Talanta 99 (2012) 426–432

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Confirmation of brand identity of a Trappist beer by mid-infrared spectroscopy coupled with multivariate data analysis

Jasper Engel [a,*], Lionel Blanchet [a], Lutgarde M.C. Buydens [a], Gerard Downey [b]

[a] Analytical Chemistry, Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands
[b] Teagasc Food Research Centre, Ashtown, Dublin 15, Ireland

ARTICLE INFO

Article history:
Received 9 February 2012
Received in revised form
31 May 2012
Accepted 4 June 2012
Available online 9 June 2012

Keywords:
Trappist beer
Authenticity
Fourier-transform infrared spectroscopy
Extended canonical variates analysis
Robust principal component analysis

ABSTRACT

Authentication of foods is of importance both to consumers and producers for e.g. confidence in label descriptions and brand protection, respectively. The authentication of beers has received limited attention and in most cases only small data sets were analysed. In this study, Fourier-transform infrared attenuated total reflectance (FT-IR ATR) spectroscopy was applied to a set of 267 beers (53 different brands) to confirm claimed identity for samples of a single beer brand based on their spectral profiles. Skewness-adjusted robust principal component analysis (ROBPCA) was deployed to detect outliers in the data. Subsequently, extended canonical variates analysis (ECVA) was used to reduce the dimensionality of the data while simultaneously achieving maximum class separation. Finally, the reduced data were used as inputs to various linear and non-linear classifiers. Work focused on the specific identification of Rochefort 8° (a Trappist beer) and both direct and indirect (using an hierarchical approach) identification strategies were studied. For the classification problems Rochefort vs. non-Rochefort, Rochefort 8° vs. non-Rochefort 8° and Rochefort 8° vs. Rochefort 6° and 10°, correct prediction abilities of 93.8%, 93.3% and 97.3%, respectively were achieved.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Food authentication is a wide-ranging issue that has received much attention in recent years. It is defined as the process by which a food is verified as complying with its label description [1]. Authentication of foods is of importance both to consumers and producers; trust in the claims made on a label is essential for consumer confidence while product authentication to prevent unfair competition with adulterated products is important for producers and processors.

The verification of claimed brand identity among Trappist beers is an example of a food authentication problem. These specialty beers, which are only brewed by monks in a Trappist monastery, enjoy a reputation of being of extremely high quality. To profit from the success of Trappist beers, commercial breweries introduced beers which mimicked the styles and names of Trappist beers (e.g. by licensing the name of a long-closed Trappist abbey). Because of this, the International Trappist Association introduced the "Authentic Trappist Product" logo [2]. This logo assures a customer that the product is a genuine Trappist beer, brewed according to tradition and specific quality standards. Moreover, the logo indicates that the economic purpose of the brewery is directed only to support of the monastery instead of financial profit. Nowadays, eight brands may carry the "Authentic Trappist Product" logo; these include six brands originating in Belgium (Achel, Chimay, Orval, Rochefort, Westmalle and Westvleteren), one from The Netherlands (La Trappe) and one from France (Mont des Cats). The Mont des Cats brand was only recently introduced (June 2011) and was therefore not considered in this work. The availability of an analytical methodology to confirm the claimed identity of any of these beers would permit brewers and regulatory authorities to uncover fraudulent labelling.

Until recently, the authentication of beer products has received limited attention, and in some cases only small data sets were analysed. For example, Weeranantanaphan et al. published a study on Trappist beer identity using ultraviolet (UV) spectroscopy and involving 52 samples [3]. Larger data sets of specialty beers (including all Trappist brands except for Mont des Cats) were analysed by means of several fingerprinting techniques as part of the EU-funded TRACE project (http://www.trace.eu.org). The results showed that metabolic profiles collected by liquid chromatography–mass spectrometry (LC–MS) can be used to classify Rochefort beers [4]. Alternatively, a fingerprint of the beer volatiles obtained by headspace solid-phase micro-extraction coupled to gas chromatography–mass spectrometry (HS-SPME-GC–MS) can also be used for this purpose [5]. The latter approach was shown to be successful in discriminating Trappist

* Corresponding author. Tel.: +31 24 3653135; fax: +31 24 3652653.
E-mail address: j.engel@science.ru.nl (J. Engel).

0039-9140/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.06.005

beers in general from other, non-Trappist, beers. A disadvantage of these chromatographic methods, however, is the high cost of the equipment [6].

In an industrial setting, the use of near-infrared (NIR) or mid-infrared (MIR) spectroscopy would be more suitable, as they are easy to use and involve relatively low purchase and running costs [6]. The value of these techniques for the authentication of Rochefort 8° beers was partially explored as a challenge to the Francophone part of Europe during the 'Chimiométrie' congress held in Paris in December 2010 [7]. In this challenge, a subset of the beers studied in the TRACE project was measured by NIR, MIR and Raman spectroscopy and subsequently analysed using four different approaches. It was shown that interpretation of sequential principal component analysis (PCA) models applied to the Raman data resulted in a perfect authentication of a particular Rochefort beer—Rochefort 8°. Since no real classification model was constructed, however, authentication was based on the interpretation of the PCA models by the operator, which makes the whole process rather subjective and, moreover, requires expert knowledge regarding PCA. Therefore, "true" classification models which allow one to objectively authenticate beer products without the need for expert knowledge were developed as well. Unfortunately, no satisfactory results were obtained when using the individual NIR, MIR or Raman data. However, Rochefort 8° beers were identified with an accuracy of 93.0% using a non-linear support vector machine (SVM) classifier coupled with data fusion to combine the classification results from the individual platforms (NIR, MIR, and Raman). This method will be referred to as SVMfusion from now on.

