BAREKENG: Journal of Mathematics and Its Applications
June 2025 Volume 19 Issue 2 Page 1279-1294
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss2pp1279-1294

GRAPHICAL REPRESENTATION AND TWO GROUPS ANALYSIS ON DATA MATRIX OF ROBUSTA GREEN CHERRIES PRODUCTION IN TWO HARVEST PERIODS

Irmeilyana[1*], Bambang Suprihatin[2], Anita Desiani[3], Ngudiantoro[4], Sri Indra Maiyanti[5]

[1,2,3,4,5]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Sriwijaya*
Jln. Lintas Timur Palembang-Prabumulih KM. 32, Indralaya, 30129, Indonesia

Corresponding author's e-mail: * **irmeilyana@unsri.ac.id**

ABSTRACT

Article History:

Received: 11th October 2024
Revised: 3rd February 2025
Accepted: 5th March 2025
Published: 1st April 2025

Keywords:

Graphical Representation;
Green Cherry Production;
PCA;
Robusta Coffee;
Two Groups Analysis.

Several factors that play a role in the productivity of Robusta coffee trees are the influence of pruning techniques and weather elements. This paper discussed the graphical analysis and comparison of two data matrices of Robusta green cherries production, which would enter the ripening process in branch categories for the harvest period in 2023 and 2024. Hypothesis testing on secondary data in the form of daily weather conditions in 2022 and 2023, which include temperature, dew, humidity, wind speed, and cloud cover for the two periods, was significantly different. However, solar radiation and precipitation were not. The data source for each harvest period was primary data, with the object being a sample of 30 trees that were sampled purposively. The research object was in Pagaralam Municipality, South Sumatra. There were 18 variables covering many branch categories based on production year, position, and shape. The PCA (Principal Component Analysis) results on each data matrix show similarities in the dominant variables representing each subspace. The first three PCs in each data matrix for 2023 and 2024 span a subspace and describe the variation of the original data of 77.3% and 68.8%, respectively. The 3rd and 1st-year production branch categories dominate the subspace of each data matrix for 2023 and 2024. Comparison of the two PC subspaces using two groups analysis in 3rd dimension space produces angles of 19.70, 28.80, and 69.10. The bisector components show that the variables that dominate the similarity of the two data matrices are the variables that tend to represent both PC subspaces dominantly. Robusta green cherry production can be represented by the number of secondary branches, which are straight in shape, along with the number of fruit clusters. This study result can be a reference for farmers when considering the composition of the number of branch categories when pruning.

How to cite this article:

Irmeilyana, B. Suprihatin, A. Desiani, Ngudiantoro and S. I. Maiyanti., "GRAPHICAL REPRESENTATION AND TWO GROUPS ANALYSIS ON DATA MATRIX OF ROBUSTA GREEN CHERRIES PRODUCTION IN TWO HARVEST PERIODS," *BAREKENG: J. Math. & App.*, vol. 19, iss. 2, pp. 1279-1294, June, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

The productivity of coffee plants is very dependent on the genotype, cultivation method, plant care and land processing, weather [1]-[3], the presence of shade trees [4], and other environmental factors. Pruning is one method of tree care [5]-[7], to increase crop production [8]-[9], especially for smallholder farmers whose cost of maintaining crop is low [10]. The results of a review of 148 kinds of literature show that there is still a lack of critical analytical functions on the precise relationship between the potential risks of climate change on coffee farming systems and the environment [11]. In facing climate change, it is necessary to adapt coffee agronomic practices [12], improving agricultural practices significantly, so that coffee harvest become less sensitive to weather [13]. The impact of climate change has a complex effect on robusta coffee productivity and shifts in the timing of Robusta coffee cultivation activities in Polewali Mandar Regency [14] It has caused a decline in coffee productivity in Tana Toraja [15]. Climate elements, i.e., temperature, solar radiation, and humidity, are related to Robusta coffee production in Lampung [16]. However, the coffee production referred to in these articles was the production of green or dry coffee beans.

As Indonesia's highest Robusta coffee-producing province, South Sumatra in 2023 also decreased by 3.78% compared to the previous year, from 770.19 kg/ha to 741.11 kg/ha [17]. South Sumatra contributes to national coffee production by 27.32% [18]. Pagaralam is one of South Sumatra's coffee producers with a Geographical Indication. In 2023, coffee productivity in Pagaralam decreased by 1.1%, from 908 kg/ha to 789 kg/ha. Based on previous literature, this decrease in production can be influenced by weather elements and pruning techniques. Rejuvenation of Robusta coffee trees is generally carried out by grafting and tree maintenance by pruning branches. These processes can affect the performance of the tree, which, of course, can be related to the number of fruits on the tree branches and external factors such as weather conditions, attacks by fungi, and coffee berry borer pests.

Based on interviews in [19], the majority of farmers stated that weather was a factor that caused fruit to fall. They are less aware that too many twigs (branches), improper pruning, pruning that is not done on unproductive branches, and inadequate plant maintenance factors can affect coffee fruit production. So, in this research, we will examine the presence of branches on coffee trees in the production of coffee cherries. This research focuses on the production of green cherries in Pagaralam robusta coffee. In this study, we assumed that the coffee cherry fruits resulted from grafted cuttings from the exact clone, and the sample trees came from a uniform tree population.

The data matrix consists of row entries as variable values for an object and column elements as variable values for each object. Data matrices can be represented in graphical form by reducing the dimensions of variable space, including the biplot of PCA (Principal Component Analysis) results and the dendrogram of cluster analysis results. PCA is an interdependent technique in multivariate analysis that can reduce data dimensions and simplify structure by maintaining as much variation as possible in the data set and eliminating original variables with a relatively small information contribution [20].

Biplot exploration of the first 2 PC PCA results includes relationships between variables, similarities between objects, and relative relationships between objects and variables. PC is a linear combination of the original variables so that the coefficients of the linear combination form a basis vector. The first few PCs resulting from reducing a data matrix using PCA are basis vectors that span a subspace. In [21], a comparison of 2 or more data matrices (known as groups) can be done by comparing the subspaces of the PCA results. [22] Two data matrices representing the characteristics of Pagaralam coffee farmers were compared based on the use of reductant herbicide by using two groups analysis and cluster analysis. The dominant variables that characterize the similarities between the two categories of farmers based on the results of the two groups analysis are variables in the same cluster. On the other hand, the dominant variables that characterize the dissimilarity of the two farmer categories are the variables from separate clusters. In another case, PCA was also developed to calculate a composite indicator of deprivation in 121 municipalities in the province of Rome by considering spatial heterogeneity [23]. PCA was also used to examine ground-based VLF wave intensity variations [24], and analyze the correlations between the presence of a multitude of migratory water bird communities and water quality metrics in Transylvania [25]. So, data visualization from PCA results is beneficial in understanding and interpreting the data matrix.

Several Arabica coffee harvest prediction modeling studies use objects from cherry production through image data [26]-[29] but without considering weather factors. In the Robusta cherry harvest period from smallholder plantations, because the production branches are different and can be classified into several categories, the object is multivariate data, so the data matrix of cherry production in each harvest period can be very diverse. Therefore, it is necessary to reduce data space and graphical exploration. In this research, the objects were tree samples, focusing on the quantitative production of green cherries in 2 harvest periods, namely the 2023 and 2024 harvest periods. This research aims to represent the data matrix graphically based on the results of variable reduction with PCA and then apply two groups analysis to analyze the similarities and dissimilarities of data matrices for 2023 and 2024. Indicators of weather elements in these 2 years were also considered as one factor related to differences in cherry production in the two harvest periods. The output of this research is that it can be seen whether the results of applying a group analysis can align with the results of mean differences hypothesis testing between the two sample groups on the 2 data matrices. Furthermore, the results of this analysis can be used as initial information on whether there is a relationship between weather changes and branch categories on the production of green cherries on Robusta coffee trees. This study result can be a reference for farmers when considering the composition of the number of branch categories that must be maintained during pruning so that cherry production can be optimal.

2. RESEARCH METHODS

Robusta coffee cherry production in Pagaralam is generally done yearly, with a 1 to 4-month harvest period. After the harvesting process, farmers prune the tree branches. The number of types of branches that are less or unproductive can also affect the optimal production of cherries. The dense branches can also create micro-weather, making it easier for fungi and coffee berry borers to attack. In this case, a group analysis was used to study whether there were differences in tree performance systems regarding cherry fruit production in branch categories for the two harvest periods. Furthermore, the study also investigated the relationship between differences in tree performance and differences in weather elements a year before each harvest period.

