_i^t\mathrm{K}\right) + \frac{1}{2}y^t K^t\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1}\left(K^t\mathbf{I}_i\mathbf{I}_i^t K\right)\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1} yK$$

$$\frac{1}{2}\left(\frac{1}{\left|K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right|}\left(K^t\mathbf{I}_i\mathbf{I}_i^t K\right)\right) = \frac{1}{2}y^t K^t\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1}\left(K^t\mathbf{I}_i\mathbf{I}_i^t K\right)\left(K^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t K\right)^{-1} yK. \tag{27}$$

Let $\mathbf{Q}_h\mathbf{V} = \mathbf{K}^t\left(\mathrm{K}^t\sum_{i=0}^{2}\sigma_i^2\mathbf{I}_i\mathbf{I}_i^t\mathrm{K}\right)^{-1}\mathbf{K}$ then (27) becomes

$$\frac{1}{2}\left(tr\left(\mathbf{Q}_h\mathbf{V}_i\right)\right) = \frac{1}{2}\left(y^t\mathbf{Q}_h\mathbf{V}_i\mathbf{Q}_h y\right). \tag{28}$$

Multiply the left hand side of (28) by $\mathbf{V}\mathbf{V}^{-1}$ we have

$$\frac{1}{2}\left(tr\left(\mathbf{Q}_h\mathbf{V}_i\right)\right)\sigma^2_{j(h+1)}\left(\mathbf{Q}_h\mathbf{V}_j\right) = \frac{1}{2}\left(y^t\mathbf{Q}_h\mathbf{V}_i\,\mathbf{Q}_h y\right), \tag{29}$$

$$\left\langle tr\left(\hat{\mathbf{Q}}_{(h)}\hat{\mathbf{V}}_i\,\hat{\mathbf{Q}}_{(h)}\hat{\mathbf{V}}_j\right)\right\rangle \times \left\langle\left(\hat{\sigma}^2_{j(h+1)}\right)\right\rangle = \left\langle\left(y^t\hat{\mathbf{Q}}_{(h)}\hat{\mathbf{V}}_i\,\hat{\mathbf{Q}}_{(h)}y\right)\right\rangle, \tag{30}$$

$$\left\langle\left(\hat{\sigma}^2_{j(h+1)}\right)\right\rangle = \left\langle tr\left(\hat{\mathbf{Q}}_{(h)}\hat{\mathbf{V}}_i\,\hat{\mathbf{Q}}_{(h)}\hat{\mathbf{V}}_j\right)\right\rangle^{-1} \times \left\langle\left(y^t\hat{\mathbf{Q}}_{(h)}\hat{\mathbf{V}}_i\,\hat{\mathbf{Q}}_{(h)}y\right)\right\rangle. \tag{31}$$

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.

7. Median Adequacy Measure (MAM) Statistics

Four proposed median adequacy measure (MAM) statistics for assessing the adequacy of linear SPD models (David et al. 2016) are used for this research to assess the adequacy of the fitted NSPD models. The four statistics used are resistant coefficient of determination (r_r^2) proposed by Kvalseth (1985), resistant prediction coefficient of determination ($\text{Pred}-r_r^2$), resistant modeling efficiency (RMEF) and median square error prediction (MedSEP). Procedures for calculating the WP and SP respective models residuals are given by Almimi et al. (2009) and David et al. (2016). These statistics are called resistant due to their ability of withstanding outliers or extreme values and not to increase or decrease unnecessarily when a variable is added or removed from the original model. The four statistics are presented as follows.

8. Resistant Coefficient of Determination (r_r^2)

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

$$r^2_{r(wp)} = 1 - \left(\underset{i=1}{\overset{n}{M}}\left(\left|\varepsilon_i\right|\right)_{WP} \Big/ \underset{i=1}{\overset{n}{M}}\left(\left|Y_i - \bar{Y}\right|\right)_{WP} \right), \tag{35}$$

$$r^2_{r(sp)} = 1 - \left(\underset{i=1}{\overset{n}{M}}\left(\left|\varepsilon_i\right|\right)_{SP} \Big/ \underset{i=1}{\overset{n}{M}}\left(\left|Y_i - \bar{Y}\right|\right)_{SP} \right), \tag{36}$$

where M is the median of the absolute values from $i = 1$ to n and e_i is the fitted models residuals. The statistic (35 and 36) 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 $-1 \le r_r^2 \le 1$.

