International Journal of Integrative Research (IJIR)
Vol.4, No.5, 2026: 253-262

Multivariate PCA Analysis for Identifying Adaptation Patterns of Sweet Potato Accessions in the High and Lowlands of Papua

Agustinus Warbaal[1*], Saraswati Prabawardani[2], Eko Agus Martanto[3], Nouke Lenda Mawikere[4], Syaifullah Rahim[5]
Papua University
Corresponding Author: Agustinus Warbaal; agustinus.warbaal@gmail.com

ARTICLEINFO

Keywords: *Biplot, Main Component, Germplasm, Source-Sink, Sweet Potato*

Received : 5 March
Revised : 23 April
Accepted: 23 Mei

©2026 Warbaal, Prabawardani, Martanto, Mawikere, Rahim: This is an open-access article distributed under the termsof the Creative Commons Atribusi 4.0 Internasional.

ABSTRACT

The Land of Papua possesses a rich genetic diversity of local sweet potato (Ipomoea batatas L.), yet its utilization and characterization remain constrained by environmental fluctuations across elevation gradients. This study aimed to analyze the phenotypic diversity structure of local Papuan sweet potato genotypes and map their adaptation profiles under contrasting agro-ecosystems in the highlands (Pegaf) and lowlands (Prafi) using multivariate Principal Component Analysis (PCA). Morpho-agronomic evaluation encompassed upper canopy vegetative growth and underground yield components. The PCA results revealed that the total phenotypic variance of the population was predominantly explained by the First Principal Component (PC1) and the Second Principal Component (PC2). Yield component traits (root diameter, economic root weight per plant, and total yield production) acted as the principal diagnostic markers, registering the highest loading factors on the positive axis of PC1. The biplot visualization detected an orthogonal relationship reflecting a competitive physiological trade-off in photoassimilate allocation (source-sink dynamics) between vegetative vigor and economic yield attributes across the altitudinal zones. Through biplot coordinate mapping, the sweet potato genotypes were objectively clustered based on their agro-morphological similarities, distinguishing the accession cluster of Madu Matawolot, Sarmi-1, and Sarmi-2 as superior in yield components. Integrating PCA ordination with performance consistency parameters identified the UM-2 genotype as elite and consistently productive in both extreme ecosystems, with a superior mean yield (46.8 tons ha^{-1}). Consequently, UM-2 is highly recommended either as a new national superior variety or as a donor parent for future sweet potato breeding programs in Papua

DOI: https://doi.org/10.59890/ijir.v4i5.184
E-ISSN 2988-2184
https://mrymultitechpublisher.my.id/index.php/ijir

INTRODUCTION

Naturally, the region of Tanah Papua harbors a very high genetic diversity of sweet potato plants (Ipomoea batatas L.), which is visually reflected in the variety of morphological characteristics of the accessions cultivated by local farmers. As emphasized by Fu (2015), understanding the genetic diversity aspects of this germplasm is crucial for increasing the efficiency of modern plant breeding programs and maintaining the sustainability of agricultural production in the future. Although common external plant identification methods are used by researchers, the use of conventional morphological markers has a significant drawback because the physical appearance of the plants is susceptible to fluctuations in environmental factors.

The challenge of characterization becomes increasingly complex when sweet potato genotypes are tested across contrasting macroclimatic environments in West Papua, namely the highland agroecosystem (Pegaf District, Arfak Mountains Regency) and the lowland agroecosystem (Aimasi Village, Prafi District, Manokwari Regency). This extreme difference in altitude triggers fluctuations in plant performance, confirming the phenomenon of plasticity in sweet potato productivity traits in response to agroecological differences, as well as the physiological capacity for assimilate allocation (source-sink). Furthermore, the analysis of relationships between agronomic traits urgently needs to be studied in depth because there are patterns indicating a potential trade-off between vegetative growth and yield formation, where excessive vegetative vigor does not always correspond with an increase in the production of economic tuber components.

