View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

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.

© 2014 Elsevier Ltd. All rights reserved.

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/© 2014 Elsevier Ltd. All rights reserved.

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

spectroscopy, can be found elsewhere (Botelho, Mendes, & Sena, 2013; Faber, Song, & Hopke, 2003; Ferreira et al., 2013).

A novel tool for validation of analytical methods is the estimate of an accuracy profile based on the β-expectation tolerance intervals (β-TI) (Rozet et al., 2007). The β-TI are used as a complementary visual decision tool to evaluate the models predictive performances. They give the guarantee that a ratio β (i.e. 95%) of all the future results will presented an error within the calculated limits. Therefore, if the β-TI is included within the acceptance limits established by validation guides, the proposed method fulfill all the requirements needed for validation. The use of β-TI has been recently extended to multivariate calibration, mainly focused on the validation of NIR methods in pharmaceutical analysis (Mantanus et al., 2010).

The β-TI can be estimated using the following equations:

$$\beta - \mathrm{TI} = RE(\%)_j \pm t\sqrt{1+\frac{1}{pnB_j^2}}RSD(\%)_j \quad (1)$$

where p is the number of series, n is the number of independent replicates per series, RE(%) is the mean relative error for the n replicates for the j_{th} level, RSD(%) is the relative standard deviation for the n replicates for the j_{th} level. t is bicaudal t-student critical values for v degrees of freedom. v is calculated according to the equation below:

$$v = \frac{(R_j+1)^2}{\frac{(R_j+\frac{1}{n})^2}{p-1}+\frac{1-\frac{1}{n}}{pn}} \quad (2)$$

where R_j is the ratio between within series variance and between series variance, and B_j is estimated using R_j

$$B_j = \sqrt{\frac{R_j+1}{nR_j+1}} \quad (3)$$

3. Results and discussion

3.1. PLS model

All the analyzed samples were orange soft beverages and contained SY as the only artificial dye. Previously to the construction of the multivariate model using real samples, an alternative strategy was tried. An attempt to build a multivariate model with standard solutions of SY for predicting the real samples did not present good results. Since all the beverages contain some additives, which make them slightly cloudy, it was not possible to predict these samples with a model built with standard solutions. In addition, several samples also contain natural juices and consequently natural dyes, as interferences that contribute for color, thus justifying the use of multivariate calibration. So, the PLS model was built with real samples and the reference values were obtained using HPLC.

Fig. 1(a) shows the RGB histograms for the 83 samples. It is possible to note that the channel 513 is the most frequent, followed by the channel 256, which indicates a predominance of the yellowish tones in the images (mixture of red and green primary colors). As these two channels showed the highest frequencies in the histograms, they were used for building individual univariate calibration models. These models were compared with the multivariate model using the estimated correlation coefficients between the reference and predicted values. Reasonable correlations were found for the univariate models (0.827 for channel 513 and −0.778 for channel 256), but which have not been considered satisfactory for quantification. The removal of some channel regions (1–150 and 260–335) without any significant signal was also tried, but no improvement was observed for the multivariate model.

For building the PLS model, the sample histograms were divided into two thirds (56) for the calibration set and one third (27) for the validation set by using the Kennard–Stone algorithm, which assured the presence of the most representative samples in the calibration set through a uniform scanning of the independent variables data set. No preprocessing other than mean centering was used. The number of latent variables (LV) was selected based on the smallest root mean square error of cross validation (RMSECV), estimated using venetian blinds cross validation (6 data splits). The best model was obtained with 4 LV. The estimated regression coefficients for this PLS model are show in Fig. 1(b). By observing this plot, it was possible to identify the channels that more contribute to SY prediction. The largest regression coefficient was observed at the channel 768, which showed strong negative influence on the SY quantification, indicating the negative contribution of a blue pattern in the sample colors. The largest positive regression coefficient was at the channel 256, but the regression coefficient at 513, the most frequent channel in the sample histograms, was not so large, indicating that the most intense signals were not necessarily the most predictive. A univariate model correlating the frequencies at the channel 768 and the SY reference concentrations was also tested, but it did not presented good results ($r = -0.539$).