This research was a case study of green cherry production on Robusta coffee trees in Kelurahan Candi Jaya, Pagaralam Municipality, South Sumatra, during two harvest periods. The primary data source for each harvest period was data, with the object being a sample of trees on one farmer's land with relatively the same plant maintenance culture. The farmer's coffee plantation consists of coffee trees that are relatively the same age, with the spacing between the trees being relatively the same. So, the assumption used was that the samples come from two independent populations. Tree sample selection was carried out purposively, referring to recommendations from **[30]**. Depending on the variance, the recommended number of sample trees is 5 to 10 in a close group. Because we used 30 sample trees in this study, we could use a normal distribution to estimate population parameters.

Pruning of Robusta coffee trees cultivated by grafting is divided into 3, namely pruning of shape, production, and rejuvenation **[30]**, so that existing branches can be categorized based on shape, year of production, and position. There were 18 variables used, including (i) the number of branch categories based on the position of the grafting on the **rootstock** plant, namely primary, secondary, and tertiary branches; (ii) the number of secondary branch categories based on production year, namely production years 0 to 3; (iii) the number of shape branches categories, namely straight, fan, and broken, and (iv) the number of clusters and total green cherries on a tree and also in each category of branch position. The notation of these variables is SPrim, SSec, STer for (i); NBP1, NBP2, NBP3, and NBP0 for (ii); Straight, Fan, and Broken for (iii); TF, TC, TF Prim, TC Prim, TC Sec, TF Sec, TC Ter, and TF Ter for (iv). We calculated the values for these variables in May 2023 and June 2024.

The initial step in this study was defining data matrix as $X = (x_{ij}); i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$, where n is the number of objects and p is the number of variables. In this research, there were two types of data matrices, namely data matrices whose objects were tree samples in 2 harvest periods in 2023 and 2024. Daily weather elements data from June 15, 2022, to June 16, 2023, and June 15, 2023, to June 16, 2024, were also considered in these two harvest periods. The weather data source is www.visualcrossing.com/weather/. The data on daily weather elements analyzed included 17 variables with numerical values and several weather conditions. These variables include temperature, air humidity, precipitation, wind conditions, and solar radiation. Each weather element data was compared for 1 year before each harvest period using the mean difference hypothesis test. Meanwhile, weather condition data for each period was described using a pie chart.

The following steps taken in the data matrix of tree samples are as follows:

a. Description of variable data in each data matrix of tree samples.

b. Carrying out mean difference hypothesis tests on variables.

c. Graphing several boxplots of variables in both data matrices.

d. Performing PCA on each data matrix [20], [31].

   i. Analyzing the first 3 PCs (Principal Components) from the PCA results on each data matrix.

   ii. Interpret the variables that dominate each 3-dimensional subspace.

   iii. Graphing PCA results in the form of a 2-dimensional biplot.

e. Performing two-groups analysis [21].

   i. Define the matrices formed from the first k PCs resulting from PCA of each data matrix, namely matrices $L^T_{(p\times k)} = (l_{ij})$ and $M^T_{(p\times k)} = (m_{ij})$, where l_{ij} and m_{ij} are coefficients from the first k PCs which are the eigenvectors of each correlation matrix of the data matrix. In this study, the comparison dimensions $k = 1, 2, 3$ are taken.

   ii. Determine the matrix $N_{(k\times k)} = LM^TML^T$.

   iii. Determine the eigenvalues λ_i of the matrix N and corresponding eigenvectors $\mathbf{a}_i$.

   iv. Determine the angles $\cos^{-1}\sqrt{\lambda_i}$ that formed between pairs of PCs from the two subspaces.

   v. Determine bisector $\boldsymbol{c}$ with the equation:

$$\boldsymbol{c}_{i\ (p\times 1)} = \{2(1+\sqrt{\lambda_i})\}^{-\frac{1}{2}}\left(I + \frac{1}{\sqrt{\lambda_i}}M^TM\right)\boldsymbol{b}_i;\ i = 1,\ \ldots,\ k \tag{1}$$

   where $\boldsymbol{b_i} = L^T\,\boldsymbol{a_i}$

f. Carrying out Steps 4 and 5 on two data matrices, with variables including the total number of fruits, the number of branches, and the fruits based on the position of the branches.

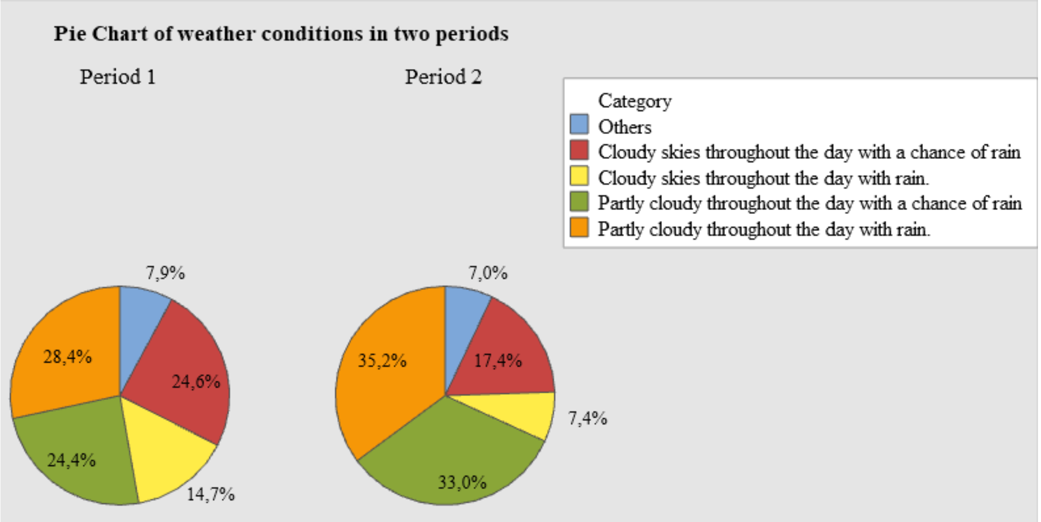
g. Interpretation of results.

h. Relating differences in weather elements before the harvest period with differences in subspace representation of the data matrix of tree samples in each harvest period.

3. RESULTS AND DISCUSSION

3.1 Description of Weather Data

Daily weather data is sourced from www.visualcrossing.com/weather/ for 2 periods, namely period 1 from June 15, 2022, to June 16, 2023, and period 2 from June 15, 2023, to June 16, 2024. Weather description data in the two periods can be seen in **Figure 1**. The pie diagram shows that the weather conditions in the two periods are quite different. The difference in the percentage of weather conditions in periods 1 and 2 is quite large, around 6% to 8%, for example, Cloudy skies throughout the day with a chance of rain of 24.6% and 17.4%. In general, rain-partially cloudy and rain-overcast conditions in period 1 were 52.8% and 39.3%, respectively, while in period 2, they were 68.2% and 24.8%, respectively.



Data source: www.visualcrossing.com/weather/

Figure 1. Weather Description in 2 Time Periods

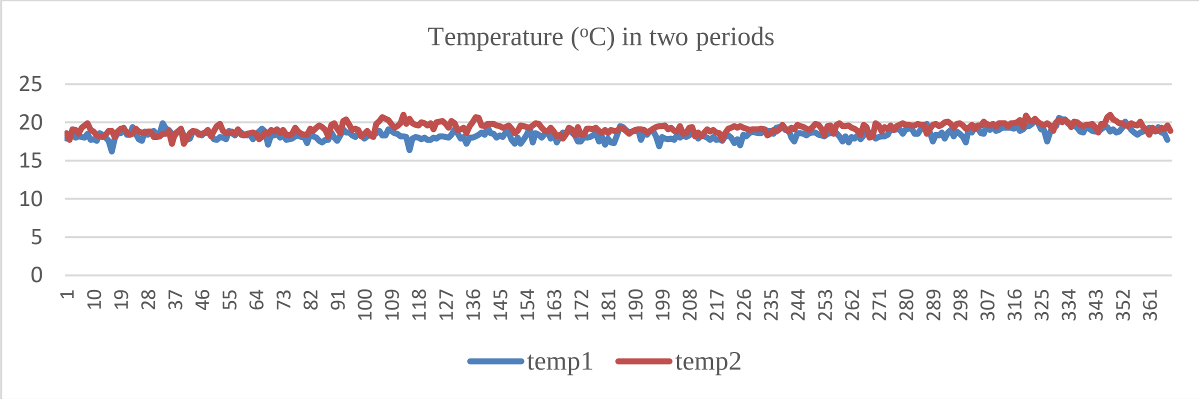
The weather elements considered in this research are numerical data from 18 variables, which included temperature, air humidity, precipitation, wind conditions, and solar radiation. Based on the mean difference test for each weather element, 10 variables were significantly different (at the significance level α = 5%) between the two periods, namely temperature (i.e., Temp max, Temp min, and Temp), Dew, Humidity, wind conditions (Wind gust, Wind speed, and Wind direction), Cloud cover, and Sea level pressure. However, solar radiation and precipitation were not significantly different. Some test results can be seen in **Table 1**. A negative Z value in the H_0 rejection test results indicates that the mean of a weather element in period 1 is significantly lower than the weather element in period 2. Vice versa for positive Z values. The difference in weather elements in these two time periods was one of the factors that influenced differences in cherry production in the following harvest period.

