International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

Multivariate Assessment of Nitrogen Fertilization Effects on Yield and Growth Traits of Sesame (*Sesamum indicum* L.) using Principal Component and Regression Analyses

Victoria I. Adamu [1], Michael Kudyo Adamu [1,*], Paul Madina [2], Sani Bala Muhammad [1] and Geoffrey I. Zaaya [1]

[1] *Department of Statistics, Joseph Sarwuan Tarka University, Makurdi – Nigeria.*
[2] *Department of Crop Production, Joseph Sarwuan Tarka University, Makurdi – Nigeria.*

International Journal of Science and Research Archive, 2026, 18(02), 386-393

Publication history: Received on 02 January 2026; revised on 09 February 2026; accepted on 11 February 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.18.2.0257

Abstract

Efficient nitrogen management is critical for improving sesame *(Sesamum indicum L.)* productivity, yet the complex interrelationships among growth, yield, nitrogen-use efficiency, and quality traits remain poorly understood. This study aimed to identify the major nitrogen-driven determinants of grain yield in sesame using a combination of principal component analysis (PCA) and regression modeling. Analysis was performed on growth traits, yield components, nitrogen-use efficiency and grain quality attributes using the field experimental data All quantitative variables were standardized prior to PCA to reduce dimensionality and elucidate dominant patterns of variation. The contributions of key agronomic variables to grain yield were quantified using the multiple linear regression. Four principal components with an eigenvalue higher than 1 was retained by PCA, which could explain 78.77% of total variation. The axis of productivity and N-response was the first principal component (43.50%), which had high scores for leaf area index, grain yield, plant biomass, nitrogen content, agronomic efficiency, and N-use efficiency. Regression analysis showed that LAI at week-6 after sowing, plant biomass, and N-use efficiency responded significantly ($p < 0.01$) to grain yield while SPAD value had the strongest correlation with grain yield; however, there was no synergistic response for any of the efficiency-related variables when they were analyzed simultaneously. Traits related to plant height (vegetative hight) were sources of structural variability, but these not main factors in the determination of yield. Generally, the applied PCA and regression methods showed that sesame yield under different nitrogen levels is influenced not only by N-rate but also by canopy growth, dry matter accumulation and utilization of N. These results emphasize the use of efficiency based-nitrogen management practices to sustain sesame yield.

Keywords: Multiple regression Analysis; Principal Component Analysis; Sesame *(Sesamum indicum L)*; Nitrogen; ANOVA

1. Introduction

In the world current, Sesame (Sesamum indicium L.) is one of the oldest oilseed crops that is cultivated and plays a vital role in nutrition, food security, and income generation, particularly subtropical and tropical regions. The presence of health promoting antioxidants, high oil content and rich protein composition has made sesame a valued crop (Ashri, 2007; Pathak et al., 2014). In many developing countries, including Nigeria, sesame has gained increasing importance as an export commodity and a source of livelihood for small holder farmers. In spite of its economic potential, sesamee productivity remains relatively low compared to its genetic yield potential, this is largely due to suboptimal agronomic management practices (FAO, 2021). Nitrogen is a key macro-nutrient influencing plant growth, canopy development, biomass accumulation, and grain yield. Adequate nitrogen supply enhances leaf area expansion and photosynthetic capacity, thereby increasingg assimilate production and yield formation (Marschner, 2012). However, excessive or poorly managed nitrogen application can result in diminishing yield returns, low nitrogen-use efficiency, environmental

* Corresponding author: Michael Kudyo Adamu

degradation, and increased production costs (Ladha et al., 2016). In sesame production systems, nitrogen management is particularly challenging due to variability in soil fertility, crop response, and interaction among growth, yield, and quality traits (Eifediyi et al., 2023). Previous studies have reported significant effects of nitrogen fertilization on sesame growth and yield components, including plant height, biomass, capsule number, and grain yield (Haruna & Aliyu, 2011; Olowe & Busari, 2015).

