<GLM2>

Bayesian Estimation of Mixed Effects Models of Fertilizer Response with Independent Skew-Normally Distributed Random Parameter

Mohammad Masjkur[1,a)] and Henk Folmer[2, b)]

[1]*Department of Statistics, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University, Indonesia*
[2]*Faculty of Spatial Sciences, University of Groningen, The Netherlands*
Department of Agricultural Economics, College of Economics and Management, Northwest A&F University, Yangling, Shaanxi, China

[a)]Corresponding author: m_masjkur@yahoo.com
[b)]h.folmer@rug.nl

Abstract. The mixed effects model has been used for modelling the fertilizer response to predict the optimum doses. However, a major restriction of this type of models is the normality assumption of the random parameter component. The purpose of this paper is to investigate the performance of random parameter models of fertilizer dosing with independent skew-normally distributed random parameter components. We compare the Linear Plateau, Spillman-Mitscherlich, and Quadratic random parameter models with different random effects distribution assumption, i.e. the normal, Skew-normal, Skew-*t*, Skew-slash, and Skew-contaminated distributions and the random errors following symmetric normal independent distributions. The method is applied to datasets of multi-location trials of potassium fertilization of soybeans. The results show respectively that the Skew-*t* Model is the best Linear Plateau Response Model, the Normal Model for Spillman-Mitscherlich Response Model, and the Skew-*t* Model for the Quadratic Response Model. However, overall the normal Spillman-Mitscherlich Response Model is the best model for soybean yield prediction.

Keywords: Bayesian estimation, Dose-response model, Random parameter model, Skew-normal independent distributions.

INTRODUCTION

Many linear and nonlinear functions have been used for describing multi-environment crop response to fertilizer, such as linear plateau and quadratic functions. The model parameters usually estimated using least squares method assuming that the model has a fixed effect and the random error terms were independent and normally distributed with a constant variances ([1]-[2]). However, this approach is unrealistic because it ignores the variability that probably exist between site-years.

An alternative model is the mixed effects approach ([3]-[6]). This approach allows the parameters to have a random effect component that represent between sites or years variability. The random parameter models have been found to outperform the fixed parameter models to model dose-response relationships ([5], [7]-[8]). Furthermore, the quadratic functional form commonly used is not always the best model. [7] and [9] showed that the stochastic linear plateau model and the Mitscherlich exponential type functions outperform the quadratic form. In a similar way, [8] showed that the stochastic linear plateau function is more adequate than the stochastic quadratic plateau function for corn response to Nitrogen fertilizer.

The random parameter components and the errors are usually taken as normally distributed random variables ([5]-[8]). However, the normality and symmetry assumptions may be too restrictive because in practice departures from normality is common. Particularly, [10] and [11] concluded that the field crop yield distributions are in general non-normal or non-lognormal. The degree of skewness and kurtosis vary by crop and by the amount of nutrients uptake. In addition, (random) weather effects could result in positively or negatively skewed probability functions. Therefore, [12] suggested the beta distribution for the random parameter component of the linear plateau function of wheat response to Nitrogen fertilizer.

Lachos et al. [13] advocated the use of the Skew-normal independent distribution for robust modeling of linear mixed models. The Skew-normal independent distribution is a class of asymmetric, heavy-tailed distributions that includes the Skew-normal distribution, Skew-*t,* Skew-slash and the Skew-contaminated normal distributions. The class of Skew-normal distributions accommodate observations with high skewness and heavy tails as well as the normal distribution.

Traditionally, fertilizer-dose response models are estimated by means of maximum likelihood estimation (ML) ([5]-[8]). However, for nonlinear models and small sample sizes ML is frequently biased ([14]). In addition, convergence can be a problem even with careful scaling and good starting values. Bayesian estimation is an alternative to ML. The advantages of Bayesian estimation are that the results are valid in small samples and that convergence in the case of nonlinear models is not an issue ([12], [14]-[15]).

The purpose of this paper is Bayesian estimation of random parameter dose (fertilization)-response (yield) models for yield data that is Skew heavy-tailed distributed.

The Normal Mixed Effects Model

In general, a Normal mixed effects model reads:

$$\boldsymbol{Y_i} = \eta(\boldsymbol{\phi_i}, \boldsymbol{X_i}) + \boldsymbol{\epsilon_i}, \qquad \boldsymbol{\phi_i} = \boldsymbol{A_i\beta} + \boldsymbol{B_i b_i}, \tag{1}$$

with

$$(\boldsymbol{b_i}, \boldsymbol{\epsilon_i}) \overset{ind}{\sim} N_{n_i+q}\left(\boldsymbol{0}, Diag(\boldsymbol{\Sigma}, \sigma_e^2 \boldsymbol{I}_{n_i})\right),$$