LITERATURE REVIEW

In examining the relationships between these plant variables, simple linear relationship analysis is generally calculated based on the statistical formula procedures proposed by Gomez & Gomez (1995). Based on this correlation modeling, a correlation coefficient value approaching +1 indicates a very strong positive relationship, whereas a value approaching zero indicates a relatively weak or insignificant linear relationship between variables. Regarding yield components, various previous studies have reported that parameters such as tuber weight, number of tubers, and tuber size have a very strong positive correlation with total yield, thus positioned as key determinants of sweet potato productivity. More specifically, Dwi Julianto et al. (2021) reported that fresh and dry tuber weight have a high closeness of relationship with overall harvest yield ($r > 0.80$). Meanwhile, in a study of local germplasm in the Papua region, Mustamu et al. (2021) successfully identified that tuber length, tuber weight per plant, and the number of tubers per plot are significantly positively associated with economic yield. This finding aligns with the confirmation by Dwi Julianto et al. (2020) in the evaluation of several superior varieties, where the number of tubers and the fresh-dry weight of tubers were very significantly positively correlated with total production. This conclusion is further reinforced by the recent report from Merahabia et al. (2025), which explains that tuber weight per plant and the number of economic tubers are the main contributors controlling variations in harvest yield.

However, if the number of morpho-agronomic characters evaluated is too large, univariate statistical analysis becomes less efficient, so the multivariate analysis method Principal Component Analysis (PCA) is specifically used to evaluate the influence and relative contribution of each character to the total diversity of the genotypes tested. According to the explanation from Irfan et al. (2025), through the PCA approach, the observed characters can be isolated based on the magnitude of their contribution to the formation of principal components, making it easier for plant breeders to identify which characters play the most significant role in distinguishing appearances among genotypes. Regarding this selection parameter, Alam et al. (2024) explain that characters with high loading factor values on the principal components reflect a major influence on genotype variation, while the direction and length of the vectors can be visualized using a biplot diagram.

This graphical visualization theory is reinforced by Geleta et al. (2024), who stated that traits with longer vectors indicate a much greater contribution in explaining genotype variation, making them dominant distinguishing factors in differentiating plant groups. Through the mapping of these principal component coordinates, tuber weight and tuber diameter traits have been shown to be the main drivers of phenotypic variation in the first principal component (PC1), while vegetative traits dominate the second principal component (PC2). This multivariate formulation is comprehensively concluded by Merahabia et al. (2025) that sweet potato genotype variation is highly influenced by agromorphological traits, particularly those directly related to yield components, which hold crucial control in dictating genotype adaptability under different ecosystem conditions.

Although the uniqueness of local Papua germplasm has been recognized, research that integrates multilocation testing in highland and lowland areas with multivariate PCA instruments to simultaneously map adaptive capacity is still very limited. Therefore, this study aims to analyze the influence of agroecosystem heterogeneity on agromorphological component traits, as well as to map adaptation patterns and establish key distinguishing traits of several local Papua sweet potato accessions by applying multivariate PCA analysis. Through a biplot graphical approach, the mutual relationships between genotypes and distinguishing variables can be objectively validated. The results of this study are expected to provide a strong scientific macro basis for modern breeding programs to improve the accuracy of indirect selection and recommend superior sweet potato varieties that are adaptive across extreme ecosystems in the Land of Papua.

METHODOLOGY

Time and Place

This field experimental research was carried out simultaneously in two contrasting types of agroecosystems, namely in the highland and lowland areas of West Papua from June 2024 to February 2025 to test the level of crop adaptability. The details of the test locations are as follows:

Table 1. Research Locations

No	**Characteristics**	**Location 1 (Highlands)**	**Location 2 (Lowland)**
1	Place	Community Field	Community Field
2	Village	Hungku	Umbui
3	District	Anggi	Prafi
4	Regency	Arfak Mountains	Manokwari
5	Province	West Papua	West Papua
6	Plain	Tall	Low
7	Height	1944 mdpl	112 mdpl
8	Coordinates	1°19′29″S 133°53′48″E	0°55′32″S 133°54′04″E

Tools and Materials

The tools used in this study included a tape measure to measure the ridges, a refractometer to measure the sweetness of the sweet potatoes, a digital scale to weigh the sweet potatoes, calipers to measure the diameter of the sweet potatoes, a tape measure/ruler to measure the length of the sweet potatoes, stakes, a hoe, a sickle, scissors, a camera, and writing utensils.

