Vol. 22, No. 2(2023) Alim2382
Revista Mexicana de Ingeniería Química

Statistical analysis on phenolic compounds dataset obtained by HPLC-MS/MS in fruit juices as a discrimination tool

Análisis estadístico del conjunto de datos de compuestos fenólicos obtenidos por HPLC-MS/MS en zumos de frutas como herramienta de discriminación

J. Llupa[1]*, E. Kalluci[2], D. Topi[3], K. Akrida-Demertzis[1], P. Demertzis[1]
[1]*Laboratory of Food Chemistry and Food Microbiology, Department of Chemistry, University of Ioannina, GR-45110 Ioannina, Greece.*
[2]*Department of Applied Mathematics, Faculty of Natural Sciences, University of Tirana, 1001, Tirana, Albania.*
[3]*Department of Chemistry, Faculty of Natural Sciences, University of Tirana, 1001, Tirana, Albania.*
Received: June 30, 2023; Accepted: July 20, 2023

Abstract
The phenolic profile patterns in selected fruit juices were investigated by conducting qualitative and quantitative analyses. Phenolic compounds comprise a broad range of secondary metabolites characterized by wide biological activity, with two of the most distinguished antioxidant and antimicrobial properties. Obtained through cold-press method, different fruits, such as apple, plum, cherry, and quince, were investigated concerning their phenolic profiles by Ultra-High-Performance Liquid Chromatography and Mass Spectroscopy. The mass data hunting identified 35 phenolic compounds. Application the Principal Component Analysis before and after phenolic group clustering enabled fruit juice discrimination according to their constituents. Principal Component Analysis effectively discriminated the main phenolic compounds and their fragments of first and second generations by resolving the regularity and discrepancy of the principal component statistically. These results indicate a strong correlation between hydroxybenzoic and phenolic acids, respectively, 89% and 54% with flavonols.
Keywords: fruit juice, phenolic compounds, UHPLC, PCA, Multivariance.

Resumen
Se investigaron los patrones del perfil fenólico en jugos de frutas seleccionados mediante la realización de análisis cualitativos y cuantitativos. Los compuestos fenólicos comprenden una amplia gama de metabolitos secundarios caracterizados por una amplia actividad biológica, con dos de las propiedades antioxidantes y antimicrobianas más distinguidas. Obtenidos a través del método de prensado en frío, diferentes frutas, como manzana, ciruela, cereza y membrillo, fueron investigadas en cuanto a sus perfiles fenólicos mediante cromatografía líquida de ultra alta resolución y espectroscopia de masas. La búsqueda masiva de datos identificó 35 compuestos fenólicos. La aplicación del Análisis de Componentes Principales antes y después de la agrupación de grupos fenólicos permitió la discriminación del jugo de frutas según sus constituyentes. El análisis de componentes principales discriminó de manera efectiva los principales compuestos fenólicos y sus fragmentos de primera y segunda generación al resolver estadísticamente la regularidad y la discrepancia del componente principal. Estos resultados indican una fuerte correlación entre los ácidos hidroxibenzoico y fenólico, respectivamente, 89% y 54% con flavonoles.
Palabras clave: jugo de frutas, compuestos fenólicos, UHPLC, PCA, Multivarianza.

*Corresponding author. *E-mail*: jonilda.llupa@gmail.com
https://doi.org/10.24275/rmiq/Alim2382
ISSN:1665-2738, *issn-e*: 2395-8472

1 Introduction

The world population growth trends have significantly increased the global demand for food in quantity and quality. Fruit crops are of crucial importance to ensuring a qualitative diet. Apple (*Malus domestica*), plum (*Prunus domestica*), cherry (*Prunus avium*), and quince (*Cydonia oblonga*) are among the most distinguished tree fruits from temperate regions. World fruit production in 2020 went to 887 million tons (t), apples 93.14 $\times 10^6$ t, plums 12.01$\times 10^6$ t, cherries 2.73$\times 10^6$ t, and quinces 0.7$\times 10^6$ t. Fruit production in the European continent counts for almost a quarter of world production, with apples 18.36 $\times 10^6$ t, plums 2.73 $\times 10^6$ t, cherries 0.74 $\times 10^6$ t, and quinces 0.03 $\times 10^6$ t, respectively (FAO, 2021).

Trade globalization and world climate diversity have increased the opportunity for consumers to access various fruits despite their natural seasonality.

In analogy with the European market, the plum tree is considered to be between native Mexican species and has a high potential for commercial large-scale production (Bautista-Banos *et al*., 2003; Vargas *et al*., 2017). Even though the area devoted to quince plantations is smaller compared with the plantations of other native fruit plantations, many research studies give evidence that the market accepts products derived from these fruits. Therefore, a better knowledge regarding their chemical composition and polyphenolic profile related to the beneficial biological processes will lead to a greater interest in cultivation and increasing annual production (Arras-Vota *et al*., 2015; Coronado-Reyes *et al*., 2020). In this context, these data and the Principal Component Analysis model can benefit breeders, researchers, and consumers.

