OPEN ACCESS · SENA · REVISTA COLOMBIANA DE INVESTIGACIONES AGROINDUSTRIALES

Development of analytical methodologies for the quality control of liquors using nearby infrared spectroscopy and multivariate analysis

Desarrollo de metodologías analíticas para el control de calidad de licores mediante espectroscopia infrarroja cercana y análisis multivariante

José Libardo Tapiero Cuellar[1] iD

Alexis Valle Mora[2] iD

[1]Universidad Nacional Sede Palmira. Colombia ✉jltapieroc@unal.edu.co

[2]Centro de Comercio y Servicios, Regional Atlántico ✉avalle@sena.edu.co

Resumen La aplicación de la espectroscopia de infrarrojo cercano NIR y el análisis multivariante es una técnica de análisis rápido en la cuantificación del contenido de metanol en licores de distribución comercial (Aguardiente), para identificar si estos licores superan los límites de metanol establecidos en la norma *AOAC 972.11* (100 mg/dm³) (*AOAC 972.11*, 1973). En una primera etapa, se realizó la aplicación de un diseño de mezclas ternarias metanol-etanol-agua (MDOE) con el fin de reducir los efectos de correlación entre las variables, además para lograr modelos multivariados que incluyan la variabilidad importante para ampliar el rango de calibración que actúa en las condiciones del proceso y en los sitios de comercialización. La segunda etapa estuvo asociada al registro de los espectros NIR de las muestras preparadas, para obtener una matriz de datos, a partir de la cual, se desarrollan los modelos multivariados utilizando diferentes algoritmos. Los resultados obtenidos demuestran que es posible desarrollar un método de análisis rápido para la cuantificación de metanol en muestras de aguardiente, utilizando la espectroscopia NIR y los modelos multivariantes construidos, en los que se obtuvo un Error Cuadrático Medio de Predicción (RMSEP) de 0,7766 % y un coeficiente de correlación (R^2) de 0,9553.

Palabras clave: adulteración, aguardiente, espectroscopia de infrarrojo cercano NIR, quimiometría, análisis multivariado.

Abstract The application of NIR near-infrared spectroscopy and multivariate analysis as a rapid analysis technique in the quantification of the content of methanol in commercially distributed liqueurs (Aguardiente), to identify if these liqueurs exceed the methanol limits established in the standard. *AOAC 972.11* (100 mg/dm³) (*AOAC 972.11*, 1973). For this, in a first stage, the application of a design of ternary mixtures methanol-ethanol-water (MDOE) was carried out in order to reduce correlation effects between the variables, in addition to achieving multivariate models that include variability important by expanding the calibration range that act under process conditions and in commercialization sites. The second stage was associated with the recording of the NIR spectra of the prepared samples, to obtain a data matrix, from which the multivariate models are developed using different algorithms. The results obtained showed that it is possible to develop a rapid analysis method for the quantification of methanol in Aguardiente samples, using NIR spectroscopy and the multivariate models constructed, in which a Mean Square Prediction Error (RMSEP) of 0.7766 % was obtained and correlation coefficient (R^2) of 0.9553.

Keywords: adulteration, aguardiente, NIR near-infrared spectroscopy, chemometry, multivariate analysis.

Introduction

Brandy as distilled alcoholic beverage comprises 98 percent by water and ethanol, (Flórez, 2009; Zuleta & Jaramillo, 2000). The difference corresponds to inorganic elements of the total mass of the product. The regulations for its consumption and distribution are governed by each country where it is marketed. During recent years the adulterations have grown considerably, becoming an expensive and difficult problem to control. Generally, adulterations are performed with methanol, resulting from blindness to death of the consumer.

In Colombia, the adulteration of products has been repeatedly presented in various markets, specifically food, medicine, fuel and liquor. Its contraband is a constant practice of groups that are dedicated to its commercialization and even with the endorsement of the final buyer for its low prices. This has led to the country having a greater tax loss for illegal alcohol, about US $ 468 million a year (EH, 2015). Falsifying or adulterating liquor has become a very lucrative business that is increasing and expanding. Access to the raw materials necessary for the creation of the product is very simple and by not paying taxes, it is possible to sell it at a price below the original, which is the main sign of a market affectation. And although it is true that many people prefer the original product already for the time, it is very difficult to identify at first glance an adulterated liquor from a commercial one.

