Available online at www.sciencedirect.com
SCIENCE DIRECT
Talanta 69 (2006) 469–480
Talanta
www.elsevier.com/locate/talanta

Determination of quality parameters of beers by the use of attenuated total reflectance-Fourier transform infrared spectroscopy

Rafael Llario [a], Fernando A. Iñón [b], Salvador Garrigues [a], Miguel de la Guardia [a,*]

[a] *Departamento de Química Analítica, Universidad de Valencia, Edificio Jeròmin Muñoz, C/Dr. Moliner, 50, 46100 Burjassot, Valencia, Spain*

[b] *Laboratorio de Análisis de Trazas, Departamento de Química Inorgánica, Analítica y Química Física, Universidad de Buenos Aires, Pabellón 2, Ciudad Universitaria, 1428 Buenos Aires, Argentina*

Received 14 March 2005; received in revised form 1 July 2005; accepted 14 October 2005
Available online 17 November 2005

Abstract

The estimation of important quality parameters of beers, such as original and real extracts and alcohol content, has been evaluated by attenuated total reflectance-Fourier transform infrared spectroscopy (ATR-FTIR) using a partial least square (PLS) calibration approach. Two sample populations, one consisting of 24 samples and other of 21 samples, obtained from the Spanish market and covering different types of beer were used. The first set was used for building and validating the model, whereas the second, measured 6 months after, was used for evaluating its robustness. The spectral range and the size of the calibration set and its suitability for building the PLS model have been evaluated.

Considering a calibration set comprised of 12 samples, selected via hierarchical cluster analysis, and a validation data set of 11 samples, the absolute mean difference (d_{x-y}) and standard deviation of mean differences (s_{x-y}) of the real extract, original extract and alcohol content were 0.009 and 0.069% (w/w), −0.021 and 0.20% (w/w) and −0.003 and 0.130% (v/v), respectively. The maximum error for the prediction of any of these three parameters for a new sample did not exceed 2.5%. These values were practically invariant for both tested data sets.

The developed methodology favourably compares with the automatic reference methodology in terms of speed and reagent consumption and waste generation.

Keywords: Alcohol content; Beers; Hierarchical cluster analysis; PLS; ATR-FTIR

1. Introduction

Beer is one of the oldest known alcoholic beverages and is obtained by fermentation of cereals germinated in water in the presence of yeast [1]. The main components of beers are water, carbohydrates and ethanol. These three parameters are commonly used in the brewing industry for quality control of the final product, under the name of real extract, original extract and ethanol content. Original and real extracts are the amounts of sugars in beer before and after producing the fermentation process. Original and real extracts are normally expressed in % (w/w) (the former can be also be found in degrees Plato, but both units are equivalent). Ethanol content is a key economic and organoleptic parameter affecting both the beer classification (in term of taxes) and its taste [1,2]. It is usually expressed in % (v/v).

The real extract is also important because it is a measure of the amount of sugars that did not underwent fermentation and remains in the beer. With this parameter, as consumers, we can know the sweetness of the product and its energetic value as source of carbohydrates. The determination of the original extract and the ethanol content together with the real extract is important for the industry for knowing the fermentation grade of the beer (efficiency factor). This parameter informs the brewing master about the fermentation yield.

The official methods of the Analytical Division of European Brewery Convention for the determination of these three parameters are based on the distillation of the beer and the measurement of the density, after making up to a standardized final volume, of the distillate and the remaining solutions [3].

* Corresponding author.
E-mail address: miguel.delaguardia@uv.es (M. de la Guardia).

0039-9140/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2005.10.016

The density values of both solutions are introduced in semi-empirical tables for obtaining the percentage of extracts and the ethanol content. These methods are time consuming and therefore interns of time/cost inefficient. Therefore, there is a need for developing fast alternative methods for routine quality control programs.

On the other hand, the most common way in which the determination of alcohol and extracts in beer is carried out in the brewing industry is through the use of autoanalysers. These instruments allow one to carry out the analysis of these parameters by triplicate analysis in about 30 min, that is, 10 min per replicate.

For the best of our knowledge the following are the most relevant works carried out up-to-date for developing alternative methodologies. Table 1 summarises the main aspects of the works found in the literature on the use of multivariate calibration and vibrational spectrometers in order to obtain a prompt information about the characteristics of beer samples. Parameters, like the number of samples and composition of the calibration and validation set and the root-mean-squared error of prediction (RMSEP), of most significant works are listed.

Mendes et al. [4] made a study on the determination of ethanol in fuel ethanol and some beverages based upon near infrared spectra evaluated by partial least squares regression (NIR-PLS) and Raman-PLS. This work shows that this determination can be successfully carried out, although the prediction capabilities for beers are far from being considered exhaustive as it was only proved with a single sample. Another work by Norgaard et al. [5] provided a chemometric study based on NIR measurements and interval partial least square (iPLS) regression for the determination of original extract in beer, with a root-mean-squared error of prediction of 0.1–0.2% (w/w). There are other works using NIR spectroscopy for determining these constituents in beer samples, using either univariate [6–8] or multivariate calibration [9,10] models, with similar or worse prediction capabilities that those reported by Norgaard et al. [5]. In particular, the work of Engelhard et al. [8] provides an interesting review on the assays used for alcohol determination and proposes a quantification method based on interpretative difference of the absorption of radiation in the NIR spectral region using univariate statistics. On the other hand, this work only deals with ethanol determination without addressing the simultaneous determination of the other quality parameters. For example, in the NIR-based study published by Maudoux et al. [9] it has been reported that alcohol, real extract, original gravity, nitrogen and polyphenols could be estimated from the transmission spectra of samples with relative prediction errors around 6% through the use of multiple linear regression and PLS on using actual beer samples for calibration. Nevertheless, mid-infrared spectrum has been used in less extension.

In the work of Malone and Flournoy [11] the determination of ethanol in synthetic aqueous mixtures by Fourier transform mid-infrared spectroscopy (FTIR) using attenuated total reflectance (ATR) was evaluated, but no application to actual samples was included. On the other hand, Kupper et al. [12] have presented different developments in mid-IR fiber-optic spectroscopy based on ATR. One of these developments was related to the deter-

Table 1
Summary of the content of previously published chemometric-spectroscopy procedures for the determination of parameters in beers

Determined parameter	Technique	Calibration		Validation		Number factor	Prediction error	Reference
		Type	Number of standards	Type	Number of standards			
Only classification	FTIR and H-RMN, PCA	Beer (ale, lager, alcohol-free)	50					Duarte et al. [13]
Ethanol	NIR and FT-Raman and PLS	Ethanol fuel standards	25	Ethanol fuel standards	25			Mendes et al. [4]
Original extract	NIR-iPLS	Beer samples	40	Beer samples	20	4	0.1–0.2% (w/w)[a]	Norgaard et al. [5]
Ethanol						6	4.29%[b]	
Real extract						7	6.53%	
Original extract	NIR-PLS	Beer Samples	70	Beer Samples	40	5	4.5%	Maudoux et al. [9]
Nitrogen							6.06%	
Polyphenols							4.74%	

Note: Different models have been built, the best one in terms of prediction capabilities was chosen.
[a] RMSEP.
[b] Coefficient of validation.

mination of ethanol in beers, but neither quantitative data nor calibration efforts were covered.