The Balkan peninsula is one of the most important regions in the European continent regarding fruit production capacity. Fruit processing has been one of the most distinguished features of civilization from antiquity. Traditional non-alcoholic products such as marmalade, resins, dried fruits, gliko, narden, molasse, and conserved fruits are specific to this region (Topi *et al*., 2022). Meanwhile, fruit juice is increasingly present in the consumer's table, avoiding the fruit shelf-life as the main drawback.

Tree fruits are considered a basic and primary source of primary metabolites such as carbohydrates and lipids (Paredes-Lopez *et al*., 2010; Franco-Vásquez *et al*., 2023) as well as an enormous range of secondary plant metabolites such as phenolic compounds, organic acids, amino acids, and minerals (Wojdyło *et al*., 2013; Morales-Tapia *et al*., 2022). Fruit juices comprise an essential source of nutrients, such as soluble sugars, glycosylated phenolics, coumarins, saponins, alkaloids, organic acids, and pigments (Felix *et al*., 2018; Franco-Vasquez *et al*., 2023). Phytochemicals like polyphenols support well-being and contribute to better health conditions; many exhibit extensive biological activity (Perez-Alonso *et al*., 2015). Phenolic compounds such as flavonoids and anthocyanins have been linked with positive health effects such as cardiovascular and anti-inflammatory effects (Cicerale *et al*., 2010; Felix *et al*., 2017), as well as antioxidant properties contributing against oxidative stress at the cellular level (Swallah *et al*., 2020; Rana *et al*., 2022). Besides the normal mutual properties, fruits own their specific and individual phytochemical ingredients. A distinguished dependence between the fruit phenolics according to variety, climate, and cultivation models has been verified (Carmona-Gomez *et al*., 2018). Five major phenolic groups are found in apple juice: hydroxycinnamic acids, flavan-3-ols, anthocyanins, flavonols, and dihydrochalcones (Wojdylo *et al*., 2008).

However, much work still exists to report information, especially on fruit juices (Nowicka *et al*., 2019; López-Fernández *et al*., 2020). The bioactive compounds extraction rate is influenced by many factors such as matrix, bioactive chemical nature, solvent type, extraction time duration, and extraction temperature. Conventional and non-conventional (assisted-) extraction technologies can be employed (Carmona-Gomezet *et al*., 2018; Franco-Vasquez *et al*., 2023).

This study gave insight into applying principal constituent analysis as a discrimination tool for fruit juice based on tree and phenolic profiles and the correlation between the fruit juice type and phenolic group clusters employed. The molecular ion and primary ions MS2 fragments dataset were produced by chemical screening and determined by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of fruit juices.

2 Material and methods

The current study employed the cold-pressing method to produce fruit juice samples. Fresh fruits in 20 kg were harvested in a small plantation of sweet cherry (*Prunus avium* L.) and quince (*Cydonia oblonga* Mill), respectively. Mixed juice fruit samples were produced by cold-pressing too. The fruits originated from the Korça region, Albania, and were harvested from different plantations to ensure representative samples. Mixed fruit sample 1 is a mixture of red apple, pomegranate (*Punica granatum* L.), and beetroot (*Beta vulgaris subsp. vulgaris*).

In contrast, sample 2 consisted of plum (*Prunus domestica subsp. Domestica R.*) and blueberry

(*Vaccinium myrtillus* L.). In contrast, the other samples are no mixture of fruit juices; sample 3 is cherry juice, and sample 4 is quince juice. During the measurements, there were established three replications per sample.

Chemicals and reagents: Analytical grade reagents were used; Methanol, Sodium Carbonate, and deionized water (Millipore, USA).

2.1 Phenolic profile determined by LC-MS/MS

Recently modernized the LC-MS/MS using the Accela 600 UHPLC system connected to an LTQ OrbiTrap XL mass spectrometer in negative ionization mode has been applied to identify and determine the phenolic compounds. The analytical column used for separation was Syncronis C18 (50 × 2.1 mm, 1.7 μm particle size). The exact UHPLC conditions and MS parameters have been reported (Vasić *et al*., 2019). Tentative identification of some compounds with no available standards was confirmed using previously reported MS fragmentation data (Stojković *et al*., 2020; Llupa *et al*., 2022). Applying the standard procedure described by (Kostić *et al*., 2019; Llupa *et al*., 2022), has been calculated the limits of detection (LOD) and quantification (LOQ).

