sciendo

DOI: 10.2478/bile-2023-0003

Biometrical Letters
Vol. 60 (2023), No. 1, 37-52

Johnson–Schumacher Split-Plot Design Modelling of Rice Yield

I. J. David[1], O. E. Asiribo[2], H. G. Dikko[3], P. O. Ikwuoche[1]

[1]Department of Mathematics and Statistics, Federal University Wukari, Nigeria
[2]Department of Statistics, Federal University of Agriculture, Abeokuta, Nigeria
[3]Department of Statistics, Ahmadu Bello University, Zaria, Nigeria, davidij@fuwukari.edu.ng

SUMMARY

In this research, an intrinsically nonlinear split-plot design model (INSPDM) is formulated and studied. It was formulated by fitting a Johnson–Schumacher (JS) function to the split-plot model mean function. The fitted model parameters are estimated using the estimated generalized least squares (EGLS) technique based on a Gauss–Newton procedure with Taylor series expansion, by minimizing the objective function of the model. The variance components for the whole plot and subplot random effects are estimated using restricted maximum likelihood estimation (REML) techniques. The adequacy of the fitted INSPDM was tested using four median adequacy measures: resistant coefficient of determination, resistant prediction coefficient of determination, the resistant modeling efficiency statistic, and the median square error prediction statistic based on the residuals of the fitted model. Akaike's Information Criterion (AIC), Corrected Akaike's Information Criterion (AICC) and Bayesian Information Criterion (BIC) statistics are used to select the best parameter estimation technique. The results obtained are compared with the techniques of ordinary least squares (OLS) and EGLS via maximum likelihood estimation (MLE). The results showed the model to be adequate, reliable, stable, and a good fit based on EGLS-REML when compared with OLS and EGLS-MLE fitted model parameter estimates.

Key words: Split-plot design; Johnson–Schumacher function; Least squares; Likelihood estimation; Goodness-of-fit measures.

1. Introduction

Sir R. A. Fisher introduced the split-plot design (SPD) to agricultural experiments in 1925, and since then it has been applied in industry and other fields as a linear model (Anderson, 2016; David, Ikwuoche, 2022; Kulachi and Menon, 2017; Huameng *et al*., 2017). However, the intrinsically nonlinear SPD model

(INSPDM), which is a class of model with parameters that are not linearizable, has received little attention. Since the SPD has two sources of random variations, that is, the whole plot error (WPE) and subplot error (SPE), traditional nonlinear regression is not a suitable option 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 an experiment on the effect of ozone (O_3) exposure (WP treatment I) on soybean yield under two watering regimes (WP treatment II) in thirty chambers arranged in three randomized blocks (each block has 10 chambers). Two cultivars (SP treatments) are within each chamber where the soybean is 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 the Logistic and Gompertz functions. A theoretical account of INSPDM has been given by Gumpertz and Pantula (1992), David *et al*. (2018), and David *et al*. (2019). Recently, a Weibull SPDM was developed and applied to rice yield by David *et al*. (2022a), and it was found that the estimated generalized least square (EGLS) via residual maximum likelihood estimation (REML) technique produced reliable and stable estimates compared with the ordinary least square approach. Also, David *et al*. (2022b) introduced the Bertalanffy–Richards SPDM for modeling rice yield, and the results obtained indicated the superiority of the model when its parameters are estimated with EGLS-REML compared with EGLS-MLE and OLS. Similarly, David *et al*. (2023) nonlinearly modeled rice yield from a split-plot designed experiment using the Chapman–Richards function, and the results indicated that the model was a good fit with reliable and stable properties for prediction.

In this research a balanced INSPDM is presented. The WP and SP are modeled using a three-parameter Johnson–Schumacher function (JSF) with fixed block effect. The variance covariance matrix **V** is estimated using the restricted maximum likelihood estimation (REML) technique for estimated generalized

least square (EGLS), and the results are compared with estimates from the ordinary least square (OLS) technique and the maximum likelihood estimation (MLE) technique for EGLS of the fitted model. All fitted models are assessed for goodness of fit using median adequacy measures (MAM) as given by David *et al*. (2016), and information criteria.

2. Materials and methods

In this section we present the INSPDM and a theoretical framework for estimating the parameters of the model using an iterative Gauss–Newton procedure with Taylor series expansion. The INSPDM with WPE and SPE is a special case of a nonlinear model with random effects (also called a nonlinear model with **V**, that is, WPE and SPE). The formulated model and assumptions are as follows.

