1. Krastena NIKOLOVA, 2. Diana TSANKOVA, 3. Tinko EVTIMOV

FLUORESCENCE SPECTROSCOPY, COLORIMETRY AND NEURAL NETWORKS IN DISTINGUISHING DIFFERENT TYPES OF HONEY

1-2.University of Food Technologies, Plovdiv, BULGARIA
3. Plovdiv University "Paisii Hilendarski", Plovdiv, BULGARIA

ABSTACT: The aim of the article is to investigate the optical properties of Bulgarian honey in regard to the potential of honey discrimination on the base of its botanical origin. Samples from three types of honey (acacia, linden, and honeydew) are measured by a fluorescence spectrometer recording emission from 350 to 800 nm with excitation at 370, 395 and 405 nm. Principal components analysis (PCA) is used for reducing the number of inputs (wavelengths) and for a proper visualization of the experimental results. A combination of fluorescence emission spectra with some colorimetric parameters (CIELab) is used for training data of a neural network (a multilayered perceptron) with back propagation learning algorithm. The good accuracy of the proposed neural network based honey classifier is confirmed by a validation test carried out in MATLAB environment.
Keywords: fluorescence spectroscopy, honey discrimination, PCA, artificial neural network

1. INTRODUCTION

Honey is a mixture of sugars (70-80 %) and water (10-20 %) containing a large number of minor constituents. The main sugars found in honey are the monosaccharides, fructose and glucose [10]. Honey is known to be rich in enzymatic and non-enzymatic antioxidants, including glucose-oxidase, catalase, flavonoids, ascorbic acid, phenolic acids and carotenoids [1, 2]. Its composition depends highly on the type of flowers as well as the climatic conditions [19].

Production of natural honey is a laborious process, which is time consuming and involves a lot of cost. Therefore honey is often subject to falsification by adding sugar and other impurities. Furthermore, the botanical and geographical declaration of the origin seems to be one of the fundamental aspects of the honey quality that affects its commercial value [13, 17]. So in order to prevent fraud in the labeling, it should be developed a means of distinguishing between different types of honey.

At the current stage of knowledge, a reliable authentication of floral origin of honey can be achieved by a global interpretation of sensory, pollen and physicochemical analyses carried out by an expert [11, 14, 15]. The content of different phenolic compounds is recognized to well reflect the type of honey and its quality, because phenolic acids and flavonoids are inherent chemical markers of the floral origin [12, 17]. Unfortunately, the most of these methods are generally too time-consuming, complex, and labor intensive for quality control application or require very specialized personnel to interpret the results.

In addition, most of the analytical techniques involve some kind of sample pre-treatment. The advantages of the technique of spectroscopy (visible, near and middle infrared, fluorescent) with respect to other methods are the non-invasive approach, the relatively easy and quick data acquisition. The principal advantages of fluorescence spectroscopy, pointed out by almost all

authors, are its rapidity and sensitivity [16] (100–1000 times more sensitive than other spectrophotometric techniques [14, 15]).

Food contains many different fluorophores, whose signals overlap and make it impossible to measure the concentration of a single compound. Nevertheless, the shape of normalized fluorescence spectra in combination with multivariate statistics can be used to characterize and identify different food [14, 15], including different types of honey.

Due to the correlation of incoming data, statistical classification methods (linear discriminant analysis, naive Bayesian classifier, and so on) encounter some computational difficulties such as 'badly scaled or close to singular matrix'. Artificial neural networks overcome these problems at processing raw data and can be used for multivariate analysis to create more accurate classifiers [6].

The purpose of this study is to investigate the optical properties of Bulgarian honey and the possibility of recognizing its botanical origin using fluorescence spectroscopy in right-angle fluorescence mode. Spectroscopic data obtained undergo subsequent statistical processing including principal components analysis (PCA), which is used for reducing the input space dimension and visualizing the clusters formed by different types of honey. An artificial neural network (NN) with Back propagation (BP) learning algorithm is proposed to classify honey in appropriate classes related to its floral origin. The performance of the neural calibration model is confirmed by leave-one-out-cross validation test in MATLAB environment.

