

AC15 - MODELADO DE DESCRIPTORES SENSORIALES DEL VINO CON FUSIÓN DE DATOS DE TÉCNICAS INSTRUMENTALES. SIMULACIÓN DE PANEL DE CATA ELECTRÓNICO

Alba Valls^{1a}, Eva Borràs^{1a}, Ricard Boqué^{1b}, Joan Ferré^{1b}, Anna Brull²,
Montserrat Mestres^{1a}, Laura Aceña^{1a}, Olga Busto^{1a}

¹Universitat Rovira i Virgili, Dpto. de Química Analítica y Química Orgánica, Campus Sescelades, 43007 Tarragona. ^aGrupo de Sensometría Instrumental; ^b Grupo de Quimiometría, Cualimetría y Nanosensores.

²VITEC, Parque Tecnológico del Vino. Crta. Porrera, Km.1, 43730 Falset (Tarragona).
alba.valls@estudiants.urv.cat

Resumen: La calidad sensorial del vino es el resultado de la percepción simultánea de diversos atributos de color, olor y gusto, resultado de combinaciones de compuestos químicos. El análisis sensorial, basado en paneles de cata, es lento, caro y hasta cierto punto subjetivo, lo que ha motivado que en los últimos años haya habido un interés creciente en la búsqueda de sistemas para relacionar los atributos sensoriales descritos por los paneles con medidas instrumentales objetivas. En este trabajo se han combinado tres instrumentos, una nariz-e HS-MS, una lengua-e FT-MIR y un ojo-e UV-Vis, para predecir descriptores sensoriales en vinos tintos de la DOC Priorat. Los conjuntos de datos resultantes se han procesado mediante técnicas quimiométricas y se han desarrollado modelos de calibración multivariante PLS para la predicción de los descriptores.

Palabras clave: vino, PLS regresión, panel electrónico, nariz electrónica, lengua electrónica, análisis multivariante. fusión de datos. sensometría

1. Introducción

El análisis sensorial implica la medida, interpretación y comprensión de las respuestas humanas a los estímulos percibidos por los sentidos de la vista, olor y gusto [1-2]. El análisis sensorial se lleva a cabo en muchos tipos de bebidas alcohólicas [1,3], siendo el vino una de las más representativas debido a su compleja composición, la cual contribuye a sus características sensoriales únicas [1]. El análisis sensorial lo realizan catadores entrenados, siendo una herramienta muy útil para la clasificación de los vinos y su control de calidad. Sin embargo, dicho proceso es lento, costoso y hasta cierto punto subjetivo [2]. Debido a ello, existe un interés creciente en el desarrollo de métodos instrumentales capaces de predecir atributos sensoriales en alimentos de una forma objetiva, económica, rápida, reproducible e ininterrumpida [1-2].

En los últimos años se han aplicado diferentes métodos instrumentales para establecer una relación entre los atributos sensoriales y la composición química del vino. La espectrometría de masas con espacio de cabeza (HS-MS) y las técnicas espectroscópicas como el infrarrojo medio (FT-MIR) y el ultravioleta-visible (UV-Vis) se han adaptado para ser utilizadas como instrumentos sensoriales [4-5]. La combinación de los resultados de diferentes técnicas instrumentales se denomina "fusión de datos" y ha surgido como una alternativa para incrementar la fiabilidad en la clasificación o la predicción de propiedades en alimentos, respecto a la utilización de técnicas instrumentales individuales. La estrategia más simple de fusión se denomina de bajo nivel (*low-level*) y en ella los datos instrumentales de las diferentes técnicas se concatenan manteniendo en común la dimensión de las muestras para dar lugar a una sola matriz [5]. Después, sobre dicha matriz se aplican los métodos quimiométricos tradicionales de clasificación y/o predicción.

