FT-MIR spectroscopy and multivariate analysis for determination of bioactive compounds and antioxidant capacity in Cabernet Sauvignon wines

Claudia Grijalva-Verdugo, Maylet Hernández-Martínez, Ofelia Gabriela Meza-Márquez, Tzayhri Gallardo-Velázquez and Guillermo Osorio-Revilla

ABSTRACT
Thirty samples of Cabernet Sauvignon Mexican wines from 5 wine regions and 19 wineries before 2004 and 2012 were characterised regarding total phenolic compounds (TPC), anthocyanins (AC), tannins (TC), flavonoids (FC) and antioxidant capacity (ABTS and DPPH assays). Samples were processed using Fourier transform mid-infrared spectroscopy combined with partial least squares regression (PLS1 and PLS2) and principal components regression algorithms to obtain one optimised chemometric model for each parameter. The spectral region selected for each model was between 1550 and 824 cm$^{-1}$. The PLS1 model showed the best predictive capacity for simultaneous prediction of the six parameters. Excellent values of the coefficient of determination in validation ($R_v^2$; 0.93–0.95) and residual predictive deviation (RPD; 4.28–6.95) were obtained for TPC, FC, ABTS and DPPH. Suitable values of $R_v^2$ (0.91–0.92) and RPD (3.75–3.84) were obtained for TC and AC.

1. Introduction
Wine is a complex beverage from a chemical point of view, being composed of water, alcohol and contains more than 500 compounds, some originating from the grapes and some metabolic by-products of yeast activity during fermentation (Soleas, Diamandis, & Goldberg, 1997). Ethanol is the main by-product, but since wine’s elaboration incorporates the seeds and skin of the grapes, both with high concentrations of bioactive compounds, red wine is also rich in polyphenolics. Polyphenols contributed either the grape pigmentation and red colour or the flavour (Ribéreau-Gayon, Glories, Maujean, & Dubourdieu, 2006).

Polyphenols have a particular importance due to their antioxidant properties (Dávalos & Lasunción, 2009). The moderate red wine consumption (100–200 mL/day) reduces the risk of cardiovascular diseases, neurodegenerative illness, some symptoms of arthritis and some cancers (Jackson, 2016; Kang, Shin, Lee, & Lee, 2011).

One of the growing wine producer countries is Mexico. Nowadays, it is among the most prominent wine-producing countries in America (Medina-Albala de León, Martínez-Carrion, & Ramón-Muñoz, 2014). The main wine-growing regions are in Mexico’s north, and Cabernet Sauvignon grape variety is the most cultivated (SAGARPA, Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, 2011). According to the literature, there is little information about Mexican wines, despite its growth in the global wine market.

Some analytical methods are available for identifying the bioactive compounds in wine, but gas chromatography and high-performance liquid chromatography are the most widely used (Porgali & Büyüktuncel, 2012). However, techniques of vibrational spectroscopy have gained popularity (Bevilacqua, Bucco, Materazzi, & Marin, 2013) due to their features: speed, reliability, effortlessness, environmentally friendly, requires minimal or no sample treatment and suitability for routine analyses because of its capacity to process large samples in a short time. Thus, they are an alternative to conventional methods.
chemical analysis. The advantage of using infrared spectra is that it represents the combined absorbance of all components in the wine matrix, rather than merely the composition of a few chemical compounds (Bevin, Ferguson, Perry, Janik, & Cozzolino, 2006). Besides, the coupling with multivariate techniques enables to develop predictive models, which allow quantitative information of sample composition, and this is known as chemometric analysis and is a powerful analysis tool.

Fourier transform mid-infrared (FT-MIR) spectroscopy coupled to chemometrics has been applied in the analysis of a wide range of enological parameters in the finished product (e.g., Ozturk, Yucesoy, & Ozen, 2012; Silva, Feliciano, Boas, & Bronze, 2014) or during winemaking (e.g., Di Egidio, Sinelli, Giovannelli, Moles, & Casiraghi, 2010; Fragoso, Aceña, Guasch, Mestres, & Busto, 2011; Preseroova, Ranc, Milde, Kubistova, & Stavek, 2015). The PLS1 algorithm is commonly used to model properties, and the PLS2 or PCR algorithms are not tested.

In the wine industry, the routine analysis is focused on analysing chemical components such as alcohol, sugars and acidity (Cozzolino, 2015). However, the investigation and prediction of bioactive compounds content and the antioxidant capacity are necessary to characterise the vinification process, know the sensorial attributes and recognise wine quality to predict its protector effect which is a significant aspect for consumers.

The scientific literature reports few studies of the prediction of antioxidant capacity in wine by FT-MIR (Patz, Bliike, Ristow, & Dietrich, 2004; Preseroova, Ranc, Milde, Vladimira, & Stavek, 2015; Silva et al., 2014; Versari, Parpinello, Scizzina, & Del Rio, 2010). To our knowledge, there is no PLS or PCR models base on FT-MIR applied to the simultaneous prediction of wine antioxidant capacity by ABTS and DPPH assays. These methods differ in their principles, and they are complementary.