where the subscript i is the subject index, $i = 1, \ldots\ldots, n$; $\boldsymbol{Y_i} = (y_{i1}, \ldots, y_{in_i})^T$ is a $n_i \times 1$ vector of n_i observed continuous responses for subject i, $\eta_i(\boldsymbol{\phi_i}, \boldsymbol{X_i}) = \{\eta(\phi_i, X_{i1}), \ldots, \eta(\phi_i, X_{in_i})\}^T$ with $\eta(.)$ the nonlinear or linear function of random parameters $\boldsymbol{\phi_i}$, and covariate vector $\boldsymbol{X_i}$, $\boldsymbol{A_i}$ and $\boldsymbol{B_i}$ are known design matrices of dimensions $n_i \times p$ and $n_i \times q$, respectively, $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed effects, $\boldsymbol{b}_i$ is the $q \times 1$ vector of random effects, and $\boldsymbol{\epsilon_i}$ is the $n_i \times 1$ vector of random errors, and $\boldsymbol{I}_{\boldsymbol{n_i}}$ denotes the identity matrix. The matrices $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\alpha})$ with unknown parameter $\boldsymbol{\alpha}$ is the $q \times q$ unstructured dispersion matrix of $\boldsymbol{b_i}$, σ_e^2 the unknown variance of the error term. When $\eta(.)$ is a nonlinear parameter function, we have the Normal NonLinear Mixed Model (N-NLMM); if $\eta(.)$ is a linear parameter function, we have the N-Linear Mixed Model (N-LMM).

It follows that

$$\boldsymbol{b_i} \overset{ind}{\sim} \boldsymbol{N_q}(\boldsymbol{0}, \boldsymbol{\Sigma}) \text{ and } \boldsymbol{\epsilon_i} \overset{ind}{\sim} \boldsymbol{N_{n_i}}(\boldsymbol{0}, \sigma_\epsilon^2 \boldsymbol{I}_{n_i})$$

and they are uncorrelated, since $Cov(\boldsymbol{b_i}, \boldsymbol{\epsilon_i}) = \boldsymbol{0}$ ([16]-[17]).

Skew-Normal Independent (SNI) Distributions

A skew-normal independent distribution is defined as the *p*-dimensional random vector $\boldsymbol{y} = \boldsymbol{\mu} + U^{1/2}\boldsymbol{Z}$, where $\boldsymbol{\mu}$ is a location vector, $\boldsymbol{Z}$ is a multivariate skew-normal random vector with location vector $\boldsymbol{0}$, scale matrix $\boldsymbol{\Sigma}$ and skewness parameter vector $\boldsymbol{\lambda}$, i.e. $\boldsymbol{Z} \sim SN_p(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$ ([12]). Furthermore, U is a positive weight random variable with cumulative distribution function (cdf) $H(u|v)$ and probability density function (pdf) $h(u|v)$, $\boldsymbol{v}$ is a scalar or vector of parameters indexing the distribution of the scale factor U. Given U, $\boldsymbol{Y}$ follows a multivariate skew-normal distribution with location vector $\boldsymbol{0}$, scale matrix $u^{-1}\boldsymbol{\Sigma}$ and skewness parameter vector $\boldsymbol{\lambda}$, i.e., $\boldsymbol{Y}|U = u \sim SN_p(\boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}, \boldsymbol{\lambda})$. Thus, the SNI distributions are scale mixtures of the skew-normal distributions denoted by $\boldsymbol{Y} \sim SNI_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, H)$. The marginal pdf of $\boldsymbol{Y}$ is

$$f(\boldsymbol{y}) = 2\int_0^\infty \phi_p(\boldsymbol{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma})\Phi\left(u^{1/2}\boldsymbol{\lambda}^T\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\right)dH(u|\boldsymbol{v}),$$

The skew-normal independent distribution is a group of asymmetric heavy-tailed distribution of robust alternative to the routinely used of normal distribution for mixed effects model ([18]-[21]). A convenient stochastic representation of $\boldsymbol{Y}$, follows from [19]-[20]:

$$\boldsymbol{Y} = \boldsymbol{\mu} + \Delta\boldsymbol{T} + \boldsymbol{\Gamma}^{1/2}\boldsymbol{T_1}, \tag{2}$$

where $\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\delta}, \boldsymbol{\Gamma} = \boldsymbol{\Sigma}^{1/2}(\boldsymbol{I} - \boldsymbol{\delta}\boldsymbol{\delta}^T)\boldsymbol{\Sigma}^{1/2} = \boldsymbol{\Sigma} - \boldsymbol{\Delta}\boldsymbol{\Delta}^T$, $\boldsymbol{I}$ denotes the identity matrix and $\boldsymbol{\delta} = \boldsymbol{\lambda}/(1 + \boldsymbol{\lambda}^T\boldsymbol{\lambda})^{1/2}$, $\boldsymbol{\lambda} = \frac{(\boldsymbol{\Gamma}+\Delta\Delta^T)^{-1/2}\Delta}{[1-\Delta^T(\boldsymbol{\Gamma}+\Delta\boldsymbol{\Delta}^T)^{-1}\Delta]^{1/2}}$, $\boldsymbol{\Sigma} = \boldsymbol{\Gamma} + \boldsymbol{\Delta}\boldsymbol{\Delta}^T, T = |T_0|, T_0 \sim N_1(0,1)$ and $\boldsymbol{T_1} \sim N_p(\boldsymbol{0}, \boldsymbol{I_p})$.

