JCARM
Journal of Contemporary Academic Research and Methodologies

An Online Peer Reviewed / Refereed Journal
Volume 1 | Issue 6 | July 2026
ISSN: 3139-7247 (Online)
Website: www.journal.jcarm.in

Original Research Article

Correlation and Principal Component Analysis on Growth, Yield and Weed Parameters of Kenaf (*Hibiscus cannabinus* L.) in Sudan Savanna, Nigeria

Yakellu Bukar Kajidu[1], Makanjuola O. D[2], Wulgo M. A[3], Yamba I[3], Usman A[3]

[1] Department of Crop Production /Faculty of Agriculture /University of Maiduguri, Nigeria.
[2] Department of Environmental Sciences, College of Environmental Studies, Kaduna Polytechnic, Kaduna, Nigeria.
[3] Department of Crop Protection, University of Maiduguri, Nigeria.

Correspondence should be addressed to Yakellu Bukar Kajidu: babykajidu@gmail.com

Article No: 028 | **Accepted**: 02 August 2026 | **Published**:10 August 2026

Abstract: This paper aims to show the effects of correlation and principal component analysis on growth, yield and weed parameters of Kenaf (*Hibiscus cannabinus* L.) in the Sudan Savanna, Nigeria. Kenaf has been used as a major source of fibres, paper, and other materials for many years. This fibre crop can be grown on poor soils with fewer inputs. Nigeria's kenaf production figure in the year 2020 is estimated to be about 1,460 tonnes out of the 15,900 tonnes of production in Africa. The global output in 2024 is estimated at approximately 195,800 tonnes. Kenaf fibre and seed yield are mostly affected by growth, yield and weed parameters. The experiment was conducted at the Department of Crop Production Teaching and Research Farm, Faculty of Agriculture, University of Maiduguri (Latitude 11°50'N: and Longitude 13°10'E at an altitude of 319 m above sea level) during the rainy seasons of 2018 to 2019 between June and November. Data were collected on growth, yield components, and weed parameters, which were subjected to correlation analysis, and PCA was analyzed using R software (R Core Team, 2023). An AI assistant (Claude, Anthropic) was used to help write the R analysis scripts. The results of the combined year study showed that Stand count at harvest (0.748) showed the strongest significant linear correlation relationship with yield, Establishment count (0.478), and Weed Control Efficiency (0.402) had a moderate positive correlation with yield. In the year 2018, PC1 (43.66%) was the main axis of variation. PC2 (15.71%) captured additional treatment differences. Together, they explained 59.37% of the total variability. Whereas, the year 2019 PC1 explained 39.69% of the variation, and PC2 explained 11.27%; together, they accounted for 50.97% of the total variation. PC1, therefore, shows crop vigour and productivity, while PC2 was mainly associated with establishment count, stand count at harvest, seed yield and fibre yield, expressing plant population and yield formation. The interpretation shows that correlation outlined relationships existed among the variables,

 https://doi.org/10.5281/zenodo.21860015

while PCA successfully reduced the dimensionality of the variables and identified the principal factors leading to treatment differences across years.

Keywords: Kenaf, Seed Yield, Weed Parameters, Correlation and Principal Component Analysis

Introduction

Correlation is a statistical measure that describes the type of relationship between two or more variables (Sebastian, 2020). It indicates how much the variables change together. while Principal Component Analysis (PCA) is a size reduction technique used to facilitate large, complicated variables. It reduces numerous correlated variables into a smaller set of uncorrelated variables called "principal components", making the maximum possible variance while reducing information loss (Gewers *et al*., 2021). Correlation in agriculture measures how powerfully two variables move together while Principal Component Analysis (PCA) is a multivariate technique that reduces a large set of correlated variables into a smaller number of new, uncorrelated variables called principal components Singh and Singh (2020) also reported that strongly correlated variable clusters usually collapse onto the same principal component, which is often how PCA is used to interpret which traits "go together" Studies have shown that Kenaf fibre yield is significantly affected by growth, yield and weed parameters.

