Microchemical Journal 118 (2015) 110–114

Contents lists available at ScienceDirect

Microchemical Journal

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

Application of digital images to determine color in honey samples from Argentina

Marina A. Dominguez, María E. Centurión *

Departamento de Química, INQUISUR (UNS-CONICET), Universidad Nacional del Sur, Av. Alem 1253 (B8000CPB), Bahía Blanca, Buenos Aires, Argentina

A R T I C L E I N F O

Article history:
Received 28 April 2014
Received in revised form 4 August 2014
Accepted 4 August 2014
Available online 10 August 2014

Keywords:
Honey
Color
Digital image
Multivariate calibration

A B S T R A C T

Honey color is the first quality attribute evaluated by consumers and is an important sensory property in the beekeeping market. The most commonly used method for color determination is based on optical comparison, employing a Pfund colorimeter. This method is laborious, time consuming and is based on a subjective assessment. In order to obtain a fast and objective measurement of honey color, an analytical method based on digital images analysis combined with multivariate calibration (Partial Least Squares) was developed. Three color models, RGB (red–green–blue), HSB (hue–saturation–brightness) and Grayscale, were used to analyze the digital images. The optimum results were obtained by using HSB color model ($r^2 = 0.97$ and RMSEP = 2.46), indicating that the digital image analysis combined with multivariate calibration is an excellent strategy for quantifying color in honey samples.

1. Introduction

In terms of production Argentina is positioned is in the sixth place, being simultaneously the second exporter of honey in the world [1]. For its extensive natural grasslands, its geographical location and climate, Argentina has a wide variety of floral resources that allow these activities to be carried out in most of the territory. However, the Pampas region located in the province of Buenos Aires concentrates more than 50% of national production [2]. In Argentina the internal market is underdeveloped due to low consumption of this product. However, Argentine honey is internationally appreciated for its organoleptic characteristics, its chemical composition and its wide variety of colors and purity [3–5]. In fact, the 95% of national production is destined for export markets which require good quality of products.

Color is the most important feature from the commercial standpoint and encourages honey export. This attribute varies according to the geographical origin of honey and depends on the pigment (carotenes and xanthophylls) and polyphenol (flavonols) contents [6–9]. The appropriate color measurement allows exporters to choose the most advantageous trading market for their products and is the only sensory examination with accurate coding within current regulations.

Internationally honeys are classified for commercialization using the Pfund color scale. The Pfund colorimeter is a simple instrument by which an observer compares the color of the honey sample with a standard colored glass [10]. The reference unit is the Pfund scale, which ranges from 0 to 140 mm, beginning with very light-colored honey and increasing up to the darkest honey.

However, this method does not distinguish between small variations of color and the determination is laborious, time consuming, requires large amounts of sample and depends on the person performing the analysis since different observers get different measurements.

Nowadays, there are few articles published in order to obtain other methods to determine honey color. Gallez et al. propose a nonlinear quadratic polynomial regression between data recorded with the Pfund and the CIELAB methods by means of multivariate analysis [11]. Other authors proposed a method to the reconstruction of honey reflectance spectra by using a characteristic vector analysis to obtain the color of honeys [12]. A potentiometric electronic tongue was used for the determination of color and other parameters for the discrimination of honey according to the botanical origin [13]. However, these methods have certain disadvantages such as they are not quite as simple as other methods, require expensive instrument which are not available in routine laboratories, tests need to be carried out by specialized analysts and required complex mathematical procedures. Therefore, it is necessary to develop alternative methods for color determination in honey samples to achieve objective measures. In a previous research work we described that digital images combined with chemometric techniques are an adequate tool for geographic origin classification of honey samples. These digital images were analyzed using different color models as RGB (Red–Green–Blue), HSB (Hue–Saturation–Brightness) and Grayscale [14].