2.2 Statistical analysis

One of the difficulties in multiple statistics is the problem of visualizing data that has many variables, and more preciously, when there are more than three variables, it is difficult to visualize their relationship. In our case, the chemical compounds in the fruit juice represent the variables, and referring to the dataset in Supplementary Work, there are 35 variables. In datasets with many variables, groups of variables often move together, and in many systems, there are only a few such driving variables governing the system's behavior. The abundance of instrumentation enables us to measure dozens of variables, and the redundancy of dimensions is considered an advantage; replacing a group of variables with a single new variable and Principal Component Analysis (PCA) is a quantitatively rigorous method for achieving this simplification. The method generates a new set of variables, each a linear combination of the original variables.

Principal component analysis was performed on the raw dataset obtained from UHPLC. The statistical and numerical calculation methods and the graphical visualization were performed using built-in functions or M-files in MATLAB software. Before grouping, the dataset was composed of 35 phenolics (variables), while after grouping, it was compressed into six factors. A classical multivariate analysis technique, PCA, is applied to reduce the dimensions. This ability simplifies posterior analysis and demonstrates visualization in fewer dimensions. The PCA was used as the starting procedure before performing multivariate analysis.

3 Results and discussion

3.1 PCA and multivariate statistical analysis

This study has investigated four fruit juices from different fruits produced by the cold-press method. Instrumental analysis using the liquid chromatography separation technique showed a detailed framework for identifying and quantifying present phenolic compounds. In MS data collected from analyzing fruit juice, it was found that various phenolic glucosides are present in all four fruit juices. This finding is supported by the characteristic fragmentation of base peaks through sugar units losing 162 Da (hexosyl) or 132 Da (pentosyl). PCA identifies essential components as a dimensionality reduction tool and visualizes sample clustering. However, other multivariate analysis techniques like cluster analysis (CA) are employed to identify distinct groups in fruit juice based on their phenolic composition.

The extraction rate of the bioactive compounds is influenced by many factors such as matrix, bioactive chemical nature, solvent type, extraction time duration, and extraction temperature. Conventional and non-conventional (assisted-) extraction technologies can be employed (Franco-Vasquez *et al*., 2023).

PCA is a well-established method where dimensionality reduction is sought. Through PCA, dimensions decrease, and it is possible to capture the 'information' contained in a data set with fewer variables than the original one (Wright *et al*., 2009; Jolliffe & Cadima, 2016). Thirty-five phenolic compounds were identified in juice samples through LC-MS/MS screening (Figure 1).

Figure 1 presents the concentration of the chemical compounds in the fruit juices, and it is evident, even from the estimated mean concentration for each chemical compound, that some of them vary in a large range and some of them have less concentration in the juice; however, their contribution in the juice is essential. For this reason, we use PCA to find new variables that are linear combinations of the original ones.

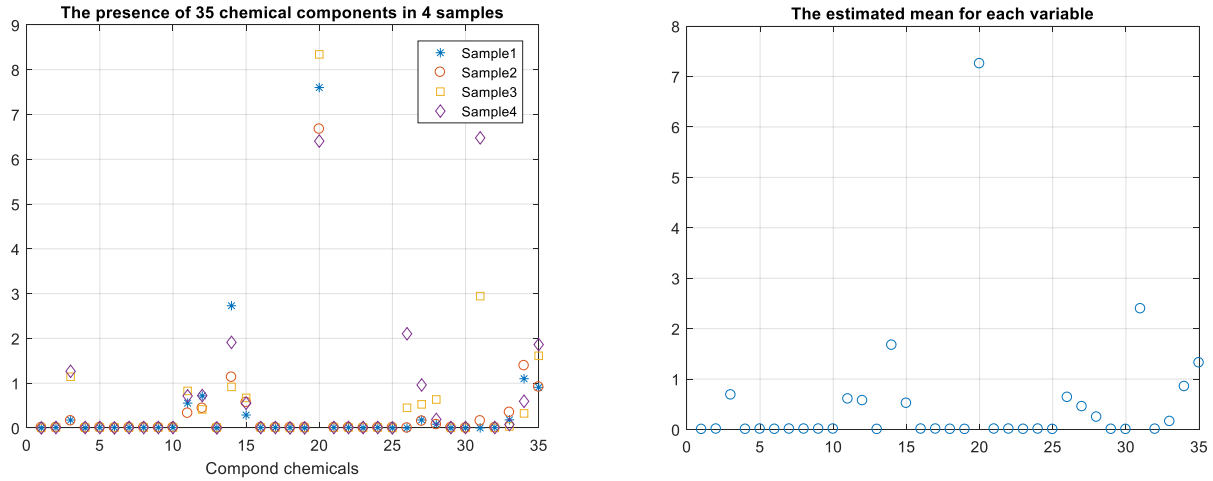


Figure 1. (On the left) The concentration of the chemical components in 4 different samples. (On the right) The estimated mean for each chemical component in 4 different samples.

Considering the chemical concentration of the individual components, the most evident were epicatechin, 5-O-caffeoylquinic acid, and vanillic acid.