Kenaf (*Hibiscus cannabinus* L.) is a fibre crop belonging to the Malvaceae family. Nigeria is one of the top producers of Kenaf in Africa. Globally, Africa is still in the beginning stages of accepting Kenaf as an economic crop. The official estimate of the total output of the Kenaf fibre crops in Nigeria in 2020 is about 1,460 tonnes out of 15,900 tonnes of production in Africa. The global output in 2024 is estimated at approximately 195,800 tonnes (FAOSTAT, 2025). Compared to Cotton value chain in Nigeria, Kenaf presents to ensure a more environmentally friendly, socially acceptable, profitable and sustainable crop to grow, process, manufacture and market in Nigeria FMARD/NAERLS (2019), they also reported the challenges that have accompanied the Cotton industry in Nigeria in the past decades have sufficiently inhibited economic growth in the fabric and associate sectors and diminished supply chain actors' income. Thus, presenting a diversification opportunity for close substitutes like Kenaf is relevant due to its yield potential and positive regenerative impact on the environment. Kenaf has been produced and used as a major source of fibres, paper, and other materials for many years. This fibre crop does not require fertilizers or other agrochemical inputs. It also requires much less water than Cotton during its production cycle FMARD/NAERLS, 2019).

Different studies were conducted on the relationship and contribution among kenaf traits. Kajidu and Gworgwor (2015) reported a positive correlation between growth and yield parameters, indicating good growth unaffected by weed infestation, and this resulted in higher growth of sorrel. A negative correlation existed between growth yield and weed parameters, which indicated that no significant correlation existed between growth at the early stage and weed parameters. Highly positive correlation of millet vigour was observed with plant count, plant height, leaf area index (LAI), days to 50% booting and blooming. These are factors that contributed to photosynthate production, which supported a positive harvest index. Negative correlation was observed on phyto toxicity ratings and number of branches of cowpea; however did not affect branching of cowpea, which confirms that cowpea is tolerant to herbicides (Joshua & Gworgwor, 2001). More so, Mostafa *et al*. (2002) in their studies on Kenaf heritability and correlation showed that positive correlation was found among growth and yield parameters, but days to 50% flowering was negatively correlated to yield parameters. Also, Abgaje *et al*. (2011) reported a positive correlation among growth and yield parameters in Kenaf in 2006, but no correlation relationship between yield and plant height, seed weight, number of capsules per plant and stem girth was observed in 2007. This paper aims to show the effects of correlation and principal component analysis on growth, yield and weed parameters of Kenaf (*Hibiscus cannabinus* L.) in the Sudan Savanna, Nigeria.

Methodology

The experiment was conducted at the Department of Crop Production Teaching and Research Farm, Faculty of Agriculture, University of Maiduguri (Latitude 11°50'N: and Longitude 13°10'E at an altitude of 319 m above sea level) during the rainy seasons of 2018 to 2019 between June and November. The Physical and chemical characteristics of the experimental sites were determined prior to the experiment. Soil samples were collected randomly from 5 points in the seasons as a composite at a depth of 0 -15 cm and 15 cm -30 cm using a soil auger. The physical and chemical properties of the soil determined for the two years showed that the soil is characterized by a sandy-clay loam texture.