El objetivo de nuestro estudio fue combinar los datos provenientes de tres técnicas instrumentales (HS-MS, MIR y UV-Vis) para mejorar los resultados de predicción alcanzados por las técnicas individuales. Se estudiaron tres atributos sensoriales: la calidad de tanino, la evolución de color y la madera en nariz. Se desarrollaron modelos de calibración multivariante utilizando el método de regresión PLS (*Partial Least Squares*) para relacionar los datos instrumentales con las puntuaciones sensoriales proporcionadas por un panel de cata de VITEC (Parque Tecnológico del Vino) para una serie de vinos de la misma vendimia y denominación de origen (DOC Priorat).

2. Material y métodos

Muestras de vino. Un total de 78 vinos tintos y envejecidos proporcionados por 41 bodegas de la DOC Priorat, todos ellos de la misma vendimia (2009), fueron catados por un panel sensorial y analizados mediante FT-MIR, HS-MS y UV-visible. Las bodegas se localizan en 12 pueblos de la comarca del Priorat (Tarragona) y los vinos de estudio tienen diferente composición varietal, con predominancia de las variedades Garnacha y Cariñena.

Panel sensorial. Hasta 24 atributos sensoriales (visuales, olfatorios y gustativos) fueron evaluados por 42 catadores. Los atributos estudiados se pueden dividir en parámetros generales como la evolución de color, la calidad de tanino, la intensidad aromática o la madera en nariz (valorados de 1 a 6) y en aromas específicos como animal, vegetal o especiado, entre otros (valorados de 1 a 4). Las puntuaciones con valores de 1 son bajas y las de 4/6 son altas, con excepción del atributo *evolución de color* donde 1 se refiere a más evolución y 6 a menos evolución.

Análisis instrumental. La composición volátil se determinó mediante espectrometría de masas con espacio de cabeza (HS-MS). Los volátiles se extrajeron y concentraron usando un espacio de cabeza estático (G1888 Network Headspace Sampler) y se detectaron con un detector selectivo de masas HP5973N. Se prepararon muestras de vino (5 ml) con 0.58 g de NaCl y se depositaron en viales de 10 ml cerrados con un septum de PRFE/silicona. Después de la extracción, se llevó a cabo una desorción térmica de la fibra. Para transferir los volátiles de la fibra al detector de MS sin separación cromatográfica, se mantuvo la columna HP-5MS en un cromatógrafo de gases a la temperatura adecuada. Los espectros de masas se obtuvieron con el modo de impacto electrónico (70 eV) en un rango de masas de 50 a 200 uma.

La composición líquida se analizó mediante espectroscopia de infrarrojo medio (MIR) con un espectrómetro FT-MIR Nexus (Thermo Nicolet, USA) equipado con un detector de sulfato de triglicina deuterado (dTGS). Los espectros de transmitancia se registraron a temperatura ambiente en el rango de 4000-400 cm^{-1} , a una resolución de 4 cm^{-1} y con 32 escaneos tanto para el fondo como para las muestras. Se utilizó el software OMNIC (versión 6.2) de Thermo Nicolet para la adquisición de los espectros, el control instrumental y la manipulación preliminar de los archivos.

Finalmente, las medidas de color se obtuvieron con un espectrofotómetro (Thermo) Helios Gamma de un solo haz equipado con tecnología de diodos en fila. Los espectros ultravioleta visible (UV-vis) se obtuvieron a temperatura ambiente en el rango de longitudes de onda de 250-700 nm, con una resolución de 1 nm. Todas las muestras se analizaron por duplicado en una cubeta de cuarzo con un camino óptico de 1 mm.