When $\boldsymbol{\lambda} = 0$, the class of SNI distributions reduces to the class of thick-tailed normal independent (NI) distributions ([22]-[24]). The probability density function (pdf) is $f_0(\boldsymbol{y}) = \int_0^\infty \phi_p(\boldsymbol{y}; \boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma})dH(u|\boldsymbol{v})$, denoted as $\boldsymbol{Y} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, H)$.

The SNI-Mixed Effects Model

Using the general framework (1), the following general SNI-Mixed Model (SNI-MM) is defined as:

$$\boldsymbol{b_i} \overset{iid}{\sim} SNI_q(\boldsymbol{0}, Diag(\boldsymbol{\Sigma}), \boldsymbol{\lambda}, H) \text{ and } \boldsymbol{\epsilon_i} \overset{ind}{\sim} NI_{n_i}(\boldsymbol{0}, \sigma_e^2\boldsymbol{I}_{n_i}, H), i = 1, \ldots\ldots, n.$$

where the random effects are assumed to have a multivariate SNI distributions and the random errors are assumed to have a NI distribution.

Prior Distributions and Joint Posterior Density

Below, we apply a Bayesian framework based on the Markov Chain Monte Carlo (MCMC) algorithm to infer posterior parameter estimates. Using the representation (2), the general mixed model can be formulated in hierarchical for $i = 1, \ldots\ldots, n,$ as follows:

$$\boldsymbol{Y_i}|\boldsymbol{b_i}, U_i = u_i \overset{ind.}{\sim} N_{n_i}\left(\eta(\boldsymbol{A_i\beta} + \boldsymbol{B_i b_i}, \boldsymbol{X_i}), u_i^{-1}\sigma_e^2\boldsymbol{I}_{n_i}\right).$$

$$\boldsymbol{b_i}|T_i = t_i, U_i = u_i \overset{ind.}{\sim} N_q(\boldsymbol{\Delta}t_i, u_i^{-1}\boldsymbol{\Gamma}).$$

$$T_i|U_i = u_i \overset{ind.}{\sim} HN_1(0, u_i^{-1}).$$

$$U_i \overset{iid.}{\sim} H(u_i|\boldsymbol{v}).$$

where $HN_1(0,\sigma^2)$ is the half- $N_1(0,\sigma^2)$ distribution, $\boldsymbol{\Delta} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\delta}$ and $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} - \boldsymbol{\Delta}\boldsymbol{\Delta}^{\boldsymbol{T}}$, with $\boldsymbol{\delta} = \boldsymbol{\lambda}/(1 + \boldsymbol{\lambda}^T\boldsymbol{\lambda})^{1/2}$ and $\boldsymbol{\Sigma}^{1/2}$ the square root of $\boldsymbol{\Sigma}$ containing $q(q+1)/2$ distinct elements (17], [19], [24]).

Let $\boldsymbol{Y} = (\boldsymbol{y}_1^T, \ldots, \boldsymbol{y}_n^T)^T$, $\boldsymbol{b} = (\boldsymbol{b}_1^T, \ldots, \boldsymbol{b}_n^T)^T$, $\boldsymbol{u} = (u_1, \ldots, u_n)^T$, $\boldsymbol{t} = (t_1, \ldots, t_n)^T$. Then, the complete likelihood function associated with $(\boldsymbol{y}^T, \boldsymbol{b}^T, \boldsymbol{u}^T, \boldsymbol{t}^T)^T$, is given by

$$L(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{b}, \boldsymbol{u}, \boldsymbol{t}) \propto \prod_{i=1}^{n}[\phi_{n_i}(\boldsymbol{y}_i; \eta(\boldsymbol{A_i\beta} + \boldsymbol{B_i b_i}, \boldsymbol{X}_i), u_i^{-1}\sigma_e^2 \boldsymbol{I}_{n_i})\phi_q(b_i; \boldsymbol{\Delta} t_i, u_i^{-1}\boldsymbol{\Gamma}) \times \ \phi_1(t_i; 0, u_i^{-1})h(u_i|\boldsymbol{v})].$$

To complete Bayesian specification, we need to consider prior distributions for all the unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \sigma_e^2, \boldsymbol{\alpha}^T, \boldsymbol{\lambda}^T, \boldsymbol{v}^T)^T$. We consider $\boldsymbol{\beta} \sim N_p(\boldsymbol{\beta}_0, \boldsymbol{S}_\beta)$, $\sigma_e^2 \sim IG(q_0/2, \lambda_0/2)$, $\boldsymbol{\Gamma} \sim IW_q(\Lambda_0^{-1}, v_0)$, $\boldsymbol{\Delta} \sim N_p(\boldsymbol{\Delta}_0, \boldsymbol{S}_\Delta)$ ([17], [19]). For v we take $v \sim Exp(\tau/2)\mathbb{I}_{(2,\infty)}$ for the Skew-t (St) model, Gamma (a, b) for the Skew-slash (SSL) model. Furthermore, $U(0,1)$ for v_1 and Beta (ρ_0, ρ_1) for v_2 for the Skew-contaminated normal (SCN) model.