The vegetation characteristics of the experimental site show to have perennial and annual vegetation; trees found there included Neem (*Azadirachta indica* Linn.), Acacia spp., Citrus spp., Cashew (*Anacardium occidentale*), and Banana (*Musa* spp). Jatropha (*Jatropha curcas*). Annual crops grown there are Onion (*Allium cepa* L.), Carrots (*Daucus carota* L.), Amaranthus (*Amaranthus* spp), Sweet potatoes (*Ipomoea batatas* (L.) Poir), Groundnut (*Arachis hypogeae* L.), Cowpea (*Vigna unguiculata* (L.) Walp), Okra (*Abelmoschus esculentus* L. Moench) and many more. The commonest weeds are (*Cyperus* spp.), (*Euphorbia hirta* Linn.), (*Cenchrus biflorus* Roxb.) and many more The climate there is characterized by hot dry season, rainfall from June-October with the peak in July-August. Low humidity in the dry months and high humidity in the wet months. detailed meteorological data from April (7.6 mm and 4.8 mm) to September (141.0 mm and 210.9 mm) with the peak at August (185.7 mm and 220.9 mm) in both years NIMET, (2020). The treatments consisted of three populations of Kenaf plant densities 221,778 stands/ha, 110,889 stands/ha 73,926 stands/ha.Each treatment was replicated three times. Gross plot size measured 3 m x 3 m (9 m^2) and net plot was 2.5 m x 2.5 m (4.5 m^2). Parameters including crop injury, stand count at harvest, number of leaves/plant at harvest, plant height at harvest, stem girth, seed yield t/ha, weed cover score at 9 WAS, WCE=weed control efficiency, 1000seed weight and fiber yield t/ha. were measured. Data collected were subjected to correlation analysis and PCA was analyzed using R software (R Core Team, 2023). An AI assistant (Claude, Anthropic) was used to help write the R analysis scripts; all statistical outputs were reviewed by the authors.

Results and Discussions

This report presents the correlation analysis and Principal Component Analysis (PCA) conducted on the two-year kenaf field trial dataset collected in 2018 and 2019. The analyses were used to identify relationships among measured traits and determine the major sources of variation among observations.

Correlation Analysis

Correlation analysis indicated positive associations among growth and yield traits, while weed-related variables generally showed negative relationships with crop performance. Improved weed control efficiency was associated with better growth and productivity.

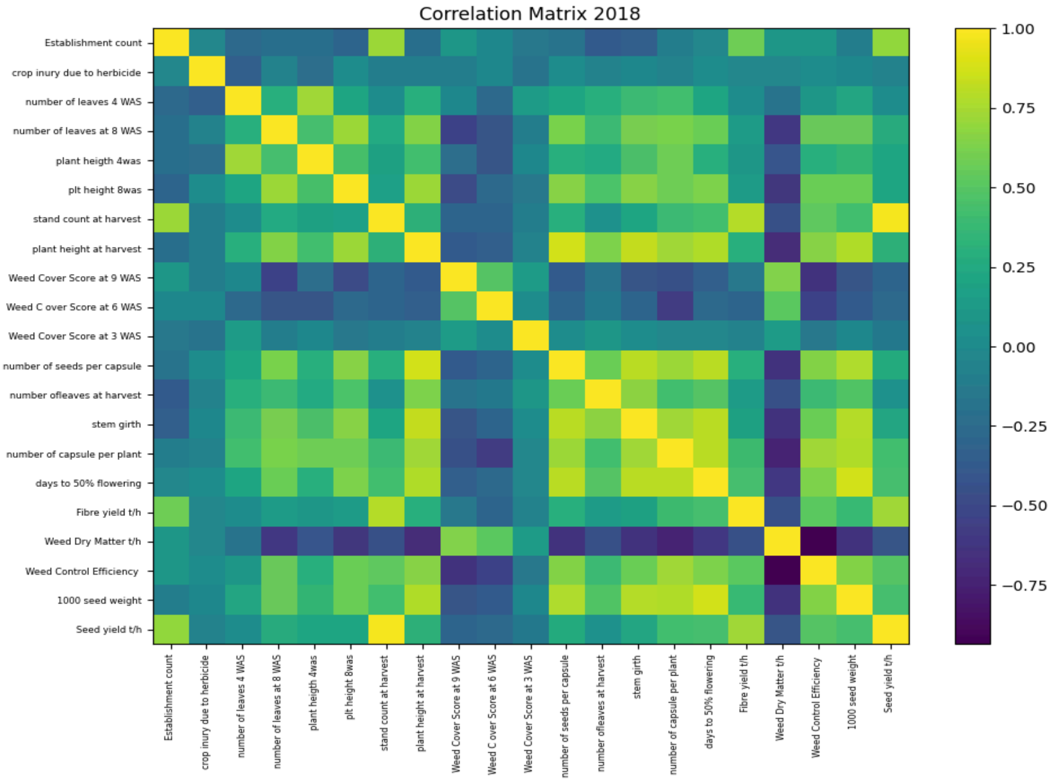

Figure 1. Correlation matrix for 2018 trial.