The use of NIR spectroscopy for the authentication of Trappist beers was more thoroughly explored by Di Egidio et al. [8]. In this study, the complete set of beers from the TRACE project was measured and subsequently analysed using several class-modelling techniques. This approach proved unsuitable for the discrimination between Trappist and non-Trappist beers in general but, in an alternative Rochefort vs. non-Rochefort classification, Rochefort beers were identified with an accuracy of 84.5%. Furthermore, a partial least squares – discriminant analysis (PLS-DA) model was found to be suitable for distinguishing Rochefort 8° from Rochefort 10° beers (93.4% accuracy).

In this study, Fourier-transform infrared (FT-IR) attenuated total reflectance (ATR) spectroscopy was applied to confirm claimed brand identity of a specific beer (Rochefort 8°) based on beer spectral profiles. Due to its greater molecular selectivity, the use of MIR spectroscopy often provides easier spectral interpretation than NIR, which is more commonly applied for the analysis of foods. Moreover, Lachenmeier et al. showed that FT-IR is a powerful tool for predicting the content of a range of compounds present in beer [9]. Our research attempted the identification both directly (i.e. in a single step) and via an hierarchical approach involving the consecutive identification or authentication of the beer sub-classes Trappist, Rochefort, and finally Rochefort 8° beers. The classification was preceded by extensive chemometric analysis including the use of various options for pre-processing of the spectra (e.g. to correct for baseline differences), robust principal component analysis (ROBPCA) for outlier detection and extended canonical variates analysis (ECVA) for supervised dimension reduction of the data. Finally, the reduced data were used as inputs to various linear and non-linear classifiers namely linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and *k*-nearest neighbours (*k*-NN).

So, our aim was to show that FT-IR can be used for authentication of Rochefort 8° beers; if successful, this would provide a relatively cheap alternative for authentication of these beers compared to previously published results. In the final part of the work, our results were compared to those obtained by NIR spectroscopy on the same sample set using previously published data [8].

2. Materials and methods

2.1. Sample preparation

A set of 267 bottles of both Trappist and non-Trappist beers was analysed in this study. The non-Trappist group mostly contained specialty beers (e.g. Abbey beers) which originated from Belgium together with two Czech beers (Primator 21 and 24). More details regarding the beers investigated in this study are included as Supplemental material. Beer samples were collected from retail outlets in two batches (set 1 and 2) to cover possible seasonal variability of the products; they were received in Ashtown, Ireland in September 2008 and January 2009, respectively and immediately placed in storage at 4 °C in the dark. In general, the same number of bottles (at least two) of each brand was collected in each batch although that was not always possible given the small and often seasonal production runs of some of the sample types. It is obvious that the limited number of samples from any given brewery cannot be representative of total production but, on the other hand, the aim of this study was to maximise the variance among beer types to fully challenge any classification model developed and this necessitated a compromise on the numbers of replicate beer samples.

One day before spectral analysis, an appropriate number of beer bottles was randomly selected, removed from storage and left on a laboratory bench out of direct sunlight to equilibrate to room temperature (21 ± 5 °C) for 24 h. Shortly before analysis, each bottle was gently uncapped, left undisturbed for 5 min to allow any foaming to subside after which an aliquot of beer was removed from the bottle using a Pasteur pipette inserted about half-way down the bottle. This procedure was adopted to facilitate the simplest possible sample handling protocol. The aliquot was immediately scanned. No major discernible problems in spectral quality were encountered as a result of gas bubbles possibly because of the ATR sample presentation option selected.

2.2. FT-IR spectroscopy

FT-IR spectra were acquired on a BIO-RAD (Philadelphia, PA, USA) Excalibur series FTS 300 FT-IR spectrometer. Each beer sample (approximately 2 ml) was transferred by pipette onto an in-compartment benchmark ATR trough plate containing a 45° germanium crystal with 11 internal reflections (Specac Ltd., Kent, UK). Care was taken to minimise the presence of bubbles. For each sample, a single beam spectrum of 64 co-added scans at a nominal resolution of 4 cm^{-1} was recorded. Subsequently, a reference background spectrum of air was subtracted. Spectra were truncated to 800–4000 cm^{-1}. Between samples, the ATR crystal surface was rinsed with distilled water and dried with soft tissue. The spectral baseline subsequently recorded by the spectrometer was examined visually to ensure that no residue from the previous sample was retained on the ATR crystal. All spectra were recorded without any nitrogen purge of the sample compartment. Beers were analysed in random order; duplicate spectra of each sample were collected and averaged before input to the data analysis steps.

2.3. Data analysis strategy

After spectral acquisition, the dataset was split into a calibration (training) set consisting of 172 of the samples and a test set

consisting of the remaining 95 samples. For this purpose every third beer measured was selected as a test object.

Next, extensive chemometric analysis was performed to construct a model that could identify Rochefort 8° beers based on spectral profiles. This analysis was part of a larger project where in general the effects of different choices/approaches for data pre-treatment, outlier detection and classification were investigated since, for example, the wrong choice of pre-treatment or the presence of outliers can heavily influence the performance of a final model. Here, only the methods and results that are relevant to the current study will be described.

