

**V CONGRESO LATINOAMERICANO DE PROFESIONALES Y
ESTUDIANTES DE CIENCIA Y TECNOLOGIA DE LOS ALIMENTOS**

CLASIFICACIÓN MULTIVARIANTE DE RONES CON EMPLEO DE LA ESPECTROSCOPIA INFRARROJA CERCANA

Autores

Estrella P. Zayas*, Kurt Varmuza**, Isneri Talavera***, Magdalena Lorenzo**** y María del Carmen Vasallo****

*Facultad de Ingeniería Química. Instituto Superior Politécnico "José Antonio Echeverría" . Cuba.

** Universidad Tecnológica de Viena. Austria.

*** Centro de Aplicaciones en Tecnologías de Avanzada. (CENATAV). Cuba.

**** Centro de Referencia de Alcoholes y Bebidas. ICIDCA. Cuba.

estrella@quimica.cujae.edu.cu

2010

Introducción

La industria de bebidas cubanas goza de gran prestigio y en especial el ron cubano se vende en el mundo entero con gran aceptación pero de forma general no existe una metodología científicamente fundamentada que permita la clasificación de los diferentes rones que se fabrican en el mundo o la detección de adulteraciones en rones. En el presente trabajo se analizan muestras de rones cubanos y no cubanos mediante la espectroscopia infrarroja cercana (NIR) y se aplica la Quimiometría en especial el análisis exploratorio y clasificadorio de datos.

Objetivo

Analizar de forma preliminar si la espectroscopia infrarroja cercana permite la diferenciación de rones cubanos de los no cubanos.

Materiales y Métodos

Se analizaron 29 muestras de ron, 24 de rones cubanos y 5 no cubanos, obteniéndose 601 lecturas de absorbancia en el intervalo de 1100 a 2300 nm en un Espectrómetro Brimrose Luminar 5030 NIR con un filtro óptico acústico ajustable, una sonda de fibra óptica, paso óptico 0,5 nm a una temperatura de 23°C. Se midieron 3 espectros uno después de otro a dos minutos entre lectura. Se utilizó el software cubano Quimiometrix para obtener modelos de clasificación multivariante por los métodos K vecinos más cercanos (KNN), Modelación suave independiente por analogía de clases (SIMCA) y Análisis lineal discriminante por mínimos cuadrados (PLS-DA) todas las clasificaciones se hicieron teniendo en cuenta dos clases, cubano (c) y no cubano (nc). Previamente se realizó el Análisis de Componentes Principales para analizar el agrupamiento de las muestras y reconocer patrones.