The PLS model was optimized through outlier detection and evaluated through the root mean squared error of prediction (RMSEP) and the relative prediction deviation (RPD), as shown in Table 1. Good models present higher RPD values. According to Brazilian and international validation guidelines (MAPA, 2011), outliers can be removed up to 22.2% (two out of nine) from the original samples. Eleven outliers were removed from the calibration set (19.6%), nine based on their large Y-residuals and two based on their high leverages. Five outliers were removed from the validation set (18.5%), all of them based on large Y-residuals. The final model was built with 4 LV, accounting for 98.8% of the variance in X and 97.2% in Y.

3.2. Multivariate analytical validation

Table 2 presents the FOM assessed for the optimized model. The plot of reference *versus* predicted values is show in Fig. 2. To confirm the method linearity, suitable statistical tests were performed, verifying the normality (Ryan–Joiner), homocedasticity (Brown–Forsythe) and absence of autocorrelation (Durbin–Watson) of the residuals, all at 95% confidence level. The results of these tests assured the randomness of the prediction residuals, especially the estimated Durbin–Watson value (1.86), which is within the acceptance limits (1.5–2.5), confirming the residual independence. Once the model was considered linear, the parameters for a linear regression were estimated (Table 2), including a correlation coefficient (r) of 0.9742.

Trueness was estimated through the absolute error parameters, such as a RMSEP of 1.3 mg L^{-1} (Table 1). Individual relative errors of prediction were also estimated, showing values between −6.2% and 9.0%, which are in accordance with the limits established by the MAPA (2011) validation guide (−20/+10%). Method precision was also evaluated at two different levels, repeatability and intermediary precision (a different analyst in a different day), through the estimate of the relative standard deviation (RSD). The method was considered precise, with repeatability RSD values ranging from 0.3% to 2.0%, and intermediary precision RSD values ranging from 2.0% to 4.1%. These values are within the limits defined by the Brazilian guidelines (MAPA, 2011), which establishes maximum acceptable RSD values of 4.9% for repeatability and of 7.3% for intermediary precision. Trueness and precision results corroborated that

