Talanta 81 (2010) 1010–1015

Contents lists available at ScienceDirect

Talanta

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

Multivariate analysis of reflectance spectra from propolis: Geographical variation in Romanian samples

Augustin Cătălin Moţ, Florin Soponar, Costel Sârbu*

Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, 11 Arany Janos St, RO-400028 Cluj-Napoca, Romania

ARTICLE INFO

Article history:
Received 11 November 2009
Received in revised form 20 January 2010
Accepted 23 January 2010
Available online 1 February 2010

Keywords:
Romanian propolis
Floral origin discrimination
Reflectance spectroscopy
Cluster analysis
Principal component analysis
Linear discriminant analysis

ABSTRACT

The present study described reflectance spectroscopy as a suitable analytical tool to discriminate the floral origin of 39 Romanian propolis samples. Relevant differences between the UV–vis reflectance spectra of the investigated propolis samples within the 220–850 nm spectral range were found. The results obtained applying cluster analysis, principal component analysis and linear discriminant analysis to the digitized data of zero order, zero order normalized and first order derivative spectra support the reliability of this technique. In addition, the application of the linear discriminant analysis to the score matrices corresponding to the first principal components appeared to be an illuminating solution. Generally, the samples have been assigned to two large groups in a good agreement with their vegetal sampling location, samples originating from predominant forest area and samples originating from meadows. Within the first group, two subgroups were identified according to the dominant type of the forest, deciduous or resinous, while within the last group three subgroups were found according to the extend and variety of the meadow.

1. Introduction

Honeybees (*Apis mellifera*) collect natural occurring resinous substances from various plants and trees and mix them partly unintentionally with beeswax, pollen and salivary secretions producing a well known sticky solid material called propolis. Propolis is used by the bees to seal the holes from the hives stabilizing the inner temperature around 35 °C, strengthening the combs but its major role is in defending their community working as a "chemical weapon" against intruders mummifying them after they were killed. The hive is kept aseptic and clean mainly due to antibacterial, antifungal, antitrypanosomal, antimicrobial activities of propolis [1–3]. Several studies proved other important biological activities such as antioxidant [4,5] antiinflamatory [6], anticancer [7], antihepatotoxic [8] or benefic effects in dental care [9], gastric ulcer [10], cartilage protection [11] and immune system [12]. Some of these activities were known from centuries all over the world, and is still used not only in folk medicine but also in modern medicine [13] as ingredient in commercial products such as vitamin C, toothpaste, medicinal syrups and so on.

The aspect, texture and chemical composition of propolis is variable and depends upon the climate, season and mainly upon the local flora which is exploited by the bees, thus on the geographic characteristics of the region [14]. Roughly, raw propolis contains 50% resin and balsam, 30% wax, 10% aromatic oils, 5% pollen, 5% other organic substances, traces of inorganic salts. The bioactive components are thought to be polyphenols, phenolic acids and their esters and flavonoids [15,16].

The chemical diversity mainly caused by the plant origin and the problem of standardization of propolis are issues that still need effort to be clarified [17]. Since the chemistry and biological activities of propolis depend upon its geographical origin a proper method to discriminate its origin is needed [18]. In his review, Bankova underlines the actual necessity of research concerning plant origin and composition–biological activity relationship of propolis types [19]. Several analytical methods based on chromatography techniques (HPLC and GC), mass spectrometry techniques (ICP-MS and GC–MS), spectroscopy (NMR and IR) were most frequently used until now to determine the geographical origin of food products [20]. Recently a high-performance liquid-chromatography method was proposed and applied to assess the geographical origin of several types of propolis [21]. Spectroscopic techniques (FT-IR and 2D IR) were also proved to be useful to achieve the same goal [22]. Nevertheless these methods are laborious, destructive, time consuming and require preparation of the samples of the raw propolis, process where some information concerning the original samples might be lost.

Recently, reflectance spectroscopy proved to be a useful tool in the discrimination of some food products such as the European wheat varieties [23] or Greek red wines [24].

The present paper discussed and successfully applied reflectance spectroscopy as a tool to discriminate the origin

* Corresponding author. Tel.: +40 264 593833; fax: +40 264 590818.
E-mail address: csarbu@chem.ubbcluj.ro (C. Sârbu).

0039-9140/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2010.01.052

of various samples of propolis from several places of Romania, developing in this order a proper methodology, including experimental approach and chemometric treatment of the reflectance digitized spectra. It was proved that the reflectance spectra derived from the raw samples of propolis carry useful information about their origin according to the color, physical and overall chemical composition.