The primary genetic material evaluated in this study consisted of various local sweet potato (Ipomoea batatas L.) accessions collected from various customary areas in Papua. These included 12 local sweet potato genotypes and one control genotype (Table 2), plastic sacks/bags, chicken manure, NPK fertilizer, raffia rope, and label paper.

Table 2. Genotype and Origin

No	**Genotype**	**Origin**
1	Amban Beach-1	Manokwari
2	Koya-1	Kota Jayapura
3	Koya-5	Kota Jayapura
4	Koya – 6	Kota Jayapura
5	Madu Matawolot	Sorong
6	PPUS-1	Sorong
7	PPUS-2	Sorong
8	Sarmi-1	Kota Jayapura
9	Sarmi-2	Kota Jayapura
10	Subsay-1	Manokwari
11	Subsay-2	Manokwari
12	Subsay-3	Manokwari
13	UM-2	Manokwari

Research Design

The field test was arranged using a single-factor Randomized Complete Block Design (RCBD) repeated three times, in line with the principles of multi-location experimental layout recommended by Mattjik & Sumertajaya (2013).

$$Y_{ij} = \mu + \tau i + \beta_j + \varepsilon_{ij}$$

Information:

- Y_{ij} = The observation or response value for the i-th treatment and j-th group.
- μ = The actual general average value (middle mean).
- τi = Effect of treatment i.
- β_j = The influence of the jth group.
- ε_{ij} = The effect of error (random error) on the i-th treatment and j-th group.

The experiment was carried out in two environments, repeated three times for each environment, so that there were 39 experimental plots. Each experimental plot was planted in one plot measuring 6 m □ 1 m.

Morphoagronomic Characterization

Morphological characterization of plants was conducted during the maximum vegetative phase and at harvest (4 months after planting in lowland areas and 6 months after planting in highland areas, according to the tuber harvest age at each altitude). The measurement and standardization procedures for morphoagronomic traits strictly adhered to the international sweet potato descriptor guidelines developed by CIP/IPGRI (1991). Observed variables were grouped into two main categories:

Vegetative Characteristics

Main stem length (cm), stem diameter (mm), leaf shape, and fresh weight of the upper stalk per plant (kg).

Yield Component Characteristics

Economic tuber weight per plant (kg), tuber diameter (mm), tuber length (cm), number of tubers per plant, and estimated total yield (tons/ha).

Multivariate Statistical Analysis (PCA)

The agromorphological data matrix from both locations was first tested for its analysis prerequisites. To address multicollinearity and differences in measurement units between variables (e.g., centimeters, millimeters, and kilograms), which could bias the analysis results, all data were transformed using the Z-score standardization method before running the PCA algorithm, following the multivariate statistical protocol formulated by Johnson & Wichern (2007).

Principal Component Analysis (PCA) was applied to reduce the dimensionality of the large data variables into several new, orthogonal principal components (PCs). The determination of the number of Principal Components retained to explain the data's diversity structure was based on the Kaiser (1960) criterion, which states that only components with eigenvalues greater than one ($\lambda > 1$) are considered significant for analysis.