Table 1. Mean Difference Test for Some Weather Conditions in Two Periods

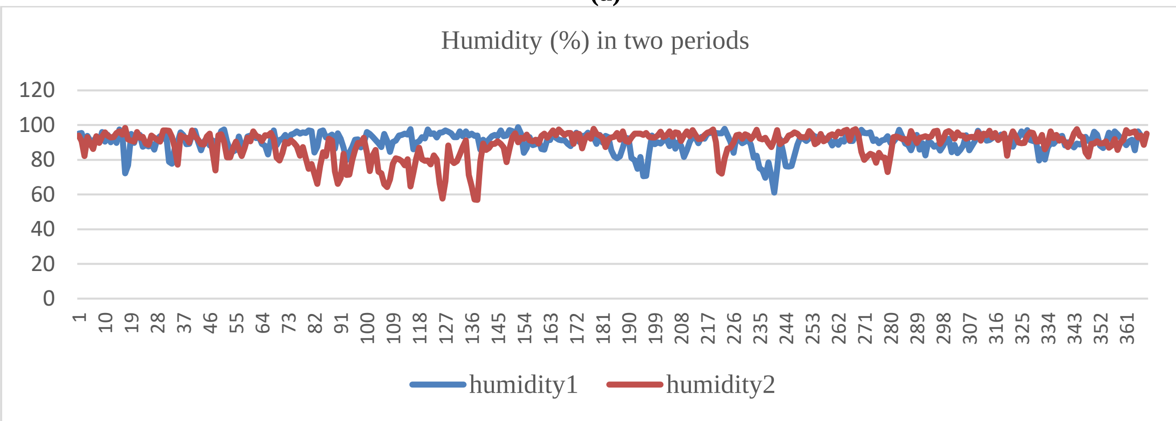
No.	Variable	Period	Mean	StDev	Z_{count}	*p-value*	Test Result	Explanation
1	Temperature	1	18.468	0.644	-15.99	0.000	Reject H_0	Statistically
		2	19.224	0.637				significant
2	Dew	1	90.43	5.35	2.54	0.011	Reject H_0	Statistically
		2	89.18	7.68				significant
3	Humidity	1	90.43	5.35	2.54	0.011	Reject H_0	Statistically
		2	89.18	7.68				significant
4	Solar radiation	1	192.8	60.4	-0.91	0.362	Accept H_0	Statistically
		2	196.7	56.7				not significant
5	Precipitation	1	10.3	11.3	-0.21	0.836	Accept H_0	Statistically
		2	10.5	11.1				not significant

***Note**: There were 367 days and 368 days of data in period 1 and 368 days in period 2. The critical Z for* $\frac{\alpha}{2} = 5\%$ *is 1.65;* $\frac{\alpha}{2} = 2.5\%$ *is 1.96. The two-tailed hypothesis test on* H_0 *states that the mean of the two populations is the same. The two populations are assumed to be independent with the Z test statistic* $Z_{count} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$.

Meanwhile, **Figure 2** describes the daily fluctuations in temperature and humidity conditions in both periods. Almost all daily temperatures for 1 year in period 1 tend to be higher in period 2. In several time intervals, the humidity in period 1 was higher than in period 2, for example, from day 66 to day 154 or around August to November 2022. In period 2, the humidity was lower, with very high fluctuations from August until November 2023. The opposite happens on days 182 to 212 (or around December 2022 to January 2023) and days 232 to 248 (or around February 2023). The humidity in period 1 was lower than in those months for period 2.



(a)



(b)

Figure 2. Daily Temperature and Humidity Conditions in Both Periods (a) Temperature in Two Periods, (b) Humidity in Two Periods

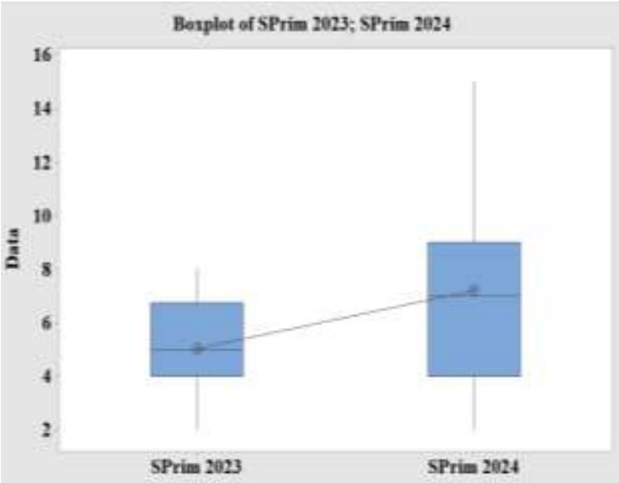
3.2. Description of Tree Sample Data

Mean difference tests were also carried out on branch category variables along with the total and number of fruits at each branch position. The test results for several variables can be seen in Table 2. Based on the mean difference test with a significance level of α = 5%, it is found that only SPrim, NBP0, and Broken had a $p-value$ < 0.05. The $t-value$ and $p-value$ on SPrim are respectively -2.83 and 0.007; NBP0 are -2.44 and 0.021; Broken are -3.61 and 0.001. The results of this test can state that the samples taken are insufficient to provide evidence that these variables have values that are not significantly different in the two periods. The average number of primary branches (SPrim variable) and the number of branches that are not yet in production (reserve branches) and 4th year of production branches (NBP0 variable) in the sample for the 2024 harvest period is greater than the 2023 harvest period. On the other hand, the number of fruits on tertiary branches (TF Ter. variable) on the sample of the 2024 harvest period is smaller than the 2023 harvest period.

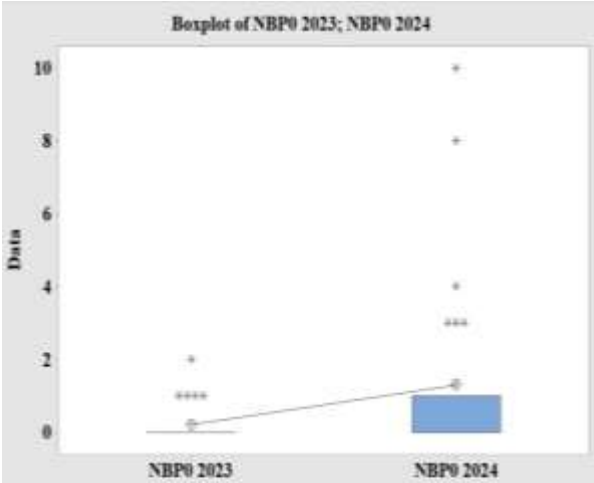
Table 2. Mean Difference Test for Some Variables in Two Harvest Periods

No.	Variable	Period	Mean	StDev	t_{count}	*p-value*	Test Result	Explanation
1	SPrim	2023	5.04	1.73	-2.83	0.007	Reject H_0	Statistically significant
		2024	7.20	3.78				
2	SSec	2023	14.71	6.37	-0.53	0.598	Accept H_0	Statistically not significant
		2024	15.73	8.20				
3	NBP0	2023	0.214	0.499	-2.44	0.021	Reject H_0	Statistically significant
		2024	1.30	2.38				
4	Broken	2023	0.71	1.18	-3.61	0.001	Reject H_0	Statistically significant
		2024	2.10	1.71				
5	TF Ter.	2023	172	223	1.70	0.097	Accept H_0	Statistically not significant
		2024	87	147				

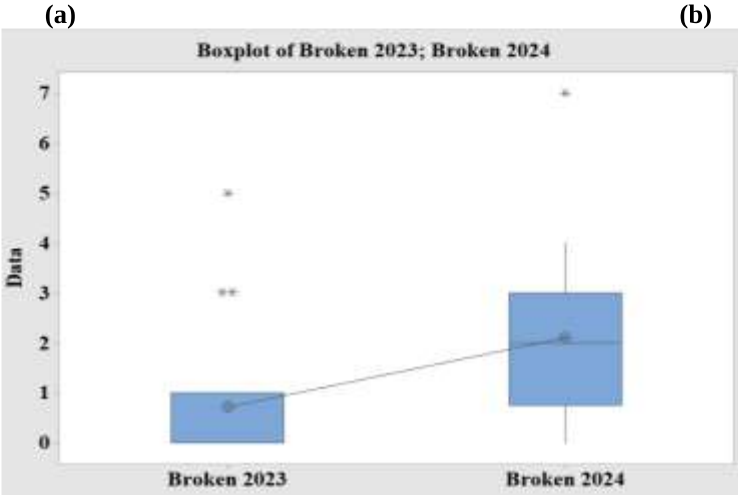
A comparison of the mean and distribution of significantly different variable values can be seen in the boxplot of Figure 3. Boxplot of the SPrim and NBP0 variables of the 2024 sample has more diverse values, where the mean is higher than the median. The opposite is true for the Broken variable. The mean difference test on other variables failed to reject the null hypothesis. In this case, the samples taken are insufficient to prove that these variables have different values in the two harvest periods. Next, the mean values of the variables that were significantly different in the two harvest periods were further analyzed to see whether they also dominated the dissimilarities between the two data matrices of tree samples.



