



Revista Ciencias Técnicas Agropecuarias

ISSN: 2071-0054

Universidad Agraria de La Habana

Fernández-Chuairey, Lucía; Rangel-Montes de Oca, Lazara; Varela-Nualles, Mario; Pino-Roque, José Antonio; del Pozo-Fernández, Jany; Lim-Chang, Nelson Ulises
Analysis of Main Components, an Effective Tool in Agricultural Technical Sciences
Revista Ciencias Técnicas Agropecuarias, vol. 31, no. 1, e10, 2022, January-March
Universidad Agraria de La Habana

Available in: <https://www.redalyc.org/articulo.oa?id=93270179010>

- ▶ How to cite
- ▶ Complete issue
- ▶ More information about this article
- ▶ Journal's webpage in redalyc.org

The Redalyc logo consists of the word 'redalyc.org' in a lowercase, sans-serif font. The 'e' is stylized with a red dot above it. To the left of the text is a small red icon resembling a stylized 'R' or a similar symbol.

Scientific Information System Redalyc

Network of Scientific Journals from Latin America and the Caribbean, Spain and Portugal

Project academic non-profit, developed under the open access initiative



<https://revistas.unah.edu.cu/index.php/rcta/article/view/1551>

VIEW POINTS | PUNTOS DE VISTA

Analysis of Main Components, an Effective Tool in Agricultural Technical Sciences

Análisis de componentes principales, una herramienta eficaz en las Ciencias Técnicas Agropecuarias

Dr.C. Lucía Fernández-Chuairey¹, MSc. Lazara Rangel-Montes de Oca¹, Mario Varela-Nualles¹, MSc. José Antonio Pino-Roque¹, Dr.MVZ Jany del Pozo-Fernández¹, MSc. Nelson Ulises Lim-Chamg¹

¹Universidad Agraria de La Habana (UNAH), San José de las Lajas, Mayabeque, Cuba.

¹Instituto Nacional de Ciências Agrícola (INCA), San José de las Lajas, Mayabeque, Cuba.

ABSTRACT. Currently there is a wide range of multivariate techniques, which are used in different areas of research. The present work focuses on the Principal Components Method and aims to establish a set of methodological criteria for the processing and interpretation of results in the use of this technique on mathematical-statistical bases. An example associated with post-harvest studies of the pineapple (variety Cayena Lisa) is developed. A sequence of steps is proposed that includes: previous analysis of correlation between variables, determination of the number of components to be selected (compromise between the different criteria), weight of variables in each component, biological interpretation and graphs that validate the results obtained in reference to components and individuals. The study had the variables: weight loss in g (PP), firmness, color index (IC), soluble solids content (SSC) and pH. The variables were grouped into two components that explain 88.36% of the variation in the data. A positive relationship was observed among PP, SSC and pH and the negative relationship of firmness with these variables. It is shown that the highest PP and pH are reached from the sixth day and the highest firmness, in the first two days, aspects to take into account in making timely decisions for storage, transportation and marketing. It is concluded that the use of multivariate techniques and, particularly, the analysis of principal components constitutes an efficient and non-destructive way in monitoring the quality of fruits in storage.

Keywords: Main Components, Agricultural Engineering, Multivariate Methods.

RESUMEN. En la actualidad existe una amplia gama de técnicas multivariadas, que se utilizan en las diferentes áreas de investigación. El presente trabajo se centra en el Método de Componentes Principales y tiene como objetivo establecer sobre bases matemático-estadísticas un conjunto de criterios metodológicos para el procesamiento e interpretación de resultado con el empleo de dicha técnica. Se desarrolla un ejemplo asociado a estudios pos cosecha de la Piña (*variedad Cayena Lisa*). Se proponen una secuencia de pasos que incluye: análisis previo de correlación entre variables, determinación de números de componentes a seleccionar (compromiso entre los diferentes criterios), peso de variables en cada componente, interpretación biológica y gráficos que validan los resultados obtenidos en sentido de las componentes e individuos. El estudio contó con las variables: pérdida de peso en g (PP), firmeza, índice de color (IC), contenido de sólidos solubles (SSC) y PH. Las variables se agruparon en dos componentes que explican el 88,36 % de la variación de los datos. Se observó una relación positiva entre PP, SSC y PH y la relación negativa de la firmeza con estas variables, se muestra que la mayor PP y PH se alcanza a partir del sexto día, y la mayor firmeza en los dos primeros días, aspectos a tener en cuenta en la toma de decisiones oportunas para el almacenaje, transporte y comercialización. Se concluye que el empleo de técnicas multivariadas y en particular el análisis de componentes principales constituye una vía eficiente y no destructiva en el monitoreo de la calidad de frutos en almacenamiento.

Palabras clave: componentes principales, ingeniería agrícola, métodos multivariados.

