ISSN 1807-1929
Revista Brasileira de Engenharia Agrícola e Ambiental
Brazilian Journal of Agricultural and Environmental Engineering
v.28, n.6, e275006, 2024
Campina Grande, PB – http://www.agriambi.com.br – http://www.scielo.br/rbeaa

DOI: http://dx.doi.org/10.1590/1807-1929/agriambi.v28n6e275006

ORIGINAL ARTICLE

Relationships between yellow and purple passion fruit variables[1]

Relações entre variáveis do maracujá amarelo e roxo

Beatriz G. Lopes[2*], Gabriela M. Rodrigues[2], Afrânio M. C. Vieira[3], Taciana V. Savian[2] & Glaucia A. Faria[4]

[1] Research developed at Universidade de São Paulo, Escola Superior de Agricultura "Luiz de Queiroz", Departamento de Ciências Exatas, Piracicaba, SP, Brazil
[2] Universidade de São Paulo/Escola Superior de Agricultura "Luiz de Queiroz"/Departamento de Ciências Exatas, Piracicaba, SP, Brazil
[3] Universidade Federal de São Carlos/Departamento de Estatística, São Carlos, SP, Brazil
[4] Universidade Estadual Paulista "Júlio de Mesquita Filho"/Faculdade de Engenharia de Ilha Solteira/Departamento de Matemática, Ilha Solteira, SP, Brazil

HIGHLIGHTS:

PCA efficiently identifies key variables, optimizing resource allocation and saving time and costs.
Young plants efficiently give information for the identification of promising variables for the assessment of passion fruit.
Reliable inferences require retain a small number of variables to explain the highest data variability.

ABSTRACT: Brazil is the world's largest producer of passion fruit, with the species *Passiflora edulis* Sims (yellow and purple passion fruit) being the most popular commercially. Due to the great economic importance of this culture, seeking high productivity, phytotechnical quality, and maintenance of existing germplasms, among others, it is necessary to carry out studies of the most diverse types of this culture. In this crop, the use of multivariate techniques has increasingly driven studies on genetic differences between passion fruit species. Therefore, this research aimed to identify the variables that deserve greater emphasis in experiments with two commercial species of *Passiflora: Passiflora edulis* Sims f. *edulis* (purple passion fruit) and *Passiflora edulis* Sims (yellow passion fruit), verifying if the species differ from each other. For this purpose, principal component analysis, multivariate analysis of variance (MANOVA), and Hotelling's T^2 test were performed. The principal component analysis proved to be effective, allowing the removal of five variables for yellow passion fruit and six variables for purple passion fruit and yellow passion fruit present contrasting differences between them, being the explanatory variables for the purple passion fruit the external and internal color and for the yellow passion fruit length, diameter, and citric acid.

Key words: *Passiflora edulis* Sims f. *edulis*, *Passiflora edulis* Sims, multivariate analysis of variance, Hotelling's T^2 test, principal component analysis

RESUMO: O Brasil é o maior produtor mundial de maracujá, sendo a espécie *Passiflora edulis* Sims (maracujá-amarelo e roxo), a de maior expressão comercial. Devido à grande importância econômica desta cultura, buscando alta produtividade, qualidade fitotécnica e manutenção dos germoplasmas existentes, entre outros, faz-se necessário estudos dos mais diversos tipos sobre esta cultura. Nesta cultura, a utilização de técnicas multivariadas tem impulsionado cada vez mais os estudos referentes a divergências genéticas entre espécies de maracujá. Diante disso, o objetivo do trabalho foi identificar as variáveis que merecem maior destaque em experimentos com duas espécies comerciais de *Passiflora*: *Passiflora edulis* Sims f. *edulis* (maracujá-roxo) e *Passiflora edulis* Sims (maracujá-amarelo), verificando se as espécies diferem entre si. Para tanto, foram realizadas a análise de componentes principais, análise multivariada da variância (MANOVA) e o teste T^2 de Hotelling. As espécies maracujá-roxo e maracujá-amarelo, apresentam diferenças contrastantes entre si; as variáveis explicativas para o maracujá-roxo são coloração externa e interna, e para maracujá-amarelo são comprimento e diâmetro dos frutos, acidez, Brix, espessura da casa e rendimento do suco. A análise de componentes principais mostrou-se eficaz, permitindo a retirada de cinco variáveis para o maracujá amarelo e seis variáveis para o maracujá roxo. *Passiflora edulis* Sims f. *edulis* (maracujá roxo) e *Passiflora edulis* Sims (maracujá amarelo) apresentam diferenças contrastantes entre si, sendo as variáveis explicativas para o maracujá roxo a cor externa e interna e para o maracujá-amarelo comprimento, diâmetro e ácido cítrico.