Let

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

be a linear SPD model with two factors **A** and **B**. The corresponding NSPD model is given as follows:

$$y_{ijk} = f(x_{ijk}, \theta) + w_{ij} + \varepsilon_{ijk}, \quad (1)$$

where y_{ijk} is the response variable; i = 1, ..., s are replicates (*Reps*) or blocks; j = 1, ..., a are levels of the *WP* factor **A**; k = 1, …, b are levels of the *SP* factor **B**; w_{ij} is the *WP* error; ε_{ijk} is the *SP* error; and $f(x_{ijk}, \theta)$ 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 $\sigma^2(\boldsymbol{F}'\boldsymbol{F})^{-1}$, where **F** is

the $n \times u$ matrix with elements $\left(\partial f\left(x_{ijk}, \theta\right)/\partial\theta'\right)$ where the u columns of the matrix are 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 $n \geq p + r + 1$.

2.1. Split-plot model with Johnson–Schumacher function (JSF) as the mean curve

The mean curve $f(x_{ijk}, \theta)$ in equation (1) is replaced with the JSF. The JSF used for this research is a three-parameter function.

Let $f(x_{ijk}, \theta)$ be a JSF. Therefore,

$$f(x_{ijk}, \theta) = \alpha_{ijk} \times exp\left[-\left(\omega\left(\lambda + x_{ijk}\right)^{-1}\right)\right], \tag{2}$$

where α_{ijk} is the asymptote, and it is of following form: $\alpha_{ijk} = \alpha + Rep_i + A_j + B_k + (AB)_{jk}$. Hence, equation (2) can be rewritten as follows:

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

The SPD model with the JSF as the mean curve is therefore given as follows:

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

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

2.2. Generalized least square technique (GLST)

The GLST estimator, $\hat{\theta}_{GLS}$, is found by minimizing the objective function with respect to θ (Gumpertz and Rawlings, 1992; David *et al*., 2019):

$$(y - f(X,\theta))^t \boldsymbol{V}^{-1}(y - f(X,\theta)),$$

and when the covariance matrix of y is known, a known positive definite covariance matrix **V** derives from the model

$$y_{ijk} = f(x_{ijk},\theta) + w_{ij} + \varepsilon_{ijk}, \tag{3}$$

where E(w_{ij}) = 0, Cov(w_{ij}) = $\sigma_w^2\mathbf{I}_N$, E(ε_{ijk}) = 0 and Cov(ε_{ijk}) = $\sigma_\varepsilon^2\mathbf{I}_N$.
Let the observations var(y) for the matrix **V** be given as **V** = $\sigma_w^2\mathbf{I}_N + \sigma_\varepsilon^2\mathbf{I}_N$, , $\sigma^2\mathbf{I}$.
By Cholesky decomposition, multiplying model (3) by L^{-1} on both sides yields

$$L^{-1}y_{ijk} = L^{-1}f(x_{ijk},\theta) + L^{-1}(w_{ij}) + L^{-1}(\varepsilon_{ijk}). \tag{4}$$

Define $\mathbf{T}_{ijk} = L^{-1}y_{ijk}$, $\mathbf{M}(x_{ijk},\theta*) = L^{-1}f(x_{ijk},\theta)$ and $\boldsymbol{\Omega}_{ijk} = L^{-1}(w_{ij}) + L^{-1}(\varepsilon_{ijk})$. Then equation (4) becomes

$$\mathbf{T}_{ijk} = \mathbf{M}(x_{ijk},\theta^*) + \boldsymbol{\Omega}_{ijk}, \tag{5}$$

where $E(\boldsymbol{\Omega}_{ijk}) = 0$ and $\mathbf{V}(\boldsymbol{\Omega}_{ijk}) = \sigma^2\mathbf{I}$. Thus the GLS model has been transmuted to an OLS model. Hence, model (5) is to be solved using the OLS technique as follows. Take the summation of both sides of (5) and square to obtain

$$\sum_i^s\sum_j^a\sum_k^b \boldsymbol{\Omega}_{ijk}^2 = \sum_i^s\sum_j^a\sum_k^b\left[\boldsymbol{T}_{ijk} - \boldsymbol{M}(x_{ijk},\theta*)\right]^2.$$