A recent work [13] has been focussed on the classification of beer types using FTIR and nuclear magnetic resonance (NMR) spectra, followed by principal component analysis (PCA). It was concluded by the authors that good estimation of the type of beer can be obtained from the FTIR spectra, whereas NMR helps to differentiate the sugar content of the samples. Nevertheless, no quantitative results were provided.

Regarding the use of multivariate or univariate statistics, it may be expected that for overlapped bands (such as sugars in both NIR and MIR) multivariate models will provide better results for total or individual prediction of constituent than univariate ones. No references on the determination of real and/or original extract combined or not to the determination of ethanol in beers by ATR-FTIR have been found in the literature in spite of the fact that it is a methodology which potentially can combine the high sample throughput achievable by spectroscopic techniques with the capability of multi-parameter determination of autoanalysers.

Therefore, this work is devoted to the development of a methodology based upon ATR-FTIR measurements and chemometric data processing for the simultaneous determination of real extract, original extract and ethanol content in beers. For increasing the application range of the multivariate model, a heterogeneous sample population has been chosen for selecting the calibration and the validation data set. The selection process has been carried out using hierarchical cluster analysis [14], which is based upon the obtained classification of beers through their spectra. The robustness of the model has been checked using a second set of sample measured 6 months after the original model was built.

Different stages for the development of a PLS-FTIR methodology for the prediction of ethanol and real and original extracts are discussed. Models were selected by means of the: (i) cross-validation (RMSECV), (ii) root-mean-square error of prediction and (iii) relative standard deviation of predicted values of an independent validation set of samples.

As it is well known, in many cases, elimination of non-informative wavenumbers results in better and more robust multivariate models than those based on the whole spectra. Therefore, direct efforts have been conduced for selecting the best spectral range for each analyte that minimises both prediction and calibration errors.

Analytical figures of merit, based on net analyte signal calculations [15], were obtained for the sensitivity and selectivity of the determination of original and real extracts and ethanol content in beers.

2. Experimental

A Nicolet Magna Series 750 model FTIR spectrometer controlled by the Omnic for Windows software from Nicolet Instrument Corp. (Madison, WI, USA) and equipped with Specaclamp IN-Compartment Contact Sampler horizontal ATR from Graseby-Specac (Orpington, UK) with a 45° crystal and six reflections ZnSe through top-plate was employed for spectra acquisition. Room and sample compartment temperatures were monitored using a mercury thermometer with a precision of ±0.5 °C. Both temperatures did not differ significantly during spectra acquisition of all samples. For sample degassing, a magnetic stirrer (Selecta, Barcelona, Spain) and a thermostatic bath (Grant, Cambridge, England) were used.

2.1. Samples

A total number of 45 beer samples contained in sealed aluminium cans were obtained from the market. Samples were analysed in two periods of time. The first and second series of measurements were carried out in September 2003 and February 2004, respectively. In the first series, 24 samples were analysed, whereas in the second stage, additional 21 samples were included. In each case, duplicate cans were collected from Spanish supermarket shelves, one was used for ATR-FTIR measurement and the other was sent to a brewery for measuring reference parameters.

Sample population contains different types of beer:

1. Normal beers (34 samples): these beers are made form barley and some other cereal (normally the most abundant around production area). For these beers, alcohol content is greater than 3% and the original extract is lower than 12.5%.
2. 100% malt beers (4 samples): for these beers the unique source of carbohydrates is malt and it does not contain any other cereal.
3. Special brewed beers (2 samples): the original extract in these beers is higher than 12.5%.
4. Alcohol free beer (1 sample): although its name may lead to confusion, the alcohol content of this beer is lower than 0.9–1%. This sample was obtained from normal beer after undergoing thermal evaporation of the main part of alcohol. The lowest amount of residual alcohol that can be achieved is 0.03%.
5. Beer with soda (2 samples): produced as a mix of normal beer with soda, which is done after the fermentation, just before the canning. This mixture produces a dilution of the beer, changing its flavour and decreasing the extract and alcohol contents.
6. Beer with lemon (1 sample): similar to the previous type, but using a lemon-based soda.
7. German beer (1 sample): apart from its certificate of origin and its flavour, it does not have significant changes in its global composition. The beer acquired for this work has a relative higher alcohol and extract contents, and darker color than a normal beer considered.

Table 2 summarises the basic characteristics of the beer samples used for this study and their labelled reference value for alcohol content expressed both in % (v/v). Standard procedures for the autoanalyser Anton Paar provided by the manufacturer has been followed for analysing the samples. Results obtained by the autoanalyser for sample number 17 were unsuccessful; therefore, this sample will be only used for qualitative analysis, but not quantitative.

Table 2
General description of samples employed in this study

Sample number	Classification	Alcohol (%, v/v)
Data set 1		
1	100% malt	5.5
2	100% malt	5.5
3	Special	5.4
4	Normal	4.5
5	Normal	4.8
6	Normal	5
7	100% malt	5
8	Normal	4.8
9	Special	5.5
10	Normal	4.8
11	Normal	4.5
12	With soda	3.9
13	Normal	4.5
14	Normal	4.5
15	German type	5.5
16	Normal	4.8
17	With lemon	2.8
18	Normal	4.5
19	Normal	4.8
20	Alcohol free	0
21	With soda	3.9
22	Normal	4.5
23	Normal	4.5
24	Normal	4.8
Data set 2		
25	Normal	4.8
26	Normal	4.8
27	100% malt	5.5
28	Normal	4.5
29	Normal	4.8
30	Normal	4.5
31	Normal	4.5
32	Normal	4.8
33	Normal	4.8
34	Normal	4.5
35	Normal	4.8
36	Normal	4.5
37	Normal	4.8
38	Normal	4.8
39	Normal	4.5
40	Normal	4.8
41	Normal	4.8
42	Normal	4.8
43	Normal	4.8
44	Normal	4.8
45	Normal	4.8

Note: Alcohol content is that labelled in the can. The actual content has been determined as explained in the text.

Although density reference values have also been provided, it was not possible to built a calibration model with good prediction capabilities. This unsuccessful result can be explained by three reasons: (i) the small range spanned by this parameter (all samples are between 1.007 and 1.011 g mL^{-1}), (ii) the range of values predicted for triplicates, which was 0.002 g mL^{-1}, and (iii) the temperature of the cell which was not strictly controlled, but monitored. Therefore, no additional comments on this parameter will be provided in this study.