Gas chromatography has been used as a reference method to determine and classify adulterated liquors (Dong *et al*., 2014; Ding *et al*., 2016). However, this technique is expensive, requires careful sample preparation, long response times, and is also destructive in nature. For this reason, in recent years, spectroscopic techniques have been proposed for the quality control of liqueurs. Due to their various characteristics, within these techniques, near-infrared spectroscopy (NIR) can be cited, used for the classification of various alcoholic beverages (Chen *et al*., 2014). This technique is non-destructive, has good reproducibility and precision, short response time, offers the advantages of simplicity of sample presentation and low cost. (Yu *et al*., 2005; Hu *et al*., 2018).

Advances in NIR instrumentation in combination with the development of chemometrics through multivariate analysis methods make this technology ideal for large volumes, rapid detection and characterization of food components down to parts per billion (ppb) levels (Ghasemi-Varnamkhasti & Forina, 2014). Given the use in quality applications and process control, the food industry has become familiar with the technology and its potential to expand into monitoring food adulteration.

NIR near-infrared spectroscopy coupled to multivariable PLS and PCR analysis methods is proposed to determine the concentration of methanol in liqueurs in a fast way, so that it can be implemented at commercial points and evaluate its possible adulteration.

Methodology

Samples were considered two samples of Firewater supplied and certified under the NTC-4118: 1997 Seventh updated by the Liquor Factory Antioquia, (FLA) Colombia. The methanol (CAS: 67-56-1, lot: X08C03, JT BAKER) and water ultrapure (water purification system type I, Thermo Smart2Pure 3 UV) used in the preparation of the mixtures.

The mixtures were prepared in the Food and Beverage laboratory of the SENA Trade and Service Center, Regional Atlántico. The samples were prepared in duplicate, Table 1.

Table 1

Percentage ratio of Aguardiente, ultrapure water and methanol, established by the design of experiments

	COMPONENTS OF THE MIX						
Mix	**Aguardiente** (%w/w)	**Water** (%w/w)	**Methanol** (%w/w)	**Mix**	**Aguardiente** (%w/w)	**Water** (%w/w)	**Methanol** (%w/w)
A-01	99,9440	0,0000	0,0560	A-26	97,4176	0,0000	2,5824
A-02	99,9294	0,0000	0,0706	A-27	96,6698	0,0000	3,3302
A-03	99,6539	0,0000	0,3461	A-28	95,8696	0,0000	4,1304
A-04	99,7086	0,0000	0,2914	A-29	94,9725	0,0000	5,0275
A-05	99,5118	0,0000	0,4882	A-30	92,4831	0,0000	7,5169
A-06	99,9772	0,0000	0,0228	A-31	91,1211	0,0000	8,8789
A-07	99,3172	0,0000	0,6828	A-32	89,9279	0,0000	10,0721
A-08	98,9859	0,0000	1,0141	A-33	87,4269	0,0000	12,5731
A-09	97,9367	0,0000	2,0633	A-34	96,5329	0,1444	3,3227
A-10	95,0760	0,0000	4,9240	A-35	80,4110	1,6019	17,9870
A-11	90,9091	0,0000	9,0909	A-36	91,1785	2,1154	6,7061
A-12	99,8409	0,0000	0,1591	A-37	93,8604	5,4344	0,7052
A-13	99,9476	0,0000	0,0524	A-38	88,9109	6,9864	4,1027
A-14	99,8868	0,0000	0,1132	A-39	79,9673	8,3930	11,6397
A-15	99,8004	0,0000	0,1996	A-40	84,1846	10,2271	5,5883
A-16	99,7621	0,0000	0,2379	A-41	74,4924	10,9420	14,5656
A-17	99,6875	0,0000	0,3125	A-42	70,2617	12,4239	17,3144
A-18	99,6391	0,0000	0,3609	A-43	86,5044	13,2305	0,2651
A-19	99,5795	0,0000	0,4205	A-44	65,1615	14,9210	19,9176
A-20	99,4743	0,0000	0,5257	A-45	81,9674	15,1279	2,9047
A-21	99,2653	0,0000	0,7347	A-46	76,4803	16,4758	7,0440
A-22	99,0419	0,0000	0,9581	A-47	70,2843	20,5133	9,2023
A-23	98,9022	0,0000	1,0978	A-48	89,8303	4,2589	5,9108
A-24	98,6726	0,0000	1,3274	A-49	85,7456	9,4469	4,8075
A-25	98,2445	0,0000	1,7555	A-50	93,5144	5,8826	0,6030

Note. Own work.