2. MATERIAL AND METHODS

Honey spectrum acquisition

Thirty-two samples of three different types of Bulgarian honey (acacia – 8 samples; linden – 10 samples; and honeydew – 14 samples) were purchased from supermarkets (Lexie, Kaufland, Piccadilly) and from private producers. Before spectral measurement, the honey samples were placed in a water container at 50^{0}C until the soluble substances fully dissolved. Then the samples were annealed at room temperature (25-26^{0}C).

The fluorescence spectral characteristics of the honey were taken with a fiber optic spectrometer (AvaSpec-2038, Avantes) with sensitivity in the (200-1100) nm range. The sources used to measure the fluorescence spectra are 370 nm, 395 nm, 405 nm light emitting diodes (LEDs). The resolution of the spectrometer is about 8 nm for a 200 μm input slit. An optical fiber with a diameter of 200 μm is used to bring light to the probe and to measure the scattered and fluorescent light. A collimator with a lens of an aperture D = 5 mm is used to gather more light and send it to the receiver.

Generally, with classical right-angle fluorescence spectroscopy, the measurements are carried out in dilute solutions where the absorbance is below 0.1 [14, 15]. At a higher absorbance rate, the fluorescence intensity decreases due to the inner filter effect. In that case the front-face fluorescence spectroscopy is more suitable for use. In the presented study, in order to measure the fluorescence spectra of honey (especially dark honeydew honey) without dilution, the cuvette holder was modified as follows. The first probe (optical fiber) was placed between two glass slides, which were fixed by a threshold, consistent with the diameter of the probe. The second probe (LEDs) was fixed on the upper glass, 90^{0}–angle to the first and the minimum distance between them. Honey was located between the two slides. The resulting emission spectra with excitation at 370 nm, 395 nm, 405 nm were normalized by dividing with the maximum intensity value of the respective excitation signal.

Colour, β-carotene and water content measuring

The following measurements were determined according to the methods of the European Honey Commission [5]. All measurements were performed at room temperature. Colorimetric study of honey was made using a software package VISIONlite ColorCalc for spectrophotometer Helios Omega. It was used mode 'Advanced', i.e. calculations were performed in the range of 380 nm - 780 nm (instead of 'Basic' mode: 400 nm – 700 nm). The honey samples were placed in a cuvette 10 mm x 10 mm (Recommendations on uniform color spaces, 1971) and the color parameters in CIELab colorimetric system were measured.

The β – carotene was calculated by using the transmission spectra in the visible region and values for color parameters by software program developed specially for Lovibond PFX 880 from the producer. The water content was determined by measuring the refractive index of honey at room temperature with Abbe's refractometer. The data were corrected at 20^{0}C and values were obtained using methods adopted by the International Honey Commission [5].

Principal components analysis [8, 9]

The aim of the method is to reduce the dimensionality of multivariate data (e.g., wavelengths) whilst preserving as much of the relevant information as possible. PCA is a linear transformation that transforms the data (observations of possibly correlated variables) to a new coordinate system such that the new set of variables, the principal components, are linear functions of the original variables. PCs are uncorrelated, and the greatest variance by any projection of the data comes to lie on the first coordinate, the second greatest variance on the second coordinate, and so on. This is achieved by computing the covariance matrix for the full data set. Then, the eigenvectors and eigenvalues of the covariance matrix are computed, and sorted according to decreasing eigenvalue [8, 9].

All the principal components are orthogonal to each other. The full set of principal components is as large as the original set of variables. Usually the sum of the variances of the first few principal components exceeds 80% of the total variance of the original data [18].

In this study, the first two PCs are used mainly for the purposes of visualization.

