Food Chemistry 159 (2014) 175–180

Contents lists available at ScienceDirect

Food Chemistry

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

Analytical Methods

Development and analytical validation of a simple multivariate calibration method using digital scanner images for sunset yellow determination in soft beverages

Bruno G. Botelho [a], Luciana P. de Assis [a], Marcelo M. Sena [a,b,*]

[a] Departamento de Química, ICEx, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil
[b] Instituto Nacional de Ciência e Tecnologia em Bioanalítica (INCT Bio), 13083-970 Campinas, SP, Brazil

ARTICLE INFO

Article history:
Received 28 November 2013
Received in revised form 21 February 2014
Accepted 9 March 2014
Available online 18 March 2014

Keywords:
RGB histograms
Image analysis
PLS
Food dyes
Soft drinks
Quality control

ABSTRACT

This paper proposed a novel methodology for the quantification of an artificial dye, sunset yellow (SY), in soft beverages, using image analysis (RGB histograms) and partial least squares regression. The developed method presented many advantages if compared with alternative methodologies, such as HPLC and UV/VIS spectrophotometry. It was faster, did not require sample pretreatment steps or any kind of solvents and reagents, and used a low cost equipment, a commercial flatbed scanner. This method was able to quantify SY in isotonic drinks and orange sodas, in the range of 7.8–39.7 mg L^{-1}, with relative prediction errors lower than 10%. A multivariate validation was also performed according to the Brazilian and international guidelines. Linearity, accuracy, sensitivity, bias, prediction uncertainty and a recently proposed tool, the β-expectation tolerance intervals, were estimated. The application of digital images in food analysis is very promising, opening the possibility for automation.

1. Introduction

Artificial dyes are constantly present in the modern lifestyle, being largely used in cosmetics, clothes, drugs and particularly in foodstuff. They have a great number of advantages if compared with natural dyes, such as higher stability to oxygen, light and pH changes, good water solubility and lower production cost (Ghoreishi, Behpour, & Golestaneh, 2012; Xing et al., 2012). The azo dyes are the largest group of artificial dyes (60–70% of all artificial dyes) and their molecular structures are characterized by the presence of an azo group (—N=N—) placed between aromatic rings. Although they provide a lot of technological benefits related to aesthetic and organoleptic characteristics of a particular foodstuff, a great number of studies have already confirmed negative effects of their consumption for human health, especially when in excess, such as allergic responses, asthma, urticarial and immunosuppression (Yadav, Kumar, Tripathi, & Das, 2013). Sunset yellow (SY), also known as evening yellow, E110 or edible yellow 3, is one of the most used azo dyes. It has an orange color, and is used in a great number of fruit products, like sodas, juices, candies and ice creams. Usually, it is the only artificial dye present in orange soft beverages. SY has also a large use in the pharmaceutical industry and in cosmetics. Nevertheless, it also causes some side effects in humans and its consumption has been related to renal failure and hepatocelular damages (Xing et al., 2012).

The great increase in the consumption of artificial dyes, mainly in products destined for children, creates an urge for methods that can monitor and quantify these dyes. ANVISA (National Health Surveillance Agency) is the governmental agency responsible for food regulation in Brazil, and it establishes the limits for artificial dyes in different products. According to the resolution R05/07, the limit for SY concentration in nonalcoholic beverages is 100 mg L^{-1} (ANVISA., 2011), and the official method for azo dyes determination is based on UV/VIS spectrophotometry, which requires sequential liquid–liquid extractions with methanol containing 5% hydroxide ammonium (IAL, 2005). Other methods involving different analytical techniques, such as chromatography (Bonan, Fedrizzi, Menotta, & Elisabetta, 2013; Vidotti, Costa, & Oliveira, 2006); potenciometry (Ghoreishi et al., 2012), voltametry (Nevado, Flores, & Llerena, 1997), immunoassays (Xing et al., 2012) and cloud point extraction with spectrophotometric detection (El-Shahawi, Hamza, Al-Sibaai, Bashammakh, & Al-Saidi, 2013) have been reported in the literature. Chemometrics strategies have also been applied for food dyes determination, mainly with UV/VIS spectrophotometry and binary and ternary mixture of dyes (Berzas

* Corresponding author at: Departamento de Química, ICEx, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil. Tel.: +55 31 34096389; fax: +55 31 34095700.
E-mail address: marcsen@ufmg.br (M.M. Sena).

http://dx.doi.org/10.1016/j.foodchem.2014.03.048
0308-8146/

Nevado, Rodrıguez Flores, Guiberteau Cabanillas, et al., 1998; El-Sheikh & Al-Degs, 2013; Nevado, Flores, Llerena, & Fariñas, 1999).

The main objective of this paper was to develop and validate a multivariate image analysis (MIA) method based on digital images obtained by a commercial flatbed scanner coupled with chemometrics for the determination of SY in non-alcoholic orange beverages (isotonic and soft drinks). This strategy has several advantages compared to the classical methods, such as the rapidity of analysis (a few seconds), no need for extraction procedures, environmentally friendly and solvent free, with no chemical waste generation, and the low cost of the equipment (around US$ 100), providing sufficient accuracy and sensitivity with less human intervention. The proposed method was also validated in accordance with Brazilian and International guidelines, corroborating that MIA is a reliable technique, that besides it may be easily automatized or used in portable equipments, can also fulfill all the statistical requirements for an official analysis.