Palavras-chave: *Passiflora edulis* Sims f. *edulis*, *Passiflora edulis* Sims, análise multivariada da variância, T^2 de Hotelling, análise de componentes principais

• Ref. 275006 – Received 22 May, 2023
* Corresponding author - E-mail: beatrizgl@usp.br
• Accepted 01 Feb, 2024 • Published 20 Feb, 2024
Editors: Lauriane Almeida dos Anjos Soares & Walter Esfrain Pereira

Introduction

The genus *Passiflora* originates from South America, belonging to the Passifloraceae family, with about 525 species, most of which are identified in Brazil, turning the country into a center of genetic diversity of the *Passiflora* genus. All species are commonly called passion fruit, which in Tupi means "food in the form of a gourd"; Brazil is the world's largest producer of passion fruit, where the species *Passiflora edulis* (yellow and purple passion fruit) stands out, producing about 690 thousand tons year^{-1}, arousing economic interest in this culture (Faleiro et al., 2019).

Principal component analysis (PCA) is a multivariate statistical technique that examines several variables to condense a large number of dimensions, components, or latent factors from the original vast dimension of the data (Santos et al., 2019). Thereby the study of correlations between variables can be applied to almost all research areas. Thus, a simple correlation can provide false results, as in each set it does not capture information that may relate to a third variable or group of variables, which is of total importance for genetic improvement since it seeks to obtain gains with variables of interest through other variables that may be correlated (Lopes et al., 2022).

Using multivariate techniques has increasingly driven studies on genetic differences between passion fruit species (Ramos et al., 2019; Devi Ramaiya et al., 2013; Janzantti & Moneteiro, 2014). Among these techniques, the analysis with principal components can be highlighted (Vargas et al., 2020; Costa et al., 2023).

Therefore, this research aimed to identify the variables that deserve greater emphasis in experiments with two commercial species of *Passiflora: Passiflora edulis* Sims f. *edulis* (purple passion fruit) and *Passiflora edulis* Sims (yellow passion fruit), verifying if the species differ from each other.

Material and Methods

This research was performed at the University of São Paulo, Luiz de Queiroz College of Agriculture from March 2021 to July 2021. This study was carried out based on two experiments under identical cultivation conditions (period, climate, location, and cultural practices). The first one uses the species purple passion fruit and was performed as a blank test in a completely randomized design with 77 plants from 12 distinct families planted in vertical espalier at 2.0 × 5.0 m spacing.

In the second experiment, in a completely block design, the yellow passion fruit was used in three blank tests, with each trial constituting a block of 36 plants from different families for the uniformity tests in each block, with each plant belonging to a single family, totaling 108 plants, which were planted in a vertical espalier with a spacing of 2.0 × 5.0 m. Both experiments were performed in the field, at Embrapa Mandioca e Fruticultura, located in the city of Cruz das Almas, in the state of Bahia, Brazil, 12°40'0" South Latitude and 39°06'0" West Longitude from Greenwich. The altitude is 200 m above sea level, with an Aw to Am climate, characterized as hot and humid tropical, according to the Köppen classification.

The variables analyzed in the two experiments were: fruit length (FL, mm), and fruit diameter (FD, mm) were measured using a graduated trough; external color (EC), and internal color (IC) were measured using a color scale, peel thickness (PT, mm), the yield of juice (YJ, mL) was measured with a graduated cylinder, soluble solids (Brix, °Brix) was determined by refractometry, citric acid (AC, % citric acid) was determined by titration, number of fruits (NF) and the total weight of fruits (TW, g) were measured with a graduated caliper.

Initially, a descriptive analysis was performed to verify the means and variances of the species. Principal component analysis (PCA) was also performed as exploratory data analysis. The PCA technique has the principle of transforming a group of original variables into another group of variables with the same dimension, which is called principal components (Salem & Hussein, 2019; Gwelo, 2019). It follows that every component represents a linear combination of the observed variables, where the main objective is to briefly seek as much information as possible about the total variation comprised in the data set, i.e., it seeks to reduce the data so that little information can be missed in the process. It is important to emphasize that, through this technique, it is possible to redistribute the variation found between the original axes to acquire uncorrelated orthogonal axes (Bernardes & Mizusaki, 2022).

Let the variables be X_1, X_2, ..., X_p, that is, p response variables. In this work, p = 10 with n = 185 observations each (yellow passion fruit and purple passion fruit). Therefore, the data array X has dimension 185 × 10. p components are necessary to explain the total variability of the data. However, a large part of this variability can often be explained by a small number of principal components k (k < p). This way, these k components can replace the original p variables, reducing the dimensionality of the set (Johnson & Wichern, 2014).

Correlation matrix of the variables was used, so that the results were not influenced by the magnitude of their units, due to the set presents quantitative, discrete and continuous variables, as well as ordinal qualitative variables. The development of the analysis is based on the calculation of eigenvalues and eigenvectors of this matrix, where the eigenvalues represent the amount of variance of the matrix that is expressed by each component, and the eigenvectors are the loadings, that is, they are the principal components. The observed variables are represented in terms of these components.