Artificial neural network based classifier

It is well known that artificial neural networks with a feed forward multilayered structure are universal function approximators [4, 7]. One classification task can be easily reduced to a task for approximation. Let the classifier of honey be implemented as a neural network with a feed forward structure and Backpropagation (BP) learning algorithm. The neural network consists of n + 3 inputs (n is the number of wavelengths included in the emission spectrum characteristics of the honey), 3 outputs and 2 hidden layers. The three additional inputs are designed for the 3 colorimetric indicators (parameters L, a and b) of the CIELab system. The proposed combination of fluorescent emission spectra with the three colorimetric parameters of CIELab system aims to increase the accuracy of predicting the floral origin of honey. The three outputs of the network correspond to the three classes of honey: acacia, linden and honeydew honey. The two hidden layers contain neurons with 'tansigmoid' activation function (hyperbolic tangent), and the activation function of the three output neurons is 'logsigmoid' [3].

The method of the PCs is applied to one of the emission spectra (Ex. 370 nm), extended by the three colorimetric parameters L, a and b (total n + 3 inputs). The resulting n + 3 PCs are used as input training samples for the neural network. The supervisor supplies the network's output with the following three combinations: '1 0 0', '0 1 0' or '0 0 1', depending on whether the input receives the data for the classes 'acacia', 'linden' or 'honeydew', respectively. BP is a gradient-based learning algorithm, which minimizes the sum squared error between the real and required input of the NN. The leave-one-out-cross-validation test is used to validate the NN-based honey classifier.

3. RESULTS AND DISCUSSION

Fluorescence spectra, color, β -carotene and water content of honey

The normalized fluorescence spectra of a random sample from the three types of honey (acacia, linden, and honeydew) with wavelengths ranging in visible domain under excitation at 370 nm, 395 nm, and 405 nm are shown in Figure 1a. The first maxima (with magnitude 1) correspond to the excitation signals, and the second maxima – to the emission spectra's significant values. The averaged fluorescence (emission) spectra of the three classes of honey (acacia, linden and honeydew) at 370 nm, 395 nm and 405 nm excitation are shown in Figure1b.

Generally, for each type of honey the emission with the highest intensity was obtained at the excitation wavelength of 370 nm, and the lowest intensity - at 405 nm. The honeydew honey had the highest intensity of emission in respect to the other types of honey, and the acacia honey – the lowest intensity. The mean values of the wavelengths with maximum emission varied from 476 nm to 502 nm, under standard deviation from 1.4 nm to 28.9 nm.

Table 1. Water content, β-carotene, and color parameters of honey

		Acacia honey		Linden honey		Honeydew	
		Average value	Standard deviation	Average value	Standard deviation	Average value	Standard deviation
Water content, %		16.30	1.17	16.50	1.08	15.97	0.95
β- carotene		7.11	2.58	15.53	10.24	18.52	21.46
CIE-Lab Values (Ill. D65 / 10 deg Observer / 380-780 nm)	L	92.45	5.46	85.35	6.63	51.27	12.80
	a	0.29	1.78	3.05	4.78	27.94	5.83
	b	32.01	15.00	57.65	16.72	74.54	11.63