2. Theoretical considerations

In reflectance spectroscopy, the Lambert–Beer law from absorption molecular spectroscopy is replaced by the Kubelka–Munk function [25] which establishes a mathematical relationship among the spectral reflection ($R\%$), absorption coefficient (K) and scattering coefficient (S).

$$F(R) = \frac{(1-R)^2}{2R} = \frac{K}{S} \quad (1)$$

According to Kubelka–Munk theory, the absorption coefficient is a measure of chemical composition and the scattering coefficient carries information about the particle size and their distribution. The value of K and S can be determined by considering that each of these two coefficients is an additive function [26] as is illustrated by the following equation:

$$F(R_M) = \frac{(1-R_M)^2}{2R_M} = \left(\frac{K}{S}\right)_M = \frac{\sum C_i K_i}{\sum C_i S_i} \quad (2)$$

where M refers to mixture, C_i is the mass fraction of component i with absorption coefficient K_i and scattering coefficient S_i. Finally, considering only two components, the sample and a white standard ($K_{st} = 0$), Eq. (3) can be obtained by suitable transformations [27]:

$$\frac{2R_M}{(1-R)^2} = \frac{C_p S_p + C_{st} S_{st}}{C_p K_p} = \frac{S_p}{K_p} + \frac{C_{st}}{C_p}\frac{S_{st}}{K_p} \quad (3)$$

This method was successfully applied to the study of influence of iron oxides on soil color [28]. Thus reflectance spectroscopy appears to be a proper tool for discrimination of geographical origin of propolis and when the spectral data are processed and interpreted by means of chemometric techniques, its objectivity and reliability is strengthened.

3. Experimental

3.1. Sampling sites, sampling and UV–vis spectra

The UV–vis reflectance spectra were recorded using a V-550 Jasco spectrophotometer equipped with integrating sphere attachment (Jasco ISV 469). The recording speed was 400 nm min^{-1} while the wavelength reproducibility was ±0.1 nm. The propolis samples were collected from several places with different flora from Romania (Table 1) by professional beekeepers (in almost cases being assisted by a specialized chemist) and were deposited in a −20 °C freezer during the research period.

The raw propolis was homogenized using a grinding mortar and a pestle; this process was accomplished as soon as it was removed from the freezer while the propolis was still brittle. A proper amount (~0.7 g) of the resulting homogenized material was pressed into a large pastille with 20–25 mm diameter and ~1 mm thickness using a manual pressing machine. The resulting pastille was used to record the UV–vis reflectance spectra. All spectra were recorded at 1 nm resolution in the range 220–850 nm. Three reflectance spectra from three different zones of the lozenge were measured and the mean spectrum was calculated and further used. In addition, the reflectance spectra of three pastilles obtained in the similar way from the same type of propolis were also recorded in order to evaluate the sample preparation variability. The first order spectra were calculated using *Spectra Analysis* option (included in Jasco spectrophotometer software) according to Savitzky–Golay algorithm [29]. In addition, the zero order spectra were normalized within the range 0–1, using Eq. (4), where y_i is the new normalized variable, x_i is the raw reflectance value at a given wavelength and x_{max} and x_{min} are the maximum and the minimum value, respectively, corresponding to each spectrum.

$$y_i = \frac{x_{max} - x_i}{x_{max} - x_{min}} \quad (4)$$

3.2. Chemometric methods and data matrices

Many problems from chemistry and other technical fields are strongly related to principal component analysis (PCA), linear discriminant analysis (LDA) and cluster analysis (CA) [30–33]. PCA is a typical projection method, which allows one to find out the structure of data by depicted the objects (samples) in a reduced

Table 1
Propolis samples code, location and description.

Types of propolis	Sample group codes	Number of samples	Origin place of samples (place, county)	Origin coordinates	Flora type	Color	Texture
T1	T11	4	Nadaselu, Cluj	46°50′N/23°27′E	Meadow, low content mixture of deciduous forest	Brown	Rigid powder
	T12	3	Sarmasu, Mures	46°45N/24°12′E		Brown	Rigid powder
	T13	4	Gledin, Bistrita-Nasaud	46°57′N/24°33′E		Yellowish brown	Waxy powder
T2	T21	1	Pestera, Constanta	44°11′N/28°07′E	Grassy meadow	Brownish yellow	Waxy
	T22	2	Livada, Arad	46°13′N/21°23′E		Yellowish brown	Waxy
	T23	1	Sanleani, Arad	46°12′N/21°23′E		Dark greenish brown	Rigid waxy
	T24	2	Carei, Satu-Mare	47°40′N/22°25′E		Brownish yellow	Rigid waxy
T3	T31	2	Cristian, Sibiu	45°48′N/24°02′E	Meadow, high content of mixture of deciduous forests	Brown	Waxy
	T32	4	Orlat, Sibiu	45°45′N/23°57′E		Brown	Waxy
T4	T41	2	Brebi, Salaj	47°13′N/23°10′E	Mixture of deciduous and resinous forests	Reddish brown	Rigid
	T42	2	Romuli, Bistrita-Nasaud	47°32′N/24°24′E		Reddish brown	Rigid powder
	T43	2	Fagaras, Brasov	45°47′N/24°58′E		Brown	Rigid waxy
T5	T51	7	Romuli, Bistrita-Nasaud	47°32′N/24°25′E	Mixture of deciduous forests	Light reddish brown	Rigid waxy
	T52	2	Gura Raului, Sibiu	45°43′N/23°58′E		Light brown	Rigid waxy
T6	T6	1	Pastoral[a]	Pastoral	Complex	Dark brown	Rigid waxy

[a] Pastoral refers to the fact that the beehive was moved to various places during a season.