The aim of this study was to apply digital image analysis combined with multivariate calibration for determining color in honey samples. From color data obtained by the reference method (Pfund) and color

* Corresponding author. Tel.: +54 291 4595100; fax: +54 291 4595160.
E-mail address: mecentur@criba.edu.ar (M.E. Centurión).

http://dx.doi.org/10.1016/j.microc.2014.08.002
0026-265X/

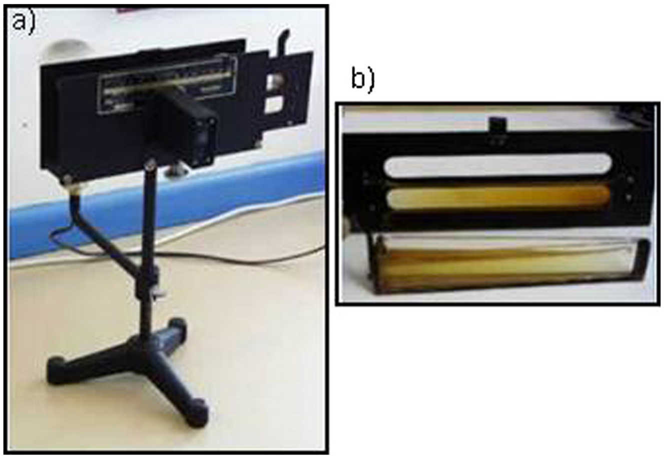

Fig. 1. a) Pfund color grader, b) wedge of amber-colored glass and wedge-shaped cell filled with the honey sample.

histograms, a multivariate calibration model using partial least squares (PLS) were constructed. By this way a simple method for color measurements, with less time analysis and human participation to assist the quality control of honey samples, was proposed.

2. Materials and methods

2.1. Samples

Eighty honey samples from southwest of the Buenos Aires province, Argentina were obtained from local markets. Fresh honey samples were stored at 4 °C in the dark until analysis. All samples were allowed to reach room temperature before analysis.

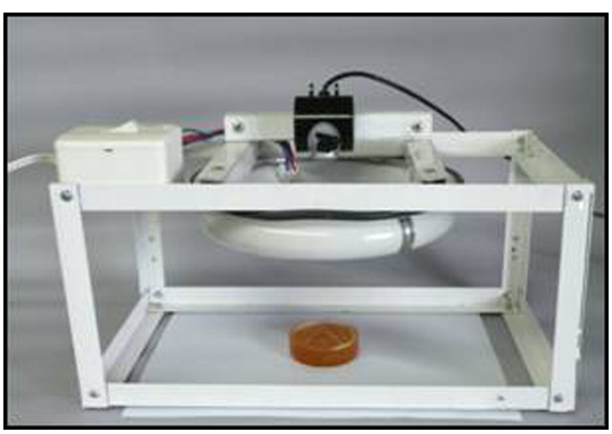

Fig. 2. Honey image capturing device.

Table 1
Calibration models results.

Model	LV[a]	RMSEC	RMSECV	r^2	Slope	Bias	Correlation
RGB	6	2.60	4.73	0.97	0.97	7.15×10^{-7}	0.98
Grayscale	4	3.14	4.33	0.97	0.97	5.88×10^{-7}	0.98
HSB	4	2.06	3.63	0.98	0.98	-1.28×10^{-6}	0.99

[a] Latent variables.

2.2. Apparatus and software

A Philips Webcam SPC900NC VGA with a CCD sensor was used to obtain digital images. These images were processed to acquire the corresponding color histogram with the ImageJ1.44p program (a free internet download). Chemometric data treatment was implemented with The Unscrambler_9.7 (CAMO S/A), and Matlab® 6.5 software (The Mathworks). A Pfund color grader Koehler (New York, USA) was employed to determine color, according to Argentina regulations.

2.3. Reference method

Color determination was carried out according to Argentina regulations using a Pfund color grader Koehler (New York, USA). This device visually compares standard amber-colored glass wedge with liquid honey contained in a wedge-shaped cell placed in a sliding panel, as can be seen in Fig. 1a). Color measurement is done by sliding the panel until the color of the honey sample matches the color of the calibrated amber glass prism. The color intensity is expressed as a distance (in mm) along the amber-wedge and usually ranged between 1 and 140 mm beginning with very light-colored honey and increasing up to the darkest honey. Ten readings must be taken for each honey sample, five from each side of the glass wedge. Pfund values were obtained by a trained analyst.

2.4. Digital image acquisition

Approximately 10 g of honey sample was heated up to 40 °C to dissolve sugar crystals and was allowed to stand to remove air bubbles. The honey sample was placed in a Petri dish and using a Philips Webcam SPC900NC VGA the image was obtained. This procedure was performed in triplicate for each sample.