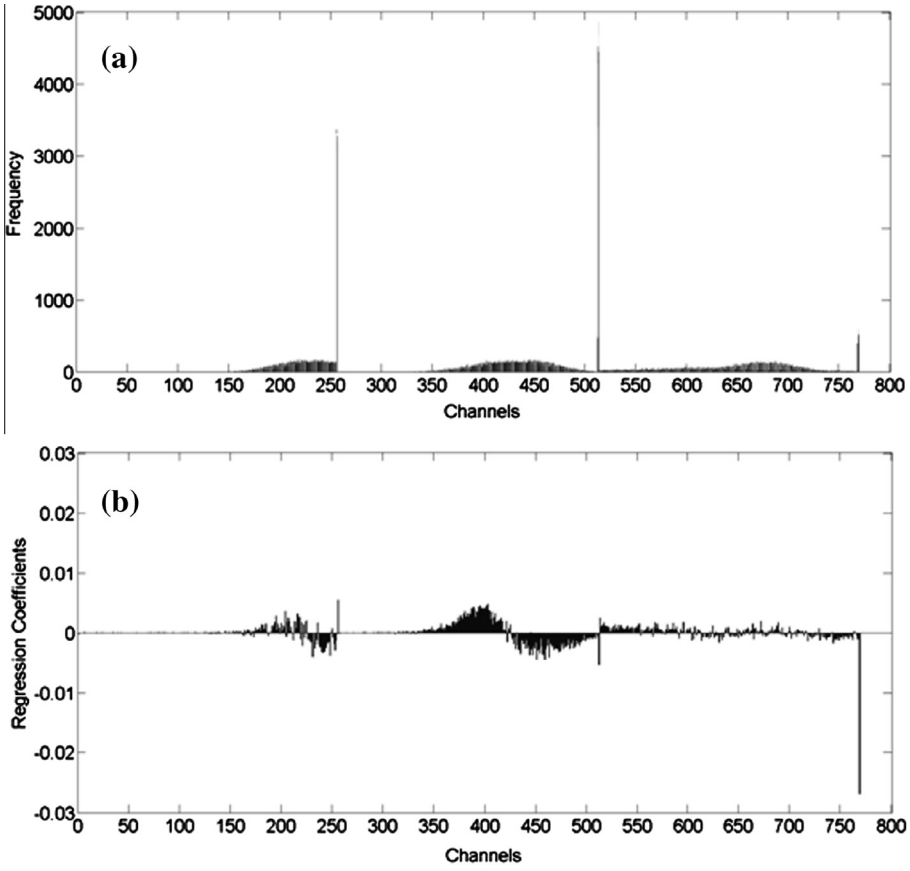

Fig. 1. (a) RGB histograms for the beverage samples containing SY. (b) Regression coefficients for the developed PLS model.

Table 1
Results for the optimization of the developed PLS model through the detection of outliers (in bold for the final model).

Model	1st	2nd	3rd	4th
Number of calibration samples	56	51	48	**45**
Number of validation samples	27	27	27	**22**
Number of latent variables	4	4	4	**4**
RMSEC (mg L^{-1})	2.7	2.1	1.7	**1.3**
RMSEP (mg L^{-1})	2.6	2.9	3.1	**1.3**
RPD calibration	2.3	3.1	3.8	**4.6**
RPD validation	3.3	3.0	2.8	**6.5**

Table 2
Parameters for evaluating the main FOM of the developed MIA method for the determination of SY in beverages.

Figures of merit	Parameter	Values
Trueness	Relative errors (max/min)	−6.2/9.0%
Precision	RSD repeatability[a]	2.7%
		1.0%
		0.3%
	RSD intermediate precision[a]	2.0%
		4.1%
		3.9%
Linearity	Durbin-Watson test	1.86
	Slope[b]	0.97 ± 0.02
	Intercept[b]	0.80 ± 0.74
	Correlation coefficient[b]	0.9702
Working range		7.8–39.7 mg L^{-1}
Selectivity		0.17
Sensitivity		154.11[c]
Analytical sensitivity (γ)		3.4 L mg^{-1}
γ^{-1}		0.3 mg L^{-1}
Bias		0.512 ± 1.210 mg L^{-1}

[a] Results for six replicates from three samples at three different content levels.
[b] Values for the line fitted to the calibration samples.
[c] Values expressed as the ratio between frequency and mg L^{-1}.

the method can be considered accurate. Considering the accuracy and linearity studies the analytical working range was defined from 7.8 to 39.7 mg L^{-1} of SY. It is possible to note in Fig. 2 that the samples did not show a homogeneous concentration distribution along the curve, with three clearly distinct groups. The first group, with the smallest concentrations, is formed only by the isotonic samples. The second group, in the middle of the curve, is formed by the most famous brands of orange soft drinks, and the third group is formed by low-cost orange sodas, which presented the highest levels of SY.

SEN and SEL were also estimated, based on the concept of net analyte signal (NAS). A requirement of a minimum value of SEL does not have practical use in multivariate validation, since unlike for univariate methods low values of SEL can be obtained even for accurate multivariate models. The SEL estimate indicated that 17% of the original analytic signal was used in the model for SY prediction. The SEN itself is not a very informative FOM, because it cannot be used for comparison with other methods. So, the analytical sensitivity (γ) was calculated from the estimate of the instrumental noise (44.7), which was obtained through ten replicates of images from an empty Petri dish. Its inverse, 0.3 mg L^{-1}, provides an estimate of the minimum concentration difference that the method is able to distinguish, considering the random instrumental noise as the only source of errors, and also defines the number of decimal places that should be used to express the prediction results.

The bias assessment (Table 2), estimated only for the validation set, shows a *t*-value (1.98) lower than the *t*-critical value (2.06, with 22 degrees of freedom and 95% confidence level), which indicates the absence of systematic errors in the model predictions.