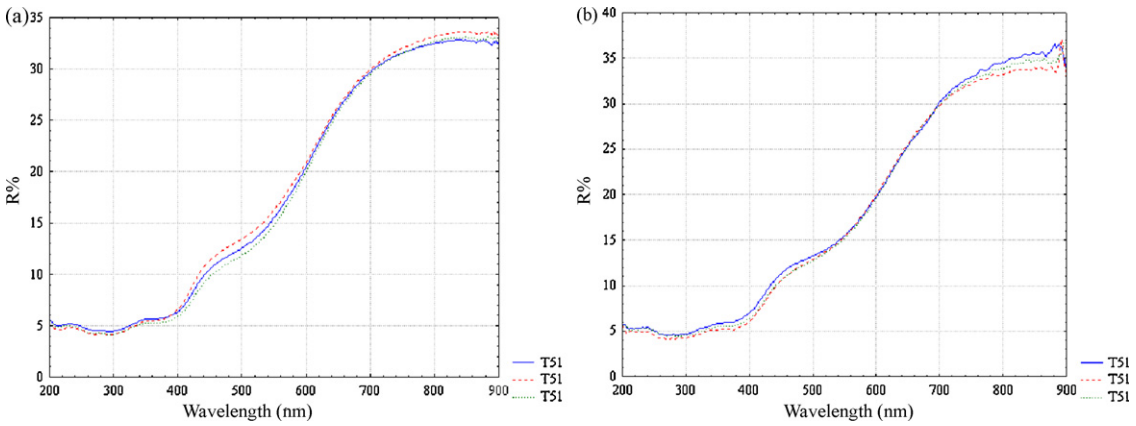

Fig. 1. UV–vis zero order spectra: (a) corresponding to three different zones of the same sample and (b) corresponding to three different pastilles of the same sample.

coordinate system with maximum possible information. Score plots, for example, are very useful as a display tool for examining the relationships between objects, looking for groups and trends, sorting out outliers. LDA is based on the extraction of linear discriminant functions of the independent variables by means of a qualitative dependent variables and several quantitative independent variables. LDA is a supervised method. We can also visualize how the functions discriminate between groups by plotting the individual scores for the discriminant functions. CA is a well known and widely used unsupervised clustering procedure with its hierarchical and non-hierarchical approaches. There is a wide variability of ways how to measure the distances between objects and how to group objects based on various distance measures. However, the Euclidean distance and *Ward's method* as linkage procedure are strongly recommended [31,32].

CA, PCA and LDA were successfully applied to the spectroscopic data matrices (39 samples × 630 variables) and/or score matrices (38 samples × 15 PCs) using *Statistica 7.1* (StatSoft, Inc., Tulsa, USA).

4. Results and discussion

As we have already mentioned in the experimental paragraph, the reflectance spectrum corresponding to three different zones of each of three pastilles prepared from the same type of propolis were measured in order to evaluate the variability. The results presented in Fig. 1a and b are quite similar and illustrate very good repeatability and reproducibility in both replicate and sample preparation variability.

The UV–vis reflectance spectra were relatively simple but significant differences between samples could be easily observed. Light colored samples exhibit higher reflectance spectra with a plateau in 850–700 nm region and a strong absorption centered on 380 nm (Fig. 2, T5).

Dark samples with a reddish tint (T4, originating from zones with mixture of deciduous and resinous forests) present a region between 850 and 650 with an almost constant slope and a very broad absorption centered at 520 nm and higher absorption at 350 nm compared to the other samples. The T1, T2 and T3 samples are pretty similar, the difference between the samples being the slope variation of the 420–850 nm spectral region. Waxy samples originated from forests (especially those from resinous forests) exhibit a strong decreasing of the reflectance starting around 480 nm, with a shoulder centered on 420, this fact being observable as a negative spectral band in first derivative spectra (Fig. 3). The waxier the samples are, the largest the spectral band is.

Within the UV region of the spectra, all the samples present a strong absorption centered on 280–310, the 350 shoulder being specific to the samples originated from forests (Fig. 2 inset).

