



The use of advanced chemometric techniques and trace element levels for controlling the authenticity of organic coffee



Rommel Melgaço Barbosa^a, Bruno Lemos Batista^{b,c}, Renan M. Varrique^b, Vinicius A. Coelho^b, Andrés D. Campiglia^d, Fernando Barbosa Jr.^{b,*}

^a Instituto de Informática, Universidade Federal de Goiás, Goiânia-Go, Brazil

^b Laboratório de Toxicologia e Essencialidade de Metais, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Avenida do Café s/n, Monte Alegre, 1404903 Ribeirão Preto-SP, Brazil

^c Universidade Federal do ABC, Centro de Ciências Naturais e Humanas, Santo André-SP, Brazil

^d Department of Chemistry, University of Central Florida, Orlando-USA

ARTICLE INFO

Article history:

Received 11 May 2013

Received in revised form 19 July 2013

Accepted 26 July 2013

Available online 7 August 2013

Keywords:

Chemometric

ICP-MS

Trace elements

Food samples

Quality control

ABSTRACT

This study examines the application of advanced data mining techniques associated with trace element concentrations for authenticity evaluation of organic coffee samples. Fifty-four coffee samples (20 organic and 34 non-organic) were evaluated according to their macro and microelement contents to establish reference ranges and identify differences in patterns of chemical elements in both crop procedures. Inductively coupled plasma mass spectrometry (ICP-MS) was used for the determination of thirty-eight chemical elements (Al, As, B, Ba, Be, Bi, Cd, Ce, Co, Cr, Cs, Cu, Dy, Eu, Er, Gd, Ge, Ho, La, Lu, Mn, Mo, Ni, Pb, Rb, Se, Sm, Sr, Yb, W, U, Ta, Tb, Th, Ti, Tl, Zr and Zn). In general, higher levels of chemical elements were found in conventional coffee compared with organic samples. The classification of samples was performed by using three well known data mining algorithms, named Multilayer Perceptron (MLP), Support Vector Machine (SVM) and Naïve Bayes (NB). When all analyzed variables were considered, a 96.3% accuracy level was obtained with the three algorithms. The level of accuracy remained at 96.3% for SVM and MLP and rose to 98.2 for NB when only eight of the investigated minerals – i.e. Zn, Se, Ba, U, Dy, Tl, Th, Mo – were used for chemometric analysis. This level of accuracy provides a strong foundation to propose the combination of trace element spectral fingerprints with data mining techniques as a valuable tool to evaluate the authenticity of organic coffee samples.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Driven by consumers' perceptions of the quality and safety of organic foods and by the positive environmental impact of organic agricultural practices, consumption and market of organic food are growing considerably. "Organic" is a labeling term that indicates food production according to specific rules, particularly those that exclude the use of pesticides, inorganic fertilizers and other chemical products that may cause harm to consumers' health (Kahl et al., 2012). Certification of organic products requires a series of precautions, such as the detoxification of the soil in areas under transition from chemical to organic agriculture, the absence of chemical fertilizers and pesticides, attention to ecological aspects, etc. (Kahl et al., 2012). The "certified label" adds market value to organic products, which makes the authentication process relevant under the perspective of possible fraud. Since the differences among organic and non-organic products are not visible to

the naked eye, systematic control of the authenticity of organic food products still relies on chemical analysis.

Coffee is highly consumed worldwide and one of the most important food commodities. About 60 tropical and subtropical countries produce coffee extensively, being in some cases the main agricultural export product (Esquivel & Jiménez, 2012). Several macro and microelements are absorbed from the soil by the coffee seeds in particular proportions during the plant growing (dos Santos & de Oliveira, 2001; Frost-Meyer, Logomarsino, & Logomarsino, 2012). The different pattern of accumulation will be influenced by the soil type, soil chemical composition and the general environmental growing conditions. Because of that, the determination of trace-element content to assure the geographical origin of food samples has been proposed (Heaton, Kelly, Hoogewerff, & Woolfed, 2008).