Let $L(\theta^*) = \sum_i^s\sum_j^a\sum_k^b \boldsymbol{\Omega}_{ijk}^2 = \sum_i^s\sum_j^a\sum_k^b\left[\mathbf{T}_{ijk} - \mathbf{M}(x_{ijk},\theta*)\right]^2$ and partially differentiate $L(\theta^*)$ w.r.t. θ^* to obtain

$$\frac{\partial L(\theta^*)}{\partial\theta_h^*} = \sum_i^s\sum_j^a\sum_k^b\left[\boldsymbol{T}_{ijk} - \boldsymbol{M}(x_{ijk},\theta^*)\right] \times \left[\frac{\partial \boldsymbol{M}(x_{ijk},\theta^*)}{\partial\theta_h^*}\right]_{\theta^*=\widehat{\theta}^*} = 0. \tag{6}$$

At this point, solutions of equation (6) in closed form do not exist. It will be solved iteratively using the Gauss–Newton method with Taylor series expansion of $\mathbf{M}(x_{ijk},\theta^*)$ to the first order. This corresponds to

$$\boldsymbol{M}(x_{ijk},\theta^*) = \boldsymbol{M}(x_{ijk},\theta_0^*) + (\theta_1^* - \theta_{10}^*)\frac{\partial \boldsymbol{M}(x_{ijk},\theta^*)}{\partial\theta_1^*}\Big|_{\theta^*=\theta_0^*} + (\theta_2^* - \theta_{20}^*)\frac{\partial \boldsymbol{M}(x_{ijk},\theta^*)}{\partial\theta_2^*}\Big|_{\theta^*=\theta_0^*} + \ldots + (\theta_h^* - \theta_{h0}^*)\frac{\partial \boldsymbol{M}(x_{ijk},\theta^*)}{\partial\theta_h^*}\Big|_{\theta^*=\theta_0^*} . \tag{7}$$

Let $\mathbf{M}(x_{ijk},\theta^*) = \eta(\theta^*)$ and $d_{ijk} = \frac{\partial \mathbf{M}(x_{ijk},\theta^*)}{\partial\theta_h^*}\Big|_{\theta^*=\theta_0^*}$ for all N cases and $\delta = \theta^* - \theta_0^*$, then (7) becomes

$$\eta(\theta^*) = \eta(\theta_0^*) + D_0\delta,$$

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

$$\boldsymbol{\Omega}(\theta^*) = \boldsymbol{T} - [\eta(\theta_0^*) + D_0\delta]_{=} z_0 - D_0\delta,$$

where $z_0 = \boldsymbol{T} - \eta(\theta_0^*)$ 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 is reached where $\hat{\eta}_1 = \eta(\theta_1^*) = \eta(\theta_0^* + \delta_0)$ should now be closer to y than $\eta(\theta_0^*)$; then move to a better parameter value $\theta_1^* = \theta_0^* + \delta_0$ and perform another iteration by calculating new residuals $z_1 = T - \eta(\theta_1^*)$, a new derivative matrix D_0, and a new increase. This process is reiterated until convergence is achieved (Bates and Watts, 1988).

2.3 REML estimation of the variance component

The REML technique does not include $\hat{\theta}^*$ in the estimation of **V**. The log-likelihood function is based on vectors in the error space, and to obtain these vectors the linear approximation of the residuals is used, that is,$z_0 = D_0\delta + \varepsilon$. Vectors of the form $K^t y$ are formed where K is chosen so that $K^t D_0 = 0$, which falls in the linear approximation to the error space. Elements of $K^t y$ are called error contrasts (Harville, 1977), that is, the part of the data that is not dependent on the values of the fixed effect estimates; K is a full rank matrix satisfying $K^t D_0 = 0$. Applying maximum likelihood to $K^t y$, the log-likelihood function of $K^t y$ is

$$ln\, L(\boldsymbol{\theta}) = -\frac{n}{2} ln\,(2\pi)\; - \frac{1}{2} ln\,|K^t\, V\, K| - \frac{1}{2}\left(K^t y - K^t f(X,\theta)\right)^t$$
$$\times\, (K^t\, V\, K)^{-1}\left(K^t y - K^t f(X,\theta)\right), \qquad (8)$$

where $\boldsymbol{\Theta}^t = (\sigma_{WP}^2, \sigma_{SP}^2)^t$, and by surface approximation and letting ln L be Γ, equation (8) becomes