9. Resistant Prediction Coefficient of Determination ($\text{Pred} - r_r^2$)

The statistic to calculate the WP and SP, $\text{Pred}-r_r^2$ values are as follows:

$${}_{WP}\text{Pred} - r_r^2 = 1 - \left(\underset{i=1}{\overset{n}{M}}\left[\left(\left|e_i\right|\right) \big/ \left(1-h_{ii}\right) \right]^2_{WP} \underset{i=1}{\overset{n}{M}}\left[\left| Y_{i(WP)} - \bar{Y}_{(WP)} \right| \right]^{-2} \right), \tag{37}$$

$${}_{SP}\text{Pred} - r_r^2 = 1 - \left(\underset{i=1}{\overset{n}{M}}\left[\left(\left|e_i\right|\right) \big/ \left(1-h_{ii}\right) \right]^2_{SP} \underset{i=1}{\overset{n}{M}}\left[\left| Y_{i(SP)} - \bar{Y}_{(SP)} \right| \right]^{-2} \right), \tag{38}$$

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 $1 \le \text{Pred} - r_r^2 \le 1$. However, for nonlinear models the prediction coefficient of determination value can be negative when the fit is worse, that is $1 \le \text{Pred} - r_r^2 \le 1$.

10. Resistant Modeling Efficiency (RMEF)

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

$$\text{RMEF}_{WP} = 1 - \left(\frac{M_{i=1}^{n}\left[\left|Y_i - f\left(X_i ..., X_p\right)_i\right|\right]_{WP}}{M_{i=1}^{n}\left(\left|Y_i - \bar{Y}\right|\right)_{WP}} \right)^2, \tag{39}$$

$$\text{RMEF}_{SP} = 1 - \left(\frac{M_{i=1}^{n}\left[\left|Y_i - f\left(X_i ..., X_p\right)_i\right|\right]_{SP}}{M_{i=1}^{n}\left(\left|Y_i - \bar{Y}\right|\right)_{SP}} \right)^2, \tag{40}$$

where M the median of the absolute is values from $i = 1$ to n and $f(X_i,...,X_p)_i$ 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 $(-\infty < \text{RMEF} \le 1)$.

11. Median Square Error Prediction (MedSEP)

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

$${}_{WP}\text{MedSEP} = \left(n\right)_{WP}^{-1} \left(M_{i=1}^{n}\left[\left|Y_i - f\left(X_i ..., X_p\right)_i\right|\right]_{WP} \right)^2, \tag{41}$$

$${}_{SP}\text{MedSEP} = \left(n\right)_{SP}^{-1} \left(M_{i=1}^{n}\left[\left|Y_i - f\left(X_i ..., X_p\right)_i\right|\right]_{SP} \right)^2, \tag{42}$$

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

12. Information Criteria 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:

$$\text{AIC} = 2f(\hat{\theta}) + 2p, \tag{43}$$

$$\text{AICC} = 2f(\hat{\theta}) + \frac{2np}{n-p-1}, \tag{44}$$

$$\text{BIC} = 2f(\hat{\theta}) + p\log(s), \tag{45}$$

where $f(\cdot)$ 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.

13. Experimental Data and Analysis Procedure

The data used for this research is a balanced $3^1 \times 4^2$ replicated mixed level SP experimental design data is used. 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, 30 kg N ha^{-1}, 60 kg N ha^{-1}, 90 kg N ha^{-1} and 120 kg N ha^{-1} on each of the four varieties of rice. The aim of the field trial was to determine irrigation effect on the yield of rice. The research was conducted by 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. Performed a traditional SP experimental design analysis. This was done to see which of the effects are significant because only the significant effects will be included for the main nonlinear model. Another reason is to avoid unnecessary inclusion of factors in the model and to decrease the number of parameter estimates. To achieve this step using SAS 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 regression model. The reason is the size of parameters to be estimated will be too large for meaningful nonlinear modeling and as well interpretation of results. At this stage, the main effects, and their significant interaction effects, the WP and SP $\mathbf{V}$ 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 main NSPD models under study.

3. The asymptote, shape and scale parameters for each of the nonlinear functions used for remodeling the traditional SPD model where estimated using Proc Nlin code in SAS.