(a)



(b)



(c)

Figure 3. Boxplot of Variables in Samples from Both Harvest Periods
(a) Boxplot of SPrim, (b) Boxplot of NBP0, (c) Boxplot of Broken

3.3. PCA of Both Data Matrices

Primary data of tree samples in the 2023 and 2024 harvest periods were arranged in each $n \times 18$ data matrix. PCA is carried out using a correlation matrix. The subspaces being compared can be represented in 2-dimensional and 3-dimensional graphs so the first 3 PCs are used. The first three PCs on the 2023 data matrix can explain 77.3% (= 37.2% + 24.4% + 15.6%) of the total variance. Meanwhile, the first three PCs on the 2024 data matrix can explain 68.8% (= 34.4%+24%+10.5%) of the total variance. The coefficients of each of the first three PCs from the PCA results on the data matrix for the 2023 and 2024 harvest periods can be seen in Table 3.

Table 3. Coefficients of PCs in PCA Results of Each Data Matrix

Variable	2023 Data Matrix			2024 Data Matrix		
	PC1	**PC2**	**PC3**	**PC1**	**PC2**	**PC3**
SPrim	0.109	0.124	**-0.425**	0.240	0.160	-0.290
SSec	**0.363**	0.092	0.103	**0.375**	-0.104	0.130
STer	0.168	**-0.370**	-0.219	-0.081	**-0.442**	0.017
NBP1	0.262	0.258	-0.069	**0.345**	0.004	0.085
NBP2	0.193	-0.080	0.233	0.285	-0.153	0.066
NBP3	0.070	**-0.306**	0.073	0.019	-0.119	0.006
NBP0	0.205	-0.030	0.037	-0.083	-0.016	0.105

Variable	2023 Data Matrix			2024 Data Matrix		
	PC1	PC2	PC3	PC1	PC2	PC3
Straight	**0.322**	0.220	0.122	**0.374**	-0.003	0.147
Fan	0.109	**-0.359**	-0.214	-0.068	**-0.441**	-0.046
Broken	-0.132	0.158	**-0.351**	-0.119	-0.138	0.121
TF	**0.356**	0.004	-0.140	0.192	-0.201	**-0.542**
TC	**0.371**	-0.033	-0.055	0.276	0.025	**0.395**
TF Prim.	0.061	**0.302**	**-0.386**	0.109	0.109	**-0.584**
TC Prim.	0.101	**0.302**	**-0.376**	0.210	0.105	0.187
TF Sec.	**0.329**	0.083	0.199	**0.336**	-0.186	-0.105
TC Sec.	**0.340**	0.059	0.226	**0.360**	-0.161	-0.005
TF Ter.	0.134	**-0.377**	-0.233	-0.083	**-0.440**	0.020
TC Ter.	0.150	**-0.366**	-0.231	-0.102	**-0.446**	0.022

Note: *Numbers in bold have higher values.*

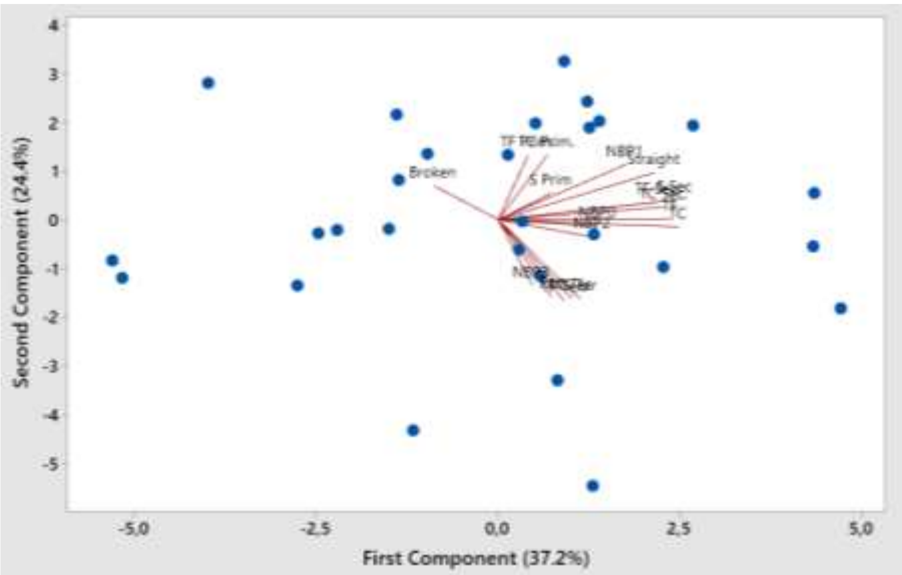
Suppose that the space resulting from reduction by PCA yields a new variable subspace or PC, which in this case can be called a group. Suppose a group (or subspace) is denoted as a matrix $L^T_{(18\times k)}$ whose columns consist of the first k PCs resulting from the reduction of the 2023 data matrix in **Table 3**. Likewise, for a group which is denoted as a matrix $M^T_{(18\times k)}$ whose columns are the first k PCs resulting from the 2024 data matrix reduction. The variables that dominate the PCs subspace can be summarized as in **Table 4**.

Table 4. Variables That Dominate First 3 PCs in Each Data Matrix

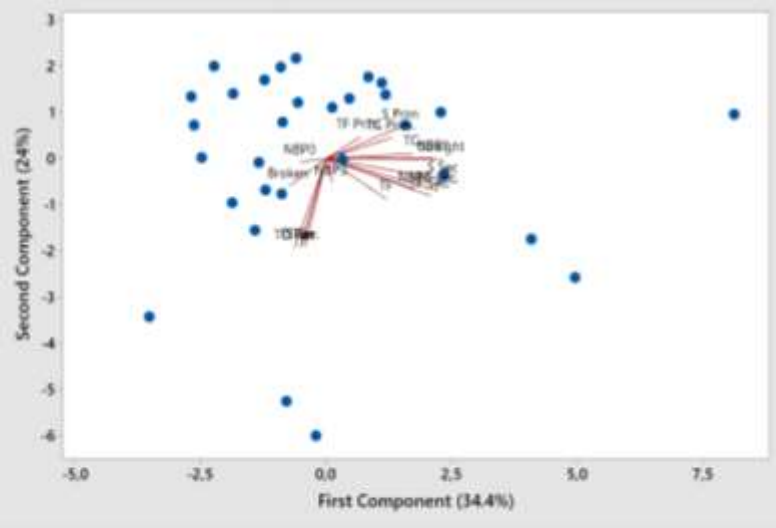
PC	2023 Data matrix	2024 Data matrix
PC1	**SSec, Straight,** TF, TC, **TF Sec, TC Sec**	**SSec,** NBP1, **Straight, TF Sec, TC Sec**
PC2	**STer**, NBP3, **Fan,** TF Prim, TC Prim, **TF Ter, TC Ter**	**STer, Fan, TF Ter, TC Ter**
PC3	SPrim, Broken, **TF Prim**, TC Prim	TF, TC, **TF Prim**

Note: *The variable printed in bold explains that the variable dominates the same PC in both data matrices. The red text states that the variable is in two PCs. The blue text states that the variable dominates the different PCs of the two data matrices.*