$$\ln L(\Theta) = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathrm{K}^t\ \mathrm{V}\ \mathrm{K}| - \frac{1}{2}\left(\mathrm{K}^t y - \mathrm{K}^t f(x,\theta)\right)^t (\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1}\left(\mathrm{K}^t y - \mathrm{K}^t f(x,\theta)\right) = C - \frac{1}{2}\ln|\mathrm{K}^t\ \mathrm{V}\ \mathrm{K}| - \frac{1}{2}\left(\mathrm{K}^t y - \mathrm{K}^t f(x,\theta)\right)^t (\mathrm{K}^t y(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1} - \mathrm{K}^t f(x,\theta)(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1})$$

$$= C - \frac{1}{2}\ln|\mathrm{K}^t\ \mathrm{V}\ \mathrm{K}| - \frac{1}{2}((\mathrm{K}^t y - \mathrm{K}^t f(x,\theta))^t \mathrm{K}^t y(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1}) + \frac{1}{2}((\mathrm{K}^t y - \mathrm{K}^t f(x,\theta))^t \mathrm{K}^t f(x,\theta)(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1}). \quad (9)$$

Expanding the third and fourth terms in equation (9) gives

$$\mathrm{In}L(\boldsymbol{\Theta}) = C - \frac{1}{2}ln\ |\mathrm{K}^t\ \mathrm{V}\ \mathrm{K}| - \frac{1}{2}(y^t \mathrm{K}^t(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1} y\mathrm{K} - f(x,\theta)^t \mathrm{K}^t(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1} y\mathrm{K}) + \frac{1}{2}(y^t \mathrm{K}^t(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1} f(x,\theta)\mathrm{K} - f(x,\theta)^t \mathrm{K}^t(\mathrm{K}^t\ \mathrm{V}\ \mathrm{K})^{-1} f(x,\theta)\mathrm{K}) \quad (10)$$

$\mathbf{V} = \sigma^2\mathbf{I} = \mathrm{K}\sum_{i=1}^{2}\sigma_i^2 \mathbf{I}_\mathrm{i}\mathbf{I}_\mathrm{i}^\mathrm{t}\,\mathrm{K}^\mathrm{t}$ and $\mathbf{V}\mathbf{V}^{-1}$ can be expressed as follows:

$$VV^{-1} = (KVK^t)(KVK^t)^{-1} = V(K(KVK^t)^{-1}K^t) = \sigma^2 I(Q_h) = \sigma^2\left(Q_h V_j\right)$$

Inserting **V** into equation (10) gives

$$\mathrm{In}L(\Theta) = C - \frac{1}{2}ln\left|\mathrm{K}^\mathrm{t}\ \sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\ \mathrm{K}\right| - \frac{1}{2}\left(y^t \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\ \sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\ \mathrm{K}\right)^{-1} y\mathrm{K} - f(x,\theta)^t \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\ \sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\ \mathrm{K}\right)^{-1} y\mathrm{K}\right) + \frac{1}{2}\left(y^t \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\ \sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\ \mathrm{K}\right)^{-1} f(x,\theta)\mathrm{K} - f(x,\theta)^t \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\ \sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\ \mathrm{K}\right)^{-1} f(x,\theta)\mathrm{K}\right). \quad (11)$$

By transformation, all other terms in (11) become zero since $\mathrm{K}^\mathrm{t}D_0 = \mathrm{K}^\mathrm{t} f(x,\theta) = \mathrm{K}f(x,\theta)^t = 0$. Partial differentiation of (11) w.r.t. σ_i^2 and equating to zero gives

$$\frac{\partial\, ln\, L(\Theta)}{\partial\sigma_i^2} = -\frac{1}{2}\frac{1}{|\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\mathrm{K}|}\left(\mathrm{K}^\mathrm{t}\mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\mathrm{K}\right) + \frac{1}{2}y^t \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\,\mathrm{K}\right)^{-1}\left(\mathrm{K}^\mathrm{t}\mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\mathrm{K}\right)\left(\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\,\mathrm{K}\right)^{-1} y\mathrm{K}\ \frac{1}{2}\left(\frac{1}{|\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\mathrm{K}|}\left(\mathrm{K}^\mathrm{t}\mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\mathrm{K}\right)\right) = \frac{1}{2}y^t \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\,\mathrm{K}\right)^{-1}\left(\mathrm{K}^\mathrm{t}\mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\mathrm{K}\right)\left(\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\,\mathrm{K}\right)^{-1} y\mathrm{K} \quad (12)$$