4. The 11 parameter estimates are used as initial estimates for the nonlinear models formulated in this research. The SAS Proc Nlmixed code is used at this stage of the research to obtain the results for EGLS. While the Proc Nlin code is used for obtaining the OLS results.

5. The residuals obtained from each fitted NSPD models are used to calculate all four median adequacy measures introduced in the research for assessing the adequacy of each fitted models so as to identify which model is a better adequate model.

14. Results

Tables 1 and 2 below present the analysis of variance tables. Table 1 shows that all main effects and two factor interaction effects are significant at 5% significance level since their respective p-values are all less than 5%. However, the three factor interaction effect is not significant because its p-value of 0.1271 is greater than 5% significance level. Based on the outcome of the analysis, the three factor interaction effect is removed and a reanalysis is performed. Table 2 presents the reanalysis which is a regression SPD analysis results. It was adopted to reduce the large treatments combinations from 48 to 11.

The results shows that all the main effects and interaction effects are significant at 5% significance level except for I*V (Irrigation by variety) and V*N (variety by nitrogen) interaction effects. This is because I*V and V*N respective p-values are greater than 5%. However, these two interaction effects are not dropped for further analysis because their respective main effects (I, V and N) are all significant at 5% level of significance. The covariance components estimates for the WP and SP are obtained based on this final regression analysis with SP errors. The two methods adopted for estimating the covariance components for this research are MLE and REML techniques. Table 3 presents their respective results.

Table 1 A 3×4^2 split-plot design ANOVA table

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Rep	1	17.1653	17.1653	2.8700	0.1184
I	2	742.9498	371.4749	62.0800	< 0.0001
V	3	118.0203	39.3401	6.5700	0.0083
I*V	6	113.7322	18.9554	3.1700	0.0467
WP Error	11	65.8245	5.9840		
N	3	198.1499	66.0500	22.0000	< 0.0001
I*N	6	156.6210	26.1035	8.6900	< 0.0001
V*N	9	187.8973	20.8775	6.9500	< 0.0001
I*V*N	18	84.0821	4.6712	1.5600	0.1271
SP Error	36	108.0991	3.0028		
Total	95	1792.5415			

Table 2 A 3×4^2 regression analysis with split-plot errors ANOVA table

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Rep	1	412.5178	412.5178	26.96	< 0.0001
I	1	342.7304	342.7304	22.40	< 0.0001
V	1	290.5514	290.5514	18.99	< 0.0001
I*V	1	0.3910	0.391008	0.03	0.8734
WP Error	1	248.3836	248.3836		
N	1	799.9866	799.9866	52.28	< 0.0001
I*N	1	407.1343	407.1343	26.61	< 0.0001
V*N	1	27.3536	27.35356	1.79	0.1847
SP Error	88	1346.5585	15.30180		
Total	95	3875.607113			

Table 3 Covariance parameter estimates

Parameter	MLE	REML
σ_δ^2	0	0.01648
σ_ε^2	16.6140	15.3018

The VC estimates presented in Table 3 above shows that the WP variance estimate for MLE is zero which is smaller than the estimates form REML (0.01648). However, for the SP variance estimate, the MLE estimate (16.6140) is larger than the estimates from REML (15.3018).

Table 4 presents the Weibull SPD model parameter estimates, standard errors and p-values from the OLS and EGLS via MLE and REML. It is shown in Table 4 that the parameter estimates obtained from OLS estimation technique is quite different from the EGLS estimation technique via MLE and REML. Also, it can be observed that the EGLS estimates via MLE are very similar to that of REML. The OLS produced a smaller mean estimate of 13.8105 compared to the EGLS mean estimates via MLE (22.3285) and REML (22.3183). Their respective p-values of 0.8630, 0.0001 and 0.0001 shows that the EGLS estimates via MLE and REML are significant at 5% significance level but not significant at 5% for the OLS estimate. However, the replicate effect estimates of 0.8458, −1.4140 and −1.4087 with p-values of 0.1503, 0.3052 and 0.0386 for OLS and EGLS via MLE and REML respectively are not significant at 5% significance level except for EGLS via REML estimate which has a p-value of 0.0386. It can be seen from Table 4 that variety effect estimates of −0.5767, −0.2242 and −0.2378 with p-values of 0.4696, 0.8685 and 0.7748 for OLS and EGLS via MLE and REML respectively are not significant at 5% significance level. Also, nitrogen fertilizer effect estimates of 0.0687, 0.0193 and 0.0197 with p-values of 0.011, 0.7819 and 0.5900 for OLS and EGLS via MLE and REML respectively are not significant at 5% significance level except for the OLS estimate whose p-value of 0.011 is less than 5%.