A correlation analysis was performed to understand how each component affects the presence of another phenolic compound in every sample, and it is noticeable that the components sorted in the range 25-35 have a significant dependence, referring to the whole dataset obtained from HLPC/MS and given in Supplementary work (*Corr_ compound*).

Among the various methods available to test the normality of all responses, the Kolmogorov-Smirnov test was employed before correlation analysis. Pearson's correlation coefficients are calculated to associate the chemical compounds of the target fruit juices. Principal Component Analysis was applied to the dataset presented in the Supplementary work (PCA) for exploratory purposes.

From the statistical point of view, 35 phenolic compounds represent 35 variables. Despite the magnitude of their concentration dominance, we wanted to reduce the dimensions without losing the effect of any of the components. Principal Component Analysis (PCA) is a classical multivariate analysis technique that enables reducing dimensions, where one variable is a function of the others and, consequently, contains redundant information, and the other case is when the observed variables are correlated. A reduced number of new variables could replace them without losing data.

After applying the PCA analysis on the dataset composed of 35 variables, we perform the score values given in the Supplementary work (*PCA*). Still, for a general overview of the percentage of each component regarding the respective principal components, we present in Table 1 the variability explained for each principal component and Kaiser's criterion.

Table 1: Variability explanation and Kaiser's criterion presented in latent eigenvalues column.

PCA	Explained	Latent_ Eigenvalues
First PCA	86.1181631	11.0744422
Second PCA	9.07913069	1.167538931
Third PCA	4.80270624	0.617608304

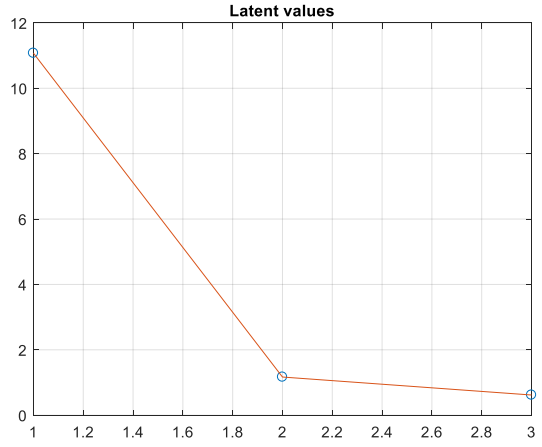


Figure 2. The latent values related to the Kaisler criterion for the PCA analysis were performed with 35 variables.

Different methods accessible to decide the number of components and the percentage of explained variability were utilized in this study. The desirable range to attain a comparably significant percentage of explained variability is 70-90%, and Kaiser's criterion excludes those components whose eigenvalues are smaller than the average eigenvalue or less than one if they are calculated from the correlation matrix (Grane & Jach, 2014). Table 1 presents the variability explained for each principal component and Kaiser's criterion.

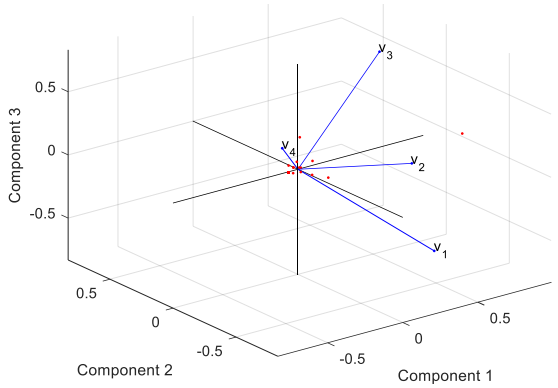


Figure 3. Graphical presentation of the coefficients of the three PCA and the scores (numerical values are in the Supplementary work (PCA)).

The visualization of Kaiser's criterion for this problem is presented in Figure 2 (Grane & Jach, 2014). From it, the first and second principal components are sufficient to represent the entire dataset.

Referring to the percentage of explained variability, the first PCA explains the data at 86.12%, the second PCA explains the data at 9.08%, and the third PCA has the expression percentage of 4.8%. Even if we consider only the first two principal components, the data variability is explained at 95.2%.

From the PCA analysis results and the scores presented in Supplementary Work (PCA), it was visible that the first principal component is remarkably correlated with the primary variable, Vanillic acid (91%). The second principal component strongly correlates with 5-O-caffeoylquinic acid (70%), and the third principal component is represented by epicatechin (82%), indicating that these phenolic compounds can discriminate the fruit juice according to fruit type.

In Figure 3, the samples are projected on the factor space ($PC1 \times PC2 \times PC3$) to perceive the relationship between the grouping acceding to the collected results.

Statistically, the variables number must be smaller than the samples number, meanwhile referring to their chemical structure characteristics, the 35 chemical compounds were grouped into six phenolic groups: hydroxybenzoic acids, flavonols, flavanols, ellagitannins, phenolic acids, and hydroxycinnamic acids. The PCA analysis with six variables and four samples was considered in the following approach.