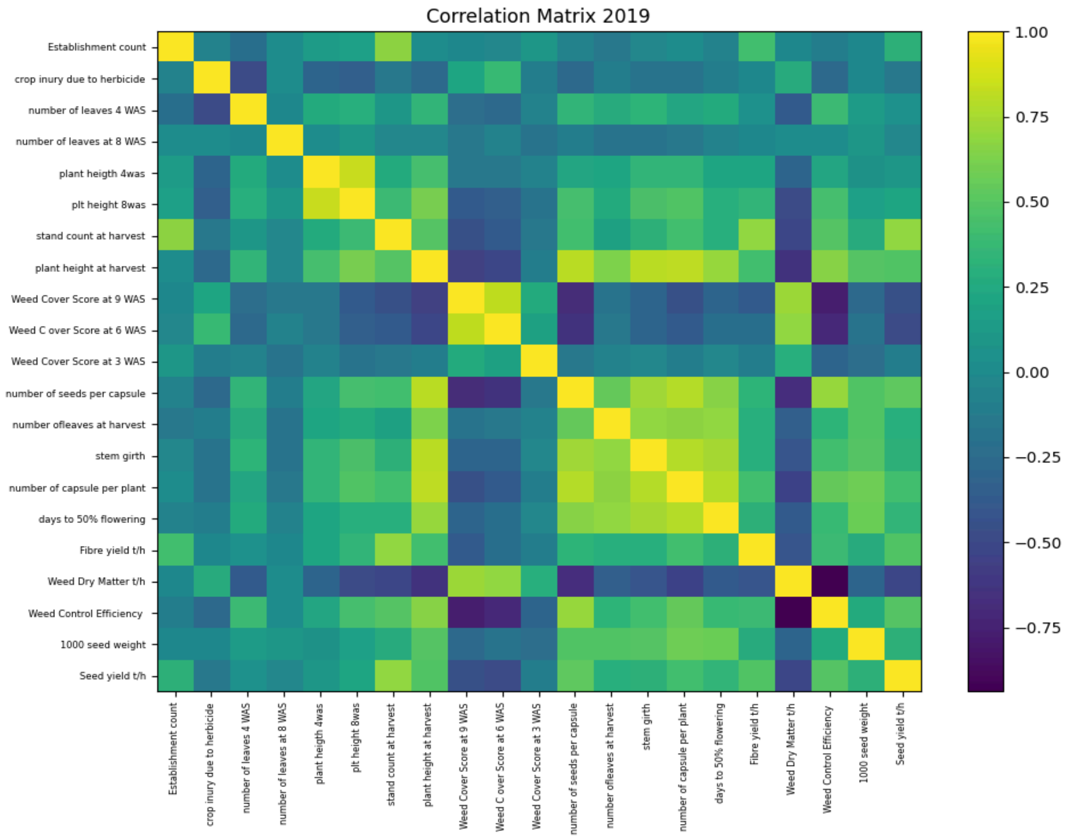

Figure 2. Correlation matrix for 2019 trial.