Experimental Design

To develop analysis establishing the considered Design Experiments of Mixture Ternaries (DEMT). Next, the NIR spectra were obtained to create the data matrix; build the multivariate calibration models and external validation of the built models. Finally, the application of the model made to commercial samples is carried out. This is performed to reduce correlation effects between the variables and achieve multivariate models including variability allowing wider range calibration and the simulation of conditions of adulteration in samples commercialized.

For the case, the concentration of the Aguardiente samples was carried out with different contents of methanol and water to generate adulteration (ternary mixtures

of Liquor/methanol/ water, L/M/A). This condition was achieved by applying the design of ternary mixtures with the restriction design approach using the statistical software JMP 13 64-bit *free trial* version. For the investigation,

50 mixtures were established in the range from 0 to 100% for liquor, from 0 to 20% for methanol and water, expressed in % v/v, Figure 1.
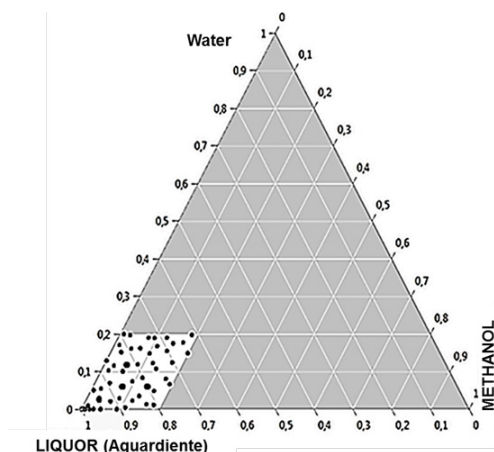

Figure 1

Ternary diagram of the design of mixtures by restrictions

Note. Own work.

Sample Analysis

The NIR readings mixtures and study samples are performed in one Spectrophotometer Ocean Optics NIRQuest 256-2.5, using the accessory liquid samples CUV Cuvette Holder. Where an external VIS-NIR light source with tungsten halogen lighting type is used. The detector is of InGaAs with size Slit 025 microns, the spectral range of 900 to 2500 nm (4000 to 11500 cm^{-1}), integration time 1 ms to 400 ms. Using *OceanView*® 1.6 .5 software and it calibrated, the transmittances were performed, in wave number (cm^{-1}), with 4 *boxcart* and an average of 50 scans per spectrum. The NIR reading of the samples were made using 2 mL quartz cuvettes and were acquired in a time of 5 seconds. The spectrums of the sample are get triplicate.

For the development of the calibration models, the prepared samples were randomly classified into two sets: a calibration set (70% of the samples-70 samples), and a validation

or prediction set (30% of the samples-30 samples) used in external validation. This random classification was made by simple random sampling using the MS Excel 2010 platform. Different mathematical pretreatments were applied to the NIR spectra obtained, in order to increase the signal/noise ratio. The applied pretreatment methods were: baseline correction using *baseline offset and linear baseline correction*, derivation with Savitzky-Golay filter with second-order polynomial, orthogonal signal correction (OSC) and SNV.

The construction of the multivariate calibration models was done using the samples of the calibration set, applying the PLS algorithm and using the method of internal cross-validation with 20 random segments. The qualitative analysis of spectra was done by applying the PCA, which allows reducing the dimensionality of the data to a small number of main components, making it possible to pool samples to determine whether

Development of analytical methodologies for the quality
control of liquors using nearby infrared spectroscopy
and multivariate analysis

they are adulterated with methanol or not. The validation of the multivariate models was done by external validation, using the data set not included in the calibration (validation set).

Statistical Analysis

The statistical parameters used to evaluate quality models in calibration sets and external validation, were the correlation coefficient (R^2) and the mean error standard Quadratic, RMSE (Root Mean Square Calibration, RMSEC; Quadratic Prediction Error (RMSEP) (Martens & Naes, 1992). The RMSE is a measure of the prediction error expressed in units of the original measurement. The calculations of the corresponding statistical parameters and graphs were performed with the 30-day trial version of The Unscrambler X 10.4 Chemometric software (free trial version–30 days).