Several atomic spectroscopic techniques have been applied to the determination of macro and microelements in coffee samples (Pohl, Stelmach, Welna, & Szymczycha-Madeja, 2013). These include atomic absorption spectrometry (AAS) (Amorim-Filho, Polito, & Neto, 2007; Ashu & Chandravanshi, 2011; dos Santos, dos Santos, Conti, dos Santos, & de Oliveira, 2009; Duran, Tuzen, & Soyak, 2010; Krivan, Barth, & Morales, 1993) atomic emission spectrometry with inductively

* Corresponding author. Tel.: +55 1636024701; fax: +55 16 4725.
E-mail address: fbarbosa@fcrp.usp.br (F. Barbosa).

coupled plasma (ICP-OES) (dos Santos & de Oliveira, 2001; Krejcová & Cernohorsky, 2003) and inductively coupled plasma mass spectrometry (ICP-MS) (Santato, Bertoldi, Perini, Camin, & Larcher, 2012). However, the techniques of choice to obtain the elemental fingerprint of food products are those with multi-element detection capability (e.g., ICP-based techniques (ICP-OES and ICP-MS)). On the other hand, ICP-MS typically provides much lower detection limits and wider linear dynamic ranges than those obtained with ICP-OES for 70 elements (in the range of ppb and ppt) (Chudzinska & Baralkiewicz, 2011; Nardi et al., 2009; Parsons & Barbosa, 2007).

Supervised and unsupervised chemometric tools have been used for food authenticity and traceability (Aguar et al., 2012; Amorim-Filho et al., 2007; Anderson & Smith, 2002; Batista et al., 2012; Borin, Ferrão, Mello, Maretto, & Poppi, 2006; Costa Freitas & Mosca, 1999; dos Santos et al., 2009; Inarejos-García, Gómez-Alonso, Fregapane, & Salvador, 2013; Ouyang, Zhao, & Chen, 2013; Ribeiro, Salva, & Ferreira, 2010; Santato et al., 2012; Santos et al., 2012; Valentin & Watling, 2013; Zaidi et al., 2006). Supervised multivariate classification techniques – such as discriminant analysis on principal components or partial least squares scores – refer to a calibration or training data set that is used to develop a mathematical model capable of classifying future samples. Unsupervised classification – known as cluster analysis – relies on pattern recognition of complex data sets. SVM, MLP and NB constitute a group of supervised learning models with associated learning algorithms. Based on a training set of examples with the correct responses, these algorithms are able to generate a correct response to all possible inputs (Aguar et al., 2012; Batista, Grotto, Carneiro, & Barbosa, 2012; Borin et al., 2006; Brereton & Loyd, 2010; Mutihac & Mutihac, 2008; Xu, Zomer, & Brereton, 2006). This property provides SVM, MLP and NB with the ability to identify the group to which a new sample belongs after preliminary group classifications (Brereton & Loyd, 2010; Mutihac & Mutihac, 2008). Then, this paper aims to examine the application of advanced data mining techniques (Multilayer Perceptron (MLP), Support Vector Machine (SVM) and Naïve Bayes (NB)) associated with trace element concentrations for authenticity evaluation of organic coffee samples. To the extent of our literature search, this is the first application of chemometric algorithms for the validation of organic coffee.

2. Material and methods

2.1. Instruments

The determination of trace elements in coffee samples was carried out by using an ICP-MS (ELAN DRCL, PerkinElmer, CT, USA) with high-purity argon (99.999%, White Martins, Brazil). The instrumental parameters and optimized conditions are provided by Sucharov and Suchara (2006).