Table 1. Correlation analysis for Variables most positively associated with seed yield

Correlation with Seed Yield	Value
Stand count at harvest	0.748***
Establishment count	0.478***
Weed Control Efficiency	0.402***
Number of capsules per plant	0.357***
Number of leaves at 8 WAS	0.326***
Number of leaves at harvest	0.324***
Fibre yield	0.320***
Stem girth	0.271***
1000-seed weight	0.225***

*** = ($p<0.0001$)

Stand count at harvest showed the strongest relationship with seed yield, indicating that maintaining plant population until harvest is one of the major determinants of productivity. The moderate positive correlation between establishment count and seed yield suggests that good crop establishment early in the season contributes substantially to final yield. Weed Control Efficiency (WCE) was positively associated with yield, implying that effective weed suppression improved crop performance. Yield components such as capsules per plant, leaf production, and stem girth also contributed positively to seed production.

The positive relationship between fibre yield and seed yield indicates that high biomass-producing plants tended to produce more seed as well similar findings were reported by Kajidu and Gworgwor (2015) who observed positive linear relationship among growth and yield parameters in sorrel.

2018 PCA Biplot

Figure 3 shows PC1 (43.66%) was the dominant axis of variation. PC2 (15.71%) captured additional treatment differences. Together they explained 59.37% of total variability (Table 2). Variables pointing in the same direction are positively correlated. Variables pointing in opposite directions are negatively correlated. Longer vectors indicate variables with stronger influence on treatment separation.

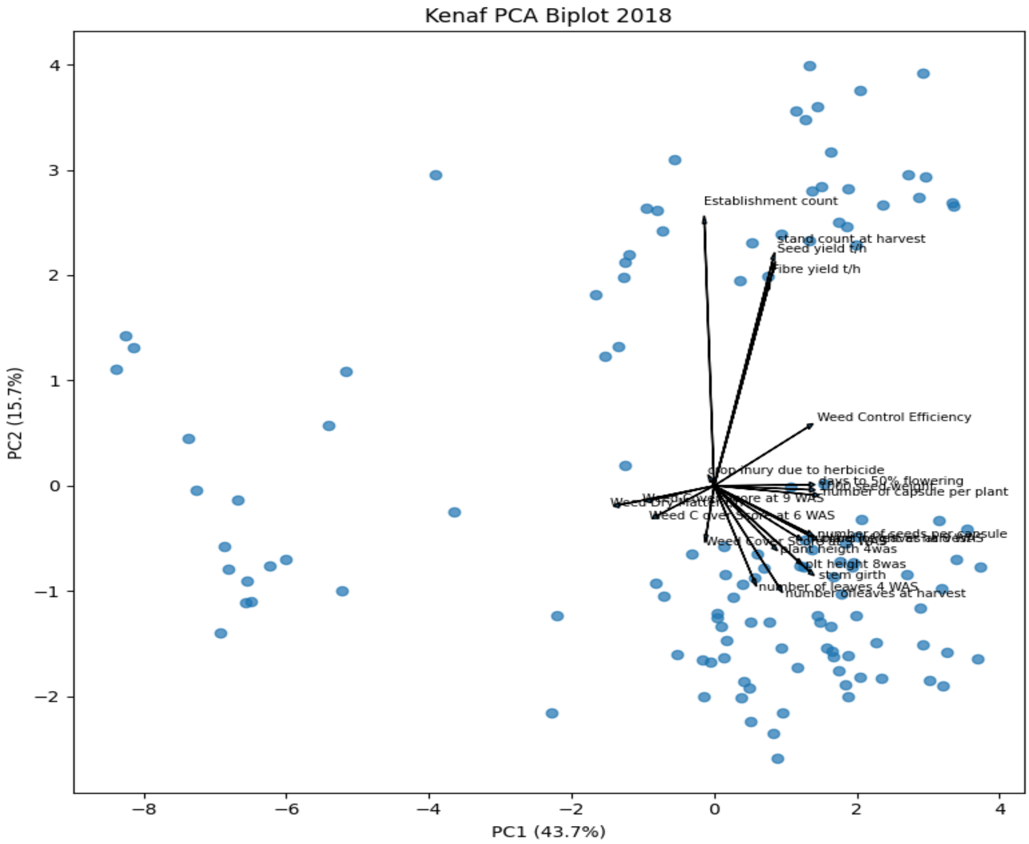

Figure 3. PCA biplot for 2018 trial.