3.2 *Phenolic profiles screening*

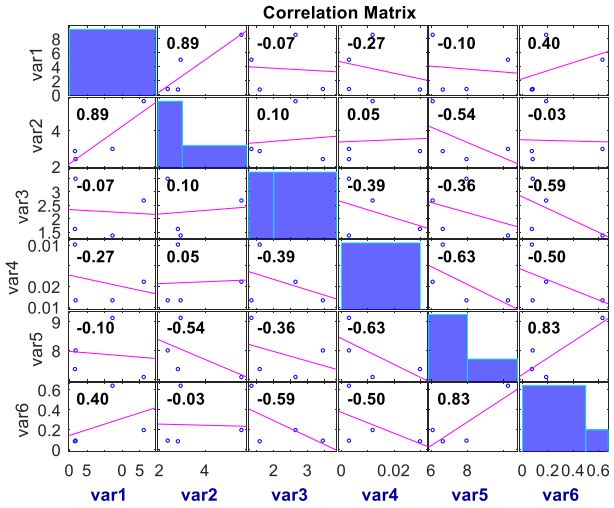
The UHPLC spectral data were confronted to the Mass spectra data bank and were classified in groups: *Flavonoids* with subgroups such as *Flavonols* (Quercetin 3-O-(6"-rhamnosyl)glucoside , Quercetin 3-O-glucoside, Quercetin 3-O-pentoxide, Quercetin 3-O-rhamnoside[a], Phloretin 2'-O-glucoside, Quercetin[a]), *Flavanols* (Procyanidin dimer B type isomer 1, Catechin[a], Epicatechina, Procyanidin dimer B type isomer 2), and *Nonflavonoids or Phenolic Acids* with subgroups such as: *Hydroxy Benzoic acids* (Gallic acid hexoside, Gallic acid hexoside, Protocatechuic acid, *p*-Hydroxybenzoic acid[a], Vanillic acid[a]), and *Hydroxycinnamic acid* (Caffeoylquinic acid hexoside, 3-*O*-Caffeoylquinic acid, Caffeic acid hexoside, 3-*O*-*p*-Coumaroylquinic acid, 4-*O*-Caffeoylquinic acid, 3-*O*-Feruoylquinic acid, Caffeic acid, Coumaric acid hexoside, Ferulic acid hexoside, 5-*O*-Caffeoylquinic acid[a], 5-*O*-*p*-Coumaroylquinic acid, Caffeoylshikimic acid, 4-*O*-Feruoylquinic acid, 4-*O*-*p*-Coumaroylquinic acid, *p*-Coumaric acid[a]), *Ellagitannins* (Ellagic acid hexoside, Ellagic acid rhamnoside, Ellagic acid) (Craizer *et al*. 2009; Jose-Motilva *et al*., 2013). Table 2 are presented the concentrations of the grouped variables according to fruit juice.

The values of each chemical compound given in Table 2 are performed by taking the average of the concentration of each prior chemical compound it is composed regarding the values given in Supplementary work (*Dataset*).

We have applied the PCA analysis using the dataset given in Table 2 and using the percentage of explained variability 87.67% for the first principal component, 8.65% for the second principal component, and Kaisler's criterion of each principal component given respectively in the first and second column of Table 3; we conclude that the first two principal components explain the dataset 96.31%.

Table 2: The concentration of the grouped variables for each fruit juice.

Classification of polyphenols into groups	3 Red (Apple, pomegranate, beetroot)	Plum and blueberry	Cherry	Quince
Hydroxy benzoic acids	0.753986	0.683137	4.937941	8.481109
Flavonols	2.3806	2.82802	2.949846	5.617925
Flavanols	3.466156	1.596088	1.351482	2.656043
Ellagitannins	0.003	0.03	0.003	0.012
Phenolic acids	7.98448	7.331256	9.125492	7.052719
Hydroxycinnamic acids	0.084194	0.077603	0.635642	0.190574



Compound name	Id_Var
Hydroxy benzoic acids	Var1
Flavonols	Var2
Flavanols	Var3
Ellagotannis	Var4
Phenolic Acids	Var5
Hydroxycinnamic acid	Var6

Figure 4. Pearson correlation matrix of the variables. Grouped components variables are coded as Var1-Var6.

Table 3: Variability explanation and Kaiser's criterion presented in latent eigenvalues column.

PCA	Explained	Latent_ Eigenvalues
First PCA	87.66906	15.86419
Second PCA	8.645155	1.564387
Third PCA	3.685785	0.666963

Table 4: Coefficients of the PCA analysis for six variables.