Table 3
Correlation between studied parameters in beer samples

	Original extract	Alcohol (%, v/v)	Alcohol (%, w/w)
Data set 1			
Real extract	0.966	0.922	0.919
Original extract	1	0.989	0.988
Alcohol (%, v/v)		1	0.998
Data set 2			
Real extract	0.817	0.626	0.618
Original extract	1	0.930	0.926
Alcohol (%, v/v)		1	1.000
Whole data set			
Real extract	0.925	0.861	0.856
Original extract	1	0.971	0.968
Alcohol (%, v/v)		1	0.998

Table 3 shows the correlation in the samples of the parameters considered for this study. As it is expected, a high linear correlation was found between the alcohol content expressed in % (w/w) and % (v/v) (alcohol (%, w/w) = 0.7852 alcohol (%, v/v), $R^2 = 0.9957$). There is also a good correlation between the extract and the original one and between the real or the original extract and the content of alcohol. It must be also noticed that the correlation between the experimentally obtained parameters has a similar trend when considering each data set separately or the whole sample population.

2.2. FTIR analysis

Original samples were placed in the same temperature controlled room where the spectrometer was located before to carry out the analysis. The sample compartment temperature was monitored and it remained stable at $26 \pm 1\,^{\circ}C$ during the acquisition of IR spectra for the first data set and $20 \pm 1\,^{\circ}C$ for the second considered set of samples.

Samples were degassed by 5 min stirring and filtered before filling the ATR cell. Three sub-samples were poured into the ATR cell and the FTIR spectra were taken as follows. Sample spectra were scanned between 4000 and 600 cm^{-1}, by averaging 25 scans per spectrum with a nominal resolution of 4 cm^{-1} (data spacing of 1.93 cm^{-1}) and using a mirror velocity of 0.6329 cm s^{-1}. The acquisition of each averaged spectrum requires 35 s.

The background and blank spectra were collected filling the ATR plate cell with Millipore Q-purified water (Bedford, MA, USA) and using the same instrumental conditions as those employed in the case of samples. Background spectrum was scanned at a seven samples interval, while blank spectrum was collected after the measurement of each sample, for assessing that memory effect of IR spectra was minimised after cleaning the ATR crystal.

The cleaning procedure consists of removing the sample by rinsing the ATR crystal, firstly, with warm (45–50 °C) MilliQ water and then with MilliQ water at room temperature. Before filling the accessory with water for measuring the following blank spectrum, the ART crystal was dried with the help of a

soft tissue paper. This drying step was repeated before filling the crystal with the following sample.

Both, sample and blank spectra, were collected in the absorbance mode. The regions between 4000 and 3050 cm^{-1} and between 840 and 600 cm^{-1} were eliminated prior the calculations as it was observed that variations in these regions cannot be ascribed to variations in sample composition. Furthermore, the absorption bands around 2382 and 2314 cm^{-1}, which are due to atmospheric CO_2, were also cut off from raw spectra. These bands could be also minimised by purging the ATR cell with N_2. Nevertheless, as the IR region in which they are located does not provide any relevant information about samples, N_2 purging was considered unnecessary and this region was simply cut off from raw spectra.

The spectra of the three sub-samples of each beer were taken by refilling the ATR cell once again.

2.3. Data analysis

Data obtained from Ommic were exported in text format and analysed using Matlab® (The Mathworks Inc., South Natick, MA, USA). First of all, internal and home-made Matlab functions were used for hierarchical cluster analysis in order to evaluate the similarity of samples in terms of their ATR-FTIR spectra and to assess the number of characteristic subsets in which the available samples could be divided. Similar criteria to that already published for milk classification have been used [16]. Multivariate calibration calculation was made with the MVC1 toolbox [17], including net analyte signal (NAS) calculations. Sensitivity is defined in the multivariate context [15] as the norm of the NAS vector (s_k^*) of the analyte k. Selectivity is defined the quotient between s_k^* and spectrum containing analyte k at unit concentration.

The following figures of model's fit to the data and predictive power have been used throughout the text. In all of the cases, the scope was to estimate the average deviation of the model from the actual data.

PRESS is the sum of squares prediction error (quadratic sum term in Eq. (1)) for the model, which includes A factors. The root-mean-square error of calibration (RMSEC) is a measure of how well the model fits the calibration data, and is defined as:

$$\mathrm{RMSEC} = \left[\left(\sum_{i=1}^{n} (C_i - \hat{C}_i)^2 \right) \cdot (\mathrm{d.f.})^{-1} \right]^{0.5} \quad (1)$$

where $\hat{C}_i$ means the values of the predicted parameter (in our case extracts and alcohol content) when all samples are included in the model building and d.f. is the number of degrees of freedom calculated as the number of calibration samples with known concentration (C_i) minus $A + 1$, the number of factors kept in the model plus one.

The root-mean-square error of cross-validation (RMSECV) is a measure of predictive ability of the model formed on part of a data set to predict the remainder of the data. The RMSECV is defined as the previous equation, except that $\hat{C}_i$ are predictions for samples *not* included in the model formulation, and d.f. is the number of times in which the cross-validation is repeated (i.e., in the leave-one-out cross-validation d.f. is equal to the number of calibration samples).

As the ability of the model to fit the calibration data is not a direct measurement of its prediction capabilities, it is mandatory to compare the values predicted for new samples not used to build the model. This can be performed by calculating the root-mean-square error of prediction when the model is applied to new data for which the reference values are known. RMSEP is calculated exactly as in Eq. (1) except that the estimates for C_i are based on a previously developed model, in which the sample concentrations of the validation set are excluded in the model-building step and the degree of freedom is the number of samples in this set.

In order to validate the FTIR methodology against the auto-analyser, different quality indicators are also given. Among them are the absolute mean difference (d_{x-y}) between FTIR predicted values ($\hat{C}_i$) and reference data (C_i), the standard deviation of mean differences (s_{x-y}), the quality coefficient (QC) and the pooled standard error of prediction for validation samples (s_{reg}) [18]. As stated by Massart et al., the QC is to be preferred over correlation coefficient (of $\hat{C}_i$ versus C_i) "not only because it gives a better idea of the spread of the data points around the fitted straight line but also because it gives some indication on the percentage error to be expected for the estimated concentration".

In order to build and select PLS models, the following iterative procedure was carried out. For building the best calibration model, a selection of the optimum number of factors, which minimise the root-square-of-cross-validation, was made based on the criterion of Haaland and Thomas [19]. To improve the prediction performance of the regression method, a search for suitable sensors was considered. In this sense, one subroutine from MVC1 toolbox was used to find the minimum PRESS, as a function of the number of factors, based on a moving spectral window strategy [17]. Several spectral windows were tested in order to evaluate their prediction capabilities for the validation set. Only most significant results will be shown here.