Based on the PCA, a biplot was created for the first two principal components of the sample points and variables. This plot displays the results efficiently and allows analyzing the interrelationships between the characteristics evaluated: the size of the vector is proportional to the variance of the variables, and the cosine of the angle between two vectors is the correlation between the variables. Regarding the samples, close points may indicate an association between them (Gabriel, 1971).

Still, in the multivariate statistical theory, considerable attention has been devoted to tests to compare mean vectors, such as those by Erichson et al. (2020), Ó et al. (2020) and the most known T^2 test, proposed by Harold Hotelling in 1947 (Mason & Young, 2002). After the exploratory analyses, the equality of the mean vectors was tested by Hotelling's T^2 test. According to Johnson & Wichern (2014), with sufficiently large n_1 - p and n_2 - p, it is possible to perform tests and build confidence regions independently of the normality assumption, since the multivariate normality is not true in this set due to the magnitude differences between the variables. Considering the continuous quantitative variables, the method of Box & Cox (1964) and Yeo & Johnson (2000) suggests the transformation of the juice yield, brix, and citric acid variables. So, to test the hypothesis, Eq. 1:

$$H_0 : \sum - \sum_2 \neq 0 \quad (1)$$

Reject H_0 if Eq. 2:

$$T^2 = \left[\underline{x_1} - \underline{x_2}\right]' \left[\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right] \left[\underline{x_1} - \underline{x_2}\right] > \chi_p^2(\alpha) \quad (2)$$

where:

S_1 and S_2 are the covariance matrices of each sample;
n_1 = 108 (number of yellow passion fruit plants); and,
n_2 = 77 (number of purple passion fruit plants).

Based on the T^2 statistic, intervals are constructed for each component of $\mu' = [\mu_1, \mu_2, \ldots, \mu_p]$ individually, with a simultaneous confidence level, called Hotelling intervals (Johnson & Wichern, 2014). For the differences $\mu_1 - \mu_2$, withdrew a = [1, -1] and the simultaneous intervals with 100 (1 - α)% confidence are given by Eq. 3:

$$a'(\underline{x_1} - \underline{x_2}) \pm \sqrt{\chi_p^2(\alpha)} \sqrt{a' \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right) a} \quad (3)$$

if it contains 0 the difference is not significant.

Also based on the T^2 statistic, the most critical linear combination that leads to the rejection of H_0 is calculated, which has a vector of coefficients given by Eq. 4:

$$\hat{a} \propto \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right)^{-1} (\bar{x}_1 - \bar{x}_2) \quad (4)$$

Nonparametric multivariate techniques are becoming a promising alternative, such as multivariate analysis of variance (MANOVA) based on approximation of the F distribution (Arias-Nava et al., 2023) and permutation tests (Anderson, 2001). This methodology becomes very interesting and advantageous in many cases, as it has less restrictive assumptions, and is more robust to deviations from multi-normality, the presence of outliers, and heterogeneity of variances compared to parametric methods (Arias-Nava et al., 2023).

Therefore, it was performed a non-parametric test based on the approximation of the F distribution for the ANOVA-type statistic. This procedure allows a combination of quantitative, ordinal, or binary response variables. It can be considered as a modification of the theory test statistics, robust to normality, based on classifications, and therefore, it is also robust to outliers (Rubarth et al., 2021).

It was denoted by H and G the matrices of the mean squares due to treatment (sum of square mean hypothesis) and the mean squares due to error, respectively, both of the p × p dimension.

The distribution of the ANOVA tr(H)/tr(G) type statistics can be approximated for distribution with degrees of freedom of the numerator and denominator, respectively, estimated by Eqs. 5 and 6:

$$\hat{f} = (a-1)\frac{\operatorname{tr}(G_3)^2}{\operatorname{tr}(G_3^2)} \quad (5)$$

$$\hat{f}_0 = \frac{a^2}{(a-1)\sum_{i=1}^{a}\frac{1}{n_i - 1}}\hat{f} \quad (6)$$

Multivariate one-way analysis of variance was also performed, based on permutation tests. Proposed by Anderson (2001), the test statistic is analogous to Fisher's F ratio and is calculated from a dissimilarity matrix, and its p-values are obtained through permutation tests. One thousand permutations were used, considered sufficient according to Anderson (2001), for $p \leq 0.05$.

The analyzes were performed using the R software (R Core Team, 2020), with the principal component analysis performed using the princomp function and both non-parametric methods using the npmv package.