Assuming independency of the parameter vector, the joint prior distribution of all unknown parameters is

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\sigma_e^2)\pi(\boldsymbol{\Gamma})\pi(\boldsymbol{\Delta})\pi(\boldsymbol{v}).$$

Combining the likelihood function and the prior distribution, the joint posterior density of all unknown parameters is

$$\pi(\boldsymbol{\beta}, \sigma_e^2, \boldsymbol{\Gamma}, \boldsymbol{\Delta}, \mathbf{b}, \mathbf{u}, \mathbf{t}|\mathbf{y}) \propto \prod_{i=1}^{n}[\phi_{n_i}(\boldsymbol{y}_i; \eta(\boldsymbol{A_i\beta} + \boldsymbol{B_i b_i}, \boldsymbol{X}_i), u_i^{-1}\sigma_e^2 \boldsymbol{I}_{n_i})\phi_q(\boldsymbol{b}_i; \boldsymbol{\Delta} t_i, u_i^{-1}\boldsymbol{\Gamma})$$

$$\times\, \phi_1(t_i; 0, u_i^{-1})h(u_i|\boldsymbol{v})]\pi(\boldsymbol{\theta}).$$

Model Comparison Criteria

The expected Akaike information criterion (EAIC) and the expected Bayesian information criterion (EBIC) are a deviance-based measure appropriate for Bayesian model selection ([25]-[26]).

Let $\boldsymbol{\theta}$ and $\boldsymbol{Y} = (y_1, \dots, y_n)^T$ be the entire model parameters and data, respectively. Define $D\,(\boldsymbol{\theta}) = -2\ln f(\boldsymbol{y}|\boldsymbol{\theta}) = -2\sum_{i=1}^{N}\ln f(\boldsymbol{y_i}|\boldsymbol{\theta})$, where $f(\boldsymbol{y_i}|\boldsymbol{\theta})$ is marginal distribution of y_i, then $E\,[D\,(\boldsymbol{\theta})]$ is a measure of fit and can be approximated by using the MCMC output in a Monte Carlo simulation. This index is given by $\overline{D} = \frac{1}{K}\sum_{k=1}^{K} D(\boldsymbol{\theta}^{(k)})$. Where $\boldsymbol{\theta}^{(k)}$ is the k^{th} iteration of MCMC chain of the model and K is the number of iterations.

The expected Akaike information criterion (EAIC) and the expected Bayesian information criterion (EBIC) define as follows

$$\widehat{EAIC} = \overline{D} + 2p, \text{ and } \widehat{EBIC} = \overline{D} + p\; log(N)$$

where $\overline{D}$ is the posterior mean of the deviance, p is the number of parameters in the model, N is the total number of observations. These criteria penalizing models with more complexity. Smaller value of EAIC and EBIC indicate a better fit ([19]).

CASE STUDY

Data

The dataset is obtained from 19 multi-location trials of potassium fertilization of soybeans. The experiments were carried out between 2002 and 2014. The soil types are Ultisols, Inceptisols, Vertisols, and Oxisols with soil potassium contents varying from very low to very high. Common soybean varieties were used. Each experiments consisted of five levels of potassium fertilization. The doses applied were 0, 40, 80,

ISSN : 2579-4361

160 and 320 kg ha^{-1} of KCl. The plots were 6 by 5 m, or 4 by 5 m arranged in a randomized complete block design with three to nine replications. The response variable was soybean yield (t ha^{-1}). The yields reported are averages over replications ([27]-[29]).

Response functions

We consider three response functions: the Linear Plateau (LP), the Spillman-Mitscherlich (SM) and the Quadratic functions (Q).

The stochastic LP is defined as follows:

$$Y_i = \min\big(\alpha_1 + (\alpha_2 + b_{2i})X_i;\ \mu_p + b_{3i}\big) + b_{1i} + \varepsilon_i \tag{3}$$

where for location i, Y_i is the soybean yield; X_i the potassium fertilizer dose; α_1 the intercept parameter; α_2 the linear response coefficient; u_p the plateau yield; b_{1i} , b_{2i}, and b_{3i} are the random effects; and ε_i is the random error term. In term of (1), $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^T$ $\boldsymbol{b}_i = (b_{1i}, b_{2i}, b_{3i})^T$; $\boldsymbol{b_i} \overset{iid}{\sim} SNI_q(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}, H)$ and $\boldsymbol{\epsilon_i} \overset{ind}{\sim} NI_{n_i}(\boldsymbol{0}, \sigma_e^2 \boldsymbol{I}_{n_i}, H)$.