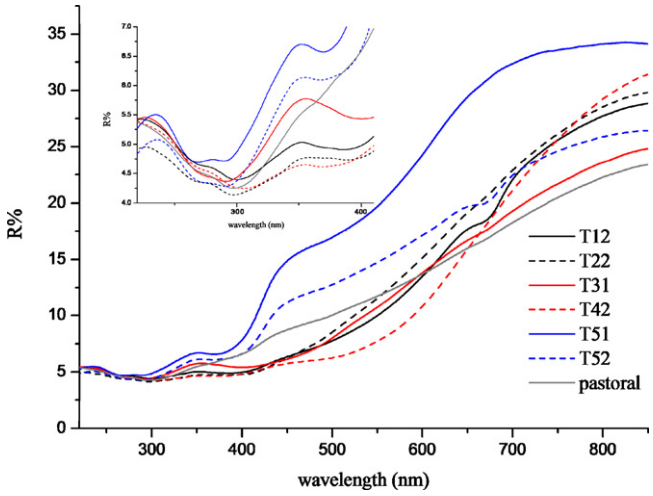

Fig. 2. UV–vis zero order spectra.

Based on their floral origin, the propolis samples were classified into five different types according to the vegetation zone (Table 1).

In the first step of digitized data analysis, a study of the structure of data by cluster analysis was carried out. In each case, a matrix (39 samples × 360 values) consisting of the squared Euclidean distance was used as similarity matrix. To obtained clusters, a hierarchical agglomerative method was employed, the Ward method. The

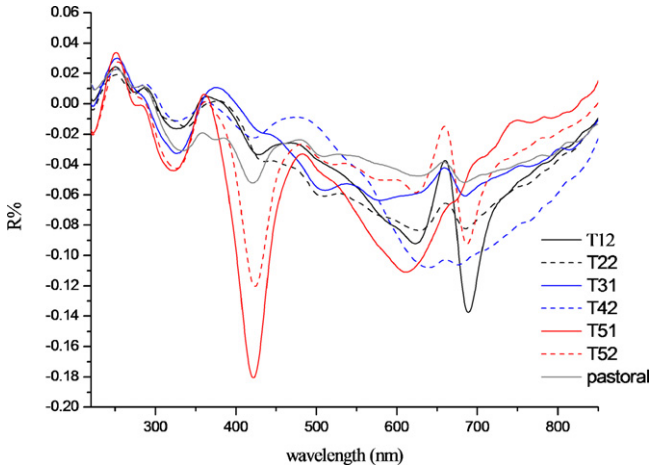

Fig. 3. UV–vis first order spectra.

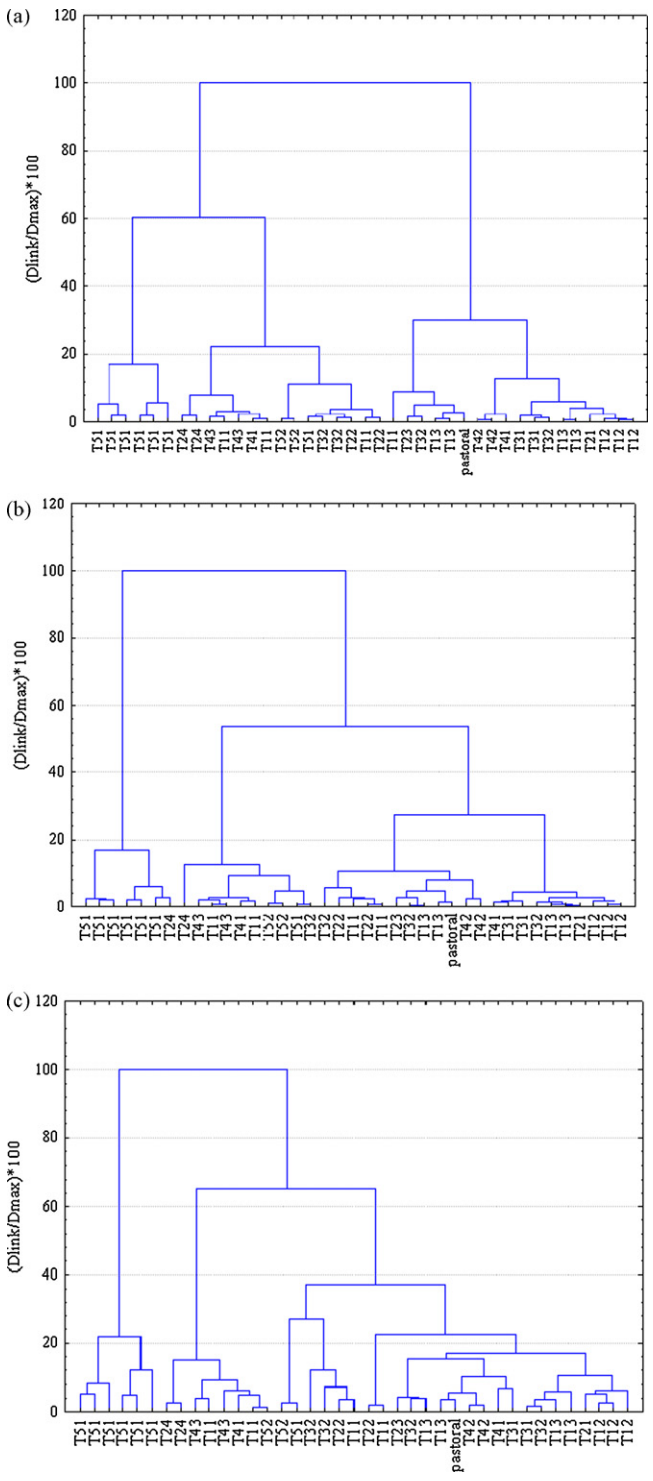

Fig. 4. Hierarchical clustering of Romanian propolis samples using spectra data: (a) zero order spectra; (b) normalized zero order spectra; and (c) first order spectra.

