Journal of Scientific Agriculture 2024, 8: 15-30
doi: 10.25081/jsa.2024.v8.9000
https://updatepublishing.com/journal/index.php/jsa

Research Article

ISSN: 2184-0261

Farmers' knowledge and perceptions of Cassava (*Manihot esculenta* Crantz) diseases and management

Dickmi Vaillam Claudette*, Tchouamo Isaac Roger

Faculty of Agronomy and Agricultural Sciences, University of Dschang, P.O. Box 96, Dschang, West Region, Cameroon

Received: February 16, 2024
Revised: June 26, 2024
Accepted: June 27, 2024
Published: July 02, 2024

***Corresponding author:**
Dickmi Vaillam Claudette
E-mail: ngangbaiclaudette@gmail.com

ABSTRACT

This study aimed to examine farmers' understanding and views on Cassava diseases and control methods. To achieve the former, *k*-means clustering and Principal Component Analysis (PCA) were used to identify and visualize response patterns for each group of variables relating to farmers' understanding of Cassava diseases and control methods, and heatmaps were used to detail the characteristics of each pattern. To achieve the latter, bar plots were used to visualize variables related to farmers' views. Out of 22 response patterns relating to causes of Cassava Mosaic Disease (CMD), 11 didn't link a virus to CMD symptoms, while only one pattern associated CMD symptoms with a virus, the whitefly, and infected cuttings, indicating a lack of farmers' knowledge on cassava viral diseases. Also, only 18.88% of farmers know about Cassava diseases and management technologies. This study highlights the urgent need for education and resources for farmers to safeguard their crops and livelihoods.

KEYWORDS: Farmers' knowledge, Farmers' perception, CMD, CMD management

INTRODUCTION

Understanding farmers' knowledge and perceptions of cassava diseases and their management is crucial for several reasons: (1) Cassava is a staple food crop that provides a major source of calories for many rural and urban households in Africa (Burns *et al.*, 2010; Morgan & Choct, 2016; Szyniszewska, 2020; Tize *et al.*, 2021; Alves *et al.*, 2022; Okike *et al.*, 2022; Adebayo, 2023; Mohidin *et al.*, 2023). However, the crop is plagued by various pests and diseases that can significantly reduce yields and income for farming families. Diseases such as cassava mosaic disease (Otim-Nape & Thresh, 1998; Legg, 2008; Ndunguru *et al.*, 2016; Rey & Vanderschuren, 2017; Doungous *et al.*, 2022; Uke *et al.*, 2022; Hareesh *et al.*, 2023; Niño-Jimenez *et al.*, 2024), root rot (Ekundayo & Daniel, 1973; Onyeka *et al.*, 2005; Brito *et al.*, 2017; Nakatumba-Nabende *et al.*, 2020; Pham & Tran, 2021; Sangpueak *et al.*, 2023; Wang *et al.*, 2023a; da Silva *et al.*, 2024; Hohenfeld *et al.*, 2024; Thepbandit *et al.*, 2024), and viral, bacterial (Fanou *et al.*, 2017; McCallum *et al.*, 2017a; Sedano *et al.*, 2017; Yoodee *et al.*, 2018; Mustarichie *et al.*, 2020; Toure *et al.*, 2020; Teixeira *et al.*, 2021; Zárate-Chaves *et al.*, 2021; Pérez *et al.*, 2022; Wydra & Verdier, 2002; Veley *et al.*, 2023), and fungal infections (Bartkowski *et al.*, 1988; Makambila, 1994; Elliot *et al.*, 2002; McCallum *et al.*, 2017b; Chavez *et al.*, 2022; Alleyne *et al.*, 2023; Fathima *et al.*, 2023; Leiva *et al.*, 2023; Owomugisha *et al.*, 2023; Pardo *et al.*, 2023; Thepbandit *et al.*, 2024) can have devastating impacts if not properly managed. (2) Farmers' knowledge and perceptions of these diseases are critical, as they are the primary decision-makers when it comes to managing pests and diseases on their farms (Van den Berg & Jiggins, 2007; Milne *et al.*, 2015; Bottrell & Schoenly, 2018; Miyittah *et al.*, 2022; Taramuel-Taramuel *et al.*, 2023; Bloom *et al.*, 2024; Phung & Dao, 2024). If farmers lack awareness or have an inaccurate understanding of the diseases affecting their crops, they are less likely to implement effective control measures. This can lead to the persistence and spread of diseases, further exacerbating yield losses and food insecurity. (3) By understanding farmers' existing knowledge, attitudes, and practices towards cassava disease management, researchers and extension agents can develop targeted interventions to address knowledge gaps and promote more effective disease control strategies. This may include improving access to disease-resistant planting materials, providing training on integrated pest and disease management, and strengthening linkages between farmers and agricultural extension services (Van den Berg & Jiggins, 2007; Milne *et al.*, 2015; Brévault & Clouvel, 2019; Laizer *et al.*, 2019; Taramuel-Taramuel *et al.*, 2023; Phung & Dao, 2024). Therefore, this study aimed to examine farmers' understanding and views on Cassava diseases and control methods.

MATERIALS AND METHODS

Dataset

Recall that the two-fold objectives of this work were to study (1) farmers' knowledge and (2) farmers' perceptions of Cassava diseases and management technologies. The dataset had 508 entries corresponding to 508 respondents from four regions of Cameroon – Adamawa, Center, East and South. It had a total of 52 variables (Tables 1 & 2), partitioned as follows – 27 binary (0=No/1=Yes) categorical variables used to assess farmers' knowledge of Cassava diseases and management technologies (Table 1), and 25 variables used to assess farmers' perceptions of Cassava diseases and management technologies (Table 2). The 27 variables used to study farmers' knowledge were subdivided into 5 groups – group one aimed to assess farmers' knowledge about various pests and diseases which they observed in their Cassava farms. Group two aimed to assess farmers' knowledge about the causes of Cassava Mosaic Disease (CMD)-related symptoms. Group three aimed to assess farmers' knowledge about the impact of the appearance of CMD-related symptoms on Cassava plants. Group four aimed to assess farmers' reactions to CMD-related symptoms. Finally, group five aimed to assess farmers' knowledge about CMD prevention and management.

Data Analysis

Farmers' knowledge of cassava diseases and management technologies

For each section of the group of variables used to study farmers' knowledge of Cassava diseases and management technologies (Table 1), response patterns were computed and visualized using both *k*-means clustering (to determine the number of unique response patterns) and Principal Component Analysis (PCA) (to reduce the dimensionality of the resulting response patterns and visualize them using the first 2 principal components). Also, the variable-wise characteristics for each response pattern were visualized as an annotated heatmap for each of the 5 sections.

k-means clustering

k-means clustering is a widely used unsupervised machine learning algorithm that partitions a set of data points into K clusters based on their similarities and dissimilarities (Capó *et al.*, 2018; Bateni *et al.*, 2024). The algorithm is particularly useful for identifying patterns and structures within large datasets where the number of clusters is unknown or difficult to determine (Kim *et al.*, 2024). It is often used in various fields such as data mining, image processing, and bioinformatics to group similar data points together and to identify outliers or anomalies (Garst & Reinders, 2024).

The k-means clustering algorithm begins by randomly selecting K initial centroids, which are the mean values of the data points in each cluster (Vardakas & Likas, 2023; Yfantis *et al.*, 2023). These centroids serve as the starting points for the clustering process (Mussabayev *et al.*, 2023). The algorithm then iteratively assigns each data point to the cluster with the closest centroid, and updates the centroids to be the mean of all data points in each cluster (Zhu *et al.*, 2021). This process is repeated until the centroids no longer change significantly, indicating that the clusters have converged (Miao *et al.*, 2023).

One of the key advantages of k-means clustering is its simplicity and efficiency. The algorithm is relatively easy to implement and can handle large datasets with high-dimensional features (Ergun *et al.*, 2022; Mohammadi *et al.*, 2022; Poggiali *et al.*, 2024). Additionally, k-means clustering is robust to noise and outliers, as it is based on the mean of the data points in each cluster, which makes it less sensitive to extreme values (Chen & Witten, 2022; Clum *et al.*, 2022).

Table 1: Description and codes of variables used to study farmers' knowledge of cassava diseases and management technologies

S. No.	Code	Variable Description
		Identification of pests and diseases **What diseases and pests do you frequently see in your cassava fields?**
1	V339	Fungal
2	V340	Bacterial
3	V341	Viral (CMD, CBSD)
4	V342	Mite damage and cassava pests
5	V343	Whiteflies
		Causes of symptoms **What is the cause of the symptoms seen in the second photo?**
1	V352	A virus
2	V353	The whitefly
3	V354	The use of infected cuttings
4	V355	Lack of rain
5	V356	Soil moisture
6	V357	Mineral deficiency
		Impact of CMD on Cassava plants **What is the impact of the appearance of these symptoms (CMD) on cassava plants/yield?**
1	V370	Poor plant growth
2	V371	Decrease in yield
3	V372	Lack of healthy plant material
4	V373	Other (None of the above)
		Reactions to CMD symptoms **How do you react when your cassava plants show the symptoms shown in photo 2?**
1	V377	Removal of infected plants
2	V378	Destruction of infected plants
3	V379	Replacement of infected plants by healthy cuttings
4	V380	Analysis of the plants concerned with the Nuru application
5	V381	Consultation with agricultural agents
6	V382	Use of inputs
7	V383	I do nothing
		CMD Prevention **What do you think can be done to prevent or combat the onset of these symptoms/disease?**
1	V387	Use of healthy plant material
2	V388	Regular monitoring of fields (removal, destruction, and replacement of infected plants)
3	V389	Regular cleaning of the fields
4	V390	Respect of the planting density
5	V391	Other (None of the above)

Table 2: Description and codes of variables used to study farmers' perceptions of cassava diseases and management technologies

S. No.	Code	Variable Description
		Farmers' Perceptions of Cassava Viral Diseases and Management
1	V412	Cassava viral diseases are caused by poor hygiene on the field
2	V413	Viral symptoms observed on cassava leaves result from the application of herbicides.
3	V414	Older plants are more attacked by cassava viral diseases
4	V415	Late planting can lead to cassava viral diseases
5	V416	Drought and high temperatures can lead to cassava viral diseases.
6	V417	Planting in muddy or waterlogged soils causes infections.
7	V418	Poor aeration promotes cassava viral diseases
8	V419	A late harvest can lead to cassava viral diseases.
9	V420	Cassava viral diseases are caused by the use of poor-quality planting
10	V421	Cassava viral diseases can be managed by breaking the affected part
11	V422	The management practices can easily be integrated into the traditional farming system
12	V423	The management practices taught by the agricultural officers are complex to understand
13	V424	The use of integrated approaches to viral diseases control in cassava is more expensive than using chemicals.
14	V425	Chemicals are effective in controlling cassava viral diseases
15	V426	No cassava variety is resistant to viral diseases
16	V427	There is no control for cassava viral diseases
17	V428	The plant infected by cassava viral diseases always recovers at the beginning of the rains
18	V429	The management practices of cassava viral diseases are not culturally accepted in my community
19	V430	The use of cassava diseases management technologies helps to reduce the incidence of cassava viral diseases.
20	V431	The use of cassava disease management technologies increases productivity
21	V432	The management technologies are accessible at all times
22	V433	Cassava viral diseases prevent rooting
23	V434	Whenever cassava plants are attacked by viral diseases, it results in poor quality tubers
24	V435	Viral diseases of cassava can lead to 100% yield loss if left untreated.
25	V436	Cassava viral diseases result in loss of planting material

However, k-means clustering also has some limitations. For instance, the algorithm is sensitive to the initial placement of the centroids, which can affect the final clustering results. Additionally, k-means clustering assumes that the clusters are spherical and well-separated, which may not always be the case in real-world datasets (Dorabiala *et al.*, 2024). Furthermore, the algorithm can be computationally expensive for large datasets, especially when the number of clusters is high. To address these limitations, various extensions and modifications of the k-means clustering algorithm have been proposed. For example, k-means++ is a variant of the algorithm that uses a more sophisticated method to select the initial centroids, which can improve the clustering results (Zhuang *et al.*, 2024). Another extension is the k-medoids algorithm, which uses medoids (objects that are representative of their cluster) instead of centroids, which can be more robust to noise and outliers (Schubert, 2023).