¹ Author for correspondence: Héctor R. de las Cuevas Milán, e-mail: cuevasm@nauta.cu ORCID iD: <https://orcid.org/0000-0002-0467-9749>

Received: 25/08/2021.

Approved: 12/11/2021.

INTRODUCTION

Historically, in agricultural sector, the need for the use of different statistical-mathematical methodologies that respond to current problems in scientific research has been present. Recently, Fernández *et al.* (2018; 2019), established criteria and evaluations on mathematical-statistical bases in the analysis and application of models that describe agrarian processes (based mainly on univariate and bivariate statistics).

Similarly, the literature reports on the use of multivariate methods, which are used to study phenomena that include the measurement of several variables and which are applied depending on the characteristics of the research. Among the most used multivariate statistical techniques are: Multiple Regression; Principal Component Analysis (PCA); Factor Analysis (AF); Discriminant Analysis (AD); the Numerical Taxonomy (CLUSTER); Multidimensional Scaling, among others, those that have been addressed by Lozares & López (1991); Robaina *et al.* (2001); Hair & Anderson (2004); Bouza & Sistachs (2006); González *et al.* (2008); Miranda (2011); Coronados *et al.* (2017); Quindemil & Rumbaut (2019); Gozá *et al.* (2020); Varela (2021) among other authors.

The objective of this work is to establish, on mathematical-statistical bases, a set of methodological criteria for the processing and interpretation of results with the use of the Principal Components method, its analysis is focused on post-harvest studies of pineapple (variety Cayenne Lisa)

DEVELOPMENT OF THE TOPIC

Theoretical Fundament

Various criteria have been given on the definition of multivariate statistical techniques. A general definition was proposed by Hair & Anderson (2004), who argue that “Multivariate analysis refers to all statistical methods that simultaneously analyze multiple measures of each individual or object under investigation and emphasize that any simultaneous analysis of more than two variables can be considered approximately as a multivariate analysis”.

These methods group a set of statistical techniques that are responsible for the analysis of data corresponding to measurements of p variables observed in n individuals; allowing the study of interrelations. The literature collects various multivariate methods, and classifies them fundamentally according to the purposes pursued in the research. In this sense Varela (2021), based on an analysis carried out, groups them into descriptive or decisional and alleges that one of the most widespread Multivariate Analysis techniques at present is the Principal Component Analysis (PCA) where the variables are quantitative, since it works with the Pearson correlation coefficient, designed to measure linear association between variables of this type, although there is the Principal Component Analysis option for categorical variables, which will be addressed in a future work.

Miranda (2011), refers that the objective of the ACP is to reduce the number of variables that intervene in an analysis

INTRODUCCIÓN

Históricamente en el sector agrario ha estado presente la necesidad, del empleo de diferentes Metodologías Estadísticas-Matemáticas que den respuesta a problemas actuales en Investigaciones Científicas. Recientemente Fernández *et al.* (2018; 2019), establecieron sobre bases matemático-estadísticas criterios y valoraciones en el análisis y aplicación de modelos que describen Procesos Agrarios (basado fundamentalmente en la estadística univariada y bivariada).

De igual forma, la literatura notifica sobre el empleo de métodos multivariados, que se emplean para estudiar fenómenos que incluyen la medición de varias variables y que se aplican en dependencia de las características de la investigación. Entre las técnicas estadísticas multivariadas más utilizadas esta: Regresión Múltiple; Análisis de Componentes Principales (ACP); Análisis de Factorial (AF); Análisis Discriminante (AD); la Taxonomía Numérica (CLUSTER); Escalamiento Multidimensional, entre otros, los que han sido abordados por Lozares & López (1991); Robaina *et al.* (2001); Hair & Anderson (2004); Bouza & Sistachs (2006); González *et al.* (2008); Miranda (2011); Coronados *et al.* (2017); Quindemil & Rumbaut (2019); Gozá *et al.* (2020); Varela (2021) entre otros autores.

El objetivo del presente trabajo es establecer sobre bases matemático-estadísticas un conjunto de criterios metodológicos para el procesamiento e interpretación de resultados con el empleo del método de Componentes Principales, se centra su análisis en estudios postcosecha de la Piña (*variedad Cayena Lisa*)

DESARROLLO DEL TEMA

Fundamentos Teóricos

Se han dado diversos criterios sobre la definición de técnicas estadísticas multivariadas, una definición general fue planteada por Hair & Anderson (2004), quienes fundamentan que “El análisis multivariado, se refiere a todos los métodos estadísticos que analizan simultáneamente medidas múltiples de cada individuo u objeto sometido a investigación y enfatizan que cualquier análisis simultáneo de más de dos variables puede ser considerado aproximadamente como un análisis multivariante.