Table 1 shows the mean values and standard deviations of the water content, β-carotene and the CIELab color parameters related to the different classes (types) of honey. Moisture content is an important quality parameter that influences the shelf life of honey [5]. It ranged between 13.92% and 18.90% for all 32 samples presented and it was below the upper limit of 20% set by the relevant EU directive.
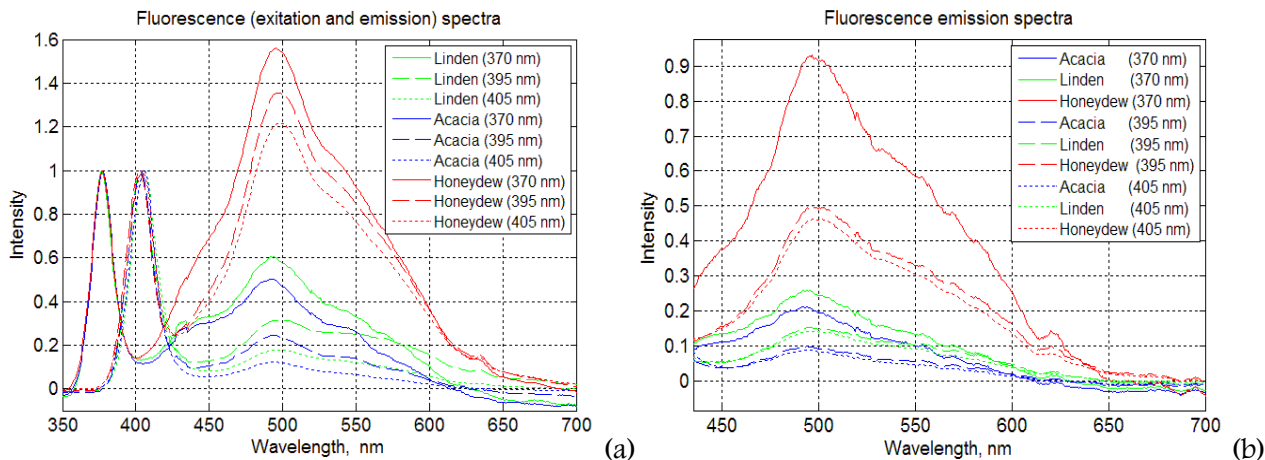

Figure 1. Fluorescence spectra at excitation of 370 nm, 395 nm and 405 nm: (a) excitation and emission spectra for one sample from the three types of honey; (b) the average emission spectra of the three classes of honey (acacia, linden and honeydew).

NN-based model for classification of honey- calibration and testing

Since the intensity of emission spectra of the three types of honey is greatest at excitation of 370 nm (Figure 1b), only the spectral characteristics at this excitation were used for the synthesis of a honey's classifier. PCA was carried out in order to visualize data from different honey samples and to identify their similarities and differences. Using PCA, the spectral dimensionality was reduced to a small number (two) of principal components. The scores scatter plot of the 1st and 2nd PCs is shown in Figure 2a, where samples from classes 'Acacia', 'Linden' and 'Honeydew' are marked with circular, triangular and squared symbol, respectively. It is evident that the samples form three clusters (acacia, linden and honeydew), which are overlapped. Here, determining the type of honey is based solely on the inscription on the label by the manufacturer, i.e. trusting the manufacturer. The two PCs suitably visualize the honey's spectra, but the information contained in them is not enough to properly distinguish different types of honey. Therefore PCA was applied to a combination of fluorescence spectra characteristics and the three indicators (L, a, b) of the colorimetric system CIELab. In this case the first two PCs explain as high as 94.27 % of variance of the combined data (76.21 % for PC-1 and 18.06 % for PC-2). The result (Figure 2b) shows a better distinguishing between different types of honey, with the exception of a few overlapping samples of classes 'acacia' and 'linden'.
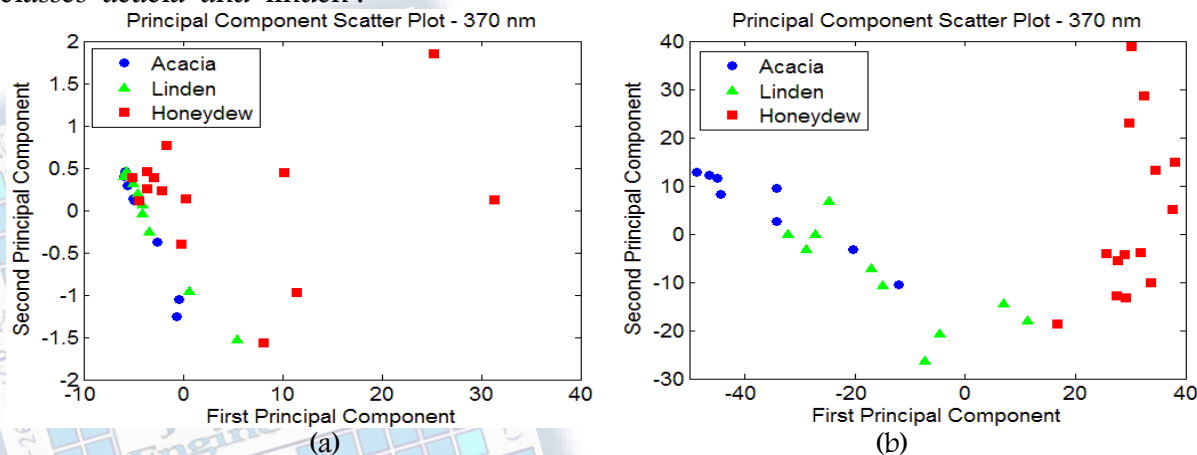