The first aim of the research was to obtain a characterisation of total phenolic compounds (TPC), total anthocyanins content (AC), tannins content (TC), flavonoids content (FC) and antioxidant capacity (ABTS and DPPH assays) of Mexican Cabernet Sauvignon wines produced in the wine-growing region of Mexico which is an emergent wine sector on the global market. The second aim was to develop chemometric models based on data obtained by FT-MIR spectroscopy and PLS1, PLS2 and PCR algorithms for simultaneous prediction of phenolic compounds (TPC, AC, TC and FC) and antioxidant capacity of red wine variety Cabernet Sauvignon produced in Mexico, to obtain a fast, reliable and easy methodology.

2. Materials and methods

2.1 Samples

Thirty commercial red wine varieties Cabernet Sauvignon produced in Mexico (Aguascalientes, Aguascalientes, n = 1; Encinillas, Chihuahua, n = 1; Ensenada, Baja California, n = 23; Parras, Coahuila, n = 3; and Tecate, Baja California, n = 2) were collected in commercial wineries in Mexico City. The samples were originated from 18 Mexican brand producers and belong to different vintages (2001, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013 and 2014) to consider the extensive and natural variability of red wines when the chemometric model is built of the calibration models. All samples were stored in a dark room, at room temperature until analysis. The chemicals used in the experiments were analytical grade.

2.2 Bioactive compounds determination

The 30 red wine samples were analysed in triplicate by spectrophotometric assays (Jenway 7305, UK). The total phenol content (TPC) was determined based on Singleton and Rossi (1965) method and was calculated as milligrams equivalents of gallic acid per litre (mg GAE/L). Total AC was measured using bisulphite bleaching method (Ribéreau-Gayon et al., 2006) and was expressed as mg malvidin-3-glucoside/L. The FC was determined based on Kramling and Singleton (1969) method and was calculated as milligrams equivalents of gallic acid per litre (mg GAE/L). The TC was measured according to Ribéreau-Gayon and Stonestreet method (Ribéreau-Gayon et al., 2006) and was expressed as g cyanidin/L wine.

2.3 Antioxidant capacity determination

The antioxidant capacity of the wine samples was analysed using DPPH radical (1,1-diphenyl-2-picrylhydrazyl; Brand-Williams, Cuvelier, & Berset, 1995) and ABTS+ radical cation (2,2’-azinobis(3-ethylbenzothiazoline-6-sulphonic acid; Re et al., 1999). The antioxidant capacity was expressed as millimoles equivalents of Trolox per litre (mmol TE/L). The samples were analysed in triplicate.

2.4 Statistical analysis

Values obtained for TPC, AC, TC, FC, DPPH and ABTS were analysed using one-way ANOVA with Tukey’s test (p < 0.05) and coefficient of variation (CV) to confirm the variability present in the data set which is necessary for a proper calibration. Moreover, principal component analysis (PCA) was performed for data set reference values before chemometric models were developed. PCA allows to examine sample groupings and detects outliers. Minitab Statistical Software version 16.1.0 (State College, PA, USA) was used for the analyses.

2.5 FT-MIR spectroscopic analysis

The wine samples were scanned using a 1600 FT-MIR spectrometer (PerkinElmer, MA, USA) within the range of 4000–550 cm⁻¹ at a resolution of 4 cm⁻¹, and the average of 64 scans was considered. Measurements were carried out with an attenuated total reflectance (ATR) accessory with a zinc selenide crystal. An aliquot of 1.5 mL of red wine was placed onto the ATR crystal with a Pasteur pipette, avoiding the bubble formation, and covering with a press accessory to prevent the evaporation of the sample. A background air spectrum was subtracted from each wine sample spectrum. Samples were analysed in triplicate and averaged.

2.6 Quantitative models

The chemometric analysis was performed using partial least squares regression (PLS1 and PLS2) and principal components regression (PCR) algorithms of Unscrambler software v9.2 (CAMO Process AS, Oslo, Norway). Each one was used to correlate spectral data with the data set of bioactive compounds and antioxidant capacity values of wine samples. The regression models were created with 30 averaged FT-MIR spectra. The samples for calibration and validation sets were chosen before the analysis of principal component 2D scatterplot defined by the two first PCA. When all samples...
were placed in a 2D space, the 20% was randomly selected for the validation set (6 samples). The selected samples were well distributed in that 2D space. The remaining 80% (24 samples) was used for the calibration set.