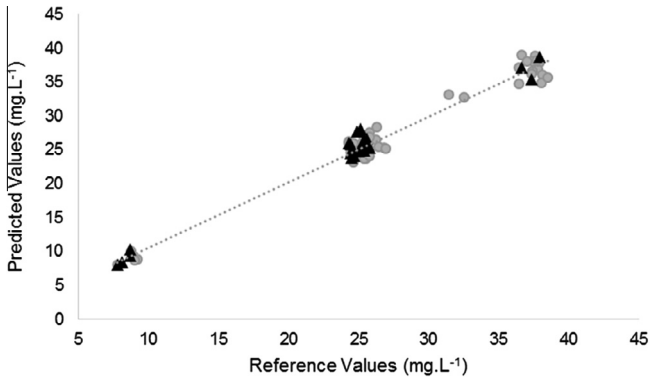

Fig. 2. Plot of the reference *versus* predicted values for the PLS model. Calibration (circles) and validation samples (triangles).

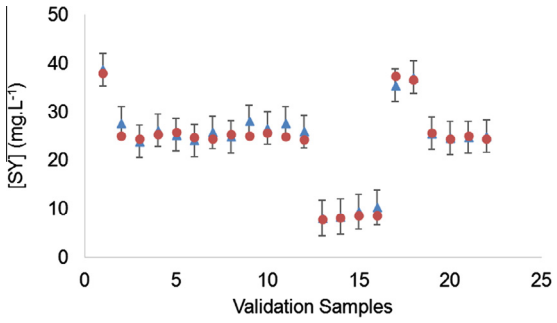

Fig. 3. Confidence intervals for the prediction of the validation samples, estimated based on the SEP. Predicted (triangles) and reference values (circles).

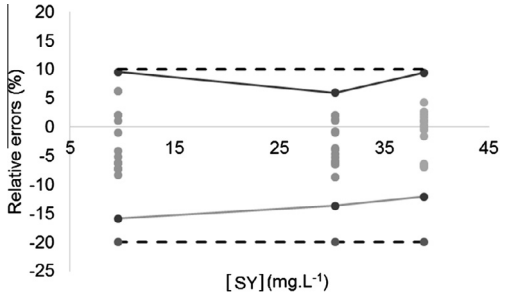

Fig. 4. β-expectation tolerance intervals calculated for the model (solid lines) and maximum relative errors established by the Brazilian validation guidelines (dashed line).

According to the literature (Botelho, Mendes, & Sena, 2013), good calibration models must have a RPD value higher than 2.4, while values between 2.4 and 1.5 are considered acceptable. Considering the values presented in Table 1, RPD estimates were satisfactory for both calibration and validation.

Standard prediction errors (SPE), calculated with the errors-in-variables (EIV) equation (Faber et al., 2003), ranged from 3.3 to 3.6 mg/L. Fig. 3 shows the reference and the predicted values with the respective confidence intervals estimated for each sample. For these calculations, the reference method uncertainty (0.03 mg L^{-1}) was obtained from intermediary precision studies. Five pseudo-degrees of freedom were used in the confidence interval estimates ($t = 2.77$). For all samples, the reference values were within the estimated confidence interval (agreement of 100%).

3.3. β-expectation tolerance intervals

The β-TI estimated for the model can be seen in Fig. 4 and were based on three series of triplicates ($p = 3$ and $n = 3$). This accuracy profile shows that their confidence limits are within the limits established by the Brazilian validation guidelines, −20/+10% (MAPA,2011). This indicates that it is expected that all the predicted values obtained using the developed model will present relative errors within these acceptance limits. For the lowest concentration (9.6 mg L^{-1}), the β-TI ranges from −15.4% to 9.6%. For the medium concentration (30.0 mg L^{-1}), β-TI ranges from −13.7% to 6.0%, and for the highest concentration (38.8 mg L^{-1}), it ranges from −12.1% to 9.4%.