2.4. Data pre-treatment

Various techniques were investigated for data pre-treatment since the optimal method could not be selected on the basis of visual inspection of the spectra. Asymmetric least squares smoothing (AsLS), Savitzky–Golay (SG) derivatisation (1st and 2nd order) and standard normal variate (SNV) transformations were used [10–12]. It must be noted that these methods do not correct for the same effects; AsLS and SG derivatisation can correct differences in offset and slope while SNV corrects for differences in offset and effects due to multiplicative scattering [13]. Therefore, the combinations AsLS+SNV and SG derivatisation+SNV were used for pre-processing of the data as well. Note that AsLS requires the optimisation of two tuning parameters λ_{AsLS} and p_{AsLS}. This was achieved by visual inspection of the corrected calibration spectra during a grid search (the AsLS parameters were varied between $10^2 \le \lambda_{AsLS} \le 10^9$ and $0.001 \le p_{AsLS} \le 0.1$ on a logarithmic grid). A window size of 11 variables followed by a second order polynomial smoothing step was used during SG derivatisation.

In addition to the various pre-processing options, a final smoothing step of the spectra was also considered. For this purpose, a SG filter (window size of 11 plus second order polynomial) was used [11]. In total, therefore, 16 forms of pre-treatment, namely the various pre-processing options (including no pre-processing whatsoever) with and without a subsequent smoothing step, were investigated.

2.5. Outlier detection

PCA (in combination with Hotelling's T^2 and the Q statistic) can be used to identify potential outliers in calibration data [14–17]. Outliers in the test data were detected, after pre-treatment of the spectra, by projection of the objects onto the PCA hyperplane calculated for the calibration data.

Unfortunately, the standard PCA algorithm is very sensitive to outliers. Consequently, this method may not detect all outlying objects while, conversely, good objects may appear to be outliers (the so-called masking and swamping effects) [18]. Robust methods such as ROBPCA aim to estimate PCs using a majority of the data (the most similar data points) and thus ignore the negative effect of outliers upon the PCA models. For these reasons, the standard PCA algorithm was replaced by a skewness-adjusted version of ROBPCA, which can handle both symmetric and skewed data distributions [19]. The default settings of the algorithm were used, except for the proportion of objects used to estimate the robust PCs - this value was set to 85%. The number of PCs selected for retention was based on the breakpoint in the corresponding scree plot.

Note that ROBPCA flags objects that differ from the majority of the data as outlier. This difference can be caused by experimental errors, but also by scatter effects (for example). The latter can be corrected for by appropriate data pre-treatment. Therefore, data was pre-treated before outlier detection in this study, and only pre-treatment methods that are not influenced by outliers were used. In this way, a maximum number of samples was retained for subsequent data analysis.

2.6. Dimension reduction and classification

After outlier detection and removal, discriminant models for five classification problems were constructed (see Table 1). Note that these models were chosen such that direct and indirect (using a hierarchical approach) identification of Rochefort 8° was possible (see results). The 4class problem was studied to see if multiple brands of beer can be classified at the same time since this could be useful for practical brand identification problems.

The models were constructed in two steps namely dimension reduction followed by classification. PCA is a commonly-used method to compress data and visualise between-class differences. However, PCA aims to describe all variance in the data which may result in suboptimal results and/or needlessly complicated interpretation of the final classification model [20,21]. For our purposes, only the component(s) related to between-class differences are required and therefore ECVA was used in this study [21]. This method does not suffer from the above mentioned pitfalls since it performs supervised dimension reduction together with classification in accordance with LDA.

The concept of ECVA very much resembles Fisher's version of LDA (F-LDA). It aims to find directions or canonical variates (**CV**s) in the data along which a maximum separation of the classes of interest is achieved while the scatter within classes is simultaneously kept to a minimum. Assuming a data matrix $\mathbf{X}$ $(n \times m)$ where the samples are divided into g groups with n_i samples in the ith group, the within-class scatter matrix $\mathbf{S_w}$ is defined as

$$\mathbf{S_w} = \frac{1}{n-g}\sum_{i=1}^{g}\sum_{j=1}^{n_i}(\mathbf{x_{ij}}-\overline{\mathbf{x}}_\mathbf{i})(\mathbf{x_{ij}}-\overline{\mathbf{x}}_\mathbf{i})' \quad (1)$$

and between-class scatter matrix $\mathbf{S_b}$ is defined as

$$\mathbf{S_b} = \frac{1}{g-1}\sum_{i=1}^{g}(x_i-\overline{\mathbf{x}})(x_i-\overline{\mathbf{x}})' \quad (2)$$

where $\mathbf{x_{ij}}$ is the jth sample belonging to the ith group, $\overline{\mathbf{x}}_i$ the mean vector in the ith group and $\overline{\mathbf{x}}$ the overall mean vector. F-LDA can

Table 1
The number of calibration and test set samples for each discrimination problem.