Estos métodos agrupan un conjunto de técnicas estadísticas que se encargan del análisis de datos correspondientes a mediciones de p variables observadas en n individuos; permitiendo el estudio de las interrelaciones. La literatura recoge diversos métodos multivariados, y los clasifican fundamentalmente atendiendo a los fines que se persiguen en la investigación, en tal sentido Varela (2021), a partir de un análisis realizado, los agrupa en descriptivo o decisional y alega que una de las técnicas del Análisis Multivariado más difundida en la actualidad es el **Análisis de Componentes Principales (ACP)** donde las variables son cuantitativas, ya que se trabaja con el coeficiente de correlación de Pearson, diseñado para medir asociación lineal entre variables de este tipo, aunque existe la opción del Análisis de Componentes Principales para variables categóricas, el cual será abordado en un próximo trabajo.

Miranda (2011), refiere que el objetivo del (ACP) es reducir el número de variables que intervienen en un análisis de un de-

of a certain process under study. And it states that the method consists of obtaining new variables (called Y_p components) that are unrelated to each other and that keep a logical order, where the first component is the one with the greatest influence on the phenomenon under study and so on, that is:

$$\text{Var}(Y_1) + \text{Var}(Y_2) + \dots + \text{Var}(Y_p) = \text{Total Variance} = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_p)$$

such that: $\text{Var}(Y_1) > \text{Var}(Y_2) > \dots > \text{Var}(Y_p)$

How to describe the information contained in a data set by a smaller set of new variables or components? When is it effective to apply the Principal Components Method?

Principal Component Analysis is more effective to the extent that initially there is a marked correlation structure between the variables. In this respect, Miranda (2011) corroborates that, when there is no association between the variables, it makes no sense to carry out these types of analysis.

This procedure is used above all in exploratory data analysis and for descriptive purposes, it manages to simplify the studies that will be made from a smaller number of variables than the original, as well as to elucidate the relationship and weight between the observed variables and, at the same time, it allows observing the formation of groups of individuals attending to their behavior from graphic representations.

The application of this method starts from the data matrix of n individuals with p variables in which $n \geq p$, where a sequence of steps that correspond to the following aspects is applied:

- Construction of the components (it should be noted that when the quantitative variables appear on the same measurement scale, the variance and covariance matrix is used, in the case that they are on a different scale, the correlation matrix (standardization) is used).
- Selection of the number of components to take into account (percentage criteria: to include sufficient criterion of principal components that give a percentage of the acceptable variance (regularly above 70%), or eigenvalue criterion with values that are greater or equal to 1, among other criteria. Practical experiences indicate working in the sense of a compromise between different criteria).
- Analysis of the variables. Relationship or weight of variables in each component.
- Biological sense of the components from their relationship with the initial variables.
- Graphic analysis (individuals), formation of possible groups

At present, there are valuable results regarding the use of these techniques, as shown in the works of Mesa *et al.* (2018) in monoclonal antibody fermentation studies, in the same way they were used in investigations associated with biopharmaceutical purification processes carried out by Goza *et al.* (2020). Their use in problems associated with causality in Biomedical Sciences is also reported, which included the determination of risk factors and prognoses (Sagaró & Zamora, 2020), as well as studies of functional dynamic mechanical systems of internal combustion engines according to Aliaga *et al.* (2021), among other applications.

terminado proceso en estudio. Y expresa que el método consiste en obtener nuevas variables (denominadas componentes Y_p) que están incorrelacionadas entre si y que guardan un orden lógico, donde la primera componente es la de mayor influencia en el fenómeno de estudio y así sucesivamente, es decir:

¿Cómo describir la información contenida en un conjunto de datos mediante un conjunto menor de nuevas variables o componentes? ¿Cuándo es efectivo aplicar el Método de Componentes Principales?

El Análisis de Componentes Principales resulta más efectivo en la medida en que inicialmente exista una estructura de correlación marcada entre las variables; en tal sentido Miranda (2011) corrobora que, cuando no existe asociación entre las variables, carece de sentido realizar estos tipos de análisis.

Este procedimiento se emplea sobre todo en análisis exploratorio de datos y con fines descriptivos, el mismo logra simplificar los estudios que se harán a partir de un número menor de variables que la original, así como dilucidar la relación y peso entre las variables observadas, y paralelamente permite observar la formación de grupos de individuos atendiendo al comportamiento de estos a partir de representaciones gráficas.

Para la aplicación de este método se parte de la matriz de datos de n individuos con p variables en el que ($n \geq p$), donde se aplica una secuencia de pasos que se corresponden con los siguientes aspectos:

- Construcción de las componentes (es de destacar que cuando las variables cuantitativas aparecen en la misma escala de medida se utiliza la matriz de varianza y covarianza, en el caso que estén en diferente escala se emplea la matriz de correlación (estandarización).
- Selección del número de componentes a tener en cuenta (criterios de porcentaje: Incluir el criterio suficiente de componentes principales que den un porcentaje de la varianza aceptable (regularmente por encima del 70%), o criterio de autovalor con valores que sean mayores o iguales a 1, entre otros criterios, las experiencias prácticas indican trabajar en el sentido de un compromiso entre diferentes criterios).
- Análisis de las variables. Relación o peso de variables en cada componente.
- Sentido biológico de las componentes a partir de su relación con las variables iniciales.
- Análisis gráfico (individuos), formación de posibles grupos.