However, these studies often relied on univariate statistical approaches that analyze individual traits in isolation. Such approaches fail to capture the complex interrelationships among multiple agronomic traits that collectively determine yield performance. Moreover, yield responses to nitrogen are not solely dependent on application rate but also on how efficiently applied nitrogen is converted into biomass and grain yield, as reflected by nitrogen-use efficiency indices (Dobermann, 2007). Multivariate statistical techniques provide a powerful framework for addressing these complexities. Principal component analysis (PCA) is widely used to reduce data dimensionality and to identify key traits that explain the majority of variation in multivariate datasets (Jolliffe & Cadima, 2016). In crop science research, PCA has been successfully applied to identify dominant yield-related traits, evaluate treatment performance, and simplify trait interrelationships under different management regimes (Yan & Rajcan, 2002; Long et al., 2024). However, while PCA identifiess underlying trait structures, it does not quantify the direct contribution of individual variables to yield. Regression analysis complements PCA by quantifying the magnitude and statistical significance of relationships between yield and its potential determinants. Multiple regression models have been widely used to identify key drivers of crop yield and to support agronomic decision-making (Montgomery et al., 2012).

When combined with PCA, regression analysis allows for both dimensional reduction and causal interpretation, providing a robust analytical framework for understanding yield formation processes under nutrient management scenarios. Despite the relevance of these methods, there is limited empirical evidence integrating PCA and regression analysis to examine nitrogen-driven yield determinants in sesame. Most existing studies focus either on mean yield responses to nitrogen levels or on simple correlations among traits, without explicitly identifying the dominant multivariate structure of agronomic traits or quantifying their relative contributions to yield. This gap limits the development of efficient, data-driven nitrogen management strategies for sustainable sesame production.

1.1. Statement of the Problem

Although nitrogen fertilization is widely recognized as a major determinant of sesame productivity, yield responses are often inconsistent due to complex interactions among growth traits, nitrogen-use efficiency, and quality attributes. Conventional univariate analyses are insufficient to disentangle these interactions and to identify the most influential nitrogen-related determinants of yield. Consequently, if farmers apply nitrogen inefficiently, this may lead to suboptimal yields, increased production costs, and environmental risks. There is therefore a need for an integrated multivariate and regression-based approach to identify the key nitrogen-driven traits that govern sesame yield performance.

1.2. Aim and Objectives of the study

The aim of this study was to identify the major nitrogen-driven determinants of grain yield in sesame *(Sesamum indicum L.)* using principal component analysis and regression modeling, with a view to achieving the following objectives:

- To examine the multivariate structure of growth, yield, nitrogen-use efficiency, and quality traits of sesame using principal component analysis;
- To identify the principal components that account for the largest proportion of variability under different nitrogen regimes;
- To quantify the effects of key nitrogen-related agronomic traits on grain yield using multiple regression analysis; and
- To determine whether sesame yield is driven more by nitrogen application rate or by nitrogen-use efficiency and biomass-related traits.

2. Materials and methods

2.1. Study Area and Experimental Design

The study was based on sesame *(Sesamum indicum L.)* yield, obtained from the Department of Plant Breeding and Seed Science, Joseph Sarwuan Tarka University, Makurdi Benue State from August to November 2020. The experiment was laid out using a factorial arrangement in a randomized complete block design (RCBD) with three replications. Treatments consisted of combinations of sesame varieties, nitrogen sources, and nitrogen application levels, allowing for the evaluation of their individual and combined effects on growth, yield, nitrogen-use efficiency, and quality traits.

2.2. Data Collection

Data were collected on a range of agronomic, efficiency, and quality traits at different growth stages and at harvest. Growth parameters included plant height measured at 4, 6, and 8 weeks after sowing (WAS) and leaf area index at corresponding growth stages. Yield-related traits measured at harvest included number of pods per plant, capsules per plant, 1000-seed weight, grain yield (kg ha^{-1}), and total aboveground biomass (kg ha^{-1}). Nitrogen-use indices such as nitrogen-use efficiency and agronomic efficiency were computed using standard agronomic formulas based on yield response relative to nitrogen application rate. Grain quality attributes, including crude protein content, oil (fat) content, and aflatoxin concentration, were also determined using standard laboratory procedures.

2.3. Data Management

Data Preparation and Standardization Prior to statistical analysis, all quantitative variables were examined for completeness and consistency. Variables measured on different scales were standardized to zero mean and unit variance to remove scale effects and ensure comparability among traits. This step was particularly necessary for multivariate analysis to prevent variables with large numerical ranges from dominating the results.