Principal component analysis (PCA)

Principal Component Analysis (PCA) is a widely used dimensionality reduction technique that simplifies a large dataset into a smaller set while preserving significant patterns and trends (Maćkiewicz & Ratajczak, 1993; Demšar *et al.*, 2013; Jolliffe & Cadima, 2016). It is a linear method that transforms the data onto a new coordinate system where the directions (principal components) capturing the largest variation in the data can be easily identified. These directions are orthogonal and constitute an orthonormal basis in which different individual dimensions of the data are linearly uncorrelated (Jolliffe, 2011; Elhaik, 2022).

The process of PCA involves several steps. First, the data is standardized to ensure that each variable has a mean of 0 and a standard deviation of 1 (Chen *et al.*, 2024; Shlens, 2014; Wang *et al.*, 2024). This is done to prevent variables with large ranges from dominating the analysis. Next, the covariance matrix is computed to identify correlations between the variables. The eigenvectors and eigenvalues of this covariance matrix are then calculated to identify the principal components (Soto-Quiros & Torokhti, 2021; Dorabiala *et al.*, 2023; Lee *et al.*, 2023; Leroux *et al.*, 2023; Hope *et al.*, 2024). The eigenvectors are the directions of the new axes, and the eigenvalues represent the amount of variance explained by each component. The principal components are ordered in decreasing order of importance, with the first component capturing the most variation in the data (Gniazdowski, 2017; Tang & Allen, 2021; Hurwitz & Hahn, 2023). The second component captures the maximum variance that is orthogonal to the first component, and so on (Gewers *et al.*, 2018; Suzen *et al.*, 2020; Marzban *et al.*, 2024). This process continues until all the variance in the data is explained. The total variance captured by all the principal components is equal to the total variance in the original dataset (van Elst, 2021).

PCA is particularly useful when many of the variables are highly correlated with each other and it is desirable to reduce their number to an independent set. It can be used for data visualization, feature selection, and data compression. In data visualization, PCA can be used to plot high-dimensional data in two or three dimensions, making it easier to interpret. In feature selection, PCA can be used to identify the most important variables in a dataset (Guo *et al.*, 2002; Song *et al.*, 2010; Mishra *et al.*, 2011; Zhang, 2019; Rahmat *et al.*, 2024). In data compression, PCA can be used to reduce the size of a dataset without losing important information. One of the key advantages of PCA is its ability to deal with multicollinearity, which is a common problem in regression analysis where two or more independent variables are highly correlated. PCA can help identify the underlying structure in the data and create new, uncorrelated variables that can be used in the regression model. Additionally, PCA can be used to reduce the noise in data by removing the principal components with low variance, which are assumed to represent noise (Statheropoulos *et al.*, 1999; Razifar *et al.*, 2009; Bailey, 2012; Du *et al.*, 2015; Ling-Qun *et al.*, 2015; Li, 2018).

However, PCA also has some limitations. The principal components created by PCA are linear combinations of the original variables, and it can be difficult to interpret them in terms of the original variables (Johnstone & Lu, 2009; Lee, 2011; Wang *et al.*, 2023b). This can make it challenging to explain the results of PCA to others. Additionally, PCA is sensitive to the scale of the data, and if the data is not properly scaled, then PCA may not work well (Görtler *et al.*, 2020).

Farmers' perceptions of cassava diseases and management technologies

This was analyzed descriptively by generating 25 annotated bar plots corresponding to each of the 25 variables used to study farmers' perceptions of Cassava diseases and management technologies (Table 2).

RESULTS AND DISCUSSION

Farmers' Knowledge of Cassava Diseases and Management Technologies

Farmers' response patterns relating to survey questions on identification of pests and diseases in their Cassava farms

Within the 508 individual responses, *k*-means analysis identified 28 distinct response patterns. However, with PCA-aided visualization, some of these response patterns were observed to be closely related (Figure 1a) – (5, 8, 21); (6, 7, 22); (16, 24, 25); (15, 18); (19, 20); (1, 27); (3, 17), *inter alia* – and others appeared to be actually unique – 0, 2, 4, 9, 10, 11, 23, *inter alia*.

Response pattern 0 (Figure 1b) was characterized by farmers who identified bacterial and viral diseases as those with the highest incidences on their farms. Response pattern 1 was characterized by farmers who identified viral diseases, mite damage, cassava pests and white flies as those with the highest incidences on their farms. Response pattern 2 was characterized by farmers who identified no pests or diseases on their farms as serious enough to take note of. Response pattern 3 was characterized by farmers who identified only viral diseases as those with the highest incidences on their farms. Response pattern 4 was characterized by farmers who identified all 4 categories of pests and diseases as equally having the highest incidences on their farms. Response pattern 5 was characterized by farmers who identified only mite damage and cassava pests as those with the highest incidences on their farms.

Response pattern 6 was characterized by farmers who identified viral diseases and whiteflies as those with the highest incidences on their farms. Response pattern 7 was characterized by farmers who identified viral diseases, mite damage and cassava pests as those with the highest incidences on their farms. Response pattern 8 was characterized by farmers who identified only whiteflies as those with the highest incidences on their farms. Response pattern 9 was characterized by farmers who identified fungal diseases and viral diseases as those with the highest incidences on their farms. Response pattern 10 was characterized by farmers who identified only fungal diseases as those with the highest incidences on their farms.

Response pattern 11 was characterized by farmers who identified mite damage, cassava pests and whiteflies as those with the highest incidences on their farms. Response pattern 12 was characterized by farmers who identified all 4 categories of pests and diseases except fungal diseases, as having the highest incidences on their farms. Response pattern 13 was characterized by farmers who identified fungal diseases, bacterial diseases and viral diseases as those with the highest incidences on their farms. Response pattern 14 was characterized by farmers who identified all 4 categories of pests and diseases except bacterial diseases, as having the highest incidences on their farms. Response pattern 15 was characterized by farmers who identified all 4 categories of pests and diseases except whiteflies, as having the highest incidences on their farms.

Response pattern 16 was characterized by farmers who identified fungal diseases, viral diseases and whiteflies as those with the highest incidences on their farms. Response pattern 17 was characterized by farmers who identified only bacterial diseases as those with the highest incidences on their farms. Response pattern 18 was characterized by farmers who identified all 4 categories of pests and diseases except mite damage and cassava pests, as having the highest incidences on their farms. Response pattern 19 was characterized by farmers who identified bacterial diseases, viral diseases and whiteflies as those with the highest incidences on their farms. Response pattern 20 was characterized by farmers who identified bacterial diseases, viral diseases and mite damage and cassava pests as those with the highest incidences on their farms.

Response pattern 21 was characterized by farmers who identified fungal diseases and mite damage and cassava pests as those with the highest incidences on their farms. Response pattern 22 was characterized by farmers who identified bacterial diseases and whiteflies as those with the highest incidences on their farms. Response pattern 23 was characterized by farmers who identified fungal diseases, mite damage and cassava pests, and whiteflies as those with the highest incidences on their farms. Response pattern 24 was characterized by farmers who identified fungal diseases, bacterial diseases and whiteflies as those with the highest incidences on their farms. Response pattern 25 was characterized by farmers who identified fungal diseases, viral diseases and mite damage and cassava pests, as those with the highest incidences on their farms. Response pattern 26 was characterized by farmers who identified all 4 categories of pests and diseases except viral diseases, as having the highest incidences on their farms. Response pattern 27 was characterized by farmers who identified bacterial diseases, mite damage and cassava pests, and whiteflies as those with the highest incidences on their farms.

Farmers' response patterns relating to survey questions on causes of CMD-related symptoms

Within the 508 individual responses, *k*-means analysis identified 22 distinct response patterns (Figure 2a). Response pattern 0 (Figure 2b) was characterized by farmers who responded positive for a virus, negative for the whitefly, positive for the use of infected cuttings, positive for lack of rain, negative for

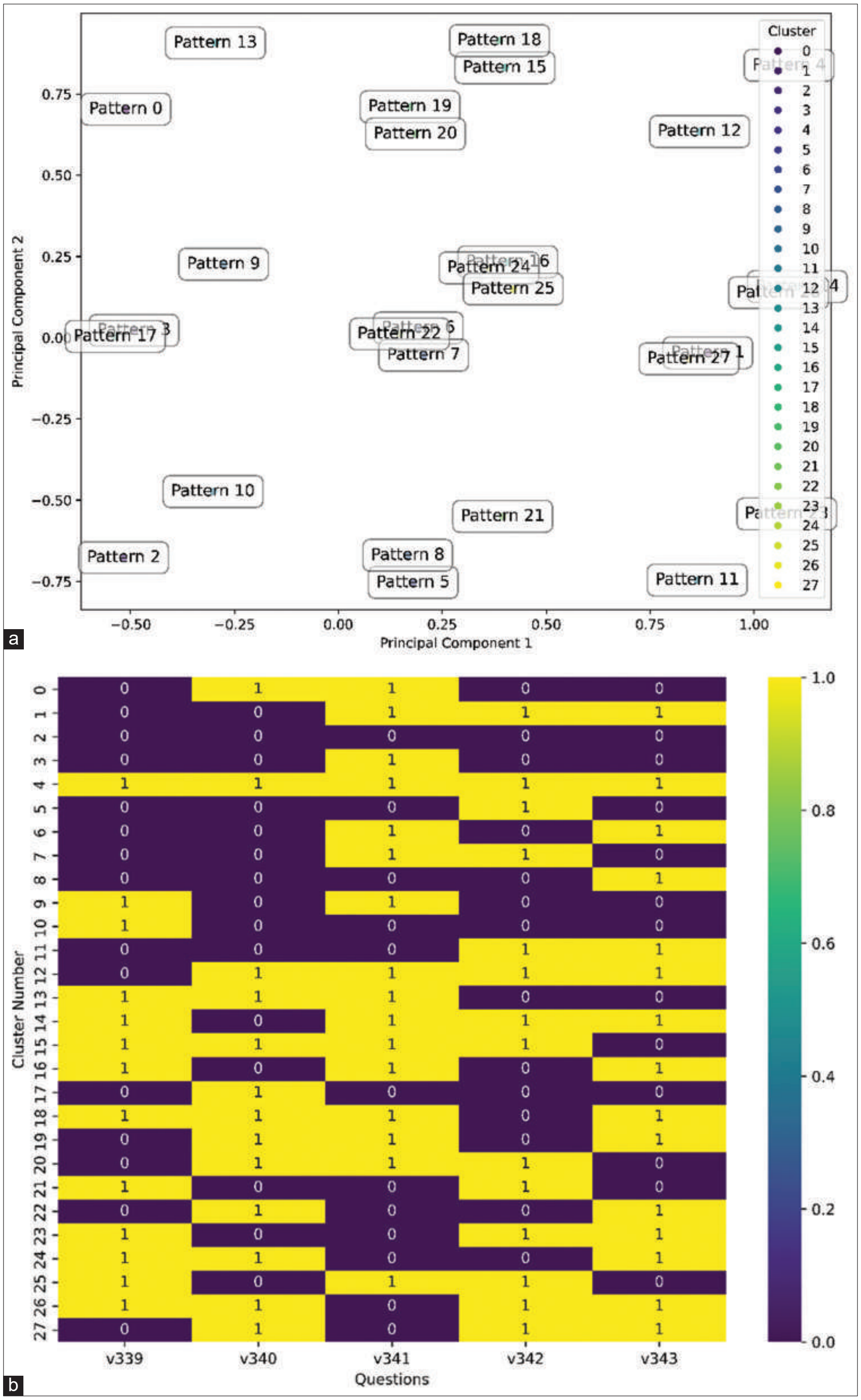

Figure 1: Farmers' response patterns relating to survey questions on identification of pests and diseases in their cassava farms. a) PCA plot showing response patters and b) heatmap showing variable-wise pattern characteristics.

