Área: ANA   **Nº de Inscrição: 00002**

Geographical traceability of South Brazilian red wines through fused spectroscopic data

Mateus P. Schneider(PG)[a], Fernanda C. Bock (PG)[a], Celito C. Guerra (PQ)[b], Leticia L. Silva (PQ)[b], Marco F. Ferrão (PQ)[a], Adriano A Gomes (PQ)[a]

araujo.gomes@ufrgs.br; mateuspschneider@gmail.com

[a] Universidade Federal do Rio Grande do Sul, Instituto de Química, Porto Alegre, RS, Brasil, 90650-001
[b] Empresa Brasileira de Pesquisa Agropecuária, Centro Nacional de Pesquisa de Uva e Vinho, LACEM, Bento Gonçalves, RS, Brasil, 95700-000
keywords: wine classification, data fusion, spectroscopic

Highlights

South Brazilian red wines. Geographical traceability. Data fusion.

Abstract

The state of Rio Grande do Sul is the largest national wine producer, being the regions of Serra Gaúcha and Campanha the most outstanding. These two zones are different in climate, altitude, soil and pluviometric indicies, thus generating different wines in aroma, flavor and quality. The traceability of wines with indication of their geographical origin identity is an important aspect for producers and consumers. Geographical traceability is not a trivial task, and an useful approach is to use signals recorded in different sources in order to achieve best results. In this work is reported classification of red wines produced in the state of Rio Grande do Sul by means of fused spectroscopic (UV-Vis and ATR-FTIR) data. Thirty-five samples (provided by EMBRAPA Grape and Wine) from different varieties of grapes (Ancellotta, Cabernet Franc, Carbenet Sauvignon, Marselan, Merlot, Pinot Noir and Terodelgo) were taken into account. All calculation was carried out in MatLab® environmental. Initially, recorded data were inspected through of principal components analysis (PCA), in order to evaluate the discriminating power of UV-Vis and IR spectra, both individually and fused at low level. This first step showed tendency of separation between the samples based on the region of the grapes. PLS-DA models were constructed for UV (16 VL) and IR (2 VL) data, then reaching correct classification rates of 37% and 75% respectively for a test set (8 samples) chosen from the pool of samples by Kernnard-Stone approach. Data was fused at low (join of the data matrices), medium (selection of the most informative variables by genetic algorithm-GA) and high levels. In the last case, selected variable by GA were used as input in the construction of potential functions for each class, these responses were combined to construct a quadratic decision boundary – see **Fig. 1**)
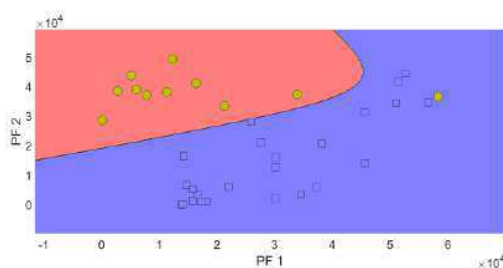
Figure 1: Quadratic decision boundary based on the potential function values of calss 1 (Serra Gaúcha-squares) versus class 2 (Campanha –balls).

For fused data at low and medium levels the correct classification rate was 81% and 85%, respectively. In high-level fusion, all samples, of the independent test set, were classified correctly. Only one sample, of the training set, was allocated in the wrong class. The findings suggest that the fusion of the spectroscopic data show synergism, achieving better results when compared to the individual data. The fusion at a high level proved to be the most appropriate.

Acknowledgments