The stochastic SM reads:

$$Y_i = \beta_1 - (\beta_2 + b_{2i}) \exp\ ((-\beta_3 + b_{3i})\, X_i) + b_{1i} + \varepsilon_i \tag{4}$$

where β_1 is the maximum yield attainable by potassium fertilization; β_2 is the yield increase; β_3 is the ratio of consecutive increments of the yield; all other parameters, variables and distributions as in (3).

The stochastic Q is defined as:

$$Y_i = \gamma_1 + (\gamma_2 + b_{2i})X_i + (\gamma_3 + b_{3i})X_i^2 + b_{1i} + \varepsilon_i \tag{5}$$

where γ_1 is the intercept parameter whose position (value) can be shifted up or down by the random effect b_{1i}; γ_2 is the linear response coefficient with random effect parameter b_{2i}; γ_3 is the quadratic response coefficient whose position can be shifted up or down by the random effect b_{3i}; $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)^T$; all other variables and distributions as in (3) ([7]-[9]).

Statistical Analysis

The datasets was used to identify the model with the best fit among the random parameter models of fertilizer dosing. Several statistical models with differing distribution in the random effects and random errors are compared. These models are:

Model 1: Skew-normal distribution for the random effects and Normal distribution for the random errors (SN-N)

Model 2: Skew-t distribution for the random effects and Student-*t* distribution for the random errors (*St-t*)

Model 3: Skew-slash distribution for the random effects and slash distribution for the random errors (SSL-SL)

Model 4: Skew-contaminated normal distribution for the random effects and contaminated normal distribution for the random errors (SCN-CN).

Model 5: Normal distribution for the random effects and random errors (N-N)

The following independent priors were considered to perform the Gibbs sampler, $\boldsymbol{\beta_k} \sim N(0, 10^3)$, $\sigma^2 \sim IG(0.1, 0.1)$, $\boldsymbol{\Gamma} \sim IG(0.1, 0.1)$, $\boldsymbol{\Delta} \sim N(0, 0.001)$, and $v \sim Exp(0.1)I(2,)$ for the Skew-*t* model,

ISSN : 2579-4361

$v \sim Gamma(0.1,0.01)$ for the skew-slash model, $v_1 \sim Beta(1,1)$ and $v_2 \sim Beta\,(2,2)$ for the skew-contaminated normal model, respectively.

For each of the models, we ran three parallel independent chains of the Gibbs sampler with size 50 000 iterations for each parameter with thinning of 5 and an initial burn in of 25 000. We monitored chain convergence using trace plots, autocorrelation plots and the Brooks-Gelman-Rubin scale reduction factor ($\hat{R}$) ([30]). To avoid non-convergence, we normalized the original doses (subtracted the mean and divided by the standard deviation) which gave: -1.06, -0.70, -0.35, 0.35, and 1.76, respectively ([31]). We fitted the models using the R2jags package available in R ([32]).

RESULTS AND DISCUSSION

Soybean yield data

Figure 1 shows the histogram and normal Q-Q plot of soybean yield data for 19 locations, while the boxplot is presented in Fig. 2. The figures indicates non-normality (skew heavy-tailed) pattern. The Q-Q plot does not show a straight line, while the boxplot shows asymmetry and an outlier. Thus, it seems appropriate to fit a skewed heavy-tailed model to the data.