Problem	Name	Calibration set (# samples)	Test set (# samples)
Four class problem	4class		
non-Trappist		94	36
Rochefort 8°		29	20
Rochefort 10°		12	13
Trappist (without Rochefort 8° and 10°)		37	26
Rochefort 8° vs. rest	R8-nonR8		
Rochefort 8°		29	20
Rest (non-Trappist and Trappist without Rochefort 8)		143	75
Rochefort 6°, 8°, 10° vs. rest	R-nonR		
Rochefort 6°, 8°, 10°		45	35
Rest (non-Trappist and Trappist without Rochefort 6°, 8°, 10°)		127	60
Rochefort 8° vs. Rochefort 6°, 10°	R8-R6R10		
Rochefort 8°		29	20
Rochefort 6°, 10°		16	15
Trappist vs. non-Trappist	T-nonT		
Trappist		78	59
non-Trappist		94	36

be formulated as the problem of finding those **CV** that maximise

$$j(\mathbf{CV}) = \frac{\mathbf{CV}'\mathbf{S_b}\mathbf{CV}}{\mathbf{CV}'\mathbf{S_w}\mathbf{CV}} \quad (3)$$

A unique solution to (3) can be found by solving the following generalised eigenvalue problem

$$\mathbf{S_w^{-1}S_bCV} = \lambda\mathbf{CV} \quad (4)$$

This equation has $a = \min(m,g-1)$ non-zero eigenvalues. This means that for high dimensional data (large m) the number of variables is reduced to the number of classes minus 1 e.g. one variable for a two-class problem. However, when the matrix $\mathbf{S_w}$ is singular it is not possible to left multiply by the inverse of this matrix as in (4). This is the cause of the breakdown of F-LDA for under sampled ($n < m$) data (for example). ECVA circumvents this problem by transforming (4) into an expression that can be solved by regression (in this case PLS)

$$\mathbf{Y} = \mathbf{S_w B} + \mathbf{E} \quad (5)$$

where **Y** contains as columns the differences ($\overline{\mathbf{x}}_\mathbf{i} - \overline{\mathbf{x}}$) and the columns of **B** are used to obtain the CVs. The matrix **E** represents the residual error. Finally, dimension reduction is achieved by multiplication of **X** with **CV** and a discriminant function is constructed by means of the LDA classifier (see below). Note that the PLS step in ECVA requires optimisation of a number of latent variables (LVs); in this study, the number was varied between 1 and 50.

As mentioned above, ECVA uses an LDA classifier. However, there is no restriction on which classifier should be applied. Therefore, the reduced data was also used as an input to a QDA and k-NN ($k = 1, 3, 5$) classifier [22] with Euclidian distance being used to determine the nearest neighbours. These popular classifiers (including LDA) perform well on amazingly large and diverse classification tasks [22]. Note that LDA fits a multivariate normal density to each class with the assumption of equal class distributions (a pooled estimate of covariance is used). In this case, different classes are identified by means of linear class boundaries (hyperplanes). QDA is a generalised version of LDA: no equal covariance is assumed and the class boundary between a pair of classes is described by a quadratic function. The k-NN classifier makes no distributional assumptions and is often successful when the decision boundary is very irregular since non-linear class boundaries can be described.

2.7. Selection of the final model

The predictive performance of each model (a combination of pre-treatment + number of LVs in ECVA + classifier) was quantified by calculating the geometric mean of correctly classified samples. This measure, instead of the arithmetic mean, was used because the number of samples in each class was unequal. For each classification problem, the most predictive model was found by a 20-fold cross-validation procedure that was applied to the calibration data. Next, this model was applied to the independent test set.

From now on, the predictive performance found for the calibration and the test data will be referred to as the recognition and prediction ability respectively. For binary classification problems, the sensitivity, and specificity of the classifier were computed as well. Here, the sensitivity and specificity were defined as the proportion of samples corresponding to respectively class one and class two which were correctly classified. In the final step of the analysis, the CVs found by ECVA for the final models were studied to identify which spectral regions were used to achieve class separation.

2.8. Software

All data analysis described in this paper was performed in MatLab (R2009b, The MathWorks, Natick, MA, USA). In general, algorithms developed in-house were used for this purpose. For the skewness-adjusted ROBPCA, however, the function from the Libra toolbox was used [23]. For ECVA, the algorithm from the ECVA toolbox was used [21].

3. Results

This study focused on the specific identification of Rochefort 8° (a Trappist beer) and both direct and indirect (using an hierarchical approach) identification strategies were studied (see Table 1). The direct, and maybe most logical, approach involves construction of the R8-nonR8 model, while hierarchical approaches sequentially applied models listed in the table (e.g. R-nonR followed by R8-R6R10). The 4class problem was studied to see if multiple brands of beer can be classified at the same time; this could be useful for practical brand identification problems.

3.1. FT-IR spectra

As a first step, the FT-IR spectra were visually inspected. A number of different labelling procedures were applied to these spectra (e.g. labelling samples as Rochefort or non-Rochefort, etc.). However, in none of these cases (Fig. 1a) was any visual differentiation between the sample types possible indicating that more extensive multivariate data analysis was required.

Based on plots of different bottles of the same beer, such as shown in Fig. 1b, differences between spectra due to multiplicative effects were expected; these were also visible in the original spectra, for instance between 3700 and 2900 cm^{-1}. Such multiplicative effects may have been caused by light scattering due to very small bubbles in the beer since the samples were not degassed before spectral analysis. Interestingly, the multiplicative effects appeared to be much stronger between samples originating from different batches although it will be shown that these effects did not influence identification of the beers. Furthermore, slight baseline differences were observed between batches which might be related to instrument drift.