2019 PCA Biplot

PC1 explained 39.69% of the variation. PC2 explained 11.27%. Combined, they accounted for 50.97% of total variation (Table 2). The wider spread of observations suggests greater environmental variation or more variable treatment responses compared with 2018 (Figure 4).

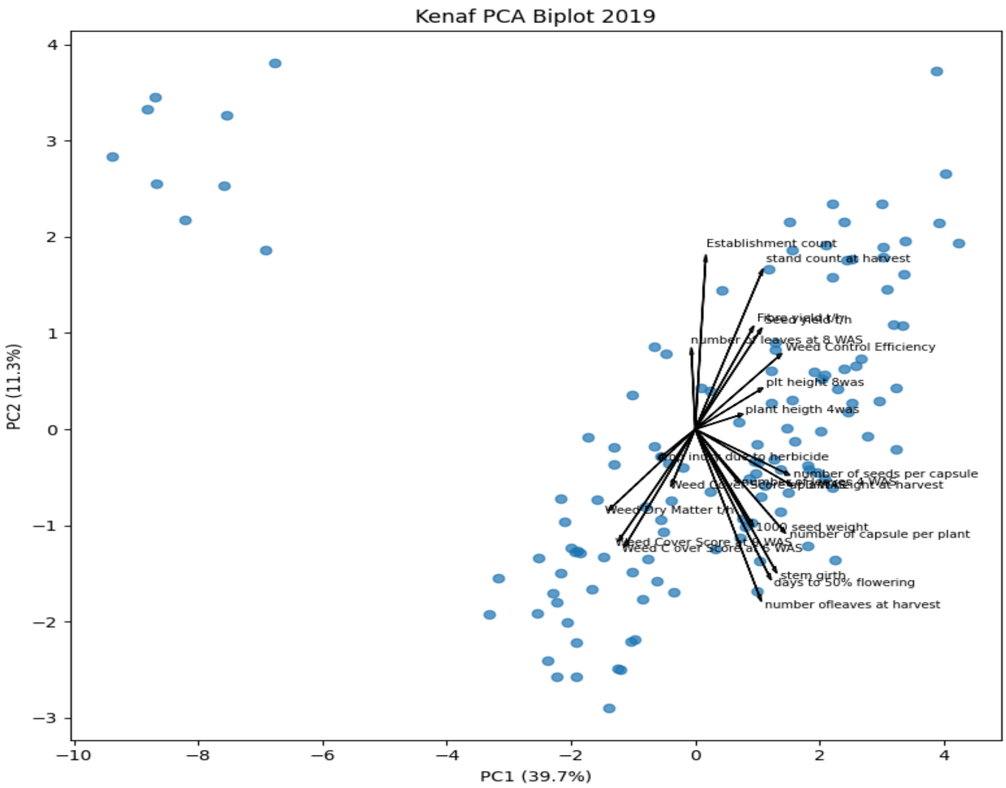

Figure 4. PCA biplot for 2019 trial.

Table 2. Principal Component Analysis

Year	PC1 (%)	PC2 (%)	PC1 + PC2 (%)
2018	43.66	15.71	59.37
2019	39.69	11.27	50.97

Combined results of Correlation and Principal Component Analysis of a Two-Year Kenaf Trial (2018 and 2019)

Variables most positively associated with seed yield were: Stand count at harvest (0.748), Establishment count (0.478), Weed Control Efficiency (0.402), Number of capsules per plant (0.357), Number of leaves at 8 WAS (0.326), Number of leaves at harvest (0.324), Fibre yield (0.320), Stem girth (0.271), and 1000-seed weight (0.225).

Stand count at harvest showed the strongest relationship with seed yield, indicating that maintaining plant population until harvest is a major determinant of productivity. Good establishment and effective weed suppression contributed substantially to final yield. Yield components such as capsules per plant, leaf production and stem girth also contributed positively. Sebastian (2020) outlined that correlation is a relationship between two or more variables.