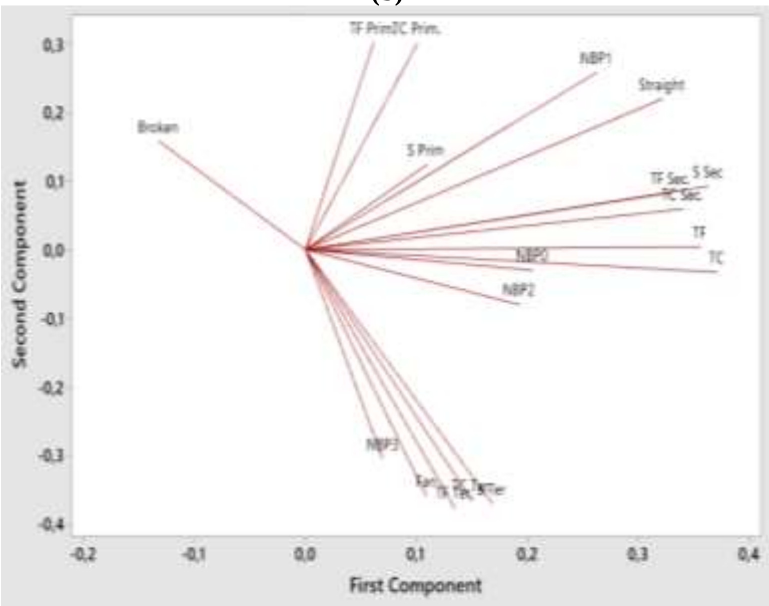
Based on **Table 4**, the PCA results in both data matrices are equally dominated by the variables Number of secondary branches along with their Total clusters and Total fruit, and also their straight shape (namely the variables Straight, SSec, TF Sec, and TC Sec); the Number of tertiary branches along with the Total clusters and Total fruits, and the Fan shape of the secondary branches (namely the variables Fan, STer, TF Ter, and TC Ter); and Total fruit on primary branches (i.e. variable TF Prim). In the 2023 data matrix, almost all variables dominate the subspaces spanned by 3 PCs. If we look at the production year of the secondary branch, the 2023 matrix subspace is dominated by the 3rd year production branch category. Meanwhile, the 2024 data matrix is dominated by the 1st year production branch category. PCA results can be represented graphically in a biplot as in **Figure 4**. Measure of goodness of fit on the two biplots are respectively 61.6% and 58.4%.



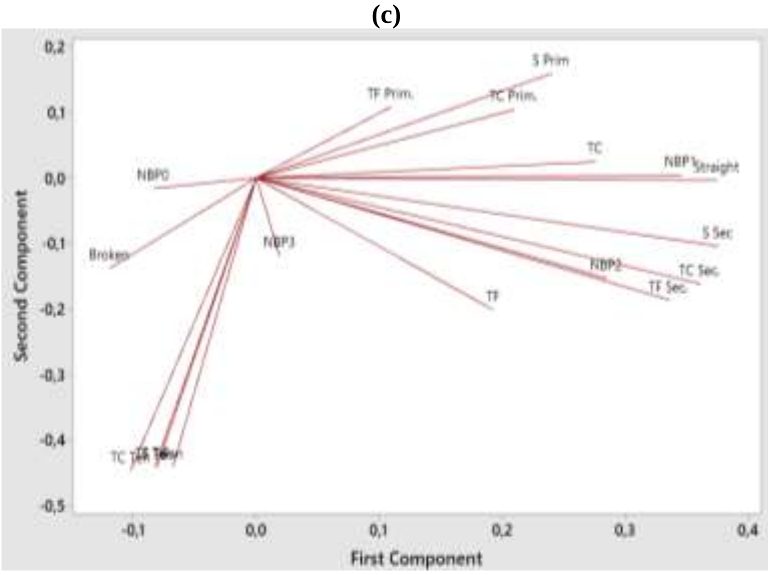
(a)



(b)



(c)



(d)

Figure 4. Graphical Representation of the Subspace of Each Data Matrix
(a) Biplot of the 2023 Data Matrix, (b) Biplot of the 2024 Data Matrix, (c) Loading Plot of the 2023 Data Matrix, (d) Loading Plot of the 2024 Data Matrix

The grouping of variables tends to be different. In the biplot of the 2023 data matrix, PC1 is more characterized by variables TF and TC (i.e., Total number of fruits and a Total number of clusters), while PC2 is more characterized by TF Prim and TC Prim (i.e., TF and TC on primary branches). Variables that are strongly correlated with each other are TF Prim and TC Prim; STer, TF Ter and TC Ter; SSec and TF; TF and TC. Some objects have high variable values.

In the 2024 data matrix biplot, the objects are more clustered, and several objects are potential outliers. The variables that are moderately and strongly correlated with each other are TF Ter, TC Ter and Fan; TF, TF Sec and TC Sec; NBP1 and Straight; SSec, TF Sec and TC Sec. The majority of objects have low variable values.

3.4. Two Groups Analysis

Two groups analysis is used to analyze the sources of variability between two subspaces of PCs, so that the similarities and dissimilarities of each data matrix can be analyzed. The results of the two groups analysis are bisectors and the angles formed between the bisector subspace and each subspace of PCs on the two data matrices. Determining bisectors uses the principle of orthogonal projection of a vector, which is an eigenvector of matrix N. Bisectors is determined by using **Equation (1)**.

In this study, because of the first 3 PCs that were analyzed, there were also three bisectors on a subspace. The results of the two groups' analysis of the comparison of 1-dimensional, 2-dimensional, and 3-dimensional subspaces can be seen in **Table 5**.

In each comparison dimension, the variables that dominate bisector 1 and bisector 2 are the same. Bisector $\mathbf{c_1}$ is dominated by the variables SSec, NBP1, Straight, TC, TF Sec, TC Sec. Bisector $\mathbf{c_2}$ is dominated by the variables STer, Fan, TF Ter, and TC Ter. The angles formed between bisector $\mathbf{c_1}$ and PC1 in 1-dimensional, 2-dimensional, and 3-dimensional comparisons are relatively small, that are 36.5^0, 20.2^0 and 19.7^0 respectively. The angles formed between bisector $\mathbf{c_2}$ and PC2 in 2-dimensional and 3-dimensional comparisons are also relatively small 32.4^0 and 28.8^0, respectively. The variables that dominate bisectors $\mathbf{c_1}$ and $\mathbf{c_2}$ tend to represent the similarity between the two data matrices.

Table 5. Results of Two Groups Analysis on Comparison of The Two Subspaces of PCs

Variable	Bisector Coefficients in Comparison of the Two Subspaces (Two Groups)					
	1 PC	2 PCs		3 PCs		
	$\mathbf{c_1}$	$\mathbf{c_1}$	$\mathbf{c_2}$	$\mathbf{c_1}$	$\mathbf{c_2}$	$\mathbf{c_3}$
S Prim	0.184	0.164	0.163	0.161	-0.092	**-0.445**
S Sec	**0.388**	**0.388**	0.002	**0.389**	0.042	0.124
S Ter	0.046	0.059	**-0.440**	-0.006	**0.462**	-0.075
NBP1	**0.319**	**0.330**	0.148	**0.340**	-0.088	-0.028
NBP2	0.251	0.245	-0.101	0.242	0.105	0.188
NBP3	0.047	0.022	-0.220	-0.001	0.206	0.091
NBP0	0.064	0.059	-0.062	0.052	0.066	0.089
Straight	**0.366**	**0.369**	0.121	**0.386**	-0.079	0.129
Fan	0.022	0.037	**-0.425**	-0.026	**0.442**	-0.112
Broken	-0.132	-0.084	0.011	-0.101	0.024	-0.161
TF	0.288	0.298	-0.117	0.273	0.161	**-0.422**
TC	**0.340**	**0.310**	-0.012	**0.302**	0.067	0.214
TF Prim.	0.090	0.106	0.214	0.111	-0.158	**-0.631**
TC Prim.	0.163	0.176	0.221	0.181	-0.143	-0.152
TF Sec.	**0.349**	**0.363**	-0.046	**0.363**	0.069	0.037
TC Sec.	**0.368**	**0.374**	-0.043	**0.376**	0.066	0.118
TF Ter.	0.027	0.040	**-0.439**	-0.025	**0.459**	-0.080
TC Ter.	0.025	0.041	**-0.441**	-0.025	**0.462**	-0.080
Angle between PCs	36.5	20.2	32.4	19.7	28.8	69.1

***Note**: The numbers printed in bold are the coefficients of the variables that dominate the bisector.*

In the 3-dimensional comparison, bisector $\mathbf{c_3}$ is dominated by the variables SPrim, TF, and TF Prim. Only TF Prim dominates PC3 in each data matrix. TF dominates PC1 in the 2023 data matrix, but it dominates PC3 in the 2024 data matrix. Meanwhile, the variable SPrim only dominates PC3 in the 2023 data matrix. Because the angle formed between bisector $\mathbf{c_3}$ and PC3 in each subspace of PCs is 69.1^0, then it can be said that these three variables tend to represent the dissimilarity of the two data matrices.

In the two harvest periods, the branch categories that were dominantly related to the total number of green cherries were the number of secondary branches, clusters, and fruit on these secondary branches. The number of primary branches and clusters and fruit on the primary branches were dominantly related to differences in green cherry fruit production in the two harvest time periods. This can be attributed to differences in the conditions of several different weather elements each year before observing green cherry production in the two harvest periods.