2.4. Cluster analysis

In hierarchical cluster analysis, the similarity between samples is calculated using the distance concept, calculated using a mathematical relationship (i.e., Euclidian norm) of numerical properties of the samples (i.e., absorbance at different wavelengths). In a successive procedure, each sample is linked to the closest sample or group of samples and a characteristic distance is used to describe this union. This distance between groups of samples can be evaluated in different ways and is the main difference among common linkage methods (Ward, complete, average, etc.). In other words, by this procedure, each sample is replaced by a group comprised of the sample and their neighbour samples located within the given similarity distance. The results are represented in a dendrogram, which shows at which normalised or re-scaled distance (i.e., each distance rationed to the maximum distance, multiplied by a factor) of a group of samples is differentiated from others, when it is read from right to left. At the far-left end each replicate of each sample comprises a group of one member, that is, each spectrum is unique. Thus, for a given

re-scale distance, different number of groups are kept. At this stage we proposed [16] to use the similarity distance between triplicates as minimum cut-off criterion. Actually, taking into account the concepts of limit of detection and quantification, we have chosen 10 times the average distance between replicates as cut-off value.

3. Results and discussion

3.1. Beer FTIR spectra

Beers are aqueous solutions and because of that only the spectral region in which water does not present significant absorption can be used for analytical purposes. This region is between 3050 and 800 cm^{-1}, except the CO_2 region comprised between 2382 and 2314 cm^{-1}. Regions compressed between 4000 and 3050 cm^{-1} and between 840 and 600 cm^{-1} presented high absorption of the incident light by components of the system such as water and the ZnSe crystal and were ignored for this study.

Fig. 1 shows the spectrum of four different types of beers. In the spectrum of the beer without alcohol a series of bands between 1200 and 950 cm^{-1} can be observed, corresponding to C–C and C–O vibration in carbohydrates, which may be correlated with the real and original extracts. The German beer is the sample with the highest extract value, and it can be seen that the band in the sugar region is the highest one. On the other hand, the sample without alcohol has the lowest extract value, and accordingly it has the lowest absorption in the aforementioned spectral region. In the region above 2600 cm^{-1} the bands observed correspond to aliphatic vibrations, which are ascribed to stretching fundamentals and deformations combinations bands of methylene groups [20].

When comparing the spectrum of the beer without alcohol and that of a normal beer, it can be observed that the main differences correspond to the bands located at 875 cm^{-1} (in-phase C–C–O stretch), 1052 cm^{-1} (out-of-phase C–C–O stretch), 1800 cm^{-1} (C–C skeletal stretching) and 2970 cm^{-1} (asymmetric stretching band of methyl group), due to the absorbance of ethanol.

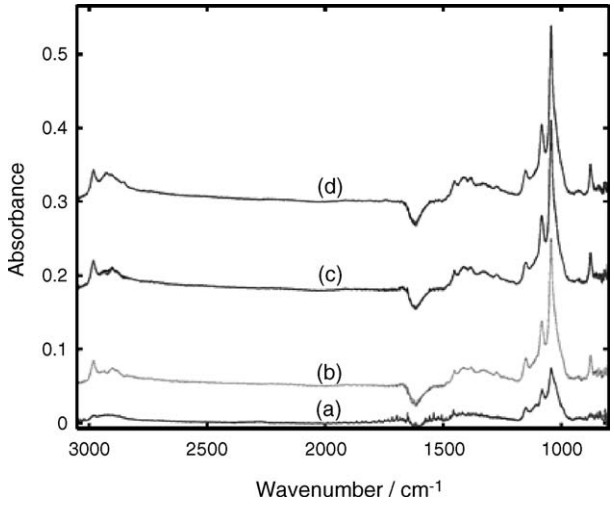

Fig. 1. ATR-FTIR spectra of main types of beer samples: (a) beer without alcohol; (b) normal beer; (c) 100% malt beer; (d) German beer. Spectra were shifted in the absorbance axis for clarity purposes.

The main wavelengths belong to carbohydrates and ethanol are present in the whole set of different beer samples.

3.2. Correction of ATR-FTIR data

When comparing all ATR-FTIR collected spectra for samples considered and in spite of the careful control of the cell cleaning a significant shift of the baseline, sometimes even within replicates of the same sample, was observed. These variations make difficult the subtraction of blanks to the corresponding sample spectra (see Fig. 2A).

Our previous experience in milk and oil samples analysis [14,16] shows that good results, evaluated as minimum misclassification of samples and better prediction capabilities, were achieved when the average absorbance in a fairly flat region was subtracted to each spectrum before data treatment for correcting additive artifacts.

In this case, the average of absorbance values between 2007 and 2056 cm^{-1} was used for performing the spectral correction and these corrected spectra were utilized for further calculations (see Fig. 2B, in which the region of the CO_2 absorption has been deleted).

As the optics involved in ATR are quite different from those used in the transmission experiments, the infrared spectrum of a sample obtained by ATR exhibits some significant differences when compared to its transmission counterpart, such as a distortion of the relative intensities of bands and the introduction of a shift to lower frequencies.

The distortion of relative peak intensities in an ATR spectrum is well known and it is due to the fact that the depth to which the sample is penetrated (d_p) by the infrared beam is directly proportional to the wavelength (λ), and function of incident angle (θ), ATR crystal refractive index (n_c) and sample refractive index (n_s).

In this work, the most simple and commonly applied correction was used, which is based on the hypothesis that the effect of the incident angle, ATR crystal refractive index and sample refractive index on the absorbance is fairly constant between samples. Therefore, as the penetration depth is inversely proportional to the wavenumber, the absorbance at a given wavenumber is corrected by multiplying the absorbance value by the wavenumber value and rationing the result by the maximum wavenumber in the spectrum. The resulting spectrum can be optionally re-scaled to the highest absorbance in the raw spectrum (see Fig. 2C).