Next, the strength of each morphoagronomic character's contribution to the formation of the Principal Components was measured based on the vector coefficient or loading factor. Referring to the practical limitations proposed by Hair et al. (2014), agronomic characters with loading values ($|\text{loading}| > 0.40$) were identified as the main character traits that control the phenotypic variability of accessions. The first principal component (PC1) has the greatest ability to explain diversity compared to other components, because it is a linear combination of observed variables that maximizes the total variance in the data, as explained by Kabacoff (2011):

$$PC_1 = a_1 X_1 + a_2 X_2 + \ldots\ldots\ldots + a_3 X_3$$

Where a1 is the characteristic vector of the matrix that corresponds to the largest characteristic root (eigenvalue), while X1 is the original variable group.

Two-dimensional graphical visualization in the form of a biplot diagram (Combination of genotype scores and character vectors on the PC1 and PC2 axes) was explored simultaneously to detect accession-specific adaptation patterns, the closeness of relationships between characters, and the separation of genotype performance across extreme ecosystems. All multivariate statistical calculations and biplot graph construction were operated using the Minitab Statistical Software 22 package.

RESULTS *AND* DISCUSSION

Principal Component Analysis (PCA) and Eigenvalue Decomposition

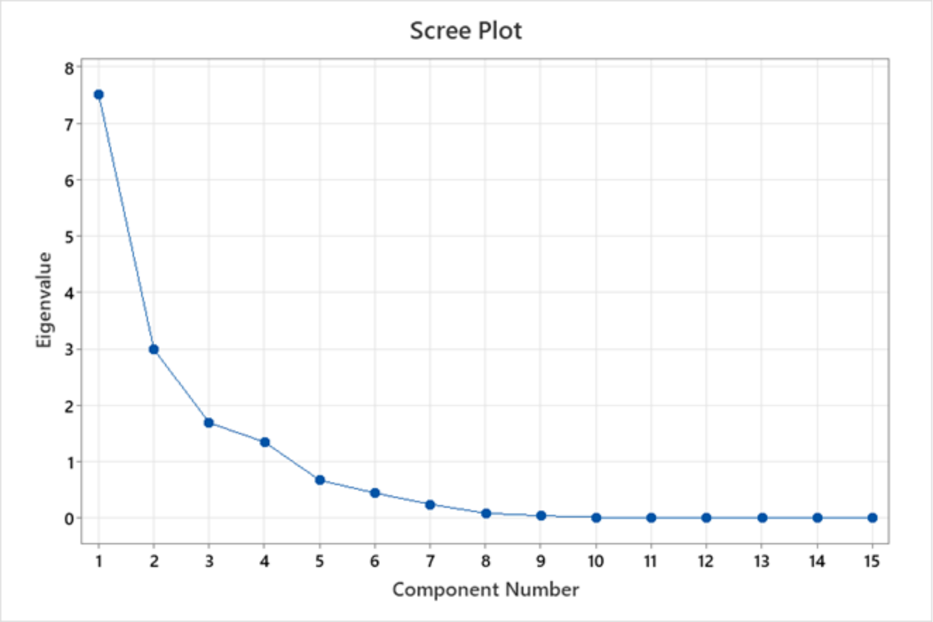

Figure 1. Scree Plot PCS

This distribution pattern demonstrates that most of the total variation in the phenotypic data from the sweet potato germplasm studied can be efficiently explained by just the first few principal components, primarily through the contribution of PC1 and PC2. This high percentage of explained variation reflects the broad genetic foundation of local sweet potato populations in Papua, influenced by differences in plant response to the test environmental conditions. This provides valuable baseline information to support the analysis of genotype adaptation and stability across various agroecosystems in West Papua, both in the highlands and lowlands.

