Talanta 190 (2018) 55–61

Contents lists available at ScienceDirect

Talanta

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

Variable selection for multivariate classification aiming to detect individual adulterants and their blends in grape nectars

Carolina Sheng Whei Miaw[a,b,c], Marcelo Martins Sena[d], Scheilla Vitorino Carvalho de Souza[a], Itziar Ruisanchez[c,*], Maria Pilar Callao[c]

[a] Department of Food Science, Faculty of Pharmacy (FAFAR), Federal University of Minas Gerais (UFMG), Av. Antônio Carlos, 6627, Campus da UFMG, Pampulha, 31270-010 Belo Horizonte, MG, Brazil
[b] CAPES Foundation, Ministry of Education of Brazil, 70040-020 Brasília, DF, Brazil
[c] Chemometrics, Qualimetric and Nanosensors Group, Department of Analytical and Organic Chemistry, Rovira i Virgili University, Marcel·lí Domingo s/n, 43007 Tarragona, Spain
[d] Department of Chemistry, Institute of Exact Sciences (ICEX), Federal University of Minas Gerais (UFMG), Campus da UFMG, Pampulha, 31270-010 Belo Horizonte, MG, Brazil

ARTICLE INFO

Keywords:
Variable selection
Multi-class methods
PLS-DA
SIMCA
Grape nectar
Food fraud

ABSTRACT

During the quality inspection control of fruit beverages, some types of adulterations can be detected, such as the addition or substitution with less expensive fruits. To determine whether grape nectars were adulterated by substitution with apple or cashew juice or by a mixture of both, a methodology based on attenuated total reflectance Fourier transform mid infrared spectroscopy (ATR-FTIR) and multivariate classification methods was proposed. Partial least squares discriminant analysis (PLS-DA) and soft independent modeling of class analogy (SIMCA) models were developed as multi-class methods (classes unadulterated, adulterated with cashew and adulterated with apple) with the full-spectra. PLS-DA presented better performance parameters than SIMCA in the classification of samples with just one adulterant, while poor results were achieved for samples with blends of two adulterants when using both classification methods. Three variable selection methods were tested in order to improve the effectiveness of the classification models: interval partial least squares (iPLS), variable importance in projection scores (VIP scores) and a genetic algorithm (GA). Variable selection methods improved the performance parameters for the SIMCA and PLS-DA methods when they were used to predict samples with only one adulterant. Only PLS-DA coupled with iPLS was able to classify samples with blends of two adulterants, providing sensitivity values between 100% and 83% at 100% specificity for the three studied classes.

1. Introduction

Adulteration or possible food fraud is a problem that affects many food products and has an important economic impact. In the case of fruit beverages, the most frequent types of adulteration include the addition of water or syrup, acidification, addition or substitution with cheaper fruits and addition of colorants or flavors. Unfortunately, fruit beverages are one of the easiest products to adulterate because of their complex chemical composition and the wide natural variation of fruits [1].

Specifically, fruit nectars are unfermented beverages intended for direct consumption, which are formulated by dilution of the edible part of the fruits or their extracts with water and added sugars [2]. The most commonly consumed flavor of fruit nectar in Brazil, and one of the most expensive, is grape. Often, consumers choose grape nectars looking for a more nutritional product, since this fruit is rich in phenolic compounds, mainly flavonoids [3]. Due to their sensory characteristics and lower cost, apple and cashew juices are likely to be used as adulterants in some of the more expensive fruit nectars. Additionally, apple and cashew are fruits being suspected to be used as fillers by fraudulent industries based on denunciations received by the Ministry of Agriculture, Livestock, and Supply – MAPA, and some evidence in fiscal activities. Since the commercialization of nectars containing more than one fruit has been expanded in the market, the declared presence of each and every single fruit must be confirmed in order to guarantee the authenticity of the single fruit nectars and to prevent adulterations [4].

Recent methods have been developed to identify and detect different fruits in fruit beverages employing various analytical techniques. The use of ultra-performance liquid chromatography–quadrupole time of flight mass spectrometry (UPLC–QToFMS) [5], high-performance

* Corresponding author.
E-mail address: itziar.ruisanchez@urv.cat (I. Ruisanchez).

https://doi.org/10.1016/j.talanta.2018.07.078
Received 11 June 2018; Received in revised form 19 July 2018; Accepted 23 July 2018
Available online 24 July 2018
0039-9140/ © 2018 Elsevier B.V. All rights reserved.

liquid chromatography [6], conventional [7] and real time polymerase chain reaction [7,8] are some examples. Unfortunately, all of these techniques are laborious and expensive, consume reagents and/or solvents, and generate a considerable amount of residue.

Vibrational spectroscopic techniques, such as Raman and near and mid infrared (NIR and MIR) spectroscopy, are simpler, faster and less expensive alternative techniques, which require little or no sample pretreatment. Methods developed based on these techniques are in accordance with the principles of green chemistry, being more environmentally friendly [9]. These techniques generate spectra that demand subsequent analysis with multivariate classification methods. Recently, the combination of vibrational spectroscopic techniques and multivariate qualitative methods has provided good results in the detection of food frauds [10,11]. The use of multivariate classification methods involves the assignment of a sample to a class previously established. In the case of food adulterations, one class corresponds to the non-adulterated food, and different additional classes are established depending on the number of adulterants to be detected. The ideal result assigns samples to the classes to which they actually belong.