Results and Discussion

Table 1 presents the averages of each variable for each species and their respective standard deviation. Differences can be observed in the averages of the variables due to the species, mainly in the citric acid variable, in which the presence of citric acid was 30.44% of citric acid for the yellow passion fruit and 3.06% of citric acid for the purple passion fruit. The citric acid

Table 1. Averages and standard deviations of fruit length (FL, mm), fruit diameter (FD, mm), external color (EC), internal color (IC), peel thickness (PT, mm), juice yield (JY, mL), soluble solids (Brix, °Brix), citric acid (AC, %), number of fruits (NF), and total fruit weight (TW, g) for both species

Variables	Yellow		Purple	
	Average	SD	Average	SD
FL	83.20	4.23	78.02	5.35
FD	76.98	3.34	69.66	4.48
EC	1.93	0.26	7.70	1.93
IC	1.18	0.38	4.32	0.79
PT	7.37	1.17	6.05	1.44
JY	493.78	103.17	477.08	92.66
Brix	14.21	0.89	13.51	2.02
AC	30.44	3.10	3.06	0.47
NF	64.84	31.51	82.27	31.19
TW	9,610.66	4,629.18	10,840.19	4,138.77

SD - standard deviation

is a variable of great interest to the food products market since, according to Pinto et al. (2003), genotypes that present citric acid greater than 1% are the most important for the industry since it is not necessary to add citric acid to better conserve the pulp, resource widely used to prevent the development of microorganisms.

There is also a large difference between the external and internal color averages. For external color, the average was 1.93 for yellow passion fruit and 7.70 for purple passion fruit and for internal color, the average was 1.18 for yellow passion fruit and 4.32 for purple passion fruit. Regarding the external coloration, the purple passion fruit presents higher values compared to the yellow passion fruit, which is expected since the passion fruit has nine external color characteristics (yellowish-red, yellowish-purple, greenish-purple, light reddish-purple, reddish-purple, dark reddish-purple, light purple, purple and dark purple), while the yellow passion fruit has only two characteristics (light and dark yellow).

According to Ocampo et al. (2013), fruits that meet the standards of different categories of consumers regarding internal and external color are of high quality; These standards are taken into account to achieve the desired quality in the markets, allowing the identification of elite accessions with high quality fruits. The internal ones are related to flavor, that is, soluble solids, citric acid, and juice yield, and the external ones are related to good appearance, that is, skin color, size, and weight.

Also in Table 1, it is observed that the following variables with the highest average values among the analyzed species are: length of the fruit *P. edulis* (yellow) - 83.20 mm and *P. edulis* (purple) - 78.02 mm) and fruit diameter *P. edulis* (yellow) - 76.98 mm and *P. edulis* (purple) - 69.66 mm), which leads to the conclusion that, for these variables, the yellow and purple passion fruit are similars.

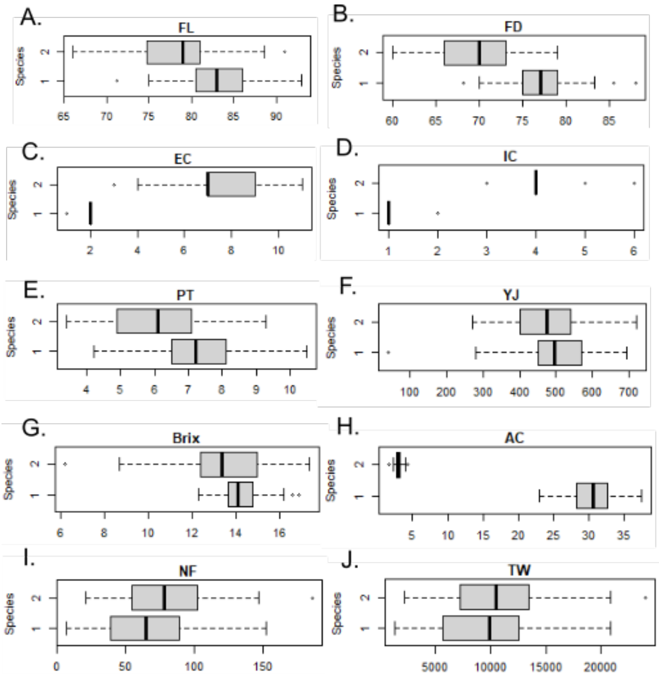
Concerning the number of fruits, the yellow passion fruit has the lowest value (NF 64.84) and 82.27 in the purple, even though yellow has the highest number of plants evaluated. Hence, yellow has the lowest mean value for total weight (9610.66 g) and 10840.19 g in the purple) (Table 1).

In addition, the average values for *P. edulis* juice yield (yellow) - 493.78 mL and *P. edulis* (purple) - 477.08 mL are observed, average values for *P. edulis* peel thickness (yellow) - 7.37 mm and *P. edulis* (purple) - 6.05; average values for Brix *P. edulis* (yellow) - 14.21 and *P. edulis* (purple) - 13.51°Brix; It can be concluded that, for these variables, the yellow passion fruit is slightly superior (Table 1).

Furthermore, the highest standard deviations are observed in juice yield and number of fruits for both species, which demonstrates that these variables are the ones with the highest variability in each species (Table 1).

According to the boxplots (Figure 1), the medians differ in some variables, especially in length (Figure 1A), diameter (Figure 1B), and peel thickness (Figure 1E), in general, they present few discrepant points and higher variations in the numbers of fruits (Figure 1I) and the total weight of fruits (Figure 1J), in agreement with the results from Table 1.