2.4. Statistical Analysis

2.4.1. Principal Component Analysis

Principal Component Analysis (PCA) was employed as a data reduction and pattern recognition technique to examine the multivariate structure of growth, yield, nitrogen-use efficiency, and quality traits. PCA was performed on the standardized data using the correlation matrix. Principal components were extracted based on the Kaiser criterion (eigenvalues > 1), and the proportion of total variance explained by each component was computed. Component loadings were used to identify variables contributing most strongly to each principal component, and a biplot of the first two principal components was generated to visualize relationships among traits and treatments. PCA was applied strictly as a descriptive multivariate method; therefore, tests associated with factor analysis, such as the Kaiser–Meyer–Olkin measure and Bartlett's test of sphericity, were not required.

2.4.2. Multiple Regression Analysis

We use the Multiple regression analysis to quantify the contribution of key agronomic and nitrogen-related traits to grain yield; multiple linear regression analysis was conducted with grain yield (kg ha^{-1}) as the dependent variable. Independent variables included nitrogen application level, leaf area index at 6 WAS, plant biomass, number of pods per plant, and nitrogen-use efficiency. These variables were selected based on agronomic relevance and their importance in the PCA results. The regression model was specified as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

Where, Y represents grain yield, β_0 is the intercept, β_i are regression coefficients, X_i are explanatory variables, and ε is the random error term.

Model parameters were estimated using the ordinary least squares (OLS) method. Statistical significance of regression coefficients was assessed at the 5% probability level. Treatment Effects on Principal Components To examine the influence of experimental treatments on overall productivity patterns, analysis of variance (ANOVA) was conducted on scores of the first principal component (PC1), which represented the dominant productivity and nitrogen-response axis. Nitrogen level, nitrogen source, and sesame variety were treated as explanatory factors.

3. Results

Table 1 Eigenvalues and Percentage Variance Explained

Principal Component	Eigenvalue	Variance Explained (%)	Cumulative Variance (%)
PC1	7.91	41.21	41.21
PC2	3.34	17.41	58.62
PC3	1.91	9.96	68.58
PC4	1.57	8.19	76.78
PC5	0.93	4.83	81.61
PC6	0.90	4.68	86.29
PC7	0.75	3.89	90.17
PC8	0.59	3.06	93.23
PC9	0.43	2.23	95.46
PC10	0.35	1.83	97.29

3.1. PCA Loadings (Trait Contributions)

Table 2 Loadings for First Three Principal Components

Variable	PC1	PC2	PC3
Nitrogen level	−0.311	0.091	0.229
Leaf area index (4–6 WAS)	−0.340	0.000	0.108
Grain yield (kg/ha)	−0.333	−0.064	−0.001
Plant biomass (kg/ha)	−0.319	0.045	0.204
Pods / Capsules per plant	−0.274	−0.214	−0.210
Agronomic efficiency	−0.288	−0.203	−0.155
Nitrogen use efficiency	−0.270	−0.225	−0.234
Plant height (4–8 WAS)	−0.12 to −0.13	≈ 0.48	−0.18
Crude protein	−0.015	0.017	−0.407
Aflatoxin	0.066	−0.017	0.327