Different spectral pre-treatments were used to improve the accuracy of the chemometric models. The pre-treatments applied were Savitzky–Golay filter with until 19 smoothing points to diminish the random noise in the spectral signal. The correction of offset to avoid changes in the baseline. The first and second order derivatives with until 13 smoothing points to increase spectral differences; and multiplicative scatter correction (MSC) or standard normal variate (SNV), which eliminate or minimise interferences caused by scattering effects. Furthermore, the correct choice of the spectral region was a necessary step.

The standard error of calibration (SEC) and the coefficient of determination in calibration ($R^2_c$) were obtained to evaluate calibration step. The standard error of prediction (SEP), the coefficient of determination in validation ($R^2_v$), Mahalanobis total distance and residual error were obtained with the Unscrambler software v9.2 to estimate the predictive capacity of the model. Additionally, in validation step, the residual predictive deviation (RPD) was calculated as the ratio between the standard deviation of the population (SD) and the SEP, and the percentage of relative difference (%RD) was obtained by Eq. (1):

$$
\text{%RD} = \frac{(\text{Reference value} - \text{Predicted value})}{\text{Reference value}} \times 100
$$

The closest to minimum SEP for each parameter at the lowest number of factors was used as a criterion for selecting the optimal number of factors (in PLS) or components (in PCR) for each parameter. The selection is necessary to avoid overfitting or underfitting of the model, which is the inclusion of unnecessary data or the lack of variance in the data, respectively.

3. Results and discussion

3.1 Bioactive compounds

Figure 1 summarises the analytical values of bioactive compounds and antioxidant capacity of Mexican Cabernet Sauvignon wines, also shows the wine region, winery’s number and vintage of each sample. The samples in each graph are ordered from the highest to the lowest values of each analysed parameter.

By making an overall analysis, it was observed that the wines from Chihuahua (H) and Coahuila (C) are between samples with high TPC, AC, TC and FC values (Figure 1). For all parameters, the wines showed a slightly significant difference between samples for each parameter, although the CV ranges from 13.65 to 57.77. The exception was the FC parameter with a highly significant difference between wine samples. Besides, taking into account the vintage, it was not found a diminish of TPC or TC values through the time. However, some oldest wines showed low values of AC and FC.

In the present investigation, the TPC ranged from 803 to 2432 mg GAE/L and the average value was 1254 mg GAE/L (Figure 1). The TPC shows a CV of 16.0%. These results are higher than those reported by others authors for Cabernet Sauvignon red wine, which concentration ranged from 0.96 to 2.34 g cyanidin/L for Spain, French and Romanian wines (Chira et al., 2011; Hosu et al., 2014; Monagas et al., 2006; Soto-Vázquez, Río-Segade, & Orriols-Fernández, 2010).

The variation in the TC may also be due to the conditions under which the maceration takes place since this process has a significant effect on the relative proportion of tannins that pass from the grape to the wine (Favre, Charamelo, & González-Neves, 2013). Another possible explanation for the variation in tannin concentration is the use of oenological tannins. They are used to facilitate the clarification of wines and musts (International Oenological Codex, 2015). For instance, the high values for some Spain wine are due to tannins addition (Soto-Vázquez et al., 2010).

The FC in samples ranged from 115.22 to 647.41 mg GAE/L, and the average value was 185.4 mg GAE/L (Figure 1). The values were lower compared to the reported values in the current literature for Cabernet Sauvignon red wine which ranged from 15.31 to 377 mg malvidin-3-glucoside/L for French, Romanian and Spain wines (Chira et al., 2011; Fragoso et al., 2011; Hosu et al., 2014).

The differences may be due to the conditions used during the maceration phase in winemaking. The extended macerations allow wines with high colour, rich in anthocyanins and tannins, due to the ripeness of the grape and ambient conditions during the wine growing (Aleixandre, 2006, and due to progressive transformation into more stable oligomeric and polymeric pigments during ageing (Monagas, Gómez-Cordovés, & Bartolomé, 2006). The TC of wines ranged from 1.93 to 4.74 g cyanidin/L, and the average value was 3.2 g cyanidin/L (Figure 1). The TC shows a CV of 20.97%. These results are higher than those reported by others authors for Cabernet Sauvignon red wine, which concentration ranged from 0.96 to 2.34 g cyanidin/L for Spain, French and Romanian wines (Chira et al., 2011; Hosu et al., 2014; Monagas et al., 2006; Soto-Vázquez, Río-Segade, & Orriols-Fernández, 2010).

The variation in grape composition or different winemaking techniques probably modify the FC (Lim et al., 2011). Moreover, the lower flavonoids concentration in some samples could be due to progressive anthocyanin-flavanol condensation reactions during ageing (Monagas et al., 2006). Thus, aged wines exhibited lower flavonoids and anthocyanins contents. Similar results were reported by Granato et al. (2010).

3.2 Antioxidant capacity

The evaluation of the antioxidant capacity of wine is an indirect index of phenolic compounds present in wine (Fotakis, Christodouleas, Zervou, Papadopoulos, & Calokerinos, 2012).
No method fully reflects all the antioxidants of a sample. Therefore, two methodologies were used in the present study to measure the antioxidant capacity of a complex system such as wine.