Principal Component Analysis

The variance explained by the principal components was: PC1 = 37.98%, PC2 = 12.23%, PC3 = 9.36%, PC4 = 7.37%, and PC5 = 6.68% (Table 3). The first two components together explained approximately 50.2% of the total variability in the dataset. PC1 was mainly associated with plant height at harvest, number of capsules per plant, number of seeds per capsule, weed control efficiency, stem girth and days to 50% flowering, while weed dry matter contributed negatively. PC1 therefore represents crop vigour and productivity. PC2 was mainly associated with establishment count, stand count at harvest, seed yield and fibre yield, representing plant population and yield formation. Plant population maintenance was the most important determinant of seed yield. Weed management strongly influenced crop performance. Plant vigour traits were closely associated with yield performance. PCA identified two dominant processes controlling variability: crop vigour and weed suppression (PC1), and stand establishment and yield formation (PC2). Treatments with good establishment, high stand survival, effective weed control and vigorous growth would be expected to produce the highest fibre and seed yields (Gewers *et al.*, 2021) earlier mentioned that PCA reduces numerous correlated variables into a smaller set of uncorrelated variables making the maximum possible variance while reducing information loss.

Table 3. Principal Component Analysis (PCA) of Kenaf Trial (2018 and 2019)

Principal Component	Percentage	Eigenvalue
PC1	37.98%	7.986
PC2	12.23%	2.57
PC3	9.36%	1.971
PC4	7.37%	1.55
PC5	6.68%	1.399
PC6	4.87%	1.022

Sampling adequacy for PCA was confirmed using the Kaiser-Meyer-Olkin (KMO) measure (0.85 for the combined dataset, 0.85 for 2018, 0.84 for 2019; all in the "meritorious" range) and Bartlett's test of sphericity ($\chi^2 = 4189.9$, df = 210, $p < 0.0001$ for the combined dataset), confirming the correlation matrix was suitable for factor extraction. Kaiser's criterion (eigenvalue > 1) supported retaining six components for the combined dataset

Table 4. Variables with the largest contributions to PC1

Trait	Scores
Plant height at harvest	+0.309
Number of capsules per plant	+0.305
Number of seeds per capsule	+0.297
Weed Control Efficiency	+0.293
Stem girth	+0.281
Days to 50% flowering	+0.265
Weed Dry Matter	−0.282

Table 5. Variables with the largest contributions to (PC2)

Trait	Scores
Establishment count	+0.538
Stand count at harvest	+0.466
Seed yield	+0.372
Fibre yield	+0.343

PC1 mainly represents a crop vigour and productivity axis: taller plants with larger stems, more capsules, and more seeds per capsule scored highly on PC1. High weed control efficiency increased PC1 scores. High weed dry matter reduced PC1 scores. This suggests that vigorous growth coupled with effective weed suppression was the major source of variation among treatments (Table 4).

PC2 represents a plant population and yield axis. High PC2 values were associated with: Better crop establishment, Higher plant survival to harvest, Increased seed yield, Increased fibre production (Table 5). This indicates that maintaining an adequate stand density was another major factor influencing productivity in the trial. Overall interpretation of the two-year Kenaf trial Plant population maintenance was the most important determinant of seed yield. Weed management strongly influenced crop performance through its effects on Weed Control Efficiency. Traits related to plant vigour (height, stem girth, capsules per plant) were closely associated with better yield performance. The PCA suggests two dominant processes controlling variation: Crop vigour and weed suppression (PC1). Stand establishment and yield formation (PC2). Treatments that achieved good establishment, high stand survival, effective weed control, and vigorous growth would be expected to produce the highest fibre and seed yields. Singh and Singh (2020) previously reported that strongly correlated variable clusters usually collapse onto the same principal component, which is often how PCA is used to interpret which traits "go together"