Regarding Table 2, the values that determine the correlation between the variables under study are highlighted, which are: the colors are correlated with each other and with the citric acid. In other words, it is possible to notice a difference between the species in terms of coloration, which occurs, since they are species with completely different colors; this also occurs with citric acid, where it is known that purple passion fruit is less acidic when compared to yellow passion fruit (Faleiro et al., 2019).



Species: 1 - Yellow and 2 - Purple

Figure 1. Boxplot of fruit length (FL, mm), fruit diameter (FD, mm), external color (EC), internal color (IC), peel thickness (PT, mm), juice yield (JY, mL), soluble solids (Brix, °Brix), citric acid (AC, %), number of fruits (NF), and total fruit weight (TW, g), for yellow and purple passion fruit

Table 2. Correlations between, fruit length (FL, mm), fruit diameter (FD, mm), external color (EC), internal color (IC), peel thickness (PT, mm), juice yield (JY, mL), soluble solids (Brix, °Brix), citric acid (AC, %), number of fruits (NF), and total fruit weight (TW, g) in: yellow and purple passion fruit

	FL	FD	EC	IC	PT	JY	Brix	AC	NF	TW
FL	1.00									
FD	**0.81**	1.00								
EC	-0.41	-0.59	1.00							
IC	-0.45	-0.65	**0.86**	1.00						
PT	0.15	0.33	-0.42	-0.45	1.00					
JY	0.42	0.38	-0.05	-0.06	-0.05	1.00				
Brix	0.05	0.12	-0.11	-0.19	0.10	0.27	1.00			
AC	0.47	0.68	**-0.90**	**-0.92**	0.44	0.13	0.25	1.00		
NF	-0.07	-0.13	0.24	0.21	-0.22	-0.10	-0.14	-0.25	1.00	
TW	0.06	0.02	0.13	0.09	-0.16	-0.03	-0.11	-0.13	**0.94**	1.00

It is also observed that the length and diameter of the fruits present a strong and positive correlation, which is expected, knowing that increasing the length of the fruit also increases its diameter. The number of fruits exhibits a strong and positive correlation with the total weight of the fruits, as expected, since the greater the number of fruits, the greater their total weight (Table 2).

Based on the correlation matrix between the species, the principal components were calculated, which provides the multiple integrations of information obtained, and thus, it is possible to identify which variables best describe the data set. From the PCA (Table 3), it turns out that the first two principal components explain 61% of the total variation. In the first component, the most important variables were the external and internal colors with a negative sign and the diameter and citric acid variables with a positive sign. In the second component, only the variables number of fruits per plant and total weight of fruits stand out positively.

Presented two components and their levels of explanation under the dataset, the researcher can decide whether the synthesis provided by this reduction in dimensionality is worthwhile or whether should consider all the variables under study. In other words, the analysis allows the removal of five variables for yellow passion fruit and six variables for purple passion fruit.

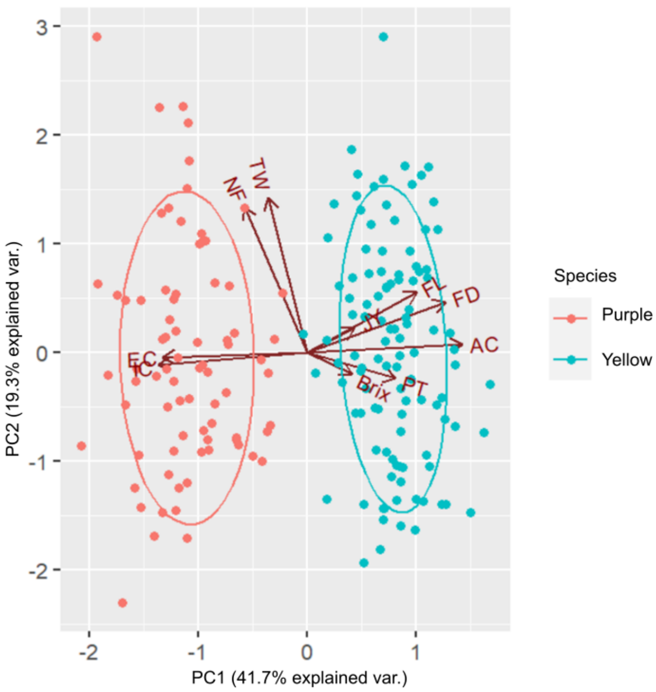
It follows that the distribution of variables, as well as the dispersion of variables concerning the first two principal components, are presented in Figure 2, in which the first component represents 41.7% and the second component represents 19.3% of the total variation. It is noteworthy that the first and second components, on axes one and two, respectively, show the highest variations of external and internal color on axis one and the number of fruits and weight of fruits on axis two. There is also a clear separation between the two passion fruit species (yellow and purple). Thus, the highest citric acid value appears in the yellow passion fruit, while the darkest colors are in the purple passion fruit, as described by Faleiro et al. (2019). Besides these, the most expressive variables for the yellow passion fruit were the length and diameter of the fruits.