In general, it was observed that the wines from Chihuahua (H) and Coahuila (C) have high antioxidant capacity values (Figure 1). There are not high statistical differences among wine samples. However, a CV of 13.65 and 20.47% was obtained for ABTS and DPPH values, respectively, and the variation was enough for the development of the chemometric models. Moreover, the ageing of the wine was not associated with lower values of antioxidant capacity (Figure 1).

In the present work, the antioxidant capacity with ABTS method ranged from 29.12 to 52.67 mmol TE/L, with an average value of 40.33 mmol TE/L (Figure 1). Other authors have been reported different values (from 1.08 to 30.9 mmol TE/L) for Romanian, Italian and German wines (Hosu et al., 2014; Patz et al., 2004; Pellegrini et al., 2000). Additionally, the antioxidant capacity with DPPH method ranged from 6.93 to 13.76 mmol TE/L, with an average value of 10.92 mmol TE/L (Table 1). Similar values have been reported for California, South Australia, Chile, Bulgaria, Italy and Macedonian wines, which values ranged from 5.71 to 11.2 mmol TE/L (Ivanova-Petropulus et al., 2015; Van Leeuw, Kever, Pincemail, Defraigne, & Dommes, 2014). However, higher values have been reported for Brazilian, Argentine and Chilean wines (Granato et al., 2010).

The ABTS and DPPH results suggested that Mexican wines had comparable antioxidant capacity with wines from others wine regions.

### 3.3 PCA and correlation analysis

The TPC, AC, TC, FC, DPPH and ABTS results were subjected to the unsupervised PCA pattern recognition method to investigate the presence of outliers and the formation of
The chemical composition of red wines can be correlated with the FT-MIR spectra of samples. The compounds have functional groups that absorb in the mid-infrared region and produce particular vibrational signals of molecular bonds. The resulting spectrum is a fingerprint of the wine.

Figure 2 depicts the raw FT-MIR spectra of Cabernet Sauvignon wines in the region of the infrared medium spectrum (4000–850 cm\(^{-1}\)). In the FT-MIR, the broad and strong band between 3640 and 3200 cm\(^{-1}\) are due to overlapping of O-H stretches vibrations for water, alcohol and phenols in the sample (Larkin, 2011). The bands at 2939 and 2923 cm\(^{-1}\) are due to methyl (-CH\(_3\)) and methylene (-CH\(_2\))- asymmetric stretching, respectively, present in the substitutions of the aromatic rings (Socrates, 2001). The peak observed at 1630–1640 cm\(^{-1}\) corresponds to the C = C stretch vibration (Socrates, 2001), that is a characteristic bond of flavonoids and non-flavonoids. The band near 1640 cm\(^{-1}\) corresponds to O-H group deformation (Larkin, 2011). The absorption at 1448–1444 cm\(^{-1}\) and 1376–1373 cm\(^{-1}\) corresponds to asymmetric and symmetric in-plane bending of -CH\(_3\), respectively (Silva et al., 2014). The bands observed in the region 1410–1310 cm\(^{-1}\) correspond to O-H deformation and C-O stretching of phenols (Socrates, 2001). The bands in the region 1330–1210 cm\(^{-1}\) and 960–850 cm\(^{-1}\) correspond to the C-O stretching and O-H out-of-plane torsions of the carboxylic acids (Coates, 2006), which are present in red wine such as coumaric, ferulic, cinnamic and caffeic acids. The absorption at 1410–1310 cm\(^{-1}\) has been assigned to O-H bending in phenols (Coates, 2006). The peak at 1270–1290 cm\(^{-1}\) is due to the = C-O-C asymmetric stretching (Socrates, 2001) and C-O stretching.

### 3.4 FT-MIR spectra

The chemical composition of red wines can be correlated with the FT-MIR spectra of samples. The compounds have functional groups that absorb in the mid-infrared region and produce particular vibrational signals of molecular bonds. The resulting spectrum is a fingerprint of the wine.