Conclusion

In conclusion, the results of the combined year study showed that a significant positive correlation at $p \leq 0.0001$ existed among the variables, where stand count at harvest (0.748) showed the strongest linear correlation relationship with yield, Establishment count (0.478), and Weed Control Efficiency (0.402) had a moderate positive correlation with yield. Principal components variance showed that PC1 = 37.98%, PC2 = 12.23%, PC3 = 9.36%, PC4 = 7.37%, PC5 = 6.68% and PC6 = 4.87%. PC1 was mainly associated with plant height at harvest, number of capsules per plant, number of seeds per capsule, weed control efficiency, stem girth and days to 50% flowering, while weed dry matter contributed negatively. PC2 values were associated with: Better crop establishment, Higher plant survival to harvest, which led to Increased seed yield, Increased fibre production.

The analyses demonstrate that kenaf seed yield was strongly influenced by crop growth attributes and weed management efficiency. PCA successfully reduced the size of the dataset and outlined the principal factors leading to treatment differences across years.

References

1. Adeniyan, O. O (2018). Adding Value to Kenaf: Panacea for Economic. Diversification. *Journal of Environmental Agriculture Internal*. 27(6):1-9; Article no JEAL 27832. ISSN l: 2457-0591. Institute of Agricultural Research and Training (IAR&T), Obafemi Awolowo University (OAU), Ibadan, Nigeria.

2. Agbaje, G. O, Aluko, O. A and Olasoji, J. O (2011). Effect of plant spacing on seed yield and yield components in Kenaf (Hibiscus cannabinus) variety, Ifeken 400. *African Journal of Plant Science*, 5(12): 718-721.

3. FMARD/NAERLS (2019) *"Agricultural Performance Survey of 2018 Wet Season in Nigeria"*, Federal Ministry of Agriculture and Rural Development (FMARD)/National Agricultural Extension and Research Liaison Services (NAERLS)

4. Gewers, Felipe L., Ferreira, Gustavo R., Arruda, Henrique F. De, Silva, Filipi N., Comin, Cesar H., Amancio, Diego R., Costa and Luciano da F. (2021). "Principal Component Analysis: A Natural Approach to Data Exploration". *ACM Comput. Surv*. 54 (4): 70:1–70:34. arXiv:1804.02502. doi:10.1145/3447755.

5. IAR and T. (2022). Institute for Agricultural Research and Training. *Market Study and Value Chain Analysis of Kenaf in Nigeria*. LINKS Catalizing Economic Growth for Northern Nigeria.

6. Joshua S. D. and Gworgwor N. A. (2001). Economic assessment of chemical weed control in cereal-legume intercrop in the savanna zone of Nigeria. *Journals of Borno* 17(18): 247-256.

7. Kajidu Y. B, Gworgwor N. A. (2015). Linear Correlation Analysis among Growth and Yield Parameters of Sorrel as Affected by Spacing and Weeding Regime. *Journal of Advanced Studies in Agricultural, Biological and Environmental Sciences (JABE)* A Peer-Reviewed & Refereed, International Open Access Journal www.jabe.in (2)2, ISSN:2394-2606.

8. Mostafa M. G., Islam M. R., Morsheed-Alam, A. T. M., Mahbub-Ali, S. M. and Mollah M. A. F. (2002). Genetic variability, heritability and correlation studies in Kenaf (*Hibiscus cannabinus* L.), *Online Journal of Biological Sciences* 2 (6) 422-424

9. NIMET (2020). Nigeria Meteorological Agency

10. Sebastian T. (2020). Correlation: A statistical measure of the relationship between two variables. *Corporate Finance Institute*. https://corporatefinanceinstitute.com/resources/data-science/correlation/

11. Singh V. and Singh A. (2020). Analysis of agriculture data using principal component analysis *International Journal of Multidisciplinary Research and Development* Online ISSN: 2349-4182, Print ISSN: 2349-5979; Impact Factor: RJIF 5.72 www.allsubjectjournal.com Volume 7; Issue 1; Page No. 34-37