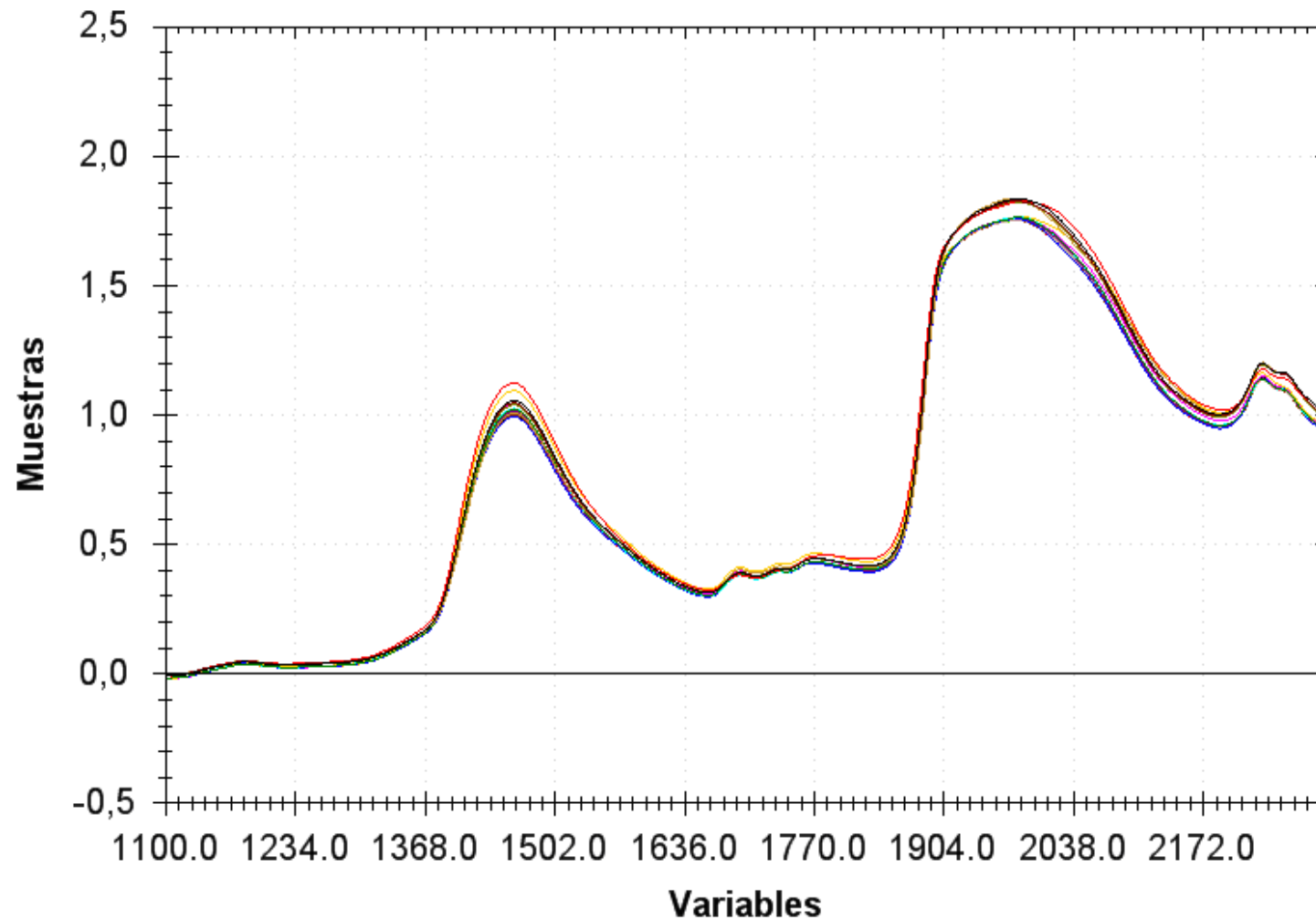
Conjuntos de calibración y validación

Calibración	Validación
No cubanos	
02_005-m3	02_004-m3
02_006-m3	
02_007-m3	
01_004-m3	
Cubanos	
03_V01-m3	03_V21-m3
03_V02-m3	03_V23-m3
03_V03-m3	03_V24-m3
03_V04-m3	03_V25-m3
03_V05-m3	03_V26-m3

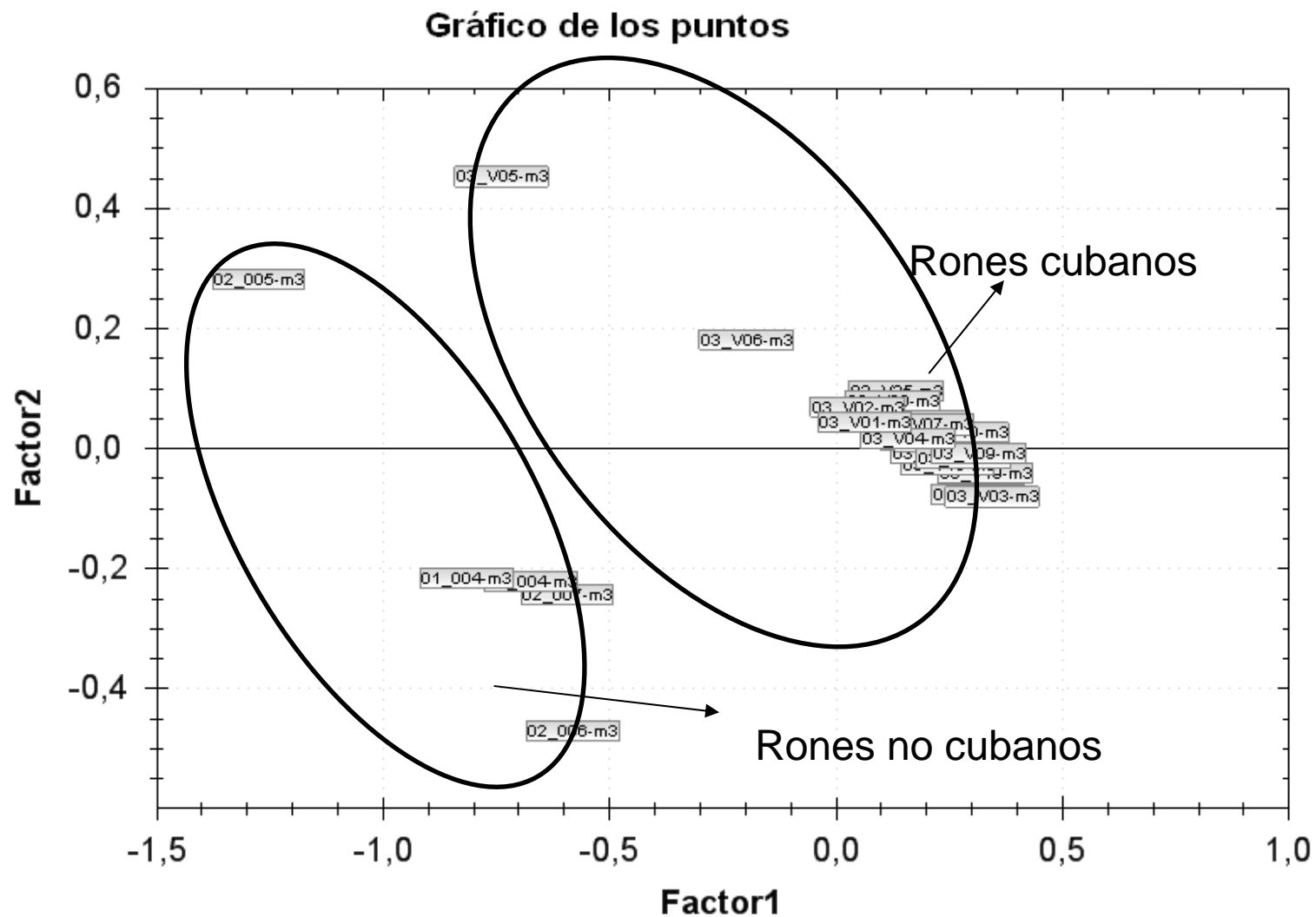
Calibración	Validación
03_V06-m3	
03_V07-m3	
03_V08-m3	
03_V09-m3	
03_V10-m3	
03_V11-m3	
03_V13-m3	
03_V14-m3	
03_V15-m3	
03_V16-m3	
03_V17-m3	
03_V18-m3	
03_V19-m3	
03_V20-m3	