Let $\mathrm{Q_h V} = \mathrm{K}^\mathrm{t}\left(\mathrm{K}^\mathrm{t}\sum_{i=0}^{2}\sigma_i^2 \mathrm{I}_\mathrm{i}\mathrm{I}_\mathrm{i}^\mathrm{t}\,\mathrm{K}\right)^{-1}\mathrm{K}$; then equation (12) becomes

$$\frac{1}{2}\left(tr(Q_h V_i)\right) = \frac{1}{2}(y^t Q_h V_i Q_h y). \quad (13)$$

Multiplying the left-hand side of equation (13) by $\mathbf{VV}^{-1}$ we have

$$\frac{1}{2}\left(tr(Q_h V_i)\right)\sigma^2_{j(h+1)}\left(Q_h V_j\right) = \frac{1}{2}(y^t Q_h V_i Q_h y),$$

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

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

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

2.4 Median Adequacy Measure (MAM) statistics

Four proposed Median Adequacy Measure (MAM) statistics for assessing the adequacy of linear SPD models (David *et al*., 2016; David *et al*., 2020) are used in this research to assess the adequacy of the fitted NSPD models. The four statistics used are resistant coefficient of determination (r_r^2) as proposed by Kvalseth (1985), resistant prediction coefficient of determination (Pred-r_r^2), resistant modeling efficiency (*RMEF*) and median square error prediction (*MedSEP*). Procedures for calculating the respective WP and SP model residuals are given by Almimi *et al*. (2006) and David *et al*. (2016). These statistics are called resistant due to their ability to withstand 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.

2.4.1 Median square error prediction (MedSEP)

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

$$WPMedSEP = (n)^{-1}_{WP}\left(M^n_{i=1}\left[\left|Y_i - f\left(X_i \ldots, X_p\right)_i\right|\right]_{WP}\right)^2,$$

$$SPMedSEP = (n)^{-1}_{SP}\left(M^n_{i=1}\left[\left|Y_i - f\left(X_i \ldots, X_p\right)_i\right|\right]_{SP}\right)^2$$

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

2.4.2 Resistant coefficient of determination (r_r^2)

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

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

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

where M is the median of the absolute values from i = 1, …, n, and e_i are the fitted model's residuals. The statistics above use the median instead of the mean to obtain a coefficient of determination value that is highly resistant to outliers, $0 \leq r_r^2 \leq 1$. However, for nonlinear models the coefficient of determination value may be negative when the fit is worse, that is, $-1 \leq r_r^2 \leq 1$.

2.4.3 Resistant prediction coefficient of determination (Pred-r_r^2)

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

$$_{WP}Pr\ e\ d - r_r^2 = 1 - \left(M_{i=1}^{n} \left[\frac{(|e_i|)}{(1 - h_{ii})} \right]_{WP}^2 M_{i=1}^{n} \left[\left| Y_{i(WP)} - \bar{Y}_{(WP)} \right| \right]^{-2} \right)$$

$$_{SP}Pr\ e\ d - r_r^2 = 1 - \left(M_{i=1}^{n} \left[\frac{(|e_i|)}{(1 - h_{ii})} \right]_{SP}^2 M_{i=1}^{n} \left[\left| Y_{i(SP)} - \bar{Y}_{(SP)} \right| \right]^{-2} \right)$$

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 $1 \leq Pr\ e\ d - r_r^2 \leq 1$. However, for nonlinear models the prediction coefficient of determination value may be negative when the fit is worse, that is, $-1 \leq Pr\ e\ d - r_r^2 \leq 1$.

2.4.4 Resistant modeling efficiency (RMEF)

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

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

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

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

2.5 Information criterion statistics

In this research, Akaike's Information Criterion (AIC), Corrected AIC (AICC) and the Bayesian Information Criterion (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.

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

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

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

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.

2.6 Data and analysis procedures

The data used for this research is a balanced 3^1 x 4^2 replicated mixed level SP experimental design data sourced from David *et al*. (2023). The WP has two factors, which are irrigation and rice varieties. The irrigation was administered for three different times, 7 days, 14 days, and 21 days, on four different rice varieties, NERICA II, NERICA III, NERICA IV, and NERICA V. 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 the effect of irrigation on the yield of rice. The procedures for the analysis are as follows.