Figure 2 depicts the raw FT-MIR spectra of Cabernet Sauvignon wines in the region of the infrared medium spectrum (4000–850 cm\(^{-1}\)). In the FT-MIR, the broad and strong band between 3640 and 3200 cm\(^{-1}\) are due to overlapping of O-H stretches vibrations for water, alcohol and phenols in the sample (Larkin, 2011). The bands at 2939 and 2923 cm\(^{-1}\) are due to methyl (-CH\(_3\)) and methylene (-CH\(_2\))- asymmetric stretching, respectively, present in the substitutions of the aromatic rings (Socrates, 2001). The peak observed at 1630–1640 cm\(^{-1}\) corresponds to the C = C stretch vibration (Socrates, 2001), that is a characteristic bond of flavonoids and non-flavonoids. The band near 1640 cm\(^{-1}\) corresponds to O-H group deformation (Larkin, 2011). The absorption at 1448–1444 cm\(^{-1}\) and 1376–1373 cm\(^{-1}\) corresponds to asymmetric and symmetric in-plane bending of -CH\(_3\), respectively (Silva et al., 2014). The bands observed in the region 1410–1310 cm\(^{-1}\) correspond to O-H deformation and C-O stretching of phenols (Socrates, 2001). The bands in the region 1330–1210 cm\(^{-1}\) and 960–850 cm\(^{-1}\) correspond to the C-O stretching and O-H out-of-plane torsions of the carboxylic acids (Coates, 2006), which are present in red wine such as coumaric, ferulic, cinnamic and caffeic acids. The absorption at 1410–1310 cm\(^{-1}\) has been assigned to O-H bending in phenols (Coates, 2006). The peak at 1270–1290 cm\(^{-1}\) is due to the = C-O-C asymmetric stretching (Socrates, 2001) and C-O stretching.

---

**Figure 2.** a) Principal components 3D scatterplot of 30 wines grouped by wine region and winery, letters and numbers are codes as in Figure 1, b) Principal components loading plot.

**Tabla 1.** Resultados de calibración para los modelos quimiométricos basados en espectroscopía FT-MIR.

| Algorithm | Parameter | Factors | \( R^2 \) \(^{a}\) | SEC \(^{b}\) | SEP \(^{c}\) |
|-----------|-----------|---------|-----------------|-----------|-----------|
| PLS1      | TPC       | 6       | 0.9436          | 1.45      | 1.51      |
|           | AC        | 6       | 0.9478          | 0.88      | 0.92      |
|           | TC        | 6       | 0.9389          | 0.74      | 0.71      |
|           | FC        | 5       | 0.9435          | 1.83      | 1.77      |
|           | ABTS      | 6       | 0.9392          | 0.89      | 1.01      |
|           | DPPH      | 5       | 0.9480          | 1.02      | 0.96      |
| PLS2      | TPC       | 8       | 0.9390          | 2.68      | 2.74      |
|           | AC        | 8       | 0.9101          | 2.52      | 2.65      |
|           | TC        | 8       | 0.9231          | 1.84      | 1.91      |
|           | FC        | 6       | 0.9397          | 3.02      | 2.99      |
|           | ABTS      | 6       | 0.9285          | 2.54      | 2.67      |
|           | DPPH      | 6       | 0.9320          | 2.98      | 3.11      |
| PCR       | TPC       | 7       | 0.9001          | 7.51      | 7.98      |
|           | AC        | 6       | 0.9127          | 7.02      | 7.32      |
|           | TC        | 5       | 0.9085          | 6.98      | 6.23      |
|           | FC        | 5       | 0.9124          | 7.31      | 6.89      |
|           | ABTS      | 7       | 0.9109          | 7.81      | 8.01      |
|           | DPPH      | 6       | 0.9227          | 6.97      | 5.87      |

\(^{a}\)Optimum number of factors (PLS) or optimum number of principal components (PCR).

\(^{b}\)SEC, standard error of calibration, should be as low as possible.

\(^{c}\)SEP, standard error of prediction, should be as low as possible.
vibration of the pyran-derived ring structure of flavonoid-based tannins (Edelmann & Lendl, 2002). Strong absorptions are also observed in the region 1000–1095 cm\(^{-1}\) due to C-O stretching vibrations of polysaccharides, also at 900–835 cm\(^{-1}\) region, the absorption peaks of \(\alpha\) and \(\beta\) pyranose compounds are identified (Socrates, 2001), such as those present in glycosylated anthocyanins such as malvidin-3-glucoside.

In the FTIR spectra, many bands that have been attributed to the same functional group of components in the complex matrix. Thus, a combination of the spectra and chemometric is required to predict the bioactive compounds’ concentrations.

### 3.5 Chemometric models

Three optimised chemometric models were obtained with the algorithms PLS1, PLS2 and PCR. They were used to associate TPC, AC, TC, FC and antioxidant capacity (ABTS and DPPH) of Cabernet Sauvignon wines with their FT-MIR spectra. The correlation allows to find components or factors from the original correlated spectroscopy data and uses only these linear combinations in the regression equation (Naes, Isaksson, Fearn, & Davies, 2004).

Different pre-treatments were applied to the models to reduce the noise and to intensify the signals of interest for the samples. The selected pre-treatments applied in the three models were smoothing of 13 points, SNV normalisation and second derivative with 9 points. The region used for the models was from 1550 to 824 cm\(^{-1}\). This region showed a suitable correlation between absorbance and concentration data to avoid X-variable outliers. Neither regions with high absorbance nor regions of very low absorbance are desirable, as they contribute to extremely high variation or may be considered to be only noise, respectively.