Results and discussion

The NIR spectra for the certified Aguardiente sample and the mixtures are presented in Figure 2. where methanol is used as an adulterating agent in alcoholic beverages. Spectral curves show the spectral difference between methanol, ethanol, water, and brandy.

Between methanol and ethanol given spectral difference very small between 5000 cm^{-1} and 5800 cm^{-1}, being very similar to the nature of the alcohol compound (-OH). Furthermore, the structure of an alcohol is similar to that of water since an alcohol comes from the substitution of one of the hydrogens in the water for an alkyl group. Water has a spectrum very similar to that of brandy, differing from these in the region of 7200 cm^{-1} to 10000 cm^{-1}.
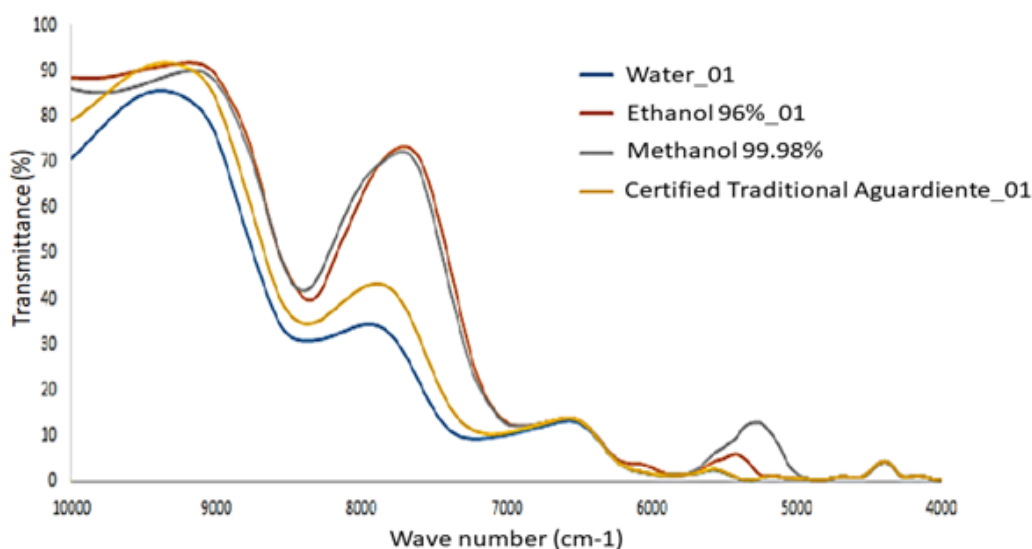

Figure 2

NIR spectra of the pure components of the mixtures
Note. Own work.

The overlays of the spectra of the evaluated samples show similarity in their path, where the region with the greatest variation is achieved between 7000 cm^{-1} and 8500 cm^{-1}, Figure 3.

In this case there is no significant variation between methanol and ethanol indicating that the value is not given by changes in the concentration of the two compounds.
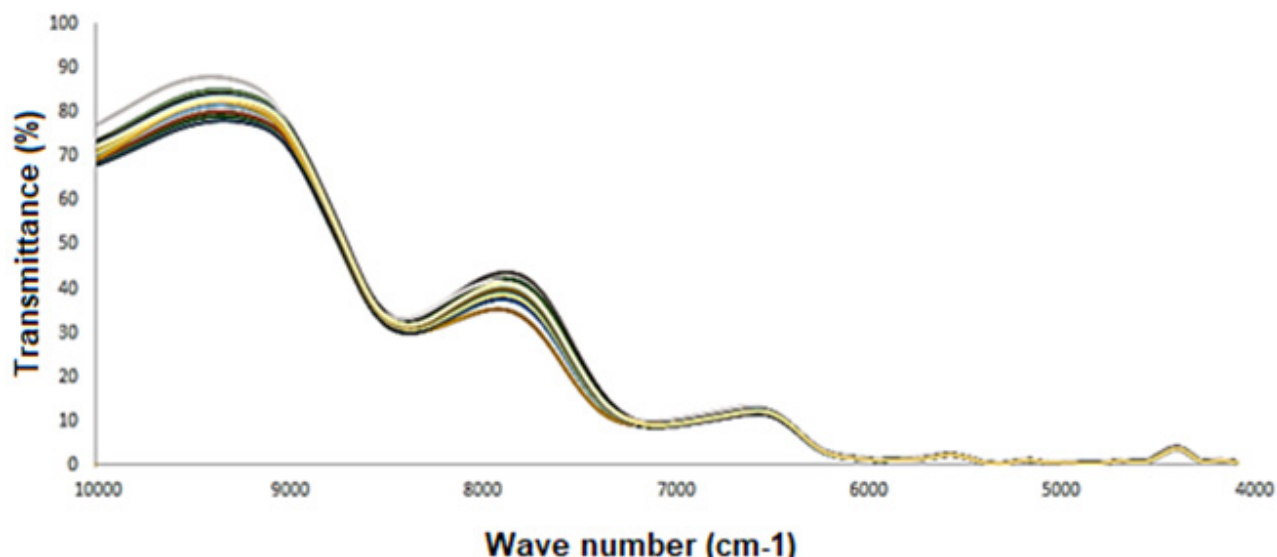
Figure 3