3.3. Clustering of beer samples from their FTIR spectra

A clustering method was carried out before PLS data treatment in order to evaluate possible classes among samples considered and allows us to select properly a representative calibration set, thus improving the prediction of unknown samples. Furthermore, differences in sample composition can

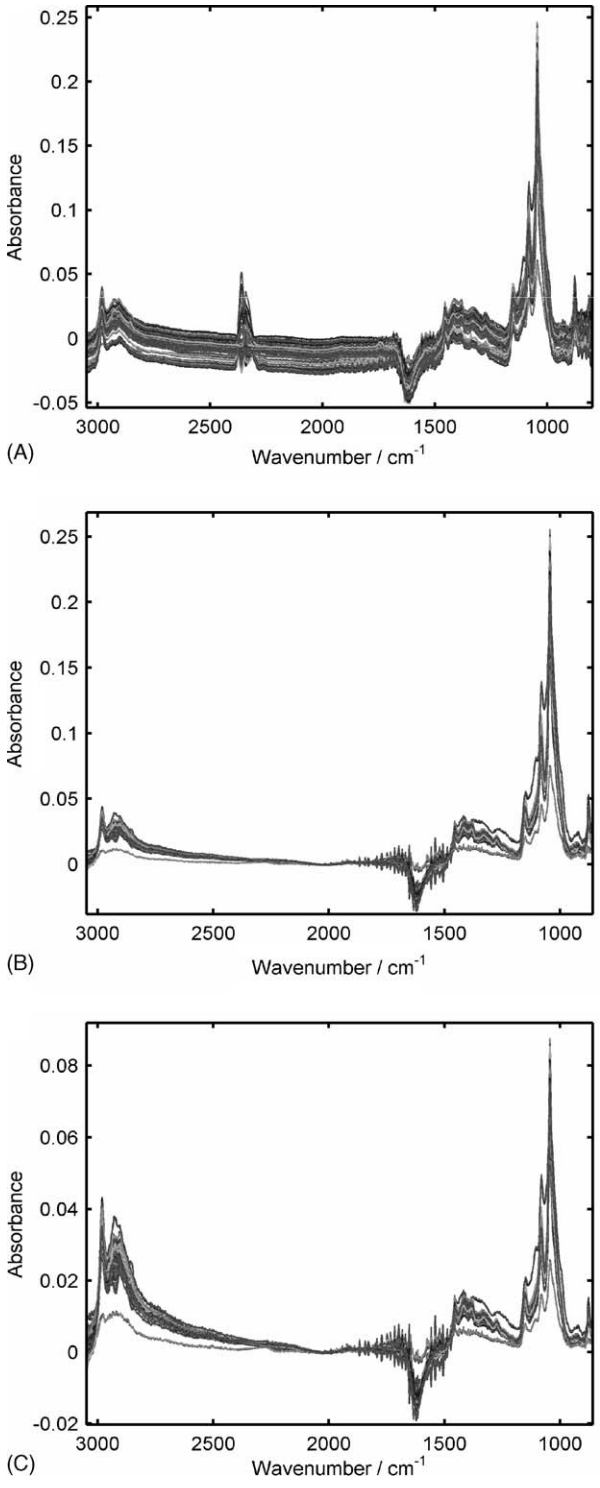

Fig. 2. ATR-FTIR spectra of the whole population of beer samples: (A) raw; (B) base-line shift corrected; (C) additionally corrected through ATR algorithm. Experimental conditions: nominal resolution of 4 cm^{-1}, 25 scans per spectra accumulated and a mirror velocity of 0.6329 $cm\,s^{-1}$. In 'B' and 'C' the CO_2 atmospheric band was eliminated as indicated in the text.

be used for the interpretation of the classification of samples based on their spectra.

In previous works, the calibration set in different complex groups of samples has been successfully selected by the use of hierarchical cluster analysis [14,16]. In the present work, this strategy has been followed using the best combination of distance measuring and linkage methods previously obtained, that is, Euclidian distance and Ward linkage. This linkage method is preferable to other because it is based on a well known statistical procedure, namely within cluster sum of squares [21]. As it has been previously shown, it has some advantages to perform this analysis using the scores of the most significant principal components selected after a principal components analysis, because of: (i) the known noise-filtering effects of PCA analysis and that (ii) PCA drastically reduces the number of variables to be considered for clustering and thus, the computer memory requirements are also reduced. In this case, three factors were enough to explain the 99.2% of the variance in the X-block (spectral data).

When considering a new group of samples, such as in this work when data set 2 (sample nos. 25–45) is considered, the dendrogram has to be constructed using the new spectral data projected over the PCA space constructed from the previous data set. Otherwise, changes in both, the scores and loadings, of the principal components will occur, resulting in noticeable changes in the dendrogram. Fig. 3A–C shows the dendrographic classification of samples obtained for the first, the second and the whole data set, respectively.

As it can be seen from these figures, there is a great overlap between clusters constructed from data set 1 and those obtained when considering all the samples. This fact should reflect the good variety of samples recollected at the first stage. It can be also seen that most samples on data set 2 are grouped in clusters already constituted by samples of data set 1. Nevertheless, some of the samples that are grouped together in Fig. 3A are split in two clusters when considering all the samples. This is a known drawback of some linkage method such as Ward, because it minimises the within sum of squares around the centre of each new formed cluster, which can vary when new samples are included. After analysing Fig. 3A and C, this drawback can be considered not significant because it is noticeable at low dissimilarity distances, but it must be realised that the total number of cluster does not change. A way to overcome this drawback is to combine the result obtained from hierarchical cluster analysis with a supervised learning method, such as discriminant analysis or artificial neural network [22]. In this work, we tested linear discriminant analysis [18] with good results, as the samples were grouped as can be predicted from Fig. 3C maintaining the number of clusters of Fig. 3A.

In all cases, the three replicates of each sample were grouped together. Next, it is possible to try to elucidate if the agglomeration level has a clear interpretation. The main clusters formed (from right to left) are directly correlated with the mean intensity of the FTIR spectra of the samples in these groups and thus, samples with high absorbance level are grouped together. As the absorbance intensity is mainly related with the total amount of carbohydrates and alcohol the groups are basically related with the similar content of these analytes between the samples. Table 4 shows the mean and the standard deviation of the main considered sample parameters.

Basically, clustering criteria seem to be firstly based on original extract and secondly on the alcohol content, which separates