soil moisture and negative for mineral deficiency. Response pattern 1 was characterized by farmers who did not believe that any of the alternatives caused CMD-related symptoms. Response pattern 2 was characterized by farmers who believed that only the use of infected cuttings caused CMD-related symptoms. Response pattern 3 was characterized by farmers

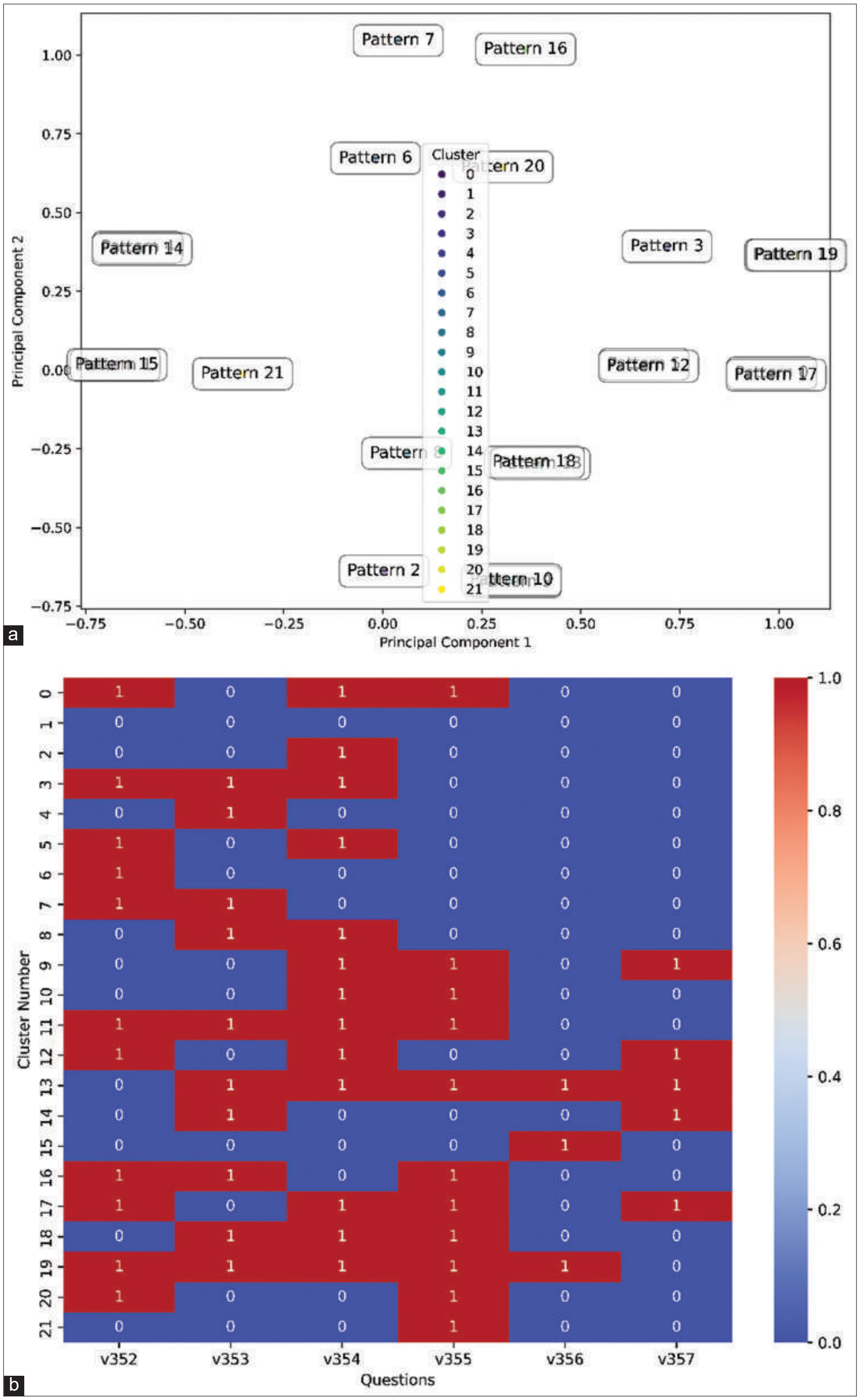

Figure 2: Farmers' response patterns relating to survey questions on causes of CMD-related symptoms. a) PCA plot showing response patters and b) heatmap showing variable-wise pattern characteristics.

who believed that a virus, the whitefly and the use of infected cuttings caused CMD-related symptoms. Response pattern 4 was characterized by farmers who believed that only the whitefly caused CMD-related symptoms. Response pattern 5 was characterized by farmers who believed that a virus and the use of infected cuttings caused CMD-related symptoms.

Response pattern 6 was characterized by farmers who believed that only a virus caused CMD-related symptoms.

Response pattern 7 was characterized by farmers who believed that a virus and the whitefly caused CMD-related symptoms. Response pattern 8 was characterized by farmers who believed that the whitefly and the use of infected cuttings caused CMD-related symptoms. Response pattern 9 was characterized by farmers who believed that the use of infected cuttings, lack of rain and mineral deficiency caused CMD-related symptoms. Response pattern 10 was characterized by farmers who believed that the use of infected cuttings and lack of rain caused CMD-related symptoms. Response pattern 11 was characterized by farmers who believed that all of the alternatives except moisture and mineral deficiency caused CMD-related symptoms. Response pattern 12 was characterized by farmers who believed that a virus, the use of infected cuttings and mineral deficiency caused CMD-related symptoms. Response pattern 13 was characterized by farmers who believed that all of the alternatives except a virus caused CMD-related symptoms.

Response pattern 14 was characterized by farmers who believed that all of the alternatives except the whitefly and mineral deficiency caused CMD-related symptoms. Response pattern 15 was characterized by farmers who believed that all of the alternatives except soil moisture caused CMD-related symptoms. Response pattern 16 was characterized by farmers who believed that a virus, the whitefly and lack of rain caused CMD-related symptoms. Response pattern 17 was characterized by farmers who believed that all of the alternatives except the whitefly and soil moisture caused CMD-related symptoms. Response pattern 18 was characterized by farmers who believed that the whitefly, the use of infected cuttings and lack of rain caused CMD-related symptoms. Response pattern 19 was characterized by farmers who believed that all of the alternatives except mineral deficiency caused CMD-related symptoms. Response pattern 20 was characterized by farmers who believed that a virus and lack of rain caused CMD-related symptoms. Response pattern 21 was characterized by farmers who believed that only a lack of rain caused CMD-related symptoms.

Worthy of note is the fact that 11 of the 22 response patterns did not associate a virus with CMD-related symptoms – response patterns 1, 2, 4, 8, 9, 10, 13, 14, 15, 18, 21. Only one response pattern uniquely associated a virus, the whitefly and the use of infected cuttings with CMD-related symptoms – response pattern 3. This suggests that the farmers are not knowledgeable on the viral diseases of Cassava.

Farmers' response patterns relating to survey questions on the impact of CMD on Cassava plants

Within the 508 individual responses, *k*-means analysis identified 9 distinct response patterns (Figure 3a). Response pattern 0 (Figure 3b) was characterized by farmers who believed that poor plant growth and decrease in yield were among the impacts of CMD on Cassava plants. Response pattern 1 was characterized by farmers who believed that neither poor plant growth, decrease in yield nor lack of healthy plant material were among the impacts of CMD on Cassava plants. Response pattern 2 was characterized by farmers who believed that only poor plant growth, decrease in yield and lack of healthy plant material were among the impacts of CMD on Cassava plants. Response pattern 3 was characterized by farmers who believed that only poor plant growth and decrease in yield were among the impacts of CMD on Cassava plants. Response pattern 4 was characterized by farmers who believed that only poor plant growth was among the impacts of CMD on Cassava plants. Response pattern 5 was characterized by farmers who believed that only a decrease in yield was among the impacts of CMD on Cassava plants. Response pattern 6 was characterized by farmers who believed that only poor plant growth and a lack of healthy plant material were among the impacts of CMD on Cassava plants. Response pattern 7 was characterized by farmers who believed that only lack of healthy plant material was among the impacts of CMD on Cassava plants. Response pattern 8 was characterized by farmers who believed that only a decrease in yield and lack of healthy plant material was among the impacts of CMD on Cassava plants. Only response pattern 1 associated neither poor plant growth, decrease in yield or lack of healthy plant material with the impacts of CMD on Cassava plants.

Farmers' response patterns relating to survey questions on farmers' reactions to CMD-related symptoms

Within the 508 individual responses, *k*-means analysis identified 32 distinct response patterns (Figure 4a). Response pattern 0 (Figure 4b) was characterized by farmers who mitigated CMD-related symptoms on their farms through the replacement of infected plants by healthy cuttings, analysis of the plants concerned with the Nuru application and consultation with agricultural agents. Response patterns 1 and 6 were characterized by farmers who did not mitigate CMD-related symptoms on their farms. Response pattern 2 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants, destruction of infected plants and replacement of infected plants by healthy cuttings. Response pattern 3 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants and destruction of infected plants. Response pattern 4 was characterized by farmers who mitigated CMD-related symptoms on their farms only through the removal of infected plants. Response pattern 5 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants and the replacement of infected plants by healthy cuttings. Response pattern 7 was characterized by farmers who mitigated CMD-related symptoms on their farms only through the replacement of infected plants by healthy cuttings. Response pattern 8 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants, destruction of infected plants and consultation with agricultural agents.