Figure 2. PCA of: (a) the honey's fluorescent spectra (b) a combination of the honey's fluorescent spectra and CIELab's parameters

The neural classifier was trained using the values of all PCs, obtained by PCA applied to the enriched data (spectral characteristics of fluorescence + colour parameters *L*, *a*, *b*). The number of neurons in the first and second hidden layers of the neural network was selected heuristically - 500 and

250, respectively. The efficiency of the neural network classifier was confirmed by leave-one-out-cross-validation test (in MATLAB environment [3]), the result of which is shown in Figure 3 and Table 2. As evident in Table 2, 2 samples from observed class 'acacia' are predicted wrong as 'linden', while 3 samples from class 'linden' – as 'acacia'. The model predicts 27 out of 32 samples correctly. 84.4% prediction accuracy (75% class 'acacia', 70% class 'linden', and 100% class 'honeydew') is achieved.

Figure 3. PCA based visualization of leave-one-out-cross-validation test of the proposed neural network based classifier of honey

Table 2. Discrimination accuracy of NN-based model

		Predicted Class by PC-NN			
		Acacia	Linden	Honeydew	
Observed Class	Acacia	6	2	0	8
	Linden	3	7	0	10
	Honeydew	0	0	14	14
		9	9	14	

4. CONCLUSION

In this article the optical properties of Bulgarian honey were investigated in regard to the potential of honey discrimination on the base of its botanical origin. The fluorescence spectra combined with the colorimetric parameters of CIELab were used for training the neural network based classifier. The classifier shows a good prediction accuracy, 84.4%, determined by the 32 leave-one-out-cross-validation tests. The neural network is open for new honey samples which will precise the clustering and maybe will improve the performance of the honey's floral origin predictor. Future work will include a comparative analysis of the proposed neural network classifier and some popular statistical classifiers, such as those based on a linear or a quadratic discriminant analysis, as well as the use of fluorescence spectra with excitation and emission in the UV region.

Note

This paper is based on the paper presented at The Vth International Conference Industrial Engineering and Environmental Protection 2015 – IIZS 2015, University of Novi Sad, Technical Faculty „Mihajlo Pupin", Zrenjanin, SERBIA, October 15-16th, 2015, referred here as[20].

References

[1] Aljadi A. M., Kamarrudin M. Y., Evaluation of the phenolic contents and antioxidant capacities of two Malaysian floral honeys, Food Chemistry, Vol.85, pp.513-518, 2004.

[2] Batrusaityte V., Venskutonis P.R., Ceksteryte, V., Radical scavenging activity of different floral origin honey and beebread phenolic extracts, Food chemistry, Vol.80, pp.249-254, 2007.

[3] Beale, M.H., Hagan, M.T., Demuth, H.B., Neural Network Toolbox™, User's Guide, The MathWorks Inc., 1992-2015. https://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf

[4] Blum, E.K., Li, L.K., Approximation theory and feedforward networks, Neural Networks, Vol.4, pp.511-515, 1991.

[5] Bogdanov S., Martin P., Lullmann C., Harmonised methods of the European honey commission, Apidologie, pp.1-59, extra issue, 1997.

[6] Chen, L., Wang, J., Ye, Z., Zhao, J., Xue, X., Vander Heyden, Y., Sun, Q., Classification of Chinese honeys according to their floral origin by near infrared spectroscopy. Food Chemistry, Elsevier Ltd, Vol.135, No.2, pp.338-342, 2012.