En la actualidad se cuenta con valiosos resultados relativos al empleo de estas técnicas, como muestras los trabajos de Mesa *et al.* (2018), en estudios de fermentación de anticuerpo monoclonal, de igual forma se emplearon en investigaciones asociadas a procesos de purificación de biofármaco realizado por Gozá *et al.* (2020), se reporta además su empleo en problemas asociados a la causalidad en las Ciencias Biomédicas, que incluyó la determinación de factores de riesgo y pronósticos Sagaró & Zamora (2020), así como también en estudios de sistemas mecánicos dinámicos funcional de motores de combustión interna según Aliaga *et al.* (2021) entre otras aplicaciones.

Example of Application of ACP in Post-Harvest Studies of Pineapple (*variety Cayena Lisa*)

Pineapple is one of the most important commercial fruit crops in the world, it is known as the queen of fruits for its excellent taste and its implication in nutrition and health (Hernández *et al.*, 2021), hence, currently the research associated with its characterization, nutritional composition, growth studies, quality, post-harvest, among other aspects, is intensified as shown in the works of Rangel *et al.* (2018) and Lorente *et al.* (2021), among others.

Luchsinger (2017) considers that one of the impacts of post-harvest studies lies in maintaining the quality of the products until their consumption, hence the importance of investigating the different indicators. The study was carried out in areas of the company of various crops located in Havana-Matanzas Plain, with a range of average annual temperature between 25 and 32 °C and high environmental humidity. The Weight Loss (PP) was carried out through the weighing of the fruits with the use of the electronic scale, during the days (1, 2, 3, 5, 6, 8 and 10) of harvest, and indicators such as PP, firmness, color index (IC), soluble solids content (SSC) and pH. It is desired to analyze the behavior of these variables (5 variables) on the different days (6 individuals).

The data were processed using statistical software (Statgraphics Centurion, 2012). A previous analysis showed that there is a marked correlation structure among this group of variables, which shows a positive and direct relationship between (PP - pH with $r = 0.84$) and (of pH -SSC with $r = 0.62$). It was also observed a negative and inverse relationship between (PP-firmness with $r = -0.80$) and (CI-firmness with $r = - 0.65$), which suggests a study using principal component analysis.

Construction and Selection of the Number of Components

Table 1 shows the selection of two components (eigenvalues above one). Note that the first two components explain 88.36% of the total variability. This indicates that, from 5 initial variables, two components can be extracted to explain the association between the variables and observations.

TABLE 1. Number of Principal Components from criteria of eigenvalues and percentage
 TABLA 1. Número de Componentes Principales a partir de criterios de autovalores y porcentaje

Components	Eigenvalue	Percentage of variance	Cumulative percentage
1	2.71	55.43	55.43
2	1.64	32.93	88.36
3	0.51	10.38	98.74
4	0.05	1.02	99.76
5	0.01	0.23	100

Relationship or Weight of Variables in Each Component

The weight of the variables in component 1 is fundamentally characterized by the variables loss of weight, pH and

Ejemplo de aplicación de ACP en estudios pos cosecha de la Piña (*variedad Cayena Lisa*)

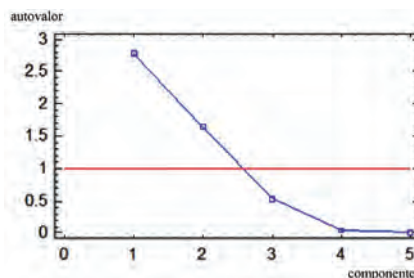
La piña es uno de los cultivos comerciales de frutas más importantes del mundo, es conocida como la reina de las frutas por su excelente sabor y su implicación en la alimentación y la salud según Hernández *et al.* (2021), de ahí que en la actualidad se intensifican las investigaciones asociadas a su caracterización, composición nutricional, estudios crecimiento, calidad, postcosecha, entre otros aspectos como muestran los trabajos de Rangel *et al.* (2018) y Lorente *et al.* (2021), entre otros.

Luchsinger (2017) considera que uno de los impactos de los estudios pos cosecha radica en mantener la calidad de los productos hasta su consumo, de ahí la importancia de indagar en los diferentes indicadores. El estudio se llevó a cabo en áreas de la empresa de cultivos varios ubicados dentro de la Llanura Habana-Matanzas, con un rango de la temperatura media anual entre los 25 y los 32 °C y elevada humedad ambiental. La Pérdida de Peso (PP) se realizó a través del pesaje de los frutos con la utilización de la balanza electrónica, durante los días (1, 2, 3, 5, 6, 8 y 10) de cosechado, y se midieron indicadores como: PP, firmeza, índice de color (IC), contenido de sólidos solubles (SSC) y PH. Se desea analizar en comportamiento de estas variables (5 variables) en los diferentes días (6 individuos).