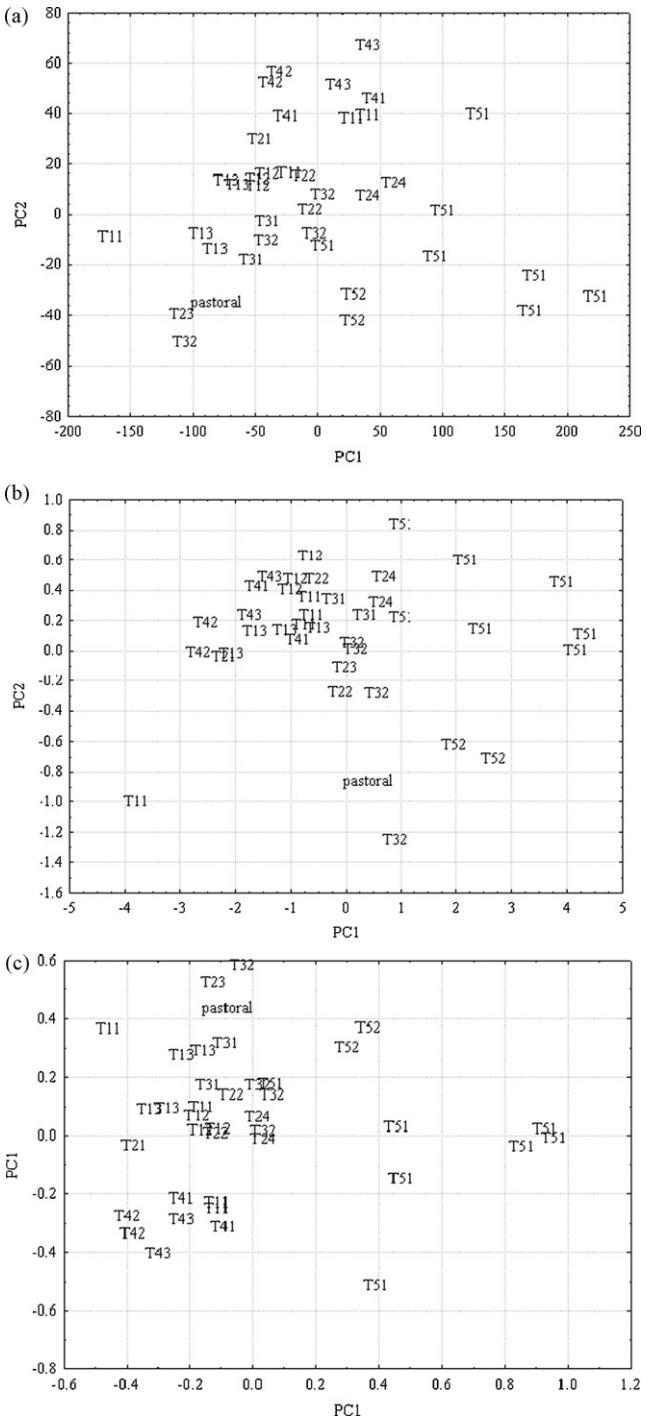

Fig. 5. PC1–PC2 score plot: (a) zero order spectra; (b) normalized zero order spectra; and (c) first order spectra.

results obtained, presented as a dendrogram (Fig. 4a and b) show the presence of clusters of propolis samples corresponding to the geographic origin rather than the vegetal sampling location (T51, T52, T24, T12, and T41). We have to remark also the high similarity of the three dendrograms (zero order spectra, normalized zero order spectra and first order spectra) with a better separation of T51 and T24 samples in the case of zero and first order spectra.

The results of PCA for each kind of spectra (centered data) are presented in Fig. 5a–c. The first PCs explain more than 99% of the total data variance in the case of zero order spectra (original and normalized) and only approximately 90% of the total data variance, respectively in the case of first order spectra. Fig. 5a–c indicates that the compression was quite effective in all cases, because the data variance is well distributed over the first two PCs and as a direct consequence the grouping of samples according to their vegetation zone can be revealed. The samples originating from forest area (T4 and T5) are well differentiated in the PC1–PC2 score plot because the floral species are totally different from the other regions; the spread of samples in the T5 group (mixtures of deciduous forest) may be explained by variability and diversity of deciduous forests.