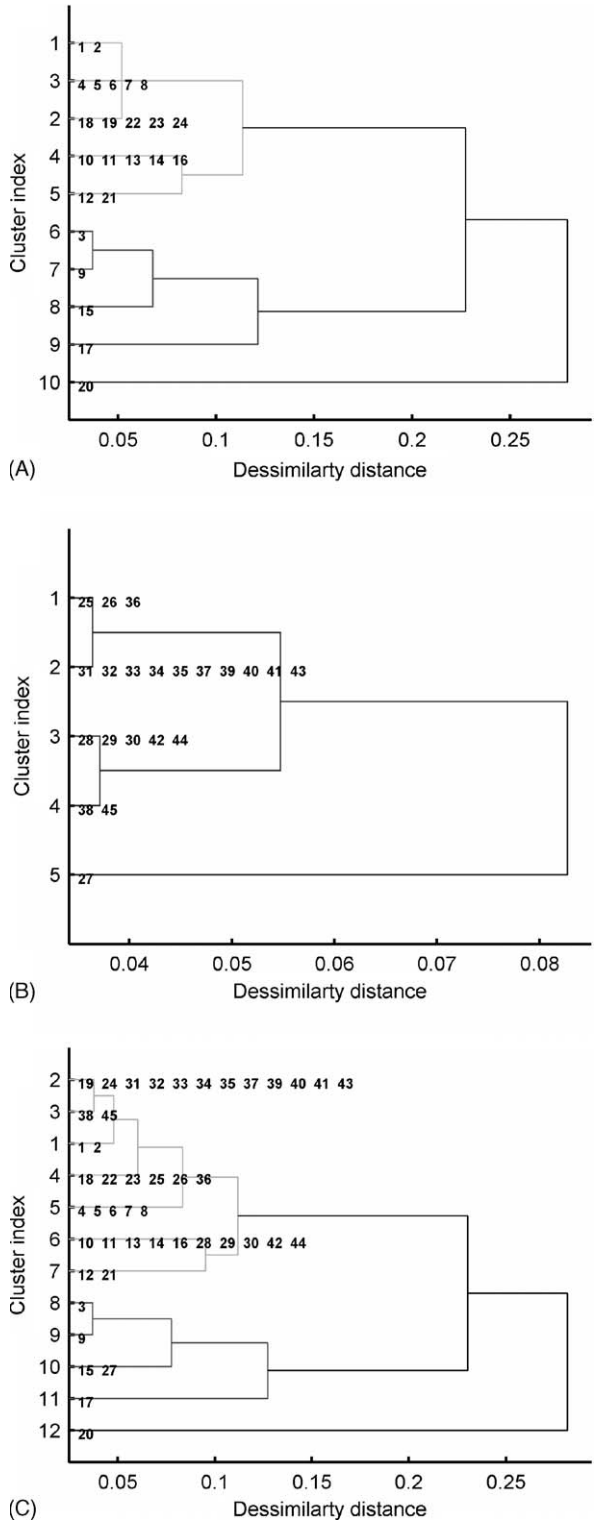

Fig. 3. Dendrographic classification of samples using the Euclidean distance after PCA analysis of ATR-FTIR spectra: (A) data set 1; (B) data set 2; (C) all samples. Cut-off distance was set to 0.035 arbitrary units (typical dissimilarity distance between replicates 0.007 arbitrary units).

samples into three main groups. The classification can set apart: (i) the beer sample with very low original extract and without alcohol (sample no. 20) from (ii) the rest samples. In this second group, another large distinction can be made: (iia) the beer with lime (sample no. 17), the special brewed beers (samples nos. 3 and 9), the beers with soda (samples nos. 12 and 21) and the German type beer (sample no. 15) and (iib) the rest of the beers. The remaining samples (clusters 1–5) are normal ones and present different contents of alcohol and extract. The agglomeration in this cluster seems to follow the direct correlation found between original (and real) extract with alcohol content in these samples, but it is not clear the limits that these values must have in order to distinguish between clusters. Normal beers with similar content of alcohol and extract are grouped together, such as a normal with 4.8% of alcohol (sample no. 5) and a normal with 4.5% (sample no. 4). Similar conclusion may be obtained analysing Fig. 3B, but as the samples in this second data set are more similar to each other, intra-cluster variability obscures the interpretation of agglomeration criterion. The type and content of different carbohydrates may have an important role in the ATR-FTIR classification of beer, but their determination has not been carried out.

3.4. Selection of the calibration set

The determination of the number and the nature of samples to be used for calibration is always a critical factor in multivariate analysis. In the present work, it was done based on the hierarchical cluster analysis results. The selection of the calibration and validation data sets was carried out using data set 1 (Fig. 3A). The selection criterion was based on the following principles:

(i) At least one sample of each cluster was selected for calibration.
(ii) If the cluster is comprised of more than one sample, the number of samples selected for calibration was approximately the root square of the total number of samples included in the cluster, while the remaining samples were integrated in the validation data set. So, the number of samples assigned to the validation was equal or higher than the number of those employed for calibration.
(iii) The samples within a given cluster were selected randomly.

Furthermore, a set composed by one sample of each cluster of the dendrogram of Fig. 3A, randomly selected for cluster comprised of more than one sample, was used to train a linear discriminate analysis (LDA) algorithm. The number of predefined classes in LDA was 10, the same number of classes that the hierarchical cluster analysis provides. The rest of the sample of data set 1 was used to validate the classification obtained by LDA, showing excellent agreement with Fig. 3A. This LDA helps to verify if new samples (in our case data set 2) are represented by the existing calibration data set.

Table 5 summarises the descriptive statistic (mean and standard deviation) of the calibration and validation sets of data set 1 and for data set 2 considering each quality parameter to be determined. As can be seen, the selection of the calibration data

Table 4
Characteristics of the samples grouped using clusters depicted in Fig. 3A

Cluster index	Number of samples	Real extract (%, w/w)		Original extract (%, w/w)		Alcohol (%, v/v)		Samples
		Mean	std	Mean	std	Mean	std	
1	2	4.0_4	0.0_3	12.2_3	0.1_2	5.4_0	0.0_6	1, 2
2	5	3.5_0	0.0_7	10.5_3	0.1_5	4.6_0	0.0_8	18, 19, 22, 23, 24
3	5	3.6_0	0.1_1	10.9_0	0.4_7	4.7_4	0.2_7	4, 5, 6, 7, 8
4	5	3.5_2	0.1_0	10.5_0	0.2_2	4.5_5	0.1_1	10, 11, 13, 14, 16
5	2	2.9_6	0.0_7	8.9_2	0.2_1	3.8_3	0.1_3	12, 21
6	1	4.5_7		12.9_5		5.5_5		3
7	1	4.8_1		13.0_0		5.4_3		9
8	1	4.1_5		12.4_7		5.4_7		15
9	1					2.8_0		17
10	1	2.1_7		4.2_5		1.3_1		20

Note: std—standard deviation.

set can be considered representative of both, the validation set and data set 2.

In all the cases, slightly better RMSEP values (around 1–3%) were obtained using spectra corrected by the ATR algorithm than when using uncorrected ones. Therefore, only model constructed with ATR corrected data will be presented.

3.5. Determination of the real extract

Different models were built and compared in terms of RMSECV and RMSEP values for both, the validation data set and the prediction of samples in data set 2, using as calibration data set: (i) that selected as explained above and (ii) the whole data set 1. It should be noticed that this second case may be regarded as a calibration model that learns from new samples predicted by the original calibration model and for which reference method values are also available. In any case (calibration and validations) outliers have been detected. Outliers detection was carried out as reported by Haaland and Thomas [19].

For the real extract determination the optimum PLS method was based on three extracted factors in the spectral range 1027–1183 cm^{-1}. This spectral region corresponds to the absorption of the hydroxyl groups of sugars and ethanol. Table 6 includes the figures of merit obtained under all aforementioned conditions. The reproducibility of the determination, established from the mean standard deviation of each triplicate and the standard error of prediction (that includes the uncertainty from the model [23,24], considering the standard deviation of reference concentration (s_c) and standard deviation of the response (s_R)), was 0.02 and 0.02% (w/w), similar to that of the reference method. Additionally, it can be seen from Table 6 that there is no significant difference in the obtained figures of merit for the original model and the extended model that also learns from validated samples. The prediction of the new samples in data set 2 has a similar quality to that for the validation data set in both cases.