Los datos fueron procesados mediante el software estadístico (Statgraphics Centurion, 2012). Un análisis previo mostró que existe una estructura de correlación marcada entre este grupo de variables, que muestra una relación positiva y directa entre (PP - PH con $r = 0.84$) y (de PH -SSC con $r = 0.62$), de igual forma se observó una relación negativa e inversa entre (PP-firmeza con $r = -0.80$) y (IC-firmeza con $r = - 0.65$), lo que sugiere un estudio con el empleo de análisis de componentes principales.

Construcción y selección del número de componentes

Se observa en la Tabla 1 la selección de dos componentes (autovalores por encima de uno). Nótese que con las dos primeras componentes se explica el 88.36% de la variabilidad total. Esto indica que, de 5 variables iniciales, se puede extraer dos componentes para explicar la asociación entre las variables y observaciones



Relación o peso de variables en cada componente

El peso de las variables en la componente 1 está caracterizada fundamentalmente por las variables pérdida de peso, PH y firmeza (Tabla 2) mientras que la componente

firmness (Table 2) while component 2 is characterized by the soluble solids content and the color index.

TABLE 2. Component weights
TABLA 2. Pesos de las componentes

	Component 1	Component 2
Weight loss	0.562	-0.048
Firmness	-0.516	0.386
(IC)	0.120	-0.728
(SSC)	0.329	0.484
pH	0.541	0.287

In the case of Component 1, with positive values in weight loss and pH, it can be stated that as the value of Component 1 increases, the weight loss and pH increase and the firmness of the fruits decreases. On the other hand, in the case of the second component, as its value increases, it indicates that the values of the contents of soluble solids increase and the color index decreases.

Considering the graphic representation (Figure 1), it can be argued that there are basically three groups in post-harvest. The first group characterized by the greatest loss of weight and pH, which occurs from the sixth day. From the physical point of view, the weight losses, associated with the water content of the fruit, indirectly decrease the concentration of hydrogen ions by doing this, that the pH rises, due to the senescence or putrefaction that it is reaching, which it does not facilitate its consumption as fresh fruit, hence the importance and timely decision-making for commercialization and industrialization.

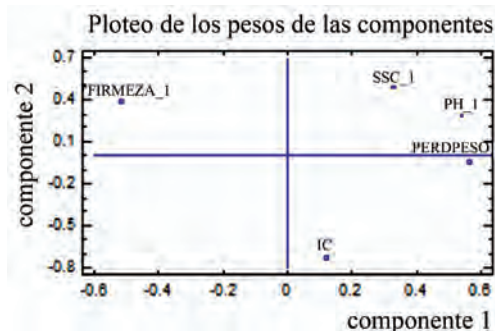
Formation of possible groups. Biological sense of the components from their relationship with the initial variables

day	Component 1	Component 2
1	-2.422	1.681
2	-0.524	0.738
3	-0.840	-0.802
5	-0.817	-1.593
6	0.463	-1.478
8	1.67	0.528
10	2.47	0.926

FIGURE 1. Graphic analysis of individuals and group formation. Principal Component values for each row.
FIGURA 1. Análisis gráficos de individuos y formación de grupos. Valores de los Componentes principales para cada fila.

In contrast to it, there is the third group, formed by the first day, where the greatest firmness is achieved, with the least loss of weight and pH. This answer is given due to the nature of the product, because once the exchange of ethylene with the surrounding environment begins, it causes increased respiration and accelerates the ripening process, a recurring phenomenon in previous investigations with this or other agricultural products (Thompson, 1998). Likewise, a gradual response is reflected in the concentration of soluble solids contents that tends to influence its acceptance by consumers and marketers. As well as the color index which allows the naked eye to discern its state of maturity, regardless of its

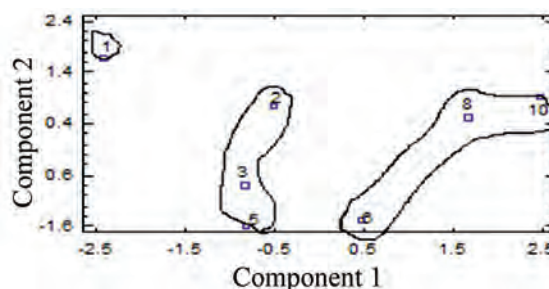
2 se caracteriza los contenidos de solidos solubles y el índice de color.



En el caso de la Componente 1, con valores positivos en pérdida de peso y PH, se puede afirmar que a medida que aumenta el valor de la componente 1, se incrementa la pérdida de peso y el PH y disminuye la firmeza de los frutos. Por otra parte, en el caso de la segunda componente a medida que su valor se incrementa, indica que los valores de los contenidos de los sólidos solubles aumentan y disminuye el índice de color.