Table 1 shows the optimal number of principal components for PCR or factors for PLS, the \(R^2\), the SEC and the SEP of the three optimised calibration models. Factors represent a new dimension with significant data that resume the original variables in the spectrum, and they are useful to explain a parameter of interest. According to Table 1, PLS1 and PLS2 models showed higher \(R^2\) values than PCR model in calibration data set. Values of \(R^2\) high to 0.9 suggest excellent quantitative information (Shenk & Westerhaus, 1996). For PLS1 and PLS2, lower SEC and SEP values were obtained, in comparison with PCR using the optimal number of factors. The SEC and SEP have the same unit as the reference value and provide an indication of the potential error in the regression; consequently, low values are desirable.

Thus, SEC and SEP values obtained by PLS1 model indicated a better calibration than PCR and PLS2 models. Moreover, the PLS1 model used fewer factors to predict, and then, a more robustness model is expected. For that reason, the PLS1 model was chosen as the better model, and the subsequent validation step was applied only for PLS1 model.

PLS1 and PLS2 are two different approaches for PLS. The first algorithm calibrates the properties one at a time and PLS2 calibrates all simultaneously, searching the maxima variance between analytic and spectral data. PCR algorithm also calibrates all at a time and selects components only according to the maxima variability in spectral data (Naes et al., 2004). PLS1 usually gives more accurate models. However, PLS2 provides models with similar predictive abilities to PLS1 when the dependent variables are strongly correlated (Pedro & Ferreira, 2007). The present work explored the prediction ability of PLS2 and PCR, but PLS1 showed the best calibration performance at less number of factors.

The validation set of 6 external samples was employed to know the performance of prediction models. The regression of predicted values versus the reference values was obtained for each parameter. Table 2 shows the \(R^2\) and RPD values for the prediction of each parameter with PLS1 model. Excellent correlations (\(R^2 > 0.91\)) between the reference and the predicted data were obtained for the six parameters.

**Figure 3.** FT-MIR spectra in the region of 4000–850 cm\(^{-1}\) of Mexican Cabernet Sauvignon wines.

**Figura 3.** Espectros FT-MIR en la región de 4000–850 cm\(^{-1}\) de vinos mexicanos Cabernet Sauvignon.
Table 2. Validation results for the PLS1 model obtained using external validation.

| Parameter | $R^2$ | $RPD^b$ | Prediction range |
|-----------|-------|---------|------------------|
| TPC (mg GAE/L) | 0.9301 | 4.78 | 855–1789 |
| AC (mg malvidin-3-glucoside/L) | 0.9184 | 3.84 | 5.17–59.06 |
| TC (g cyanidin/L) | 0.9147 | 3.75 | 2.03–4.98 |
| FC (mg GAE/L) | 0.9502 | 6.03 | 107.93–655.01 |
| ABTS (mmol TE/L) | 0.9396 | 4.28 | 28.87–54.02 |
| DPPH (mmol TE/L) | 0.9460 | 6.95 | 7.01–14.15 |

$R^2$, coefficient of determination of validation.

$RPD$, residual predictive deviation of validation.

Table 3. Comparative results of six external samples quantify by reference methods and PLS1 model based on FT-MIR spectroscopy.

| Wine's code | Parameter | Reference value | Predicted value | Mahalanobis total distance $^a$ | Residual ratio $^b$ |
|-------------|-----------|-----------------|-----------------|------------------------------|---------------------|
| C2_2011     | TPC       | 1440            | 1463            | 0.3184                       | 0.9853              |
|             | AC        | 22.96           | 22.94           |                              |                     |
|             | TC        | 4.33            | 4.02            |                              |                     |
|             | FC        | 493.67          | 493.98          |                              |                     |
|             | ABTS      | 50.94           | 51.30           |                              |                     |
|             | DPPH      | 12.61           | 12.45           |                              |                     |
| E4_2005     | TPC       | 1306            | 1309            | 0.3204                       | 1.0843              |
|             | AC        | 17.21           | 17.52           |                              |                     |
|             | TC        | 3.35            | 3.29            |                              |                     |
|             | FC        | 331.52          | 329.87          |                              |                     |
|             | ABTS      | 42.78           | 42.55           |                              |                     |
|             | DPPH      | 10.66           | 10.87           |                              |                     |
| E5_2008     | TPC       | 1222            | 1206            | 0.1876                       | 1.0923              |
|             | AC        | 11.67           | 11.61           |                              |                     |
|             | TC        | 2.71            | 2.75            |                              |                     |
|             | FC        | 302.56          | 303.20          |                              |                     |
|             | ABTS      | 39.47           | 39.55           |                              |                     |
|             | DPPH      | 9.11            | 9.87            |                              |                     |
| E8_2011     | TPC       | 1423            | 1436            | 0.2873                       | 0.9831              |
|             | AC        | 12.21           | 12.28           |                              |                     |
|             | TC        | 3.99            | 4.09            |                              |                     |
|             | FC        | 484.44          | 475.31          |                              |                     |
|             | ABTS      | 40.48           | 40.84           |                              |                     |
|             | DPPH      | 11.56           | 11.61           |                              |                     |
| E3_2010     | TPC       | 1085            | 1089            | 0.3419                       | 1.0213              |
|             | AC        | 16.38           | 17.25           |                              |                     |
|             | TC        | 3.41            | 3.37            |                              |                     |
|             | FC        | 444.44          | 445.22          |                              |                     |
|             | ABTS      | 39.75           | 40.10           |                              |                     |
|             | DPPH      | 10.08           | 10.43           |                              |                     |
| E3_2011     | TPC       | 1079            | 1085            | 0.1830                       | 1.0089              |
|             | AC        | 15.46           | 15.99           |                              |                     |
|             | TC        | 2.70            | 2.65            |                              |                     |
|             | FC        | 499.26          | 500.28          |                              |                     |
|             | ABTS      | 38.49           | 39.78           |                              |                     |
|             | DPPH      | 9.57            | 10.83           |                              |                     |