Coeff*	PCA_ 1	PCA_ 2	PCA_ 3
Hydroxy benzoic acids	0.939476	0.165327	0.135877
Flavonols	0.339463	-0.41817	-0.29605
Flavanols	-0.01253	-0.57197	0.816766
Ellagitannins	-0.00073	-0.00315	-0.01439
Phenolic acids	-0.03812	0.657627	0.473544
Hydroxycinnamic acid	0.023199	0.195375	0.048306

The following table is presented the variability explained and Kaisler's criterion for each principal component in the case of grouped variables. Comparing the PCA results obtained in the case with 35 variables with the ones with grouped variables, we obtain a slightly increased percentage of explained variability, and respectively the first principal component explains the original dataset at 87.7%. The second principal component explains the original dataset 9%.

As the analysis with grouped variables has better accuracy, we will analyze the compounds that better correlate for each of the principal components. Table 4 presents the core values of the PCA analysis for the grouped variables. From it, we observe that the first principal component shows a strong correlation with the original variable Hydroxybenzoic acids (0.939476), meaning that it increases only with the increase of Hydroxybenzoic acids, and the second contribution is given from the flavonols (0.339463). The second principal component strongly correlates with the original Phenolic acids (0.657627) and Flavanols (-0.57197). This component increases with increasing of Phenolic acids and decreasing of Flavanols. The third principal component also strongly correlates with Flavanols (0.816766) and Phenolic acids (0.473544).

As a conclusion, the PCA analysis applied to the grouped variables is more efficient because, firstly, even with two principal components is a larger explained variability. Secondly, we have obtained three new variables. The first new variable is hydroxybenzoic acids, the second is the linear combination of phenolic acids and flavanols, and the third is the linear combination of flavanols and phenolic acid.

In order to understand the chemical interaction of the grouped compounds, cross-correlation were performed by revealing flavonols have a strong correlation, respectively 89% with hydroxybenzoic acids and -54% with phenolic acids.

Referring to the correlation results, we have performed a multivariate regression of the Flavonols concerning hydroxybenzoic and phenolic acids. The model constructed with statistical significance *p*-value = 0.000263 is given with the following linear combination

$$y = 7.8761 + 0.3321x_1 - 0.7196x_2 \qquad (1)$$

where x_1 represents the Hydroxy benzoic, x_2 represents the Phenolic acids, and y represents the Flavanols. Figure 5 presents the visual interpretation of the model constructed, and we observe that the nodes, which in our case are the four samples, perfectly fit the model constructed, meaning that using

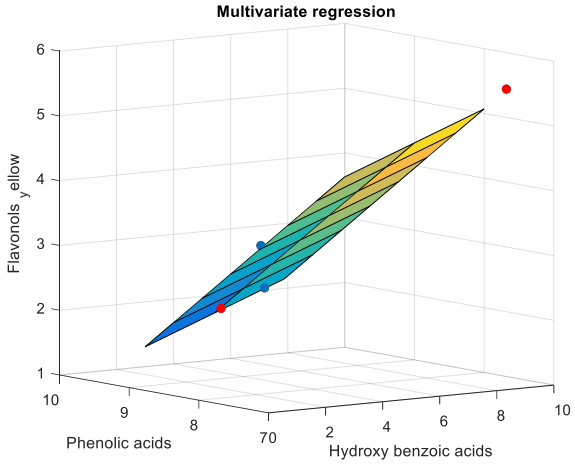


Figure 5. Multivariate linear regression of Flavonols_ yellow concerning Hydroxy benzoic acids and -54% with Phenolic acids.

this model, we can perform confidence intervals for each of the respective compounds.

Multiple linear regression was applied to conclude qualitatively how Hydroxy benzoic and Phenolic acids affect the Flavonols. The magnitude of the coefficients of the equations allows for comparing the relative contribution of each independent variable in the prediction of the dependent variable, and through this model, giving values to the variables x_1 and x_2 could be predicted as the corresponding values of y.

This work aimed to expand the findings collected through LC-MS/MS by applying supervised multivariate statistical data analysis on fruit juice's phenolic profiles. Multivariate analysis refers to a broader set of statistical techniques used to analyze datasets with multiple variables simultaneously. It can be applied to gain further insights and explore various aspects of the sample dataset.

Conclusions

In this work, the PCA was performed with a dataset of 35 phenolic compounds obtained by analytical technique, and using the chemical structure similarity, being selected the compounds into six groups. Comparing the results obtained in both cases, where the first PCA was applied in a dataset of 35 variables and the second PCA was applied with a grouped dataset of 6 variables, it was concluded that the second PCA analysis gives a more significant percentage of variance explanation, respectively 96.31%, with two principal components, which is considered pretty good. It was also distinguished that three out of six compound clusters have the most considerable influence, Hydroxybenzoic acids for the first principal component; meanwhile, Flavanols and Phenolic acids present considerable influence for the second and third principal components.

Moreover, from the PCA performed in the 35 variables dataset, the most significant compounds were vanillic acid for PC1, 5-O-Caffeoylquinic acids, and epicatechin for PC2.