Table 4 Weibull 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
α_0	13.8105	22.3285	22.3183	79.8294	3.3316	1.3541	0.8630	**<0.0001**	**<0.0001**
α_1	0.8458	−1.4140	−1.4087	0.5833	1.3715	0.6718	0.1503	0.3052	**0.0386**
α_2	−0.5767	−0.2242	-0.2378	0.7943	1.3506	0.8290	0.4696	0.8685	0.7748
α_3	0.0687	0.0193	0.0197	0.0265	0.0694	0.0364	**0.0110**	0.7819	0.5900
α_4	−0.0439	−0.1057	−0.1046	0.0337	0.0252	0.0243	0.1963	**<0.0001**	**<0.0001**
α_5	−0.0047	−0.0044	−0.0044	0.0011	0.0017	0.0011	**<0.0001**	**0.0132**	**<0.0001**
α_6	0.0159	0.0239	0.0239	0.0078	0.0128	0.0069	**0.0428**	0.0634	**0.0009**
ω	−45.6810	−19.3449	−19.3722	1087.35	3.6365	1.4363	0.9666	**<0.0001**	**<0.0001**
λ	21.3869	0.7678	0.7695	1835.19	0.2509	0.0798	0.9907	**0.0029**	**<0.0001**
$\hat{\sigma}_\delta^2$	10.9228	3.6297	3.9387	79.8300	5.3432	3.1451	0.8915	0.4986	0.2135
$\hat{\sigma}_\varepsilon^2$	2.8574	19.7562	18.7310	0.2062	21.2235	8.8027	**<0.0001**	0.3543	**0.0359**

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

Looking at the interaction effects (Irrigation×Variety [I*V], Irrigation×Nitrogen [I*N] and Variety×Nitrogen [V*N]) parameter estimates, the EGLS via MLE and REML produced similar estimates of −0.1057 and −0.1046 respectively for I*V with respective p-values of 0.0001 which implies their estimates are significant at 5%. However, the OLS estimate of -0.0439 with a p-value of 0.1963 is not significant at 5% significance level. While I*N interaction effect parameter estimates of −0.0047, −0.0044 and −0.0044 with p-values of 0.0001, 0.0132 and 0.0001 from OLS and EGLS via MLE and REML respectively are all significant at 5% significance level. Similarly, V*N interaction effect parameter estimates of 0.0159, 0.0239 and 0.0239 with p-values of 0.0428, 0.0634 and 0.0009 for OLS and EGLS via MLE and REML are significant at 5% except for EGLS via MLE whose p-value of 0.0634 is greater than 5% significance level.

The OLS estimates for scale parameter, ω (–45.681) is smaller compared to the EGLS estimates via MLE (–19.3449) and REML (–19.3722). Their respective p-values of 0.9666, 0.0001 and 0.0001 indicates a significant parameter estimate from EGLS via MLE and REML but the OLS estimate is not significant at 5% significance level. Similarly, the shape parameter, λ estimates from OLS (21.3869) is greater than that of the estimates from EGLS via MLE (0.7678) and REML (0.7695) but their individual p-values of 0.9907, 0.0029 and 0.0001 indicates that the EGLS estimates via MLE and REML are significant at 5% significance level but the OLS estimate is not significant because its p-value is greater than 5% significance level.

The final whole plot variance component ($\hat{\sigma}_\delta^2$) parameter estimate from OLS (10.9228) is larger than the EGLS via MLE (3.6297) and REML (3.9387) estimates. However, these estimates are not significant because their respective p-values of 0.8915, 0.4986 and 0.2135 are all greater than 5% significance level. While the split-plot variance component ($\hat{\sigma}_\varepsilon^2$) estimate from OLS (2.8574) is smaller compared to the EGLS via MLE (19.7562) and REML (18.731) estimates however, their p-values of 0.0001, 0.3543 and 0.0359 indicates that the OLS and EGLS via REML estimates are significant at 5% significance level but the estimate from EGLS via MLE is not significant since its p-value is greater than 5% significance level. Table 4 reveals clearly that the covariance estimate from OLS is larger for the WP and smaller for the SP. However, this is not the case for the EGLS via MLE and REML where their estimates are smaller for the WP but larger for the SP.