Loading Plot Projection and Inter-Trait Associations

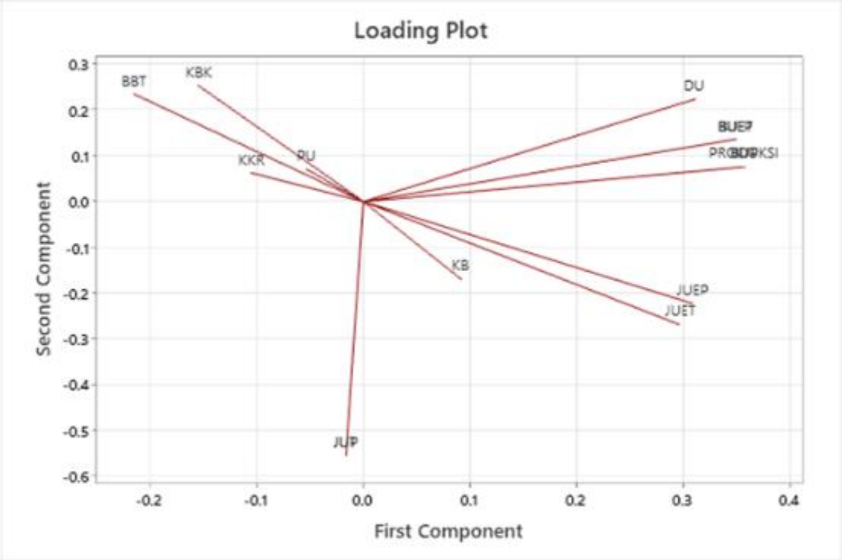

Figure 2. PCA Loading Plot

Analysis of the loading plot provides a visual representation of the contribution of each morphoagronomic trait to the principal components. The vector length reflects the strength of the trait's contribution to total variability, while the vector angle indicates the observed pattern of relationships between traits (Figure 2). Traits with high loading values on PC1, such as tuber diameter (DU), economic tuber weight per plant (BUEP), and total yield, were the main variables contributing most to genotypic variation. This finding confirms that yield component variables play a crucial and essential role in distinguishing performance between sweet potato genotypes.

Vectors that are in the same direction or form an acute angle indicate a strong positive correlation between the variables, where a simple linear correlation coefficient approaching +1 indicates a very close directional relationship. This finding aligns with various previous studies reporting that tuber weight, tuber number, and tuber physical dimensions have a strong positive correlation with total yield, thus determinants of sweet potato productivity. The highly significant positive correlation between yield components and total production was also confirmed by Dwi Julianto et al. (2020) on superior varieties, Mustamu et al. (2021) on local Papuan genotypes, and Merahabia et al. (2025) found that tuber weight per plant and number of economic tubers were the main contributors to yield variation.

Conversely, traits on the negative side of PC1, such as plant stump weight (BBT) and dry matter content (DMC), showed the opposite contribution compared to traits on the positive side of PC1. Meanwhile, the number of tubers per plot (JUP) had a dominant loading on PC2, indicating its important role in explaining variation in the second principal component. The negative or low correlation between vegetative growth traits and yield (ranging from r = –0.30 to –0.60) indicates a potential physiological trade-off between vegetative growth of the aboveground canopy and underground tuber formation, where increased vegetative vigor is not always accompanied by increased tuber production capacity. This association pattern demands careful selection of balanced plant architectural traits by plant breeders. These results align with Demelie & Aragaw (2016) who stated that, although often used as an indicator of vegetative vigor, this trait does not always reflect productivity levels. Therefore, the balance between vegetative (source) and productive (sink) components is a key factor in optimizing photosynthetic capacity and efficient assimilate allocation to tubers, thus supporting plant productivity.

Genotype Score Distribution and Biplot Interpretation

The use of score plots and biplots allows for simultaneous detection of genotype groupings based on similarities in expressed morphoagronomic traits. Sweet potato genotypes that are close together in coordinate space indicate relatively similar characteristics, while those located far apart reflect significant differences in phenotypic traits. Based on the ordination diagram, local genotypes such as Madu Matawolot, Sarmi-1, and Sarmi-2 cluster together on the positive side of PC1. This indicates that these genotype groups are strongly influenced by yield and production component variables that have high contribution loadings on PC1.