3.5. Using Two Groups Analysis on the Data Matrix of Green Cherry Production in the Branch Position Category

Previously, based on the results of the mean difference test, SPrim was a different variable in the two data matrices. Below, a two-groups analysis is carried out to compare the relationship between green cherry production based on the number of branch position categories. The data matrix variables consist of the total number of fruit (or denoted by TF), the total number of fruit clusters (or TC), the number of branch position categories, namely SPrim, SSec, STer, and the number of fruits on each branch. The data matrix for each harvest period is $n \times 8$.

Next, PCA was applied to each data matrix for the 2023 and 2024 harvest periods. The first three PCs of PCA on each data matrix produced a goodness of fit of 90.6% (= 49.3% + 23.7% + 17.5%) and and 86.1% (= 37.6% + 28.6% + 20.0%). The coefficients of each of the first three PCs from the PCA results can be seen in **Table 6**.

Table 6. Coefficients of PCs in PCA Results of Each Data Matrix

Variable	2023 Data Matrix			2024 Data Matrix		
	PC1	PC2	PC3	PC1	PC2	PC3
SPrim	0.179	-0.231	**0.630**	**0.421**	0.242	-0.184
SSec	**0.433**	-0.203	-0.254	**0.476**	-0.124	0.370
STer	0.278	**0.545**	0.209	-0.072	**-0.618**	-0.073
TF	**0.487**	-0.067	0.042	**0.402**	-0.251	**-0.436**
TC	**0.488**	0.016	-0.076	0.306	0.046	**0.573**
TF Prim.	0.114	**-0.467**	**0.516**	0.320	0.182	**-0.534**
TF Sec.	0.398	-0.210	**-0.404**	**0.477**	-0.227	0.132
TF Ter.	0.239	**0.585**	0.241	-0.079	**-0.627**	-0.058

Note: *Numbers in bold have higher values.*

Suppose that the space resulting from reduction by PCA yields a new variable subspace or PC, which in this case can be called a group. Suppose two groups (or subspaces) which are denoted as matrix $L^T_{(8\times k)}$ and matrix $M^T_{(8\times k)}$ whose columns are respectively the first k PCs resulting from the 2023 and 2024 data matrices reduction in **Table 6**. The variables that dominate the PCs subspace can be summarized as in **Table 7**.

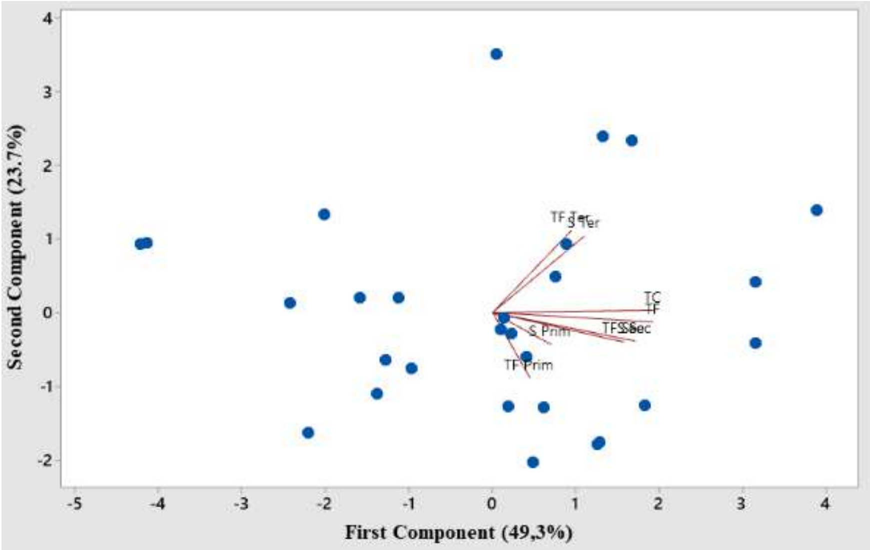
Table 7. Variables That Dominate First 3 PCs in Each Data Matrix

PC	2023 Data matrix	2024 Data matrix
PC1	**SSec**, **TF**, TC	SPrim, **SSec**, **TF**, TF Sec
PC2	**STer**, TF Prim, **TF Ter**	**STer**, **TF Ter**
PC3	SPrim, **TF Prim**, TF Ter	TF, TC, **TF Prim**

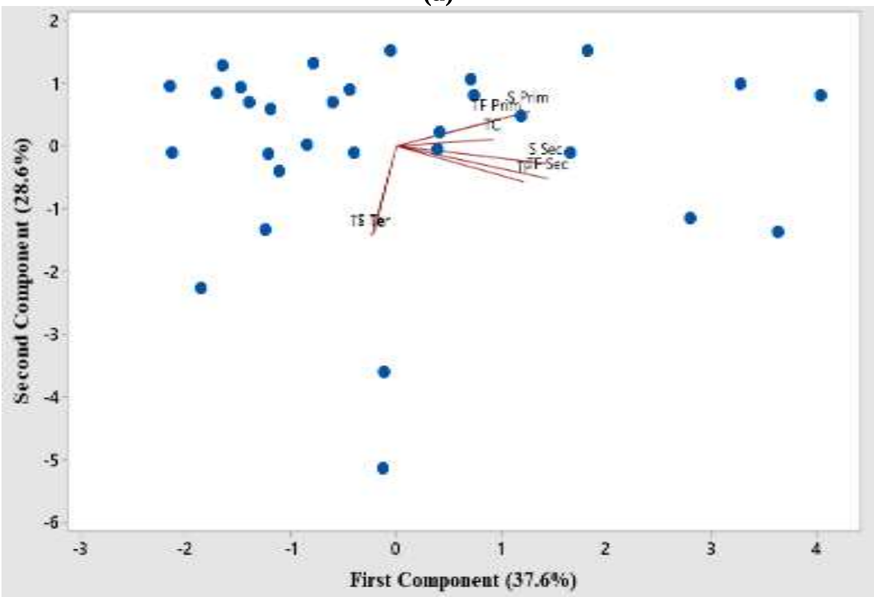
Note: *The variable printed in bold explains that the variable dominates the same PC in both data matrices. The red text states that the variable is in two PCs. The blue text states that the variable dominates the different PCs of the two data matrices. The green text states that the variable only dominates in a PC of one data matrix.*

Based on **Table 7**, all variables except TF Sec dominate the 3-dimensional subspace of both data matrices. TF Sec only dominates the subspace of the 2024 data matrix. In both data matrices, TF and SSec dominate PC1, STer and TF Ter dominate PC2, and TF Prim dominates PC3. In the 2023 data matrix, TC is one of the variables that dominates PC1, but in the 2024 data matrix, TC dominates PC3. On the other hand, SPrim dominates PC3 in the 2023 data matrix, but SPrim dominates PC1 in the 2024 data matrix. In addition, TF Prim in the 2023 data matrix dominates PC2 and PC3. Meanwhile, TF in the 2024 data matrix dominates PC1 and PC3.

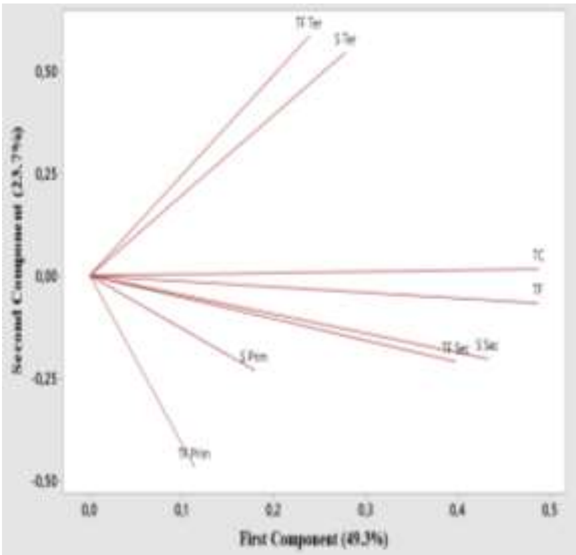
The grouping of variables tends to be different. In biplot of the 2023 data matrix, PC1 is more characterized by variables TF, TC, and SSec (i.e., total number of fruits, total number of clusters, and the number of secondary branches), while PC2 is more characterized by TF Ter, STer, and TF Prim (i.e., TF on tertiary branches, The number of tertiary branches, and TF on primary branches). Variables strongly correlated with each other are TC and TF, TF Sec and SSec, and TF Ter and STer. Some objects have high variable values. The biplot of **Figure 5a** cannot represent 27% of the information.