4. Conclusions

A simple multivariate calibration method based on RGB histograms from digital images was developed and validated for sunset yellow determination in orange beverages (isotonic and soft drinks). It uses a very low cost equipment (a commercial flatbed scanner), does not require sample pretreatment nor use reagents or solvents, and is also much faster than the reference method (less than 1 min against a 15 min of a chromatographic run). Additionally, the developed method may be used for online automatization of industrial processes with a higher sampling frequency, and in portable equipments. This method was throroughly validated in accordance with the Brazilian and international guidelines, being considered linear, accurate, unbiased, and suitable for use as an official methodology for artificial dyes determination in beverages.

Acknowledgements

B.G.B. thanks CAPES and CNPq for fellowships. L.P.A. thanks "Pró-Noturno" (Pró-Reitoria de Graduação, UFMG) for a fellowship. We also thank Julio César C. Silva and Prof. Rodinei Augusti (ICEx, UFMG) for the support with the chromatographic analysis.

References

Acevedo, C. A., Skurtys, O., Young, M. E., Enrione, J., Pedreschi, F., & Osorio, F. (2009). A non-destructive digital imaging method to predict immobilized yeast-biomass. *LWT – Food Science and Technology, 42*, 1444–1449.

ANVISA. (2011). Agência Nacional de Vigilância Sanitária. Compêndio da Legislação Brasileira de Aditivos Alimentares (p. 163), Brasilia.

Berzas Nevado, J. J., Rodrıguez Flores, J., Guiberteau Cabanillas, C., et al. (1998). Resolution of ternary mixtures of Tartrazine, sunset yellow and Ponceau 4R by derivative spectrophotometric ratio spectrum-zero crossing method in commercial foods. *Talanta, 46*, 933–942.

Bonan, S., Fedrizzi, G., Menotta, S., & Elisabetta, C. (2013). Simultaneous determination of synthetic dyes in foodstuffs and beverages by high-performance liquid chromatography coupled with diode-array detector. *Dyes and Pigments, 99*, 36–40.

Borin, A., Ferrão, M. F., Mello, C., Cordi, L., Pataca, L. C. M., Durán, N., et al. (2007). Quantification of *Lactobacillus* in fermented milk by multivariate image analysis with least-squares support-vector machines. *Analytical and Bioanalytical Chemistry, 387*, 1105–1112.

Botelho, B. G., Mendes, B. A. P., & Sena, M. S. (2013). Development and analytical validation of robust near-infrared multivariate calibration models for the quality control of mozzarella cheese. *Food Analytical Methods, 6*, 881–891.

EC (2002). European Community Directive 657. Comission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results, Official Journal, L221 (pp. 8–36), Brussels.

El-Shahawi, M. S., Hamza, A., Al-Sibaai, A. A., Bashammakh, A. S., & Al-Saidi, H. M. (2013). A new method for analysis of sunset yellow in food samples based on cloud point extraction prior to spectrophotometric determination. *Journal of Industrial and Engineering Chemistry, 19*, 529–535.

El-Sheikh, A. H., & Al-Degs, Y. S. (2013). Spectrophotometric determination of food dyes in soft drinks by second order multivariate calibration of the absorbance spectra-pH data matrices. *Dyes and Pigments, 97*, 330–339.

Faber, N. M., Song, X. H., & Hopke, P. K. (2003). Sample-specific standard error of prediction for partial least squares regression. *TrAC – Trends in Analytical Chemistry, 22*, 330–334.

Ferreira, M. H., Braga, J. W. B., & Sena, M. M. (2013). Development and validation of a chemometric method for direct determination of hydrochlorothiazide in pharmaceutical samples by diffuse reflectance near infrared spectroscopy. *Microchemical Journal, 109*, 158–164.

Foca, G., Masino, F., Antonelli, A., & Ulrici, A. (2011). Prediction of compositional and sensory characteristics using RGB digital images and multivariate calibration techniques. *Analytica Chimica Acta, 706*, 238–245.

Geladi, P., Wold, S., & Esbensen, K. (1986). Image analysis and chemical information in images. *Analytica Chimica Acta, 191*, 473–480.