Fig. 2 shows the image capturing device [15]. The webcam was placed in a fixed position in the center of a circular daylight fluorescent lamp (22 W, temperature color 6400 K) over the honey sample. By this way the illumination and the distance between the sample and the camera remained constant. The sample holder used in this study was placed in a sealed box to prevent the passage of light. In order to avoid light scattering the interior walls of the box were covered with white paper.

In Fig. 3 digital images obtained from four different honey samples can be observed.

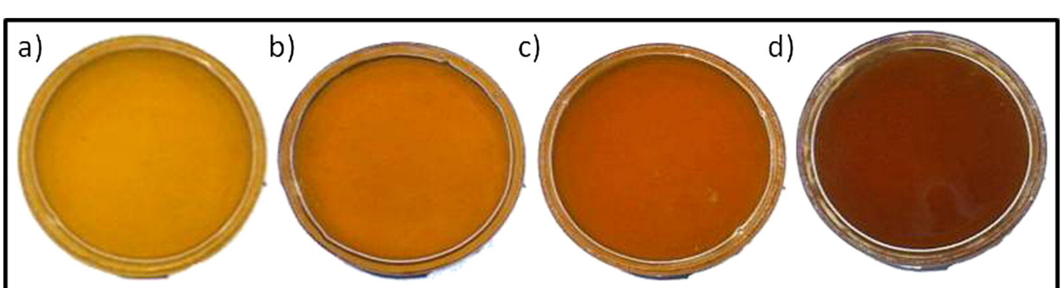

Fig. 3. Digital images obtained from four samples of honey with different color values. a) 14.7 mm Pfund, b) 35.0 mm Pfund, c) 57.1 mm Pfund and d) 95.0 mm Pfund.

Table 2
Final results of the HSB model.

HSB Model	LV[a]	Samples	Variables	RMSE	r^2	Slope	Bias	Correlation
Calibration	4	59	322	1.92[a]	0.98	0.98	-5.17×10^{-7}	0.99
Validation	4	59	322	3.07[b]	0.96	0.93	0.0173	0.98

[a] RMSEC.
[b] RMSECV.

2.5. Histograms and data analysis

To perform the analysis of the photos a circular region of 205 × 205 pixels was selected at the center of each one. Using the selected regions, histograms for each color channel (red, green, blue, hue, saturation, brightness and Grayscale) were performed. The histogram of an image is a graph representing the color intensity levels with respect to the number of pixels. Separate histograms of each color for the three replicas of the sample to obtain a "mean histogram" of each one (red, green, blue, hue, saturation, brightness and Grayscale) were obtained (Fig. 4). In this work the histograms were created using a free downloadable software ImageJ 1.44p. Each color component is composed of 256 tones, which are used as analytical information.

2.6. Chemometrics

The overall set of 80 honey samples were distributed into calibration (60) and prediction (20) subsets by using the SPXY (sample set partitioning based on joint x–y distances) algorithm in the Matlab® 6.5 software (The Mathworks) [16].

Partial Least Squares (PLS) method was employed to construct the chemometric models with The Unscrambler 9.7 software. Histograms from the three color models were set as independent variables (X) and color of honey samples expressed in mm Pfund was set as the dependent variable or response (Y). The optimal number of latent variables that gives the smallest prediction error to build the calibration model was used. Cross-validation (CV) was used to validate the model. In the PLS procedure, there are two steps: in the first one the calibration set is employed to establish the relationship between color histogram and mm Pfund values to build the models. In the second step the results of the calibration model are used to estimate the color of the prediction sample set.

In this work three different color models were employed to evaluate the influence of each one: (a) RGB, (b) Grayscale and (c) HSB.

3. Results and discussion

The evaluation of the calibration performance is estimated by computing the root mean square error of calibration (RMSEC). This value (expressed in mm Pfund) is a measurement of the average difference between predicted and measured response values, at the calibration stage. The root mean square error of cross validation (RMSECV) is a measurement of the average difference between predicted and measured response values at the validation stage. This parameter explains the ability of the model to predict a new sample. The bias can be defined as the mean of the errors. Table 1 shows the obtained results for the calibration models of honey samples in the test set using the RGB, HSB and Grayscale color models. As can be seen, the best calibration model was achieved employing the HSB color model.