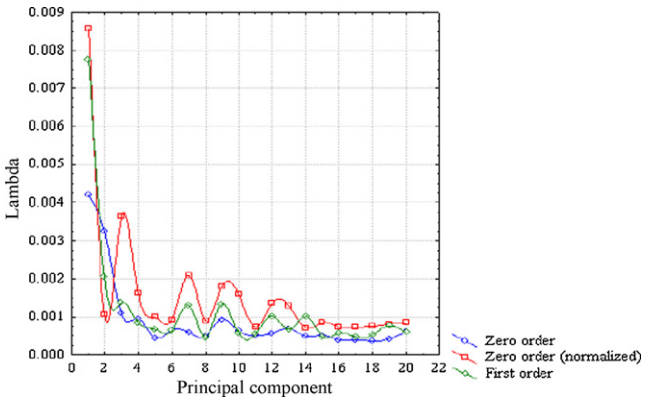

Fig. 6. Profile variation of Wilks' λ as a function of principal components number.

The groups of samples originating from the other area (T1, T2 and T3) which present also low or high level of meadow are more or less overlapped according with the large variety and high similarity of floral species. The position of the pastoral sample in the overlapping regions illustrates clearly the influence of different floral species. The distinction between samples originating from forest area and those with intense meadow is more evident in the case of normalized zero order and first order spectra (Fig. 4b and c).

The results obtained using HCA and PCA showed that the samples from all five groups fell into discernible, but partially overlapping groups according to their respective vegetal sampling location and/or geographic zone. In order to get more information concerning the similarities and differences of the investigated propolis samples, LDA was used to perform reallocation of samples to their respective groups, based on the scores corresponding to the first principal components. The quality of discrimination and the selection of the most discriminant independent variables can be evaluated by applying different criteria. The Wilks' lambda *F* test is used to test whether the discriminant model as a whole is significant; the larger the lambda, the more likely it is significant [34]. Following the variation of the Wilks' λ as a function of the number of principal components (variables) (Fig. 6) one can be easily observed that around 15 variables (principal components) should be enough for a relevant discrimination between the samples.

Table 2
Classification matrix of Romanian propolis (38 samples) using scores corresponding to the first 15 principal components.

Spectrum	Group	Correct classification (%)	Observed classification				
			G1	G2	G3	G4	G5
Zero order	T1	100.00	11	0	0	0	0
	T2	100.00	0	6	0	0	0
	T3	83.33	1	0	5	0	0
	T4	100.00	0	0	0	6	0
	T5	100.00	0	0	0	0	9
	Total	97.37	12	6	5	6	9
Zero order (normalized)	T1	100.00	11	0	0	0	0
	T2	100.00	0	6	0	0	0
	T3	83.33	1	0	5	0	0
	T4	100.00	0	0	0	6	0
	T5	100.00	0	0	0	0	9
	Total	97.37	12	6	5	6	9
First order	T1	100.00	11	0	0	0	0
	T2	100.00	0	6	0	0	0
	T3	100.00	0	0	6	0	0
	T4	100.00	0	0	0	6	0
	T5	100.00	0	0	0	0	9
	Total	100.00	11	6	6	6	9