As it can be seen from d_{x-y}, s_{x-y} and QC, there is no significant difference between results provided by the proposed ATR-FTIR methodology and that provided by the reference methodology. The maximum percentage error to be expected for a new estimated real extract should be 2.9%.

Sensitivity and selectivity parameters were obtained from the net analyte signal [15] and evidences that the real extract determination is quite selective (39%) and provides a sensitivity of the order of 2.65×10^{-2} in all the cases.

3.6. Determination of the original extract

A similar procedure to that followed for real extract determination in beers was made for building a calibration–prediction model for original extract. In this case, the optimum PLS model was also based on three extracted factors, but the optimum spectral region is a little bit different from 1008 to 1202 cm^{-1}. This region also corresponds to the absorption of the hydroxyl groups of sugars and ethanol. It may be thought that both, the original and the real extracts, may be associated with the same spectral features. Evidence to prove or to reject this hypothesis can be obtained from the net sensitivity vector associated for each PLS method (see Fig. 4). As it can be seen, there is a significant difference in the range 1130–1040 cm^{-1} showing that in spite of the correlation observed between the two parameters (see Table 3) there is a different spectral feature for both extracts.

Table 5
Descriptive statistic of calibration and validation data set

Cluster index	Number of samples	Real extract (%, w/w)		Original extract (%, w/w)		Alcohol (%, v/v)	
		Mean	std	Mean	std	Mean	std
Calibration	12	3.61	0.71	10.52	2.34	4.51	1.14
Validation	11	3.58	0.19	10.76	0.61	4.68	0.30
Data set 2	21	3.65	0.24	10.77	0.34	4.70	0.23

Note: Sample number 17 was not included in Validation data set, because no reference data were available. std: standard deviation.

Table 6
Prediction capabilities of PLS-ATR-FTIR for real and original extract determination in beers

Data set	Real extract			Original extract		
	Validation	Data set 2	Data set 2*	Validation	Data set 2	Data set 2*
RMSECV (%, w/w)	0.069	0.069	0.077	0.17	0.17	0.19
RMSEP (%, w/w)	0.075	0.107	0.103	0.20	0.20	0.20
RRMSEP (%)	2.1	2.9	2.8	1.9	1.9	1.9
d_{x-y} (%, w/w)	0.009	−0.006	−0.003	−0.021	0.083	0.079
s_{x-y} (%, w/w)	0.069	0.106	0.104	0.20	0.19	0.19
QC (%)	1.9	2.9	2.9	1.8	2.0	1.9
s_{trip} (%, w/w)	0.04	0.02	0.02	0.09	0.05	0.04
s_{reg} (%, w/w)	0.02	0.02	0.02	0.04	0.04	0.04
Selectivity (%)	39.2	39.2	39.1	53.3	53.3	52.5
Sensitivity	2.64E−02	2.64E−02	2.66E−02	1.30E−02	1.30E−02	1.28E−02

Note: Three factors were used in all the cases. Spectral ranges for real extract and original extract were 1027–1183 and 1008–1202 cm^{-1}, respectively. Data sets 2 and 2* refer to the prediction of new samples in data set 2 using, as calibration set, the calibration set selected form HCA and the whole data set 1, respectively. RRMSEP is the RMSEP divided by the mean value of the original and real extracts. $s_c = 0.05\%$ (w/w); $s_R = 0.0005$; s_{trip} and s_{reg} are the standard deviation and the standard error of prediction, respectively. For additional details see the text.

Regarding prediction capabilities, Table 6 summarises different characteristics of the optimum PLS model built for this parameter using both calibration sets assayed. It should be also said that in any case (calibration and validation) outliers have been detected. As it can be seen, there is neither significant difference in the main prediction indicator values for the prediction of samples in the validation data set nor in the whole data set 2. Regardless which calibration set is used, the same prediction capabilities were obtained, with only three extracted factors. This is an important fact as the robustness of multivariate models decreases as the number of factors increases.

The reproducibility of the determination established from the pooled standard deviation of each triplicate and from the standard error of prediction (have in mind that this latter parameter includes the uncertainty of the model) was 0.05 and 0.04% (w/w), respectively. Therefore, the uncertainty of the calibration model is not significantly different to the uncertainty of the measurement. As it can be also seen from Table 6, there is no significant difference between predicted values by the proposed ATR-FTIR methodology and that provided by the autoanalyser for original extract in the samples assayed, being observed that the d_{x-y} value is significantly smaller than the s_{x-y} value, thus evidencing the absence of systematic errors in the PLS-ATR-FTIR methodology. The maximum percentage error to be expected for a new estimated real extract should be 1.9%.

The selectivity for the PLS-ATR-FTIR determination of original extract is high than that obtained for real extract with a value around 53%, which means that more than the half part of the signal is due to the parameter of interest. Regarding the sensitivity, that found for original extract is two times lower than that found for real extract.

3.7. Determination of alcohol

The alcohol content expressed in % (w/w) and % (v/v) is closely related, so only a model for the latter one was built. Similar procedure as described above was followed for building a PLS model for this parameter. The optimum model was based on two extracted factors in the spectral range from 1055 to 997 cm^{-1}. This region is characteristic of the absorption from –OH alcoholic groups, particularly the band centred at 1045 cm^{-1}. The limitation of the spectral region only to this band yielded, in all the cases, to a model with a slightly higher RMSEP value (around 2% higher) to that provided by the aforementioned region.

Table 7 summarises the prediction capabilities of the optimum PLS model built for this parameter using both calibration sets assayed. In any case (calibration and validations) outliers have been detected. As it can be seen, there is no significant difference in the main prediction capabilities for the three evaluated cases. Regardless of which calibration set is used, the same prediction capabilities are obtained, with only two extracted factors, showing good robustness of model when applied to data set 2.

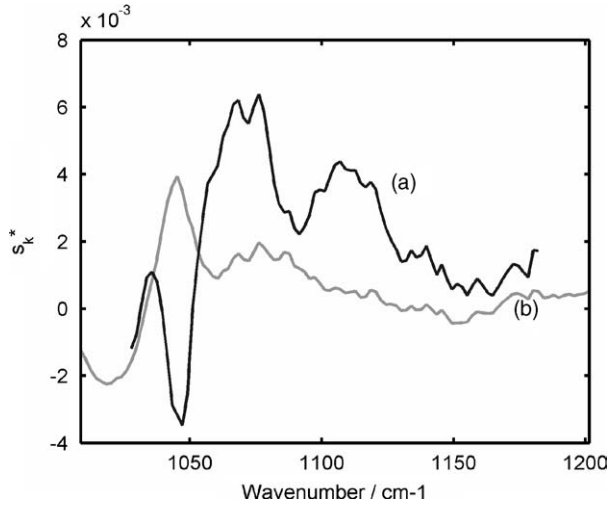

Fig. 4. Net analyte sensitivity (s_k^*) vector for: (a) real extract and (b) original extract.