In order to identify and delete noisy variables a Variable Importance in the Projection (VIP) based on the regression coefficients was applied [17–19]. Besides an outlier analysis was performed by means of an influence plot and an X–Y relation outlier plot. From this analysis an outlier sample was found and it was removed from the final model. Then, 446 variables and 1 sample were eliminated. The calibration and validation parameters of the resulting models are shown in Table 2. The quality of the calibration model was assessed based on the root mean standard error of cross-validation (3.07), correlation coefficient r^2 (0.96) and BIAS (0.0173) that shows acceptable values. Also it can be observed that in both cases there is a good correlation between the experimental and reference values.

The optimal number of components is obtained from Fig. 5 (cross validated error versus number of latent factors).

Fig. 6 shows a plot of residuals versus the color values of the calibration set samples. As can be seen the residuals are randomly distributed.

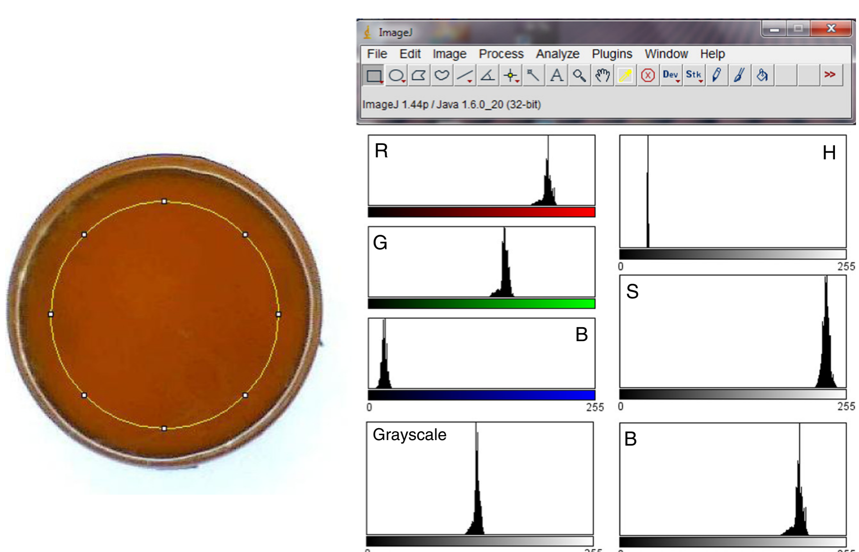

Fig. 4. a) Circular region selected in the center of the image. b) The ImageJ software and the corresponding histograms of Red (R), Green (G), Blue (B), Grayscale, Hue (H), Saturation (S) and Brightness (B).

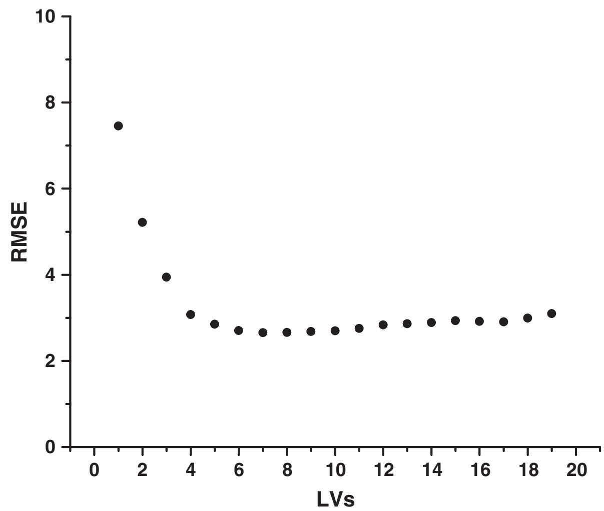

Fig. 5. Plot of cross validated error versus number of latent factors.

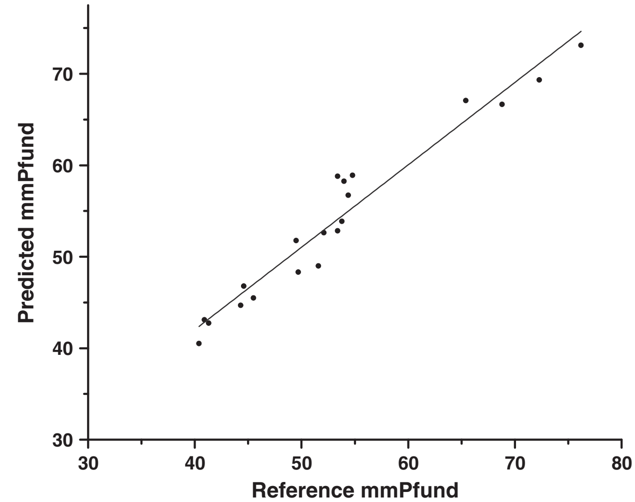

Fig. 7. Scatter plot showing the correlation between reference measurement by Pfund colorimeter and the prediction results by the digital image analysis with HSB color model.