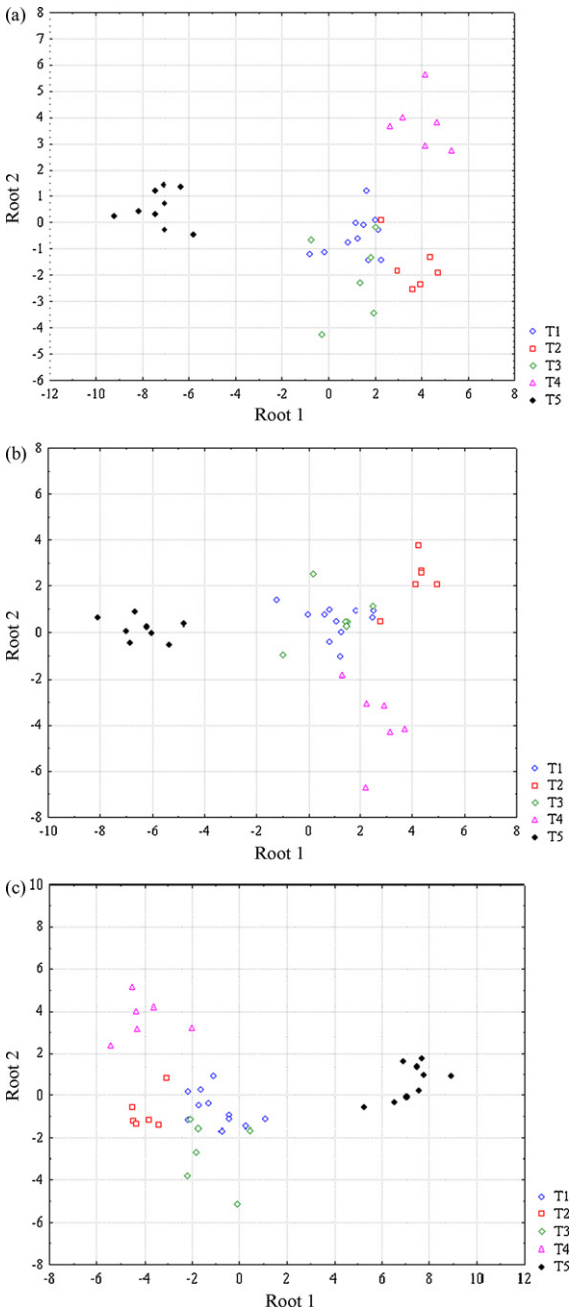

Fig. 7. R1–R2 canonical score plot: (a) zero order spectra; (b) normalized zero order spectra; and (c) first order spectra.

The results summarized in Table 2 clearly support our supposition and show a total separation (100%) in the case of first order spectra. All the statements above are also well illustrated in the R1–R2 score plots presented in Fig. 7a–c. The vegetation distinctiveness of propolis was confirmed and all samples were reassigned to their floral origin with significant differences between forest and meadow groups. In addition, by applying Eq. (3) it possible to determine the level of the absorption coefficient (K) and scattering coefficient (S) for propolis. In this order, the sample was mixed with variable amounts of a white standard ($BaSO_4$) whose $K_{st} = 0$ and for simplicity S_{st} was set to unity and thus relative S_p rather than absolute was obtained [27]. By plotting the $1/F(R_M)$ against C_{st}/C_p for different mixtures, K_p can be calculated from the slope and the intercept gives the values of S_p. The results obtained are depicted in Fig. 8a and b. The spectral behavior of the scattering coefficients

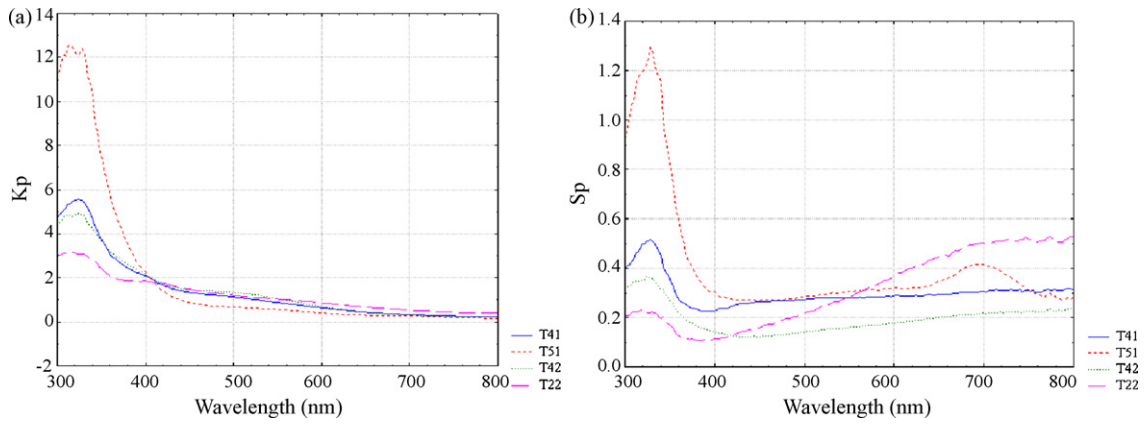

Fig. 8. The spectral profile of the absorption coefficient level (a) and scattering coefficient level (b) obtained for some propolis samples.

for samples belonging to the same group are similar (T41 and T42) indicating similar mean particle size and similar particle size distribution while the spectral pattern of the absorption coefficients of all four samples presents a similar band around 320 nm most likely due to the polyphenolics which are major components in all the samples [35]. Nevertheless, noticeable differences at numerous wavelengths can be observed due to the manifold of constituents.

5. Conclusions

The research work developed in this paper describes, for the first time, the proper methodology, sampling technique and data treatment of the reflectance measurements which were efficiently applied to the discrimination of the origin of propolis samples from various floral places. Thirty-nine samples from 12 places from Romania have been analyzed and 5 different types of propolis have been identified. The most important variable which influences the quality of propolis was found to be, as suspected, the vegetation origin. The obtained results demonstrate that the reflectance spectroscopy assisted by multivariate methods (cluster analysis, principal component analysis and discriminant analysis) is a proper tool for discrimination of geographical origin of propolis. Two large types of propolis have been indentified, those originated from forest area and those with intense meadows or mixture of meadows with forests. Within the first group there are two subgroups depending on the type of forests, deciduous or resinous while within the last group the samples were categorized into three subgroups according to content of meadow, not present at all, low and high.