1. A traditional SP experimental design analysis is performed. 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 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 a regression model. The reason is that the size of parameters to be estimated will be too large for meaningful nonlinear modeling and interpretation of the results. At this stage, the main effects and their significant interaction effects, the WP and SP **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. 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 are 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 to obtain the results for EGLS, while the Proc Nlin code is used to obtain the OLS results.
5. The residuals obtained from each fitted NSPD model are used to calculate all four median adequacy measures introduced above, to assess the adequacy of all fitted models so as to identify a more adequate model.

3. Results and discussion

Table 1 presents the JSSPD model parameter estimates, standard errors and P-values from OLS, EGLS-MLE, and EGLS-REML. It is seen from Table 1 that the parameter estimates obtained by the OLS and EGLS-REML estimation

techniques are quite similar, but the EGLS-REML values are much different from those obtained by the other two techniques. The model mean estimates, 24.4307, 25.9908 and 26.2591 respectively from the three techniques, OLS, EGLS-MLE, and EGLS-REML, are similar, but their respective p-values of 0.6753, 0.9344 and 0.0004 show that only the EGLS-REML mean estimate is significant at a 5% significance level.

In turn, the replicate parameter estimates of 0.8988, 5.5832 and 0.5908 with p-values of 0.1150, 0.9517 and 0.3253 for OLS, EGLS-MLE, and EGLS-REML respectively are not significant at the 5% significance level. Similarly, the variety parameter estimates of –0.3009, 5.5299 and –0.5694 with p-values of 0.6910, 0.9955 and 0.9171 for OLS, EGLS-MLE, and EGLS-REML respectively are not significant at the 5% significance level. For the nitrogen fertilizer effect the parameter estimates of 0.06717, –0.2413 and 0.0636 with p-values of 0.0071, 0.9961 and 0.6415 for OLS, EGLS-MLE, and EGLS-REML respectively are not significant at the 5% significance level except for the OLS estimate, whose p-value is less than 5%.

In Table 1, I×V interaction parameter estimates from OLS (–0.04420), EGLS-MLE (0.01518), and EGLS-REML (–0.00407) are not significant because their p-values of 0.1588, 0.9997 and 0.99 are all greater than the 5% significance level. Correspondingly, I×N interaction parameter estimates of –0.00414, –0.2863 and –0.00333 with p-values of 0.0001, 0.9110 and 0.732 from OLS, EGLS-MLE, and EGLS-REML respectively are not significant except for the OLS estimate, whose p-value is less than 5%. Also, V×N interaction effect parameter estimates of 0.01286, 2.4796 and 0.01335 with p-values of 0.0909, 0.2017 and 0.0993 from OLS, EGLS-MLE, and EGLS-REML are not significant at the 5% significance level. The OLS estimate for the scale parameter (ω) of –0.5619 is smaller than the EGLS estimates obtained via MLE (28.1247) and REML (2.7119). Their respective p-values of 0.7549, 0.1697 and 0.6008 indicate that the scale parameter estimates are not significant at the 5% significance level. Similarly, the shape parameter (λ) estimate from OLS (7.5687) is larger than the estimates from EGLS-MLE (–7.0) and EGLS-REML (5.5521), but their

respective p-values of 0.8120, 0.3707 and 0.6257 imply that the shape parameter estimates are not significant at 5%.