Response pattern 9 was characterized by farmers who mitigated CMD-related symptoms on their farms only through the destruction of infected plants. Response pattern 10 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected

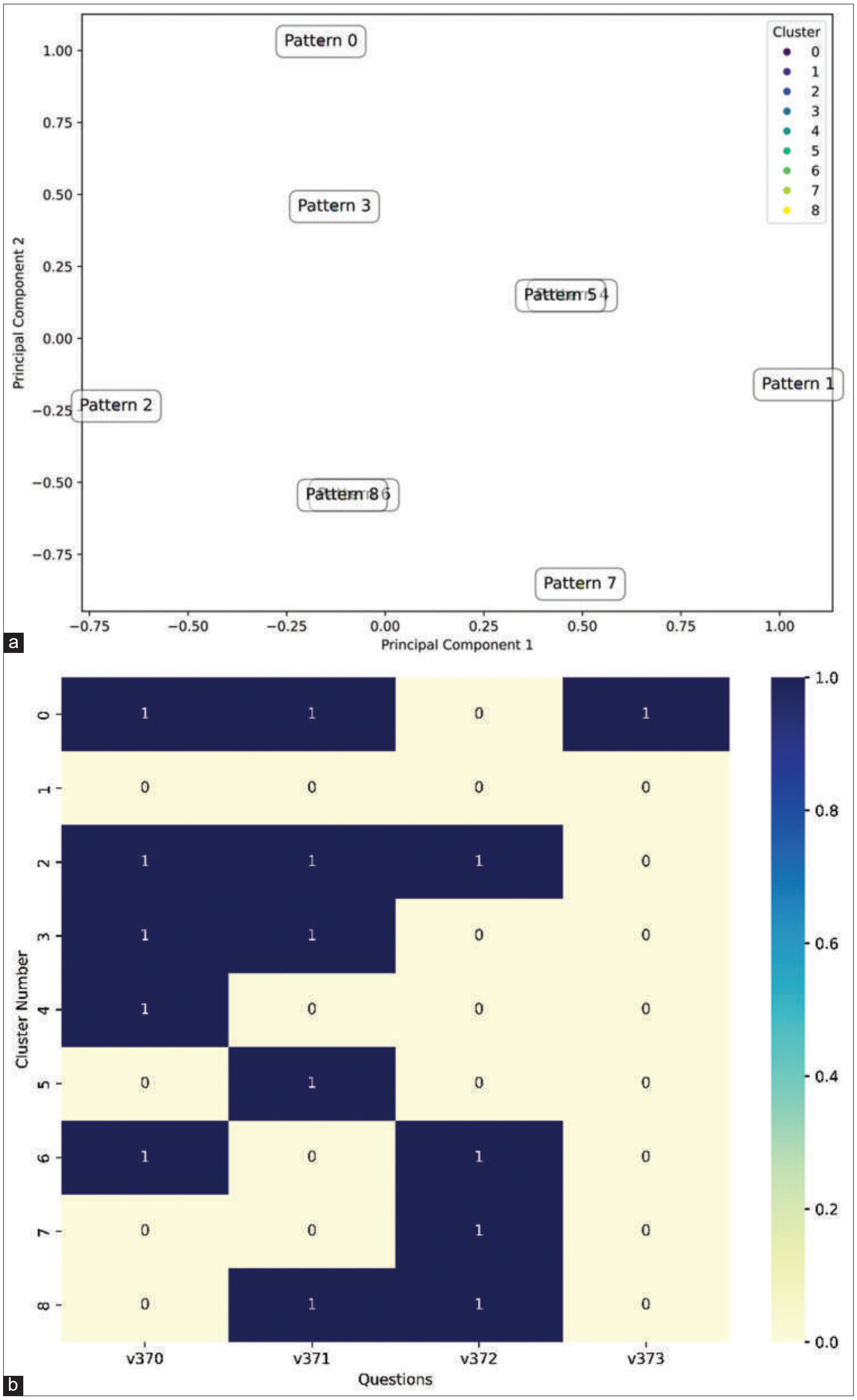

Figure 3: Farmers' response patterns relating to survey questions on the impact of CMD on cassava plants. a) PCA plot showing response patters and b) heatmap showing variable-wise pattern characteristics.

plants, destruction of infected plants and analysis of the plants concerned with the NURU application. Response pattern 11 was characterized by farmers who mitigated CMD-related symptoms on their farms through the replacement of infected plants by healthy cuttings, consultation with agricultural agents and the use of inputs. Response pattern 12 was characterized by

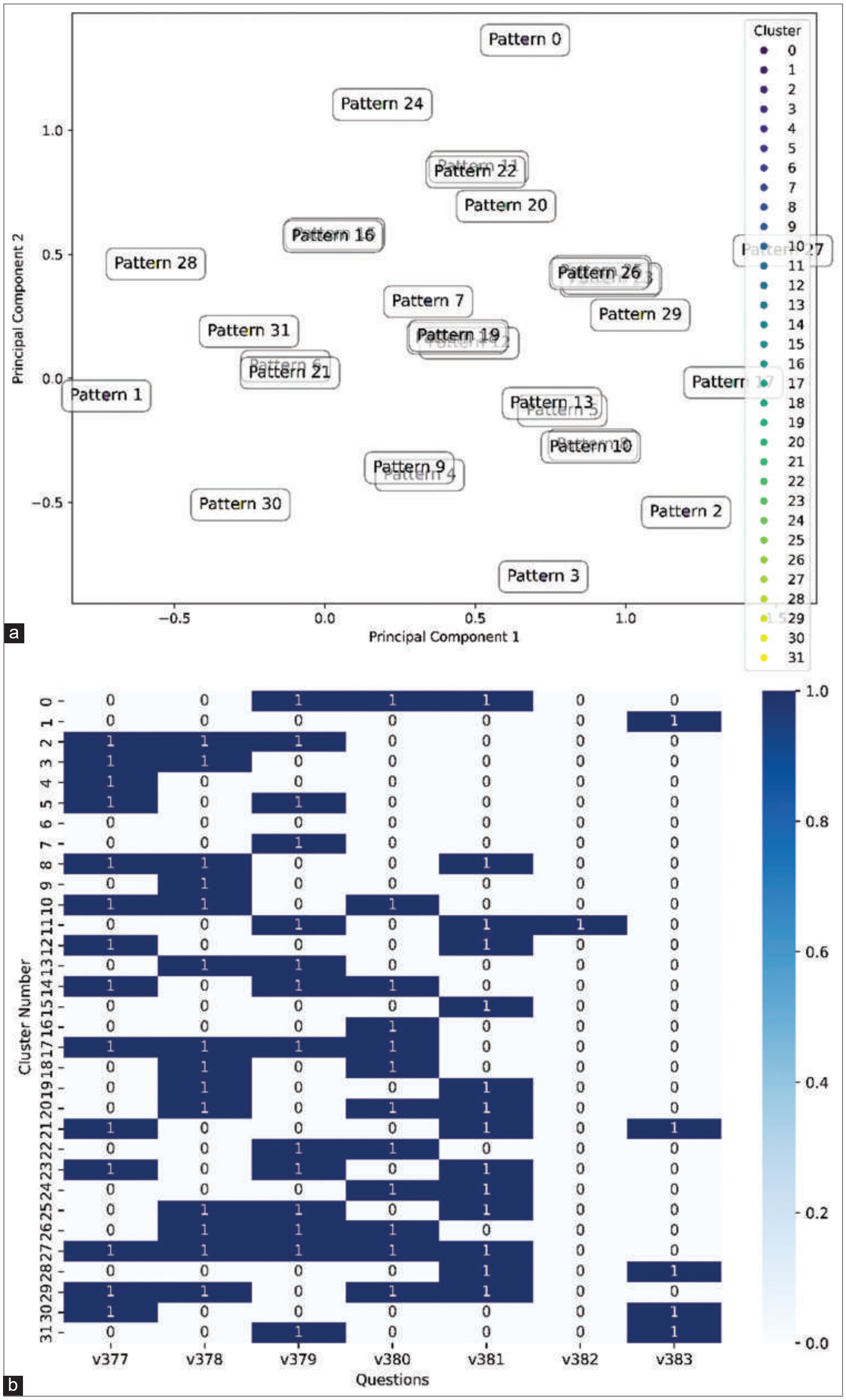

Figure 4: Farmers' response patterns relating to survey questions on farmers' reactions to CMD-related symptoms. a) PCA plot showing response patters and b) heatmap showing variable-wise pattern characteristics.

farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants and consultation with agricultural agents. Response pattern 13 was characterized by farmers who mitigated CMD-related symptoms on their farms through the destruction of infected plants and the replacement of infected plants by healthy cuttings. Response pattern 14 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants, replacement of infected plants by healthy cuttings and analysis of the plants concerned with the NURU application. Response pattern 15 was characterized by farmers who mitigated CMD-related symptoms on their farms only through consultation with agricultural agents.

Response pattern 16 was characterized by farmers who mitigated CMD-related symptoms on their farms only through the analysis of the plants concerned with the NURU application. Response pattern 17 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants, destruction of infected plants, and replacement of infected plants by healthy cuttings and analysis of the plants concerned with the NURU application. Response pattern 18 was characterized by farmers who mitigated CMD-related symptoms on their farms through the destruction of infected plants and analysis of the plants concerned with the NURU application. Response pattern 19 was characterized by farmers who mitigated CMD-related symptoms on their farms through the destruction of infected plants and consultation with agricultural agents. Response pattern 20 was characterized by farmers who mitigated CMD-related symptoms on their farms through the destruction of infected plants, analysis of the plants concerned with the NURU application and consultation with agricultural agents. Response pattern 21 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants and consultation with agricultural agents. Response pattern 22 was characterized by farmers who mitigated CMD-related symptoms on their farms through the replacement of infected plants by healthy cuttings and analysis of the plants concerned with the NURU application. Response pattern 23 was characterized by farmers who mitigated CMD-related symptoms on their farms through the removal of infected plants, replacement of infected plants by healthy cuttings and consultation with agricultural agents.

Response pattern 24 was characterized by farmers who mitigated CMD-related symptoms on their farms through the analysis of the plants concerned with the NURU application and consultation with agricultural agents. Response pattern 25 was characterized by farmers who mitigated CMD-related symptoms on their farms through the destruction of infected plants, and replacement of infected plants by healthy cuttings and consultation with agricultural agents. Response pattern 26 was characterized by farmers who mitigated CMD-related symptoms on their farms through the destruction of infected plants, replacement of infected plants by healthy cuttings and analysis of the plants concerned with the NURU application. Response pattern 27 was characterized by farmers who mitigated CMD-related symptoms on their farms using all of the six action categories except the use of inputs. Response pattern 28 was characterized by farmers who mitigated CMD-related symptoms on their farms only through consultation with agricultural agents. Response pattern 29 was characterized by farmers who mitigated CMD-related symptoms on their farms using all of the six action categories except the replacement of infected plants by healthy cuttings and the use of inputs. Response pattern 30 was characterized by farmers who mitigated CMD-related symptoms on their farms only through the removal of infected plants. Response pattern 31 was characterized by farmers who mitigated CMD-related symptoms on their farms only through the replacement of infected plants by healthy cuttings. Only response patterns 1 and 6 did not mitigate CMD-related symptoms on their farms.

Farmers' response patterns relating to survey questions on CMD prevention

Within the 508 individual responses, *k*-means analysis identified 16 distinct response patterns (Figure 5a). Response pattern 0 (Figure 5b) was characterized by farmers who indicated that CMD could be prevented through regular monitoring of fields (removal, destruction, and replacement of infected plants) and regular cleaning of the fields. Response pattern 1 was characterized by farmers who indicated that CMD could be prevented only through the use of healthy plant material. Response pattern 2 was characterized by farmers who indicated that CMD could be prevented using all of the five action categories. Response pattern 3 was characterized by farmers who indicated not knowing how to combat or prevent the onset of CMD. Response pattern 4 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material and regular monitoring of fields (removal, destruction, and replacement of infected plants). Response pattern 5 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material, regular monitoring of fields (removal, destruction, and replacement of infected plants) and regular cleaning of the fields. Response pattern 6 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material and regular cleaning of the fields. Response pattern 7 was characterized by farmers who indicated that CMD could be prevented only through regular cleaning of the fields.