The unit circle makes it possible to visually identify which variables are related to the cases under study. According to

Table 3. Result of the principal components analysis: eigenvectors (correlation between the variables and the components), percentage of explained variance, and accumulated percentage by the components

	Component 1	Component 2
FL	0.32 (0.66)	0.26 (0.36)
FD	0.41 (0.83)	0.22 (0.30)
EC	-0.43 (-0.87)	-0.02 (-0.03)
IC	-0.44 (0.90)	-0.05 (0.08)
PT	0.26 (0.53)	-0.11 (-0.15)
JY	0.14 (0.28)	0.11 (0.15)
Brix	0.13 (0.27)	-0.09 (-0.13)
AC	0.46 (0.93)	0.04 (0.05)
NF	-0.18 (-0.38)	0.63 (0.87)
TW	-0.12 (-0.24)	0.67 (0.93)
Variances (%)	41.7	19.3
Total variance (%)	41.7	61.0

FL - Fruit length (mm); FD - Fruit diameter (mm); EC - External color; IC - Internal color; PT- Peel thickness (mm); JY - Juice yield (mL); Brix - Soluble solids (°Brix); AC- Citric acid (%); NF - Number of fruits, and TW- Total fruit weight (g)



FL - Fruit length (mm); FD - Fruit diameter (mm); EC - External color; IC - Internal color; PT - Peel thickness (mm); JY- Juice yield (mL); Brix - Soluble solids (°Brix); AC - Citric acid (%); NF - Number of fruits; TW - Total fruit weight (g)

Figure 2. Biplot of variables of the species yellow and purple passion fruit

Figure 2, some variables are very close to each other such as external and internal color, number of fruits, fruit weight, and fruit length and fruit diameter, as previously noted; it shows that these have identical representations in the graph. Another important fact is that some variables are very close to the unit circle (center) which shows that they have a higher contribution than the variables that are further away.

The main characteristic evaluated for fresh trade is the quality of the fruits, which involves the external appearance of the fruits, where the loss of mass and subsequent wilting, which gives the fruit a wrinkled appearance, is one of the concerns highlighted by the production chain for its commercialization (Rinaldi et al., 2017); characteristics evaluated by the length, diameter, and weight of the fruit and also the peel thickness, which makes the evaluation of these variables of great importance for national and international trade, both for yellow and purple passion fruit; since, a fruit that presents these qualities, leads to greater consumer interest.

When marketing the fruit, the weight of the fresh fruit is also considered, therefore, knowledge of the same is essential (Rotili et al., 2013). It is worth mentioning that heavier fruits are also the largest, which makes them of greater interest to the consumers of fresh fruit (Santos et al., 2015); where, in this work, the correlation between the length and diameter of the fruit is seen previously in Table 2.

Still, in Figure 2, as well be concluded that the arrows or vectors that represent the variables are leading the direction of possible changes of the variables under study within an ordering space. The relevance of the interpretation of the variance designed on the axis is obtained by the length of the arrows, i.e., the size of the arrows is proportional to the explanation of the variation (Silva et al., 2019).

According to results of the Hotelling's test (statistics T^2 = 11,157.13 and χ^2_{10} (0.05) = 18.30), it was possible to verify that the species differ from each other with a p-value < 0.0001, i.e., the null hypothesis of equality of means is rejected. The simultaneous confidence intervals for the difference in averages can verify in Table 4; note that only the peel thickness and Brix variables did not obtain a significant difference, since the confidence interval contains zero.

In addition, Table 4 presents the vector of coefficients referring to the most critical linear combination that leads to the rejection of H_0. Corroborating to this result, MANOVA performed via the permutation test showed that the species considered as a factor was significant (p-value < 0.001) for the joint explanation of the vector of response variables considered as the estimated parameters.

In a study with *Passiflora cincinnata*, Dutra et al. (2019) to distinguish morphoagronomic and variability descriptors of 53 individuals from the base population, needed two components to explain more than 70% of the total variety, which corroborates the results of this study. Santos et al. (2011), found similar results, who used principal component analysis to reduce the 14 descriptors for two main components that accounted for 84% of the total variance in their study to estimate genetic parameters by PCA of two species of *Passiflora*, *P. sub-lanceolata* (Killip) and *P. foetida*, such as their hybrids with ornamental potential.

Table 4. Hotelling's simultaneous confidence intervals at the p ≤ 0.05 and coefficients refering to the most critical linear combination that leads to the rejection of H_0

	Lower	Upper	Coefficient
FL	2.04	8.31	29.20
FD	4.18	10.46	13.99
EC	-8.91	-2.64	-127.09
IC	-6.28	-0.01	-370.04
PT	-1.81	4.46	39.90
JY	13.56	19.83	-3.06
Brix	-2.44	3.84	68.11
AC	24.24	30.51	339.29
NF	-20.57	-14.29	-16.59
TW	-1232.67	-1226.40	0.10

FL - Fruit length (mm); FD - Fruit diameter (mm); EC - External color; IC - Internal color; PT - Peel thickness (mm); JY- Juice yield (mL); Brix - Soluble solids (°Brix); AC - Citric acid (%); NF - Number of fruits; TW - Total fruit weight (g)

Conclusions

1. The principal component analysis proved to be effective, allowing the removal of five variables for yellow passion fruit and six variables for purple passion fruit.