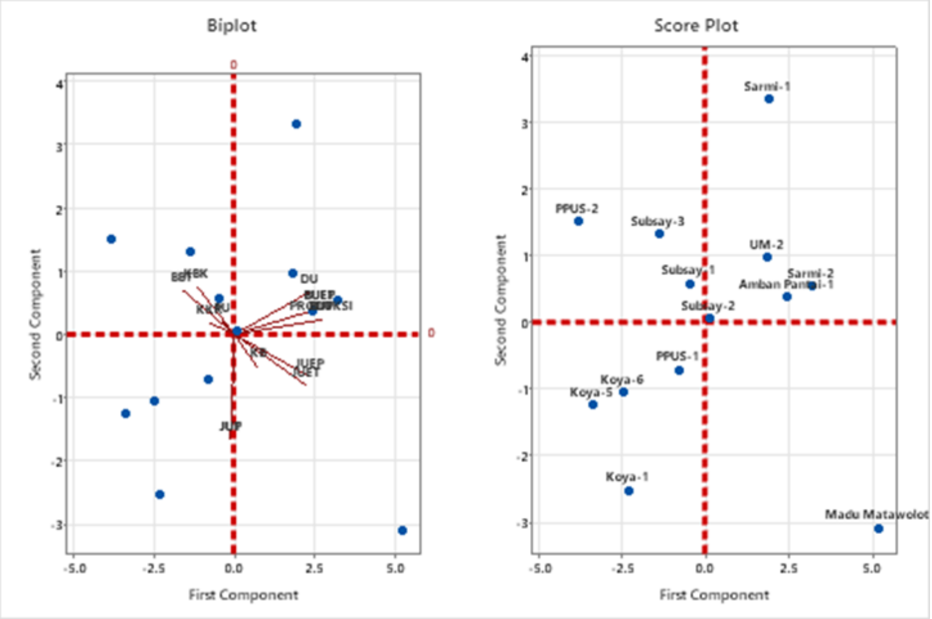

Figure 3. PCA Biplot and Score Plot

A comprehensive integration of multivariate PCA analysis shows that genotypes with consistent performance also exhibit highly favorable agronomic component profiles in the field. The UM-2 genotype is capable of producing superior average yields of approximately 46.8 tons ha^{-1}. This figure is considered very high and competitive within the distribution of other genotype data, which ranges from 31.4–52.5 tons ha^{-1}. The optimal combination of high yield potential and consistent sweet potato production performance across both test environments positions the UM-2 genotype as a highly prospective local genotype. This genotype has great potential to be developed into a new national superior variety with broad adaptability across ecosystems.

CONCLUSIONS AND RECOMMENDATIONS

Principal Component Analysis (PCA) showed that the total phenotypic variance of the sweet potato population was dominated by PC1 and PC2. Yield and physiological sink capacity components, such as tuber diameter, economic tuber weight per plant, and total production yield, were the principal markers because they had the highest loading factor on the positive axis of PC1 in differentiating between genotypes. The biplot detected an inverse (orthogonal) relationship between the vegetative traits of the upper canopy (stove weight) and the economic yield components on PC1. This geometric pattern demonstrated a competitive trade-off in the allocation of assimilable plant energy (source-sink dynamics) across elevations between the Pegaf highlands and the Prafi lowlands.

Through biplot coordinate mapping, sweet potato genotypes were objectively grouped based on their agromorphological traits, with the Madu Matawolot, Sarmi-1, and Sarmi-2 accessions having absolute superiority in yield component variables. Furthermore, the integration of multivariate PCA ordination and the high and consistent performance of sweet potato production successfully identified the UM-2 genotype as a superior genotype and had consistent performance in both extreme agroecosystems with a superior average production (± 46.8 tons ha^{-1}). Therefore, the UM-2 genotype is highly prospective to be recommended as a new national superior variety or donor parent for future sweet potato breeding programs in Papua.

FURTHER STUDY

This research still has limitations, so it is necessary to conduct further research related to the topic of Multivariate PCA Analysis for Identifying Adaptation Patterns of Sweet Potato Accessions in the High and Lowlands of Papua in order to perfect this research and increase insight for readers.