2. Materials and methods

2.1. Instruments and software

The images were obtained using a commercial flatbed scanner CanoScan LiDE 110 (Tokyo, Japan). Data were handled using MATLAB software, version 7.13 (The MathWorks, Natick, MA, USA). The PLS routine came from the PLS Toolbox, version 6.5 (Eigenvector Technologies, Manson, WA, USA), images were treated with the Image Processing Toolbox, version 8.0 (The MathWorks), and a homemade routine was also employed for the detection of outliers (Ferreira, Braga, & Sena, 2013).

2.2. Samples

Eighty-three samples of commercial beverages (orange soda and isotonic drinks) containing SY from different brands (25) and production batches were purchased in local markets, and stored under refrigeration at 4 °C until analysis.

2.3. Procedure

The samples were allowed to rest for 30 min for thermal equilibrium before starting the measurements. 30 mL from each sample were collected in a 50 mL beaker and degassed using an ultrasonic bath for 5 min. After degassing, 1 mL was used for chromatographic quantification of SY (reference values), and 5 mL were used for the image acquisition.

The acquisition of images was performed using a small Petri dish (5.0 cm radius $\times$ 1.5 cm height) filled with the sample and positioned in the corner of the scanner. A white screen was used to block the light from external sources. All images were digitized in the 24-bit RGB system, with 16.8 million colors and 300 dpi resolution, in ".tif" format. The conversion of the images in RBG histograms was carried out in MatLab environment. Firstly, a 100 $\times$ 100 pixel size area was selected from the central area of the dish, in a homogeneous part of the image. This selected area was then treated with a digital filter (unsharp) for noise reduction, and decomposed in a RGB histogram. After all treatments, a histogram containing 768 channels (256 for each RGB color) was obtained for each sample. Each sample was scanned three times and the average histograms were used for building PLS models.

2.4. Chromatographic analysis

The chromatographic analysis were based on a chromatograph manufacturer's method (Pedjie, 2012) and performed in a Finnigan Surveyor HPLC System (Thermo Fisher Scientific, San Jose, USA) with diode array detection (HPLC-DAD), using a Shimadzu Shim-Pack XR-ODS (3.0 mm I.D. $\times$ 150 mm L) C-18 column. Gradient elution was employed with a mobile phase composed of ammonium acetate 20 mM aqueous phase and acetonitrile/methanol (80:20, v/v) as organic phase. A flow rate of 1.2 mL min^{-1} and detection at 484 nm were used. The chromatographic run lasted 15 min, with SY retention time around 7.5 min.

2.5. Multivariate image analysis

Digital images have been used as source of analytical information since last century. The first published paper describing the use of digital imaging has employed an early version of a scanner for converting medical image exams into digital data (Ledley, 1964). More than twenty years later, Geladi and coworkers published the first paper concerning exclusively image analysis and chemometrics (Geladi, Wold, & Esbensen, 1986). Since then, mainly in the last years, a great variety of papers have been published, using different kinds of instruments, like cell phones, webcams, flatbed scanners and "point-and-shoot" digital cameras, for developing multivariate classification and calibration models applied to the analysis of food products and other matrices (Acevedo et al., 2009; Borin et al., 2007; Foca, Masino, Antonelli, & Ulrici, 2011; Godinho, Oliveira, & Sena, 2010; Iqbal & Bjorklund, 2011; Oliveira et al., 2013; Santos, Wentzell, & Pereira-Filho, 2012).

The most common way to extract the information from digital images is their decomposition in a color system, such as RGB. The RGB system is an additive system, which uses the combination of the colors Red, Green and Blue to form a wide variety of color tones. Each pixel (basic unity of a digital image) is formed by a combination of these colors. The intensity of each color in the RGB system is measured in channels. Channel 0 means complete absence of a color and channel 255 means the maximum intensity of a color. The combination of the RGB channels creates the different colors (256^3 possible combinations). After the decomposition of all the pixels of the image, the frequency of each channel of each color is counted, resulting in a frequency histogram. This histogram can be treated as spectral data and used for developing chemometric models. Alternatively, RGB variables can be fused with other color parameters, such as hue, saturation, intensity and lightness, resulting in colourgrams. Recently, several papers have developed multivariate calibration methods based on RGB histograms or colourgrams (Acevedo et al., 2009; Borin et al., 2007; Oliveira et al., 2013; Santos et al., 2012). Another strategy, which requires more complex mathematical handling, is the use of Fourier transform for obtaining congruent images and generating three-dimensional data arrays, which can be treated by multi-way methods (Godinho et al., 2010). This work chose to use the simplest strategy, combining RGB histograms with partial least squares (PLS) for the determination of SY in commercial samples of soft beverages.

2.6. Multivariate analytical validation

The analytical validation of multivariate methods is still not a completely well-established subject. Concerning food analysis, neither Brazilian nor international validation guidelines even mention multivariate statistics, completely ignoring its utilization (EC, 2002; MAPA, 2011; Thompson, Ellison, & Wold, 2002). As the importance and application of these methods have grown very quickly, it is necessary a harmonization between the validation aspects of univariate and multivariate methods. The establishment of validation procedures for multivariate calibration is very important because it is the first step for the recognition of these methods for official analysis. Further information on the state of the art of multivariate analytical validation, mainly focused on near infrared