Finally an examination of the linearity of the relationship between each PLS factor and the dependent variable was done. The corresponding values of the r^2 were 0.86 (PC1), 0.94 (PC2), 0.97 (PC3) and 0.98 (PC4).

Fig. 7 shows the correlation between the values determined by the Pfund method and the predicted values by digital image analysis, on the total prediction honey sample set. As can be seen, there is no evidence of systematic errors, since the points are randomly distributed around the bisector line along the entire range of y-values. The figure shows a good correlation between the obtained values by the reference method and those calculated by digital images by the HSB color model. The obtained values from the mean square prediction error (RMSEP) and correlation coefficient (r^2) were 2.46 mm Pfund and 0.97 respectively.

The obtained results can be explained taking into account that HSB is more representative of the way humans perceive colors. The HSB color model is defined from RGB model but specify colors using coordinates related with perceptual attributes. Then, it is expected to be HSB the most suitable color model to predict mm Pfund since in the calibration step the relationship between color histogram and a human perception of the color of the honey (Pfund method) is established.

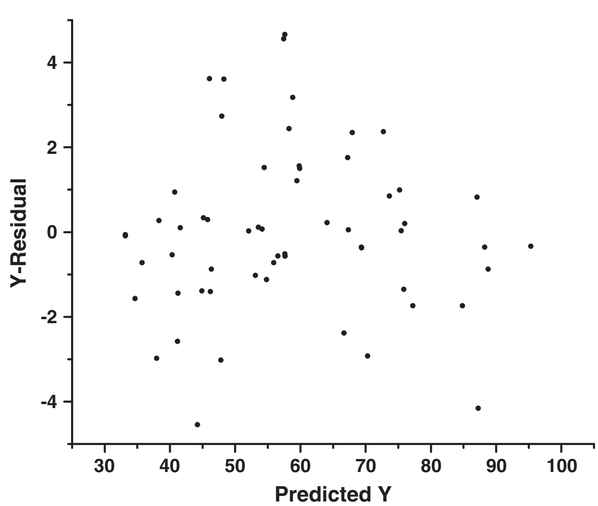

Fig. 6. Plot of residuals versus the color values of the training set samples.

4. Conclusion

In this paper a new method to determine the color of honey samples from southwest of the Buenos Aires province, Argentina was proposed. This method employ as analytical information the histograms obtained from digital image and a basic chemometric tool as PLS. Three color models, RGB, HSB and Grayscale, were used to analyze the digital images.

The obtained results employing HSB color model, indicate that the use of digital images combined with multivariate calibration is an excellent strategy analysis for quantifying color in these kinds of samples. The proposed method is simple, fast, require small amount of sample and it is simple to be automatized.

Furthermore, responds to the goals of green chemistry taking into account that not requires sample treatment and reagents and does not generate waste. Another advantage is that it allows an objective color evaluation, reducing errors due to human participation and enable examining chromaticity differences between samples regarding the Pfund method. Then the proposed method is a good alternative to the quality control of honey samples.

Acknowledgments

M.A. Dominguez acknowledge CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas) scolarship for support. The authors acknowledge financial support from Proyecto Grupo de Investigación (PGI 24/Q049) granted for Secretaría General de Ciencia y Tecnología, Universidad Nacional del Sur (Argentina). The authors thank Dr. Marcelo F. Pistonesi (Universidad Nacional del Sur) for his support and assistance during this paper.

References

[1] Plan estratégico Argentina apícola 2017. Secretaría de Agricultura, Ganadería, Pesca y Alimentos, SAGPyA, 2014. [<http://www.minagri.gob.ar> [Accessed March 2014]].
[2] Informe de Coyuntura N°2 — Sector Apícola. Área de Sectores Alimentarios — Dirección de Agroalimentos. Lic. Carolina Blengino. 2012.
[3] M.I. Isla, A. Craiga, R. Ordoñez, C. Zampini, J. Sayago, E. Bedascarrasbure, A. Alvarez, V. Salomón, L. Maldonado, Physico chemical and bioactive properties of honeys from Northwestern Argentina, LWT Food Sci. Technol. 44 (2011) 1922–1930.
[4] M.S. Finola, M.C. Lasagno, J.M. Marioli, Microbiological and chemical characterization of honeys from central Argentina, Food Chem. 100 (2007) 1649–1653.