Análisis estadístico de los datos. Los espectros obtenidos en primer lugar se pre-procesaron usando métodos quimiométricos, para eliminar el ruido y la información irrelevante. Como análisis exploratorio preliminar, se usó PCA (Análisis de Componentes Principales) para revelar posibles agrupaciones de muestras y para detectar muestras discrepantes. A continuación, se construyeron los modelos de calibración para predecir tres atributos sensoriales generales: evolución de color, madera en nariz y calidad de tanino, mediante las técnicas individuales y a partir de las técnicas combinadas (fusionadas). El método utilizado fue la regresión PLS (Mínimos Cuadrados Parciales) [5,6]. Los espectros de

cada técnica se trataron de forma específica dependiendo de sus características. Los datos fusionados necesitaron un pre-procesado adicional con el objetivo de compensar las diferentes escalas de medida y la variabilidad de cada técnica [5]. Las regiones espectrales más informativas se seleccionaron escogiendo los valores de los índices de las variables importantes de proyección (VIP) mayores que uno [7]. La validación de los modelos se realizó dividiendo 10 veces al azar los conjuntos de datos en conjuntos de calibración y de validación (70:30). La calidad de los modelos se comprobó mediante el valor de RMSECV (*Root Mean Squared Error of Cross-Validation*) y el coeficiente de determinación de predicción (r^2_p). El procesado de los datos y la construcción de los modelos se llevó a cabo usando rutinas de Matlab v.7.8 propias (Mathworks, MA, USA) y el software PLS Toolbox v.6.2 (Eigenvector Research, Manson, WA, USA).

3. Resultados

Los espectros de cada una de las técnicas instrumentales se organizaron en matrices de dimensiones 78x151 (HS-MS), 78x246 (FT-MIR) y 78x410 (UV-visible). Para desarrollar la estrategia de fusión de datos de nivel bajo, los espectros de las tres técnicas se concatenaron obteniendo de este modo una única matriz.

Respecto a cada técnica individual, se aplicó el Análisis de Componentes Principales (PCA) para detectar muestras discrepantes, para comprobar la repetitividad de las medidas y para reconocer patrones en los datos, básicamente relacionados con el pueblo de origen y con la composición varietal. No se observaron patrones y algunos valores atípicos fueron identificados.

A continuación, se calcularon los modelos de regresión PLS con los espectros de cada técnica individual y también con los datos fusionados para cada uno de los tres descriptores. Los parámetros de validación de los modelos finales se encuentran sintetizados en la *Tabla 1*.

Tabla 1. Parámetros de validación de los modelos de regresión PLS de los datos individuales y combinados para cada descriptor estudiado (*). Las técnicas en negrita son las que proporcionan los mejores resultados.

DESCRIPTOR	TÉCNICA	RMSEP		r^2 de predicción		Número factores PLS
		promedio	D.S	promedio	D.S	promedio
Evolución de color	FT-MIR	0,27	0,04	0,38	0,20	5
	HS-MS	0,28	0,05	0,23	0,12	2
	DF: MIR-MS	0,20	0,04	0,66	0,11	7
	UV-visible	0,15	0,011	0,79	0,05	7
	DF: MS-MIR-UVV	0,15	0,017	0,81	0,05	4
Madera en nariz	FT-MIR	0,34	0,06	0,05	0,05	1
	HS-MS	0,26	0,04	0,38	0,17	4
	UV-visible	0,30	0,014	0,19	0,10	2
	DF: MS-UVV	0,27	0,04	0,45	0,12	4
Calidad de tanino	FT-MIR	0,22	0,03	0,42	0,13	6
	HS-MS	0,24	0,03	0,23	0,15	2
	DF: MIR-MS	0,19	0,04	0,59	0,15	3
	UV-visible	0,25	0,03	0,23	0,10	2

(*) Los resultados se presentan como promedios y desviaciones estándares (D.S.) de los 10 modelos.

MS: headspace/espectrometría de masas, MIR: espectroscopia infra-rojos, UV-vis: espectroscopia UV-visible;

Se probaron diferentes pre-procesados, regiones espectrales y número de factores retenidos por los modelos PLS. En el caso de HS-MS los espectros fueron pre-procesados mediante "row profile" y logaritmo en base diez. Los espectros de FT-MIR y UV-visible se pre-procesaron mediante SNV (*Standard Normal Variate*), corrección de línea base o primera derivada, o

combinación de ellos. Antes de construir el modelo de regresión PLS todos los espectros se centraron y escalaron (*autoscaling o Pareto scaling*). Las regiones espectrales más influyentes fueron identificadas estudiando los valores de los índices VIP. Por su lado, los datos fusionados necesitaron escalado y pre-procesado adicional. Finalmente, se seleccionaron las mejores condiciones buscando obtener los valores óptimos de RMSECV y r^2_p .