Atendiendo a la representación gráfica (Figura 1), se puede plantear que existen fundamentalmente en post cosecha, tres grupos. El *primer grupo* caracterizado por la mayor pérdida de peso y pH, que sucede a partir del sexto día. Desde el punto de vista físico las pérdidas de peso, asociadas con los contenidos de agua de la fruta, indirectamente disminuyen la concentración de iones hidrogeno haciendo esto, que el pH se eleve, producto a la senescencia o putrefacción que este está alcanzando, lo cual no facilita su consumo como fruto fresco, de ahí la importancia y la toma de decisiones oportuna para la comercialización e industrialización.

Formación de posibles grupos. Sentido biológico de las componentes a partir de su relación con las variables iniciales



En contraposición con él, está el *tercer grupo*, formado por el primer día, donde se alcanza la mayor firmeza, menor pérdida de peso y de pH. Esta respuesta está dada a causa de la naturaleza del producto, pues una vez que comience el intercambio del etileno con el medio circundante, provoca el aumento de la respiración y acelera el proceso de maduración, fenómeno recurrente en investigaciones precedentes con este u otros productos agrícolas (Thompson, 1998). Igualmente se refleja una respuesta paulatina en la concentración de contenidos de solidos solubles que suelen incidir en su aceptación ante consumidores y comercializadores. Así como el índice de color el cual permite discernir a simple vista su estado de madurez, independientemente que su menor valor a

lowest value to be reached, is included in the first day after harvest as reflected in component 2.

The result obtained of the quality of pineapple represented by these groups constitutes a valuable tool that avoids from carrying out an exhaustive control of these properties during their commercialization, transport or storage and even to make up for the lack of instrumentation for their determination. This largely makes it a non-destructive tool to monitor the quality of the fruit in storage. One of the main purposes and curiosities of this research is also satisfied. This would enhance in this time range the timely decision-making in relation to its storage, transport and commercialization. This reaffirms the criterion that quality is sought from the field and is modulated post-harvest.

alcanzar, está comprendido en el primer día pos cosecha tal y como lo refleja la componente 2.

El resultado obtenido de calidad de la piña representada por estos agrupamientos constituye una valiosa herramienta que nos exonera de llevar a cabo un control exhaustivo de estas propiedades durante su comercialización, transporte o almacenaje e incluso suplir la falta de instrumentación para su determinación. Ello en buena medida la convierte en una herramienta no destructiva para monitorear la calidad del fruto en almacenamiento. También se satisface uno de los principales propósitos y curiosidades de esta investigación. Esto potenciaría en este rango de tiempo la toma de decisiones oportunas en relación a su almacenaje, transporte y comercialización. Ello reafirma el criterio que la calidad se procura desde el campo y se modula pos cosecha.

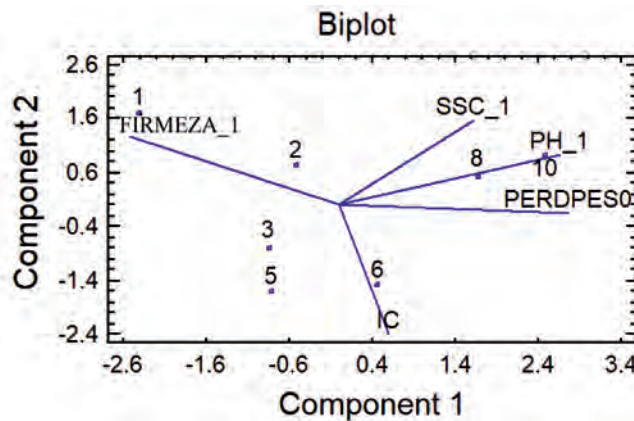


FIGURE 2. Biplot graph.
FIGURA 2. Gráfico de Biplot.

Finally, the Biplot graph (Figure 2), allowed the joint analysis of variables and individuals. The positive relationship among SSC, pH and weight loss and the negative relationship of firmness with the previous variables can be appreciated; corresponding to days 8 and 10 the highest values of SSC, pH and weight loss and the lowest values of firmness in contrast to day 1. Similarly, it is observed (by means of perpendicular to the firmness axis) that the greatest firmness is reached in the first two days.

CONCLUSIONS

- It is concluded that the use of multivariate techniques, on methodological bases and emphasis on the interpretation of the results, increases the quality of scientific research in agricultural and related processes.
- The use of principal component analysis is an alternative analysis tool in post-harvest studies and constitutes an efficient and non-destructive way to monitor the quality of fruits in storage.