Table 1. Johnson–Schumacher split-plot design (JSSPD) model parameter estimates

Parameter	OLS	EGLS (MLE)	EGLS (REML)	Std. Error a	Std. Error b	Std. Error c
$\boldsymbol{\alpha_0}$	24.4307	25.9908	26.2591	58.1364	314.75	7.1753
$\boldsymbol{\alpha_1}$	0.8988	5.5832	0.5908	0.5651	91.8495	0.5975
$\boldsymbol{\alpha_2}$	-0.3009	5.5299	-0.5694	0.7547	981.49	5.4567
$\boldsymbol{\alpha_3}$	0.06717	-0.2413	0.0636	0.02443	49.6582	0.1362
$\boldsymbol{\alpha_4}$	-0.04420	0.01518	-0.00407	0.03112	46.9679	0.3235
$\boldsymbol{\alpha_5}$	-0.00414	-0.2863	-0.00333	0.000965	2.5535	0.009693
$\boldsymbol{\alpha_6}$	0.01286	2.4796	0.01335	0.007532	1.9290	0.008018
$\boldsymbol{\omega}$	-0.5619	28.1247	2.7119	1.7943	20.3298	5.1654
$\boldsymbol{\lambda}$	7.5687	-7.0000	5.5521	31.7379	7.7834	11.3443
$\hat{\boldsymbol{\sigma}}_{\boldsymbol{\delta}}^{\boldsymbol{2}}$	9.964E-6	0.6267	3.142	59.8104	0.1919	23.883
$\hat{\boldsymbol{\sigma}}_{\boldsymbol{\varepsilon}}^{\boldsymbol{2}}$	1.6851	7.1192	6.2981	0.06054	0.2227	1.1719

Parameter	P-value a	P-value B	P-value c
$\boldsymbol{\alpha_0}$	0.6753	0.9344	**0.0004**
$\boldsymbol{\alpha_1}$	0.1150	0.9517	0.3253
$\boldsymbol{\alpha_2}$	0.6910	0.9955	0.9171
$\boldsymbol{\alpha_3}$	**0.0071**	0.9961	0.6415
$\boldsymbol{\alpha_4}$	0.1588	0.9997	0.99
$\boldsymbol{\alpha_5}$	**<.0001**	0.9110	0.732
$\boldsymbol{\alpha_6}$	0.0909	0.2017	0.0993
$\boldsymbol{\omega}$	0.7549	0.1697	0.6008
$\boldsymbol{\lambda}$	0.8120	0.3707	0.6257
$\hat{\boldsymbol{\sigma}}_{\boldsymbol{\delta}}^{\boldsymbol{2}}$	1.0000	**0.0015**	0.8956
$\hat{\boldsymbol{\sigma}}_{\boldsymbol{\varepsilon}}^{\boldsymbol{2}}$	**<.0001**	**<.0001**	**<.0001**

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

The final whole plot variance component ($\hat{\sigma}_{\delta}^{2}$) parameter estimate of 3.142 from EGLS-REML is larger than the OLS estimate of 9.96E-6 and the EGLS-MLE estimate of 0.6267. However, the OLS and EGLS-REML estimates are not significant because their p-values of 1.0000 and 0.8956 are greater than the 5% significance level, although the EGLS-MLE p-value of 0.0015 implies significance at the 5% level. The SP variance component estimate ($\hat{\sigma}_{\varepsilon}^{2}$) from OLS (1.6851) is also smaller than the EGLS-MLE (7.1192) and EGLS-REML (6.2981) estimates, and their p-values of 0.0001 indicate a significant covariance parameter estimate for the subplot at the 5% significance level. Generally, the standard errors for each of the estimates from OLS, EGLS-MLE and EGLS-

REML in Table 4 show that EGLS-MLE produced larger standard errors than those of the OLS and EGLS-REML parameter estimates. This gives a preliminary indication that either the OLS or EGLS-REML estimates for the JSSPD model are estimated adequately with better stability. Hence, one of the OLS or EGLS-REML techniques is more efficient than REML-MLE. The OLS, EGLS-MLE and EGLS-REML estimated fitted models for the JSSPD model are as follows:

$$y_{ijkl} = [24.4307 + 0.8988Rep - 0.3009V + 0.06717N - 0.0442IV - 0.00414IN + 0.01286VN] \times exp\left[-\left(\frac{-0.5619}{I+7.5687}\right)\right].$$

$$y_{ijkl} = [25.9907 + 5.5832Rep + 5.5299V - 0.2413N + 0.01518IV - 0.2863IN + 2.4796VN] \times exp\left[-\left(\frac{28.1247}{I-7}\right)\right].$$

$$y_{ijkl} = [26.2591 + 0.5908Rep - 0.5694V + 0.0636N - 0.00407IV - 0.00333IN + 0.01335VN] \times exp\left[-\left(\frac{2.7119}{I+5.5521}\right)\right].$$

The estimated (OLS, EGLS-MLE and EGLS-REML) fitted JSSPD model adequacy measures for the WP and SP subdesign models are also presented in Table 2. The results reveal that the EGLS-MLE fitted models for both the WP and SP subdesign models are worse than in the case of the OLS and EGLS-REML WP and SP subdesign models.