Table 7
Prediction capabilities of PLS-ATR-FTIR for alcohol content (%, v/v) determination in beers

Data set	Alcohol		
	Validation	Data set 2	Data set 2*
RMSECV (%, v/v)	0.12	0.12	0.13
RMSEP (%, v/v)	0.14	0.12	0.12
RRMSEP (%)	3.0	2.5	2.5
d_{x-y} (%, v/v)	−0.003	0.059	0.058
s_{x-y} (%, v/v)	0.130	0.103	0.105
QC (%)	2.8	2.6	2.6
s_{trip} (%, v/v)	0.07	0.04	0.03
s_{reg} (%, v/v)	0.03	0.03	0.02
Selectivity (%)	42.6	42.6	45.1
Sensitivity	2.04E−02	2.04E−02	2.15E−02

Note: Two factors were extracted in all the cases using the spectral range from 997 to 1055 cm^{-1}. For additional details see the footnote of Table 6 and/or the text.

It should be remembered that these good results were obtained without applying any calibration transfer algorithm, even though the spectra of these samples were acquired nearly 6 months after those of data set 1 and at a different temperature.

The reproducibility of the determination, established from the pooled standard deviation of each triplicate and from the standard error of prediction, was 0.04 and 0.03% (v/v), respectively. Therefore, the uncertainty of the calibration model is not significantly different to the uncertainty of the measurement. As it can be also seen from Table 7, there is no significant difference between alcohol content values predicted by the proposed ATR-FTIR methodology and that provided by reference technique. This conclusion is based on the fact that the d_{x-y} value (0.06%, v/v) is nearly the half of the s_{x-y} value (0.11%, v/v), thus evidencing the absence of systematic errors in the PLS-ATR-FTIR methodology. The maximum percentage error to be expected for a new estimated real extract should be 2.8%.

Regarding sensitivity and selectivity of the alcohol content determination through PLS-ATR it was found, by using net analyte signal treatment, values of 2×10^{-2} and 42.6%, respectively, with a trend of increasing the selectivity on using an extended calibration model (45.1%).

4. Conclusions

In this work, different aspects for the ATR-FTIR estimation of original and real extracts and alcohol content in beers were discussed. The use of ATR measurements in front of the use of transmission ones can be justified by the good signal to noise ratio obtained in a wide range of wavenumbers, being less affected by the presence of water than transmission measurements. It is also interesting to note that for all these parameters, the most important IR spectral region is that of the OH features, namely between 1200 and 1000 cm^{-1}.

Hierarchical cluster analysis has been used for selecting samples for calibration data set, using triplicate measurements distance as internal distance to compare how different the clusters are. The calibration set selected from this analysis proved to be a good representation of the whole population of samples analysed. Moreover, a rough classification of samples, from the mid-IR spectra, can be achieved, being mainly related, in first place, to the original extract and secondly on the alcohol content. It is supposed that clustering of most similar samples may be to the type and content of different carbohydrates in the beer, but their determination has not been carried out.

The performance of the developed methodology favourably compares with that obtained by the reference methodology.

On the other hand, as compared with previously published works on chemometric-based vibrational procedures, it is the single precedent on the ATR-FTIR determination of alcohol, real and original extracts on beers with a relative prediction error of the order of 1.5% for ethanol, 2.8% for real extract and 1.9% for original extract which are better than those reported by PLS-NIR [9].

Acknowledgments

The financial support of the Oficina de Ciència I Tecnològia de la Conselleria d′ Inovació i Competetivitat de la Generalitat-Valenciana (Project GV 04B-247, Grupos 03–118 and invited Professor FAI grant), Universitat de València (Project UV-AE-20050203) and University of Buenos Aires (X013) is acknowledged.

References

[1] A.H. Varnam, J.P. Sutherland, Bebidas: Tecnología, Química y Microbiología, Editorial Acribia, S.A., Zaragoza, 1996.
[2] Council Directive 92/84/EEC: OJ L 316, 1992, 31.10.1992 Bulletin, 10-1992.
[3] European Brewery Convention, Analytica EBC, fourth ed., Brauerei und Getränke Rundschau, Zurich, 1987.
[4] L.S. Mendes, F.C.C. Oliveira, P.A.Z. Suarez, J.C. Rubim, Anal. Chim. Acta 493 (2003) 219.
[5] L. Norgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl. Spectrosc. 54 (2000) 413.
[6] M. Gallignani, S. Garrigues, M. de la Guardia, Anal. Chim. Acta 296 (1994) 155.
[7] M. Gallignani, S. Garrigues, M. de la Guardia, Analyst 118 (1993) 1167.
[8] S. Engelhard, H.G. Lohmannsroben, F. Schael, Appl. Spectrosc. 58 (2004) 1205.
[9] M. Maudoux, S.H. Yan, S. Collin, J. Near Infrared Spectrosc. 6 (1998) A363.
[10] F. Westad, H. Martens, J. Near Infrared Spectrosc. 8 (2000) 117.
[11] C.P. Malone, P.A. Flournoy, Spectrochim. Acta 21 (1965) 1361.
[12] L. Kupper, H.M. Heise, L.N. Butvina, J. Mol. Struct. 563–564 (2001) 173.
[13] I.F. Duarte, A. Barros, C. Almeida, M. Spraul, A.M. Gil, J. Agric. Food Chem. 52 (2004) 1031.
[14] F.A. Inon, J.M. Garrigues, S. Garrigues, A. Molina, M. de la Guardia, Anal. Chim. Acta 489 (2003) 59.
[15] A. Lorber, K. Faber, B.R. Kowalski, Anal. Chem. 69 (1997) 1620.
[16] F.A. Inon, S. Garrigues, M. de la Guardia, Anal. Chim. Acta 513 (2004) 401.
[17] A.C. Olivieri, H.C. Goicoechea, F.A. Inon, Chemom. Intell. Lab. Syst. 73 (2004) 189.
[18] D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. de Jong, P.J. Lewi, J. Verbeke, Handbook of Chemometrics and Qualimetrics (Part A and B), Elsevier Science B.V., Amsterdam, 1997.
[19] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193.

[20] D. Lin-Vien, N.B. Colthup, W.B. Fateley, J.G. Graselli, The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules, Academic Press, Boston, 1991.
[21] B. Everitt, Cluster Analysis, second ed., Heinemann Educational Books Ltd., London, 1981.
[22] B.D. Ripley, Pattern Classification and Neural Networks, Cambridge University Press, Cambridge, 1996.
[23] K. Faber, B.R. Kowalski, J. Chemom. 11 (1997) 181.
[24] A.C. Olivieri, J. Chemom. 16 (2002) 207.