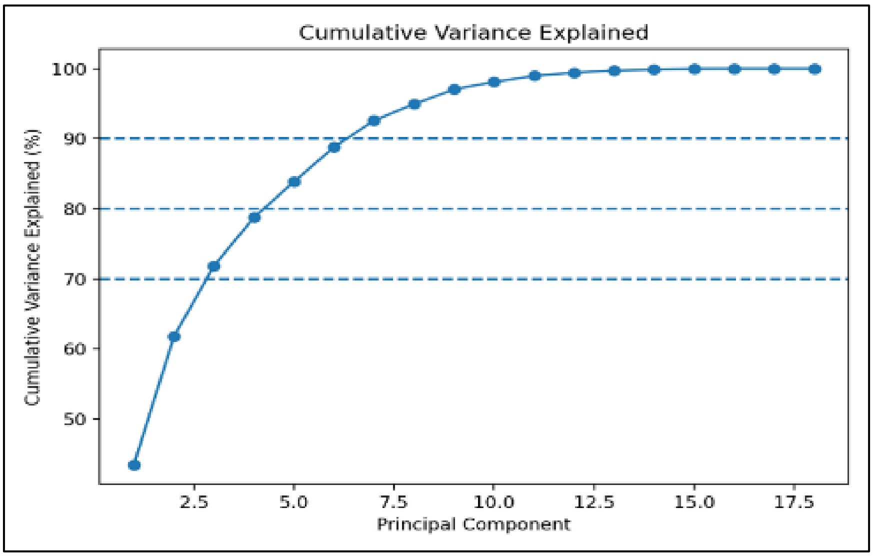

Figure 1 Plot of eigenvalue explaining cumulative variance

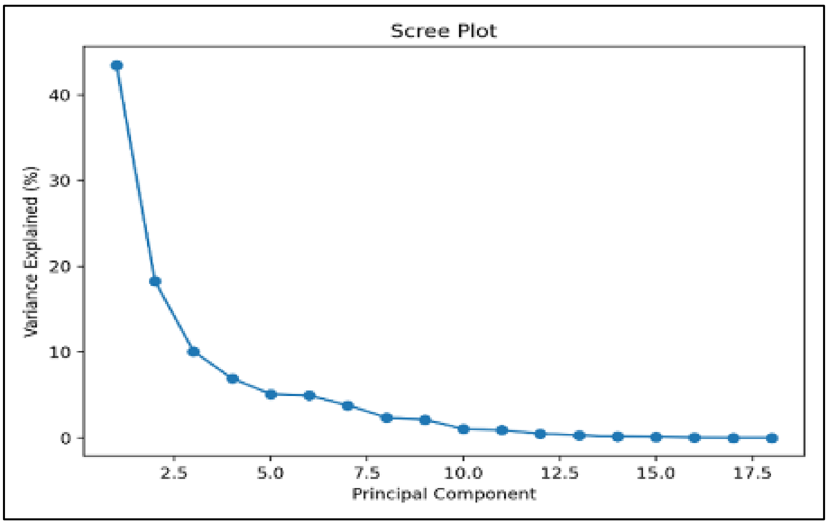

Figure 2 The scree plot

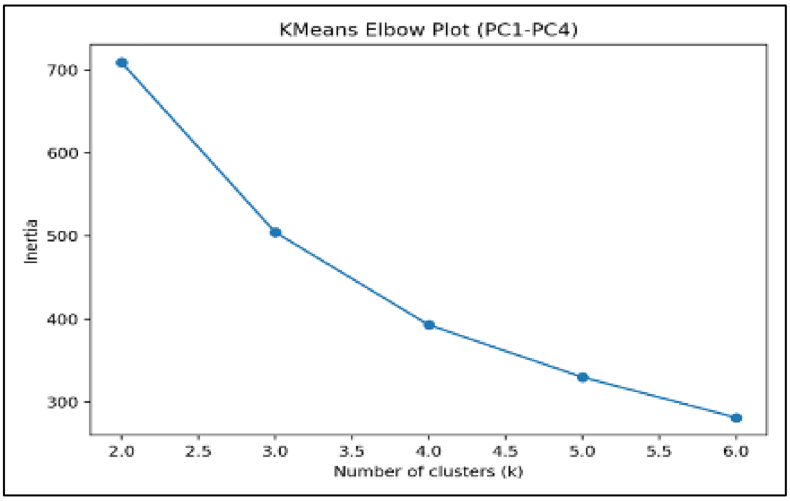

Figure 3 The biplot plot

Table 3 Model Fit Statistics and Diagnostic Tests for the Multiple Regression Model

Category	Statistic	Value	Decision / Interpretation
Model Fit	R^2	0.873	High explanatory power
	Adjusted R^2	0.866	Model remains strong after adjustment
Diagnostic Tests			
Normality of Residuals	Shapiro–Wilk statistic	0.991	—
	Shapiro–Wilk p-value	0.786	Residuals normally distributed
Homoscedasticity	Breusch–Pagan statistic	5.463	—
	Breusch–Pagan p-value	0.362	No heteroscedasticity detected
Multicollinearity (VIF)	Nitrogen level	16.65	Moderate multicollinearity
	Leaf area index (6 WAS)	12.38	Moderate multicollinearity
	Plant biomass (kg ha^{-1})	14.48	Moderate multicollinearity
	Pods per plant	2.96	Low multicollinearity
	Nitrogen-use efficiency	3.18	Low multicollinearity

Table 4 Multiple Linear Regression Results for Determinants of Grain Yield

Dependent variable: Grain yield (kg ha^{-1})

Predictor	Coefficient (β)	Std. Error	t-value	p-value
Intercept	−60.84	62.58	−0.97	0.334
Nitrogen level	−6.87	1.42	−4.83	<0.001
Leaf area index (6 WAS)	2549.96	506.35	5.04	<0.001
Plant biomass (kg ha^{-1})	0.041	0.007	5.55	<0.001
Pods per plant	0.26	0.31	0.85	0.395
Nitrogen-use efficiency	4.70	1.56	3.02	0.003