Superimposed NIR spectra of the prepared mixtures

Note. Own work.

With the intention of increasing the signal-to-noise ratio, the first order Savitzky-Golay derivative and baseline correction pretreatments were applied. The first derivative is used for the exact determination of the absorption maximum, especially when dealing with wide peaks. Figure 4 shows the NIR spectra of the mixtures after applying the first derivative. It was observed that in the range between 7500 cm^{-1} and 9000 cm^{-1} is given the greatest variation, showing one minimum displacement spectra derivatives along the horizontal axis.
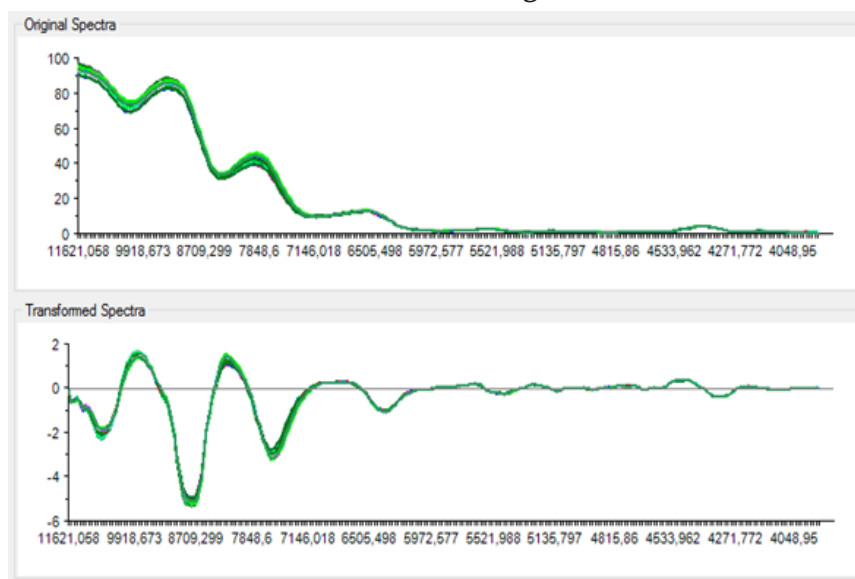

Figure 4

Transformation of the spectra with the first derivative of Savitzky-Golay, 7 window points and 2nd polynomial order

Note. Own work.

Development of analytical methodologies for the quality
control of liquors using nearby infrared spectroscopy
and multivariate analysis