[5] M.V. Baroni, C. Arrua, M.L. Nores, P. Fayé, M.P. Díaz, G.A. Chiabrando, D.A. Wunderlin, Composition of honey from Córdoba (Argentina): assessment of North/South provenance by chemometrics, Food Chem. 114 (2009) 727–733.
[6] I.K. Karabagias, A. Badeka, S. Kontakos, S. Karabournioti, M.G. Kontominas, Characterization and classification of *Thymus capitatus* (L.) honey according to geographical origin based on volatile compounds, physicochemical parameters and chemometrics, Food Res. Int. 55 (2014) 363–372.
[7] M.F. Silvano, M.S. Varela, M.A. Palacio, S. Ruffinengo, D.K. Yamul, Physicochemical parameters and sensory properties of honeys from Buenos Aires region, Food Chem. 152 (2014) 500–507.
[8] C.I.G. Tuberoso, I. Jerkovic, G. Sarais, F. Congiu, Z. Marijanovic, P. Marek Kus, Color evaluation of seventeen European unifloral honey types by means of spectrophotometrically determined CIE $L^*C^*_{ab}$ h^0_{ab} chromaticity coordinates, Food Chem. 145 (2014) 284–291.
[9] M.O. Iurlina, A.I. Saiz, R. Fritz, G.D. Manrique, Major flavonoids of Argentinean honeys. Optimization of the extraction method and analysis of their content in relationship to the geographical source of honeys, Food Chem. 115 (2009) 1141–1149.
[10] IRAM 15941–2, Honey. Pfund Color Determination, 2007.
[11] L.M. Gallez, A. Marconi, E. Tourn, L. Gonzalez-Miret, F.J. Heredia, Color of Honeys From the Southwestern Pampas Region Relationship Between the Pfund Color Scale and CIELAB Coordinates, in: J.L. Caivano, M.P. Buera (Eds.), Color in Food:Technological and Psychophysical Aspects CRC Press, 2012, pp. 133–142.
[12] M.L. González-Miret, F. Ayala, A. Terrab, J.F. Echávarri, A.I. Negueruela, F.J. Heredia, Simplified method for calculating colour of honey by application of the characteristic vector method, Food Res. Int. 40 (2007) 1080–1086.
[13] I. Escriche, M. Kadar, E. Domenech, L. Gil-Sanchez, A potentiometric electronic tongue for the discrimination of honey according to the botanical origin. Comparison with traditional methodologies: physicochemical parameters and volatile profile, J. Food Eng. 109 (2012) 449–456.
[14] M.A. Domínguez, P.H.G. Dias Diniz, M.S. Di Nezio, M.C.U. Araújo, M.E. Centurión, Geographical origin classification of Argentinean honeys using a digital image-based flow-batch system, Microchem. J. 112 (2014) 104–108.
[15] P.H.G.D. Diniz, H.V. Dantas, K.D.T. Melo, M.F. Barbosa, D.P. Harding, E.C.L. Nascimento, M.F. Pistonesi, B.S.F. Band, M.C.U. Araújo, Using a simple digital camera and SPA-LDA modeling to screen teas, Anal. Methods 4 (2012) 2648–2652.
[16] R.K.H. Galvão, M.C.U. Araujo, G.E. José, M.J.C. Pontes, E.C. Silva, T.C.B. Saldanha, A method for calibration and validation subset, Talanta 67 (2005) 736–740.
[17] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in partial least squares regression, Chemom. Intell. Lab. Syst. 118 (2012) 62–69.
[18] A. Wang, L. Xie, Technology using near infrared spectroscopic and multivariate analysis to determine the soluble solids content of citrus fruit, J. Food Eng. 143 (2014) 17–24.
[19] F. Liu, Y. He, L. Wang, Determination of effective wavelengths for discrimination of fruit vinegars using near infrared spectroscopy and multivariate analysis, Anal. Chim. Acta. 615 (2008) 10–17.