Significance level: $p < 0.05$

The fitted multiple linear regression model for grain yield is given as:

$$Y = -60.84 - 6.87X_1 + 2549.96X_2 + 0.041X_3 + 0.26X_4 + 4.70X_5 + \varepsilon$$

where: Y= Grain yield (kg ha^{-1}), X_1= Nitrogen application level, X_2= Leaf area index at 6 weeks after sowing, X_3= Plant biomass (kg ha^{-1}), X_4= Number of pods per plant, X_5= Nitrogen-use efficiency and ε is the random error term assumed to be independently and normally distributed with mean zero and constant variance.

Table 5 ANOVA Table for the Multiple Regression Model

Dependent variable: Grain yield (kg ha^{-1})

Source of Variation	df	Sum of Squares (SS)	Mean Square (MS)	F-statistic	p-value
Regression	5	1,429,136.05	285,827.21	123.40	< 0.001
Residual (Error)	90	208,459.08	2,316.21		
Total	95	1,637,595.13			

4. Discussion

From table 1, Principal component analysis revealed that the first four components accounted for 78.77% of the total variability in the dataset, indicating that a limited number of underlying processes govern most of the observed variation in sesame agronomic performance. The dominance of PC1, which explained 43.50% of the variance, underscores the central role of nitrogen-related productivity traits in sesame yield formation. High loadings of leaf area index, plant biomass, grain yield, nitrogen level, agronomic efficiency, and nitrogen-use efficiency on PC1 in table 2, suggested that these traits operate in a coordinated manner to determine the crop productivity. This result is consistent with earlier studies reporting that nitrogen availability enhances canopy development and dry matter accumulation, which in turn support higher yield potential in sesame and other oilseed crops (Haruna & Aliyu, 2011; Marschner, 2012). Madina (2020) reported that inherent genetic make couple with the type and time of nitrogen fertilizer type affect both vegetative and reproductive characters in sesame. From figure 3, the clustering of yield, biomass, and nitrogen-use efficiency vectors in the PCA biplot further confirms that yield performance is strongly associated with how effectively plants convert applied nitrogen into photosynthetically active canopy and biomass, rather than nitrogen rate alone.

In the same vein, the second principal component (PC2), dominated by plant height at different growth stages, captured substantial variability related to vegetative vigor. However, the regression analysis results in table 4, demonstrated that plant height did not significantly contribute once biomass and canopy-related variables were included in the model. This finding indicated that taller plants do not necessarily translate into higher yield, supporting the view that structural growth alone is a poor predictor of productivity. Similar observations have been reported in sesame and other crops, where excessive vegetative growth can occur without proportional increases in yield, particularly under high nitrogen conditions (Olowe & Busari, 2015).

The regression model explained 87.3% of yield variation in the data and the analysis of variance showed that the regression model was highly significant ($F(5, 90) = 123.40$, $p < 0.001$), which indicated that the explanatory variables jointly explained a significant proportion of the variation in grain yield. Leaf area index, plant biomass, and nitrogen-use efficiency were significant positive predictors of grain yield, whereas nitrogen rate showed diminishing returns when efficiency variables were included. These findings align closely with the PCA results, reinforcing the interpretation of PC1 as a productivity and nitrogen-response axis. The strong influence of nitrogen-use efficiency emphasizes that yield improvement in sesame depends not only on nitrogen supply but also on the plant's capacity to utilize applied nitrogen efficiently. The negative coefficient associated with nitrogen level in the regression model, when efficiency-related variables were included, suggests diminishing returns at higher nitrogen rates. This result likely reflects multicollinearity between nitrogen level and nitrogen-use efficiency and indicates that excessive nitrogen application may not translate into proportional yield gains. Similar patterns of diminishing nitrogen response have been documented in oilseed crops, where excessive nitrogen can promote vegetative growth at the expense of reproductive development and efficiency (Dobermann, 2007; Ladha et al., 2016). Diagnostic tests confirmed model adequacy.