For development of the multivariate model, one initially performed Principal Component Analysis (PCA), which indicated by g to FICO score the relationship between the study samples and by graphic loading the spectral region which presents the greatest variation, Figure 5. The *score* indicates the main source of variability with respect to the first two main components; marked according to their methanol content. The first factor shows 80% of the total variability and the second factor contains 16%. The variation that these two factors present has no relation to the methanol content in the study samples.
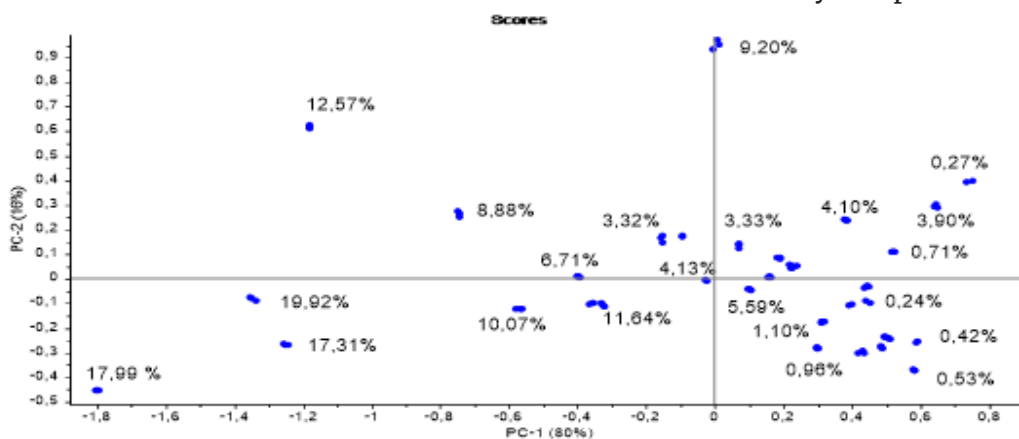

Figure 5

A CP score graph of the first two factors with the samples marked according to their methanol content

Note. Own work.

Figure 6 shows the *loading* graph of the first factor. Spectral region that shows the greatest variation is between 7000 cm⁻¹ and 8500 cm⁻¹. Coinciding what reported in the visual analysis of NIR spectra of Figure 3.

The construction of the calibration models was performed using the Partial Least Squares algorithm (PLS algorithm) where the spectral ranges and the factors studied were considered.
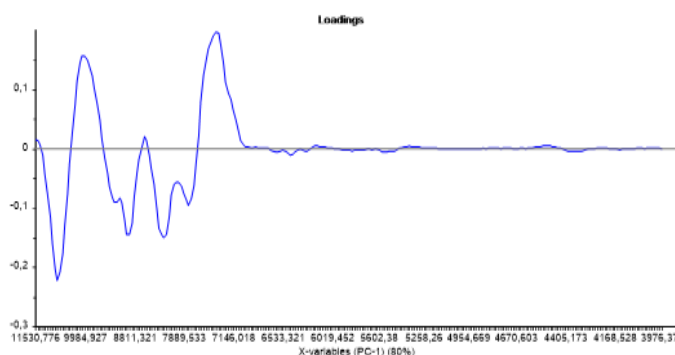
The perishability of the built models was tested against brandy samples added with methanol and water in different proportions of known composition (validation set). Table 2. shows the values of the Minimum Cross–Validation Quadratic Value. (MCVQV) (Obtained in the internal validation), the Mean Square Error of Prediction (MSEP) (obtained in the external validation) and R^2 obtained for each of the models constructed.

Figure 6

Graphic loading of PCA loading graph of the first factor

Note. Own work.

Table 2

RMSECV, RMSEP and R² values obtained in the multivariate models constructed

Spectral ranges [1/cm]	Parameters	Pre treatments		
		1° Derivada (1)	BC (2)	BC/1°Derivada
4000–11600	F*	5	5	5
	R²	0,9808	0,9794	0,9804
	MCVQV	0,7792	0,8038	0,7624
	MSEP	1,0786	7,3837	1,0786
7000–1000	F*	3	3	3
	R²	0,9697	0,9535	0,9684
	MCVQV	0,9790	1,1995	0,9802
	MSEP	2,5190	3,7073	2,5190
7000–8500	F*	4	4	4
	R²	0,9740	0,9716	0,9741
	MCVQV	0,8940	0,9328	0,8930
	MSEP	0,7766	1,5743	0,7767

*Note. * F: Factors. (1) 1st Der: first derivative with Savitzky–Golay filter and 3 window points. (2) BC: Baseline Offset Correction. Own work.*