$^a$ Mahalanobis total distance should be as low as possible, not to exceed 1
$^b$ Residual ratio should be as low as possible, not to exceed 3; the wine's code, TPC, AC, TC, FC, ABTS and DPPH as in Figure 1.

Nevertheless, less satisfactory calibrations had been reported for wine's components in low concentration and some studies have addressed this problem (Di Egidio et al., 2010; Fernandez & Agosín, 2007; Fragoso et al., 2011; Laghi, Versari, Parpinello, Nakaji, & Boulton, 2011; Ozturk et al., 2012). In general, in those studies, TPC yielded excellent and FC good predictions. However, AC and TC showed low correlations. The last studies had been considered the fingerprint region as the best option for predictions, although Fernández & Agosín (2007) suggested the full range of 4000–650 cm$^{-1}$ for improved TC prediction. In the present work, a broad spectral range was tested, and TC prediction was not improved. Future works are required to improve the accuracy of AC and TC predictions.

As far as antioxidant capacity, few chemometric model based on FT-MIR had been developed. Presereova et al. (2015) and Versari et al. (2010) predicted antioxidant capacity using ferric reducing-antioxidant (FRAP) assay, and they reported $R^2$ of cross-validation of 0.88 and 0.85, respectively. Silva et al. (2014) also predicted antioxidant capacity both as FRAP and as % inhibition of DPPH, and they reported $R^2$ of...
cross-validation of 0.383 and 0.367, respectively, and rough estimates for DPPH and FRAP antioxidant capacities, especially for FRAP. Finally, Patz et al. (2004) predicted antioxidant capacity based on decolourisation of ABT+ radical, and they reported a $R^2$ of 0.9204.

The present work tested both DPPH and ABTS assays to improve prediction of antioxidant capacity, and the result was satisfactory since DPPH and ABTS predictions showed $R^2$ higher than 0.93 and RPD values higher than three which indicate good prediction models. It is worth mentioning that, to our knowledge, no previous published studies in wine had modelled simultaneously DPPH and ABTS assays using FT-MIR and PLS algorithm.

The prediction of antioxidant capacity is of paramount importance since it is a way to estimate the presence of phenolic compounds with radical scavenging activity due to polyphenols from red wine have been reported to exert potent antioxidant effects (Dávalos & Lasunción, 2009). The last statement is in agreement with the correlation between antioxidant capacity and TPC found in the present work.

### 4. Conclusions

The present study indicates an overview of the Mexican Cabernet Sauvignon wines’ composition of main bioactive compounds and antioxidant capacity. A high correlation among DPPH, ABTS, TPC and TC was observed in oppose to the weak correlation between AC and TPC, DPPH or ABTS. The TPC values were comparable with the reported in the current literature, but lower AC and FC values and higher TC values were observed respect previously published results.

The use of the PLS1 algorithm, instead of PLS2 and PCR algorithm, shows a more satisfactory model for the prediction of Cabernet Sauvignon wine’s properties. The statistical results demonstrated that PLS1 model was suitable to simultaneous prediction of TPC, AC, TC, FC and antioxidant capacity by ABTS and DPPH methods. Excellent predictions of TPC, FC, ABTS and DPPH were provided, and a good prediction of AC and TC values was obtained. The proposed model is a convenient and reliable instrument for quality control or authenticity studies when a considerable number of samples must be analysed.

### Acknowledgements

Financial support from the Consejo Nacional de Ciencia y Tecnología (CONACyT) and the Secretaría de Estudios de Posgrado e Investigación del Instituto Politécnico Nacional de México (SIP-IPN) is greatly appreciated.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

Financial support from the Consejo Nacional de Ciencia y Tecnología (CONACyT) and the Secretaría de Estudios de Posgrado e Investigación del Instituto Politécnico Nacional de México (SIP-IPN) is greatly appreciated.