The broad band in the spectra around 3360 cm^{-1}, and the bands around 2130 and 1640 cm^{-1} were due to water, while that at 2348 cm^{-1} was due to CO_2 [24]. Additional bands were located mostly in the fingerprint region of the spectrum, the major ones being around 2983, 1151, 1084, 1045 and 878 cm^{-1}. A shoulder at 1026 cm^{-1} which showed large variation between the beer samples was also visible. Ascription of these peaks is difficult in biological matrices given their complexity but absorptions around 1053, 1099 and 1149 cm^{-1} have been identified as the fundamental absorption bands of primary, secondary and tertiary alcohols [25].

Most of the above mentioned peaks, such as the water peaks, were not expected to be useful in discriminating the different beers and the spectra could therefore be truncated to the fingerprint region between 800 and 1500 cm^{-1}. However, our models should automatically ignore useless peaks since they were constructed in a supervised manner and truncation was therefore not considered in this study.

3.2. Outlier detection

After spectral pre-treatment, ROBPCA for skewed data was used for outlier detection and removal [19]. In general, the first 2–6 robust PCs, explaining > 95% of the variance in the data, were

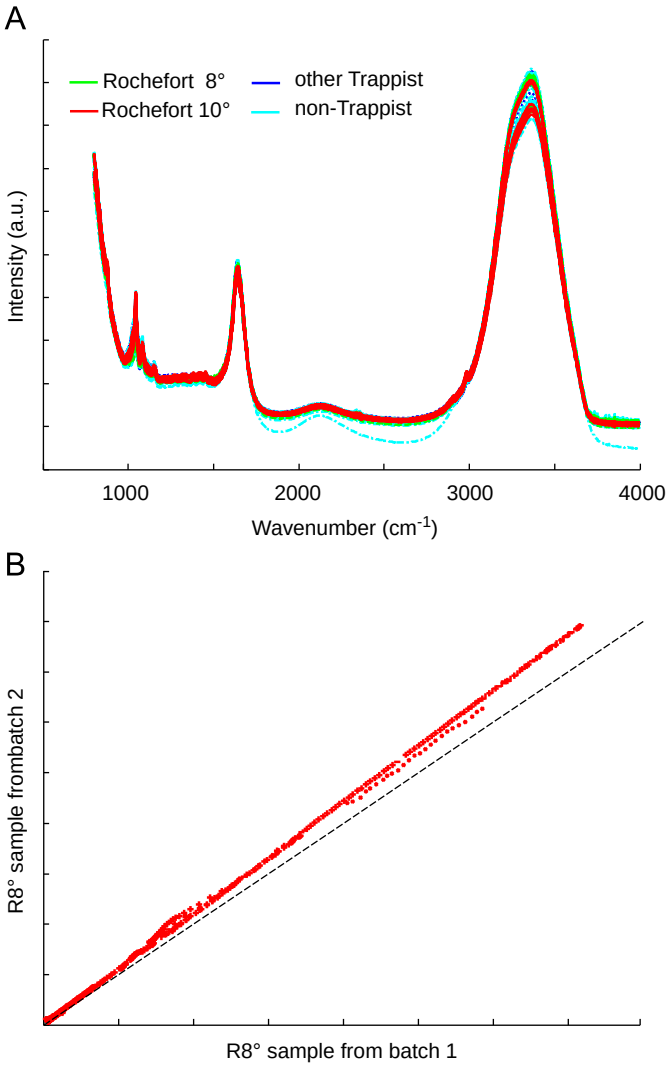

Fig. 1. (a) Raw FT-IR spectra of all samples and (b) raw spectra of two Rochefort 8° beers from different batches plotted against each other. The dashed black line in (b) represents the ideal relationship between identical spectra.

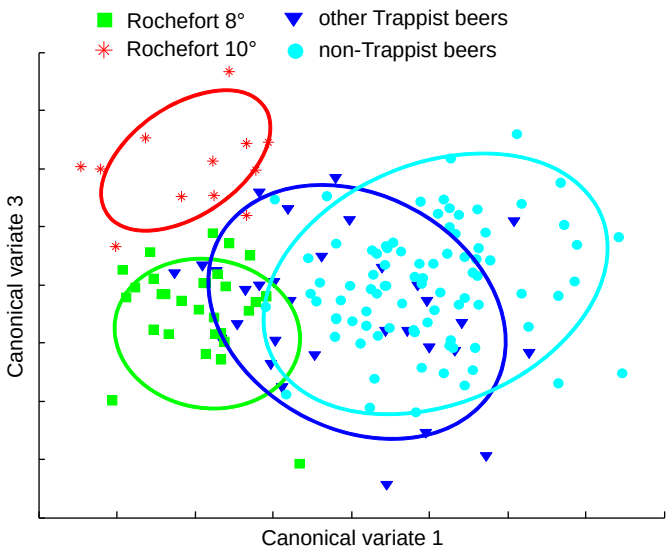

Fig. 2. Projections of the MIR data onto the most discriminative canonical variates.

Table 2
Summary of discriminant model performance for each classification problem. Note that two strategies were used for the R8-nonR8 problem: (1) direct discrimination and (2) sequential application of the R-nonR model followed by the R8-R6R10 model. The abbreviations gmet, sens and spec respectively indicate the geometric mean, sensitivity and specificity.