When applying the Baseline Offset Correction pretreatment followed by the first derivative, the results are very similar to those obtained only with the first derivative, that is, the BC does not contribute significantly to the prediction error. These models present an RMSEP value of 0.776 and an R² correlation of 0.974, Table 2. Therefore, it is established as the multivariate model with the best predictive capacity, the one in which it was applied as the first derivative pretreatment, using the spectral range of 7000 cm⁻¹ to 8500 cm⁻¹ and 4 factors. Figures 7 and 8 show the calibration curve and the reference graph against the preset value for the model.
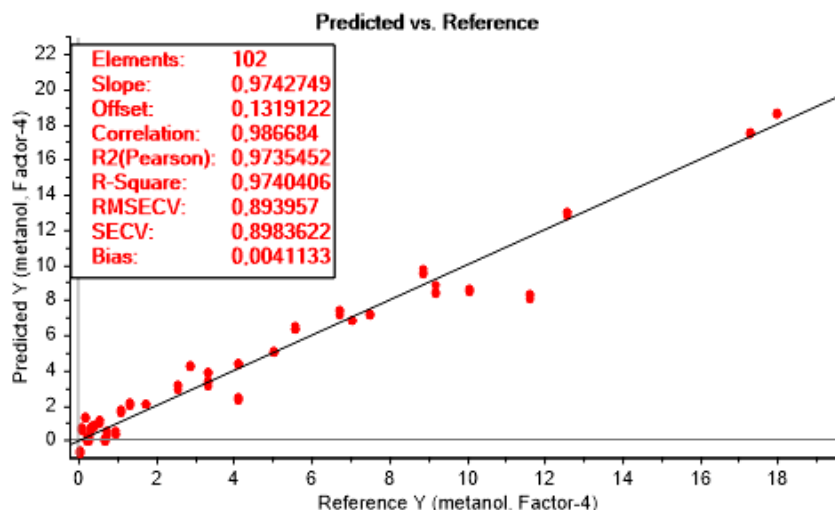

Figure 7

Calibration curve of the multivariate model with better predictability
Note. Own work.

Development of analytical methodologies for the quality
control of liquors using nearby infrared spectroscopy
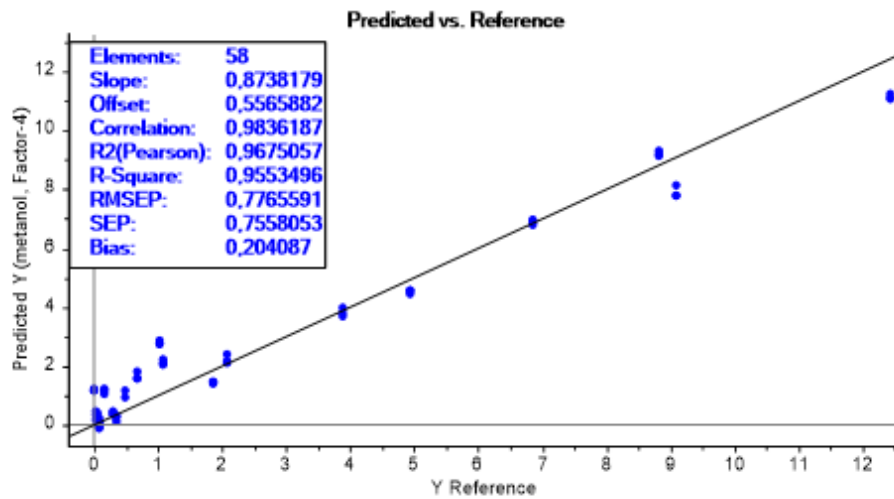and multivariate analysis

Figure 8

Graph Reference opposite to the preset value established of the multivariate model with improving prediction

Note. Own work.

Figure 9 shows the comparison of the values of methanol in Aguardiente pre-established by NIR model and the reference values for the samples used in the validation of the model.
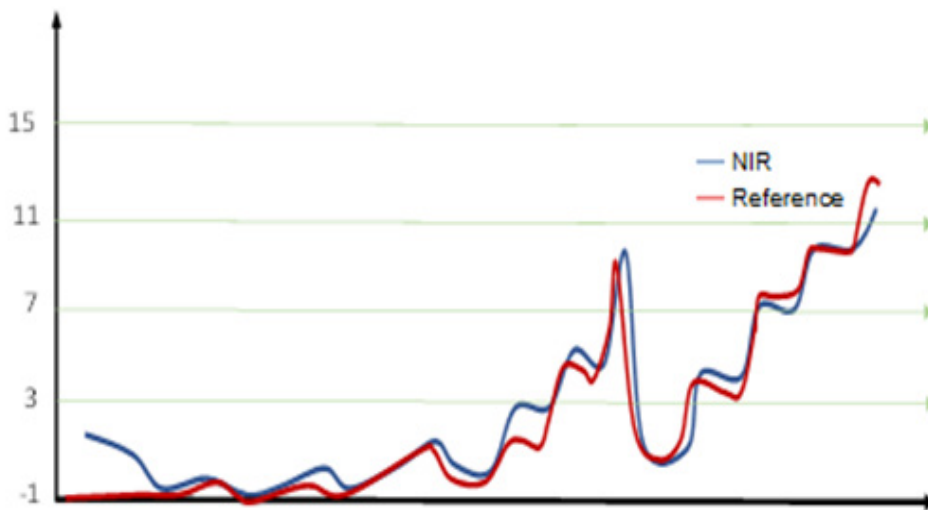