2. *Passiflora edulis* Sims f. *edulis* (purple passion fruit) and *Passiflora edulis* Sims (yellow passion fruit) present contrasting differences between them, being the explanatory variables for the purple passion fruit the external and internal color and for the yellow passion fruit length, diameter, and citric acid.

Acknowledgements

The authors would like to thank the financial support provided by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES - Financing Code 001 and the authors would also like to thank Embrapa Mandioca e Fruticultura (Cruz das Almas, BA, Brazil) for technical support.

Literature Cited

Anderson, M. J. A new method for non-parametric multivariate analysis of variance. Austral Ecology, v.26, p.32-46, 2001.

Arias-Nava, E. H.; Valles-Rosales, D. J.; Sullivan, B. P. Biopolymer non-parametric analysis: A degradation study under accelerated destructive tests. Polymers, v.15, p.1-14, 2023. https://doi.org/10.3390/polym15030620

Bernardes, A. K.; Mizusaki, A. M. P. Statistical methods applied to geochemical analysis of data from the Guaritas Allogroup, Camaquã Basin-RS, Brazil. Journal of South American Earth Sciences, v.116, e103732, 2022. https://doi.org/10.1016/j.jsames.2022.103732

Box, G. E. P.; Cox, D. R. An analysis of transformations. Journal of the Royal Statistical Society: Series B (Methodological), v.26, p.211-243, 1964. https://www.jstor.org/stable/2984418

Costa, T. F.; Lopes, B. G.; Faria, G. A.; Ribeiro, O. A. P. da S.; Suares-Rocha, P.; Lima, J. F. de; Felizardo, L. M.; Furlani Junior, E. Selection of desirable characters for papaya genetic improvement programs associated with hydric and thermal stress. Revista Brasileira de Engenharia Agrícola e Ambiental, v.27, p.422-428, 2023. https://doi.org/10.1590/1807-1929/agriambi.v27n5p422-428

Devi Ramaiya, S.; Bujang, J. S.; Zakaria, M. H.; King, W. S.; Shaffiq Sahrir, M. A. Sugars, ascorbic acid, total phenolic content and total antioxidant activity in passion fruit (Passiflora) cultivars. Journal of the Science of Food and Agriculture, v.93, p.1198-1205, 2013. https://doi.org/10.1002/jsfa.5876

Dutra, J. A.; Oliveira, A. C. de; Porto, A. C. M.; Mathias, J. L. M. Characterization and selection of "Maracujá-do-mato" (*Passiflora cincinnata* Mast) morphoagronomic descriptors. Revista Brasileira de Fruticultura, v.41, e060, 2019. http://dx.doi.org/10.1590/0100-29452019060

Erichson, N. B.; Zheng, P.; Manohar, K.; Brunton, S. L.; Kutz, J. N.; Aravkin, A. Y. Sparse principal component analysis via variable projection. SIAM Journal on Applied Mathematics, v.80, p.977-1002, 2020. https://doi.org/10.1137/18M1211350

Faleiro, F.; Oliveira, J. da S.; Junqueira, N. Banco ativo de germoplasma de passiflora 'flor da paixão': Aspectos históricos e a importância da conservação e caracterização de recursos genéticos. In: Faleiro, F. G.; Oliveira, J. da S.; Junqueira, N; Santos, R. S. Banco de Germoplasma de *Passiflora* L. 'Flor da Paixão' no portal Alelo Recursos Genéticos. Brasília: Embrapa Cerrados, 2019. Cap.1, p.12-22, 2019.

Gabriel, K. R. The biplot graphic display of matrices with application to principal component analysis. Biometrika, v.58, p.453-467, 1971. https://doi.org/10.2307/2334381

Gwelo, A. S. Principal components to overcome multicollinearity problem. *Oradea* Journal of Business and Economics, v.4, p.79-91, 2019. https://doi.org/10.47535/1991ojbe062

Janzantti, N. S.; Monteiro, M. Changes in the aroma of organic passion fruit (*Passiflora edulis* Sims f. flavicarpa Deg.) during ripeness. LWT-Food Science and Technology, v.59, p.612-620, 2014. https://doi.org/10.1016/j.lwt.2014.07.044

Johnson, R. A.; Wichern, D. W. Applied multivariate statistical analysis. London: Pearson, 2014. 800p.

Lopes, B. G.; Otoboni, M. E. F.; Vargas, P. F.; Pavan, B. E.; Oliveira, D. J. L. S. F. de. Correlação canônica e efeito direto e indireto na seleção de acessos de batata-doce coletados no Vale do Ribeira. Research, Society and Development, v.11, e493111134022, 2022. https://doi.org/10.33448/rsd-v11i11.34022

Mason, R. L.; Young, J. C. Multivariate statistical process control with industrial applications. Philadelphia, Pennsylvania: SIAM, 2002. 259p.