Response pattern 8 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material, regular cleaning of the fields and respect of the planting density. Response pattern 9 was characterized by farmers who indicated that CMD could be prevented only through regular monitoring of fields (removal, destruction, and replacement of infected plants). Response pattern 10 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material and respect of the planting density. Response pattern 11 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material, regular monitoring of fields (removal, destruction, and replacement of infected plants) and respect of the planting density. Response pattern 12 was characterized by farmers who indicated that CMD could be prevented through regular monitoring of fields (removal, destruction, and replacement of infected plants), regular cleaning of the fields and respect of the planting density. Response pattern 13 was characterized by farmers who indicated

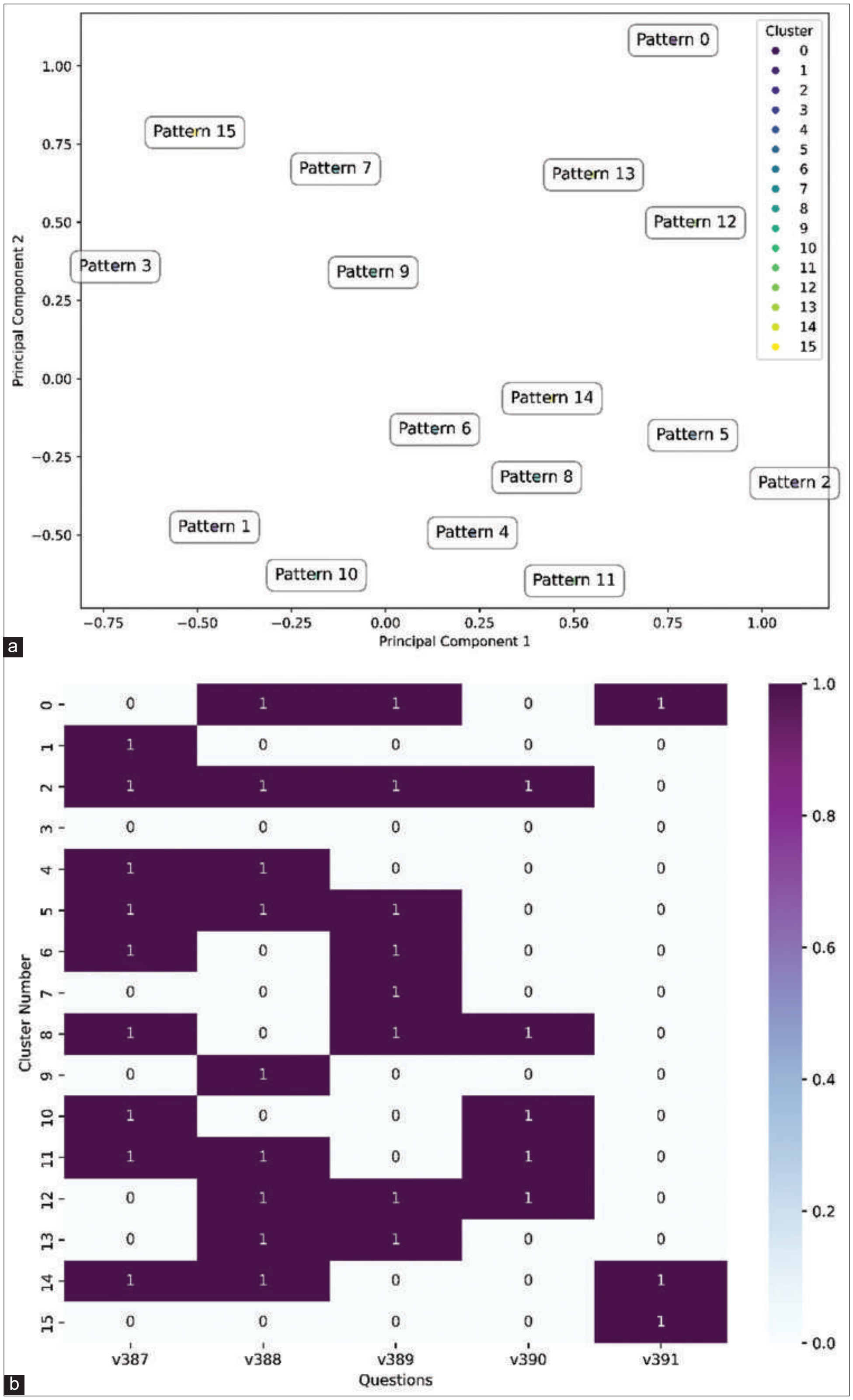

Figure 5: Farmers' response patterns relating to survey questions on CMD prevention. a) PCA plot showing response patters and b) heatmap showing variable-wise pattern characteristics.

that CMD could be prevented through regular monitoring of fields (removal, destruction, and replacement of infected plants) and regular cleaning of the fields. Response pattern 14 was characterized by farmers who indicated that CMD could be prevented through the use of healthy plant material and regular monitoring of fields (removal, destruction, and replacement of infected plants). Response pattern 15 was characterized by farmers who indicated other means (not amongst the action categories) by which CMD could be prevented. Only response pattern 3 reported not knowing how to combat or prevent the onset of CMD.

Farmers' perceptions of cassava diseases and management technologies

Table 3 and Figure 6 numerically and graphically summarize farmers' perceptions of CMD and management technologies, respectively.

39.76% of respondents did not believe that cassava viral diseases are caused by poor hygiene on the field, while 37.80% were unsure and 22.44% agreed. 49.80% of respondents did not believe that viral symptoms observed on cassava leaves result from the application of herbicides, while 34.45% were unsure and 15.75% agreed. 45.87% of respondents did not believe that older plants are more attacked by cassava viral diseases, while 43.11% were unsure and 11.02% agreed. 40.16% of respondents did not believe that late planting can lead to cassava viral diseases, while 35.63% were unsure and 24.21% agreed. 37.40% of respondents did not believe that drought and high temperatures can lead to cassava viral diseases, while 35.24% were unsure and 27.36% agreed.

45.28% of respondents did not believe that planting in muddy or waterlogged soils causes infections, while 32.28% were unsure and 22.24% agreed. 43.70% of respondents did not believe that poor aeration promotes cassava viral diseases, while 34.84% were unsure and 21.46% agreed. 48.62% of respondents did not believe that a late harvest can lead to cassava viral diseases, while 37.99% were unsure and 13.39% agreed. 54.72% of respondents did not believe that cassava viral diseases are caused by the use of poor-quality planting, while 39.57% were unsure and 5.71% agreed. 55.51% of respondents did not believe that cassava viral diseases can be managed by breaking the affected part, while 34.45% were unsure and 10.04% agreed.

41.34% of respondents did not believe that the management practices can easily be integrated into the traditional farming system, while 31.30% were unsure and 27.36% agreed. 52.56% of respondents did not believe that the management practices taught by the agricultural officers are complex to understand, while 25.39% were unsure and 22.05% agreed. 42.91% of respondents did not believe that the use of integrated approaches to viral disease control in cassava is more expensive than using chemicals, while 29.13% were unsure and 27.95% agreed. 47.64% of respondents did not believe that chemicals are effective in controlling cassava viral diseases, while 31.69% were unsure and 20.67% agreed. 44.69% of respondents did not believe that no cassava variety is resistant to viral diseases, while 31.50% were unsure and 23.82% agreed.

42.32% of respondents did not believe that there is no control for cassava viral diseases, while 38.39% were unsure and 19.29% agreed. 42.72% of respondents did not believe that the plant infected by cassava viral diseases always recovers at the beginning of the rains, while 33.07% were unsure and 24.21% agreed. 51.38% of respondents did not believe that the management practices of cassava viral diseases are not culturally accepted in my community, while 25.59% were unsure and 23.03%

Table 3: Farmers' perceptions of CMD and its management

Variable	Description	% Disagree	% Don't know	% Agree
V412	cassava viral diseases are caused by poor hygiene on the field	39.76	37.80	22.44
V413	viral symptoms observed on cassava leaves result from the application of herbicides	49.80	34.45	15.75
V414	older plants are more attacked by cassava viral diseases	45.87	43.11	11.02
V415	late planting can lead to cassava viral diseases	40.16	35.63	24.21
V416	drought and high temperatures can lead to cassava viral diseases	37.40	35.24	27.36
V417	planting in muddy or waterlogged soils causes infections	45.28	32.28	22.24
V418	poor aeration promotes cassava viral diseases	43.70	34.84	21.46
V419	a late harvest can lead to cassava viral diseases	48.62	37.99	13.39
V420	cassava viral diseases are caused by the use of poor-quality planting	54.72	39.57	5.71
V421	cassava viral diseases can be managed by breaking the affected part	55.51	34.45	10.04
V422	the management practices can easily be integrated into the traditional farming system	41.34	31.30	27.36
V423	the management practices taught by the agricultural officers are complex to understand	52.56	25.39	22.05
V424	the use of integrated approaches to viral diseases control in cassava is more expensive than using chemicals	42.91	29.13	27.95
V425	chemicals are effective in controlling cassava viral diseases	47.64	31.69	20.67
V426	no cassava variety is resistant to viral diseases	44.69	31.50	23.82
V427	there is no control for cassava viral diseases	42.32	38.39	19.29
V428	the plant infected by cassava viral diseases always recovers at the beginning of the rains	42.72	33.07	24.21
V429	the management practices of cassava viral diseases are not culturally accepted in my community	51.38	25.59	23.03
V430	the use of cassava diseases management technologies helps to reduce the incidence of cassava viral diseases	50.20	32.28	17.52
V431	the use of cassava disease management technologies increases productivity	48.81	41.73	9.45
V432	the management technologies are accessible at all times	37.40	37.20	25.39
V433	cassava viral diseases prevent rooting	45.28	31.89	22.83
V434	whenever cassava plants are attacked by viral diseases, it results in poor quality tubers	53.74	40.55	5.71
V435	viral diseases of cassava can lead to 100% yield loss if left untreated	40.16	32.68	27.17
V436	cassava viral diseases result in loss of planting material	58.27	39.57	2.17