AUTHOR CONTRIBUTIONS:

Conceptualization: L. Fernández Data curation. L. Fernández, L.R. Montes de Oca Formal analysis L. Fernández, J. A. Pino, J.del Pozo, N. U. Lim, Investigation L. Fernández, L.R. Montes de Oca ,M Varela., J. A. Pino , J.del Pozo N. U. Lim Methodology: Resources : L. Fernández , L.R. Montes de Oca Roles/Writing, original draft L. Fernández, Writing, review & editingL. Fernández, L.R. Montes de Oca , M Varela., J. A. Pino , J.del

Por último, el grafico de Biplot (Figura 2), permitió el análisis conjunto de variables e individuos. Se puede apreciar la relación positiva entre SSC, PH y pérdida de peso y la relación negativa de la firmeza con las variables anteriores; correspondiendo a los días 8 y 10 los valores más altos de SSC, Ph y pérdida de peso y los valores más bajos de firmeza en contraposición con el día 1. De igual forma se observa (mediante perpendiculares al eje de firmeza) que es en los dos primeros días es donde se alcanza la mayor firmeza.

CONCLUSIONES

- Se concluye que el uso de técnicas multivariadas, sobre bases metodológicas y énfasis en la interpretación de los resultados, eleva la calidad de las investigaciones científicas en procesos agrarios y afines.
- El uso de análisis de componentes principales resulta una herramienta alternativa de análisis en estudios de post-cosecha y constituye una vía eficiente y no destructiva en el monitoreo de la calidad de frutos en almacenamiento.

AUTHOR CONTRIBUTIONS:

Conceptualization: L. Fernández Data curation. L. Fernández, L.R. Montes de Oca Formal analysis L. Fernández, J. A. Pino, J.del Pozo, N. U. Lim, Investigation L. Fernández, L.R. Montes de Oca ,M Varela., J. A. Pino , J.del Pozo N. U. Lim Methodology: Resources : L. Fernández , L.R. Montes de Oca Roles/Writing, original draft L. Fernández, Writing, review & editingL. Fernández, L.R. Montes de Oca , M Varela., J. A. Pino , J.del