REFERENCES

Alam, M. S., et al. (2024). Multivariate analysis and principal component evaluation of morpho-agronomic traits in tuber crops. Journal of Agronomy and Crop Science, 210(2), 145-158.

CIP/IPGRI. (1991). Descriptors for Sweet Potato. Huaman, Z. (ed). International Potato Center (CIP), Lima, Peru / International Plant Genetic Resources Institute (IPGRI), Rome, Italy.

Demelie, M., & Aragaw, A. (2016). Genetic variability of sweet potato on yield and yield related traits at Werer Agricultural Research Center, Ethiopia. Electronic Journal of Plant Breeding, 7(2), 326–370. https://doi.org/10.5958/0975-928X.2016.00046.1

Dwi Julianto, R., et al. (2020). Korelasi fenotipik dan analisis sidik lintas komponen hasil terhadap total produksi pada beberapa varietas unggul ubijalar (Ipomoea batatas L.). Jurnal Budidaya Pertanian Tropis, 16(1), 45-52.

Dwi Julianto, R., et al. (2021). Evaluasi karakter morfologi dan hubungannya dengan bobot segar umbi pada genotipe ubijalar lokal. Agrikultura, 32(3), 211-220.

Dwi Julianto, T., et al. (2020). Correlation and path coefficient analysis for yield components in improved varieties of sweet potato (Ipomoea batatas L.). Journal of Tropical Crop Science, 7(2), 45–52.

Dwi Julianto, T., et al. (2021). Phenotypic plasticity and character association of local sweet potato germplasm under varying agro-ecological zones. Indonesian Journal of Agricultural Science, 22(1), 12–20.

Fu, Y. B. (2015). Understanding crop genetic diversity under modern plant breeding. Theoretical and Applied Genetics, 128(11), 2131-2142.

Geleta, N., et al. (2024). Characterization of root crops germplasm using vector analysis and multi-environment biplots. Crop Science Reports, 41(3), 202-215.

Gomez, K. A., & Gomez, A. A. (1995). Statistical Procedures for Agricultural Research (2nd ed.). John Wiley & Sons, New York.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). Multivariate Data Analysis (7th ed.). Pearson Education Limited, Harlow, Essex.

Irfan, M., et al. (2025). Application of principal component analysis (PCA) for germplasm characterization and indirect selection criteria in root crops. International Journal of Plant Breeding, 19(1), 12-25.

Irfan, M., et al. (2025). Principal component analysis as a tool for genetic diversity characterization in root and tuber crops. Genetic Resources and Crop Evolution, 72(1), 89–102.

Johnson, R. A., & Wichern, D. W. (2007). Applied Multivariate Statistical Analysis (6th ed.). Pearson Prentice Hall, Upper Saddle River, New Jersey.

Kabacoff, R. I. (2011). R In Action : Data analysis and graphics with R. Manning Publications.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20(1), 141-151.

Mattjik, A. A., & Sumertajaya, I. M. (2013). Perancangan Percobaan dengan Aplikasi SAS dan SPSS. IPB Press, Bogor.

Merahabia, A., et al. (2025). Phenotypic plasticity and source-sink relationship of local sweet potato landraces across different agro-ecological zones. Tropical Plant Biology, 18(2), 89-101.

Merahabia, J., et al. (2025). Sink-source relationship and component marker analysis in indigenous root crops using principal component regressions. Journal of Agronomy and Crop Science, 191(1), 34–47.

Mustamu, Y. A., Prabawardani, S., & Lestari, P. (2021). Morphological characteristics and yield components of local sweet potato (Ipomoea batatas L.) accessions in West Papua, Indonesia. Biodiversitas Journal of Biological Diversity, 22(8), 3125-3135.

Mustamu, Y., et al. (2021). Keragaman karakter morfoagronomi dan komponen hasil beberapa genotipe ubijalar lokal Papua. Jurnal Agroekoteknologi, 13(2), 78-87.