Figure 9

Curve comparison between the values pre-established by NIR model and the reference values for methanol Aguardiente

Note. Own work.

The results obtained show the correlation between the reference values and the NIR measurements in the ranges of methanol studied. Although the region between 7000 cm^{-1} and 8500 cm^{-1} is a region where methanol and ethanol are very similar; the multivariate model allows finding predictive values with low levels of uncertainty.

Conclusions

The application of NIR spectroscopy and chemometry is a rapid analytical method suitable for determining the methanol content of spirits. The design of ternary mixtures used in preparing samples Aguardiente additivated with methanol and water, permit and develop multivariate models with good predictive capability of the methanol content in brandy multivariate model with better predictability was where applied first derived as pretreatment. Generating a quick alternative to determine the alterations of these liqueurs in our country.

References

Chen, H., Tan, C., Wu, T., Wang, L. & Zhu, W. (2014). Discrimination between authentic and adulterated liquors by near-infrared spectroscopy and ensemble classification. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 130, 245–249. https://doi.org/10.1016/j.saa.2014.03.091

Ding, X., Wu, C., Huang, J. & Zhou, R. (2016). Characterization of interphase volatile compounds in Chinese Luzhou-flavor liquor fermentation cellar analyzed by head space-solid phase micro extraction coupled with gas chromatography mass spectrometry (HS-SPME / GC / MS). *LWT - Food Science and Technology*, 66, 124–133. https://doi.org/10.1016/j.lwt.2015.10.024

Dong, D., Zheng, W., Wang, W., Zhao, X., Jiao, L. & Zhao, C. (2014). A new volatiles-based differentiation method of Chinese spirits using longpath gas-phase infrared spectroscopy. *Food Chemistry*, 155, 45–49. https://doi.org/10.1016/j.foodchem.2014.01.025

EH (09 de febrero de 2015). Alcohol ilegal y contrabando, un negocio redondo en América Latina. *El Heraldo.* https://www.elheraldo.co/economia/alcohol-ilegal-y-contrabando-un-negocio-redondo-en-america-latina-183583

Flórez, B. R.A. (2009). Sugar cane and brandy in the sovereign state of Bolívar, 1857-1886. *Cuadernos de Desarrollo Rural*, 6 (63). http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0122-14502009000200003

Ghasemi-Varnamkhasti, M. & Forina, M. (2014). NIR spectroscopy coupled with multivariate computational tools for qualitative characterization of the aging of beer. *Computers and Electronics in Agriculture*, 100, 34–40. https://doi.org/10.1016/j.compag.2013.10.001

Hu, L., Yin, C., Ma, S. & Liu, Z. (2018). Rapid detection of three quality parameters and classification of wine based on Vis-NIR spectroscopy with wavelength selection by ACO and CARS algorithms. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 205, 574–581. https://doi.org/10.1016/j.saa.2018.07.054

Martens, H. & Naes, T. (1992). *Multivariate calibration.* John Wiley & Sons.

Yu, H., Ying, Y., Fu, X., & Lu, H. (2005). PLS-NIR determination of five parameters in different types of Chinese rice wine. *Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality*, 5996. https://doi.org/10.1117/12.630574

Zuleta L. A. & Jaramillo L. (2000). *The liquor industry in Colombia.* FEDESARROLLO. https://repository.fedesarrollo.org.co/bitstream/handle/11445/986/Repor_Octubre_2000_Zuleta_y_Jaramillo.pdf?sequence=1isAllowed=y