5. Conclusion

This study applied a principal component and regression approach to identify nitrogen-driven determinants of grain yield in sesame *(Sesamum indicum L.)*. The results demonstrated that yield variation was primarily governed by canopy development, biomass accumulation, and nitrogen-use efficiency rather than nitrogen application rate alone. The first principal component, representing a productivity and nitrogen-response axis, explained the largest proportion of variability and was strongly associated with leaf area index, plant biomass, and efficiency-related traits. The regression analysis confirmed that leaf area index at 6 weeks after sowing, plant biomass, and nitrogen-use efficiency were significant positive predictors of grain yield, while excessive nitrogen rates showed diminishing returns when efficiency variables were considered. Overall, the findings highlight the importance of efficiency-based nitrogen management strategies for improving sesame productivity.

Recommendations

On the basis of the finding from this study, nitrogen management in sesame production should prioritize practices that enhance nitrogen-use efficiency and early canopy development rather than solely increasing fertilizer application rates. Agronomic intervention that promotes biomass accumulation, such as appropriate nitrogen timing and balanced nutrient supply, should be emphasized to maximize grain yield.

- **Strength and limitations of the study**

A major strength of this study lies in the integration of multivariate and regression analyses, which allowed for both dimensional reduction and causal iinterpretation of yield determinants. However, the study was based on aggregated experimental data from a single season, and environmental variability across seasons and locations was not explicitly modelled.

Future research

Future research should incorporate multi-season and multi-location trials to validate the stability of the identified yield determinants under diverse agro-ecological conditions. Further studies should examine the interaction between nitrogen management and grain quality traits to develop integrated strategies that optimize both productivity and food safety also, mixed-effects modelling could further strengthen the generalizability of the findings.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Ashri, A. (2007). Sesame (*Sesamum indicum* L.). In R. J. Singh (Ed.), *Genetic resources, chromosome engineering, and crop improvement* (pp. 231–289). CRC Press.

[2] Dobermann, A. (2007). Nutrient use efficiency—measurement and management. *Proceedings of the International Fertilizer Industry Association Workshop*, 1–28.

[3] Eifediyi, E., Ogunkoya, F. R., Ogedegbe, F. O., Olugbemi, O., & Olayode, A. (2023). Response of sesame (Sesamum indicum L.) to moisture regime and nitrogen fertilizer in Ilorin, Nigeria. Agrosearch, 22(1), 59–69. https://doi.org/10.4314/agrosh.v22i1.6

[4] FAO. (2021). FAOSTAT statistical database. Food and Agriculture Organization of the United Nations.

[5] Haruna, I. M., & Aliyu, L. (2011). Yield and economic returns of sesame as influenced by nitrogen and phosphorus fertilization. American Journal of Experimental Agriculture, 1(2), 41–49.

[6] Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. Philosophical Transactions of the Royal Society A, 374(2065), 20150202.

[7] Ladha, J. K., Tirol-Padre, A., Reddy, C. K., Cassman, K. G., Verma, S., Powlson, D. S., & Pathak, H. (2016). Global nitrogen budgets in cereals: A 50-year assessment. Scientific Reports, 6, 19355.

[8] Long, Y., Zeng, Y., Liu, X., & Yang, Y. (2024). Multivariate Analysis of Grain Yield and Main Agronomic Traits in Different Maize Hybrids Grown in Mountainous Areas. Agriculture, 14(10), 1703. https://doi.org/10.3390/agriculture14101703

[9] Madina P. (2020). Performance of Ten (10) Varieties of Sesame (Sesamum indicum) Grown in Billiri, Gombe State, Nigeria. Fudma Journal of Sciences (FJS) ISSN online: 2616-1370 ISSN print: 2645 - 2944 Vol. 4 No. 3, September, 2020, pp 453 – 456 DOI: https://doi.org/10.33003/fjs-2020-0403-412

[10] Marschner, H. (2012). Marschner's mineral nutrition of higher plants (3rd ed.). Academic Press.

[11] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to linear regression analysis (5th ed.). Wiley.

[12] Olowe, V. I. O., & Busari, L. D. (2015). Growth and yield response of sesame to nitrogen fertilization in the southern Guinea savanna of Nigeria. Tropical Agriculture, 92(1), 45–54.

[13] Pathak, N., Rai, A., Kumari, R., Thapa, A. and Bhat, K. (2014) Sesame Crop: An Underexploited Oilseed Holds Tremendous Potential for Enhanced Food Value. Agricultural Sciences, 5, 519-529. doi: 10.4236/as.2014.56054.

[14] Yan, W., & Rajcan, I. (2002). Biplot analysis of test sites and trait relations of soybean in Ontario. *Crop Science*, 42(1), 11–20.