2.2. Reagents

All reagents used were of analytical-reagent grade except for HNO₃, which was previously purified in a quartz sub-boiling still (Kürner Analysentechnik) before use. High purity deionized water (resistivity 18.2 MΩ cm) obtained using a Milli-Q water purification system (Millipore, Bedford, MA, USA) was used. Rhodium, iron, magnesium, zinc and copper (1000 mg L⁻¹) and multi-element (10 mg L⁻¹) solutions were obtained from PerkinElmer (Shelton, CT, USA). Triton® X-100 and tetramethylammonium hydroxide solution (TMAH) 25% (w/v) in water were purchased from Sigma-Aldrich (St. Louis, USA).

2.3. Sampling and analytical procedures

Certified organic (n = 20) and non organic coffee (n = 34) samples were obtained from Minas Gerais, São Paulo, Bahia and Espírito Santo

states. All organic coffee samples were certified by the Brazilian IBD-Agricultural and Food Inspections and Certifications that is accredited by the International Federation of Organic Agriculture Movements. After sampling, samples were put in propylene metal-free Falcon® tubes (Becton Dickinson) before analysis, frozen at -80 °C and freeze-dried. The method proposed by Sucharov and Suchara (2006) was used to determine trace elements in coffee samples. Thirty-eight chemical elements named Al, As, B, Ba, Be, Bi, Cd, Ce, Co, Cr, Cs, Cu, Dy, Eu, Er, Gd, Ge, Ho, La, Lu, Mn, Mo, Ni, Pb, Rb, Se, Sm, Sr, Yb, W, U, Ta, Tb, Th, Ti, Tl, Zr and Zn were determined. A solution of rhodium 10 µg L⁻¹ was used as the internal standard. The nonparametric Mann-Whitney test was applied to evaluate possible differences between chemical elements in organic and conventional samples. Differences were considered significant when p < 0.05.

2.4. Analytical quality control

The quality of the determination of trace elements in coffee was guaranteed by analyzing NIST Standard Reference Material NIST SRM 1515 apple leaves and NIST 1547 peach leaves acquired from the National Institute of Standards and Technology (NIST, USA). Results for the determination of chemical elements in the certified reference materials by applying the ICP-MS methodology (Table 1) were in good agreement with the target/certified values provided by the National Institute of Standards and Technology (NIST, USA), since no statistical differences were found after applying the paired *t*-test at the 99% level.

2.5. Data mining techniques for organic and non-organic classification

Machine learning consists of using mathematical, statistical and computational methods to make classifications based on known information. Usually the known information is electronic data collected and made accessible for analysis for the purpose of classification. Classification involves identifying to which category (in our case organic or non-organic) a new sample belongs, based on a training set of data having features (observations) whose category is already labeled. The set of attributes associated to an example is called features, represented as a vector (in this case the minerals are the features) (Mohri, Rostamizadeh, & Talwalkar, 2012).

Learning algorithms for classification have been successfully deployed in a variety of applications, such as: message classification (spam or not spam), fraud detection, and medical diagnosis, to name a few.

The learning stages of our classification problem are as follows. Starting with an existing collection of labeled examples, we randomly divide the data into a training set and a test set. The size of the training set and test set depends on several different considerations (Mohri et al., 2012). If the labeled sample is relatively small, often a decision is made to use a larger amount of training data than that of the test data since the learning process of an algorithm depends on the training set. Then we choose the most important variables to the examples. All features are not equal in their usefulness. While uninformative features can mislead the learning algorithms, useful ones guide us to greater accuracy. Then the features selected are used to train our learning algorithms. The algorithm selects a different hypothesis out of the hypothesis set (i.e., a set of function mapping features to the set of label C). Among the hypotheses, we choose the one resulting in the best performance on the validation sample. Finally, the labels in the test sample are predicted. The performance of an algorithm is evaluated based on its test errors (Mohri et al., 2012).

Three well known classification algorithms were used in the present study. The software Weka (Waikato Environment for Knowledge Analysis) was used to perform these algorithms. Weka contains a collection of visualization tools and algorithms for data analysis and classification.

Download English Version:

<https://daneshyari.com/en/article/6396432>

Download Persian Version:

<https://daneshyari.com/article/6396432>

[Daneshyari.com](https://daneshyari.com)