Given that in multi-class classification methods the number of predefined classes corresponds to the number of known adulterants present in the samples plus one (authentic class), the developed multivariate models are commonly built to detect the presence or absence of only a single adulterant per sample. Although less common in the literature, adulteration with blends of two or more adulterants is a possibility in some real situations. Therefore, it is important to check whether the developed methods are able to correctly classify samples adulterated with more than one adulterant. These samples should be assigned to all the classes established for the respective adulterants [12]. Since this problem is rarely addressed in the literature, the present study represents an important contribution to the food science community and can be easily extended to other types of frauds involving other products or matrices.

The amount of information generated with mid-infrared (MIR) analysis often comprises hundreds or thousands of variables, and a number of them may be noisy and/or irrelevant to the problem under study. The selection of a limited number of informative predictors/variables can improve the effectiveness of the classification models, reduce their complexity and increase their robustness [13]. Therefore, prior to the application of the developed classification methods, variable selection strategies will be employed to optimize these methods. Recent analytical methods have been found in the literature applying variable selection with Fourier transform infrared (FT-IR) spectra in classification problems [14,15]. There are different methods of variable selection for removing noisy and irrelevant information and other methods aiming to select the most discriminatory variables when working with different groups of samples. Due to the broad diversity of variable selection methods, choosing the most appropriate method is not a simple matter [16].

Multi-class methods using partial least squares discriminant analysis (PLS-DA) and soft independent modeling of class analogy (SIMCA) were previously developed by the authors, showing good performance for classification of nectar samples with only one adulterant [17].

Thus, the objective of this work was to apply the developed strategy to detect blends of two adulterants in grape nectar using total reflectance attenuated Fourier transform mid infrared spectroscopy (ATR-FTIR). To improve the ability of these methods to detect apple, cashew and blends of both of these adulterants in grape nectars, three variable selection methods, namely, interval partial least squares (iPLS) [18], variable importance in projection scores (VIP scores) [19] and a genetic algorithm (GA) [20], were implemented. The strategy proposed in this work is described in Fig. 1, which schematically shows its different steps, including sampling, classification methodology (PLS-DA and SIMCA), analytical validation and variable selection strategies.

2. Materials and methods

2.1. Formulation of nectar samples

In this study, we chose to manufacture all of the analyzed nectar samples starting from reliable raw materials and rigorously meeting the established regulations [21–23]. This is due to the lack of the reliability about information provided on the composition of commercial Brazilian nectars related to the minimum required amounts of pulps.

Isabel grapes were supplied by EMBRAPA (Brazilian Agricultural Research Corporation) Grape & Wine, located in Petrolina, PE, Brazil. Red cashews and Fuji apples were purchased at Minas Gerais Supply Center (CEASA), located in Contagem, MG, Brazil. The fruits without physical and phytopathological damage were stored in the refrigerator (4–7 °C) until the preparation of the nectars.

The components in the final produced nectars were grape juice, pulps of the respective adulterants (apple, cashew or both), sugar syrup (water and sugar) and permitted additives (added within the permitted limits) [23], such as ascorbic acid, citric acid and guar gum (Pryme Foods, Sorocaba, SP, Brazil). Eq. (1) was applied to calculate the quantities of main fruit, adulterant fruit(s) and syrup:

$$\frac{a \times A}{100} + \frac{b \times B}{100} + \frac{c \times C}{100} + \frac{d \times D}{100} = \frac{m \times (A + B + C + D)}{100} \quad (1)$$

where “a”, “b”, “c” and “d” denote the Brix of the main fruit, syrup, fruit adulterant 1 and fruit adulterant 2, respectively; “A”, “B”, “C” and “D” represent the percentages of the main fruit, which is established as 50% for grape nectars [21,22], syrup, fruit adulterant 1 and fruit adulterant 2, respectively; "m" is the final nectar's Brix; and “A + B + C + D” is equal to 100 [17]. If no adulterant is added, this formula will only include “A” and “B”. If only one adulterant is added, this formula will include “A”, “B” and “C”.

Grape nectar samples were prepared for each of the three studied classes as follows:

- unadulterated (UN) - formulated with 50% grape juice, 50% sugar syrup and additives;
- adulterated with cashew (CAS) - formulated with 40% grape juice, 10% cashew pulp, 50% sugar syrup and additives;
- adulterated with apple (APP) - formulated with 40% grape juice, 10% apple pulp, 50% sugar syrup and additives;

Additionally, two external data sets were prepared, in which samples were adulterated with blends of cashew and apple at two different concentration levels:

- external data set 1 - adulterated with cashew and apple (CAS + APP) - formulated with 40% grape juice, 5% cashew pulp, 5% apple pulp, 50% sugar syrup and additives;
- external data set 2 - adulterated with cashew and apple (CAS + APP) - formulated with 30% grape juice, 10% cashew pulp, 10% apple pulp, 50% sugar syrup and additives;

The adulteration level has been fixed considering that too low concentrations (lower than 10%) are not economically advantageous for the fraudulent industry and too high concentrations could be perceived by sensorial evidences.

Fig. 2 shows a schematic of the experimental design for preparing the samples, resulting in 42 representative samples of each class, as well as 15 samples for the external data set 1 and 42 for the external data set 2, totaling 183 samples. Spectra from each class were split into training (28 samples) and test sets (14 samples) using the Kennard-Stone algorithm, selecting representative samples distributed homogeneously into the multivariate space [24].

2.2. Instrumentation and software

The analysis was conducted at controlled temperature (20.0 ± 0.5 °C). ATR-FTIR spectra were obtained using an IRAffinity-1 FTIR (Shimadzu, Kyoto, Japan) spectrophotometer with a DLATGS

Download English Version:

https://daneshyari.com/en/article/7675041

Download Persian Version:

https://daneshyari.com/article/7675041

Daneshyari.com