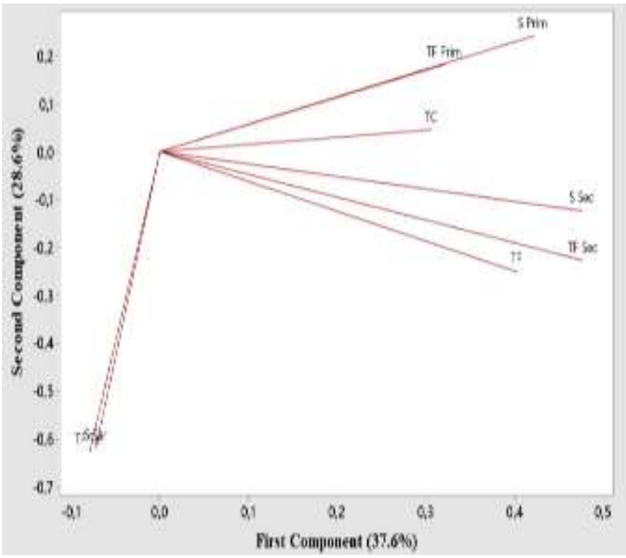


(a)



(b)





(c) **(d)**

Figure 5. Graphical Representation of the Subspace of Each Data Matrix
(a) Biplot of the 2023 Data Matrix, (b) Biplot of the 2024 Data Matrix, (c) Loading Plot of the 2023 Data Matrix, (d) Loading Plot of the 2024 Data Matrix

In the 2024 data matrix biplot, the objects are more clustered, and several objects are potential outliers. The biplot of **Figure 5b** cannot represent 33.9% of the information. PC1 is more characterized by variables SPrim, SSec, and TF Sec (i.e., the number of the primary, secondary, and TF on secondary branches). At the same time, PC2 is more characterized by STer and TF Ter (i.e., The number of tertiary branches and TF on tertiary branches). The variables that have a strong correlation are TF Sec and SSec, TF Ter, and STer. The objects in the second quadrant have low variable values.

Table 8. Results of Two Groups Analysis on Comparison of The Two Subspaces of PCs

Variable	Bisector Coefficients in Comparison of the Two Subspaces (Two Groups)					
	1 PC	2 PCs		3 PCs		
	$\mathbf{c_1}$	$\mathbf{c_1}$	$\mathbf{c_2}$	$\mathbf{c_1}$	$\mathbf{c_2}$	$\mathbf{c_3}$
S Prim	0.315	0.359	-0.152	**-0.463**	-0.126	0.335
S Sec	**0.477**	**0.470**	0.126	-0.337	0.390	-0.298
S Ter	0.108	-0.105	**0.623**	0.261	**0.446**	**0.422**
TF	**0.467**	**0.411**	0.261	-0.343	0.280	0.328
TC	**0.417**	0.346	0.141	-0.217	0.376	-0.303
TF Prim	0.228	0.352	-0.241	**-0.503**	-0.259	**0.418**
TF Sec	**0.459**	**0.456**	0.165	-0.317	0.392	-0.234
TF Ter	0.084	-0.136	**0.632**	0.291	**0.440**	**0.438**
Angle between PCs	35.6	10.5	25.3	8.4	15.0	54.7

Note: *The numbers printed in bold are the coefficients of the variables that dominate the bisector.*

Based on **Table 8**, when comparing 1-dimensional and 2-dimensional subspaces, the variables that dominate bisector $\mathbf{c_1}$ include SSec, TF, and TF Sec. In comparing 2-dimensional and 3-dimensional subspaces, STer and TF Ter are the variables that both dominate bisector $\mathbf{c_2}$. In the 3-dimensional subspace comparison, these two variables and TF Prim are dominant.

The angles formed in the 3-dimensional subspace are relatively small, namely 8.4^0 and 15^0. The variables that dominate the bisectors of both data matrices are SPrim, TF Prim, S Ter, and TF Ter. The angle formed in bisector $\mathbf{c_3}$ tends to be large, namely 54.7^0, with variables that also dominate bisectors $\mathbf{c_1}$ and $\mathbf{c_2}$, namely, STer, TF Prim, and TF Ter. Because the 3-dimensional subspace represents the higher variation of data, the variables SPrim, TF Prim, S Ter, and TF Ter are the dominant variables representing the similarity between the two data matrices.

The variables SPrim, TF Prim, S Ter, and TF Ter are the dominant variables representing the similarity of two data matrices with 8 variables (namely: total fruit, number of branches based on branch position category, and number of fruits on each branch). This result differs from the dominant variables, representing the similarity of the data matrix with 18 variables, including the number of branches based on shape category and year of production and the number of fruit clusters in each branch position category. The dominant variables represent the dissimilarities between the two data matrices. In this case, comparing similarities and dissimilarities between two data matrices depends on the type and number of variables being analyzed.

In comparing two data matrices with 18 variables, the variables significantly different from the mean difference test results do not dominate the bisector in the 2-dimensional subspace. Because the angle formed between the 3rd PC in a 3-dimensional subspace comparison is relatively large (69.10), SPrim is dominant in determining the dissimilarity between the two data matrices. Apart from the number of primary branches, the number of fruits on the primary branches and the total number of fruits also differ between the two data matrices. One factor influencing this dissimilarity is the difference in weather elements during the year before the two harvest periods.

4. CONCLUSIONS

Daily weather elements in two periods showed differences, including temperature (temp, temp min, and temp max), dew, humidity, wind conditions (wind gust, wind direction, wind speed), cloud cover, and sea level pressure. In the two harvest periods, 2023 and 2024, the mean difference test showed that the mean values of the variables significantly different in the tree sample data matrix were SPrim, NBP0, and Broken. Meanwhile, the two groups' analysis results showed that only SPrim was dominant in representing the dissimilarity of the two data matrices. One factor influencing this dissimilarity is the difference in weather elements during the year before the two harvest periods.

Graphical representation by biplots and the results of two groups analysis show that similar variables dominate the subspaces of the first 2 PCs in both data matrices for the 2023 and 2024 harvest periods. The variables of the number of secondary branches, clusters, and fruits on these secondary branches tend to represent the similarity between the two data matrices. Meanwhile, the number of primary branches, green cherry fruits on primary branches, and the total number of green cherry fruits dominate the dissimilarity of the two data matrices. On the other hand, in comparing two data matrices of green cherry production in the branch position category, the number of primary branches and the number of fruits on primary branches determine the similarity of the two subspaces of both data matrices. In this case, comparing similarities and dissimilarities between two data matrices depends on the type and number of variables being analyzed. The research results show the role of branch category composition on the production of Robusta coffee cherries and the relationship between the tree performance and external factors such as weather conditions.

Differences in the conditions of several weather elements before the time of observation are one of the factors related to differences in green cherry production in the two harvest periods. For further research, it can be studied how the influence of daily weather elements is one of the factors related to the production of Robusta red cherries. Future research can be more devoted to analyzing the relationship between weather conditions from when the cherries are green until they become red cherries ready to be harvested. The analysis can be expanded by considering the influence of crop maintenance culture on two or more farmers' fields.

ACKNOWLEDGMENT

The research of this article was funded by the DIPA of the Public Service Agency of Universitas Sriwijaya 2024, SP DIPA-023.17.2.677515/2024, on November 24, 2023. By the Rector's Decree Number: 0013/UN9/LP2M.PT/2024, on May 20, 2024. The authors would also like to thank Mas Ismail and Mas Wahid as partner farmers and PT. Pandawa Agri Indonesia (PAI), especially the South Sumatra team, Mr. Ahmad Surkati, Rici Wijaya, Kania, Riduwan, and Nando, assisted in obtaining data in the field.

REFERENCES

[1] J. Kath, V. Mittahalli, Byrareddy, S. Mushtaq, A. Craparo, and P. Mario, "TEMPERATURE AND RAINFALL IMPACTS ON ROBUSTA COFFEE BEAN CHARACTERISTICS," *Clim. Risk Manag.*, vol. 32, no. 100281, pp. 1–15, January 2021.

[2] T. R. Moreira, S. F. da Silva, N. B. da Silva, G. M. A. D. A. dos Santos, and A. R. dos Santos, "GLOBAL WARMING AND THE EFFECTS OF CLIMATE CHANGE ON COFFEE PRODUCTION," *Food Engineering Series: Quality Determinants In Coffee Production*, 1st ed. Switzerland: Springer Cham, 2021, 65–100.

[3] E. Rahn, P. Vaastd, P. Läderachb, P. van Astenc, L. Jassognec, and J. Ghazoul, "EXPLORING ADAPTATION STRATEGIES OF COFFEE PRODUCTION TO CLIMATE CHANGE USING A PROCESS-BASED MODEL," *Ecol. Modell.*, vol. 371, pp. 76–89, January 2018.

[4] N. Van Long, N. Q. Ngoc, N. N. Dung, P. Kristiansen, I. Yunusa, and C. Fyfe, "THE EFFECTS OF SHADE TREE TYPES ON LIGHT VARIATION AND ROBUSTA COFFEE PRODUCTION IN VIETNAM," *Engineering*, vol. 7, pp. 742–753, November 2015.