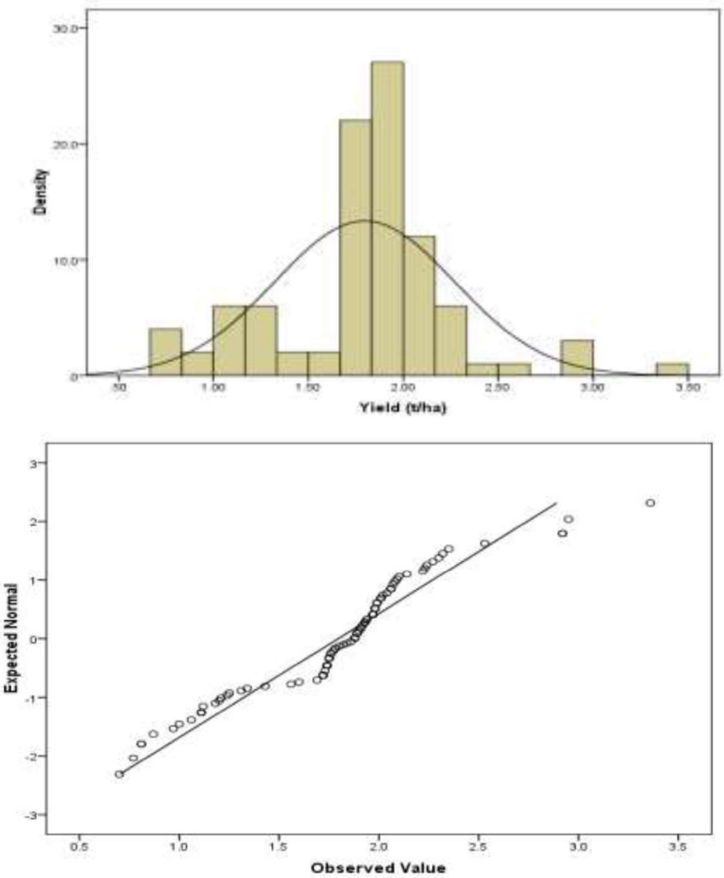

FIGURE 1. Histogram and Normal Q-Q plot of soybean yield data

ISSN : 2579-4361

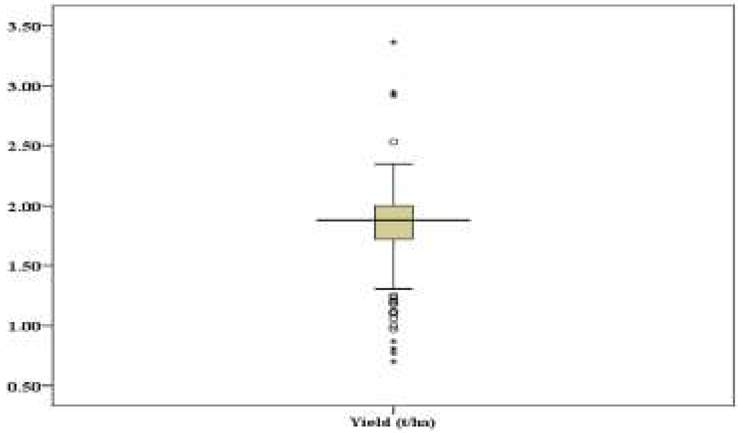

FIGURE 2. Boxplot of soybean yield data

Linear Plateau Response Models

Based on the EAIC and the EBIC in table 1, we find that among the SNI models the Skew-*t* (S*t*-*t*) Model gives the best fit, followed by the Skew-slash (SSL-SL), Skew-contaminated normal (SCN-CN) and Skew-normal (SN-N) Model. We furthermore find that the S*t*-*t* Model outperform the normal distributions. Thus, the S*t*-*t* Model is the best Linear Plateau Response Model.

TABLE 1. The Linear Plateau Models

Parameter	N-N		SN-N		S*t*-*t*		SSL-SL		SCN-CN	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
α_1	1.473	0.114	1.471	0.113	1.533	0.143	1.469	0.151	1.509	0.159
α_2	39.968	20.482	29.193	19.320	29.417	19.604	29.657	19.531	28.982	18.929
μ_p	1.878	0.129	1.880	0.134	1.843	0.217	1.826	0.234	1.882	0.246
σ^2_ε	0.139	0.014	0.020	0.004	0.015	0.004	0.013	0.004	0.014	0.005
d_1	0.467	0.095	0.250	0.114	0.173	0.089	0.148	0.076	0.160	0.093
d_2	13.135	11.964	49.074	237.868	30.677	129.636	30.903	121.481	26.718	92.546
d_3	0.306	0.081	0.121	0.074	0.071	0.045	0.062	0.039	0.069	0.046
λ_1			0.056	0.467	-0.006	0.262	-0.024	0.265	0.017	0.288
λ_2			0.059	0.654	0.005	0.338	-0.015	0.325	0.003	0.367
λ_3			0.175	1.036	-0.007	0.516	-0.039	0.529	0.036	0.573
ν (ν_1)					5.834	3.018	2.846	1.748	0.473	0.265
ν_2									0.513	0.215
EAIC	-93.56		-85.78		**-98.03**		-89.47		-88.35	
EBIC	-93.71		-86.00		**-98.28**		-89.71		-88.61	

Table 1 furthermore shows that for the St-t Model, all the fixed effects, i.e., the intercept parameter (α_1), the linear response coefficient (α_2), the plateau yield u_p and the random effects (d_1, d_2, d_3) are significant.

Spillman-Mitscherlich Response Models

Based on the EAIC and EBIC in table 2 we find the following rankings of the SNI models: S*t*-*t* < SSL-SL < SN-N < SCN-CN. However, we observe that the normal distributions outperform the S*t*-*t* Model, and that the asymmetry parameters $(\lambda_1, \lambda_2, \lambda_3)$ of the S*t*-*t* Model are not significant. Therefore, the N-N Model is the best Spillman-Mitscherlich Response Model.

TABLE 2. The Spillman-Mitscherlich Models

Parameter	N-N		SN-N		S*t*-*t*		SSL-SL		SCN-CN	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
β_1	1.950	0.111	1.982	0.110	1.924	0.095	1.949	0.106	1.960	0.109
β_2	0.032	0.013	0.073	0.025	0.059	0.033	0.074	0.036	0.073	0.033
β_3	2.495	0.415	1.762	0.294	1.771	0.311	1.699	0.310	1.740	0.313
σ^2_ε	0.104	0.010	0.013	0.003	0.010	0.003	0.009	0.003	0.011	0.003
d_1	0.465	0.087	0.220	0.085	0.148	0.072	0.133	0.064	0.170	0.077
d_2	0.008	0.006	0.005	0.005	0.004	0.002	0.004	0.002	0.004	0.002
d_3	0.623	0.199	0.364	0.289	0.283	0.287	0.275	0.240	0.315	0.269
λ_1			0.008	0.101	-0.002	0.071	0.000	0.069	0.001	0.063
λ_2			0.061	0.785	-0.013	0.441	0.003	0.423	0.004	0.434
λ_3			0.006	0.105	-0.001	0.086	0.000	0.067	0.001	0.067
ν (ν_1)					5.413	2.733	3.197	1.769	0.317	0.235
ν_2									0.549	0.205
EAIC	**-147.72**		-128.10		-144.88		-130.58		-127.70	
EBIC	**-147.88**		-128.33		-145.13		-130.83		-127.96	