Ghoreishi, S. M., Behpour, M., & Golestaneh, M. (2012). Simultaneous determination of sunset yellow and Tartrazine in soft drinks using gold nanoparticles carbon paste electrode. *Food Chemistry, 132*, 637–641.

Godinho, M. S., Oliveira, A. E., & Sena, M. M. (2010). Determination of interfacial tension of insulating oils by using image analysis and multi-way calibration. *Microchemical Journal, 96*, 42–45.

IAL (2005). *Métodos Físicos-Químicos para Análise de Alimentos* (4th ed.). São Paulo: Instituto Adolfo Lutz (IAL) (p. 1020).

Iqbal, Z., & Bjorklund, R. B. (2011). Assessment of a mobile phone for use as a spectroscopic analytical tool for foods and beverages. *International Journal of Food Science and Technology, 46*, 2428–2436.

Ledley, R. S. (1964). High-speed automatic analysis of biomedical pictures. *Science, 146*, 216–223.

Mantanus, J., Ziemons, E., Rozet, E., Streel, B., Klinkenberg, R., Evrard, B., et al. (2010). Building the quality into pellet manufacturing environment – Feasibility study and validation of an in-line quantitative near infrared (NIR) method. *Talanta, 83*, 305–311.

MAPA (2011). Secretaria de Defesa Agropecuária. Ministério da Agricultura, Pecuária e Abastecimento. Manual da Garantia da Qualidade (p. 227), Brasilia.

Nevado, J. J. B., Flores, J. R., & Llerena, M. J. V. (1997). Square wave adsorptive voltammetric determination of sunset yellow. *Talanta, 44*, 467–474.

Nevado, J. J. B., Flores, J. R., Llerena, M. J. V., & Fariñas, N. R. (1999). Simultaneous spectrophotometric determination of tartrazine, patent blue V, and indigo carmine in commercial products by partial least squares and principal component regression methods. *Talanta, 48*, 895–903.

Oliveira, L. F., Canevari, N. T., Guerra, M. B. B., Pereira, F. M. V., Schaefer, C. E. G. R., & Pereira-Filho, E. R. (2013). Proposition of a simple method for chromium (VI) determination in soils from remote places applying digital images: A case study from Brazilian Antarctic Station. *Microchemical Journal, 109*, 165–169.

Pedjie, N. (2012). Analysis of color additives in beverages with the PerkinElmer Flexar FX-15 system equipped with a PDA detector. USA: Application Note, PerkinElmer, Waltham. <http://www.perkinelmer.com/CMSResources/Images/44-134896APP_Food_Additives_in_Beverages.pdf>.

Rozet, E., Wascotte, V., Lecouturier, N., Préat, V., Dewé, W., Boulanger, B., et al. (2007). Improvement of the decision efficiency of the accuracy profile by means of desirability function for analytical methods validation. Application to a diacetyl-monoxime colorimetric assay used for the determination of urea in transdermal iontophoretic extracts. *Analytica Chimica Acta, 591*, 239–247.

Santos, P. M., Wentzell, P. D., & Pereira-Filho, E. R. (2012). Scanner digital images combined with color parameters: A case study to detect adulterations in liquid cow's milk. *Food Analytical Methods, 5*, 89–95.

Thompson, M., Ellison, S. L. R., & Wold, R. (2002). Harmonised guidelines for single-laboratory validation methods of analysis. *Pure and Applied Chemistry, 74*, 835–855.

Vidotti, E. C., Costa, W. F., & Oliveira, C. C. (2006). Development of a green chromatographic method for determination of colorants in food samples. *Talanta, 68*, 516–521.

Xing, Y., Meng, M., Xue, H., Zhang, T., Yin, Y., & Xi, R. (2012). Development of a polyclonal antibody-based enzyme-linked immunosorbent assay (ELISA) for detection of sunset yellow FCF in food samples. *Talanta, 99*, 125–131.

Yadav, A., Kumar, A., Tripathi, A., & Das, M. (2013). Sunset yellow FCF, a permitted food dye, alters functional responses of splenocytes at non-cytotoxic dose. *Toxicology Letters, 217*, 197–204.