From the performed calculations and the statistical evaluation of the two methods, Principal Component Analysis and Multivariate regression, grouping the compounds related to their chemical structures is an effective procedure.

The study results reveal that PCA of MS data obtained by UHPLC could be used to identify specific phenolics able to discriminate fruit juice according to the fruit type. Moreover, the results showed that flavonols have a strong positive correlation, respectively 89% with Hydroxy benzoic acids and negative 54% with phenolic acids, indicating a correlation between the phenolic groups and fruit type.

References

Arras-Vota A.M.D.G., Lopez-Dia J.C., Bustillos-Nunez G., (2015). The market and transformation of quince in several regions of Mexico and Spain. *International Review of Social Sciences 3*, 10. ISSN 2309-0081.

Bautista-Baños, S.; Díaz-Pérez, J.C.; Barrera-Necha, L.L.; Bravo-Luna, L. (2003) Postharverst study of red-mombin (*Spondias pupurea*) fruit during storage. *Revista Iberoamericana de Tecnología Postcosecha, Hermosillo 5*, 2, 82- 85.

Carmona-Gomez, L.A., Piovesana, A., Norena, C.P.Z., Brandelli, A. (2018). Separation of polyphenolic compounds by ultrafiltration of bordo grape (*Vitis labrusca var. Bordo*) skin extract. *Revista Mexicana de Ingenieria Química 17*, 1, 203-213.

Cicerale, S., Lucas, L., & Keast, R. (2010). Biological activities of phenolic compounds present in virgin olive oil. *International Journal of Molecular Sciences 11*(2), 458-479. doi: 10.3390/ijms11020458.

Coronado-Reyes J.A., Tinoco-Salazar J., Guisa-Morales M.L., Cortes-Penagos C.D.J., Gonzales-Hernandez J.C.(2020). Morphology about the varieties of Mexican pomegranate fruit (*Punica granatum*) and its ripening stage. *Journal of Food Science and Technology (Campinas. 41 (4)* DOI: 10.1590/fst.28620

Crozier, A., Jaganath, I. B., & Clifford, M. N. (2009). Dietary phenolics: chemistry, bioavailability,

and effects on health. *Natural Product Reports 26*(8), 1001. doi: 10.1039/b802662a.

FAO. 2021. Production: Crops and livestock products. In: FAO. Rome. Cited March 2022. https://www.fao.org/faostat/en/#data/QCL.

Felix A.C.S., Alvarez L.D.G., Santana R.A., Valasques Junior G.L., Bezerra M.A., de Olivera Neto N.M., de Olivera Lima E., de Olivera Filho A.A., Franco M., do Nascimento Junior B.B., (2018). Application of experimental designs to evaluate the total phenolics content and antioxidant activity of cashew apple bagasse. *Revista Mexicana de Ingeniería Química 17*, 1, 165-175.

Franco-Vásquez D.C., Carreón-Hidalgo J.P., GómezLinton D.R., Román-Guerrero A., FrancoVásquez A.M., Arreguín-Espinosa R., Alavez S. and Pérez-Flores L.J. (2023). Conventional and non-conventional extraction of functional compounds from flotilla (Escontria chinchilla) fruits and evaluation of their antioxidant activity. *Revista Mexicana de Ingeniería Química 22*(1), Alim2963. https://doi.org/10.24275/rmiq/Alim296

Ghosh, S. & Mitra, J. (2020). Application of Multivariate Statistical Analysis for Quality Control of Food Products. In: *Mathematical and Statistical Application in Food Engineering*, Eds. Sevda S. & Singh A. CRS Press, p. 107.

Grane, A. & A. Jach, (2014). Applications of principal component analysis (PCA) in food science and technology. *Mathematical and Statistical Methods in Food Science and Technology*, First Edition, John Wiley & Sons.

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

Kostić, A., Milinčić, D., Gašić, U., Nedić, N., Stanojević, S., Tešić, Ž., & Pešić, M. (2019). Polyphenolic profile and antioxidant properties of bee-collected pollen from sunflower (*Helianthus Annuus* L.) plant. *LWT - Food Science and Technology 112*, 108224.

Llupa, J., Gašić, U., Brčeski, I., Demertzis, P., Tešević, V., Topi, D. (2022). LC-MS/MS characterization of phenolic compounds in the quince (*Cydonia oblonga* Mill.) and sweet cherry (*Prunus avium* L.) fruit juices. *Agriculture and Forestry 68* (2), 193-205. doi: 10.17707/AgricultForest.68.2.14.

López-Fernández, O., Domínguez, R., Pateiro, M., Munekata, P.S., Rocchetti, G., Lorenzo, J.M. (2020). Determination of polyphenols using liquid chromatography-tandem mass spectrometry technique (LC-MS/MS): A review. *Antioxidants (Basel). 2020 Jun 2; 9*(6), 479. doi: 10.3390/antiox9060479.