### References

Aleixandre, J. (2003). La cultura del vino, cata y degustación [Wine culture and tasting] (pp. 361). Valencia, ES: Universitat Politecnica de Valencia, Servicio de publicaciones.

Atanackovic, M. (2013). Analysis of phenolic content and antioxidative capacity of red wines from Serbia. In L. Č. Popović, M. Vidakovic, & D. S. Kostić (Eds.), Resources of Danubian Region: The possibility of cooperation and utilization (pp. 10–16). Belgrade, RS: Humboldt-Club Serbia.

Brivallacqua, M., Buccì, R., Materazzi, S., & Marinì, F. (2013). Application of near infrared (NIR) spectroscopy coupled to chemometrics for dried egg-pasta characterization and egg content quantification. Food Chemistry, 140(4), 726–734. doi:10.1016/j.foodchem.2012.11.018

Bevin, C. J., Fergusson, A. J., Perry, W. B., Janik, L. J., & Cozzolino, D. (2006). Development of a rapid “fingerprinting” system for wine authenticity by mid-infrared spectroscopy. Journal of Agricultural and Food Chemistry, 54(26), 9713–9718. doi:10.1021/jf062265o

Brand-Williams, W., Cuvelier, M. E., & Beres, C. (1995). Use of a free radical method to evaluate antioxidative activity. LWT–Food Science and Technology, 28(1), 25–30. doi:10.1016/0023-6438(95)80008-5

Chira, K., Pacella, N., Jourdes, M., & Teissèdre, P. L. (2011). Chemical and sensory evaluation of Bordeaux wines (Cabernet-Sauvignon and Merlot) and correlation wine age. Food Chemistry, 126(4), 1971–1977. doi:10.1016/j.foodchem.2010.12.056

Coates, J. (2006). Interpretation of infrared spectra: A practical approach. In R. A. Meyers (Ed.), Encyclopedia of analytical chemistry (pp. 1–23). Chichester, UK: John Wiley and Sons.

Cozzolino, D. (2015). Sample presentation, sources of error and future perspectives on the application of vibrational spectroscopy in the wine industry. Journal of the Science of Food and Agriculture, 95, 861–868. doi:10.1002/jsfa.6733

Dávalos, A., & Lasunción, M. A. (2009). Health-promoting effects of wine phenolics. In M. V. Moreno-Aribas & M. C. Polo (Eds.), Wine chemistry and biochemistry (pp. 571–591). New York, NY: Springer.

Di Egidio, V., Sinelli, N., Giovanelli, G., Moles, A., & Casiraghi, E. (2010). NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. European Food Research and Technology, 230(6), 947–955. doi:10.1007/s00217-010-1227-5

Edelmann, A., & Lendl, B. (2002). Toward the optical tongue: Flow-through sensing of tannin-protein interactions based on FTIR spectroscopy. Journal of the American Chemical Society, 124(49), 14741–14747. doi:10.1021/ja026309y

Favre, G., Charamelo, D., & González-Neves, G. (2013). Empleo de taninos enológicos y maceración premermativa en frío en una experiencia de elaboración de vinos Tannat [Use of oenological tannins and cold premermativeness maceration in a Tannat winemaking experiment]. Agrociencia Uruguay, 17(1), 65–73. Retrieved from http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S0023-6438(95)80008-5

Fotakis, C., Christodouleas, D., Zervou, M., Papadopoulos, K., & Cozzolino, D. (2012). Toward the optical tongue: Flow-through sensing of tannin-protein interactions based on FTIR spectroscopy. Journal of the American Chemical Society, 124(49), 14741–14747. doi:10.1021/ja026309y

Fragoso, S., Aceña, L., Guasch, J., Mestres, M., & Busto, O. (2011). Quantification of phenolic compounds during red winemaking using FT-MIR spectroscopy and PLS-regression. Journal of Agricultural and Food Chemistry, 59(20), 10795–10802. doi:10.1021/jf201973e

Granato, D., Katayama, F. C. U., & Castro, I. A. (2010). Assessing the association between phenolic compounds and the antioxidant activity of Brazilian red wines using chemometrics. LWT–Food Science and Technology, 43(10), 1542–1549. doi:10.1016/j.lwt.2010.05.031

Hosu, A., Vasile-Mircea, C., & Cimpoiu, C. (2014). Analysis of total phenolic, flavonoids, anthocyanins and tannins content in Romanian red wines: Prediction of antioxidative activities and classification of wines using artificial neural networks. Food Chemistry, 150, 113–118. doi:10.1016/j.foodchem.2013.10.153

International Oenological Codex. (2015). Oenological tannins. COEI-1-TANINS: 2015. OIV-Deno 554–2015. Paris, FR: Organisation Internationale de la Vigne et du Vin (OIV).