For the N-N Model, the fixed effects, i.e., the maximum yield coefficient (β_1), the increase in yield (β_2), the ratio of successive increment (β_3) and the random effects (d_1, d_2, d_3) are significant.

The Quadratic Response Models

Comparison of the EAIC and EBIC in table 3 leads to the following rankings: S*t*-*t* < SSL-SL < SCN-CN < SN-N. The results furthermore show that the Skew heavy-tailed distributions outperform the skew normal and normal distribution, and that overall the S*t*-*t* Model is the best Quadratic Response Model.

TABLE 3. The Quadratic Models

Parameter	N-N		SN-N		S*t*-*t*		SSL-SL		SCN-CN	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
γ_1	1.796	0.107	1.794	0.126	1.810	0.088	1.786	0.102	1.788	0.101
γ_2	0.510	0.072	0.506	0.101	0.351	0.063	0.401	0.080	0.393	0.081
γ_3	-0.386	0.072	-0.389	0.101	-0.261	0.059	-0.298	0.075	-0.293	0.074
σ^2_ε	0.033	0.006	0.031	0.021	0.016	0.005	0.012	0.005	0.014	0.006
d_1	0.445	0.085	0.203	0.233	0.130	0.066	0.096	0.050	0.110	0.064
d_2	0.046	0.030	0.021	0.227	0.008	0.004	0.007	0.004	0.007	0.004
d_3	0.030	0.022	0.018	0.232	0.006	0.005	0.005	0.003	0.006	0.003
λ_1			0.014	0.288	0.000	0.064	0.000	0.065	0.000	0.068
λ_2			0.026	0.699	0.000	0.267	-0.001	0.249	0.000	0.257
λ_3			0.021	0.666	0.000	0.300	-0.001	0.277	-0.001	0.286
ν (ν_1)					3.764	1.335	1.682	0.956	0.365	0.165
ν_2									0.275	0.152
EAIC	-45.32		-42.49		**-91.03**		-76.20		-73.51	
EBIC	-45.47		-42.71		**-91.27**		-76.45		-73.78	

For the S*t*-*t* Model, all the fixed effects, i.e., the intercept parameter (γ_1), the linear response coefficient (γ_2), the quadratic response coefficient (γ_3), and the variance component (d_1, d_2, d_3) are significant.

Comparing the Linear Plateau, Spillman-Mitscherlich and Quadratic models

Comparing the Linear Plateau (LP), Spillman-Mitscherlich (SM) and Quadratic (Q) models under five distributional assumptions, we find that the N-N Spillman-Mitscherlich model has the smallest EAIC and EBIC values among the competing models indicating that this is the best fit model for the soybean yield data (table 4).

TABLE 4. Comparison of LP, SM and Q models

Distribution	LP		SM		Q	
	EAIC	**EBIC**	**EAIC**	**EBIC**	**EAIC**	**EBIC**
SN	-85.78	-86.00	-128.10	-128.33	-42.49	-42.71
S*t*	**-98.03**	**-98.28**	-144.88	-145.13	**-91.03**	**-91.27**
SSL	-89.47	-89.71	-130.58	-130.83	-76.20	-76.45
SCN	-88.35	-88.61	-127.70	-127.96	-73.51	-73.78
N	-93.56	-93.71	**-147.72**	**-147.88**	-45.32	-45.47

ISSN : 2579-4361

CONCLUSION

We investigated the performance of linear and nonlinear mixed response models with Skew normal independent (SNI) distributions of random effects. We applied the Bayesian estimation framework to datasets of multi-location trials of potassium fertilization of soybeans. We compared the Linear Plateau, Spillman-Mitscherlich, and Quadratic random parameter models with different distributions of the random parameter component, i.e. the Skew-normal, Skew-*t*, Skew-slash, and Skew-contaminated normal distributions and also the normal distribution with the errors following their symmetric normal independent distributions.

The overall results showed that for Linear Plateau and Quadratic models of fertilizer dosing, the Skew-*t* distributions outperform the normal ones. However, for Spillman-Mitscherlich model the normal distribution is better than the Skew-*t* distribution. The best model for soybean yield prediction is the Normal Spillman-Mitscherlich Response Model.