Motilva MJ, Serra A, Macià A. (2013). Analysis of food polyphenols by ultra-high-performance liquid chromatography coupled to mass spectrometry: an overview. *Journal of Chromatography A 1292*, 66-82. doi: 10.1016/j.chroma.2013.01.012.

Morales-Tapia A.A., Gonzales-Jimenez F.E., Vivar-Vava G., del Angel-Zumaya J.A., Reyes-Reyes M., Alamilla-Beltran L., Barajas-Zavaleta E., Cooper-Bribiesca B.L., Jimenez-Guzman J. (2022), Use of freeze-drying and convection as drying methods of the xoconostle by-product and the effect on its antioxidant properties. *Revista Mexicana de Ingeniería Química 21*, (2) (2022). http://rmiq.org/iqfvp/Vol21No2.html

Nowicka, P., Wojdyło, A. & Laskowski, P. (2019). Principal component analysis (PCA) of physicochemical compounds' content in different cultivars of peach fruits, including qualification and quantification of sugars and organic acids by HPLC. *European Food Research Technology 245*, 929-938. doi.org/10.1007/s00217-019-03233-z.

Paredes-López, O., Cervantes-Ceja, M. L., Vigna-Pérez, M., & Hernández-Pérez, T. (2010). Berries: Improving human health and healthy aging, and promoting quality life-a review. *Plant Foods and Human Nutrition 65*(3), 299-308. doi:10.1007/s11130-010-0177-1

Perez-Alonso, C., Campos-Montiel, R.G., Morales-Luna, E., Reyes-Munguia, A., Aguirre-Alvarez, G. Pimentel-Gonzalez, D.J. (2015). Stabilization of phenolic compounds from *Opuntia oligacantha* Först by microencapsulation with agave sap (Aguamiel). *Revista Mexicana de Ingenieria Química 14*(3), 579-588.

Rana, A., Samtiya, M., Dhewa, T., Mishra, V., Aluko, R.E. 2022). Health benefits of polyphenols: A concise review. *Journal of Food Biochemistry 46*(10), e14264. doi: 10.1111/jfbc.14264.

Swallah, M. S., Fu, H., Sun, H., Affoh, R., & Yu, H. (2020). The impact of polyphenol on general nutrient metabolism in the monogastric gastrointestinal tract. *Journal of Food Quality*, 1-12, doi: 10.1155/2020/5952834.

Stojković, D., Drakulić, D., Gašić, U., Zengin, G., Stevanović, M., Rajčević, N., & Soković, M. (2020). *Ononis spinose* L., an edible and medicinal plant: UHPLCLTQ-Orbitrap/MS chemical and biological activities of the herbal extract. *Food & Function 11*, 7138-7151.

Topi, D., Guclu, G., Kelebek H. & Selli, S. (2022). LC_ DAD-ESI-MS/MS characterization of phenolic compounds in wines from Vitis vinifera '*Shesh i bardhë*' and '*Vlosh*' cultivars. *Journal of Food Processing and Preservation 46*, e16157. https://doi.org/10.1111/jfpp.16157.

Vargas, A. S., Juárez-López, P., López-Martínez, V., Flores, L. J. P., Sánchez, D. G., & Alia-Tejacal, I. (2017). Botânica e fisiologia/ botany and physiology antioxidant activity and physicochemical parameters in "Cuernavaqueña" Mexican Plum (*Spondias purpurea L.*) at different ripening stages. *Revista Brasileira de Fruticultura 39*(4). doi: 10.1590/0100-29452017787

Vasić, V., Gašić, U., Stanković, D., Lušić, D., Vukić-Lušić, D., Milojković-Opsenica, D., *et al*. (2019). Towards better quality criteria of European honeydew honey: Phenolic profile and antioxidant capacity. Food Chemistry, 274, 629-641.

Wojdylo, A., Osmianski, J., Laskowski, P. (2008). Polyphenolic compounds and antioxidant activity of new and old apple varieties. *Journal of Agricultural and Food Chemistry 56*(15), 6520-6530. doi: 10.1021/jf800510j.

Wojdylo, A., Osmianski, J., Bielicki, P. (2013). Polyphenolic composition, antioxidant activity, and polyphenol oxidase (PPO) activity of quince (*Cydonia oblonga* Miller) varieties. *Journal of Agricultural and Food Chemistry 61*(11), 2762-2772. doi: 10.1021/jf304969b.

Wright, J., Peng, Y., Ma, Y., Ganesh, A., Rao, S. (2009). Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization. In Proceedings. of Neural Information Processing Systems 2009 (NIPS 2009), Vancouver, BC, Canada, 7-10 December 2009. See http://papers.nips.cc/paper/3704-robust-principal-component-analysis-exact-recovery-of-corrupted-low-rank-matrices-via-convex-optimization.pdf.