Problem name	Calibration set	Test set					
		Without outliers			With outliers		
	gmet	gmet	sens	spec	gmet	sens	spec
4class	61.3	62.7	x	x	60.4	x	x
R-nonR	91.8	93.8	96.7	91.1	91.6	91.4	91.6
R8-R6R10	93.1	97.3	94.7	100.0	97.5	95.0	100.0
R8-nonR8 —direct	89.1	89.7	100.0	79.4	87.9	95.0	81.3
R8-nonR8 —sequential	–	93.3	97.1	89.5	93.0	96.0	90.0
T-nonT	81.7	76.8	88.6	66.7	75.7	88.9	64.4

selected based on the first breakpoint in a scree plot. Between 2 and 8% of the objects were flagged as outliers in each of the differently pre-treated data sets. The flagged outliers included objects that were clearly seen to differ from the majority by visual inspection of the spectra; they could not be related to a specific brand of beer.

3.3. *Dimension reduction and classification*

Visualisation of the data after dimension reduction was used to see if chemometrics (ECVA) could help to identify the beers compared to visual inspection of the data. It was found that the application of ECVA resulted in clear separation between the different groups of interest. In Fig. 2, a score plot of the two canonical variates (CV1 vs. CV3) for which the greatest separation was found for the 4class problem is shown as an example. Note that the data was smoothed and 17 LVs were used in ECVA to obtain this separation. Based on Fig. 2, it can be seen that in general the Trappist and non-Trappist beers overlapped heavily while Rochefort 8° and 10° beers were somewhat separated from the rest. This pattern was also observed for the other classification problems indicated in Table 1. Furthermore, the shape and size of the four classes seemed to differ slightly, indicating an unequal covariance structure. This is not unexpected since the scatter in the Rochefort classes arose from variation between the different bottles of beer while the scatter in the other classes can also be partly due to variation between different brands of beer.

The projection of data onto canonical variates is a useful visualisation tool, but it does not necessarily provide an objective means of authentication of new samples. Therefore, the data were used as inputs to various classifiers after dimension reduction by ECVA. An LDA classifier was used by default because both ECVA and LDA are optimal for homoscedastic class distributions [21,22]. Since the distributions of the classes investigated, as judged by visual inspection of Fig. 2, did not seem to be homoscedastic, the reduced data was also used as an input to the QDA and *k*-NN classifiers [22]. As mentioned above, these classifiers make different assumptions about the data compared to LDA and might be able to deal more efficiently with the data at hand.

In Table 2, the specifics of the model for which the highest recognition ability was achieved are shown for each classification problem studied. The models for which a predictive performance above 90% was observed are deemed suitable for industrial use. As shown in the table, a correct classification rate slightly below 90% was achieved when directly trying to discriminate the Rochefort 8° beers from the other brands. However, a second

strategy to discriminate these beers could also be applied. This involved a sequential approach in which beers were first classified as either Rochefort or non-Rochefort (by the R-nonR model) after which the Rochefort beers were subsequently classified as Rochefort 8° or Rochefort 6°, 10° (by the R8-R6R10 model). This model resulted in a better performance with a prediction ability of 93.3%. Note that a hierarchical approach could also make use of the T-nonT model but, as suspected by the large overlap between the Trappist and non-Trappist classes evident in Fig. 2, this approach was not viable: many Trappist beers were misclassified as non-Trappist and vice versa. The low prediction ability of the 4class model indicated that a multi-group approach was not possible in this case either. Interestingly, all models deemed suitable for industrial use may be regarded as robust given that the accuracy for both test sets (with and without outliers) differed by less than 2%. For none of the classification problems could misclassified beers be related to any specific brand.

After the "best" models were selected, canonical weight vectors obtained by ECVA were studied to determine which areas of the MIR spectra contained important information for the discrimination of the beers. These weight vectors are analogous to PCA loadings and can be interpreted in a similar manner. Note that this interpretation was only performed for the models that were acceptable for industrial use. As shown in Fig. 3a, the canonical weight vector that was obtained for the R8-R6R10 problem revealed that the different Rochefort beers were mostly identified by features around 1000–1200 cm^{-1}. This was confirmed by repeating the analysis (including pre-treatment and outlier detection) using spectra truncated to this region. Unfortunately, as mentioned above, ascription of these peaks is difficult for biological matrices. For the R-nonR problem, interpretation of the canonical weight vector was much more difficult since the absolute values of weights for the baseline and peaks were more similar (see Fig. 3b). This noisy structure of the weight vector was reflected in the large number of LVs that was used in ECVA (see Table 3). In general, all major peaks in the spectrum seemed to have a comparable weight which was slightly higher compared to baseline areas. Therefore, no discriminative region could be identified for this particular problem.