References

[1] M. Kartal, S. Yıldız, S. Kaya, S. Kurucu, G. Topçu, J. Ethnopharmacol. 86 (2003) 69–73.
[2] E. Prytzyk, A.P. Dantas, K. Salomao, A.S. Pereira, V.S. Bankova, S.L. De Castro, F.R. Aquino Neto, J. Ethnopharmacol. 88 (2003) 189–193.
[3] K. Hikmet, M. Nazime, Afr. J. Biotechnol. 5 (2006) 1151–1153.
[4] F. Sun, S. Hayami, S. Haruna, Y. Ogiri, K. Tanaka, Y. Yamada, K. Ikeda, H. Yamada, H. Sugimoto, N. Kawai, S. Kojo, J. Agric. Food Chem. 48 (2000) 1462–1465.
[5] J. Ivona, B. Mirza, M. Ana, B. Erim, B. Kajo, M.S. Marica, Mol. Online 12 (2007) 1006–1021.
[6] H. Miyataka, M. Nishiki, H. Matsumoto, T. Fujimoto, M. Matsuka, T. Satoh, Biol. Pharmaceut. Bull. 20 (1997) (1997) 496–501.
[7] G.A. Burdock, Food Chem. Toxicol. 36 (1998) 341–363.
[8] A.H. Banskota, Y. Tezuka, I.K. Adnyana, Phytomedicine 8 (2001) 16–23.
[9] H. Koo, P.L. Rosalen, J.A. Cury, Y.K. Park, W.H. Bowen, Antimicrob. Agents Chemother. 46 (2002) 1302–1309.
[10] P.B. Muriel, P.B.S. Joao, K.B. Jairo, F.A. Sergio, J. Ethnopharmacol. 110 (2007) 567–571.
[11] V. Cardilea, A. Panicob, B. Gentileb, F. Borrellic, A. Russod, Life Sci. 73 (2003) 1027–1035.
[12] J.M. Sforcin, J. Ethnopharmacol. 113 (2007) 1–14.
[13] C. Stefano, C. Francesco, Fitoterapia 73 (2002) 1–6.
[14] K.R. Markham, K.A. Mitchell, A.L. Wilkins, J.A. Daldy, Y. Lu, Phytochemistry 42 (1996) 205–211.
[15] M.C. Marcucci, Apidologie 26 (1995) 83–99.
[16] V.S. Bankova, S.S. Popov, N.I. Marekov, J. Nat. Prod. 46 (1983) 471–474.
[17] V.S. Bankova, J. Ethnopharmacol. 100 (2005) 114–117.
[18] A. Salatino, E.W. Teixeira, G. Negri, D. Message, Evid. Base Complement. Altern. Med. 2 (2005) 33–38.
[19] V.S. Bankova, Evid. Base Complement. Altern. Med. 2 (2005) 29–32.
[20] M.L. Dion, M.R. Saskia, Food Chem. 107 (2007) 897–911.
[21] J. Zhou, Y. Li, J. Zhao, X. Xue, L. Wu, F. Chen, Food Chem. 108 (2008) 749–759.
[22] Y.W. Wu, S.Q. Sun, J. Zhao, Y. Li, Q. Zhou, J. Mol. Struct. 884 (2008) 48–54.
[23] C. Miralbes, Food Chem. 106 (2008) 386–389.
[24] P.A. Tarantilisa, V.E. Troianoub, C.S. Pappasa, Y.S. Kotseridisb, M.G. Polissioua, Food Chem. 1 (2008) 192–196.
[25] P. Kubelka, F. Munk, Z. Technol. Phys. 12 (1931) 593–601.
[26] D.R. Duncan, J. Oil Colour Chem. Assoc. 32 (1949) 296–391.
[27] V. Barron, J. Torrent, Methods of Soil Analysis, Soil Science Society of America, Madison, Wisconsin, 2008, pp. 367–385.
[28] V. Barron, J. Torrent, J. Soil Sci. 37 (1986) 499–510.
[29] A. Savitzky, M.J.E. Golay, Anal. Chem. 36 (1964) 1627–1639.
[30] T. Cundari, J. Deng, H. Pop, C. Sârbu, J. Chem. Inf. Comput. Sci. 40 (2000) 1052.
[31] J.W. Einax, H.W. Zwanziger, S. Geiß, Chemometrics in Environmental Analysis, Wiley, Chichester, 1997.
[32] M. Otto, Chemometrics, Wiley-VCH, Weinheim, 1998.
[33] R.C. Brereton, Applied Chemometrics for Scientists, John Wiley & Sons Ltd., Chichester, 2007.
[34] C. Sârbu, H.F. Pop, R.S. Elekes, G. Covaci, Rev. Chim. (Bucharest) 59 (2008) 1237–1241.
[35] A.C. Mot, G. Damian, C. Sârbu, R. Silaghi-Dumitrescu, Redox Rep. 14 (2009) 267–274.