[5] N. Khayati, A. Wachjar, and Sudarsono, "PENGELOLAAN PEMANGKASAN TANAMAN KOPI ARABIKA (COFFEA ARABICA L.) DI KEBUN KALISAT JAMPIT, PT PERKEBUNAN NUSANTARA XII (PERSERO), BONDOWOSO, JAWA TIMUR," *Bul. Agrohorti*, vol. 7, no. 3, pp. 295–301, May 2019.

[6] E. Mustari, "PERBEDAAN...MEMBUKA PELUANG UNTUK PENGEMBANGAN KOPI LEBIH BERMANFAAT," Webinar Pengurus Perkumpulan Enterprener Ganesha, STIH ITB, 2021.

[7] S. Tahmrin, J. Ashan, Muh Dzulkifly Junaedi, M. N. I. Ilham, and M. Maslam, "PENERAPAN TEKNOLOGI BUDIDAYA TANAMAN KOPI SECARA BERKELANJUTAN BAGI PETANI DI KABUPATEN GOWA," *J. Apl. Teknol. Rekayasa dan Inov.*, vol. 2, no. 1, pp. 34–41, April 2023.

[8] Z. Zulkarnain, Rahmaddiansyah, R. Alpian, and B. Bagio, "TINGKAT PRODUKTIVITAS DAN PENDAPATAN PETANI KOPI ARABIKA YANG MELAKUKAN TEKNIK PEMANGKASAN RUTIN DAN YANG TIDAK DI KECAMATAN BENER KELIPAH," *Agrifo*, vol. 5, no. 1, pp. 78–86, June 2020.

[9] R. Subantoro and M. A. Azis, "TEKNIK PEMANGKASAN TANAMAN KOPI (COFFEA sp)," *MEDIAGRO*, vol. 15, no. 1, pp. 52–65, November 2019.

[10] IDH (The Sustainable Trade Initiative), SCP (The Sustainable Coffee Program), and TechnoServe, "A BUSINESS CASE FOR SUSTAINABLE COFFEE PRODUCTION," *IDH*, 2014. [Online]. Available: http://exchange.growasia.org/indonesia-business-case-sustainable-coffee-production [Accessed on November 2, 2023].

[11] C. Bilen, D. El Chami, V. Mereu, A. Trabucco, S. Marras, and D. Spano, "A SYSTEMATIC REVIEW ON THE IMPACTS OF CLIMATE CHANGE ON COFFEE AGROSYSTEMS," *Plants*, vol. 12, no. 102, pp. 1–20, June 2023.

[12] S. G. Mbwambo, S. K. Mourice, and A. J. P. Tarimo, "THE IMPACTS OF CURRENT CLIMATE VARIABILITY ON COFFEE PRODUCTION IN THE NORTHERN AND SOUTHERN HIGHLANDS OF TANZANIA," *J. Agric. Sci.*, vol. 14, no. 3, pp. 78–99, February 2022.

[13] T. L. A. Dinh, F. Aires, and E. Rahn, "STATISTICAL ANALYSIS OF THE WEATHER IMPACT ON ROBUSTA COFFEE YIELD IN VIETNAM," *Front. Environ. Sci.*, vol. 10, no. 820916, pp. 1–15, June 2022.
[14] A. W. Angka and S. Dewi, "DAMPAK PERUBAHAN IKLIM TERHADAP PRODUKTIVITAS KOPI ROBUSTA DI DESA KURRAK KECAMATAN TAPANGO KABUPATEN POLEWALI MANDAR," *Media Agribisnis*, vol. 5, no. 2, pp. 133–139, November 2021.
[15] F. H. Assamha and R. Boer, "DAMPAK PERUBAHAN IKLIM TERHADAP PRODUKTIVITAS TANAMAN KOPI DI KABUPATEN TANA TORAJA," Thesis, IPB University, Bogor, Indonesia, 2017.
[16] D. P. Widiyani and J. S. S. Hartono, "STUDI EKSPLORASI AGROKLIMAT TANAMAN KOPI ROBUSTA (COFFEA CANEPHORA) KABUPATEN TANGGAMUS, LAMPUNG," *J. Agroteknologi dan Agribisnis*, vol. 5, no. 1, pp. 20–29, March 2021.
[17] BPS-Statistics Indonesia, *STATISTICAL YEARBOOK OF INDONESIA 2024*. Jakarta: BPS-Statistics Indonesia, 2024.
[18] Directorate of Food Crops, Horticulture, and Estate Crops Statistics, *INDONESIAN COFFEE STATISTICS 2023*. Jakarta: BPS-Statistics Indonesia, 2024.
[19] Irmeilyana, Ngudiantoro, S. I. Maiyanti, and S. Makhalli, "LOGISTIC REGRESSION MODELING OF THE USE OF REDUCTANT HERBICIDE IN PAGARALAM COFFEE FARMING," *Barekeng*, vol. 17, no. 4, pp. 1957–1968, December 2023.
[20] I. T. Jolliffe, *PRINCIPAL COMPONENT ANALYSIS*, 2nd ed. New York: Springer Verlag, 2002.
[21] W. J. Krzanowski, *PRINCIPLE OF MULTIVARIATE ANALYSIS, A USER'S PERSPECTIVE*. New York: Oxford University Press, 1990.
[22] Irmeilyana, Ngudiantoro, and S. I. Maiyanti, "APPLICATION OF TWO GROUPS ANALYSIS AND CLUSTER ANALYSIS ON COMPARISON OF CHARACTERISTICS OF PAGARALAM COFFEE FARMERS CATEGORIES," *Barekeng*, vol. 17, no. 1, pp. 101–112, March 2023.
[23] P. Postiglione, A. Artone, M. S. Andreano, and R. Benedetti, "CONSTRAINED OPTIMIZATION FOR ADDRESSING SPATIAL HETEROGENEITY IN PRINCIPAL COMPONENT ANALYSIS: AN APPLICATION TO COMPOSITE INDICATORS," *Stat. Methods Appt.*, vol. 32, pp. 1539–1561, May 2023.
[24] B. Bezděková, F. Němec, and J. Manninen, "GROUND-BASED VLF WAVE INTENSITY VARIATIONS INVESTIGATED BY THE PRINCIPAL COMPONENT ANALYSIS," *Earth, Planets Sp.*, vol. 74, no. 30, pp. 1–15, February 2022.
[25] P. Tóth, B. L. Tarcsay, Z. Kovács, D. T. Ionescu, S. Németh, and · Endre Domokos, "ASSESSMENT OF THE CORRELATION BETWEEN THE NUTRIENT LOAD FROM MIGRATORY BIRD EXCREMENT AND WATER QUALITY BY PRINCIPAL COMPONENT ANALYSIS IN A FRESHWATER HABITAT," *Environ. Sci. Pollut. Res.*, vol. 30, pp. 66033–66049, April 2023.
[26] P. J. Ramos, F. A. Prieto, E. C. Montoya, and C. E. Oliveros, "AUTOMATIC FRUIT COUNT ON COFFEE BRANCHES USING COMPUTER VISION," *Computers and Electronics in Agriculture*, vol. 137, pp. 9–22, May 2017.
[27] P. J. Ramos, J. Avendaño, and F. A. Prieto, "MEASUREMENT OF THE RIPENING RATE ON COFFEE BRANCHES BY USING 3D IMAGES IN OUTDOOR ENVIRONMENTS," *Computers in Industry*, vol. 99, pp. 83–95, August 2018.
[28] J. P. Rodríguez, D. C. Corrales, J. N. Aubertot, and J. C. Corrales, "A COMPUTER VISION SYSTEM FOR AUTOMATIC CHERRY BEANS DETECTION ON COFFEE TREES," *Pattern Recognition Letters*, vol. 136, pp. 142–153, August 2020.
[29] J. P. R. Uez, D. C. Corrales, D. Griol, Z. Callejas, Corrales, and J. Carlos, "A NON-DESTRUCTIVE TIME SERIES MODEL FOR THE ESTIMATION OF CHERRY COFFEE PRODUCTION," *Computers, Materials & Continua*, vol. 70, no. 3, pp. 4725–4743, January 2022.
[30] B. Haryanto, A. Thohar, H. Basri, D. Widodo, N. S. Wibowo, and J. Juniawan, KURIKULUM NASIONAL DAN MODUL PELATIHAN BUDIDAYA BERKELANJUTAN (GOOD AGRICULTURAL PRACTICES-GAP) DAN PASCAPANEN (POST-HARVEST) KOPI ROBUSTA. Jakarta: BPP SDM Pertanian, Kementerian Pertanian RI, 2019.
[31] N. Trendafilov and M. Gallo, "PRINCIPAL COMPONENT ANALYSIS (PCA)," MULTIVARIATE DATA ANALYSIS ON MATRIX MANIFOLDS (WITH MANOPT*)*. Switzerland: Springer Cham, 2021, 89-139.