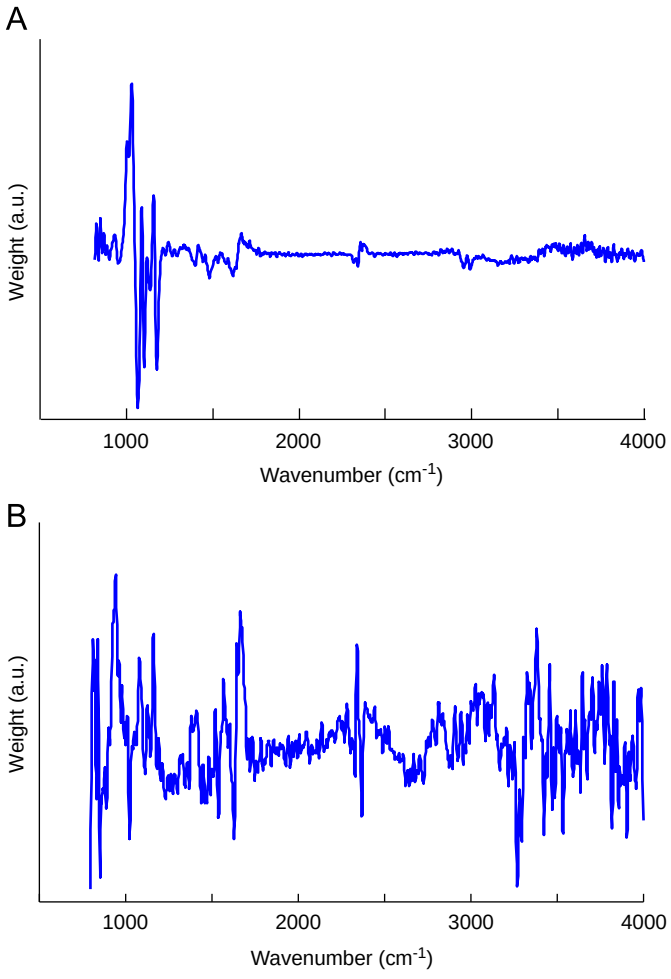

Fig. 3. Canonical weights obtained when ECVA was applied to the MIR data for (a) the R8-R6R10 and (b) the R-nonR problems, respectively.

Table 3
Summary of the pre-treatment, classifier and number of LVs used in the most discriminative models.

Problem name	Smoothing	Pre-processing	Classifier	# LV
4class	yes	None	QDA	17
R8-nonR8	yes	SNV	LDA	24
R-nonR	yes	None	LDA	25
R8-R6R10	yes	SG-deriv (1^{st})+SNV	1-NN	2
T-nonT	yes	None	5-NN	18

Table 3 provides details of the settings for the selected models e.g. which classifier was used. Interestingly, some of the selected models (4class, R-nonR and T-nonT) did not use any data pre-treatment; apparently, the ECVA model was not influenced by the observed multiplicative effects (Fig. 1b) and/or between-batch baseline differences. In the discussion, it will be argued that pre-treatment was not required for the R8-nonR8 and R8-R6R10 problems either. The large number of LVs selected in the ECVA models is a property of this technique that is often observed and does not necessarily indicate over-fitting of the training data [21]. However, as mentioned above it did seem to have implications regarding the interpretability of the models.

4. Discussion

As shown above, MIR spectroscopy in combination with ECVA was able to separate Rochefort 8° from non-Rochefort 8° beers with a prediction ability of 93.3%. This was achieved by the sequential application of the Rochefort vs. non-Rochefort and the Rochefort 8° vs. Rochefort 6°, 10° models (prediction abilities of 93.8% and 97.3%, respectively). For the Trappist vs. non-Trappist beer classification problem, a prediction ability of only 76.8% was achieved. It should be noted that the models considered in this study were optimised to maximise overall prediction ability. For certain applications, maximising the sensitivity or specificity, instead of the prediction ability, might be a more appropriate goal.

For the discrimination of Rochefort beers, our reported method is very competitive with the LC-MS, GC–MS and SVMfusion approaches that were mentioned previously. Compared to the MS methods, the advantage of MIR spectroscopy for the authentication of the beers is the relative low cost of the equipment. The same can be said when comparing our approach to the SVMfusion method since the latter requires the combination of NIR, MIR and Raman spectroscopy to achieve a similar accuracy. Moreover, the use of less complex chemometric techniques, as in this study, also allows for interpretation of the model, a task which is more complex for SVM models [26].

In contrast to the use of HS-SPME-GC–MS [5], our approach with MIR spectroscopy could not successfully discriminate Trappist from non-Trappist beers; the same negative result was observed when NIR spectroscopy was used for this purpose [8].

This can be related to lower information content in both the NIR and MIR data but may also be due to the chemometric technique(s) applied in the data analysis. Even though various linear and non-linear classifiers were investigated, the results may have been largely determined by the preceding application of the linear dimension reduction with ECVA. Identification of the beers may possibly be improved by using a kernel approach, such as SVM, that can deal with non-linear class boundaries [22] although a brief study involving an SVM with a Gaussian RBF kernel did not show any improvement over the approach that was reported in this report (results not shown). Therefore, we suggest that the imperfect discrimination between Trappist and non-Trappist beers was related to a lack of sufficient and relevant information in the data.

To compare the power of MIR to NIR spectroscopy for classification of the Rochefort beers, the chemometric analysis described in this work was applied to previously-published NIR data of the same set of beers [8]. This was done since NIR spectroscopy is a commonly-used technique for food authentication and seems intuitively more suitable for this particular problem due to a lower sensitivity to water and carbon dioxide [27]. However, it was found that the results obtained were similar for both spectroscopic methods; only small differences in predictive ability of at most 3% were observed. Note that the results obtained from the recalculation using the NIR data from [8] were not published.