Generally, the standard errors for each of the estimates from the OLS and EGLS via MLE and REML in Table 4 shows 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 a pre-confirmation that the EGLS via REML estimates for the Weibull SPD model are estimated adequately with better stability. Hence, the technique is more proficient than the OLS and REML via MLE. The OLS, EGLS-MLE and EGLS-REML estimated fitted models for the Weibull SPD model are presented as follows:

$$y_{ijkl} = \begin{bmatrix} 13.8105 + 0.8458 Rep_i - 0.5767 V_j + 0.06869 N_k \\ -0.04386 (IV)_{jl} - 0.00473 (IN)_{kl} + 0.01596 (VN)_{jk} \end{bmatrix} \times \exp\left[-\left((-45.681)^{-1} I_{ijkl} \right)^{21.3869} \right], \quad (46)$$

$$y_{ijkl} = \begin{bmatrix} 22.3285 - 1.414 Rep_i - 0.2242 V_j + 0.01926 N_k \\ -0.1057 (IV)_{jl} - 0.00441 (IN)_{kl} + 0.02394 (VN)_{jk} \end{bmatrix} \times \exp\left[-\left((-19.345)^{-1} I_{ijkl} \right)^{0.7678} \right], \quad (47)$$

$$y_{ijkl} = \begin{bmatrix} 22.3183 - 1.4087 Rep_i - 0.2378 V_j + 0.01968 N_k \\ -0.1046 (IV)_{jl} - 0.00442 (IN)_{kl} + 0.02389 (VN)_{jk} \end{bmatrix} \times \exp\left[-\left((-19.3722)^{-1} I_{ijkl} \right)^{0.7678} \right]. \quad (48)$$

The estimated (OLS and EGLS via MLE and REML) fitted Weibull SPD model adequacy measures for the WP and SP sub design models are also presented in Table 5. The results revealed that the OLS fitted model for both the WP and SP sub design models are adequately better than the EGLS-MLE and EGLS-REML WP and SP sub design models. However, the WP sub design models had larger adequacy measure values compared to the SP sub design models for r_r^2, $\text{Pred} - r_r^2$ and RMEF. While the MedSEP values for the WP sub design models is smaller compared to the SP sub design models. Although, all values for the fitted Weibull SPD models show that a large proportion of variability is explained in the data, high prediction power, better model efficiency and better error prediction strength for the OLS model

Table 5 Median adequacy measures results

	r_r^2		$\text{Pred} - r_r^2$		RMEF		MedSEP	
	WP	SP	WP	SP	WP	SP	WP	SP
OLS	**0.9969727**	**0.7379945**	**0.9965739**	**0.600662**	0.996823	0.754624	**5.55E-06**	**0.073296**
MLE	0.9659263	0.3971088	0.9614376	0.081099	**0.999227**	**0.786195**	0.000703	0.388096
REML	0.9691298	0.3938414	0.9650632	0.076119	0.98899	0.760389	0.000577	0.392314

However, Table 6 below presents the goodness of fit results for the fitted models and it showed that OLS produced the lowest AIC, AICC and BIC values of 496.0, 499.2 and 524.2 respectively. This implies that the OLS estimation technique produces reliable and stable estimates compared to EGLS via MLE and EGLS via REML parameters estimates.

Table 6 Model goodness of fit test results

Method	AIC	AICC	BIC
OLS	496.0	499.2	524.2
MLE	597.6	600.8	587.7
REML	596	599.2	586.1

12. Conclusions

Based on the research results from the analysis on a balanced 3^1 x 4^2 replicated mixed Level SP experimental design data, it was observed that the fitted Weibull SPD model is a good fit for the OLS estimated model. This is because all the four MAM for assessing the adequacy of the fitted model produced larger r_r^2, $\text{Pred}-r_r^2$ and RMEF values and smaller MedSEP values for the WP and SP sub models and as well smaller AIC, AICC and BIC values compared to the EGLS-MLE and EGLS-REML estimated models respectively. Also, all the respective OLS, EGLS-MLE and EGLS-REML estimated fitted intrinsically NSPD models for the WP sub models produced better adequacy measure values compared to the SP sub models. However, the OLS, EGLS-MLE and EGLS-REML respective parameter estimates standard errors showed that some of the OLS parameter estimates may not be stable and reliable because of high standard errors compared to EGLS-MLE and EGLS-REML parameter estimates standard errors. Also, only four out of the eleven parameter estimates for the OLS were significant at 5% significance level compared to EGLS-REML estimated fitted models which produced the lowest standard errors with eight out of the eleven parameter estimates significant at 5% significance level indicating better stability and reliability. Therefore, it can be concluded that despite the OLS fitted model may be adequate and a good fit it may not be stable and reliable for prediction.