The WP subdesign model had larger adequacy measure values than the SP subdesign model for $\boldsymbol{r_r^2}$, Pred-r_r^2 and *RMEF*, while the *MedSEP* value for the WP sub design model is smaller than for the SP sub design model. However, all values for the fitted JSSPD models show that a large proportion of variability is explained in the data, with high prediction power, better model efficiency and better error prediction strength for the OLS and EGLS-REML fitted models.

Table 2. Median adequacy measure results

	$\boldsymbol{r_r^2}$		$\boldsymbol{Pred - r_r^2}$		**RMEF**		***MedSEP***	
	WP	SP	WP	SP	WP	SP	WP	SP
OLS	**0.9994**	**0.7293**	**0.9993**	**0.5874**	**0.9985**	**0.7713**	**2E-07**	**0.0782**
MLE	0.2422	-0.0505	0.1423	-0.601	**0.9959**	**0.7372**	0.3479	1.1781
REML	**0.9984**	**0.7825**	**0.9982**	**0.6686**	**0.9984**	**0.7769**	**1.6E-06**	**0.0505**

Table 3 below presents the goodness of fit results for the fitted model, and shows that EGLS-REML produced the lowest AIC, AICC and BIC values of 500.4, 503.6 and 490.5 respectively. This implies that the EGLS-REML estimation technique produces reliable and stable estimates compared with OLS and EGLS-MLE parameter estimates.

Table 3. Model goodness of fit test results

Method	AIC	AICC	BIC
OLS	515.8	518.9	524.0
MLE	940.2	943.3	930.2
REML	**500.4**	**503.6**	**490.5**

4. Conclusions

Based on the research results from the analysis of data on a balanced 3^1 x 4^2 replicated mixed level SP experimental design, it was observed that the fitted JSSPD model is a good fit for the EGLS-REML estimated model. This is indicated by all the four MAM values for assessing the adequacy of the fitted model: largerr_r^2, Pred-r_r^2 and *RMEF* values and smaller *MedSEP* values were obtained for the WP and SP submodels, as well as smaller AIC, AICC and BIC values, compared with the EGLS-MLE and OLS estimates. Also, all of the OLS, EGLS-MLE, and EGLS-REML estimated fitted INSPDMs for the WP sub models produced better adequacy measure values than in the case of the SP submodels. However, the respective standard errors of the OLS, EGLS-MLE, and EGLS-REML parameter estimates show that the EGLS-MLE parameter estimates are not stable and reliable, because of the high standard errors compared with those of the OLS and EGLS-REML parameter estimates.

Acknowledgement: The Editor and Reviewer are well appreciated for their insightful corrections that improved the work.

REFERENCES

Almimi A.A., Kulahci M., Montgomery D.C. (2006): Checking the adequacy of fit of models from split-plot designs. Journal of Quality Technology 38: 58–190.

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.

Bates D.M., Watts D.G. (1988): Nonlinear Regression Analysis and Its Applications, New York: Wiley.

Blankenship E.E., Stroup W.W., Evans S.P., Knezevic S.Z. (2003): Statistical Inference for Calibration Points in Nonlinear Mixed Effects Models. American Statistical Association and the International Biometric Society Journal of Agricultural, Biological, and Environmental Statistics, 8(4): 455–468.

David I.J., Adubisi O.D., Ogbaji O.E., Eghwerido J.T., 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., Asiribo O.E., 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., 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., 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., Asiribo O.E., Dikko H.G. (2022): A Bertalanffy-Richards Split-Plot Design Model and Analysis. Journal of Statistical Modeling and Analysis, 4(1): 56–71.

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

David I.J., Asiribo O.E., Dikko H.G. (2023): Nonlinear split-plot design modeling and analysis of rice varieties yield. Scientific African, 19, e01444.

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

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

Harville D.A. (1977): Maximum likelihood approaches to variance components estimation and to related problems. Journal of the American Statistical Association, 72: 320–338.

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

Klotz J.H. (2006): A computational approach to statistics. Madison: University of Wisconsin.

Knezevic S.Z., Evans S.P., Blankenship E.E., Van Acker R.C., Lindquist J.L. (2002): Critical period for weed control: the concept and data analysis. Agronomy – Faculty Publications. Paper 407.

Kulachi M.A. Menon (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. The American Statistician, 39: 279–285.