Regarding the selected classification models for MIR data, further investigation revealed that similar recognition abilities ($<2\%$ difference) were obtained when pre-treatments and classifiers other than those listed in Table 3 were employed. This indicates that pre-processing of the data was not a critical step in the analysis. More specifically, neither smoothing nor corrections for multiplicative effects and baseline differences were required per se; the pre-treatment needed depended on the classifier applied. Furthermore, most classification problems (4class, T-nonT, R-nonR, and R8-nonR8) could actually be solved by an LDA classifier. This indicates that, even for heteroscedastic data, the application of ECVA with an LDA classifier is the best option in most cases and is most likely due to the fact that both techniques make the same assumptions regarding the data. In some cases, however, application of a non-LDA classifier after ECVA might be more fruitful, as shown by the R8-R6R10 problem for which the recognition ability increased by 8% this way.

5. Conclusion

The investigated combination of FT-IR ATR spectroscopy and subsequent chemometric data analysis could be applied as an effective tool for identification of Rochefort 8° beers using Rochefort vs. non-Rochefort and Rochefort 8° vs. Rochefort 6°, 10° models, resulting in an overall correct prediction ability of 93.3%. Prediction abilities of the two sub models are 93.8% and 97.3%, respectively. These results are very competitive with previously published work that used much more expensive analytical tools. Furthermore, our data analysis strategy allowed us to identify spectral regions which were important for the authentication of Rochefort beers (1000–1200 cm^{-1}). It is important to bear in mind that, while this work demonstrates the potential of this approach for beer brand confirmation, additional research into model repeatability and stability are required before any commercial or regulatory deployment would be warranted.

Identification of Trappist from non-Trappist beers with MIR spectroscopy does not seem to be a viable option: a prediction ability of only 76.8% was achieved. Therefore, GC–MS is clearly a more suitable analytical platform for this classification problem [5].

Acknowledgements

This research was initiated under the EU TRACE project (http://www.trace.eu.org). The information contained in this paper reflects the authors' views; the European Commission is not liable for any use of the information contained therein.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2012.06.005.

References

[1] M.J. Dennis, Analyst 123 (1998) 151R–156R.
[2] International Trappist Association. 〈http://www.trappist.be/indexjs.cfm?v=02&taal=en〉, 2011 (accessed November 2011).
[3] J. Weeranantanaphan, G. Downey, J. Inst. Brew. 116 (2010) 56–61.
[4] E. Mattarucchi, M. Stocchero, J.M. Moreno-Rojas, G. Giordano, F. Reniero, C. Guillou, J. Agric. Food Chem. 58 (2010) 12089–12095.
[5] T. Cajka, K. Riddellova, M. Tomaniova, J. Hajslova, J. Chromatogr. A 1217 (2010) 4195–4203.
[6] L.M. Reid, C.P. O'Donnell, G. Downey, Trends Food Sci. Technol. 17 (2006) 344–353.
[7] J.A. Fernández Pierna, L. Duponchel, C. Ruckebush, D. Bertrand, V. Baeten, P. Dardenne, Chemom. Intell. Lab. 113 (2012) 2–9.
[8] V. Di Egidio, P. Oliveri, T. Woodcock, G. Downey, Food Res. Int. 44 (2011) 544–549.
[9] D.W. Lachenmeier, Food Chem. 101 (2007) 825–832.
[10] P.H.C. Eilers, Anal. Chem. 76 (2004) 404–411.
[11] A. Savitzky, M.J.E. Golay, Anal. Chem. 36 (1964) 1627–1639.
[12] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Appl. Spectrosc. 43 (1989) 772–777.
[13] A.M.C. Davies, T. Fearn, Spectrosc. Eur. 19 (2007) 24–28.
[14] D. Cozzolino, H.E. Smyth, M. Gishen, J. Agric. Food Chem. 51 (2003) 7703–7708.
[15] C.B.Y. Cordella, J.S.L.T. Militão, M.-C. Clément, D. Cabrol-Bass, J. Agric. Food Chem. 51 (2003) 3234–3242.
[16] J. He, L.E. Rodriguez-Saona, M.M. Giusti, J. Agric. Food Chem. 55 (2007) 4443–4452.
[17] J.F. MacGregor, T. Kourti, Control Eng. Pract. 3 (1995) 403–414.
[18] S.F. Møller, J. von Frese, R. Bro, J. Chemometrics 19 (2005) 549–563.
[19] M. Hubert, P. Rousseeuw, T. Verdonck, Comput. Stat. Data Anal. 53 (2009) 2264–2274.
[20] P. Howland, H. Park, IEEE. Trans. Pattern Anal. Mach. Intell. 26 (2004) 995–1006.
[21] L. Nørgaard, R. Bro, F. Westad, S.B. Engelsen, J. Chemom. 20 (2006) 425–435.
[22] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning. Data Mining, Inference, and Prediction, first ed., Springer, New York, 2001.
[23] S. Verboven, M. Hubert, Chemom. Intell. Lab. 75 (2005) 127–136.
[24] J. Workman, The Handbook of Organic Compounds, Academic Press, London, 2001, pp. 209–242.
[25] D. Bertrand, in: D. Bertrand, E. Dufour (Eds.), La spectroscopie infrarouge at ses applications analytiques, 2nd ed.,Editions Tec & Doc, Paris, 2000, pp. 93–105.
[26] P.W.T. Krooshof, B.I. Üstün, G.J. Postma, L.M.C. Buydens, Anal. Chem. 82 (2010) 7000–7007.
[27] D.A. Skoog, F.J. Holler, T.A. Nieman, Principles of Instrumental Analysis, 5th ed., Saunders College Publishing, Fort Worth, 1992, pp. 93–105.