Ocampo, J.; Urrea, R.; Wyckhuys, K.; Salazar, M. Exploración de la variabilidad genética del maracuyá (*Passiflora edulis* f. flavicarpa Degener) como base para un programa de fitomejoramiento en Colombia. Acta Agronómica, v.62, p.352-360, 2013. https://doi.org/10.1590/0100-2945-101/14

Ó, L. M. G. do; Cova, A. M. W.; Silva, N. D. da; Azevedo Neto, A. D. de.; Silva, P. C. C.; Gheyi, H. R. Crescimento inicial de minimelancia cv. sugar baby irrigada com águas salobras. Revista Brasileira de Agricultura Irrigada, v.14, p.4086-4096, 2020. https://doi.org/10.7127/rbai.v14n101168

Pinto, W. S; Dantas, A. C. V. L.; Fonseca, A. A. O.; Ledo, C. A. S.; Jesus, S. C.; Calafange, P. L. P.; Andrade, E. M. Caracterização física, físico-química e química de frutos de genótipos de cajazeiras. Pesquisa Agropecuária Brasileira, v.38, p.1059-1066, 2003. https://doi.org/10.1590/S0100-204X2003000900006

R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2020. Available on: < https://www.r-project.org/ > . Accessed on: Feb. 2021.

Ramos, A. M. R.; Amorim, B. M. B.; Freire, C. T. M.; Lima, D. L. F. A. Atributos físicos do solo em sistema consorciado e monocultivo do maracujá *(Passiflora edulis* sims). Revista Brasileira de Engenharia de Biossistemas, v.13, p.80-87, 2019. https://doi.org/10.18011/bioeng2019v13n1p80-87

Rinaldi, M. M.; Costa, A. M.; Faleiro, F. G.; Junqueira, N. T. V. Conservação pós-colheita de frutos de *Passiflora setacea* DC. submetidos a diferentes sanitizantes e temperaturas de armazenamento. Brazilian Journal of Food Technology, v.20, e2016046, 2017. https://doi.org/10.1590/1981-6723.4616

Rotili, M. C. C.; Vorpagel, J. A.; Braga, G. C.; Kuhn, O. J.; Salibe, A. B. Atividade antioxidante, composição química e conservação do maracujá-amarelo embalado com filme PVC. Revista Brasileira de Fruticultura, v.35, p.942-952, 2013. https://doi.org/10.1590/S0100-29452013000400004

Rubarth, K.; Sattler, P.; Zimmermann, H. G.; Konietschke, F. Estimation and testing of Wilcoxon-Mann-Whitney effects in factorial clustered data designs. Symmetry, v.14, p.1-34, 2021. https://doi.org/10.3390/sym14020244

Salem, N.; Hussein, S. Data dimensional reduction and principal components analysis. Procedia Computer Science, v.1, p.292-299, 2019. https://doi.org/10.1016/j.procs.2019.12.111

Santos, E. A.; Souza, M. M.; Viana, A. P.; Almeida, A. A. F.; Freitas, J. C. O.; Lawinscky, P. R. Multivariate analysis of morphological characteristics of two species of passion flower with ornamental potential and of hybrids between them. Genetics and Molecular Research, v.10, p.2457-2471, 2011. http://dx.doi.org/10.4238/2011.October.13.3

Santos, J. A.; Nasser, F. A. D. C. M.; Nasser, M. D. Qualidade dos frutos de goiabeiras submetidas à poda e aplicação de nitrato de potássio. Revista Cultura Agronômica, v.24, p.271-280, 2015. https://doi.org/10.32929/2446-8355.2015v24n4p271-280

Santos, R. D. O.; Gorgulho, B. M.; Castro, M. A. D.; Fisberg, R. M.; Marchioni, D. M.; Baltar, V. T. Principal component analysis and factor analysis: Differences and similarities in nutritional epidemiology application. Revista Brasileira de Epidemiologia, v.22, p.1-14, 2019. https://doi.org/10.1590/1980-549720190041

Silva, L. K. dos S.; Alves, M. C. J. L.; Costa, R. N; Silva, D. M. R.; Santos, J. C. C. dos; Moura, F. de B. P.; Silva Júnior, J. M. da; Silva, J. V. Gas Exchange and photochemical efficiency of caatinga plants submitted to different water management strategies. Journal of Agricultural Science, v.11, p.53-69, 2019. https://doi.org/10.5539/jas.v11n11p53

Vargas, P. F.; Otoboni, M. E. F.; Lopes, B. G.; Pavan, B. E. Prediction of genetic gains through selection of sweet potato accessions. Horticultura Brasileira, v.38, p.387-393, 2020. https://doi.org/10.1590/s0102-0536202004008

Yeo, I. K.; Johnson, R. A. A new family of power transformations to improve normality or symmetry. Biometrika, v.87, p.954-959, 2000. https://www.jstor.org/stable/2673623