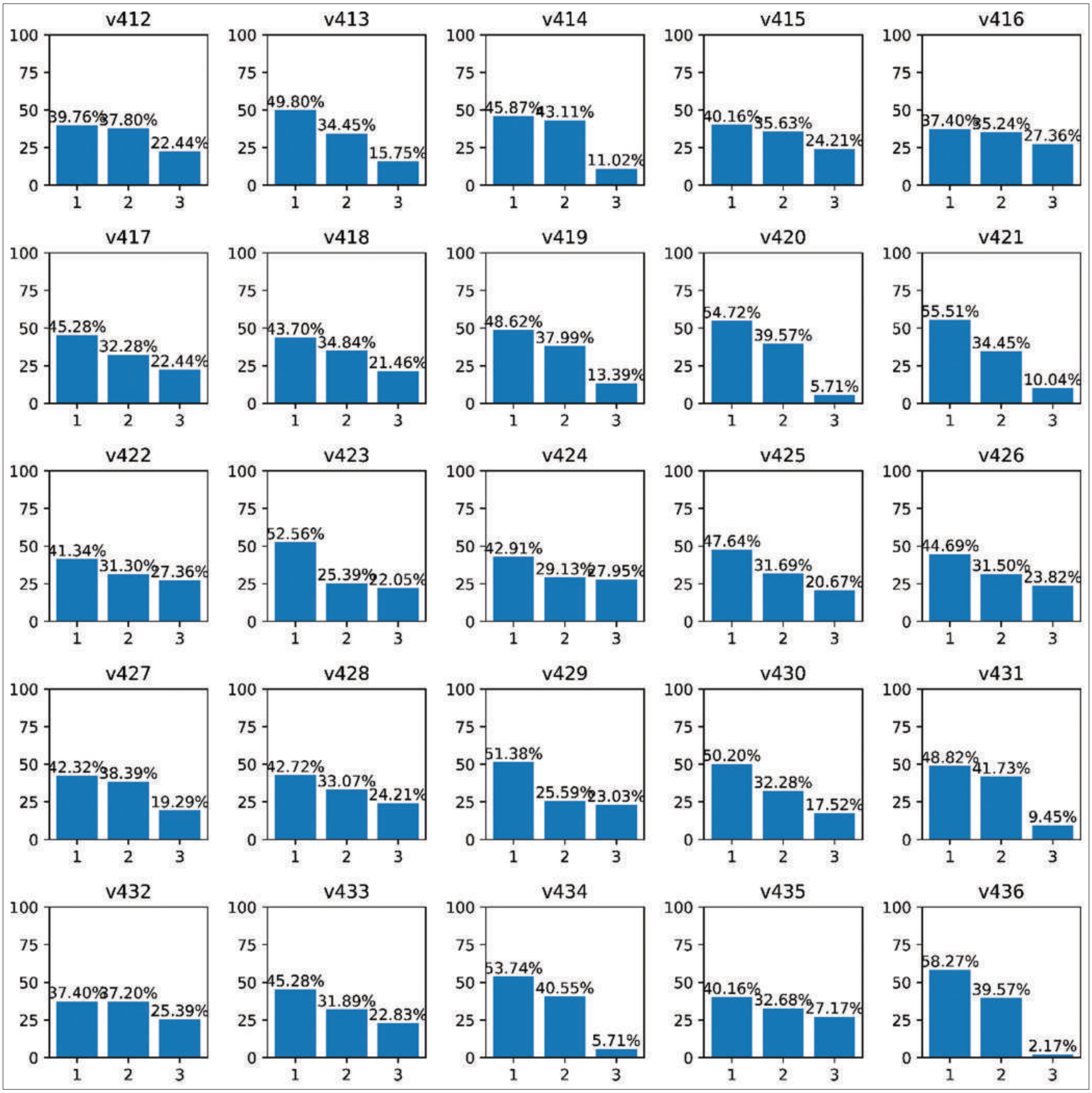

Figure 6: Farmers' perceptions of CMD and its management. 1-Disagree, 2-Don't know and 3-Agree.

agreed. 50.20% of respondents did not believe that the use of cassava diseases management technologies helps to reduce the incidence of cassava viral diseases., while 32.28% were unsure and 17.52% agreed. 48.81% of respondents did not believe that the use of cassava disease management technologies increases productivity, while 41.73% were unsure and 9.45% agreed.

37.40% of respondents did not believe that the management technologies are accessible at all times, while 37.20% were unsure and 25.39% agreed. 45.28% of respondents did not believe that cassava viral diseases prevent rooting, while 31.89% were unsure and 22.83% agreed. 53.74% of respondents did not believe that whenever cassava plants are attacked by viral diseases, it results in poor quality tubers, while 40.55% were unsure and 5.71% agreed. 40.16% of respondents did not believe that viral diseases of cassava can lead to 100% yield loss if left untreated, while 32.68% were unsure and 27.17% agreed. 58.27% of respondents did not believe that cassava viral diseases result in loss of planting material, while 39.57% were unsure and 2.17% agreed.

Overall, 18.88% of farmers (or 96 out of 508 farmers) know about cassava diseases and management technologies. This finding highlights a significant gap in agricultural education and extension services. This low level of awareness could lead to ineffective disease control, reduced crop yields, and economic losses, considering cassava's importance as a staple food and income source in Cameroon. It suggests an urgent need for improved dissemination of information and training for farmers to enhance their capacity to manage cassava diseases effectively, which could ultimately contribute to food security and sustainable agricultural practices in the region.

CONCLUSION

This study has highlighted the urgency of the critical need for agricultural extension services to provide education and resources to farmers, ensuring they are equipped with the necessary skills and knowledge to protect their crops and livelihoods effectively.

REFERENCES

Adebayo, W. G. (2023). Cassava production in africa: A panel analysis of the drivers and trends. *Heliyon, 9*(9), e19939. https://doi.org/10.1016/j.heliyon.2023.e19939

Alleyne, A., Mason, S., & Vallès, Y. (2023). Characterization of the Cassava Mycobiome in Symptomatic Leaf Tissues Displaying Cassava Superelongation Disease. *Journal of Fungi, 9*(12), 12. https://doi.org/10.3390/jof9121130

Alves, A. A. C., de Oliveira, L. A., & da Silva Motta, J. (2022). Transferring Cassava Processing Technology from Brazil to Africa. In G. Thiele, M. Friedmann, H. Campos, V. Polar & J. W. Bentley (Eds.), *Root, Tuber and Banana Food System Innovations: Value Creation for Inclusive Outcomes* (pp. 207-239) Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-92022-7_7

Bailey, S. (2012). Principal Component Analysis with Noisy and/or Missing Data. *Publications of the Astronomical Society of the Pacific, 124*(919), 1015-1023. https://doi.org/10.1086/668105

Bartkowski, J., Odindo, M. O., & Otieno, W. A. (1988). Some Fungal Pathogens of the Cassava Green Spider Mites *Mononychellus* spp. (Tetranychidae) in Kenya. *International Journal of Tropical Insect Science, 9*(4), 457-459. https://doi.org/10.1017/S174275840001095X

Bateni, M., Cohen-Addad, V., Epasto, A., & Lattanzi, S. (2024). A Scalable Algorithm for Individually Fair K-means Clustering. *arXiv*, arXiv:2402.06730. https://doi.org/10.48550/arXiv.2402.06730

Bloom, E. H., Atallah, S. S., & Casteel, C. L. (2024). Motivating organic farmers to adopt practices that support the pest-suppressive microbiome relies on understanding their beliefs. *Renewable Agriculture and Food Systems, 39*, e8. https://doi.org/10.1017/S174217052400005X

Bottrell, D. G., & Schoenly, K. G. (2018). Integrated pest management for resource-limited farmers: Challenges for achieving ecological, social and economic sustainability. *The Journal of Agricultural Science, 156*(3), 408-426. https://doi.org/10.1017/S0021859618000473

Brévault, T., & Clouvel, P. (2019). Pest management: Reconciling farming practices and natural regulations. *Crop Protection, 115*, 1-6. https://doi.org/10.1016/j.cropro.2018.09.003

Brito, A. C., Oliveira, S. A. S., & Oliveira, E. J. (2017). Genome-wide association study for resistance to cassava root rot. *The Journal of Agricultural Science, 155*(9), 1424-1441. https://doi.org/10.1017/S0021859617000612

Burns, A., Gleadow, R., Cliff, J., Zacarias, A., & Cavagnaro, T. (2010). Cassava: The Drought, War and Famine Crop in a Changing World. *Sustainability, 2*(11), 3572-3607. https://doi.org/10.3390/su2113572

Capó, M., Pérez, A., & Lozano, J. A. (2018). An efficient K -means clustering algorithm for massive data. *arXiv*, arXiv:1801.02949. https://doi.org/10.48550/arXiv.1801.02949

Chavez, V. A., Milne, A. E., van den Bosch, F., Pita, J., & McQuaid, C. F. (2022). Modelling cassava production and pest management under biotic and abiotic constraints. *Plant Molecular Biology, 109*, 325-349. https://doi.org/10.1007/s11103-021-01170-8

Chen, E., Chen, X., Jing, W., & Zhang, Y. (2024). Distributed Tensor Principal Component Analysis. *arXiv,* arXiv:2405.11681. https://doi.org/10.48550/arXiv.2405.11681

Chen, Y. T., & Witten, D. M. (2022). Selective inference for k-means clustering. *arXiv,* arXiv:2203.15267. https://doi.org/10.48550/arXiv.2203.15267

Clum, C., Mixon, D. G., Villar, S., & Xie, K. (2022). Sketch-and-solve approaches to k-means clustering by semidefinite programming. *arXiv,* arXiv:2211.15744. https://doi.org/10.48550/arXiv.2211.15744

da Silva, J. S. A., Alves, V. C. S., da Silva, S. F., do Nascimento Barbosa, R., de Souza, C. A. F., da Costa, D. P., Machado, A. R., de Medeiros, E. V., & de Souza-Motta, C. M. (2024). *Diaporthe ueckeri* causing cassava root rot in Pernambuco, Brazil. *Crop Protection, 184*, 106811. https://doi.org/10.1016/j.cropro.2024.106811

Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A. S., & McLoone, S. (2013). Principal Component Analysis on Spatial Data: An Overview. *Annals of the Association of American Geographers, 103*(1), 106-128. https://doi.org/10.1080/00045608.2012.689236

Dorabiala, O., Aravkin, A., & Kutz, J. N. (2023). Ensemble Principal Component Analysis. *arXiv,* arXiv:2311.01826. https://doi.org/10.48550/arXiv.2311.01826

Dorabiala, O., Dabke, D. V., Webster, J., Kutz, N., & Aravkin, A. (2024). Spatiotemporal k-means. *arXiv,* arXiv:2211.05337. https://doi.org/10.48550/arXiv.2211.05337

Doungous, O., Masky, B., Levai, D. L., Bahoya, J. A. L., Minyaka, E., Mavoungou, J. F., Mutuku, J. M., & Pita, J. S. (2022). Cassava mosaic disease and its whitefly vector in Cameroon: Incidence, severity and whitefly numbers from field surveys. *Crop Protection, 158*, 106017. https://doi.org/10.1016/j.cropro.2022.106017

Du, L., Wang, B., Wang, P., Ma, Y., & Liu, H. (2015). Noise Reduction Method Based on Principal Component Analysis With Beta Process for Micro-Doppler Radar Signatures. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(8), 4028-4040. https://doi.org/10.1109/JSTARS.2015.2451004

Ekundayo, J. A., & Daniel, T. M. (1973). Cassava rot and its control. *Transactions of the British Mycological Society, 61*(1), 27-32. https://doi.org/10.1016/S0007-1536(73)80084-1

Elhaik, E. (2022). Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Scientific Reports, 12*, 14683. https://doi.org/10.1038/s41598-022-14395-4

Elliot, S. L., Mumford, J. D., de Moraes, G. J., & Sabelis, M. W. (2002). Age-Dependent Rates of Infection of Cassava Green Mites by a Fungal Pathogen in Brazil. *Experimental & Applied Acarology, 27*, 169-180. https://doi.org/10.1023/A:1021644321360

Ergun, J. C., Feng, Z., Silwal, S., Woodruff, D. P., & Zhou, S. (2022). Learning-Augmented k-means Clustering. *arXiv,* arXiv:2110.14094. https://doi.org/10.48550/arXiv.2110.14094

Fanou, A. A., Zinsou, V. A., & Wydra, K. (2017). Cassava Bacterial Blight: A Devastating Disease of Cassava. In V. Y. Waisundara (Eds.), *Cassava* London, UK: IntechOpen. https://doi.org/10.5772/intechopen.71527

Fathima, A. A., Sanitha, M., Tripathi, L., & Muiruri, S. (2023). Cassava (Manihot esculenta) dual use for food and bioenergy: A review. *Food and Energy Security, 12*(1), e380. https://doi.org/10.1002/fes3.380

Garst, S., & Reinders, M. (2024). Federated K-means Clustering. *arXiv,* arXiv:2310.01195. https://doi.org/10.48550/arXiv.2310.01195

Gewers, F. L., Ferreira, G. R., de Arruda, H. F., Silva, F. N., Comin, C. H., Amancio, D. R., & Costa, L. da F. (2018). Principal Component Analysis: A Natural Approach to Data Exploration. *ACM Computing Surveys, 54*(4), 70. https://doi.org/10.1145/3447755