REFERENCES

1. G. E. Sain and M. A. Jauregui, Agron. J. **85**, 934-937 (1993).
2. J. Shen, R. Li, F. Zhang, Z. Rengel and C. Tang, Nutr. Cycl. Agroecosyst. **65**, 243–251 (2003).
3. D. Wallach, Biometrics **51,** 338–346 (1995).
4. D. Makowski, D. Wallach and J.M. Meynard, Agron. J. **93**, 531–539 (2001).
5. D. Makowski and D. Wallach, Agronomie **22**, 179–89 (2002).
6. D. Makowski and M. Lavielle, J. Agric. Biol. Environ. Stat. **11**, 45-60 (2006).
7. E. Tumusiime, B.W. Brorsen, J. Mosali, J. Johnson, J. Locke and J. T. Biermacher, J. Agr. Appl. Econ. **43**, 541–552 (2011).
8. C. N. Boyer, J. A. Larson, R. K. Roberts, A.T. McClure, D. D. Tyler and V. Zhou, J. Agric. Appl. Econ. **45**, 669-681 (2013).
9. S. C. Park, B. W. Brorsen, A. L. Stoecker and J. A. Hattey, J. Agr. Appl. Econ. **44**, 593–606 (2012).
10. H. R. Day, J. Farm Econ. **47**, 713-741 (1965).
11. O. A. Ramirez, S. Misra and J. Field, Amer. J. Agr. Econ. **85**, 108-120 (2003).
12. F. B. Ouedraogo and B.W. Brorsen, "Bayesian estimation of optimal nitrogen rates with a non-normally distributed stochastic plateau function," *Selected Paper,* (the Southern Agricultural Economics Association (SAEA) Annual Meeting, Dallas, Texas, 2014).
13. V. H. Lachos, P. Ghosh and R. B. Arellano-Valle, Statist. Sinica. **20,** 303-322 (2010).
14. B.W. Brorsen, "Using bayesian estimation and decision theory to determine the optimal level of nitrogen in cotton," *Selected Paper,* (Southern Agricultural Economics Association, Orlando, Florida, 2013).
15. M. Masjkur and H. Folmer, "Bayesian Estimation of Linear Mixed Response Models of Skew-Normally Distributed Stimuli: Evidence from Monte Carlo Simulation," in *Enhancing Innovation in Science for Sustainable Development*, Proc. Basic Sci. 6, edited by A. Effendi *et al.* (Faculty of Mathematics & Sciences, Brawijaya University, Malang, 2016), pp. 372–376.
16. R. B Arellano-Valle, H. Bolfarine and V. H. Lachos, J. Appl. Stat. **34**, 663-682 (2007).
17. V. H. Lachos, D. K. Dey and V. G. Cancho, J. Statist. Plann. Inference **139,** 4098-4110 (2009).
18. D. Bandyopadhyay, V. H. Lachos, C. A. Abanto-Valle and P. Ghosh, Stat. Med. **29**, 2643–2655 (2010).
19. D. Bandyopadhyay, V. H. Lachos, L. M. Castro, and D. K. Dey, Biom. J. **54**, 405–425 (2012).
20. D. Bandyopadhyay, L. M. Castro, V. H. Lachos and H. P. Pinheiro, J. Agric. Biol. Environ. Stat. **20**, 121–139 (2015).
21. Y. Huang, R. Chen, G. Dagne, Y. Zhu and H. Chen, J. Biopharm. Stat. **25**, 373-396 (2015).
22. K. Lange and J. Sinsheimer, J. Comput. Graph. Stat. **2,** 175–198 (1993).
23. G. J. M. Rosa, C. R. Padovani and D. Gianola, Biom. J. **45**, 573–590 (2003).
24. V. H. Lachos, L. M. Castro and D. K. Dey, Comput. Statist. Data Anal. **64**, 237–252 (2013).
25. R. B. Arellano-Valle, H. Bolfarine, and V. H. Lachos, J. Appl. Stat. **34**, 663-682 (2007).
26. T. Baghfalaki, and M. Ganjali, *REVSTAT – Stat. J.* **13**, 169–191 (2015).
27. D. Nursyamsi, M.T. Sutriadi dan U. Kurnia, J. Trop. Soils. **10,** 1-9 (2004).
28. D. Nursyamsi, Jurnal Ilmu Tanah dan Lingkungan **6,** 71-81 (2006).
29. M. Masjkur and W. Hartatik, "Pengembangan rekomendasi pemupukan P dan K pada kedelai berbasis model," Penelitian Institusi, Institut Pertanian Bogor, 2014.
30. A. Gelman, J. Carlin and D. Rubin, *Bayesian Data Analysis* (Chapman & Hall/CRC, New York, 2006).

ISSN : 2579-4361

31. M. Kéry, *Introduction to WinBUGS for Ecologists: A Bayesian approach to regression, ANOVA, mixed models and related analyses* (Academic Press, Elsevier Inc., Amsterdam, The Netherlands, 2010).
32. Y. S. Su and M. Yajima, R2jags: A package for running jags from R (R package version 0.5-7, 2015).