Los mejores resultados de predicción para las técnicas individuales se obtuvieron con espectroscopia de UV-visible para *evolución de color* (RMSEP = 0.15), HS-MS para *madera en nariz* (RMSEP = 0.26) y FT-MIR para *calidad de tanino* (RMSEP = 0.22). Las estrategias de fusión de datos de nivel bajo mejoraron los resultados de predicción para *madera en nariz* y *calidad de tanino*. La fusión de datos MS-MIR-UVV proporcionó una ligera mejora de los resultados de *evolución de color*, pero no de forma suficiente para justificar el uso de los tres instrumentos. Es remarcable enfatizar que los valores de RMSEP resultantes son comparables a las desviaciones estándar ponderadas del método de referencia (panel de cata), que son 0.14 para la *evolución de color*, 0.16 para la *madera en nariz* y 0.15 para la *calidad de tanino*.

4. Conclusiones

Este estudio muestra que es posible predecir de forma aceptable descriptores del vino usando técnicas instrumentales y que la estrategia de fusión de datos de nivel bajo mejora los resultados de predicción. Los mejores modelos de regresión PLS se obtuvieron con UV-visible para la *evolución de color*, con datos fusionados de MS-UVV para la *madera en nariz* y datos fusionados MIR-MS para *calidad de tanino*. Futuros estudios, usando estrategias de fusión de datos de nivel medio y alto deberían ser considerados para mejorar los resultados de predicción. Finalmente, la mejora en la exactitud de los resultados proporcionados por el método de referencia (panel de cata), por ejemplo a través de procesos de acreditación, también mejoraría los errores de predicción de los modelos.

5. Bibliografía

- [1] Cozzolino, D., Cowey, G., Lattey, K. A., Godden, P., Cynkar, W. U., Damberg, R. G., Gishen, M. (2008). Relationship between wine scores and visible-near-infrared spectra of Australian red wines. *Analytical and Bioanalytical Chemistry*, 391, 975–981.
- [2] Lozano, J., Santos, J., Arroyo, T., Aznar, M., Cabellos, J., Gil, M., Horrillo, M. (2007). Correlating e-nose responses to wine sensorial descriptors and gas chromatography–mass spectrometry profiles using partial least squares regression analysis. *Sensors and Actuators B: Chemical*, 127(1), 267–276.
- [3] Vera, L., Aceña, L., Guasch, J., Boqué, R., Mestres, M., Busto, O. (2011). Characterization and classification of the aroma of beer samples by means of an MS e-nose and chemometric tools. *Analytical and Bioanalytical Chemistry*, 399, 2073–2081.
- [4] Vera, L., Aceña, L., Boqué, R., Guasch, J., Mestres, M., Busto, O. (2010). Application of an electronic tongue based on FT-MIR to emulate the gustative mouthfeel “tannin amount” in red wines. *Analytical and Bioanalytical Chemistry*, 397, 3043–3049.
- [5] Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., Busto, O. (2015). Data Fusion Methodologies for Food and Beverage Authentication and Quality Assessment – a Review. *Analytica Chimica Acta*.
- [6] Cozzolino, D., Smyth, H. E., Lattey, K. A., Cynkar, W., Janik, L., Damberg, R. G., Gishen, M. (2006). Combining mass spectrometry based electronic nose, visible-near infrared spectroscopy and chemometrics to assess the sensory properties of Australian Riesling wines. *Analytica Chimica Acta*, 563, 319–324.
- [7] Mehmood, T., Liland, K. H., Snipen, L., Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62–69.

6. Agradecimientos

El Ministerio Español de Ciencia y Tecnología (Project AGL2011-26456) es reconocido con agradecimiento por el soporte financiero.