Gniazdowski, Z. (2017). New Interpretation of Principal Components Analysis. *Zeszyty Naukowe WWSI, 11*(16), 43-65. https://doi.org/10.26348/znwwsi.16.43

Görtler, J., Spinner, T., Streeb, D., Weiskopf, D., & Deussen, O. (2020). Uncertainty-Aware Principal Component Analysis. *IEEE Transactions on Visualization and Computer Graphics, 26*(1), 822-831. https://doi.org/10.1109/TVCG.2019.2934812

Guo, Q., Wu, W., Massart, D. L., Boucon, C., & de Jong, S. (2002). Feature selection in principal component analysis of analytical data. *Chemometrics and Intelligent Laboratory Systems, 61*(1-2), 123-132. https://doi.org/10.1016/S0169-7439(01)00203-9

Hareesh, P. S., Resmi, T. R., Sheela, M. N., & Makeshkumar, T. (2023). Cassava mosaic disease in South and Southeast Asia: Current status and prospects. *Frontiers in Sustainable Food Systems, 7*, 1086660. https://doi.org/10.3389/fsufs.2023.1086660

Hohenfeld, C. S., de Oliveira, S. A. S., Ferreira, C. F., Mello, V. H., Margarido, G. R. A., Passos, A. R., & de Oliveira, E. J. (2024). Comparative analysis of infected cassava root transcriptomics reveals candidate genes for root rot disease resistance. *Scientific Reports, 14*, 10587. https://doi.org/10.1038/s41598-024-60847-4

Hope, T. M. H., Price, C. J., Halai, A., Salvi, C., Crinion, J., Keijsers, M., Sperber, C., & Bowman, H. (2024). Estimating the construct validity of Principal Components Analysis. *arXiv,* arXiv:2401.12905. https://doi.org/10.48550/arXiv.2401.12905

Hurwitz, R. M., & Hahn, G. (2023). Penalized Principal Component Analysis using Nesterov Smoothing. *ArXiv,* arXiv:2309.13838v1. https://doi.org/10.48550/arXiv.2309.13838

Johnstone, I. M., & Lu, A. Y. (2009). On Consistency and Sparsity for Principal Components Analysis in High Dimensions. *Journal of the American Statistical Association, 104*(486), 682-693. https://doi.org/10.1198/jasa.2009.0121

Jolliffe, I. (2011). Principal Component Analysis. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 1094-1096) Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-642-04898-2_455

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Kim, K., Kim, J., & Kennedy, E. H. (2024). Causal K-Means Clustering. *arXiv,* arXiv:2405.03083. https://doi.org/10.48550/arXiv.2405.03083

Laizer, H. C., Chacha, M. N., & Ndakidemi, P. A. (2019). Farmers' Knowledge, Perceptions and Practices in Managing Weeds and Insect Pests of Common Bean in Northern Tanzania. *Sustainability, 11*(15), 4076. https://doi.org/10.3390/su11154076

Lee, J., Cho, H., Yun, S.-Y., & Yun, C. (2023). Fair Streaming Principal Component Analysis: Statistical and Algorithmic Viewpoint. *arXiv,* arXiv:2310.18593. https://doi.org/10.48550/arXiv.2310.18593

Lee, S. (2011). Drawbacks of Principal component analysis. *arXiv,* arXiv:1005.1770. https://doi.org/10.48550/arXiv.1005.1770

Legg, J. (2008). African cassava mosaic disease. In F. Claude (Eds.), *Encyclopedia of virology* (pp. 30-36) Oxford: Elsevier. https://doi.org/10.1016/b978-012374410-4.00693-2

Leiva, A. M., Pardo, J. M., Arinaitwe, W., Newby, J., Vongphachanh, P., Chittarath, K., Oeurn, S., Hang, L. T., Gil-Ordóñez, A., Rodriguez, R., & Cuellar, W. J. (2023). *Ceratobasidium* sp. Is associated with cassava witches' broom disease, a re-emerging threat to cassava cultivation in Southeast Asia. *Scientific Reports, 13*, 22500. https://doi.org/10.1038/s41598-023-49735-5

Leroux, A., Crainiceanu, C., & Wrobel, J. (2023). Fast Generalized Functional Principal Components Analysis. *arXiv,* arXiv:2305.02389. https://doi.org/10.48550/arXiv.2305.02389

Li, B. (2018). A Principal Component Analysis Approach to Noise Removal for Speech Denoising. 2018 International Conference on Virtual Reality and Intelligent Systems (pp. 429-432). IEEE. https://doi.org/10.1109/ICVRIS.2018.00111

Ling-Qun, W., Bing-Bing, L., Jun, L., Bin, X., Qi, W., Yu-Qi, C., & Kai-Guang, Z. (2015). Noise Removal Based on Filtered Principal Component Reconstruction. *Chinese Journal of Geophysics, 58*(5), 589-598. https://doi.org/10.1002/cjg2.20197

Maćkiewicz, A., & Ratajczak, W. (1993). Principal components analysis (PCA). *Computers & Geosciences, 19*(3), 303-342. https://doi.org/10.1016/0098-3004(93)90090-R

Makambila, C. (1994). The fungal diseases of cassava in the republic of Congo, central Africa. *African Crop Science Journal, 2*(4), 4.

Marzban, C., Yurtsever, U., & Richman, M. (2024). Principal Component Analysis for Equation Discovery. *arXiv,* arXiv:2401.04797. https://doi.org/10.48550/arXiv.2401.04797

McCallum, E. J., Anjanappa, R. B., & Gruissem, W. (2017a). Tackling agriculturally relevant diseases in the staple crop cassava (Manihot esculenta). *Current Opinion in Plant Biology, 38*, 50-58. https://doi.org/10.1016/j.pbi.2017.04.008

Miao, S., Zheng, L., Liu, J., & Jin, H. (2023). K-means Clustering Based Feature Consistency Alignment for Label-free Model Evaluation. *arXiv,* arXiv:2304.09758. https://doi.org/10.48550/arXiv.2304.09758

Milne, A. E., Bell, J. R., Hutchison, W. D., van den Bosch, F., Mitchell, P. D., Crowder, D., Parnell, S., & Whitmore, A. P. (2015). The Effect of Farmers' Decisions on Pest Control with Bt Crops: A Billion Dollar Game of Strategy. *PLoS Computational Biology, 11*(12), e1004483. https://doi.org/10.1371/journal.pcbi.1004483

Mishra, D., Dash, R., Rath, A. K., & Acharya, M. (2011). Feature selection in gene expression data using principal component analysis and rough set theory. *Advances in Experimental Medicine and Biology, 696*, 91-100. https://doi.org/10.1007/978-1-4419-7046-6_10

Miyittah, M. K., Kosivi, R. K., Tulashie, S. K., Addi, M. N., & Tawiah, J. Y. (2022). The need for alternative pest management methods to mitigate risks among cocoa farmers in the Volta region, Ghana. *Heliyon, 8*(12), e12591. https://doi.org/10.1016/j.heliyon.2022.e12591

Mohammadi, S. O., Kalhor, A., & Bodaghi, H. (2022). K-Splits: Improved K-Means Clustering Algorithm to Automatically Detect the Number of Clusters. In A. P. Pandian, X. Fernando & W. Haoxiang (Eds.), *Computer Networks, Big Data and IoT* (Vol. 117, pp. 197-213) Singapore: Springer. https://doi.org/10.1007/978-981-19-0898-9_15

Mohidin, S. R. N. S. P., Moshawih, S., Hermansyah, A., Asmuni, M. I., Shafqat, N., & Ming, L. C. (2023). Cassava (*Manihot esculenta* Crantz): A Systematic Review for the Pharmacological Activities, Traditional Uses, Nutritional Values, and Phytochemistry. *Journal of Evidence-Based Integrative Medicine, 2023*, 28. https://doi.org/10.1177/2515690X231206227

Morgan, N. K., & Choct, M. (2016). Cassava: Nutrient composition and nutritive value in poultry diets. *Animal Nutrition, 2*(4), 253-261. https://doi.org/10.1016/j.aninu.2016.08.010

Mussabayev, R., Mladenovic, N., Jarboui, B., & Mussabayev, R. (2023). How to Use K-means for Big Data Clustering? *Pattern Recognition, 137*, 109269. https://doi.org/10.1016/j.patcog.2022.109269

Mustarichie, R., Sulistyaningsih, S., & Runadi, D. (2020). Antibacterial Activity Test of Extracts and Fractions of Cassava Leaves (*Manihot esculenta* Crantz) against Clinical Isolates of *Staphylococcus epidermidis* and *Propionibacterium acnes* Causing Acne. *International Journal of Microbiology, 2020*, 1975904. https://doi.org/10.1155/2020/1975904

Nakatumba-Nabende, J., Akera, B., Tusubira, J. F., Nsumba, S., & Mwebaze, E. (2020). A dataset of necrotized cassava root cross-section images. *Data in Brief, 32*, 106170. https://doi.org/10.1016/j.dib.2020.106170

Ndunguru, J., De León, L., Doyle, C. D., Sseruwagi, P., Plata, G., Legg, J. P., Thompson, G., Tohme, J., Aveling, T., Ascencio-Ibáñez, J. T., & Hanley-Bowdoin, L. (2016). Two Novel DNAs That Enhance Symptoms and Overcome CMD2 Resistance to Cassava Mosaic Disease. *Journal of Virology, 90*(8), 4160-4173. https://doi.org/10.1128/jvi.02834-15

Niño-Jimenez, D.-P., López-López, K., & Cuervo-Ibáñez, M. (2024). Quantitative detection of cassava common mosaic virus for health certification of cassava (*Manihot esculenta* Crantz) germplasm using qPCR analysis. *Heliyon, 10*(6), e27604. https://doi.org/10.1016/j.heliyon.2024.e27604

Okike, I., Wigboldus, S., Samireddipalle, A., Naziri, D., Adesehinwa, A. O. K., Adejoh, V. A., Amole, T., Bordoloi, S., & Kulakow, P. (2022). Turning Waste to Wealth: Harnessing the Potential of Cassava Peels for Nutritious Animal Feed. In G. Thiele, M. Friedmann, H. Campos, V. Polar, & J. W. Bentley (Eds.), *Root, Tuber and Banana Food System Innovations: Value Creation for Inclusive Outcomes* (pp. 173-206) Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-92022-7_6

Onyeka, T. J., Dixon, A. G. O., & Ekpo, E. J. A. (2005). Identification of levels of resistance to cassava root rot disease (*Botryodiplodia theobromae*) in African landraces and improved germplasm using in vitro inoculation method. *Euphytica, 145*(3), 281-288. https://doi.org/10.1007/s10681-005-1646-8

Otim-Nape, G. W., & Thresh, J. M. (1998). The current pandemic of cassava mosaic virus disease in Uganda. In D. G. Jones (Eds.), *The Epidemiology of Plant Diseases* (pp. 423-443) Netherlands: Springer. https://doi.org/10.1007/978-94-017-3302-1_21

Owomugisha, G., Nakatumba-Nabende, J., Dhikusooka, J. J., Taravera, E., Nuwamanya, E., & Mwebaze, E. (2023). A labeled spectral dataset with cassava disease occurrences using virus titre determination protocol. *Data in Brief, 49*, 109387. https://doi.org/10.1016/j.dib.2023.109387

Pardo, J. M., Chittarath, K., Vongphachanh, P., Hang, L. T., Oeurn, S., Arinaitwe, W., Rodriguez, R., Sophearith, S., Malik, A. I., & Cuellar, W. J. (2023). Cassava Witches' Broom Disease in Southeast Asia: A Review of Its Distribution and Associated Symptoms. *Plants, 12*(11), 2217. https://doi.org/10.3390/plants12112217