[7] Hornik, K., Stinchcombe, M., White, H., Multilayer feedforward networks are universal approximators, Neural Networks, Vol.2, pp.359-366, 1989.

[8] Hotelling, H., Analysis of a Complex of Statistical Variables into Principal Components. Journal of Educational Psychology, Vol.24, No.6 & 7, pp.417–441 & pp.498–520, 1933.

[9] Jolliffe, I. T., Principal Component Analysis. Second ed. Springer Series in Statistics. New York: Springer-Verlag New York, 2002.
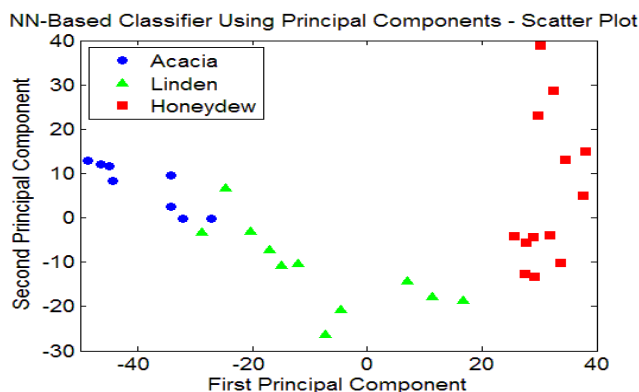
[10] Nagai T., Inoue R., Inoue H., Suzuki N., Scavenging capacities of pollen extracts from Cistus ladaniferus on autoxidation superoxide radicals, hydroxyl radicals and DPPH radicals. Nutrition Research, Vol.22, pp.519-526, 2002.

[11] Persano Oddo, L., Bogdanov, S., Determination of Honey Botanical Origin: Problem and Issues. Apidologie, Vol.35 (special issue), pp.2-3, 2004.

[12] Pyrzynska, K., Biesaga, M., Analysis of phenolic acids and flavonoids in honey. Trends in Analytical Chemistry, Vol.28, pp.893–902, 2009.

[13] Robbins, R. J., Phenolis acids in foods: An overview of analytical methodology. Journal of Agricultural and Food Chemistry, Vol.51, pp.2866–2887, 2003.

[14] Ruoff, K., Authentication of the Botanical Origins of Honey, A Dissertation for the degree of Doctor of Sciences, University of Helsinky, p.32, 2006.

[15] Ruoff, K., Karoui, R., Dufour, E., Luginbuhl, W., Bosset, J.O., Bogdanov, S., Amado, R., Authentication of the botanical origin of honey by front-face fluorescence spectroscopy. A preliminary study. J Agric Food Chem. Vol.53, No.5, pp.1343-1347, 2005.

[16] Sádecká J., Tóthová J., Fluorescence spectroscopy and chemometrics in the food classification − a review. Czech J. Food Sci., Vol.25, pp.159–173, 2007.

[17] Sergiel, I., Pohl, P., Biesaga, M., Mironczyk, A., Suitability of three-dimensional synchronous fluorescence spectroscopy for fingerprint analysis of honey samples with reference to their phenolic profiles, Food Chemistry, Vol.145, pp.319–326, 2014. http://dx.doi.org/10.1016/j.foodchem.2013.08.069

[18] Statistics Toolbox™ User's Guide, R2014a, The MathWorks, Inc., 1993–2014.

[19] Tarboush, A., Kahtani, A., El-Sarrage, M.S., Floral type identification and quality of some honey types, Food Chemistry, Vol.46, pp.13-17, 1993.

[20] Krastena Nikolova, Diana Tsankova, Tinko Evtimov, Fluorescence spectroscopy, colorimetry and neural networks in distinguishing different types of honey, Proceedings of the Vth International Conference Industrial Engineering and Environmental Protection 2015 – IIZS 2015

ANNALS of Faculty Engineering Hunedoara
– International Journal of Engineering