REFERENCES

- ALIAGA, N.R.; DE LA TORRE, S.F.; RODRÍGUEZ, S.A.A.; GUILLÉN, G.J.: “Análisis de componentes principales en los motores de combustión interna Hyundai 1.7 MW”, *Revista Ingeniería Energética*, 42(1), 2021, ISSN: 1815-5901.
- BOUZA, C.N.; SISTACHS, V.: *Estadística, teoría básica y ejercicios*, Ed. Editorial Félix Varela, La Habana, Cuba, 2006, ISBN: 959-258-373-0.
- CORONADOS, Y.; VILTRES, V.; SISTACH, V.: “Aplicación de técnicas estadísticas multivariantes en el análisis de datos”, *Revista Cubana de Medicina Física y Rehabilitación*, 9(2): 1-12, INFOMED., 2017.
- FERNÁNDEZ, C.L.; GUERRA, B.C.W.; DE CALZADILLA, P.J.; CHANG, L.N.U.: “Desarrollo de la modelación estadístico-matemática en las ciencias agrarias. Retos y perspectivas”, *Investigación Operacional*, 38(5): 462-467, 2018, ISSN: 2224-5405.
- FERNÁNDEZ, C.L.; RANGEL, M. de O.L.; GUERRA, B.C.W.; DEL POZO, F.J.: “Modelación Estadístico-Matemática en Procesos Agrarios. Una aplicación en la Ingeniería Agrícola”, *Revista Ciencias Técnicas Agropecuarias*, 28(2): 72-79, 2019, ISSN: 1010-2760, e-ISSN: 2071-0054.
- GONZÁLEZ, Á.L.; SOLANO, H.L.; TILANO, J.: “Análisis multivariado aplicando componentes principales al caso de los desplazados”, *Ingeniería y desarrollo*, (23): 119-142, 2008, ISSN: 0122-3461.
- GOZÁ, L.O.; FERNÁNDEZ, A.M.; RODRÍGUEZ, G.R.H.; OJITO, M.E.: “Aplicación del Análisis de Componentes Principales en el proceso de purificación de un biofármaco”, *Vaccimonitor*, 29(1): 5-13, 2020, ISSN: 1025-028X.
- HAIR, J.F.; ANDERSON, R.E.: *Multivariate data analysis*, Ed. Pearson Prentice Hall, 5ª ed., Madrid, España, 2004, ISBN: 84-8322-035-0.
- HERNÁNDEZ, R.G.; ORTEGA, I.E.; ORTEGA, I.I.H.: “Composición nutricional y compuestos fitoquímicos de la piña (Ananas comosus) y su potencial emergente para el desarrollo de alimentos funcionales”, *Boletín de Ciencias Agropecuarias del ICAP*, 9(14): 24-28, 2021, ISSN: 2448-5357.
- LORENTE, G.Y.; RODRÍGUEZ, H.D.; CAMACHO, R.L.; CARVAJAL, O.C.C.; DE ÁVILA, G.R.; GONZÁLEZ, O.J.; RODRÍGUEZ, S.R.: “Efecto de la aplicación de Biobras-16 sobre el crecimiento y calidad de frutos de piña ‘MD-2’”, *Revista de Cultivos Tropicales*, 42(2), 2021, ISSN: 0258-5936.
- LOZARES, C.C.; LÓPEZ, R.P.: “El análisis multivariado: definición, criterios y clasificación”, 1991.
- LUCHSINGER, L.: *Impacto de la postcosecha en la calidad de frutos de exportación. [en línea]*, Perú, Redagráfica, 2017, Disponible en: <https://www.redagricola.com/pe/impacto-de-la-postcosecha-en-la-calidad-de-frutas-de-exportacion>, [Consulta: 9 de julio de 2021].
- MESA, R.L.; GOZÁ, L.O.; URANGA, M.M.; TOLEDO, R.A.; GÁLVEZ, T.Y.: “Aplicación del Análisis de Componentes Principales en el proceso de fermentación de un anticuerpo monoclonal”, *Vaccimonitor*, 27(1): 8-15, 2018, ISSN: 1025-028X, e-ISSN: 1025-0298.
- MIRANDA, I.: *Estadística Aplicada a la Sanidad Vegetal*, Inst. Centro Nacional de Sanidad Agropecuaria (CENSA), folleto, San José de las Lajas, mayabeque, Cuba, 173 p., 2011.
- QUINDEMI, T.E.M.; RUMBAUT, L.F.: “Análisis de componentes principales para obtener indicadores reducidos de medición en la búsqueda de información”, *Revista Cubana de Información en Ciencias de la Salud*, 30(3), 2019, ISSN: 2307-2113.
- RANGEL, M. de O.L.; MONZÓN, M.L.L.; GARCIA, C.J.; GARCIA, P.A.: “Técnicas matemáticas para inferir cambios poscosecha en las propiedades de productos agrícolas”, *Revista Ciencias Técnicas Agropecuarias*, 27(4): 42-54, 2018, ISSN: 1010-2760, e-ISSN: 2071-0054.
- ROBAINA, C.G.R.; MEDINA, P.; MANUEL, J.; MORALES, R.J.M.; ROBAINA, C.R.E.: “Análisis multivariado de factores de riesgo de prematuridad en Matanzas”, *Revista Cubana de obstetricia y ginecología*, 27(1): 62-69, 2001, ISSN: 0138-600X.
- SAGARÓ, D.C.N.M.; ZAMORA, M.L.: “Técnicas estadísticas multivariadas para el estudio de la causalidad en Medicina”, *Revista Ciencias Médicas*, 24(2), 2020, ISSN: 1561-3194.
- STATGRAPHICS CENTURION: *Statgraphics Centurion, X: “Version 16.1. 17”*, Statpoint Technologies, Inc., 2012.
- THOMPSON, K.A.: *Tecnología post-cosecha de frutas y hortalizas*, Ed. Kinesis Ltda., Colombia, 268 p., 1998.
- VARELA, M.: *Análisis multivariado, [en línea]*, Ediciones INCA, 2021, Disponible en: <http://ediciones.inca.edu.cu/files/folleto/analisismultivariado.pdf>, [Consulta: 30 de abril de 2021].

Lucía Fernández-Chuairey, Profesor Titular, Universidad Agraria de La Habana (UNAH), Departamento de Matemática y Física, e-mail: lucia@unah.edu.cu ORCID iD: <https://orcid.org/0000-0003-2439-1176>

Lazara Rangel-Montes de Oca, Profesor Asistente, (UNAH), Departamento de Ingeniería Agrícola, e-mail: lazarar@unah.edu.cu ORCID iD: <https://orcid.org/0000-0003-1640-1987>

Mario Varela Nualles, Investigador Titular, Instituto Nacional de Ciencia agrícola (INCA), e-mail: varela@inca.edu.cu ORCID iD: <https://orcid.org/orcid.org/0000-0002-5953-7561>

José Antonio Pino Roque, Profesor Auxiliar (UNAH), Departamento de Matemática y Física, e-mail: pino@unah.edu.cu ORCID iD: <https://orcid.org://orcid.org/0000-0001-9728-6700>

Jany del Pozo-Fernández, Instructor, Universidad Agraria de La Habana (UNAH), Facultad de Medicina Veterinaria, e-mail: janydelpozo@gmail.com ORCID iD: <http://orcid.org/0000-0001-9138-1563>

Nelson Ulises Lim Chamg, Profesor Auxiliar (UNAH), Departamento de Matemática y Física e-mail: limc@unah.edu.cu ORCID iD: <http://orcid.org/0000-0002-7000-802X>

The authors of this work declare no conflict of interests.

This item is under license Reconocimiento-NoComercial de Creative Commons 4.0 Internacional (CC BY-NC 4.0).

The mention of trademarks of specific equipment, instruments or materials is for identification purposes, there being no promotional commitment in relation to them, neither by the authors nor by the publisher.