Pérez, D., Duputié, A., Vernière, C., Szurek, B., & Caillon, S. (2022). Biocultural Drivers Responsible for the Occurrence of a Cassava

Bacterial Pathogen in Small-Scale Farms of Colombian Caribbean. *Frontiers in Ecology and Evolution, 10*, 841915. https://doi.org/10.3389/fevo.2022.841915

Pham, C. V., & Tran, H. T. (2021). Cunninghamella elegans causing cassava root rot in Vietnam. *Australasian Plant Disease Notes, 16*, 14. https://doi.org/10.1007/s13314-021-00427-x

Phung, Q. A., & Dao, N. (2024). Farmers' perceptions of sustainable agriculture in the Red River Delta, Vietnam. *Heliyon, 10*(7), e28576. https://doi.org/10.1016/j.heliyon.2024.e28576

Poggiali, A., Berti, A., Bernasconi, A., Del Corso, G. M., & Guidotti, R. (2024). Quantum Clustering with k-Means: A Hybrid Approach. *Theoretical Computer Science, 992*, 114466. https://doi.org/10.1016/j.tcs.2024.114466

Rahmat, F., Zulkafli, Z., Ishak, A. J., Abdul Rahman, R. Z., Stercke, S. D., Buytaert, W., Tahir, W., Ab Rahman, J., Ibrahim, S., & Ismail, M. (2024). Supervised feature selection using principal component analysis. *Knowledge and Information Systems, 66*, 1955-1995. https://doi.org/10.1007/s10115-023-01993-5

Razifar, P., Muhammed, H. H., Engbrant, F., Svensson, P.-E., Olsson, J., Bengtsson, E., Långström, B., & Bergström, M. (2009). Performance of Principal Component Analysis and Independent Component Analysis with Respect to Signal Extraction from Noisy Positron Emission Tomography Data—A Study on Computer Simulated Images. *The Open Neuroimaging Journal, 3*, 1-16. https://doi.org/10.2174/1874440000903010001

Rey, C., & Vanderschuren, H. (2017). Cassava Mosaic and Brown Streak Diseases: Current Perspectives and Beyond. *Annual Review of Virology, 4*, 429-452. https://doi.org/10.1146/annurev-virology-101416-041913

Sangpueak, R., Duchanee, S., Saengchan, C., Papathoti, N. K., Hoang, N. H., Thanh, T. L., Phansak, P., Buensanteai, N., Sangpueak, R., Duchanee, S., Saengchan, C., Papathoti, N. K., Hoang, N. H., Thanh, T. L., Phansak, P., & Buensanteai, N. (2023). Identification of cassava black stem and root rot agents in Thailand. *Chilean Journal of Agricultural Research, 83*(1), 70-82. https://doi.org/10.4067/S0718-58392023000100070

Schubert, E. (2023). Stop using the elbow criterion for k-means and how to choose the number of clusters instead. *ACM SIGKDD Explorations Newsletter, 25*(1), 36-42. https://doi.org/10.1145/3606274.3606278

Shlens, J. (2014). A Tutorial on Principal Component Analysis. *arXiv,* arXiv:1404.1100 https://doi.org/10.48550/arXiv.1404.1100

Song, F., Guo, Z., & Mei, D. (2010a). Feature Selection Using Principal Component Analysis. *Engineering Design and Manufacturing Informatization 2010 International Conference on System Science, 1*, 27-30. https://doi.org/10.1109/ICSEM.2010.14

Sedano, J. C. S., Moreno, R. E. M., Mathew, B., Léon, J., Gómez Cano, F. A., Ballvora, A., & López Carrascal, C. E. (2017). Major Novel QTL for Resistance to Cassava Bacterial Blight Identified through a Multi-Environmental Analysis. *Frontiers in Plant Science, 8*, 1169. https://doi.org/10.3389/fpls.2017.01169

Soto-Quiros, P., & Torokhti, A. (2021). Extended Principal Component Analysis. *arXiv,* arXiv:2111.03040. https://doi.org/10.48550/arXiv.2111.03040

Statheropoulos, M., Pappa, A., Karamertzanis, P., & Meuzelaar, H. L. C. (1999). Noise reduction of fast, repetitive GC/MS measurements using principal component analysis (PCA). *Analytica Chimica Acta, 401*(1), 35-43. https://doi.org/10.1016/S0003-2670(99)00494-8

Suzen, N., Gorban, A., Levesley, J., & Mirkes, E. (2020). Principal Components of the Meaning. *arXiv,* arXiv:2009.08859. https://doi.org/10.48550/arXiv.2009.08859

Szyniszewska, A. M. (2020). CassavaMap, a fine-resolution disaggregation of cassava production and harvested area in Africa in 2014. *Scientific Data, 7*(1), 159. https://doi.org/10.1038/s41597-020-0501-z

Tang, T. M., & Allen, G. I. (2021). Integrated Principal Components Analysis. *arXiv,* arXiv:1810.00832. https://doi.org/10.48550/arXiv.1810.00832

Taramuel-Taramuel, J. P., Montoya-Restrepo, I. A., & Barrios, D. (2023). Drivers linking farmers' decision-making with farm performance: A systematic review and future research agenda. *Heliyon, 9*(10), e20820. https://doi.org/10.1016/j.heliyon.2023.e20820

Teixeira, J. H. dos S., Guimarães, M. A. S., Cardoso, S. C., Brito, A. dos S., Diniz, R. P., de Oliveira, E. J., & de Oliveira, S. A. S. (2021). Evaluation of resistance to bacterial blight in Brazilian cassava germoplasm and disease-yield relationships. *Tropical Plant Pathology, 46*, 324-335. https://doi.org/10.1007/s40858-021-00419-3

Thepbandit, W., Papathoti, N. K., Hoang, N. H., Siriwong, S., Sangpueak, R., Saengchan, C., Laemchiab, K., Kiddeejing, D., Tonpho, K., & Buensanteai, K. (2024b). Bio-synthesis and characterization of silver nanoparticles from *Trichoderma* species against cassava root rot disease. *Scientific Reports, 14*(1), 12535. https://doi.org/10.1038/s41598-024-60903-z

Tize, I., Fotso, A. K., Nukenine, E. N., Masso, C., Ngome, F. A., Suh, C., Lendzemo, V. W., Nchoutnji, I., Manga, G., Parkes, E., Kulakow, P., Kouebou, C., Fiaboe, K. K. M., & Hanna, R. (2021). New cassava germplasm for food and nutritional security in Central Africa. *Scientific Reports, 11*(1), 7394. https://doi.org/10.1038/s41598-021-86958-w

Toure, H. M. A. C., Ehui, K. J. N., Abo, K., & Kone, D. (2020). Four years assessment of Cassava Bacterial Blight expression according to weather conditions in Côte d'Ivoire. *SN Applied Sciences, 2*(7), 1301. https://doi.org/10.1007/s42452-020-3135-z

Uke, A., Tokunaga, H., Utsumi, Y., Vu, N. A., Nhan, P. T., Srean, P., Hy, N. H., Ham, L. H., Lopez-Lavalle, L. A. B., Ishitani, M., Hung, N., Tuan, L N., Van Hong, N., Huy, N. Q., Hoat, T. X., Takasu, K., Seki, M., & Ugaki, M. (2022b). Cassava mosaic disease and its management in Southeast Asia. *Plant Molecular Biology, 109*(3), 301-311. https://doi.org/10.1007/s11103-021-01168-2

Van den Berg, H., & Jiggins, J. (2007). Investing in Farmers—The Impacts of Farmer Field Schools in Relation to Integrated Pest Management. *World Development, 35*(4), 663-686. https://doi.org/10.1016/j.worlddev.2006.05.004

van Elst, H. (2021). *Tutorial on principal component analysis, with applications in R*. https://doi.org/10.13140/RG.2.2.20075.16168/2

Vardakas, G., & Likas, A. (2023). Global k-means$++$: An effective relaxation of the global k-means clustering algorithm. *arXiv,* arXiv:2211.12271. https://doi.org/10.48550/arXiv.2211.12271

Veley, K. M., Elliott, K., Jensen, G., Zhong, Z., Feng, S., Yoder, M., Gilbert, K. B., Berry, J. C., Lin, Z.-J. D., Ghoshal, B., Gallego-Bartolomé, J., Norton, J., Motomura-Wages, S., Carrington, J. C., Jacobsen, S. E., & Bart, R. S. (2023). Improving cassava bacterial blight resistance by editing the epigenome. *Nature Communications, 14*(1), 85. https://doi.org/10.1038/s41467-022-35675-7

Wang, C., Chen, Y., Chen, S., Min, Y., Tang, Y., Ma, X., Li, H., Li, J., & Liu, Z. (2023a). Spraying chitosan on cassava roots reduces postharvest deterioration by promoting wound healing and inducing disease resistance. *Carbohydrate Polymers, 318*, 121133. https://doi.org/10.1016/j.carbpol.2023.121133

Wang, G., Lou, M., & Pananjady, A. (2023b). Do algorithms and barriers for sparse principal component analysis extend to other structured settings? *arXiv,* arXiv:2307.13535. https://doi.org/10.48550/arXiv.2307.13535

Wang, S. G. W., Patilea, V., & Klutchnikoff, N. (2024). Adaptive functional principal components analysis. *arXiv,* arXiv:2306.16091. https://doi.org/10.48550/arXiv.2306.16091

Wydra, K., & Verdier, V. (2002). Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics. *Agriculture, Ecosystems & Environment, 93*(1-3), 211-226. https://doi.org/10.1016/S0167-8809(01)00349-8

Yfantis, V., Wagner, A., & Ruskowski, M. (2023). Federated K-Means Clustering via Dual Decomposition-based Distributed Optimization. *arXiv,* arXiv:2307.13267. https://doi.org/10.48550/arXiv.2307.13267

Yoodee, S., Kobayashi, Y., Songnuan, W., Boonchird, C., Thitamadee, S., Kobayashi, I., & Narangajavana, J. (2018). Phytohormone priming elevates the accumulation of defense-related gene transcripts and enhances bacterial blight disease resistance in cassava. *Plant Physiology and Biochemistry, 122*, 65-77. https://doi.org/10.1016/j.plaphy.2017.11.016

Zárate-Chaves, C. A., Gómez de la Cruz, D., Verdier, V., López, C. E., Bernal, A., & Szurek, B. (2021). Cassava diseases caused by *Xanthomonas phaseoli* pv. *Manihotis* and *Xanthomonas cassavae*. *Molecular Plant Pathology, 22*(12), 1520-1537. https://doi.org/10.1111/mpp.13094

Zhang, J. (2019). Machine Learning With Feature Selection Using Principal Component Analysis for Malware Detection: A Case Study. *arXiv,* arXiv:1902.03639. https://doi.org/10.48550/arXiv.1902.03639

Zhu, B., Bedeer, E., Nguyen, H. H., Barton, R., & Henry, J. (2021). Improved Soft-k-Means Clustering Algorithm for Balancing Energy Consumption in Wireless Sensor Networks. *IEEE Internet of Things Journal, 8*(6), 4868-4881. https://doi.org/10.1109/JIOT.2020.3031272

Zhuang, Y., Chen, X., Yang, Y., & Zhang, R. Y. (2024). Statistically Optimal K-means Clustering via Nonnegative Low-rank Semidefinite Programming. *arXiv,* arXiv:2305.18436. https://doi.org/10.48550/arXiv.2305.18436