References

Almimi AA, Kulahci M, Montgomery DC. Checking the adequacy of fit of models from split-plot designs. J Qual Technol. 2009; 41(3): 272-284.

Anderson MJ. Design of Experiments (DOE): how to handle hard-to-change factors using a split plot. Chem Eng. 2016; 123(9): 12-18.

Anderson MJ, Whitcomb PJ. Employing power to 'right-size' design of experiments. ITEA J. 2014; 35: 40-44.

Bates DM, Watts DG. Nonlinear regression analysis and its applications. New York: John Wiley & Sons. 1988.

Blankenship EE, Stroup WW, Evans SP, Knezevic SZ. Statistical inference for calibration points in nonlinear mixed effects models. J Agric Biol Environ Stat. 2003; 8(4): 455-468.

David IJ, Asiribo OE, Dikko HG. Resistant measures in assessing the adequacy of split-plot design models. Int J Data Sci. 2016; 1(4): 382-396.

David IJ, Asiribo OE, Dikko HG. Nonlinear split-plot design model in parameters estimation using estimated generalized least square-maximum likelihood estimation. ComTech. 2018; 9(2): 65-71.

David IJ, Asiribo OE, Dikko HG. Parameter estimation of nonlinear split-plot design models: a theoretical framework. J Reliab Stat Stud. 2019; 12(1): 117-129.

Gao H, Yang F, Shi L. Split plot design and data analysis in SAS. In AIP Conference Proceedings 2017; 1834(1): 030024. https://doi:10.1063/1.4981589.

Gumpertz ML, Rawlings JO. Nonlinear regression with variance components: modeling effects of ozone on crop yield. Crop Sci. 1992; 32(1): 219-224.

Gumpertz ML, Pantula SG. Nonlinear regression with variance components. J Am Stat Assoc. 1992; 87(417): 201-209.

Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. J Am Stat Assoc. 1977; 72(358): 320-338.

Jones B, Nachtsheim CJ. Split-plot designs: what, why, and how. J Qual Technol. 2009; 41(4): 340-361.

Jones B, Goos P. I-optimal versus D-optimal split-plot response surface designs. J Qual Technol. 2012; 44(2): 85-101.

Klotz JH. A computational approach to statistics. Madison: University of Wisconsin; 2006.

Knezevic SZ, Evans SP, Blankenship EE, Van Acker RC, Lindquist JL. Critical period for weed control: the concept and data analysis. Weed Sci. 2002; 50(6): 773-786.

Kulahci M, Menon A. Trellis plots as visual aids for analyzing split plot experiments. Qual Eng. 2017; 29(2): 211-225.

Kvålseth TO. Cautionary note about R^2. Am Stat. 1985; 39(4): 279-285.

Lu L, Anderson-Cook CM, Robinson TJ. Optimization of designed experiments based on multiple criteria utilizing a Pareto frontier. Technometrics. 2011; 53(4): 353-365.

Lu L, Anderson-Cook CM. Rethinking the optimal response surface design for a first-order model with two-factor interactions, when protecting against curvature. Qual Eng. 2012; 24(3): 404-422.

Lu L, Anderson-Cook CM, Robinson TJ. A case study to demonstrate a Pareto Frontier for selecting a best response surface design while simultaneously optimizing multiple criteria. Appl Stoch Models Bus Ind. 2012; 28(3): 206-221.

Lu L, Robinson TJ, Anderson-Cook CM. A case study to select an optimal split-plot design for a mixture-process experiment based on multiple objectives. Qual Eng. 2014; 26(4): 424-439.

Lu L, Anderson-Cook CM. Balancing multiple criteria incorporating cost using Pareto front optimization for split-plot designed experiments. Qual Reliab Eng Int. 2014; 30(1): 37-55.

Myers RH, Montgomery DC, Anderson-Cook CM. Response surface methodology: Process and product optimization using designed experiments. New York: John Wiley & Sons. 2009.