Resultados y Discusión

Espectros de las muestras



De espectros semejantes debe extraerse la información vital para diferenciar los roncs cubanos de los no cubanos.



Hay dos grupos bien definidos a la izquierda agrupan los rones no cubanos y a la derecha están los cubanos El factor 1 el que da la procedencia.

Resultados y Discusión

Análisis de varianza del Análisis de Componentes Principales

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Varianza	0,187	0,026	0,00516	0,0006	0,0004
%	85,05	11,98	2,3411	0,2724	0,18
% acumulativo	85,05	97,03	99,37	99,64	99,82
PRESSCal	0,9228	0,1835	0,0390	0,0222	0,011
PRESSVal	1,3167	0,4601	0,1506	0,14	0,073
RMSECV	1,1475	0,6783	0,3881	0,3741	0,2701

De la tabla anterior se observa como tres componentes describen el 99,4% de la varianza y que con solo dos se abarca hasta el 97%.

Clasificación PLS-DA

En este método no presenta ningún error ni en el entrenamiento ni en la validación.

Resultados y Discusión

Clasificación KNN

Se trabajó con la distancia Euclideana con procesamiento centrado por la media con un vecino más cercano como máximo en la comparación que al final fue el óptimo. Las clases fueron (c) y (nc).

ERRORES DE CLASIFICACIÓN (ENTRENAMIENTO DEL MODELO)		
	PREDICHO COMO nc	PREDICHO COMO c
CLASE nc	3	1
CLASE c	0	19

ERRORES DE PREDICCIÓN (VALIDACIÓN DEL MODELO)		
	PREDICHO COMO nc	PREDICHO COMO c
CLASE nc	1	0
CLASE c	0	5

Existe un error de clasificación cuando se construye el modelo pero sin embargo al clasificar con el conjunto de validación se obtuvo un 100% de clasificación correcta.

Resultados y Discusión

Clasificación SIMCA

ERRORES DE CLASIFICACIÓN (ENTRENAMIENTO DEL MODELO)			
	PREDICHO COMO nc	PREDICHO COMO c	NO CLASIFICADAS
CLASE nc	4	0	0
CLASE c	0	19	0

ERRORES DE PREDICCIÓN (VALIDACIÓN DEL MODELO)			
	PREDICHO COMO nc	PREDICHO COMO c	NO CLASIFICADAS
CLASE nc	1	0	0
CLASE c	0	5	0

No existen errores ni en la clasificación con el conjunto de entrenamiento ni en la predicción con el conjunto de validación

Conclusiones

- La Espectroscopia infrarroja cercana (NIR) es una técnica que permite, en el intervalo estudiado, la clasificación de muestras de ron en cubanos y no cubanos.
- El uso de la Quimiometría en este campo posibilitará la obtención de certificados de origen y procedencia para rones y otros productos alimenticios.
- De los métodos de clasificación aplicados los mejores son: SIMCA y PLS-DA con 100% de clasificación correcta tanto en el entrenamiento como en la validación.

Recomendaciones

- El número de muestras debe ampliarse para que los modelos se hagan cada vez mejores y más robustos.
- Debe analizarse la posibilidad de construir modelos para predecir las concentraciones de sustancias características de los rones y que actualmente se determinan por cromatografía de gases.