Poster No 153

Wine chemical characterization by Near Infrared Spectroscopy and chemometric analysis

Amaral, AL[1,2], Genisheva, Z[2], Quintelas, C[2], Mesquita, DP[2], Ferreira, EC[2] and Oliveira, JM[2]

[1] Instituto Politécnico de Coimbra, ISEC, DEQB, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal,
[2] CEB - Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

The recent developments in both chemometrics and instrumentation have resulted in rapid methods for predicting the concentration of specific chemical constituents and helped to reduce the demand for traditional analysis. Near infrared spectroscopy (NIR) is such a rapid and non-destructive technique and generally requires minimal sample processing prior to analysis. The aim of the present work was to examine the potential of NIR spectroscopy to determine the concentration of 10 different compounds in white wines.

The collected NIR spectra used in this analysis ranged from 5435 cm^{-1} to 6357 cm^{-1}. Initially a boxplot analysis, regarding the dependent variables (Y), was performed resulting in Y outliers identification and removal. Next, a PCA–X analysis was carried out, regarding the independent variables (X) for the identification of distinct clusters and possible X outliers. This led to four different datasets fed to the PLS analysis: [a] – ensemble dataset with no X outliers removed; [b] – ensemble dataset with X outliers removed; [c] – dataset divided in 3 clusters (1, 2 and 3) with no X outliers removed; and [d] – dataset divided in 3 clusters with X outliers removed. Hence, in the performed analysis, the total number of samples varied between 90 and 100, with 2/3 used for modelling (test) purposes and 1/3 for validation.

Regarding the chemometric approach, an iterative method was applied, first determining the weights of each wavelength for the entire wavelength values PLS, next grouping the wavelength values together according to the weights similarity and, finally, recalculating the PLS with the averaged wavelength values. For all PLS analyses, the maximum number of PLS components allowed was set at half of the test data. For all the studied compounds, the best coefficient of determination (R^2) was above 0.94 (Table 1).

Table 1 – Coefficient of determination (R^2) and number of X variables (n) of the PLS analysis

Compound	[a]		[b]		[c]		[d]	
	R^2	n	R^2	n	R^2	n_1, n_2, n_3	R^2	n_1, n_2, n_3
ethyl acetate	0.88	25	0.79	15	**0.95**	**10, 10, 9**	0.91	13, 9, 7
methanol	0.90	18	0.92	24	**0.96**	**8, 10, 9**	0.92	11, 7, 7
2-methyl-1-butanol	0.96	23	**0.96**	**19**	0.92	9, 9, 9	0.91	11, 8, 7
3-methyl-1-butanol	**0.96**	**20**	0.95	18	0.92	9, 9, 8	0.91	9, 8, 7
2-phenylethanol	**0.97**	**19**	0.94	17	0.92	10, 7, 8	0.91	11, 9, 7
3-methylbutyl acetate	**0.96**	**21**	0.96	27	0.95	10, 9, 8	0.83	11, 9, 6
ethyl lactate	0.90	18	0.94	21	0.95	9, 10, 7	**0.96**	**12, 8, 7**
ethyl octanoate	0.91	18	**0.94**	**19**	0.92	8, 8, 9	0.93	11, 9, 7
diethyl succinate	**0.97**	**20**	0.94	22	0.92	8, 8, 10	0.92	8, 9, 7
diethyl malate	**0.95**	**16**	0.95	20	0.93	7, 8, 9	0.95	9, 7, 7